Article

Separable MSE-based design of two-way multiple-relay cooperative MIMO 5G networks

Donatella Darsena[1,4], Giacinto Gelli[2,4], Ivan Iudice[3], and Francesco Verde[2,4]

[1] Department of Engineering, Parthenope University, Naples I-80143, Italy
[2] Department of Electrical Engineering and Information Technology, University of Naples Federico II, Italy
[3] Italian Aerospace Research Centre (CIRA), Capua I-81043, Italy
[4] National Inter-University Consortium for Telecommunications (CNIT)
* Correspondence: f.verde@unina.it (F.V.)

Version October 31, 2020 submitted to Sensors

Abstract: While the combination of multi-antenna and relaying techniques has been extensively studied for Long Term Evolution Advanced (LTE-A) and Internet of Things (IoT) applications, it is expected to still play an important role in 5th Generation (5G) networks. However, the expected benefits of these technologies cannot be achieved without a proper system design. In this paper, we consider the problem of jointly optimizing terminal precoders/decoders and relay forwarding matrices on the basis of the sum mean square error (MSE) criterion in multiple-input multiple-output (MIMO) two-way relay systems, where two multi-antenna nodes mutually exchange information via multi-antenna amplify-and-forward relays. This problem is nonconvex and a local optimal solution is typically found by using iterative algorithms based on alternating optimization. We show how the constrained minimization of the sum-MSE can be relaxed to obtain two separated subproblems which, under mild conditions, admit a closed-form solution. Compared to iterative approaches, the proposed design is more suited to be integrated in 5G networks, since it is computationally more convenient and its performance exhibits a better scaling in the number of relays.

Keywords: Amplify-and-forward (non-regenerative) relays; minimum-mean-square-error criterion; multiple-input multiple-output (MIMO) systems; optimization; two-way relaying

1. Introduction

Cooperative multiple-input multiple-output (MIMO) communication techniques, wherein data exchange between MIMO terminal nodes is assisted by one or multiple MIMO relays, have been studied for Long Term Evolution Advanced (LTE-A) cellular systems [1–3], since they assure significant performance gains in terms of coverage, reliability, and capacity. Relay technology has been also considered for Internet of Things (IoT) applications, by allowing in particular the support of the massive access for fog and social networking services [4–6]. One of the main changes when going from LTE-A to 5th generation (5G) systems is the spectrum use at radically higher frequencies in the millimeter-wave (mmWave) range [7]. However, mmWave signals are highly susceptible not only to blockages from large-size structures, e.g., buildings, but they are also severely attenuated by the presence of small-size objects, e.g., human bodies and foliage [8]. In this regard, cooperative MIMO technology additionally represents a possible approach for circumventing the unreliability of mmWave channels [9] in 5G networks.

In addition, 5G systems have stringent requirements in terms of spectral efficiency. Many relaying protocols operate in *half-duplex* mode [10–13], where two time slots are required to perform a single transmission, due to the inability of the relays to receive and transmit at the same time. To overcome the inherent halving of spectral efficiency, a possible remedy for 5G applications is to adopt *two-way* relaying [14] (see Fig. 1), which works as follows: (i) in the first slot, the two terminal nodes simultaneously transmit their precoded signals to the relays; (ii) in the second slot, the relays precode and forward the received signals to the terminals. Since each terminal knows its own transmitted signal, the effects of self-interference can be subtracted from the received signal at the terminals, and the data of interest can be decoded. On the other side of the coin, with respect to the one-way relaying setting, the optimization of two-way cooperative networks is complicated by fact that terminal precoders/decoders and relay forwarding matrices are coupled among themselves.

Design and performance analysis of two-way cooperative MIMO networks encompassing multiple *amplify-and-forward* (AF) or *non-regenerative* relays has been considered in [15–19]. Compared with the single-relay case [20], the multiple-relay scenario generally leads to more challenging *nonconvex* constrained optimization problems, which are usually solved by burdensome iterative procedures. In [15], by adopting a weighted sum-mean-square-error (MSE) or a sum-rate cost function, iterative gradient descent optimization algorithms are proposed, with transmit-power constraints imposed at both the terminals and the relays. A similar scenario is considered in [16] and [17]. In [16], the original constrained minimum sum-MSE nonconvex optimization problem is iteratively solved. Specifically, the algorithm of [16] starts by randomly choosing the terminal precoders and the relay forwarding matrices satisfying the transmission power constraints at the source terminals and the relay nodes. In each iteration, the terminal precoders, the relay forwarding matrices, and the decoders are alternatingly updated in [16] through solving convex subproblems: first, with given precoders and relay forwarding matrices, the optimal decoders are obtained in closed-form by solving an unconstrained convex problem; second, with fixed precoders and decoders, the relaying matrix of all the relays are updated in closed-form one-by-one by freezing the relaying matrices of the other relays; finally, given the decoders and relaying matrices, the precoders are updated by solving a convex quadratically constrained quadratic programming problem. A different iterative optimization procedure is proposed in [17], based on the matrix conjugate gradient algorithm, which is shown to converge faster than conventional gradient descent methods. Finally, some recent papers [18,19] propose architectures for two-way relaying based on relay/antenna selection strategies.

In this paper, we propose an optimization algorithm for two-way AF MIMO relaying 5G networks, where terminal precoders/decoders and relay forwarding matrices are jointly derived under power constraints on the transmitted/received power at the terminals. Rather than attempting to solve it iteratively, we derive a relaxed version of the original minimum sum-MSE nonconvex optimization, which allows one to decompose it in two separate problems that admit a closed-form, albeit suboptimal, solution. We show by Monte Carlo trials that our closed-form approach performs comparably or better than representative iterative approaches proposed in the literature for the same scenario with a reduced computational complexity, especially for increasing values of the number of relays.

2. Network model and basic assumptions

We consider the two-way MIMO 5G network configuration of Fig. 1, where bidirectional communication between two terminals, equipped with $N_{\mathrm{T},1}$ and $N_{\mathrm{T},2}$ antennas, respectively, is assisted by N_{C} half-duplex relays, each equipped with N_{R} antennas. We assume that there is no direct link between the two terminals, due to high path loss values or obstructions. Even though our approach can be generalized, for simplicity, the considered physical layer is that of a single-carrier cooperative system where all the channel links are quasi static and experience flat fading.

