

Distributed Pareto Optimal Beamforming for the MISO Multi-band Multi-cell Downlink

M. Vishnu Narayanan and Srikrishna Bhashyam¹

¹Dept. of Electrical Engineering, IIT Madras, Chennai 600036, India.
skrishna@ee.iitm.ac.in

Abstract—In this paper, we consider a multi-cell multi-band downlink where the base station (BS) in each cell has multiple transmit antennas. Each cell has one active mobile station (MS) with a single receive antenna and treats interference from the other cells as noise. There is a sum transmit power constraint for each BS over all the bands. An alternating maximization (AM) algorithm is proposed to determine the optimal power allocation among the bands and the optimal beamforming vectors for each BS in each band. This algorithm can be implemented in a distributed manner with limited exchange of interference constraints between the BSs, and only local channel state information at each BS. The proposed algorithm alternates between: (1) weighted sum-rate (WSR) optimization for the beamformers in each band for a given power allocation, and (2) optimal power allocation across bands for a given set of beamformers. For the 2-cell and 3-cell settings the WSR optimization in each band is significantly simplified using analytical solutions for the sub-problems. The power allocation across bands for a given set of beamformers is obtained analytically in all cases. Numerical results show good convergence properties and significant performance gain using the proposed AM algorithm compared to: (i) equal power allocation across bands and weighted sum-rate optimization in each band, (ii) zero-forcing (ZF) beamforming, and (iii) maximal ratio transmission (MRT) beamforming.

Index Terms—Multiple-Input-Single-Output (MISO) Interference Channel, beamforming, pareto boundary, rate region, alternating maximization, weighted sum-rate maximization.

I. INTRODUCTION

Modern cellular networks allow full frequency reuse across cells to support the increasing traffic demand. However, full frequency reuse requires the use of advanced interference management techniques. One promising technique is partial or full coordination between base-stations on the downlink

[2]. In [3], [4], a multicell downlink where the base-stations (BSs) with multiple antennas cooperatively design beamforming vectors to their respective active mobile stations (MSs) is considered. Each cell has one active MS with a single receive antenna and the receivers treat interference from other cells as noise. The achievable rate region of such a multicell system, i.e., the set of all achievable (R_1, R_2, \dots, R_M) , where R_k is the rate in the k^{th} cell, is studied. In [3], [5], the Pareto boundary of the achievable rate region is characterized using an explicit parameterization. On the Pareto boundary, the rate in any cell cannot be increased without decreasing the rate in at least one other cell. In [4], it was shown that any point on the Pareto boundary can be obtained using a distributed optimization in each cell with a limited exchange of some interference parameters between cells.

The work in [3], [4] in the multi-input-single-output interference channel (MISO-IC) setting has been extended in several ways. A centralized solution to maximize weighted sum-rate is presented in [6] with full channel state information. In [7], [8], the use of more advanced receivers at each MS based on successive interference cancellation (SIC) has been analyzed. In [9]–[12], extensions to a MIMO-IC setting where the MSs have multiple receive antennas have been studied. While [10]–[12] restrict attention to single-stream transmission even in the MIMO setting, some conditions for Pareto optimality for the general MIMO-IC are obtained in [9]. The effect of partial, imperfect or quantized channel state information (CSI) has been studied in [13]–[16]. Various algorithms for weighted sum-rate optimization in the MISO-IC setting have been proposed in [17]–[19]. In [17], more than one active MSs are allowed in each cell simultaneously. The algorithms in [18], [19] are based on monotonic optimization. An iterative

This work was presented in part at GLOBECOM 2017 [1].

approach to solve problems posed by the non-convexity of weighted sum-rate maximization is studied in [20] in a different setting of wirelessly powered communications. In [21] and [22], multi-cell beamforming is studied under constraints on backhaul communication. In [23], machine learning-based selection of MISO beamforming is used to choose between zero-forcing (ZF) and maximal ratio transmission (MRT). In all these works, a single band is assumed for downlink transmission with a flat fading or Gaussian IC model.

We consider a multicell multi-band downlink with a transmit power constraint on the sum power of all the bands. While each band is modeled as flat fading or Gaussian, the multi-band system can model scenarios where multiple downlink bands are available in the same cell, or scenarios where frequency-selective fading can be well approximated using multiple band-wise flat-fading channels. In scenarios with multiple bands with per-band power constraints, maximum sum-rate is obtained when full power is utilized [24]. However, since we have multiple bands and a sum power constraint, we need also determine the optimal power allocation across bands in addition to determining the optimal beamformers for each band for each cell. Joint optimization of power allocation and beamforming is difficult.

In this work, we propose an alternating maximization (AM) algorithm to determine the power allocation and the beamformers for each band in each cell to maximize the overall weighted sum-rate (WSR) of the multi-cell multi-band MISO-IC system. As in [4], each transmitter in the BS consists of N antennas, the receiver at the MS consists of a single antenna, and interference from other cells is treated as the noise at each receiver. The proposed AM algorithm consists of two main steps. In step 1 of the AM algorithm, we fix a power allocation across the bands and maximize the weighted sum-rate for each band. This optimization for each band is solved in each cell in a distributed manner while exchanging only the interference temperature constraints between cells. For the optimal interference constraints determined in Step 1, we optimize the power allocation across bands in step 2 of the AM algorithm analytically. We make the following contributions:

- For the 2-cell single-band case, we derive an analytical solution for maximizing the rate in

each cell subject to interference temperature constraints.

- For the 3-cell single-band case, we reduce the problem of maximizing the rate in each cell subject to interference temperature constraints to a single parameter search.
- Taking advantage of the analytical results for the 2-cell and 3-cell settings, we use a gradient ascent algorithm to determine the maximum weighted sum-rate and the corresponding interference constraints.
- The AM algorithm to determine the optimal beamforming vectors and the power allocation across bands is proposed for the M -cell K -band setting.

Finally, we present numerical results for various settings and observe that: (a) the proposed AM algorithm converges to the weighted sum-rate optimal rate vector, (b) the optimal beamformers and power allocation can achieve all points on the Pareto boundary corresponding to WSR optimal rate vectors, (c) there is a significant gain in performance compared to equal power allocation across bands and WSR optimization in each band, and (d) there is significant performance gain over maximal ratio transmission (MRT) beamforming in each band with equal power allocation across bands, zero-forcing (ZF) with equal power allocation across bands (ZF-EPA) and ZF with optimal power allocation across bands (ZF-OPA). The MRT scheme does not require any coordination between cells. The ZF schemes require each BS to know the channel state information of the channel from itself to all the users in all cells. The proposed AM scheme requires the exchange of interference temperature constraints and the channel state information as in the ZF schemes.

In [25], a distributed solution to the 2-cell single-band case is provided using a Signal-to-leakage-noise ratio approach. However, the sum-rate of this scheme matches the sum-rate of the ZF scheme at high SNR. Our AM algorithm achieves rates better than ZF. In [26], a pricing-based distributed non-cooperative game approach is proposed for the multi-cell OFDMA setting. This approach is shown to converge to a Nash equilibrium and satisfies the Karush-Kuhn-Tucker (KKT) conditions for the network utility maximization problem. The information exchange in this method is similar to our proposed method. However, the iterative approach used is different and uses dual decomposition. The problem

is not decomposed into WSR optimization for each band and power allocation as in our approach. In [27], the pricing-based approach in [26] is used for distributed beamforming design in a dynamic time-division-duplex system.

The rest of this paper is organized as follows. The system model and the solution approach are described in Section II. The analytical solutions for the 2-cell single-band and 3-cell single-band cases under interference temperature constraints are in Sections III and IV. Then, this solution is used in a gradient ascent algorithm for WSR maximization. Section V presents the proposed AM algorithm for the M -cell K -band setting. Results relating the Pareto optimality in each band with overall Pareto optimality are also presented in this section. Convergence and complexity issues are discussed in Section VI. The simulation results are in Section VII. Some of the simulation details and the details of the solution in the Section IV are presented in Appendices A and B.

Notations used: Boldface lower case letters are used to represent vectors. Γ represents the set of scalars $\Gamma_{ijk} \forall i, j, k$ as defined in Section III. \mathbf{x}^H and $\|\mathbf{x}\|$ represent the Hermitian transpose and the norm of vector \mathbf{x} . a^* , $\arg(a)$ and $|a|$ respectively represents the conjugate, the complex angle and the absolute value of the scalar a .

II. SYSTEM MODEL AND PRELIMINARIES

Consider a M cell system with K distinct frequency bands of operation. Let each cell have one active mobile-station (MS). Each base-station (BS) transmitter has N antennas and each MS has a single antenna. Let \mathbf{h}_{ijk} represent the $N \times 1$ channel vector from BS i to MS j in band k . A 2-cell 2-band channel is shown in Fig. 1.

For $i, j \in \{1, 2, \dots, M\}$ and $k \in \{1, 2, \dots, K\}$, the discrete time baseband signal y_{jk} received at the j^{th} MS in the k^{th} band is given by

$$y_{jk} = \sum_i \mathbf{h}_{ijk}^H \mathbf{w}_{ik} s_{ik} + z_{jk}, \quad (1)$$

where s_{ik} is the transmitted symbol (with unit average energy) from the i^{th} BS in the k^{th} band, $z_{jk} \in \mathcal{CN}(0, \sigma_{jk}^2)$ is the additive Gaussian noise, and \mathbf{w}_{ik} is the complex $N \times 1$ transmit beamforming vector used at the i^{th} BS in the k^{th} band. Note that we use subscripts i, j and k , for BS, MS and frequency band, respectively.

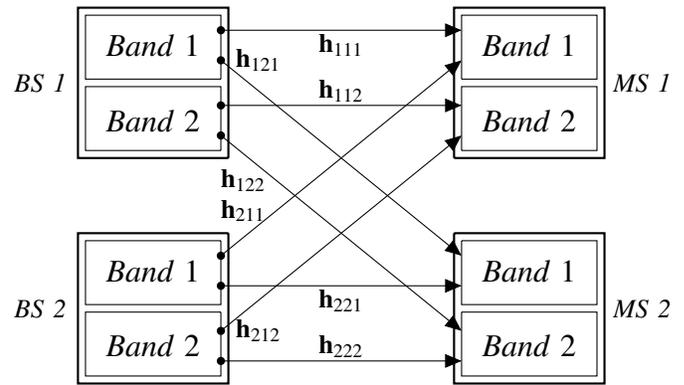


Fig. 1: 2-user 2-band multiple-input-single-output interference channel (MISO-IC) model. The multiple antennas at the BS are not shown separately.

