

# Decentralized Linear Transceiver Design in Multicell MIMO Broadcast Channels

Khaled Ardah, Yuri C. B. Silva and Francisco R. P. Cavalcanti

**Abstract**—This paper studies linear transceiver design in multicell MIMO Broadcast Channels (BCs). In this context, previous works have tried to enhance the conventional Block Diagonalization (cBD) algorithm, such as the proposal of iterative BD (iBD), which has less dimensionality restrictions and accounts for the presence of inter-cell interference (ICI). However, both approaches become interference-limited when the ICI has strong power. In this paper, we take a different direction by using the Weighted Sum-Rate (WSR) as the transceiver design criterion. For that, three different novel algorithms are proposed in this paper, which are based on the alternating optimization technique and guaranteed to converge to a local WSR-optimum. The first algorithm is an interference pricing approach, where each cell maximizes its own utility, which is formed by the local users' WSR minus the priced ICI leakage. The second algorithm designs transceivers that maximize the network-wide WSR. Interestingly, we prove that the WSR maximization via interference pricing can be made equivalent to the network-wide WSR maximization, whenever the mobile stations are equipped with single-antennas. The third algorithm is an implicit interference pricing approach, where each cell self-prices its ICI leakage and, thus, does not require the feedback of variables from other cells. To facilitate the algorithms' implementation, a novel Over-the-Air (OTA) signaling scheme based on Time Division Duplex (TDD) mode is proposed, which reduces the signaling overhead and requires no backhaul feedback, as compared to existing schemes.

**Index Terms**—MIMO systems, block diagonalization, weighted sum rate maximization, interference-pricing.

## I. INTRODUCTION

Multiple-Input Multiple-Output (MIMO) technology has great potential to eliminate/manage interference, achieve higher throughput, and enhance system capacity [1]. Using multiple antennas, the Base Stations (BSs) can transmit to multiple users simultaneously using linear or non-linear transmission techniques [2] to achieve a linear increase of system throughput in the number of BS antennas. In single-cell networks, the non-linear Dirty Paper Coding (DPC) technique [3] is known to achieve the channel capacity. However, it is widely considered that DPC has limited practical applications, due to its high complexity. Therefore, linear transmission techniques (also called beamforming) have gained more interest and were proven to achieve the same sum rate scaling law as DPC [2], while maintaining low complexity.

A notable scheme in this area is called Block Diagonalization (BD) [4]. In single-cell networks, conventional Block Diagonalization (cBD) completely eliminates intra-cell interference by forcing each user to transmit on the null space of

the other users. However, in multicell networks, cBD would ignore the inter-cell interference (ICI), which would affect the users' performance. For that purpose, the authors in [5] have proposed enhanced BD (eBD), which uses a whitening filter to reduce ICI effects. Nevertheless, both cBD and eBD algorithms have high dimensionality restrictions, since they both rely on transmit beamforming to eliminate intra-cell interference and ignore the receive beamforming. Motivated by the last observation, we have proposed iterative BD (iBD) in [6], which eliminates the intra-cell interference by jointly optimizing the transmit and receive beamforming matrices and also accounts for the ICI presence. We have shown that iBD has better sum rate performance than both cBD and eBD, while significantly reducing dimensionality restrictions.

However, it is also shown in [6] that all BD approaches become interference-limited in the presence of high ICI power. The main limitation of BD is that each user has an *altruistic* behavior with regard to other users in the same cell (since the intra-cell interference is completely eliminated) and an *egoistic* behavior with regard to users in adjacent cells (since nothing is done to reduce the ICI). Thus, the BD approach cannot achieve a good balance with regard to the users' beamforming behavior, which prevents it from achieving the optimum in terms of sum rate [7].

An alternative approach is to jointly design the transmit beamforming of all users in all cells. This approach is named as Coordinated Multi-Point (CoMP) in the literature and can be classified into Joint Processing (JP) [8] and Coordinated Beamforming (CBF) techniques [9]. In contrast to JP, each user in a CBF system is served by a single BS and thus, the BSs do not need to share the users' data or to be time and phase synchronized. Therefore, CBF has gained a lot of attention and has been extensively studied in the literature with different optimization criteria. For example, in [9]–[11] for sum-power minimization, in [12] for Signal-to-Interference-plus-Noise Ratio (SINR) balancing, in [13] for sum Mean-Square Error (MSE) minimization, and in [14]–[25] for Weighted Sum-Rate (WSR) maximization.

In this paper, we are particularly interested in the WSR maximization problem. The problem is non-convex and NP-hard [20], for which only local optima can be guaranteed via practical methods. Nevertheless, it has some desirable properties such as 1) it can prioritize the users and achieve some fairness among them by adjusting the weights, 2) it has an implicit users and streams selection, since, at convergence, the number of active streams is almost always less than or equal to the number of BS antennas, and 3) it is always feasible when only constrained by transmit power. Therefore, the WSR maximization problem has gained a lot of attention and has

The authors are with Federal University of Ceará, Wireless Telecommunications Research Group (GTEL), Fortaleza, Brazil, emails {khaledardah, yuri, rodrigo}@gtel.ufc.br.

been extensively studied in the literature, considering different optimization tools and system models.

For single-cell MIMO Broadcast Channels (BCs), the authors in [14] reformulate the problem into an equivalent problem that incorporates a weighted sum-MSE and establish a weighted sum-MSE duality that is solved iteratively using a Geometric Program (GP) formulation. For multicell Multiple-Input Single-Output (MISO) BC, the problem was addressed in [16], [18], [19]. In [16], the authors derived the Karush-Kuhn-Tucker (KKT) conditions of the problem and then devised an iterative algorithm to solve them, without the need of resorting to convex optimization methods. In [18], an iterative pricing algorithm was proposed based on game-theory, which is guaranteed to converge to an interference equilibrium that corresponds to a KKT point for the original WSR maximization problem. Among all, the global optimum solution is guaranteed only in [19], where the problem was solved using a branch-reduce-and-bound algorithm.

On the other hand, the authors in [17] considered the multicell MISO-Interference Channel (IC) and established a relation between the WSR and the Virtual Signal-to-Interference-plus-Noise Ratio (VSINR) by applying the KKT conditions, which led to a distributed and iterative algorithm. Recently, the authors in [20] considered a single-cell MIMO-BC and established a relation between the WSR maximization problem and the Weighted Minimum-Mean-Square Error (WMMSE) minimization problem by applying the KKT conditions. As a result, an iterative algorithm called WSR-WMMSE was proposed, which is based on the alternating optimization technique [26] and solves the quite hard WSR maximization problem indirectly by solving the easier WMMSE minimization problem. This later relation has inspired many extensions, such as to the multicell MIMO-IC [21] and to the multicell MIMO-BC [22]–[25].

In this paper, we consider a multicell MIMO-BC system model and propose three different decentralized and novel WSR maximization algorithms, which are based on the alternating optimization technique [26] and are guaranteed to converge to a local WSR-optimum. The proposed algorithms are summarized as follows. The first algorithm uses an interference pricing approach, the same as in [18], where each BS maximizes its own utility that is formed by the local users' WSR minus the priced-ICI leakage. In [18], the authors assumed single-antenna users and formulated the problem as a relaxed Semidefinite Programming (SDP), whose solution requires each BS to first obtain the transmit covariance matrices, followed by an operation to guarantee and extract the rank-one transmit beamforming vectors. Different from [18], we consider multi-antenna users and the transmit beamforming matrices are obtained directly by investigating the KKT conditions of the problem. The main ingredient is given by Lemma 1, which makes it possible to solve the transmit beamforming directly from the problem cost function, in contrast to the WSR-WMMSE from [20]–[25]. Through computer simulations, it is shown that the proposed algorithm can achieve a comparable sum rate performance to WSR-WMMSE, while using fewer iterations.

The second algorithm designs the transmit beamforming

that maximizes the network-wide WSR by generalizing the solution steps of the first algorithm. Interestingly, it is proven that the WSR maximization via interference pricing can be made equivalent to the network-wide WSR maximization whenever the Mobile Stations (MSs) have single-antenna, i.e., in the multicell MISO BC. However, the interference pricing approach is shown to have some performance loss when the MSs have multiple antennas, as compared to the network-wide approach.

The third algorithm is an implicit interference pricing approach, where each BS self-prices its ICI leakage to other cells. Through computer simulations, it is shown that the self-pricing approach has negligible performance loss, as compared to the network-wide approach, when the BSs have enough Degrees of Freedom (Dof). In this case, the self-pricing approach is more appealing for practical systems, since it does not require feedback of variables from other cells.

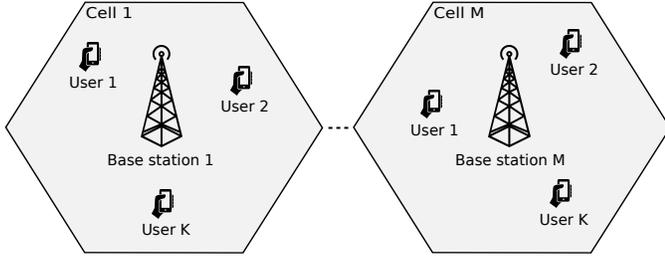
The proposed algorithms are decentralized in the sense that each BS can solve for its transmit beamforming independently, as soon as it has the required information. Here, we assume that each BS can acquire the local Channel State Information (CSI) between itself and all the MSs in the system, as in [20]–[25]. An effective technique for obtaining this CSI is the Time Division Duplex (TDD) operation, where uplink training in conjunction with reciprocity simultaneously provides the BSs with downlink and uplink channel estimates [27]. Furthermore, the TDD mode is more applicable than Frequency Division Duplex (FDD) to local area deployments and small cells, where the transmit powers, mobile speeds, and channel propagation delays are relatively low [27]–[29]. In this paper, we further propose a novel Over-the-Air (OTA) signaling scheme based on TDD mode to facilitate the algorithms' implementation. In contrast to some existing signaling schemes in [23]–[25], the proposed scheme reduces the signaling overhead and requires no feedback of variables between BSs.

The rest of this paper is organized as follows. In section II we present the system model. In section III we review the Block Diagonalization (BD) approach from [6]. The proposed algorithms and the Over-the-Air (OTA) signaling scheme for WSR maximization are presented in sections IV and V, respectively. Finally, in section VI we present numerical results and then conclude the paper in section VII.

Notations: Upper/lower boldface letters are used for matrices/vectors. The notations  $(\cdot)^H$ ,  $\|\cdot\|$ ,  $(\cdot)^\dagger$ ,  $\text{Tr}(\cdot)$ ,  $\log(\cdot)$ , and  $|\cdot|$  denote the complex conjugate transpose, the standard Euclidean norm, the pseudo-inverse, the trace, the logarithm of base 2, and the determinant, respectively.  $\mathbb{E}(\cdot)$  denotes the statistical expectation.  $\text{Bdiag}\{\cdot\}$  denotes the block-diag operator of a given vector/matrix.  $[\mathbf{A}]_{[1:N]}$  selects the first  $N$  vectors of a matrix  $\mathbf{A}$ , while  $[\mathbf{A}]_i$  selects the  $i$ -th vector.

## II. SYSTEM MODEL

We consider a multicell MIMO BC wireless network consisting of  $M$  cells, as in Fig.1. In each cell, there is one BS equipped with  $N_t$  antennas and  $K$  MSs, each equipped with  $N_r$  antennas. We denote the BS of the  $n$ -th cell as  $\text{BS}_n$  and the  $k$ -th MS in the  $n$ -th cell as  $\text{MS}_{nk}$ . Let  $\mathcal{M} \stackrel{\text{def}}{=} \{1, \dots, M\}$


 Fig. 1. Multicell MIMO BC system diagram ( $M$  cells and  $K$  users per cell).

and  $\mathcal{K} \stackrel{\text{def}}{=} \{\mathcal{K}_1, \dots, \mathcal{K}_M\}$  denote the sets of all BSs and MSs, respectively, whereas  $\mathcal{K}_n$  denotes the set of MSs associated with BS $_n$ . The  $M$  BSs are assumed to operate over a common frequency channel and communicate with their  $K$  respective MSs using linear transmit beamforming. The scenario under consideration assumes that each MS is served by only one BS. The received signal at MS $_{nk}$  is given as

$$\mathbf{y}_{nk} = \mathbf{H}_{nnk} \mathbf{T}_{nk} \mathbf{s}_{nk} + \underbrace{\sum_{i \in \mathcal{K}_n \setminus k} \mathbf{H}_{nnk} \mathbf{T}_{ni} \mathbf{s}_{ni}}_{\text{intra-cell interference}} + \underbrace{\sum_{m \in \mathcal{M} \setminus n} \sum_{j \in \mathcal{K}_m} \mathbf{H}_{mnk} \mathbf{T}_{mj} \mathbf{s}_{mj}}_{\text{inter-cell interference}} + \mathbf{z}_{nk}, \quad (1)$$

where  $\mathbf{H}_{mnk} \in \mathbb{C}^{N_r \times N_t}$  denotes the MIMO channel matrix from BS $_m$  to MS $_{nk}$ , whose coefficients are independent and identically distributed (i.i.d) complex Gaussian random variables,  $\mathbf{T}_{nk} \in \mathbb{C}^{N_t \times N_s}$  denotes the transmit beamforming, with  $N_s$  being the number of data streams,  $\mathbf{s}_{nk} \in \mathbb{C}^{N_s}$  denotes the transmitted data vector that is statistically independent with zero mean and  $\mathbb{E}(\mathbf{s}_{nk} \mathbf{s}_{nk}^H) = \mathbf{I}, \forall k \in \mathcal{K}$ , and  $\mathbf{z}_{nk} \in \mathbb{C}^{N_r}$  denotes the i.i.d complex Gaussian noise vector with zero mean and variance  $\sigma_{nk}^2$ . To decode the desired signal, each MS $_{nk}$  multiplies its received signal vector  $\mathbf{y}_{nk}$  by the receive beamforming matrix  $\mathbf{R}_{nk} \in \mathbb{C}^{N_s \times N_r}$ . Thus, the received data vector  $\hat{\mathbf{s}}_{nk}$  at MS $_{nk}$  is given as  $\hat{\mathbf{s}}_{nk} = \mathbf{R}_{nk} \mathbf{y}_{nk}$ .

