Jointly Optimizing Stream Allocation, Beamforming and Combining Weights for the MIMO Interference Channel

Luis Miguel Cortés-Peña, John R. Barry, *Senior Member, IEEE*, and Douglas M. Blough, *Senior Member, IEEE*

***Abstract*—We propose an algorithm whose goal is to maximize the sum rate of a set of interfering multiple-input multiple-output (MIMO) links by jointly optimizing which subset of transmitters should transmit, the number of streams for each transmitter (if any), and the beamforming and combining weights that support those streams. We present numerical results to illustrate that our algorithm achieves a sum rate higher than previously reported algorithms at high interference, and that it achieves comparable performance to the top-performing algorithms at medium and low interference. In one high-interference example with many links, our algorithm achieves a 65% higher sum rate than previously reported algorithms.**

***Index Terms*—Joint transceiver, minimum weighted sum mean-squared-error, sum rate.**

I. Introduction

THE performance of a wireless link can be improved without the need of additional spectrum or power by equipping the transmitter and the receiver with multiple antennas [1], leading to a multiple-input multiple-output (MIMO) link. A single MIMO link, in the absence of interference, can perform spatial multiplexing to transmit multiple streams in parallel. In the MIMO interference channel, where multiple interfering MIMO links are active simultaneously, MIMO links can perform a combination of spatial multiplexing and interference suppression, so that each transmitter can send multiple streams that can be decoded reliably and independently by their intended receivers. Computing the beamforming and combining weights that maximize the aggregate performance in the MIMO interference channel is, however, complicated by their inherent interdependence.

Manuscript received August 3, 2013; revised December 24, 2013, May 25, 2014, and October 8, 2014; accepted December 2, 2014. Date of publication December 18, 2014; date of current version April 7, 2015. This research was supported in part by Grants CNS-1017248 and CNS-1319455 from the National Science Foundation. The associate editor coordinating the review of this paper and approving it for publication was X. Zhou.

L. M. Cortés-Peña is with the Government Communications Systems Division, Harris Corporation, Melbourne, FL 32902-0037 USA (e-mail: cortes@gatech.edu).

J. R. Barry and D. M. Blough are with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0250 USA (e-mail: barry@ece.gatech.edu; doug.blough@ece.gatech.edu).

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

Digital Object Identifier 10.1109/TWC.2014.2384021

In this paper, we tackle the problem of maximizing the sum rate of a set of interfering MIMO links. This problem includes the problems of determining which subset of transmitters should transmit, how many streams each transmitter should send (if any), and the corresponding beamforming and combining weights for each stream.

Most prior work has focused on a subset of these problems. The Max-SINR algorithm from [2], the bilateral algorithm from [3], and the Max-SINR algorithm from [4] are examples of algorithms that only compute the beamforming and combining weights; they require *a priori* specification of which transmitters should transmit as well as how many streams each should transmit.

Other work has focused on the joint problem of determining the beamforming and combining weights (or transmit covariance matrices), and the number of streams that each transmitter should transmit. Examples include the MWSR algorithm from [5], the GP algorithm from [6], the greedy algorithm from [7], and the BR algorithm from [8]. All of these works require a separate algorithm to determine which transmitters should transmit.

Other work has focused on determining which transmitters should transmit, and how many streams each should transmit. The works of [9]–[15], for example, propose either linear programming formulations or heuristic algorithms that determine which transmitters should transmit and how many streams each should transmit. The proposed solutions of these works require a separate algorithm to compute the beamforming and combining weights and do not necessarily maximize the sum rate, since the decisions are not based on the beamforming and combining weights of the links, which ultimately determine the performance.

Our main contributions are as follows.

- As a stepping stone, we first extend the seminal results of Sampath *et al.* [16] to design, for a single link in the presence of interfering links, the joint optimal beamforming and combining weights that minimize the sum weighted mean-squared-error (MSE) across all streams in the network given the beamforming and combining weights of the interfering links. The transceiver has the following characteristics:
  - — it has the ability to deactivate (or activate) itself if doing so minimizes the weighted sum MSE;
  - — it diagonalizes the MIMO channel; and

— it reduces to the optimal eigen-mode transmission with power allocated through waterfilling for the special case where the transmitter causes no interference to any receiver.

- We provide an interpretation to the structure of the transceiver and show that the optimal combiner is a minimum mean-squared-error (MMSE) combiner, and that the optimal beamformer can be viewed as the MMSE combiner of a virtual network.
- Using our transceiver and the results from [5], [16], [17] that relate the minimum weighted sum MSE to the maximum weighted sum rate, we design a suboptimal algorithm that finds a local maximum on the sum rate. The algorithm has the following characteristics:
  - — it determines which subset of transmitters should transmit;
  - — it uses only local information, a required property for any distributed implementation;
  - — it outperforms previously reported algorithms at high interference; and
  - — it achieves comparable performance to the top-performing previously reported algorithms at medium and low interference.

This paper is organized as follows. In Section II, we define our model for the physical layer. In Section III, we design the joint beamforming and combining weights for a single link. In Section IV, we review the relationship between the sum rate and the minimum weighted sum MSE. In Section V, we propose our algorithm for computing the beamforming and combining weights for all links. In Section VI, we present numerical results. Finally, in Section VII, we present our conclusions. A preliminary version of this paper was published as a conference paper [18].

II. Physical-Layer Model

Consider a set of M half-duplex links. Let d_k be the number of multiplexed streams on link k, and let n_{t_k} and n_{r_k} be the number of antenna elements at the transmitter and receiver of link k, respectively. Let $\boldsymbol{H}_{kl} \in \mathbb{C}^{n_{r_k} \times n_{t_l}}$ be the matrix of complex channel gains between the antennas of transmitter l and those of receiver k.

The received vector at receiver k is given by

$$\boldsymbol{y}_k = \boldsymbol{H}_{kk}\boldsymbol{V}_k\boldsymbol{x}_k + \underbrace{\sum_{l=1,l\neq k}^{M} \boldsymbol{H}_{kl}\boldsymbol{V}_l\boldsymbol{x}_l + \boldsymbol{n}_k}_{\boldsymbol{z}_k}, \tag{1}$$

where $\boldsymbol{V}_k \in \mathbb{C}^{n_{t_k} \times d_k}$ is the beamforming matrix of transmitter k; $\boldsymbol{x}_k \in \mathbb{C}^{d_k}$ is the transmit signal vector from transmitter k, assumed to be independently encoded Gaussian codebook symbols with unit-energy so that $\mathbb{E}[\boldsymbol{x}_k\boldsymbol{x}_k^\dagger] = \boldsymbol{I}$, where $(\cdot)^\dagger$ is the conjugate transpose of $(\cdot)$; $\boldsymbol{n}_k \in \mathbb{C}^{n_{r_k}}$ is a vector of Gaussian noise elements with covariance matrix $\mathbb{E}[\boldsymbol{n}_k\boldsymbol{n}_k^\dagger] = \boldsymbol{R}_{\boldsymbol{n}_k}$; and $\boldsymbol{z}_k \in \mathbb{C}^{n_{r_k}}$ is the total received interference plus noise with covariance matrix

$$\boldsymbol{R}_{\bar{k}} = \mathbb{E}\left[\boldsymbol{z}_k\boldsymbol{z}_k^\dagger\right] = \sum_{l=1,l\neq k}^{M} \boldsymbol{H}_{kl}\boldsymbol{V}_l\boldsymbol{V}_l^\dagger\boldsymbol{H}_{kl}^\dagger + \boldsymbol{R}_{\boldsymbol{n}_k}. \tag{2}$$

In order to meet a power constraint of p_k, the beamforming weights for transmitter k must satisfy $\mathrm{tr}(\boldsymbol{V}_k\boldsymbol{V}_k^\dagger) \leq p_k$. The instantaneous capacity in bits/sec/Hz of link k before combining is given by [19]

$$C_k = \log_2\left|\boldsymbol{I} + \boldsymbol{R}_{\bar{k}}^{-1}\boldsymbol{H}_{kk}\boldsymbol{V}_k\boldsymbol{V}_k^\dagger\boldsymbol{H}_{kk}^\dagger\right|. \tag{3}$$

After combining, the received signal at receiver k is given by

$$\hat{\boldsymbol{x}}_k = \boldsymbol{U}_k^\dagger\boldsymbol{y}_k, \tag{4}$$

where $\boldsymbol{U}_k \in \mathbb{C}^{n_{r_k} \times d_k}$ is the combining matrix of receiver k, and the instantaneous capacity in bits/sec/Hz is given by

$$\hat{C}_k = \log_2\left|\boldsymbol{I} + \left(\boldsymbol{U}_k^\dagger\boldsymbol{R}_{\bar{k}}\boldsymbol{U}_k\right)^{-1}\boldsymbol{U}_k^\dagger\boldsymbol{H}_{kk}\boldsymbol{V}_k\boldsymbol{V}_k^\dagger\boldsymbol{H}_{kk}^\dagger\boldsymbol{U}_k\right|. \tag{5}$$

III. Jointly Computing the Beamforming and Combining Weights for a Single Link

We begin by optimizing the beamforming and combining weighs for a single link in the presence of interfering links as a stepping stone towards computing the beamforming and combining weights for all links. In this section, we design the joint transceiver that minimizes the weighted sum MSE across all streams in the network. We choose the weighted sum MSE criterion because, as we will review in Section IV, it reduces to maximizing the sum rate as a special case [5], [16], [17]. Later, in Section V, we will use the joint transceiver to design an algorithm that computes the beamforming and combining weights of all links.

A. The Minimum Weighted Sum MSE Problem

We formulate the weighted sum MSE optimization for the beamforming and combining weights of link k as

$$(\boldsymbol{V}_k^*, \boldsymbol{U}_k^*) = \arg\min_{\boldsymbol{V}_k, \boldsymbol{U}_k} \sum_{l=1}^{M} \mathrm{tr}(\boldsymbol{W}_l\boldsymbol{E}_l) \quad \text{such that } \mathrm{tr}\left(\boldsymbol{V}_k\boldsymbol{V}_k^\dagger\right) \leq p_k, \tag{6}$$

where

$$\boldsymbol{E}_k = \mathbb{E}\left[(\hat{\boldsymbol{x}}_k - \boldsymbol{x}_k)(\hat{\boldsymbol{x}}_k - \boldsymbol{x}_k)^\dagger\right] \tag{7}$$

is the error covariance matrix of link k and contains the MSE of the streams of link k in the diagonal. In (6), the error weight matrix $\boldsymbol{W}_k \in \mathbb{R}^{d_k \times d_k}$ is a diagonal matrix of nonnegative weights associated with the MSE of the streams of link k.

The objective function in (6) is similar to that of [5]. In this paper, however, we take a unique approach to solving (6) in that we solve for the beamforming and combining weights simultaneously, so that neither is a function of the other. Together

with the power inequality constraint $\mathrm{tr}(\boldsymbol{V}_k\boldsymbol{V}_k^\dagger) \le p_k$ in (6), the resulting transceiver can jointly optimize whether the link should be activated or deactivated, the number of streams it should transmit (if any), and the supporting beamforming and combining weights. As we will see using numerical results in Section VI, deactivating links is desirable at high interference since interference caused by one link highly affects the performance of all other links in the network.

