Improving Relay Matrices for MIMO Multi-Relay Communication Using Gradient Projection

Apriana Toding
Dept. Electrical Engineering.
Universitas Kristen Indonesia Paulus, Makassar, Sul-Sel 90245, Indonesia
Email: aprianatoding@ukipaulus.ac.id

***Abstract*—In this paper, we design the optimal relay matrices for multiple-input multiple-output (MIMO) relay communication systems with parallel relay nodes using the projected gradient (PG) approach. We show that the optimal relay amplifying matrices have a beamforming structure. Using the optimal structure, the relay power loading algorithm is developed to minimize the mean-squared error (MSE) of the signal waveform estimation at the destination node. Simulation result demonstrate the effectiveness of the proposed relay amplifying matrix with multiple parallel relay nodes using the PG approach in the system bit-error-rate performance.**

***Index Terms*—MIMO relay, parallel relay network, beamforming, non-regenerative relay, projected gradient.**

I. Introduction

In order to establish a reliable wireless communication link, one needs to compensate for the effects of signal fading and shadowing. An efficient way to address this issue is to transmit signals through one or more relays [1]. This can be accomplished via a wireless network consisting of geographically separated nodes.

When nodes in the relay system are installed with multiple antennas, we call such system multiple-input multiple-output (MIMO) relay communication system. Recently, MIMO relay communication systems have attracted much research interest and provided significant improvement in terms of both spectral efficiency and link reliability. In [3]-[6], the authors have studied the optimal relay amplifying matrix design for the source-relay-destination channel. In [3] and [4], the optimal relay amplifying matrix maximizing the mutual information (MI) between the source and destination nodes was derived assuming that the source covariance matrix is an identity matrix. In [5] and [6], the relay amplifying matrix was designed to minimize the mean-squared error (MSE) of the signal waveform estimation at the destination. In [7], the author investigated the joint source and relay optimization for MIMO relay networks using projected gradient (PG) approach. However, in [2]-[7], the authors investigated the optimal relay amplifying matrix design for two-hop MIMO relay networks with a single relay node. In [8], some linear relaying strategies are presented for multiple relays in MIMO relay networks by making use of local CSI. In [9], the authors investigated the optimal relay amplifying matrices for two-hop MIMO relay networks with multiple parallel relay nodes with sum relay power constraints at the output of the second hop channel.

In this paper, we propose the optimal relay matrices for MIMO relay communication systems with parallel relay nodes using the projected gradient (PG) approach which significantly reduces the computational complexity of the optimal design. We show that the optimal relay amplifying matrices have a beamforming structure. In addition to the PG approach, we constrain the power at each relay node which is more practical compared to the constraints in [9] as that constraint may exceed the available power budget at the relay nodes. Simulation result demonstrate the effectiveness of the proposed relay amplifying matrix with multiple parallel relay nodes using the PG approach in the system bit-error-rate performance.

The rest of this paper is organized as follows. In Section II, we introduce the system model of MIMO relay communication system with parallel relay nodes. The relay matrices design algorithm is developed in Section III. In Section IV, we show some numerical simulations. Conclusions are drawn in Section V.

II. System Model

Fig. 1 illustrates a two-hop MIMO relay communication system consisting of one source node, K parallel relay nodes, and one destination node. We assume that the source and destination nodes have N_s and N_d antennas, respectively, and each relay node has N_r antennas. The generalization to the system with different number of antennas at each node is straightforward. To efficiently exploit the system hardware, each relay node uses the same antennas to transmit and receive signals. Due to its merit of simplicity, we consider the amplify-and-forward scheme at each relay.

The communication process between the source and destination nodes is completed in two time slots. In the first time slot, the $N_s \times 1$ source signal vector $\mathbf{s}$ is transmitted to relay nodes. The received signal at the ith relay node can be written as

$$\mathbf{y}_{r,i} = \mathbf{H}_{sr,i}\mathbf{s} + \mathbf{v}_{r,i}, \qquad i = 1, \cdots, K \tag{1}$$

where $\mathbf{H}_{sr,i}$ is the $N_r \times N_s$ MIMO channel matrix between the source and the ith relay node, $\mathbf{y}_{r,i}$ and $\mathbf{v}_{r,i}$ are the received signal and the additive Gaussian noise vectors at the ith relay node, respectively.

In the second time slot, the source node is silent, while each relay node transmits the amplified signal vector to the destination node as

$$\mathbf{x}_{r,i} = \mathbf{F}_i \mathbf{y}_{r,i}, \qquad i = 1, \cdots, K \tag{2}$$

where $\mathbf{F}_i$ is the $N_r \times N_r$ amplifying matrix at the ith relay node. Thus the received signal vector at the destination node can be written as

$$\mathbf{y}_d = \sum_{i=1}^{K} \mathbf{H}_{rd,i} \mathbf{x}_{r,i} + \mathbf{v}_d \tag{3}$$

where $\mathbf{H}_{rd,i}$ is the $N_d \times N_r$ MIMO channel matrix between the ith relay and the destination node, $\mathbf{y}_d$ and $\mathbf{v}_d$ are the received signal and the additive Gaussian noise vectors at the destination node, respectively. Substituting (1)-(2) into (3), we have

$$\begin{aligned} \mathbf{y}_d &= \sum_{i=1}^{K} \left(\mathbf{H}_{rd,i} \mathbf{F}_i \mathbf{H}_{sr,i} \mathbf{s} + \mathbf{H}_{rd,i} \mathbf{F}_i \mathbf{v}_{r,i} \right) + \mathbf{v}_d \\ &= \mathbf{H}_{rd} \mathbf{F} \mathbf{H}_{sr} \mathbf{s} + \mathbf{H}_{rd} \mathbf{F} \mathbf{v}_r + \mathbf{v}_d = \tilde{\mathbf{H}} \mathbf{s} + \tilde{\mathbf{v}} \end{aligned} \tag{4}$$

