Joint Optimization of Source Beamformer and Relay Coefficients Using MSE Criterion *

Batu K. Chalise, Yimin D. Zhang and Moeness G. Amin

Center for Advanced Communications, Villanova University
PA 19085, USA

ABSTRACT

The joint optimization of source beamformer and distributed relay coefficients is considered for a cooperative network that uses the output mean-square error (MSE) as the performance metric. Two methods are proposed for solving this nonconvex optimization problem. The first approach iteratively optimizes the source beamformer and relay coefficients, whereas the second method optimizes the relay coefficients keeping the source beamformer as the maximum-ratio-transmitter (MRT). Both approaches are reformulated as relaxed semidefinite programming problems. The optimality of MRT is proven analytically under some conditions. The superiority of the proposed methods is verified using numerical examples.

Keywords: Beamforming, relays, convex optimization

1. INTRODUCTION

Recently, cooperative communications with multi-antenna nodes have received a lot of research interest[1]. The optimal designs of a single and multiple (joint optimization) amplify-and-forward (AF) multiple-input multiple-output (MIMO) relays have been investigated in[2,3] and[4], respectively, for a point-to-point communication scheme. However, these works do not consider the problem of also optimizing the linear processing at the source, which is required for exploiting the transmit diversity. The problems of jointly optimizing the source precoder and MIMO relay processing matrix have been solved in[5,6]. In[5] only a particular function such as the mutual information is maximized while in[6] a unified framework based on majorization theory has been presented for solving a wide range of optimization problems with Schur convex and concave objective functions. The common assumption among all of these works is that the relay processing matrix is a full-matrix due to the fact that the relay is a MIMO node. Consequently, in such cases, the optimal source precoder and MIMO relay turn to be those which jointly diagonalize the MSE matrix, for instance, if the objective function is Schur concave (e.g., sum mean-square error (MSE) minimization problem). The resulting optimization problem reduces to the problem of allocating powers at the source and multi-antenna relay. The joint optimization of the source beamformer, MIMO relay and the destination receiver has been solved in[7] for both the systems with and without the direct link. For the latter system model (i.e., without direct link), the optimal source and receive beamformers, and the relay matrix have been shown to yield a *matching* solution. However, the considered problems[5–7] become different and also difficult to solve, if the relay matrix is constrained to be a diagonal matrix, which is the case when the relay is not a MIMO node but a number of distributed single-antenna nodes. Intuitively, it can be said that, since the relay processing matrix is not a full matrix, there is no sufficient degrees of freedom to obtain the *matching* solution as in the case of[7]. Moreover, since it is difficult to use multiple antennas at each relay node (due to implementation cost) and multi-antenna relays suffer more from antenna correlations as the number of antennas increases, systems with distributed single-antenna relay nodes[8] are of higher practical interest.

In this paper, we consider a point-to-point communication system supported by multiple single-antenna AF relays. The source and destination nodes are equipped with multiple antennas. Considering that the CSI of the *source-relay* and *relay-destination* channels are known perfectly, we solve the problem of jointly optimizing

*This work was supported by DARPA under contract HR0011-10-1-0062. The content of the information does not necessarily reflect the position or the policy of the Government, and no official endorsement should be inferred. Approved for Public Release, Distribution Unlimited.

E-mail: {batu.chalise, yimin.zhang, moeness.amin}@villanova.edu

Wireless Sensing, Localization, and Processing VII, edited by Sohail A. Dianat, Michael David Zoltowski,
Proc. of SPIE Vol. 8404, 84040H · © 2012 SPIE · CCC code: 0277-786X/12/$18 · doi: 10.1117/12.920479

the source beamformer and relay coefficients. The objective is to minimize the MSE[†] under source and relay sum-power constraints[‡]. This optimization problem is nonconvex whose solution is difficult to be obtained in closed-form. We prove that, for the jointly optimal source beamformer and relay coefficients, the power constraints for the source and relay (sum-power) are satisfied with equality. We then propose to solve the considered nonconvex optimization problem using the following two approaches. a) Iterative method: Keeping the source beamformer fixed, we optimize the relay coefficients by reformulating the corresponding optimization problem as a rank-relaxed semidefinite programming (SDP) problem. The best approximate rank-one solutions are obtained from the relaxed solution. For the given relay coefficients, the source beamformer is another nonconvex optimization problem which is also solved with the semidefinite relaxation. The abovementioned two steps of optimization are repeated until required convergence accuracy is achieved. b) Relay optimization with the maximal-ratio transmitter (MRT) source beamformer: We select the source beamformer (MRT solution) which maximizes the total power radiated by the source beamformer towards the relays given that the available source power is fixed. By taking such MRT solution, the remaining problem of optimizing the relay coefficients is again solved using the SDP rank relaxation technique. We prove analytically that the MRT solution for source beamformer is optimal for some conditions. Our numerical results show that the iterative method yields equal or slightly improved performance compared to the non-iterative method, if the source beamformer of former method is initialized with the MRT solution. As a result, the non-iterative approach is more preferable than iterative method when system design demands computationally efficient method at the expense of some minor loss in performance.

Notations: Upper (lower) bold face letters will be used for matrices (vectors); $(\cdot)^T$, $(\cdot)^H$, $\mathrm{E}\{\cdot\}$, $\mathbf{I}_n$ and $\mathrm{diag}(\mathbf{x})$ denote transpose, Hermitian transpose, mathematical expectation, $n \times n$ identity matrix and the diagonal matrix formed from $\mathbf{x}$, respectively. $\mathrm{tr}(\cdot)$, $\mathcal{C}^{M\times M}$, $\odot$ and $\mathbf{A} \succeq 0$ denote the matrix trace operator, space of $M \times M$ matrices with complex entries, the Schur-Hadamard product and positive semidefiniteness of $\mathbf{A}$, respectively.

2. SYSTEM MODEL

We consider the scenario where signal transmission from a multi-antenna source (with N_s antennas) to a multi-antenna destination (with N_d antennas) is supported by R single-antenna relays. Each relay amplifies the signal received from the source and forwards the resulting signal to the destination. Note that the direct link between the source and destination is not considered since we assume that the direct link experiences relatively larger path attenuation compared to the links via relays. We consider that the relays operate in half-duplex mode which means that signal transmissions from source to the relays and from the relays to the destination take place in two time slots. In the first time slot, the source sends its signal to the relays and in the second time slot, the relays send their processed signals to the destination. We assume that the instantaneous channel state information (CSI) of all channels are perfectly known. The channels between the source and the relays as well as between the relays and destination are assumed to be flat fading. The signal received by the ith relay is given by

$$y_i = \mathbf{h}_i^T \mathbf{w} s + n_{\mathrm{r},i} \tag{1}$$

where $\mathbf{h}_i^T \in \mathcal{C}^{1\times N_\mathrm{s}}$ is the channel vector between the source and the ith relay, $\mathbf{w} \in \mathcal{C}^{N_\mathrm{s}\times 1}$ is the source beamforming vector, $n_{\mathrm{r},i} \in \mathcal{C}^{1\times 1}$ is the additive white Gaussian noise at the ith relay and s is the random source signal with zero-mean and a unit variance. The signal received by all relays can be expressed in the vector form as

$$\mathbf{y}_\mathrm{r} = \mathbf{H}\mathbf{w} s + \mathbf{n}_\mathrm{r} \tag{2}$$

where $\mathbf{H} = [\mathbf{h}_1, \cdots, \mathbf{h}_R]^T \in \mathcal{C}^{R\times N_\mathrm{s}}$ and $\mathbf{n}_\mathrm{r} = [n_{\mathrm{r},1}, \cdots, n_{\mathrm{r},R}]^T$. We consider $\mathbf{n}_\mathrm{r} \sim \mathcal{N}_C(\mathbf{0}, \sigma_{n_\mathrm{r}}^2 \mathbf{I}_R)$, i.e., the entries of $\mathbf{n}_\mathrm{r}$ are i.i.d. zero-mean circularly symmetric complex Gaussian (ZMCSCG) with the variance $\sigma_{n_\mathrm{r}}^2$. The ith relay

[†] A suboptimal approach for maximizing the lower bound of the signal-to-noise ratio is recently proposed in[8].

