Sum-Rate Maximization in the Multicell MIMO Multiple-Access Channel with Interference Coordination

Duy H. N. Nguyen, *Student Member, IEEE*, and Tho Le-Ngoc, *Fellow, IEEE*

***Abstract*—This paper is concerned with the maximization of the weighted sum-rate (WSR) in the multicell MIMO multiple access channel (MAC). We consider a multicell network operating on the same frequency channel with multiple mobile stations (MS) per cell. Assuming the *interference coordination* mode in the multicell network, each base-station (BS) only decodes the signals for the MSs within its cell, while the inter-cell transmissions are treated as noise. Nonetheless, the uplink precoders are jointly optimized at MSs through the coordination among the cells in order to maximize the network weighted sum-rate (WSR). Since this WSR maximization problem is shown to be nonconvex, obtaining its globally optimal solution is rather computationally complex. Thus, our focus in this work is on low-complexity algorithms to obtain at least locally optimal solutions. Specifically, we propose two iterative algorithms: one is based on successive convex approximation and the other is based on iterative minimization of weighted mean squared error. Both solution approaches shall then reveal the structure of the optimal uplink precoders. In addition, we also show that the proposed algorithms can be implemented in a distributed manner across the coordinated cells. Simulation results show a significant improvement in the network sum-rate by the proposed algorithms, compared to the case with no interference coordination.**

***Index Terms*—Multicell, interference coordination, coordinated multipoint transmission/reception, multiple-input multiple-output, multiple access channel, convex optimization, MMSE.**

I. Introduction

IN the latest 3GPP LTE-Advanced Release, coordinated multi-point transmission/reception (CoMP) has been considered as a technology to improve the system's coverage, throughput, and efficiency [1]. In the downlink direction, CoMP coordinates the simultaneous information transmissions from multiple base-stations (BS) to the mobile-stations (MS), especially to the ones in the cell-edge region. In the uplink direction, CoMP allows the system to take advantage of the multiple receptions at the multiple cells to jointly decode the uplink signal from the MSs. Since CoMP actively utilizes the inter-cell transmissions, instead of treating them as interference, the system performance can be significantly improved [1]. Note that the concept of CoMP is equivalent to that of *network MIMO*, where the multiple BSs are fully coordinated to form a large single antenna array with distributed antenna elements [2]. Thus, existing precoding techniques in single-cell MIMO transmissions can be straightforwardly adopted to this network MIMO configuration. While extracting the most performance from the multicell network, network MIMO comes at the expense of high complexity in joint precoding/decoding and ideal backhaul transmissions among the BSs for data and control signaling exchange [3].

In a lesser level of coordination, namely *interference coordination*, each BS needs to encode or decode only the signals to/from the MSs within its cell. This encoding/decoding strategy relieves the need of data exchange in the backhaul links among the coordinated BSs. The inter-cell transmissions, now being considered as the source of inter-cell interference (ICI), are still fully controlled through the coordination among the cells. In this multicell system with *interference coordination*, transmission/reception techniques for a single-cell MIMO system are no longer applicable and need a rework. In particular, in the downlink direction, the precoders from different BSs have to be coordinately designed to control the ICI. Compared to the multicell system with no interference coordination, significant power reduction or rate enhancement can be obtained by such a joint precoding design across the coordinated BSs [4], [5]. Similarly, in the uplink direction, it is expected that the system performance can be also improved by exploiting interference coordination among transmitting MSs. However, to the best of our knowledge, no work in the literature has addressed this coordinated precoding design to realize its performance enhancement in the uplink transmissions. In contrast to the downlink direction, where the coordinated precoders are designed at the BSs, it may be desirable for the uplink counterpart that each MS is able to determine its precoder distributively with local information only. In this case, the role of the BSs is to exchange useful control signaling to the MSs so that each MS can optimize its precoder on its own. On the other hand, the precoder at each MS has to be devised in a coordinated manner, in order to maximize the link performance to its connected BS while minimizing its induced ICI to other BSs.

In this paper, we examine a coordinated multicell system

Manuscript received June 5, 2012; revised October 24 and March 24, 2013; accepted April 11, 2013. The associate editor coordinating the review of this paper and approving it for publication was L. Deneire.

The work presented in this paper is partly supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Program and the NSERC Collaborative Research and Development Grant with InterDigital Canada. This paper was presented in part at the IEEE Wireless Communications and Networking Conference (IEEE WCNC), Paris, France, Apr 2012.

The authors are with the Department of Electrical Engineering, Department of Electrical and Computer Engineering, McGill University, 3480 University Street, Montreal, QC, Canada, H3A 0E9 (e-mail: huu.n.nguyen@mail.mcgill.ca, tho.le-ngoc@mcgill.ca).

Digital Object Identifier 10.1109/TWC.2013.052213.120802

in a general setting with multiple MSs per cell, where each MS is equipped with multiple transmit antennas. At each cell, the multiple MSs concurrently transmit information signals to its connected BS, which indeed emulates a MIMO multiple-access channel (MAC) system. Per the *interference coordination* mode, the BS then only decodes the signals for its connected MSs by implementing the capacity-achieving decoding technique, namely successive interference cancelation (SIC). The main interest of this paper is the study on how to design the uplink precoders at the MSs with the objective of maximizing the network weighted sum-rate (WSR). Since this WSR maximization problem is shown to be nonconvex, it is generally difficult and computationally complex to find its globally optimal solution. Thus, the main focus of this work is on proposing low-complexity algorithms to break down the nonconvex WSR maximization into a sequence of simpler convex problems.

A. Related Works

It is known that the resource allocation problem (power allocation, precoder design, etc) for maximizing the WSR in an interference network is a challenging task. Even if there is only one MS per cell, where the multicell system is known as an interference channel (IC), the WSR maximization problem turns out to be nonconvex [6]. Several works in literature have examined different numerical techniques to design the transmit precoders to maximize the WSR. Specifically, the gradient projection method was applied in [6] to search for a locally optimal transmit strategy. The works in [7], [8] applied the successive convex approximation technique to decompose the original nonconvex problem into multiple convex problems, which can be solved separately at the transmitters. In particular, each transmitter optimizes its precoder to maximize its link data rate with an *interference-penalty* term on the interference induced to other links [7], [8]. This approach, being referred to as iterative linear approximation (ILA) [9], can be traced back to earlier works in difference of convex (DC) programming [10]–[12], where the nonconvex parts are linearly approximated into the penalty terms. In [6]–[8], by considering only one single MS per cell, the decomposed problem can be readily solved in a closed-form solution at its corresponding BS [8].

One other distributed approach to locally solve the nonconvex WSR maximization problem in an interference network is the weighted mean squared error (WMMSE) algorithm. The main concept of the WMMSE is the transformation of the WSR maximization problem into an equivalent WMMSE problem with some specially chosen weight matrices [13]. The WMMSE problem is then solved by alternating optimizing the weight matrices, the precoders, and the minimum mean squared error (MSE) decoders. Initially proposed in [13] for the single-cell MIMO broadcast downlink channel, the WMMSE algorithm was considered in [14] to maximize the IC's WSR with one data stream per link. Recently, the WMMSE approach has been extended to the multicell downlink system with multiple MSs per cell in [9], where the linear precoders are optimized for the multicell throughput maximization. Compared to the sequential update by the ILA algorithm, the WMMSE algorithm may converge faster due to its distributively and simultaneously updating procedure [9].

B. Contributions of This Work

The main contribution of this work is the development of low-complexity algorithms to solve the nonconvex WSR maximization problem in the multicell MIMO-MAC. The two approaches, namely ILA and WMMSE, are considered in order to maximize the network WSR. In addition, this work presents the distributed implementation of each algorithm, which allows certain operations in the algorithm to be performed in a distributed manner among the coordinated BSs and MSs.

When applying the ILA algorithm to the multicell MIMO-MAC, the approximation and decomposition step converts the nonconvex problem into a sequence of multiple MAC sum-rate maximization problems with *interference-penalty* terms, where each problem corresponds to the MAC at each cell. We then show that each decomposed problem is a convex program, which facilitates the finding of its optimal solution at its corresponding BS. However, due to the consideration of multiple MSs per cell, a closed-form optimal solution to the decomposed problem is not readily available. Instead, by exploring the inherently decoupled constraints for the transmit covariance matrix of each MS, we derive an equivalent optimization problem that can be solved sequentially over each variable matrix by a fast-converging algorithm. Interestingly, the decomposition in the ILA algorithm then reveals the structure of the optimal uplink precoders. In addition, the ILA solution approach also reveals the message signaling mechanism to facilitate its distributed implementation among the coordinated cells.

When applying the WMMSE algorithm to the multicell MIMO-MAC, we show that the network WSR maximization problem can also be transformed into an equivalent WMMSE problem. Taking SIC into consideration, we then show how to optimally determine the precoders at the MSs, the minimum MSE decoders, and the weight matrices at the BSs. In addition, we present the message passing mechanism among the BSs themselves and between the BS and its connected MSs in the multicell MIMO-MAC that facilitates the distributed implementation of the WMMSE algorithm. For both ILA and WMMSE algorithms, monotonic convergence to at least local optimal solutions is subsequently proven. Simulation results show that the proposed algorithms can significantly improve the network WSR, in comparison of the multicell system with no interference coordination among the BSs.[1] The simulations also confirm the convergence analysis of the proposed algorithms.

Notations: $(\mathbf{X})^T$ and $(\mathbf{X})^H$ denote the transpose and conjugate transpose (Hermitian operator) of the matrix $\mathbf{X}$, respectively; $[\mathbf{X}]_{m,n}$ stands for the (m,n)th entry of the matrix $\mathbf{X}$; $[\mathbf{X}]^+$ denotes the component-wise operation $\max\{[\mathbf{X}]_{m,n}, 0\}$; $\mathrm{Tr}\{\mathbf{X}\}$, $|\mathbf{X}|$, and $\mathrm{rank}\{\mathbf{X}\}$ denote the trace, determinant, and

[1] Without interference coordination, the BSs are said to be in *competitive* mode where each BS selfishly maximizes the sum-rate for its connected MSs only. This mode is sometimes referred to as the *interference aware* mode in literature.

rank of the matrix $\mathbf{X}$, respectively; and $x^\star$ denotes the optimal value of the variable x.

II. System Model and Problem Formulation

We consider the multiuser uplink transmission of a multicell system with Q separate cells operating on the same frequency channel. At each cell, multiple MSs, each equipped with multiple transmit antennas, are sending independent data streams to its connected multiple-antenna BS. For the simplicity of presentation, it is assumed that the numbers of antennas at the BS and MS are M and N, respectively, and the number of MSs in each cell is K. Since the multicell system operates on the same frequency channel, the intended signal from a MS to a BS is now subject to the intra-cell interference from other MSs in the same cell, as well as the inter-cell interference from the MSs in other cells. In the *interference coordination* design of this multicell system, the precoders at each MS are jointly optimized to fully manage both the inter-cell and intra-cell interferences.

Consider the MAC at a particular cell, say cell-q, the received signal $\mathbf{y}_q$ at its BS can be modeled as

$$\mathbf{y}_q = \sum_{i=1}^{K} \mathbf{H}_{qq_i}\mathbf{x}_{q_i} + \sum_{r\neq q}^{Q}\sum_{i=1}^{K} \mathbf{H}_{qr_i}\mathbf{x}_{r_i} + \tilde{\mathbf{z}}_q, \tag{1}$$

where $\mathbf{x}_{r_i} \in \mathbb{C}^{N\times 1}$ is the transmitted vector from the N-antenna of MS-i in the rth cell, $\mathbf{H}_{qr_i}$ models the channel from MS-i of cell-r to the qth base station, and $\tilde{\mathbf{z}}_q$ is the zero-mean additive Gaussian noise vector with the covariance matrix $\mathbf{Z}_q$.

