

# Joint Precoding Optimization for Multiuser Multi-Antenna Relaying Downlinks Using Quadratic Programming

Wei Xu, *Member, IEEE*, Xiaodai Dong, *Senior Member, IEEE*, and Wu-Sheng Lu, *Fellow, IEEE*

**Abstract**—This paper studies the optimization problem for joint precoding design in a multi-antenna downlink channel using relaying. We formulate the joint source and relay precoding design by aiming at sum capacity maximization. Since this problem is in general nonconvex, we first convert this problem into standard convex quadratic programs, and then propose an iterative joint precoding optimization algorithm by utilizing efficient quadratic programming approaches. We observe that the iterative method always yields optimal precoding matrices which diagonalize the compound channel of the backward (source-to-relay) and the forward (relay-to-destination) links at high SNR regimes. Motivated by this observation, we further develop an efficient structured precoding design. Simulation results are presented to verify the effectiveness of our proposed precoding schemes.

**Index Terms**—Amplify-and-forward relay, multiple-input multiple-output (MIMO), joint precoding optimization, multiuser downlink channels, quadratic programming.

## I. INTRODUCTION

RELAYING technology has recently attracted a great deal of interests due to its ability to extend the coverage in future wireless networks, such as long-term evolution and IEEE 802.16m [2]–[4]. Meanwhile, multiple-input multiple-output (MIMO) is a well-known technology for current systems to significantly improve the spectral efficiency and link reliability [5]. For these reasons, a number of studies have focused on investigating different transmission technologies for relay-assisted multi-antenna systems [6]–[8].

In single-user MIMO systems, it is known that the optimal pre-processing strategy is the singular-value decomposition (SVD)-based precoding with a conventional water-filling power allocation scheme [5]. When a single relay is utilized between the source and the destination, however, the optimization of source precoding and relaying scheme becomes more

complex. From an information theoretic perspective, [6] has studied the capacity bounds of MIMO systems with a single relay. For practical implementation, studies in [7], [8] have investigated the optimal precoding at the relay in an amplify-and-forward relaying MIMO system, where it is shown that the optimal relaying matrix follows a similar structure as the SVD-based precoding in MIMO systems without relays. Further, with source precoding exploited, the authors in [9] have dealt with the joint optimization problem for both source and relay precoding.

When multiple antennas are deployed at the transmitter, however, multiple users can be scheduled at a time for simultaneous transmissions. Therefore, the multiuser MIMO downlink system also emerges as one of the most essential scenarios among various wireless applications. In [10], the information theoretic limit for multiuser MIMO systems has been investigated. Dirty paper coding (DPC) is proven to be an optimal transmission strategy for a MIMO downlink channel to achieve its full channel capacity. However, the implementation complexity of DPC is prohibitively high for most wireless applications. Therefore, a number of less-complex linear precoding technologies are developed in [11]–[14].

Recently, increasing attention has been given to the issue of using relay stations in MIMO networks to deal with remote users, or equivalently to enlarge the coverage, [15]–[18]. In [15], the precoding design for multiple access channels (MACs) using regenerative bidirectional relaying has been addressed. Concerning a single-direction relay-assisted MAC, the outage performance has been analyzed in [16]. For broadcast channels (BCs), it is still hard to obtain the optimal precoding for a MIMO relay. Several suboptimal methods are developed in [17], [18]. More specifically, sum capacity bounds are derived in [17] by exploiting nonlinear precoding at the source and linear processing at the relay. For implementation efficiency, [18] has studied linear joint precoding strategies and developed an iterative joint precoding algorithm by aiming at transmit power minimization under quality-of-service (QoS) requirements.

In this paper, we consider the joint optimization of linear pre-processing at both the source and relay. Different from [18], we optimize the joint precoding strategy by maximizing the achievable sum capacity under fixed transmit power constraints. Since the direct use of the sum capacity as an objective function leads to a nonconvex optimization problem, we resort to solving this problem by converting it to standard (convex) quadratic programs in an iterative manner. Although the proposed method generally converges to some locally optimal solutions due to the nonconvexity of the original optimization problem, numerical results demonstrate that the

Paper approved by I. Lee, the Editor for Wireless Communication Theory of the IEEE Communications Society. Manuscript received November 21, 2009; revised June 24, 2010 and October 5, 2010.

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada through grant 349722-07, the International Cooperation Project in China under grant 2010DFB10410, the National Science & Technology Projects of China under grant 2010ZX03002-012, and the Research Fund of the National Mobile Communications Research Laboratory, Southeast University (No. 2011B01). W.-S. Lu was supported by a Discovery Grant from NSERC of Canada. Part of this manuscript was presented at the IEEE International Conference on Communications (ICC), Cape Town, South Africa, May 2010 [1].

W. Xu is with the National Mobile Communications Research Lab, Southeast University, Nanjing 210096, China (e-mail: wxu@seu.edu.cn).

X. Dong and W.-S. Lu are with the Department of Electrical and Computer Engineering, University of Victoria, Victoria B.C. V8W 3P6, Canada (e-mail: {xdong, wslu}@ece.uvic.ca).

Color versions of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TCOMM.2011.041111.090713

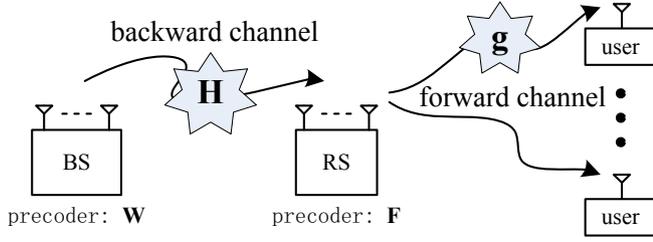


Fig. 1. System model of a relay-assisted multi-antenna downlink channel.

method is able to produce significant performance gains. Moreover, in order to further provide some tradeoffs between the complexity and performance, we also propose an efficient source and relay precoding structure with adaptive power allocation.

*Notation:* Symbols  $\mathbf{A}^T$ ,  $\mathbf{A}^H$  and  $\text{tr}(\mathbf{A})$  denote the transpose, conjugate transpose, and trace of a matrix  $\mathbf{A}$ , respectively.  $\text{vec}(\mathbf{A})$  returns the vectorization of  $\mathbf{A}$ .  $\|\mathbf{a}\|$  represents the Euclidean norm of a complex vector  $\mathbf{a}$ .  $\mathbb{R}^{x \times y}$  denotes the space of  $x \times y$  matrices with real-valued elements, while  $\mathbb{C}^{x \times y}$  represents the complex-valued matrix space. Operators  $\Re\{\cdot\}$  and  $\Im\{\cdot\}$  return the real and imaginary part of a complex input, respectively.  $\mathbb{E}\{\cdot\}$  represents the expectation.

## II. SYSTEM MODEL AND PROBLEM FORMULATION

Consider a multiuser downlink channel with an  $N_s$ -antenna base station (BS) serving  $L$  single-antenna users through an  $N_r$ -antenna relay station (RS). A two-hop protocol is employed by the RS and a half-duplex scheme is utilized. The direct links between the source and remote users are neglected due to large path loss and severe shadowing effects. The system structure of the relay-assisted multi-antenna downlink channel is illustrated in Fig. 1. Since we focus on the problem of precoding design, like in [15]–[18], user scheduling process is not considered in this study, i.e.,  $L = \min(N_s, N_r)$ .