Let $\mathbf{s}_1 \in \mathbb{C}^{N_{\mathrm{S},1}}$ and $\mathbf{s}_2 \in \mathbb{C}^{N_{\mathrm{S},2}}$ denote the symbol vectors to be transmitted by terminal 1 and 2, respectively. In the first time slot, each terminal precodes its symbols with matrix $\mathbf{P}_i \in \mathbb{C}^{N_{\mathrm{T},i} \times N_{\mathrm{S},i}}$, for $i \in \{1,2\}$, before transmitting it to the relays, which thus receive $\mathbf{y}_k = \sum_{i=1}^{2} \mathbf{H}_{i,k} \mathbf{P}_i \, \mathbf{s}_i + \mathbf{w}_k$, for $k \in \{1,2,\ldots,N_{\mathrm{C}}\}$, where $\mathbf{H}_{i,k} \in \mathbb{C}^{N_{\mathrm{R}} \times N_{\mathrm{T},i}}$ is the *first-hop* channel matrix (from terminal i to relay k),

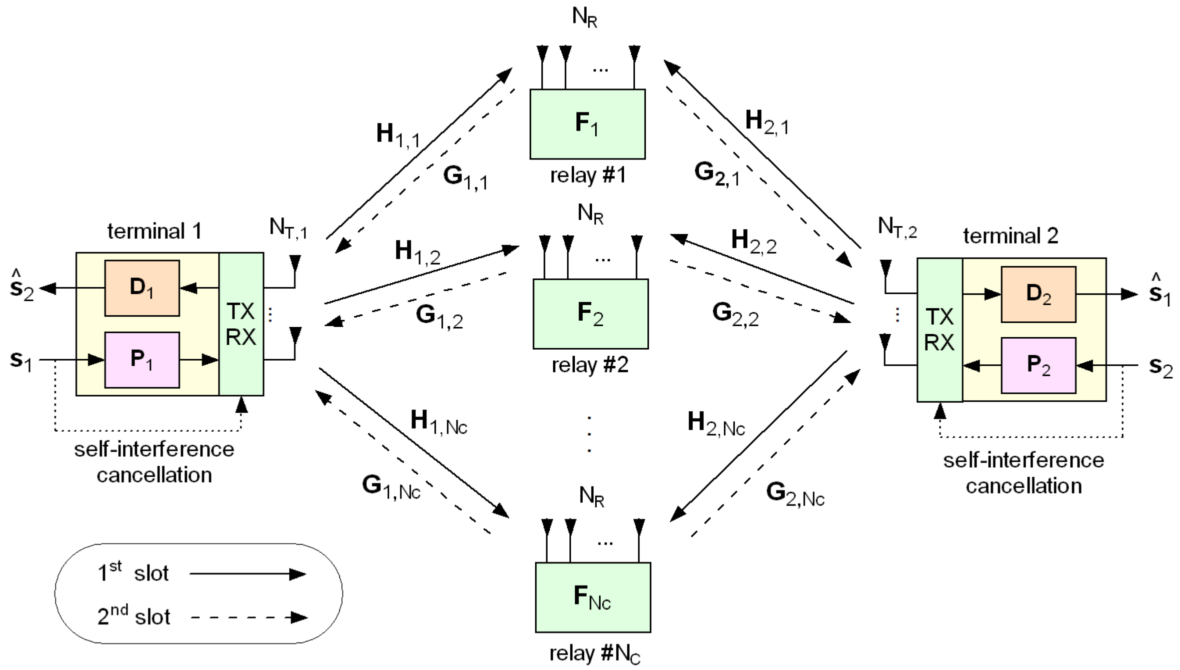

Figure 1. Model of the considered two-way relaying MIMO 5G network.

and $\mathbf{w}_k \in \mathbb{C}^{N_\mathrm{R}}$ models additive noise at kth relay. By defining $\mathbf{y} \triangleq [\mathbf{y}_1^T, \mathbf{y}_2^T, \ldots, \mathbf{y}_{N_\mathrm{C}}^T]^T \in \mathbb{C}^{N_\mathrm{C} N_\mathrm{R}}$, the overall signal received by the relays can be compactly written as

$$\mathbf{y} = \sum_{i=1}^{2} \mathbf{H}_i \mathbf{P}_i \, \mathbf{s}_i + \mathbf{w} \tag{1}$$

where $\mathbf{H}_i \triangleq [\mathbf{H}_{i,1}^T, \mathbf{H}_{i,2}^T \ldots, \mathbf{H}_{i,N_\mathrm{C}}^T]^T \in \mathbb{C}^{N_\mathrm{C} N_\mathrm{R} \times N_{\mathrm{T},i}}$ gathers all first-hop channels and the vector $\mathbf{w} \triangleq [\mathbf{w}_1^\mathrm{T}, \mathbf{w}_2^\mathrm{T}, \ldots, \mathbf{w}_{N_\mathrm{C}}^\mathrm{T}]^\mathrm{T} \in \mathbb{C}^{N_\mathrm{C} N_\mathrm{R}}$ gathers all the noise samples.

In the second time slot, the kth relay forwards its received signal $\mathbf{y}_k \in \mathbb{C}^{N_\mathrm{R}}$, by using the relaying matrix $\mathbf{F}_k \in \mathbb{C}^{N_\mathrm{R} \times N_\mathrm{R}}$, thus transmitting $\mathbf{z}_k = \mathbf{F}_k \, \mathbf{y}_k$. The received signal at each terminal can be written, for $i \in \{1, 2\}$, as

$$\mathbf{r}_i = \sum_{k=1}^{N_\mathrm{C}} \mathbf{G}_{i,k} \mathbf{F}_k \, \mathbf{y}_k + \mathbf{n}_i = \mathbf{G}_i \, \mathbf{F} \, \mathbf{y} + \mathbf{n}_i \tag{2}$$

where $\mathbf{G}_{i,k} \in \mathbb{C}^{N_{\mathrm{T},i} \times N_\mathrm{R}}$ is the *second-hop* channel matrix (from relay k to terminal i), and the vector $\mathbf{n}_i \in \mathbb{C}^{N_{\mathrm{T},i}}$ is additive noise at terminal i. Additionally, we have defined in (2) the extended matrices $\mathbf{G}_i \triangleq [\mathbf{G}_{i,1}, \mathbf{G}_{i,2} \ldots, \mathbf{G}_{i,N_\mathrm{C}}] \in \mathbb{C}^{N_{\mathrm{T},i} \times N_\mathrm{C} N_\mathrm{R}}$ and $\mathbf{F} \triangleq \mathrm{diag}(\mathbf{F}_1, \mathbf{F}_2, \ldots, \mathbf{F}_{N_\mathrm{C}}) \in \mathbb{C}^{N_\mathrm{C} N_\mathrm{R} \times N_\mathrm{C} N_\mathrm{R}}$. Moreover, by taking into account (1), the vector $\mathbf{r}_i$ can also be directly written in terms of $\mathbf{s}_1$ and $\mathbf{s}_2$ as

$$\mathbf{r}_i = \sum_{j=1}^{2} \mathbf{C}_{i,j} \, \mathbf{s}_j + \mathbf{v}_i \tag{3}$$

where $\mathbf{C}_{i,j} \triangleq \mathbf{G}_i \, \mathbf{F} \, \mathbf{H}_j \, \mathbf{P}_j \in \mathbb{C}^{N_{\mathrm{T},i} \times N_{\mathrm{S},j}}$ is the *dual-hop* matrix from terminal j to i, for $i, j \in \{1, 2\}$, and vector $\mathbf{v}_i \triangleq \mathbf{G}_i \, \mathbf{F} \, \mathbf{w} + \mathbf{n}_i \in \mathbb{C}^{N_{\mathrm{T},i}}$ is the overall noise.

