

RESEARCH

Open Access

Joint source and relay design for two-hop amplify-and-forward relay networks with QoS constraints

Jafar Mohammadi¹, Feifei Gao^{2*}, Yue Rong³ and Wen Chen⁴

Abstract

In this paper, we consider the joint design of source precoding matrix and the relay precoding matrix in a two-hop multiple-input multiple-output relay network. The goal is to find a pair of matrices in order to minimize the power consumption and at the same time meet pre-selected quality of service constraints that are defined as the mean square error of each data stream. Using majorization theory, we simplify the matrix-valued optimization problem into a scalar-valued one. We then propose a lower bound and an upper bound of the original problem, both in convex forms. Specifically, the latter is solved by a multi-level water-filling algorithm that is much efficient than directly applying the interior point method. Numerical examples corroborate the proposed studies and also demonstrate the tightness of both bounds to the original problem.

Keywords: Joint precoding design, MIMO relay networks, Quality of service, Majorization, Convex optimization

1 Introduction

Relay networks have recently attracted much attention because of their promising capability in achieving reliable communication and wide coverage for the next generation of wireless systems [1,2]. Different types of relaying strategies, e.g., *amplify and forward* (AF), *decode and forward* (DF), and *compress and forward* (CF) were introduced in [3-5], respectively. DF and CF decode data before retransmission and are, thus, also known as regenerative strategies; AF relay amplifies the received data only and is known as non-regenerative strategy. The computational simplicity in AF relay makes it highly attractive and a strong candidate for the real-time application. On the other side, multi-input multi-output (MIMO) technique [6] can enhance the data transmission rate by introducing the spatial diversity gain. Therefore, combining relaying and MIMO becomes a natural way to further advance the future wireless communication systems.

Most of the existing works focus on the AF MIMO relay networks, where a certain performance criterion is optimized subject to power constraints at both source and

relay. For example, the mutual information and the total mean square error (MSE) criteria are selected as objective functions in [7,8], respectively. The authors in [9] investigated the diversity multiplexing tradeoff for MIMO relays. Moreover, there are some works on beamforming design for special types of AF MIMO relays; for instance, the authors in [10] considered codebook design for half-duplex AF MIMO relays, while the author in [11] considered beamforming design for rather vast types of MIMO relays. Applying the majorization theory [12], the authors in [13] proposed a unified framework for most performance criteria. The extension of [13] to the multiple relay case was introduced in [14].

All of the mentioned methods above considered enhancing the system performance by maximizing or minimizing a certain objective function constraint on power consumption at some or all nodes. Although this model improves the system performance, it does not guarantee a certain quality of service (QoS) requirement for an individual user. The importance of considering QoS becomes more obvious in practical applications supporting several types of service each with a different reliability requirement. One of the pioneering works considering QoS in MIMO point-to-point systems is [15], where the

*Correspondence: feifeigao@ieee.org

²Tsinghua National Laboratory for Information Science and Technology, Beijing, 100084, China

Full list of author information is available at the end of the article

authors optimized the transmission power subject to predefined sets of QoS, e.g., individual MSE, signal-to-noise ratio (SNR), and bit error rate. In [16], the authors investigated the optimization of source beamforming and relay weighting matrix in order to minimize the total power subject to a given set of QoS for multi-input single-output broadcast channel. On the aspect of relay network, AF MIMO relay network with QoS constraints has been investigated in [17]. Applying majorization theory, the author in [17] proposed a unified framework for the design of the optimal structure of the source precoding and relay amplifying matrices. The author applied a successive geometric programming method to obtain the optimal power loading among data streams. Unfortunately, the computational complexity of the proposed solution in [17] compromises its suitability for practical implementation. With similar assumptions, the authors in [18] considered a simplified version of the problem in which only the relay power is minimized. Then, the minimization is executed over a convex lower bound of the objective function. In a rather more general setup, the authors in [19] studied the joint relay and source power minimization and applied the majorization theory to reduce the problem to a scalar one. Then, using QoS convex relaxation, an upper bound and a lower bound on the optimal results are presented.

In this paper, building upon the results in [19], we take a specific look into a dual-hop^a AF MIMO relay network. We first jointly design the source and the relay precoding matrices such that the overall transmission power is minimized subject to a given set of QoS constraints. Applying the majorization theory, we reduce the original matrix-valued problem to a scalar-valued one and then propose two new convex optimization problems whose objective values serve as the lower bound and the upper bound of the original problem. While both new problems can be handled by the existing convex optimization tools, e.g., CVX [20], we specifically design a multi-level water-filling (MLWF) algorithm to solve the upper bound problem that can further reduce the computational complexity. Compared with the successive geometric programming approach developed in [17], the MLWF algorithm does not require any optimization tool and thus is easier to implement for practical relay systems. Numerical results corroborate the proposed studies and clearly demonstrate the tightness of the proposed lower bound and upper bound, especially over low MSE region.

The rest of this paper is arranged as follows: Section 2 presents the system model and formulates the optimization problem in matrix form. In Section 3, the optimization is simplified to a scalar-valued problem from the majorization theory. Two suboptimal problems whose objectives serve as the upper bound and the lower bound of the original problems are derived in the same section.

In Section 4, the upper bound problem is solved from a multi-level water-filling algorithm coupled with decomposition methods. The simulation results are presented in Section 5, and conclusions are drawn in Section 6.

1.1 Notations

Vectors and matrices are boldface small and capital letters, respectively; the transpose, complex conjugate, Hermitian, inverse, and pseudo-inverse of \mathbf{A} are denoted by \mathbf{A}^T , \mathbf{A}^* , \mathbf{A}^H , \mathbf{A}^{-1} , and \mathbf{A}^\dagger , respectively; $\|\mathbf{a}\|$ denotes the Euclidean norm of the vector \mathbf{a} ; $\text{diag}\{\mathbf{a}\}$ is the diagonal matrix with diagonal elements given by \mathbf{a} , while $\text{diag}\{\mathbf{A}\}$ is a vector with entries taken from the diagonal elements of \mathbf{A} ; \mathbf{I} is the identity matrix; and $E\{\cdot\}$ is the statistical expectation. Moreover, basic notations and definitions of majorization theory can be found in Appendix 1.

2 System model

As shown in Figure 1, we consider a three-node multi-antenna relay network that consists of a source node, a relay node, and a destination node, equipped with M , N , and P antennas, respectively. We assume that the direct link between the source and the destination is weak and, therefore, can be neglected. Denote the baseband MIMO channels between the source and relay as the $N \times M$ matrix \mathbf{H}_s , while that between the relay and destination as the $P \times N$ matrix \mathbf{H}_r . We further assume that the channel state information is perfectly known at all nodes. Suppose that L data streams, denoted by \mathbf{x} , are precoded by an $M \times L$ precoding matrix \mathbf{B} at source node. We require $L \leq \min\{M, N, P\}$ so that the data streams can be detected with linear method at the destination. With an $N \times N$ precoding matrix \mathbf{F} at relay, the received signal at destination is as follows:

$$\mathbf{z} = \mathbf{H}_r \mathbf{F} \mathbf{H}_s \mathbf{B} \mathbf{x} + \mathbf{H}_r \mathbf{F} \mathbf{v}_r + \mathbf{v}_d, \quad (1)$$

where \mathbf{v}_r and \mathbf{v}_d are additive white complex Gaussian noise at relay and destination, respectively, i.e., $\mathbf{v}_r \in \mathcal{CN}(0, \sigma_r^2 \mathbf{I}_N)$ and $\mathbf{v}_d \in \mathcal{CN}(0, \sigma_d^2 \mathbf{I}_N)$. Without loss of generality, we set $\sigma_d^2 = \sigma_r^2 = 1$. Since the correlation and power can be designed through \mathbf{B} , the data streams from the source can be assumed independent from each other, i.e., $E\{\mathbf{x}\mathbf{x}^H\} = \mathbf{I}$.

We consider the minimum mean square error detection at destination with a $P \times L$ decoding matrix \mathbf{T} . The estimated data can be expressed as follows:

$$\hat{\mathbf{x}} = \mathbf{T}^H \mathbf{z}, \quad (2)$$

with the error matrix

$$\mathbf{C} = E\{(\hat{\mathbf{x}} - \mathbf{x})(\hat{\mathbf{x}} - \mathbf{x})^H\}. \quad (3)$$

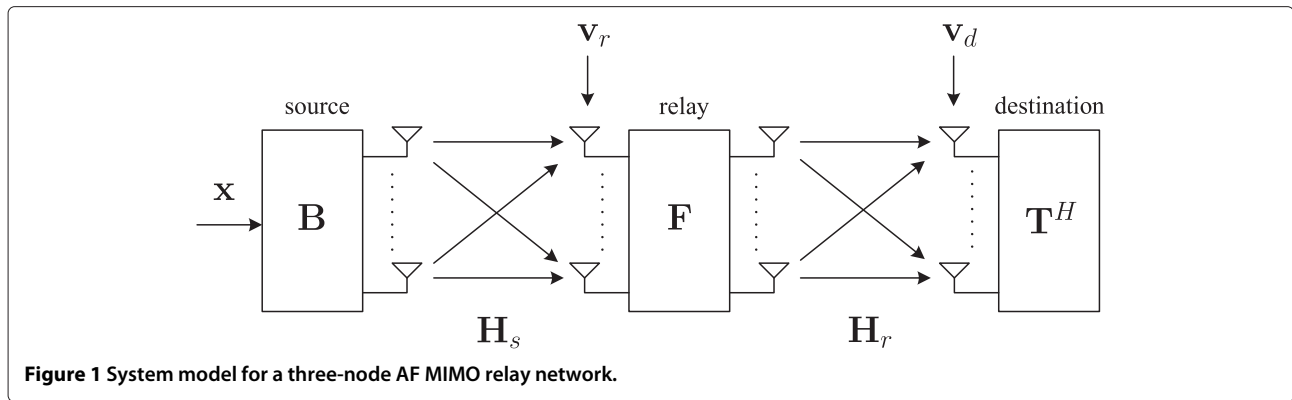


Figure 1 System model for a three-node AF MIMO relay network.