Let  $\mathbf{s} \in \mathbb{C}^{L \times 1}$  be the transmit symbol vector at the BS. Before transmitting  $\mathbf{s}$  to the RS, the BS first pre-processes  $\mathbf{s}$  by an  $N_s \times L$  precoding matrix  $\mathbf{W}$ , thus the received symbol vector at the RS is

$$\mathbf{r} = \sqrt{\rho_1} \mathbf{H} \mathbf{W} \mathbf{s} + \mathbf{n} \quad (1)$$

where  $\mathbf{H} \in \mathbb{C}^{N_r \times N_s}$  is the backward channel matrix from BS to RS,  $\mathbf{n}$  is the complex Gaussian noise with zero-mean and unit variance, and  $\rho_1$  is a scaling factor to make the total power constraint  $P_1$  satisfied. We normalize the input symbols  $\mathbf{s}$ , i.e.,  $\mathbb{E}\{\mathbf{s} \mathbf{s}^H\} = \mathbf{I}_L$ , hence the BS transmit power constraint is

$$\text{tr}(\mathbf{W} \mathbf{W}^H) = P_1 / \rho_1. \quad (2)$$

Assume that the backward channel follows the Rayleigh distribution, and the value of  $P_1$  has taken into account the path loss effect of the channel. Thus, the entries of the channel matrix  $\mathbf{H}$  are independent complex Gaussian random variables with zero-mean and unit variance.

After receiving the symbols, RS pre-processes  $\mathbf{r}$  by an  $N_r \times N_r$  precoding matrix  $\mathbf{F}$ , and then broadcasts the precoded symbols to distributed users. The RS forwards the symbol

$$\mathbf{s}_r = \sqrt{\rho_1 \rho_2} \mathbf{F} \mathbf{H} \mathbf{W} \mathbf{s} + \sqrt{\rho_2} \mathbf{F} \mathbf{n} \quad (3)$$

where  $\rho_2$  is a scaling factor at the relay to guarantee the relay power constraint  $P_2$ :

$$\rho_1 \text{tr}(\mathbf{F} \mathbf{H} \mathbf{W} \mathbf{W}^H \mathbf{H}^H \mathbf{F}^H) + \text{tr}(\mathbf{F} \mathbf{F}^H) = P_2 / \rho_2. \quad (4)$$

Under these circumstances, we can express the received symbol at user  $k$  as

$$y_k = \sqrt{\rho_1 \rho_2} \mathbf{g}_k^H \mathbf{F} \mathbf{H} \mathbf{W} \mathbf{s} + \sqrt{\rho_2} \mathbf{g}_k^H \mathbf{F} \mathbf{n} + z_k \quad (5)$$

where  $\mathbf{g}_k^H$  is the  $1 \times N_r$  channel vector between the RS and the  $k$ th user terminal, and  $z_k$  is the complex Gaussian noise at the  $k$ th user terminal with zero mean and unit variance. Without loss of generality, we define the  $k$ th entry  $s_k$  in the transmit vector  $\mathbf{s}$  as the desired information for user  $k$ .

By representing the BS precoding matrix as  $\mathbf{W} = [\mathbf{w}_1 \cdots \mathbf{w}_L]$  where  $\mathbf{w}_k$  is the beamformer for data  $s_k$  to user  $k$ , the SINR at user  $k$  is found to be

$$\gamma_k = \frac{|\mathbf{g}_k^H \mathbf{F} \mathbf{H} \mathbf{w}_k|^2}{\sum_{j \neq k} |\mathbf{g}_k^H \mathbf{F} \mathbf{H} \mathbf{w}_j|^2 + \frac{1}{\rho_1} \|\mathbf{g}_k^H \mathbf{F}\|^2 + \frac{1}{\rho_1 \rho_2}}. \quad (6)$$

Thus, aiming at sum capacity maximization, the joint precoding optimization problem is formulated by

$$\underset{\mathbf{W}, \mathbf{F}}{\text{maximize}} \quad \mathcal{R} = \frac{1}{2} \sum_{k=1}^L \log(1 + \gamma_k) \quad (7a)$$

subject to:

$$\gamma_k = \frac{|\mathbf{g}_k^H \mathbf{F} \mathbf{H} \mathbf{w}_k|^2}{\sum_{j \neq k} |\mathbf{g}_k^H \mathbf{F} \mathbf{H} \mathbf{w}_j|^2 + \frac{1}{\rho_1} \|\mathbf{g}_k^H \mathbf{F}\|^2 + \frac{1}{\rho_1 \rho_2}} \quad (7b)$$

$$k = 1, 2, \dots, L$$

$$\text{tr}(\mathbf{W} \mathbf{W}^H) = \frac{P_1}{\rho_1} \quad (7c)$$

$$\rho_1 \text{tr}(\mathbf{F} \mathbf{H} \mathbf{W} \mathbf{W}^H \mathbf{H}^H \mathbf{F}^H) + \text{tr}(\mathbf{F} \mathbf{F}^H) = \frac{P_2}{\rho_2} \quad (7d)$$

where factor 1/2 in (7a) results from the fact that data is transmitted over two time-slots. It is not hard to verify that this problem is nonconvex and thus difficult to find the globally optimal solution. In the following section, we will present an iterative algorithm which exploits efficient quadratic programming approaches to obtain locally optimal solutions.

## III. JOINT PRECODING OPTIMIZATION USING QUADRATIC PROGRAMMING

There exist several efficient methods to solve general optimization problems, e.g. gradient descent method for unconstrained problems and interior-point methods for constrained problems [20]–[22]. Especially for convex constrained problems, globally optimal solutions can be obtained using convex programming algorithms. Since our original problem (7) is nonconvex, we resort to a sequential programming approach [20] to deal with its non-convexity. In brief terms, the method we propose iteratively updates a current iterate within a *small* vicinity centered at it and, in doing so, the principles in calculus allow to replace the highly nonlinear and nonconvex functions involved in problem (7) by their low-order counterparts with reasonable accuracy, leading to a valid low-order convex sub-problem. The solution of this sub-problem yields a new iterate and the process repeats until

the iterate sequence so generated converges to a local solution of problem (7).

Because the constraints in (7c)–(7d) are all *equalities*, their approximations have to be of first-order (i.e., linear) in order for the feasible region of the sub-problem to be convex. Meanwhile, for the objective function in (7a), we implement quadratic approximation and obtain the quadratic term with a positive semi-definite quadratic coefficient matrix. It is well documented (e.g., p. 24 of [22]) that a general function expanded about a local solution is approximated well by a quadratic function, thus methods based on quadratic models should have rapid rate of convergence; and even if the present iterate is remote from a local solution, quadratic information is an effective way of predicting direction along which substantial progress can be made. In this way, the entire problem is transformed to a convex quadratic program which can be secured by fast convex programming algorithms [20]–[23].