[‡] Our proposed algorithms can be easily modified to solve the MSE minimization problem with individual power constraints for relays.

multiplies its received signal y_i with a complex coefficient z_i and forwards the resulting signal to the destination. The signal received by the destination can be written as

$$\mathbf{y}_{\mathrm{d}} = \mathbf{G}\mathbf{Z}\mathbf{y}_{\mathrm{r}} + \mathbf{n}_{\mathrm{d}} \tag{3}$$

where $\mathbf{Z} = \mathrm{diag}([z_1, \cdots, z_R]) \in \mathcal{C}^{R\times R}$, $\mathbf{G} = [\mathbf{g}_1, \cdots, \mathbf{g}_R]$, $\{\mathbf{g}_i\}_{i=1}^{R} \in \mathcal{C}^{N_{\mathrm{d}}\times 1}$ are the channels between the relays and the destination, and $\mathbf{n}_{\mathrm{d}} \in \mathcal{C}^{N_{\mathrm{d}}\times 1}$ is the additive noise at the destination. It is assumed that $\mathbf{n}_{\mathrm{d}} \sim \mathcal{N}_C(\mathbf{0}, \sigma_{n_{\mathrm{d}}}^2 \mathbf{I}_{N_{\mathrm{d}}})$. The destination uses a linear operator $\mathbf{d} \in \mathcal{C}^{N_{\mathrm{d}}\times 1}$ to recover the source signal s. The output of the destination is given by

$$\hat{s} = \mathbf{d}^H \mathbf{y}_{\mathrm{d}} \tag{4}$$

which with the help of (2) and (3) can be re-expressed as

$$\hat{s} = \mathbf{d}^H [\mathbf{G}\mathbf{Z}\mathbf{H}\mathbf{w}s + \mathbf{G}\mathbf{Z}\mathbf{n}_{\mathrm{r}} + \mathbf{n}_{\mathrm{d}}]. \tag{5}$$

The MSE $f = \mathrm{E}\left\{|s - \hat{s}|^2\right\}$ can be written as

$$\begin{aligned} f \;=\; & 1 - \mathbf{w}^H \mathbf{H}^H \mathbf{Z}^H \mathbf{G}^H \mathbf{d} - \mathbf{d}^H \mathbf{G}\mathbf{Z}\mathbf{H}\mathbf{w} + \mathbf{d}^H \cdot \\ & \underbrace{[\mathbf{G}\mathbf{Z}\mathbf{H}\mathbf{w}\mathbf{w}^H \mathbf{H}^H \mathbf{Z}^H \mathbf{G}^H + \sigma_{n_{\mathrm{r}}}^2 \mathbf{G}\mathbf{Z}\mathbf{Z}^H \mathbf{G}^H + \sigma_{n_{\mathrm{d}}}^2 \mathbf{I}_{N_{\mathrm{d}}}]}_{\mathbf{A}} \mathbf{d}. \end{aligned} \tag{6}$$

Setting the first-order derivative of (6) w.r.t. $\mathbf{d}$ to zero yields the optimal receive beamformer as

$$\mathbf{d} = \mathbf{A}^{-1}\mathbf{G}\mathbf{Z}\mathbf{H}\mathbf{w}. \tag{7}$$

Substituting (7) into (6), we can obtain the following minimium-mean-square error (MMSE) function

$$f = 1 - \mathbf{w}^H \mathbf{H}^H \mathbf{Z}^H \mathbf{G}^H \mathbf{A}^{-1} \mathbf{G}\mathbf{Z}\mathbf{H}\mathbf{w}. \tag{8}$$

The instantaneous sum-power of the relays can be given by

$$\begin{aligned} P_{\mathrm{r}} &= \mathrm{E}\left\{(\mathbf{Z}\mathbf{H}\mathbf{w}s + \mathbf{Z}\mathbf{n}_{\mathrm{r}})^H (\mathbf{Z}\mathbf{H}\mathbf{w}s + \mathbf{Z}\mathbf{n}_{\mathrm{r}})\right\} \\ &= \mathbf{w}^H \mathbf{H}^H \mathbf{Z}^H \mathbf{Z}\mathbf{H}\mathbf{w} + \sigma_{n_{\mathrm{r}}}^2 \mathrm{tr}(\mathbf{Z}^H \mathbf{Z}) \end{aligned} \tag{9}$$

and that of the source as $P_{\mathrm{s}} = \mathbf{w}^H \mathbf{w}$.

3. PROPOSED OPTIMIZATION

In this section, we consider the problem of jointly optimizing the source beamforming vector $\mathbf{w}$ and the diagonal matrix $\mathbf{Z}$ that consists of relay weights under the source power and relay sum-power constraints. This optimization problem can be mathematically formulated as

$$\min_{\mathbf{w},\mathbf{Z}} \; f \text{ s.t. } \; P_{\mathrm{r}} \le P_{\mathrm{R}}^m, \; P_{\mathrm{s}} \le P_{\mathrm{S}}^m \tag{10}$$

where P_{R}^m and P_{S}^m are the maximum powers available for the relays and source, respectively. The closed-form solution for this joint optimization problem is not known. Notice that if $\mathbf{Z}$ had not been a diagonal matrix, the closed-form optimal structures of $\mathbf{w}$ and $\mathbf{Z}$ could be directly taken from[7].

Let us define $\mathbf{v} = \mathbf{G}\mathbf{Z}\mathbf{H}\mathbf{w}$ and $\tilde{\mathbf{A}} = \sigma_{n_{\mathrm{r}}}^2 \mathbf{G}\mathbf{Z}\mathbf{Z}^H \mathbf{G}^H + \sigma_{n_{\mathrm{d}}}^2 \mathbf{I}_{N_{\mathrm{d}}}$ so that $\mathbf{A} = \mathbf{v}\mathbf{v}^H + \tilde{\mathbf{A}}$. Applying the following Sherman-Morrison formula

$$[\mathbf{v}\mathbf{v}^H + \tilde{\mathbf{A}}]^{-1} = \tilde{\mathbf{A}}^{-1} - \frac{\tilde{\mathbf{A}}^{-1}\mathbf{v}\mathbf{v}^H \tilde{\mathbf{A}}^{-1}}{1 + \mathbf{v}^H \tilde{\mathbf{A}}^{-1}\mathbf{v}} \tag{11}$$

into (8) and after some straightforward steps, we obtain

$$f = \frac{1}{1+\mathbf{v}^H\tilde{\mathbf{A}}^{-1}\mathbf{v}} = \frac{1}{1+\mathbf{w}^H\mathbf{H}^H\mathbf{Z}_G\mathbf{H}\mathbf{w}} \tag{12}$$

where

$$\mathbf{Z}_G = \mathbf{Z}^H\mathbf{G}^H[\sigma_{n_\mathrm{r}}^2\mathbf{G}\mathbf{Z}\mathbf{Z}^H\mathbf{G}^H + \sigma_{n_\mathrm{d}}^2\mathbf{I}_{N_\mathrm{d}}]^{-1}\mathbf{G}\mathbf{Z}. \tag{13}$$

Due to the formulation (12), the power constrained MSE minimization problem (10) can be written as[§]

$$\max_{\mathbf{w},\mathbf{Z}} \mathbf{w}^H\mathbf{H}^H\mathbf{Z}_G\mathbf{H}\mathbf{w} \text{ s. t. } P_\mathrm{r} \le P_\mathrm{R}^m,\ P_\mathrm{s} \le P_\mathrm{S}^m. \tag{14}$$

Before, solving the optimization problem (14), we present the following proposition to get some useful insights on (14).

