WIDELY-LINEAR TRANSCEIVER DESIGN FOR AMPLIFY-AND-FORWARD MIMO RELAYING

Donatella Darsena[†]*, Giacinto Gelli*[⋆]*, Ivan Iudice*[‡]*, and Francesco Verde*[⋆]

[†]Parthenope University, Naples, Italy (e-mail: darsena@uniparthenope.it)
[⋆]University Federico II, Naples, Italy [e-mail: (gelli, f.verde)@unina.it]
[‡]Italian Aerospace Research Centre, Capua (CE), Italy (e-mail: i.iudice@cira.it)

ABSTRACT

This paper deals with the problem of transceiver optimization for amplify-and-forward multiple-input multiple-output (MIMO) cooperative networks with multiple relays. We consider the case where precoding at the source, forwarding at the relays, and decoding at the destination are all carried out by means of widely-linear (WL) structures. On the basis of the minimum-mean-square-error criterion with appropriate power constraints, we propose a closed-from design for the WL-MIMO source precoder, relaying matrices, and destination decoder. Numerical results, in terms of bit-error rate, demonstrate that, when the disturbance exhibits improper or noncircularity features, the proposed design ensures significant performance gain with respect to its linear counterpart.

1. INTRODUCTION

Multiple-input multiple-output (MIMO) wireless relaying can increase system throughput, overcoming the adverse effects of fading and expanding network coverage [1–7]. Such benefits, however, can be fully achieved only if all the cooperating transceivers are judiciously optimized. Constrained optimization of cooperative networks with *amplify-and-forward (AF)* relays has tackled a great bulk of research activity (see, e.g., [8–13] and references therein), by assuming that full channel state information (CSI) on the relevant MIMO links is known to all the cooperating nodes.

While optimization of the *one source–one relay–one destination (1S-1R-1D)* network configuration, with power constraints at both the source and the relay, can be carried out in closed-form as in traditional MIMO systems [14], the same techniques cannot be used for the more general *one source–multiple relays–one destination (1S-MR-1D)* topology, due to the block-diagonal form of its constituent building blocks. Henceforth, with reference to the latter case, most existing works rely on iterative schemes to jointly optimize the source precoder, the forwarding matrices, and the destination decoder, with power constraints at the source, as well as at the relays [11, 13] or at the destination [8]. Moreover, *linear* complex-valued transceiver structures are considered in all the aforementioned papers, under the assumption that the transmitted symbols and disturbances are modeled as *proper* [15] or *circular* [16] random processes.

The aim of this paper is to generalize linear MIMO designs for 1S-MR-1D network configurations, by considering instead *widely-linear (WL)* MIMO complex processing [17–20] at the source, at the multiple relays, and at the destination. As a matter of fact, WL precoders and decoders for MIMO non-cooperative (i.e., without relaying) systems have been originally proposed in [21] and, subsequently, in [22]. Optimization of a WL cooperative 1S-1R-1D network with single-antenna nodes has been carried out in [23]. Herein, adopting the minimum-mean-square-error (MMSE) criterion, under the assumption that the disturbances are *improper* or *noncircular*, we develop closed-form expressions for the WL-MIMO source precoder, WL-MIMO forwarding matrices at the relays, and WL-MIMO destination decoder. The effectiveness of the proposed approach is corroborated by the results of performance analysis, carried out by means of Monte Carlo computer simulations.

2. NETWORK MODEL AND BASIC ASSUMPTIONS

Let us consider a 1S-MR-1D network model with K_{R} relays, where all nodes are equipped with multiple antennas, with $(N_{\text{S}}, N_{\text{R}}, N_{\text{D}})$ denoting the number of antennas at the source, relays, and destination. The relays adopt the AF protocol and operate in half-duplex mode; moreover, no direct link exists between the source and the destination. All the wireless links are modeled as quasi-static frequency-flat channels.

In each transmission period, the source transmits a block $\mathbf{s} \in \mathbb{C}^{N_{\text{B}}}$ of $N_{\text{B}} \leq N_{\text{S}}$ proper symbols, with mean $\mathbb{E}(\mathbf{s}) = \mathbf{0}_{N_{\text{B}}}$, correlation matrix $\mathbb{E}(\mathbf{s}\,\mathbf{s}^{\text{H}}) = \sigma_s^2\,\mathbf{I}_{N_{\text{B}}}$, and conjugate correlation matrix $\mathbb{E}(\mathbf{s}\,\mathbf{s}^{\text{T}}) = \mathbf{O}_{N_{\text{B}}}$. We consider WL signal processing [17] at each node: at the source, in particular, vector $\mathbf{s}$ and its conjugate version $\mathbf{s}^*$ are jointly processed

$$\mathbf{x} = \mathbf{B}_1\,\mathbf{s} + \mathbf{B}_2\,\mathbf{s}^* = \mathbf{B}\,\tilde{\mathbf{s}} \tag{1}$$

The work of I. Iudice was partially supported by the Italian National Project TECVOL-II.

by means of the *precoding* matrices $\mathbf{B}_1 \in \mathbb{C}^{N_S \times N_B}$ and $\mathbf{B}_2 \in \mathbb{C}^{N_S \times N_B}$, where $\mathbf{B} \triangleq [\mathbf{B}_1, \mathbf{B}_2] \in \mathbb{C}^{N_S \times 2N_B}$ and $\tilde{\mathbf{s}} \triangleq [\mathbf{s}^T, \mathbf{s}^H]^T \in \mathbb{C}^{2N_B}$.

The transmission is divided into two phases: in *Phase I*, the source directly transmits $\mathbf{x}$ to the relays; in *Phase II*, the relays precode the received data and forward the resulting blocks towards the destination. Let $\mathcal{K}_R \triangleq \{1, 2, \ldots, K_R\}$, the received signal at relay $\ell \in \mathcal{K}_R$ during Phase I is given by

$$\mathbf{z}_\ell = \mathbf{H}_\ell \, \mathbf{x} + \mathbf{w}_\ell \tag{2}$$

where $\mathbf{H}_\ell \in \mathbb{C}^{N_R \times N_S}$ is the *first-hop* or *backward* MIMO channel matrix between the source and the ℓth relay, whereas $\mathbf{w}_\ell \in \mathbb{C}^{N_R}$ models the disturbance at the ℓth relay. We assume that $\mathbb{E}[\mathbf{w}_\ell] = \mathbf{0}_{N_R}$, $\mathbb{E}[\mathbf{w}_\ell \, \mathbf{w}_\ell^H] = \sigma^2_{w_\ell} \, \mathbf{I}_{N_R}$, and $\mathbb{E}[\mathbf{w}_\ell \, \mathbf{w}_\ell^T] = \rho_{w_\ell} \, \mathbf{I}_{N_R}$, with $\rho_{w_\ell} \in \mathbb{C}$. Moreover, $\mathbf{w}_{\ell_1}$ is statistically independent of $\mathbf{x}$ and of $\mathbf{w}_{\ell_2}$, for $\ell_1 \neq \ell_2 \in \mathcal{K}_R$; it is proper if and only if (iff) $\rho_{w_\ell} = 0$, otherwise it is improper [16].