Without loss of generality, we first consider the  $(n+1)$ th iteration of the proposed algorithm, where the precoding matrices obtained at the  $n$ th iteration are denoted by  $\mathbf{W}^{(n)}$  and  $\mathbf{F}^{(n)}$ . In this iteration, the problem is to find the optimal updates  $\Delta_w^{(n+1)}$  and  $\Delta_f^{(n+1)}$  for the precoding matrices, namely

$$\mathbf{W}^{(n+1)} = \mathbf{W}^{(n)} + \Delta_w^{(n+1)} \quad \text{and} \quad \mathbf{F}^{(n+1)} = \mathbf{F}^{(n)} + \Delta_f^{(n+1)} \quad (8)$$

so as to maximize the sum capacity. From (7) and (8), it follows that the optimization problem in the  $(n+1)$ th iteration becomes

$$\underset{\Delta_w^{(n+1)}, \Delta_f^{(n+1)}}{\text{maximize}} \quad \sum_{k=1}^L \log \left( 1 + \gamma_k^{(n+1)} \right) \quad (9a)$$

subject to:

$$\gamma_k^{(n+1)} = \frac{|\mathbf{g}_k^H \mathbf{F}^{(n+1)} \mathbf{H} \mathbf{w}_k^{(n+1)}|^2}{\sum_{j \neq k} |\mathbf{g}_k^H \mathbf{F}^{(n+1)} \mathbf{H} \mathbf{w}_j^{(n+1)}|^2 + \frac{1}{\rho_1} \|\mathbf{g}_k^H \mathbf{F}^{(n+1)}\|^2 + \frac{1}{\rho_1 \rho_2}} \quad (9b)$$

$$\text{tr} \left( \mathbf{W}^{(n+1)} \mathbf{W}^{(n+1)H} \right) = \frac{P_1}{\rho_1} \quad (9c)$$

$$\rho_1 \text{tr} \left( \mathbf{F}^{(n+1)} \mathbf{H} \mathbf{W}^{(n+1)} \mathbf{W}^{(n+1)H} \mathbf{H}^H \mathbf{F}^{(n+1)H} \right) + \text{tr} \left( \mathbf{F}^{(n+1)} \mathbf{F}^{(n+1)H} \right) = \frac{P_2}{\rho_2}. \quad (9d)$$

However, the problem in (9) remains nonconvex and highly nonlinear. As mentioned above, the approach we take in this paper is to develop an iterative method, in that each iteration examines the problem in a small vicinity of the current iterate and makes a move for increasing the objective function while approximately satisfying the constraints. Because the functions involved in (9) are all continuously differentiable, the principle of calculus allows to use linear functions to approximate the equality constraints and a quadratic function to approximate the objective function as long as one is confined within a local region around the current iterate. Using this approach, the logarithmic objective function ensures that the approximation gives a convex quadratic programming problem. We stress that

the smallness constraints on  $\Delta_w^{(n+1)}$  and  $\Delta_f^{(n+1)}$  are critical for the proposed algorithm to converge. Following the details provided in Appendix A, we can convert the optimization problem (9) into the following quadratic program:

$$\underset{\mathbf{x} \in \mathbb{R}^{2(N_s L + N_k^2) \times 1}}{\text{minimize}} \quad \frac{1}{2} \mathbf{x}^T \left( \sum_{k=1}^L \frac{\mathbf{P}_k \mathbf{P}_k^T}{\left(1 + \gamma_k^{(n)}\right)^2} \right) \mathbf{x} - \left( \sum_{k=1}^L \frac{\mathbf{P}_k^T}{1 + \gamma_k^{(n)}} \right) \mathbf{x} \quad (10a)$$

$$\text{subject to:} \quad \mathbf{A}^T \mathbf{x} = \mathbf{e} \quad (10b)$$

$$-\tau \mathbf{1} \leq \mathbf{x} \leq \tau \mathbf{1} \quad (10c)$$

where  $\mathbf{p}_k$  is given by (25), and  $\mathbf{A}$  and  $\mathbf{e}$  are from (31) and (32), respectively. The desired variable  $\mathbf{x}$  is a stacked vector of  $\Delta_w^{(n+1)}$  and  $\Delta_f^{(n+1)}$ , defined in (15). Note that the bound constraint with a small predetermined  $\tau > 0$  is imposed to guarantee the assumption of a small  $\mathbf{x}$ , i.e., small precoding update steps. Moreover, it is not hard to verify that the coefficient matrix of the quadratic term in the objective function of (10) is positive semidefinite. Consequently, the quadratic problem (10) is a convex problem, and we can solve this problem to obtain its globally optimal solution efficiently by using numerically stable optimization methods [21]. In our simulations, we directly use `quadprog` in MATLAB to solve the problem in (10). A step-by-step description of the proposed iterative joint precoding optimization method is summarized in Algorithm 1.

---

#### Algorithm 1 : Iterative Joint Precoding Optimization

---

- 1: Select initial  $\mathbf{W}^{(0)}$  and  $\mathbf{F}^{(0)}$ , a small  $\tau > 0$ , and a convergence tolerance  $\epsilon$ . Set  $n = 0$ .
  - 2: Solve problem (10), and obtain  $\Delta_w^{(n+1)}$  and  $\Delta_f^{(n+1)}$  by reshaping  $\mathbf{x}$  according to (15).
  - 3: Find  $\alpha^*$ , a value of  $\alpha \in [0, 1]$  that maximizes the objective function in (7) with  $\mathbf{W} = \mathbf{W}^{(n)} + \alpha \Delta_w^{(n+1)}$  and  $\mathbf{F} = \mathbf{F}^{(n)} + \alpha \Delta_f^{(n+1)}$ .
  - 4: Update  $\mathbf{W}^{(n+1)} = \mathbf{W}^{(n)} + \alpha^* \Delta_w^{(n+1)}$  and  $\mathbf{F}^{(n+1)} = \mathbf{F}^{(n)} + \alpha^* \Delta_f^{(n+1)}$ .
  - 5: Set  $n = n + 1$  and repeat from *Step 2* until  $\|\alpha^* \mathbf{x}\|^2 \leq \epsilon$ .
  - 6: Scale the solution of  $\mathbf{W}$  and  $\mathbf{F}$  according to the constraints in (7c) and (7d), respectively.
- 

We now conclude this section with several remarks.

- *Step 3* in Algorithm 1 is known as line search as it is merely a one-dimensional optimization procedure [20]. The inclusion of this line search step is found crucial in order to ensure the algorithm's convergence. The numerical value of the optimal  $\alpha^*$  in the large part of the optimization procedure remains to be practically unity until the algorithm gets close to the solution point with  $\alpha^*$  quickly approaching to zero.
- We apply linear approximations to obtain the affine transmit power constraints in (10b). Although we have restricted that  $\mathbf{x}$  is very small, the original power constraints in (7c) and (7d) can only be satisfied within an acceptable

tolerance. *Step 6* is used to guarantee the original transmit power constraints are precisely satisfied.

- Since the original problem in (7) is nonconvex, it is generally difficult to obtain the globally optimum solution by the proposed algorithm. Although this implies that the proposed iterative optimization algorithm converges to some locally optimal solutions, numerical results still demonstrate significant performance gains. We will come to this point in Section V with numerical evidences.

#### IV. PARALLEL TRANSMISSIONS WITH ADAPTIVE POWER ALLOCATION

In the above section, we have proposed an iterative algorithm which solves the joint precoding problem in (10). The performance of Algorithm 1 depends on selected initial points. Fortunately, using different initial points, we have observed that the obtained precoding matrices always diagonalize the compound channel at high SNRs. This motivates us to further develop a simplified precoding design strategy by compound channel diagonalization as alternatives to balance the complexity and performance.