We assume customarily [14,18] that each terminal can estimate and subtract the self-interference deriving from its own symbols. To do this, terminal i has to first acquire the matrix $\mathbf{C}_{i,i}$, which can be obtained by resorting to standard training-based identification methods. Specifically, each data transmission can be preceded by a training period, wherein the two terminals transmit orthogonal pilot sequences to the relays. In this case, by redefining $\mathbf{r}_i$ with a slight abuse of notation as $\mathbf{r}_i - \mathbf{C}_{i,i}\, \mathbf{s}_i$, for $i \in \{1,2\}$, we write explicitly

$$\mathbf{r}_i = \mathbf{C}_{i,\underline{i}}\, \mathbf{s}_{\underline{i}} + \mathbf{v}_i = \mathbf{G}_i \mathbf{F} \mathbf{H}_{\underline{i}}\, \mathbf{P}_{\underline{i}}\, \mathbf{s}_{\underline{i}} + \mathbf{v}_i \tag{4}$$

where $\underline{i} = 2$ when $i = 1$, whereas $\underline{i} = 1$ when $i = 2$.

At terminal i, vector $\mathbf{r}_i$ is subject to linear equalization through matrix $\mathbf{D}_i \in \mathbb{C}^{N_{\mathrm{S},\underline{i}} \times N_{\mathrm{T},i}}$, thus yielding a soft estimate $\hat{\mathbf{s}}_{\underline{i}} \triangleq \mathbf{D}_i\, \mathbf{r}_i$ of the symbols $\mathbf{s}_{\underline{i}}$ transmitted by terminal $\underline{i}$, whose entries are then subject to minimum-distance hard decision.

In the sequel, we consider the common assumptions: (a1) $\mathbf{s}_1$ and $\mathbf{s}_2$ are mutually independent zero-mean circularly symmetric complex (ZMCSC) random vectors, with $\mathbb{E}[\mathbf{s}_i\, \mathbf{s}_i^{\mathrm{H}}] = \mathbf{I}_{N_{\mathrm{S},i}}$, for $i \in \{1,2\}$; (a2) the entries of $\mathbf{H}_i$ and $\mathbf{G}_i$ are independent identically distributed ZMCSC Gaussian unit-variance random variables, for $i \in \{1,2\}$; (a3) the noise vectors $\mathbf{w}$, $\mathbf{n}_1$ and $\mathbf{n}_2$ are mutually independent ZMCSC Gaussian random vectors, statistically independent of $\{\mathbf{s}_i, \mathbf{H}_i, \mathbf{G}_i\}_{i=1}^2$, with $\mathbb{E}[\mathbf{w}\mathbf{w}^{\mathrm{H}}] = \sigma_w^2\, \mathbf{I}_{N_{\mathrm{C}} N_{\mathrm{R}}}$ and $\mathbb{E}[\mathbf{n}_i \mathbf{n}_i^{\mathrm{H}}] = \sigma_{n,i}^2\, \mathbf{I}_{N_{\mathrm{T},i}}$, for $i \in \{1,2\}$.

Full channel-state information (CSI) is assumed to be available at both the terminals and the relays. Particularly, we assume that: (i) $\{\mathbf{H}_i\}_{i=1}^2$ are known at the terminals and at the relays; (ii) the kth second-hop channel matrices $\mathbf{G}_{1,k}$ and $\mathbf{G}_{2,k}$ are known only to the kth relay, for $k \in \{1, 2, \ldots, N_{\mathrm{C}}\}$; (iii) the dual-hop channel matrix $\{\mathbf{C}_{i,\underline{i}}\}$ and the covariance matrix[1]

$$\mathbf{K}_{\mathbf{v}_i \mathbf{v}_i} \triangleq \mathbb{E}[\mathbf{v}_i\, \mathbf{v}_i^{\mathrm{H}}] = \sigma_w^2\, \mathbf{G}_i\, \mathbf{F}\, \mathbf{F}^{\mathrm{H}}\, \mathbf{G}_i^{\mathrm{H}} + \sigma_{n,i}^2\, \mathbf{I}_{N_{\mathrm{T},i}} \tag{5}$$

of $\mathbf{v}_i$ are known at the ith terminal, for $i \in \{1,2\}$.

3. The proposed closed-form design

With reference to model (4), the problem at hand is to find optimal values of $\{\mathbf{P}_i\}_{i=1}^2$, $\mathbf{F}$, and $\{\mathbf{D}_i\}_{i=1}^2$ for recovering $\mathbf{s}_1$ and $\mathbf{s}_2$ according to a certain cost function and subject to suitable power constraints at the terminals and relays.

A common performance measure of the accuracy in recovering the symbol vector $\mathbf{s}_i$ at terminal $\underline{i}$ is the mean-square value of the error $\mathbf{e}_i \triangleq \hat{\mathbf{s}}_i - \mathbf{s}_i$: $\mathrm{MSE}_i \triangleq \mathbb{E}[\|\mathbf{e}_i\|^2] = \mathrm{tr}(\mathbf{K}_{\mathbf{e}_i \mathbf{e}_i})$, where $\mathbf{K}_{\mathbf{e}_i \mathbf{e}_i} \triangleq \mathbb{E}[\mathbf{e}_i \mathbf{e}_i^{\mathrm{H}}]$ is the error covariance matrix, which depends on $(\mathbf{P}_i, \mathbf{F}, \mathbf{D}_{\underline{i}})$. As a global cost function for the overall two-way transmission, we consider as in [15–18] the *sum-MSE*, defined as $\mathrm{MSE}(\{\mathbf{P}_i\}_{i=1}^2, \mathbf{F}, \{\mathbf{D}_i\}_{i=1}^2) = \mathrm{MSE}_1 + \mathrm{MSE}_2$. It is well-known that, for fixed values of $\{\mathbf{P}_i\}_{i=1}^2$ and $\mathbf{F}$, the matrices $\{\mathbf{D}_i\}_{i=1}^2$ minimizing the sum-MSE are the Wiener filters

$$\mathbf{D}_{i,\mathrm{mmse}} = \mathbf{C}_{i,\underline{i}}^{\mathrm{H}} (\mathbf{C}_{i,\underline{i}}\, \mathbf{C}_{i,\underline{i}}^{\mathrm{H}} + \mathbf{K}_{\mathbf{v}_i \mathbf{v}_i})^{-1} \tag{6}$$

for $i \in \{1,2\}$, thus yielding

$$\begin{aligned} \mathrm{MSE}(\{\mathbf{P}_i\}_{i=1}^2, \mathbf{F}) &\triangleq \mathrm{MSE}(\{\mathbf{P}_i\}_{i=1}^2, \mathbf{F}, \{\mathbf{D}_{i,\mathrm{mmse}}\}_{i=1}^2) \\ &= \sum_{i=1}^{2} \mathrm{tr}[(\mathbf{I}_{N_{\mathrm{S},i}} + \mathbf{C}_{\underline{i},i}^{\mathrm{H}} \mathbf{K}_{\mathbf{v}_{\underline{i}} \mathbf{v}_{\underline{i}}}^{-1} \mathbf{C}_{\underline{i},i})^{-1}]\,. \end{aligned} \tag{7}$$

It is noteworthy that the variables $\mathbf{P}_1$, $\mathbf{P}_2$, and $\mathbf{F}$ are coupled in (7) and, hence, the two terms in (7) cannot be minimized independently. Herein, we relax the original problem so as to *separate* the minimization of the two terms in (7).