Let R_{ik} be the rate from BS i to MS i in the k^{th} band. Then, $R_i = \sum_k R_{ik}$ is the rate from BS i to MS i . Under the assumption that the BSs encode data independently, and the MSs treat interference from the other cell as noise, the achieved rate at receiver i in band k for a given set of beamformers $\{\mathbf{w}_{1k}\}, \{\mathbf{w}_{2k}\}, \dots, \{\mathbf{w}_{Mk}\}$ can be written as [4]:

$$R_{ik}(\mathbf{w}_{1k}, \mathbf{w}_{2k}, \dots, \mathbf{w}_{Mk}) = \log \left(1 + \frac{|\mathbf{h}_{ik}^H \mathbf{w}_{ik}|^2}{\sum_{j \neq i} |\mathbf{h}_{jik}^H \mathbf{w}_{jk}|^2 + \sigma_{ik}^2} \right).$$

BS i has a power constraint P_i on the sum transmit power over all bands and all antennas, i.e.,

$$\sum_k \|\mathbf{w}_{ik}\|^2 \leq P_i, \forall i. \quad (2)$$

The achievable rate region \mathcal{R} for this system is given by

$$\mathcal{R} = \{(R_1, R_2, \dots, R_M) : R_i = \sum_k R_{ik}(\mathbf{w}_{1k}, \mathbf{w}_{2k}, \dots, \mathbf{w}_{Mk}) \text{ for } i = 1, 2, \dots, M; \text{ and } \mathbf{w}_{1k}, \mathbf{w}_{2k}, \dots, \mathbf{w}_{Mk} \text{ satisfy (2)}\}. \quad (3)$$

A rate vector (R_1, R_2, \dots, R_M) is *Pareto optimal* if there is no other rate vector $(R'_1, R'_2, \dots, R'_M)$ such that $R'_i \geq R_i$ for all i , i.e., it is not possible to increase any component of a Pareto optimal rate vector without decreasing at least one of the remaining components. The collection of all such Pareto optimal rate vectors is referred to as the *Pareto Boundary* of the rate region. A beamforming vector that results in a rate vector on the Pareto Boundary is referred to as a *Pareto beamformer*.

Also note that any rate vector that maximizes the weighted sum-rate is also Pareto optimal.

In this work, we determine the Pareto beamformer that maximizes the weighted sum-rate, i.e., the beamforming vectors that solve the following optimization problem:

$$\begin{aligned} \max_{\{\mathbf{w}_{ik}\}} \quad & \sum_i \beta_i \sum_k \log \left(1 + \frac{|\mathbf{h}_{ik}^H \mathbf{w}_{ik}|^2}{\sum_{j \neq i} |\mathbf{h}_{jik}^H \mathbf{w}_{jk}|^2 + \sigma_{ik}^2} \right) \\ \text{s.t.} \quad & \sum_k \|\mathbf{w}_{ik}\|^2 \leq P_i \quad \forall i, \end{aligned} \quad (4)$$

where β_i is the weight given to i^{th} BS. Note that number of variables to be determined in the above optimization problem is MNK .

A. Solution Approach

We obtain the solution to the optimization problem in (4) by iteratively optimizing smaller sub-problems. The simplified alternating maximization approach is summarized below. The details are presented in later sections.

First we define new variables P_{ik} for $i = 1, 2, \dots, M$, $k = 1, 2, \dots, K$ to denote the power for the i^{th} BS in the k^{th} band. Now, the problem in (4) can be written as:

$$\begin{aligned} \max_{\{\mathbf{w}_{ik}\}, \{P_{ik}\}} \quad & \sum_i \beta_i \sum_k \log \left(1 + \frac{|\mathbf{h}_{ik}^H \mathbf{w}_{ik}|^2}{\sum_{j \neq i} |\mathbf{h}_{jik}^H \mathbf{w}_{jk}|^2 + \sigma_{ik}^2} \right) \\ \text{s.t.} \quad & \|\mathbf{w}_{ik}\|^2 \leq P_{ik} \quad \forall i, k, \\ & \sum_k P_{ik} = P_i \quad \forall i. \end{aligned} \quad (5)$$

For a given $\{P_{ik}\}$, the above problem splits into K independent sub-problems, one for each band. The problem for the k^{th} band is:

$$\begin{aligned} \max_{\{\mathbf{w}_{ik}\}} \quad & \sum_i \beta_i \log \left(1 + \frac{|\mathbf{h}_{ik}^H \mathbf{w}_{ik}|^2}{\sum_{j \neq i} |\mathbf{h}_{jik}^H \mathbf{w}_{jk}|^2 + \sigma_{ik}^2} \right) \\ \text{s.t.} \quad & \|\mathbf{w}_{ik}\|^2 \leq P_{ik} \quad \forall i. \end{aligned} \quad (6)$$

We solve the sub-problem for each band in Sections III and IV for the 2-cell and 3-cell cases. The general M -cell case can also be solved using the approach used for the 2 and 3 cell cases. After solving the sub-problems for each band, we optimize the powers $\{P_{ik}\}$ separately in each cell i . These optimized powers are then used to solve the sub-problems in (6) for each band again. This

process is repeated till convergence. The general alternating maximization algorithm for the M -cell K -band setting is presented in Section V.

III. TWO-CELL SINGLE-BAND MISO-IC

In this section, we consider the single band case ($K = 1$). For the two cell MISO-IC system with $K = 1$, the optimization problem (4) for sum-rate ($\beta_1 = \beta_2 = 1$) can be re-formulated as [4]

$$\begin{aligned} \max_{\{\mathbf{w}_i\}} \quad & \sum_{i=1}^2 \log \left(1 + \frac{|\mathbf{h}_i^H \mathbf{w}_i|^2}{\sum_{j \neq i} \Gamma_{ji} + \sigma_i^2} \right) \\ \text{s.t.} \quad & |\mathbf{h}_{ij}^H \mathbf{w}_i|^2 \leq \Gamma_{ij}, \quad \forall j \neq i \\ & \|\mathbf{w}_i\|^2 \leq P_i, \quad \forall i, \end{aligned} \quad (7)$$

where additional interference temperature (IT) constraint variables $\mathbf{\Gamma} = \{\Gamma_{ij}\}$ have been introduced. An IT constraint $|\mathbf{h}_{ij}^H \mathbf{w}_i|^2 \leq \Gamma_{ij}$ limits the maximum interference caused at j^{th} MS by the i^{th} BS. For a given set of IT constraints, note that each term in the objective function depends only on the beamforming vector for the corresponding cell and, therefore, can be maximized separately. From [4], we know that for any rate vector on the Pareto Boundary, there exists a corresponding set of IT constraints $\mathbf{\Gamma}$ tight at the boundary. Thus, given the IT constraints $\mathbf{\Gamma}$, the optimization problem to be solved in cell i is:

$$\begin{aligned} \max_{\mathbf{w}_i} \quad & \log \left(1 + \frac{|\mathbf{h}_i^H \mathbf{w}_i|^2}{\sum_{j \neq i} \Gamma_{ji} + \sigma_i^2} \right) \\ \text{s.t.} \quad & |\mathbf{h}_{ij}^H \mathbf{w}_i|^2 \leq \Gamma_{ij}, \quad \forall j \neq i \\ & \|\mathbf{w}_i\|^2 \leq P_i. \end{aligned} \quad (8)$$

Thus, the Pareto beamformer that maximizes the sum-rate can be obtained in a distributed manner by just exchanging the right IT constraints between the cells. Furthermore, each BS needs to know only the channel from itself to the mobiles. For example, for the 2-cell case, BS 1 needs to know only \mathbf{h}_{11} and \mathbf{h}_{12} . However, finding the right IT constraints $\mathbf{\Gamma}$ corresponding to a given point on the Pareto boundary is not easy. In [4], an iterative algorithm that updates the IT constraints and the beamforming vectors is proposed. This algorithm is numerically observed to converge to a rate vector on the Pareto boundary. However, we cannot tell apriori which boundary point will be obtained. This limitation is overcome in our solution.

In the rest of this section, we: (1) solve the 2-cell optimization problem in (8) for a given $\mathbf{\Gamma}$ analytically, (2) propose a gradient ascent algorithm to find the Pareto optimal rate vector that maximizes weighted sum-rate and, (3) find the beamformer corresponding to the optimal $\mathbf{\Gamma}$.

A. Proposed analytical solution for a given $\mathbf{\Gamma}$

Without loss of generality, we will solve for $i = 1$, i.e., BS 1. The solution for BS 2 can be obtained in a similar fashion. From [3], we know that any optimal beamformer will be a linear combination of the channel vectors emerging out from that BS. Using this, \mathbf{w}_1 can be represented as a linear combination \mathbf{h}_{11} and \mathbf{h}_{12} . First, we represent the channel vectors using basis vectors obtained from Gram-Schmidt ortho-normalization as follows.

$$\begin{aligned} \mathbf{h}_{11} &= a_{11}\mathbf{u}_{11}, \\ \mathbf{h}_{12} &= a_{12}^{(1)}\mathbf{u}_{11} + a_{12}^{(2)}\mathbf{u}_{12}, \end{aligned} \quad (9)$$

with one real a_{11} and two other complex coefficients, $a_{12}^{(1)}$ and $a_{12}^{(2)}$. Then, for some b_{11} and b_{12} , we can write \mathbf{w}_1 as

$$\mathbf{w}_1 = b_{11}\mathbf{u}_{11} + b_{12}\mathbf{u}_{12}. \quad (10)$$

From (9) and (10) we get: $|\mathbf{h}_{11}^H \mathbf{w}_1|^2 = |a_{11}|^2 |b_{11}|^2$, $|\mathbf{h}_{12}^H \mathbf{w}_1|^2 = |b_{11}(a_{12}^{(1)})^* + b_{12}(a_{12}^{(2)})^*|^2$, and $\|\mathbf{w}_1\|^2 = |b_{11}|^2 + |b_{12}|^2$. Substituting these in (8) for $i = 1$, we can rewrite the problem (8) as

$$\begin{aligned} \max_{b_{11}, b_{12}} \quad & \log \left(1 + \frac{|b_{11}|^2 |a_{11}|^2}{\Gamma_{21} + \sigma_1^2} \right) \\ \text{s.t.} \quad & |b_{11}(a_{12}^{(1)})^* + b_{12}(a_{12}^{(2)})^*|^2 \leq \Gamma_{12}, \\ & |b_{11}|^2 + |b_{12}|^2 \leq P_1. \end{aligned} \quad (11)$$

Now, define $|b_{11}| = \gamma_1$, $|b_{12}| = \delta_1$, $\arg(b_{11}(a_{12}^{(1)})^*) = \theta_1$ and $\arg(b_{12}(a_{12}^{(2)})^*) = \phi_1$. Substituting these, and using the monotonicity of the log function, the above problem can be written as

$$\begin{aligned} \max_{\gamma_1, \delta_1, \theta_1, \phi_1} \quad & \gamma_1 \\ \text{s.t.} \quad & \gamma_1^2 |a_{12}^{(1)}|^2 + \delta_1^2 |a_{12}^{(2)}|^2 \\ & + 2\gamma_1 \delta_1 |a_{12}^{(1)}| |a_{12}^{(2)}| \cos(\theta_1 - \phi_1) \leq \Gamma_{12} \\ & \gamma_1^2 + \delta_1^2 \leq P_1, \gamma_1 \geq 0, \delta_1 \geq 0, \end{aligned} \quad (12)$$

where we have used $|\gamma_1 |a_{12}^{(1)}| e^{j\theta_1} + \delta_1 |a_{12}^{(2)}| e^{j\phi_1}|^2 = \gamma_1^2 |a_{12}^{(1)}|^2 + \delta_1^2 |a_{12}^{(2)}|^2 + 2\gamma_1 \delta_1 |a_{12}^{(1)}| |a_{12}^{(2)}| \cos(\theta_1 - \phi_1)$.