For each $\ell \in \mathcal{K}_R$, the vector $\mathbf{z}_\ell$ in (2) is subject to a WL transformation through the *forwarding* matrices $\mathbf{F}_{1,\ell} \in \mathbb{C}^{N_R \times N_R}$ and $\mathbf{F}_{2,\ell} \in \mathbb{C}^{N_R \times N_R}$, hence yielding

$$\mathbf{y}_\ell = \mathbf{F}_{1,\ell} \, \mathbf{z}_\ell + \mathbf{F}_{2,\ell} \, \mathbf{z}_\ell^* \, . \tag{3}$$

A compact form of the K_R relations (3) can be obtained as

$$\mathbf{y} \triangleq [\mathbf{y}_1^T, \mathbf{y}_2^T, \ldots, \mathbf{y}_{K_R}^T]^T = \mathbf{F}_1 \, \mathbf{z} + \mathbf{F}_2 \, \mathbf{z}^* = \mathbf{F} \, \tilde{\mathbf{z}} \tag{4}$$

where, for $i \in \{1, 2\}$, $\mathbf{F}_i \triangleq \text{diag}(\mathbf{F}_{i,1}, \mathbf{F}_{i,2}, \ldots, \mathbf{F}_{i,K_R}) \in \mathbb{C}^{N_C \times N_C}$, and $N_C \triangleq K_R \, N_R$, $\mathbf{F} \triangleq [\mathbf{F}_1, \mathbf{F}_2] \in \mathbb{C}^{N_C \times 2N_C}$. According to (2), the following vectors have been also defined

$$\mathbf{z} \triangleq [\mathbf{z}_1^T, \mathbf{z}_2^T, \ldots, \mathbf{z}_{K_R}^T]^T = \mathbf{H} \, \mathbf{x} + \mathbf{w} \tag{5}$$

$$\tilde{\mathbf{z}} \triangleq [\mathbf{z}^T, \mathbf{z}^H]^T = \widetilde{\mathbf{H}} \, \tilde{\mathbf{x}} + \tilde{\mathbf{w}} \tag{6}$$

where $\mathbf{H} \triangleq [\mathbf{H}_1^T, \mathbf{H}_2^T, \ldots, \mathbf{H}_{K_R}^T]^T \in \mathbb{C}^{N_C \times N_S}$, $\mathbf{w} \triangleq [\mathbf{w}_1^T, \mathbf{w}_2^T, \ldots, \mathbf{w}_{K_R}^T]^T \in \mathbb{C}^{N_C}$, $\tilde{\mathbf{w}} \triangleq [\mathbf{w}^T, \mathbf{w}^H]^T \in \mathbb{C}^{2N_C}$, and

$$\widetilde{\mathbf{H}} \triangleq \begin{bmatrix} \mathbf{H} & \mathbf{O}_{N_C \times N_S} \\ \mathbf{O}_{N_C \times N_S} & \mathbf{H}^* \end{bmatrix} \in \mathbb{C}^{2N_C \times 2N_S} \, . \tag{7}$$

Accounting for (1), we additionally observe that

$$\tilde{\mathbf{x}} \triangleq [\mathbf{x}^T, \mathbf{x}^H]^T = \widetilde{\mathbf{B}} \, \tilde{\mathbf{s}} \tag{8}$$

where

$$\widetilde{\mathbf{B}} \triangleq [\mathbf{B}^T, \mathbf{\Pi}_{N_B} \, \mathbf{B}^H]^T = \begin{bmatrix} \mathbf{B}_1 & \mathbf{B}_2 \\ \mathbf{B}_2^* & \mathbf{B}_1^* \end{bmatrix} \in \mathbb{C}^{2N_S \times 2N_B} \, . \tag{9}$$

The matrix $\mathbf{\Pi}_n \in \mathbb{R}^{2n \times 2n}$ in (9) is a permutation matrix

$$\mathbf{\Pi}_n \triangleq \begin{bmatrix} \mathbf{O}_{n \times n} & \mathbf{I}_n \\ \mathbf{I}_n & \mathbf{O}_{n \times n} \end{bmatrix} \tag{10}$$

satisfying $\mathbf{\Pi}_n^2 = \mathbf{I}_{2n}$ and $\mathbf{\Pi}_n^T = \mathbf{\Pi}_n$.

In Phase II, at the destination, the received signal is

$$\mathbf{r} = \mathbf{G} \, \mathbf{y} + \mathbf{n} \tag{11}$$

where $\mathbf{G} \triangleq [\mathbf{G}_1, \mathbf{G}_2, \ldots, \mathbf{G}_{K_R}] \in \mathbb{C}^{N_D \times N_C}$, with $\mathbf{G}_\ell \in \mathbb{C}^{N_D \times N_R}$ being the *second-hop* or *forward* MIMO channel matrix between the ℓth relay and the destination, and $\mathbf{n} \in \mathbb{C}^{N_D}$ is the disturbance vector. We assume that $\mathbf{n}$ is statistically independent of $\mathbf{y}$, having $\mathbb{E}[\mathbf{n}] = \mathbf{0}_{N_D}$, $\mathbb{E}[\mathbf{n} \, \mathbf{n}^H] = \sigma_n^2 \, \mathbf{I}_{N_D}$, and $\mathbb{E}[\mathbf{n} \, \mathbf{n}^T] = \rho_n \, \mathbf{I}_{N_D}$, with $\rho_n \in \mathbb{C}$. The disturbance vector $\mathbf{n}$ is proper iff $\rho_n = 0$, otherwise it is improper [16].