The efficient scheme implements adaptive power allocation under a fixed precoding structure given below. Start with the SVD of  $\mathbf{H}$  as  $\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$ , where  $\mathbf{\Sigma}$  is a diagonal matrix with singular values arranged in a non-increasing order. Since it is assumed that  $L = \min(N_s, N_r)$ , we let  $\overline{\mathbf{V}} \in \mathbb{C}^{N_s \times L}$  and  $\overline{\mathbf{U}} \in \mathbb{C}^{N_r \times L}$  represent the first  $L$  columns of  $\mathbf{V}$  and  $\mathbf{U}$ , respectively. The proposed structures of the precoding matrices are then specified as

$$\mathbf{W} = \overline{\mathbf{V}}\mathbf{\Xi}_w \quad \text{and} \quad \mathbf{F} = \widehat{\mathbf{F}}\mathbf{\Xi}_f\overline{\mathbf{U}}^H \quad (11)$$

where we define  $\widehat{\mathbf{F}} = [\mathbf{f}_1 \cdots \mathbf{f}_L]$  with  $\mathbf{f}_k$  being the normalized  $k$ th column of the matrix  $\mathbf{G}^H (\mathbf{G}\mathbf{G}^H)^{-1}$ , and  $\mathbf{\Xi}_w$  and  $\mathbf{\Xi}_f$  are diagonal matrices representing the power allocation patterns at the source and the relay, respectively. By substituting the proposed precoders into (5), the received symbols at users become

$$\mathbf{y} = \sqrt{\rho_1\rho_2}\mathbf{\Lambda}\mathbf{\Xi}_f\overline{\mathbf{\Sigma}}\mathbf{\Xi}_w\mathbf{s} + \sqrt{\rho_2}\mathbf{\Lambda}\mathbf{\Xi}_f\tilde{\mathbf{n}} + \mathbf{z} \quad (12)$$

where  $\mathbf{y} = [y_1 \cdots y_L]^T$  is the concatenated received symbols,  $\mathbf{\Lambda} = \mathbf{G}\widehat{\mathbf{F}}$  is a diagonal matrix,  $\overline{\mathbf{\Sigma}}$  is the primary  $L \times L$  diagonal block of  $\mathbf{\Sigma}$ , and  $\tilde{\mathbf{n}} \in \mathbb{C}^{L \times 1}$  is the equivalent Gaussian noise. Define

$$\begin{aligned} \mathbf{\Lambda}^2 &= \text{diag}(\lambda_k), \quad \overline{\mathbf{\Sigma}}^2 = \text{diag}(\sigma_k), \quad \mathbf{\Xi}_f^2 = \text{diag}(f_k), \\ \mathbf{\Xi}_w^2 &= \text{diag}(w_k), \quad k = \{1, \cdots, L\} \end{aligned} \quad (13)$$

where  $\lambda_k$ ,  $\sigma_k$ ,  $f_k$ , and  $w_k$  are all no-negative real values. Then by substituting (11)–(13) into (7), the problem can be simplified as

$$\underset{f_k \geq 0, w_k \geq 0}{\text{maximize}} \quad \frac{1}{2} \sum_{k=1}^L \log(1 + \gamma_k) \quad (14a)$$

$$\text{subject to:} \quad \gamma_k = \rho_1 \lambda_k \sigma_k f_k w_k / (\lambda_k f_k + 1 / \rho_2), \quad k = 1, \cdots, L \quad (14b)$$

$$\sum_{k=1}^L w_k = \frac{P_1}{\rho_1} \quad (14c)$$

$$\sum_{k=1}^L f_k (\rho_1 \sigma_k w_k + 1) = \frac{P_2}{\rho_2}. \quad (14d)$$

It is not hard to see that the above power allocation problem is nonconvex. Reference [19] has dealt with a similar power allocation problem across different subcarriers in a single-user MIMO-OFDM relay system, while it utilizes uniform power allocation across different eigenmodes. The optimization problem (14) is different from [19] in that we consider the power allocation across parallel eigenmode transmission of the spatial channel. This yields a more complex formulation for the transmit power constraint at the relay, which does not allow us to solve this problem by directly utilizing the method in [19].

In order to solve (14), a sequential programming approach as Algorithm 1 is utilized by following similar techniques described in the above section. We iteratively optimize the unknown variables in (14) within a small vicinity in each step and the optimization problem in each step is formulated as a convex quadratic program. Due to the page limitation, the details of the algorithm is not elaborated here. We stress that the adaptive power allocation algorithm for solving (14) is more efficient than Algorithm 1 although both methods have similar procedures. This is because we here deal with an unknown vector of size  $2L$ , which is much smaller than that in Algorithm 1.

#### V. SIMULATION RESULTS

This section presents simulation results of the proposed methods. We average the sum capacity over 2000 random channel realizations. For comparison, we implement several different precoding strategies as follows:

- 1) Iterative joint optimization method summarized in Algorithm 1. We implemented the iterative joint precoding algorithm with  $\tau = 0.05$  and  $\epsilon = 10^{-5}$ . Note that we used the closed-form precoding design in 3) with permutation matrix  $\mathbf{\Pi} = \mathbf{I}$  as the initial point.
- 2) Adaptive power allocation scheme in Section IV.
- 3) Parallel transmission with eigenmode matching in [1].
- 4) SVD-MF (matched-filter) based design:  $\mathbf{W} = \overline{\mathbf{V}}$ , and  $\mathbf{F} = \widehat{\mathbf{F}}\overline{\mathbf{U}}^H$  where  $\widehat{\mathbf{F}} = [\mathbf{g}_1/\|\mathbf{g}_1\| \cdots \mathbf{g}_L/\|\mathbf{g}_L\|]$ . We set  $\rho_1 = \frac{P_1}{L}$  and chose the value of  $\rho_2$  by satisfying the power constraint in (7d).
- 5) Identity-based precoding:  $\mathbf{W}$  was an  $N_s \times L$  matrix with 1 as its diagonal elements and zeros elsewhere, and  $\mathbf{F} = \mathbf{I}_{N_r}$ .  $\rho_1$  and  $\rho_2$  were determined by reinforcing the power constraints (7c)(7d).
- 6) Channel inversion based precoding [18] with equal power.
- 7) Nonlinear ZF-DPC relaying [17].

Figures 2–4 compare the ergodic sum capacity obtained by the above precoding strategies as a function of channel average SNRs under different antenna configurations. From the results, we observe a noticeable performance gain achieved by using properly designed precoding methods. The joint precoding optimization scheme achieves the highest sum capacity at a cost of relatively high computational complexity. Its achievable sum capacity can also serve as an upper bound

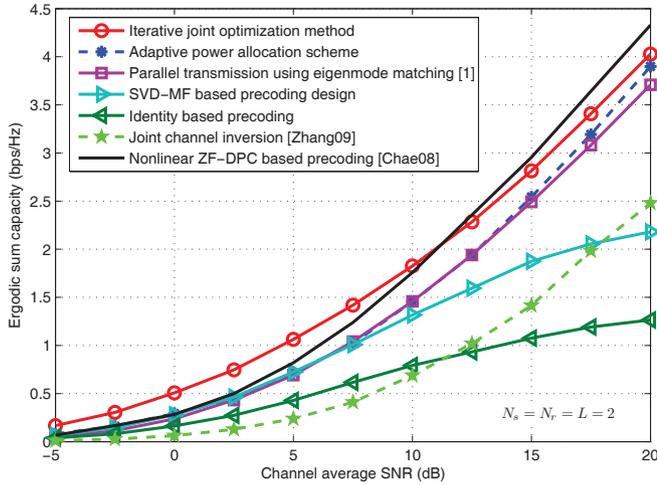


Fig. 2. Comparison of ergodic sum capacity achieved by different precoding schemes as a function of  $P_1 = P_2 = \text{SNR}$  in dB with  $N_s = N_r = L = 2$ .

