Title
Optimal Design of Spatial Source-and-Relay Matrices for a Non-Regenerative Two-Way MIMO Relay System

Journal
The IEEE Transactions on Wireless Communications, 10(5)

ISSN
1536-1276

Authors
Xu, Shengyang
Hua, Yingbo

Publication Date
2011-03-24

Peer reviewed

Optimal Design of Spatial Source-and-Relay Matrices for a Non-Regenerative Two-Way MIMO Relay System

Shengyang Xu and Yingbo Hua, *Fellow, IEEE*

***Abstract*—This paper considers a system where two users exchange information via a non-regenerative half-duplex two-way MIMO relay and each of the two users and the relay is equipped with multiple antennas. We study the design of the spatial source covariance matrices (or source matrices) at the two users and the spatial transformation matrix (or relay matrix) at the relay to maximize the achievable weighted sum rate of the system. The source matrices and the relay matrix are optimized alternately until convergence. If the relay matrix is given, we show that the optimal design of the source matrices (for uniformly weighted sum rate) follows a generalized water filling (GWF) algorithm. If the source matrices are given, we show two search algorithms to optimize the relay matrix. The first algorithm is a hybrid gradient method which adaptively switches between the (steepest) gradient descent and the Newton's search. The second is an iterative weighted minimum mean square error (WMMSE) method which alternately refines the MMSE equalizers at the users and the relay matrix. We compare the convergence behaviors of the two algorithms and demonstrate their advantage over prior algorithms. We also show an optimal structure of the relay matrix, which is useful to reduce the search complexity.**

***Index Terms*—Smart relays, multi-antenna relays, two-way relays, power scheduling, eigen-beamforming, gradient methods, weighted minimum mean square errors, generalized water-filling algorithm.**

I. Introduction

THE radio frequency spectrum for mobile wireless communication is becoming increasingly crowded especially in large cities. Developing technologies for efficient spectral usage is becoming more important. Wireless relays are such methods that can be rapidly deployed to enhance the coverage, reliability and throughput of a wireless network subject to power and spectral constraints. Wireless relays equipped with multiple antennas, also called MIMO (multiple input multiple output) relays, are particularly useful for scattering rich and non-line-of-sight environment. This paper considers a non-regenerative two-way MIMO relay system where two users concurrently exchange information via a single half-duplex MIMO relay.

The idea of two-way relay has been studied in [2], [3], [4], [5], [6], [7], [8]. Central to this idea is that two users served by a relay can transmit their information to the relay simultaneously in one channel and the relay forwards a combined information to both users in another channel. And since each user knows its own information, it can remove its self-interference from the signal received from the relay provided that the required channel state information is available. Although the relay is half-duplex, i.e., requiring two channels to relay a signal, the two-way scheme allows two users to share both channels concurrently. This leads to a high spectral efficiency.

The authors of [2] appear to be the first who proposed the idea of two-way relay, where they considered a system with a single-antenna relay and single-antenna users. The work [3] studied coding schemes for a regenerative MIMO relay system. In [4], several two-way relay schemes were proposed and their capacity regions were explored. Insights were obtained for single-antenna users and single-antenna relay. In [5], the authors considered a power minimization problem for single-antenna users and a non-regenerative MIMO relay. For multi-antenna users, they proposed a heuristic sub-optimal method called dual channel matching strategy. The work [6] considered the maximization of the sum rates of two single-antenna users assisted by a non-regenerative MIMO relay under a high SNR assumption. In [7], a convex optimization algorithm was formulated to compute the capacity region of a relay system same as in [6] but without the high SNR assumption. The algorithm developed in [7] is however not applicable for multi-antenna users. In [8], suboptimal power allocation methods were proposed to maximize the system throughput of a non-regenerative MIMO relay system.

In this paper, we consider the problem of jointly designing/computing the spatial source covariance matrices (source matrices) and the spatial relay transformation matrix (relay matrix) for a non-regenerative two-way MIMO relay system where all nodes (i.e., the relay and the two users) have multiple antennas. In practice, one of the two "users" can be an access point, and the other a user equipment. In many situations, both users are already equipped with multiple antennas. When a relay is placed between them to improve the power and spectral efficiency, the relay should be a MIMO relay and hence each of the nodes in the system is a MIMO node. It

Manuscript received July 1, 2010; revised November 12, 2010 and January 26, 2011; accepted February 6, 2011. The associate editor coordinating the review of this paper and approving it for publication was M. Valenti.

The authors are with the Department of Electrical Engineering, University of California, Riverside, CA, 92521 (e-mail: {sxu, yhua}@ee.ucr.edu).

This work was supported in part by the National Science Foundation under Grant No. TF-0514736, the Army Research Laboratory under the Collaborative Technology Alliance Program, and the Army Research Office under the MURI Grant No. W911NF-04-1-0224. Part of this work was presented at IEEE ICASSP'2010 [1].

Digital Object Identifier 10.1109/TWC.2011.030911.101173

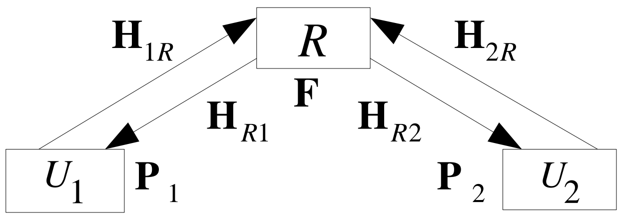


Fig. 1. A non-regenerative two-way MIMO relay system where the two users are denoted by U_1 and U_2 respectively, and the relay node by R.

is not hard to imagine that the source matrix (also called transmit covariance matrix) used at each user and the relay matrix used at the relay all affect the system capacity. It is therefore important to understand how to optimally design these matrices and how much capacity gain such a design can yield. This problem has not received enough attention from researchers.

It is known that a two-user (half-duplex) MIMO relay system as in Fig. 1 can be treated as *two* one-way (half-duplex) MIMO relay systems, and each one-way MIMO relay system can be treated as in [9]. Partially due to the known optimal structure of the source and relay matrices as revealed in [9], the complexity of designing the source and relay matrices for the one-way scheme is much lower than that for the two-way scheme. But for two users to exchange their information via a half-duplex relay, the one-way scheme has a factor-4 loss of spectral efficiency while the two-way scheme has only a factor-2 loss. Therefore, the two-way scheme is approximately twice as spectrally efficient as the one-way scheme.

We will present efficient algorithms for computing the source matrices and the relay matrix that maximize a (weighted) sum rate of the two-way system. Our algorithms are based on an alternate optimization approach where the source matrices and the relay matrix are optimized alternately until convergence. When the relay matrix is fixed, finding the optimal source matrices for the two multi-antenna users is a convex problem, and for uniformly weighted sum rate, it can be solved by the generalized water-filling algorithm developed in [10]. When the source matrices are fixed, we develop a hybrid gradient algorithm and an iterative weighted minimum mean square error (WMMSE) algorithm to find the best relay matrix. The hybrid gradient algorithm combines the gradient descent search and the Newton's search. The iterative WMMSE algorithm is based on an alternate convex search of the relay matrix and the MMSE equalizer at each user. This idea was inspired by the work [11] on source precoding for MIMO broadcasting. We will also show an optimal structure of the relay matrix, which is useful to reduce the computational complexity when the number of antennas at the relay is more than twice the number of the antennas at each user. This result is a generalization of one shown in [7] from single-antenna users to multi-antenna users. Furthermore, we will demonstrate that when applied to the case of single-antenna users, our search algorithms for the relay matrix yield the maximum sum rate much faster than the convex method developed in [7].

It is important to note that while the idea of gradient search is well known, its application to the problem addressed in this paper is the first. The problem structure is also exploited to simplify the expression and computation of the gradient vector and Hessian matrix. The resulting algorithm provides a new and useful perspective of other alternative algorithms such as one in [7] and WMMSE in this paper. This study reveals a capacity potential of a two-way MIMO relay system that no other existing work has been able to reveal.

The rest of the paper is organized as follows. In Section II, we describe in more details the non-regenerative two-way MIMO relay system, and formulate the optimization problems we aim to solve. In Section III, we present an optimal structure of the relay matrix under the condition that the number of antennas at the relay is larger than twice of that at each of the users. In IV, we present the hybrid gradient method to compute the optimal relay matrix with fixed source covariance matrices. In Section V, the iterative WMMSE method is presented. In Section VI, the procedure of finding the optimal source covariance matrices with fixed relay matrix is described. Section VII illustrates the performance of our algorithms. The conclusion is given in Section VIII.

II. Problem Formulation

The two-way relay system under consideration is illustrated in Fig. 1, which involves two users and one relay. The two users are denoted by U_1 and U_2 each with N antennas. Since the two users exchange information with each other, they are also referred to as sources and destinations. The relay is denoted by R with M antennas. We focus on a single carrier with a narrow enough bandwidth so that the channels between nodes are flat fading.

The two-way relay scheme has two phases. In the first phase, U_1 and U_2 transmit $\mathbf{x}_1 = \mathbf{P}_1\mathbf{s}_1$ and $\mathbf{x}_2 = \mathbf{P}_2\mathbf{s}_2$, respectively. Here, each of $\mathbf{s}_1$ and $\mathbf{s}_2$ consists of N independent complex symbols, $\mathbf{s}_i \in \mathbb{C}^{N\times 1}$ and $\mathbf{E}[\mathbf{s}_i\mathbf{s}_i^H] = \mathbf{I}$, $i = 1, 2$ and $\mathbf{P}_1$ and $\mathbf{P}_2$ are two square precoding matrices. The signal received at R can be expressed as

$$\mathbf{y}_R = \mathbf{H}_{1R}\mathbf{x}_1 + \mathbf{H}_{2R}\mathbf{x}_2 + \mathbf{n}_R \tag{1}$$

where $\mathbf{H}_{iR} \in \mathbb{C}^{M\times N}$, $i = 1, 2$, denotes the channel matrices (channel state information) from U_i to R, and $\mathbf{n}_R \in \mathbb{C}^{M\times 1}$ is the noise. We assume that the noise is complex white Gaussian, i.e., $\mathbf{n}_R \sim \mathcal{CN}(0, \mathbf{I})$.

In the second phase, R transmits $\mathbf{y}'_R = \mathbf{F}\mathbf{y}_R$ where $\mathbf{F} \in \mathbb{C}^{M\times M}$ is the relay transformation matrix. The transmit power consumed at the relay is

$$p_R = \mathrm{Tr}\left(\mathbf{F}\mathbf{H}_{1R}\mathbf{R}_{x_1}\mathbf{H}_{1R}^H\mathbf{F}^H + \mathbf{F}\mathbf{H}_{2R}\mathbf{R}_{x_2}\mathbf{H}_{2R}^H\mathbf{F}^H + \mathbf{F}\mathbf{F}^H\right) \tag{2}$$

where $\mathbf{R}_{x_i} = \mathbf{E}[\mathbf{x}_i\mathbf{x}_i^H] = \mathbf{P}_i\mathbf{P}_i^H$, $i = 1, 2$, denotes the source covariance matrix at U_i. Note that $\mathbf{P}_i$ is a matrix square root of $\mathbf{R}_{x_i}$. The signals received at U_i, $i = 1, 2$, are

$$\mathbf{y}_i = \mathbf{H}_{Ri}\mathbf{F}\mathbf{H}_{iR}\mathbf{x}_i + \mathbf{H}_{Ri}\mathbf{F}\mathbf{H}_{\bar{i}R}\mathbf{x}_{\bar{i}} + \mathbf{H}_{Ri}\mathbf{F}\mathbf{n}_R + \mathbf{n}_i \tag{3}$$

where $\mathbf{H}_{Ri} \in \mathbb{C}^{N\times M}$ denote the channel matrices from R to U_i, and $\mathbf{n}_i \in \mathbb{C}^{N\times 1}$ is the noise assumed to be $\mathcal{CN}(0, \mathbf{I})$. Here, $\bar{i} = 2$ for $i = 1$, and $\bar{i} = 1$ for $i = 2$. We see that $\mathbf{H}_{R1}\mathbf{F}\mathbf{H}_{1R}\mathbf{x}_1$ is the self-interference at U_1, and $\mathbf{H}_{R2}\mathbf{F}\mathbf{H}_{2R}\mathbf{x}_2$ is the self-interference at U_2. With a perfect knowledge of $\mathbf{H}_{R1}\mathbf{F}\mathbf{H}_{1R}$

at U_1 and $\mathbf{H}_{R2}\mathbf{F}\mathbf{H}_{2R}$ at U_2, each of the two users can cancel out its self-interference. After removing the self-interferences, the signals at U_1 and U_2 are

$$\mathbf{y}_i' = \mathbf{H}_{Ri}\mathbf{F}\mathbf{H}_{\bar{i}R}\mathbf{x}_{\bar{i}} + \mathbf{H}_{Ri}\mathbf{F}\mathbf{n}_R + \mathbf{n}_i \tag{4}$$

It is useful to note that the two phases required here may correspond to two frequency channels or two time slots. If two time slots are used, the symbol vector $\mathbf{y}_R$ needs to be stored at R for at least one symbol interval. If two frequency channels are used, $\mathbf{y}_R$ needs not to be stored at R, and the radio frequency output of R can be simply a frequency translation of its radio frequency input with the transformation $\mathbf{F}$. The implementation of $\mathbf{F}$ can be done digitally although no packet decoding or encoding is needed here. Both time division and frequency division for the two phases appear feasible.