[1] Hereinafter all the ensemble averages are evaluated for fixed values of the first- and second-hop channel matrices.

As a first step, we observe that minimizing (7) is complicated by the presence of $\mathbf{K}_{\mathbf{v}_{\underline{i}}\mathbf{v}_{\underline{i}}}^{-1}$, which depends non-trivially on $\mathbf{F}$. For such a reason, we consider instead minimization of the following high signal-to-noise ratio (SNR) approximation:

$$\text{MSE}(\{\mathbf{P}_i\}_{i=1}^2, \mathbf{F}) \approx \sum_{i=1}^{2} \text{tr}[(\mathbf{I}_{N_{S,i}} + \sigma_{n,\underline{i}}^{-2}\, \mathbf{C}_{\underline{i},i}^{\text{H}} \mathbf{C}_{\underline{i},i})^{-1}] \tag{8}$$

which turns out to be accurate when $\sigma_w^2 \ll \min(\sigma_{n,\underline{i}}^2, \mu_{\min})$, where $\mu_{\min}$ is the smallest eigenvalue of $\mathbf{G}_{\underline{i}}\, \mathbf{F}\, \mathbf{F}^{\text{H}}\, \mathbf{G}_{\underline{i}}^{\text{H}}$. Suitable constraints must be set to avoid trivial solutions in minimizing (8). It is customary to impose power constraints to limit the average transmit power at the terminals:

$$\mathbb{E}[\|\mathbf{P}_i \mathbf{s}_i\|^2] = \text{tr}(\mathbf{P}_i\, \mathbf{P}_i^{\text{H}}) \le \mathcal{P}_{\text{T},i} > 0 \tag{9}$$

for $i \in \{1,2\}$. In order to limit $\mathbf{F}$, we impose a constraint on the average power received at the terminals in the second time slot, i.e., with reference to (2), we attempt to limit, for $i \in \{1,2\}$, the following quantities:

$$\mathbf{E}[\|\mathbf{G}_i \mathbf{F}\, \mathbf{y}\|^2] = \text{tr}(\mathbf{G}_i \mathbf{F} \mathbf{K}_{\mathbf{yy}} \mathbf{F}^{\text{H}} \mathbf{G}_i^{\text{H}}) \tag{10}$$

where $\mathbf{K}_{\mathbf{yy}} \triangleq \mathbb{E}[\mathbf{y}\mathbf{y}^{\text{H}}] = \sum_{i=1}^{2} \mathbf{H}_i \mathbf{P}_i \mathbf{P}_i^{\text{H}} \mathbf{H}_i^{\text{H}} + \sigma_w^2 \mathbf{I}_{N_\text{C} N_\text{R}}$ is the covariance matrix of $\mathbf{y}$. It is noteworthy that (10) is typically limited in those scenarios where a target performance has to be achieved and per-node fairness is not of concern [10,12]. Moreover, the average power received at the terminals is an important metric measuring the human exposure to radio frequency (RF) fields generated by transmitters operating at mmWave frequencies [21] and, with respect to traditional per-relay transmit power constraints, it is more easily related to regulatory specifications [22]. To simplify (10), we exploit the following chain of inequalities:

$$\begin{aligned}
\text{tr}(\mathbf{G}_i \mathbf{F} \mathbf{K}_{\mathbf{yy}} \mathbf{F}^{\text{H}} \mathbf{G}_i^{\text{H}}) &\le \text{tr}(\mathbf{G}_i \mathbf{F} \mathbf{F}^{\text{H}} \mathbf{G}_i^{\text{H}})\text{tr}(\mathbf{K}_{\mathbf{yy}}) \\
&\le \text{tr}(\mathbf{G}_i \mathbf{F} \mathbf{F}^{\text{H}} \mathbf{G}_i^{\text{H}}) \left[\sum_{i=1}^{2} \text{tr}(\mathbf{H}_i \mathbf{H}_i^{\text{H}}) \mathcal{P}_{\text{T},i} + \sigma_w^2 N_\text{C} N_\text{R} \right] \\
&\lesssim \text{tr}(\mathbf{G}_i \mathbf{F} \mathbf{F}^{\text{H}} \mathbf{G}_i^{\text{H}}) N_\text{C} N_\text{R} \left(\sum_{i=1}^{2} N_{\text{T},i}\, \mathcal{P}_{\text{T},i} + \sigma_w^2 \right)
\end{aligned} \tag{11}$$

where the last approximate inequality holds noting that, for fixed values of $N_{\text{T},i}$, by the law of large numbers one has $\mathbf{H}_i^{\text{H}}\, \mathbf{H}_i / (N_\text{C} N_\text{R}) \to \mathbf{I}_{N_{\text{T},i}}$ almost surely as $N_\text{C} N_\text{R}$ gets large. Therefore, if we impose $\text{tr}(\mathbf{G}_i \mathbf{F} \mathbf{F}^{\text{H}} \mathbf{G}_i^{\text{H}}) \le \widetilde{\mathcal{P}}_{\text{R},i} > 0$, we get the upper bound:

$$\text{tr}(\mathbf{G}_i \mathbf{F} \mathbf{K}_{\mathbf{yy}} \mathbf{F}^{\text{H}} \mathbf{G}_i^{\text{H}}) \lesssim \underbrace{\widetilde{\mathcal{P}}_{\text{R},i} N_\text{C} N_\text{R} \left(\sum_{i=1}^{2} N_{\text{T},i} \mathcal{P}_{\text{T},i} + \sigma_w^2 \right)}_{\triangleq \mathcal{P}_{\text{R},i}} . \tag{12}$$