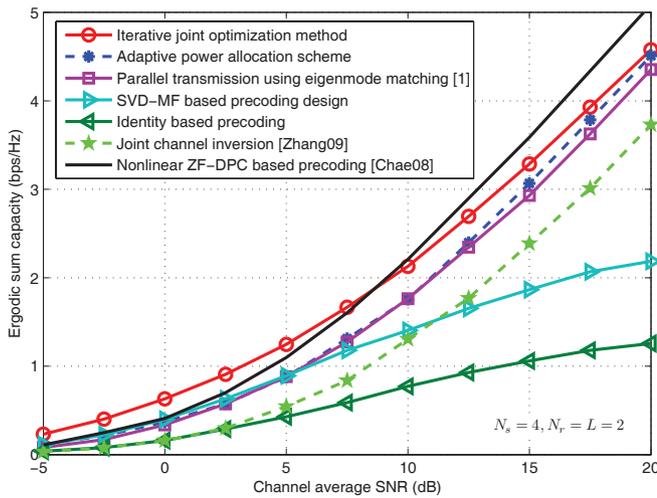


Fig. 3. Comparison of ergodic sum capacity achieved by different precoding schemes as a function of  $P_1 = P_2 = \text{SNR}$  in dB with  $N_s = 4$  and  $N_r = L = 2$ .

for the performance of other efficiently-designed precoding approaches. Moreover, by comparing the results in the three figures, it shows that the iterative joint optimization scheme provides more performance gains for systems equipped with more antennas and serving more users. This is because a larger number of antennas provides more degrees of freedom for precoding optimization, hence resulting in a larger potential performance gain.

For fair comparison, we here discuss the complexity of these algorithms in a quantitative manner. The computational complexity of the eigenmode-matching method [1] is mainly determined by the required SVD operations which are of order  $\mathcal{O}(N_s^2 N_r)$  [25]. While for the proposed iterative methods, the complexity depends on both the amount of computation in each iteration and the number of iterations. Since both iterative algorithms are based on quadratic programming, the complexity of each iteration is approximated by  $\mathcal{O}(M^{3.5})$  where  $M$

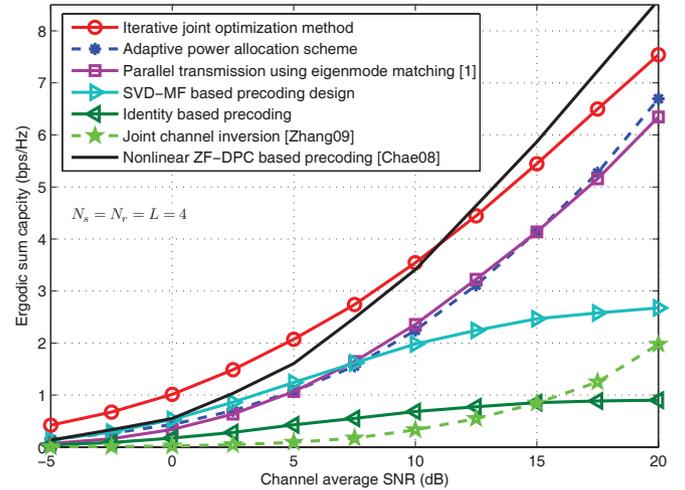


Fig. 4. Comparison of ergodic sum capacity achieved by different precoding schemes as a function of  $P_1 = P_2 = \text{SNR}$  in dB with  $N_s = N_r = L = 4$ .

is the size of the quadratic program problem<sup>1</sup>. Accordingly for any fixed iteration numbers, the computational complexity of Algorithm 1 is  $\mathcal{O}((N_s L)^{3.5})$ . Concerning the power allocation algorithm in Section IV, because it requires SVD operations for calculating the structured precoding matrices, its complexity is approximated by  $\mathcal{O}(L^{3.5} + N_s^2 N_r)$ . It shows that the proposed power allocation scheme generally requires less computational complexity than Algorithm 1 because it deals with a smaller number of variables. Note that some numerical examples will be shown later to evaluate the effect of iteration numbers on their achieved sum capacity. Comparing the four non-iterative schemes listed as 3)–6), the identity-based scheme is obviously the most computationally efficient one. The other three methods have comparable complexity of  $\mathcal{O}(N_s^2 N_r)$  because they all require SVD operations which dominate the overall complexity.

Since the proposed algorithms are based on iterative mechanisms, it is also necessary to illustrate their convergence by some numerical examples. Taking Algorithm 1 for instance, we demonstrate in Fig. 5 the convergence of the proposed iterative joint precoding design algorithm. It shows the relative achieved ergodic sum rate of the algorithm as a function of the number of iterations. From this figure, we find that generally 5 iterations can achieve a significant part of the overall performance gain by the joint precoding optimization method. As in the systems with more antennas and more users, the algorithm needs more iterations while it of course generates more relative performance gains. Moreover, for some practical scenarios with temporal channel correlations, the convergence of the algorithm can be accelerated by using the precoding matrices obtained in the former time-slot as initializations.

The results above assume that both the backward and the forward channels have the same transmit power constraints, i.e.,  $P_1 = P_2$ . Figure 6 plots the sum rate versus the transmit power constraint at the BS with fixed power  $P_2 = 10$  dB

<sup>1</sup>From [21], for computing each step of a quadratic program with size  $M$  by using interior-point methods, the complexity is of  $\mathcal{O}(M^3)$ , and the number of iteration steps for solving a quadratic program is upper bounded by  $\mathcal{O}(\sqrt{M})$ .

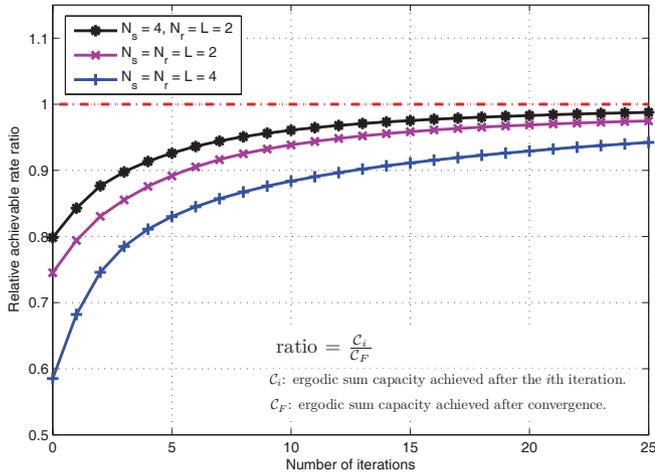


Fig. 5. Relative capacity by iterative joint precoding optimization algorithm versus the number of iterations with  $P_1 = P_2 = \text{SNR} = 10$  dB.

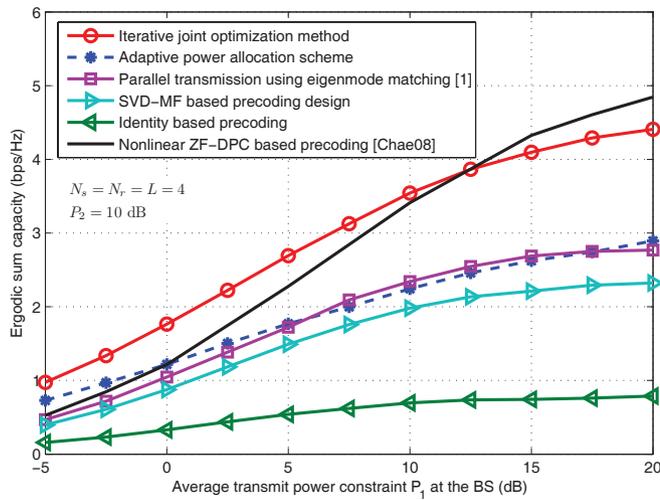


Fig. 6. Comparison of ergodic sum capacity achieved by different precoding schemes as a function of  $P_1$  in dB with  $N_s = N_r = L = 4$  and fixed  $P_2 = 10$  dB.

at the relay. Both the SVD-MF precoding and the parallel transmission strategy are found to produce similar sum rate performance. Moreover, the channel inversion strategy is found to outperform the SVD-MF based one in the forward broadcast channel at high (relay) SNRs. This is similar to that in MIMO systems without relay.