### III. BLOCK DIAGONALIZATION APPROACH

Theoretically, the cBD algorithm from [4] can be interpreted as the equivalent Zero-Forcing (ZF) algorithm for the MIMO system. The main objective is to completely eliminate intra-cell interference by forcing each user to transmit on the null space of all other users in the same cell. The cBD optimization problem of BS $_n, \forall n \in \mathcal{M}$ , can be written as

$$\begin{aligned} \max_{\mathbf{T}_n} \quad & r_n^{\text{cBD}} = \log \left| \mathbf{I} + \frac{\mathbf{R}_n \mathbf{H}_n \mathbf{T}_n \mathbf{T}_n^H \mathbf{H}_n^H \mathbf{R}_n^H}{\mathbf{R}_n \mathbf{Y}_n \mathbf{R}_n^H} \right|, \quad (\mathcal{P}_{\text{cBD}}) \\ \text{s.t.} \quad & \sum_{j \in \mathcal{K}_n \setminus k} \mathbf{R}_{nk} \mathbf{H}_{nnk} \mathbf{T}_{nj} = 0, \forall k \in \mathcal{K}_n, \\ & \text{Tr}[\mathbf{T}_n \mathbf{T}_n^H] = p_n, \end{aligned}$$

where  $p_n$  is the transmit power threshold,  $r_n^{\text{cBD}}$  is the BS $_n$  achievable rate,  $\mathbf{H}_n = [\mathbf{H}_{nn1}^T, \dots, \mathbf{H}_{nnK}^T]^T$ ,  $\mathbf{T}_n = [\mathbf{T}_{n1}, \dots, \mathbf{T}_{nK}]$ ,  $\mathbf{R}_n = \text{Bdiag}\{\mathbf{R}_{n1}, \dots, \mathbf{R}_{nK}\}$ , and  $\mathbf{Y}_n =$

$\text{Bdiag}\{\mathbf{Y}_{n1}, \dots, \mathbf{Y}_{nK}\}$ , where  $\mathbf{Y}_{nk}$  denotes the ICI plus noise covariance matrix of MS $_{nk}$ , which is given as

$$\mathbf{Y}_{nk} = \sum_{m \in \mathcal{M} \setminus n} \sum_{j \in \mathcal{K}_m} \mathbf{H}_{mnk} \mathbf{T}_{mj} \mathbf{T}_{mj}^H \mathbf{H}_{mnk}^H + \sigma_{nk}^2 \mathbf{I}_{N_r}. \quad (2)$$

As one can notice, problem  $\mathcal{P}_{\text{cBD}}$  does nothing to deal with the ICI that is being received from the other cells or is leaking to the other cells, since its main objective is to maximize each cell's achievable rate such that all intra-cell interference is eliminated. The main advantage, though, is that  $\mathcal{P}_{\text{cBD}}$  is completely distributed between  $M$  cells and has a closed-form solution as follows. The transmit beamforming matrix of MS $_{nk}$  is given as

$$\mathbf{T}_{nk} = \mathbf{G}_{nk} \mathbf{F}_{nk} \mathbf{P}_{nk}^{\frac{1}{2}}, \quad (3)$$

where  $\mathbf{P}_{nk}$  holds on its diagonal the power allocation,  $\mathbf{G}_{nk}$  holds the orthogonal basis vectors of the null space of the *intra-cell users' channels*, and  $\mathbf{F}_{nk}$  holds the right singular vectors of the *effective channel* of MS $_{nk}$ . Let the *intra-cell users' channels* of MS $_{nk}$  be given as

$$\mathbf{H}_{nk}^{-k} = [\mathbf{H}_{nnj}^T, \forall j \in \mathcal{K}_n \setminus k]^T. \quad (4)$$

Then, to calculate  $\mathbf{G}_{nk}$ , define the Singular Value Decomposition (SVD) of  $\mathbf{H}_{nk}^{-k}$  as

$$\mathbf{H}_{nk}^{-k} = \mathbf{U}_{nk}^{-k} \mathbf{\Sigma}_{nk}^{-k} [\mathbf{V}_{nk}^{-k} \quad \mathbf{G}_{nk}], \quad (5)$$

where  $\mathbf{G}_{nk}$  is the last  $(N_t - l_{nk}^{-k})$  right singular vectors and  $l_{nk}^{-k}$  is the rank of  $\mathbf{H}_{nk}^{-k}$ . Further, let the *effective channel* of MS $_{nk}$  be given as

$$\mathbf{H}_{nk}^e = \mathbf{H}_{nnk} \mathbf{G}_{nk}. \quad (6)$$

Then, to calculate  $\mathbf{F}_{nk}$ , define the SVD of  $\mathbf{H}_{nk}^e$  as

$$\mathbf{H}_{nk}^e = \mathbf{U}_{nk}^e \begin{bmatrix} \mathbf{\Sigma}_{nk}^e & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} [\mathbf{V}_{nk}^{e(1)} \quad \mathbf{V}_{nk}^{e(0)}], \quad (7)$$

where  $\mathbf{\Sigma}_{nk}^e$  is an  $[l_{nk}^e \times l_{nk}^e]$  diagonal matrix,  $\mathbf{V}_{nk}^{e(1)}$  contains the first  $l_{nk}^e$  singular vectors, and  $l_{nk}^e$  is the rank of  $\mathbf{H}_{nk}^e$ . Therefore, assuming the values of  $\mathbf{\Sigma}_{nk}^e$  are in a decreasing order, we choose  $\mathbf{F}_{nk}$  and  $\mathbf{R}_{nk}$  to be the first  $N_s$  vectors of  $\mathbf{V}_{nk}^{e(1)}$  and  $\mathbf{U}_{nk}^e$ , respectively, i.e.,

$$\mathbf{F}_{nk} = [\mathbf{V}_{nk}^{e(1)}]_{[1:N_s]} \quad \text{and} \quad \mathbf{R}_{nk} = [\mathbf{U}_{nk}^e]_{[1:N_s]}^H. \quad (8)$$

With transmit and receive beamforming matrices calculated as above, the BS $_n$  rate function is reduced to

$$r_n^{\text{cBD}} = \max_{\mathbf{P}_n} \log \left| \mathbf{I} + \frac{\mathbf{\Sigma}_n^e \mathbf{P}_n}{\mathbf{R}_n \mathbf{Y}_n \mathbf{R}_n^H} \right|, \quad (9)$$

where  $\mathbf{\Sigma}_n = \text{Bdiag}\{\mathbf{\Sigma}_{n1}^e, \dots, \mathbf{\Sigma}_{nK}^e\}$  and  $\mathbf{P}_n$  is a diagonal matrix that holds the optimal power loading found using water-filling method [30] on the  $\mathbf{\Sigma}_n$  diagonal elements<sup>1</sup>.

From above, it can be seen that one of the main issues with cBD is that it does nothing to reduce the effects of the

<sup>1</sup>Note that the water-filling method is applied individually on each sub-matrix of  $\mathbf{\Sigma}_n$  assuming that total transmit power of the BS is divided equally between its  $K$  users. Clearly, the equal power allocation is a suboptimal solution. However, this is done to make sure that all  $K$  users are allocated for transmission.

ICI each user is receiving. This issue has been considered in [5], where the authors proposed the eBD algorithm to account for the ICI presence. The eBD algorithm is summarized as follows. First, to suppress the ICI effects,  $\text{MS}_{nk}$  uses the whitening matrix  $\mathbf{W}_{nk} = \mathbf{\Upsilon}_{nk}^{-\frac{1}{2}}$  at the received signal. Let  $\mathbf{W}_n = \text{Bdiag}[\mathbf{W}_{n1}, \dots, \mathbf{W}_{nK}]$ . Then, the  $\text{BS}_n$  rate function  $r_n^{\text{eBD}}$  can be written as

$$\begin{aligned} r_n^{\text{eBD}} &= \max_{\hat{\mathbf{T}}_n} \log \left| \mathbf{I} + \frac{\hat{\mathbf{R}}_n \mathbf{W}_n \mathbf{H}_n \hat{\mathbf{T}}_n \hat{\mathbf{T}}_n^H \mathbf{H}_n^H \mathbf{W}_n^H \hat{\mathbf{R}}_n^H}{\hat{\mathbf{R}}_n \mathbf{W}_n \mathbf{\Upsilon}_n \mathbf{W}_n^H \hat{\mathbf{R}}_n^H} \right| \\ &= \max_{\hat{\mathbf{T}}_n} \log \left| \mathbf{I} + \hat{\mathbf{R}}_n \hat{\mathbf{H}}_n \hat{\mathbf{T}}_n \hat{\mathbf{T}}_n^H \hat{\mathbf{H}}_n^H \hat{\mathbf{R}}_n^H \right|, \end{aligned} \quad (10)$$

where  $\hat{\mathbf{H}}_n = \mathbf{W}_n \mathbf{H}_n$  (thus,  $\hat{\mathbf{H}}_{nk} = \mathbf{W}_{nk} \mathbf{H}_{nnk}$ ). As in the case of cBD, the transmit beamforming matrix of  $\text{MS}_{nk}$  is given as  $\hat{\mathbf{T}}_{nk} = \hat{\mathbf{G}}_{nk} \hat{\mathbf{F}}_{nk} \hat{\mathbf{P}}_{nk}^{\frac{1}{2}}$ , where  $\hat{\mathbf{G}}_{nk}$  is calculated similar to (5) from

$$\hat{\mathbf{H}}_{nk}^{-k} = [\hat{\mathbf{H}}_{nj}^{-k}, \forall j \in \mathcal{K}_n \setminus k]^T. \quad (11)$$

The  $\hat{\mathbf{F}}_{nk}$  and  $\hat{\mathbf{R}}_{nk}$  matrices are calculated similar to (8) from the  $\text{MS}_{nk}$  effective channel

$$\hat{\mathbf{H}}_{nk}^e = \hat{\mathbf{H}}_{nk} \hat{\mathbf{G}}_{nk} = \mathbf{W}_{nk} \mathbf{H}_{nnk} \hat{\mathbf{G}}_{nk}. \quad (12)$$

Consequently, the  $\text{BS}_n$  rate function  $r_n^{\text{eBD}}$  given by (10) is reduced to

$$r_n^{\text{eBD}} = \max_{\hat{\mathbf{P}}_n} \log \left| \mathbf{I} + \hat{\Sigma}_n^2 \hat{\mathbf{P}}_n \right|. \quad (13)$$

From above, one can see that both BD approaches, cBD and eBD, have the same dimensionality restrictions. The expressions given by (4) and (11) have dimension of  $[(K-1)N_r \times N_t]$  with rank of  $[N_t - (K-1)N_r]$ . Therefore, to have  $N_s$  columns in the null space,  $[N_t - (K-1)N_r]$  should be larger than  $N_s$ , i.e.,  $[N_t - (K-1)N_r] \geq N_s$ . Moreover, it is important to note that both approaches use only the transmit beamforming  $\mathbf{T}_n$  to eliminate the intra-cell interference, i.e., the receive beamforming  $\mathbf{R}_n$  is not utilized. Motivated by the last observation, one possible way to reduce the dimensionality restrictions is to utilize the receive beamforming matrix when calculating the transmit beamforming matrix [6]. To achieve this end, the receive beamforming matrix  $\tilde{\mathbf{R}}_{nk}$  can be included in (11), then we have

$$\tilde{\mathbf{H}}_{nk}^{-k} = [[\tilde{\mathbf{R}}_{nj} \hat{\mathbf{H}}_{nj}]^T, \forall j \in \mathcal{K}_n \setminus k]^T. \quad (14)$$

Note that  $\tilde{\mathbf{H}}_{nk}^{-k}$  has dimension of  $[(K-1)N_s \times N_t]$ , which is no longer in function of  $N_r$ . Calculating the null space from  $\tilde{\mathbf{H}}_{nk}^{-k}$  is always satisfied if, and only if, the number of data streams transmitted by a BS is less than or equal to its number of transmit antennas, i.e., the condition of  $[N_t - (K-1)N_s] \geq N_s$  should be satisfied. The following steps are much similar to the ones above. The transmit beamforming is given as  $\tilde{\mathbf{T}}_{nk} = \tilde{\mathbf{G}}_{nk} \tilde{\mathbf{F}}_{nk} \tilde{\mathbf{P}}_{nk}^{\frac{1}{2}}$ , where  $\tilde{\mathbf{G}}_{nk}$  is calculated similarly from  $\tilde{\mathbf{H}}_{nk}^{-k}$ . The  $\tilde{\mathbf{F}}_{nk}$  and  $\tilde{\mathbf{R}}_{nk}$  matrices are calculated from the  $\text{MS}_{nk}$  effective channel

$$\tilde{\mathbf{H}}_{nk}^e = \hat{\mathbf{H}}_{nk} \tilde{\mathbf{G}}_{nk} = \mathbf{W}_{nk} \mathbf{H}_{nnk} \tilde{\mathbf{G}}_{nk}. \quad (15)$$

Since the transmit and receive beamforming matrices are now coupled, the BS is required to conduct some iterations in

---

**Algorithm 1** iterative BD (iBD).
 

---

- 1: Initialize  $\tilde{\mathbf{R}}_{nk}^{(1)}, \tilde{\mathbf{T}}_{nk}^{(1)}, \forall k \in \mathcal{K}$  and set  $t = 1$ .
  - 2:  $\text{BS}_n, \forall n$ : Transmit data using  $\tilde{\mathbf{T}}_{nk}^{(t)}, \forall k \in \mathcal{K}_n$ .
  - 3:  $\text{MS}_{nk}, \forall k$ : Calculate  $\mathbf{\Upsilon}_{nk}^{(t)}$  and feed it back to  $\text{BS}_n$ .
  - 4:  $\text{BS}_n, \forall n$ : Calculate  $\tilde{\mathbf{R}}_{nk}^{(t)}$  and  $\tilde{\mathbf{T}}_{nk}^{(t)}, \forall k \in \mathcal{K}_n$  as:
    - Construct  $\tilde{\mathbf{H}}_{nk}^{-i(t)}$  using  $\tilde{\mathbf{R}}_{nk}^{(t)}$ .
    - Calculate  $\tilde{\mathbf{G}}_{nk}^{(t)}$  from  $\tilde{\mathbf{H}}_{nk}^{-i(t)}$ .
    - Construct  $\tilde{\mathbf{H}}_{nk}^{e(t)}$  using  $\tilde{\mathbf{G}}_{nk}^{(t)}$ .
    - Calculate  $\tilde{\mathbf{F}}_{nk}^{(t+1)}$  and  $\tilde{\mathbf{R}}_{nk}^{(t+1)}$  from  $\tilde{\mathbf{H}}_{nk}^{e(t)}$ .
  - 5: Repeat steps 2-4 (until convergence)
- 

order to achieve BD. Therefore, we refer to this approach as iterative BD (iBD) and summarize it in Algorithm 1.