Vector $\mathbf{r}$ is subject to WL equalization by means of the *decoding* matrices $\mathbf{D}_1 \in \mathbb{C}^{N_B \times N_D}$ and $\mathbf{D}_2 \in \mathbb{C}^{N_B \times N_D}$, which yields an estimate

$$\mathbf{e} \triangleq \mathbf{D}_1 \, \mathbf{r} + \mathbf{D}_2 \, \mathbf{r}^* = \mathbf{D} \, \tilde{\mathbf{r}} \tag{12}$$

of the source block $\mathbf{s}$, whose entries are quantized to the nearest symbols of the considered constellation set; moreover, $\mathbf{D} \triangleq [\mathbf{D}_1, \mathbf{D}_2] \in \mathbb{C}^{N_B \times 2N_D}$, and, by virtue of (11) and (4),

$$\tilde{\mathbf{r}} \triangleq [\mathbf{r}^T, \mathbf{r}^H]^T = \widetilde{\mathbf{G}} \, \tilde{\mathbf{y}} + \tilde{\mathbf{n}} \tag{13}$$

$$\tilde{\mathbf{y}} \triangleq [\mathbf{y}^T, \mathbf{y}^H]^T = \widetilde{\mathbf{F}} \, \tilde{\mathbf{z}} \tag{14}$$

with $\tilde{\mathbf{n}} \triangleq [\mathbf{n}^T, \mathbf{n}^H]^T \in \mathbb{C}^{2N_D}$, and

$$\widetilde{\mathbf{G}} \triangleq \begin{bmatrix} \mathbf{G} & \mathbf{O}_{N_D \times N_C} \\ \mathbf{O}_{N_D \times N_C} & \mathbf{G}^* \end{bmatrix} \in \mathbb{C}^{2N_D \times 2N_C} \tag{15}$$

$$\widetilde{\mathbf{F}} \triangleq [\mathbf{F}^T, \mathbf{\Pi}_{N_C} \, \mathbf{F}^H]^T = \begin{bmatrix} \mathbf{F}_1 & \mathbf{F}_2 \\ \mathbf{F}_2^* & \mathbf{F}_1^* \end{bmatrix} \in \mathbb{C}^{2N_C \times 2N_C} \, . \tag{16}$$

Taking into account (6), (8), (13), (14), one obtains

$$\tilde{\mathbf{r}} = \widetilde{\mathbf{G}} \, \widetilde{\mathbf{F}} \, \widetilde{\mathbf{H}} \, \widetilde{\mathbf{B}} \, \tilde{\mathbf{s}} + \widetilde{\mathbf{G}} \, \widetilde{\mathbf{F}} \, \tilde{\mathbf{w}} + \tilde{\mathbf{n}} = \widetilde{\mathbf{C}} \, \tilde{\mathbf{s}} + \tilde{\mathbf{v}} \tag{17}$$

where $\widetilde{\mathbf{C}} \triangleq \widetilde{\mathbf{G}} \, \widetilde{\mathbf{F}} \, \widetilde{\mathbf{H}} \, \widetilde{\mathbf{B}} \in \mathbb{C}^{2N_D \times 2N_B}$ represents the *dual-hop channel matrix* and $\tilde{\mathbf{v}} \triangleq \widetilde{\mathbf{G}} \, \widetilde{\mathbf{F}} \, \tilde{\mathbf{w}} + \tilde{\mathbf{n}} \in \mathbb{C}^{2N_D}$ is the overall disturbance term. It is noteworthy that the constituent building blocks of the considered 1S-MR-1D MIMO network with WL processing exhibit special structures. In particular, $\widetilde{\mathbf{G}}$, $\widetilde{\mathbf{F}}$, $\widetilde{\mathbf{H}}$, and $\widetilde{\mathbf{B}}$ are *centrohermitian* 2*-by-*2 *blockwise matrices* [24], whereas $\tilde{\mathbf{s}}$, $\tilde{\mathbf{w}}$, and $\tilde{\mathbf{n}}$ are *conjugate symmetric vectors* [21].

Hereinafter, the aim is to jointly optimize the matrices $\mathbf{B}$, $\mathbf{F}$, and $\mathbf{D}$ under the assumption that full CSI is available at each node in the network.

3. WIDELY-LINEAR TRANSCEIVER DESIGN

Most performance measures in vector optimization problems involve the mean square error (MSE) matrix $\mathbf{E}(\mathbf{B}, \mathbf{F}, \mathbf{D}) \triangleq \mathbb{E}[(\mathbf{e} - \mathbf{s}) \, (\mathbf{e} - \mathbf{s})^H]$. Indeed, for $n \in \{1, 2, \ldots, N_B\}$, the MSE on the nth symbol stream is the nth diagonal element $\text{MSE}_n(\mathbf{B}, \mathbf{F}, \mathbf{D}) \triangleq \{\mathbf{E}(\mathbf{B}, \mathbf{F}, \mathbf{D})\}_{nn}$ of $\mathbf{E}(\mathbf{B}, \mathbf{F}, \mathbf{D})$. To assess the global performance of the network, one might use any function of the *per-stream* MSEs, i.e., $\{\text{MSE}_n(\mathbf{B}, \mathbf{F}, \mathbf{D})\}_{n=1}^{N_B}$, that is increasing in each argument [14]. Herein, we consider as global performance measure the sum of the diagonal elements of $\mathbf{E}(\mathbf{B},\mathbf{F},\mathbf{D})$, i.e.,

$$\mathrm{MSE}(\mathbf{B},\mathbf{F},\mathbf{D}) \triangleq \mathrm{tr}[\mathbf{E}(\mathbf{B},\mathbf{F},\mathbf{D})] = \sum_{n=1}^{N_{\mathrm{B}}} \mathrm{MSE}_n(\mathbf{B},\mathbf{F},\mathbf{D}) \,. \tag{18}$$

Minimization of (18) with respect to $\mathbf{B}$, $\mathbf{F}$, and $\mathbf{D}$ must be carried out under appropriate power constraints involving $\mathbf{B}$ and $\mathbf{F}$; the choice of such constraints greatly influences the possibility of obtaining closed-form solutions.