The MSE is defined as $\text{tr}(\mathbf{C})$. For the given \mathbf{B} and \mathbf{F} , one can easily find the optimal \mathbf{T} from the standard approach [21]:^b

$$\mathbf{T} = (\mathbf{H}_r \mathbf{F} \mathbf{H}_s \mathbf{B} (\mathbf{H}_r \mathbf{F} \mathbf{H}_s \mathbf{B})^H + \mathbf{H}_r \mathbf{F} (\mathbf{H}_r \mathbf{F})^H + \mathbf{I})^{-1} \mathbf{H}_r \mathbf{F} \mathbf{H}_s \mathbf{B}, \quad (4)$$

and the corresponding error matrix is

$$\mathbf{C} = [\mathbf{I} + (\mathbf{H}_r \mathbf{F} \mathbf{H}_s \mathbf{B})^H (\mathbf{H}_r \mathbf{F} (\mathbf{H}_r \mathbf{F})^H + \mathbf{I})^{-1} \mathbf{H}_r \mathbf{F} \mathbf{H}_s \mathbf{B}]^{-1}. \quad (5)$$

As in [15], the QoS measurement here is taken as the MSE of each individual data stream. Let us denote the QoS vector as $\boldsymbol{\rho} = [\rho_1, \dots, \rho_L]^T$, i.e., the MSE of the i th data stream is required to be smaller than ρ_i . Note that $\rho_i < 1$ is necessary to avoid trivial solutions.

From (5), the QoS constraints are stated as follows:

$$\mathbf{C}_{ii} \leq \rho_i \quad \text{or} \quad \text{diag}\{\mathbf{C}\} \leq \boldsymbol{\rho}, \quad (6)$$

where ' \leq ' denotes the element-wise operation if used between vectors.

The average power consumed by the source is computed from

$$E\{\text{tr}(\mathbf{B} \mathbf{x} \mathbf{x}^H \mathbf{B}^H)\} = \text{tr}(\mathbf{B} \mathbf{B}^H), \quad (7)$$

while that spent by the relay is

$$\begin{aligned} E\{\text{tr}((\mathbf{F} \mathbf{H}_s \mathbf{B} \mathbf{x} + \mathbf{F} \mathbf{v}_r)(\mathbf{F} \mathbf{H}_s \mathbf{B} \mathbf{x} + \mathbf{F} \mathbf{v}_r)^H)\} \\ = \text{tr}(\mathbf{F} (\mathbf{H}_s \mathbf{B} \mathbf{B}^H \mathbf{H}_s^H + \mathbf{I}) \mathbf{F}^H). \end{aligned}$$

3 Optimization problem

The goal is to find the optimal \mathbf{B} and \mathbf{F} to minimize the overall power spent by the source and relay and at the same time meet the QoS requirements.

Define

$$\mathbf{M} = \mathbf{H}_r \mathbf{F} \mathbf{H}_s \mathbf{B}, \quad (8)$$

$$\mathbf{R} = ((\mathbf{H}_r \mathbf{F}) (\mathbf{H}_r \mathbf{F})^H + \mathbf{I})^{-1}. \quad (9)$$

The optimization problem is then expressed as follows:

$$(P1): \quad \min_{\mathbf{B}, \mathbf{F}} \quad \text{tr}(\mathbf{B} \mathbf{B}^H) + \text{tr}(\mathbf{F} (\mathbf{H}_s \mathbf{B} \mathbf{B}^H \mathbf{H}_s^H + \mathbf{I}) \mathbf{F}^H) \quad (10)$$

$$\text{subject to} \quad \text{diag}\{(\mathbf{I} + \mathbf{M}^H \mathbf{R} \mathbf{M})^{-1}\} \leq \boldsymbol{\rho}.$$

Unfortunately, the problem is non-convex and cannot be solved in an efficient way.

3.1 Equivalent problem

Let us define a new optimization:

$$(P2): \quad \min_{\tilde{\mathbf{B}}, \mathbf{F}} \quad \text{tr}(\tilde{\mathbf{B}} \tilde{\mathbf{B}}^H) + \text{tr}(\mathbf{F} (\mathbf{H}_s \tilde{\mathbf{B}} \tilde{\mathbf{B}}^H \mathbf{H}_s^H + \mathbf{I}) \mathbf{F}^H) \quad (11)$$

$$\text{subject to} \quad \text{diag}\{(\mathbf{I} + \tilde{\mathbf{M}}^H \mathbf{R} \tilde{\mathbf{M}})^{-1}\} \succ^w \boldsymbol{\rho}$$

$$\tilde{\mathbf{M}} = \mathbf{H}_r \mathbf{F} \mathbf{H}_s \tilde{\mathbf{B}}$$

$$\tilde{\mathbf{M}}^H \mathbf{R} \tilde{\mathbf{M}} \text{ is diagonal,}$$

where ' \succ^w ' means weak majorization, and the details can be found in Appendix 1.

Theorem 1. *Problems (P1) and (P2) are equivalent.*

Proof. The idea is to prove that for each feasible point of (P1), there is a corresponding feasible point in (P2) that yields the same objective value and vice versa.

(P1) \rightarrow (P2): For any feasible \mathbf{B} in (P1), construct a new matrix $\tilde{\mathbf{B}} = \mathbf{B} \mathbf{Q}$, where \mathbf{Q} is the unitary eigenmatrix of $\mathbf{M}^H \mathbf{R} \mathbf{M}$. Then, $\tilde{\mathbf{M}}^H \mathbf{R} \tilde{\mathbf{M}}$ is a diagonal matrix. It can be readily checked that the objective value of (P2) with $(\tilde{\mathbf{B}}, \mathbf{F})$ is the same as the objective value of (P1) with (\mathbf{B}, \mathbf{F}) . Moreover, since $\mathbf{I} + \tilde{\mathbf{M}}^H \mathbf{R} \tilde{\mathbf{M}}$ is a diagonal matrix, there is

$$\begin{aligned} \text{diag}\{(\mathbf{I} + \tilde{\mathbf{M}}^H \mathbf{R} \tilde{\mathbf{M}})^{-1}\} &= \boldsymbol{\lambda}\{(\mathbf{I} + \tilde{\mathbf{M}}^H \mathbf{R} \tilde{\mathbf{M}})^{-1}\} \\ &= \boldsymbol{\lambda}\{(\mathbf{I} + \mathbf{M}^H \mathbf{R} \mathbf{M})^{-1}\} = \boldsymbol{\lambda}\{\mathbf{Z}\}, \end{aligned} \quad (12)$$

where $\mathbf{Z} \triangleq (\mathbf{I} + \mathbf{M}^H \mathbf{R} \mathbf{M})^{-1}$. From $\mathbf{Z}_{ii} \leq \rho_i$, one can conclude that $\text{diag}\{\mathbf{Z}\} \succ^w \boldsymbol{\rho}$. We further know from Lemma 2 in Appendix 1 that $\boldsymbol{\lambda}\{\mathbf{Z}\} \succ^w \text{diag}\{\mathbf{Z}\}$. Therefore, $\boldsymbol{\lambda}\{\mathbf{Z}\} \succ^w \boldsymbol{\rho}$.

ρ and we achieve the conclusion that for any feasible point (\mathbf{B}, \mathbf{F}) in (P1), there is always a corresponding feasible point $(\tilde{\mathbf{B}}, \mathbf{F})$ in (P2) that gives the same objective value.

(P2) \rightarrow (P1): Define $\tilde{\mathbf{Z}} = (\mathbf{I} + \tilde{\mathbf{M}}^H \mathbf{R} \tilde{\mathbf{M}})^{-1}$ and assume that $(\tilde{\mathbf{B}}, \mathbf{F})$ is a feasible point of (P2). From (P2), we know that $\text{diag}\{\tilde{\mathbf{Z}}\} = \lambda\{\tilde{\mathbf{Z}}\} \succ^w \rho$ holds. From Lemma 3 in Appendix 1, we know that there exists a vector \mathbf{c} satisfying both $\lambda\{\tilde{\mathbf{Z}}\} \succ \mathbf{c}$ and $\mathbf{c} \leq \rho$. From Lemma 2, we know that for each $\mathbf{c} \prec \lambda\{\tilde{\mathbf{Z}}\}$, there exists a matrix \mathbf{W} with $\text{diag}\{\mathbf{W}\} = \mathbf{c}$ and $\lambda\{\mathbf{W}\} = \lambda\{\tilde{\mathbf{Z}}\}$. Let $\mathbf{W} = \mathbf{Q}^H \tilde{\mathbf{Z}} \mathbf{Q}$ and define $\mathbf{B} = \tilde{\mathbf{B}} \mathbf{Q}$. Then, $\text{diag}\{(\mathbf{I} + \mathbf{M}^H \mathbf{R} \mathbf{M})^{-1}\} = \text{diag}\{\mathbf{W}\} \leq \rho$. Moreover, the objective function of (P1) with (\mathbf{B}, \mathbf{F}) is the same as the objective function of (P2). \square

3.2 Processing matrix variables

Based on Theorem 1, we can solve (P2) instead of (P1).