where $\mathbf{H}_{sr} \triangleq [\mathbf{H}_{sr,1}^T, \mathbf{H}_{sr,2}^T, \cdots, \mathbf{H}_{sr,K}^T]^T$ is a $KN_r \times N_s$ channel matrix between the source node and all relay nodes, $\mathbf{H}_{rd} \triangleq [\mathbf{H}_{rd,1}, \mathbf{H}_{rd,2}, \cdots, \mathbf{H}_{rd,K}]$ is an $N_d \times KN_r$ channel matrix between all relay nodes and the destination node, $\mathbf{F} \triangleq \mathrm{bd}[\mathbf{F}_1, \mathbf{F}_2, \cdots, \mathbf{F}_K]$ is the $KN_r \times KN_r$ block diagonal equivalent relay matrix, $\mathbf{v}_r \triangleq \left[\mathbf{v}_{r,1}^T, \mathbf{v}_{r,2}^T, \cdots, \mathbf{v}_{r,K}^T\right]^T$ is obtained by stacking the noise vectors at all the relays, $\tilde{\mathbf{H}} \triangleq \mathbf{H}_{rd} \mathbf{F} \mathbf{H}_{sr}$ as the effective MIMO channel matrix of the source-relay-destination link, and $\tilde{\mathbf{v}} \triangleq \mathbf{H}_{rd} \mathbf{F} \mathbf{v}_r + \mathbf{v}_d$ as the equivalent noise vector. Here $(\cdot)^T$ denotes the matrix (vector) transpose, and $\mathrm{bd}[\cdot]$ stands for a block-diagonal matrix. We assume that all noises are independent and identically distributed (i.i.d.) Gaussian noise with zero mean and unit variance. The transmission power consumed by each relay node (2) can be expressed as

$$\mathrm{E}[\mathrm{tr}(\mathbf{x}_{r,i}\mathbf{x}_{r,i}^H)] = \mathrm{tr}\left(\mathbf{F}_i \left[\mathbf{H}_{sr,i}\mathbf{H}_{sr,i}^H + \mathbf{I}_{N_r}\right]\mathbf{F}_i^H\right), i = 1, \cdots, K \tag{5}$$

where $\mathrm{E}[\cdot]$ denotes statistical expectation, $\mathrm{tr}(\cdot)$ stands for the matrix trace, and $(\cdot)^H$ denotes the matrix (vector) Hermitian transpose.

Using a linear receiver, the estimated signal waveform vector at the destination node is given by $\hat{\mathbf{s}} = \mathbf{W}^H \mathbf{y}_d$, where $\mathbf{W}$ is an $N_d \times N_s$ weight matrix. The minimal MSE (MMSE) approach tries to find a weight matrix $\mathbf{W}$ that minimizes the statistical expectation of the signal waveform estimation given by

$$\begin{aligned} \mathrm{MSE} &= \mathrm{tr}\left(\mathrm{E}\left[\left(\hat{\mathbf{s}} - \mathbf{s}\right)\left(\hat{\mathbf{s}} - \mathbf{s}\right)^H\right]\right) \\ &= \mathrm{tr}\left(\left(\mathbf{W}^H\tilde{\mathbf{H}} - \mathbf{I}_{N_s}\right)\left(\mathbf{W}^H\tilde{\mathbf{H}} - \mathbf{I}_{N_s}\right)^H + \mathbf{W}^H\tilde{\mathbf{C}}\mathbf{W}\right) \end{aligned} \tag{6}$$

where $\tilde{\mathbf{C}}$ is the equivalent noise covariance matrix given by $\tilde{\mathbf{C}} = \mathrm{E}\left[\tilde{\mathbf{v}}\tilde{\mathbf{v}}^H\right] = \mathbf{H}_{rd}\mathbf{F}\mathbf{F}^H\mathbf{H}_{rd}^H + \mathbf{I}_{N_d}$. The weight matrix $\mathbf{W}$ which minimizes (6) is the Wiener filter and can be written as

$$\mathbf{W} = (\tilde{\mathbf{H}}\tilde{\mathbf{H}}^H + \tilde{\mathbf{C}})^{-1}\tilde{\mathbf{H}} \tag{7}$$

where $(\cdot)^{-1}$ denotes the matrix inversion. Substituting (7) back into (6), it can be seen that the MSE is a function of $\mathbf{F}$ can be written as

$$\mathrm{MSE} = \mathrm{tr}\left(\left[\mathbf{I}_{N_s} + \tilde{\mathbf{H}}^H\tilde{\mathbf{C}}^{-1}\tilde{\mathbf{H}}\right]^{-1}\right) \tag{8}$$

III. Minimal MSE Relay Design

In this section, we address the relay amplifying matrices optimization problem for systems with a linear receiver at the destination node. In particular, we show that the optimal relay matrices has a general beamforming structure. Base on (5) and (8), the relay amplifying matrices optimization problem can be formulated as

$$\min_{\{\mathbf{F}_i\}} \quad \mathrm{tr}\left(\left[\mathbf{I}_{N_s} + \tilde{\mathbf{H}}^H\tilde{\mathbf{C}}^{-1}\tilde{\mathbf{H}}\right]^{-1}\right) \tag{9}$$

$$\text{s.t.} \quad \mathrm{tr}\left(\mathbf{F}_i\left[\mathbf{H}_{sr,i}\mathbf{H}_{sr,i}^H + \mathbf{I}_{N_r}\right]\mathbf{F}_i^H\right) \leq P_{r,i}, i = 1, \cdots, K \tag{10}$$

where (10) is the power constraint at the relay node, and $P_{r,i} > 0$ is the corresponding power budget availabe at the ith relay.

A. Optimal Relay Design Using Projected Gradient (PG) Approach

Let us introduce the following singular value decompositions (SVD)

$$\mathbf{H}_{sr,i} = \mathbf{U}_{s,i}\mathbf{\Lambda}_{s,i}\mathbf{V}_{s,i}^H, \quad \mathbf{H}_{rd,i} = \mathbf{U}_{r,i}\mathbf{\Lambda}_{r,i}\mathbf{V}_{r,i}^H \tag{11}$$

where $\mathbf{\Lambda}_{s,i}$ and $\mathbf{\Lambda}_{r,i}$ are $R_s \times R_s$ and $R_r \times R_r$ diagonal matrix. Here $R_s \triangleq \mathrm{rank}(\mathbf{H}_{sr,i})$, $R_r \triangleq \mathrm{rank}(\mathbf{H}_{rd,i})$, $\mathrm{rank}(\cdot)$ denotes the rank of a matrix. The following theorem states the structure of the optimal $\mathbf{F}_i$.

THEOREM 1: The optimal structure of $\mathbf{F}_i$ as the solution to the problem (9)-(10) is given by

$$\mathbf{F}_i = \mathbf{V}_{r,1}\mathbf{A}_i\mathbf{U}_{s,1}^H, \quad i = 1, \cdots, K \tag{12}$$

where $\mathbf{A}_i$ is an $R \times R$ diagonal matrix and $R \triangleq \min(R_s, R_r)$..