B. The Minimum Weighted Sum MSE Solution

The solution to (6) can be expressed in terms of the following compact singular-value decomposition (SVD):

$$\boldsymbol{R}_{\bar{k}}^{-1/2}\boldsymbol{H}_{kk}\boldsymbol{P}_{\bar{k}}^{-1/2} = \boldsymbol{F}_k\boldsymbol{D}_k\boldsymbol{G}_k^\dagger, \tag{8}$$

where $\boldsymbol{D}_k \in \mathbb{R}^{d_k\times d_k}$ is a diagonal matrix containing the nonzero singular values of $\boldsymbol{R}_{\bar{k}}^{-1/2}\boldsymbol{H}_{kk}\boldsymbol{P}_{\bar{k}}^{-1/2}$ ordered in decreasing order from top left to bottom right; $\boldsymbol{F}_k \in \mathbb{C}^{n_{r_k}\times d_k}$ and $\boldsymbol{G}_k \in \mathbb{C}^{n_{t_k}\times d_k}$ have orthonormal column vectors that correspond to the left and right eigenvectors of $\boldsymbol{R}_{\bar{k}}^{-1/2}\boldsymbol{H}_{kk}\boldsymbol{P}_{\bar{k}}^{-1/2}$ with nonzero singular values, respectively; and $\boldsymbol{P}_{\bar{k}}$ is

$$\boldsymbol{P}_{\bar{k}} = \underbrace{\sum_{l=1,l\neq k}^{M} \boldsymbol{H}_{lk}^\dagger\boldsymbol{U}_l\boldsymbol{W}_l\boldsymbol{U}_l^\dagger\boldsymbol{H}_{lk}}_{\boldsymbol{\Pi}_k} + \mu_k\boldsymbol{I}, \tag{9}$$

where $\mu_k \ge 0$ is a Lagrange multiplier that must satisfy the Karush-Kuhn-Tucker (KKT) conditions for the optimization of (6) (see Appendix A).

The SVD in (8) requires that $\boldsymbol{P}_{\bar{k}}$ be invertible, which is clearly true whenever $\mu_k > 0$. In the following, we present the beamforming and combining weights that solve (6) by assuming that we already know μ_k and that $\boldsymbol{P}_{\bar{k}}$ is invertible. Later in this section, we show how to obtain μ_k, and show that $\boldsymbol{P}_{\bar{k}}$ is always invertible.

Theorem 1: The joint beamforming and combining weights that solve (6) are given by

$$\boldsymbol{V}_k = \boldsymbol{P}_{\bar{k}}^{-1/2}\boldsymbol{G}_k\boldsymbol{\Theta}_k, \tag{10}$$

$$\boldsymbol{U}_k = \boldsymbol{R}_{\bar{k}}^{-1/2}\boldsymbol{F}_k\boldsymbol{\Phi}_k, \tag{11}$$

where

$$\boldsymbol{\Theta}_k = \left(\boldsymbol{W}_k^{1/2}\boldsymbol{D}_k^{-1} - \boldsymbol{D}_k^{-2}\right)_+^{1/2} \tag{12}$$

$$\boldsymbol{\Phi}_k = \boldsymbol{W}_k^{-1/2}\boldsymbol{\Theta}_k, \tag{13}$$

and $(\cdot)_+$ is the matrix $(\cdot)$ with the negative entries replaced with zeros.

Proof: See Appendix A. □

Using Theorem 1, we can rewrite (4) as

$$\hat{\boldsymbol{x}}_k = \boldsymbol{\Phi}_k(\boldsymbol{D}_k\boldsymbol{\Theta}_k\boldsymbol{x}_k + \hat{\boldsymbol{n}}_k), \tag{14}$$

where $\hat{\boldsymbol{n}}_k = \boldsymbol{F}_k^\dagger\boldsymbol{R}_{\bar{k}}^{-1/2}\boldsymbol{z}_k$ is a vector of white Gaussian noise satisfying

$$\mathbb{E}\left[\hat{\boldsymbol{n}}_k\hat{\boldsymbol{n}}_k^\dagger\right] = \boldsymbol{F}_k^\dagger\boldsymbol{R}_{\bar{k}}^{-1/2}\boldsymbol{R}_{\bar{k}}\boldsymbol{R}_{\bar{k}}^{-1/2}\boldsymbol{F}_k = \boldsymbol{I}.$$

1: **if** $\boldsymbol{\Pi}_k$ is invertible **then**
2: Set $\mu_k = 0$;
3: **if** $\mathrm{tr}(\boldsymbol{V}_k\boldsymbol{V}_k^\dagger) \le p_k$ **then**
4: **return** μ_k;
5: **end if**
6: **end if**
7: Find $\mu_k > 0$ such that $\mathrm{tr}(\boldsymbol{V}_k\boldsymbol{V}_k^\dagger) = p_k$;
8: **return** μ_k;

Fig. 1. Pseudocode for finding μ_k.

Therefore, the beamforming and combining weights in (10) and (11) diagonalize the MIMO channel.

To complete the solution to (6), we must find μ_k. Because a closed-form solution for μ_k is unknown, and because $\mathrm{tr}(\boldsymbol{V}_k\boldsymbol{V}_k^\dagger)$ is a decreasing function of μ_k [4], [20], we search for the value of μ_k as follows.

First, let us consider the case in which $\boldsymbol{\Pi}_k$ in (9) is invertible. For this case, we first test for $\mu_k = 0$. If the beamforming weights satisfy $\mathrm{tr}(\boldsymbol{V}_k\boldsymbol{V}_k^\dagger) \le p_k$, then the search is done because all KKT conditions are satisfied. If $\mathrm{tr}(\boldsymbol{V}_k\boldsymbol{V}_k^\dagger) > p_k$, then we search for the $\mu_k > 0$ such that $\mathrm{tr}(\boldsymbol{V}_k\boldsymbol{V}_k^\dagger) = p_k$, thereby satisfying all KKT conditions.

Now, let us consider the case in which $\boldsymbol{\Pi}_k$ is singular. For this case, we state the following lemma.

Lemma 1: If $\boldsymbol{\Pi}_k$ in (9) is singular, then the limit of $\mathrm{tr}(\boldsymbol{V}_k\boldsymbol{V}_k^\dagger)$ using $\boldsymbol{V}_k$ in (10) as μ_k approaches zero from the right is

$$\lim_{\mu_k\to 0^+} \mathrm{tr}\left(\boldsymbol{V}_k\boldsymbol{V}_k^\dagger\right) = \infty. \tag{15}$$

Proof: See Appendix B. □

Lemma 1 together with the fact that $\mathrm{tr}(\boldsymbol{V}_k\boldsymbol{V}_k^\dagger)$ is a decreasing function of μ_k suggest that whenever $\boldsymbol{\Pi}_k$ is singular, there exist a $\mu_k > 0$ such that $\mathrm{tr}(\boldsymbol{V}_k\boldsymbol{V}_k^\dagger) = p_k$, which satisfies all KKT conditions. Therefore, $\boldsymbol{P}_{\bar{k}}$ is always invertible.

Fig. 1 summarizes the algorithm for finding μ_k for the general case. In our simulations, we use the bisection method to perform the search for $\mu_k > 0$.

C. Interpreting the Solution

The beamforming and combining weights in (10) and (11) have three components that can be inter-related through the use of a virtual network in which receivers become virtual transmitters and transmitters become virtual receivers. The concept of a virtual network has been previously used to aid the design of the transmitter's beamforming weights in the works of [2], [3], [21].

To build the virtual network that relates (10) and (11), let us define $\overleftarrow{\boldsymbol{H}}_{lk} = \boldsymbol{H}_{kl}^\dagger$ as the virtual MIMO channel between the virtual transmitter of link k and the virtual receiver of link l; $\overleftarrow{\boldsymbol{V}}_k = \boldsymbol{U}_k\boldsymbol{W}_k^{1/2}$ as the virtual beamforming weights of link k; $\overleftarrow{\boldsymbol{U}}_k = \boldsymbol{V}_k\boldsymbol{W}_k^{-1/2}$ as the virtual combining weights of link k; and $\overleftarrow{\boldsymbol{R}}_{n_k} = \mu_k\boldsymbol{I}$ as the virtual noise covariance of the virtual receiver of link k.

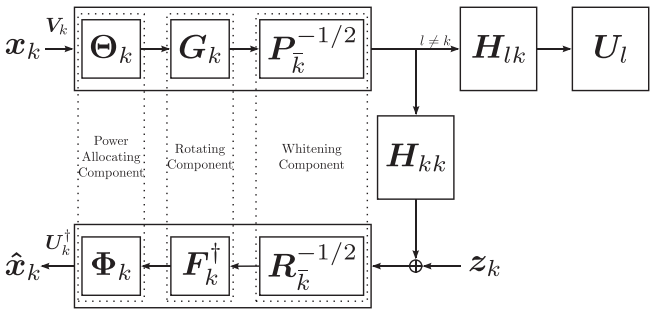

Fig. 2. Block diagram of the components of the joint beamforming and combining weights.

Fig. 2 shows a block diagram of the transmit and receive structure for link k and highlights the three components of the joint transceiver using dotted boxes. The three components and their functions are as follows:

- *Whitening Component*—The first component of the transceiver is a whitening component. At the receiver side, the receiver whitens the interference plus noise of the received signal. At the transmitter side, this whitening component performs a similar function by whitening a "virtual" interference plus noise with covariance matrix given by

$$\overleftarrow{\boldsymbol{R}}_{\bar{k}} = \sum_{l=1, l\neq k}^{M} \overleftarrow{\boldsymbol{H}}_{kl} \overleftarrow{\boldsymbol{V}}_l \overleftarrow{\boldsymbol{V}}_l^{\dagger} \overleftarrow{\boldsymbol{H}}_{kl}^{\dagger} + \overleftarrow{\boldsymbol{R}}_{\boldsymbol{n}_k} = \boldsymbol{P}_{\bar{k}}. \tag{16}$$

- *Rotating Component*—The second component is a rotating component. Using this component, both the transmitter and the receiver rotate their signal so as to diagonalize their MIMO channel. The rotating matrices are chosen based on the SVD of the cascade of the whitening components and the MIMO channel ($\boldsymbol{R}_{\bar{k}}^{-1/2} \boldsymbol{H}_{kk} \boldsymbol{P}_{\bar{k}}^{-1/2}$).
- *Power Allocating Component*—The third component is a power allocating component that scales each element of the signal vector. Due to the $(\cdot)_+$ operator in (12), the power allocating component at the transmitter can potentially prevent some, if not all, streams from being transmitted. On the virtual network, the receiver's power allocating component acts similarly to the transmitter's power allocating component by scaling some signal elements and even reducing the number of streams on the virtual link.