In order to limit the average transmit power of the source (in units of energy per transmission) $\mathcal{P}_{\mathrm{S}}(\mathbf{B}) \triangleq \mathbb{E}[\|\mathbf{x}\|^2] = \sigma_s^2\, \mathrm{tr}(\widetilde{\mathbf{B}}\,\widetilde{\mathbf{B}}^{\mathrm{H}})$, the constraint $\mathrm{tr}(\widetilde{\mathbf{B}}\,\widetilde{\mathbf{B}}^{\mathrm{H}}) \le \gamma_{\mathrm{S}}$, where $\gamma_{\mathrm{S}} > 0$, is typically used. As to $\mathbf{F}$, a viable approach [8, 10] consists of imposing a constraint on the average power received at the destination in Phase II which, accounting for (13) and (14), can be upper bounded as

$$\begin{aligned}\mathcal{P}_{\mathrm{D}}(\mathbf{F}) \triangleq \mathbb{E}[\|\widetilde{\mathbf{G}}\,\widetilde{\mathbf{F}}\,\tilde{\mathbf{z}}\|^2] &= \mathrm{tr}(\widetilde{\mathbf{G}}\,\widetilde{\mathbf{F}}\,\mathbf{R}_{\tilde{\mathbf{z}}\tilde{\mathbf{z}}}\,\widetilde{\mathbf{F}}^{\mathrm{H}}\,\widetilde{\mathbf{G}}^{\mathrm{H}}) \\ &\le \mathrm{tr}(\mathbf{R}_{\tilde{\mathbf{z}}\tilde{\mathbf{z}}})\,\mathrm{tr}(\widetilde{\mathbf{G}}\,\widetilde{\mathbf{F}}\,\widetilde{\mathbf{F}}^{\mathrm{H}}\,\widetilde{\mathbf{G}}^{\mathrm{H}})\end{aligned} \tag{19}$$

where $\mathbf{R}_{\tilde{\mathbf{z}}\tilde{\mathbf{z}}} \triangleq \mathbb{E}[\tilde{\mathbf{z}}\,\tilde{\mathbf{z}}^{\mathrm{H}}] = \sigma_s^2\,\widetilde{\mathbf{H}}\,\widetilde{\mathbf{B}}\,\widetilde{\mathbf{B}}^{\mathrm{H}}\,\widetilde{\mathbf{H}}^{\mathrm{H}} + \mathbf{R}_{\tilde{\mathbf{w}}\tilde{\mathbf{w}}}$, $\mathbf{R}_{\tilde{\mathbf{w}}\tilde{\mathbf{w}}} \triangleq \mathbb{E}[\tilde{\mathbf{w}}\,\tilde{\mathbf{w}}^{\mathrm{H}}] \in \mathbb{C}^{2N_{\mathrm{C}}\times 2N_{\mathrm{C}}}$, and the inequality stems from the matrix trace inequality for positive semidefinite matrices. By virtue of (19), instead of directly limiting $\mathcal{P}_{\mathrm{D}}(\mathbf{F})$, we impose the constraint $\mathrm{tr}(\widetilde{\mathbf{G}}\,\widetilde{\mathbf{F}}\,\widetilde{\mathbf{F}}^{\mathrm{H}}\,\widetilde{\mathbf{G}}^{\mathrm{H}}) \le \gamma_{\mathrm{R}}$, with $\gamma_{\mathrm{R}} > 0$, which ensures that $\mathcal{P}_{\mathrm{D}}(\mathbf{F}) \le \gamma_{\mathrm{R}}\,[\gamma_{\mathrm{S}}\,\sigma_s^2\,\mathrm{tr}(\widetilde{\mathbf{H}}\,\widetilde{\mathbf{H}}^{\mathrm{H}}) + \mathrm{tr}(\mathbf{R}_{\tilde{\mathbf{w}}\tilde{\mathbf{w}}})]$. The aforementioned constraint allows one to simplify the derivation of the optimal precoding and forwarding matrices. Therefore, the proposed design problem can be formulated as

$$\min_{\mathbf{B},\mathbf{F},\mathbf{D}} \mathrm{MSE}(\mathbf{B},\mathbf{F},\mathbf{D}) \text{ s.t. } \begin{cases}\mathrm{tr}(\widetilde{\mathbf{B}}\,\widetilde{\mathbf{B}}^{\mathrm{H}}) \le \gamma_{\mathrm{S}}\,; \\ \mathrm{tr}(\widetilde{\mathbf{G}}\,\widetilde{\mathbf{F}}\,\widetilde{\mathbf{F}}^{\mathrm{H}}\,\widetilde{\mathbf{G}}^{\mathrm{H}}) \le \gamma_{\mathrm{R}}\,.\end{cases} \tag{20}$$

The following theorem (whose proof is omitted) simplifies the search for the solution of (20).

Theorem 1 *The matrices $\mathbf{B}_{opt}$, $\mathbf{F}_{opt}$, and $\mathbf{D}_{opt}$ are solutions of the original problem (20) iff the augmented matrices*

$$\tilde{\mathbf{B}}_{opt} \triangleq [\mathbf{B}_{opt}^T, \mathbf{\Pi}_{N_B}\,\mathbf{B}_{opt}^H]^T \in \mathbb{C}^{2N_S\times 2N_B} \tag{21}$$
$$\tilde{\mathbf{F}}_{opt} \triangleq [\mathbf{F}_{opt}^T, \mathbf{\Pi}_{N_C}\,\mathbf{F}_{opt}^H]^T \in \mathbb{C}^{2N_C\times 2N_C} \tag{22}$$
$$\tilde{\mathbf{D}}_{opt} \triangleq [\mathbf{D}_{opt}^T, \mathbf{\Pi}_{N_D}\,\mathbf{D}_{opt}^H]^T \in \mathbb{C}^{2N_B\times 2N_D} \tag{23}$$

are solutions of the constrained minimization problem

$$\min_{\widetilde{\mathbf{B}},\widetilde{\mathbf{F}},\widetilde{\mathbf{D}}} \tilde{MSE}(\widetilde{\mathbf{B}},\widetilde{\mathbf{F}},\widetilde{\mathbf{D}}) \text{ s.t. } \begin{cases} tr(\widetilde{\mathbf{B}}\widetilde{\mathbf{B}}^H) \le 2\,\gamma_S\,; \\ tr(\widetilde{\mathbf{G}}\widetilde{\mathbf{F}}\,\widetilde{\mathbf{F}}^H\,\widetilde{\mathbf{G}}^H) \le 2\,\gamma_R\,;\end{cases} \tag{24}$$

where $\tilde{MSE}(\widetilde{\mathbf{B}},\widetilde{\mathbf{F}},\widetilde{\mathbf{D}}) \triangleq tr[\widetilde{\mathbf{E}}(\widetilde{\mathbf{B}},\widetilde{\mathbf{F}},\widetilde{\mathbf{D}})]$ is the augmented MSE matrix, with