Note that the objective function does not depend on θ_1 and ϕ_1 . However, the constraints depend on $\theta_1 - \phi_1$. $\theta_1 - \phi_1$ can be chosen to be any value in $[0, 2\pi]$ by appropriately choosing the argument of b_{11} and b_{12} . Since the feasible set depends on $\theta_1 - \phi_1$, the choice that results in the largest feasible set is optimal. Observe that, since $\cos(\theta_1 - \phi_1) \geq -1$,

$$\begin{aligned} & \gamma_1^2 |a_{12}^{(1)}|^2 + \delta_1^2 |a_{12}^{(2)}|^2 - 2\gamma_1 \delta_1 |a_{12}^{(1)}| |a_{12}^{(2)}| \\ & \leq \gamma_1^2 |a_{12}^{(1)}|^2 + \delta_1^2 |a_{12}^{(2)}|^2 + 2\gamma_1 \delta_1 |a_{12}^{(1)}| |a_{12}^{(2)}| \cos(\theta_1 - \phi_1), \end{aligned}$$

Therefore, the choice $\theta_1 - \phi_1 = \pi$, i.e. $\cos(\theta_1 - \phi_1) = -1$, results in the largest feasible set. This constraint can now be rewritten as:

$$\left| \gamma_1 |a_{12}^{(1)}| - \delta_1 |a_{12}^{(2)}| \right|^2 \leq \Gamma_{12}$$

or, equivalently

$$\frac{-\sqrt{\Gamma_{12}} + \delta_1 |a_{12}^{(2)}|}{|a_{12}^{(1)}|} \leq \gamma_1 \leq \frac{\sqrt{\Gamma_{12}} + \delta_1 |a_{12}^{(2)}|}{|a_{12}^{(1)}|} \quad (13)$$

Thus, the feasible region for γ_1 , δ_1 in problem (12) is described by the constraints $\gamma_1^2 + \delta_1^2 \leq P_1$, $\gamma_1 \geq 0$, $\delta_1 \geq 0$, and (13). This region is shown as the shaded region in Fig. 2. The power constraint $\gamma_1^2 + \delta_1^2 \leq P_1$ corresponds to the circular region in Fig. 2, and the straight lines correspond to inequalities (i) and (ii) in (13). Now, we have two cases. The circular

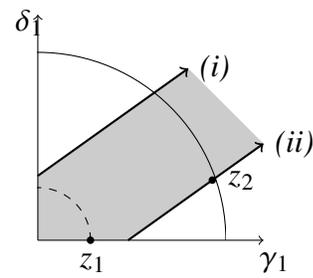


Fig. 2: Feasible region and solution of (12).

power constraint region can be either inside or outside the strip as represented by the dotted and solid circles, respectively. The solution to problem (12) in the first case is clearly the point z_1 , which corresponds to $\gamma_1 = \sqrt{P_1}$. In the second case, the solution to the problem (12) is the x-coordinate of the point z_2 . We can solve for the intersection of the circle and the line segment to get the point z_2 , whose x-coordinate corresponds to the optimum

γ_1 . Putting both cases together, we can write this solution formally as

$$\gamma_1 = \begin{cases} \frac{\sqrt{P_1}}{|a_{12}^{(1)}|\sqrt{\Gamma_{12}} + |a_{12}^{(2)}|\sqrt{P_1 d_1 - \Gamma_{12}}} & \text{if } \Gamma_{12} > P_1 |a_{12}^{(1)}|^2 \\ \text{else} & \end{cases} \quad (14)$$

where $d_1 = |a_{12}^{(1)}|^2 + |a_{12}^{(2)}|^2$. Note that in both the cases, the solution is on the circle $\gamma_1^2 + \delta_1^2 = P_1$ and satisfies the power constraint with equality. Therefore, we get $\delta_1 = \sqrt{P_1 - \gamma_1^2}$. Note that this optimum corresponds to the maximum rate that can be achieved with a particular $\mathbf{\Gamma}$, and may not correspond to a point on the Pareto boundary. We will use this solution in a gradient ascent algorithm to maximize the weighted sum-rate and obtain the corresponding point on the Pareto boundary in the next subsection. The maximum achievable rate $R_1(\mathbf{\Gamma})$ for a given $\mathbf{\Gamma}$ can be written as:

$$R_1(\mathbf{\Gamma}) = \log \left(1 + \frac{\gamma_1^2 a_{11}^2}{\Gamma_{21} + \sigma_1^2} \right) \quad (15)$$

Similarly, we can solve (8) for BS 2 ($i = 2$) to get $R_2(\mathbf{\Gamma})$.

B. Weighted Sum Rate (WSR) Maximization

In the previous subsection, we determined the maximum achievable rates $R_1(\mathbf{\Gamma})$ and $R_2(\mathbf{\Gamma})$ for a given $\mathbf{\Gamma}$. Here, we propose to maximize the weighted sum-rate $R_w = \beta R_1(\mathbf{\Gamma}) + (2 - \beta)R_2(\mathbf{\Gamma})$ with respect to $\mathbf{\Gamma}$ for a fixed weight β . This results in a point which is the foot of a tangent to the rate region, and is on the line $\beta x + (2 - \beta)y = C$, where $C = \max_{\mathbf{\Gamma}} \{\beta R_1(\mathbf{\Gamma}) + (2 - \beta)R_2(\mathbf{\Gamma})\}$. Choosing $\beta = 1$ corresponds to maximizing sum-rate.

To maximize the WSR, we use a simple coordinate ascent or gradient ascent algorithm [28], where in each step we move closer to the optimum by changing $\mathbf{\Gamma}$ in the positive gradient direction. For example, Γ_{12} is updated as:

$$\Gamma_{12,n+1} = \Gamma_{12,n} + \Delta \left[\text{sign} \left(\frac{\partial R_w}{\partial \Gamma_{12,n}} \right) \right], \quad (16)$$

where $\Delta > 0$ is a step-size parameter. Since we have an analytical solution for the rates for a given $\mathbf{\Gamma}$, the gradient can be calculated by the differentiation of

(15). The sign of the derivative of R_1 when $\Gamma_{12} \leq P_1 |a_{12}^{(1)}|^2$ is given by

$$\text{sign} \left(\frac{\partial R_1}{\partial \Gamma_{12}} \right) = \text{sign} \left(\frac{|a_{12}^{(1)}|}{\sqrt{\Gamma_{12}}} - \frac{|a_{12}^{(2)}|}{\sqrt{P_1 d_1 - \Gamma_{12}}} \right) \quad (17)$$

Similarly, we can update Γ_{21} as well. Using gradient ascent, we determine point on the Pareto boundary corresponding to the optimal weighted sum-rate. Although the iterative algorithm in [4] converges to a point on the Pareto boundary, it is not known a priori which point on the boundary is obtained. In our method we choose the boundary point by choosing the weights in the weighted sum-rate.

C. Finding the beamformer that maximizes WSR

Once we obtain the $\mathbf{\Gamma}$ and γ_1 corresponding to the Pareto boundary rate vector that maximizes WSR, we can find the beamformer that achieves this rate vector as follows.

From Section III-A, we know the optimal solution for γ_1 , δ_1 , and $\theta_1 - \phi_1$. Substituting this solution in (10), we get the optimal beamformer. From (10) and the definition of γ_1 and δ_1 , we have

$$\begin{aligned} \mathbf{w}_1 &= \gamma_1 \mathbf{u}_{11} e^{j \arg(b_{11})} + \delta_1 \mathbf{u}_{12} e^{j \arg(b_{12})} \\ &= \left(\gamma_1 \mathbf{u}_{11} + \delta_1 \mathbf{u}_{12} e^{j(\arg(b_{12}) - \arg(b_{11}))} \right) e^{j \arg(b_{11})}. \end{aligned}$$

Observe that the achieved rate (see objective function in (8) or (26)) is invariant to a constant phase rotation of all the beamformer components, i.e., \mathbf{w}_1 and $\mathbf{w}_1 e^{j\theta}$ result in the same rate for any θ . Therefore, we can set the phase term $e^{j \arg(b_{11})}$ to be 1. Finally, since $\theta_1 - \phi_1 = \pi$, we have

$$\theta_1 - \phi_1 = \arg(b_{11}^* a_{12}^{(1)}) - \arg(b_{12}^* a_{12}^{(2)}) = \pi$$

or, equivalently

$$\arg(b_{12}) - \arg(b_{11}) = \arg(a_{12}^{(2)}) - \arg(a_{12}^{(1)}) + \pi.$$

Therefore, we have

$$\mathbf{w}_1 = \gamma_1 \mathbf{u}_{11} + \delta_1 \mathbf{u}_{12} e^{j(\arg(a_{12}^{(2)}) - \arg(a_{12}^{(1)}) + \pi)}. \quad (18)$$

Similarly, we can obtain the beamformer for BS 2 \mathbf{w}_2 also. It is worth mentioning that: (1) the base stations only need to share the IT constraints and can then calculate their WSR beamformer in a distributed manner, and (2) each BS requires the channel state information only from itself to the MSs.

IV. THREE-CELL SINGLE-BAND MISO-IC

In this section, we solve the 3-cell single-band case for a given set of IT constraints. In Section III-A, we derived an analytical solution for the 2-cell single-band case where we have one IT constraint in each cell. In the 3-cell case, we have 2 IT constraints in each cell. The optimization problem to be solved in each cell i is given in (8). Here, we focus on the solution for cell 1. The solutions for cells 2 and 3 can be obtained in a similar manner. The optimization problem to be solved in cell 1 is:

$$\begin{aligned} \max_{\mathbf{w}_1} \log \left(1 + \frac{|\mathbf{h}_{11}^H \mathbf{w}_1|^2}{\sum_{j=2}^3 \Gamma_{j1} + \sigma_1^2} \right) \quad (19) \\ \text{s.t. } |\mathbf{h}_{1j}^H \mathbf{w}_1|^2 \leq \Gamma_{1j}, j = 2, 3 \\ \|\mathbf{w}_1\|^2 \leq P_1. \end{aligned}$$

Using the property that, every beamforming vector can be represented in terms of the channel vectors [3], we find the basis vectors by the Gram-Schmidt ortho-normalization method. Now, we can represent the beamformer and channel vectors as

$$\begin{aligned} \mathbf{h}_{11} &= a_{11} \mathbf{u}_{11}, \\ \mathbf{h}_{12} &= a_{12}^{(1)} \mathbf{u}_{11} + a_{12}^{(2)} \mathbf{u}_{12}, \\ \mathbf{h}_{13} &= a_{13}^{(1)} \mathbf{u}_{11} + a_{13}^{(2)} \mathbf{u}_{12} + a_{13}^{(3)} \mathbf{u}_{13}, \\ \mathbf{w}_1 &= b_{11} \mathbf{u}_{11} + b_{12} \mathbf{u}_{12} + b_{13} \mathbf{u}_{13}, \end{aligned} \quad (20)$$

where a_{11} is some real coefficient and remaining coefficients are complex. Substituting these in (19) and using the monotonicity of the log function, we get

$$\begin{aligned} \max_{b_{11}, b_{12}, b_{13}} |b_{11}|^2 a_{11}^2 \\ \text{s.t. } |b_{11}(a_{12}^{(1)})^* + b_{12}(a_{12}^{(2)})^*|^2 \leq \Gamma_{12}, \\ |b_{11}(a_{13}^{(1)})^* + b_{12}(a_{13}^{(2)})^* + b_{13}(a_{13}^{(3)})^*|^2 \leq \Gamma_{13}, \\ |b_{11}|^2 + |b_{12}|^2 + |b_{13}|^2 \leq P_1. \end{aligned} \quad (21)$$