The three components operate jointly to determine the number of streams to allocate to the link (if any). After whitening the channel, the rotating components provide the appropriate rotating matrices to use for communication through the MIMO channel. Then, the power allocating components allocate non-zero power to the link's i^{th} stream if $(\boldsymbol{W}_k)_{ii}^{1/2} > (\boldsymbol{D}_k)_{ii}^{-1}$ and allocate zero power otherwise, where $(\cdot)_{ij}$ is the element at the i^{th} row and j^{th} column of $(\cdot)$. Clearly, the higher the values of $(\boldsymbol{W}_k)_{ii}^{1/2}$ and $(\boldsymbol{D}_k)_{ii}$ are, the more likely that $(\boldsymbol{W}_k)_{ii}^{1/2} > (\boldsymbol{D}_k)_{ii}^{-1}$ will be satisfied. Note that the values of $\boldsymbol{D}_k$ in (8) are determined according to the whitening components. Also note that decreasing the values of $\boldsymbol{R}_{\bar{k}}$ and $\boldsymbol{P}_{\bar{k}}$ results in higher values of $\boldsymbol{D}_k$. Therefore, the threshold values of $\boldsymbol{D}_k$ used to determine whether or not to allocate zero power are determined by the received interference plus noise at the receiver ($\boldsymbol{R}_{\bar{k}}$), the interference caused by the transmitter to other receivers ($\boldsymbol{\Pi}_k$), and the power available to the beamformer (determined by μ_k).

Using the MMSE combining weights for link k, as given by [4], [5], [20]

$$\boldsymbol{U}_k^{\mathrm{MMSE}} = \left(\boldsymbol{H}_{kk} \boldsymbol{V}_k \boldsymbol{V}_k^{\dagger} \boldsymbol{H}_{kk}^{\dagger} + \boldsymbol{R}_{\bar{k}}\right)^{-1} \boldsymbol{H}_{kk} \boldsymbol{V}_k, \tag{17}$$

we can further relate the joint beamforming and combining weights, as described by the following lemmas.

Lemma 2: If the beamforming weights for link k are given by (10), then the MMSE combining weights for link k are equal to (11).

Lemma 3: If the combining weights for link k are given by (11), then the MMSE combining weights for link k of the virtual network are $\overleftarrow{\boldsymbol{U}}_k^{\mathrm{MMSE}} = \boldsymbol{P}_{\bar{k}}^{-1/2} \boldsymbol{G}_k \boldsymbol{\Phi}_k$, and so the beamforming weights of the real network are given by (10).

Proof: See Appendix C for the proof of Lemma 3. We omit the proof of Lemma 2 because of its similarity to the proof of Lemma 3. □

Remark 1: By the data processing theorem from information theory, we know that $C_k \geq \hat{C}_k$ for any choice of $\boldsymbol{U}_k$. Using the matrix inversion lemma, it is easy to show that the MMSE combining weights in (17) are information lossless, so that $C_k = \hat{C}_k$ [5]. Therefore, from Lemma 2, it is clear that the combining weights given by (11) are also information lossless.

IV. The Weighted Sum MSE and the Sum Rate

We have chosen the weighted sum MSE as our objective function because, with a proper choice of the error weight matrix $\boldsymbol{W}_k$, minimizing the weighted sum MSE also maximizes the sum rate. This relationship was exploited for the single MIMO link in the absence of interference by Sampath *et al.* in [16], for the MIMO broadcast channel by Christensen *et al.* in [17], and for the MIMO interference channel by Negro *et al.* in [5]. In the following sections, we discuss how to choose the error weight matrices that maximize the sum rate for two separate cases.

A. Case 1: The MIMO Link That Causes No Interference, but Receives Interference

In [16], Sampath *et al.* showed that for a single MIMO link in the absence of interference, minimizing the weighted MSE will also maximize the rate on the link when the error weights $\boldsymbol{W}$ are chosen appropriately. In this section, we generalize this result to include the case in which the receiver is subject to interference from other links.

Consider the case where we wish to maximize the rate on a MIMO link k in which the transmitter causes no interference to any receiver, but in which the receiver is interfered by other transmitters. For this case, the solution is to transmit through the eigen modes of the whitened channel ($\boldsymbol{R}_{\bar{k}}^{-1/2} \boldsymbol{H}_{kk}$) and allocate power through waterfilling the same way that a greedy transmitter does [6], [7]. We find that for this scenario, the $\boldsymbol{W}_k$ that maximizes the rate on the link when optimizing (6) is given

by $\boldsymbol{W}_k = \alpha_k \boldsymbol{\Lambda}_k$, where $\boldsymbol{\Lambda}_k^{1/2} \in \mathbb{R}^{d_k \times d_k}$ is a diagonal matrix containing the singular values of $\boldsymbol{R}_{\bar{k}}^{-1/2}\boldsymbol{H}_{kk}$ in decreasing order from top left to bottom right and α_k is any positive scalar. The following proposition summarizes this result.

Proposition 1: For the case when the transmitter of link k causes no interference to other receivers but the receiver of link k is subject to interference from other transmitters and $\boldsymbol{W}_k = \alpha_k \boldsymbol{\Lambda}_k$ for any scalar $\alpha_k > 0$, the beamforming and combining weights in (10) and (11) reduce to the optimal eigenmode transmission with power allocated through waterfilling.

Proof: See Appendix D. □

B. Case 2: The MIMO Link That Causes and Receives Interference

Consider the general case of the MIMO interference channel where all transmitters interfere with all receivers. In [5], the authors find that the gradient of the sum rate and the gradient of the weighted sum MSE are equal if

$$\boldsymbol{W}_k = \boldsymbol{I} + \boldsymbol{B}_k^{\dagger}\boldsymbol{V}_k^{\dagger}\boldsymbol{H}_{kk}^{\dagger}\boldsymbol{R}_{\bar{k}}^{-1}\boldsymbol{H}_{kk}\boldsymbol{V}_k\boldsymbol{B}_k, \tag{18}$$

where $\boldsymbol{B}_k$ is an arbitrary unitary matrix.

Notice that $\boldsymbol{W}_k$ in (18) is a function of $\boldsymbol{V}_k$, which is itself one of the variables to optimize. To solve this interdependency, the authors of [5] propose to compute $\boldsymbol{W}_k$ and $\boldsymbol{V}_k$ in separate steps in an iterative algorithm. We follow the same approach.

In our formulation of (6), we require that $\boldsymbol{W}_k$ be diagonal. To guarantee that (18) is always diagonal, we choose $\boldsymbol{B}_k$ in (18) from the following SVD:

$$\boldsymbol{A}_k\boldsymbol{S}_k\boldsymbol{B}_k^{\dagger} = \boldsymbol{R}_{\bar{k}}^{-1/2}\boldsymbol{H}_k\boldsymbol{V}_k, \tag{19}$$

where $\boldsymbol{S}_k \in \mathbb{R}^{d_k \times d_k}$ is a diagonal matrix containing the singular values of $\boldsymbol{R}_{\bar{k}}^{-1/2}\boldsymbol{H}_k\boldsymbol{V}_k$ ordered in decreasing order from top left to bottom right; $\boldsymbol{B}_k \in \mathbb{C}^{d_k \times d_k}$ is a unitary matrix; and $\boldsymbol{A}_k \in \mathbb{C}^{n_{r_k} \times d_k}$ has orthonormal column vectors. This way, (18) becomes

$$\boldsymbol{W}_k = \boldsymbol{I} + \boldsymbol{S}_k^2. \tag{20}$$

This choice of $\boldsymbol{B}_k$ was used by Christensen *et al.* in [17] to design the WSRBF-WMMSE-D algorithm with diagonal weighting matrix for the MIMO broadcast channel.

V. Computing the Beamforming and Combining Weights for All Links

We now propose an algorithm whose goal is to maximize the sum rate of a set of interfering MIMO links by jointly optimizing the number of streams (if any) on each link as well as their beamforming and combining weights.

In the following sections, we will first describe the proposed algorithm, then we will describe the computational complexity of the proposed algorithm.

1: Initialize $d_k = \text{rank}\,(\boldsymbol{H}_{kk})$, $\boldsymbol{V}_k = \sqrt{\frac{p_k}{d_k}}\boldsymbol{I}_{n_{t_k} \times d_k}$ for all $k \in \{1, \ldots, M\}$;
2: **for** $iteration \leftarrow 1$ **to** N_{max} **do**
3: Compute $\boldsymbol{R}_{\bar{k}}$ using (2), $\boldsymbol{W}_k$ using (20), and $\boldsymbol{U}_k^{\text{MMSE}}$ using (17) for all $k \in \{1, \ldots, M\}$;
4: Compute $\boldsymbol{V}_k$ using (10) for all $k \in \{1, \ldots, M\}$;
5: Stop if the maximum absolute value of the difference of elements between the previous $\boldsymbol{Q}_k = \boldsymbol{V}_k\boldsymbol{V}_k^{\dagger}$ and the newly computed $\boldsymbol{Q}_k$ is less than ϵ for all $k \in \{1, \ldots, M\}$;
6: **end for**

Fig. 3. Pseudocode for computing the beamforming and combining weights of all links.

A. Algorithm Description

The proposed algorithm is summarized in Fig. 3. Our algorithm begins by activating all links and allocating the maximum number of streams at every transmitter, as shown in Line 1 of Fig. 3, where $\boldsymbol{I}_{n_{t_k} \times d_k}$ is an $n_{t_k} \times d_k$ matrix with ones on its diagonal and zeros elsewhere. Then, in lines 2–6, the algorithm computes the beamforming and combining weights iteratively. In Line 3, the receivers compute their interference-plus-noise covariance, error weights, and combining weights using the previously computed beamforming weights. In Line 4, the transmitters compute their beamforming weights using the previously computed interference-plus-noise covariance, error weights, and combining weights.

The proposed algorithm has three desirable properties. First, the proposed algorithm will deactivate, or re-activate, links if it determines that doing so improves the overall performance. Second, the computation of the beamforming or combining weights at a node requires only local information. That is, the computation of the beamforming weights at a node requires information only from the desired receiver and from those receivers it causes interference to. Similarly, the computation of the combining weights at a node requires information only from the desired transmitter and from those transmitters that it receives interference from. Requiring only local information is a desired property for any distributed implementation. Third, it is easy to prove using the technique from [17] that our proposed algorithm is guaranteed to converge, since at every iteration the algorithm moves monotonically towards a bounded objective.

For the MIMO interference channel, it is well known that the problem of maximizing the sum rate is non-convex [6]. Therefore, the proposed algorithm in Fig. 3 can only guarantee finding a local maximum on the sum rate. Additionally, different initial conditions on the beamforming weights in Line 1 of Fig. 3 will lead to different solutions. Similar to [4], we find that random initializations are just as good as any smart initialization we have tried, such as initializing the beamforming weights to the optimal SVD weights in the absence of interference. In the results presented in the following section, we choose the initialization stated in Line 1 of Fig. 3.

While we have no way of measuring how close the proposed algorithm is to optimal, our numerical results in Section VI

demonstrate that it outperforms previous algorithms, particularly in high-interference cases.