Assuming linear precoding at each MS, the transmitted signal from MS-i in cell-q can be expressed as

$$\mathbf{x}_{q_i} = \mathbf{V}_{q_i}\mathbf{s}_{q_i}, \tag{2}$$

where $\mathbf{s}_{q_i} \in \mathbb{C}^{D\times 1}$ represents the information signal vector for MS-i, and $\mathbf{V}_{q_i} \in \mathbb{C}^{N\times D}$ is the precoder matrix for MS-i, and $D = \min(M, N)$ is the maximum number of spatial data streams can be supported by MS-i. Without loss of generality, columns of the precoding matrix $\mathbf{V}_{q_i}$ may be set to zero if the corresponding streams are not active. While our formulation allows each MS to multiplex up to D independent spatial streams., it may not be clear *a priori* how many spatial streams are actually active (capable of carrying information data). Instead, after the optimization process to determine the optimal precoder $\mathbf{V}_{q_i}$ for MS-i, the actual number of data streams for MS-i can be determined easily by assessing the rank of $\mathbf{V}_{q_i}$. It is assumed that $\mathbb{E}\left[\mathbf{s}_{q_i}\mathbf{s}_{q_i}^H\right] = \mathbf{I}$. In addition, the precoder $\mathbf{V}_{q_i}$ is constrained by transmit power limit P_{q_i}, i.e.,

$$\mathrm{Tr}\left\{\mathbf{V}_{q_i}\mathbf{V}_{q_i}^H\right\} \leq P_{q_i}. \tag{3}$$

Let $\mathbf{X}_{q_i} = \mathbb{E}\left[\mathbf{x}_{q_i}\mathbf{x}_{q_i}^H\right] = \mathbf{V}_{q_i}\mathbf{V}_{q_i}^H$ be the transmit covariance matrix of MS-i of cell-q. Since $\mathrm{rank}\{\mathbf{X}_{q_i}\} = \mathrm{rank}\{\mathbf{V}_{q_i}\}$, should the optimization be carried over $\mathbf{X}_{q_i}$, assessing the rank of $\mathbf{X}_{q_i}$ then reveals the number of active streams supported by MS-i of cell-q. Denote $\mathbf{X}_q = \{\mathbf{X}_{q_i}\}_{i=1}^K$ as the uplink precoder profile of the K users in cell-q. Likewise, denote $\mathbf{X}_{-q} = (\mathbf{X}_1, \ldots, \mathbf{X}_{q-1}, \mathbf{X}_{q+1}, \ldots, \mathbf{X}_Q)$ as the precoding profile of all cells except cell-q. Denote

$$\mathbf{z}_q = \sum_{r\neq q}^{Q}\sum_{i=1}^{K} \mathbf{H}_{qr_i}\mathbf{x}_{r_i} + \tilde{\mathbf{z}}_q \tag{4}$$

as the total ICI plus noise (IPN) at BS-q, and

$$\mathbf{R}_q = \mathbb{E}\left[\mathbf{z}_q\mathbf{z}_q^H\right] = \sum_{r\neq q}^{Q}\sum_{i=1}^{K} \mathbf{H}_{qr_i}\mathbf{X}_{r_i}\mathbf{H}_{qr_i}^H + \mathbf{Z}_q \tag{5}$$

as the covariance matrix of the IPN at BS-q.

Per the *interference coordination* design being considered, each BS only attempts to decode the signals from its connected MSs using the capacity-achieving multiuser decoding technique, namely successively interference cancelation (SIC) [15]. For instance, BS-q employs SIC for the transmissions from the K MSs within cell-q. Assuming the decoding order from MS-1 (first) to MS-K (last), for a certain IPN covariance $\mathbf{R}_q$, the achievable rate of user-i in cell-q can be expressed as

$$R_{q_i}(\mathbf{X}_q, \mathbf{X}_{-q}) = \log \frac{\left|\mathbf{R}_q + \sum_{j=i}^{K}\mathbf{H}_{qq_j}\mathbf{X}_{q_j}\mathbf{H}_{qq_j}^H\right|}{\left|\mathbf{R}_q + \sum_{j>i}^{K}\mathbf{H}_{qq_j}\mathbf{X}_{q_j}\mathbf{H}_{qq_j}^H\right|}, \tag{6}$$

where the intra-cell interference from user-1 to user-$(i-1)$ has been suppressed. Collectively, the MAC sum-rate of all K users in cell-q is given by [16]

$$\begin{aligned} R_q\left(\mathbf{X}_q, \mathbf{X}_{-q}\right) &= \sum_{i=1}^{K} R_{q_i} \\ &= \log\left|\mathbf{I} + \mathbf{R}_q^{-1}\left(\sum_{i=1}^{K}\mathbf{H}_{qq_i}\mathbf{X}_{q_i}\mathbf{H}_{qq_i}^H\right)\right|, \end{aligned} \tag{7}$$

where the denominators inside the log function are sequentially eliminated. Note that this sum-rate is obtained when BS-q does not decode the transmissions from the users in other cells. The network WSR is then given by $\sum_{q=1}^{Q} w_q R_q(\mathbf{X}_q, \mathbf{X}_{-q})$, where w_q denoted the nonnegative weight of cell-q. To maximize the network WSR, let us consider the following optimization

$$\begin{aligned} \underset{\mathbf{X}_1,\ldots,\mathbf{X}_Q}{\text{maximize}} \quad & \sum_{q=1}^{Q} w_q \log\left|\mathbf{I} + \mathbf{R}_q^{-1}\left(\sum_{i=1}^{K}\mathbf{H}_{qq_i}\mathbf{X}_{q_i}\mathbf{H}_{qq_i}^H\right)\right| \\ \text{subject to} \quad & \mathrm{Tr}\{\mathbf{X}_{q_i}\} \leq P_{q_i},\ \forall i, \forall q \\ & \mathbf{X}_{q_i} \succeq 0,\ \forall i, \forall q. \end{aligned} \tag{8}$$

It is observed that problem (8) is clearly a nonconvex problem due to the presence of $\mathbf{X}_{q_i}$'s in interference terms $\mathbf{R}_r$'s, $r \neq q$. Thus, obtaining the globally optimal solution to the problem is computationally complex and intractable for practical applications. It may also require a centralized solver unit to obtain such a solution. In this case, designing low-complexity algorithms with distributed implementation to compute local optimizers becomes a more attractive option. To this end, we examine two simple and fast converging algorithms with the goals: (i) obtaining at least locally optimal solutions to the problem, and (ii) distributed implementation which does not require a centralized unit nor full CSI across the coordinated cells.

III. The Iterative Linear Approximation (ILA) Algorithm

A. *The Iterative Linear Approximation Solution Approach*

This section presents a solution approach to the original nonconvex problem (8) by considering it as a DC program [10]–[12]. Specifically, by iteratively isolating and approximating the nonconvex part of the objective function, the DC program shall be decomposed into multiple convex optimization problems, which can be solved distributively at each MS with low complexity. In addition, the iterative procedure allows the MSs to continuously refine and improve their uplink precoders, which eventually yields a local optimal solution of the original problem.

Denote $f_q(\mathbf{X}_q, \mathbf{X}_{-q}) = \sum_{r\neq q}^{Q} w_r R_r(\mathbf{X}_q, \mathbf{X}_{-q})$ as the WSR of all other cells except cell-q so that the network WSR can be expressed as $w_q R_q(\mathbf{X}_q, \mathbf{X}_{-q}) + f_q(\mathbf{X}_q, \mathbf{X}_{-q})$. Since $f_q(\mathbf{X}_q, \mathbf{X}_{-q})$ is not a concave function in $\mathbf{X}_{q_i}$, we take the approximation to this term. At a given value $(\bar{\mathbf{X}}_q, \bar{\mathbf{X}}_{-q})$, taking the Taylor expansion of f_q around $\bar{\mathbf{X}}_{q_i}$, $i = 1, \ldots, K$, and retaining the first linear term

$$f_q(\mathbf{X}_q, \bar{\mathbf{X}}_{-q}) \approx f_q(\bar{\mathbf{X}}_q, \bar{\mathbf{X}}_{-q}) - \sum_{i=1}^{K} \mathrm{Tr}\left\{\mathbf{A}_{q_i}\left(\mathbf{X}_{q_i} - \bar{\mathbf{X}}_{q_i}\right)\right\}, \quad (9)$$

where $\mathbf{A}_{q_i}$ is the negative partial derivative of f_q with respect to $\mathbf{X}_{q_i}$, evaluated at $\mathbf{X}_{q_i} = \bar{\mathbf{X}}_{q_i}$, given at the bottom of this page.

Using (9), the network WSR around $\bar{\mathbf{X}}_q$, $w_q R_q(\mathbf{X}_q, \bar{\mathbf{X}}_{-q}) + f_q(\mathbf{X}_q, \bar{\mathbf{X}}_{-q})$, can be approximated as $w_q R_q(\mathbf{X}_q, \bar{\mathbf{X}}_{-q}) - \sum_{i=1}^{K} \mathrm{Tr}\{\mathbf{A}_{q_i}\mathbf{X}_{q_i}\} + \left[f_q(\bar{\mathbf{X}}_q, \bar{\mathbf{X}}_{-q}) + \sum_{i=1}^{K} \mathrm{Tr}\left\{\mathbf{A}_{q_i}\bar{\mathbf{X}}_{q_i}\right\}\right]$. Since the term $f_q(\bar{\mathbf{X}}_q, \bar{\mathbf{X}}_{-q}) + \sum_{i=1}^{K} \mathrm{Tr}\left\{\mathbf{A}_{q_i}\bar{\mathbf{X}}_{q_i}\right\}$ is now fixed and consequently does not affect the maximization of the network WSR. As a result, it can be omitted in the objective function, i.e., maximizing $w_q R_q(\mathbf{X}_q, \bar{\mathbf{X}}_{-q}) + f_q(\mathbf{X}_q, \bar{\mathbf{X}}_{-q})$ is equivalent to maximizing $w_q R_q(\mathbf{X}_q, \bar{\mathbf{X}}_{-q}) - \sum_{i=1}^{K} \mathrm{Tr}\{\mathbf{A}_{q_i}\mathbf{X}_{q_i}\}$, and the nonconvex problem (8) can be approximated as

$$\begin{aligned} \underset{\mathbf{X}_{q_1},\ldots,\mathbf{X}_{q_K}}{\text{maximize}} \quad & w_q \log\left|\mathbf{R}_q + \sum_{i=1}^{K} \mathbf{H}_{qq_i}\mathbf{X}_{q_i}\mathbf{H}_{qq_i}^H\right| - \sum_{i=1}^{K}\mathrm{Tr}\{\mathbf{A}_{q_i}\mathbf{X}_{q_i}\} \\ \text{subject to} \quad & \mathrm{Tr}\{\mathbf{X}_{q_i}\} \le P_{q_i}, \ \forall i \\ & \mathbf{X}_{q_i} \succeq \mathbf{0}. \end{aligned} \quad (11)$$

which can be solved solely at cell-q. In other words, the optimization problem (8) can be approximately solved as Q per-cell separate optimization problems (11).

It is observed that the approximated problem (11) is similar to the MAC sum-rate maximization problem, studied in [16]. The difference here is the presence of the penalty term $\sum_{i=1}^{K} \mathrm{Tr}\{\mathbf{A}_{q_i}\mathbf{X}_{q_i}\}$, which encourages cell-$q$ to adopt a more cooperative precoding strategy by limiting the ICI to other cells. Should this term be not presented, the multicell system is said to be in *competitive* mode where each cell would selfishly maximize the sum-rate for its connected users only. This results in a noncooperative game among the cells, similar to the game studied in [17] for the case of one MS per cell. We shall present some numerical results for this noncooperative design in comparison to the considered coordinated design.