Such a choice allows one to considerably simplify the system design. In summary, the optimization problem to be solved can be expressed as

$$\begin{aligned}
&\min_{\{\mathbf{P}_i\}_{i=1}^2, \mathbf{F}} \sum_{i=1}^{2} \text{tr}[(\mathbf{I}_{N_{S,i}} + \sigma_{n,\underline{i}}^{-2}\, \mathbf{C}_{\underline{i},i}^{\text{H}} \mathbf{C}_{\underline{i},i})^{-1}] \\
&\quad \text{s.to} \quad \begin{cases} \text{tr}(\mathbf{P}_i\, \mathbf{P}_i^{\text{H}}) \le \mathcal{P}_{\text{T},i} \\ \text{tr}(\mathbf{G}_{\underline{i}} \mathbf{F} \mathbf{F}^{\text{H}} \mathbf{G}_{\underline{i}}^{\text{H}}) \le \widetilde{\mathcal{P}}_{\text{R},\underline{i}} \end{cases} \quad i \in \{1,2\} .
\end{aligned} \tag{13}$$

In order to find a closed-form solution of (13), we introduce the matrix $\mathbf{B}_i \triangleq \mathbf{G}_i \mathbf{F} \in \mathbb{C}^{N_{\mathrm{T},i} \times N_{\mathrm{C}} N_{\mathrm{R}}}$, with $i \in \{1,2\}$, and rewrite (13) as follows

$$\min_{\{\mathbf{P}_i\}_{i=1}^2, \{\mathbf{B}_i\}_{i=1}^2} \sum_{i=1}^{2} \mathrm{tr}[(\mathbf{I}_{N_{\mathrm{S},i}} + \sigma_{n,\underline{i}}^{-2}\, \mathbf{P}_i^{\mathrm{H}} \mathbf{H}_i^{\mathrm{H}} \mathbf{B}_{\underline{i}}^{\mathrm{H}} \mathbf{B}_{\underline{i}} \mathbf{H}_i \mathbf{P}_i)^{-1}]$$

$$\text{s.to} \quad \begin{cases} \mathrm{tr}(\mathbf{P}_i\, \mathbf{P}_i^{\mathrm{H}}) \leq \mathcal{P}_{\mathrm{T},i} \\ \mathrm{tr}(\mathbf{B}_{\underline{i}} \mathbf{B}_{\underline{i}}^{\mathrm{H}}) \leq \widetilde{\mathcal{P}}_{\mathrm{R},\underline{i}} \end{cases} \quad i \in \{1,2\}\,. \tag{14}$$

Remarkably, the cost function is the sum of two terms: the former one depends only on the variables $\{\mathbf{P}_1, \mathbf{B}_2\}$, whereas the latter one involves only the variables $\{\mathbf{P}_2, \mathbf{B}_1\}$. Therefore, (14) can be decomposed in two problems involving $\{\mathbf{P}_1, \mathbf{B}_2\}$ and $\{\mathbf{P}_2, \mathbf{B}_1\}$ separately, which can be solved in parallel in a closed-form manner. Indeed, capitalizing on such a decomposition, the solution of (14) can be characterized by the following theorem.

Theorem 1. *Assume that: (a4) $\mathbf{P}_i \in \mathbb{C}^{N_{T,i} \times N_{S,i}}$ is full-column rank, i.e., rank$(\mathbf{P}_i) = N_{S,i} \leq N_{T,i}$, $i \in \{1,2\}$; (a5) $\mathbf{B}_{\underline{i}} \mathbf{H}_i \in \mathbb{C}^{N_{T,\underline{i}} \times N_{T,i}}$ is full-column rank, i.e., rank$(\mathbf{B}_{\underline{i}} \mathbf{H}_i) = N_{T,i} \leq N_{T,\underline{i}}$, for $i \in \{1,2\}$.[2] Moreover, let $\mathbf{H}_i = \mathbf{U}_{h,i} \mathbf{\Lambda}_{h,i} \mathbf{V}_{h,i}^H$ denote the singular value decomposition (SVD) of $\mathbf{H}_i$, where $\mathbf{U}_{h,i} \in \mathbb{C}^{N_C N_R \times N_C N_R}$ and $\mathbf{V}_{h,i} \in \mathbb{C}^{N_{T,i} \times N_{T,i}}$ are the unitary matrices of left/right singular vectors, and $\mathbf{\Lambda}_{h,i} \in \mathbb{C}^{N_C N_R \times N_{T,i}}$ is the rectangular diagonal matrix of the corresponding singular values arranged in increasing order. Then, the solution of (14) has the following general form:*

$$\mathbf{P}_i = \mathbf{V}_{h,i,right}\, \mathbf{\Omega}_i \tag{15}$$

$$\mathbf{B}_{\underline{i}} = \mathbf{Q}_{\underline{i}}\, \mathbf{\Delta}_{\underline{i}} \mathbf{U}_{h,i,right}^H \tag{16}$$

where $\mathbf{V}_{h,i,right}$ contains the $N_{S,i}$ rightmost columns of $\mathbf{V}_{h,i}$, $\mathbf{U}_{h,i,right}$ contains the $N_{T,i}$ rightmost columns of $\mathbf{U}_{h,i}$, the diagonal matrices $\mathbf{\Omega}_i \in \mathbb{R}^{N_{S,i} \times N_{S,i}}$ and $\mathbf{\Delta}_{\underline{i}} \in \mathbb{R}^{N_{T,i} \times N_{T,i}}$ will be specified soon after, and $\mathbf{Q}_{\underline{i}} \in \mathbb{C}^{N_{T,\underline{i}} \times N_{T,i}}$ is an arbitrary semi-unitary matrix, i.e., $\mathbf{Q}_{\underline{i}}^H \mathbf{Q}_{\underline{i}} = \mathbf{I}_{N_{T,i}}$.

Proof. See Appendix A. □

Remark 1. *(a4) implies that $N_{S,i} \leq N_{T,i}$, $i \in \{1,2\}$.*

Remark 2. *(a5) implies that $N_{T,1} = N_{T,2}$ and, hence, in the following we set $N_T \triangleq N_{T,1} = N_{T,2}$.*

Under (a4) and (a5), the dual-hop channel matrices $\{\mathbf{C}_{i,\underline{i}} = \mathbf{B}_i \mathbf{H}_{\underline{i}} \mathbf{P}_{\underline{i}}\}_{i=1}^2$ are full-column rank, i.e., rank$(\mathbf{C}_{i,\underline{i}}) = N_{\mathrm{S},\underline{i}} \leq N_{\mathrm{T},\underline{i}}$, for $i = 1,2$: this ensures perfect recovery of the source symbol vectors $\{\mathbf{s}_i\}_{i=1}^2$ at the terminals in the absence of noise by means of linear equalizers. Although Theorem 1 holds for any value of $N_{\mathrm{S},1}$ and $N_{\mathrm{S},2}$, we will assume herein that $N_{\mathrm{S},1} = N_{\mathrm{S},2} = N_{\mathrm{T}}$, which allows the terminals to transmit as many symbols as possible with an acceptable performance in practice.