At the first step, Algorithm 1 initializes the transmit and receive beamforming matrices for all users. For instance,  $\tilde{\mathbf{T}}_{nk}^{(1)}$  can be initialized using the Maximum Ratio Transmission (MRT) approach and  $\tilde{\mathbf{R}}_{nk}^{(1)} = \mathbf{I}$ . At the  $t$ -th iteration, each BS transmits pilot signals precoded with  $\tilde{\mathbf{T}}_{nk}^{(t)}$  at step-2 so that each  $\text{MS}_{nk}$  can calculate the ICI covariance matrix, i.e.,  $\mathbf{\Upsilon}_{nk}^{(t)}$ , and feed it back to its serving BS via feedback channels. After that, each BS updates the transmit and receive beamforming of its users at step-4. The aforementioned steps are repeated until convergence. Note that the transmit and receive beamforming matrices of all users are calculated at the BSs. Therefore, at convergence, each BS would forward the receive beamforming matrices to its users using the feedforward channels. With the transmit and receive beamforming matrices calculated as given by Algorithm 1, the  $\text{BS}_n$  rate function  $r_n^{\text{eBD}}$  given by (10) is reduced to

$$r_n^{\text{iBD}} = \max_{\tilde{\mathbf{P}}_n} \log \left| \mathbf{I} + \tilde{\Sigma}_n^2 \tilde{\mathbf{P}}_n \right|. \quad (16)$$

Note that both equations (12) and (15) have the same structure. The following theorem indicates their relation.

*Theorem 1:* If the number of data streams transmitted to any user is equal to the number of its receive antennas, i.e., if  $N_s = N_r$ , then, both eBD and iBD are equivalent and have the same exact performance.

*Proof 1:* Please refer to Appendix A.

It's worth noting that if the system has only one cell, then the eBD algorithm is equivalent to cBD. In this case, the iBD algorithm is also equivalent to cBD, only if  $N_s = N_r$ , which is a straightforward result of Theorem 1.

#### IV. WEIGHTED SUM RATE MAXIMIZATION APPROACH

In this section, we use WSR as the transceiver design criterion. We assume that each MS employs single-user detection by treating the interference as additive noise. Therefore, the achievable rate of  $\text{MS}_{nk}$  can be written as

$$r_{nk} = \log \left| \mathbf{I}_{N_s} + \mathbf{T}_{nk}^H \mathbf{H}_{nnk}^H \mathbf{\Phi}_{nk}^{-1} \mathbf{H}_{nnk} \mathbf{T}_{nk} \right|, \quad (17)$$

where  $\mathbf{\Phi}_{nk}$  denotes the received interference plus-noise covariance matrix for  $\text{MS}_{nk}$ , which is given as

$$\mathbf{\Phi}_{nk} = \sum_{i \in \mathcal{K}_n \setminus k} \mathbf{H}_{nnk} \mathbf{T}_{ni} \mathbf{T}_{ni}^H \mathbf{H}_{nnk}^H + \mathbf{\Upsilon}_{nk}, \quad (18)$$

whereas  $\mathbf{T}_{nk}$  is given by (2), which denotes the ICI plus noise covariance matrix of  $\text{MS}_{nk}$ . Here, we assume that each  $\text{MS}_{nk}$  uses Minimum-Mean-Square Error (MMSE) receive beamforming, which is given as [20]

$$\begin{aligned} \mathbf{R}_{nk}^{\text{MMSE}} &= \arg \min_{\mathbf{R}_{nk}} \mathbb{E}[\|\mathbf{R}_{nk} \mathbf{y}_{nk} - \mathbf{s}_{nk}\|^2] \\ &= \mathbf{T}_{nk}^H \mathbf{H}_{nnk}^H \mathbf{\Omega}_{nk}^{-1}, \end{aligned} \quad (19)$$

where  $\mathbf{\Omega}_{nk} = \mathbf{H}_{nnk} \mathbf{T}_{nk} \mathbf{T}_{nk}^H \mathbf{H}_{nnk}^H + \mathbf{\Phi}_{nk}$ . Using (19), the MSE-matrix of  $\text{MS}_{nk}$  is given as [23]

$$\begin{aligned} \mathbf{E}_{nk} &= \mathbb{E}[\|\mathbf{R}_{nk}^{\text{MMSE}} \mathbf{y}_{nk} - \mathbf{s}_{nk}\|^2] \\ &= \mathbf{I}_{N_s} - \mathbf{R}_{nk}^{\text{MMSE}} \mathbf{H}_{nnk} \mathbf{T}_{nk}, \end{aligned} \quad (20)$$

which can be equivalently expressed as [20]

$$\mathbf{E}_{nk} = (\mathbf{I}_{N_s} + \mathbf{T}_{nk}^H \mathbf{H}_{nnk}^H \mathbf{\Phi}_{nk}^{-1} \mathbf{H}_{nnk} \mathbf{T}_{nk})^{-1}. \quad (21)$$

The latter form of  $\mathbf{E}_{nk}$  in (21) shows that the rate function given by (17) can be equivalently expressed as

$$r_{nk} = \log \left| \mathbf{E}_{nk}^{-1} \right|. \quad (22)$$

Note that  $\mathbf{E}_{nk}$  must be Hermitian, since from (21),  $\mathbf{E}_{nk}$  equals a quantity (right-hand side) that is Hermitian, which means that  $\mathbf{E}_{nk} = \mathbf{E}_{nk}^H$ . Furthermore, the following lemma is needed throughout the rest of the paper.

*Lemma 1:* Given the MSE matrix  $\mathbf{E}_{nk}$  as in (20), or equivalently as in (21), the receive beamforming matrix  $\mathbf{R}_{nk}$  can be written as

$$\mathbf{R}_{nk} = \mathbf{E}_{nk} \mathbf{T}_{nk}^H \mathbf{H}_{nnk}^H \mathbf{\Phi}_{nk}^{-1}. \quad (23)$$

*Proof 2:* Please refer to Appendix B.

### A. Per-Cell WSR maximization via Interference Pricing

In this section, we consider an interference pricing approach for designing the transmit beamforming. The main idea is to manage the ICI received by a user by pricing the interfering BSs. Similar to [18], [31], we define the interference price as the marginal decrease in the user rate due to a marginal increase in the received interference. Mathematically, the  $\text{MS}_{nk}$  interference price is given as

$$\pi_{nk} = \nabla_{\mathbf{\Phi}_{nk}} r_{nk}. \quad (24)$$

Using the result of  $\nabla \log |\mathbf{X}| = \text{Tr}(\mathbf{X}^{-1} \nabla \mathbf{X})$ , where  $\mathbf{X}$  is a matrix [32], then  $\pi_{nk}$  is given as

$$\pi_{nk} = \text{Tr}(\mathbf{E}_{nk} \mathbf{T}_{nk}^H \mathbf{H}_{nnk}^H \mathbf{\Phi}_{nk}^{-2} \mathbf{H}_{nnk} \mathbf{T}_{nk}). \quad (25)$$

By observing (25) and the Lemma 1 result, we have the following corollary.

*Corollary 1:* The  $\text{MS}_{nk}$  interference price  $\pi_{nk}$  given by (25) can be equivalently written as

$$\pi_{nk} = \text{Tr}(\mathbf{R}_{nk}^H \mathbf{E}_{nk}^{-1} \mathbf{R}_{nk}). \quad (26)$$

*Proof 3:* According to Lemma 1, the receive beamforming can be written as  $\mathbf{R}_{nk} = \mathbf{E}_{nk} \mathbf{T}_{nk}^H \mathbf{H}_{nnk}^H \mathbf{\Phi}_{nk}^{-1}$ . Then, the  $\text{MS}_{nk}$  interference price  $\pi_{nk}$  given by (25) is reduced to

$$\begin{aligned} \pi_{nk} &= \text{Tr}(\mathbf{E}_{nk} \mathbf{T}_{nk}^H \mathbf{H}_{nnk}^H \mathbf{\Phi}_{nk}^{-2} \mathbf{H}_{nnk} \mathbf{T}_{nk}) \\ &\stackrel{(a)}{=} \text{Tr}(\mathbf{R}_{nk} \mathbf{R}_{nk}^H \mathbf{E}_{nk}^{-1}) \\ &\stackrel{(b)}{=} \text{Tr}(\mathbf{R}_{nk}^H \mathbf{E}_{nk}^{-1} \mathbf{R}_{nk}) \end{aligned}$$

where (a) is obtained by substituting  $\mathbf{R}_{nk}$  into the first equality and (b) is obtained by using the results of  $\text{Tr}(\mathbf{X}\mathbf{Y}\mathbf{Z}) = \text{Tr}(\mathbf{Y}\mathbf{Z}\mathbf{X}) = \text{Tr}(\mathbf{Z}\mathbf{X}\mathbf{Y})$  [32], which completes the proof.

Let  $\boldsymbol{\pi}_n = \{\pi_{mj}, \forall m \in \mathcal{M} \setminus n, \forall j \in \mathcal{K}_m\}$  denote the vector that collects all interference prices of all users in the system except  $\text{BS}_n$  users. Then, define the following MS-specific function

$$f_{nk}(\boldsymbol{\pi}_n) = \mu_{nk} r_{nk} - \text{Tr}(\mathbf{L}_{nk}), \quad (27)$$

where  $\mu_{nk} > 0$  denotes the weight associated to  $\text{MS}_{nk}$  and  $\mathbf{L}_{nk}$  defines the priced-ICI caused by the  $\text{MS}_{nk}$  beamforming  $\mathbf{T}_{nk}$ , which is given as

$$\mathbf{L}_{nk} = \sum_{m \in \mathcal{M} \setminus n} \sum_{j \in \mathcal{K}_m} \pi_{mj} \mathbf{H}_{nmj} \mathbf{T}_{nk} \mathbf{T}_{nk}^H \mathbf{H}_{nmj}^H. \quad (28)$$

Afterwards, each  $\text{BS}_n, \forall n \in \mathcal{M}$ , updates its transmit beamforming  $\mathbf{T}_{nk}, \forall k \in \mathcal{K}_n$ , as the solution to the following interference-priced WSR maximization problem

$$\begin{aligned} \max_{\substack{\mathbf{T}_{nk}, \\ \forall k \in \mathcal{K}_n}} \sum_{k \in \mathcal{K}_n} f_{nk}(\boldsymbol{\pi}_n) \quad (\mathcal{P}_{\text{WSRP}}) \\ \text{s.t.} \sum_{k \in \mathcal{K}_n} \text{Tr}(\mathbf{T}_{nk} \mathbf{T}_{nk}^H) = p_n, \end{aligned}$$

where we have used an equality power constraint rather than the often used  $\sum_{k \in \mathcal{K}_n} \text{Tr}(\mathbf{T}_{nk} \mathbf{T}_{nk}^H) \leq p_n$ , since the WSR optimum is reached at maximum transmit power [20]. From problem  $\mathcal{P}_{\text{WSRP}}$ , one can see that this approach is different from the BD approach, in the sense that the ICI received by a user is being managed by the interfering BSs and not the serving BS.

In [18], the authors addressed problem  $\mathcal{P}_{\text{WSRP}}$  from a game-theoretic view-point assuming single-antenna users, where the function  $f_{nk}(\boldsymbol{\pi}_n)$  is interpreted as a user utility function that penalizes the user rate by the ICI that he is leaking. The problem in [18], however, was formulated as a relaxed SDP and its solution would require each BS to obtain first the transmit covariance matrices, i.e.,  $\mathbf{Q}_{nk} \stackrel{\text{def}}{=} \mathbf{t}_{nk} \mathbf{t}_{nk}^H \succeq 0$ , followed by operations to guarantee and extract the rank-one transmit beamforming vectors, i.e.,  $\mathbf{t}_{nk}$ . It was proven in [18] that problem  $\mathcal{P}_{\text{WSRP}}$  is guaranteed to converge to an equilibrium point that corresponds to a KKT point for the original problem  $\mathcal{P}_{\text{WSRP}}$ . In the following, we present a different solution to problem  $\mathcal{P}_{\text{WSRP}}$ . The solution is obtained by investigating the KKT conditions of problem  $\mathcal{P}_{\text{WSRP}}$  and with the help of the Lemma 1 result. The solution of  $\mathcal{P}_{\text{WSRP}}$  w.r.t. transmit beamforming for  $\text{MS}_{nk}, \forall k \in \mathcal{K}$ , is given by Proposition 1.

*Proposition 1:* Let the receive beamforming  $\mathbf{R}_{nk}$  and MSE matrix  $\mathbf{E}_{nk}$  for  $\text{MS}_{nk}$  be given by (19) and (20),

respectively, and by utilizing the Lemma 1 result, then the solution of problem  $\mathcal{P}_{\text{WSRP}}$  w.r.t. transmit beamforming  $\mathbf{T}_{nk}$  for  $\text{MS}_{nk}, \forall k \in \mathcal{K}_n$ , is given as

$$\mathbf{T}_{nk}^{\text{WSRP}} = (\mathbf{A}_{nk} + \mathbf{B}_n + \lambda_n \mathbf{I}_{N_t})^{-1} \mathbf{H}_{nnk}^H \mathbf{R}_{nk}^H \mu_{nk}, \quad (29)$$

where  $\lambda_n, \forall n \in \mathcal{M}$ , are the Lagrange multipliers associated with the  $\mathcal{P}_{\text{WSRP}}$  constraint functions,  $\mathbf{A}_{nk}$  and  $\mathbf{B}_n$  are given as

$$\mathbf{A}_{nk} = \sum_{i \in \mathcal{K}_n \setminus k} \mu_{ni} \mathbf{H}_{nni}^H \mathbf{R}_{ni}^H \mathbf{E}_{ni}^{-1} \mathbf{R}_{nk} \mathbf{H}_{nni}, \quad (30)$$

$$\mathbf{B}_n = \sum_{m \in \mathcal{M} \setminus n} \sum_{j \in \mathcal{K}_m} \pi_{mj} \mathbf{H}_{nmj}^H \mathbf{H}_{nmj}. \quad (31)$$

*Proof 4:* Please refer to Appendix C.