An algorithm based on similar concepts to those of the proposed algorithm is the WMMSE algorithm [22]. The WMMSE algorithm was independently developed for a more general problem and is based on the results of [17]. Although the algorithm of [22] can allocate low power to some links, the algorithm is unable to deactivate links. The reason is that the beamforming and combining weights of a link, as designed in [22], are functions of each other. Thus, the beamformer in [22] will be active if the combiner is active and vise versa. Our proposed algorithm is, on the other hand, based on the joint transceiver discussed in Section III-B. By performing the joint optimization, the decision on whether to activate or deactivate the link is based on the current state of the network and not on whether the link was active or inactive previously.

Note that because the WMMSE algorithm from [22] can gradually decrease the power allocated to some links as the algorithm progresses, the algorithm can asymptotically deactivate links. For a finite number of iterations, however, the WMMSE algorithm never completely deactivates links. A modified version that deactivates links below some threshold would also differ from our proposed algorithm because such an algorithm would never reactivate a deactivated link, which is possible with our algorithm. In the numerical results presented in Section VI-D, we show that when the number of iterations is small, the proposed algorithm achieves significantly better sum rate on average than the WMMSE algorithm.

B. Complexity Analysis

In terms of complexity per iteration of the algorithm, we perform the following analysis. For simplicity of notation, we drop the subscript k and assume that each transmitter is equipped with n_t transmit antennas and that each receiver is equipped with n_r receive antennas. For each iteration, the proposed algorithm requires that each receiver computes $\boldsymbol{R}$, $\boldsymbol{W}$, and $\boldsymbol{U}^{\mathrm{MMSE}}$. Computing $\boldsymbol{R}$ requires M matrix additions of size $n_r \times n_r$, which is $O(Mn_r^2)$. Computing $\boldsymbol{W}$ requires computing $\boldsymbol{R}^{-1/2}$ (an inverse and a square root) and the singular values of the $n_r \times d$ matrix in (19). Assuming that computing the SVD of an $m \times n$ matrix is $O(m^2n + n^3)$ and that computing the square root and inverse of an $n \times n$ matrix is $O(n^3)$, then computing $\boldsymbol{W}$ is $O(n_r^2 d + d^3 + n_r^3)$. Computing $\boldsymbol{U}^{\mathrm{MMSE}}$ requires computing the inverse of an $n_r \times n_r$ matrix, which is $O(n_r^3)$. At the transmitter side, computing $\boldsymbol{V}$ requires a search for μ as well as computing $\boldsymbol{P}^{-1/2}$ and computing the SVD of the $n_r \times n_t$ matrix in (8). Computing $\boldsymbol{P}^{-1/2}$ requires roughly M matrix additions of size $n_t \times n_t$ and a square root and inverse operation. The overhead of searching for μ can be reduced by searching around the value of μ obtained from the previous iteration. Ignoring the search for μ, computing $\boldsymbol{P}^{-1/2}$ is $O(Mn_t^2 + n_t^3)$ and the SVD operation in (8) is $O(n_t^2 n_r + n_r^3)$. Since $n_r \geq d$, the complexity of the proposed algorithm per iteration becomes $O(M^2n_r^2 + Mn_r^3 + M^2n_t^2 + Mn_t^3 + Mn_t^2 n_r)$. As the number of links M becomes large, the complexity of the algorithm becomes $O(M^2)$, which is equal to that of previously reported algorithms in its class [4], [5], [20], [22].

VI. Numerical Results

In this section, we present numerical results comparing the sum rate and complexity of the proposed algorithm to those of the previously reported algorithms in various levels of interference. For all simulations, we assume a "quasi-static" flat-fading Rayleigh channel model where the channel is assumed constant for the duration of a burst, but random between bursts [23]. Also, we assume that noise at each receiver is white, satisfying $\boldsymbol{R}_{\boldsymbol{n}_k} = \boldsymbol{I}$ for all $k \in \{1, \ldots, M\}$. We set the reference signal-to-noise ratio (SNR) and interference-to-noise ratio (INR) at one meter to 65.3 dB and assume a path-loss exponent of three. We set each node to have four antenna elements. Unless otherwise specified, we fix the distance between the transmitter and its corresponding receiver to 50 meters. We uniformly distribute the center of each link within a circle of a given radius. Also, we uniformly distribute the angles from the horizontal axis to the line that goes through the transmitter and receiver of every link from zero to 2π.

For all algorithms, we initialize the beamforming weights as shown in Line 1 of Fig. 3, and use the convergence criterion as shown in Line 5 of Fig. 3 with $\epsilon = 0.0001$. Additionally, we set the maximum number of iterations to $N_{max} = 10000$. If an algorithm reaches the maximum number of iterations, the algorithm stops and we record the sum rate.

In the following, we show results for the MWSR algorithm from [5] and the GP algorithm from [6]. We also show results for the MMSE algorithm from [4], [20], since, as we will see in Section VI-B, it has good performance when the number of links is high. We do not include a comparison against the SDP algorithm from [24], because on the few sample runs we attempted, we found the execution time of the SDP algorithm to be about two to three orders of magnitude higher than other methods. Additionally, we do not show a comparison with the linear-approximation-based algorithm from [25] since it was shown in [22] to have higher complexity than and comparable sum rate to approaching the problem through the weighted sum MSE criterion.

The rest of this section is organized as follows. In Section VI-A, we fix the number of links and vary the radius of the circle in which the links are placed. In Section VI-B, we fix the radius of the circle and vary the number of links within the circle. In Section VI-C, we study the effect on the sum rate and stream allocation when initializing the proposed algorithm using random initial conditions. Finally, in Section VI-D, we compare the complexity of our proposed algorithm with that of previously reported algorithms.

A. Sum Rate Versus Circle Radius

We consider ten MIMO links and vary the radius of the circle in which these links are placed. In Fig. 4, we show the sum rate, averaged over 100 trials, plotted as a function of the radius of the circle. In this experiment, a large radius corresponds to a sparse scenario where interference from other links is low (resulting in low INR between links). In contrast, a small radius corresponds to a dense scenario where interference from other links is high (resulting in high INR between links).

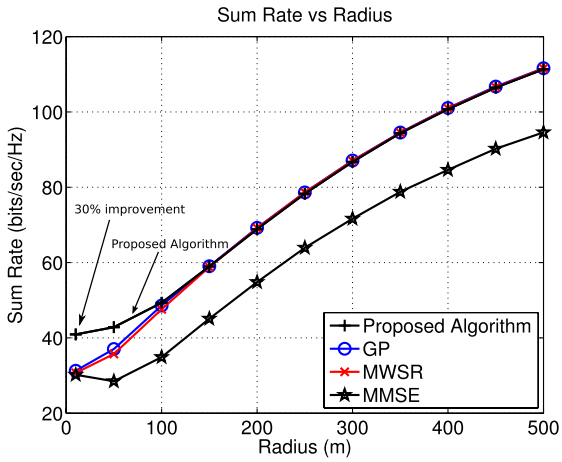

Fig. 4. Sum rate as a function of the radius of the circle where the center of the ten links are placed.

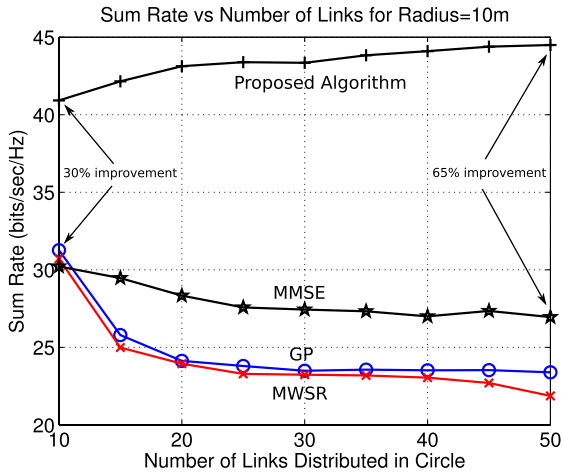

Fig. 5. Sum rate as a function of the number of links placed within the circle when the radius of the circle is fixed to ten meters. All links placed within the circle are considered by the algorithms, but an algorithm may choose to deactivate some links.

Note that because the distance between a transmitter and its corresponding receiver is fixed at all times, the SNR is equal and fixed for every link, independent of the radius of the circle.

The results of Fig. 4 show that our proposed algorithm achieves at least 30% higher sum rate at the lowest radius tested (ten meters), as compared to the other algorithms. In this range, interference is high and the degrees-of-freedom available from the multiple antennas on the nodes are not enough to support high performance on all links. The proposed algorithm achieves higher performance at high interference because, not only does it optimize the number of streams on each transmitter and the corresponding beamforming and combining weights for each stream, but it also optimizes which subset of transmitters should transmit. On average, the proposed algorithm activated slightly more than half the links within the circle at high interference and almost all links at low interference. The MMSE algorithm from [4], [20] minimizes the unweighted sum MSE and, therefore, does not necessarily achieve a high sum rate.

B. Sum Rate Versus Number of Links at High Interference

In this section, we fix the radius of the circle to ten meters and vary the number of links placed within this circle. Note again that the SNR is equal and fixed for every link, independent of the number of links placed within the circle. Fig. 5 shows the sum rate, averaged over 100 trials, plotted as a function of the maximum number of active links when the radius of the circle is fixed to ten meters. These results show that as the number of links increases, the sum rate of the GP, MWSR, and MMSE algorithms decrease while the sum rate of the proposed algorithm increases. Our algorithm achieves a sum rate that is at least 65% better than the sum rate of the other algorithms at the largest number of links tested. When the number of links is large, the proposed algorithm achieves a sum rate that is at least 90% better than that of the GP and MWSR. The proposed algorithm achieves high performance, and its performance increases as the number of links increases, since the algorithm has diversity on which links to activate and deactivate. On average, for the results shown on Fig. 5, our proposed algorithm activates between six to eight links.

When the number of antenna elements at every node is increased, the performance gap between the proposed algorithm and both the GP and MWSR is narrowed, especially when the number of links in the circle is low. This is expected since increasing the number of antenna elements at every node increases the degrees of freedom available to support more links. For example, when the number of antenna elements at every node is increased from four to eight and the number of links is high (50 links), the proposed algorithm achieves a sum rate that is 73% (as opposed to 90% when every node has four antennas) higher than that of the GP and the MWSR. However, when the number of antenna elements at every node is increased from four to eight and the number of links is low (ten links), the proposed algorithm no longer has a 30% performance improvement over the other algorithms and instead, all algorithms except for the MMSE algorithm achieve comparable performance. The MMSE algorithm achieved a lower sum rate than all other algorithms when all nodes are equipped with eight antenna elements.

C. Sum Rate Using Random Initial Conditions

In this section, we investigate the performance of the proposed algorithm using random initial conditions. Fig. 6 plots a histogram of the sum rate of the proposed algorithm for a random trial of the experiment of Section VI-B using 1000 random initial conditions for the cases where the number of links in the circle is $M = 10$ and $M = 50$. As expected, the results show that in both cases, the sum rate of the proposed algorithm varies depending on the initial conditions. Interestingly, the standard deviation of the sum rate does not increase as the number of links increases from $M = 10$ to $M = 50$. Instead, only the mean sum rate increases as the number of links increases from $M = 10$ to $M = 50$.