PROOF:

Without loss of generality, $\mathbf{F}_i$ can be written as

$$\mathbf{F}_i = \left[\begin{array}{cc} \mathbf{V}_{r,1} & \mathbf{V}_{r,1}^{\perp} \end{array}\right] \left[\begin{array}{cc} \mathbf{A}_i & \mathbf{X}_i \\ \mathbf{Y}_i & \mathbf{Z}_i \end{array}\right] \left[\begin{array}{c} \mathbf{U}_{s,1}^H \\ (\mathbf{U}_{s,1}^{\perp})^H \end{array}\right] \quad i = 1, \cdots, K \tag{13}$$

where $\mathbf{V}_{r,1}^{\perp}(\mathbf{V}_{r,1}^{\perp})^H = \mathbf{I}_{N_r} - \mathbf{V}_{r,1}\mathbf{V}_{r,1}^H$, $\mathbf{U}_{s,1}^{\perp}(\mathbf{U}_{s,1}^{\perp})^H = \mathbf{I}_{N_r} - \mathbf{U}_{s,1}\mathbf{U}_{s,1}^H$, such that $[\mathbf{V}_{r,1}, \mathbf{V}_{r,1}^{\perp}]$ and $[\mathbf{U}_{s,1}, \mathbf{U}_{s,1}^{\perp}]$ are unitary matrices. The matrices $\mathbf{A}_i, \mathbf{X}_i, \mathbf{Y}_i, \mathbf{Z}_i$ are arbitrary matrices with dimensions of $R \times R$, $R \times (N_r - R)$, $(N_r - R) \times R$, $(N_r - R) \times (N_r - R)$, respectively. Substituting (13) back into (9), we obtain that $\mathbf{H}_{rd,i}\mathbf{F}_i\mathbf{H}_{sr,i} = \mathbf{U}_{r,i}\mathbf{\Lambda}_{r,i}\mathbf{A}_i\mathbf{\Lambda}_{s,i}\mathbf{V}_{s,i}^H$ and $\mathbf{H}_{rd,i}\mathbf{F}_i\mathbf{F}_i^H\mathbf{H}_{rd,i}^H = \sum_{i=1}^K \mathbf{U}_{r,i}\mathbf{\Lambda}_{r,i}(\mathbf{A}_i\mathbf{A}_i^H + \mathbf{X}_i\mathbf{X}_i^H)\mathbf{\Lambda}_{r,i}^H\mathbf{U}_{r,i}^H$. Thus we can rewrite equation (9) as

$$\mathrm{MSE} = \mathrm{tr}\left(\left[\mathbf{I}_{N_s} + \sum_{i=1}^{K}\mathbf{V}_{s,i}\mathbf{\Lambda}_{s,i}^H\mathbf{A}_i^H\mathbf{\Lambda}_{r,i}^H\mathbf{U}_{r,i}^H\left(\sum_{i=1}^{K}\mathbf{U}_{r,i}\mathbf{\Lambda}_{r,i}(\mathbf{A}_i\mathbf{A}_i^H + \mathbf{X}_i\mathbf{X}_i^H)\mathbf{\Lambda}_{r,i}^H\mathbf{U}_{r,i}^H + \mathbf{I}_{N_d}\right)^{-1}\sum_{i=1}^{K}\mathbf{U}_{r,i}\mathbf{\Lambda}_{r,i}\mathbf{A}_i\mathbf{\Lambda}_{s,i}\mathbf{V}_{s,i}^H\right]^{-1}\right)$$
$$i = 1, \cdots, K. \tag{14}$$

TABLE I
PROCEDURE OF APPLYING THE PROJECTED GRADIENT ALGORITHM TO SOLVE THE PROBLEM (15) - (16)

1) Initialize the algorithm at a feasible $\mathbf{A}_i^{(0)}$ for $i = 1, \cdots, K$; Set $n = 0$.
2) Compute the gradient of (15) $\nabla f(\mathbf{A}_i^{(n)})$;
   Project $\tilde{\mathbf{A}}_i^{(n)} = \mathbf{A}_i^{(n)} - s_n \nabla f(\mathbf{A}_i^{(n)})$ to obtain $\bar{\mathbf{A}}_i^{(n)}$.
   Update $\mathbf{A}_i$ with $\mathbf{A}_i^{(n+1)} = \mathbf{A}_i^{(n)} + \delta_n(\bar{\mathbf{A}}_i^{(n)} - \mathbf{A}_i^{(n)})$
3) if max abs $\|\mathbf{A}_i^{(n+1)} - \mathbf{A}_i^{(n)}\| \le \varepsilon$, then end.
   Otherwise, let $n := n + 1$ and go to step 2).

Substituting (11) back into the left-hand-side of the transmission power constraint (10), we have $\mathrm{tr}\big(\mathbf{A}_i(\mathbf{\Lambda}_{s,i}^2 + \mathbf{I}_{N_r})\mathbf{A}_i^H + \mathbf{Y}_i(\mathbf{\Lambda}_{s,i}^2 + \mathbf{I}_{N_r})\mathbf{Y}_i + \mathbf{X}_i\mathbf{X}_i^H + \mathbf{Z}_i\mathbf{Z}_i^H\big)$. From (13), we find that $\mathbf{X}_i = 0_{R\times(N_r-R)}, \mathbf{Y}_i = 0_{(N_r-R)\times R}$, and $\mathbf{Z}_i = 0_{(N_r-R)\times(N_r-R)}$, minimize the power consumption. Thus we have $\mathbf{F}_i = \mathbf{V}_{r,i}\mathbf{A}_i\mathbf{U}_{s,i}^H$. □

The remaining task is to find the optimal $\mathbf{A}_i, i = 1, \cdots, K$. From (14), we can write the optimization problem as

$$\min_{\mathbf{A}_i} \ \mathrm{tr}\left(\left[\mathbf{I}_{N_s} + \sum_{i=1}^{K}\mathbf{V}_{s,i}\mathbf{\Lambda}_{s,i}^H\mathbf{A}_i^H\mathbf{\Lambda}_{r,i}^H\mathbf{U}_{r,i}^H\left(\sum_{i=1}^{K}\mathbf{U}_{r,i}\mathbf{\Lambda}_{r,i}\mathbf{A}_i \mathbf{A}_i^H\mathbf{\Lambda}_{r,i}^H\mathbf{U}_{r,i}^H + \mathbf{I}_{N_d}\right)^{-1}\sum_{i=1}^{K}\mathbf{U}_{r,i}\mathbf{\Lambda}_{r,i}\mathbf{A}_i\mathbf{\Lambda}_{s,i}\mathbf{V}_{s,i}^H\right]^{-1}\right) \quad (15)$$

$$\text{s.t.} \ \mathrm{tr}\big(\mathbf{A}_i(\mathbf{\Lambda}_{s,i}^2 + \mathbf{I}_{N_r})\mathbf{A}_i^H \le P_{r,i}\big), \quad i = 1, \cdots, K. \quad (16)$$