Define two new positive semi-definite matrices as well as their singular value decomposition (SVD) as follows:

$$\mathbf{X} = \mathbf{H}_s \tilde{\mathbf{B}} \tilde{\mathbf{B}}^H \mathbf{H}_s^H = \mathbf{U}_X \Lambda_X \mathbf{U}_X^H, \quad (13)$$

$$\mathbf{Y} = \mathbf{H}_r \mathbf{F} (\mathbf{X} + \mathbf{I}) \mathbf{F}^H \mathbf{H}_r^H = \mathbf{U}_Y \Lambda_Y \mathbf{U}_Y^H, \quad (14)$$

where \mathbf{U}_X and \mathbf{U}_Y are the $N \times L$ and $P \times L$ orthonormal matrices, respectively. Throughout this paper, we always sort the singular values and eigenvalues in an increasing order. Note that it is still possible that Λ_X or Λ_Y contains some zero entries.

The matrices $\mathbf{H}_s \tilde{\mathbf{B}}$ and $\mathbf{H}_r \mathbf{F}$ can be represented as follows:

$$\mathbf{H}_s \tilde{\mathbf{B}} = \mathbf{U}_X \Lambda_X^{\frac{1}{2}} \mathbf{V}_X^H, \quad (15)$$

$$\mathbf{H}_r \mathbf{F} = \mathbf{U}_Y \Lambda_Y^{\frac{1}{2}} \mathbf{V}_Y^H (\mathbf{X} + \mathbf{I})^{-\frac{1}{2}}, \quad (16)$$

where \mathbf{V}_X can be any $L \times L$ unitary matrix, and \mathbf{V}_Y can be any $N \times L$ orthonormal matrices. These two matrices will be designed later to fulfill the optimality requirement.

Let

$$\mathbf{H}_s = \mathbf{U}_s \Lambda_s \mathbf{V}_s^H \quad \text{and} \quad \mathbf{H}_r = \mathbf{U}_r \Lambda_r \mathbf{V}_r^H \quad (17)$$

be the SVD of \mathbf{H}_s and \mathbf{H}_r respectively. We further express the singular matrices as $\mathbf{U}_s \triangleq [\mathbf{U}_{s,2}, \mathbf{U}_{s,1}]$ and $\mathbf{U}_r \triangleq [\mathbf{U}_{r,2}, \mathbf{U}_{r,1}]$, in which $\mathbf{U}_{s,1}$ and $\mathbf{U}_{r,1}$ contain L first columns.

One can rewrite (15) and (16) as follows:

$$\begin{bmatrix} \Sigma_s & \mathbf{0}_{L \times (M-L)} \\ \mathbf{0}_{(N-L) \times L} & \mathbf{0}_{(N-L) \times (M-L)} \end{bmatrix} \mathbf{V}_s^H \tilde{\mathbf{B}} = \mathbf{U}_s^H \mathbf{U}_X \Lambda_X^{\frac{1}{2}} \mathbf{V}_X^H, \quad (18)$$

$$\begin{bmatrix} \Sigma_r & \mathbf{0}_{L \times (N-L)} \\ \mathbf{0}_{(P-L) \times L} & \mathbf{0}_{(P-L) \times (N-L)} \end{bmatrix} \mathbf{V}_r^H \mathbf{F} = \mathbf{U}_r^H \mathbf{U}_Y \Lambda_Y^{\frac{1}{2}} \mathbf{V}_Y^H (\mathbf{X} + \mathbf{I})^{-\frac{1}{2}}.$$

There are three cases to be discussed:

- If $N = M = L$, then $\tilde{\mathbf{B}}$ can be re-expressed from (18) as

$$\tilde{\mathbf{B}} = \mathbf{V}_s \Sigma_s^{-1} \mathbf{U}_s^H \mathbf{U}_X \Lambda_X^{\frac{1}{2}} \mathbf{V}_X^H. \quad (19)$$

- If $M > N = L$, (18) can be simplified as

$$[\Sigma_s^{-1} \mathbf{0}_{L \times (M-L)}] \mathbf{V}_s^H \tilde{\mathbf{B}} = \mathbf{U}_{s,1}^H \mathbf{U}_X \Lambda_X^{\frac{1}{2}} \mathbf{V}_X^H. \quad (20)$$

The solution for $\tilde{\mathbf{B}}$ is not unique, and the one which minimizes the objective function should be chosen.

Basically, we need to solve

$$\min_{\tilde{\mathbf{B}}} \text{tr}(\tilde{\mathbf{B}} \tilde{\mathbf{B}}^H) \quad (21)$$

$$\text{subject to } [\Sigma_s^{-1} \mathbf{0}_{L \times (M-L)}] \mathbf{V}_s^H \tilde{\mathbf{B}} = \mathbf{U}_{s,1}^H \mathbf{U}_X \Lambda_X^{\frac{1}{2}} \mathbf{V}_X^H.$$

Note that the second term in the objective function is not included in (21) since it will be taken care by \mathbf{F} . From Lemma 9 in [14], the optimal structure of $\tilde{\mathbf{B}}$ is as follows:

$$\tilde{\mathbf{B}} = \mathbf{V}_s [\Sigma_s^{-1} \mathbf{0}_{L \times (M-L)}]^T \mathbf{U}_{s,1}^H \mathbf{U}_X \Lambda_X^{\frac{1}{2}} \mathbf{V}_X^H. \quad (22)$$

- If $M > L$ and $N > L$, then (20) holds if and only if $\mathbf{U}_{s,2}^H \mathbf{U}_X = \mathbf{0}_{((N-L) \times N)}$. In this case, (22) is still the optimal structure for $\tilde{\mathbf{B}}$. Please refer to [14] for the detailed derivation.

Following the same procedure, the optimal structure of \mathbf{F} can be computed as follows:

$$\mathbf{F} = \mathbf{V}_r [\Sigma_r^{-1} \mathbf{0}_{L \times (N-L)}]^T \mathbf{U}_{r,1}^H \mathbf{U}_Y \Lambda_Y^{\frac{1}{2}} \mathbf{V}_Y^H (\mathbf{X} + \mathbf{I})^{-\frac{1}{2}}. \quad (23)$$

We then proceed to the optimization over new variables $\mathbf{U}_X, \mathbf{V}_X, \mathbf{U}_Y, \mathbf{V}_Y, \Lambda_X$, and Λ_Y . Substituting (22) and (23) into (P2) yields the new objective function

$$\text{tr}(\Lambda_s^\dagger \mathbf{U}_{s,1}^H \mathbf{U}_X \Lambda_X \mathbf{U}_X^H \mathbf{U}_{s,1} \Lambda_s^\dagger) + \text{tr}(\Lambda_r^\dagger \mathbf{U}_{r,1}^H \mathbf{U}_Y \Lambda_Y \mathbf{U}_Y^H \mathbf{U}_{r,1} \Lambda_r^\dagger), \quad (24)$$

where the pseudo-inverses of Λ_s^\dagger and Λ_r^\dagger are $[\Sigma_s^{-1} \mathbf{0}_{L \times (N-L)}]^T$ and $[\Sigma_r^{-1} \mathbf{0}_{L \times (M-L)}]^T$, respectively.

We apply the inversion lemma [22],

$$(\mathbf{A}_1 + \mathbf{A}_2 \mathbf{A}_3 \mathbf{A}_4)^{-1} = \mathbf{A}_1^{-1} - \mathbf{A}_1^{-1} \mathbf{A}_2 (\mathbf{A}_3^{-1} + \mathbf{A}_4 \mathbf{A}_1^{-1} \mathbf{A}_2)^{-1} \times \mathbf{A}_4 \mathbf{A}_1^{-1}, \quad (25)$$

for the constraint and obtain

$$(\mathbf{I} + \tilde{\mathbf{M}}^H \mathbf{R} \tilde{\mathbf{M}})^{-1} = \mathbf{I} - \underbrace{\tilde{\mathbf{M}}^H (\mathbf{R}^{-1} + \tilde{\mathbf{M}} \tilde{\mathbf{M}}^H)^{-1} \tilde{\mathbf{M}}}_{\mathbf{G}}, \quad (26)$$

where \mathbf{G} is defined as the corresponding term. Substituting (22) and (23) into \mathbf{G} yields

$$\begin{aligned} \mathbf{G} &= \mathbf{V}_X \Lambda_X^{\frac{1}{2}} (\Lambda_X + \mathbf{I})^{-\frac{1}{2}} \mathbf{U}_X^H \mathbf{V}_Y \Lambda_Y^{\frac{1}{2}} [\Lambda_Y + \mathbf{I}]^{-1} \Lambda_Y^{\frac{1}{2}} \mathbf{V}_Y^H \mathbf{U}_X \\ &\times (\Lambda_X + \mathbf{I})^{-\frac{1}{2}} \Lambda_X^{\frac{1}{2}} \mathbf{V}_X^H. \end{aligned} \quad (27)$$

Remark 1. Representing the problem in terms of these new variables decreases the dependency between the objective function and constraints and thus facilitates the

optimization procedure. For example, \mathbf{U}_Y is only involved in objective function, but not in the constraints. Hence, we can adjust \mathbf{U}_Y to change the objective without affecting the constraint. Reversely, \mathbf{V}_X and \mathbf{V}_Y only affect the constraint, but not the objective function.