Proposition 1: For the jointly optimal $\mathbf{w}$ *and* $\mathbf{Z}$*, the relay and source power constraints are satisfied with equality.*

Proof: In order to prove this proposition, we use the following properties from matrix theory[9].

Property 1: For any two positive definite Hermitian matrices $\mathbf{A}$ and $\mathbf{B}$, $\mathbf{A} \succeq \mathbf{B}$ if and only if $\mathbf{A}^{-1} \preceq \mathbf{B}^{-1}$.

Property 2: For any two Hermitian matrices $\mathbf{A}$ and $\mathbf{B}$, if $\mathbf{A} \succeq \mathbf{B}$, then $\mathbf{Q}^H\mathbf{A}\mathbf{Q} \succeq \mathbf{Q}^H\mathbf{B}\mathbf{Q}$ where $\mathbf{Q}$ is an arbitrary matrix.

We first show that the optimal $\mathbf{w}$ is the one for which the source power constraint $\mathbf{w}^H\mathbf{w} \le P_\mathrm{S}^m$ is satisfied with equality. Notice that the relay sum-power (9) can be rewritten as

$$\begin{aligned} P_\mathrm{r} &= \sum_{i=1}^{R}\left(\underbrace{[\mathbf{H}\mathbf{w}\mathbf{w}^H\mathbf{H}^H]_{i,i}}_{s_i(\mathbf{w})} + \sigma_{n_\mathrm{r}}^2\right)|z_i|^2 \\ &= \sum_{i=1}^{R} z_i\sqrt{s_i(\mathbf{w})+\sigma_{n_\mathrm{r}}^2}\, z_i^*\sqrt{s_i(\mathbf{w})+\sigma_{n_\mathrm{r}}^2} \end{aligned} \tag{15}$$

where $s_i(\mathbf{w}) = [\mathbf{H}\mathbf{w}\mathbf{w}^H\mathbf{H}^H]_{i,i}$ is the ith diagonal element of $\mathbf{H}\mathbf{w}\mathbf{w}^H\mathbf{H}^H$. Defining another optimization variable as $\bar{z}_i = z_i\sqrt{s_i(\mathbf{w})+\sigma_{n_\mathrm{r}}^2}$, the relay sum-power constraint can be expressed as $\sum_{i=1}^{R}|\bar{z}_i|^2 \le P_\mathrm{R}^m$. Thus, (14) can be written as

$$\begin{aligned} &\max_{\mathbf{w},\bar{\mathbf{Z}}} \mathbf{w}^H\mathbf{H}^H\underbrace{\tilde{\bar{\mathbf{Z}}}^H\mathbf{G}^H[\sigma_{n_\mathrm{r}}^2\mathbf{G}\tilde{\bar{\mathbf{Z}}}\tilde{\bar{\mathbf{Z}}}^H\mathbf{G}^H + \sigma_{n_\mathrm{d}}^2\mathbf{I}_{N_\mathrm{d}}]^{-1}\mathbf{G}\tilde{\bar{\mathbf{Z}}}}_{\bar{\mathbf{Z}}_G}\mathbf{H}\mathbf{w} \\ &\text{s. t. } \sum_{i=1}^{R}|\bar{z}_i|^2 \le P_\mathrm{R}^m,\ \mathbf{w}^H\mathbf{w} \le P_\mathrm{S}^m, \end{aligned} \tag{16}$$

where $\bar{\mathbf{Z}} = \mathrm{diag}(\bar{z}_1, \cdots, \bar{z}_R)$ and $\tilde{\bar{\mathbf{Z}}}$ is

$$\tilde{\bar{\mathbf{Z}}} = \mathrm{diag}\left(\frac{\bar{z}_1}{\sqrt{s_1(\mathbf{w})+\sigma_{n_\mathrm{r}}^2}}, \cdot\cdot, \frac{\bar{z}_R}{\sqrt{s_R(\mathbf{w})+\sigma_{n_\mathrm{r}}^2}}\right) \triangleq \mathbf{B}\bar{\mathbf{Z}} \tag{17}$$

Let $\mathbf{w}_\mathrm{o}$ and $\bar{\mathbf{Z}}_\mathrm{o}$ be the optimal solutions of (16), and $f_\mathrm{o}(\mathbf{w}_\mathrm{o}, \bar{\mathbf{Z}}_\mathrm{o})$ be the optimal objective value. Suppose that, for $\mathbf{w} = \mathbf{w}_\mathrm{o}$, the source power constraint is satisfied with inequality, i.e., $\mathbf{w}_\mathrm{o}^H\mathbf{w}_\mathrm{o} < P_\mathrm{S}^m$. Then, we can scale $\mathbf{w}_\mathrm{o}$

[§]We propose to solve this optimization problem at the destination which is assumed to have knowledge of all instantaneous channels. The destination is considered to send the optimized $\mathbf{w}$ and $\mathbf{Z}$ to the source and relays, respectively. The feedback requirement can be reduced by optimizing $\mathbf{w}$ and $\mathbf{Z}$ on the basis of the second-order statistics of the channels rather than their instantaneous values.

by $\sqrt{\alpha}$ with $\alpha > 1$, and make the power constraint active, i.e., $\alpha \mathbf{w}_{\rm o}^H \mathbf{w}_{\rm o} = P_{\rm S}^m$. For $\tilde{\mathbf{w}}_{\rm o} \triangleq \sqrt{\alpha}\mathbf{w}_{\rm o}$, the objective function in (16) can be given by

$$\tilde{f}_{\rm o}(\tilde{\mathbf{w}}_{\rm o}, \bar{\mathbf{Z}}_{\rm o}) = \mathbf{w}_{\rm o}^H \mathbf{H}^H \mathbf{D}_1^H \mathbf{G}^H \left[\sigma_{n_{\rm r}}^2 \mathbf{G}\mathbf{D}_2\mathbf{G}^H + \sigma_{n_{\rm d}}^2 \mathbf{I}_{N_{\rm d}}\right]^{-1} \mathbf{G}\mathbf{D}_1\mathbf{H}\mathbf{w}_{\rm o}. \tag{18}$$

where

$$\begin{aligned} \mathbf{D}_1 &\triangleq \operatorname{diag}\left(\frac{\bar{z}_1\sqrt{\alpha}}{\sqrt{\alpha s_1(\mathbf{w}_{\rm o}) + \sigma_{n_{\rm r}}^2}}, \cdots, \frac{\bar{z}_1\sqrt{\alpha}}{\sqrt{\alpha s_R(\mathbf{w}_{\rm o}) + \sigma_{n_{\rm r}}^2}}\right) \\ \mathbf{D}_2 &\triangleq \operatorname{diag}\left(\frac{|\bar{z}_1|^2}{\alpha s_1(\mathbf{w}_{\rm o}) + \sigma_{n_{\rm r}}^2}, \cdots, \frac{|\bar{z}_R|^2}{\alpha s_R(\mathbf{w}_{\rm o}) + \sigma_{n_{\rm r}}^2}\right). \end{aligned} \tag{19}$$