Both the problem (9)-(10) and the problem (15)-(16) have matrix optimization variable. However, in the former problem, the optimization variable $\mathbf{F}_i$ is an $N_r \times N_r$ matrix. In general, the problem (15) - (16) is nonconvex and globally optimal solution is difficult to obtain with a reasonable computational complexity. Fortunately, we can resort to numerical methods, such as the projected gradient algorithm [11] to find (at least) a locally optimal solution of (15) - (16). The procedure of the projected gradient algorithm is listed in Table I, where δ_n and s_n denote the step size parameters at the nth iteration. max abs$\|\cdot\|$ denote the maximum among the absolute value of all elements in a matrix, and ε is a positive constant close to 0.

THEOREM 2: If $f(\mathbf{A}_i) = \mathrm{tr}\left(\left[\mathbf{I}_{N_s} + \tilde{\mathbf{H}}^H\tilde{\mathbf{C}}^{-1}\tilde{\mathbf{H}}\right]^{-1}\right)$ is chosen as the objective function, then its gradient $\nabla f(\mathbf{A}_i)$ with respect to $\mathbf{A}_i$ can be calculated by using results on derivatives of matrices in [13] as

$$\begin{aligned}\nabla f(\mathbf{A}_i) \ &= \ 2\left([\mathbf{M}_i\mathbf{R}_i]^T[\mathbf{S}_i\mathbf{C}_i]^T + [\mathbf{M}_i\mathbf{R}_i]^T[\mathbf{D}_i]^T \right.\\ &\quad \left. -[\mathbf{E}_i^H\mathbf{G}_i^{-1}\mathbf{R}_i]^T[\mathbf{S}_i]^T\right)^* \\ &\quad i = 1, \cdots, K. \end{aligned} \quad (17)$$

PROOF: See Appendix A.

The projection of $\tilde{\mathbf{A}}_i$ onto the feasible set of $\bar{\mathbf{A}}_i$ given by (16) is performed by solving the following optimization problem

$$\min_{\bar{\mathbf{A}}_i} \quad \mathrm{tr}\big((\bar{\mathbf{A}}_i - \tilde{\mathbf{A}}_i)(\bar{\mathbf{A}}_i - \tilde{\mathbf{A}}_i)^H\big) \quad (18)$$

$$\text{s.t.} \quad \mathrm{tr}\big(\bar{\mathbf{A}}_i(\mathbf{\Lambda}_{s,i}^2 + \mathbf{I}_{N_r})\bar{\mathbf{A}}_i^H\big) \le P_{r,i}. \quad (19)$$

By using the Lagrange multiplier method, the solution to the problem (18)-(19) is given by

$$\bar{\mathbf{A}}_i = \tilde{\mathbf{A}}_i[(\lambda + 1)\mathbf{I}_{N_r} + \lambda\mathbf{\Lambda}_{s,i}^2]^{-1}$$

where $\lambda > 0$ is the solution to the nonlinear equation

$$\mathrm{tr}\Big(\tilde{\mathbf{A}}_i[(\lambda+1)\mathbf{I}_{N_r} + \lambda\mathbf{\Lambda}_{s,i}^2]^{-1}(\mathbf{\Lambda}_{s,i}^2 + \mathbf{I}_{N_r}) [(\lambda+1)\mathbf{I}_{N_r} + \lambda\mathbf{\Lambda}_{s,i}^2]^{-1}\tilde{\mathbf{A}}_i^H\Big) = P_{r,i}. \quad (20)$$

Equation (20) can be efficiently solved by the bisection method [11]. The step size parameters δ_n and s_n are determined by the Armijo rule [11], i.e., $s_n = s$ is a constant through all iterations, while at the nth iteration, δ_n is set to be γ^{m_n}. Here m_n is the terminal nonnegative integer that satisfies the following inequality $\mathrm{MSE}(\mathbf{A}_i^{(n+1)}) - \mathrm{MSE}(\mathbf{A}_i^{(n)}) \le \alpha\gamma^{m_n}\mathrm{realtr}\Big(\big(\nabla f(\mathbf{A}_i)^{(n)}\big)^H(\bar{\mathbf{A}}_i^{(n)} - \mathbf{A}_i^{(n)})\Big)$, where α and γ are constants. According to [11], usually α is chosen close to 0, for example $\alpha\varepsilon[10^{-5}, 10^{-1}]$, while a proper choice of γ is normally from 0.1 to 0.5.

B. Simplified Design

By introducing

$$\bar{\mathbf{F}} \triangleq \mathbf{H}_{rd}\mathbf{F}. \quad (21)$$

The received signal vector at the destination can be equivalently written as $\mathbf{y}_d = \bar{\mathbf{H}}\mathbf{s} + \bar{\mathbf{v}}$, where $\bar{\mathbf{H}} \triangleq \bar{\mathbf{F}}\mathbf{H}_{sr}$, and $\bar{\mathbf{v}} \triangleq \bar{\mathbf{F}}\mathbf{v}_r + \mathbf{v}_d$. Considering (2) and (21), the transmission power consumed at the output of $\mathbf{H}_{rd}$ can be expressed as

$$\begin{aligned}&\mathrm{E}[\mathrm{tr}((\mathbf{H}_{rd}\mathbf{x}_r)(\mathbf{H}_{rd}\mathbf{x}_r)^H)] = \mathrm{tr}\big(\bar{\mathbf{F}}\big[\mathbf{H}_{sr}\mathbf{H}_{sr}^H + \mathbf{I}_{KN_r}\big]\bar{\mathbf{F}}^H\big)\\ &\le \mathrm{tr}(\mathbf{H}_{rd,i}\mathbf{H}_{rd,i}^H)\mathrm{tr}\big(\mathbf{F}_i\big[\mathbf{H}_{sr,i}\mathbf{H}_{sr,i}^H + \mathbf{I}_{N_r}\big]\mathbf{F}_i^H\big). \end{aligned} \quad (22)$$