In (29), the  $\lambda_n, \forall n \in \mathcal{M}$ , are calculated to satisfy the power constraint at  $\text{BS}_n, \forall n \in \mathcal{M}$ , by using the KKT condition  $\lambda_n (\sum_{k \in \mathcal{K}_n} \text{Tr}(\mathbf{T}_{nk} \mathbf{T}_{nk}^H) - p_n) = 0$  and by utilizing the fact that the transmit power is monotonically decreasing with respect to increasing  $\lambda_n$  [21]. The closed-form solution can be obtained by readapting the approach shown in [20] as

$$\begin{aligned} \mathbf{T}_{nk}^{\text{WSRP}} &= \beta_n \bar{\mathbf{T}}_{nk}, \quad (32) \\ \begin{cases} \bar{\mathbf{T}}_{nk} = (\mathbf{A}_{nk} + \mathbf{B}_n + \alpha_n \mathbf{I}_{N_t})^{-1} \mathbf{H}_{nnk}^H \mathbf{R}_{nk}^H \mu_{nk}, \\ \alpha_n = \sum_{i \in \mathcal{K}_n} \text{Tr}(\mu_{ni} \mathbf{E}_{ni}^{-1} \mathbf{R}_{ni}^H \mathbf{R}_{ni}) / p_n, \\ \beta_n = \sqrt{p_n / \sum_{i \in \mathcal{K}_n} \text{Tr}(\bar{\mathbf{T}}_{ni} \bar{\mathbf{T}}_{ni}^H)}. \end{cases} \end{aligned}$$

### B. Network-Wide WSR Maximization

In this section, we consider the general network-wide WSR maximization problem. Mathematically, the WSR maximization problem can be written as [21]

$$\begin{aligned} \max_{\mathbf{T}_{nk}} & \sum_{n \in \mathcal{M}} \sum_{k \in \mathcal{K}_n} \mu_{nk} r_{nk} \quad (\mathcal{P}_{\text{WSRM}}) \\ \text{s.t.} & \sum_{k \in \mathcal{K}_n} \text{Tr}(\mathbf{T}_{nk} \mathbf{T}_{nk}^H) = p_n, \forall n \in \mathcal{M}. \end{aligned}$$

Problem  $\mathcal{P}_{\text{WSRM}}$  has been addressed in [20]–[25]. For all the algorithms presented in these references,  $\mathcal{P}_{\text{WSRM}}$  was solved by exploiting its relation to the WMMSE minimization problem, which was initially shown in [20]. Different from all, in the following, we propose a novel solution that directly solves  $\mathcal{P}_{\text{WSRM}}$ . Similar to  $\mathcal{P}_{\text{WSRP}}$ , the solution is obtained by investigating the KKT conditions of problem  $\mathcal{P}_{\text{WSRM}}$  and with the help of the Lemma 1 result. The solution of  $\mathcal{P}_{\text{WSRM}}$  w.r.t. transmit beamforming for  $\text{MS}_{nk}, \forall n \in \mathcal{M}, \forall k \in \mathcal{K}_n$ , is given by Proposition 2.

*Proposition 2:* Let the receive beamforming  $\mathbf{R}_{nk}$  and the MSE matrix  $\mathbf{E}_{nk}$  for  $\text{MS}_{nk}$  be given by (19) and (20), respectively, and by utilizing the Lemma 1 result, then the solution to  $\mathcal{P}_{\text{WSRM}}$  w.r.t. transmit beamforming for  $\text{MS}_{nk}, \forall n \in \mathcal{M}, \forall k \in \mathcal{K}_n$ , is given as

$$\mathbf{T}_{nk}^{\text{WSRM}} = (\mathbf{A}_{nk} + \mathbf{C}_n + \lambda_n \mathbf{I}_{N_t})^{-1} \mathbf{H}_{nnk}^H \mathbf{R}_{nk}^H \mu_{nk}, \quad (33)$$

where  $\lambda_n, \forall n \in \mathcal{M}$ , are the Lagrange multipliers associated with the  $\mathcal{P}_{\text{WSRM}}$  constraint functions,  $\mathbf{A}_{nk}$  is given by (30), and  $\mathbf{C}_n$  is given as

$$\mathbf{C}_n = \sum_{m \in \mathcal{M} \setminus n} \sum_{j \in \mathcal{K}_m} \mu_{mj} \mathbf{H}_{nmj}^H \mathbf{R}_{mj}^H \mathbf{E}_{mj}^{-1} \mathbf{R}_{mj} \mathbf{H}_{nmj}. \quad (34)$$

*Proof 5:* The proof can be shown by generalizing the derivation steps shown in Appendix C and thus we omit them here for brevity.

By observing (33), we can see that it is closely related to (29). Theorem 2 shows the connection between both equations in a special case.

*Theorem 2:* Both equations  $\mathbf{T}_{nk}^{\text{WSRP}}$  and  $\mathbf{T}_{nk}^{\text{WSRM}}$ , given by (29) and (33), respectively, are equal if  $N_r = 1$  and the  $\text{MS}_{nk}$  interference price  $\pi_{nk}$  given by (26) is replaced by  $\tilde{\pi}_{nk}$  that is given as

$$\tilde{\pi}_{nk} = \mu_{nk} \pi_{nk}. \quad (35)$$

*Proof 6:* When  $N_r = 1$ , the interference price  $\pi_{nk}$  in (26) reduces to  $\pi_{nk} = \mathbf{R}_{nk}^H \mathbf{E}_{nk}^{-1} \mathbf{R}_{nk}$ , since both terms  $\mathbf{R}_{nk}$  and  $\mathbf{E}_{nk}$  are scalars. By substituting  $\pi_{nk}$  into the  $\mathbf{A}_{nk}$  and  $\mathbf{C}_n$  terms, we have

$$\begin{aligned} \mathbf{A}_{nk} &= \sum_{i \in \mathcal{K}_n \setminus k} \mu_{ni} \pi_{ni} \mathbf{H}_{nni}^H \mathbf{H}_{nni}. \\ \mathbf{C}_n &= \sum_{m \in \mathcal{M} \setminus n} \sum_{j \in \mathcal{K}_m} \mu_{mj} \pi_{mj} \mathbf{H}_{nmj}^H \mathbf{H}_{nmj}. \end{aligned}$$

Since  $\mathbf{A}_{nk}$  is common in both, the only difference is between the  $\mathbf{B}_n$  and  $\mathbf{C}_n$  terms. Now, comparing  $\mathbf{B}_n$  to  $\mathbf{C}_n$ , we can see that both terms are equal if each interference price in  $\mathbf{B}_n$  is replaced by  $\tilde{\pi}_{mj} = \mu_{mj} \pi_{mj}$ , which completes the proof.

The result of Theorem 2 establishes a relation between problems  $\mathcal{P}_{\text{WSRM}}$  and  $\mathcal{P}_{\text{WSRP}}$ . When  $N_r = 1$ , the problems  $\mathcal{P}_{\text{WSRP}}$  and  $\mathcal{P}_{\text{WSRM}}$  are exactly equivalent. In this case, the receive beamforming and MSE terms are scalars and directly specify the interference prices of the MSs. However, when  $N_r > 1$ ,  $\mathcal{P}_{\text{WSRP}}$  would provide a suboptimal solution to  $\mathcal{P}_{\text{WSRM}}$ . In this latter case, the interference prices cannot exploit the spatial dimension that the receive beamforming brings, since  $\mathbf{R}_{nk}^H \mathbf{E}_{nk}^{-1} \mathbf{R}_{nk}$  has a dimension of  $N_r \times N_r$ , irrespective of the number of data streams  $N_s$ , whereas the interference-price  $\pi_{nk} = \text{Tr}(\mathbf{R}_{nk}^H \mathbf{E}_{nk}^{-1} \mathbf{R}_{nk})$  is represented by a scalar. Therefore, when  $N_r > 1$ , the  $\mathbf{C}_n$  term given by (34) contains extra information, as compared to  $\mathbf{B}_n$  given by (31), which can be exploited by the BSs to reshape the interference.

Similar to (29), the  $\lambda_n, \forall n \in \mathcal{M}$ , in (33) are calculated to satisfy the power constraint at  $\text{BS}_n, \forall n \in \mathcal{M}$ . The closed-form solution of (33) can be obtained similar to (32), by replacing the  $\mathbf{B}_n$  term with  $\mathbf{C}_n$ .

### C. WSR Maximization Based on Self-pricing

In this approach, we consider a different strategy in the sense that each BS would self-price the ICI it is leaking to other cells. In this regard, when compared to  $\mathcal{P}_{\text{WSRP}}$ , the BSs do not need to collect the interference prices from the users when calculating the transmit beamforming.

In order to show this, let us assume for a moment that the rate function of a given user, say  $\text{MS}_{mj}, j \in \mathcal{K}_m$ , is mostly degraded by the transmit beamforming from a single interfering BS, say  $\text{BS}_n, n \neq m$ . This can be translated to many scenarios, such as the BSs other than  $\text{BS}_n$  have a mutual interference that is negligible, or they are using a transmit

beamforming strategy that eliminates ICI by any means. Therefore, the interference plus noise covariance matrix of  $\text{MS}_{m_j}$  can be approximated as

$$\Psi_{m_j} \approx \sum_{k \in \mathcal{K}_n} \mathbf{H}_{nm_j} \mathbf{T}_{nk} \mathbf{T}_{nk}^H \mathbf{H}_{nm_j}^H + \mathbf{I}_{N_r}, \quad (36)$$

where the approximation is used due to the assumption that the intra-cell interference as well as the ICI from other interfering BSs than  $\text{BS}_n$  are negligible. By (36), the achievable rate of  $\text{MS}_{m_j}$  is given as

$$r_{m_j} = \log \left| \mathbf{I}_{N_s} + \mathbf{T}_{m_j}^H \mathbf{H}_{mm_j}^H \Psi_{m_j}^{-1} \mathbf{H}_{mm_j} \mathbf{T}_{m_j} \right|. \quad (37)$$

Considering the high Signal-to-Noise Ratio (SNR) regime, the rate function of  $\text{MS}_{m_j}$  (37) can be approximated as

$$\begin{aligned} r_{m_j} &\approx \log \left| \mathbf{T}_{m_j}^H \mathbf{H}_{mm_j}^H \Psi_{m_j}^{-1} \mathbf{H}_{mm_j} \mathbf{T}_{m_j} \right| \\ &= \log \left| \mathbf{T}_{m_j}^H \mathbf{H}_{mm_j}^H \mathbf{H}_{mm_j} \mathbf{T}_{m_j} \right| - \log \left| \Psi_{m_j} \right|. \end{aligned} \quad (38)$$

From (38), it can be noticed that the second term in the right-hand side represents the performance loss of  $\text{MS}_{m_j}$  in terms of rate due to the beamforming at  $\text{BS}_n$ . An important point to observe is that this rate loss at  $\text{MS}_{m_j}$  is already known to  $\text{BS}_n$ , as it denotes the interference leakage from  $\text{BS}_n$  to  $\text{MS}_{m_j}$ . Therefore,  $\text{BS}_n$  can consider an implicit interference pricing approach to reduce the interference leakage towards the MSs of other cells. Let  $\Psi_n = \{\Psi_{m_j}, \forall m \in \mathcal{M} \setminus n, \forall j \in \mathcal{K}_m\}$  and define the following  $\text{BS}_n$  specific function

$$g_n(\Psi_n) = \sum_{k \in \mathcal{K}_n} \mu_{nk} r_{nk} - \sum_{m \in \mathcal{M} \setminus n} \sum_{j \in \mathcal{K}_m} \log \left| \Psi_{m_j} \right|. \quad (39)$$

Using (39), each  $\text{BS}_n, \forall n \in \mathcal{M}$ , updates the transmit beamforming  $\mathbf{T}_{nk}, \forall k \in \mathcal{K}_n$ , as solution to the following optimization problem

$$\begin{aligned} &\max_{\substack{\mathbf{T}_{nk}, \\ \forall k \in \mathcal{K}_n}} g_n(\Psi_n) && (\mathcal{P}_{\text{WSRH}}) \\ &\text{s.t.} \sum_{k \in \mathcal{K}_n} \text{Tr}(\mathbf{T}_{nk} \mathbf{T}_{nk}^H) = p_n. \end{aligned}$$

The solution of  $\mathcal{P}_{\text{WSRH}}$  w.r.t. transmit beamforming for  $\text{MS}_{nk}, \forall n \in \mathcal{M}, \forall k \in \mathcal{K}_n$ , is given by Proposition 3.

*Proposition 3:* Let the receive beamforming  $\mathbf{R}_{nk}$  and MSE matrix  $\mathbf{E}_{nk}$  for  $\text{MS}_{nk}$  be given by (19) and (20), respectively, and by utilizing the Lemma 1 result, then the solution to  $\mathcal{P}_{\text{WSRH}}$  w.r.t. transmit beamforming for  $\text{MS}_{nk}, \forall n \in \mathcal{M}, \forall k \in \mathcal{K}_n$ , is given as

$$\mathbf{T}_{nk}^{\text{WSRH}} = (\mathbf{A}_{nk} + \mathbf{D}_n + \lambda_n \mathbf{I}_{N_t})^{-1} \mathbf{H}_{nnk}^H \mathbf{R}_{nk}^H \mu_{nk}, \quad (40)$$

where  $\mathbf{A}_{nk}$  is given by (30) and

$$\mathbf{D}_n = \sum_{m \in \mathcal{M} \setminus n} \sum_{j \in \mathcal{K}_m} \mathbf{H}_{nm_j}^H \Psi_{m_j}^{-1} \mathbf{H}_{nm_j}. \quad (41)$$

*Proof 7:* The derivation steps are similar to the ones shown in Appendix C and thus omitted here for brevity.

Similar to (29),  $\lambda_n, \forall n \in \mathcal{M}$ , in (40) are calculated to satisfy the power constraint at  $\text{BS}_n, \forall n \in \mathcal{M}$ . The closed-form solutions of (40) can be obtained similar to (32), by replacing the  $\mathbf{B}_n$  term with  $\mathbf{D}_n$  given by (41).

#### D. Algorithm Design and Convergence Analysis

To solve either problem of  $\mathcal{P}_{\text{WSRP}}, \mathcal{P}_{\text{WSRM}}$ , or  $\mathcal{P}_{\text{WSRH}}$ , an algorithm based on alternating optimization can be used [21]–[25]. The basic idea is to optimize each problem with respect to one variable at a time, while keeping the rest of the variables fixed. The proposed algorithm to solve either optimization problem is summarized in Algorithm 2. We refer to this algorithm as WSRP when solving  $\mathcal{P}_{\text{WSRP}}$ , as WSRM when solving  $\mathcal{P}_{\text{WSRM}}$ , and as WSRH when solving  $\mathcal{P}_{\text{WSRH}}$ .