$$\widetilde{\mathbf{D}} \triangleq [\mathbf{D}^T, \mathbf{\Pi}_{N_D}\,\mathbf{D}^H]^T = \begin{bmatrix}\mathbf{D}_1 & \mathbf{D}_2 \\ \mathbf{D}_2^* & \mathbf{D}_1^*\end{bmatrix} \in \mathbb{C}^{2N_B\times 2N_D}$$

$$\begin{aligned}\widetilde{\mathbf{E}}(\widetilde{\mathbf{B}},\widetilde{\mathbf{F}},\widetilde{\mathbf{D}}) &\triangleq \mathbb{E}[(\tilde{\mathbf{e}}-\tilde{\mathbf{s}})\,(\tilde{\mathbf{e}}-\tilde{\mathbf{s}})^H] \\ &= \sigma_s^2\,(\widetilde{\mathbf{D}}\,\widetilde{\mathbf{C}} - \mathbf{I}_{2N_B})(\widetilde{\mathbf{D}}\,\widetilde{\mathbf{C}} - \mathbf{I}_{2N_B})^H + \widetilde{\mathbf{D}}\,\mathbf{R}_{\tilde{\mathbf{v}}\tilde{\mathbf{v}}}\,\widetilde{\mathbf{D}}^H\end{aligned} \tag{25}$$

$\tilde{\mathbf{e}} \triangleq [\mathbf{e}^T, \mathbf{e}^H]^T \in \mathbb{C}^{2N_B}$, whereas $\mathbf{R}_{\tilde{\mathbf{v}}\tilde{\mathbf{v}}} \triangleq \mathbf{E}(\tilde{\mathbf{v}}\,\tilde{\mathbf{v}}^H) = \widetilde{\mathbf{G}}\,\widetilde{\mathbf{F}}\,\mathbf{R}_{\tilde{\mathbf{w}}\tilde{\mathbf{w}}}\,\widetilde{\mathbf{F}}^H\,\widetilde{\mathbf{G}}^H + \mathbf{R}_{\tilde{\mathbf{n}}\tilde{\mathbf{n}}}$ and $\mathbf{R}_{\tilde{\mathbf{n}}\tilde{\mathbf{n}}} \triangleq \mathbf{E}(\tilde{\mathbf{n}}\,\tilde{\mathbf{n}}^H) \in \mathbb{C}^{2N_D\times 2N_D}$.

Theorem 1 allows one to reformulate the original optimization problem (20) as in (24), which admits an easier solution. It is well-known (see, e.g., [14]) that, for fixed matrices $\widetilde{\mathbf{B}}$ and $\widetilde{\mathbf{F}}$, problem (24) becomes convex quadratic in the matrix $\widetilde{\mathbf{D}}$ and the corresponding closed-form optimal solution is given by the Wiener filtering matrix

$$\begin{aligned}\widetilde{\mathbf{D}}_{\mathrm{mmse}} &= \sigma_s^2\,\widetilde{\mathbf{C}}^{\mathrm{H}}\,(\sigma_s^2\,\widetilde{\mathbf{C}}\,\widetilde{\mathbf{C}}^{\mathrm{H}} + \mathbf{R}_{\tilde{\mathbf{v}}\tilde{\mathbf{v}}})^{-1} \\ &= \sigma_s^2\,(\mathbf{I}_{2N_{\mathrm{B}}} + \sigma_s^2\,\widetilde{\mathbf{C}}^{\mathrm{H}}\,\mathbf{R}_{\tilde{\mathbf{v}}\tilde{\mathbf{v}}}^{-1}\,\widetilde{\mathbf{C}})^{-1}\,\widetilde{\mathbf{C}}^{\mathrm{H}}\,\mathbf{R}_{\tilde{\mathbf{v}}\tilde{\mathbf{v}}}^{-1}\end{aligned} \tag{26}$$

where the second expression comes from the matrix inversion lemma. Substituting (26) into (25) yields

$$\widetilde{\mathbf{E}}(\widetilde{\mathbf{B}},\widetilde{\mathbf{F}}) \triangleq \mathbf{E}(\widetilde{\mathbf{B}},\widetilde{\mathbf{F}},\widetilde{\mathbf{D}}_{\mathrm{mmse}}) = \sigma_s^2\,(\mathbf{I}_{2N_{\mathrm{B}}} + \sigma_s^2\,\widetilde{\mathbf{C}}^{\mathrm{H}}\mathbf{R}_{\tilde{\mathbf{v}}\tilde{\mathbf{v}}}^{-1}\widetilde{\mathbf{C}})^{-1}\,. \tag{27}$$

Therefore, let $\tilde{\mathrm{MSE}}(\widetilde{\mathbf{B}},\widetilde{\mathbf{F}}) \triangleq \mathrm{tr}[\widetilde{\mathbf{E}}(\widetilde{\mathbf{B}},\widetilde{\mathbf{F}})]$, problem (24) boils down to minimizing $\tilde{\mathrm{MSE}}(\widetilde{\mathbf{B}},\widetilde{\mathbf{F}})$ under the power constraints $\mathrm{tr}(\widetilde{\mathbf{H}}\,\widetilde{\mathbf{B}}\,\widetilde{\mathbf{B}}^{\mathrm{H}}\,\widetilde{\mathbf{H}}^{\mathrm{H}}) \le 2\,\gamma_{\mathrm{S}}$ and $\mathrm{tr}(\widetilde{\mathbf{G}}\,\widetilde{\mathbf{F}}\,\widetilde{\mathbf{F}}^{\mathrm{H}}\,\widetilde{\mathbf{G}}^{\mathrm{H}}) \le 2\,\gamma_{\mathrm{R}}$. In order to derive closed-form expressions for such a problem, we rely on Lemma 1 (whose proof is omitted).