Remark 2. On the other hand, the diagonality constraint in (P2) can be satisfied by only adjusting \mathbf{V}_X so that the other constraint and the objective functions will not be affected.

3.3 Simplification

In the following subsections, we derive the optimum structures for $\mathbf{V}_X, \mathbf{V}_Y, \mathbf{U}_X$, and \mathbf{U}_Y , respectively.

3.3.1 Optimal \mathbf{V}_X and \mathbf{V}_Y

The functionality of the optimal \mathbf{V}_X is to make \mathbf{G} diagonal. Moreover, if there is a \mathbf{V}_Y such that all components in $\text{diag}\{\mathbf{I} - \mathbf{G}\}$ are minimized simultaneously, then this \mathbf{V}_Y must be optimal because it provides the largest possible freedom to optimize over the rest variables, i.e., $\mathbf{U}_X, \mathbf{U}_Y, \Lambda_X$, and Λ_Y . Applying (9.H.2 in [12]) on \mathbf{G} , we obtain

$$\lambda\{\mathbf{G}\} \prec_w \text{diag}\{\hat{\mathbf{G}}\}, \quad (28)$$

where

$$\hat{\mathbf{G}} = \Lambda_X^{\frac{1}{2}}(\Lambda_X + \mathbf{I})^{-\frac{1}{2}} \Lambda_Y^{\frac{1}{2}}[\Lambda_Y + \mathbf{I}]^{-1} \Lambda_Y^{\frac{1}{2}}(\Lambda_X + \mathbf{I})^{-\frac{1}{2}} \Lambda_X^{\frac{1}{2}} \quad (29)$$

is the multiplication of the diagonal matrices in (27). According to the definition of submajorization and supermajorization in Appendix 1, (28) can also be written as follows:

$$-\lambda\{\mathbf{G}\} \prec^w -\text{diag}\{\hat{\mathbf{G}}\} \quad (30)$$

and is equivalent to

$$\mathbf{1} - \lambda\{\mathbf{G}\} \prec^w \mathbf{1} - \text{diag}\{\hat{\mathbf{G}}\}, \quad (31)$$

where $\mathbf{1}$ is an all one vector. Moreover, (31) can be re-expressed as follows:

$$\lambda\{\mathbf{I} - \mathbf{G}\} \prec^w \text{diag}\{\mathbf{I} - \hat{\mathbf{G}}\}. \quad (32)$$

Since one can always find \mathbf{V}_X that makes \mathbf{G} diagonal, there is $\lambda\{\mathbf{I} - \mathbf{G}\} = \text{diag}\{\mathbf{I} - \mathbf{G}\}$. Moreover, from Lemma 2, we can rewrite (32) as follows:

$$\text{diag}\{\mathbf{I} - \mathbf{G}\} \prec^w \text{diag}\{\mathbf{I} - \hat{\mathbf{G}}\}, \quad (33)$$

which indicates that $\text{diag}\{\mathbf{I} - \hat{\mathbf{G}}\}$ is a simultaneous lower bound for all the elements in $\text{diag}\{\mathbf{I} - \mathbf{G}\}$. Therefore, $\mathbf{G} = \hat{\mathbf{G}}$ must hold at the optimal point, and this can be achieved if $\mathbf{V}_Y = \mathbf{U}_X$ and $\mathbf{V}_X = \mathbf{I}$.

3.3.2 Optimal \mathbf{U}_X and \mathbf{U}_Y

With the optimal \mathbf{V}_Y , the constraint is independent from both \mathbf{U}_Y and \mathbf{U}_X . Therefore, one can find the optimal structure of \mathbf{U}_X and \mathbf{U}_Y purely from

$$\begin{aligned} \min_{\mathbf{U}_X, \mathbf{U}_Y} \quad & \text{tr}(\Lambda_s^\dagger \mathbf{U}_{s,1}^H \mathbf{U}_X \Lambda_X \mathbf{U}_X^H \mathbf{U}_{s,1} \Lambda_s^\dagger) \\ & + \text{tr}(\Lambda_r^\dagger \mathbf{U}_{r,1}^H \mathbf{U}_Y \Lambda_Y \mathbf{U}_Y^H \mathbf{U}_{r,1} \Lambda_r^\dagger). \end{aligned} \quad (34)$$

We need the following matrix inequality (9.H.1.h in [12]) to proceed: Given two $L \times L$ positive semi-definite matrices \mathbf{A}_1 and \mathbf{A}_2 with eigenvalues $\lambda_l(\mathbf{A}_1)$ and $\lambda_l(\mathbf{A}_2)$ arranged in the increasing order, there is

$$\sum_{l=1}^L \lambda_l(\mathbf{A}_1) \lambda_{L-l+1}(\mathbf{A}_2) \leq \text{tr}(\mathbf{A}_1 \mathbf{A}_2) \leq \sum_{l=1}^L \lambda_l(\mathbf{A}_1) \lambda_l(\mathbf{A}_2). \quad (35)$$

Then, the first part in (34) can be lower bounded by the following:

$$\text{tr}(\Lambda_s^\dagger \Lambda_X \Lambda_s^\dagger) \leq \text{tr}(\Lambda_s^\dagger \mathbf{U}_{s,1}^H \mathbf{U}_X \Lambda_X \mathbf{U}_X^H \mathbf{U}_{s,1} \Lambda_s^\dagger). \quad (36)$$

Obviously, the minimum value can be achieved when $\mathbf{U}_X = \mathbf{U}_s$, since the diagonal elements of Λ_s and Λ_X are all arranged in increasing order. Similar discussion holds for the second term in (34), namely

$$\text{tr}(\Lambda_r^\dagger \Lambda_Y \Lambda_r^\dagger) \leq \text{tr}(\Lambda_r^\dagger \mathbf{U}_{r,1}^H \mathbf{U}_Y \Lambda_Y \mathbf{U}_Y^H \mathbf{U}_{r,1} \Lambda_r^\dagger), \quad (37)$$

and the minimum is achieved when $\mathbf{U}_Y = \mathbf{U}_{r,1}$.

Substituting the optimal $\mathbf{U}_X, \mathbf{U}_Y, \mathbf{V}_X$, and \mathbf{V}_Y into the objective function and the constraints of (P2), we obtain the following:

$$\begin{aligned} \min_{\Lambda_X, \Lambda_Y} \quad & \text{tr}(\Lambda_s^\dagger \Lambda_X \Lambda_s^\dagger) + \text{tr}(\Lambda_r^\dagger \Lambda_Y \Lambda_r^\dagger) \\ \text{subject to} \quad & \mathbf{I} - \Lambda_X \Lambda_Y (\Lambda_X + \mathbf{I})^{-1} (\Lambda_Y + \mathbf{I})^{-1} \succ^w \rho. \end{aligned} \quad (38)$$

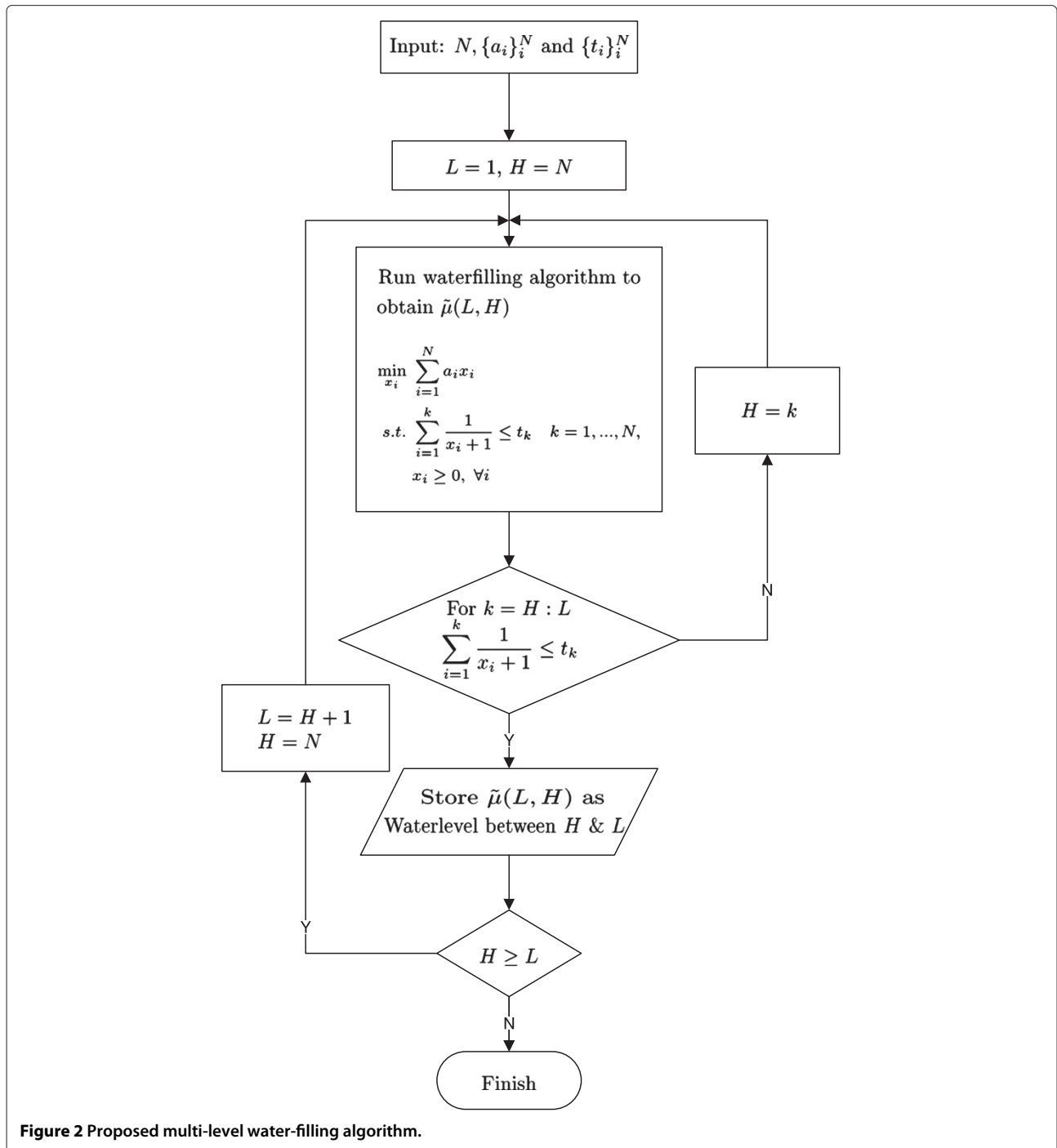
Define a_i and b_i as the i th diagonal entries of $(\Lambda_s^\dagger)^2$ and $(\Lambda_r^\dagger)^2$, respectively. Further, define x_i and y_i as the i th diagonal entries of Λ_X and Λ_Y , respectively. Then, problem (P2) is converted to a scalar form:

$$(P3): \quad \min_{x_i, y_i} \sum_{i=1}^L a_i x_i + b_i y_i \quad (39)$$

$$\text{subject to} \quad \sum_{i=1}^k \frac{y_i + x_i + 1}{y_i + x_i + y_i x_i + 1} \leq \sum_{i=1}^k \rho_i \quad k=1, \dots, L, \quad (40)$$

$$x_i \geq 0, y_i \geq 0, \quad \forall i.$$

Unfortunately, the constraint (40) is non-convex, and the problem cannot be solved efficiently. We then propose



the following two convex bounds for each summand on the left-hand side of (40):

$$\frac{y_i + x_i + 1}{y_i + x_i + y_i x_i + 1} \leq \frac{y_i + x_i + 2}{y_i + x_i + y_i x_i + 1}, \quad (41)$$

$$\frac{y_i + x_i + 1}{y_i + x_i + y_i x_i + 1} \geq \frac{y_i + x_i}{y_i + x_i + y_i x_i}. \quad (42)$$

Replacing the corresponding term in (40) by the right-hand side (RHS) in (41) or (42) while keeping the same objective function, we can obtain the upper or the lower bound of the original problem, respectively. Both the lower bound and the upper bound problems can be solved by the existing convex optimization tools based on the interior point method, e.g., CVX, [20,23].

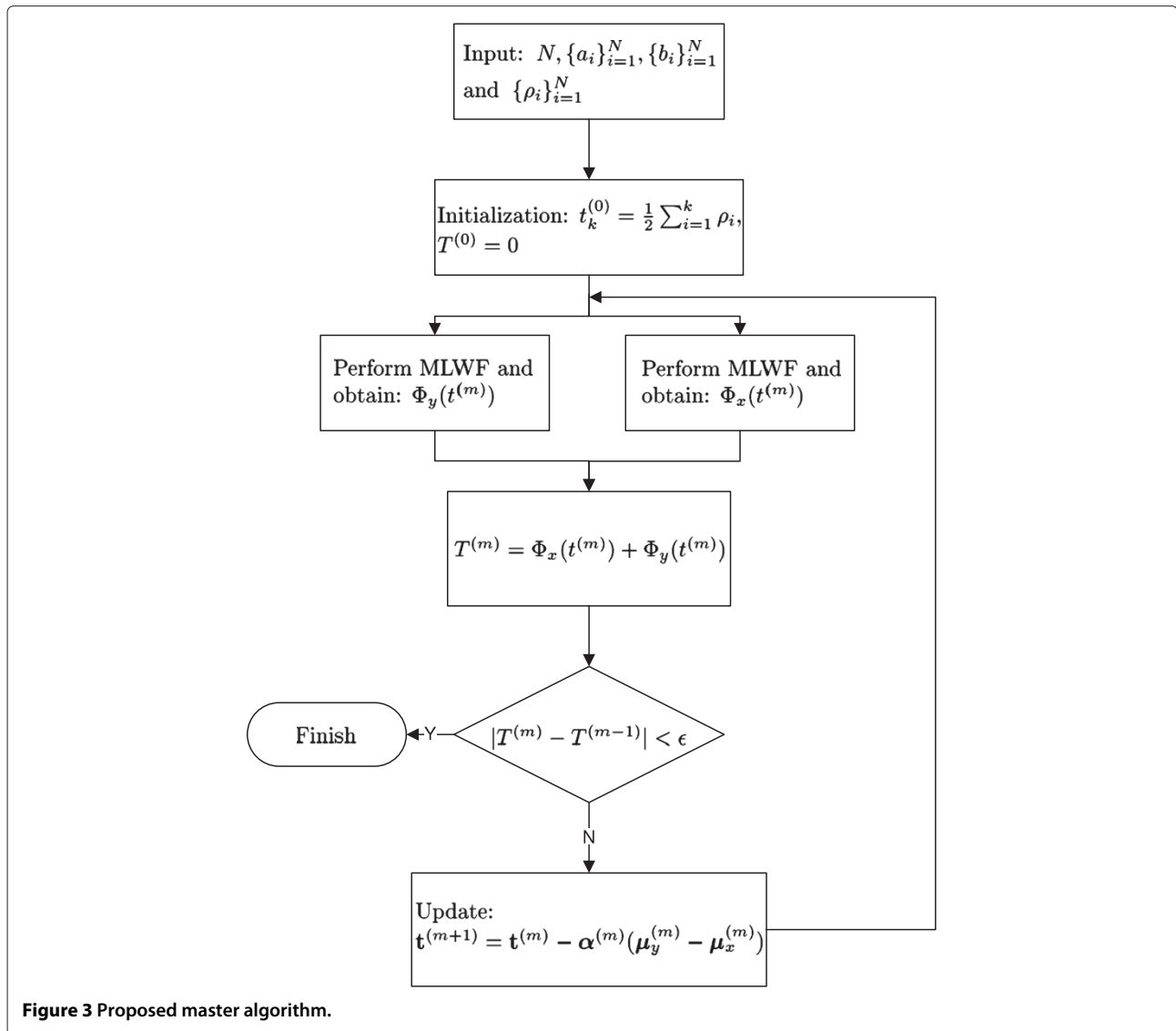


Figure 3 Proposed master algorithm.

4 Algorithm for the upper bound problem

Power minimization is ensured by considering the minimization of the upper bound of the objective value. Therefore, we take a detailed look into the upper bound problem and target at a more efficient solution. The upper bound problem is restated as follows:

$$\begin{aligned}
 & \min_{x_i, y_i} \sum_{i=1}^N a_i x_i + b_i y_i & (43) \\
 & \text{subject to } \sum_{i=1}^k \frac{1}{x_i + 1} + \frac{1}{y_i + 1} \leq \sum_{i=1}^k \rho_i \quad k = 1, \dots, N, \\
 & \quad x_i \geq 0, y_i \geq 0, \quad \forall i.
 \end{aligned}$$

It is hard to obtain the closed form solution even with the Karush-Kuhn-Tucker (KKT) conditions. Observing

the symmetry, we can apply the primal decomposition method [23] to break down the problem into two simpler subproblems.

4.1 The decomposition method

Problem (43) can be decomposed into two parallel subproblems together with a master problem connecting as a bridge [23]. Defining a new auxiliary vector $\mathbf{t} = [t_1, t_2, \dots, t_N]^T$, the subproblems are as follows:

$$\begin{aligned}
 \Phi_x(\mathbf{t}) : & \min_{x_i} \sum_{i=1}^N a_i x_i \\
 & \text{subject to } \sum_{i=1}^k \frac{1}{x_i + 1} \leq t_k \quad k = 1, \dots, N, \\
 & \quad x_i \geq 0, \quad \forall i
 \end{aligned}$$

$$\begin{aligned} \Phi_y(\mathbf{t}) : \min_{y_i} & \sum_{i=1}^N b_i y_i \\ \text{subject to} & \sum_{i=1}^k \frac{1}{y_i + 1} \leq \sum_{i=1}^k \rho_i - t_k \quad k = 1, \dots, N, \\ & y_i \geq 0, \forall i \end{aligned}$$

and the master problem is

$$\min_{\mathbf{t}} \Phi_x(\mathbf{t}) + \Phi_y(\mathbf{t}). \quad (44)$$

In many cases, the subproblems do not have closed-form solution. Therefore, searching the optimum \mathbf{t} for the master problem can be iteratively done from the subgradient method. Note that the convergence of the decomposed problem to the global optimal point can be guaranteed since (43) is convex.