Substituting (10) into (22) we have

$$\mathrm{tr}\big(\bar{\mathbf{F}}\big[\mathbf{H}_{sr}\mathbf{H}_{sr}^H + \mathbf{I}_{KN_r}\big]\bar{\mathbf{F}}^H\big) \le \sum_{i=1}^{K}P_{r,i}\sum_{i=1}^{K}\mathrm{tr}(\mathbf{H}_{rd,i}\mathbf{H}_{rd,i}^H). \quad (23)$$

Here $\sum_{i=1}^{K}P_{r,i} \triangleq P_r$, is the total transmission power budget available to all K relay nodes. Using (23), the relaxed relay optimization problem can be written as

$$\min_{\bar{\mathbf{F}}} \quad \mathrm{tr}\left(\left[\mathbf{I}_{N_s} + \bar{\mathbf{H}}^H\bar{\mathbf{C}}^{-1}\bar{\mathbf{H}}\right]^{-1}\right) \quad (24)$$

$$\text{s.t.} \quad \mathrm{tr}\big(\bar{\mathbf{F}}\big[\mathbf{H}_{sr}\mathbf{H}_{sr}^H + \mathbf{I}_{KN_r}\big]\bar{\mathbf{F}}^H\big) \le \bar{P}_r, i = 1, \cdots, K \quad (25)$$

where $\bar{P}_r \triangleq P_r\mathrm{tr}(\mathbf{H}_{rd}\mathbf{H}_{rd}^H)$. Let $\mathbf{H}_{sr} = \mathbf{U}_s\mathbf{\Lambda}_s\mathbf{V}_s^H$ denote the singular value decomposition (SVD) of $\mathbf{H}_{sr}$, where the dimensions of $\mathbf{U}_s$, $\mathbf{\Lambda}_s$, $\mathbf{V}_s$ are $KN_r \times KN_r$, $KN_r \times N_s$, $N_s \times N_s$, respectively. We assume that the main diagonal elements of $\mathbf{\Lambda}_s$ is arranged in a decreasing order. Using Theorem 1 in [10], the optimal structure of $\bar{\mathbf{F}}$ as the solution to the problem (24)-(25) is given by

$$\bar{\mathbf{F}} = \mathbf{Q}\mathbf{\Lambda}_f\mathbf{U}_{s,1}^H \quad (26)$$

where $\mathbf{Q}$ is any $N_d \times N_s$ semi-unitary matrix with $\mathbf{Q}^H\mathbf{Q} = \mathbf{I}_{N_s}$, $\mathbf{U}_{s,1}$ contain the leftmost N_b columns of $\mathbf{U}_s$, and $\mathbf{\Lambda}_f$ is an $N_s \times N_s$ diagonal matrix. The proof of (26) is similar to the proof of Theorem 1 in [10]. From (26), we see that the optimal $\bar{\mathbf{F}}$ has a beamforming structure. In fact, the optimal $\bar{\mathbf{F}}$ diagonalizes the source-relay-destination channel $\bar{\mathbf{H}}$ up to a rotation matrix $\mathbf{Q}$. Using (26), the relay optimization problem (24)-(25) becomes

$$\min_{\mathbf{\Lambda}_f} \quad \mathrm{tr}\left(\left[\mathbf{I}_{N_s} + \left(\mathbf{\Lambda}_f\mathbf{\Lambda}_s\right)^2\left[\mathbf{\Lambda}_f^2 + \mathbf{I}_{N_s}\right]^{-1}\right]^{-1}\right) \tag{27}$$

$$\text{s.t.} \quad \mathrm{tr}\left(\mathbf{\Lambda}_f^2\left[\mathbf{\Lambda}_s^2 + \mathbf{I}_{N_s}\right]\right) \le \bar{P}_r. \tag{28}$$

Let us denote $\lambda_{f,i}, \lambda_{s,i}, i = 1, \cdots, N_s$, as the main diagonal elements of $\mathbf{\Lambda}_f$, $\mathbf{\Lambda}_s$, respectively, and introduce

$$a_i \triangleq \lambda_{s,i}^2, \qquad y_i \triangleq \lambda_{f,i}^2\left[\lambda_{s,i}^2 + 1\right], \qquad i = 1, \cdots, N_s. \tag{29}$$

The optimization problem (27)-(28) can be equivalently rewritten as

$$\min_{\mathbf{y}} \quad \sum_{i=1}^{N_s} \frac{a_i x_i + y_i + 1}{a_i x_i y_i + a_i x_i + y_i + 1} \tag{30}$$

$$\text{s.t.} \quad \sum_{i=1}^{N_s} y_i \le \bar{P}_r \qquad y_i \ge 0, \qquad i = 1, \cdots, N_s \tag{31}$$

where $\mathbf{y} \triangleq [y_1, y_2, \cdots, y_{N_s}]^T$. The problem (30)-(31) can be solved by an iterative method developed in [10], where in iteration, $\mathbf{y}$ is updated alternatingly by fixing the other vector. After the optimal $\mathbf{y}$ is found, $\lambda_{f,i}$ can be obtained from (29) as

$$\lambda_{f,i} = \sqrt{\frac{y_i}{\lambda_{s,i}^2 x_i + 1}}, \qquad i = 1, \cdots, N_s. \tag{32}$$

Using (21) and the optimal structure of $\bar{\mathbf{F}}$ in (26), we have $\mathbf{H}_{rd,i}\mathbf{F}_i = \mathbf{Q}\mathbf{\Lambda}_f\mathbf{\Phi}_i$, where matrix $\mathbf{\Phi}_i$ contains the $(i-1)N_r + 1$ to iN_r columns of $\mathbf{U}_{s,1}^H$. Then we obtain

$$\mathbf{F}_i = \mathbf{H}_{rd,i}^{\dagger}\mathbf{Q}\mathbf{\Lambda}_f\mathbf{\Phi}_i, \qquad i = 1, \cdots, K \tag{33}$$

where $(\cdot)^{\dagger}$ denotes matrix pseudo-inverse. Finally, we scale $\mathbf{F}_i$ in (33) to satisfy the power constraint (10) at each relay node as

$$\tilde{\mathbf{F}}_i = \alpha_i\mathbf{F}_i, \qquad i = 1, \cdots, K \tag{34}$$

where the scaling factor α_i is given by $\alpha_i = \sqrt{P_{r,i}/\mathrm{tr}(\mathbf{F}_i[\mathbf{H}_{sr,i}\mathbf{H}_{sr,i}^H + \mathbf{I}_{N_r}]\mathbf{F}_i^H)}$, $i = 1, \cdots, K$.