Note that the decomposed problem (11), corresponding to the precoder design at cell-q, is now a convex program, unlike the original problem (8). Thus, it can be readily solved by any efficient convex optimization techniques [18]. However, these direct solution approaches may require a centralized solver unit at the BS, and hence are not suitable for distributed implementation at the MSs for the MAC. Fortunately, it is observed that the constraints for each transmit covariance matrix $\mathbf{X}_{q_i}$ are inherently decoupled in problem (11). By exploring this decoupled structure, the optimization (11) can be solved sequentially over each variable matrix, like the MAC sum-rate maximization problem in [16]. More importantly, the optimization process over each variable can be performed at the corresponding MS in a fully distributed manner. We elaborate these observations in the following theorem and later propose a fast-converging and distributed algorithm to solve problem (11).

Theorem 1. *For the* K*-user problem* (11), $\left\{\mathbf{X}_{q_i}\right\}_{i=1}^{K}$ *is an optimal solution if and only if* $\mathbf{X}_{q_i}$ *is the solution of the following optimization problem*

$$\begin{aligned} \underset{\mathbf{X}_{q_i}}{\text{maximize}} \quad & w_q \log\left|\mathbf{I} + \mathbf{R}_{q_i}^{-1}\mathbf{H}_{qq_i}\mathbf{X}_{q_i}\mathbf{H}_{qq_i}^H\right| - \mathrm{Tr}\{\mathbf{A}_{q_i}\mathbf{X}_{q_i}\} \\ \text{subject to} \quad & \mathrm{Tr}\{\mathbf{X}_{q_i}\} \le P_{q_i}, \mathbf{X}_{q_i} \succeq \mathbf{0}, \end{aligned} \quad (12)$$

where

$$\mathbf{R}_{q_i} = \mathbf{R}_q + \sum_{j\neq i}^{K} \mathbf{H}_{qq_j}\mathbf{X}_{q_j}\mathbf{H}_{qq_j}^H \quad (13)$$

is considered as noise.

Proof: Please refer to Appendix A. ■

Since the structure of the optimal solution to problem (11) has been revealed in Theorem 1, one can easily obtain its optimal solution by sequentially solving problem (12) for each user, i.e., MS-1 to MS-K in cell-q, until convergence. We note that each problem (12) can be effectively solved by the water-filling process, as presented in [8]. This sequential optimization at cell-q accounts for the *inner-loop* iterative precoder updates of the K MSs in cell-q.

$$\begin{aligned} \mathbf{A}_{q_i} = -\left.\frac{\partial f_q}{\partial \mathbf{X}_{q_i}}\right|_{\mathbf{X}_{q_i}=\bar{\mathbf{X}}_{q_i}} &= -\sum_{r\neq q}^{Q} w_r \left.\frac{\partial R_r}{\partial \mathbf{X}_{q_i}}\right|_{\mathbf{X}_{q_i}=\bar{\mathbf{X}}_{q_i}} \\ &= \sum_{r\neq q}^{Q} w_r \mathbf{H}_{rq_i}^H \left[\mathbf{R}_r^{-1} - \left(\mathbf{R}_r + \sum_{j=1}^{K}\mathbf{H}_{rr_j}\mathbf{X}_{r_j}\mathbf{H}_{rr_j}^H\right)^{-1}\right]\mathbf{H}_{rq_i}\Bigg|_{\mathbf{X}_{q_i}=\bar{\mathbf{X}}_{q_i}}. \end{aligned} \quad (10)$$

For the problem of Q cells (8), the proposed ILA algorithm requires each cell-q, $q = 1, \ldots, Q$ to continuously update the parameters $\{\mathbf{A}_{q_i}\}_{i=1}^{K}$ and take turns to solve its corresponding optimization (11). This sequential procedure accounts for the *outer-loop* iterative updates across the Q cells. We summarize the ILA algorithm for the multicell MIMO-MAC in Algorithm 1. The convergence of the proposed ILA algorithm is given in the following theorem.

Algorithm 1: ILA Algorithm for Multicell MIMO-MAC

Initialize $\{\mathbf{X}_{q_i}\}_{\forall q, \forall i}$, such that $\mathrm{Tr}\{\mathbf{X}_{q_i}\} = P_{q_i}$;
repeat
  $\bar{\mathbf{X}}_{q_i} \leftarrow \mathbf{X}_{q_i}$;
  for $q = 1, 2, \ldots, Q$ **do**
    Compute $\mathbf{R}_q$ with $\bar{\mathbf{X}}_{q_i}$ at BS-q and exchange among the BSs;
    At BS-q, update the pricing matrix $\mathbf{A}_{q_i}$ at MS-i and perform ;
    repeat
      for $i = 1, 2, \ldots, K$ **do**
        Compute $\mathbf{R}_{q_i}$ at the BS and pass it to MS-i;
        Perform $\underset{\mathbf{X}_{q_i}}{\text{maximize}} \;\; w_q \log \left| \mathbf{I} + \mathbf{R}_{q_i}^{-1} \mathbf{H}_{qq_i} \mathbf{X}_{q_i} \mathbf{H}_{qq_i}^{H} \right| - \mathrm{Tr}\left\{\mathbf{A}_{q_i} \mathbf{X}_{q_i}\right\}$, with $\mathrm{Tr}\{\mathbf{X}_{q_i}\} \leq P_{q_i}$ at MS-i
      end
    until *convergence*;
  end
until *convergence*;

Theorem 2. *The Gauss-Seidel (sequential) iterative update always improves network WSR and is guaranteed to converge to at least a local maximum.*

Proof: Please refer to Appendix B. ■

It is worth mentioning that the proposed ILA algorithm can be executed by a central controller, which then passes the local optimal precoders to the corresponding MSs. In this case, the central controller must possess the full CSI knowledge of all channels in the network. On the other hand, it is possible to implement the proposed ILA algorithm in a distributed manner by assigning certain optimization steps in the algorithm to be performed each coordinated BS and MS. To this end, we shall next present the interpretation to the ILA algorithm that allows its distributed implementation. The complexity in implementing the algorithm will be then discussed in Section V.

B. Distributed Implementation of The Proposed ILA Algorithm

In order to realize the distributed implementation of the proposed ILA algorithm, we make the following assumptions:

- *Assumption 1:* Each MS knows the channel matrices $\mathbf{H}_{rq_i}$'s to all the BS-r's in the network. This assumption allows the MS to control its induced ICI to other cells.
- *Assumption 2:* The coordinated BSs have reliable backhaul channels to exchange control information among themselves.
- *Assumption 3:* The channels are in block-fading or vary sufficiently slow such that they can be considered fixed during the optimization being performed.

It is noted that Algorithm 1 involves two levels of computations. At the inner-loop level, assuming $\mathbf{R}_q$ known at BS-q and $\mathbf{A}_{q_i}$ known at MS-i, cell-q performs the corresponding optimization (11) autonomously using the result from Theorem 1. The role of BS-q is to measure the total signaling plus noise $\mathbf{R}_q + \sum_{i=1}^{K} \mathbf{H}_{qq_i}\mathbf{X}_{q_i}\mathbf{H}_{qq_i}^{H}$, then pass this value to its connected MSs. MS-i in cell-q, having known its channel to its BS $\mathbf{H}_{qq_i}$, can compute the noise component $\mathbf{R}_{q_i}$. MS-i is then required to update its uplink covariance matrix by solving the optimization (12). This process, which corresponds to the inner-loop iterations, is performed until convergence in cell-q.

At the outer-loop level, each BS needs to exchange the data to compute the parameters $\{\mathbf{A}_{q_i}\}_{i=1}^{K}$ for next update. It is observed from equation (10) that MS-i in cell-q needs to know the channels $\mathbf{H}_{rq_i}$'s to all the BSs (per Assumption 1), as well as the pricing matrix

$$\mathbf{B}_r = \mathbf{R}_r^{-1} - \left(\mathbf{R}_r + \sum_{j=1}^{K} \mathbf{H}_{rr_j}\mathbf{X}_{r_j}\mathbf{H}_{rr_j}^{H} \right)^{-1}, \tag{14}$$

in order to compute $\mathbf{A}_{q_i}$. Thus, it is required that each BS needs to compute its corresponding price $\mathbf{B}_q, q = 1, \ldots, Q$, using local measurements on the IPN $\mathbf{R}_q$ and the total signal plus IPN $\mathbf{R}_q + \sum_{i=1}^{K} \mathbf{H}_{qq_i}\mathbf{X}_{q_i}\mathbf{H}_{qq_i}^{H}$. These factors are then exchanged among the BSs. Using the messages received from other cells, BS-q then can easily pass $\{\mathbf{B}_r\}_{r \neq q}$ to its connected MSs before the inner-loop iterative procedure. The outer-loop iteration is performed until the WSR reaches to a local maximum, as stated in Theorem 2.

Remark 1: It is shown in Theorem 2 that the proposed ILA algorithm allows the uplink precoders to be refined and improved after each Gauss-Seidel update, which ultimately converges to a local maximum. However, this update mechanism requires all the BSs to compute the pricing matrices $\mathbf{B}_q$'s and exchange them within the network, after one cell updates its precoding matrices. To reduce the amount of information exchanges among the coordinated cells, the proposed algorithm can be also implemented using the Jacobi (simultaneous) iterative update. In particular, after the exchange of the pricing matrices, all the cells *simultaneously* update their precoding matrices. Specifically, the inner-loop iterations can be performed independently and concurrently in each cell. Although the convergence of the Jacobi update is not analytically proved, most of the numerical simulations confirm its rapid convergence rate. Thus, in our simulations for the ILA algorithm, we utilize the Jacobi update instead of the Gauss-Seidel to reduce the computational time.

Remark 2: When the multicell MIMO-MAC system operates under the *competitive* mode, the update of the precoders across the Q cells also involve two levels of iterations. In an outer-loop iteration, each BS, say BS-q, needs to measure its IPN covariance matrix $\mathbf{R}_q$. In an inner-loop iteration, BS-q needs to continuously measure and pass its total signal plus IPN $\mathbf{R}_q + \sum_{i=1}^{K} \mathbf{H}_{qq_i}\mathbf{X}_{q_i}\mathbf{H}_{qq_i}^{H}$ to its K connected MSs, while the MSs at cell-q take turns to selfishly maximize the MAC sum-rate of cell-q by the iterative water-filling procedure [16]. Compared to this *competitive* mode, the ILA algorithm requires the inter-BS signaling in each outer-loop iteration for exchanging the pricing matrices $\mathbf{B}_q$'s. However, the pricing matrices $\mathbf{B}_r, r \neq q$ are required to be passed from BS-q to its K connected MSs only once time before the inner-loop iterations. Thus, the ILA algorithm does require a rather similar amount of intra-cell BS-MS signaling as the *competitive* mode.

IV. The Weighted Minimum Mean Squared Error (WMMSE) Algorithm

A. The Weighted Minimum Mean Squared Error Solution Approach

In Section III, we have examined a linear convex approximation technique to solve the nonconvex optimization problem (8) by successively improving the uplink covariance matrices at the MSs. In this section, we examine a different approach to solve this optimization problem by relating it to a matrix-weighted sum-MSE minimization problem. In particular, the new problem of interest is to alternately find the transmit beamformers $\mathbf{V}_{q_i}$'s and their corresponding receive beamformers $\mathbf{U}_{q_i}$'s, which shall be characterized shortly.