Given the above two-way relay scheme and the signal model (4), the achievable data rate received at U_i is known to be $r_i = \frac{1}{2}\log_2\det\mathbf{X}_i$ (under ideal encoding and decoding) where

$$\mathbf{X}_i = \mathbf{I} + (\mathbf{H}_{Ri}\mathbf{F}\mathbf{F}^H\mathbf{H}_{Ri}^H + \mathbf{I})^{-1}\mathbf{H}_{Ri}\mathbf{F}\mathbf{H}_{\bar{i}R}\mathbf{R}_{x_{\bar{i}}}\mathbf{H}_{\bar{i}R}^H\mathbf{F}^H\mathbf{H}_{Ri}^H \tag{5}$$

The achievable (weighted) sum rate of this relay system is $r_{sum}(\mathbf{F}, \mathbf{R}_{x_1}, \mathbf{R}_{x_2}) = \mu_1 r_1 + \mu_2 r_2$, where the (non-negative) weights μ_1 and μ_2 can be chosen in any way subject to $\mu_1 + \mu_2 = 2$. For uniform weighting, we have $\mu_1 = \mu_2 = 1$.

The sum rate is a function of the source covariance matrices $\mathbf{R}_{x_1}$ and $\mathbf{R}_{x_2}$ and the relay matrix $\mathbf{F}$. One important goal is to develop fast algorithms to determine these matrices to maximize the sum rate subject to power constraints at U_1, U_2 and R. This problem can be expressed as follows:

$$\begin{aligned} \min_{\mathbf{F},\mathbf{R}_{x_1},\mathbf{R}_{x_2}} \quad & -r_{sum}(\mathbf{F}, \mathbf{R}_{x_1}, \mathbf{R}_{x_2}) \\ s.t. \quad & p_R(\mathbf{F}, \mathbf{R}_{x_1}, \mathbf{R}_{x_2}) \le P_R, \mathrm{Tr}(\mathbf{R}_{x_i}) \le P_i, i = 1, 2 \end{aligned} \tag{6}$$

This problem was not solved in the previous works [5], [6] and [7] except for the special case $N = 1$. When $N = 1$, there is no issue about the source matrices as $\mathbf{R}_{x_1}$ and $\mathbf{R}_{x_2}$ reduce to scalars. With $N > 1$ and $M > 1$, this problem is a much harder (non-convex) problem and there is currently no globally optimal solution.

In this paper, we propose to optimize the pair $\mathbf{R}_{x_1}$ and $\mathbf{R}_{x_2}$ and the single matrix $\mathbf{F}$ alternately. If we fix $\mathbf{R}_{x_1} = \mathbf{R}_{x_1}^{(k-1)}$ and $\mathbf{R}_{x_2} = \mathbf{R}_{x_2}^{(k-1)}$ where k be the index of the kth iteration, the problem of finding $\mathbf{F}$ is

$$\begin{aligned} \min_{\mathbf{F}} \quad & -r_{sum}(\mathbf{F}, \mathbf{R}_{x_1}^{(k-1)}, \mathbf{R}_{x_2}^{(k-1)}) \\ s.t. \quad & p_R(\mathbf{F}, \mathbf{R}_{x_1}^{(k-1)}, \mathbf{R}_{x_2}^{(k-1)}) \le P_R \end{aligned} \tag{7}$$

where P_R is the upper bound on the transmit power at the relay. We will refer to this sub-problem as the relay optimization problem.

If we fix $\mathbf{F} = \mathbf{F}^{(k)}$, the problem of finding $\mathbf{R}_{x_1}$ and $\mathbf{R}_{x_2}$ is

$$\begin{aligned} \min_{\mathbf{R}_{x_1},\mathbf{R}_{x_2}} \quad & -r_{sum}(\mathbf{F}^{(k)}, \mathbf{R}_{x_1}, \mathbf{R}_{x_2}) \\ s.t. \quad & p_R(\mathbf{F}^{(k)}, \mathbf{R}_{x_1}, \mathbf{R}_{x_2}) \le P_R \\ & \mathrm{Tr}(\mathbf{R}_{x_i}) \le P_i, \ \mathbf{R}_{x_i} \ge 0, \ i = 1, 2 \end{aligned} \tag{8}$$

where P_i is the upper bound on the transmit power at U_i, and $\mathbf{R}_{x_i} \ge 0$ means that the matrix $\mathbf{R}_{x_i}$ is positive semi-definite. We will refer to this sub-problem as the source optimization problem.

Our approach to finding $\mathbf{R}_{x_1}$, $\mathbf{R}_{x_2}$ and $\mathbf{F}$ is based on the alternation of the optimizations of the above two sub-problems until convergence. At the time of this writing, we do not know whether there exists a more effective approach. This alternate optimization between relay and sources is similar to one previously used for a one-way MIMO relay system of two or more hops [12], [9], [13]. But the problem here is much more complex and, to our knowledge, does not have the diagonalization property as available for the one-way MIMO relays.

Unlike the one-way (two-hop) scheme, the relay optimization problem here is still non-convex, for which we will present two algorithms in Section IV. The source optimization is convex, for which we will present an algorithm in Section VI based on the generalized water-filling algorithm developed in [10]. In the next section, we show an optimal structure of the relay matrix, which is useful to reduce the complexity of the problem.

III. Optimal Structure of the Relay Matrix

Theorem 1: Assume $M \ge 2N$. Define the following two QR decompositions:

$$\begin{bmatrix}\mathbf{H}_{R1}^H & \mathbf{H}_{R2}^H\end{bmatrix} = \mathbf{V}_1\mathbf{R}_1 \tag{9}$$

$$\begin{bmatrix}\mathbf{H}_{1R} & \mathbf{H}_{2R}\end{bmatrix} = \mathbf{U}_2\mathbf{R}_2 \tag{10}$$

where $\mathbf{R}_1, \mathbf{R}_2 \in \mathbb{C}^{2N\times 2N}$ are upper triangular matrices, and $\mathbf{V}_1, \mathbf{U}_2 \in \mathbb{C}^{M\times 2N}$ are orthonormal matrices. Then, the optimal relay matrix $\mathbf{F}$ that maximizes the sum rate in problem (7) has the following structure:

$$\mathbf{F} = \mathbf{V}_1\mathbf{A}\mathbf{U}_2^H \tag{11}$$

where $\mathbf{A} \in \mathbb{C}^{2N\times 2N}$.

Proof: See Appendix A. ■

Note that this optimal structure of $\mathbf{F}$ is governed by the channel matrices only, which is not affected by the source covariance matrices. It is also clear from the proof that this structure of $\mathbf{F}$ remains optimal for any weights μ_1 and μ_2.

If $M > 2N$, then the M^2 unknowns in $\mathbf{F}$ are effectively reduced to $4N^2$ unknowns in $\mathbf{A}$ for the relay optimization problem (7). If $M = 2N$, this theorem does not seem important. With this theorem, we can now write

$$\mathbf{F} = \mathbf{S}\mathbf{A}\mathbf{T}^H \text{ where } \begin{cases} \mathbf{S} = \mathbf{V}_1, \mathbf{T} = \mathbf{U}_2 & \text{if } M > 2N \\ \mathbf{S} = \mathbf{T} = \mathbf{I} & \text{if } M \le 2N \end{cases} \tag{12}$$

We also define $\mathbf{G}_{Ri} = \mathbf{H}_{Ri}\mathbf{S}$ and $\mathbf{G}_{iR} = \mathbf{T}^H\mathbf{H}_{iR}$. Then, (5) becomes

$$\mathbf{X}_i = \mathbf{I} + (\mathbf{G}_{Ri}\mathbf{A}\mathbf{A}^H\mathbf{G}_{Ri}^H + \mathbf{I})^{-1}\mathbf{G}_{Ri}\mathbf{A}\mathbf{G}_{\bar{i}R}\mathbf{R}_{x_{\bar{i}}}\mathbf{G}_{\bar{i}R}^H\mathbf{A}^H\mathbf{G}_{Ri}^H \tag{13}$$

and (2) becomes

$$p_R = \mathrm{Tr}\left(\mathbf{A}\mathbf{G}_{1R}\mathbf{R}_{x_1}\mathbf{G}_{1R}^H\mathbf{A}^H + \mathbf{A}\mathbf{G}_{2R}\mathbf{R}_{x_2}\mathbf{G}_{2R}^H\mathbf{A}^H + \mathbf{A}\mathbf{A}^H\right) \tag{14}$$

Although this theorem reduces the complexity (dimension of search space) of the problem when $M > 2N$, it does not change the structure of the remaining problem for finding $\mathbf{A}$.

In other words, the problem (7) with respect to $\mathbf{F}$ has the same structure as the equivalent problem with respect to $\mathbf{A}$. For convenience, we will treat $\mathbf{F}$ and $\mathbf{A}$ in (7) interchangeable.

It remains an open problem to find additional structure in the optimal $\mathbf{A}$. This difficulty is caused by the fact that the matrix $\mathbf{A}$ is weighted by different matrices in $\mathbf{X}_1$ and $\mathbf{X}_2$. In the next two sections, we present two algorithms for finding the optimal $\mathbf{A}$.

IV. Relay Optimization by Gradient Method

In this section, we will present a hybrid gradient method that dynamically switches between steepest gradient descent and Newton's search. We need to solve the relay optimization problem (7), i.e., $\min_{\mathbf{A}} -r_{sum}$ subject to $p_R \le P_R$. This is a non-convex problem because $-r_{sum}$ is not a convex function of $\mathbf{A}$. However, the set defined by $p_R \le P_R$ with respect to $\mathbf{A}$ is a convex set, i.e., if $\mathbf{A}_1$ and $\mathbf{A}_2$ satisfy $p_R \le P_R$, so does $\alpha\mathbf{A}_1 + (1-\alpha)\mathbf{A}_2$ where $0 \le \alpha \le 1$. Therefore, we can apply the interior point method to convert the hard constraint $p_R \le P_R$ into a soft constraint in the following problem:

$$\min_{\mathbf{A}} f(\mathbf{A}) = -r_{sum} - \frac{1}{t}\ln(P_R - p_R) \qquad (15)$$

where the second term in $f(\mathbf{A})$ is the soft constraint known as logarithmic barrier function [14], and t is the barrier factor. By increasing t gradually, the soft constraint becomes harder and harder. For each given t and a previous choice of $\mathbf{A}$ in the interior region (satisfying $p_R < P_R$), we can apply a gradient method to optimize $\mathbf{A}$. The loop of increasing t is called the outer loop, the gradient search under each fixed t is called the inner loop. There is no guarantee that this algorithm leads to the globally optimal solution unless the problem is convex. Since $-r_{sum}$ is non-convex of $\mathbf{A}$, multiple initializations and multiple runs of the algorithm are desirable to increase the likelihood of finding the globally optimal solution.