Although this study focuses on linear precoding design, for reference, we also provide the sum rate performance derived in [17] which is obtained by utilizing a ZF-DPC relaying scheme. ZF-DPC is a nonlinear precoding method developed in [26] and it achieves a noticeable gain with respect to conventional linear precoding schemes like ZFBF. In general, when system SNR grows larger, the sum rate gain by ZF-DPC becomes more pronounced [26]. From Figs. 2-4,6, the performance in [17] is found better than other linear precoding schemes at high SNRs due to the advantage of ZF-DPC. Nevertheless, since the implementation of DPC is prohibitively high and practical design of dirty paper codes is still unknown [24], the performance in [17] can be treated as an upper bound to linear schemes at this case. For low to moderate SNRs,

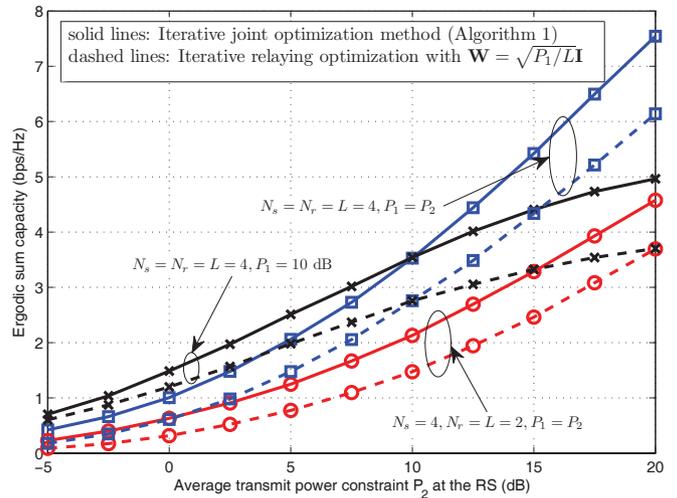


Fig. 7. Comparison of the proposed iterative joint optimization method and its simplified iterative relaying optimization method with identity based BS precoding.

the proposed joint precoding schemes outperforms the ZF-DPC relaying method. This is because our proposed scheme jointly optimizes the source and relay precoding without any structure constraint while the performance in [17] is calculated by using precoding schemes with fixed structure separately predetermined at both source and relay. Even though ZF-DPC generates better performance for the relay-to-user links, the merit of joint precoding optimization as proposed in our scheme dominates the performance gain of an entire two-hop system at low to moderate SNRs.

Finally, in order to explicitly show the performance gain achieved by joint precoding optimization, Fig. 7 compares the proposed joint optimization method with a simplified method with only relay precoding optimization. To isolate the effect of source precoding optimization, we simplify the joint precoding optimization algorithm by letting  $\mathbf{W} = \sqrt{P_1/L}\mathbf{I}$ .<sup>2</sup> With this variation, this simplified method follows the same procedure as Algorithm 1 but solves a different quadratic problem simplified from (10). The simplified quadratic problem is obtained by setting the first  $2N_sL$  elements of  $\mathbf{x}$  in (10) to zero and defining the rest of  $\mathbf{x}$  as a new variable set for optimization. By comparing the sum rate results of both schemes in Fig. 7, we find that joint optimization provides a noticeable contribution to the performance gain under different system configurations.

## VI. CONCLUSION

We have proposed an iterative algorithm for jointly optimizing the precoding at the source and relay in a multiuser downlink channel. The precoding matrices generated by the joint optimization algorithm always diagonalize the compound channel at high SNRs. Based on this observation, a parallel transmission structure with adaptive power allocation has also been proposed. The joint optimization algorithm is found to be able to provide the best performance at the cost of increased complexity.

<sup>2</sup>For the case with  $N_s \neq L$ , the source precoding  $\mathbf{W}$  is not a square matrix. We construct  $\mathbf{W}$  with a scaled identity matrix at the top block and zeros elsewhere.

## APPENDIX A

## QUADRATIC PROGRAM REFORMULATION OF (9)

A standard quadratic program is a problem minimizing a quadratic objective under a set of affine inequality and equality constraints [20], [21]. Under the smallness assumption for  $\Delta_w$  and  $\Delta_f$  in each iteration, it is theoretically legitimate to apply quadratic and linear approximations to the objective function and the constraints in (9), respectively. For notational brevity, we redefine the unknown variables in (9) by a new vector  $\mathbf{x} \in \mathbb{R}^{2(N_s L + N_r^2) \times 1}$  as

$$\mathbf{x} = [\Re\{\text{vec}(\Delta_w)^T\}, \Im\{\text{vec}(\Delta_w)^T\}, \Re\{\text{vec}(\Delta_f)^T\}, \Im\{\text{vec}(\Delta_f)^T\}]^T. \quad (15)$$

Let

$$x_k = |\mathbf{g}_k^H \mathbf{F} \mathbf{H} \mathbf{w}_k|^2 \quad (16)$$

and

$$y_k = \sum_{j=1, j \neq k}^L |\mathbf{g}_k^H \mathbf{F} \mathbf{H} \mathbf{w}_j|^2 + \frac{1}{\rho_1} \|\mathbf{g}_k^H \mathbf{F}\|^2 + \frac{1}{\rho_1 \rho_2}. \quad (17)$$

By substituting (8) into the above equations, we can express  $x_k^{(n+1)}$  and  $y_k^{(n+1)}$  at the  $(n+1)$ th step as

$$\begin{aligned} & x_k^{(n+1)} \\ &= \mathbf{g}_k^H \mathbf{F}^{(n+1)} \mathbf{H} \mathbf{w}_k^{(n+1)} \mathbf{w}_k^{(n+1)H} \mathbf{H}^H \mathbf{F}^{(n+1)H} \mathbf{g}_k \\ &= x_k^{(n)} + 2\Re \left\{ \text{vec} \left( \mathbf{H}^H \mathbf{F}^{(n)} \mathbf{g}_k \mathbf{g}_k^H \mathbf{F}^{(n)} \mathbf{H} \mathbf{w}_k^{(n)} \right)^H \Delta_{w,k} \right\} \\ & \quad + 2\Re \left\{ \text{vec} \left( \mathbf{g}_k \mathbf{g}_k^H \mathbf{F}^{(n)} \mathbf{H} \mathbf{w}_k^{(n)} \mathbf{w}_k^{(n)H} \mathbf{H}^H \right)^H \text{vec}(\Delta_f) \right\} \\ & \quad + O(\|\mathbf{x}\|^2) \\ & \approx x_k^{(n)} + \mathbf{d}_{x,k}^T \mathbf{x} \end{aligned} \quad (18)$$

and

$$\begin{aligned} & y_k^{(n+1)} \\ &= \sum_{j=1, j \neq k}^L \mathbf{g}_k^H \mathbf{F}^{(n+1)} \mathbf{H} \mathbf{w}_j^{(n+1)} \mathbf{w}_j^{(n+1)H} \mathbf{H}^H \mathbf{F}^{(n+1)H} \mathbf{g}_k \\ & \quad + \frac{1}{\rho_1} \mathbf{g}_k^H \mathbf{F}^{(n+1)} \mathbf{F}^{(n+1)H} \mathbf{g}_k + \frac{1}{\rho_1 \rho_2} \\ &= y_k^{(n)} + 2 \sum_{j=1, j \neq k}^L \Re \left\{ \text{vec} \left( \mathbf{H}^H \mathbf{F}^{(n)} \mathbf{g}_k \mathbf{g}_k^H \mathbf{F}^{(n)} \mathbf{H} \mathbf{w}_j^{(n)} \right)^H \Delta_{w,j} \right\} \\ & \quad + 2\Re \left\{ \text{vec} \left( \mathbf{g}_k \mathbf{g}_k^H \mathbf{F}^{(n)} \left( \frac{1}{\rho_1} \mathbf{I}_{N_r} + \sum_{j=1, j \neq k}^L \mathbf{H} \mathbf{w}_j^{(n)} \mathbf{w}_j^{(n)H} \mathbf{H}^H \right) \right)^H \right. \\ & \quad \left. \times \text{vec}(\Delta_f) \right\} + O(\|\mathbf{x}\|^2) \\ & \approx y_k^{(n)} + \mathbf{d}_{y,k}^T \mathbf{x}, \end{aligned} \quad (19)$$