With $\mathbf{V}_{q_i}$'s as the variables to be optimized, the optimization problem (8) can be restated as

$$\begin{aligned} \underset{\{\mathbf{V}_{q_i}\}_{\forall i, \forall q}}{\text{maximize}} \quad & \sum_{q=1}^{Q} w_q \sum_{i=1}^{K} R_{q_i} \\ \text{subject to} \quad & \mathrm{Tr}\{\mathbf{V}_{q_i}\mathbf{V}_{q_i}^H\} \leq P_{q_i},\ \forall i, \forall q, \end{aligned} \tag{15}$$

where the achievable rate R_{q_i}, given in (6), can be rewritten as

$$R_{q_i} = \log \frac{\left|\mathbf{R}_q + \sum_{j=i}^{K} \mathbf{H}_{qq_j}\mathbf{V}_{q_j}\mathbf{V}_{q_j}^H\mathbf{H}_{qq_j}^H\right|}{\left|\mathbf{R}_q + \sum_{j>i}^{K} \mathbf{H}_{qq_j}\mathbf{V}_{q_j}\mathbf{V}_{q_j}^H\mathbf{H}_{qq_j}^H\right|} \tag{16}$$

with $\mathbf{R}_q = \sum_{r\neq q}^{Q}\sum_{j=1}^{K} \mathbf{H}_{qr_j}\mathbf{V}_{r_j}\mathbf{V}_{r_j}^H\mathbf{H}_{qr_j}^H + \mathbf{Z}_q$.

It is noted that this achievable rate can be stated as a function of the error covariance matrix after the MMSE receive filtering. Since SIC is applied at each BS to its connected MSs, the signal from user-i is not corrupted by the intra-cell interference from user-i to user-$(i-1)$. Thus, while treating the interference as noise, the estimated signal for user-i in BS-q is then given by

$$\hat{\mathbf{s}}_{q_i} = \mathbf{U}_{q_i}^H \left(\sum_{j=i}^{K} \mathbf{H}_{qq_j}\mathbf{V}_{q_j}\mathbf{s}_{q_j} + \mathbf{z}_q\right), \tag{17}$$

where $\mathbf{U}_{q_i}^H$ is the linear receive beamformer for user-i. This receive beamformer is designed such that the MSE for the data streams from user-i of cell-q is minimized. As the MSE matrix $\mathbf{E}_{q_i}$ is obtained from

$$\begin{aligned} \mathbf{E}_{q_i} &= \mathbb{E}\left[(\hat{\mathbf{s}}_{q_i} - \mathbf{s}_{q_i})(\hat{\mathbf{s}}_{q_i} - \mathbf{s}_{q_i})^H\right] \\ &= \left(\mathbf{I} - \mathbf{U}_{q_i}^H\mathbf{H}_{qq_i}\mathbf{V}_{q_i}\right)\left(\mathbf{I} - \mathbf{U}_{q_i}^H\mathbf{H}_{qq_i}\mathbf{V}_{q_i}\right)^H \\ &\quad + \sum_{j>i}^{K} \mathbf{U}_{q_i}^H\mathbf{H}_{qq_j}\mathbf{V}_{q_j}\mathbf{V}_{q_j}^H\mathbf{H}_{qq_j}^H\mathbf{U}_{q_i} \\ &\quad + \sum_{r\neq q}^{Q}\sum_{j=1}^{K} \mathbf{U}_{q_i}^H\mathbf{H}_{qr_j}\mathbf{V}_{r_j}\mathbf{V}_{r_j}^H\mathbf{H}_{qr_j}^H\mathbf{U}_{q_i} + \mathbf{U}_{q_i}^H\mathbf{Z}_q\mathbf{U}_{q_i}, \end{aligned} \tag{18}$$

the optimal receive beamformer is indeed the well-known MMSE filter

$$\begin{aligned} \mathbf{U}_{q_i} &= \arg\min_{\mathbf{U}_{q_i}} \|\mathbf{E}_{q_i}\| \\ &= \left(\sum_{j=i}^{K} \mathbf{H}_{qq_j}\mathbf{V}_{q_j}\mathbf{V}_{q_j}^H\mathbf{H}_{qq_j}^H \right. \\ &\quad \left. + \sum_{r\neq q}^{Q}\sum_{j=1}^{K} \mathbf{H}_{qr_j}\mathbf{V}_{r_j}\mathbf{V}_{r_j}^H\mathbf{H}_{qr_j}^H + \mathbf{Z}_q\right)^{-1} \mathbf{H}_{qq_i}\mathbf{V}_{q_i}. \end{aligned} \tag{19}$$

Consequently, the minimum MSE for user-i in cell-q is given by

$$\begin{aligned} \mathbf{E}_{q_i}^{\mathrm{MMSE}} &= \mathbf{I} - \mathbf{U}_{q_i}^H\mathbf{H}_{qq_i}\mathbf{V}_{q_i} \\ &= \left[\mathbf{I} + \mathbf{V}_{q_i}^H\mathbf{H}_{qq_i}^H \left(\sum_{j>i}^{K} \mathbf{H}_{qq_j}\mathbf{V}_{q_j}\mathbf{V}_{q_j}^H\mathbf{H}_{qq_j}^H \right.\right. \\ &\quad \left.\left. + \sum_{r\neq q}^{Q}\sum_{j=1}^{K} \mathbf{H}_{qr_j}\mathbf{V}_{r_j}\mathbf{V}_{r_j}^H\mathbf{H}_{qr_j}^H + \mathbf{Z}_q\right)^{-1} \mathbf{H}_{qq_i}\mathbf{V}_{q_i}\right]^{-1} \end{aligned} \tag{20}$$

Given (16) and (20), the well-known relationship between the data rate and the MMSE matrix can be stated as

$$R_{q_i} = \log\left|\left(\mathbf{E}_{q_i}^{\mathrm{MMSE}}\right)^{-1}\right|. \tag{21}$$

Utilizing this relationship, we establish the equivalence between the WSR maximization problem in the multicell MIMO-MAC and the matrix-weighted sum-MSE minimization in the following theorem.

Theorem 3. *The multicell MIMO-MAC WSR maximization problem* (15) *is equivalent to the following matrix weighted sum-MSE minimization*

$$\begin{aligned} \underset{\mathbf{W}_{q_i}, \mathbf{V}_{q_i}, \mathbf{U}_{q_i}}{\text{minimize}} \quad & \sum_{q=1}^{Q} w_q \sum_{i=1}^{K} \left[\mathrm{Tr}\left\{\mathbf{W}_{q_i}\mathbf{E}_{q_i}\right\} - \log\left|\mathbf{W}_{q_i}\right|\right] \\ \text{subject to} \quad & \mathrm{Tr}\left\{\mathbf{V}_{q_i}\mathbf{V}_{q_i}^H\right\} \leq P_{q_i}, \forall q, \forall i, \end{aligned} \tag{22}$$

where $\mathbf{W}_{q_i} \succeq \mathbf{0}$ is the weight matrix for MS-i at cell-q. In particular, the globally optimal solutions $\{\mathbf{V}\}_{\forall q, \forall i}$ are identical for the two problems.

Proof: The proof for this theorem follows the same spirit as in [9], [13] for the case of downlink transmission. First, it is noticed that there is no constraints to the weight matrices $\mathbf{W}_{q_i}$'s and the receive beamformers $\mathbf{U}_{q_i}$'s in problem (22). Fixing all other variables, the optimal weight matrices $\mathbf{W}_{q_i}$'s are given by

$$\mathbf{W}_{q_i} = \mathbf{E}_{q_i}^{-1} = \mathbf{I} - \mathbf{U}_{q_i}^H\mathbf{H}_{qq_i}\mathbf{V}_{q_i}. \tag{23}$$

On the other hand, the optimal receive beamformers $\mathbf{U}_{q_i}$'s to problem (22) are the MMSE receivers, given in (19). Thus, with the optimal $\mathbf{W}_{q_i}$'s and $\mathbf{U}_{q_i}$'s, the following optimization problem is equivalent to problem (22)

$$\begin{aligned} \underset{\{\mathbf{V}_{q_i}\}_{\forall i, \forall q}}{\text{minimize}} \quad & \sum_{q=1}^{Q} w_q \sum_{i=1}^{K} \log\left|\mathbf{E}_{q_i}^{\mathrm{MMSE}}\right| \\ \text{subject to} \quad & \mathrm{Tr}\left\{\mathbf{V}_{q_i}\mathbf{V}_{q_i}^H\right\} \leq P_{q_i}, \forall q, \forall i. \end{aligned} \tag{24}$$

This problem is indeed the WSR maximization problem for the multicell MIMO-MAC (15), due to the connection between the MMSE matrix $\mathbf{E}_{q_i}^{\mathrm{MMSE}}$ and the rate $\mathbf{R}_{q_i}$ given in (21). ■

Having established the equivalence between the two optimization problems (15) and (22), we now proceed to solve the latter problem, which then prompts the solution to the former one. Note that the objective function in (22) is convex in each of the optimization variables $\mathbf{U}_{q_i}, \mathbf{V}_{q_i}, \mathbf{W}_{q_i}$. Thus, it is possible to solve problem (22) by alternately optimizing one of the variables while fixing the other two until convergence. First, with fixed transmit beamformers $\mathbf{V}_{q_i}$'s, the receive beamformers $\mathbf{U}_{q_i}$'s are optimally designed as in (19). Second, fixing the transmit and receive beamformers $\mathbf{V}_{q_i}$'s and $\mathbf{U}_{q_i}$'s, the weighted-matrices $\mathbf{W}_{q_i}$'s are updated in a closed form solution, given in (23). Finally, by decomposing the objective function in (22), the update transmit beamformers $\mathbf{V}_{q_i}$'s are carried by solving decoupled optimization problems across the MSs

$$\begin{aligned}\underset{\mathbf{V}_{q_i}}{\text{minimize}}\quad & w_q \mathrm{Tr}\left\{\mathbf{W}_{q_i}\left(\mathbf{I}-\mathbf{U}_{q_i}^H\mathbf{H}_{qq_i}\mathbf{V}_{q_i}\right)\left(\mathbf{I}-\mathbf{U}_{q_i}^H\mathbf{H}_{qq_i}\mathbf{V}_{q_i}\right)^H\right\}\\ & +w_q\sum_{j=1}^{i-1}\mathrm{Tr}\left\{\mathbf{W}_{q_j}\mathbf{U}_{q_j}^H\mathbf{H}_{qq_i}\mathbf{V}_{q_i}\mathbf{V}_{q_i}^H\mathbf{H}_{qq_i}^H\mathbf{U}_{q_j}\right\}\\ & +w_r\sum_{r\neq q}^{Q}\sum_{j=1}^{K}\mathrm{Tr}\left\{\mathbf{W}_{r_j}\mathbf{U}_{r_j}^H\mathbf{H}_{rq_i}\mathbf{V}_{q_i}\mathbf{V}_{q_i}^H\mathbf{H}_{rq_i}^H\mathbf{U}_{r_j}\right\} \quad (25)\\ \text{subject to}\quad & \mathrm{Tr}\left\{\mathbf{V}_{q_i}\mathbf{V}_{q_i}^H\right\}\leq P_{q_i}.\end{aligned}$$

It is noted that this optimization can be carried independently and simultaneously across the MSs. As stated in [9], [13], this problem is a convex quadratic program, which can be optimally solved by standard optimization techniques. Using the Lagrangian duality, the optimal solution to (25) can be derived as

$$\begin{aligned}\mathbf{V}_{q_i} = w_q & \left(\sum_{r\neq q}^{Q}\sum_{j=1}^{K} w_r\mathbf{H}_{rq_i}^H\mathbf{U}_{r_j}\mathbf{W}_{r_j}\mathbf{U}_{r_j}^H\mathbf{H}_{rq_i}\right.\\ & \left.+\sum_{j=1}^{i} w_q\mathbf{H}_{qq_i}^H\mathbf{U}_{q_j}\mathbf{W}_{q_j}\mathbf{U}_{q_j}^H\mathbf{H}_{qq_i}+\mu_{q_i}^{\star}\mathbf{I}\right)^{-1}\mathbf{H}_{qq_i}^H\mathbf{U}_{q_i}\mathbf{W}_{q_i},\end{aligned} \quad (26)$$

where $\mu_{q_i}^{\star}$ is the optimal Lagrangian multiplier associated with the power constraint at the MS. It is noticed that in case of $\mu_{q_i}^{\star}=0$ resulting in $\mathrm{Tr}\left\{\mathbf{V}_{q_i}\mathbf{V}_{q_i}^H\right\}<P_{q_i}$, the corresponding MS effectively does not transmit at full power. Otherwise, $\mu_{q_i}^{\star}$ can be easily obtained by the bisection method such that the power constraint is met with equality.