Since $\alpha > 1$, it is easy to see

$$\mathbf{D}_2 \preceq \mathbf{D}_3 \triangleq \tilde{\bar{\mathbf{Z}}}\tilde{\bar{\mathbf{Z}}}^H \tag{20}$$

which, according to the property 2, means that $\sigma_{n_{\rm r}}^2 \mathbf{G}\mathbf{D}_2\mathbf{G}^H + \sigma_{n_{\rm d}}^2 \mathbf{I}_{N_{\rm d}} \preceq \sigma_{n_{\rm r}}^2 \mathbf{G}\mathbf{D}_3\mathbf{G}^H + \sigma_{n_{\rm d}}^2 \mathbf{I}_{N_{\rm d}}$. Using property 1, the latter relation can be expressed as $[\sigma_{n_{\rm r}}^2 \mathbf{G}\mathbf{D}_2\mathbf{G}^H + \sigma_{n_{\rm d}}^2 \mathbf{I}_{N_{\rm d}}]^{-1} \succeq [\sigma_{n_{\rm r}}^2 \mathbf{G}\mathbf{D}_3\mathbf{G}^H + \sigma_{n_{\rm d}}^2 \mathbf{I}_{N_{\rm d}}]^{-1}$. Once again, using property 2, we can see that $\mathbf{D}_4 \triangleq \mathbf{G}^H[\sigma_{n_{\rm r}}^2 \mathbf{G}\mathbf{D}_2\mathbf{G}^H + \sigma_{n_{\rm d}}^2 \mathbf{I}_{N_{\rm d}}]^{-1}\mathbf{G} \succeq \mathbf{G}^H[\sigma_{n_{\rm r}}^2 \mathbf{G}\mathbf{D}_3\mathbf{G}^H + \sigma_{n_{\rm d}}^2 \mathbf{I}_{N_{\rm d}}]^{-1}\mathbf{G}$. Based on these discussions and noting that $\frac{\alpha}{\alpha s_i(\mathbf{w}_{\rm o}) + \sigma_{n_{\rm r}}^2} \geq \frac{1}{s_i(\mathbf{w}_{\rm o}) + \sigma_{n_{\rm r}}^2}$ for $\alpha \geq 1$, we can find that the new objective function $\tilde{f}_{\rm o}(\tilde{\mathbf{w}}_{\rm o}, \bar{\mathbf{Z}}_{\rm o}) = \mathbf{w}_{\rm o}^H \mathbf{H}^H \mathbf{D}_1^H \mathbf{D}_4 \mathbf{D}_1 \mathbf{H}\mathbf{w}_{\rm o} > f_{\rm o}(\mathbf{w}_{\rm o}, \bar{\mathbf{Z}}_{\rm o})$. Since the objective function is increased with the scaled $\mathbf{w}_{\rm o}$ that satisfies the source power constraint with equality, it can be concluded that, for the optimal solutions of $\mathbf{w}$ and $\bar{\mathbf{Z}}$, the source power constraint is active at optimality.

In order to show that the relay power constraint is also satisfied with equality at optimality, we use matrix inversion Lemma and rewrite $\bar{\mathbf{Z}}_G$ as

$$\bar{\mathbf{Z}}_G = \frac{1}{\sigma_{n_{\rm r}}^2}\mathbf{I}_R - \frac{\sigma_{n_{\rm d}}^2}{\sigma_{n_{\rm r}}^2}[\sigma_{n_{\rm r}}^2 \bar{\mathbf{Z}}^H \mathbf{B}^H \mathbf{G}^H \mathbf{G}\mathbf{B}\bar{\mathbf{Z}} + \sigma_{n_{\rm d}}^2 \mathbf{I}_R]^{-1}. \tag{21}$$

Suppose at the optimal solution $\bar{\mathbf{Z}} = \bar{\mathbf{Z}}_{\rm o}$, the relay power constraint is inactive, i.e., $\operatorname{tr}(\bar{\mathbf{Z}}_{\rm o}^H \bar{\mathbf{Z}}_{\rm o}) < P_{\rm R}^m$. Using a scaling factor $\sqrt{\beta}$ where $\beta > 1$, and new $\bar{\mathbf{Z}}_{\rm o}$ which is $\sqrt{\beta}\bar{\mathbf{Z}}_{\rm o}$, the inactive relay power constraint can be satisfied with equality. With this scaled (new) $\bar{\mathbf{Z}}_{\rm o}$, the objective function of (16) can be written as

$$\tilde{f}_{\rm o}(\mathbf{w}_{\rm o}, \sqrt{\beta}\bar{\mathbf{Z}}_{\rm o}) = \frac{\mathbf{w}_{\rm o}^H \mathbf{H}^H \mathbf{H}\mathbf{w}_{\rm o}}{\sigma_{n_{\rm r}}^2} - \frac{\sigma_{n_{\rm d}}^2}{\sigma_{n_{\rm r}}^2}\mathbf{w}_{\rm o}^H \mathbf{H}^H \left[\beta \bar{\mathbf{Z}}_{\rm o}^H \mathbf{B}^H \mathbf{G}^H \mathbf{G}\mathbf{B}\bar{\mathbf{Z}}_{\rm o}\sigma_{n_{\rm r}}^2 + \sigma_{n_{\rm d}}^2 \mathbf{I}_R\right]^{-1} \mathbf{H}\mathbf{w}_{\rm o} \tag{22}$$

Clearly, since $\beta > 1$, $\beta \bar{\mathbf{Z}}_{\rm o}^H \mathbf{B}^H \mathbf{G}^H \mathbf{G}\mathbf{B}\bar{\mathbf{Z}}_{\rm o} \succeq \bar{\mathbf{Z}}_{\rm o}^H \mathbf{B}^H \mathbf{G}^H \mathbf{G}\mathbf{B}\bar{\mathbf{Z}}_{\rm o}$. Using the properties 1 and 2, we can thus easily see that the second term of (22) decreases with the scaled $\bar{\mathbf{Z}}_{\rm o}$. This means that $\tilde{f}_{\rm o}(\mathbf{w}_{\rm o}, \sqrt{\beta}\bar{\mathbf{Z}}_{\rm o}) > f_{\rm o}(\mathbf{w}_{\rm o}, \bar{\mathbf{Z}}_{\rm o})$. As a result, the relay power constraint also must be satisfied with equality.

In the following, we propose two methods for optimizing $\mathbf{Z}$ and $\mathbf{w}$.

3.1 Method I-Iterative Approach

In this method, the relay coefficients and the source beamformer are alternatively optimized within a iterative framework.

Optimization over Z for the given $\mathbf{w} = \mathbf{w}_{\rm f}$: Again using matrix inversion lemma, we can rewrite (13) as

$$\mathbf{Z}_G = \frac{1}{\sigma_{n_{\rm r}}^2}\mathbf{I}_R - \frac{\sigma_{n_{\rm d}}^2}{\sigma_{n_{\rm r}}^2}[\sigma_{n_{\rm r}}^2 \mathbf{Z}^H \mathbf{G}^H \mathbf{G}\mathbf{Z} + \sigma_{n_{\rm d}}^2 \mathbf{I}_R]^{-1}. \tag{23}$$

Substituting (23) into (14), for the given $\mathbf{w} = \mathbf{w}_{\rm f}$ that satisfies the source power constraint, we can write the optimization problem (14) w.r.t. $\mathbf{Z}$ as

$$\begin{aligned} &\max_{\mathbf{Z}} \left\{t - a\mathbf{w}_{\rm f}^H \mathbf{H}^H [\sigma_{n_{\rm r}}^2 \mathbf{Z}^H \mathbf{G}^H \mathbf{G}\mathbf{Z} + \sigma_{n_{\rm d}}^2 \mathbf{I}_R]^{-1} \mathbf{H}\mathbf{w}_{\rm f}\right\} \\ &\text{s.t. } \mathbf{w}_{\rm f}^H \mathbf{H}^H \mathbf{Z}^H \mathbf{Z}\mathbf{H}\mathbf{w}_{\rm f} + \sigma_{n_{\rm r}}^2 \operatorname{tr}(\mathbf{Z}^H\mathbf{Z}) \leq P_{\rm R}^m \end{aligned} \tag{24}$$

where $t = \frac{\mathbf{w}_{\mathrm{f}}^H \mathbf{H}^H \mathbf{H} \mathbf{w}_{\mathrm{f}}}{\sigma_{n_{\mathrm{r}}}^2}$ and $a = \frac{\sigma_{n_{\mathrm{d}}}^2}{\sigma_{n_{\mathrm{r}}}^2}$. Since $\mathbf{Z}$ is a diagonal matrix, we can write