This problem has an additional constraint compared to the problem in (11). Define $|b_{11}| = \gamma_1$, $|b_{12}| = \delta_1$, $|b_{13}| = \zeta_1$, and $\theta_k = \arg(b_{1k})$, for $k \in 1, 2, 3$. Let

$\arg(a_{12}^{(k)}) = A_k$, for $k = 1, 2$, and $\arg(a_{13}^{(k)}) = B_k$, for $k = 1, 2, 3$. We can now rewrite the problem (21) as

$$\begin{aligned} \max_{\gamma_1, \delta_1, \zeta_1, \theta_1, \theta_2, \theta_3} \gamma_1 \\ \text{s.t. } \gamma_1^2 + \delta_1^2 + \zeta_1^2 \leq P_1, \\ \gamma_1^2 |a_{12}^{(1)}|^2 + \delta_1^2 |a_{12}^{(2)}|^2 \\ + 2\gamma_1 \delta_1 |a_{12}^{(1)}| |a_{12}^{(2)}| \cos(\theta_1 - \theta_2 - A_1 + A_2) \leq \Gamma_{12}, \\ \gamma_1^2 |a_{13}^{(1)}|^2 + \delta_1^2 |a_{13}^{(2)}|^2 + \zeta_1^2 |a_{13}^{(3)}|^2 \\ + 2\gamma_1 \delta_1 |a_{13}^{(1)}| |a_{13}^{(2)}| \cos(\theta_1 - \theta_2 - B_1 + B_2) \\ + 2\delta_1 \zeta_1 |a_{13}^{(2)}| |a_{13}^{(3)}| \cos(\theta_2 - \theta_3 - B_2 + B_3) \\ + 2\gamma_1 \zeta_1 |a_{13}^{(1)}| |a_{13}^{(3)}| \cos(\theta_1 - \theta_3 - B_1 + B_3) \leq \Gamma_{13}. \end{aligned} \quad (22)$$

Note again that the objective function does not depend on θ_1 , θ_2 , and θ_3 . Therefore, θ_1 , θ_2 , and θ_3 , should be chosen to result in the largest feasible set as in Section III-A. For a given choice of $\theta_1 - \theta_2$, the optimal choice of $\theta_2 - \theta_3$ is derived in closed-form in Appendix B. Note also that $\theta_1 - \theta_3 = (\theta_1 - \theta_2) + (\theta_2 - \theta_3)$. Using the solution in Appendix B, the problem (22) reduces to a simple 1-dimensional search for γ_1 . For each candidate γ_1 in this search, we only need to check if the feasible region is non-empty. Once we find the largest γ_1 for which the feasible region is non-empty, we also have δ_1 , ζ_1 , $\theta_1 - \theta_2$ and $\theta_2 - \theta_3$. As in Section III-C, θ_1 can be chosen to be 0. The optimal beamformer is then calculated as:

$$\mathbf{w}_1 = \gamma_1 \mathbf{u}_{11} + \delta_1 e^{j\theta_2} \mathbf{u}_{12} + \zeta_1 e^{j\theta_3} \mathbf{u}_{13}. \quad (23)$$

Using the optimum γ_1 , we can get the corresponding rate as:

$$R_1(\mathbf{\Gamma}) = \log \left(1 + \frac{\gamma_1^2 a_{11}^2}{\Gamma_{21} + \Gamma_{31} + \sigma_1^2} \right) \quad (24)$$

Similarly, we can solve for $i = 2$ and $i = 3$ to get $R_2(\mathbf{\Gamma})$ and $R_3(\mathbf{\Gamma})$. The $\mathbf{\Gamma}$ corresponding to the rate vector that maximizes WSR can be determined using gradient ascent as in Section III-B.

V. MULTI-CELL MULTI-BAND MISO-IC

In this section, we consider a M -cell MISO-IC with K bands. Our aim is to maximize the weighted sum-rate $\beta_1 R_1 + \beta_2 R_2 + \dots + \beta_M R_M$ (see (4)), where $R_i = \sum_{k=1}^K R_{ik}$, for $i = 1, 2, \dots, M$. For each BS, we have a sum power constraint over all bands, P_i for BS i . Let P_{ik} be the power used by BS i in

band k . Note that due to the orthogonality of non overlapping frequency bands, if all P_{ik} 's are known, the problem can be split into K independent single band problems. However, we have a sum power constraint $\sum_k P_{ik} \leq P_i$. We need to find the optimal power allocation, i.e. the P_{ik} 's with $\sum_k P_{ik} \leq P_i$, such that the weighted sum-rate is maximized. For each BS i , we define a $N \times 1$ power allocation vector $\alpha_i = [\alpha_{i1} \ \alpha_{i2} \ \dots \ \alpha_{iK}]^T$ such that $\sum_k \alpha_{ik} = 1$ and we can write $P_{ik} = \alpha_{ik} P_i$.

In the rest of this section, we (1) first identify and prove a property of the Pareto boundary for the multi-band problem, and (2) propose an Alternating Maximization (AM) algorithm to obtain the optimal beamformers and the optimal power allocation between the bands that maximize weighted sum-rate.

A. Relationship between the Pareto boundary for the multi-band case and the single-band case

Proposition 1: Consider a point (R_1, R_2, \dots, R_M) on the Pareto boundary for the multi-band case and the corresponding power allocation $\{P_{ik}\}$. Then, for each band k , $(R_{1k}, R_{2k}, \dots, R_{Mk})$, lies on the Pareto boundary for that band, for an allocated power of P_{ik} in band k of BS i .

Proof. Note that we have $R_i = \sum_{k=1}^K R_{ik}$, for $i = 1, 2, \dots, M$. Suppose $(R_{1k}, R_{2k}, \dots, R_{Mk})$ is not on the Pareto boundary for band k . Then, at least one of R_{ik} can be increased without decreasing the others. This implies that we can increase one of the R_i without decreasing the others. This is a contradiction since (R_1, R_2, \dots, R_M) is assumed to be on the Pareto boundary for the overall rate over the K bands. \square

For a fixed power allocation α_i , $i = 1, 2, \dots, M$, we observe that

$$\begin{aligned} \max_{\{\mathbf{w}_{ik}\}} \sum_i \beta_i R_i &= \max_{\{\mathbf{w}_{ik}\}} \sum_i \beta_i \sum_k R_{ik} \\ &= \max_{\{\mathbf{w}_{ik}\}} \sum_k \sum_i \beta_i R_{ik} \\ &= \sum_k \max_{\mathbf{w}_{1k}, \mathbf{w}_{2k}, \dots, \mathbf{w}_{Mk}} \left\{ \sum_i \beta_i R_{ik} \right\} \end{aligned} \quad (25)$$

This is because, for a fixed α_i , $i = 1, 2, \dots, M$, the rates R_{ik} in band k depend only on the beamforming vectors \mathbf{w}_{ik} for the same band k . Therefore, for a fixed power allocation across bands, overall WSR

maximization reduces to WSR maximization for each band with the same weights $\{\beta_i\}$.

B. Proposed Alternating Maximization (AM) algorithm for WSR maximization

Ideally, we need to find the optimal $\{\alpha_i\}$ and $\{\mathbf{w}_{ik}\}$ jointly. However, this is difficult. Therefore, based on the observations in the previous subsection, we now propose an alternating maximization algorithm to obtain the optimal power allocation across bands and the optimal beamforming vectors for each BS for each band.

First, we fix $\{\alpha_i\}$ and maximize the WSR in each band k . To solve the sub-problem for each band k , we use the approach in Sections III and IV where IT constraints Γ are introduced. Given the IT constraints Γ , the optimization problem to be solved in cell i for band k is:

$$\begin{aligned} \max_{\mathbf{w}_{ik}} \quad & \log \left(1 + \frac{|\mathbf{h}_{iik}^H \mathbf{w}_{ik}|^2}{\sum_{j \neq i} \Gamma_{jik} + \sigma_{ik}^2} \right) \\ \text{s.t} \quad & |\mathbf{h}_{ijk}^H \mathbf{w}_{ik}|^2 \leq \Gamma_{ijk}, \forall j \neq i \\ & \|\mathbf{w}_{ik}\|^2 \leq P_{ik} = \alpha_{ik} P_i. \end{aligned} \quad (26)$$

This problem is solved independently in each cell i , $i = 1, 2, \dots, M$. An analytical solution for $M = 2$ was obtained in Section III. In Section IV, for $M = 3$, this problem is reduced to a single parameter optimization, where the at each search step only feasibility needs to be checked. For any M , this problem is convex and can be solved using standard convex optimization tools. The optimal IT constraints corresponding to the WSR optimal rate vector can be found using the gradient ascent algorithm in Section III-B for any M .

Next, for the optimal Γ obtained by WSR maximization in each band, we optimize the $\{\alpha_i\}$. The rate R_i in cell i after optimizing Γ is:

$$\begin{aligned} R_i &= \sum_{k=1}^K \log \left(1 + \frac{|\mathbf{h}_{iik}^H \mathbf{w}_{ik}|^2}{\sum_{j \neq i} \Gamma_{jik} + \sigma_{ik}^2} \right) \\ &= \sum_{k=1}^K \log \left(1 + \frac{|\mathbf{h}_{iik}^H \tilde{\mathbf{w}}_{ik}|^2 \alpha_{ik} P_i}{\sum_{j \neq i} \Gamma_{jik} + \sigma_{ik}^2} \right), \end{aligned} \quad (27)$$

where $\tilde{\mathbf{w}}_{ik} = \mathbf{w}_{ik} / \sqrt{\alpha_{ik} P_i}$ is the direction of the beamformer \mathbf{w}_{ik} , and the WSR is:

$$\sum_i \beta_i R_i = \sum_i \beta_i \sum_{k=1}^K \log \left(1 + \frac{|\mathbf{h}_{iik}^H \tilde{\mathbf{w}}_{ik}|^2 \alpha_{ik} P_i}{\sum_{j \neq i} \Gamma_{jik} + \sigma_{ik}^2} \right). \quad (28)$$

Now, fixing the directions of the beamformers and fixing the Γ , we can optimize $\{\alpha_i\}$. This reduces to optimizing the rate R_i in each cell i over the power allocation over bands in that cell, resulting in a water-filling solution for power allocation, i.e., we have:

$$\alpha_{ik} P_i = \left(\lambda - \frac{\sum_{j \neq i} \Gamma_{jik} + \sigma_{ik}^2}{|\mathbf{h}_{ik}^H \tilde{\mathbf{w}}_{ik}|^2} \right)^+, \quad (29)$$

with λ being a constant such that $\sum_k \alpha_{ik} P_i = P_i$. Now, this power allocation $\{\alpha_i\}$ can be used in the first step to solve the WSR maximization in each band again.

We alternate between these 2 optimization steps until we have convergence. Thus, we have an *Alternating Maximization* (AM) algorithm given in Algorithm.1.