---

#### Algorithm 2 WSR Max. via Alternate Optimization.

---

- 1: Initialize  $\mathbf{T}_{nk}^{(1)}, \forall k \in \mathcal{K}$ .
  - 2:  $\text{MS}_{nk}, \forall k$ : Calculates  $\mathbf{R}_{nk}^{(t)}$  using (19).
  - 3:  $\text{MS}_{nk}, \forall k$ : Calculates  $\mathbf{E}_{nk}^{(t)}$  using (20).
  - 4: **if**  $\mathcal{P}_{\text{WSRP}}$  **then**
  - 5:      $\text{BS}_n, \forall n$ : Updates  $\mathbf{T}_{nk}^{(t+1)}, \forall k \in \mathcal{K}_n$ , using (29).
  - 6: **end if**
  - 7: **if**  $\mathcal{P}_{\text{WSRM}}$  **then**
  - 8:      $\text{BS}_n, \forall n$ : Updates  $\mathbf{T}_{nk}^{(t+1)}, \forall k \in \mathcal{K}_n$ , using (33).
  - 9: **end if**
  - 10: **if**  $\mathcal{P}_{\text{WSRH}}$  **then**
  - 11:      $\text{BS}_n, \forall n$ : Updates  $\mathbf{T}_{nk}^{(t+1)}, \forall k \in \mathcal{K}_n$ , using (40).
  - 12: **end if**
  - 13: Repeat steps 2-12 (until convergence)
- 

In step 1, Algorithm 2 initializes the transmit beamforming matrices for all users in the system by any means. Afterwards, the algorithm alternates between the following three steps. In steps 2 and 3, all MSs calculate, in parallel, their receive beamforming and MSE matrices, respectively, for the given transmit beamforming. Next, all BSs calculate, in parallel, their transmit beamforming, using either approach, for the given receive beamforming and MSE matrices. If this iterative process converges, it converges to a fixed point that is a stationary point of the WSR-objective function [20]. It is worth noting that in a single-cell scenario, all proposed algorithms coincide, where each algorithm differs from the other two in the ICI handling.

For alternating optimization, monotonic convergence of the objective to a stationary (locally optimal) point is guaranteed, if each step has a unique optimum [33, Proposition 2.7.1]. The requirement for the transmit beamforming optimization to be unique is that the matrix to be inverted in (29), (33) or (40) is invertible, i.e.,  $(\mathbf{A}_{nk} + \mathbf{B}_n/\mathbf{C}_n/\mathbf{D}_n + \lambda_n \mathbf{I})^{-1}$  does exist. One sufficient condition for invertibility is that all the power constraints are active, i.e.,  $\sum_{k \in \mathcal{K}_n} \text{Tr}(\mathbf{T}_{nk} \mathbf{T}_{nk}^H) = p_n$  so that we always have  $\lambda_n > 0$ , which is the case in the formulation of our problems, since the WSR optimum is reached at maximum transmit power [20]. Another condition is that there are at least  $N_t$  active vectors whose effective channels are linearly independent, i.e.,  $\text{rank}(\mathbf{A}_{nk} + \mathbf{B}_n/\mathbf{C}_n/\mathbf{D}_n) \geq N_t$ . In practice, the cases when the matrix is non-invertible, and the optimal beamforming solution is not unique, are very rare [25]. Nevertheless, if the matrix is not invertible, the pseudo-inverse may be used to get a solution. However, since the original problems  $\mathcal{P}_{\text{WSRP}}, \mathcal{P}_{\text{WSRM}}$ , and  $\mathcal{P}_{\text{WSRH}}$  are non-convex, a globally optimal point cannot be found, in general,

via alternating optimization. Moreover, different initializations and iteration orders may converge to different local WSR-optima [21]–[25].

## V. WSR MAXIMIZATION SIGNALING SCHEMES

The proposed algorithms above are decentralized, where each BS can calculate its transmit beamforming locally once it has the required information. Here, we assume that each  $\text{BS}_n, \forall n$ , has access to the local CSI, i.e.,  $\mathbf{H}_{nmj}, \forall m, \forall j$ , as in [23]–[25]. TDD operation is an effective technique for obtaining this CSI, where uplink training in conjunction with reciprocity provides the BSs with downlink and uplink CSI simultaneously [27], [28]. In the following, we propose a novel OTA signaling scheme based on TDD mode to facilitate the algorithm implementation.

We assume that 1) each BS and MS has orthogonal pilot symbols (training) in the downlink and uplink direction, respectively, for the OTA signaling, 2) each TDD frame is divided into two parts; signaling and data parts, as shown in Fig. 2, where the signaling part is further divided into downlink and uplink sub-parts to facilitate the variables' exchange between BSs and MSs, and 3) all exchanged variables are globally estimated at each iteration.

At the downlink, we assume that each  $\text{BS}_n$  transmits pilot signals that are precoded with the transmit beamforming  $\mathbf{T}_{nk}, \forall k \in \mathcal{K}_n$ . Thus, each  $\text{MS}_{nk}$  can estimate the downlink equivalent channels  $\mathbf{H}_{mnk}\mathbf{T}_{mj}, \forall m, \forall j$ , and update its receive beamforming  $\mathbf{R}_{nk}$  and MSE-matrix  $\mathbf{E}_{nk}$ . On the other hand, to update the transmit beamforming, each algorithm has different signaling needs. Therefore, we propose the following two signaling schemes.

### A. Signaling Scheme A

In this scheme, we assume that each  $\text{MS}_{nk}$  transmits an uplink pilot signal that is precoded with the receive beamforming  $\mathbf{R}_{nk}$ . Thus, each  $\text{BS}_n$  can estimate the uplink equivalent channels  $\mathbf{R}_{mj}\mathbf{H}_{nmj}, \forall m, \forall j$ , and calculate  $\mathbf{E}_{nk}, \forall k \in \mathcal{K}_n$ , using (20). This information is sufficient to calculate  $\mathbf{A}_{nk}, \forall k \in \mathcal{K}_n$ , which is common in the three algorithms.

For WSRH,  $\mathbf{A}_{nk}, \forall k \in \mathcal{K}_n$ , is all that is needed to update the transmit beamforming  $\mathbf{T}_{nk}, \forall k \in \mathcal{K}_n$ , where the second term  $\mathbf{D}_n$  can be calculated locally. However, for WSRP, each  $\text{BS}_n$  would require vector  $\boldsymbol{\pi}_n$  to calculate  $\mathbf{B}_n$ , which collects all interference prices from the users of other cells. The direct approach, as assumed in [18], is to let each  $\text{MS}_{nk}$  calculate its interference price  $\pi_{nk}$  and feed it back to its serving BS, i.e.,  $\text{BS}_n$ . Then, all BSs perform broadcast-and-gather operation of their interference prices using the backhaul. Different from [18], we assume that each  $\text{BS}_n$  first recalculates the receive beamforming as  $\mathbf{R}_{nk} = (\mathbf{I} - \mathbf{E}_{nk})(\mathbf{H}_{nnk}\mathbf{T}_{nk})^{-1}, \forall k \in \mathcal{K}_n$ , using local information (see Appendix B), and then calculates the interference prices  $\pi_{nk} = \text{Tr}[\mathbf{R}_{nk}^H \mathbf{E}_{nk}^{-1} \mathbf{R}_{nk}], \forall k \in \mathcal{K}_n$ , and exchanges them with the other BSs using the backhaul. Thus, we do not need the feedback from the MSs. On the other hand, for WSRM, each  $\text{BS}_n$  would require  $\mu_{mj}\mathbf{E}_{mj}, \forall j \in \mathcal{K} \setminus \mathcal{K}_n$ , to calculate  $\mathbf{C}_n$ . Using this signaling scheme, one possible way,

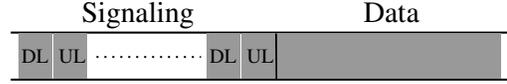


Fig. 2. Considered TDD frame structure.

as assumed in [25], is to let the BSs exchange them using the backhaul.

From above, we can see that signaling Scheme A is best applicable to WSRH, since no further variables feedback is required. However, for WSRM (WSRP), the feedback of matrices (scalars) between BSs is required. To reduce the signaling overhead of WSRM and WSRP, we further propose the following signaling scheme.

### B. Signaling Scheme B

In this scheme, we assume that each  $\text{MS}_{nk}$  transmits a pilot signal that is precoded with  $\sqrt{\mu_{nk}}\mathbf{E}_{nk}^{-\frac{1}{2}}\mathbf{R}_{nk}$  [23], [25]. Thus, each  $\text{BS}_n$  can estimate the uplink equivalent channels  $\sqrt{\mu_{mj}}\mathbf{E}_{mj}^{-\frac{1}{2}}\mathbf{R}_{mj}\mathbf{H}_{nmj}, \forall m, \forall j$ , which are sufficient to calculate  $\mathbf{A}_{nk}, \forall k \in \mathcal{K}_n$ , and  $\mathbf{B}_n$  or  $\mathbf{C}_n$ . However, with WSRM and WSRP, each  $\text{BS}_n$  still requires  $\mathbf{R}_{nk}, \forall k \in \mathcal{K}_n$ , to calculate the uplink equivalent channels  $\mu_{nk}\mathbf{R}_{nk}\mathbf{H}_{nnk}, \forall k \in \mathcal{K}_n$ . One possible way, as proposed in [23], is to let each  $\text{MS}_{nk}$  transmit two consecutive uplink pilot signals; one precoded with  $\sqrt{\mu_{nk}}\mathbf{E}_{nk}^{-\frac{1}{2}}\mathbf{R}_{nk}$  and another precoded with  $\mathbf{R}_{nk}$ . However, this approach would unnecessarily increase the signaling overhead. In the following, we propose an alternative approach, where the main idea is to let each  $\text{BS}_n$  recalculate  $\mathbf{R}_{nk}, \forall k \in \mathcal{K}_n$ , using only local information and thus reduce the signaling overhead.

Let  $\mathbf{X}_{nk} = \sqrt{\mu_{nk}}\mathbf{E}_{nk}^{-\frac{1}{2}}\mathbf{R}_{nk}\mathbf{H}_{nnk}$  denote the uplink equivalent channel with  $\text{MS}_{nk}$  estimated at  $\text{BS}_n$ . Substitute  $\mathbf{R}_{nk} = (\mathbf{I} - \mathbf{E}_{nk})(\mathbf{H}_{nnk}\mathbf{T}_{nk})^{-1}$  into  $\mathbf{X}_{nk}$  and simplify the resulting expression, then we have

$$\mathbf{X}_{nk} = (\sqrt{\mu_{nk}}\mathbf{E}_{nk}^{-\frac{1}{2}} - \sqrt{\mu_{nk}}\mathbf{E}_{nk}^{\frac{1}{2}})\mathbf{Y}_{nk}, \quad (42)$$

where  $\mathbf{Y}_{nk} = (\mathbf{H}_{nnk}\mathbf{T}_{nk})^{-1}\mathbf{H}_{nnk}$ . Right multiply both sides of the latter equation by  $\mathbf{Y}_{nk}^\dagger$  and again simplify the resulting expression, then we have

$$\mathbf{E}_{nk}^{-\frac{1}{2}} = \frac{1}{\sqrt{\mu_{nk}}}\mathbf{M}_{nk} + \mathbf{E}_{nk}^{\frac{1}{2}}, \quad (43)$$

where  $\mathbf{M}_{nk} = \mathbf{X}_{nk}\mathbf{Y}_{nk}^\dagger$  (where  $\mathbf{Y}_{nk}^\dagger = \mathbf{Y}_{nk}^H(\mathbf{Y}_{nk}\mathbf{Y}_{nk}^H)^{-1}$ ), which is formed using local information. Then (43) can be solved for  $\mathbf{E}_{nk}$  iteratively as given by Algorithm 3.

---

#### Algorithm 3 Recalculating $\mathbf{R}_{nk}$ at the BS for Scheme B.

---

- 1: Construct  $\mathbf{M}_{nk}$  and initialize  $\mathbf{E}_{nk}^{(1)}$  randomly.
  - 2: Set  $\bar{\mathbf{E}}_{nk}^{(t)} = \frac{1}{\sqrt{\mu_{nk}}}\mathbf{M}_{nk} + \mathbf{E}_{nk}^{\frac{1}{2}(t)}$ .
  - 3: Set  $\mathbf{E}_{nk}^{(t+1)} = \bar{\mathbf{E}}_{nk}^{-2(t)}$ .
  - 4: Repeat steps 2-3 (until convergence).
  - 5: At convergence, solve  $\mathbf{R}_{nk} = (\mathbf{I} - \mathbf{E}_{nk})(\mathbf{H}_{nnk}\mathbf{T}_{nk})^{-1}$ .
- 

In step 1, Algorithm 3 constructs the local matrix  $\mathbf{M}_{nk}$  and randomly initializes  $\mathbf{E}_{nk}^{(1)}$ . Given those initial matrices, the

algorithm alternates between steps 2 and 3 at each iteration. At the  $t$ -th iteration, the algorithm solves for  $\mathbf{E}_{nk}^{-\frac{1}{2}(t)}$  given  $\mathbf{E}_{nk}^{\frac{1}{2}(t)}$  at step 2 (denoted as  $\bar{\mathbf{E}}_{nk}^{(t)}$ ). Then, at step 3, the algorithm solves for  $\mathbf{E}_{nk}^{(t+1)}$  given  $\bar{\mathbf{E}}_{nk}^{(t)}$ . Those two steps are repeated until convergence. If matrix  $\mathbf{M}_{nk}$  is assumed perfect, the algorithm is able to recalculate the MSE-matrix  $\mathbf{E}_{nk}$  perfectly. Then, using  $\mathbf{E}_{nk}$  at step 5 we can calculate for the receive beamforming matrix  $\mathbf{R}_{nk}$ . Algorithm 3 convergence behavior is shown numerically in the next section. A proof of convergence is open and we leave it for a future work.

## VI. NUMERICAL RESULTS

In this section we evaluate the performance of the proposed algorithms by means of simulation. We consider a flat Rayleigh fading scenario with uncorrelated channels between antennas, i.e., each element of  $\mathbf{H}_{nnk}, \forall n \in \mathcal{M}, \forall k \in \mathcal{K}_n$ , is an i.i.d. complex Gaussian random variable with zero mean and unit variance. For each simulated algorithm, we initialize the transmit beamforming matrices using the MRT approach, i.e.,  $\mathbf{T}_{nk}^{(1)} = [\mathbf{V}_{nk}]_{[1:N_s]} \sqrt{p_n/KN_s}, \forall n \in \mathcal{M}, \forall k \in \mathcal{K}_n$ , where  $\mathbf{V}_{nk}$  denotes the matrix holding in its columns the right singular vectors of  $\mathbf{H}_{nnk}$  arranged in a decreasing order w.r.t their singular values. Moreover, we assume that the noise variance  $\sigma_{nk}^2 = \sigma^2 = 1$ .