$$\mathbf{Z}^H \mathbf{G}^H \mathbf{G} \mathbf{Z} = [\mathbf{G}^H \mathbf{G}] \odot \tilde{\mathbf{Z}}^T \tag{25}$$

where $\tilde{\mathbf{Z}} = [z_1, \cdots, z_R]^T [z_1^*, \cdots, z_R^*]$ is a rank-one matrix. The relay sum-power constraint can be written in terms of $\tilde{\mathbf{Z}}$ as

$$\mathbf{w}_{\mathrm{f}}^H \mathbf{H}^H \mathbf{Z}^H \mathbf{Z} \mathbf{H} \mathbf{w}_{\mathrm{f}} + \sigma_{n_{\mathrm{r}}}^2 \mathrm{tr}(\mathbf{Z}^H \mathbf{Z}) = \mathrm{tr}\left(\underbrace{[\mathbf{H}\mathbf{w}_{\mathrm{f}} \mathbf{w}_{\mathrm{f}}^H \mathbf{H}^H + \sigma_{n_{\mathrm{r}}}^2 \mathbf{I}_R]}_{\tilde{\mathbf{H}}} (\mathbf{I}_R \odot \tilde{\mathbf{Z}}) \right) \leq P_{\mathrm{R}}^m. \tag{26}$$

The resulting optimization problem in terms of $\tilde{\mathbf{Z}}$ becomes

$$\begin{aligned} &\min_{\tilde{\mathbf{Z}}} \mathbf{w}_{\mathrm{f}}^H \mathbf{H}^H \left[[\mathbf{G}^H \mathbf{G}] \odot \tilde{\mathbf{Z}}^T \sigma_{n_{\mathrm{r}}}^2 + \sigma_{n_{\mathrm{d}}}^2 \mathbf{I}_R \right]^{-1} \mathbf{H} \mathbf{w}_{\mathrm{f}} \\ &\text{s.t. } \mathrm{tr}\left(\tilde{\mathbf{H}} (\mathbf{I}_R \odot \tilde{\mathbf{Z}}) \right) \leq P_{\mathrm{R}}^m, \ \mathrm{rank}(\tilde{\mathbf{Z}}) = 1. \end{aligned} \tag{27}$$

Introducing auxiliary variable $r \geq 0$, (27) can be written as

$$\begin{aligned} &\min_{\tilde{\mathbf{Z}}, r} r \text{ s. t. } r \geq \mathbf{w}_{\mathrm{f}}^H \mathbf{H}^H \left[[\mathbf{G}^H \mathbf{G}] \odot \tilde{\mathbf{Z}}^T \sigma_{n_{\mathrm{r}}}^2 + \sigma_{n_{\mathrm{d}}}^2 \mathbf{I} \right]^{-1} \mathbf{H} \mathbf{w}_{\mathrm{f}}, \\ &\mathrm{tr}\left(\tilde{\mathbf{H}} (\mathbf{I}_R \odot \tilde{\mathbf{Z}}) \right) \leq P_{\mathrm{R}}^m, \ \mathrm{rank}(\tilde{\mathbf{Z}}) = 1. \end{aligned} \tag{28}$$

Using the Schur-complement theorem[14], (28) can be written as

$$\begin{aligned} &\min_{\tilde{\mathbf{Z}}, r} r \\ &\text{s.t. } \begin{bmatrix} [\mathbf{G}^H \mathbf{G}] \odot \tilde{\mathbf{Z}}^T \sigma_{n_{\mathrm{r}}}^2 + \sigma_{n_{\mathrm{d}}}^2 \mathbf{I}_R & \mathbf{H} \mathbf{w}_{\mathrm{f}} \\ \mathbf{w}_{\mathrm{f}}^H \mathbf{H}^H & r \end{bmatrix} \succeq 0, \\ &\mathrm{tr}\left(\tilde{\mathbf{H}} (\mathbf{I}_R \odot \tilde{\mathbf{Z}}) \right) \leq P_{\mathrm{R}}^m, \ \mathrm{rank}(\tilde{\mathbf{Z}}) = 1. \end{aligned} \tag{29}$$

The objective and constraints of (29) are all convex except the rank-one constraint. Relaxing the latter constraint, we get the following convex optimization problem:

$$\begin{aligned} &\min_{\tilde{\mathbf{Z}}, r} r \\ &\text{s.t. } \begin{bmatrix} [\mathbf{G}^H \mathbf{G}] \odot \tilde{\mathbf{Z}}^T \sigma_{n_{\mathrm{r}}}^2 + \sigma_{n_{\mathrm{d}}}^2 \mathbf{I}_R & \mathbf{H} \mathbf{w}_{\mathrm{f}} \\ \mathbf{w}_{\mathrm{f}}^H \mathbf{H}^H & r \end{bmatrix} \succeq 0, \\ &\mathrm{tr}\left(\tilde{\mathbf{H}} (\mathbf{I}_R \odot \tilde{\mathbf{Z}}) \right) \leq P_{\mathrm{R}}^m, \ \tilde{\mathbf{Z}} \succeq 0. \end{aligned} \tag{30}$$

Notice that, due to rank relaxation, by solving (30), we only solve the convex approximation of the original problem (29). If the rank of $\tilde{\mathbf{Z}}$ is one, the solution of (30) is also optimal for (29). In this case, $\mathbf{Z}$ can be recovered from the eigen decomposition (ED) of $\tilde{\mathbf{Z}}$. However, if $\tilde{\mathbf{Z}}$ is not rank-one, the best rank-one solution should be approximated from $\tilde{\mathbf{Z}}$. In the literature[10–12], for various type of applications, it has been shown that Gaussian randomization technique gives very good approximations of rank-one solutions. Let the ED of $\tilde{\mathbf{Z}}$ be given by $\tilde{\mathbf{Z}} = \mathbf{U}_1 \mathbf{\Lambda}_1 \mathbf{U}_1^H$ and consider $\tilde{\mathbf{v}} \sim \mathcal{N}_C(\mathbf{0}, \mathbf{I}_R)$ i.e., the entries of $\tilde{\mathbf{v}}$ are i.i.d. ZMCSCG with unit variance. We generate the candidate vector $\mathbf{z}^{\mathrm{c}}$ as $\mathbf{z}^{\mathrm{c}} = \mathbf{U}_1 \mathbf{\Lambda}_1^{\frac{1}{2}} \tilde{\mathbf{v}}$. The vector $\mathbf{z}^{\mathrm{c}}$ requires to be properly scaled. Using $\sqrt{\mu} \mathbf{z}^{\mathrm{c}}$ in (27), where $\sqrt{\mu} > 0$, the optimization problem w.r.t. μ can be given as

$$\begin{aligned} &\min_{\mu} \mathbf{w}_{\mathrm{f}}^H \mathbf{H}^H \left[[\mathbf{G}^H \mathbf{G}] \odot \mu (\mathbf{z}^{\mathrm{c}} (\mathbf{z}^{\mathrm{c}})^H)^T \sigma_{n_{\mathrm{r}}}^2 + \sigma_{n_{\mathrm{d}}}^2 \mathbf{I}_R \right]^{-1} \mathbf{H} \mathbf{w}_{\mathrm{f}} \\ &\text{s.t. } \mu \mathrm{tr}\left(\tilde{\mathbf{H}} (\mathbf{I}_R \odot (\mathbf{z}^{\mathrm{c}} (\mathbf{z}^{\mathrm{c}})^H)) \right) \leq P_{\mathrm{R}}^m. \end{aligned} \tag{31}$$

The objective of (31) decreases monotonically with μ. Thus, it is very easy to see that the optimal scaling factor μ is

$$\mu = \frac{P_{\mathrm{R}}^m}{\mathrm{tr}\left(\tilde{\mathbf{H}}(\mathbf{I}_R \odot (\mathbf{z}^{\mathrm{c}}(\mathbf{z}^{\mathrm{c}})^H))\right)}. \tag{32}$$

The candidate vector $\sqrt{\mu}\mathbf{z}^{\mathrm{c}}$ is obtained for a particular realization of $\mathbf{v}$. The process is repeated for a number of realizations, and that candidate vector $\mathbf{z}^{\mathrm{c}}$ and the corresponding scaling factor $\sqrt{\mu}$ which give the minimum objective value are chosen.