We summarize the proposed WMMSE algorithm for the multicell MIMO-MAC as in Algorithm 2. In this algorithm, at each outer-loop iteration, the updates can be performed simultaneously across the Q coordinated cells. At *each* cell-q, the corresponding BS-q *sequentially* updates the receive beamformers $\mathbf{U}_{q_i}$ and the weight matrices $\mathbf{W}_{q_i}$ for its MS-i for i from 1 to K. After that, the transmit beamformers $\mathbf{V}_{q_i}$'s can be updated *simultaneously* by its MS-i. The iterative process is performed until each variable converges to a local optimal solution. We summarize the convergence and the optimality of the proposed WMMSE algorithm in the following theorem.

Algorithm 2: WMMSE Algorithm for Multicell MIMO-MAC

```
Initialize {V_{q_i}}_{∀q,∀i}, such that Tr{V_{q_i} V_{q_i}^H} = P_{q_i};
repeat
    Set V̄_{q_i} ← V_{q_i}, ∀q,∀i;
    Simultaneously update across Q cells;
    for q = 1,...,Q do
        At the BS, sequentially update the receive beamformers
        and weight matrices;
        for i = 1,...,K do
            Update U_{q_i} as in (19);
            Update W_{q_i} as in (23);
        end
        At the K MSs, simultaneously update the transmit
        matrices V_{q_i}, ∀i as in (26);
    end
until convergence;
```

Theorem 4. *The alternating minimization process in the proposed WMMSE algorithm results in a monotonic decreasing of the objective function in problem* (22). *For any limit point* $(\mathbf{W}_{q_i}^{\star}, \mathbf{U}_{q_i}^{\star}, \mathbf{V}_{q_i}^{\star})$ *as the minimizer of problem* (22), $\mathbf{V}_{q_i}^{\star}$ *is also a local minimizer of the original problem* (15).

Proof: Please refer to Appendix C. ■

Like the ILA algorithm, the WMMSE algorithm can be implemented either at a central controller or distributively across the BSs and MSs, as being discussed in the following section. The complexity in implementing the WMMSE algorithm shall be addressed in Section V.

B. Distributed Implementation of the Proposed WMMSE Algorithm

Note that the proposed WMMSE algorithm is executed by alternately computing the receive beamformers and the weight matrices at the BSs and computing the transmit beamformers at the MSs. Under the same assumptions stated in Section III-B, the WMMSE algorithm can be implemented in a distributed manner as follows.

At the receiving end, each BS, say BS-q, needs to locally measure its IPN covariance matrix $\mathbf{R}_q$ and estimate the transmit beamformers $\mathbf{V}_{q_i}$ from its connected MSs. Knowing the decoding order, BS-q updates its receive beamformers $\mathbf{U}_{q_i}$'s and the weight matrices $\mathbf{W}_{q_i}$'s with local information as stated in (19) and (23). BS-q then computes the matrix $w_q\sum_{i=1}^{K}\mathbf{U}_{q_i}\mathbf{W}_{q_i}\mathbf{U}_{q_i}^H$ and exchange it to the other coordinated BSs in the network. Subsequently, BS-q passes the updated matrices $w_r\sum_{j=1}^{K}\mathbf{U}_{r_j}\mathbf{W}_{r_j}\mathbf{U}_{r_j}^H$'s as well as the matrix $w_q\sum_{j=1}^{i}\mathbf{U}_{q_j}\mathbf{W}_{q_j}\mathbf{U}_{q_j}^H$ to its ith connected MS. At the transmitting end, MS-i then optimizes its transmit beamforming $\mathbf{V}_{q_i}$ within its power limit as stated in Equation (25) and feeds back the updated $\mathbf{V}_{q_i}$ to BS-q.

Remark 3: In comparing to the sequential update in the ILA algorithm, the proposed WMMSE algorithm allows simultaneous updates across the coordinated BSs and across the MSs. This is due to the fact that the updating steps for the receive beamformers $\mathbf{U}_{q_i}$'s and the weight matrices $\mathbf{W}_{q_i}$'s are decoupled among the BSs, whereas the updating steps for the transmit beamformers $\mathbf{V}_{q_i}$'s are decoupled among

TABLE I
COMPLEXITY OF THE PROPOSED ALGORITHMS

Algorithm	Operation	Message Type	Complexity	Number of Operations	Total Complexity
ILA	$\mathbf{R}_q$ in Eq. (5)	BS-Local	$\mathcal{O}(KQL^3)$	Q	$\mathcal{O}(KQ^2L^3)$
	$\mathbf{B}_q$ in Eq. (14)	BS $\leftrightarrow$ BS BS $\leftrightarrow$ MS	$\mathcal{O}(KL^3)$	Q	$\mathcal{O}(KQL^3)$
	$\mathbf{R}_{q_i}$ in Eq. (13)	BS $\leftrightarrow$ MS	$\mathcal{O}(KL^3)$	K	$\mathcal{O}(K^2L^3)$
	$\mathbf{A}_{q_i}$ in Eq. (10)	MS-Local	$\mathcal{O}(QL^3)$	K	$\mathcal{O}(KQL^3)$
	$\mathbf{X}_{q_i}$ in Eq. (27)	MS-Local	$\mathcal{O}(L^3)$	K	$\mathcal{O}(KL^3)$
ILA - Gauss-Seidel					$\mathcal{O}(KQ^3L^3 + K^2QL^3)$
ILA - Jacobi					$\mathcal{O}(KQ^2L^3 + K^2QL^3)$
WMMSE	$\mathbf{R}_q$ in Eq. (5)	BS-Local	$\mathcal{O}(KQL^3)$	Q	$\mathcal{O}(KQ^2L^3)$
	$\mathbf{U}_{q_i}$ in Eq. (19) sequentially	BS-Local	$\mathcal{O}(L^3)$	KQ	$\mathcal{O}(KQL^3)$
	$\mathbf{W}_{q_i}$ in Eq. (23) sequentially	BS-Local	$\mathcal{O}(L^3)$	KQ	$\mathcal{O}(KQL^3)$
	$w_q \sum_{i=1}^{K} \mathbf{U}_{q_i}\mathbf{W}_{q_i}\mathbf{U}_{q_i}^H$	BS $\leftrightarrow$ BS BS $\leftrightarrow$ MS	$\mathcal{O}(KL^3)$	Q	$\mathcal{O}(KQL^3)$
	$w_q \sum_{j=1}^{i} \mathbf{U}_{q_i}\mathbf{W}_{q_i}\mathbf{U}_{q_i}^H$	BS $\leftrightarrow$ MS	$\mathcal{O}(L^3)$	KQ	$\mathcal{O}(KQL^3)$
	$\mathbf{V}_{q_i}$ in Eq. (26)	MS-Local	$\mathcal{O}(L^3)$	KQ	$\mathcal{O}(KQL^3)$
WMMSE					$\mathcal{O}(KQ^2L^3)$

the MSs. Between the two periods of simultaneous updates at BSs and at MSs, the BSs need to exchange the matrices $w_q \sum_{i=1}^{K} \mathbf{U}_{q_i}\mathbf{W}_{q_i}\mathbf{U}_{q_i}^H$'s.

V. COMPLEXITY OF THE PROPOSED ALGORITHMS

In this section, we analyze the complexity in implementing the proposed ILA and WMMSE algorithms in a multicell system. Similar to the approach used in [9], the complexity of each algorithm is analyzed in each outer-loop iteration. Note that an outer-loop iteration can be also defined as an instance of signaling exchange among the BSs. To simplify the complexity analysis, let us denote $L = \max\{M, N\}$. In Table I, we enumerate the complexity in undertaking the operations (by computing the listed variables) at each iteration as well as the total complexity of each algorithm. In addition, if an algorithm is to be implemented in a distributed manner, Column 3 "Message Type" in Table I classifies whether an operation is taken at the BS, i.e., BS-Local, or at the MS, i.e., MS-Local. The column also classifies whether a variable obtained from its corresponding operation is passed as a signaling message among the BSs, i.e., BS $\leftrightarrow$ BS, or between the BS and its connected MSs, i.e., BS $\leftrightarrow$ MS. We elaborate the content of Table 1 in the following.

In the ILA algorithm, at each iteration, the covariance matrix of the ICI plus noise $\mathbf{R}_q$ must be computed at BS-q. As given in (5), $\mathbf{R}_q$ is approximately the summation of KQ components, where matrix multiplication in each component $\mathbf{H}_{qr_i}\mathbf{X}_{r_i}\mathbf{H}_{qr_i}$ yields the complexity of $\mathcal{O}(L^3)$. Thus, the complexity of computing $\mathbf{R}_q$ is $\mathcal{O}(KQL^3)$. Similarly, the calculation of each pricing matrix $\mathbf{B}_q$ in (14), involves two matrix inversions and a summation of K components $\mathbf{H}_{qq_i}\mathbf{X}_{q_i}\mathbf{H}_{qq_i}$, yields the complexity of $\mathcal{O}(KL^3)$. The same technique can be applied to the calculations of $\mathbf{R}_{q_i}$, $\mathbf{A}_{q_i}$, and $\mathbf{X}_{q_i}$. With the Gauss-Seidel update, only one cell at the time, say cell-q, updates its K matrices $\mathbf{A}_{q_i}, i = 1, \ldots, K$, which yields the complexity of $\mathcal{O}(KQL^3)$. The optimization at cell-q then involves the calculation of the noise matrices $\mathbf{R}_{q_i}$'s and their inverses. Thus, ignoring the few iterations by the bisection step in optimizing the transmit covariance at each MS, this optimization shall take the complexity of $\mathcal{O}(K^2L^3)$. Consequently, in order to sequentially update all the transmit covariances across all Q cells, one has the complexity of $\mathcal{O}(KQ^3L^3 + K^2QL^3)$. On the contrary, with the Jacobi update, all the precoders can be updated simultaneously across the Q cells, after each instance of calculating the IPN matrices $\mathbf{R}_q$'s and the pricing matrices $\mathbf{B}_q$'s. In this case, the complexity of the ILA algorithm with the Jacobi update is $\mathcal{O}(KQ^2L^3 + K^2QL^3)$, which is lower than that with the Gauss-Seidel update by a factor of Q.

Similar to the complexity analysis of the ILA algorithm, the complexity of the WMMSE algorithm can be found to be $\mathcal{O}(KQ^2L^3)$. First, calculating the IPN covariance matrices $\mathbf{R}_q, q = 1, \ldots, Q$, and the signaling messages $w_q \sum_{i=1}^{K} \mathbf{U}_{q_i}\mathbf{W}_{q_i}\mathbf{U}_{q_i}^H, q = 1, \ldots, Q$, yields the overall complexity of $\mathcal{O}(KQ^2L^3)$. Second, utilizing the calculated IPN matrix $\mathbf{R}_q$ at BS-q, the receive beamformers $\mathbf{U}_{q_i}$'s and weight matrices $\mathbf{U}_{q_i}$'s can be updated sequentially from MS-1 (first) to MS-K (last), which yields the overall complexity of $\mathcal{O}(KQL^3)$. Third, the complexity of updating KQ transmit beamformers $\mathbf{V}_{q_i}$'s is $\mathcal{O}(KQL^3)$. Thus, per outer-loop iteration, in order to update the receive beamformers, the weight matrices, and the transmit beamformers for Q cells in the network, the complexity of the WMMSE algorithm is $\mathcal{O}(KQ^2L^3)$, which is comparable to that of the ILA algorithm with the Jacobi update.