For comparison, we show simulation results of the WSR-WMMSE algorithm from [21]–[25], MRT, and the Matrix Orthogonal Projection (MOP) approach [34]. For the MOP approach, the transmit beamforming for  $\mathbf{MS}_{nk}$  is given as

$$\mathbf{T}_{nk}^{\text{MOP}} = [\mathbf{t}_1^{\text{MOP}} \dots \mathbf{t}_{N_s}^{\text{MOP}}] \sqrt{\mathbf{P}_{nk}}, \quad (44)$$

where  $\mathbf{t}_i^{\text{MOP}} = \frac{\mathbf{\Pi}_{nk}[\mathbf{V}_{nk}]_{[i]}}{\|\mathbf{\Pi}_{nk}[\mathbf{V}_{nk}]_{[i]}\|}$ ,  $\mathbf{P}_{nk}$  is an  $[N_s \times N_s]$  diagonal matrix holding the  $N_s$  power allocations found using the water-filling method [30] over the largest  $N_s$  singular values,  $\mathbf{\Pi}_{nk} = \mathbf{I}_{N_t} - \mathbf{H}_{nk}^{-k} ((\mathbf{H}_{nk}^{-k})^H \mathbf{H}_{nk}^{-k})^{-1} (\mathbf{H}_{nk}^{-k})^H$ , whereas  $\mathbf{H}_{nk}^{-k} = [\{\mathbf{H}_{nmj}^H, \forall j \in \mathcal{K} \setminus k\}]$  denotes all channel matrices from BS $_n$  to all MSs except MS $_{nk}$ .

### A. Example 1: Algorithm 3 Convergence

In this example, we show simulation results to evaluate the convergence behavior of Algorithm 3. Fig. 3 shows the log-scale convergence results of Algorithm 3 for the first user, i.e., MS $_{11}$ , in terms of the Absolute Error that is defined as

$$\text{Absolute Error} = \left| \sum_{i,j} (\mathbf{E}_{11}^{(t)} - \mathbf{E}_{11}^*) \right|, \quad (45)$$

where  $\mathbf{E}_{11}^*$  is the perfect MSE-matrix of MS $_{11}$  and  $\mathbf{E}_{11}^{(t)}$  is the obtained MSE-matrix at the  $t$ -th iteration. Each simulated point is averaged over 1,000 channel realizations.

From Fig. 3, it can be seen that Algorithm 3 has a fast convergence rate, where it is able to obtain the perfect MSE-matrix using a few iterations. Note that, when  $N_t$  increases,  $N_s$  decreases, and  $MK$  decreases, the algorithm has faster convergence rate. We note that all simulated channel realizations have converged to the perfect MSE-matrix, although the convergence of some channel realizations is not necessarily monotonic.

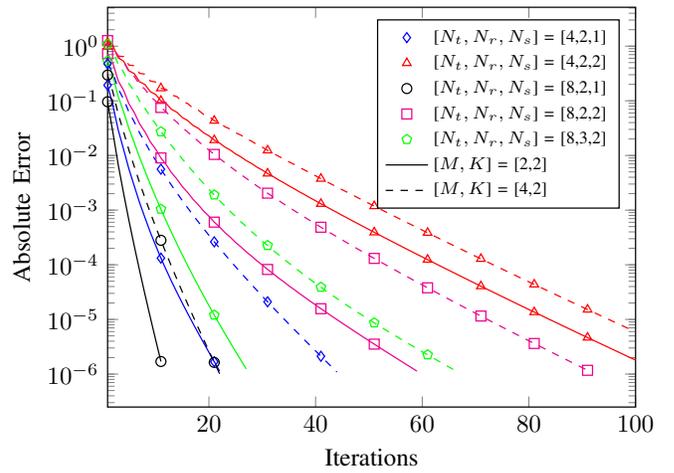


Fig. 3. Algorithm 3 Log-scale convergence behavior [ $M = 2, K = 2$ ].

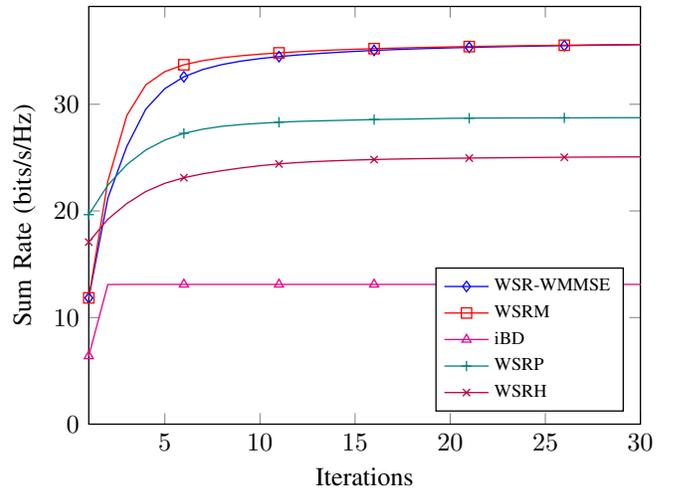


Fig. 4. Convergence behavior of Algorithms 1 and 2, SNR = 10 dB and  $[M, K, N_t, N_r, N_s] = [3, 3, 9, 2, 2]$ .

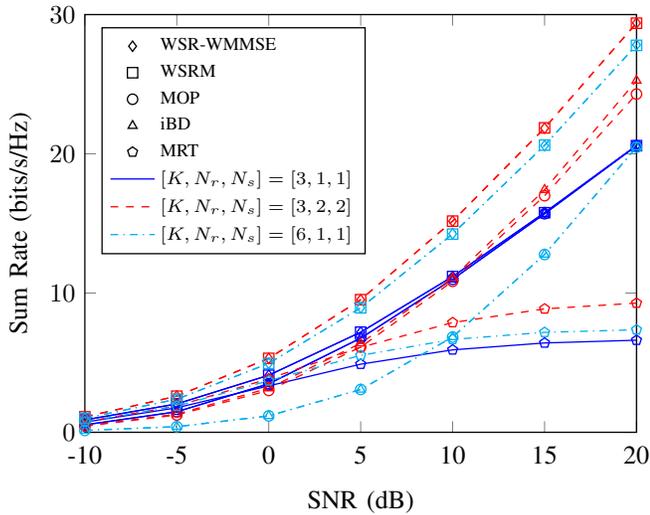
### B. Example 2: Convergence behavior of Algorithms 1 and 2

In this example, we show simulation results to evaluate the convergence behavior of Algorithms 1 and 2. Fig. 4 shows the averaged sum rate convergence results assuming SNR = 10 dB.

From Fig. 4, we can see that all iterative algorithms have a fast convergence rate, within 1-to-2 iterations for iBD and within 10-to-15 iterations for the other algorithms. It's worth noting that WSRM has a slightly faster convergence rate than WSR-WMMSE, although both algorithms seem to converge almost to the same point. However, the convergence speed of either algorithm varies for the individual channel realizations. For some channel realizations, WSRM appears to converge slightly faster and to a higher sum rate than WSR-WMMSE and vice-versa for the other channel realizations.

### C. Example 3: Sum Rate Performance for Single-Cell Case

In this example, we show simulation results to evaluate the sum rate performance of Algorithms 1 and 2 in the single cell case, i.e.,  $M = 1$ . Fig. 5 shows the average sum rate results for a range of SNR values, assuming  $\mu_{nk} = \mu = 1$ . Note


 Fig. 5. Sum rate performance for single cell case  $[M, N_t] = [1, 6]$ .

that, when  $M = 1$ , algorithms WSRM, WSRP, and WSRH are all equivalent, since the  $\mathbf{B}_n$ ,  $\mathbf{C}_n$  and  $\mathbf{D}_n$  terms are all identity matrices and all algorithms share the  $\mathbf{A}_{nk}$  calculation. Therefore, for this example, we only show WSRM results.

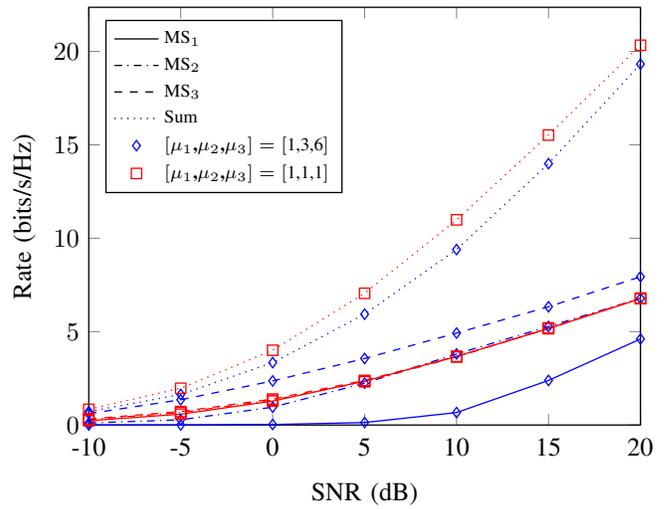
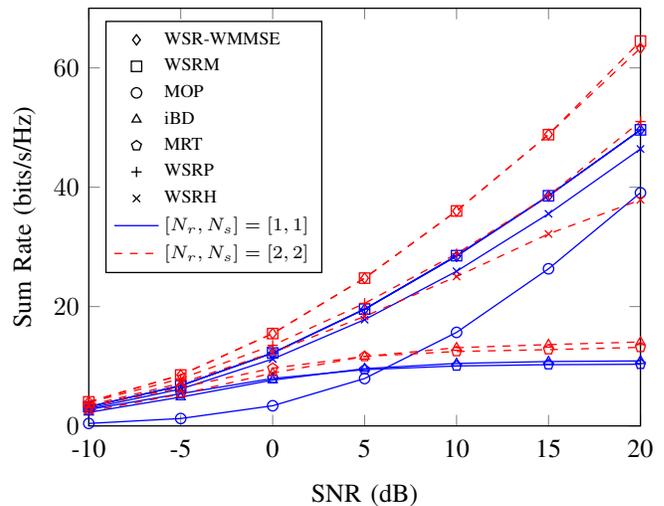
From Fig. 5, it can be seen that when  $[K, N_r, N_s] = [3, 1, 1]$  (solid-lines), both iBD and MOP have very close performance to both WSR-WMMSE and WSRM, which seem to have the same sum rate performance. However, for the other simulated scenarios, when  $K$  and/or  $N_s$  increases while keeping  $N_t$  fixed, both iBD and MOP have large performance loss as compared to WSR-WMMSE and WSRM, although they maintain the same multiplexing gain. None of the iBD, MOP, and MRT algorithms can achieve a good balance between the altruistic and egoistic behaviors of users. While iBD and MOP have complete altruistic behavior, MRT has complete egoistic behavior. Consequently, they have performance loss, as compared to WSR-WMMSE and WSRM.

To examine the impact of user-weights, Fig. 6 shows the simulation results for a system with equal and unequal user-weights. From Fig. 6, it can be seen that when all users have equal weights, they achieve equal performance, in average. However, when a user has a larger weight than others,  $MS_3$  in this case, the algorithm favors him and, thus, achieves better performance. In terms of sum rate performance, the system with equal users' weights has better performance than otherwise. The reason behind this is that when the algorithm favors one user over the others, the user(s) with lower weight would have a degradation in his(their) performance,  $MS_1$  in this case. In general, the increase of one user's rate does not compensate the loss of the other users' rate. Thus, the algorithm would lose in terms of sum rate.

#### D. Example 4: Sum Rate Performance for Multicell Case

In this example, we show simulation results to evaluate the sum rate performance of Algorithms 1 and 2 in the multicell case, i.e.,  $M > 1$ , assuming  $\mu_{nk} = \mu = 1$ . Fig. 7 shows the average sum rate results for a range of SNR values.

From Fig. 7, we can see that WSRM and WSRP have the exact sum rate performance when  $N_r = N_s = 1$  (solid-


 Fig. 6. Sum rate performance with equal and different user-weights,  $[M, K, N_t, N_r, N_s] = [1, 3, 6, 1, 1]$ .

 Fig. 7. Sum rate performance for multicell case,  $[M, K, N_t] = [3, 3, 9]$ .

lines), since both algorithms are equivalent as it is shown in Theorem 2. However, when  $N_r = N_s = 2$ , we can see that WSRP has some performance loss, as compared to WSRM, and the performance loss increases as the SNR value increases. Furthermore, MOP (not feasible when  $N_r = N_s = 2$ ) has the same multiplexing gain as WSRM, but with very large performance loss. On the other hand, both iBD and MRT have a flat performance as the SNR increases, due to severe ICI. For WSRH, we can see that it has close performance to WSRM for the entire SNR range with a small performance loss when  $N_r = N_s = 1$ . However, increasing  $N_s$ , the performance loss increases, as well, since  $N_t$  is fixed. WSRH is a self-pricing algorithm that is distributed between cells. Thus, the algorithm has less transmit coordination than WSRM and WSR-WMMSE.

In Fig. 8, we show sum rate performance while varying the number of MS antennas  $N_r$  and fixing the other parameters. From Fig. 8, we can see that all algorithms have better sum rate performance as  $N_r$  increases. However, WSRP has a much

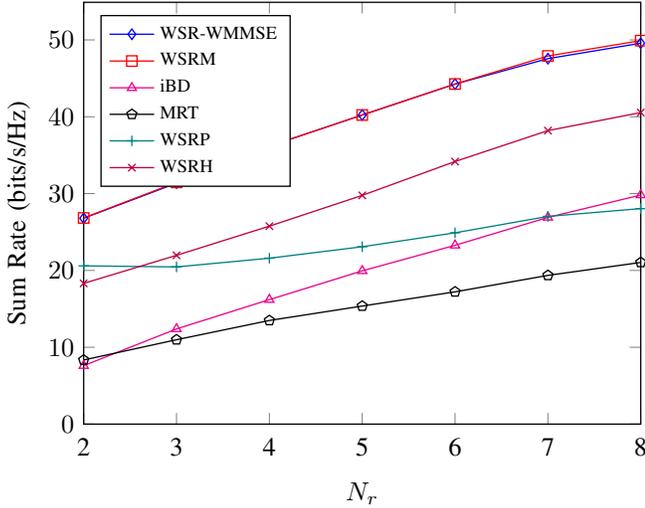


Fig. 8. Sum rate performance for multicell case while varying number of BS antennas  $N_r$ , SNR = 10 dB and  $[M, K, N_t, N_r, N_s] = [3, 3, 6, N_r, 2]$ .