The most common gradient methods are the (steepest) gradient descent method and the Newton's method. Both of these methods are well documented in textbooks on optimization, such as Algorithm 9.3 (gradient descent method) and Algorithm 9.5 (Newton's method) in [14]. The key component in the gradient descent method is the computation of the gradient vector ∇f of $f(\mathbf{A})$. In the Newton's method, we need the gradient vector ∇f as well as the Hessian matrix $\nabla^2 f$ of $f(\mathbf{A})$. The vector of independent variables here is $\mathbf{x} = [Re(\mathbf{a})^T, Im(\mathbf{a})^T]^T$ where $\mathbf{a} = \mathrm{vec}(\mathbf{A})$. Hence, $\nabla f = \frac{\partial f}{\partial \mathbf{x}}$ and $\nabla^2 f = \frac{\partial^2 f}{\partial \mathbf{x}\partial \mathbf{x}^T}$. Alternatively, we can write $\nabla f = \left[\mathrm{vec}^T\left(\frac{\partial f}{\partial Re(\mathbf{A})}\right), \mathrm{vec}^T\left(\frac{\partial f}{\partial Im(\mathbf{A})}\right)\right]^T$ and

$$\nabla^2 f = \nabla(\nabla f)^T = \begin{bmatrix} \dfrac{\partial(\nabla f)^T}{\partial \mathrm{vec}(Re(\mathbf{A}))} \\ \dfrac{\partial(\nabla f)^T}{\partial \mathrm{vec}(Im(\mathbf{A}))} \end{bmatrix}$$

Next, we explain the key steps needed to derive and compute ∇f and $\nabla^2 f$.

A. Computation of Gradient

It follows from (15) that

$$\partial f(\mathbf{A}) = -\mu_1 \partial r_1 - \mu_2 \partial r_2 + \frac{1}{t(P_r - p_R)}\partial p_R \qquad (16)$$

where $\partial r_i = \frac{1}{2}(\log_2 e)\mathrm{Tr}(\mathbf{X}_i^{-1}\partial\mathbf{X}_i)$. For basic facts of matrix differential calculus, see [15]. Using (13), the chain rule of differentials, $\partial A^{-1} = -A^{-1}\partial A A^{-1}$ and $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$, one can verify

$$\begin{aligned}
\mathrm{Tr}(\mathbf{X}_i^{-1}\partial\mathbf{X}_i) = \\
&-\mathrm{Tr}\left(\mathbf{A}^H\mathbf{G}_{Ri}^H\mathbf{D}_i^{-1}\mathbf{N}_{\bar{i}}\mathbf{X}_i^{-1}\mathbf{D}_i^{-1}\mathbf{G}_{Ri}\partial\mathbf{A}\right) \\
&-\mathrm{Tr}\left(\mathbf{G}_{Ri}^H\mathbf{D}^{-1}\mathbf{N}_{\bar{i}}\mathbf{X}_i^{-1}\mathbf{D}_1^{-1}\mathbf{G}_{Ri}\mathbf{A}\partial\mathbf{A}^H\right) \\
&+\mathrm{Tr}\left(\mathbf{G}_{\bar{i}R}\mathbf{R}_{x_{\bar{i}}}\mathbf{G}_{\bar{i}R}^H\mathbf{A}^H\mathbf{G}_{Ri}^H\mathbf{X}_i^{-1}\mathbf{D}_i^{-1}\mathbf{G}_{Ri}\partial\mathbf{A}\right) \\
&+\mathrm{Tr}\left(\mathbf{G}_{Ri}^H\mathbf{X}_i^{-1}\mathbf{D}_i^{-1}\mathbf{G}_{Ri}\mathbf{A}\mathbf{G}_{\bar{i}R}\mathbf{R}_{x_{\bar{i}}}\mathbf{G}_{\bar{i}R}^H\partial\mathbf{A}^H\right) \quad (17)
\end{aligned}$$

where $\mathbf{D}_i = \mathbf{G}_{Ri}\mathbf{A}\mathbf{A}^H\mathbf{G}_{Ri}^H + \mathbf{I}$ and $\mathbf{N}_{\bar{i}} = \mathbf{G}_{Ri}\mathbf{A}\mathbf{G}_{\bar{i}R}\mathbf{R}_{x_{\bar{i}}}\mathbf{G}_{\bar{i}R}^H\mathbf{A}^H\mathbf{G}_{Ri}^H$. With these notations, we also have $\mathbf{X}_i = \mathbf{I} + \mathbf{D}_i^{-1}\mathbf{N}_{\bar{i}}$. Using the fact $A(I+BA)^{-1} = (I+AB)^{-1}A$ (which follows from the matrix inversion lemma), one can verify that $(\mathbf{N}_{\bar{i}}\mathbf{X}_i^{-1})^H = \mathbf{N}_{\bar{i}}\mathbf{X}_i^{-1}$. One can also verify that if $\partial f = \mathrm{Tr}(A\partial X) + \mathrm{Tr}(A^H\partial X^H)$ then $\frac{\partial f}{\partial Re(X)} = 2Re(A)^T$ and $\frac{\partial f}{\partial Im(X)} = -2Im(A)^T$. We call the two terms in $\mathrm{Tr}(A\partial X) + \mathrm{Tr}(A^H\partial X^H)$ a conjugate differential pair. The first two terms in (17) are a conjugate differential pair, and so are the last two terms in (17). Therefore,

$$\begin{aligned}
\frac{\partial r_i}{\partial Re(\mathbf{A})} &= -(\log_2 e)Re\left(\mathbf{A}^H\mathbf{G}_{Ri}^H\mathbf{D}_i^{-1}\mathbf{N}_{\bar{i}}\mathbf{X}_i^{-1}\mathbf{D}_i^{-1}\mathbf{G}_{Ri}\right)^T \\
&+(\log_2 e)Re\left(\mathbf{G}_{\bar{i}R}\mathbf{R}_{x_{\bar{i}}}\mathbf{G}_{\bar{i}R}^H\mathbf{A}^H\mathbf{G}_{Ri}^H\mathbf{X}_i^{-1}\mathbf{D}_i^{-1}\mathbf{G}_{Ri}\right)^T \quad (18)
\end{aligned}$$

and $\frac{\partial r_i}{\partial Im(\mathbf{A})}$ is the same as the above except that Re is replaced by $-Im$. Clearly,

$$\nabla r_i = \left[\mathrm{vec}^T\left(\frac{\partial r_i}{\partial Re(\mathbf{A})}\right), \mathrm{vec}^T\left(\frac{\partial r_i}{\partial Im(\mathbf{A})}\right)^T\right]^T.$$

Using the same technique, we can find

$$\frac{\partial p_R}{\partial Re(\mathbf{A})} = 2Re(\mathbf{G}_{1R}\mathbf{R}_{x_1}\mathbf{G}_{1R}^H\mathbf{A}^H + \mathbf{G}_{2R}\mathbf{R}_{x_2}\mathbf{G}_{2R}^H\mathbf{A}^H + \mathbf{A}^H)^T$$

and $\frac{\partial p_R}{\partial Im(\mathbf{A})}$ is the same but with Re replaced by $-Im$. Using the above results, the gradient $\nabla f = -\mu_1\nabla r_1 - \mu_2\nabla r_2 + \frac{1}{t(P_r - p_R)}\nabla p_R$ can be determined.

B. Computation of Hessian

To compute the Hessian $\nabla^2 f$, we need to compute each term in $\nabla^2 f = -\mu_1\nabla^2 r_1 - \mu_2\nabla^2 r_2 + \frac{1}{t(P_r - p_R)}\nabla^2 p_R$. The procedure of finding $\nabla^2 p_R$ is similar to that of finding $\nabla^2 r_i$ for $i = 1, 2$. To find $\nabla^2 r_i = \nabla(\nabla r_i)^T$, we need to find the elements $\frac{\partial\left(\frac{\partial r_i}{\partial Re(\mathbf{A})}\right)_{m,n}}{\partial Re(\mathbf{A})_{k,l}}$, $\frac{\partial\left(\frac{\partial r_i}{\partial Re(\mathbf{A})}\right)_{m,n}}{\partial Im(\mathbf{A})_{k,l}}$, $\frac{\partial\left(\frac{\partial r_i}{\partial Im(\mathbf{A})}\right)_{m,n}}{\partial Re(\mathbf{A})_{k,l}}$ and $\frac{\partial\left(\frac{\partial r_i}{\partial Im(\mathbf{A})}\right)_{m,n}}{\partial Im(\mathbf{A})_{k,l}}$ where $1 \le m, n, k, l \le M_0 \doteq \min(2N, M)$. In other words, each entry of the $2M_0^2 \times 2M_0^2$ matrix $\nabla^2 r_i$ is determined by one of the four elements for some m, n, k, l.

To find $\frac{\partial\left(\frac{\partial r_i}{\partial Re(\mathbf{A})}\right)_{m,n}}{\partial Re(\mathbf{A})_{k,l}}$, we need to take the first order differential of (18) to obtain

$$\begin{aligned}
&\partial \frac{\partial r_i}{\partial Re(\mathbf{A})} = \\
&-(\log_2 e) Re\left(\partial \mathbf{A}^H \mathbf{G}_{Ri}^H \mathbf{D}_i^{-1} \mathbf{N}_{\bar{i}} \mathbf{X}_i^{-1} \mathbf{D}_i^{-1} \mathbf{G}_{Ri}\right)^T \\
&+(\log_2 e) Re\left(\mathbf{A}^H \mathbf{G}_{Ri}^H \mathbf{D}_i^{-1} \partial \mathbf{D}_i \mathbf{D}_i^{-1} \mathbf{N}_{\bar{i}} \mathbf{X}_i^{-1} \mathbf{D}_i^{-1} \mathbf{G}_{Ri}\right)^T \\
&-(\log_2 e) Re\left(\mathbf{A}^H \mathbf{G}_{Ri}^H \mathbf{D}_i^{-1} \partial \mathbf{N}_{\bar{i}} \mathbf{X}_i^{-1} \mathbf{D}_i^{-1} \mathbf{G}_{Ri}\right)^T \\
&+(\log_2 e) Re\left(\mathbf{A}^H \mathbf{G}_{Ri}^H \mathbf{D}_i^{-1} \mathbf{N}_{\bar{i}} \mathbf{X}_i^{-1} \partial \mathbf{X}_i \mathbf{X}_i^{-1} \mathbf{D}_i^{-1} \mathbf{G}_{Ri}\right)^T \\
&+(\log_2 e) Re\left(\mathbf{A}^H \mathbf{G}_{Ri}^H \mathbf{D}_i^{-1} \mathbf{N}_{\bar{i}} \mathbf{X}_i^{-1} \mathbf{D}_i^{-1} \partial \mathbf{D}_i \mathbf{D}_i^{-1} \mathbf{G}_{Ri}\right)^T \\
&+(\log_2 e) Re\left(\mathbf{G}_{\bar{i}R} \mathbf{R}_{x_{\bar{i}}} \mathbf{G}_{\bar{i}R}^H \partial \mathbf{A}^H \mathbf{G}_{Ri}^H \mathbf{X}_i^{-1} \mathbf{D}_i^{-1} \mathbf{G}_{Ri}\right)^T \\
&-(\log_2 e) Re\left(\mathbf{G}_{\bar{i}R} \mathbf{R}_{x_{\bar{i}}} \mathbf{G}_{\bar{i}R}^H \mathbf{A}^H \mathbf{G}_{Ri}^H \mathbf{X}_i^{-1} \partial \mathbf{X}_i \mathbf{X}_i^{-1} \mathbf{D}_i^{-1} \mathbf{G}_{Ri}\right)^T \\
&-(\log_2 e) Re\left(\mathbf{G}_{\bar{i}R} \mathbf{R}_{x_{\bar{i}}} \mathbf{G}_{\bar{i}R}^H \mathbf{A}^H \mathbf{G}_{Ri}^H \mathbf{X}_i^{-1} \mathbf{D}_i^{-1} \partial \mathbf{D}_i \mathbf{D}_i^{-1} \mathbf{G}_{Ri}\right)^T
\end{aligned} \tag{19}$$

where $\partial \mathbf{D}_i$, $\partial \mathbf{N}_{\bar{i}}$ and $\partial \mathbf{X}_i$ can be determined similarly such that the resulting expression is a linear function of $\partial \mathbf{A}$ and $\partial \mathbf{A}^H$. To obtain $\frac{\partial\left(\frac{\partial r_i}{\partial Re(\mathbf{A})}\right)_{m,n}}{\partial Re(\mathbf{A})_{k,l}}$ from $\partial \frac{\partial r_i}{\partial Re(\mathbf{A})}$, we simply choose the (m,n)th element of $\partial \frac{\partial r_i}{\partial Re(\mathbf{A})}$, replace $\partial \mathbf{A}$ by $\mathbf{e}_k \mathbf{e}_l^T$ and replace $\partial \mathbf{A}^H$ by $\mathbf{e}_l \mathbf{e}_k^T$, where all entries of the vector $\mathbf{e}_k$ are zero except that its kth entry is one.