Algorithm 1 AM Algorithm

- 1: Initialize: $\alpha_{ik} = 1/K$ for all i, k
 - 2: **repeat**
 - 3: **for** Band $k = 1, 2, \dots, K$ **do**
 - 4: $P_{ik} = \alpha_{ik} P_i \forall i$
 - 5: **repeat**
 - 6: Initialize: $\{\Gamma_{ijk}\}$ for band k
 - 7: Solve (26)
 - 8: Update $\{\Gamma_{ijk}\}$ using gradient ascent
 - 9: **until** Convergence
 - 10: **end for**
 - 11: **for** BS $i = 1, 2, \dots, M$ **do**
 - 12: Update $\alpha_{ik} \forall k$ using (29)
 - 13: **end for**
 - 14: **until** Convergence
-

In Algorithm 1:

- We initialize with equal power-sharing between the bands in all cells.
- The analytical solution for (26) in Line 7 is provided in Section III for $M = 2$. For $M = 3$, problem (26) is reduced to a single parameter search in Section IV. For general M , standard convex optimization tools can be used. The solutions provided for $M = 2, 3$ are much simpler than using standard tools.
- In this gradient ascent step, we use the equation (16) in Section III-B for band k . For $M = 2$, the derivatives are easily obtained using equation (17). For general M , numerical evaluation is used.

- In Lines 9 and 14, convergence is decided by checking if the change in rate during the previous iteration is less than a specified threshold ϵ .

VI. DISCUSSION

A. Convergence

In the AM algorithm, we have two steps: WSR optimization for each band (Step 1) and power optimization for each cell (Step 2). Step 1 is a gradient ascent algorithm wherein each iteration rate optimization in each cell with IT constraints is solved. The rate optimization with IT constraints is convex and the solution is obtained analytically (for $M = 2, 3$). The gradient ascent iterations can possibly converge to a local maximum since the overall WSR optimization is non-convex. However, in our simulations, we did not observe this. Fig. 3b shows the convergence of the gradient ascent averaged over 1000 channel realizations. Various choice of step sizes for the gradient ascent step were studied. The convergence performance for three choices of step size Δ , namely $\Delta = \frac{\Gamma_{max}}{2^{n+1}}$, $\Delta = \frac{\Gamma_{max}}{3n}$, and $\Delta = \frac{\Gamma_{max}}{6n}$, are shown in Fig. 3b. Here, n is the iteration number and Γ_{max} is the maximum possible value for the IT constraint. Overall, we observe that good convergence behavior is observed for all 3 step sizes in less than 10 iterations.

Step 2 is a convex optimization step with a closed-form solution. The overall convergence of the AM algorithm is shown in Fig. 3a. Fig. 3a shows the convergence of AM algorithm for the 2-band and 3-band cases. There is a slight reduction in the convergence rate with the increase in one extra dimension (band). However, in both cases, the convergence is quite fast and requires very few iterations of AM.

In addition to the equal power allocation initialization, we tried several other initial power allocations for the AM algorithm. However, no change was observed in the results.

B. Complexity

The number of variables to be optimized in the overall WSR optimization problem (4) is MNK . In the proposed algorithm this is solved by breaking this down into simple sub-problems. In Step 1, MK sub-problems, each to find one $N \times 1$ beamforming

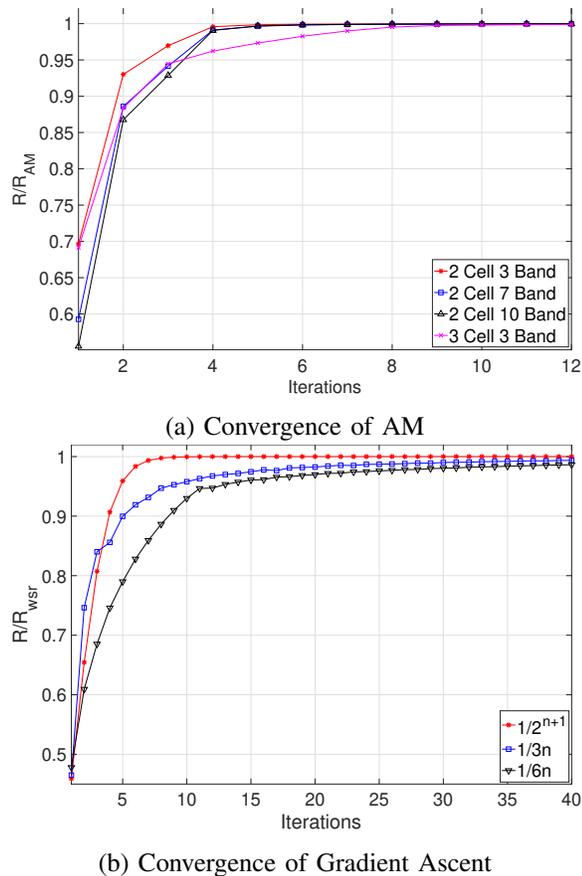


Fig. 3: (a):Convergence of AM algorithm with iterations in terms of the ratio of the rate at a particular iteration to the final rate, averaged over 1000 channel realizations.

(b):Convergence of Gradient Ascent algorithm for different approaches for the step size. Ratio of objective (sum-rate) at a particular iteration to the optimum value of objective, averaged over 1000 realizations is plotted against the iterations

vector, are solved. For $M = 2$, an analytical solution is provided, and for $M = 3$, it is reduced to a single parameter search. In Step 2, M sub-problems with a simple water-filling analytical solution are solved. Thus the complexity of the overall joint optimization problem is significantly reduced by iteratively solving simple problems with analytical solutions.

C. Information exchange overhead for coordination

Information is exchanged between cells in each iteration of the gradient ascent step for each band. The number of parameters sent by each cell per band is $2(M - 1)$ in each gradient ascent step, i.e.,

2 parameters (IT constraint, gradient with respect to IT constraint) to each of the other $(M - 1)$ cells. Therefore, the overall information exchange for the whole AM algorithm for a K band system is $2(M - 1)KN_{AM}N_{GA}$ real scalars, where N_{AM} and N_{GA} are the number of AM iterations and the number of gradient ascent iterations, respectively for each band. Reduction in exchange is possible by: (1) optimizing the number of iterations for each band, and (2) taking advantage of the fact that only the sign of the gradient is used in the gradient ascent step.

The proposed algorithm can also be implemented at a centralized server by sending all the channel information to the server and returning the solution for the beamformers to the base-stations. This requires MK complex channel vectors of dimension N to be sent to the server from each cell and K beamforming vectors to be sent back to each cell, resulting in a total exchange of $2(M + 1)KN$ real scalars by each cell. Note that this information exchange increases linearly with the number of antennas. The centralized implementation will be useful when the coherence time of the channels is large. The distributed implementation has two advantages: (1) The intermediate solutions can also be used as the solution is converging. When the channels are changing slowly, this distributed implementation can operate in tracking mode and allows for continuous updates of local channel information, and (2) the delay between the time at which channel is estimated and the time at which the corresponding beamformer is used will be lesser than in the centralized implementation.

In case of a system with multiple users with slot allocation in a round-robin fashion, the information needs to be exchanged in every slot. The amount of information exchanged over time in such a system can be drastically reduced by providing memory to the base stations so that they do not have to run the algorithm again if the same set of users are paired up again within a coherence time.

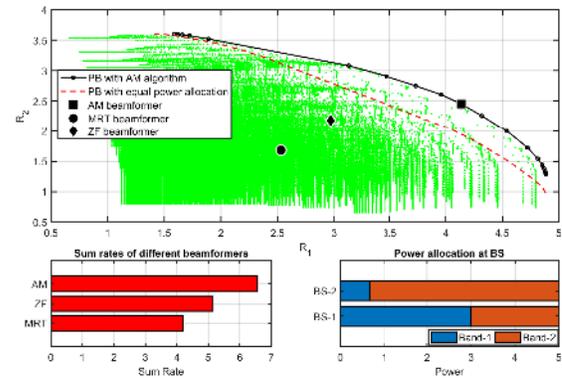
D. Reducing complexity for a system with a large number of sub-carriers

The computational complexity of the proposed AM algorithm increases linearly in the number of bands. In an OFDM system with a large number of sub-carriers, considering each sub-carrier as a

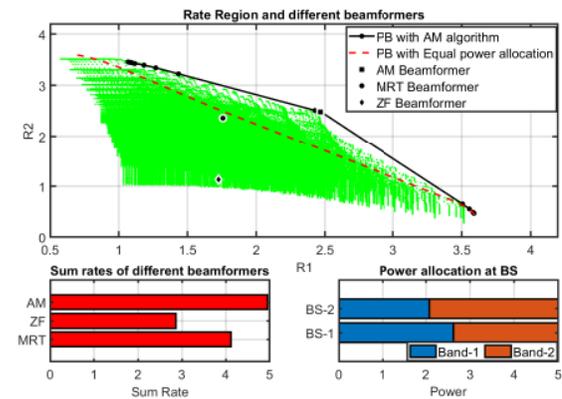
band might be computationally expensive and unnecessary. We can reduce the number of bands by grouping multiple sub-carriers within the coherence bandwidth into a single band. The same beamforming solution can be used for all sub-carriers in the same band. Suppose we have a multipath channel with L taps using N_c sub-carriers, and we form K groups each containing N_k sub-carriers with $\sum_k N_k = N_c$. The grouping is done in such a way that the maximum number of sub-carriers are grouped together while keeping the channel almost frequency-flat within each band. The channel responses of each band are taken as the average of the channel responses of the containing sub-carriers. Intuitively, when the number of paths is small, a higher number of sub-carriers can be grouped together without significant rate reduction. After performing beamforming on the K bands, we can use the same optimal beamformers with power adjusted to match the total power of the band. Let, \mathbf{w}_{ik} be the optimal beamformer for band k in cell i , then the optimal beamformer for all the N_k sub-carriers within the band is given by $\frac{\mathbf{w}_{ik}}{\sqrt{N_k}}$.

E. Multiple users in each cell

Thus far, we have considered that each cell contains only one active user. This active user could be selected in a system with multiple users per cell by using some scheduling method. For scheduling, a simple round-robin scheduler or sophisticated schedulers which take the user demand into account can be used depending on the necessity. For example, each user can be allowed to transmit in a particular time slot scheduled to it and can be given full access to all available frequency bands in the system during this slot. The AM algorithm has to be repeated in each slot to find out the beamforming vectors corresponding to the users selected in each cell. In situations where the same users have paired again within a coherence time, the recalculation of beamforming vectors can be avoided by storing the previous result in memory. This is particularly useful in situations where the slot duration is small compared to the coherence time. We could also use scheduling in the frequency domain (where one or more bands are allocated to a particular user) or scheduling in a combination of frequency and time domains.



(a) Convex Rate Region



(b) Non-convex Rate Region

Fig. 4: Rate region, Pareto Boundary(PB) and optimal sum-rate point using AM algorithm. Rate points corresponding to Zero-forcing and Maximal Ratio Transmission beamformers and a Pareto Boundary with equal power allocation across the bands are shown for comparison. $P_i = 5$ and $\sigma_{ik}^2 = 1 \forall i, k$. (a) Convex rate region (b) Non-Convex rate region

VII. SIMULATION RESULTS

In this section, we first present simulation results for the 2-cell K -band downlink to show the effectiveness of our proposed distributed Pareto optimal beamforming using alternating maximization (AM). The power allocation across bands obtained from the AM algorithm is shown to perform significantly better than equal power allocation across bands. We also compare our results with two other well-known beamforming approaches - zero-forcing (ZF) beamforming and maximal-ratio-transmission (MRT) in each cell - in terms of the achieved rate vectors and the sum-rate. For the MRT scheme we consider equal power allocation across bands, and for the zero-forcing scheme we consider both

equal power allocation (ZF-EPA) and optimal power allocation across bands (ZF-OPA). Then, we present simulation results to show how the grouping of sub-carriers in an OFDM system can be used to reduce the number of frequency bands, and hence the complexity. We show that multiple sub-carriers can be grouped without a significant rate reduction. Later, we show results for the 3-cell multiband case. Finally, we simulate a scenario where multiple users are present in each cell but one user is scheduled per cell. In all these cases we compare the AM algorithm with the ZF and the MRT beamformers.