respectively, where  $\Delta_{w,j}$  is  $j$ th column of  $\Delta_w$ . In (18) and (20), we neglect the second-order term of  $\|\mathbf{x}\|$  for element-wise small  $\mathbf{x}$ , and define

$$\mathbf{d}_{x,k}^T = 2 \left[ \mathbf{0}_{1 \times N_s(k-1)}, \Re\{\mathbf{a}_{k,k}^T\}, \mathbf{0}_{1 \times N_s(L-1)}, \Im\{\mathbf{a}_{k,k}^T\}, \mathbf{0}_{1 \times N_s(L-k)}, \Re\{\mathbf{b}_{k,k}^T\}, \Im\{\mathbf{b}_{k,k}^T\} \right] \quad (21)$$

and

$$\mathbf{d}_{y,k}^T = 2 \left[ \Re\{\tilde{\mathbf{a}}_k^T\}, \Im\{\tilde{\mathbf{a}}_k^T\}, \Re \left\{ \mathbf{c}_k^T + \sum_{j \neq k} \mathbf{b}_{k,j}^T \right\}, \Im \left\{ \mathbf{c}_k^T + \sum_{j \neq k} \mathbf{b}_{k,j}^T \right\} \right], \quad (22)$$

where  $\mathbf{a}_{k,j} = \text{vec} \left( \mathbf{H}^H \mathbf{F}^{(n)H} \mathbf{g}_k \mathbf{g}_k^H \mathbf{F}^{(n)} \mathbf{H} \mathbf{w}_j^{(n)} \right)$ ,  $\mathbf{b}_{k,j} = \text{vec} \left( \mathbf{g}_k \mathbf{g}_k^H \mathbf{F}^{(n)} \mathbf{H} \mathbf{w}_j^{(n)} \mathbf{w}_j^{(n)H} \mathbf{H}^H \right)$ ,  $\mathbf{c}_k = \frac{1}{\rho_1} \text{vec} \left( \mathbf{g}_k \mathbf{g}_k^H \mathbf{F}^{(n)} \right)$ , and  $\tilde{\mathbf{a}}_k = [\mathbf{a}_{k,1}^T, \dots, \mathbf{a}_{k,(k-1)}^T, \mathbf{0}_{1 \times N_s}, \mathbf{a}_{k,(k+1)}^T, \dots, \mathbf{a}_{k,L}^T]^T$ .

Then, using (18) and (20), we can rewrite the SINR  $\gamma_k$  in (9) as

$$\begin{aligned} & \gamma_k^{(n+1)} \\ &= \frac{x_k^{(n+1)}}{y_k^{(n+1)}} \\ & \approx \frac{x_k^{(n)}}{y_k^{(n)}} + (\mathbf{d}_{x,k}^T \mathbf{x}) \frac{\partial \gamma}{\partial x} \Big|_{x_k^{(n)}, y_k^{(n)}} + (\mathbf{d}_{y,k}^T \mathbf{x}) \frac{\partial \gamma}{\partial y} \Big|_{x_k^{(n)}, y_k^{(n)}} \\ &= \gamma_k^{(n)} + \left( \frac{1}{y_k^{(n)}} \mathbf{d}_{x,k} - \frac{x_k^{(n)}}{y_k^{(n)2} \mathbf{d}_{y,k}} \right)^T \mathbf{x}. \end{aligned} \quad (23)$$

Define

$$\mathbf{p}_k = \frac{1}{y_k^{(n)}} \mathbf{d}_{x,k} - \frac{x_k^{(n)}}{y_k^{(n)2} \mathbf{d}_{y,k}}. \quad (25)$$

By utilizing Taylor expansion with respect to  $\gamma_k$  in the objective function of (9) and neglecting the high-order terms, we obtain

$$\begin{aligned} & \sum_{k=1}^L \log \left( 1 + \gamma_k^{(n+1)} \right) \\ & \approx \sum_{k=1}^L \log \left( 1 + \gamma_k^{(n)} + \mathbf{p}_k^T \mathbf{x} \right) \\ &= \left( \sum_{k=1}^L \frac{1}{1 + \gamma_k^{(n)}} \mathbf{p}_k^T \right) \mathbf{x} - \mathbf{x}^T \left( \sum_{k=1}^L \frac{1}{2 \left( 1 + \gamma_k^{(n)} \right)^2} \mathbf{p}_k \mathbf{p}_k^T \right) \mathbf{x} \\ & \quad + \kappa_1. \end{aligned} \quad (26)$$

where  $\kappa_1 = \sum_{k=1}^L \log \left( 1 + \gamma_k^{(n)} \right)$  is a constant. The problem of maximizing the objective function in (9) can now be converted to

$$\begin{aligned} & \underset{\mathbf{x}}{\text{minimize}} \quad \mathbf{x}^T \left( \sum_{k=1}^L \frac{1}{2 \left( 1 + \gamma_k^{(n)} \right)^2} \mathbf{p}_k \mathbf{p}_k^T \right) \mathbf{x} \\ & \quad - \left( \sum_{k=1}^L \frac{1}{1 + \gamma_k^{(n)}} \mathbf{p}_k^T \right) \mathbf{x} \end{aligned} \quad (27)$$

with the quadratic coefficient matrix being positive semidefinite, that is, a convex objective function.

Subsequently, we will deal with the two power constraints by using linear approximations. Concerning the BS transmit

$$\mathbf{A} = \begin{bmatrix} 2\Re \left\{ \text{vec} \left( \mathbf{W}^{(n)} \right) \right\} & 2\rho_1 \Re \left\{ \text{vec} \left( \mathbf{H}^H \mathbf{F}^{(n)H} \mathbf{F}^{(n)} \mathbf{H} \mathbf{W}^{(n)} \right) \right\} \\ 2\Im \left\{ \text{vec} \left( \mathbf{W}^{(n)} \right) \right\} & 2\rho_1 \Im \left\{ \text{vec} \left( \mathbf{H}^H \mathbf{F}^{(n)H} \mathbf{F}^{(n)} \mathbf{H} \mathbf{W}^{(n)} \right) \right\} \\ \mathbf{0}_{N_r^2 \times 1} & 2\rho_1 \Re \left\{ \text{vec} \left( \mathbf{F}^{(n)} \left( \mathbf{H} \mathbf{W}^{(n)} \mathbf{W}^{(n)H} \mathbf{H}^H + \frac{1}{\rho_1} \mathbf{I}_{N_r} \right) \right) \right\} \\ \mathbf{0}_{N_r^2 \times 1} & 2\rho_1 \Re \left\{ \text{vec} \left( \mathbf{F}^{(n)} \left( \mathbf{H} \mathbf{W}^{(n)} \mathbf{W}^{(n)H} \mathbf{H}^H + \frac{1}{\rho_1} \mathbf{I}_{N_r} \right) \right) \right\} \end{bmatrix} \quad (31)$$

$$\mathbf{e} = \begin{bmatrix} \frac{P_1}{\rho_1} - \text{tr} \left( \mathbf{W}^{(n)} \mathbf{W}^{(n)H} \right) \\ \frac{P_2}{\rho_2} - \rho_1 \text{tr} \left( \mathbf{F}^{(n)} \mathbf{H} \mathbf{W}^{(n)} \mathbf{W}^{(n)H} \mathbf{H}^H \mathbf{F}^{(n)H} \right) - \text{tr} \left( \mathbf{F}^{(n)} \mathbf{F}^{(n)H} \right) \end{bmatrix} \quad (32)$$

power constraint, we have

$$\text{tr} \left( \mathbf{W}^{(n+1)} \mathbf{W}^{(n+1)H} \right) = \text{tr} \left( \mathbf{W}^{(n)} \mathbf{W}^{(n)H} \right) + 2\Re \left\{ \text{tr} \left( \Delta_w \mathbf{W}^{(n)H} \right) \right\} + \text{tr} \left( \Delta_w \Delta_w^H \right). \quad (28)$$