4.2 Solving the subproblems

The Lagrangian corresponding to $\Phi_x(\mathbf{t})$ is as follows:

$$L = \sum_{k=1}^N a_k x_k + \sum_{k=1}^N \mu_k \left(\sum_{i=1}^k \frac{1}{1+x_i} - t_k \right) - \sum_{k=1}^N \gamma_k x_k, \quad (45)$$

and the corresponding KKT conditions are the following:

$$a_k - \sum_{i=k}^N \mu_i \frac{1}{(1+x_k)^2} - \gamma_k = 0, \quad (46)$$

$$\mu_k \left(\sum_{i=1}^k \frac{1}{1+x_i} - t_k \right) = 0, \quad (47)$$

$$\sum_{i=1}^k \frac{1}{1+x_i} - t_k \leq 0, \quad (48)$$

$$\gamma_k x_k = 0, \quad (49)$$

$$x_k \geq 0, \mu_k \geq 0, \gamma_k \geq 0. \quad (50)$$

For simplicity, we define $\tilde{\mu}_k = \sum_{i=k}^N \mu_i$. Multiplying both sides of (46) by x_k and combining with (49) give the following:

$$x_k a_k = \tilde{\mu}_k \frac{x_k}{(1+x_k)^2} \implies x_k (a_k (1+x_k)^2 - \tilde{\mu}_k) = 0. \quad (51)$$

After a straightforward calculation, we obtain the following:

$$x_k = \begin{cases} (\tilde{\mu}_k^{1/2} a_k^{-1/2} - 1)^+ & \tilde{\mu}_k \geq a_k \\ 0 & \tilde{\mu}_k < a_k \end{cases}. \quad (52)$$

Based on (52), the KKT conditions are reduced to the following:

$$x_k = \begin{cases} (\tilde{\mu}_k^{1/2} a_k^{-1/2} - 1)^+ & \tilde{\mu}_k \geq a_k \\ 0 & \tilde{\mu}_k < a_k, \end{cases} \quad (53)$$

$$\tilde{\mu}_k \leq \tilde{\mu}_{k-1}, \quad (54)$$

$$\sum_{i=1}^k \frac{1}{1+x_i} - t_k \leq 0. \quad (55)$$

Next, we propose an efficient algorithm that could find the solution that satisfies (53), (54), and (55) simultaneously.

Algorithm 1 Multi-level water-filling algorithm

Input: Number of all positive eigenvalues, all inverse eigenvalues $\{a_i\}_i^N$, and a positive feasible initial vector for $\mathbf{t} = [t_1, \dots, t_N]^T$.

Output: Allocated powers $\{x_i\}_i^N$ and water levels $\{\mu_i\}_i^N$.

- 1: Set N equal to the number of variables.
 - 2: Set $H = N$ and $L = 1$ as the higher and the lower indices of the condition that we consider in the next step.
 - 3: Use standard water-filling algorithm to solve the problem subjected to only one condition, $\sum_{l=L}^H \frac{1}{x_l+1} \leq \sum_{l=L}^H \rho_l$, in which $\sum_{l=1}^k \rho_l = t_k, k = 1, \dots, L$. Denote the water level from the standard water-filling algorithm at this step as $\tilde{\mu}(L, H)$.
 - 4: Check if for each k ($L \leq k \leq H$), the conditions (55) are satisfied, starting from $k = H$ to $k = L$.
 - 5: If a condition in (55) is not satisfied for a certain l_0 ($L \leq l_0 \leq H$), set $H = l_0$ then go to step 3. If (55) is satisfied for all $L \leq k \leq H$, then $\tilde{\mu}(L, H)$ is the water level for all these indices.
 - 6: If $H \geq L$, put $H = N$ and $L = H + 1$ and go to step 3, or else, finish the algorithm.
-

The standard water-filling algorithm embedded in step 3 can be found in Appendix 2. For an explicit demonstration, we present the algorithm in a diagram in Figure 2.

We next need to prove the optimality of the above MLWF. Since the problem is convex, the output of the algorithm is the optimal solution if and only if all of the KKT conditions are satisfied. The conditions (55) are simultaneously satisfied from the step 4 in the algorithm. Condition (53) is itself satisfied by the nature of the water-filling algorithm, that is the water levels are always non-negative. The above MLWF algorithm satisfies (54) as well, as seen from the following lemma.

Lemma 1. *Successive water levels achieved by the proposed MLWF algorithm are ordered decreasingly.*

Proof. The algorithm has two loops: one is the inner loop ‘steps 3 → 4 → 5 → 3’ and the other is the outer loop ‘steps 3 → 4 → 5 → 6 → 3’. Each time when the inner loop finishes, one water level $\tilde{\mu}(L, H)$ will be achieved. Then, we proceed to compute other water levels. In the inner loop, we first adopt the water level given by standard water-filling algorithm. Our hypothesis is that the aforementioned water level satisfies all of the conditions (55). Then, we check our hypothesis by searching whether there is a $k = l_0$ for which the corresponding condition in (55) is violated.

Assume l_0 with $1 \leq L < l_0 < H \leq N$ is the point at which the inner loop finishes, i.e., $\tilde{\mu}(L, l_0)$ satisfies all the conditions from L to l_0 . At this point, $\sum_{i=L}^{l_0} \frac{1}{x_i+1} = \sum_{i=L}^{l_0} \rho_i$ is true due to the fact that the constraint holds with equality in the standard water-filling algorithm. Moreover, $\tilde{\mu}(L, H) < \tilde{\mu}(L, l_0)$ holds since $\tilde{\mu}(L, H)$ is not large enough to satisfy (55) with $k = l_0$, though it satisfies all of the conditions from $k = H$ to $k = l_0 + 1$. On the other hand, we have

$$\tilde{\mu}(L, l) < \tilde{\mu}(L, H), \quad l_0 + 1 \leq l \leq H, \quad (56)$$

since, based on our assumption, conditions (55) with index greater than l_0 are satisfied with the water level introduced by $\tilde{\mu}(L, H)$. From (56) and $\tilde{\mu}(L, H) < \tilde{\mu}(L, l_0)$, we conclude that

$$\tilde{\mu}(L, l) < \tilde{\mu}(L, l_0), \quad l_0 + 1 \leq l \leq H. \quad (57)$$

The search for the next water level (of course in the following inner loop), is between $H = N$ and $L = l_0 + 1$. We then only need to prove $\tilde{\mu}(l, l_0 + 1) < \tilde{\mu}(l, l_0)$, $l_0 + 1 \leq l \leq H$. First, we know $\tilde{\mu}(L, l)$ satisfies

$$\sum_{i=L}^l \frac{1}{x_i+1} = \sum_{i=L}^l \rho_i. \quad (58)$$

From the last inner loop, we know that applying $\tilde{\mu}(L, l)$ on the condition $\sum_{i=L}^{l_0} \frac{1}{x_i+1} \leq \sum_{i=L}^{l_0} \rho_i$, does not work, i. e.,

$$\sum_{i=1}^{l_0} \frac{1}{x_i+1} > \sum_{i=1}^{l_0} \rho_i. \quad (59)$$

Therefore, using $\tilde{\mu}(L, l)$ in both (58) and (59) yields $\sum_{i=l_0+1}^l \frac{1}{x_i+1} < \sum_{i=l_0+1}^l \rho_i$. On the other hand, $\tilde{\mu}(l_0+1, l)$ gives $\sum_{i=l_0+1}^l \frac{1}{x_i+1} = \sum_{i=l_0+1}^l \rho_i$. Therefore, we infer $\tilde{\mu}(l_0+1, l) < \tilde{\mu}(L, l)$. Together with (57), we conclude $\tilde{\mu}(l_0+1, l) < \tilde{\mu}(L, l_0)$. \square

Meanwhile, the subproblem $\Phi_y(\mathbf{t})$ can be solved with the same MLWF, which will not be restated here.

Algorithm 2 The master algorithm

Input: The number of all positive eigenvalues (N), all inverted eigenvalues $\{a_i\}_{i=1}^N$ and $\{b_i\}_{i=1}^N$, and all power $\{\rho_i\}_{i=1}^N$.

Output: Allocated powers $\{x_i\}_{i=1}^N$ and $\{y_i\}_{i=1}^N$, and water levels $\{\mu_{i,x}\}_{i=1}^N$ and $\{\mu_{i,y}\}_{i=1}^N$ obtained from MLWF.

- 1: Choose a positive increasing vector for \mathbf{t} , e.g., $t_k = \frac{1}{2} \sum_{i=1}^k \rho_i$. Set the initial objective value $T^{(0)} = 0$;
- 2: Obtain the solution and water levels of $\Phi_x(\mathbf{t})$ and $\Phi_y(\mathbf{t})$ from MLWF algorithm;
- 3: Calculate the objective function in the m th iteration as $T^{(m)} = \Phi_x(\mathbf{t}) + \Phi_y(\mathbf{t})$. If $|T^{(m)} - T^{(m-1)}| < \epsilon$, then terminate the algorithm; otherwise, go to the next step.
- 4: Use subgradient of $\Phi_x(\mathbf{t}) + \Phi_y(\mathbf{t})$ to update \mathbf{t} as $\mathbf{t}^{(m+1)} = \mathbf{t}^{(m)} - \alpha^{(m)}(\boldsymbol{\mu}_y^{(m)} - \boldsymbol{\mu}_x^{(m)})$, where $\alpha^{(m)}$ is the step size of the m th iteration, while $\boldsymbol{\mu}_y^{(m)} = [\mu_{1,y}, \dots, \mu_{N,y}]^T$ and $\boldsymbol{\mu}_x^{(m)} = [\mu_{1,x}, \dots, \mu_{N,x}]^T$. Go to step 2.