Optimization over w for the given $\mathbf{Z} = \mathbf{Z}_{\mathrm{f}}$: For the given $\mathbf{Z} = \mathbf{Z}_{\mathrm{f}}$, (14) turns to the problem of maximizing a quadratic convex objective function with two quadratic convex inequality constraints. Due to the fact that the convex objective function has to be maximized, the optimization problem for the known $\mathbf{Z}$ is not a convex problem and cannot be solved using standard convex optimization tools. We propose to use semidefinite relaxation technique to solve (14) for the given $\mathbf{Z}$. Defining $\mathbf{W} = \mathbf{w}\mathbf{w}^H$ and relaxing the rank-one constraint on $\mathbf{W}$, we get the following SDP optimization problem

$$\begin{aligned} &\max_{\mathbf{W}} \mathrm{tr}\left(\mathbf{W}\mathbf{H}^H\mathbf{Z}_{G_{\mathrm{f}}}\mathbf{H}\right) \\ &\text{s.t. } \mathrm{tr}\left(\mathbf{W}\mathbf{H}^H\mathbf{Z}_{\mathrm{f}}^H\mathbf{Z}_{\mathrm{f}}\mathbf{H}\right) \leq P_{\mathrm{R}}^m - \sigma_{n_{\mathrm{r}}}^2 \mathrm{tr}(\mathbf{Z}_0^H\mathbf{Z}_0), \\ &\mathrm{tr}(\mathbf{W}) \leq P_{\mathrm{S}}^m,\ \mathbf{W} \succeq 0 \end{aligned} \tag{33}$$

where $\mathbf{Z}_{G_{\mathrm{f}}}$ is obtained after substituting $\mathbf{Z} = \mathbf{Z}_{\mathrm{f}}$ into $\mathbf{Z}_G$.
By analyzing the complementary slackness conditions[13] for the primal (33) and its dual problem, it can be shown that the gap between the optimal value of (33) and that of the corresponding non-relaxed problem is zero. The rank-one solution can be obtained from (33) using the method described for (30). Let $\sqrt{\gamma}\mathbf{w}^{\mathrm{c}}$ be the candidate vector generated from the randomization step. The scaling factor γ can be determined from the following expression:

$$\gamma = \min\left\{\frac{P_{\mathrm{S}}^m}{(\mathbf{w}^{\mathrm{c}})^H\mathbf{w}^{\mathrm{c}}}, \frac{P_{\mathrm{R}}^m - \sigma_{n_{\mathrm{r}}}^2 \mathrm{tr}(\mathbf{Z}_{\mathrm{f}}^H\mathbf{Z}_{\mathrm{f}})}{\mathrm{tr}\left(\mathbf{w}^{\mathrm{c}}(\mathbf{w}^{\mathrm{c}})^H\mathbf{H}^H\mathbf{Z}_{\mathrm{f}}^H\mathbf{Z}_{\mathrm{f}}\mathbf{H}\right)}\right\}. \tag{34}$$

The optimization problems (30) and (33) can be solved alternatively in an iterative manner as shown in **Algorithm 1**. After each update of $\tilde{\mathbf{Z}}$ and/or $\mathbf{W}$, the objective function decreases and, since the MSE function is

Algorithm 1 Iterative algorithm for optimizing $\mathbf{w}$ and $\mathbf{Z}$.

1. Set $n = 1$ and initialize $\mathbf{w}$ such that $\mathbf{w}^H\mathbf{w} = P_{\mathrm{S}}^m$.
2. **repeat**
   - Compute $\tilde{\mathbf{Z}}$ by solving (30).
   - Compute ED of $\tilde{\mathbf{Z}}$. If $\mathrm{rank}(\tilde{\mathbf{Z}}) = 1$, set $\mathbf{Z} = \sqrt{\alpha_z}\mathrm{diag}(\mathbf{u}_z)$ where $\mathbf{u}_z$ is the principal eigenvector and α_z is the corresponding eigenvalue. Otherwise, use the randomization technique for getting the best rank-one approximation for $\tilde{\mathbf{Z}}$.
   - Compute $\mathbf{W}$ by solving (33).
   - Compute the ED of $\mathbf{W}$. If its rank is one, then set $\mathbf{w} = \sqrt{\alpha_w}\mathbf{u}_w$ where $\mathbf{u}_w$ is the principal eigenvector of $\mathbf{W}$ and α_w is the eigenvalue. Otherwise, use randomization technique to get $\sqrt{\gamma}\mathbf{w}^{\mathrm{c}}$.
   - $n = n + 1$;

   until $|f_n - f_{n+1}| \leq \epsilon$, where ϵ is the required precision.

bounded above zero, the algorithm can be guaranteed to converge.

3.2 Method II - Relay optimization with MRT beamformer

The iterative optimization algorithm can be computationally costly when the numbers of relays and antennas at the source and destination increase. Here, we present an alternative approach where the source beamformer (also called MRT solution) maximizes $\mathbf{w}^H\mathbf{H}^H\mathbf{H}\mathbf{w}$, i.e., the total power radiated by the source towards the relays. It can be easily shown that the corresponding solution is

$$\mathbf{w} \triangleq \mathbf{w}_{\mathrm{MRT}} = \sqrt{P_{\mathrm{S}}^m}\mathbf{v}_{H,1} \tag{35}$$

where $\mathbf{v}_{H,1}$ is the column vector of $\mathbf{V}_H$ corresponding to the largest singular value σ_h^1 of $\mathbf{H}$ whose singular value decomposition is given by $\mathbf{H} = \mathbf{U}_H\mathbf{\Sigma}_H\mathbf{V}_H^H$. Here, $\mathbf{U}_H \in \mathcal{C}^{R\times R}$ and $\mathbf{V}_H \in \mathcal{C}^{N_{\mathrm{s}}\times N_{\mathrm{s}}}$ are the unitary matrices and $\mathbf{\Sigma}_H$ is the diagonal matrix with non-zero elements σ_h^i $(i = 1, \cdots, \min(R, N_{\mathrm{s}}))$ in decreasing order. Substituting (35) into (14), noting that $\sigma_h^1\mathbf{u}_{H,1} = \mathbf{U}_H\mathbf{\Sigma}_H\mathbf{V}_H\mathbf{v}_{H,1}$, and introducing an auxiliary variable $\tau \geq 0$, we can write (14) as

$$\begin{aligned}
&\max_{\mathbf{Z},\tau} \tau \quad \text{s.t.}\\
&\tau \leq \frac{1}{\sigma_{n_{\mathrm{r}}}^2} - a\mathbf{u}_{H,1}^H\left[\mathbf{Z}^H\mathbf{G}^H\mathbf{G}\mathbf{Z}\sigma_{n_{\mathrm{r}}}^2 + \sigma_{n_{\mathrm{d}}}^2\mathbf{I}_R\right]^{-1}\mathbf{u}_{H,1}\\
&\mathrm{tr}\left([P_{\mathrm{S}}^m(\sigma_h^1)^2\mathbf{u}_{H,1}\mathbf{u}_{H,1}^H + \sigma_{n_{\mathrm{r}}}^2\mathbf{I}_R]\mathbf{Z}^H\mathbf{Z}\right) \leq P_{\mathrm{R}}^m.
\end{aligned} \tag{36}$$