The results when $M = 10$ show that a high number of random initial conditions achieves sum rates near 36.7 bits/sec/Hz and near 41.2 bits/sec/Hz, corresponding to the two peaks in Fig. 6. This suggests that, depending on the initial conditions,

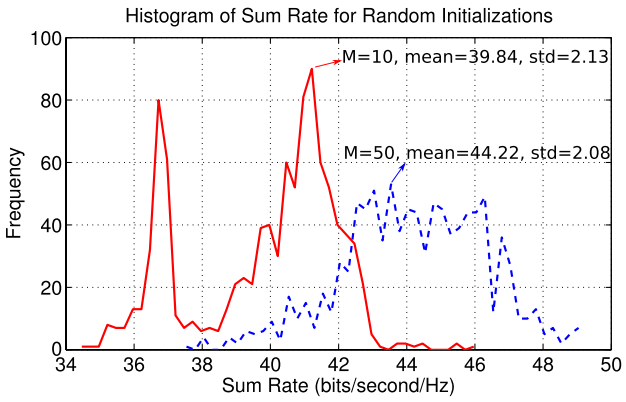

Fig. 6. Histogram of sum rate achieved by the proposed algorithm for 1000 random initial conditions for a random trial of the experiment of Section VI-B when the number of links within the circle is set to $M = 10$ and $M = 50$.

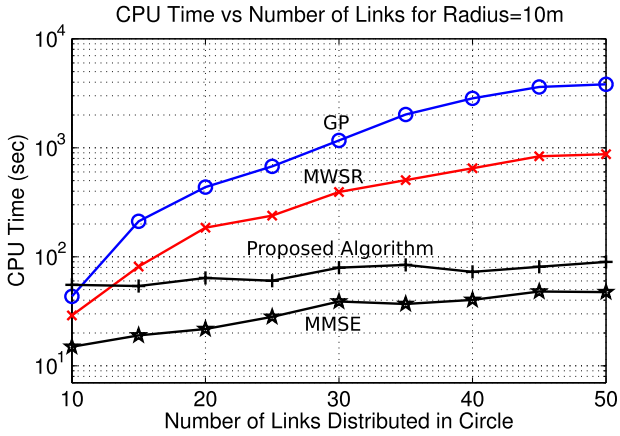

Fig. 7. Average CPU time as a function of the number of links placed within the circle when the radius of the circle is fixed to ten meters. All links placed within the circle are considered by the algorithms, but an algorithm may choose to deactivate some links.

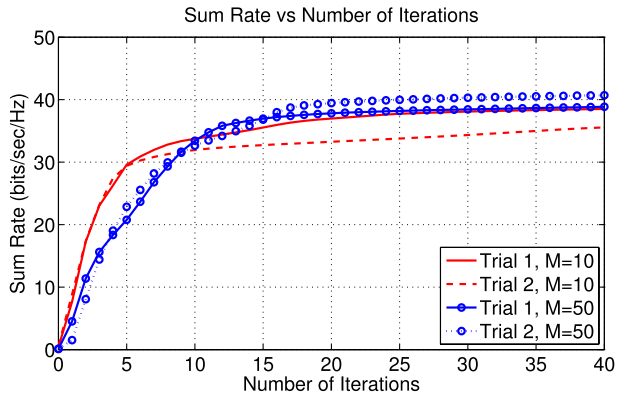

Fig. 8. Sum rate achieved by proposed algorithm as a function of the number of iterations for four different trials.

the proposed algorithm finds two large groups of solutions for this trial. Investigating the number of active links on these two groups, we find that those random initial conditions that achieve a sum rate near 36.7 bits/sec/Hz activate five of the ten available links on average. On the other hand, those random initial conditions that achieve a sum rate near 41.2 bits/sec/Hz activate six of the ten available links on average. When $M = 50$, on the other hand, the proposed algorithm produced solutions that are concentrated between 42.5 to 46 bits/sec/Hz with a shape that is almost uniform.

It is not always the case that the solutions by the proposed algorithm that activate the highest number of links also produce the highest sum rate. For $M = 50$, for example, the 100 random initial conditions that produce the highest sum rate enable 7.1 links on average, whereas the 100 random initial conditions that produce the lowest sum rate enable 8.6 links on average. Therefore, solutions that activate a large number of links do not necessarily achieve better sum rate than those that activate fewer number of links.

In terms of the number of streams allocated to the links when $M = 10$, the best three initial conditions that produce the highest sum rate allocate the same number of streams per link. The fourth best initial condition when $M = 10$ produce a different stream allocation, but its sum rate is only 3.7% lower than that of the best initial condition and 2.1% lower than that of the third best initial condition. Similarly, when $M = 50$, the best four initial conditions that produce the highest sum rate allocate the same number of streams per link. The fifth best initial condition when $M = 50$ produce a different stream allocation with comparable sum rate to that of the best four initial conditions.

D. Complexity

Now, we compare the complexity of our proposed algorithm in terms of the CPU running time for the experiment of Section VI-B in which we fix the radius of the circle to ten meters and vary the number of links placed within this circle. Fig. 7 shows the average CPU time by the algorithms as implemented in MATLAB and run on an i7-2700K Intel CPU rated at 3.5 GHz. These results show that the running time of the GP and the MWSR algorithms increases greatly as the number of active links increases. The running times of the MMSE algorithm and the proposed algorithm, however, increases only slightly as the number of active links increases. The MMSE algorithm has the shortest running time of all algorithms and the running time of our proposed algorithm was consistently about four times that of the MMSE algorithm.

We also analyze the complexity of the proposed algorithm in terms of sum rate versus number of iterations. Fig. 8 shows the sum rate achieved by the proposed algorithm as a function of the number of iterations for two random trials of the experiment of Section VI-B for the cases when $M = 10$ and $M = 50$ links are placed within a circle whose radius is ten meters. The results show that, for Trial 1, $M = 10$, the proposed algorithm is within 6% of its final sum rate after 20 iterations. For Trial 2, $M = 10$, the proposed algorithm requires slightly more iterations to reach a high-sum-rate solution; however, after 25 iterations, the proposed algorithm is within 10% of the final sum-rate. Similar results can be shown when the number of links is increased to $M = 50$. For Trial 1, $M = 50$ and Trial 2, $M = 50$, the sum rate after 25 iterations is within 5% and 6% of the final sum-rate, respectively.

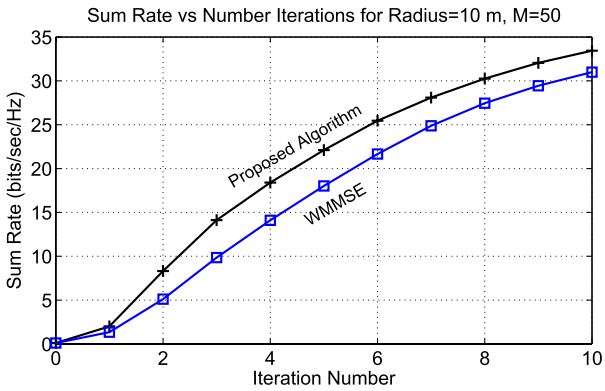

Fig. 9. Sum rate for the proposed algorithm and the WMMSE algorithm as a function of the number of iterations, when the radius is ten meters and $M = 50$.

The proposed algorithm can, at each iteration, deactivate links or reactivate previously deactivated links. For this reason, when the number of iterations is limited, the proposed algorithm can achieve better performance than the WMMSE algorithm from [22], which can only asymptotically deactivate links as the number of iterations goes to infinity. Fig. 9 shows the sum rate, averaged over 100 trials, plotted as a function of the number of iterations for the experiment of Section VI-B for the case when $M = 50$ links are placed within a circle whose radius is ten meters. The results show that the proposed algorithm can achieve up to 65% higher sum rate (when the number of iterations is three) than the WMMSE algorithm. Note, however, that as the number of iterations increases, the gap between the proposed algorithm and the WMMSE algorithm decreases.

VII. Conclusion

We have presented an algorithm whose goal is to maximize the sum rate of a set of interfering MIMO links. Our algorithm jointly optimizes which links should be active, the number of streams (if any) on each link, and their corresponding beamforming and combining weights. Our Simulation results showed that the proposed algorithm is able to achieve higher sum rate at high interference and comparable sum rate at medium and low interference than previously reported algorithms. Also, our simulation results showed that at high interference, the sum rate of our proposed algorithm increases as the number of links increases, because the proposed algorithm can deactivate links and has diversity on which links to deactivate. Finally, our results showed that the proposed algorithm also has lower time complexity than most algorithms tested at high interference with a large number of links.

Appendix A
Proof of Theorem 1

Proof: The structure of the proof of Theorem 1 follows closely the proof of Lemma 1, Lemma 2, and Theorem 1 in [16], with several differences due to the presence of interfering links in our setting. In the following, we provide a sketch of the proof.

Before we begin, let us expand (8) to include the zero singular values as follows. Let

$$\boldsymbol{R}_{\bar{k}}^{-1/2}\boldsymbol{H}_{kk}\boldsymbol{P}_{\bar{k}}^{-1/2} = \begin{bmatrix}\boldsymbol{F}_k & \widetilde{\boldsymbol{F}}_k\end{bmatrix}\begin{bmatrix}\boldsymbol{D}_k & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{bmatrix}\begin{bmatrix}\boldsymbol{G}_k & \widetilde{\boldsymbol{G}}_k\end{bmatrix}^{\dagger} \tag{21}$$

by SVD, where $\widetilde{\boldsymbol{F}}_k \in \mathbb{C}^{n_{r_k}\times(n_{r_k}-d_k)}$ and $\widetilde{\boldsymbol{G}}_k \in \mathbb{C}^{n_{t_k}\times(n_{t_k}-d_k)}$ have orthonormal column vectors that correspond to the left and right eigenvectors of $\boldsymbol{R}_{\bar{k}}^{-1/2}\boldsymbol{H}_{kk}\boldsymbol{P}_{\bar{k}}^{-1/2}$ with zero singular values, respectively.