In the simulations, we set the noise variance $\sigma_{ik}^2 = 1 \forall i, k$. The transmit power constraint for each cell is set to be equal, i.e., $P_i = P/M$ for all i , where P is the total power. When P is fixed for a plot, we set $P = 10$ units. Except in Section VII-C, each channel coefficient in \mathbf{h}_{ijk} (the channel vector from BS i to MS j in band k) is independently generated to be circularly symmetric complex Gaussian with zero mean and variance σ_{ijk}^2 , i.e., $CN(0, \sigma_{ijk}^2)$. For $i = j$, i.e., direct channels, σ_{ijk}^2 is set to 1. For $i \neq j$, i.e., cross channels, σ_{ijk}^2 is set to be the reciprocal of the cross-channel pathloss. Except in Fig. 9a, cross-channel pathloss is set to 1 (or 0dB). For the L -tap channel in Section VII-C, we generate independent zero-mean equal power circularly symmetric complex Gaussian random variables for each tap (for each transmit-receive antenna pair). Then, we get the N_c subcarrier frequency domain channel by taking the N_c -point discrete Fourier transform (DFT) of the L tap coefficients.

A. 2-cell 2-band downlink

Each BS has two antennas. The channel vectors used are given in Table.I and Table.II in Appendix A. In Figs. 4a and 4b, we show the achievable rate region for all possible beam-forming vectors satisfying the power constraints, the rate vector achieved by our proposed AM algorithm for different choices of weight β , and the rate vectors achieved by ZF and MRT beamforming. The rate vectors achieved by the proposed AM algorithm are much better than those achieved by ZF beamforming and MRT beamforming. The rates achieved with equal power allocation across the bands and only WSR maximization in each band are also shown for comparison (PB with equal power allocation). There is a significant improvement achieved using

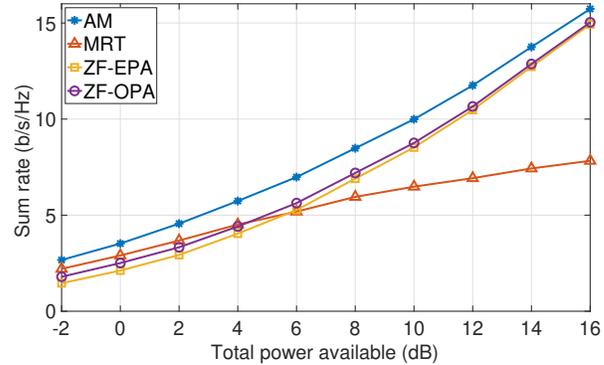
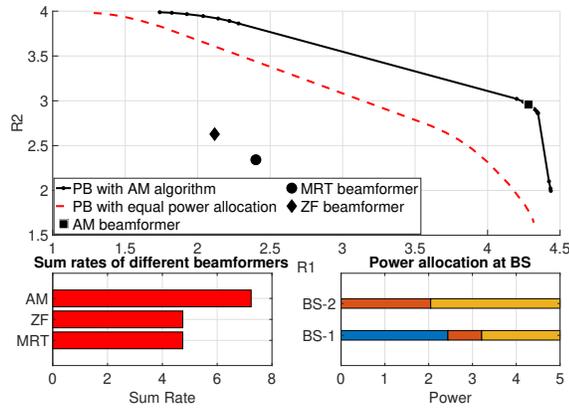


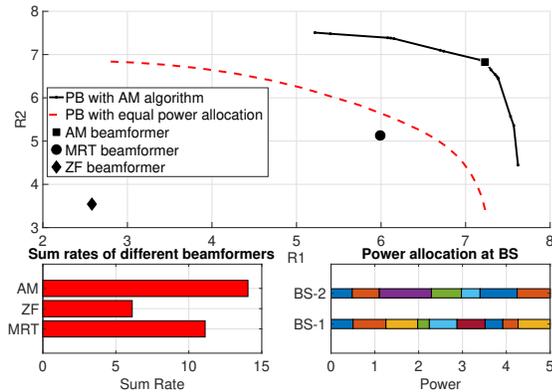
Fig. 5: Sum rate vs. Total transmit power for different beamforming schemes: 2-Band case

the AM algorithm over equal power allocation. Bar charts containing the optimal power allocation and the improvement in the sum-rates with respect to ZF and MRT are also shown below the rate region plot.

For a better understanding of the Pareto boundary, the achievable rate region is plotted by a brute-force sweep of the Γ and evaluation of the rate vector for each case (green points in the figure). Note that, this rate region obtained by brute-force is only used for giving an idea about the rate region and is no way related to the efficiency of our proposed algorithm. The range for Γ_{ijk} is $[0, \|h_{ijk}\|^2 \cdot P_i]$, which is 0 to 5 in our case. We vary each parameter with a step size of 0.1. Note that rate region in Fig. 4a is convex whereas the rate region in Fig. 4b is non-convex. The rate vector achieved by the AM algorithm for 20 different β values equally spaced in $[0, 2]$ are shown (PB with AM algorithm). It can be seen that these are points on the Pareto boundary corresponding to the maximum weighted sum-rate for that weight β . The point corresponding to the maximum sum-rate is highlighted (AM beamformer). In Fig. 4b, it can be seen that the weighted sum-rate optimal points are only on the intersection of the Pareto boundary and the convex hull of the rate region. Note that the boundary obtained by connecting the WSR optimal points gives the convex hull of the actual rate region. In Fig. 5, the average sum-rate (over 1000 channel realizations) achieved by the proposed beamformer (found using the AM algorithm) is compared with the average sum-rate achieved by MRT beamforming with equal power allocation across bands, ZF-EPA and ZF-OPA. It can be seen that the proposed AM beamformer is



(a) 3-Band



(b) 10-Band

Fig. 6: Rate region, Pareto Boundary (PB) and optimal sum rate point using AM algorithm. Rate points corresponding to ZF and MRT beamformers and the Pareto Boundary with equal power allocation across the bands are shown for comparison. $P_i = 5$ and $\sigma_{ik}^2 = 1 \forall i, k$. (a) 3-Band (b) 10-Band

significantly better than the ZF-OPA, ZF-EPA and MRT beamformers. As expected, the MRT beamformer suffers from interference at higher power (SNR) and the ZF-EPA and ZF-OPA beamformers suffer from noise enhancement.

B. 2-cell multi-band downlink

Each BS has 2 antennas and K bands. The channel vectors are generated as described earlier and one random realization of channel vectors used for the plots here is given in Appendix A.

Fig. 6a and Fig. 6b show the rate region for $K = 3$ and $K = 10$, respectively. As in the two-band case, the ZF and MRT beamformers, and the Pareto boundaries using WSR maximization with

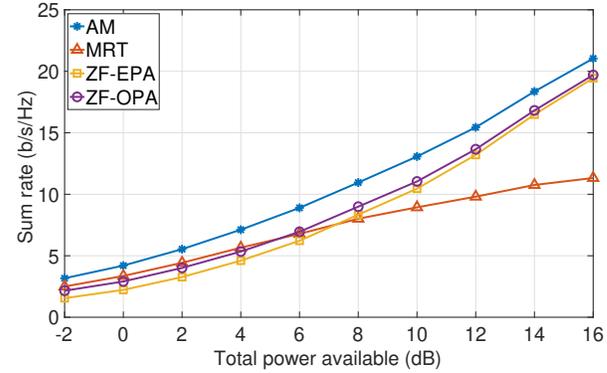


Fig. 7: Sum rate vs. Total transmit power for different beamforming schemes: 3-Band case.

equal power allocation are used for comparison. The proposed method using the AM algorithm is significantly better than the other schemes. The bar graphs at the bottom show the improvement in sum-rates obtained by AM algorithm as well as the optimum power allocation in each BS.

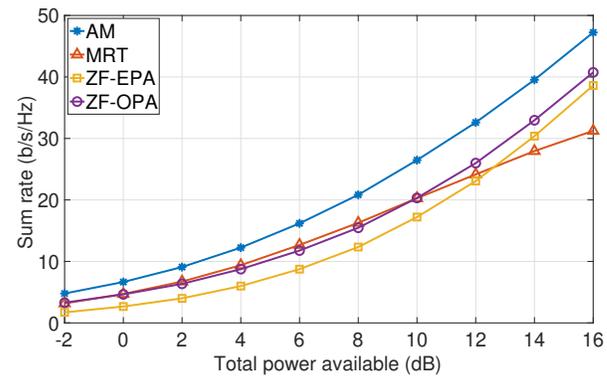
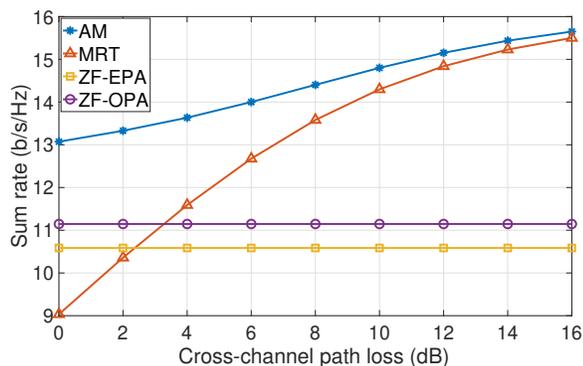


Fig. 8: Sum rate vs. Total transmit power for different beamforming schemes: 10-band case.

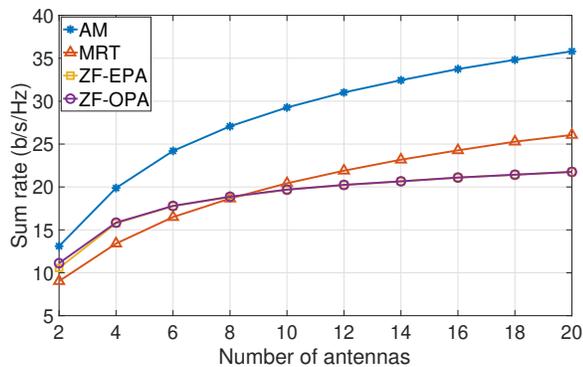
In Figs. 7 and 8, we show the behavior of the sum-rate obtained by the AM beamformer in comparison with the other beamformers (MRT, ZF-EPA and ZF-OPA) for a 3-band MISO IC and a 10-band MISO IC. The sum-rate plotted against a varying available power is the average value of 1000 random channel realizations. It is observed that gains obtained by the proposed algorithm over ZF and MRT beamformers are similar to the two-band case in both low SNR and high SNR regimes. Note that in the low SNR region, though the curves are close to each other, we achieve a good percentage improvement in the rates. We can also observe from these two figures and Fig. 5 for $K = 2$ that

the MRT beamformer performs better than the ZF beamformer for a larger range of SNR with the increase in the number of bands.