Remark 4: In term of total computational complexity, the distributed implementation of each algorithm is roughly the same as the its centralized implementation. In term of implementation, the distributed approach requires certain message passing among the coordinated BSs and MSs, as detailed in Table I. In addition, Table I also quantifies the number of messages that need to be exchanged at each outer-loop iteration. In the distributed implementation structure, the computation load in the optimization process can be shared across the coordinated BSs and MSs.

VI. SIMULATION RESULTS

This section presents simulation results on the achievable sum-rate of a multicell system in the uplink transmission

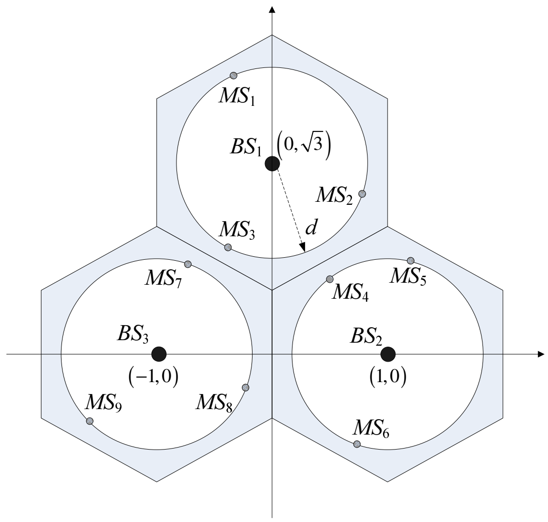

Fig. 1. A multiuser multicell system with 3 cells, 3 MSs per cell. Each MS is randomly located at a distance d from its connected BS.

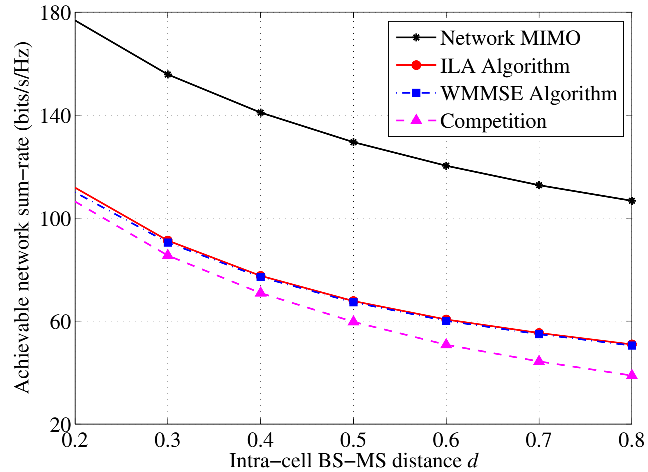

Fig. 2. Network sum-rates under the considered operating modes.

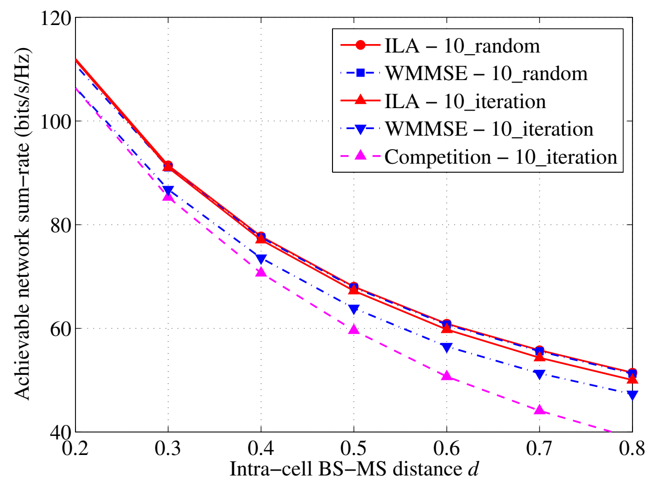

Fig. 3. Network sum-rates under the *interference coordination* mode, obtained from the ILA and WMMSE algorithms with 10 random starting points or with 10 outer-loop iterations.

under various levels of coordination and on the convergence behaviors of the proposed ILA and WMMSE algorithms. We compare the results for three operating modes: (i) the *interference coordination* mode obtained from the proposed ILA and WMMSE algorithms (with equal weight for each cell), (ii) the *competitive* mode where each cell selfishly maximizes the sum-rate for its connected MS only, and (iii) the *network MIMO* mode where the whole system is a single large MIMO MAC channel. Considered is a 3-cell system, where the distance between any two BSs is normalized to 2, as illustrated in Fig. 1. The number of MSs is set to 3 per cell, unless stated otherwise, and each MS is randomly located on a circle at distance d from its connected BS. The BS and MS are equipped with 4 and 2 antennas, respectively. The transmit power of each MS is limited to 1 W. The intra-cell and inter-cell channel coefficients are generated as products of two components: the first components account for the large-scale path loss with a path loss exponent of 3, and the second components represent the small-scale fading using i.i.d. complex Gaussian random variables with zero mean and unity variance. The AWGN power spectral density σ^2 is set at 0.01 W/Hz.

We first investigate the achievable network sum-rates versus the intra-cell MS-BS distance d of the various algorithms, which are run until fully converged. As d is varied, $10,000$ channel realizations at each value of d are used to obtain the average network sum-rates plotted in Fig. 2. As shown in the figure, when the distance d becomes smaller, the network sum-rate increases in all 3 operating modes. This is due to the increase in strength of intra-cell channels and the reduction in the strength of inter-cell channels. Out of the 3 operating modes, *network MIMO* obtains the largest sum-rate, as it is the upper bound for any uplink multicell transmission scheme. It is also observed that by implementing the *interference coordination* among the cells using the proposed algorithms (ILA and WMMSE), one can improve the network sum-rate by 5 to 15 b/s/Hz over the *competitive* mode, especially at high ICI region (large d). Note that the performances of the *interference coordination* mode are obtained from the same initialization with $\mathbf{V}_{q_i}\mathbf{V}_{q_i}^H = \mathbf{X}_{q_i} = (P_{q_i}/N)\mathbf{I}$ for both ILA and WMMSE algorithms. Fig. 2 shows that there is literally no difference in the performances obtained from the ILA and WMMSE algorithms with the same *identity matrix* starting point. In other words, they both converge to the same local maximum most of the time.

As the proposed ILA and WMMSE algorithms do not guarantee a globally optimal performance, it is interesting to investigate the effects of the starting point on their achieved network sum-rate. For this, we run simulations for 10 different randomly generated starting points and record the best sum-rate result out of the 10 fully converged maxima for each algorithm. The resulting plots for the 2 proposed algorithms designated by *ILA-10_random*, and *WMMSE-10_random*, in Fig. 3 show a negligible performance difference between the two proposed algorithms, and a slightly increased network sum-rate as compared to the case with *identity matrix* starting point in Fig. 2. This close performance indicates that the

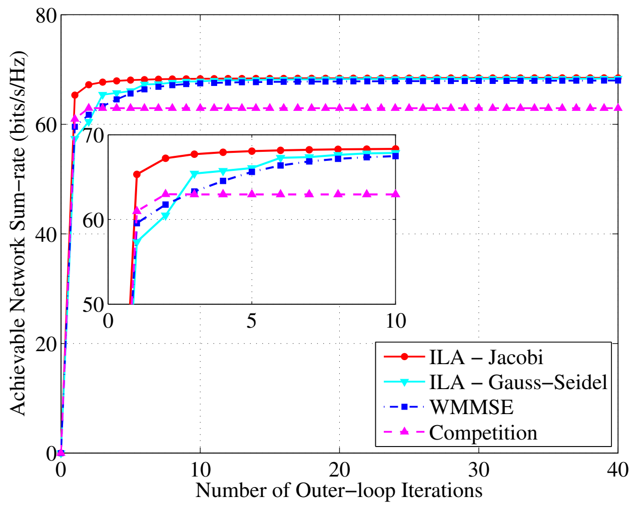

Fig. 4. Convergence of the proposed ILA and WMMSE algorithms to maximize the network sum-rate with the *interference coordination*.

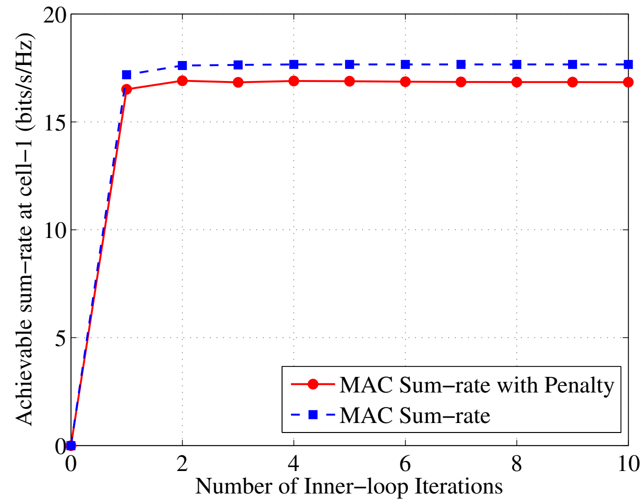

Fig. 5. Convergence of the proposed iterative algorithm to solve Problem (11).

identity matrix is a good and simple choice for the starting point. As previously discussed, both the proposed ILA and WMMSE algorithms enjoy the monotonic convergence, but require the coordinated cells to exchange signaling information after each outer-loop iteration. For practical implementation, it may be desirable to limit the number of iterations in order to reduce the amount of signaling exchange and it is interesting to understand the effect of number of iterations on their performance. For illustration, we include in Fig. 3 the plots *ILA-10_iteration*, *WMMSE-10_iteration*, and *Competition-10_iteration* representing the achieved network sum-rates after 10 outer-loop iterations of the ILA, WMMSE algorithms and competitive mode, respectively. The simulation results in Fig. 3 indicate that the ILA algorithm is faster to achieve its converged performance than the WMMSE algorithm in terms of the number of outer-loop iterations. Both algorithms outperform the *competitive* mode after just 10 iterations. Note that the ILA algorithm in this simulation is implemented with the Jacobi update. Thus, it requires a similar amount of inter-BS signaling as the WMMSE algorithm.

To investigate further their convergence behavior, we obtain simulation results for a specific channel realization with $d = 0.5$ and plot the network sum-rates achieved after each outer-loop iteration by ILA using Jacobi and Gauss-Seidel updates, WMMSE, and competition (for comparison). As observed in Fig. 4, the *interference coordination* mode, obtained by the ILA and WMMSE algorithms, offers higher network sum-rate than the *competitive* mode. It is also observed that the ILA algorithm does monotonically converge with both Jacobi and Gauss-Seidel updates. The network sum-rate is also improved monotonically by the WMMSE algorithm. These convergence behaviors of the ILA and WMMSE algorithms agree with our analysis. Interestingly, Fig. 4 confirms that the ILA algorithm has a faster convergence than the WMMSE algorithm in terms of number of outer-loop iterations. This observation explains why the ILA algorithm obtains better sum-rate performance than the WMMSE algorithm after 10 outer-loop iterations as shown in Fig. 3.

As previously discussed, WMMSE algorithm has only outer-loop iterations while ILA has both inner-loop and outer-loop iterations. To investigate its inner-loop convergence, in Fig. 5, we plot the sum-rate achieved at cell-1 after each number of inner-loop iterations by the ILA, for the same channel realization used in Fig. 4, and at outer-loop iteration #2 (when $\mathbf{A}_{q_i}$'s are now non-zero). Note that the inner-loop iterations in the ILA algorithm involve the proposed algorithm in Theorem 1 to maximize the MAC sum-rate with penalty terms (11). For comparison, the convergence of inner-loop iterations in the *competitive* mode, i.e., the iterative water-filling algorithm for MAC sum-rate maximization [16], is also illustrated. Fig. 5 indicates that, at the outer-loop iteration #2, due to the penalty terms, the ILA inner-loop algorithm achieves lower sum-rate in cell-1 than the competitive mode (without the penalty terms). However, as the cell adopts a more cooperative strategy by limiting the ICI to other cells, the overall network sum-rate performance is indeed improved, as shown in Fig. 4.