By neglecting the quadratic term of  $\Delta_w$  and utilizing the equality  $\text{tr}(\mathbf{A}^H \mathbf{B}) = \text{vec}(\mathbf{A})^H \text{vec}(\mathbf{B})$ , the BS transmit power constraint becomes

$$2\Re \left\{ \text{vec} \left( \mathbf{W}^{(n)} \right)^H \text{vec} \left( \Delta_w \right) \right\} = \frac{P_1}{\rho_1} - \text{tr} \left( \mathbf{W}^{(n)} \mathbf{W}^{(n)H} \right). \quad (29)$$

With similar manipulations, we can convert the RS transmit power constraint in (9) to

$$\begin{aligned} & 2\rho_1 \Re \left\{ \text{vec} \left( \mathbf{F}^{(n)} \left( \mathbf{H} \mathbf{W}^{(n)} \mathbf{W}^{(n)H} \mathbf{H}^H + \frac{1}{\rho_1} \mathbf{I}_{N_r} \right) \right)^H \right. \\ & \left. \times \text{vec} \left( \Delta_f \right) + \text{vec} \left( \mathbf{H}^H \mathbf{F}^{(n)H} \mathbf{F}^{(n)} \mathbf{H} \mathbf{W}^{(n)} \right)^H \text{vec} \left( \Delta_w \right) \right\} \\ & = \frac{P_2}{\rho_2} - \rho_1 \text{tr} \left( \mathbf{F}^{(n)} \mathbf{H} \mathbf{W}^{(n)} \mathbf{W}^{(n)H} \mathbf{H}^H \mathbf{F}^{(n)H} \right) \\ & \quad - \text{tr} \left( \mathbf{F}^{(n)} \mathbf{F}^{(n)H} \right). \end{aligned} \quad (30)$$

Then from (29) and (30), by defining  $\mathbf{A}$  and  $\mathbf{e}$  in (31) and (32), respectively, we can rewrite the power constraints (29) and (30) concisely by

$$\mathbf{A}^T \mathbf{x} = \mathbf{e}. \quad (33)$$

With the reformulated objective function in (27), power constraints in (31), and an affine constraint on smallness of  $\mathbf{x}$ , we obtain (10) as the quadratic program reformulation of (9).

#### ACKNOWLEDGMENT

The authors would like to thank the editor and the anonymous reviewers for their constructive comments that have helped improve the quality of this paper.

#### REFERENCES

- [1] W. Xu, X. Dong, and W.-S. Lu, "On joint optimization of source and relay precoding design for MIMO downlink channels," in *Proc. IEEE ICC*, May 2010, pp. 1–5.
- [2] R. Pabst, B. H. Walke, D. C. Schultz, *et al.*, "Relay-based deployment concepts for wireless and mobile broadband radio," *IEEE Commun. Mag.*, vol. 42, no. 9, pp. 80–89, Sep. 2004.
- [3] S. W. Peters, A. Y. Panah, K. T. Truong, and R. W. Heath, Jr., "Relay architectures for 3GPP LTE-advanced," *EURASIP J. Wireless Commun. Netw.*, vol. 2009, article ID 618787, 14 pages, 2009.

- [4] S. Sesia, I. Toufik, and M. Baker, *LTE, The UMTS Long Term Evolution: From Theory to Practice*. Wiley & Sons, 2009.
- [5] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. Telecommun.*, vol. 10, pp. 585–598, Nov. 1999.
- [6] B. Wang, J. Zhang, and A. Høst-Madsen, "On the capacity of MIMO relay channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 29–43, Jan. 2005.
- [7] X. Tang and Y. Hua, "Optimal design of non-regenerative MIMO wireless relays," *IEEE Trans. Wireless Commun.*, vol. 6, no. 4, pp. 1398–1407, Apr. 2007.
- [8] O. Muñoz-Medina, J. Vidal, and A. Agustín, "Linear transceiver design in nonregenerative relays with channel state information," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2593–2604, June 2007.
- [9] Y. Huang, L. Yang, M. Bengtsson, and B. Ottersten, "A limited feedback joint precoding for amplify-and-forward relaying," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1347–1357, Mar. 2010.
- [10] H. Weingarten, Y. Stenberg, and S. Shamai, "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3936–3964, Sep. 2006.
- [11] P. Viswanath, D. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1277–1294, June 2002.
- [12] Q. H. Spencer, A. Lee, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 461–471, Feb. 2004.
- [13] K. K. Wong, R. D. Murch, and K. B. Letaief, "A joint-channel diagonalization for multiuser MIMO antenna systems," *IEEE Trans. Wireless Commun.*, vol. 2, no. 4, pp. 773–786, July 2003.
- [14] C. Zhang, W. Xu, and M. Chen, "Hybrid zero-forcing beamforming/orthogonal beamforming with user selection for MIMO broadcast channels," *IEEE Commun. Lett.*, vol. 13, no. 1, pp. 10–12, Jan. 2009.
- [15] C. Eslí and A. Wittneben, "Multiuser MIMO two-way relaying for cellular communications," in *Proc. IEEE PIMRC*, Sep. 2008.
- [16] F. A. Onat, H. Yanikomeroglu, and S. Periyalwar, "Relay-assisted spatial multiplexing in wireless fixed relay networks," in *Proc. IEEE GLOBECOM*, Nov. 2006.
- [17] C. B. Chae, T. Tang, R. W. Heath Jr., and S. Cho, "MIMO relaying with linear processing for multiuser transmission in fixed relay networks," *IEEE Trans. Signal Process.*, vol. 56, no. 2, pp. 727–738, Feb. 2008.
- [18] R. Zhang, C. C. Chai, and Y. C. Liang, "Joint beamforming and power control for multiantenna relay broadcast channel with QoS constraints," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 726–737, Feb. 2009.
- [19] I. Hammerström and A. Wittneben, "Power allocation schemes for amplify-and-forward MIMO-OFDM relay links," *IEEE Trans. Wireless Commun.*, vol. 6, no. 8, pp. 2798–2802, Aug. 2007.
- [20] A. Antoniou and W.-S. Lu, *Practical Optimization: Algorithms and Engineering Applications*. Springer, 2007.
- [21] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [22] R. Fletcher, *Practical Methods of Optimization*, 2nd edition. Wiley, 1987.
- [23] Y. Nesterov and A. Nemirovskii, *Interior-Point Polynomial Algorithms in Convex Programming*. SIAM, 1994.
- [24] J. Lee and N. Jindal, "High SNR analysis for MIMO broadcast channels: dirty paper coding versus linear precoding," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4787–4792, Dec. 2007.
- [25] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd edition. The John Hopkins University Press, 1996.
- [26] G. Caire and S. Shamai, "On the achievable throughput of a multi-antenna Gaussian broadcast channel," *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1691–1706, July 2003.