IV. Simulations

In this section, we study the performance of the proposed optimal relay beamforming algorithms for parallel MIMO relay systems with linear MMSE receiver. All simulations are conducted in a flat Rayleigh fading environment where the channel matrices have zero-mean entries with variance σ_s^2/N_s and $\sigma_r^2/(KN_r)$ for $\mathbf{H}_{sr}$ and $\mathbf{H}_{rd}$, respectively. The BPSK constellations are used to modulate the source symbols, and all noise are i.i.d Gaussian with zero mean and unit variance. We define $\mathrm{SNR_s} = \sigma_s^2 P_s K N_r/N_s$ and $\mathrm{SNR_r} = \sigma_r^2 P_r N_d/(KN_r)$ as the signal-to-noise ration (SNR) for the source-relay link and the relay-destination link, respectively. We transmit $N_s \times 1000$ randomly generated bits in each channel realization, and all simulation results are averaged over 200 channel realizations. In all simulations, the MMSE linear receiver in (7) is employed at the destination for symbol detection.

In our example, a parallel MIMO relay system with $K = 2$ relay nodes, $N_s = N_d = 5$, and $N_r = 4$ are simulated. We compare the BER performance of the propose optimal relay matrices using Projected Gradient (ORP) algorithm in (12) with ZF algorithm in [8], MMSE algorithm in [8], and the naive amplify-and-forward (NAF) Algorithm. While Fig. 2 demonstrates BER versus $\mathrm{SNR_s}$ for $\mathrm{SNR_r}$ fixed at 20 dB. It can be seen that the propose algorithm outperforms all competing algorithms in the whole $\mathrm{SNR_s}$ range.

V. Conclusions

In this paper, we have derived the general structure of the optimal relay amplifying matrices for parallel MIMO relay communication systems using the projected gradient approach. The proposed algorithm has less computational complexity compared to the existing techniques. Simulation result shows the effectiveness of the proposed algorithm.

VI. Appendix

Base on (11) and (12), we have $\mathbf{H}_{sr,i} = \mathbf{U}_{s,i}\mathbf{\Lambda}_{s,i}\mathbf{V}_{s,i}^H$, $\mathbf{H}_{rd,i} = \mathbf{U}_{r,i}\mathbf{\Lambda}_{r,i}\mathbf{V}_{r,i}^H$, $\mathbf{F}_i = \mathbf{V}_{r,i}\mathbf{A}_i\mathbf{U}_{s,i}^H$, $\sum_{i=1}^K \mathbf{H}_{rd,i}\mathbf{F}_i\mathbf{H}_{sr,i} = \sum_{i=1}^K \mathbf{U}_{r,i}\mathbf{\Lambda}_{r,i}\mathbf{A}_i\mathbf{\Lambda}_{s,i}\mathbf{V}_{s,i}^H$, and $\sum_{i=1}^K \mathbf{H}_{rd,i}\mathbf{F}_i\mathbf{F}_i^H\mathbf{H}_{rd,i}^H = \sum_{i=1}^K \mathbf{U}_{r,i}\mathbf{\Lambda}_{r,i}\mathbf{A}_i\mathbf{A}_i^H\mathbf{\Lambda}_{r,i}^H\mathbf{U}_{r,i}^H$. Thus $f(\mathbf{A}_i)$ can be written as

$$f(\mathbf{A}_i) = \mathrm{tr}\Bigg(\Bigg[\mathbf{I}_{N_s} + \sum_{i=1}^K \mathbf{V}_{s,i}\mathbf{\Lambda}_{s,i}^H\mathbf{A}_i^H\mathbf{\Lambda}_{r,i}^H\mathbf{U}_{r,i}^H\Bigg(\sum_{i=1}^K \mathbf{U}_{r,i}\mathbf{\Lambda}_{r,i}\mathbf{A}_i \mathbf{A}_i^H\mathbf{\Lambda}_{r,i}^H\mathbf{U}_{r,i}^H + \mathbf{I}_{N_d}\Bigg)^{-1}\sum_{i=1}^K \mathbf{U}_{r,i}\mathbf{\Lambda}_{r,i}\mathbf{A}_i\mathbf{\Lambda}_{s,i}\mathbf{V}_{s,i}^H\Bigg]^{-1}\Bigg) \tag{35}$$

Let us define $\mathbf{Z}_i^H \triangleq \sum_{j=1,j\neq i}^K \mathbf{V}_{s,j}\mathbf{\Lambda}_{s,j}^H\mathbf{A}_j^H\mathbf{\Lambda}_{r,j}^H\mathbf{U}_{r,j}^H$, and $\mathbf{Y}_i \triangleq \sum_{j=1,j\neq i}^K \mathbf{U}_{r,j}\mathbf{\Lambda}_{r,j}\mathbf{A}_j\mathbf{A}_j^H\mathbf{\Lambda}_{r,j}^H\mathbf{U}_{r,j}^H + \mathbf{I}_{N_d}$. Then $f(\mathbf{A}_i)$ can be written as

$$f(\mathbf{A}_i) = \mathrm{tr}\Big(\big[\mathbf{I}_{N_s} + (\mathbf{Z}_i^H + \mathbf{V}_{s,i}\mathbf{\Lambda}_{s,i}^H\mathbf{A}_i^H\mathbf{\Lambda}_{r,i}^H\mathbf{U}_{r,i}^H)(\mathbf{Y}_i + \mathbf{U}_{r,i}\mathbf{\Lambda}_{r,i} \mathbf{A}_i\mathbf{A}_i^H\mathbf{\Lambda}_{r,i}^H\mathbf{U}_{r,i}^H)^{-1}(\mathbf{U}_{r,i}\mathbf{\Lambda}_{r,i}\mathbf{A}_i\mathbf{\Lambda}_{s,i}\mathbf{V}_{s,i}^H + \mathbf{Z}_i)\big]^{-1}\Big) \tag{36}$$