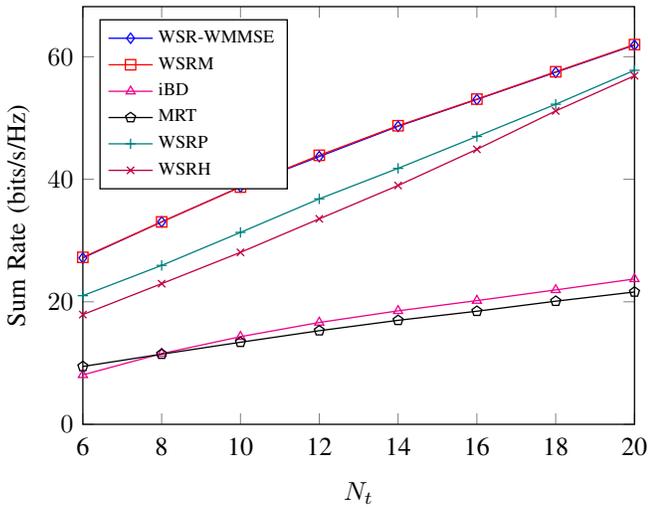


Fig. 9. Sum rate performance for multicell case while varying number of BS antennas  $N_t$ , SNR = 10 dB and  $[M, K, N_t, N_r, N_s] = [3, 3, N_t, 2, 2]$ .

slower increase rate than others, which is translated to a higher rate loss, as compared to WSRM. On the other hand, as  $N_r$  increases, iBD starts to have better sum rate performance than MRT, as compared to results from Fig. 7. The reason behind this is that when  $N_r$  increases, the interference whitening method has better impact on reducing ICI effects, and thus, better performance.

In Fig. 9, we show sum rate performance while varying the number of BS antennas and fixing the other parameters. As discussed above, we can see from Fig. 9 that all algorithms have better sum rate performance as  $N_t$  increases. Different from Fig. 8, WSRP and WSRH sum rate increases as  $N_t$  increases, thus reducing the rate loss as compared to WSRM. Meanwhile, iBD has a slower sum rate increase as compared to results from Fig. 8.

## VII. CONCLUSIONS

We have considered the Weighted Sum-Rate (WSR) maximization problem in multicell MIMO Broadcast Channel (BC)

and proposed three different algorithms, which are based on the alternating optimization technique and are guaranteed to converge to a local WSR-optimum. For all algorithms, the transmit beamforming matrices are obtained by investigating the Karush-Kuhn-Tucker (KKT) conditions of the problems with the help of Lemma 1. In contrast to the WSR-WMMSE algorithm from [21]–[25], which solves the WSR maximization problem indirectly by solving the Weighted Minimum-Mean-Square Error (WMMSE) minimization problem, the proposed algorithms in this paper provide a direct solution to the WSR maximization problem. Using computer simulations, it was shown that the proposed algorithms achieve comparable sum rate performance to the WSR-WMMSE algorithm, while using fewer iterations. Further, it was shown that the network-wide WSR maximization can be equivalently solved using an interference pricing approach if 1) each Mobile Station (MS) is equipped with a single-antenna and 2) the users' weights are included in the interference prices. Furthermore, two different signaling schemes based on Time Division Duplex (TDD) mode were also proposed to facilitate the implementation of the algorithms. Different from existing schemes, the proposed signaling schemes reduce the signaling overhead and require no feedback of variables between Base Stations (BSs).

## APPENDIX

### A. Proof of Theorem 1

At first, one can note that  $\tilde{\mathbf{H}}_{nk}^{-k}$  given by (14) can be written in function of  $\hat{\mathbf{H}}_{nk}^{-k}$  given by (11). To show this, let us define  $\tilde{\mathbf{R}}_n^{-k} = \mathbf{B} \text{diag}[\tilde{\mathbf{R}}_{nj}, \forall j \in \mathcal{K}_n \setminus k]$ . Then we can write  $\tilde{\mathbf{H}}_{nk}^{-k} = \tilde{\mathbf{R}}_n^{-k} \hat{\mathbf{H}}_{nk}^{-k}$ . Note that  $\tilde{\mathbf{R}}_n^{-k}$  is an orthogonal unitary matrix. Therefore, if  $N_s = N_r$ , then we have  $\tilde{\mathbf{R}}_n^{\text{H}-k} \tilde{\mathbf{R}}_n^{-k} = \mathbf{I}$ , otherwise, if  $N_s < N_r$ , then  $\tilde{\mathbf{R}}_n^{\text{H}-k} \tilde{\mathbf{R}}_n^{-k} \neq \mathbf{I}$ . Assuming  $N_s = N_r$ , we have

$$\tilde{\mathbf{H}}_{nk}^{\text{H}-k} \tilde{\mathbf{H}}_{nk}^{-k} = \hat{\mathbf{H}}_{nk}^{\text{H}-k} \tilde{\mathbf{R}}_n^{\text{H}-k} \tilde{\mathbf{R}}_n^{-k} \hat{\mathbf{H}}_{nk}^{-k} = \hat{\mathbf{H}}_{nk}^{\text{H}-k} \hat{\mathbf{H}}_{nk}^{-k}.$$

Therefore, we have  $\tilde{\mathbf{H}}_{nk}^{-i} \propto \hat{\mathbf{H}}_{nk}^{-i}$ . This end result proves that both matrices are proportional to each other. Consequently, their individual null spaces are also proportional to each other, i.e.,  $\tilde{\mathbf{G}}_{nk} \propto \hat{\mathbf{G}}_{nk}$ . Therefore, the singular values calculated using (7) assuming  $\tilde{\mathbf{H}}_{nk}^e$  given by (15) are exactly equal to the ones calculated assuming  $\hat{\mathbf{H}}_{nk}^e$  given by (12), which completes the proof.

### B. Proof of Lemma 1

In the lemma, we claim that  $\mathbf{R}_{nk} = \mathbf{E}_{nk} \mathbf{T}_{nk}^{\text{H}} \mathbf{H}_{nnk}^{\text{H}} \Phi_{nk}^{-1}$ . To prove this, assume  $N_s = N_r$  and solve for  $\mathbf{R}_{nk}$  from (20) as  $\mathbf{R}_{nk} = (\mathbf{I} - \mathbf{E}_{nk})(\mathbf{H}_{nnk} \mathbf{T}_{nk})^{-1}$ . Then, our claim is that

$$\begin{aligned} \mathbf{E}_{nk} \mathbf{T}_{nk}^{\text{H}} \mathbf{H}_{nnk}^{\text{H}} \Phi_{nk}^{-1} &\stackrel{\text{def}}{=} (\mathbf{I} - \mathbf{E}_{nk})(\mathbf{H}_{nnk} \mathbf{T}_{nk})^{-1} \\ \mathbf{E}_{nk} \mathbf{T}_{nk}^{\text{H}} \mathbf{H}_{nnk}^{\text{H}} \Phi_{nk}^{-1} &= (\mathbf{H}_{nnk} \mathbf{T}_{nk})^{-1} - \mathbf{E}_{nk} (\mathbf{H}_{nnk} \mathbf{T}_{nk})^{-1} \\ \mathbf{T}_{nk}^{\text{H}} \mathbf{H}_{nnk}^{\text{H}} \Phi_{nk}^{-1} &\stackrel{(a)}{=} \mathbf{E}_{nk}^{-1} [(\mathbf{H}_{nnk} \mathbf{T}_{nk})^{-1} - \mathbf{E}_{nk} (\mathbf{H}_{nnk} \mathbf{T}_{nk})^{-1}] \\ \mathbf{T}_{nk}^{\text{H}} \mathbf{H}_{nnk}^{\text{H}} \Phi_{nk}^{-1} &\stackrel{(b)}{=} \mathbf{E}_{nk}^{-1} (\mathbf{H}_{nnk} \mathbf{T}_{nk})^{-1} - (\mathbf{H}_{nnk} \mathbf{T}_{nk})^{-1}, \end{aligned}$$

where (a) is obtained by left-multiplying both sides by  $\mathbf{E}_{nk}^{-1}$  and (b) is obtained by simplifying (a). From (21),  $\mathbf{E}_{nk}^{-1} =$

$(\mathbf{I}_{N_r} + \mathbf{T}_{nk}^H \mathbf{H}_{nnk}^H \Phi_{nk}^{-1} \mathbf{H}_{nnk} \mathbf{T}_{nk})$ . Substitute  $\mathbf{E}_{nk}^{-1}$  into (b) and when simplifying the resulting expression we have

$$\mathbf{T}_{nk}^H \mathbf{H}_{nnk}^H \Phi_{nk}^{-1} = \mathbf{T}_{nk}^H \mathbf{H}_{nnk}^H \Phi_{nk}^{-1},$$

which completes the proof.

### C. Proof of Proposition 1

From the KKT conditions, a local optimum must satisfy  $\nabla_{\mathbf{T}_{nk}} \mathcal{L} = 0, \forall k \in \mathcal{K}_n$ , where  $\nabla_{\mathbf{T}_{nk}} \mathcal{L}$  defines the complex gradient operator of  $\mathcal{L}$  with respect to  $\mathbf{T}_{nk}$  and  $\mathcal{L}$  defines the Lagrangian function of problem  $\mathcal{P}_{\text{WSRP}}$ , which is given as

$$\begin{aligned} \mathcal{L} &= \sum_{k \in \mathcal{K}_n} f_{nk}(\boldsymbol{\pi}_n) - \lambda_n \left( \sum_{k \in \mathcal{K}_n} \text{Tr}(\mathbf{T}_{nk} \mathbf{T}_{nk}^H) - p_n \right) \\ &= \sum_{k \in \mathcal{K}_n} (\mu_{nk} r_{nk} - \mathbf{L}_{nk}) - \lambda_n \left( \sum_{k \in \mathcal{K}_n} \text{Tr}(\mathbf{T}_{nk} \mathbf{T}_{nk}^H) - p_n \right). \end{aligned} \quad (46)$$

The gradient is a matrix with the  $[p, q]$ -th element defined as  $[\nabla_{\mathbf{T}_{nk}} \mathcal{L}]_{[p, q]} = \nabla_{[\mathbf{T}_{nk}]_{[p, q]}} \mathcal{L}$ . In order to calculate  $\nabla_{\mathbf{T}_{nk}} \mathcal{L}$ , we need to calculate first  $\nabla_{\mathbf{T}_{nk}} r_{nk}, \nabla_{\mathbf{T}_{nk}} r_{nj}, \forall j \in \mathcal{K}_n \setminus k$ , and  $\nabla_{\mathbf{T}_{nk}} \mathbf{L}_{nk}$  by utilizing the following results from [32]:  $\nabla \log |\mathbf{X}| = \text{Tr}(\mathbf{X}^{-1} \nabla \mathbf{X})$  and  $\nabla(\mathbf{X}^{-1}) = -\mathbf{X}^{-1}(\nabla \mathbf{X})\mathbf{X}^{-1}$ , where  $\mathbf{X}$  is a matrix.

First,  $\nabla_{[\mathbf{T}_{nk}]_{[p, q]}} r_{nk} = \text{Tr}(\mathbf{E}_{nk} \nabla_{[\mathbf{T}_{nk}]_{[p, q]}} \mathbf{E}_{nk}^{-1})$ . Here,  $\nabla_{[\mathbf{T}_{nk}]_{[p, q]}} \mathbf{E}_{nk}^{-1} = \mathbf{e}_p \mathbf{e}_p^H \mathbf{H}_{nnk}^H \Phi_{nk}^{-1} \mathbf{H}_{nnk} \mathbf{T}_{nk}$ , where  $\mathbf{e}_p$  ( $\mathbf{e}_q$ ) is a vector of  $N_t$  ( $N_r$ ) dimension with one at the  $p$  ( $q$ )-th element and zeros elsewhere. Then, we have

$$\begin{aligned} \nabla_{[\mathbf{T}_{nk}]_{[p, q]}} r_{nk} &= \text{Tr}(\mathbf{E}_{nk} \mathbf{e}_p \mathbf{e}_p^H \mathbf{H}_{nnk}^H \Phi_{nk}^{-1} \mathbf{H}_{nnk} \mathbf{T}_{nk}) \\ &= \mathbf{e}_p^H \mathbf{H}_{nnk}^H \Phi_{nk}^{-1} \mathbf{H}_{nnk} \mathbf{T}_{nk} \mathbf{E}_{nk} \mathbf{e}_p. \end{aligned}$$

Since  $\nabla_{[\mathbf{T}_{nk}]_{[p, q]}} r_{nk} = [\nabla_{\mathbf{T}_{nk}} r_{nk}]_{[p, q]}$ , then

$$\nabla_{\mathbf{T}_{nk}} r_{nk} = \mathbf{H}_{nnk}^H \Phi_{nk}^{-1} \mathbf{H}_{nnk} \mathbf{T}_{nk} \mathbf{E}_{nk}. \quad (47)$$

Furthermore,  $\nabla_{[\mathbf{T}_{nk}]_{[p, q]}} r_{nj}, \forall j \in \mathcal{K}_n \setminus k$ , is given as

$$\nabla_{[\mathbf{T}_{nk}]_{[p, q]}} r_{nj} = \text{Tr}(\mathbf{E}_{nj} \nabla_{[\mathbf{T}_{nk}]_{[p, q]}} \mathbf{E}_{nj}^{-1}).$$

First,  $\nabla_{[\mathbf{T}_{nk}]_{[p, q]}} \mathbf{E}_{nj}^{-1}$  is given as

$$\begin{aligned} \nabla_{[\mathbf{T}_{nk}]_{[p, q]}} \mathbf{E}_{nj}^{-1} \\ = \mathbf{T}_{nj}^H \mathbf{H}_{nnj}^H [-\Phi_{nj}^{-1} [\nabla_{[\mathbf{T}_{nk}]_{[p, q]}} \Phi_{nj}] \Phi_{nj}^{-1}] \mathbf{H}_{nnj} \mathbf{T}_{nj}, \end{aligned}$$

whereas  $\nabla_{[\mathbf{T}_{nk}]_{[p, q]}} \Phi_{nj} = \mathbf{H}_{nnj} \mathbf{T}_{nk} \mathbf{e}_q \mathbf{e}_q^H \mathbf{H}_{nnj}^H$ . Combining all results together, we have

$$\begin{aligned} \nabla_{[\mathbf{T}_{nk}]_{[p, q]}} r_{nj} \\ = \text{Tr}(\mathbf{E}_{nj} \mathbf{T}_{nj}^H \mathbf{H}_{nnj}^H [-\Phi_{nj}^{-1} \mathbf{H}_{nnj} \mathbf{T}_{nk} \mathbf{e}_q \mathbf{e}_q^H \mathbf{H}_{nnj}^H \Phi_{nj}^{-1}] \mathbf{H}_{nnj} \mathbf{T}_{nj}) \\ = -\mathbf{e}_p^H \mathbf{H}_{nnj}^H \Phi_{nj}^{-1} \mathbf{H}_{nnj} \mathbf{T}_{nj} \mathbf{E}_{nj} \mathbf{T}_{nj}^H \mathbf{H}_{nnj}^H \Phi_{nj}^{-1} \mathbf{H}_{nnj} \mathbf{T}_{nk} \mathbf{e}_q. \end{aligned}$$