[2] (a5) implies that $\mathbf{H}_i$ is full-column rank too, i.e., rank$(\mathbf{H}_i) = N_{\mathrm{T},i}$.

Algorithm 1: The proposed design algorithm

Input quantities: $\{\mathbf{H}_i,\, \mathbf{G}_i,\, \sigma^2_{n,i},\, \mathcal{P}_{\mathrm{T},i},\, \widetilde{\mathcal{P}}_{\mathrm{R},i}\}_{i=1}^{2}$

Output quantities: $\{\mathbf{P}_i,\, \mathbf{D}_{i,\mathrm{mmse}}\}_{i=1}^{2}$ and $\{\mathbf{F}_k\}_{k=1}^{N_\mathrm{C}}$

1. Choose arbitrary $\{\mathbf{Q}_i\}_{i=1}^{2}$ such that $\mathbf{Q}_i^{\mathrm{H}}\mathbf{Q}_i = \mathbf{I}_{N_{\mathrm{S},i}}$.
2. Perform the SVD of $\{\mathbf{H}_i\}_{i=1}^{2}$. and collect the $\{N_{\mathrm{S},i}\}_{i=1}^{2}$ largest singular values and the corresponding left/right singular vectors.
3. Solve the convex problem (17) in the disjoint subsets of variables $\{z_{1,\ell},\, w_{2,\ell}\}_{\ell=1}^{N_{\mathrm{S},1}}$ and $\{z_{2,\ell},\, w_{1,\ell}\}_{\ell=1}^{N_{\mathrm{S},2}}$ separately.
4. From the solution of step 3, build the matrices $\{\mathbf{\Omega}_i, \mathbf{\Delta}_i\}_{i=1}^{2}$.
5. Build the matrices $\{\mathbf{P}_i,\, \mathbf{B}_i\}_{i=1}^{2}$ according to (15) and (16).
6. Calculate $\{\mathbf{F}_k\}_{k=1}^{N_\mathrm{C}}$ according to (19).
7. Calculate $\{\mathbf{D}_{i,\mathrm{mmse}}\}_{i=1}^{2}$ according to (6).

Theorem 1 allows one to rewrite the optimization problem (14) in a simpler scalar form:

$$\min_{\substack{\{z_{1,\ell},w_{2,\ell}\}_{\ell=1}^{N_\mathrm{T}} \\ \{z_{2,\ell},w_{1,\ell}\}_{\ell=1}^{N_\mathrm{T}}}} \sum_{i=1}^{2}\sum_{\ell=1}^{N_\mathrm{T}} \frac{1}{1+\sigma_{n,i}^{-2}\,\lambda_\ell^2(\mathbf{H}_i) z_{i,\ell}\, w_{i,\ell}}$$

$$\text{s.to} \quad \begin{cases} \displaystyle\sum_{\ell=1}^{N_\mathrm{T}} z_{i,\ell} \le \mathcal{P}_{\mathrm{T},i} & \\ \displaystyle\sum_{\ell=1}^{N_\mathrm{T}} w_{i,\ell} \le \widetilde{\mathcal{P}}_{\mathrm{R},i} & i \in \{1,2\} \\ w_{i,\ell}, z_{i,\ell} > 0 \qquad \forall \ell \in \{1,2,\ldots,N_{\mathrm{S},i}\} & \end{cases} \tag{17}$$

with $z_{i,\ell}$ and $w_{i,\ell}$ representing the ℓth squared diagonal entry of $\mathbf{\Omega}_i$ and $\mathbf{\Delta}_i$, respectively, whereas $\lambda_\ell(\mathbf{H}_i)$ denotes the ℓth nonzero singular value of $\mathbf{H}_i$, for $\ell \in \{1,2,\ldots,N_\mathrm{T}\}$. Similarly to (14), problem (17) can be decomposed into two separate problems involving disjoint subsets of variables.

It can be shown, with straightforward manipulations, that the objective function in (17) is convex if and only if

$$z_{i,\ell}\, w_{i,\ell} \ge \frac{\sigma_{n,i}^2}{3\lambda_\ell^2(\mathbf{H}_i)} \tag{18}$$

$\forall \ell \in \{1,2,\ldots,N_{\mathrm{S},i}\}$, with $i \in \{1,2\}$. It is also seen that, based on (a2), one has $\lambda_{\min}(\mathbf{H}_i) \gg 1$ in the large $N_\mathrm{C} N_\mathrm{R}$ limit, with $i \in \{1,2\}$. Thus, condition (18) boils down to $z_{i,\ell},\, w_{i,\ell} > 0$, for all $\ell \in \{1,2,\ldots,N_{\mathrm{S},i}\}$, with $i \in \{1,2\}$, which is already included in the constraints of (17). Therefore, convex programming can be used to find a global minimum of (17).

To calculate the relaying matrices, let us partition solution (16) as $\mathbf{B}_i = [\mathbf{B}_{i,1}, \mathbf{B}_{i,2}, \cdots, \mathbf{B}_{i,N_\mathrm{C}}]$, with $\mathbf{B}_{i,k} \in \mathbb{C}^{N_\mathrm{T}\times N_\mathrm{R}}$, $i \in \{1,2\}$. Defining $\widetilde{\mathbf{G}}_k \triangleq [\mathbf{G}_{1,k}^\mathrm{T}, \mathbf{G}_{2,k}^\mathrm{T}]^\mathrm{T} \in \mathbb{C}^{2N_\mathrm{T}\times N_\mathrm{R}}$ and $\widetilde{\mathbf{B}}_k \triangleq [\mathbf{B}_{1,k}^\mathrm{T}, \mathbf{B}_{2,k}^\mathrm{T}]^\mathrm{T} \in \mathbb{C}^{2N_\mathrm{T}\times N_\mathrm{R}}$, and assuming that $\widetilde{\mathbf{G}}_k$ is full-row rank, i.e., $\mathrm{rank}(\widetilde{\mathbf{G}}_k) = 2N_\mathrm{T} \le N_\mathrm{R}$, with $k \in \{1,2,\ldots,N_\mathrm{C}\}$, the kth relay can construct its own relaying matrix by solving the matrix equation $\widetilde{\mathbf{G}}_k \mathbf{F}_k = \widetilde{\mathbf{B}}_k$, whose minimum-norm solution is given by

$$\mathbf{F}_k = \widetilde{\mathbf{G}}_k^\dagger \widetilde{\mathbf{B}}_k \tag{19}$$

where the superscript † denotes the Moore-Penrose inverse.