Applying $\left[\mathbf{I}_{N_s} + \mathbf{A}^H\mathbf{C}^{-1}\mathbf{A}\right]^{-1} = \mathbf{I}_{N_s} - \mathbf{A}^H(\mathbf{A}\mathbf{A}^H + \mathbf{C})^{-1}\mathbf{A}$. Then, (36) can be written as

$$f(\mathbf{A}_i) = \mathrm{tr}\big[\mathbf{I}_{N_s} - (\mathbf{Z}_i^H + \mathbf{V}_{s,i}\mathbf{\Lambda}_{s,i}^H\mathbf{A}_i^H\mathbf{\Lambda}_{r,i}^H\mathbf{U}_{r,i}^H)((\mathbf{U}_{r,i}\mathbf{\Lambda}_{r,i} \mathbf{A}_i\mathbf{\Lambda}_{s,i}\mathbf{V}_{s,i}^H + \mathbf{Z}_i)(\mathbf{Z}_i^H + \mathbf{V}_{s,i}\mathbf{\Lambda}_{s,i}^H\mathbf{A}_i^H\mathbf{\Lambda}_{r,i}^H\mathbf{U}_{r,i}^H) + (\mathbf{Y}_i + \mathbf{U}_{r,i}\mathbf{\Lambda}_{r,i}\mathbf{A}_i\mathbf{A}_i^H\mathbf{\Lambda}_{r,i}^H\mathbf{U}_{r,i}^H))^{-1} (\mathbf{U}_{r,i}\mathbf{\Lambda}_{r,i}\mathbf{A}_i\mathbf{\Lambda}_{s,i}\mathbf{V}_{s,i}^H + \mathbf{Z}_i)\big]. \tag{37}$$

Let us now define $\mathbf{E}_i \triangleq \mathbf{U}_{r,i}\mathbf{\Lambda}_{r,i}\mathbf{A}_i\mathbf{\Lambda}_{s,i}\mathbf{V}_{s,i}^H + \mathbf{Z}_i$, $\mathbf{K}_i \triangleq \mathbf{Y}_i + \mathbf{U}_{r,i}\mathbf{\Lambda}_{r,i}\mathbf{A}_i\mathbf{A}_i^H\mathbf{\Lambda}_{r,i}^H\mathbf{U}_{r,i}^H$, and $\mathbf{G}_i \triangleq \mathbf{E}_i\mathbf{E}_i^H + \mathbf{K}_i$. We can rewrite (37) as

$$f(\mathbf{A}_i) = \mathrm{tr}\left[\mathbf{I}_{N_s} - \mathbf{E}_i^H\mathbf{G}_i^{-1}\mathbf{E}_i\right] = \mathrm{tr}\left[\mathbf{I}_{N_s} - \mathbf{E}_i\mathbf{E}_i^H\mathbf{G}_i^{-1}\right]. \quad (38)$$

Then the derivative of $f(\mathbf{A}_i)$ with respect to $\mathbf{A}_i$ is given by

$$\begin{aligned}\frac{\partial}{\partial \mathbf{A}_i} f(\mathbf{A}_i) = & -\frac{\partial}{\partial \mathbf{A}_i}\mathrm{tr}\left[\mathbf{E}_i\mathbf{E}_i^H\mathbf{G}_i^{-1}\right] \\ = & \frac{\partial}{\partial \mathbf{A}_i}\mathrm{tr}\left[\mathbf{G}_i^{-1}\mathbf{E}_i\mathbf{E}_i^H\mathbf{G}_i^{-1}\mathbf{G}_i\right] \\ & -\frac{\partial}{\partial \mathbf{A}_i}\mathrm{tr}[\mathbf{E}_i^H\mathbf{G}_i^{-1}\mathbf{U}_{r,i}\mathbf{\Lambda}_{r,i}\mathbf{A}_i\mathbf{\Lambda}_{s,i}\mathbf{V}_{s,i}^H] \\ = & \frac{\partial}{\partial \mathbf{A}_i}\mathrm{tr}\left[\mathbf{G}_i^{-1}\mathbf{E}_i\mathbf{E}_i^H\mathbf{G}_i^{-1}\left((\mathbf{U}_{r,i}\mathbf{\Lambda}_{r,i}\mathbf{A}_i\mathbf{\Lambda}_{s,i}\right.\right. \\ & \mathbf{V}_{s,i}^H + \mathbf{Z}_i)(\mathbf{Z}_i^H + \mathbf{V}_{s,i}\mathbf{\Lambda}_{s,i}^H\mathbf{A}_i^H\mathbf{\Lambda}_{r,i}^H\mathbf{U}_{r,i}^H) \\ & \left.\left.+(\mathbf{Y}_i + \mathbf{U}_{r,i}\mathbf{\Lambda}_{r,i}\mathbf{A}_i\mathbf{A}_i^H\mathbf{\Lambda}_{r,i}^H\mathbf{U}_{r,i}^H)\right)\right] \\ & -[\mathbf{E}_i^H\mathbf{G}_i^{-1}\mathbf{U}_{r,i}\mathbf{\Lambda}_{r,i}]^T[\mathbf{\Lambda}_{s,i}\mathbf{V}_{s,i}^H]^T. \qquad (39)\end{aligned}$$

Defining $\mathbf{M}_i \triangleq \mathbf{G}_i^{-1}\mathbf{E}_i\mathbf{E}_i^H\mathbf{G}_i^{-1}$, $\mathbf{C}_i \triangleq \mathbf{E}_i^H$, and $\mathbf{D}_i \triangleq \mathbf{A}_i^H\mathbf{\Lambda}_{r,i}^H\mathbf{U}_{r,i}^H$, we can rewrite (39) as

$$\begin{aligned}\frac{\partial}{\partial \mathbf{A}_i} f(\mathbf{A}_i) = & \frac{\partial}{\partial \mathbf{A}_i}\mathrm{tr}\left[\mathbf{M}_i(\mathbf{U}_{r,i}\mathbf{\Lambda}_{r,i}\mathbf{A}_i\mathbf{\Lambda}_{s,i}\mathbf{V}_{s,i}^H + \mathbf{Z}_i)\mathbf{C}_i\right. \\ & \left.+\mathbf{M}_i(\mathbf{Y}_i + \mathbf{U}_{r,i}\mathbf{\Lambda}_{r,i}\mathbf{A}_i\mathbf{D}_i)\right] \\ & -[\mathbf{E}_i^H\mathbf{G}_i^{-1}\mathbf{U}_{r,i}\mathbf{\Lambda}_{r,i}]^T[\mathbf{\Lambda}_{s,i}\mathbf{V}_{s,i}^H]^T. \end{aligned} \quad (40)$$