Lemma 1 *The matrix $\tilde{MSE}(\widetilde{\mathbf{B}},\widetilde{\mathbf{F}})$ is lower bounded as*

$$\begin{aligned}\tilde{MSE}(\widetilde{\mathbf{B}},\widetilde{\mathbf{F}}) &= tr\left[(\mathbf{I}_{2N_B} + \sigma_s^2\,\widetilde{\mathbf{C}}^H\,\mathbf{R}_{\tilde{\mathbf{v}}\tilde{\mathbf{v}}}^{-1}\,\widetilde{\mathbf{C}})^{-1}\right] \\ &\ge tr\left[(\mathbf{I}_{2N_B} + \sigma_s^2\,\widetilde{\mathbf{C}}^H\,\widetilde{\mathbf{C}})^{-1}\right] \triangleq \tilde{MSE}_{LB}(\widetilde{\mathbf{B}},\widetilde{\mathbf{F}})\,.\end{aligned} \tag{28}$$

By virtue of Lemma 1, and substituting the expression of $\widetilde{\mathbf{C}}$ [see (17)] in $\tilde{\mathrm{MSE}}_{\mathrm{LB}}(\widetilde{\mathbf{B}},\widetilde{\mathbf{F}})$, the precoding and forwarding matrices are synthesized as

$$\begin{aligned}&\min_{\widetilde{\mathbf{B}},\widetilde{\mathbf{F}}} \mathrm{tr}\left[(\mathbf{I}_{2N_{\mathrm{B}}} + \sigma_s^2\,\widetilde{\mathbf{B}}^{\mathrm{H}}\,\widetilde{\mathbf{H}}^{\mathrm{H}}\,\widetilde{\mathbf{F}}^{\mathrm{H}}\,\widetilde{\mathbf{G}}^{\mathrm{H}}\,\widetilde{\mathbf{G}}\,\widetilde{\mathbf{F}}\,\widetilde{\mathbf{H}}\,\widetilde{\mathbf{B}})^{-1}\right] \\ &\qquad \text{s.t.} \begin{cases}\mathrm{tr}(\widetilde{\mathbf{B}}\,\widetilde{\mathbf{B}}^{\mathrm{H}}) \le 2\,\gamma_{\mathrm{S}}\,; \\ \mathrm{tr}(\widetilde{\mathbf{G}}\,\widetilde{\mathbf{F}}\,\widetilde{\mathbf{F}}^{\mathrm{H}}\,\widetilde{\mathbf{G}}^{\mathrm{H}}) \le 2\,\gamma_{\mathrm{R}}\,.\end{cases}\end{aligned} \tag{29}$$

Following [14], it can be proven that $\tilde{\mathrm{MSE}}_{\mathrm{LB}}(\widetilde{\mathbf{B}},\widetilde{\mathbf{F}})$ as a function of $\widetilde{\mathbf{B}}$ and $\widetilde{\mathbf{Q}} \triangleq \widetilde{\mathbf{G}}\,\widetilde{\mathbf{F}} \in \mathbb{C}^{2N_{\mathrm{D}}\times 2N_{\mathrm{C}}}$, which are dense matrices, is both Schur-concave and Schur-convex and, then, the channel-diagonalizing structure of the precoding-forwarding processing is optimal. Consequently, the solution of the problem (29) is obtained [14] by parameterizing $\widetilde{\mathbf{B}}$ and $\widetilde{\mathbf{Q}}$ in terms of the singular value decomposition (SVD) of the augmented overall backward channel matrix $\widetilde{\mathbf{H}}$, thus having

$$\widetilde{\mathbf{B}} = \widetilde{\mathbf{V}}\,\widetilde{\boldsymbol{\Sigma}} \quad \text{and} \quad \widetilde{\mathbf{Q}} = \widetilde{\boldsymbol{\Delta}}\,\widetilde{\mathbf{U}}^{\mathrm{H}} \tag{30}$$

where $\widetilde{\mathbf{V}} \in \mathbb{C}^{2N_\mathrm{S}\times\tilde{L}}$ has as columns the right-singular vectors of $\widetilde{\mathbf{H}}$ corresponding to its largest $\tilde{L} \triangleq \min\{2N_\mathrm{B}, \mathrm{rank}(\widetilde{\mathbf{H}})\}$ singular values arranged in increasing order,

$$\widetilde{\boldsymbol{\Sigma}} = [\mathrm{diag}(\tilde{\sigma}_1, \tilde{\sigma}_2, \ldots, \tilde{\sigma}_{\tilde{L}}), \mathbf{O}_{\tilde{L}\times(2N_\mathrm{B}-\tilde{L})}] \in \mathbb{R}^{\tilde{L}\times 2N_\mathrm{B}} \tag{31}$$

with $\tilde{\sigma}_1, \tilde{\sigma}_2, \ldots, \tilde{\sigma}_{\tilde{L}}$ being scalar variables to be further optimized,[1] $\widetilde{\mathbf{U}} \in \mathbb{C}^{2N_\mathrm{C}\times\tilde{L}}$ collects the left-singular vectors of $\widetilde{\mathbf{H}}$ corresponding to its largest $\tilde{L}$ singular values, and

$$\widetilde{\boldsymbol{\Delta}} = [\mathrm{diag}(\tilde{\delta}_1, \tilde{\delta}_2, \ldots, \tilde{\delta}_{\tilde{L}}), \mathbf{O}_{(2N_\mathrm{D}-\tilde{L})\times\tilde{L}}]^{\mathrm{T}} \in \mathbb{R}^{2N_\mathrm{D}\times\tilde{L}} \tag{32}$$

with $\tilde{\delta}_1, \tilde{\delta}_2, \ldots, \tilde{\delta}_{\tilde{L}}$ being real (see footnote 1) scalar variables to be further optimized. Since $\widetilde{\mathbf{H}}$ is a centrohermitian 2-by-2 blockwise matrix [24], its singular vectors can be assumed to be conjugate symmetric [21], i.e., $\boldsymbol{\Pi}_{N_\mathrm{S}}\widetilde{\mathbf{V}}^* = \widetilde{\mathbf{V}}$ and $\boldsymbol{\Pi}_{N_\mathrm{C}}\widetilde{\mathbf{U}}^* = \widetilde{\mathbf{U}}$. In the following, we assume for simplicity that $N_\mathrm{S} \leq N_\mathrm{C}$ and the overall backward channel matrix $\mathbf{H}$ is full-column rank. Hence, one has $\mathrm{rank}(\widetilde{\mathbf{H}}) = \mathrm{rank}(\mathbf{H}) + \mathrm{rank}(\mathbf{H}^*) = 2N_\mathrm{S}$, which in its turn implies that $\tilde{L} = 2N_\mathrm{B}$, since $N_\mathrm{B} \leq N_\mathrm{S}$ by assumption.