With reference to the step-by-step description of the proposed design algorithm reported at the top of this page, the following comments are in order. The convex optimization in step 3) can be efficiently carried out using standard techniques, such as the interior-point method. We observe that the worst-case theoretical complexity of the interior-point method is proportional to $\sqrt{N_\mathrm{T}}$. Hence, for a realistic setting of the system parameters, the computational complexity of the proposed algorithm, is dominated by the SVD computation (in step 2), which is of order $\mathcal{O}(N_\mathrm{C}\, N_\mathrm{R}\, N_\mathrm{T}^2)$ and, thus, it

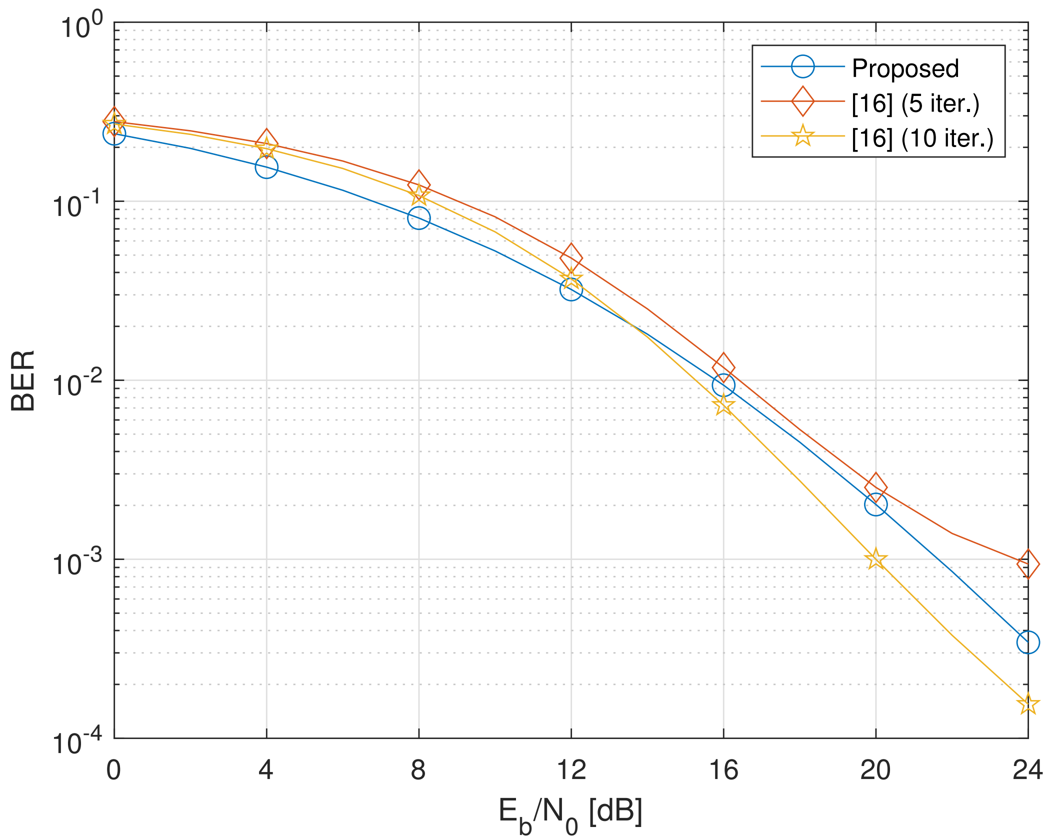

Figure 2. BER versus E_b/N_0 of the proposed design versus the iterative method of [16] ($N_C = 2$).

linearly grows with the number N_C of relays. It is noteworthy that, even though the alternating algorithm proposed in [16] allows to solve a nonconvex problem by solving convex subproblems, it is more complex than calculating the solution of (17); moreover, it requires proper initialization to monotonically converge to (at least) a local optimum.

4. Simulation results

In this section, to assess the performance of the considered design, we present the results of Monte Carlo computer simulations, aimed at evaluating the average (with respect to channel realizations) bit-error-rate (BER) of the proposed cooperative two-way MIMO system. We consider a network encompassing two terminals equipped with $N_T = 2$ antennas, and transmitting QPSK symbols with $N_{S,1} = N_{S,2} = 2$. The N_C relays are equipped with $N_R = 4$ antennas. We also assume that $\mathcal{P}_{T,1} = \mathcal{P}_{T,2} = \mathcal{P}_k = \mathcal{P}$, for all $k \in \{1, 2, \ldots, N_C\}$, where $\mathcal{P}_k$ represents the average transmitted power at the kth relay, and set $\sigma_w^2 = \sigma_{n,1}^2 = \sigma_{n,2}^2 = 1$. Consequently, the energy per bit to noise power spectral density ratio E_b/N_0 is a measure of the per-antenna link quality of both the first- and second-hop transmissions. The BER is evaluated by carrying out 10^3 independent Monte Carlo trials, with each run using independent sets of channel realizations and noise, and an independent record of 10^6 source symbols.

We compare the performances of our design (labeled as "Proposed") to those of the iterative technique proposed in [16], which has been shown [16] in its turn to outperform other iterative techniques, such as the gradient-descent technique of [15]. It is worthwhile to note that both the strategies under comparison require the same amount of CSI. Furthermore, since the method of [16] imposes different power constraints on the design of the relaying matrices, our solutions for $\{\mathbf{F}_k\}_{k=1}^{N_C}$