Finally, Fig. 6 compares the CPU running time of the proposed ILA and WMMSE algorithms versus the number of MSs at each cell under the same termination criterion. Although the CPU running time is rather relative, all algorithms are programmed with the same code implementation and run on the same platform, which allows us to realize the relative trend in computational complexity of each algorithm. As shown in the figure, the ILA algorithm with the Jacobi update shall improve the running time over that with the Gauss-Seidel update roughly by a factor of $Q = 3$. Interestingly, while the WMMSE requires more outer-loop iterations, i.e., more inter-BS signaling exchange, it has less running time than the ILA algorithm. Intuitively, the WMMSE does not require the inner-loop iterations, while the ILA algorithm does for sequentially optimizing precoders at a particular cell. Thus, the WMMSE algorithm imposes less intra-cell BS-MS operation and signaling than the ILA algorithm. In fact, the running time of the WMMSE algorithm is almost linear with K, as our complexity analysis indicates.

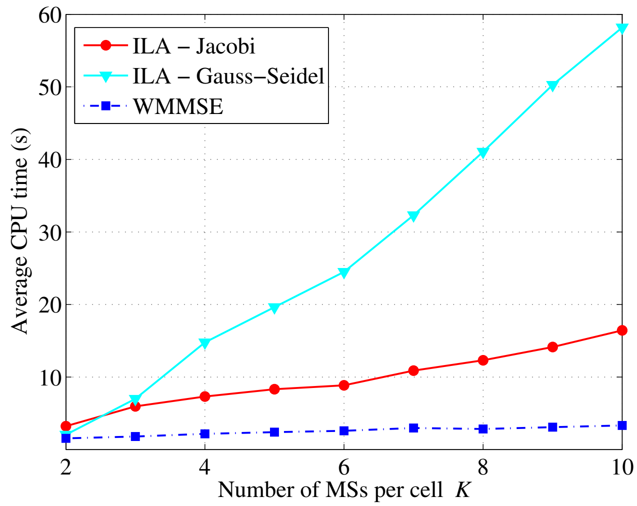

Fig. 6. Average CPU time versus the number of MSs per cell.

VII. Conclusion

This paper examined the problem of WSR maximization in the multicell MIMO MAC. Under the *interference coordination* mode among the multiple cells, the network WSR maximization problem was shown to be nonconvex. The paper then proposed two solution approaches, namely ILA and WMMSE, to approximate and transform the original nonconvex problem into convex optimization ones. In the ILA approach, the nonconvex optimization problem is successively approximated and decomposed into a set of convex problems, which can be solved distributively at each MS. In the WMMSE algorithm, by transforming the original problem into a weighted MSE minimization problem, the network WSR is maximized by alternatively optimizing the weight matrices and MSE decoders at each BS and the precoder at each MS. Simulations confirmed the convergence analysis of the proposed algorithm and showed a significant enhancement in the network sum-rate as compared to competitive design.

Appendix A
Proof to Theorem 1

The proof for this theorem follows the similar approach used in [16] for the sum-rate maximization in the MAC without the penalty components $\sum_{i=1}^{K} \mathrm{Tr}\{\mathbf{A}_{q_i}\mathbf{X}_{q_i}\}$. Before proceeding to the main part of the proof, we briefly revisit the solution of problem (12), which was previously given in [8].

Given μ_{q_i} as the Lagrangian multiplier associated with the power constraint $\mathrm{Tr}\{\mathbf{X}_{q_i}\} \le P_{q_i}$, it was shown in [8] that the optimal solution $\mathbf{X}^\star_{q_i}$ must be in the form of

$$\mathbf{X}^\star_{q_i} = \mathbf{G}_{q_i}\mathbf{P}_{q_i}\mathbf{G}^H_{q_i}, \tag{27}$$

where $\mathbf{G}_{q_i}$ is the (normalized) generalized eigen-matrix of the pair of matrices of $\mathbf{H}^H_{qq_i}\mathbf{R}^{-1}_{q_i}\mathbf{H}_{qq_i}$ and $(\mathbf{A}_{q_i} + \mu_{q_i}\mathbf{I})$. The matrix $\mathbf{P}_{q_i}$ is a nonnegative diagonal matrix, obtained from the following water-filling solution

$$\mathbf{P}_{q_i} = \left[w_q {\boldsymbol{\Sigma}^{(2)}_{q_i}}^{-1} - {\boldsymbol{\Sigma}^{(1)}_{q_i}}^{-1}\right]^+, \tag{28}$$

where $\boldsymbol{\Sigma}^{(1)}_{q_i}$ and $\boldsymbol{\Sigma}^{(2)}_{q_i}$ are diagonal matrices given by

$$\begin{aligned}\boldsymbol{\Sigma}^{(1)}_{q_i} &= \mathbf{G}^H_{q_i}\mathbf{H}^H_{qq_i}\mathbf{R}^{-1}_{q_i}\mathbf{H}_{qq_i}\mathbf{G}_{q_i}\\ \boldsymbol{\Sigma}^{(2)}_{q_i} &= \mathbf{G}^H_{q_i}(\mathbf{A}_{q_i} + \mu_{q_i}\mathbf{I})\mathbf{G}_{q_i}.\end{aligned}$$

In this solution, the dual variable μ_{q_i}, behaving as the water-level in the water-filling process, is adjusted to enforce the power constraint. Utilizing this optimal solution, we now show that the optimal solution to the multiuser problem (11) is indeed a collection of solutions to individual single-user problems (12).

If part: Let $\{\bar{\mathbf{X}}_{q_i}\}_{i=1}^{K}$ be the optimal solution of the original K-user problem (11). Suppose that $\bar{\mathbf{X}}_{q_i}$ is not the optimal solution of the corresponding problem (12) while treating $\mathbf{R}_{q_i} = \mathbf{R}_q + \sum_{j\neq i}\mathbf{H}_{qq_j}\bar{\mathbf{X}}_{q_j}\mathbf{H}^H_{qq_j}$ as noise. Then fixing all other covariance matrices $\bar{\mathbf{X}}_{q_j}, \forall j \neq i$, solving problem (12) obtains the optimal solution $\mathbf{X}^\star_{q_i}$. Clearly, $\mathbf{X}^\star_{q_i}$ strictly increases the objective function of the original problem (11). Thus, this contradicts with assumption on the optimality of $\{\bar{\mathbf{X}}_{q_i}\}_{i=1}^{K}$.

Only if part: Consider the partial Lagrangian of problem (11)

$$\begin{aligned}\mathcal{L}(\mathbf{X}_{q_i}, \boldsymbol{\mu}_q) = &\sum_{i=1}^{K}\mu_{q_i}P_{q_i} + w_q \log\left|\mathbf{R}_q + \sum_{i=1}^{K}\mathbf{H}_{qq_i}\mathbf{X}_{q_i}\mathbf{H}^H_{qq_i}\right|\\ &-\sum_{i=1}^{K}\mathrm{Tr}\left\{(\mathbf{A}_{q_i} + \mu_{q_i}\mathbf{I})\mathbf{X}_{q_i}\right\},\end{aligned} \tag{29}$$

where $\boldsymbol{\mu}_q = [\mu_{q_1}, \ldots, \mu_{q_K}]^T$ are the Lagrangian dual variables associated with the power constraints. For optimality, the solution of problem (11) must satisfy the following Karush-Kuhn-Tucker (KKT) conditions

$$\begin{aligned}w_q\mathbf{H}^H_{qq_i}\left(\mathbf{R}_q + \sum_{i=1}^{K}\mathbf{H}_{qq_i}\mathbf{X}_{q_i}\mathbf{H}^H_{qq_i}\right)^{-1}\mathbf{H}_{qq_i} &= \mathbf{A}_{q_i} + \mu_{q_i}\mathbf{I}, \forall i\\ \mathrm{Tr}\left\{\mathbf{X}_{q_i}\right\} &= P_{q_i}, \forall i\\ \mu_{q_i} &\ge 0, \forall i.\end{aligned}$$

For the case of $K = 1$, it is straightforward to verify that the optimal solution of problem (12), $\mathbf{X}^\star_{q_i} = \mathbf{G}_{q_i}\mathbf{P}_{q_i}\mathbf{G}^H_{q_i}$, with $\mathbf{P}_{q_i}$ given in (28), satisfies the above KKT conditions. However, the KKT conditions for the single-user case are different from that for the multiuser case by the additional noise term $\sum_{j\neq i}\mathbf{H}_{qq_j}\mathbf{X}_{q_i}\mathbf{H}^H_{qq_j}$. Thus, if each $\mathbf{X}^\star_{q_i}$ satisfies the single-user condition while treating the signals of other MSs as noise, then collectively, the set of $\{\mathbf{X}^\star_{q_i}\}_{i=1}^{K}$ must satisfy the above KKT conditions for the multiuser case. Then, $\{\mathbf{X}^\star_{q_i}\}_{i=1}^{K}$ must be the optimal solution to the original problem (11).

Appendix B
Proof to Theorem 2

Similar to the approach in [7], [8], the proof for this theorem is established by showing that the network sum-rate is strictly nondecreasing after an update at any given cell. Suppose that $\mathbf{X}_q = \bar{\mathbf{X}}_q = \left\{\bar{\mathbf{X}}_{q_i}\right\}_{i=1}^{K}, \forall q$ from the previous outer-loop iteration, and $\mathbf{X}^\star_q = \left\{\mathbf{X}^\star_{q_i}\right\}_{i=1}^{K}$ as the optimal solution obtained at cell-q after the current outer-loop iteration.

Similar to the technique applied in [6], [8], it can be easily shown that $f_q(\mathbf{X}_q, \mathbf{X}_{-q})$ is a convex function with respect to $\mathbf{X}_{q_i} \in \mathcal{S}_{q_i} \triangleq \{\mathbf{X}_{q_i} | \mathbf{X}_{q_i} \succeq \mathbf{0}, \mathrm{Tr}\{\mathbf{X}_{q_i}\} \leq P_{q_i}\}$. Thus, by the first-order condition for the convex function f_q [18], one has

$$f_q(\mathbf{X}_q^\star, \bar{\mathbf{X}}_{-q}) \geq f_q(\bar{\mathbf{X}}_q, \bar{\mathbf{X}}_{-q}) - \sum_{i=1}^{K} \mathrm{Tr}\{\mathbf{A}_{q_i}(\mathbf{X}_{q_i}^\star - \bar{\mathbf{X}}_{q_i})\} \tag{30}$$

with $\mathbf{A}_{q_i}$ being defined in (10) at $\bar{\mathbf{X}}_{q_i}$.

After one Gauss-Seidel iteration, the network weighted sum-rate is updated such that

$$\begin{aligned}
&\sum_{q=1}^{Q} w_q R_q(\mathbf{X}_q^\star, \bar{\mathbf{X}}_{-q}) \\
&= w_q R_q(\mathbf{X}_q^\star, \bar{\mathbf{X}}_{-q}) + f_q(\mathbf{X}_q^\star, \bar{\mathbf{X}}_{-q}) \\
&\geq w_q R_q(\mathbf{X}_q^\star, \bar{\mathbf{X}}_{-q}) + f_q(\bar{\mathbf{X}}_q, \bar{\mathbf{X}}_{-q}) - \sum_{i=1}^{K} \mathrm{Tr}\{\mathbf{A}_{q_i}(\mathbf{X}_{q_i}^\star - \bar{\mathbf{X}}_{q_i})\} \\
&\geq w_q R_q(\bar{\mathbf{X}}_q, \bar{\mathbf{X}}_{-q}) + f_q(\bar{\mathbf{X}}_q, \bar{\mathbf{X}}_{-q}) - \sum_{i=1}^{K} \mathrm{Tr}\{\mathbf{A}_{q_i}(\bar{\mathbf{X}}_{q_i} - \bar{\mathbf{X}}_{q_i})\} \\
&= \sum_{q=1}^{Q} w_q R_q(\bar{\mathbf{X}}_q, \bar{\mathbf{X}}_{-q}),
\end{aligned}$$

where the first inequality is due to the one in (30), and the second inequality is due to $\mathbf{X}_q^\star$ as the optimal solution of problem (11). Since the network sum-rate is upper-bounded and nondecreasing after each update, the sequential optimization (11) generates a Cauchy sequence that must converge to one of the local maxima.