The Lagrangian for (6) is given by

$$L(\boldsymbol{V}_k, \boldsymbol{U}_k, \mu_k) = \sum_{k=1}^{M}\operatorname{tr}(\boldsymbol{W}_k\boldsymbol{E}_k) + \mu_k\left(\operatorname{tr}\left(\boldsymbol{V}_k\boldsymbol{V}_k^{\dagger}\right) - p_k\right) \tag{22}$$

where μ_k is the Lagrange multiplier for link k, and

$$\boldsymbol{E}_k = \boldsymbol{U}_k^{\dagger}\boldsymbol{H}_{kk}\boldsymbol{V}_k\boldsymbol{V}_k^{\dagger}\boldsymbol{H}_{kk}^{\dagger}\boldsymbol{U}_k + \boldsymbol{U}_k^{\dagger}\boldsymbol{R}_{\bar{k}}\boldsymbol{U}_k - \boldsymbol{U}_k^{\dagger}\boldsymbol{H}_{kk}\boldsymbol{V}_k - \boldsymbol{V}_k^{\dagger}\boldsymbol{H}_{kk}^{\dagger}\boldsymbol{U}_k + \boldsymbol{I}, \tag{23}$$

by expanding (7) using (4). The Karush-Kuhn-Tucker (KKT) conditions to solve the optimization problem in (6) are

$$\nabla_{\boldsymbol{U}_k^{\dagger}} L = \boldsymbol{0}, \tag{24}$$

$$\operatorname{tr}\left(\boldsymbol{V}_k\boldsymbol{V}_k^{\dagger}\right) - p_k \le 0, \tag{25}$$

$$\nabla_{\boldsymbol{V}_k^{\dagger}} L = \boldsymbol{0}, \tag{26}$$

$$\mu_k\left(\operatorname{tr}\left(\boldsymbol{V}_k\boldsymbol{V}_k^{\dagger}\right) - p_k\right) = 0, \tag{27}$$

$$\mu_k \ge 0. \tag{28}$$

Setting the gradient of L with respect to $\boldsymbol{U}_k^{\dagger}$ to zero, we get

$$\boldsymbol{H}_{kk}\boldsymbol{V}_k = \boldsymbol{H}_{kk}\boldsymbol{V}_k\boldsymbol{V}_k^{\dagger}\boldsymbol{H}_{kk}\boldsymbol{U}_k + \boldsymbol{R}_{\bar{k}}\boldsymbol{U}_k, \tag{29}$$

and the gradient of L with respect to $\boldsymbol{V}_k^{\dagger}$ to zero, we get

$$\boldsymbol{H}_{kk}^{\dagger}\boldsymbol{U}_k\boldsymbol{W}_k = \boldsymbol{H}_{kk}^{\dagger}\boldsymbol{U}_k\boldsymbol{W}_k\boldsymbol{U}_k^{\dagger}\boldsymbol{H}_{kk}\boldsymbol{V}_k + \sum_{l=1,l\neq k}^{M}\boldsymbol{H}_{lk}^{\dagger}\boldsymbol{U}_l\boldsymbol{W}_l\boldsymbol{U}_l^{\dagger}\boldsymbol{H}_{lk}\boldsymbol{V}_k + \mu_k\boldsymbol{V}_k. \tag{30}$$

We can prove that the joint beamforming and combining weights have the structure given in (10) and (11), where $\boldsymbol{\Phi}_k$ and $\boldsymbol{\Theta}_k$ are arbitrary $d_k \times d_k$ matrices. To show this, first assume the most general expression for the beamforming and combining weights of link k, as given by

$$\boldsymbol{U}_k = \boldsymbol{R}_{\bar{k}}^{-1/2}\boldsymbol{F}_k\boldsymbol{\Phi}_k + \boldsymbol{R}_{\bar{k}}^{-1/2}\widetilde{\boldsymbol{F}}_k\widetilde{\boldsymbol{\Phi}}_k, \tag{31}$$

$$\boldsymbol{V}_k = \boldsymbol{P}_{\bar{k}}^{-1/2}\boldsymbol{G}_k\boldsymbol{\Theta}_k + \boldsymbol{P}_{\bar{k}}^{-1/2}\widetilde{\boldsymbol{G}}_k\widetilde{\boldsymbol{\Theta}}_k, \tag{32}$$

where $\widetilde{\boldsymbol{\Phi}}_k$ is any $(n_{r_k} - d_k) \times d_k$ matrix and $\widetilde{\boldsymbol{\Theta}}_k$ is any $(n_{t_k} - d_k) \times d_k$ matrix. Then, applying the structure of the proof of Lemma 1 in [16], we find that $\widetilde{\boldsymbol{\Phi}}_k = \boldsymbol{0}$ and $\widetilde{\boldsymbol{\Theta}}_k = \boldsymbol{0}$, and,

therefore, the joint beamforming and combining weights have the structure given in (10) and (11).

Next, premultiplying (30) with $\boldsymbol{V}_k^\dagger$, and simplifying using (21) we get

$$\boldsymbol{\Phi}_k^\dagger \boldsymbol{D}_k \boldsymbol{\Theta}_k = \boldsymbol{\Phi}_k^\dagger \boldsymbol{D}_k \boldsymbol{\Theta}_k \boldsymbol{\Theta}_k^\dagger \boldsymbol{D}_k \boldsymbol{\Phi}_k + \boldsymbol{\Phi}_k^\dagger \boldsymbol{\Phi}_k, \tag{33}$$

$$\boldsymbol{\Phi}_k^\dagger \boldsymbol{D}_k \boldsymbol{\Theta}_k \boldsymbol{W}_k = \boldsymbol{\Theta}_k^\dagger \boldsymbol{D}_k \boldsymbol{\Phi}_k \boldsymbol{W}_k \boldsymbol{\Phi}_k^\dagger \boldsymbol{D}_k \boldsymbol{\Theta}_k + \boldsymbol{\Theta}_k^\dagger \boldsymbol{\Theta}_k. \tag{34}$$

Then, using the technique in the proof of Lemma 2 in [16], we find that

$$\boldsymbol{\mathcal{D}}_1 = \boldsymbol{\Phi}_k^\dagger \boldsymbol{D}_k \boldsymbol{\Theta}_k, \tag{35}$$

$$\boldsymbol{\Phi}_k = \boldsymbol{\mathcal{D}}_2^{1/2} \succeq 0, \tag{36}$$

$$\boldsymbol{\Theta}_k = \boldsymbol{\mathcal{D}}_3^{1/2} \succeq 0, \tag{37}$$

where $(\cdot) \succeq 0$ denotes that $(\cdot)$ is a positive semidefinite matrix, and where $\boldsymbol{\mathcal{D}}_1, \boldsymbol{\mathcal{D}}_2,$ and $\boldsymbol{\mathcal{D}}_3$ are diagonal matrices. This proves that $\boldsymbol{\Phi}_k$ and $\boldsymbol{\Theta}_k$ are diagonal matrices with real nonnegative entries.

Finally, we can derive (12) and (13). Simplifying (33) and (34) using (35), and then plugging in (36) and (37) into the resulting expressions and into (35), we get

$$\boldsymbol{\mathcal{D}}_1 = \boldsymbol{\mathcal{D}}_1^2 + \boldsymbol{\mathcal{D}}_2, \tag{38}$$

$$\boldsymbol{\mathcal{D}}_1 \boldsymbol{W}_k = \boldsymbol{\mathcal{D}}_1 \boldsymbol{W}_k \boldsymbol{\mathcal{D}}_1 + \boldsymbol{\mathcal{D}}_3, \tag{39}$$

$$\boldsymbol{\mathcal{D}}_1 = \boldsymbol{\mathcal{D}}_2^{1/2} \boldsymbol{D}_k \boldsymbol{\mathcal{D}}_3^{1/2}. \tag{40}$$

Using a similar analysis as the one presented in the proof of Theorem 1 in [16], we solve for $\boldsymbol{\mathcal{D}}_2$ and $\boldsymbol{\mathcal{D}}_3$, and plug the resulting expressions into (36) and (37) to get (13) and (12), respectively. □

APPENDIX B
PROOF OF LEMMA 1

Proof: Let $\boldsymbol{\Pi}_k = \boldsymbol{J}_k \boldsymbol{\Sigma}_k \boldsymbol{J}_k^\dagger$ by eigenvalue decomposition, where $\boldsymbol{\Sigma}_k \in \mathbb{R}^{n_{t_k} \times n_{t_k}}$ is diagonal and $\boldsymbol{J}_k \in \mathbb{C}^{n_{t_k} \times n_{t_k}}$ is unitary. Then, we can rewrite (9) as

$$\begin{aligned} \boldsymbol{P}_{\bar{k}} &= \boldsymbol{J}_k \boldsymbol{\Sigma}_k \boldsymbol{J}_k^\dagger + \mu_k \boldsymbol{I} \\ &= \boldsymbol{J}_k (\boldsymbol{\Sigma}_k + \mu_k \boldsymbol{I}) \boldsymbol{J}_k^\dagger. \end{aligned} \tag{41}$$

Using (41), we rewrite (10) as

$$\begin{aligned} \boldsymbol{V}_k &= \left(\boldsymbol{J}_k (\boldsymbol{\Sigma}_k + \mu_k \boldsymbol{I}) \boldsymbol{J}_k^\dagger \right)^{-1/2} \boldsymbol{G}_k \boldsymbol{\Theta}_k \\ &= \left(\boldsymbol{J}_k (\boldsymbol{\Sigma}_k + \mu_k \boldsymbol{I})^{\frac{1}{2}} \boldsymbol{J}_k^\dagger \right)^{-1} \boldsymbol{G}_k \boldsymbol{\Theta}_k \\ &= \boldsymbol{J}_k (\boldsymbol{\Sigma}_k + \mu_k \boldsymbol{I})^{-1/2} \boldsymbol{J}_k^\dagger \boldsymbol{G}_k \boldsymbol{\Theta}_k. \end{aligned} \tag{42}$$

Similarly, we rewrite the SVD on (8) as

$$\boldsymbol{R}_{\bar{k}}^{-1/2} \boldsymbol{H}_{kk} \boldsymbol{J}_k (\boldsymbol{\Sigma}_k + \mu_k \boldsymbol{I})^{-1/2} \boldsymbol{J}_k^\dagger = \boldsymbol{F}_k \boldsymbol{D}_k \boldsymbol{G}_k^\dagger, \tag{43}$$

and the power used at the transmitter of link k as

$$\begin{aligned} \operatorname{tr}\left(\boldsymbol{V}_k \boldsymbol{V}_k^\dagger \right) &= \operatorname{tr}\Big(\boldsymbol{J}_k (\boldsymbol{\Sigma}_k + \mu_k \boldsymbol{I})^{-1/2} \boldsymbol{J}_k^\dagger \boldsymbol{G}_k \\ &\qquad \boldsymbol{\Theta}_k^2 \boldsymbol{G}_k^\dagger \boldsymbol{J}_k (\boldsymbol{\Sigma}_k + \mu_k \boldsymbol{I})^{-1/2} \boldsymbol{J}_k^\dagger \Big) \\ &= \operatorname{tr}\left(\boldsymbol{J}_k (\boldsymbol{\Sigma}_k + \mu_k \boldsymbol{I})^{-1} \boldsymbol{J}_k^\dagger \boldsymbol{G}_k \boldsymbol{\Theta}_k^2 \boldsymbol{G}_k^\dagger \right), \end{aligned} \tag{44}$$

where we have used the identity $\operatorname{tr}(\boldsymbol{XY}) = \operatorname{tr}(\boldsymbol{YX})$ to obtain (44). If $\boldsymbol{\Pi}_k$ is singular then at least one diagonal element in $\boldsymbol{\Sigma}_k$ is zero. Then, in the limit as $\mu_k \to 0^+$, at least the first element of $\boldsymbol{D}_k$ in (43) will approach infinity, and so the corresponding element in $\boldsymbol{\Theta}_k^2 = (\boldsymbol{W}_k^{1/2} \boldsymbol{D}_k^{-1} - \boldsymbol{D}_k^{-2})_+$ of (44) will approach zero. Notice that, as $\mu_k \to 0^+$, the $(\cdot)_+$ operator has no effect on the first element of $\boldsymbol{\Theta}_k^2$ since the first element of $\boldsymbol{D}_k^{-2}$ approaches zero faster than that of $\boldsymbol{D}_k^{-1}$. Also, in the limit as $\mu_k \to 0^+$, at least one element of $(\boldsymbol{\Sigma}_k + \mu_k \boldsymbol{I})^{-1}$ in (44) approaches infinity. Because this element of $(\boldsymbol{\Sigma}_k + \mu_k \boldsymbol{I})^{-1}$ increases with μ_k^{-1} and the smallest nonzero value of $\boldsymbol{\Theta}_k^2$ decreases with $\mu_k^{-1/2}$ as $\mu_k \to 0^+$ in (44), then the power used by the transmitter of link k approaches infinity as $\mu_k \to 0^+$. Therefore,

$$\lim_{\mu_k \to 0^+} \operatorname{tr}\left(\boldsymbol{V}_k \boldsymbol{V}_k^\dagger \right) = \infty. \tag{45}$$