Here $\mathbf{R}_i \triangleq \mathbf{U}_{r,i}\mathbf{\Lambda}_{r,i}$, and $\mathbf{S}_i \triangleq \mathbf{\Lambda}_{s,i}\mathbf{V}_{s,i}^H$, we can rewrite (40) as

$$\begin{aligned}\frac{\partial}{\partial \mathbf{A}_i} f(\mathbf{A}_i) = & \frac{\partial}{\partial \mathbf{A}_i}\mathrm{tr}\left[\mathbf{M}_i\mathbf{R}_i\mathbf{A}_i\mathbf{S}_i\mathbf{C}_i + \mathbf{M}_i\mathbf{R}_i\mathbf{A}_i\mathbf{D}_i\right] \\ & -[\mathbf{E}_i^H\mathbf{G}_i^{-1}\mathbf{R}_i]^T[\mathbf{S}_i]^T. \end{aligned} \quad (41)$$

Finally, the gradient is given by

$$\begin{aligned}\nabla f(\mathbf{A}_i) = & 2\big([\mathbf{M}_i\mathbf{R}_i]^T[\mathbf{S}_i\mathbf{C}_i]^T + [\mathbf{M}_i\mathbf{R}_i]^T[\mathbf{D}_i]^T \\ & -[\mathbf{E}_i^H\mathbf{G}_i^{-1}\mathbf{R}_i]^T[\mathbf{S}_i]^T\big)^*. \end{aligned} \quad (42)$$

□

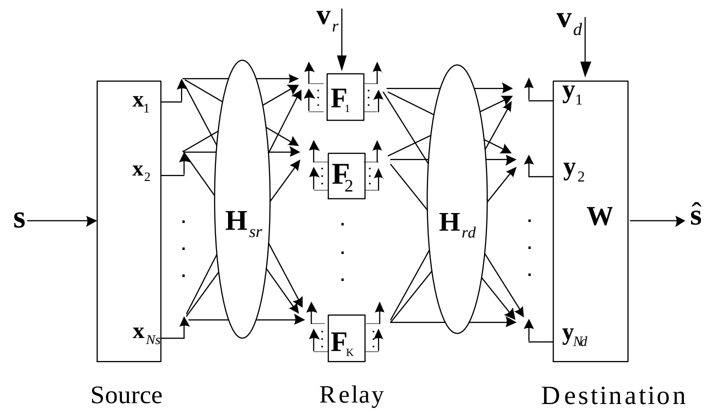

Fig. 1. Block diagram of a parallel MIMO relay communication system.

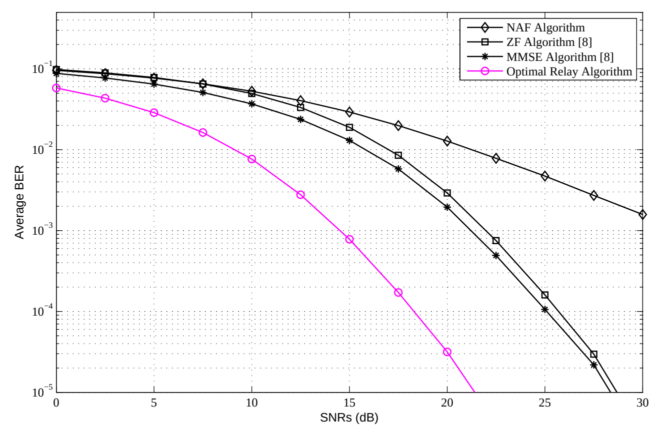

Fig. 2. BER versus $\mathrm{SNR_s}$ while fixing $\mathrm{SNR_r} = 20\mathrm{dB}$. $N_s = N_d = 5$, and $N_r = 4$.

References

[1] R. Pabst, B. H. Walke, D. C. Schultz, D. C. Herhold, H. Yanikomeroglu, S. Mukherjee, H. Viswanathan, M. Lott, W. Zirwas, M. Dohler, H. Aghvami, D. D. Falconer, and G. P. Fettweis, "Relay-based deployment concepts for wireless and mobile broadband radio," *IEEE Commun. Mag.*, vol. 42, pp. 80-89, Sep. 2004.

[2] B. Wang, J. Zhang, and A. Høst-Madsen, "On the capacity of MIMO relay channels," *IEEE Trans. Inf. Theory*, vol. 51, pp. 29-43, Jan. 2005.

[3] X. Tang and Y. Hua, "Optimal design of non-regenerative MIMO wireless relays," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 1398-1407, Apr. 2007.

[4] O. Muñoz-Medina, J. Vidal, and A. Agustín, "Linear transceiver design in nonregenerative relays with channel state information," *IEEE Trans. Signal Process.*, vol. 55, pp. 2593-2604, Jun. 2007.

[5] W. Guan and H. Luo, "Joint MMSE transceiver design in non-regenerative MIMO relay systems," *IEEE Commun. Lett.*, vol. 12, pp. 517-519, Jul. 2008.

[6] G. Li, Y. Wang, T. Wu, and J. Huang, "Joint linear filter design in multi-user cooperative non-regenerative MIMO relay systems," *EURASIP J. Wireless Commun. Networking*, vol. 2009, Article ID 670265.

[7] Y. Rong, "Joint source and relay optimization for two-way linear non-regenerative MIMO relay communications," *IEEE Trans. Wireless Commun.*, vol. 8, pp. 6068-6077, Dec. 2009.

[8] O. Oyman and A. J. Paulraj, "Design and analysis of linear distributed MIMO relaying algorithms," *IEE Proc. Inst. Elect. Eng., Commun.*, pp. 565-572, Aug. 2006.

[9] A. S. Behbahani, R. Merched, and A. M. Eltawil, "Optimizations of a MIMO relay network," *IEEE Trans. Signal Process.*, vol. 56, pp. 5062-5073, Oct. 2008.

[10] Y. Rong, X. Tang, and Y. Hua, "A unified framework for optimizing linear non-regenerative multicarrier MIMO relay communication systems," *IEEE Trans. Signal Process.*, vol. 57, pp. 4837-4851, Dec. 2009.

[11] D. P. Bertsekas, *Nonlinear Programming*. 2nd. Edition, Athena Scientific, Belmont, Massachusetts, 1999.

[12] S. Boyd and L, Vandenberghe, *Convex Optimizationh*. Cambridge, U.K.: Cambridge University Press, 2004.

[13] K. B. Petersen and M. S. Petersen, *The Matrix Cookbook*. [Online]. Available: http://matrixcookbook.com.