It is also to be noted that for a particular available power, the percentage improvement provided by AM algorithm is higher with $K > 2$ than that of the 2 band case and can be attributed to the fact that as the dimension increases, the equal power point moves farther from the edge of the feasible region (any corner of the hyper-cube which is adjacent to the origin) and the optimum point can possibly lie farther from the equal power point.



(a) Sum rate vs. Cross-channel path loss



(b) Sum rate vs. Number of transmit antennas

Fig. 9: Sum rate obtained for different beamforming schemes: 3-band case. $P_i = 5$ for each i .

So far, we have taken the cross-channel path loss to be 0 dB. Fig. 9a shows the sum rate as a function of the cross-channel path loss for a 2-cell 3-band setting with 2 transmit antennas. It can be observed that the proposed AM scheme always performs better than the MRT, ZF-EPA and ZF-OPA schemes. As the cross-channel pathloss increases, inter-cell interference reduces and MRT starts to perform better as expected. ZF performance does not depend on cross-channel path loss as the interference is

forced to zero irrespective of its strength in this strategy.

In Fig. 9b, the variation of sum rate with respect to the number of transmit antennas is shown. The performance of all the schemes improve with number of antennas, but the AM scheme performs significantly better than the other schemes.

C. Grouping sub-carriers in an OFDM system

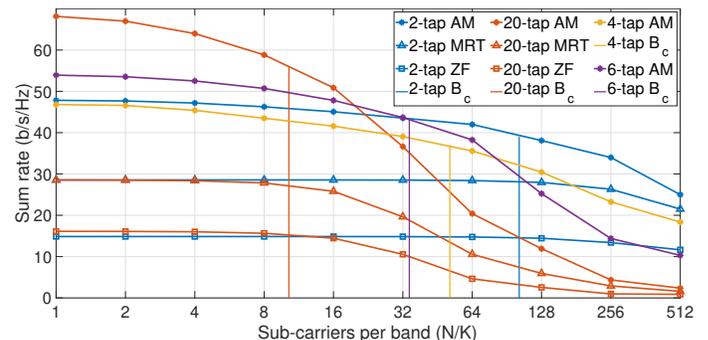


Fig. 10: Sum rate vs. band size for a 1024 sub-carrier channel. Realizations for $M = 2, 4, 6$ and 20 taps are considered. The number of sub-carriers equivalent to the coherence bandwidth (B_c) is also shown for reference. ZF and MRT beamformers are also shown for $M = 2$ and 20.

Here we simulate grouping of multiple sub-carriers into bands as discussed in Section VI-D. L -tap multi-path channels with equal power taps are considered. The N_c sub-carriers are grouped into K bands (or sub-carrier groups) with N_c/K sub-carriers each. Simulation results in Fig. 10 show the variation of sum rate with N_c/K for channels with different number of multi-path components (L) for $N_c = 1024$ sub-carriers. As in [29, Sec. 3.3.2], we define B_c , the coherence bandwidth, as the bandwidth within which the frequency correlation function is above 0.5 using the expression $B_c = 0.2/T_d$, where T_d is the delay spread. The coherent bandwidth for each channel is also shown using a vertical line in terms of the equivalent number of sub-carriers, N_{B_c} . We can see that the rate loss due to grouping is insignificant for group sizes $(N_c/K) < N_{B_c}$. Also, in this regime, the AM algorithm performs significantly better than ZF and MRT.

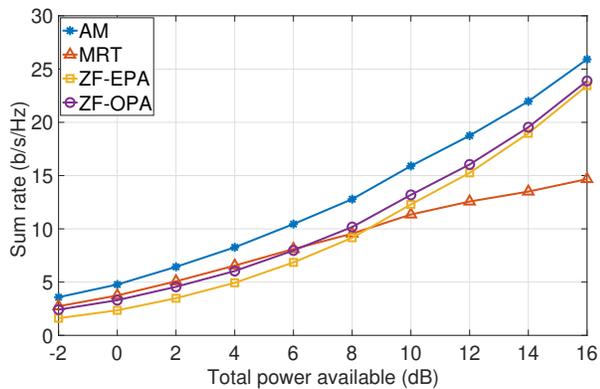


Fig. 11: Sum rate vs. Total transmit power: Average sum rate per slot for 2-cell 4-band MISO-IC with 10 users in each cell scheduled in a Round Robin fashion.

D. Multiple users in each cell

In Fig. 11, we show simulation results for a scenario where multiple users are present in each cell. For illustration, we use simple round-robin scheduling within a cell. More sophisticated schedulers can also be used along with our beamforming scheme. Each user is given full access to all frequency bands in the system during the allotted slot. The AM algorithm is repeated within each slot to find out the beamforming vectors. Fig. 11 shows the sum-rate obtained per time slot averaged over 100 slots and 1000 channel realizations. The sum rate is plotted against total power in the system. A 4-band, 2-cell MISO-IC is considered with 10 users in each cell. As expected, the use of the proposed AM beamformer provides a significant gain in sum rate.

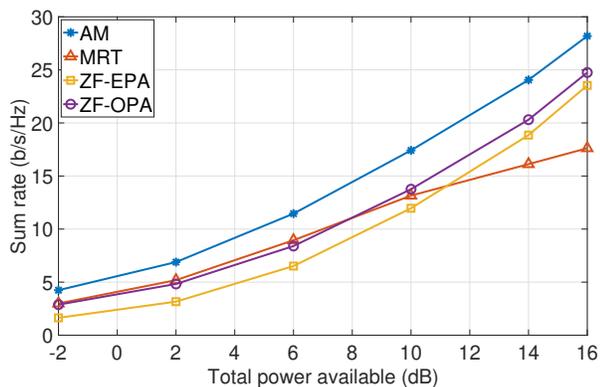


Fig. 12: Sum rate vs. Total transmit power for different beamforming schemes: 3-cell 3-band case

E. Three-Cell MISO-IC

In Fig. 12, we show the simulation results for a three-cell three-band MISO IC with 3 antennas at each BS. We use the solution in Section IV to find the optimum rate for a given Γ . The Γ that produces the rate on the boundary and the optimum power allocations are found out using the AM algorithm. Here, we plot the sum-rate as a function of the total power with different beamformers averaged over 30 channel realizations. Similar to what we saw with a 2-cell MISO-IC, we see that the AM beamformer has a significant advantage over the ZF and MRT beamformers. The ZF beamformer suffers from noise enhancement for low power and then outperforms the MRT for a higher power.

VIII. CONCLUDING REMARKS

In this paper, we considered a multi-band multi-cell MISO downlink, where the receivers treat interference from other cells as noise. We proposed the AM algorithm to determine the optimal beamformers for each band in each cell and the optimal power allocation across bands that achieve the weighted sum-rate optimal points on the Pareto boundary of the achievable rate region. The AM algorithm alternates between: (1) beamforming for WSR maximization in each band for a given power allocation across bands, and (2) power allocation across bands for a given set of beamformers. The algorithm can be implemented in a distributed manner with a limited exchange of interference constraint information among the BSs. For an M -cell system, the required exchange is proportional to $M - 1$. Furthermore, each BS needs to know only the channel state information for the links from itself to all the mobile stations. To develop the AM algorithm, we make use of the observation that overall Pareto optimality implies Pareto optimality in each band. The algorithm for the 2-cell and 3-cell settings is simplified significantly using analytical solutions to sub-problems. Significant performance gains are shown using simulations over equal power allocation across bands, ZF beamforming, and MRT beamforming. Multiple BSs can simultaneously proceed with the beamforming without waiting for the other BSs and can share the required information as and when they are ready. As every iteration of the AM algorithm satisfies the constraints and improves the achievable rate, the BS can also start

communicating with the MS without waiting for the completion of the algorithm. In a fading scenario, the optimal beamformers should be recalculated once every coherence interval. The complexity of the proposed AM algorithm can significantly be reduced for a multicarrier system with a large number of sub-carriers using sub-carrier grouping.

Even when the number of cells is more than three, the proposed method with analytical simplifications can still be used if the coordination cluster size is two or three. For coordination clusters with more than three cells, the AM algorithm can be used by solving the individual cell rate optimization under interference constraints numerically using standard convex optimization tools. Some interesting extensions of the work are to the MIMO setting, the setting with more than one active user in each cell, the setting with imperfect channel state information, and the setting with successive interference cancellation receivers at the mobile stations.

APPENDIX A

CHANNEL VECTORS USED FOR SIMULATION

See Tables I, II, III, and IV.

APPENDIX B

SIMPLIFICATION OF PROBLEM (22)

Consider the second and third constraints in (22). For a given choice of $\theta_1 - \theta_2 = \Delta$, we will derive the optimal choice ϕ for $\theta_2 - \theta_3$ that results in the largest feasible region for (22). One possible choice for Δ is

$$\Delta = \pi + A_1 - A_2, \quad (30)$$

which gives the most relaxed feasibility condition corresponding to the second constraint in (22) alone. However, in combination with the third constraint in (22), a different choice for Δ could give the best overall feasible region. Now, let

$$\begin{aligned} -B_2 + B_3 &= C, \\ \Delta - B_1 + B_3 &= D. \end{aligned} \quad (31)$$

Using the above definitions, the angles corresponding to the three cosine terms in the third constraint can be written, in order, as

$$\begin{aligned} (i) \quad \theta_1 - \theta_2 - B_1 + B_2 &= \Delta - B_1 + B_2, \\ (ii) \quad \theta_2 - \theta_3 - B_2 + B_3 &= \phi + C, \\ (iii) \quad \theta_1 - \theta_3 - B_1 + B_3 &= \phi + D. \end{aligned} \quad (32)$$

Note that the first term is a constant. To get the best feasible region, we need to minimize

$$\begin{aligned} &a \cos(\phi + C) + b \cos(\phi + D) \\ &= a \cos(\phi) \cos(C) - a \sin(\phi) \sin(C) \\ &\quad + b \cos(\phi) \cos(D) - b \sin(\phi) \sin(D) \\ &= \cos(\phi)(a \cos(C) + b \cos(D)) \\ &\quad - \sin(\phi)(a \sin(C) + b \sin(D)) \end{aligned}$$

where $a = 2\delta_1\zeta_1|a_{13}^{(2)}||a_{13}^{(3)}|$ and $b = 2\gamma_1\zeta_1|a_{13}^{(1)}||a_{13}^{(3)}|$ are the coefficients of the last two cosine terms. Define $a \cos(C) + b \cos(D) = x$, and $a \sin(C) + b \sin(D) = y$. After dividing by the constant $\sqrt{x^2 + y^2}$, now we need to minimize

$$\begin{aligned} &\frac{x}{\sqrt{x^2+y^2}} \cos(\phi) - \frac{y}{\sqrt{x^2+y^2}} \sin(\phi) \\ &= \cos\left(\phi + \cos^{-1}\left(\frac{x}{\sqrt{x^2+y^2}}\right)\right) \end{aligned}$$

We get the minimum value of -1 at $\phi = \pi - \cos^{-1}\left(\frac{x}{\sqrt{x^2+y^2}}\right)$. Thus, we have found the optimal ϕ for a given Δ .