This algorithm is presented by a diagram in Figure 3.

5 Simulation results

In this section, we numerically examine the performance of the proposed method. For all examples, we assume that the channel matrices (\mathbf{H}_r and \mathbf{H}_s) have independent and identically distributed Gaussian entries with zero mean and variance 1. We consider $6 \times 6 \times 6$ MIMO relay system, and 1000 Monte-Carlo runs are taken for average. The upper bound optimization will be solved both by the proposed decomposition algorithms and by the CVX convex optimization toolbox [20], while the lower-bounded optimization will only be solved by CVX. Moreover, we also compare the proposed solutions with a suboptimal method and a trivial method listed in the following:

1. *Diagonalization method.* We reduce the problem into a scalar problem using SVD of channel matrices. If we consider the structure of $\mathbf{B} = \mathbf{U}_s \boldsymbol{\Lambda}_B$ and $\mathbf{F} = \mathbf{U}_r \boldsymbol{\Lambda}_F \mathbf{V}_s^H$, we can reduce the problem into a simple scalar problem. In this method we basically choose the precoding matrices to match the channel matrices. The optimization problem of (10) is simplified to the following:

$$(P4): \quad \min_{x_i, y_i} \sum_{i=1}^L a_i x_i + b_i y_i \quad (60)$$

subject to $\frac{y_i + x_i + 1}{y_i + x_i + y_i x_i + 1} \leq \rho_i \quad \forall i$

$x_i \geq 0, y_i \geq 0, \quad \forall i,$

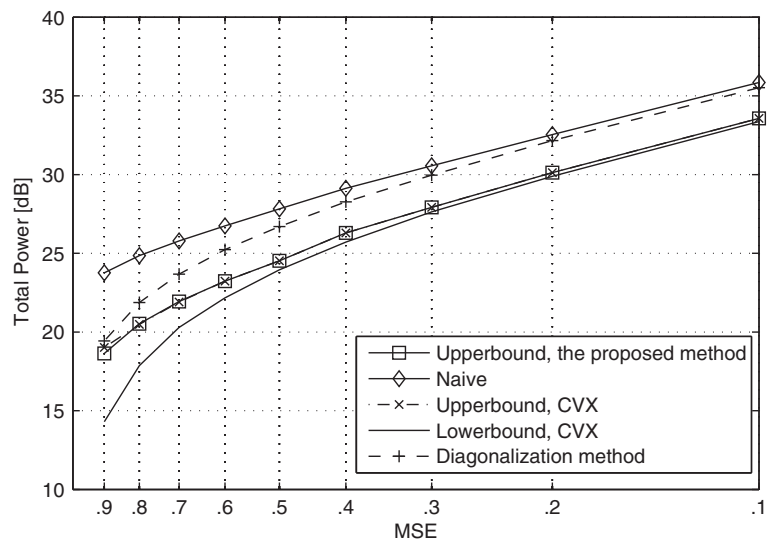


Figure 4 Performance with equal QoS constraint.

where $a_i, b_i, x_i,$ and y_i are the i th diagonal entries of $(\Lambda_s^\dagger)^2, (\Lambda_r^\dagger)^2, \Lambda_B,$ and $\Lambda_F,$ respectively. Note that the problem in (60) is a special case of (P3) in (39).

2. *Naive method.* We consider a solution that satisfies all of the constraints in (43) with equality. In this method, x and y are given by the following:

$$x = y = 2/\rho - 1.$$

In the first example, we take equal QoS requirements for all data streams. The total consumed powers versus MSE are depicted in Figure 4 for the upper bound with

the proposed method, the upper bound with CVX, the lower bound, and the suboptimal diagonalization method with CVX, as well as the naive method. It can be observed that the performance of the decomposition method is exactly the same as the one obtained from CVX. We can also observe that the lower and the upper bounds are quite tight at high SNR when they meet each other. The numerical results from the diagonalization method depicts weaker performance than the proposed method with roughly 2 dB more power consumption. The naive method, on the other hand, has the poorest performance and requires more power than the proposed method (roughly 3 dB).

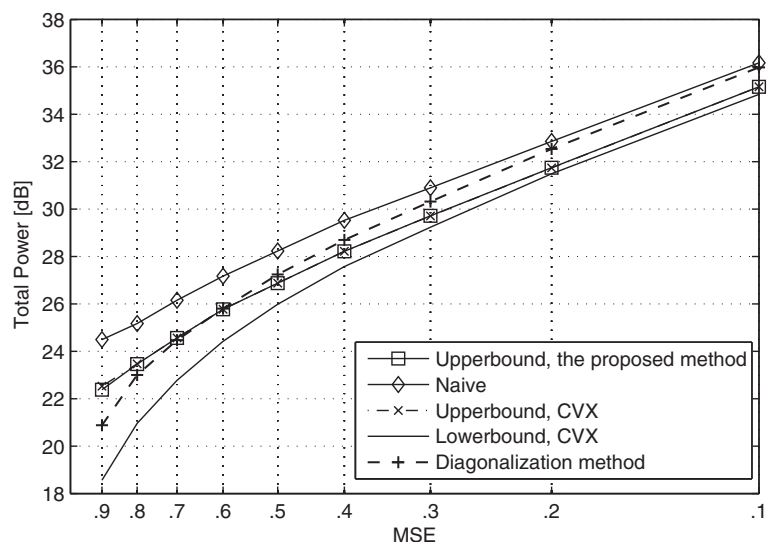


Figure 5 Performance with unequal QoS constraint $3\rho_1 = 2\rho_2 = \frac{3}{2}\rho_3 = \frac{6}{5}\rho_4 = \rho_5 = \rho_6.$

Table 1 Average time spent by each approach

Case	Proposed decomposition (s)	CVX (s)
Equal MSE	0.4377	3.2094
Unequal MSE	0.2237	2.6747

The performances with unequal QoS constraints are presented in Figure 5, where the QoS vector is chosen as $\rho = \rho[1, 1.5, 2, 2.5, 3, 3]^T$. One may notice that all curves are about 2 dB higher than those in Figure 4.

This is because the smallest QoS constraint drags down the overall performance and thus demands for more power. Once again, we find that the upper bound and the lower bound are quite close to each other especially at high SNR values. Meanwhile, the proposed method coincides with the upper bound from the CVX. Moreover, the advantage of the proposed method over the suboptimal diagonalization method and the naive method is fairly explicit.

We also measure the average running time for the proposed method and the upper bound with CVX, measured in MATLAB. The results are shown in Table 1.

Remark 3. Note that resorting to simulation cannot serve as a rigorous measurement of the complexity analysis. Nevertheless, the large difference in the running time fairly indicates the much higher efficiency of the proposed decomposition method over the regular convex optimization tool.

The effect of the number of antennas on each node is also investigated here. We set the QoS requirements for all

data streams as 0.2 and vary the number of antennas from 2 to 9. The results are displayed in Figure 6.

6 Conclusion

In this paper, we considered the joint design of source precoding matrix and relay precoding matrix in a two-hop AF MIMO relay network. We minimized the power consumption subject to a set of predefined QoS constraints of each data stream. Using matrix calculus and majorization theory, we simplified the original matrix-valued problem to a relatively simpler scalar one and proposed two bounding problems that are convex and can be solved efficiently. We specifically designed a primary decomposition method to solve the upper bound problem that has less complexity than directly applying the interior point method. Numerical examples are provided to corroborate the proposed studies.

Appendix 1

Basics of majorization theory

Here, we briefly introduce the basics of majorization theory while the more comprehensive discussion can be found in [12].

Definition 1. If for any vector $x = [x_1, x_2, \dots, x_n] \in \mathbb{R}^n$, let

$$x_{(1)} \leq \dots \leq x_{(n)}$$

represents the elements of x in an increasing order. Similarly, assume

$$x_{[1]} \geq \dots \geq x_{[n]}$$

represents the elements of x in a decreasing order.

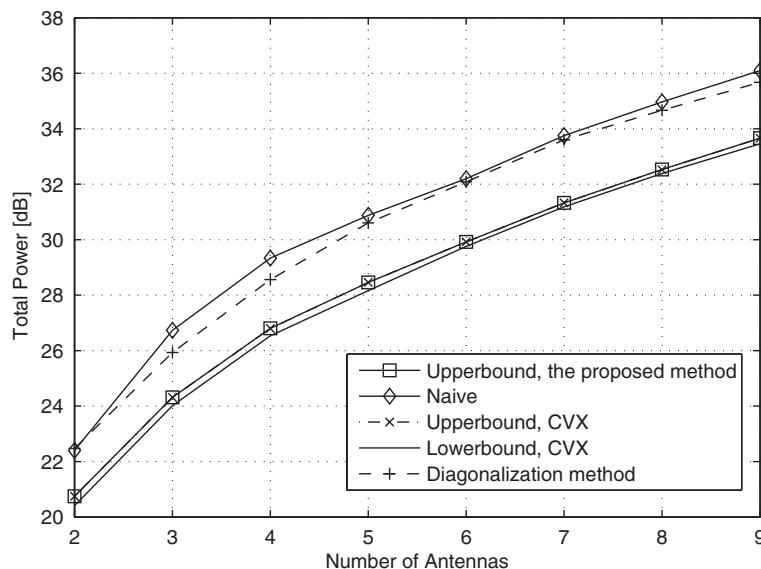


Figure 6 Performance of the system with different number of antennas. With fixed identical QoS for all of the data streams $\rho = 0.2$.