The expression of $\partial \frac{\partial r_i}{\partial Im(\mathbf{A})}$ is also given by (19) except that the operator Re should be replaced by $-Im$. Then, finding $\frac{\partial\left(\frac{\partial r_i}{\partial Im(\mathbf{A})}\right)_{m,n}}{\partial Re(\mathbf{A})_{k,l}}$ from $\partial \frac{\partial r_i}{\partial Im(\mathbf{A})}$ follows the same procedure as above.

The derivation for $\frac{\partial\left(\frac{\partial r_i}{\partial Re(\mathbf{A})}\right)_{m,n}}{\partial Im(\mathbf{A})_{k,l}}$ and $\frac{\partial\left(\frac{\partial r_i}{\partial Im(\mathbf{A})}\right)_{m,n}}{\partial Im(\mathbf{A})_{k,l}}$ is the same as for $\frac{\partial\left(\frac{\partial r_i}{\partial Re(\mathbf{A})}\right)_{m,n}}{\partial Re(\mathbf{A})_{k,l}}$ and $\frac{\partial\left(\frac{\partial r_i}{\partial Im(\mathbf{A})}\right)_{m,n}}{\partial Re(\mathbf{A})_{k,l}}$, respectively, except that for the former pair we should replace $\partial \mathbf{A}$ by $j\mathbf{e}_k \mathbf{e}_l^T$ and $\partial \mathbf{A}^H$ by $-j\mathbf{e}_l \mathbf{e}_k^T$.

The details of those expressions are tedious and omitted to save space. However, for computational efficiency, it is important to program the computations by starting from the sparse vector $\mathbf{e}_k$.

C. Hybrid Gradient Method

Since the function $f(\mathbf{A})$ is non-convex, there is no guarantee that the Hessian $\nabla^2 f$ is positive definite everywhere while a positive definite $\nabla^2 f$ is a necessary condition for the Newton's method to converge to even a local minimum. In fact, our simulation results show that the Newton's method does not always converge to a local minimum, and as a consequence the gradient descent method can sometimes yield a better result. Therefore, we propose a hybrid gradient method that combine the two gradient methods. The procedure of the hybrid gradient method is simple. At each iteration of the inner loop (under a fixed t), we compute both the gradient ∇f and the Hessian $\nabla^2 f$. If $\nabla^2 f$ is singular or the Newton's decrement $\lambda^2 = \nabla f \left(\nabla^2 f\right)^{-1} \nabla f$ is less than or equal to zero, we follow the gradient descent method. Otherwise, we choose the search (either Newton's or gradient descent) that produces a larger descent of $f(\mathbf{A})$. The iterations of the Newton's method stop when λ^2 is small enough. When $\mathbf{A}$ is close enough to a local minimum of $f(\mathbf{A})$, $\nabla^2 f$ is typically positive definite and in this case the Newton's method is known to have a quadratic convergence rate. For the gradient descent method, the convergence rate is known to be linear. In simulation, the hybrid gradient method has consistently produced either the same or better results than the pure Newton's method and the pure gradient descent method.

The hybrid gradient method is summarized as follows:

1) Choose $\epsilon > 0$, $\eta > 1$, $\alpha \in (0.01, 0.3)$, $\beta \in (0.1, 0.8)$. Initialize a feasible $\mathbf{A}$ or its equivalent $\mathbf{x}$. Set $t = t_0$. Note that the choice of ϵ governs the precision of the final result. See step 6). The choice of η (as long as larger than one) is not critical. See page 570 of [14]. The choices of $\alpha \in (0.01, 0.3)$ and $\beta \in (0.1, 0.8)$ are empirical as recommended on page 466 of [14].
2) Compute ∇f and $\nabla^2 f$.
3) If $\nabla^2 f$ is singular or $\lambda^2 = (\nabla f)^T (\nabla^2 f)^{-1} \nabla f \leq 0$, set $s = 1$ and $\Delta \mathbf{x}^{(1)} = -\nabla f$. Otherwise, set $s = 2$, $\Delta \mathbf{x}^{(1)} = -\nabla f$, $\Delta \mathbf{x}^{(2)} = -(\nabla^2 f)^{-1} \nabla f$. Set $t^{(i)} = 1$ for $i = 1, s$.
4) Line backtracking: for $i = 1, s$, while $f(\mathbf{x} + t^{(i)} \Delta \mathbf{x}^{(i)}) > f(\mathbf{x}) + \alpha t^{(i)} (\nabla f)^T \Delta \mathbf{x}^{(i)}$, $t^{(i)} := \beta t^{(i)}$
5) Update: if $f(\mathbf{x} + t^{(2)} \Delta \mathbf{x}^{(2)}) \leq f(\mathbf{x} + t^{(1)} \Delta \mathbf{x}^{(1)})$, let $\mathbf{x} := \mathbf{x} + t^{(2)} \Delta \mathbf{x}^{(2)}$. Otherwise, let $\mathbf{x} := \mathbf{x} + t^{(1)} \Delta \mathbf{x}^{(1)}$.
6) Go to Step 2 until $-(\nabla f)^T \Delta \mathbf{x} < \epsilon$.
7) Set $t := \eta t$. Go to Step 2 until $1/t < \epsilon$.

For convenience, we have used the same tolerance for both Steps 6 and 7 although different choices can be made. The inverse of $\nabla^2 f$ needs not and should not to be computed explicitly in order to compute $(\nabla^2 f)^{-1} \nabla f$ efficiently. It can be found by solving efficiently the linear equation $(\nabla^2 f)\mathbf{y} = \nabla f$. In simulation, we will choose $\epsilon = 10^{-3}$, $\eta = 3$, $\alpha = 0.01$, $\beta = 0.5$ and $t_0 = 1$. Only a locally optimal solution can be found by the hybrid gradient method with a given initial choice of $\mathbf{A}$. To increase the likelihood of finding the globally optimal solution, multiple initializations and multiple runs of the search algorithm are useful.

V. Relay Optimization by Iterative WMMSE Method

In this section, we present an alternative algorithm for the relay optimization problem (7) with respect to $\mathbf{A}$. Recall $r_{sum} = \frac{1}{2} \sum_{i=1}^{2} \mu_i \log_2 \det(\mathbf{X}_i)$ where $\mathbf{X}_i$ is given in (13). Since $\det(I + AB) = \det(I + BA)$, we can write $r_{sum} = \frac{1}{2} \sum_{i=1}^{2} \mu_i \log_2 \det(\mathbf{Y}_i)$ where

$$\begin{aligned}
\mathbf{Y}_i \;=\; & \mathbf{I} + \mathbf{R}_{x_{\bar{i}}}^{H/2} \mathbf{G}_{\bar{i}R}^H \mathbf{A}^H \mathbf{G}_{Ri}^H \cdot \\
& \left(\mathbf{G}_{Ri} \mathbf{A} \mathbf{A}^H \mathbf{G}_{Ri}^H + \mathbf{I}\right)^{-1} \mathbf{G}_{Ri} \mathbf{A} \mathbf{G}_{\bar{i}R} \mathbf{R}_{x_{\bar{i}}}^{1/2}
\end{aligned} \tag{20}$$

We define $J_i = \mathrm{Tr}\left(\mathbf{W}_i \mathbf{E}_i\right) - \log\det(\mathbf{W}_i) - N$ where $\mathbf{E}_i = E\left(\left(\mathbf{Q}_i \mathbf{y}_i' - \mathbf{s}_{\bar{i}}\right)\left(\mathbf{Q}_i \mathbf{y}_i' - \mathbf{s}_{\bar{i}}\right)^H\right)$ and $\mathbf{y}_i'$ is given in (4) or equivalently

$$\mathbf{y}_i' = \mathbf{G}_{Ri} \mathbf{A} \mathbf{G}_{\bar{i}R} \mathbf{R}_{x_{\bar{i}}}^{1/2} \mathbf{s}_{\bar{i}} + \mathbf{G}_{Ri} \mathbf{A} \mathbf{n}_R + \mathbf{n}_i \tag{21}$$

We next consider the following alternative problem

$$\begin{aligned}
\min_{\mathbf{A}, \mathbf{Q}_i, \mathbf{W}_i, i=1,2} \quad & J = \mu_1 J_1 + \mu_2 J_2 \\
s.t. \quad & p_R(\mathbf{A}) \leq P_R
\end{aligned} \tag{22}$$

It is easy to see that the optimal solution of $\mathbf{Q}_i$ from (22) is $\mathbf{Q}_i = E(\mathbf{s}_{\bar{i}} \mathbf{y}_i'^H) \left(E(\mathbf{y}_i' \mathbf{y}_i'^H)\right)^{-1}$. Using (21), one can verify that

$$\mathbf{Q}_i = \mathbf{R}_{x_{\bar{i}}}^{\frac{H}{2}} \mathbf{G}_{\bar{i}R}^H \mathbf{A}^H \mathbf{G}_{Ri}^H \cdot \left(\mathbf{G}_{Ri}\mathbf{A}\mathbf{G}_{\bar{i}R}\mathbf{R}_{x_{\bar{i}}}\mathbf{G}_{\bar{i}R}^H\mathbf{A}^H\mathbf{G}_{Ri}^H + \mathbf{G}_{Ri}\mathbf{A}\mathbf{A}^H\mathbf{G}_{Ri}^H + \mathbf{I}\right)^{-1}$$
$$= \Big(\mathbf{I} + \mathbf{R}_{x_{\bar{i}}}^{\frac{H}{2}}\mathbf{G}_{\bar{i}R}^H\mathbf{A}^H\mathbf{G}_{Ri}^H\left(\mathbf{G}_{Ri}\mathbf{A}\mathbf{A}^H\mathbf{G}_{Ri}^H + \mathbf{I}\right)^{-1}\mathbf{G}_{Ri}\mathbf{A}\,\mathbf{G}_{\bar{i}R}\mathbf{R}_{x_{\bar{i}}}^{\frac{1}{2}}\Big)^{-1}\mathbf{R}_{x_{\bar{i}}}^{\frac{H}{2}}\mathbf{G}_{\bar{i}R}^H\mathbf{A}^H\mathbf{G}_{Ri}^H\left(\mathbf{G}_{Ri}\mathbf{A}\mathbf{A}^H\mathbf{G}_{Ri}^H + \mathbf{I}\right)^{-1} \quad (23)$$

where the last equation follows from the matrix inverse lemma. With the definition of $\mathbf{Y}_i$ in (20) and the optimal $\mathbf{Q}_i$ in (23), it is easy to verify that $\mathbf{E}_i = \mathbf{Y}_i^{-1}$. It is also an easy task to verify that the optimal solution $\mathbf{W}_i$ from (22) is simply $\mathbf{W}_i = \mathbf{Y}_i$. We see that with the optimal $\mathbf{Q}_i$ and $\mathbf{W}_i$, $J_i = -\log\det(\mathbf{Y}_i)$ and hence $J = -2(\log 2)r_{sum}$. Therefore, we have shown that the optimal solution $\mathbf{A}$ to (22) is the same as that to the original problem (7).

Furthermore, it is important to notice that if we fix $\{\mathbf{Q}_i,\ i = 1,2\}$ and $\{\mathbf{W}_i,\ i = 1,2\}$, the problem of (22) with respect to $\mathbf{A}$ is a quadratic convex problem. Therefore, we can try to find the solution to (22) by optimizing $\{\mathbf{Q}_i,\ i = 1,2\}$, $\{\mathbf{W}_i,\ i = 1,2\}$ and $\mathbf{A}$ in a cyclic fashion, which is the following iterative WMMSE algorithm:

Given $\forall\mathbf{A}^{(0)}$;

Repeat Update $k := k+1$;

1) Compute $\mathbf{Q}_i^{(k)}$ based on $\mathbf{A}^{(k-1)}$ (see (23));
2) Compute $\mathbf{W}_i^{(k)} = \mathbf{Y}_i^{(k)}$ based on $\mathbf{A}^{(k-1)}$ (see (20))
3) Compute $\mathbf{A}^{(k)}$ by solving (22) and fixing $\mathbf{Q}_i = \mathbf{Q}_i^{(k)}$ and $\mathbf{W}_i = \mathbf{W}_i^{(k)}$.