Appendix C
Proof to Theorem 4

Denote

$$f(\{\mathbf{V}_{q_i}\}) = \sum_{q=1}^{Q} w_q \sum_{i=1}^{K} \log \left| \mathbf{E}_{q_i}^{\mathrm{MMSE}} \right| \tag{31}$$

and

$$\begin{aligned}
&g(\{\mathbf{U}_{q_i}\}, \{\mathbf{W}_{q_i}\}, \{\mathbf{V}_{q_i}\}) \\
&\quad = \sum_{q=1}^{Q} w_q \sum_{i=1}^{K} \left[\mathrm{Tr}\left\{ \mathbf{W}_{q_i} \mathbf{E}_{q_i} \right\} - \log |\mathbf{W}_{q_i}| \right]
\end{aligned} \tag{32}$$

as the cost functions of the original problem (15) and the restated WMMSE problem (22).

Since the constraint set of problem (22) is decoupled for the variables $\mathbf{U}_{q_i}, \mathbf{W}_{q_i}, \mathbf{V}_{q_i}$, applying the block coordinate descent method by alternative minimizing over $\mathbf{U}_{q_i}, \mathbf{W}_{q_i}, \mathbf{V}_{q_i}$ must decrease its cost function monotonically [19]. In addition, the power constraint on $\mathbf{V}_{q_i}$ is upper-bounded, the cost function (32) is lower-bounded. Thus, the proposed WMMSE must monotonically converge to at least a local minimum of the cost function (32). Note that the cost function of the original sum-rate maximization problem (15) does not necessarily improve after each iteration. However, given $(\mathbf{U}_{q_i}^\star, \mathbf{W}_{q_i}^\star, \mathbf{V}_{q_i}^\star)$ as a local minimizer of problem (22) obtained from the WMMSE algorithm, it can be proved that $\mathbf{V}_{q_i}^\star$ is also a local minimizer of the original problem (15) as follows.

First, we need to show that the gradients of $f(\cdot)$ and $g(\cdot)$ with respect to $\mathbf{V}_{q_i}$ are the same at $\mathbf{V}_{q_i}^\star$. Similar to the approach in [9], [13], evaluating the gradients of $f(\cdot)$ and $g(\cdot)$ at the (m, n)-element of $\mathbf{V}_{q_i}$, one has

$$\begin{aligned}
&\left. \frac{\partial f(\{\mathbf{V}_{q_i}\})}{\partial [\mathbf{V}_{q_i}]_{m,n}} \right|_{\mathbf{V}_{q_i} = \mathbf{V}_{q_i}^\star} \\
&= \sum_{r=1}^{Q} w_r \sum_{j=1}^{K} \left. \frac{\partial \log \left| \mathbf{E}_{r_j}^{\mathrm{MMSE}}(\mathbf{V}_{q_i}) \right|}{\partial [\mathbf{V}_{q_i}]_{m,n}} \right|_{\mathbf{V}_{q_i} = \mathbf{V}_{q_i}^\star} \\
&= \sum_{r=1}^{Q} w_r \sum_{j=1}^{K} \mathrm{Tr}\Bigg\{ \left(\mathbf{E}_{r_j}^{\mathrm{MMSE}}(\mathbf{V}_{q_i}^\star) \right)^{-1} \\
&\qquad\qquad \times \left. \frac{\partial \mathbf{E}_{r_j}^{\mathrm{MMSE}}(\mathbf{V}_{q_i})}{\partial [\mathbf{V}_{q_i}]_{m,n}} \right|_{\mathbf{V}_{q_i} = \mathbf{V}_{q_i}^\star} \Bigg\},
\end{aligned} \tag{33}$$

and

$$\begin{aligned}
&\left. \frac{\partial g(\{\mathbf{U}_{q_i}\}, \{\mathbf{W}_{q_i}\}, \{\mathbf{V}_{q_i}\})}{\partial [\mathbf{V}_{q_i}]_{m,n}} \right|_{\mathbf{V}_{q_i} = \mathbf{V}_{q_i}^\star, \mathbf{U}_{q_i} = \mathbf{U}_{q_i}^\star} \\
&= \sum_{r=1}^{Q} w_r \sum_{j=1}^{K} \mathbf{W}_{r_j} \left. \frac{\partial \mathbf{E}_{r_j}(\mathbf{U}_{q_i}, \mathbf{V}_{q_i})}{\partial [\mathbf{V}_{q_i}]_{m,n}} \right|_{\mathbf{V}_{q_i} = \mathbf{V}_{q_i}^\star, \mathbf{U}_{q_i} = \mathbf{U}_{q_i}^\star} \\
&= \sum_{r=1}^{Q} w_r \sum_{j=1}^{K} \mathbf{W}_{r_j}^\star \left. \frac{\partial \mathbf{E}_{r_j}(\mathbf{U}^\star{}_{q_i}, \mathbf{V}_{q_i})}{\partial [\mathbf{V}_{q_i}]_{m,n}} \right|_{\mathbf{V}_{q_i} = \mathbf{V}_{q_i}^\star}.
\end{aligned} \tag{34}$$

Since $\mathbf{W}_{r_j}^\star = \left(\mathbf{E}_{r_j}^{\mathrm{MMSE}}(\mathbf{V}_{q_i}^\star) \right)^{-1} = \left(\mathbf{E}_{r_j}(\mathbf{U}_{q_i}^\star, \mathbf{V}_{q_i}^\star) \right)^{-1}$, which yields the equivalence between (33) and (34). In addition, because $(\mathbf{U}_{q_i}^\star, \mathbf{W}_{q_i}^\star, \mathbf{V}_{q_i}^\star)$ is a local minimizer of problem (22), it must satisfy the stationarity condition:

$$\mathrm{Tr}\left\{ \nabla_{\mathbf{V}_{q_i}} g(\mathbf{U}_{q_i}^\star, \mathbf{W}_{q_i}^\star, \mathbf{V}_{q_i}^\star)^H (\mathbf{V}_{q_i} - \mathbf{V}_{q_i}^\star \right\} \leq 0, \forall \mathbf{V}_{q_i}. \tag{35}$$

Conversely, due to the equivalence between (33) and (34), $\mathbf{V}_{q_i}^\star$ also satisfies the stationarity condition:

$$\mathrm{Tr}\left\{ \nabla_{\mathbf{V}_{q_i}} f(\mathbf{V}_{q_i}^\star)^H (\mathbf{V}_{q_i} - \mathbf{V}_{q_i}^\star \right\} \leq 0, \forall \mathbf{V}_{q_i}. \tag{36}$$

Thus, $\mathbf{V}_{q_i}$ must be also a local minimizer to the original problem (15).

Acknowledgment

The authors would like to thank the anonymous reviewers for their insightful and constructive comments, which helped to improve the presentation of this paper.

References

[1] 4G Americas, 4G Mobile Broadband Evolution: 3GPP Release 10 and Beyond, Feb. 2011.

[2] M. Karakayali, G. Foschini, and R. Valenzuela, "Network coordination for spectrally efficient communications in cellular systems," *IEEE Wireless Commun.*, vol. 13, no. 4, pp. 56–61, Aug. 2006.

[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, no. 9, pp. 1380–1408, Dec. 2010.

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

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

[6] S. Ye and R. S. Blum, "Optimized signaling for MIMO interference systems with feedback," *IEEE Trans. Signal Process.*, vol. 51, no. 11, pp. 2839–2848, Nov. 2003.
[7] C. Shi, R. A. Berry, and M. L. Honig, "Monotonic convergence of distributed interference pricing in wireless networks," in *Proc. 2009 IEEE Int. Symp. Inf. Theory*, pp. 1619–1623.
[8] S.-J. Kim and G. B. Giannakis, "Optimal resource allocation for MIMO ad hoc cognitive radio networks," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 3117–3131, May 2011.
[9] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He, "An iteratively weighted MMSE approach to distributed sum-utility maximization for MIMO interfering broadcast channel," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4331–4340, Sept. 2011.
[10] L. An, "D.C. programming for solving a class of global optimization problems via reformulation by exact penalty," vol. 2861, pp. 87–101, 2003.
[11] L. An and P. Tao, "The DC (difference of convex functions) programming and DCA revisited with DC models of real world nonconvex optimization problems," *Annals Operations Research*, vol. 133, pp. 23–46, 2005.
[12] R. Horst and N. V. Thoai, "DC programming: overview," *J. Optim. Theory Appl.*, vol. 103, no. 1, pp. 1–43, Oct. 1999.
[13] S. S. Christensen, R. Argawal, E. de Carvalho, and J. M. Cioffi, "Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 4792-4799, Dec. 2008.
[14] D. Schmidt, C. Shi, R. A. Berry, M. L. Honig, and W. Honig, "Minimum mean square error interference alignment," in *2009 Assilomar Conf. Signals, Syst. Comput.*
[15] A. Goldsmith, *Wireless Communications*. Cambridge University Press, 2004.
[16] W. Yu, W. Rhee, S. Boyd, and J. M. Cioffi, "Iterative water-filing for Gaussian multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 1, pp. 145–152, Jan. 2004.
[17] G. Scutari, S. Barbarossa, and D. P. Palomar, "Competitive design of multiuser MIMO system based on game theory: a unified view," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 9, pp. 1089–1102, Sept. 2008.
[18] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
[19] D. P. Bertsekas, *Nonlinear Programming*. Athena Scientific, 1995.

Duy H. N. Nguyen (S'07) received the B.Eng. degree (with First Class Honours) from Swinburne University of Technology, Hawthorn, Victoria, Australia, in 2005 and the M.Sc. degree from University of Saskatchewan, Saskatoon, SK, Canada, in 2009, all in electrical engineering. Since September 2009, he has been with Department of Electrical and Computer Engineering, McGill University, Montreal, QC, Canada, where he is currently working toward the Ph.D. degree.

His research interests include resource allocation in wireless networks, signal processing for communications, convex optimization, and game theory.

Mr. Nguyen was a recipient of the Australian Development Scholarship (ADS) for his undergraduate study in Australia, a recipient of the McGill Engineering Doctoral Award, and a recipient of the Fonds de recherche du Québec - Nature et technologies (FRQNT) Doctoral Award.

Tho Le-Ngoc (F'97) obtained his B.Eng. (with Distinction) in Electrical Engineering in 1976, his M.Eng. in 1978 from McGill University, Montreal, and his Ph.D. in Digital Communications in 1983 from the University of Ottawa, Canada. During 1977-1982, he was with Spar Aerospace Limited and involved in the development and design of satellite communications systems. During 1982-1985, he was an Engineering Manager of the Radio Group in the Department of Development Engineering of SRTelecom Inc., where he developed the new point-to-multipoint DA-TDMA/TDM Subscriber Radio System SR500. During 1985-2000, he was a Professor at the Department of Electrical and Computer Engineering of Concordia University. Since 2000, he has been with the Department of Electrical and Computer Engineering of McGill University. His research interest is in the area of broadband digital communications. He is a Fellow of the Institute of Electrical and Electronics Engineers (IEEE), the Engineering Institute of Canada (EIC), the Canadian Academy of Engineering (CAE) and the Royal Society of Canada (RSC). He is the recipient of the 2004 Canadian Award in Telecommunications Research, and recipient of the IEEE Canada Fessenden Award 2005. He holds a Canada Research Chair (Tier I) on Broadband Access Communications, and a Bell Canada/NSERC Industrial Research Chair on Performance & Resource Management in Broadband xDSL Access Networks.