After using (25), the Schur-complement theorem[14] and rank relaxation, (36) can be finally given by

$$\begin{aligned}
&\max_{\tilde{\mathbf{Z}},\tau} \tau \quad \text{s.t.}\\
&\begin{bmatrix} [\mathbf{G}^H\mathbf{G}]\odot\tilde{\mathbf{Z}}^T\sigma_{n_{\mathrm{r}}}^2 + \sigma_{n_{\mathrm{d}}}^2\mathbf{I}_R & \sqrt{a}\mathbf{u}_{H,1}\\ \sqrt{a}\mathbf{u}_{H,1}^H & \frac{1}{\sigma_{n_{\mathrm{r}}}^2} - \tau \end{bmatrix} \succeq 0\\
&\mathrm{tr}\left([P_{\mathrm{S}}^m(\sigma_h^1)^2\mathbf{u}_{H,1}\mathbf{u}_{H,1}^H + \sigma_{n_{\mathrm{r}}}^2\mathbf{I}_R](\mathbf{I}_R\odot\tilde{\mathbf{Z}})\right) \leq P_{\mathrm{R}}^m,\ \tilde{\mathbf{Z}} \succeq 0
\end{aligned} \tag{37}$$

which is the SDP problem that can be solved efficiently using convex optimization tools (e.g.[15]). The remaining problem of approximating the best-rank one solution from $\tilde{\mathbf{Z}}$ of (37) is solved using Gaussian randomization method as described for the optimization problem (30).

Remark 1: The optimality of MRT solution (35) can be proven under some conditions. To this end, we propose following proposition.

Proposition 2: If signal transmissions from the relays to the destination occur over R orthogonal channels, under the approximation that $\sigma_{n_{\mathrm{d}}}^2\sigma_{n_{\mathrm{r}}}^2 << |\bar{z}_i|^2|\mathbf{g}_i|^2\sigma_{n_{\mathrm{r}}}^2 + \mathbf{w}^H\mathbf{h}_i^*\mathbf{h}_i^H\mathbf{w}\sigma_{n_{\mathrm{d}}}^2$, the MRT solution (35) is optimal if $||\mathbf{g}_i||^2 = ||\mathbf{g}||^2, \forall i \in \{1, \cdots, R\}$.

Proof: After substituting $\mathbf{G}^H\mathbf{G} = \mathrm{diag}(||\mathbf{g}_1||^2, \cdots, ||\mathbf{g}_R||^2)$ into (16), and using the approximation $\sigma_{n_{\mathrm{d}}}^2\sigma_{n_{\mathrm{r}}}^2 << |\bar{z}_i|^2||\mathbf{g}_i||^2\sigma_{n_{\mathrm{r}}}^2 + \mathbf{w}^H\mathbf{h}_i^*\mathbf{h}_i^H\mathbf{w}\sigma_{n_{\mathrm{d}}}^2$, we get

$$\begin{aligned}
&\max_{\mathbf{w},\tilde{\mathbf{Z}}} \sum_{i=1}^{R} \frac{1}{\frac{\sigma_{n_{\mathrm{r}}}^2}{\mathbf{w}^H\mathbf{h}_i^*\mathbf{h}_i^H\mathbf{w}} + \frac{\sigma_{n_{\mathrm{d}}}^2}{|\bar{z}_i|^2||\mathbf{g}_i||^2}} \quad \text{s. t.}\\
&\sum_{i=1}^{R}|\bar{z}_i|^2 = P_{\mathrm{R}}^m,\ \mathbf{w}^H\mathbf{w} = P_{\mathrm{S}}^m
\end{aligned} \tag{38}$$

It is clear that the phases of $\{\bar{z}_i\}_{i=1}^R$ do not affect the objective function and the constraints of the optimization problem (38). Thus, w.l.o.g, $\{\bar{z}_i\}_{i=1}^R$ can be considered to be real. For the given $\mathbf{w}$ with $\mathbf{w}^H\mathbf{w} = P_{\mathrm{S}}^m$, we can solve the optimization problem (38) w.r.t. $x_i \triangleq \bar{z}_i^2$ using Lagrangian multiplier method. For the case $||\mathbf{g}_i||^2 = ||\mathbf{g}||^2$, it is easy to show

$$x_i = \begin{cases} \frac{\sigma_{n_{\mathrm{d}}}b_i}{\sigma_{n_{\mathrm{r}}}^2||\mathbf{g}||^2}\left(\frac{||\mathbf{g}||}{\sqrt{\lambda}} - \sigma_{n_{\mathrm{d}}}\right) & \text{for } \lambda \leq \frac{||\mathbf{g}||^2}{\sigma_{n_{\mathrm{d}}}^2}\\ 0 & \text{otherwise} \end{cases} \tag{39}$$

where $b_i = \mathbf{w}^H \mathbf{h}_i^* \mathbf{h}_i^H \mathbf{w}$ and λ is Lagrangian multiplier. Substituting (39) into $\sum_{i=1}^{R} x_i = P_{\mathrm{R}}^m$, λ can be expressed as

$$\sqrt{\lambda} = \frac{||\mathbf{g}||\sigma_{n_{\mathrm{d}}} \sum_{i=1}^{R} b_i}{P_{\mathrm{R}}^m \sigma_{n_{\mathrm{r}}}^2 ||\mathbf{g}||^2 + \sigma_{n_{\mathrm{d}}}^2 \sum_{i=1}^{R} b_i}. \tag{40}$$

With the help of (40), x_i of (39) for $\lambda \leq \frac{||\mathbf{g}||^2}{\sigma_{n_{\mathrm{d}}}^2}$ can be expressed as

$$x_i = \frac{b_i P_{\mathrm{R}}^m}{\sum_{i=1}^{R} b_i} \tag{41}$$

whereas, for $\lambda > \frac{||\mathbf{g}||^2}{\sigma_{n_{\mathrm{d}}}^2}$, $x_i = 0$. Let $\tilde{f}$ be the objective function of the optimization problem (38). Substituting x_i from (41) into $\tilde{f}$, we obtain

$$\tilde{f} = \frac{P_{\mathrm{R}}^m ||\mathbf{g}||^2 \sum_{i=1}^{R} b_i}{||\mathbf{g}||^2 P_{\mathrm{R}}^m \sigma_{n_{\mathrm{r}}}^2 + \sigma_{n_{\mathrm{d}}}^2 \sum_{i=1}^{R} b_i}. \tag{42}$$

Since $\sum_{i=1}^{R} b_i = \mathbf{w}^H \sum_{i=1}^{R} \mathbf{h}_i^* \mathbf{h}_i^H \mathbf{w} = \mathbf{w}^H \mathbf{H}^H \mathbf{H} \mathbf{w}$, and $\tilde{f}$ monotonically increases with $\sum_{i=1}^{R} b_i$, $\tilde{f}$ is maximized when $\sum_{i=1}^{R} b_i$ is maximized. Thus, (35) is optimal.