Until Convergence

Although there is no proof or disproof that the iterative WMMSE algorithm yields the optimal $\mathbf{A}$ of the original problem (7), this algorithm is guaranteed to converge locally due to the minimization of J with respect to $\{\mathbf{Q}_i,\ i = 1,2\}$, $\{\mathbf{W}_i,\ i = 1,2\}$ and $\mathbf{A}$ cyclically.

We next show how to implement Step 3 of the above algorithm. First, using (21) and performing the expectation in $\mathbf{E}_i$, one can verify that, after removing the constant term $\log\det(\mathbf{W}_i) + N$,

$$\begin{aligned} J_i \;=\; & \mathrm{Tr}(\mathbf{W}_i\mathbf{Q}_i\mathbf{G}_{Ri}\mathbf{A}\mathbf{G}_{\bar{i}R}\mathbf{R}_{x_{\bar{i}}}\mathbf{G}_{\bar{i}R}^H\mathbf{A}^H\mathbf{G}_{Ri}^H\mathbf{Q}_i^H) \\ & + \mathrm{Tr}(\mathbf{W}_i\mathbf{Q}_i\mathbf{G}_{Ri}\mathbf{A}\mathbf{A}^H\mathbf{G}_{Ri}^H\mathbf{Q}_i^H) + \mathrm{Tr}(\mathbf{W}_i\mathbf{Q}_i\mathbf{Q}_i^H) \\ & + \mathrm{Tr}(\mathbf{W}_i) - \mathrm{Tr}(\mathbf{W}_i\mathbf{Q}_i\mathbf{G}_{Ri}\mathbf{A}\mathbf{G}_{\bar{i}R}\mathbf{R}_{x_{\bar{i}}}^{\frac{1}{2}}) \\ & - \mathrm{Tr}(\mathbf{W}_i\mathbf{R}_{x_{\bar{i}}}^{\frac{H}{2}}\mathbf{G}_{\bar{i}R}^H\mathbf{A}^H\mathbf{G}_{Ri}^H\mathbf{Q}_i^H) \end{aligned} \quad (24)$$

Recall $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$, $\mathrm{Tr}(AB^H) = \mathrm{vec}^H(B)\mathrm{vec}(A)$, $\mathrm{Tr}(ABA^HC^H) = \mathrm{vec}^H(CA)\mathrm{vec}(AB) = \mathrm{vec}(A)^H(I \bigotimes C)^H (B^T \bigotimes I)\mathrm{vec}(A) = \mathrm{vec}^H(A)\ (B^T \bigotimes C^H)\mathrm{vec}(A)$. It follows that

$$J_i = \mathbf{a}^H\mathbf{A}_i\mathbf{a} - \mathbf{c}_i^H\mathbf{a} - \mathbf{a}^H\mathbf{c}_i + d_i \quad (25)$$

where $\mathbf{a} = \mathrm{vec}(\mathbf{A})$ and

$$\begin{aligned} \mathbf{A}_i \;=\; & (\mathbf{G}_{\bar{i}R}\mathbf{R}_{x_{\bar{i}}}\mathbf{G}_{\bar{i}R}^H)^T \otimes (\mathbf{G}_{Ri}^H\mathbf{Q}_i^H\mathbf{W}_i\mathbf{Q}_i\mathbf{G}_{Ri}) \\ & + \mathbf{I} \otimes (\mathbf{G}_{Ri}^H\mathbf{Q}_i^H\mathbf{W}_i\mathbf{Q}_i\mathbf{G}_{Ri}) \end{aligned} \quad (26)$$

$$\mathbf{c}_i = \mathrm{vec}(\mathbf{G}_{Ri}^H\mathbf{Q}_i^H\mathbf{W}_i\mathbf{R}_{x_{\bar{i}}}^{\frac{H}{2}}\mathbf{G}_{\bar{i}R}^H) \quad (27)$$

$$d_i = \mathrm{Tr}(\mathbf{W}_i\mathbf{Q}_i\mathbf{Q}_i^H) + \mathrm{Tr}(\mathbf{W}_i) \quad (28)$$

Also, we can write (14) as $p_R = \mathbf{a}^H\mathbf{G}_R\mathbf{a}$ with $\mathbf{G}_R = (\mathbf{G}_{1R}\mathbf{R}_{x_1}\mathbf{G}_{1R}^H + \mathbf{G}_{2R}\mathbf{R}_{x_2}\mathbf{G}_{2R}^H + \mathbf{I})^T \bigotimes \mathbf{I}$

Therefore, step 3 is equivalent to

$$\begin{aligned} \min_{\mathbf{a}} \quad & \mathbf{a}^H\mathbf{G}_0\mathbf{a} - \mathbf{c}^H\mathbf{a} - \mathbf{a}^H\mathbf{c} \\ s.t. \quad & \mathbf{a}^H\mathbf{G}_R\mathbf{a} - P_R \le 0 \end{aligned} \quad (29)$$

where $\mathbf{G}_0 = \mu_1\mathbf{A}_1 + \mu_2\mathbf{A}_2$ and $\mathbf{c} = \mu_1\mathbf{c}_1 + \mu_2\mathbf{c}_2$. This is a simple convex problem which can be solved by the KKT method [14]. Specifically, the optimal $\mathbf{a}$ is uniquely determined by the following equations:

$$\begin{cases} \mathbf{G}_0\mathbf{a} - \mathbf{c} + \xi\mathbf{G}_R\mathbf{a} = \mathbf{0} \\ \xi(\mathbf{a}^H\mathbf{G}_R\mathbf{a} - P_R) = 0 \\ \mathbf{a}^H\mathbf{G}_R\mathbf{a} - P_R \le 0 \end{cases} \quad (30)$$

where $\xi \in \mathbb{R}$ and $\xi \ge 0$. There are two possible cases for the optimal solution $\mathbf{a}$. The first is when the constraint is not active, i.e., $\mathbf{a}^H\mathbf{G}_R\mathbf{a} < P_R$ or equivalently $\xi = 0$, and the second is when the constraint is active, i.e., $\mathbf{a}^H\mathbf{G}_R\mathbf{a} = P_R$ or equivalently $\xi > 0$. In the first case, the solution is $\mathbf{a} = \mathbf{G}_0^{-1}\mathbf{c}$. If this solution does not meet the power constraint, we can simply abandon it and consider the second case. In the second case, we have $\mathbf{a} = (\mathbf{G}_0 + \xi\mathbf{G}_R)^{-1}\mathbf{c}$ where $\xi > 0$ is such that

$$h(\xi) = \mathbf{c}^H(\mathbf{G}_0 + \xi\mathbf{G}_R)^{-1}\mathbf{G}_R(\mathbf{G}_0 + \xi\mathbf{G}_R)^{-1}\mathbf{c} = P_R \quad (31)$$

which can be solved by a 1-D search such as the bisection method since $h(\xi)$ is monotonically decreasing function of ξ. In simulation, we will choose the error tolerance $\epsilon = 10^{-3}$ for the 1-D search.

The iterative WMMSE method was inspired by the work [11] where a similar idea was used for MIMO broadcast. In Section VII, we will compare the iterative WMMSE method with the hybrid gradient method.

VI. Source Optimization by Generalized Water Filling

We now consider the source optimization problem (8) with any fixed $\mathbf{F}$. To highlight $\mathbf{R}_{x_1}$ and $\mathbf{R}_{x_2}$ which are the variables of interest now, we first reformulate (8) as (ignoring the factor 1/2):

$$\begin{aligned} \min_{\mathbf{R}_{x_1},\mathbf{R}_{x_1}} \quad & -\mu_2\log_2\det(\mathbf{I} + \mathbf{H}_1\mathbf{R}_{x_1}\mathbf{H}_1^H) \\ & -\mu_1\log_2\det(\mathbf{I} + \mathbf{H}_2\mathbf{R}_{x_2}\mathbf{H}_2^H) \\ s.t. \quad & \mathrm{Tr}(\mathbf{R}_{x_i}) \le P_i, \quad \mathbf{R}_{x_i} \ge 0, \quad i = 1,2 \\ & \mathrm{Tr}(\mathbf{G}_1\mathbf{R}_{x_1}\mathbf{G}_1^H + \mathbf{G}_2\mathbf{R}_{x_2}\mathbf{G}_2^H) \le P_3 \end{aligned} \quad (32)$$

where $\mathbf{H}_i = \left(\mathbf{H}_{iR}^H\mathbf{F}^H\mathbf{H}_{R\bar{i}}^H(\mathbf{H}_{R\bar{i}}\mathbf{F}\mathbf{F}^H\mathbf{H}_{R\bar{i}}^H + \mathbf{I})^{-1}\mathbf{H}_{R\bar{i}}\mathbf{F}\mathbf{H}_{iR}\right)^{\frac{1}{2}}$, $\mathbf{G}_i = \mathbf{F}\mathbf{H}_{iR}$, $P_3 = P_R - \mathbf{F}\mathbf{F}^H$. For any μ_1 and μ_2, the above problem is convex and can be solved by (general purpose) semi-definite programming (SDP) such as in CVX [16] . But if $\mu_1 = \mu_2$, there is an alternative method shown below.

With the uniform weights, we can write (32) as

$$\begin{aligned} \min_{\mathbf{R}_x \ge 0} \quad & -\log_2\det(\mathbf{I} + \mathbf{H}\mathbf{R}_x\mathbf{H}^H) \\ s.t. \quad & \mathrm{Tr}(\mathbf{B}_i\mathbf{R}_x\mathbf{B}_i^H) \le P_i, \quad i = 1,2,3 \end{aligned} \quad (33)$$

where $\mathbf{H} = \mathrm{diag}(\mathbf{H}_1, \mathbf{H}_2)$, $\mathbf{B}_1 = \mathrm{diag}(\mathbf{I}, \mathbf{0})$, $\mathbf{B}_2 = \mathrm{diag}(\mathbf{0}, \mathbf{I})$, $\mathbf{B}_3 = \mathrm{diag}(\mathbf{G}_1, \mathbf{G}_2)$. It is clear that if the problem

(33) has the diagonal constraint $\mathbf{R}_x = \mathrm{diag}(\mathbf{R}'_{x_1}, \mathbf{R}'_{x_2})$ where $\mathbf{R}'_{x_1}$ and $\mathbf{R}'_{x_2}$ have the same dimensions as $\mathbf{R}_{x_1}$ and $\mathbf{R}_{x_2}$, then (32) and (33) are exactly the same problem and hence have the same solution, i.e., $\mathbf{R}'_{x_1} = \mathbf{R}_{x_1}$ and $\mathbf{R}'_{x_2} = \mathbf{R}_{x_2}$.

Next, we show that $\mathbf{R}_x = \mathrm{diag}(\mathbf{R}'_{x_1}, \mathbf{R}'_{x_2})$ is always the form of the solution to (33). Without loss of generality, we can write

$$\mathbf{R}_x = \begin{bmatrix} \mathbf{R}'_{x_1} & \mathbf{R} \\ \mathbf{R}^H & \mathbf{R}'_{x_2} \end{bmatrix}$$

It is easy to see that due to the structures of $\mathbf{H}$, $\mathbf{B}_1$, $\mathbf{B}_2$ and $\mathbf{B}_3$, the off-diagonal block $\mathbf{R}$ of $\mathbf{R}_x$ has no effect on the constraints in (33). Now, we focus on the objective function of (33), for which we know, due to the Fischer's inequality [17], that

$$\begin{aligned} \det(\mathbf{I} + \mathbf{H}\mathbf{R}_x\mathbf{H}^H) &= \det \begin{bmatrix} \mathbf{I} + \mathbf{H}_1\mathbf{R}'_{x_1}\mathbf{H}_1^H & \mathbf{H}_1\mathbf{R}\mathbf{H}_2^H \\ \mathbf{H}_2\mathbf{R}^H\mathbf{H}_1^H & \mathbf{I} + \mathbf{H}_2\mathbf{R}'_{x_2}\mathbf{H}_2^H \end{bmatrix} \\ &\le \det \begin{bmatrix} \mathbf{I} + \mathbf{H}_1\mathbf{R}'_{x_1}\mathbf{H}_1^H & \mathbf{0} \\ \mathbf{0} & \mathbf{I} + \mathbf{H}_2\mathbf{R}'_{x_2}\mathbf{H}_2^H \end{bmatrix} \\ &= \det(\mathbf{I} + \mathbf{H}\mathrm{diag}(\mathbf{R}'_{x_1}, \mathbf{R}'_{x_2})\mathbf{H}^H) \qquad (34) \end{aligned}$$

Therefore, the solution to (33) must be block diagonal.