Relying on (30), the complicated matrix optimization problem (29) can be simplified to the following one

$$\min_{\tilde{\alpha}_1,\ldots,\tilde{\alpha}_{\tilde{L}},\tilde{\beta}_1,\ldots,\tilde{\beta}_{\tilde{L}}} f_0(\tilde{\alpha}_1, \ldots, \tilde{\alpha}_{\tilde{L}}, \tilde{\beta}_1, \ldots, \tilde{\beta}_{\tilde{L}})$$
$$\text{s.t.} \begin{cases} \sum_{\ell=1}^{\tilde{L}} \tilde{\alpha}_\ell \leq 2\,\gamma_\mathrm{S}\,; \\ \sum_{\ell=1}^{\tilde{L}} \tilde{\beta}_\ell \leq 2\,\gamma_\mathrm{R}\,; \\ \tilde{\alpha}_\ell > 0, \quad \tilde{\beta}_\ell > 0\,; \end{cases} \tag{33}$$

where $\tilde{\alpha}_\ell \triangleq \tilde{\sigma}_\ell^2$, $\tilde{\beta}_\ell \triangleq \tilde{\delta}_\ell^2$, and

$$f_0(\tilde{\alpha}_1, \ldots, \tilde{\alpha}_{\tilde{L}}, \tilde{\beta}_1, \ldots, \tilde{\beta}_{\tilde{L}}) \triangleq \sum_{\ell=1}^{\tilde{L}} \left(1 + \sigma_s^2\,\tilde{\lambda}_\ell\,\tilde{\alpha}_\ell\,\tilde{\beta}_\ell\right)^{-1} \tag{34}$$

with $\tilde{\lambda}_1 \geq \tilde{\lambda}_2 \geq \cdots \geq \tilde{\lambda}_{\tilde{L}}$ denoting the largest $\tilde{L}$ singular values of $\widetilde{\mathbf{H}}$. Convexity of the optimization problem (33) can be characterized by proving that the objective function (34) is a sum of convex functions [25]. By symmetry, one can infer that $\tilde{\alpha}_\ell = \tilde{\beta}_\ell$, for $\ell = \{1, 2, \ldots, \tilde{L}\}$, when $\gamma_\mathrm{S} = \gamma_\mathrm{R} \triangleq \gamma$. In this case, it can be readily shown that $(1 + \sigma_s^2\,\tilde{\lambda}_\ell\,\tilde{\alpha}_\ell^2)^{-1}$ are convex decreasing functions when $\tilde{\alpha}_\ell > (\sqrt{3}\,\sigma_s\,|\tilde{\lambda}_\ell|)^{-1}$, for each $\ell = \{1, 2, \ldots, \tilde{L}\}$. It is worth noting that such a condition boils down to $w_\ell > 0$ for moderate-to-high signal-to-noise ratio (SNR) values, i.e., when σ_s^2 is sufficiently large.

Therefore, the optimization problem (33) can be efficiently solved using the interior-point method [25]. At this point, by exploiting the centrohermitian 2-by-2 blockwise structures of the matrices $\widetilde{\mathbf{B}}$ and $\widetilde{\mathbf{Q}}$ given by (30), and assuming that the channel matrices $\mathbf{G}_\ell$ are full-row rank for each $\ell \in \mathcal{K}_\mathrm{R}$, one can obtain the corresponding closed-from solution for $\mathbf{B}$ and $\mathbf{F}$.

[1] Such variables are supposed to be real without loss of generality.

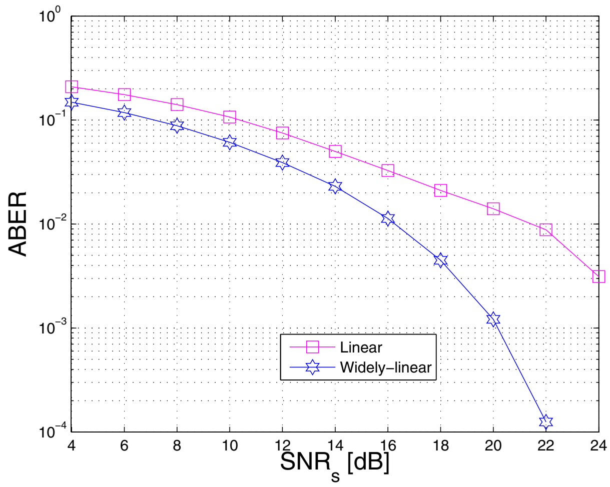

Fig. 1. Average BER versus SNR_s.

4. NUMERICAL ANALYSIS

We considered a cooperative network with $(N_\mathrm{S}, N_\mathrm{R}, N_\mathrm{D}) = (2, 2, 2)$. The source transmits QPSK symbols, with $\sigma_s^2 = 1$ and $N_\mathrm{B} = 2$, over a Rayleigh fading channel: hence, the entries of the channel matrices $\mathbf{H}$ and $\mathbf{G}$ were modeled as independent and identically-distributed (i.i.d.) zero-mean circularly symmetric complex Gaussian (CSCG) random variables, with variances σ_s^2/N_S and σ_r^2/N_C, respectively.

The noise at the relays $\mathbf{w}$ and at the destination $\mathbf{n}$ was generated according to the linear model $\mathbf{d} = \boldsymbol{\Upsilon}\,\mathbf{i} + \mathbf{u}$ [26], i.e., as the sum of improper interference and proper thermal noise $\mathbf{u}$, where $\mathbf{i} \in \mathbb{C}^m$ is composed of BPSK symbols and is independent of $\mathbf{u}$, $\boldsymbol{\Upsilon} \in \mathbb{C}^{n\times m}$ collects the channel coefficients between the interference source and the receiver, whose entries are i.i.d. zero-mean unit-variance CSCG random variables, and m and n are properly chosen. It can be shown that the resulting correlation matrix is $\mathbf{R}_{\mathbf{dd}} = \mathbf{R}_{\mathbf{uu}} + \boldsymbol{\Upsilon}\,\mathbf{R}_{\mathbf{ii}}\,\boldsymbol{\Upsilon}^{\mathrm{H}}$, where $\mathbf{R}_{\mathbf{uu}} \triangleq \mathbb{E}[\mathbf{u}\,\mathbf{u}^{\mathrm{H}}] \in \mathbb{C}^{n\times n}$ and $\mathbf{R}_{\mathbf{ii}} \triangleq \mathbb{E}[\mathbf{i}\,\mathbf{i}^{\mathrm{H}}] \in \mathbb{C}^{m\times m}$, whereas the conjugate correlation matrix is $\mathbf{R}_{\mathbf{dd}^*} = \boldsymbol{\Upsilon}\,\mathbf{R}_{\mathbf{ii}^*}\,\boldsymbol{\Upsilon}^{\mathrm{T}}$. Moreover, we defined $\mathrm{SNR}_s = \sigma_s^2\gamma_s N_\mathrm{C}/N_\mathrm{S}$ and $\mathrm{SNR}_r = \sigma_r^2\gamma_r N_\mathrm{D}/N_\mathrm{C}$ as the SNR for the source-to-relays link and the relays-to-destination link, respectively, and set SNR_r to 10 dB.