Therefore,  $\nabla_{\mathbf{T}_{nk}} r_{nj}, \forall j \in \mathcal{K}_n \setminus k$ , is given as

$$\begin{aligned} \nabla_{\mathbf{T}_{nk}} r_{nj} \\ = -\mathbf{H}_{nnj}^H \Phi_{nj}^{-1} \mathbf{H}_{nnj} \mathbf{T}_{nj} \mathbf{E}_{nj} \mathbf{T}_{nj}^H \mathbf{H}_{nnj}^H \Phi_{nj}^{-1} \mathbf{H}_{nnj} \mathbf{T}_{nk}. \end{aligned}$$

Finally,  $\nabla_{\mathbf{T}_{nk}} \mathbf{L}_{nk}$  is given as

$$\nabla_{\mathbf{T}_{nk}} \mathbf{L}_{nk} = \sum_{m \in \mathcal{M} \setminus n} \sum_{j \in \mathcal{K}_m} \mu_{mj} \mathbf{H}_{nmj}^H \mathbf{H}_{nmj} \mathbf{T}_{nk}.$$

From above,  $\nabla_{\mathbf{T}_{nk}} \mathcal{L}$  is given as

$$\begin{aligned} \nabla_{\mathbf{T}_{nk}} \mathcal{L} &= \mu_{nk} \mathbf{H}_{nnk}^H \Phi_{nk}^{-1} \mathbf{H}_{nnk} \mathbf{T}_{nk} \mathbf{E}_{nk} - \\ &\quad \tilde{\mathbf{A}}_{nk} \mathbf{T}_{nk} - \mathbf{B}_n \mathbf{T}_{nk} - \lambda_n \mathbf{T}_{nk}, \end{aligned} \quad (48)$$

where

$$\begin{aligned} \tilde{\mathbf{A}}_{nk} &= \sum_{i \in \mathcal{K}_n \setminus k} \mu_{ni} \mathbf{H}_{nni}^H \Phi_{ni}^{-1} \mathbf{H}_{nni} \mathbf{T}_{ni} \mathbf{E}_{ni} \mathbf{T}_{ni}^H \mathbf{H}_{nni}^H \Phi_{ni}^{-1} \mathbf{H}_{nni}, \\ \mathbf{B}_n &= \sum_{m \in \mathcal{M} \setminus n} \sum_{j \in \mathcal{K}_m} \mu_{mj} \mathbf{H}_{nmj}^H \mathbf{H}_{nmj}. \end{aligned}$$

From (48), it can be seen that it is not possible to solve for  $\mathbf{T}_{nk}$  directly. However, according to Lemma 1, we can write  $\mathbf{R}_{nk}^H = \Phi_{nk}^{-1} \mathbf{H}_{nnk} \mathbf{T}_{nk} \mathbf{E}_{nk}$  (note that  $\mathbf{E}_{nk} = \mathbf{E}_{nk}^H$ ). Then, the gradient function (48) can be written as

$$\nabla_{\mathbf{T}_{nk}} \mathcal{L} = \mu_{nk} \mathbf{H}_{nnk}^H \mathbf{R}_{nk}^H - \mathbf{A}_{nk} \mathbf{T}_{nk} - \mathbf{B}_n \mathbf{T}_{nk} - \lambda_n \mathbf{T}_{nk}, \quad (49)$$

where

$$\mathbf{A}_{nk} = \sum_{i \in \mathcal{K}_n \setminus k} \mu_{ni} \mathbf{H}_{nni}^H \mathbf{R}_{ni}^H \mathbf{E}_{ni}^{-1} \mathbf{R}_{ni} \mathbf{H}_{nni}.$$

From (49), we can solve for  $\mathbf{T}_{nk}$  directly as

$$\mathbf{T}_{nk} = (\mathbf{A}_{nk} + \mathbf{B}_n + \lambda_n \mathbf{I}_{N_t})^{-1} \mathbf{H}_{nnk}^H \mathbf{R}_{nk}^H \mu_{nk}. \quad (50)$$

Thus, we have the result given in proposition 1.

### ACKNOWLEDGMENTS

This work was supported by the Innovation Center, Ericsson Telecomunicações S.A., Brazil, under EDB/UFC.42 and EDB/UFC.45 Technical Cooperation Contracts. Khaled Ardah would like to acknowledge CAPES for the scholarship support. Yuri C. B. Silva and Francisco R. P. Cavalcanti acknowledge the support of CNPq.

### REFERENCES

- [1] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003, doi: 10.1109/TIT.2003.810646.
- [2] E. T. Ar and I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, pp. 585–595, Jun 1999, doi: 10.1002/ett.4460100604.
- [3] H. Weingarten, Y. Steinberg, and S. S. Shamai, "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3936–3964, Sep 2006, doi: 10.1109/TIT.2006.880064.
- [4] K. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 461–471, Feb 2004, doi: 10.1109/TSP.2003.821107.
- [5] S. Shim, J. S. Kwak, R. W. Heath, and J. G. Andrews, "Block diagonalization for multi-user MIMO with other-cell interference," *IEEE Trans. Wireless Commun.*, vol. 7, no. 7, pp. 2671–2681, Jul 2008, doi: 10.1109/TWC.2008.070093.
- [6] K. Ardah, Y. C. B. Silva, and F. R. P. Cavalcanti, "Block diagonalization for multicell multiuser MIMO systems with other-cell interference," in *Proc. Simpósio Brasileiro de Telecomunicações e Processamento de Sinais (SBTr)*, Sep 2016, pp. 1–5.
- [7] E. Björnson, M. Bengtsson, and B. Ottersten, "Optimal multiuser transmit beamforming: A difficult problem with a simple solution structure [lecture notes]," *IEEE Signal Process. Mag.*, vol. 31, no. 4, pp. 142–148, Jul 2014, doi: 10.1109/MSP.2014.2312183.
- [8] 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, doi: 10.1109/JSAC.2010.101202.

- [9] 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, doi: 10.1109/TWC.2010.05.090936.
- [10] B. Song, R. L. Cruz, and B. D. Rao, "Network duality for multiuser MIMO beamforming networks and applications," *IEEE Trans. Commun.*, vol. 55, no. 3, pp. 618–630, Mar 2007, doi: 10.1109/TCOMM.2006.888889.
- [11] S. Joshi, M. Codreanu, and M. Latva-aho, "Distributed resource allocation for MISO downlink systems via the alternating direction method of multipliers," in *Proc. Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, Nov 2012, pp. 488–493, doi: 10.1109/ACSSC.2012.6489052.
- [12] —, "Distributed SINR balancing for MISO downlink systems via the alternating direction method of multipliers," in *Proc. 11th International Symposium and Workshops on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt)*, May 2013, pp. 318–325.
- [13] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell TDD systems," *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2640–2651, Aug 2011, doi: 10.1109/TWC.2011.060711.101155.
- [14] T. E. Bogale and L. Vandendorpe, "Weighted sum rate optimization for downlink multiuser MIMO systems with per antenna power constraint: Downlink-uplink duality approach," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Mar 2012, pp. 3245–3248, doi: 10.1109/ICASSP.2012.6288607.
- [15] M. Codreanu, A. Tolli, M. Juntti, and M. Latva-aho, "MIMO downlink weighted sum rate maximization with power constraints per antenna groups," in *Proc. IEEE Vehicular Technology Conference - Spring*, Apr 2007, pp. 2048–2052, doi: 10.1109/VETECS.2007.424.
- [16] 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, doi: 10.1109/TWC.2010.04.090553.
- [17] S. H. Park, H. Park, H. Kong, and I. Lee, "New beamforming techniques based on virtual sinr maximization for coordinated multi-cell transmission," *IEEE Trans. Wireless Commun.*, vol. 11, no. 3, pp. 1034–1044, Mar 2012, doi: 10.1109/TWC.2012.011012.110156.
- [18] J. J. E. Garzás, M. Hong, A. Garcia, and A. García-Armada, "Interference pricing mechanism for downlink multicell coordinated beamforming," *IEEE Trans. Commun.*, vol. 62, no. 6, pp. 1871–1883, Jun 2014, doi: 10.1109/TCOMM.2014.2315197.
- [19] E. Björnson, G. Zheng, M. Bengtsson, and B. Ottersten, "Robust monotonic optimization framework for multicell MISO systems," *IEEE Trans. Signal Process.*, vol. 60, no. 5, pp. 2508–2523, May 2012, doi: 10.1109/TSP.2012.2184099.
- [20] S. S. Christensen, R. Agarwal, E. de Carvalho, and J. M. Cioffi, "Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design," in *Proc. IEEE International Conference on Communications*, Jun 2009, pp. 1–6, doi: 10.1109/ICC.2009.5199574.
- [21] J. Shin and J. Moon, "Weighted-sum-rate-maximizing linear transceiver filters for the K-user MIMO interference channel," *IEEE Trans. Commun.*, vol. 60, no. 10, pp. 2776–2783, Oct 2012, doi: 10.1109/TCOMM.2012.091012.110110.
- [22] Q. Shi, M. Razaviyayn, Z. Q. Luo, and C. He, "An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4331–4340, Sep 2011, doi: 10.1109/TSP.2011.2147784.
- [23] P. Jayasinghe, A. Tölli, J. Kaleva, and M. Latva-aho, "Bi-directional signaling for dynamic TDD with decentralized beamforming," in *IEEE International Conference on Communication Workshop (ICCW)*, Jun 2015, pp. 185–190, doi: 10.1109/ICCW.2015.7247176.
- [24] S. Lagen, A. Agustín, and J. Vidal, "Decentralized coordinated precoding for dense TDD small cell networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 8, pp. 4546–4561, Aug 2015, doi: 10.1109/TWC.2015.2422704.
- [25] P. Komulainen, A. Tölli, and M. Juntti, "Effective CSI signaling and decentralized beam coordination in TDD multi-cell MIMO systems," *IEEE Trans. Signal Process.*, vol. 61, no. 9, pp. 2204–2218, May 2013, doi: 10.1109/TSP.2013.2247597.
- [26] R. Hunger, W. Utschick, D. A. Schmidt, and M. Joham, "Alternating optimization for MMSE broadcast precoding," in *IEEE International Conference on Acoustics Speech and Signal Processing Proceedings*, vol. 4, May 2006, pp. 757–760, doi: 10.1109/ICASSP.2006.1661079.
- [27] S. Chen, S. Sun, Y. Wang, G. Xiao, and R. Tamrakar, "A comprehensive survey of TDD-based mobile communication systems from TD-SCDMA 3G to TD-LTE(A) 4G and 5G directions," *China Communications*, vol. 12, no. 2, pp. 40–60, Feb 2015, doi: 10.1109/CC.2015.7084401.
- [28] M. Kobayashi, N. Jindal, and G. Caire, "Training and feedback optimization for multiuser MIMO downlink," *IEEE Trans. Commun.*, vol. 59, no. 8, pp. 2228–2240, Aug 2011, doi: 10.1109/TCOMM.2011.051711.090752.
- [29] M. Biguesh and A. B. Gershman, "Training-based MIMO channel estimation: a study of estimator tradeoffs and optimal training signals," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 884–893, Mar 2006, doi: 10.1109/TSP.2005.863008.
- [30] D. P. Palomar and J. R. Fonollosa, "Practical algorithms for a family of waterfilling solutions," *IEEE Trans. Signal Process.*, vol. 53, no. 2, pp. 686–695, Feb 2005, doi: 10.1109/TSP.2004.840816.
- [31] C. Shi, D. A. Schmidt, R. A. Berry, M. L. Honig, and W. Utschick, "Distributed interference pricing for the MIMO interference channel," in *Proc. IEEE International Conference on Communications*, Jun 2009, pp. 1–5, doi: 10.1109/ICC.2009.5198927.
- [32] K. B. Petersen and M. S. Pedersen, "The matrix cookbook," Nov 2012. [Online]. Available: <http://www2.imm.dtu.dk/pubdb/p.php?3274>
- [33] D. Bertsekas, *Nonlinear Programming*. Athena Scientific, 1999.
- [34] J. Park, G. Lee, Y. Sung, and M. Yukawa, "Coordinated beamforming with relaxed zero forcing: The sequential orthogonal projection combining method and rate control," *IEEE Trans. Signal Process.*, vol. 61, no. 12, pp. 3100–3112, Jun 2013, doi: 10.1109/TSP.2013.2258343.



**Khaled Ardah** received his M.Sc. (Tech.) degree in Mobile Systems from Luleå University of Technology, Luleå, Sweden in 2013. He is currently working towards the D.Sc. (Tech.) degree at the Wireless Telecom Research Group (GTEL) at the Federal University of Ceará, Fortaleza, Brazil. His research interests are in radio resource management for wireless communications systems with special focus on interference coordination and decentralized algorithms in cellular networks.



**Yuri C. B. Silva** received the B.Sc. and M.Sc. degrees from the Federal University of Ceará, Fortaleza, Brazil, in 2002 and 2004, respectively, and the Dr.-Ing. degree from the Technische Universität Darmstadt, Germany, in 2008, all in Electrical Engineering. From 2001 to 2004 he was with the Wireless Telecom Research Group (GTEL), Fortaleza, Brazil. In 2003 he was a visiting researcher at Ericsson Research, Stockholm, Sweden, where he developed advanced radio resource management solutions for the GSM/EDGE standard. From 2005

to 2008 he was with the Communications Engineering Lab of the Technische Universität Darmstadt and since 2010 he is a Professor at the Federal University of Ceará. His main research interests are in the areas of wireless communications systems, multi-antenna processing, interference alignment, multicast services, and cooperative communications.



**Francisco R. P. Cavalcanti** received his B.Sc. and M.Sc. degrees in electrical engineering from UFC in 1994 and 1996, respectively, and his D.Sc. degree in electrical engineering from the State University of Campinas, São Paulo, Brazil, in 1999. Upon graduation, he joined UFC, where he is currently an associate professor and holds the Wireless Communications Chair with the Department of Teleinformatics Engineering. In 2000, he founded and since then has directed GTEL, which is a research laboratory based in Fortaleza that focuses on the

advancement of wireless telecommunications technologies. At GTEL, he manages a program of research projects in wireless communications sponsored by the Ericsson Innovation Center in Brazil and Ericsson Research in Sweden. He has produced a varied body of work including two edited books, conference and journal papers, international patents, and computer software dealing with subjects such as radio resource allocation, cross-layer algorithms, quality of service provisioning, radio transceiver architectures, signal processing, and project management. He is a distinguished researcher of the Brazilian Scientific and Technological Development Council for his technology development and innovation record. He also holds a Leadership and Management professional certificate from the Massachusetts Institute of Technology, Cambridge.