For (33), we can directly apply the generalized water filling (GWF) theorem from [10], which yields

$$\mathbf{R}_x = \mathbf{K}^{-H}\mathbf{V}(\mathbf{I} - \mathbf{\Sigma}^{-2})^+\mathbf{V}^H\mathbf{K}^{-1} \qquad (35)$$

where $\mathbf{K} = (\sum_{i=1}^3 \theta_i \mathbf{B}_i^H\mathbf{B}_i)^{\frac{1}{2}}$, $\mathbf{V}$ and $\mathbf{\Sigma}$ are given by the SVD $\mathbf{H}\mathbf{K}^{-H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$, and $x^+ = \max(x, 0)$ is applied to each diagonal element in $(\mathbf{I} - \mathbf{\Sigma}^{-2})^+$. The Lagrangian variables $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3)$ are the solution to the dual problem

$$\begin{aligned} &\max_{\boldsymbol{\theta} \ge \mathbf{0}} - \log_2 \det(\mathbf{I} + \mathbf{H}\mathbf{R}_x\mathbf{H}^H) + \sum_{i=1}^{3} \theta_i \left(\mathrm{Tr}(\mathbf{B}_i\mathbf{R}_x\mathbf{B}_i^H) - P'_i\right) \\ &s.t. \mathbf{R}_x = \mathbf{K}^{-H}\mathbf{V}(\mathbf{I} - \mathbf{\Sigma}^{-2})^+\mathbf{V}^H\mathbf{K}^{-1} \qquad (36) \end{aligned}$$

The dual problem is convex and can be solved by the Newton's method [10]. In our Matlab simulation, the GWF algorithm is much faster than CVX to solve (33).

Finally, we would like to add that the source optimization problem (32) differs from a conventional single-link MIMO channel problem such as in [18]. For the former, the optimal source matrices are not necessarily diagonal in general.

VII. Simulation Results

For simulation, we choose the uniform weights $\mu_1 = \mu_2 = 1$, and the elements in $\mathbf{H}_{1R}$, $\mathbf{H}_{2R}$, $\mathbf{H}_{R1}$, and $\mathbf{H}_{R2}$ as independent complex circular Gaussian random variables of zero mean and unit variance, i.e., $\mathcal{CN}(0, 1)$. For illustration of the convergence behaviors of our algorithms, we will use one realization of the channel matrices. For illustration of the averaged sum rate, we will use multiple (100) realizations of the channel matrices. We assume that all noise vectors are complex white Gaussian, i.e., $\mathbf{n}_R \sim \mathcal{CN}(0, \mathbf{I})$ and $\mathbf{n}_i \sim \mathcal{CN}(0, \mathbf{I})$ where $i = 1, 2$. We denote $\mathrm{SNR}_R = P_R/M$ and $\mathrm{SNR}_i = P_i/N$ where $i = 1, 2$.

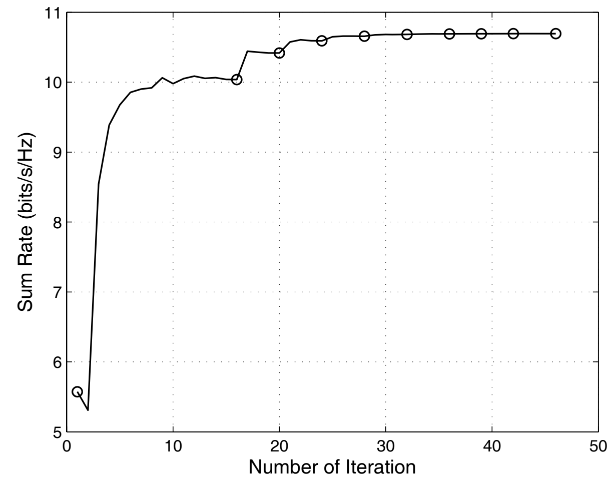


Fig. 2. The sum rate by the hybrid gradient method versus its iteration. The circles indicate the iteration steps where t is increased by the factor three, and the iterations between two adjacent circles are the iterations of either the (steepest) gradient descent method or Newton's method for a fixed t. $N = 2$ and $M = 6$. $\mathrm{SNR}_1 = \mathrm{SNR}_2 = \mathrm{SNR}_R = 10dB$.

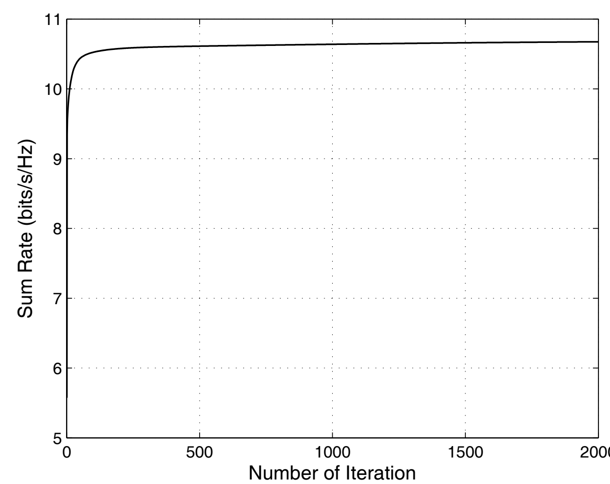


Fig. 3. The sum rate by the iterative WMMSE method versus its iteration. $N = 2$ and $M = 6$. $\mathrm{SNR}_1 = \mathrm{SNR}_2 = \mathrm{SNR}_R = 10dB$.

A. Relay Optimization

In this subsection, we assume that $\mathbf{R}_{\mathbf{x}_1}$ and $\mathbf{R}_{\mathbf{x}_2}$ are diagonal matrix with identical entries subject to $\mathrm{Tr}(\mathbf{R}_{\mathbf{x}_1}) = P_1$ and $\mathrm{Tr}(\mathbf{R}_{\mathbf{x}_2}) = P_2$. We randomly generate the initial $\mathbf{A}$ satisfying $p_R(\mathbf{A}) = P_R - \theta$ where θ is a (small) positive value so that $-\ln(P_R - p_R(\mathbf{A}))$ is not too large. This condition is required for the gradient methods although it is not necessary for the iterative MMSE method.

Note that the hybrid gradient method is always much faster than the gradient descend method and yields the same or better result than the Newton's method. Only the comparison between the hybrid gradient method and the iterative WMMSE method is shown next.

Under all common initializations of $\mathbf{A}$, the hybrid gradient method and the iterative WMMSE method have consistently

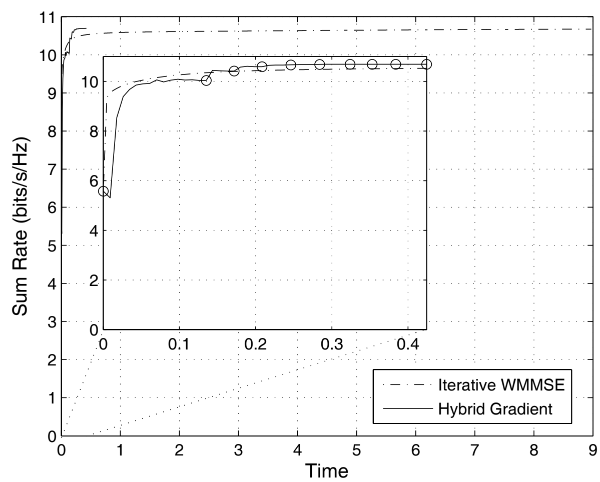


Fig. 4. The time-capacity curves of the hybrid gradient method and the iterative WMMSE method. The time is in seconds. $N = 2$ and $M = 6$. $\text{SNR}_1 = \text{SNR}_2 = \text{SNR}_R = 10dB$.

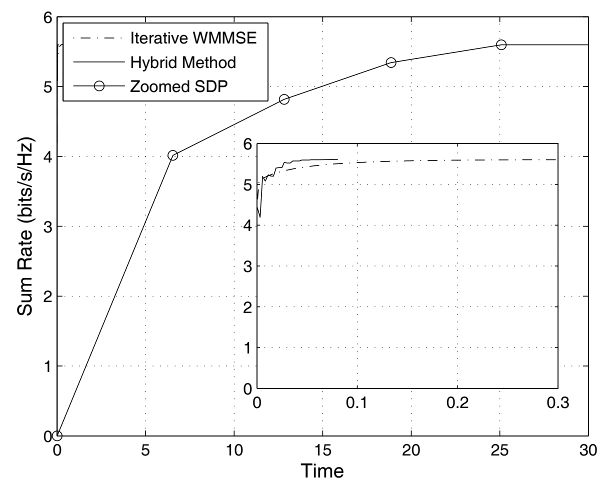


Fig. 5. The time-capacity curves of the zoomed SDP method based on [7], the hybrid gradient method and the iterative WMMSE method. $N = 1$ and $M = 5$. The time is in seconds. $\text{SNR}_1 = \text{SNR}_2 = \text{SNR}_R = 10dB$. The zoomed SDP is approximately 1000 times slower than the hybrid gradient.

yielded the same maximized sum rate upon convergence. Fig. 2 illustrates the convergence behavior of the hybrid gradient method, which is the value of the sum rate versus the iteration index. The circles indicate the locations along the iteration process where the barrier constant $t = 3^n$ is increased by the factor three. Fig 3 illustrates the convergence behavior of the iterative WMMSE method versus its iteration cycle of renewing $\mathbf{Q}_1$, $\mathbf{Q}_2$, $\mathbf{W}_1$ and $\mathbf{W}_2$ in the cost $J_1 + J_2$.

To provide a more precise comparison of the speed of the two methods, Fig. 4 illustrates the sum rate by the hybrid gradient method and the iterative WMMSE method versus the actual computational time (in Matlab on the same computer) along the iteration process. We call such curves as time-capacity curves. (Undoubtedly, these curves would be affected by how each algorithm is implemented. But we believe that these curves are meaningful to reflect the major performance gaps between algorithms.) Each point on a time-capacity curve is a pair of values of time and sum rate. The time is the time an algorithm takes to produce the corresponding sum rate. For each of the two methods, we embedded stopwatch check points in the programs, and the times (in seconds) and the corresponding values of intermediate $\mathbf{A}$ are collected. For each $\mathbf{A}$, there is a corresponding value of the sum rate. We see that when the sum rate is close to the optimal, the hybrid gradient method is much faster. Note that the Newton's method has a quadratic convergence rate near the optimal point. On the other hand, when the sum rate is far from the optimal, the iterative WMMSE method converges faster initially.

We like to mention that the method shown in [8] is a non-iterative sub-optimal method. This method is faster than a single iteration of the hybrid gradient or the iterative WMMSE method. However, using the method in [8], we obtained the sum rate at about 8.5 bits/s/Hz, which is significantly smaller than the maximum sum rate shown in Fig. 4.

The work [7] considered a special case of the two-way MIMO relay system where each user has a single antenna. In this case, there is no need for source optimization but only relay optimization. In [7], an algorithm based on SDP was developed to compute the capacity region of this special case. This algorithm can be easily adopted to compute the maximum sum rate as well. In Appendix B, a "zoomed" SDP algorithm is formulated to compute the maximum sum rate, which is a modified (and faster) version of the original SDP algorithm shown in [7]. Upon convergence, the zoomed SDP method has consistently produced the same sum rates as the hybrid gradient and iterative WMMSE methods under the same conditions. Fig. 5 illustrates the time-capacity curves of the three methods for $N = 1$ and $M = 5$. In this figure, the zoomed SDP method is about 1000 times slower than the hybrid gradient method to yield the same maximized sum rate.

B. Joint Source-Relay Optimization

When $N > 1$ and $M > 1$, both the relay optimization and the source optimization are important. We now let $\text{SNR}_1 = \text{SNR}_2 = \text{SNR}_R = \text{SNR}$. Fig. 6 shows the averaged sum rates achieved by different schemes where SNR varies from 10dB to 50dB. We used 100 channel realizations for each of the averaged sum rates.