Definition 2. For any $x, y \in \mathbb{R}^n$, x is majorized by y if

$$\sum_{i=1}^k x_{(i)} \geq \sum_{i=1}^k y_{(i)}, \quad 1 \leq k \leq n-1 \quad (61)$$

$$\sum_{i=1}^n x_{(i)} = \sum_{i=1}^n y_{(i)}, \quad (62)$$

and is denoted by $x \prec y$ or $y \succ x$.

Definition 3. [12] For any $x, y \in \mathbb{R}^n$, x is weakly supermajorized by y if

$$\sum_{i=1}^k x_{(i)} \geq \sum_{i=1}^k y_{(i)}, \quad 1 \leq k \leq n. \quad (63)$$

We denote this with $x \prec^w y$ (or equivalently $y \succ^w x$). Also, x is weakly submajorized by y if

$$\sum_{i=1}^k x_{[i]} \leq \sum_{i=1}^k y_{[i]}, \quad 1 \leq k \leq n. \quad (64)$$

We denote this with $x \prec_w y$ (or equivalently $y \succ_w x$).

Lemma 2. (9.B.1 in [12]) Let M be an $n \times n$ Hermitian matrix with a vector $\text{diag}\{M\}$ denoting its diagonal elements and let vector $\lambda\{M\}$ contain its eigenvalues. Then,

$$\bar{\text{diag}}\{M\} \prec \text{diag}\{M\} \prec \lambda\{M\}$$

where $\bar{\text{diag}}\{M\}_i = \text{mean}(\text{diag}\{M\})$.

Reversely, given vectors a and b with $a \prec b$, then there exists an $n \times n$ Hermitian matrix M whose diagonal elements are a and eigenvalues are b .

Lemma 3. (5.A.9 in [12]) For any a and b satisfying $a \prec^w b$, there exists a vector x such that

$$x \prec b, \quad a \geq x.$$

Appendix 2

Standard water-filling algorithm

A standard water-filling algorithm will be used to solve the following convex problem which appeared in our MLWF algorithm:

$$\begin{aligned} \min_{x_i} \quad & \sum_{i=1}^N a_i x_i \\ \text{subject to} \quad & \sum_{i=1}^N \frac{1}{x_i + 1} \leq \rho_k, \quad x_i \geq 0, \quad \forall i. \end{aligned} \quad (65)$$

The water-filling algorithm that yields the optimum x_i 's is given by the following:

Input: The number of all positive eigenvalues N , all inverse eigenvalues $\{a_i\}_i^N$, and a positive feasible initial vector for ρ_i .

Output: Allocated powers $\{x_i\}_i^N$ and the optimum water level $\{\mu\}$.

- Sort all of the eigenvalues in the increasing order and $a_{N+1} = \infty$. Set $L = N$.
- If $a_L = a_{L+1}$, then $L = L - 1$. Set $\mu = a_N$.
- If $\mu \leq \frac{\sum_{i=1}^N a_i^{1/2}}{(\rho - (N-L))^2}$, then the optimum solution is $\mu = \frac{\sum_{i=1}^N a_i^{1/2}}{(\rho - (N-L))^2}$ and $x_i = \mu^{1/2} a_i^{-1/2} - 1$, or else, go to the next step.
- Set $L = L - 1$ and $\mu = a_L$ go back to step 3.

Endnotes

^aDual-hop relay network is of particular importance and has been the most frequently studied type in the past decades.

^bWe assume perfect channel knowledge in this paper, while the channel estimation can be performed by the method in [24].

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

This work was supported in part by the National Basic Research Program of China (973 Program) under grants 2013CB336600 and 2012CB316102, the National Natural Science Foundation of China under grant 61201187, and the Tsinghua University Initiative Scientific Research Program under grant 20121088074. Jafar Mohammadi's work was also supported by Helmholtz Research School on Security Technologies. Y. Rong's work was supported under Australian Research Council's *Discovery Projects* funding scheme (project numbers DP110100736, DP110102076).

Author details

¹School of Engineering and Sciences, Jacobs University Bremen, Bremen, 28759, Germany. ²Tsinghua National Laboratory for Information Science and Technology, Beijing, 100084, China. ³Electrical and Computer Engineering, Curtin University of Technology, Bentley, WA 6102, Australia. ⁴Department of Electronics Engineering, Shanghai Jiaotong University, Shanghai, 200240, China.

Received: 31 May 2012 Accepted: 25 March 2013

Published: 19 April 2013

References

- JN Laneman, GW Wornell, Distributed space time block coded protocols for exploiting cooperative diversity in wireless networks. *IEEE Trans. Inform. Theory*. **49**, 2415–2425 (2003)
- JN Laneman, DNC Tse, GW Wornell, Cooperative diversity in wireless networks: efficient protocols and outage behavior. *IEEE Trans. Inform. Theory*. **50**, 3062–3080 (2004)
- EC Van De Meulen, Three-terminal communication channels. *Adv. Appl. Prob.* **3**, 120–154 (1971)
- H Sato, Information transmission through a channel with relay, Technical Report B7-67, The Aloha System, University of Hawaii, Honolulu (1976)
- T Cover, A El Gamal, Capacity theorems for the relay channel. *IEEE Trans. Inform. Theory*. **25**, 572–584 (1979)
- IE Telatar, Capacity of multi-antenna Gaussian channels. *Eur. Trans. Telecom.* **10**, 585–595 (1999)
- X Tang, Y Hua, Optimal design of non-regenerative MIMO wireless relays. *IEEE Trans. Wireless Commun.* **6**(4), 1398–1407 (2007)
- I Hammerstrom, A Wittneben, Power allocation schemes for amplify-and-forward MIMO-OFDM relay links. *IEEE Trans. Wireless Commun.* **6**(8), 2798–2802 (2007)

9. S Yang, JC Belfiore, Towards the optimal amplify-and-forward cooperative diversity scheme. *IEEE Trans. Inform. Theory*. **53**(9), 3114–126 (2007)
10. B Khoshnevis, W Yu, R Adve, Grassmannian beamforming for MIMO amplify-and-forward relaying. *IEEE J. Sel. Areas Commun.* **26**, 1397–407 (2008)
11. Y Hua, in *Proceedings of IEEE Military Communications Conference*. An overview of beamforming and power allocation for MIMO relays, San Jose, CA, 2010), pp. 375–380
12. AW Marshall, I Olkin, *Inequalities: Theory of Majorization and Its Applications*. (Academic, New York, 1979)
13. Y Rong, X Tang, Y Hua, A unified framework for optimizing linear non-regenerative multicarrier MIMO relay communication systems. *IEEE Trans. Signal Process.* **57**, 4837–4851 (2009)
14. Y Rong, Y Hua, Optimality of diagonalization of multi-hop MIMO relays. *IEEE Trans. Wireless Commun.* **8**, 6068–6077 (2009)
15. DP Palomar, MA Lagunas, JM Cioffi, Optimum linear joint transmit-receive processing for MIMO channels with QoS constraints. *IEEE Trans. on Signal Processing*. **52**(5), 1179–1197 (2004)
16. R Zhang, CC Chai, Y Liang, Joint beamforming power control for multiantenna relay broadcast channel with QoS constraints. *IEEE Trans. Sig. Proc.* **57**(2), 726–737 (2009)
17. Y Rong, Multi-hop non-regenerative MIMO relays-QoS considerations. *IEEE Trans. Sig. Proc.* **59**(1), 290–303 (2011)
18. Y Fu, L Yang, W Zhu, C Liu, Optimum linear design of two-hop MIMO relay network with QoS requirements. *IEEE Trans. Sig. Proc.* **59**(5), 2257–2269 (2011)
19. J Mohammadi, F Gao, Y Rong, in *Proceedings of the Global Telecommunications Conference (GLOBECOM 2010)*. Design of amplify and forward MIMO relay networks with QoS constraint (Miami, USA, 2010), pp. 6–10
20. M Grant, S Boyd, CVX: Matlab software for disciplined convex programming, version 2.0 beta (2012). <http://cvxr.com/cvx>
21. SM Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. (Prentice-Hall, Englewood Cliffs, 1993)
22. KB Petersen, MS Pedersen, *The Matrix Cookbook* (2007). Accessed September 2007
23. S Boyd, L Vandenberghe, *Convex Optimization*. (Cambridge University Press, New York, 2004). Available (online) at <http://www.stanford.edu/~boyd/cvxbook.html>
24. F Gao, T Cui, A Nallanathan, On channel estimation and optimal training design for amplify and forward relay network. *IEEE Trans. Wireless Commun.* **7**(5), 1907–1916 (2008)

doi:10.1186/1687-1499-2013-108

Cite this article as: Mohammadi et al.: Joint source and relay design for two-hop amplify-and-forward relay networks with QoS constraints. *EURASIP Journal on Wireless Communications and Networking* 2013 **2013**:108.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com