4. NUMERICAL RESULTS

For all numerical simulations, we take $R = 4$, $N_{\mathrm{d}} = 5$, $P_{\mathrm{S}}^m = P_{\mathrm{R}}^m = P_0$ and all channel coefficients are taken as i.i.d. ZMCSCG variables with the unit variance. The average MSE is determined from 200 independent random channel realizations. In all cases, the relaxed optimization problems (30), (33) and (37) are solved using CVX software[15]. If the optimum solutions for $\tilde{\mathbf{Z}}$ in (30) and (37), and the optimum solution for $\mathbf{W}$ in (33) are rank-one, their principal components are used to determine $\mathbf{Z}$ and $\mathbf{w}$, respectively, otherwise the best rank-one solutions are obtained using the randomization technique. For iterative optimization (Method I), we take $\epsilon = 10^{-4}$ and we have noticed that the algorithm converges within 5-8 iterations. The average MSE as a function of $\mathrm{SNR}_{\mathrm{rd}} = \frac{P_0}{\sigma_{n_{\mathrm{d}}}^2}$,

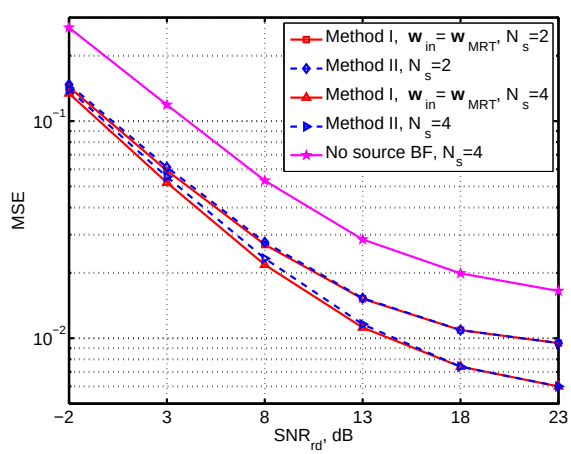

Figure 1. MSE as a function of $\mathrm{SNR}_{\mathrm{rd}} = \frac{P_0}{\sigma_{n_{\mathrm{d}}}^2}$.

is displayed in Fig. 1 for both the methods I and II using different values of N_{s}. The performance of method I is shown for the initialization $\mathbf{w}_{\mathrm{in}} = \mathbf{w}_{\mathrm{MRT}}$ (see (35)). For Fig. 1, we fix the relay noise power to -10 dBW, take $P_0 = 3$ dBW and vary $\sigma_{n_{\mathrm{d}}}^2$. It can be seen from Fig. 1 that the iterative method with $\mathbf{w}_{\mathrm{in}} = \mathbf{w}_{\mathrm{MRT}}$ provides better performance than method II. However, there is almost no difference between the two methods for high

$\mathrm{SNR_{rd}}$. This observation can be explained from the fact that, for a high $\mathrm{SNR_{rd}}$ (given that $\mathrm{SNR_{sr}} = \frac{P_0}{\sigma^2_{n_\mathrm{r}}}$ is fixed), the *source-relay* links dominate the MSE at the destination, and thus, MRT tends to be optimal solution. Fig. 2 shows the average MSE versus $\mathrm{SNR_{sr}} = \frac{P_0}{\sigma^2_{n_\mathrm{r}}}$, for both methods with different values of N_s. In Fig. 2,

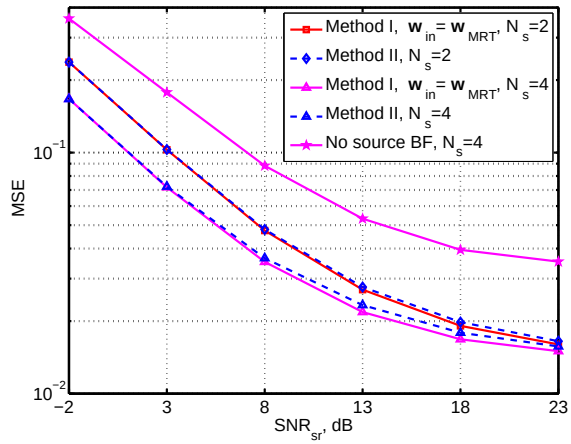

Figure 2. MSE as a function of $\mathrm{SNR_{sr}} = \frac{P_0}{\sigma^2_{n_\mathrm{r}}}$.

we take $P_0 = 3$ dBW, keep $\sigma^2_{n_\mathrm{d}}$ to -5 dBW and vary $\sigma^2_{n_\mathrm{r}}$. It can be observed from Fig. 2 that the difference between method I with $\mathbf{w}_\mathrm{in} = \mathbf{w}_\mathrm{MRT}$ and method II is negligible for low values of $\mathrm{SNR_{sr}}$ but visible for higher values of $\mathrm{SNR_{sr}}$. The latter observation can be explained from the fact that, for a low $\mathrm{SNR_{sr}}$ (given $\mathrm{SNR_{rd}}$ is fixed), the *source-relay* links dominate the end-to-end MSE and thus, MRT approaches to optimality. It can also be observed from Figs. 1 and 2, that the proposed methods significantly outperform the system that employs no beamforming at the source. The MSE error floor in both figures can be easily explained from the fact that either $\sigma^2_{n_\mathrm{r}}$ or $\sigma^2_{n_\mathrm{d}}$ is kept fixed. Although, method II gives same or slightly inferior performance compared to method I with the initialization $\mathbf{w}_\mathrm{in} = \mathbf{w}_\mathrm{MRT}$, the former requires solving only one relaxed SDP problem (which is (37)). This means that method II is more preferable than method I, if computational cost is a major concern for system design.

5. CONCLUSIONS

The joint optimization problem of the source beamformer and relay coefficients has been solved for a point-to-point system that consists of multi-antenna source and destination, and multiple single-antenna relays. An iterative optimization technique which alternatively optimizes the source and relay beamformers at a time, has been shown to have relaxed SDP formulation. The considered numerical examples show that, by keeping the source beamformer equal to the MRT solution of the source-relay channel and solving the optimization w.r.t. relay coefficients, the computational cost can be decreased significantly with a reasonably low performance loss.

REFERENCES

[1] Wang, B., Zhang, J., and Host-Madsen, A. "On the capacity of MIMO relay channels," *IEEE Trans. Inf. Theo.*, 51(1), 29-43 (2005).

[2] Munoz, O., Vidal, J., and Augustin, A. "Linear transceiver design in nonregenerative relays with channel state information," *IEEE Trans. Sig. Proc.*, 55(6), 2593-2604 (2007).

[3] Tang, X. and Hua, Y. "Optimal design of non-regenerative MIMO wireless relays," *IEEE Trans. Wireless Commun.*, 6(4), 1398-1407 (2007).

[4] Behbahani, A. S., Merched, R., and Eltawil, A. M. "Optimizations of a MIMO Relay Network," *IEEE Trans. Sig. Proc.*, 56(2), 5062-5073 (2008).

[5] Hammerstrom, H. and Wittneben, A. "Power allocation for amplify and forward MIMO-OFDM relay links," *IEEE Trans. Wireless Commun.*, 6(5), 1774-1786 (2007).

[6] Rong, Y., Tang, X., and Hua, Y. "A unified framework for optimizing linear nonregenerative multicarrier MIMO relay systems," *IEEE Trans. Sig. Proc.*, 57(12), 4837-4851 (2009).

[7] Khoshnevis, B., Yu, W., and Adve, R. "Grassmannian beamforming for MIMO amplify-and-forward relaying," *IEEE J. Sel. A. Commun.*, 26(8), 1397-1407 (2008).

[8] Li, C., Wang, X., Yang, L., and Zhu, W.-P. "Joint source-and-relay beamforming for multiple-input multiple-output systems with single-antenna distributed relays," *IET Communications*, 6(1), 62-68 (2012).

[9] Horn, R. A. and Johnson, C. A. [*Matrix Analysis*], Cambridge University Press (1985).

[10] Sidiropoulos, N. D., Davidson, T. N., and Luo, Z.-Q. "Transmit beamforming for physical layer multicasting," *IEEE Trans. Sig. Proc.*, 54(6), 2239-2251 (2006).

[11] Chalise, B. K. and Vandendorpe, L. "MIMO relay design for multipoint-to-multipoint communications with imperfect channel state information," *IEEE Trans. Sig. Proc.*, 57(7), 2785 - 2796 (2009).

[12] Chalise, B. K. and Vandendorpe, L. "Optimization of MIMO relays for multipoint-to-multipoint communications: Nonrobust and robust designs," *IEEE Trans. Sig. Proc.*, 58(12), 6355-6368 (2010).

[13] Huang, Y. and Zhang, S. "Complex matrix decomposition and quadratic programming," *Mathematics of Operations Research*, 32(3), 758768 (2007).

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

[15] M. Grant and S. Boyd, *CVX: Matlab Software for Disciplined Convex Programming*, http://stanford.edu/ boyd/cvx.