The scheme of "no optimization (F)" means that both the source matrices, $\mathbf{R}_{\mathbf{x}_1}$ and $\mathbf{R}_{\mathbf{x}_2}$, and the relay matrix $\mathbf{F}$ were chosen randomly but meet power constraints at source and relay. The scheme of "no optimization (A)" means that the relay matrix $\mathbf{F}$ meets the optimal structure (11) but otherwise is randomly chosen. For "no optimization (F)" and "no optimization (A)", the same source matrices were used. The scheme of "source only (F)" means that the source matrices were optimized but the relay matrix $\mathbf{F}$ was chosen as in "no optimization(F)". The scheme of "source only (A)" means that the source matrices were optimized and $\mathbf{A}$ was chosen as in "no optimization(A)". The scheme of "relay only" means that the source matrices were chosen as in "no optimization" but the relay matrix $\mathbf{F}$ was completely optimized. The scheme of "joint source-relay optimization" means that both the source

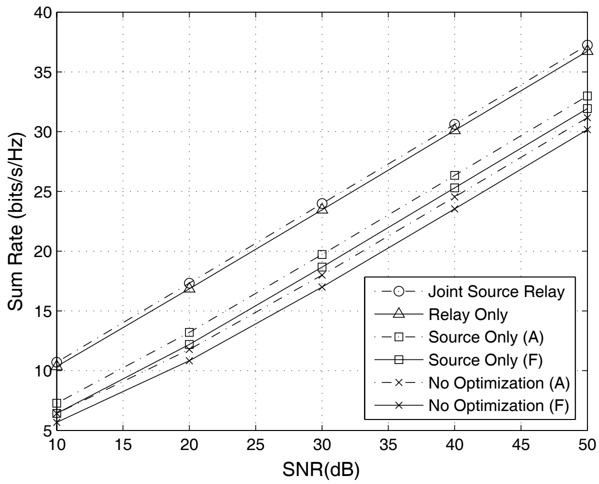


Fig. 6. Average sum rate under different schemes versus SNR $=$ $\mathrm{SNR}_1 = \mathrm{SNR}_2 = \mathrm{SNR}_R$. Averaged over 100 randomly generated channels. $N = 2$ and $M = 6$.

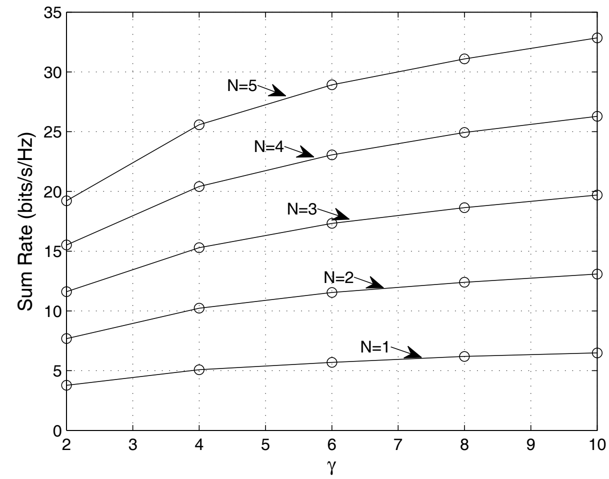


Fig. 7. Averaged sum rate of the two-way relay system under joint source-and-relay optimization versus $\gamma = M/N$. $\mathrm{SNR}_1 = \mathrm{SNR}_2 = 10dB$ and $\mathrm{SNR}_R = 13dB$.

matrices and the relay matrix were optimized alternately until convergence.

For the joint optimization, about 5-10 alternations were needed until convergence with the stopping criterion 10^{-3}. Both WMMSE and hybrid gradient methods were used (in separate runs) for finding the relay matrix at each alternation. The same result was obtained. This is because WMMSE and hybrid gradient yielded the same locally converged result.

In all cases, the power constraints at the sources and the relay were met with equality. The order of these curves is as expected. However, the relay only optimization yields the largest gain of the sum rate. This is because the number of antennas at the relay (M) is significantly larger the number of antennas at the users (N). Here, $M = 6$ and $N = 2$.

Fig. 7 illustrates the effect of the number of antennas on the maximum sum rate under the joint source-relay optimization. We define $\gamma = M/N$ in this figure. We see that both N and M have a significant effect on the sum rate. The effect of M on the sum rate becomes more significant as N becomes larger. This property makes Theorem 1 important in reducing the complexity.

VIII. Conclusion

In this paper, we have shown a study of the optimal design of the source and relay matrices for a non-regenerative two-way MIMO relay system. Although the two-way scheme is spectrally efficient for a half-duplex MIMO relay, the design of the optimal source matrices and the optimal relay matrix is not trivial especially when all nodes have multiple antennas. This study has shown an optimal structure for the relay matrix, which is useful for reducing the complexity of the optimal design when the relay has much more antennas than the users. For relay matrix optimization, we have developed the hybrid gradient method and the iterative WMMSE method, both of which have consistently yielded the same results. The hybrid gradient method is faster to converge, and the iterative WMMSE method is easier to implement. Both of these methods are much faster than the SDP based optimization method in [7]. The latter is only applicable to single-antenna users. We have established that for any given relay matrix, the optimal design of the source matrices (for uniformly weighted sum rate) follows the generalized water filling algorithm in [10]. We have demonstrated that by alternating the relay optimization and the source optimization, our joint source-and-relay optimization method can yield much improved system capacity especially when the number of antennas at all users is large.

Appendix A
Proof of Theorem 1

Without loss of generality, we can express $\mathbf{F}$ as

$$\mathbf{F} = \begin{bmatrix} \mathbf{V}_1 & \mathbf{V}_1^{\perp} \end{bmatrix} \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{U}_2^H \\ \mathbf{U}_2^{\perp H} \end{bmatrix} \tag{A.37}$$

where $\mathbf{V}_1^{\perp} \in \mathbb{C}^{M \times (M-2N)}$ is such that $\begin{bmatrix} \mathbf{V}_1 & \mathbf{V}_1^{\perp} \end{bmatrix}$ is unitary, $\mathbf{U}_2^{\perp} \in \mathbb{C}^{M \times (M-2N)}$ is such that $\begin{bmatrix} \mathbf{U}_2 & \mathbf{U}_2^{\perp} \end{bmatrix}$ is unitary. In other words, for any given $\mathbf{F}$, there is a unique set of $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{D}$, and vice versa. Clearly, $\mathbf{V}_1^H \mathbf{V}_1^{\perp} = \mathbf{0}$ and $\mathbf{U}_2^{\perp H} \mathbf{U}_2 = \mathbf{0}$. Thus, from (9) and (10), we have

$$\begin{bmatrix} \mathbf{H}_{R1} \\ \mathbf{H}_{R2} \end{bmatrix} \mathbf{V}_1^{\perp} = \begin{bmatrix} \mathbf{H}_{R1} \mathbf{V}_1^{\perp} \\ \mathbf{H}_{R2} \mathbf{V}_1^{\perp} \end{bmatrix} = \mathbf{0} \tag{A.38}$$

$$\mathbf{U}_2^{\perp H} \begin{bmatrix} \mathbf{H}_{1R} & \mathbf{H}_{2R} \end{bmatrix} = \begin{bmatrix} \mathbf{U}_2^{\perp H} \mathbf{H}_{1R} & \mathbf{U}_2^{\perp H} \mathbf{H}_{2R} \end{bmatrix} = \mathbf{0} \tag{A.39}$$

Using the above properties of $\mathbf{F}$ in (5), we have

$$\mathbf{X}_i = \mathbf{I} + (\mathbf{H}_{Ri}\mathbf{V}_1\mathbf{A}\mathbf{A}^H\mathbf{V}_1^H\mathbf{H}_{Ri}^H + \mathbf{H}_{Ri}\mathbf{V}_1\mathbf{B}\mathbf{B}^H\mathbf{V}_1^H\mathbf{H}_{Ri}^H + \mathbf{I})^{-1} \cdot \mathbf{H}_{Ri}\mathbf{V}_1\mathbf{A}\mathbf{U}_2^H\mathbf{H}_{\bar{i}R}\mathbf{R}_{x_{\bar{i}}}\mathbf{H}_{\bar{i}R}^H\mathbf{U}_2\mathbf{A}^H\mathbf{V}_1^H\mathbf{H}_{Ri}^H \tag{A.40}$$

We see that the matrices $\mathbf{C}$ and $\mathbf{D}$ do not affect $\mathbf{X}_i$ and hence the rates r_i. We now apply the properties of $\mathbf{F}$ to (2). We then have

$$\begin{aligned} p_R = &\operatorname{Tr}(\mathbf{A}\mathbf{U}_2^H\mathbf{H}_{1R}\mathbf{R}_{x_1}\mathbf{H}_{1R}^H\mathbf{U}_2\mathbf{A}^H) \\ &+ \operatorname{Tr}(\mathbf{C}\mathbf{U}_2^H\mathbf{H}_{1R}\mathbf{R}_{x_1}\mathbf{H}_{1R}^H\mathbf{U}_2\mathbf{C}^H) \\ &+ \operatorname{Tr}(\mathbf{A}\mathbf{U}_2^H\mathbf{H}_{2R}\mathbf{R}_{x_2}\mathbf{H}_{2R}^H\mathbf{U}_2\mathbf{A}^H) \\ &+ \operatorname{Tr}(\mathbf{C}\mathbf{U}_2^H\mathbf{H}_{2R}\mathbf{R}_{x_2}\mathbf{H}_{2R}^H\mathbf{U}_2\mathbf{C}^H) + \operatorname{Tr}(\mathbf{A}^H\mathbf{A}) \\ &+ \operatorname{Tr}(\mathbf{B}^H\mathbf{B}) + \operatorname{Tr}(\mathbf{C}^H\mathbf{C}) + \operatorname{Tr}(\mathbf{D}^H\mathbf{D}) \end{aligned} \tag{A.41}$$

which shows that $\mathbf{C}$ and $\mathbf{D}$ only increase the transmit power at the relay unless they are zero. Therefore, the optimal choice of $\mathbf{C}$ and $\mathbf{D}$ is simply $\mathbf{C} = 0$ and $\mathbf{D} = 0$. So, we can now write

$$\begin{aligned} p_R = &\operatorname{Tr}(\mathbf{A}\mathbf{U}_2^H\mathbf{H}_{1R}\mathbf{R}_{x_1}\mathbf{H}_{1R}^H\mathbf{U}_2\mathbf{A}^H) + \operatorname{Tr}(\mathbf{A}^H\mathbf{A}) \\ &+ \operatorname{Tr}(\mathbf{A}\mathbf{U}_2^H\mathbf{H}_{2R}\mathbf{R}_{x_2}\mathbf{H}_{2R}^H\mathbf{U}_2\mathbf{A}^H) + \operatorname{Tr}(\mathbf{B}^H\mathbf{B}) \end{aligned} \tag{A.42}$$

The remaining task is to prove that the optimal $\mathbf{B}$ is also zero. Define the positive semidefinite matrices

$$\mathbf{J}_i(\mathbf{X}) = \mathbf{H}_{Ri}\mathbf{V}_1\mathbf{X}\mathbf{X}^H\mathbf{V}_1^H\mathbf{H}_{Ri}^H$$

$$\mathbf{L}_i(\mathbf{X}) = \mathbf{H}_{Ri}\mathbf{V}_1\mathbf{X}\mathbf{U}_2^H\mathbf{H}_{\bar{i}R}\mathbf{R}_{x_{\bar{i}}}\mathbf{H}_{\bar{i}R}^H\mathbf{U}_2\mathbf{X}^H\mathbf{V}_1^H\mathbf{H}_{Ri}^H$$

Then, $r_i = \frac{1}{2}\log_2\det(\mathbf{X}_i)$ becomes

$$r_i = \frac{1}{2}\log_2\det\left(\mathbf{I} + (\mathbf{J}_i(\mathbf{A}) + \mathbf{J}_i(\mathbf{B}) + \mathbf{I})^{-1}\mathbf{L}_i(\mathbf{A})\right) \tag{A.43}$$

By Lemma 1 shown later, (A.43) implies

$$r_i \le \frac{1}{2}\log_2\det\left(\mathbf{I} + (\mathbf{J}_i(\mathbf{A}) + \mathbf{I})^{-1}\mathbf{L}_i(\mathbf{A})\right) \tag{A.44}$$

where the upper bound is achieved when $\mathbf{B} = 0$. We also know that $\mathbf{B}$ only increases p_R unless $\mathbf{B} = 0$. Therefore, the optimal $\mathbf{B}$ is zero.