REFERENCES

- [1] M. V. Narayanan and S. Bhashyam, "Pareto optimal distributed beamforming for the multi-band multi-cell downlink," in *GLOBECOM 2017 - 2017 IEEE Global Communications Conference*, Dec 2017, pp. 1–6.
- [2] E. Bjornson and E. Jorswieck, "Optimal resource allocation in coordinated multi-cell systems," *Foundations and Trends in Communications and Information Theory*, vol. 9, no. 2-3, pp. 113–381, 2013. [Online]. Available: <http://dx.doi.org/10.1561/01000000069>
- [3] E. Jorswieck, E. Larsson, and D. Danev, "Complete characterization of the pareto boundary for the MISO interference channel," *Signal Processing, IEEE Transactions on*, vol. 56, no. 10, pp. 5292–5296, Oct 2008.
- [4] R. Zhang and S. Cui, "Cooperative interference management with MISO beamforming," *IEEE Transactions on Signal Processing*, vol. 58, no. 10, pp. 5450–5458, Oct 2010.
- [5] J. Lindblom, E. Karipidis, and E. G. Larsson, "Closed-form parameterization of the pareto boundary for the two-user miso interference channel," in *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2011, pp. 3372–3375.
- [6] Q. Zhang, C. He, and L. Jiang, "Achieving maximum weighted sum-rate in multicell downlink miso systems," *IEEE communications letters*, vol. 16, no. 11, pp. 1808–1811, 2012.
- [7] J. Lindblom, E. Karipidis, and E. G. Larsson, "Efficient computation of pareto optimal beamforming vectors for the MISO interference channel with successive interference cancellation," *IEEE Transactions on Signal Processing*, vol. 61, no. 19, pp. 4782–4795, Oct 2013.
- [8] H. Song, J. Y. Ryu, and W. Choi, "Characterization of the pareto boundary for the two-user symmetric gaussian interference channel," *IEEE Transactions on Communications*, vol. 62, no. 8, pp. 2812–2824, Aug 2014.

TABLE I: Fig. 4a

Band	\mathbf{h}'_{11}	\mathbf{h}'_{12}	\mathbf{h}'_{21}	\mathbf{h}'_{22}
Band 1	[-0.66-0.93i -0.01+0.28i]	[-0.15+0.81i -0.74-0.46i]	[0.86+0.75i -0.31-0.35i]	[-0.15-0.70i 0.08+0.05i]
Band 2	[-0.97+0.59i -0.07+0.98i]	[-0.95-0.45i 0.03+0.71i]	[-0.70-0.36i -0.78-0.85i]	[0.18+0.93i -0.77-0.47i]

TABLE II: Fig. 4b

Band	\mathbf{h}'_{11}	\mathbf{h}'_{12}	\mathbf{h}'_{21}	\mathbf{h}'_{22}
Band 1	[0.14+0.66i 0.03-0.59i]	[-0.76-0.25i 0.15+0.19i]	[0.51+0.63i -0.07-0.56i]	[-0.26-0.36i 0.20+0.99i]
Band 2	[-0.79+0.12i 0.91-0.43i]	[0.98-0.47i -0.40+0.37i]	[-0.41-0.47i 0.81+0.45i]	[0.25-0.05i -0.88-0.42i]

TABLE III: Fig. 6a

Band	\mathbf{h}'_{11}	\mathbf{h}'_{12}	\mathbf{h}'_{21}	\mathbf{h}'_{22}
Band 1	[0.53+0.55i -0.47-0.83i]	[-0.32-0.94i -0.85+0.03i]	[-0.03+0.93i -0.11-0.09i]	[-0.36+0.39i -0.39-0.05i]
Band 2	[0.06-0.44i -0.09-0.69i]	[0.92-0.81i -0.76+0.94i]	[-0.21-0.05i -0.003-0.68i]	[0.57+0.24i -0.05-0.84i]
Band 3	[-0.61-0.18i -0.31+0.24i]	[-0.79-0.35i -0.17+0.16i]	[-0.16+0.21i -0.47+0.19i]	[-0.19+0.46i -0.47-0.13i]

TABLE IV: Fig. 6b

Band	\mathbf{h}_{11}	\mathbf{h}_{12}	\mathbf{h}_{21}	\mathbf{h}_{22}
Band 1	[0.06 + 0.78i 0.98 0.74i]	[-0.27 - 0.13i 0.52 0.69i]	[-0.84 + 0.53i -0.39 - 0.4i]	[0.07 - 0.24i 0.05 - 0.4i]
Band 2	[-0.91 - 0.77i -0.70 0.67i]	[0.46 - 0.25i -0.04 0.79i]	[-0.26 + 0.17i 0.29 + 0.6i]	[0.54 - 0.70i -0.57 - 0.0i]
Band 3	[-0.60 - 0.67i -0.36 0.97i]	[-0.73 - 0.95i 0.89 0.08i]	[0.24 + 0.19i 0.98 - 0.5i]	[0.54 - 0.92i 0.91 - 0.1i]
Band 4	[-0.89 + 0.74i 0.68 0.84i]	[-0.92 + 0.03i -0.73 0.54i]	[0.71 + 0.29i -0.33 - 0.1i]	[0.27 - 0.30i 0.69 + 0.1i]
Band 5	[0.57 - 0.89i 0.26 0.30i]	[0.01 + 0.42i 0.01 0.27i]	[0.24 - 0.43i -0.10 - 0.3i]	[0.56 + 0.38i -0.05 + 0.0i]
Band 6	[-0.71 + 0.31i -0.94 0.83i]	[0.22 - 0.37i 0.99 0.89i]	[-0.07 + 0.17i 0.41 + 0.1i]	[0.14 - 0.67i 0.39 - 0.9i]
Band 7	[0.00 - 0.42i 0.85 0.01i]	[0.06 - 0.51i -0.21 0.46i]	[-0.41 - 0.53i 0.78 - 0.2i]	[0.37 - 0.87i 0.87 - 0.1i]
Band 8	[0.83 + 0.11i 0.21 0.52i]	[-0.29 + 0.96i -0.20 0.09i]	[0.30 - 0.56i 0.55 - 0.6i]	[0.94 - 0.49i -0.99 - 0.6i]
Band 9	[-0.94 + 0.66i -0.90 0.38i]	[0.56 + 0.12i 0.70 0.83i]	[0.65 + 0.85i 0.72 + 0.3i]	[0.45 - 0.32i 0.73 + 0.6i]
Band 10	[0.67 + 0.64i 0.46 0.35i]	[0.95 - 0.28i 0.31 0.00i]	[0.76 - 0.83i -0.40 - 0.4i]	[0.74 - 0.53i 0.38 - 0.8i]

- [9] J. Park and Y. Sung, "On the pareto-optimal beam structure and design for multi-user MIMO interference channels," *IEEE Transactions on Signal Processing*, vol. 61, no. 23, pp. 5932–5946, Dec 2013.
- [10] R. Mochaourab, P. Cao, and E. Jorswieck, "Alternating rate profile optimization in single stream MIMO interference channels," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, May 2013, pp. 4834–4838.
- [11] P. Cao, E. A. Jorswieck, and S. Shi, "Pareto boundary of the rate region for single-stream MIMO interference channels: Linear transceiver design," *IEEE Transactions on Signal Processing*, vol. 61, no. 20, pp. 4907–4922, Oct 2013.
- [12] R. Mochaourab, P. Cao, and E. Jorswieck, "Alternating rate profile optimization in single stream MIMO interference channels," *IEEE Signal Processing Letters*, vol. 21, no. 2, pp. 221–224, Feb 2014.
- [13] E. Bjornson, G. Zheng, M. Bengtsson, and B. Ottersten, "Robust monotonic optimization framework for multicell MISO systems," *IEEE Transactions on Signal Processing*, vol. 60, no. 5, pp. 2508–2523, May 2012.
- [14] E. Stathakis, J. Jalden, L. K. Rasmussen, and M. Skoglund, "Outage region characterization for beamforming in MISO interference networks with imperfect CSI," *IEEE Signal Processing Letters*, vol. 22, no. 12, pp. 2378–2382, Dec 2015.
- [15] Y. Noam, N. Kimelfeld, and B. M. Zaidel, "On the two-user MISO interference channel with single user decoding and partial CSIT," in *2016 IEEE International Symposium on Information Theory (ISIT)*, July 2016, pp. 720–724.
- [16] Y. Noam and B. M. Zaidel, "On the two-user MISO interference channel with single-user decoding: Impact of imperfect CSIT and channel dimension reduction," *IEEE Transactions on Signal Processing*, vol. 67, no. 10, pp. 2608–2623, May 2019.
- [17] P. C. Weeraddana, M. Codreanu, M. Latva-aho, and A. Ephremides, "Multicell MISO downlink weighted sum-rate maximization: A distributed approach," *IEEE Transactions on Signal Processing*, vol. 61, no. 3, pp. 556–570, Feb 2013.
- [18] Q. Zhang, C. He, and L. Jiang, "Achieving maximum weighted sum-rate in multicell downlink MISO systems," *IEEE Communications Letters*, vol. 16, no. 11, pp. 1808–1811, November 2012.
- [19] L. Liu, R. Zhang, and K. C. Chua, "Achieving global optimality for weighted sum-rate maximization in the K-User gaussian interference channel with multiple antennas," *IEEE Transactions on Wireless Communications*, vol. 11, no. 5, pp. 1933–1945, May 2012.
- [20] H. Kim, H. Lee, L. Duan, and I. Lee, "Sum-rate maximization methods for wirelessly powered communication networks in interference channels," *IEEE Transactions on Wireless Communications*, vol. 17, no. 10, pp. 6464–6474, Oct 2018.
- [21] O. Simeone, O. Somekh, H. V. Poor, and S. Shamai (Shitz), "Downlink multicell processing with limited-backhaul capacity," *EURASIP Journal on Advances in Signal Processing*, vol. 2009, no. 1, p. 840814, Jun 2009. [Online]. Available: <https://doi.org/10.1155/2009/840814>
- [22] Y. Kim, H. J. Yang, and H. Jwa, "Multicell downlink beamforming with limited backhaul signaling," *IEEE Access*, vol. 6,

- pp. 64 122–64 130, 2018.
- [23] H. J. Kwon, J. H. Lee, and W. Choi, “Machine learning-based beamforming in two-user MISO interference channels,” in *2019 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC)*, Feb 2019, pp. 496–499.
 - [24] X. Yang and A. L. Swindlehurst, “Limited rate feedback for two-user MISO gaussian interference channel with and without secrecy,” *IEEE Transactions on Signal Processing*, vol. 66, no. 18, pp. 4884–4897, Sep. 2018.
 - [25] C. Li, C. He, and L. Jiang, “Distributed beamforming design of the pareto boundary for miso interference channels,” in *2014 IEEE/CIC International Conference on Communications in China (ICCC)*. IEEE, 2014, pp. 748–752.
 - [26] W. Xu and X. Wang, “Pricing-based distributed downlink beamforming in multi-cell ofdma networks,” *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 9, pp. 1605–1613, 2012.
 - [27] F. R. Guimaraes, G. Fodor, W. C. Freitas, and Y. C. Silva, “Pricing-based distributed beamforming for dynamic time division duplexing systems,” *IEEE Transactions on Vehicular Technology*, vol. 67, no. 4, pp. 3145–3157, 2017.
 - [28] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
 - [29] A. Goldsmith, *Wireless communications*. Cambridge university press, 2005.