□

APPENDIX C
PROOF OF LEMMA 3

Proof: It is easy to see that the MMSE combining weights for the virtual receiver of link k are given by

$$\overleftarrow{\boldsymbol{U}}_k^{\mathrm{MMSE}} = \left(\overleftarrow{\boldsymbol{H}}_{kk} \overleftarrow{\boldsymbol{V}}_k \overleftarrow{\boldsymbol{V}}_k^\dagger \overleftarrow{\boldsymbol{H}}_{kk}^\dagger + \overleftarrow{\boldsymbol{R}}_{\bar{k}} \right)^{-1} \overleftarrow{\boldsymbol{H}}_{kk} \overleftarrow{\boldsymbol{V}}_k. \tag{46}$$

Equation (46) can be rewritten as

$$\begin{aligned} \boldsymbol{U}_k^{\mathrm{MMSE}} &= \left(\overleftarrow{\boldsymbol{R}}_{\bar{k}}^{\frac{1}{2}} \left(\overleftarrow{\boldsymbol{R}}_{\bar{k}}^{-1/2} \overleftarrow{\boldsymbol{H}}_{kk} \overleftarrow{\boldsymbol{V}}_k \overleftarrow{\boldsymbol{V}}_k^\dagger \overleftarrow{\boldsymbol{H}}_{kk}^\dagger \overleftarrow{\boldsymbol{R}}_{\bar{k}}^{-1/2} \boldsymbol{I} \right) \right. \\ &\qquad \left. + \overleftarrow{\boldsymbol{R}}_{\bar{k}}^{\frac{1}{2}} \right)^{-1} \overleftarrow{\boldsymbol{H}}_{kk} \overleftarrow{\boldsymbol{V}}_k \\ &= \overleftarrow{\boldsymbol{R}}_{\bar{k}}^{-1/2} \left(\overleftarrow{\boldsymbol{R}}_{\bar{k}}^{-1/2} \overleftarrow{\boldsymbol{H}}_{kk} \overleftarrow{\boldsymbol{V}}_k \overleftarrow{\boldsymbol{V}}_k^\dagger \overleftarrow{\boldsymbol{H}}_{kk}^\dagger \overleftarrow{\boldsymbol{R}}_{\bar{k}}^{-1/2} \right. \\ &\qquad \left. + \boldsymbol{I} \right)^{-1} \overleftarrow{\boldsymbol{R}}_{\bar{k}}^{-1/2} \overleftarrow{\boldsymbol{H}}_{kk} \overleftarrow{\boldsymbol{V}}_k. \end{aligned} \tag{47}$$

Substituting $\overleftarrow{\boldsymbol{R}}_{\bar{k}} = \boldsymbol{P}_{\bar{k}}$, $\overleftarrow{\boldsymbol{H}}_{lk} = \boldsymbol{H}_{kl}^\dagger$, and $\overleftarrow{\boldsymbol{V}}_k = \boldsymbol{U}_k \boldsymbol{W}_k^{1/2}$ into (47) we get

$$\begin{aligned} \overleftarrow{\boldsymbol{U}}_k^{\mathrm{MMSE}} &= \boldsymbol{P}_{\bar{k}}^{-1/2} \left(\boldsymbol{P}_{\bar{k}}^{-1/2} \boldsymbol{H}_{kk}^\dagger \boldsymbol{U}_k \boldsymbol{W}_k \boldsymbol{U}_k^\dagger \boldsymbol{H}_{kk} \boldsymbol{P}_{\bar{k}}^{-1/2} + \boldsymbol{I} \right)^{-1} \\ &\qquad \times \boldsymbol{P}_{\bar{k}}^{-1/2} \boldsymbol{H}_{kk}^\dagger \boldsymbol{U}_k \boldsymbol{W}_k^{\frac{1}{2}}. \end{aligned} \tag{48}$$

Using (11) and (8) we see that

$$\begin{aligned}\boldsymbol{U}_k^{\dagger}\boldsymbol{H}_{kk}\boldsymbol{P}_{\bar{k}}^{-1/2} &= \boldsymbol{\Phi}_k\boldsymbol{F}_k^{\dagger}\boldsymbol{R}_{\bar{k}}^{-1/2}\boldsymbol{H}_{kk}\boldsymbol{P}_{\bar{k}}^{-1/2}\\ &= \boldsymbol{\Phi}_k\boldsymbol{D}_k\boldsymbol{G}_k^{\dagger} = \boldsymbol{D}_k\boldsymbol{\Phi}_k\boldsymbol{G}_k^{\dagger} \end{aligned} \tag{49}$$

and so (48) becomes

$$\overleftarrow{\boldsymbol{U}}_k^{\mathrm{MMSE}} = \boldsymbol{P}_{\bar{k}}^{-1/2}\left(\boldsymbol{G}_k\boldsymbol{\Phi}_k\boldsymbol{D}_k^2\boldsymbol{W}_k\boldsymbol{\Phi}_k\boldsymbol{G}_k^{\dagger}+\boldsymbol{I}\right)^{-1}\boldsymbol{G}_k\boldsymbol{\Phi}_k\boldsymbol{D}_k\boldsymbol{W}_k^{\frac{1}{2}}. \tag{50}$$

Using the matrix inversion lemma we can rewrite (50) as

$$\begin{aligned}\overleftarrow{\boldsymbol{U}}_k^{\mathrm{MMSE}} &= \boldsymbol{P}_{\bar{k}}^{-1/2}\left(\boldsymbol{I}-\boldsymbol{G}_k\boldsymbol{\Phi}_k\right.\\ &\quad\left.\times\left(\boldsymbol{D}_k^{-2}\boldsymbol{W}_k^{-1}+\boldsymbol{\Phi}_k^2\right)^{-1}\boldsymbol{\Phi}_k\boldsymbol{G}_k^{\dagger}\right)\boldsymbol{G}_k\boldsymbol{\Phi}_k\boldsymbol{D}_k\boldsymbol{W}_k^{1/2}\\ &= \boldsymbol{P}_{\bar{k}}^{-1/2}\left(\boldsymbol{G}_k\boldsymbol{\Phi}_k\boldsymbol{D}_k\boldsymbol{W}_k^{1/2}-\boldsymbol{G}_k\boldsymbol{\Phi}_k\right.\\ &\quad\left.\times\left(\boldsymbol{D}_k^{-2}\boldsymbol{W}_k^{-1}+\boldsymbol{\Phi}_k^2\right)^{-1}\boldsymbol{\Phi}_k^2\boldsymbol{D}_k\boldsymbol{W}_k^{1/2}\right)\\ &= \boldsymbol{P}_{\bar{k}}^{-1/2}\boldsymbol{G}_k\boldsymbol{\Phi}_k\left(\boldsymbol{D}_k\boldsymbol{W}_k^{1/2}-\boldsymbol{D}_k\boldsymbol{W}_k^{1/2}\right.\\ &\quad\left.\times\left(\boldsymbol{D}_k^{-1}\boldsymbol{W}_k^{-1/2}+\boldsymbol{\Phi}_k^2\boldsymbol{D}_k\boldsymbol{W}_k^{1/2}\right)^{-1}\boldsymbol{\Phi}_k^2\boldsymbol{D}_k\boldsymbol{W}_k^{1/2}\right). \end{aligned} \tag{51}$$

Applying the matrix inversion lemma to (51) and substituting (13) into the resulting expression we get

$$\begin{aligned}\overleftarrow{\boldsymbol{U}}_k^{\mathrm{MMSE}} &= \boldsymbol{P}_{\bar{k}}^{-1/2}\boldsymbol{G}_k\boldsymbol{\Phi}_k\left(\boldsymbol{D}_k^{-1}\boldsymbol{W}_k^{-1/2}+\boldsymbol{D}_k\boldsymbol{W}_k^{1/2}\boldsymbol{\Phi}_k^2\right)^{-1}\\ &= \boldsymbol{P}_{\bar{k}}^{-1/2}\boldsymbol{G}_k\boldsymbol{\Phi}_k\left(\boldsymbol{D}_k^{-1}\boldsymbol{W}_k^{-1/2}+\boldsymbol{D}_k\boldsymbol{W}_k^{-1/2}\boldsymbol{\Theta}_k^2\right)^{-1}\\ &= \boldsymbol{P}_{\bar{k}}^{-1/2}\boldsymbol{G}_k\boldsymbol{\Phi}_k\left(\boldsymbol{D}_k^{-1}\boldsymbol{W}_k^{-1/2}\right.\\ &\quad\left.+\boldsymbol{D}_k\boldsymbol{W}_k^{-1/2}\left(\boldsymbol{W}_k^{1/2}\boldsymbol{D}_k^{-1}-\boldsymbol{D}_k^{-2}\right)_+\right)^{-1}\\ &= \boldsymbol{P}_{\bar{k}}^{-1/2}\boldsymbol{G}_k\boldsymbol{\Phi}_k. \end{aligned} \tag{52}$$

Since $\overleftarrow{\boldsymbol{U}}_k = \boldsymbol{V}_k\boldsymbol{W}_k^{-1/2}$, then we have $\boldsymbol{V}_k = \overleftarrow{\boldsymbol{U}}_k^{\mathrm{MMSE}}\boldsymbol{W}_k^{1/2}$ and so the beamforming weights are given by (10). □

APPENDIX D
PROOF OF PROPOSITION 1

Proof: For this case, $\boldsymbol{P}_{\bar{k}} = \mu_k\boldsymbol{I}$ and (8) simplifies to

$$\begin{aligned}\mu_k^{-1/2}\boldsymbol{R}_{\bar{k}}^{-1/2}\boldsymbol{H}_{kk} &= \boldsymbol{F}_k\boldsymbol{D}_k\boldsymbol{G}_k^{\dagger},\\ \boldsymbol{R}_{\bar{k}}^{-1/2}\boldsymbol{H}_{kk} &= \boldsymbol{F}_k\left(\mu_k^{1/2}\boldsymbol{D}_k\right)\boldsymbol{G}_k^{\dagger}. \end{aligned} \tag{53}$$

From (53), it is clear that $\boldsymbol{\Lambda}_k^{1/2} = \mu_k^{1/2}\boldsymbol{D}_k$ contains the singular values of the whitened channel $\boldsymbol{R}_{\bar{k}}^{-1/2}\boldsymbol{H}_{kk}$ in decreasing order from top left to bottom right, and that $\boldsymbol{F}_k$ and $\boldsymbol{G}_k$ correspond to the left and right eigenvectors of $\boldsymbol{R}_{\bar{k}}^{-1/2}\boldsymbol{H}_{kk}$ with positive singular values, respectively. Plugging in $\boldsymbol{P}_{\bar{k}} = \mu_k\boldsymbol{I}$ into (10), the beamforming weights simplify to

$$\boldsymbol{V}_k = \boldsymbol{G}_k\left(\mu_k^{-1/2}\boldsymbol{\Theta}_k\right), \tag{54}$$

and so the transmitter of link k is transmitting through the eigen-modes of the whitened channel.