The average bit-error-rate (ABER) versus SNR_s is plotted in Fig. 1. Results show that the WL transceiver significantly outperforms the linear one, assuring performance gain of about 5 dB in the moderate-to-high SNR_s range.

5. REFERENCES

[1] H. Bolcskei, R.U. Nabar, O. Oyman, and A.J. Paulraj, "Capacity scaling laws in MIMO relay networks", *IEEE Trans. Wireless Commun.*, vol. 5, pp. 1433-1444, June 2006.

[2] Y.J. Fan and J. Thompson, "MIMO configurations for relay channels: Theory and practice", *IEEE Trans. Wireless Commun.*, vol. 6, pp. 1774-1786, May 2007.

[3] F. Verde and A. Scaglione, "Randomized space-time block coding for distributed amplify-and-forward cooperative relays", *Proc. of IEEE Int. Conf. Acoustics, Speech and Signal Process. (ICASSP)*, Dallas, Texas, Mar. 2010, pp. 3030-3033.

[4] F. Verde, "Design of randomized space-time block codes for amplify-and-forward cooperative relaying", *Proc. of 5th Int. Symp. on Commun. Control and Signal Process. (ISCCSP)*, Rome, Italy, May 2012, pp. 1-5.

[5] F. Verde, "Performance analysis of randomized space-time block codes for amplify-and-forward cooperative relaying," *IET Commun.*, vol. 67, pp. 1–16, Dec. 2013.

[6] D. Darsena, G. Gelli, F. Melito, and F. Verde, "Error probability analysis of a MIMO system with noncoherent relaying and MMSE reception", *IEEE Commun. Lett.*, vol. 18, pp. 1198-1201, July 2014.

[7] D. Darsena, G. Gelli, F. Melito, and F. Verde, "Performance analysis of amplify-and-forward multiple-relay MIMO systems with ZF reception", *IEEE Trans. Veh. Technol.*, vol. 64, July 2015.

[8] A.S. Behbahani, R. Merched, and A.M. Eltawil, "Optimizations of a MIMO relay network," *IEEE. Trans. Signal Process.*, vol. 56 no. 10, pp. 5062–5073, Oct. 2008.

[9] Y. Rong, X. Tang, and Y. Hua, "A unified framework for optimizing linear nonrigenerative multicarrier MIMO relay communication systems," *IEEE. Trans. Signal Process.*, vol. 57 no. 12, pp. 4837–4851, Dec. 2009.

[10] A. Toding, M. Khandaker, and Y. Rong, "Joint source and relay optimization for parallel MIMO relay networks," *EURASIP J. Advances Signal Process.*, 2012, 2012:174.

[11] C. Zhao and B. Champagne, "Joint design of multiple non-regenerative MIMO relaying matrices with power constraints," *IEEE Trans. Signal Process.*, vol. 61, no. 10, pp. 4861–4873, Oct. 2013.

[12] D. Darsena, G. Gelli, F. Melito, and F. Verde, "Optimization of coherent amplify-and-forward cooperative transmissions in spatially-multiplexed MIMO-OFDM systems", in *Proc. of European Conf. on Networks and Commun. (EUCNC)*, Bologna, Italy, June 2014, pp. 1-5.

[13] C. Zhao and B. Champagne, "A unified approach to optimal transceiver design for nonregenerative MIMO relaying," *IEEE Trans. Veh. Technol.*, vol. 64, no. 7, pp. 2938–2951, July 2015.

[14] D. Palomar, J. Cioffi, and M.-A. Lagunas, "Joint tx-rx beamforming design for multicarrier MIMO channels: a unified framework for convex optimization," *IEEE. Trans. Signal Process.*, vol. 51 no. 9, pp. 2381–2401, Sep. 2003.

[15] B. Picinbono, "On circularity", *IEEE Trans. Signal Process.*, pp. 3473–3482, Dec. 1994.

[16] F.D. Neeser and J.L. Massey, "Proper complex random processes with applications to information theory," *IEEE Trans. Information Theory*, pp. 1293–1302, July 1993.

[17] B. Picinbono and P. Chevalier, "Widely linear estimation with complex data", *IEEE Trans. Signal Process.*, pp. 2030–2033, Aug. 1995.

[18] W.H. Gerstacker, R. Schober, and A. Lampe, "Receivers with widely linear processing for frequency-selective channels," *IEEE Trans. Commun.*, vol. 51, pp. 1512–1523, Sep. 2003.

[19] D. Darsena, G. Gelli, L. Paura, and F. Verde, "Widely-linear equalization and blind channel identification for interference-contaminated multicarrier systems," *IEEE Trans. Signal Processing*, pp. 1163–1177, March 2005.

[20] D. Darsena, G. Gelli, and F. Verde, "Universal linear precoding for NBI-proof widely-linear equalization in MC systems," *Eurasip Journal on Wireless Communications and Networking*, vol. 2008, Article ID 321450.

[21] D. Darsena, G. Gelli, and F. Verde, "Widely-linear precoders and decoders for MIMO channels", in *Proc. of the Tenth International Symposium on Wireless Communication Systems (ISWCS)*, Ilmenau, Germany, Aug. 2013, pp. 587-591.

[22] S. Lagen, A. Agustin, and J. Vidal, "Improper Gaussian signaling for the Z-interference channel", *Proc. of IEEE Int. Conf. Acoustics, Speech and Signal Process. (ICASSP)*, Florence, Italy, May 2014, pp. 1140-1144.

[23] Z. Ho, A. Zappone, and E. Jorswieck, "Optimal nonregenerative relay processing with improper signals", in *Proc. of the Tenth International Symposium on Wireless Communication Systems (ISWCS)*, Ilmenau, Germany, Aug. 2013, pp. 567-571.

[24] A. Lee, "Centrohermitian and skew-centrohermitian matrices," *Linear Algebra and its Applications*, vol. 29, pp. 205-210, 1980.

[25] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge University Press, 2004.

[26] D. Darsena, G. Gelli, L. Paura, F. Verde, "A constrained maximum-SINR NBI-resistant receiver for OFDM systems," *IEEE Trans. Signal Process.*, vol. 55, pp. 3032-3047, June 2007.