Lemma 1: Given the conjugate symmetric matrices $\mathbf{X} \ge 0$, $\mathbf{Y} > 0$ and $\mathbf{Z} \ge 0$, it follows that

$$\det\left(\mathbf{I} + (\mathbf{Y}+\mathbf{Z})^{-1}\mathbf{X}\right) \le \det(\mathbf{I} + \mathbf{Y}^{-1}\mathbf{X})$$

Proof: We can write $\det\left(\mathbf{I} + (\mathbf{Y}+\mathbf{Z})^{-1}\mathbf{X}\right) = \det\left(\mathbf{I} + \mathbf{X}^{\frac{1}{2}}\mathbf{Y}^{-\frac{1}{2}}(\mathbf{I}+\mathbf{Y}^{-\frac{H}{2}}\mathbf{Z}\mathbf{Y}^{-\frac{1}{2}})^{-1}\mathbf{Y}^{-\frac{H}{2}}\mathbf{X}^{\frac{H}{2}}\right)$ and $\det(\mathbf{I}+\mathbf{Y}^{-1}\mathbf{X}) = \det\left(\mathbf{I} + \mathbf{X}^{\frac{1}{2}}\mathbf{Y}^{-\frac{1}{2}}\mathbf{Y}^{-\frac{H}{2}}\mathbf{X}^{\frac{H}{2}}\right)$. We know $\mathbf{I} + \mathbf{Y}^{-\frac{H}{2}}\mathbf{Z}\mathbf{Y}^{-\frac{1}{2}} \ge \mathbf{I}$, $(\mathbf{I}+\mathbf{Y}^{-\frac{H}{2}}\mathbf{Z}\mathbf{Y}^{-\frac{1}{2}})^{-1} \le \mathbf{I}$ and $\mathbf{X}^{\frac{1}{2}}\mathbf{Y}^{-\frac{1}{2}}(\mathbf{I}+\mathbf{Y}^{-\frac{H}{2}}\mathbf{Z}\mathbf{Y}^{-\frac{1}{2}})^{-1}\mathbf{Y}^{-\frac{H}{2}}\mathbf{X}^{\frac{H}{2}} \le \mathbf{X}^{\frac{1}{2}}\mathbf{Y}^{-\frac{1}{2}}\mathbf{Y}^{-\frac{H}{2}}\mathbf{X}^{\frac{H}{2}}$. Since $\mathbf{A}' \le \mathbf{B}'$ implies $\det(\mathbf{A}') \le \det(\mathbf{B}')$, therefore $\det\left(\mathbf{I}+(\mathbf{Y}+\mathbf{Z})^{-1}\mathbf{X}\right) \le \det(\mathbf{I}+\mathbf{Y}^{-1}\mathbf{X})$. ■

Appendix B
Zoomed SDP Algorithm

In [7], a special case of the two-way MIMO relay system was considered where the two users are each with a single antenna. For this case, there is no need for source optimization since $\mathbf{R}_{x_1}$ and $\mathbf{R}_{x_2}$ are simply P_1 and P_2 respectively. And to maximize the sum rate $r_1 + r_2$, the relay optimization problem is reduced to

$$\begin{aligned} \max_{\mathbf{F}} \quad & \frac{1}{2}\log_2\left(1 + \frac{|\mathbf{h}_{R1}^T\mathbf{F}\mathbf{h}_{2R}|^2P_2}{||\mathbf{F}^H\mathbf{h}_{R1}^*||^2+1}\right) \\ & + \frac{1}{2}\log_2\left(1 + \frac{|\mathbf{h}_{R2}^T\mathbf{F}\mathbf{h}_{1R}|^2P_1}{||\mathbf{F}^H\mathbf{h}_{R2}^*||^2+1}\right) \\ s.t. \quad & ||\mathbf{F}\mathbf{h}_{1R}||^2P_1 + ||\mathbf{F}\mathbf{h}_{2R}||^2P_2 + \operatorname{Tr}(\mathbf{F}\mathbf{F}^H) \le P_R \end{aligned} \tag{B.45}$$

Since each user has a single antenna, the channel matrices $\mathbf{H}_{1R}$, $\mathbf{H}_{R1}$, $\mathbf{H}_{2R}$ and $\mathbf{H}_{R2}$, are now reduced to channel vectors $\mathbf{h}_{1R}$, $\mathbf{h}_{R1}$, $\mathbf{h}_{2R}$ and $\mathbf{h}_{R2}$. This problem is still non-convex. In order to use a convex optimization program to solve this problem, the authors of [7] proposed to fix the ratio between r_1 and r_2. Specifically, they introduced the rate profile vector $\boldsymbol{\alpha} = [\alpha_1, \alpha_2]$ where $\alpha_1 = \frac{r_1}{r_{sum}}$, $\alpha_2 = 1 - \alpha_1 = \frac{r_2}{r_{sum}}$ and $r_{sum} = r_1 + r_2$. For each fixed $\boldsymbol{\alpha}$, the sum rate r_{sum} is maximized by solving the following problem

$$\begin{aligned} \max_{\mathbf{F}} \quad & r_{sum} \\ s.t. \quad & \frac{1}{2}\log_2\left(1 + \frac{|\mathbf{h}_{R1}^T\mathbf{F}\mathbf{h}_{2R}|^2P_2}{||\mathbf{F}^H\mathbf{h}_{R1}^*||^2+1}\right) \ge \alpha_1 r_{sum} \\ & \frac{1}{2}\log_2\left(1 + \frac{|\mathbf{h}_{R2}^T\mathbf{F}\mathbf{h}_{1R}|^2P_1}{||\mathbf{F}^H\mathbf{h}_{R2}^*||^2+1}\right) \ge \alpha_2 r_{sum} \\ & ||\mathbf{F}\mathbf{h}_{1R}||^2P_1 + ||\mathbf{F}\mathbf{h}_{2R}||^2P_2 + \operatorname{Tr}(\mathbf{F}\mathbf{F}^H) \le P_R \end{aligned} \tag{B.46}$$

which in turn can be solved by a general purpose SDP algorithm such as in CVX [16]. For more details, see [7] .

To find the maximum sum rate, we have to solve the additional problem $\max_{0\le\alpha_1\le 1} r_{sum}$. One method is the brute force search within $0 \le \alpha_1 \le 1$, which is too costly. To do it more efficiently, we formulate a zoomed SDP algorithm:

1) Let L be an even integer larger than or equal to 4. Choose a small number ϵ. Partition $[0, 1]$ into L uniform segments each of length $\delta = 1/L$, which yields $L-1$ interior uniform sample points $0 < \alpha_1^{(1)}, \cdots, \alpha_1^{(L-1)} < 1$.
2) Run the above SDP algorithm to compute the maximum of r_{sum} for each of the $L-1$ sample points.
3) Determine the best sample point: $\alpha_1^* = \arg\max_{\alpha_1^{(1)},\cdots,\alpha_1^{(L-1)}} r_{sum}$.
4) Partition $[\alpha_1^* - \delta, \alpha_1^* + \delta]$ into L uniform segments (i.e., "zooming"), which resets the $L-1$ uniform sample points $\alpha_1^{(1)}, \cdots, \alpha_1^{(L-1)}$. Also reset $\delta := 2\delta/L$
5) Go to Step 2 until $\delta < \epsilon$.

By choosing L to be an even integer, we ensure that each new set of $L-1$ sample points include α_1^* in Step 4. We also need $L = 2m$ with $m \ge 2$. Otherwise, if $L = 2$, δ would stay the same and the algorithm would not work. The zoomed SDP algorithm is used to compare with the hybrid gradient algorithm and the iterative WMMSE algorithm developed in this paper. We choose $\epsilon = 10^{-3}$ and $L = 4$ in simulations.

References

[1] S. Xu and Y. Hua, "Source-relay optimization for a two-way MIMO relay system," in *Proc. IEEE ICASSP*, Mar. 2010.

[2] B. Rankov and A. Wittneben, "Spectral efficient protocols for half-duplex fading relay channels," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 2, pp. 379–389, Feb. 2007.

[3] I. Hammerstrom, M. Kuhn, C. Esli, J. Zhao, A. Wittneben, and G. Bauch, "MIMO two-way relaying with transmit CSI at the relay," in *IEEE Signal Processing Advances in Wireless Communications*, June 2007, 5 pages.

[4] S. J. Kim, N. Devroye, P. Mitran, and V. Tarokh, "Achievable rate regions for bi-directional relaying," *IEEE Trans. Inf. Theory*, submitted, available online.

[5] R. Vaze and R. W. Heath Jr., "On the capacity and diversity-multiplexing tradeoff of the two-way relay channel," *IEEE Trans. Inf. Theory*, to appear, available online.

[6] N. Lee, H. J. Yang, and J. Chun, "Achievable sum-rate maximizing af relay beamforming scheme in two-way relay channels," *IEEE International Conference on Communications Workshops*, May 2008.

[7] R. Zhang, Y.-C. Liang, C. C. Chai, and S. Cui, "Optimal beamforming for two-way multi-antenna relay channel with analogue network coding," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 5, pp. 699–712, June 2009.

[8] T. Unger and A. Klein, "Duplex schemes in multiple antenna two-hop relaying," *EURASIP J. Advances in Signal Process.*, volume 2008, article ID 128592, 14 pages.

[9] Y. Rong, X. Tang, and Y. Hua, "A unified framework for optimizing linear non-regenerative multicarrier MIMO relay communication systems," *IEEE Trans. Signal Process.*, vol. 57, no. 12, pp. 4837–4852, Dec. 2009.

[10] Y. Yu and Y. Hua, "Power allocation for a MIMO relay system with multiple-antenna users," *IEEE Trans. Signal Process.*, vol. 58, no. 5, pp. 2823–2835, May 2010.

[11] S. S. Christensen, R. Agarwal, E. de Carvalho, and J. M. Cioffi, "Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 4792–4799, Dec. 2008.

[12] Z. Fang, Y. Hua, and J. C. Koshy, "Joint source and relay optimization for a non-regenerative MIMO relay," in *Proc. IEEE Workshop on Sensor Array Multichannel Signal Processing*, pp. 239–243, July 2006.

[13] Y. Rong and Y. Hua, "Optimality of diagonalization of multi-hop MIMO relays," *IEEE Trans. Wireless Commun.*, vol. 8, no. 12, pp. 6068–6077, Dec. 2009.

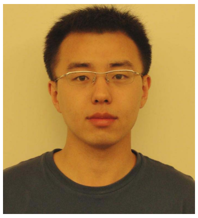
[14] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[15] J. Magnus and H. Neudecker, *Matrix Differential Calculus with Applications in Statistics and Econometrics*. John Wiley & Sons, 1999.

[16] M. Grant and S. Boyd, CVX: Matlab software for disciplined convex programming. Available: http://stanford.edu/ boyd/cvx, 2008.

[17] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1990.

[18] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, 1999.



Shengyang Xu received the B.E. degree from Dalian University of Technology, Dalian, Liaoning, China, and the M.S. degree from University of California, Riverside, both in electrical engineering, in 2007 and 2009, respectively. He is currently working towards the Ph.D. degree in electrical engineering at University of California, Riverside. His research interests include cooperative wireless communications and multiuser MIMO systems.

Yingbo Hua (S'86-M'88-SM'92-F'02) is a senior full professor at the University of California, Riverside, where he joined in 2001. Prior to that from 1990, he was a professor at University of Melbourne, Australia. He was a visiting professor with Hong Kong University of Science and Technology in 1999-2000. He consulted with Microsoft Research, WA, in 2000. He received Ph.D. and M.S. degrees from Syracuse University, Syracuse, NY, and B.S. degree from Nanjing Institute of Technology (Southeast University), Nanjing, China. He has published over 280 articles, with more than 2000 citations, in the fields of signal processing, sensing and wireless communications. He has edited two books, served as editor for five IEEE and EURASIP journals, as member on several IEEE SPS Technical Committees, and other technical committees for numerous international conferences. He was elevated to Fellow of IEEE in 2002.