Now we show that power is allocated through waterfilling. Using (54), the inequality constraint $\mathrm{tr}(\boldsymbol{V}_k\boldsymbol{V}_k^{\dagger}) \le p_k$ becomes

$$\begin{aligned}\mathrm{tr}\left(\boldsymbol{G}_k\left(\mu_k^{-1}\boldsymbol{\Theta}_k^2\right)\boldsymbol{G}_k^{\dagger}\right) &\le p_k\\ \mathrm{tr}\left(\mu_k^{-1}\boldsymbol{\Theta}_k^2\right) &\le p_k \qquad (55)\\ \mathrm{tr}\left(\mu_k^{-1}\left(\boldsymbol{W}_k^{1/2}\boldsymbol{D}_k^{-1}-\boldsymbol{D}_k^{-2}\right)_+\right) &\le p_k\\ \mathrm{tr}\left(\mu_k^{-1/2}\boldsymbol{W}_k^{1/2}\left(\mu_k^{1/2}\boldsymbol{D}_k\right)^{-1}-\left(\mu_k\boldsymbol{D}_k^2\right)^{-1}\right)_+ &\le p_k\\ \mathrm{tr}\left(\mu_k^{-1/2}\boldsymbol{W}_k^{1/2}\boldsymbol{\Lambda}_k^{-1/2}-\boldsymbol{\Lambda}_k^{-1}\right)_+ &\le p_k, \qquad (56)\end{aligned}$$

where we have used the fact that $\mathrm{tr}(\boldsymbol{X}\boldsymbol{Y}) = \mathrm{tr}(\boldsymbol{Y}\boldsymbol{X})$ to obtain (55). To maximize the rate on its link, the transmitter of link k transmits with maximum power and the inequality in (56) is treated with equality. Finally, with the choice of $\boldsymbol{W}_k = \alpha_k\boldsymbol{\Lambda}_k$ for any scalar $\alpha_k > 0$, (56) becomes the waterfilling solution [6], [7], [26]

$$\mathrm{tr}\left(\gamma_k\boldsymbol{I}-\boldsymbol{\Lambda}_k^{-1}\right)_+ = p_k, \tag{57}$$

where $\gamma_k = \sqrt{\alpha_k/\mu_k}$ is the waterfilling level. □

REFERENCES

[1] A. Molisch, *Wireless Communications*. Hoboken, NJ, USA: Wiley, 2005.

[2] K. Gomadam, V. Cadambe, and S. Jafar, "A distributed numerical approach to interference alignment and applications to wireless interference networks," *IEEE Trans. Inf. Theory*, vol. 57, no. 6, pp. 3309–3322, Jun. 2011.

[3] L. Cortés-Peña, J. Barry, and D. Blough, "The performance loss of unilateral interference cancellation," in *Proc. IEEE ICC*, Jun. 2012, pp. 4181–4186.

[4] S. Peters and R. Heath, "Cooperative algorithms for MIMO interference channels," *IEEE Trans. Veh. Technol.*, vol. 60, no. 1, pp. 206–218, Jan. 2011.

[5] F. Negro, S. Shenoy, I. Ghauri, and D. Slock, "On the MIMO interference channel," in *Proc. Inf. Theory Appl. Workshop*, Feb. 2010, pp. 1–9.

[6] S. Ye and R. Blum, "Optimized signaling for MIMO interference systems with feedback," *IEEE Trans. Signal Process.*, vol. 51, no. 11, pp. 2839–2848, Nov. 2003.

[7] M. Demirkol and M. Ingram, "Power-controlled capacity for interfering MIMO links," in *Proc. IEEE Veh. Technol. Conf.*, May 2001, vol. 1, pp. 187–191.

[8] G. Arslan, M. Demirkol, and Y. Song, "Equilibrium efficiency improvement in MIMO interference systems: A decentralized stream control approach," *IEEE Trans. Wireless Commun.*, vol. 6, no. 8, pp. 2984–2993, Aug. 2007.

[9] B. Hamdaoui and K. Shin, "Characterization and analysis of multi-hop wireless MIMO network throughput," in *Proc. ACM Int. Symp. MobiHoc Netw. Comput.*, Sep. 2007, pp. 120–129.

[10] J. Liu, Y. Shi, and Y. Hou, "A tractable and accurate cross-layer model for multi-hop MIMO networks," in *Proc. IEEE INFOCOM*, Mar. 2010, pp. 1–9.

[11] Y. Shi, J. Liu, C. Jiang, C. Gao, and Y. Hou, "An optimal link layer model for multi-hop MIMO networks," in *Proc. IEEE INFOCOM*, Apr. 2011, pp. 1916–1924.

[12] R. Srinivasan, D. Blough, L. Cortés-Peña, and P. Santi, "Maximizing throughput in MIMO networks with variable rate streams," in *Proc. EW Conf.*, Apr. 2010, pp. 551–559.
[13] D. Blough, G. Resta, P. Santi, R. Shrinivasan, and L. Cortés-Peña, "Optimal one-shot scheduling for MIMO networks," in *Proc. IEEE Int. Conf. SECON*, Jun. 2011, pp. 404–412.
[14] R. Bhatia and L. Li, "Throughput optimization of wireless mesh networks with MIMO links," in *Proc. IEEE INFOCOM*, May 2007, pp. 2326–2330.
[15] S. Chu and X. Wang, "Opportunistic and cooperative spatial multiplexing in MIMO ad hoc networks," *IEEE/ACM Trans. Netw.*, vol. 18, no. 5, pp. 1610–1623, Oct. 2010.
[16] H. Sampath, P. Stoica, and A. Paulraj, "Generalized linear precoder and decoder design for MIMO channels using the weighted MMSE criterion," *IEEE Trans. Commun.*, vol. 49, no. 12, pp. 2198–2206, Dec. 2001.
[17] S. Christensen, R. Agarwal, E. Carvalho, and J. Cioffi, "Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 4792–4799, Dec. 2008.
[18] L. Cortés-Peña, J. Barry, and D. Blough, "Joint optimization of stream allocation and beamforming and combining weights for the MIMO interference channel," in *Proc. IEEE GLOBECOM*, Dec. 2013, pp. 4012–4018.
[19] R. Blum, "MIMO capacity with interference," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 5, pp. 793–801, Jun. 2003.
[20] H. Shen, B. Li, M. Tao, and X. Wang, "MSE-based transceiver designs for the MIMO interference channel," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3480–3489, Nov. 2010.
[21] F. Rashid-Farrokhi, K. Liu, and L. Tassiulas, "Transmit beamforming and power control for cellular wireless systems," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1437–1450, Oct. 1998.
[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.
[23] G. Foschini and M. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Pers. Commun.*, vol. 6, no. 3, pp. 311–335, Mar. 1998.
[24] M. Razaviyayn, M. Sanjabi, and Z. Luo, "Linear transceiver design for interference alignment: Complexity and computation," *IEEE Trans. Inf. Theory*, vol. 58, no. 5, pp. 2896–2910, May 2012.
[25] S.-J. Kim and G. B. Giannakis, "Optimal resource allocation for MIMO ad hoc cognitive radio networks," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 3117–3131, Apr. 2011.
[26] E. Telatar, "Capacity of multi-antenna gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, Dec. 1999.

Luis Miguel Cortés-Peña received the B.S. degree in computer engineering (with highest honors), and the M.S. and Ph.D. degrees in electrical and computer engineering from the Georgia Institute of Technology, Atlanta, GA, USA, in 2006, 2008, and 2014, respectively. His doctoral research involved the network-wide optimization of MIMO links. He was a Goizueta Foundation Fellow from 2008 to 2012. He worked as an intern at BlackBerry, Sunrise, FL, USA, and at Digital Receiver Technology, Germantown, MD, USA, during the summers of 2009 and 2010, respectively. Currently, he is with the Government Communications Systems Division at Harris Corporation, Melbourne, FL, USA.

John R. Barry (S'85–M'86–SM'04) received the B.S. degree (summa cum laude) from the State University of New York at Buffalo in 1986, and the M.S. and Ph.D. degrees from the University of California at Berkeley in 1987 and 1992, respectively, all in electrical engineering. His doctoral research explored the feasibility of broadband wireless communications using diffuse infrared radiation. Since 1985, he has held engineering positions in the fields of communications and radar systems at Bell Communications Research, IBM T.J. Watson Research Center, Hughes Aircraft Company, and General Dynamics. Currently, he is serving as a Guest Editor for a special issue of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS. He is a coauthor of *Digital Communication* (3rd ed., Kluwer, 2004). He is a co-editor of *Advanced Optical Wireless Communication Systems* (Cambridge Univ. Press, 2010). He is the author of *Wireless Infrared Communications* (Kluwer, 1994). He is a Senior Member of the IEEE. He served as a Technical Program Chair for IEEE Globecom 2013. He received the 1992 David J. Griep Memorial Prize and the 1993 Eliahu Jury Award from U.C. Berkeley, a 1993 Research Initiation Award from NSF, and a 1993 IBM Faculty Development Award.

Douglas M. Blough (S'87–M'88–SM'00) received the B.S. degree in electrical engineering and the M.S. and Ph.D. degrees in computer science from The Johns Hopkins University, Baltimore, MD, USA, in 1984, 1986, and 1988, respectively. Since Fall 1999, he has been Professor of Electrical and Computer Engineering at the Georgia Institute of Technology, where he also holds a joint appointment in the School of Computer Science. From 1988 to 1999, he was on the faculty of Electrical and Computer Engineering at the University of California, Irvine, CA, USA. His research interests include distributed systems and wireless multihop networks. He was General Chair for the 2014 International Conference on Dependable Systems and Networks (DSN), and was Program Chair for DSN 2000 and the 1995 Pacific Rim International Symposium on Fault-Tolerant Systems. He has served as Associate Editor for IEEE TRANSACTIONS ON COMPUTERS (1995 through 2000), IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS (2001 through 2005), and IEEE TRANSACTIONS ON MOBILE COMPUTING (2008–2013).