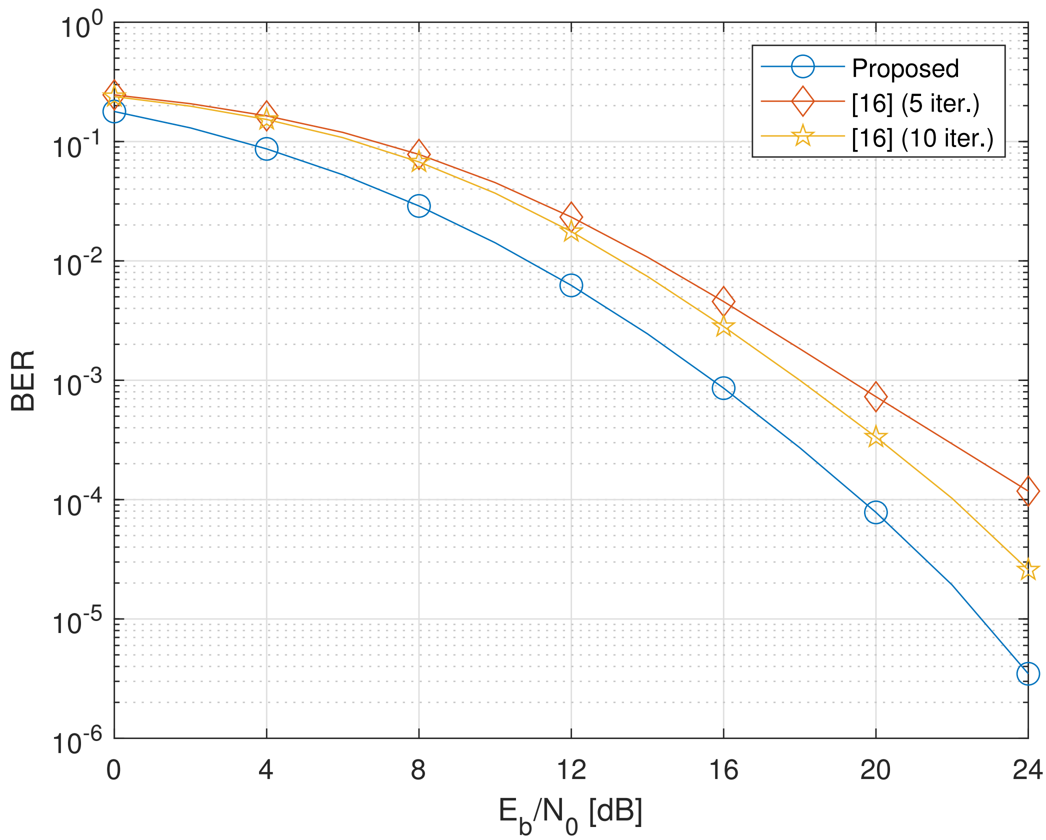

Figure 3. BER versus E_b/N_0 of the proposed design versus the iterative method of [16] ($N_C = 3$).

are properly scaled so as to ensure that the average power transmitted by each relay is the same for both methods.

In Figs. 2–4, we report the BER for different values of the number $N_C \in \{2, 3, 4\}$ of relays. Results in Fig. 2 for $N_C = 2$ show that the proposed closed-form design, based on the solution of the relaxed problem (14), exhibits performances comparable with the iterative solution of [16] in the considered range of E_b/N_0 values only when the latter employs more than 5 iterations. Specifically, when the method of [16] employs 10 iterations, a crossover can be observed in Fig. 2 between the BER curve of the proposed algorithm and that of [16]. This behavior is due to the fact that the rate of convergence of [16] strongly depends on the SNR. Figs. 3 and 4 show that, as the number of relays increases, the proposed method clearly outperforms the method of [16] even when the latter employs 10 iterations.[3]

In a nutshell, although the alternating iterative procedure [16] attempts to solve the nonconvex original two-way constrained minimum sum-MSE problem, its convergence behaviors are affected in practice by both the operative SNR and number of relays: in the low-SNR region and/or when the number of relays is sufficiently large, convergence to a local minimizer is not guaranteed in a reasonable number of iterations for all possible initializations. This is the price to pay for swapping a difficult joint optimization with a sequence of easier problems involving subsets of the variables. On the other hand, the proposed optimization strategy gives up the idea of solving the original nonconvex problem, by resorting to suitable relaxations of both the cost function and the relaying power constraint. This allows us to jointly optimize all the variables, without using burdensome iterative algorithms.

[3] Performance improvement of [16] is negligible after 10 iterations.

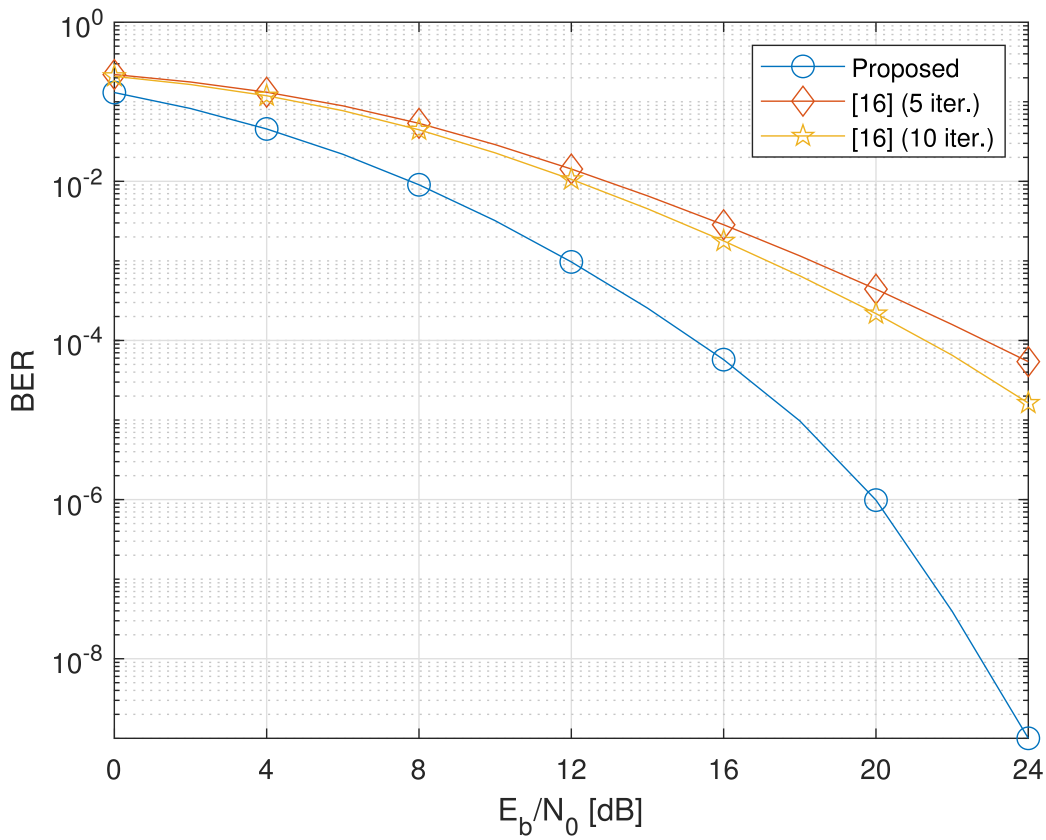

Figure 4. BER versus E_b/N_0 of the proposed design versus the iterative method of [16] ($N_C = 4$).

5. Discussion and directions for future work

We tackled the joint sum-MSE design of terminal precoders/decoders and relay forwarding matrices for two-way AF MIMO 5G systems. We showed that a relaxed version of such a problem can be separated into two simpler ones, which can be solved in parallel by admitting closed-form solutions. The proposed technique exhibits a performance gain over the iterative method of [16], exhibiting a better scaling with the number of relays and a reduced computational complexity.

In this paper, we assumed the availability of full-CSI at both terminals and the relays. In this respect, an interesting research subject consists of considering the use of partial CSI to extend network lifetime and reduce the complexity burden. Moreover, since channel estimation errors occur in practical situations, an additional research issue is to develop robust optimization designs.

Appendix A. Proof of Theorem 1

We focus on the optimization (14) with indexes $i = 1$ and $\underline{i} = 2$, i.e., we consider

$$\min_{\mathbf{P}_1,\mathbf{B}_2} \quad \mathrm{tr}[(\mathbf{I}_{N_{\mathrm{S},1}} + \sigma_{n,2}^{-2}\, \mathbf{P}_1^{\mathrm{H}}\mathbf{H}_1^{\mathrm{H}}\mathbf{B}_2^{\mathrm{H}}\mathbf{B}_2\mathbf{H}_1\mathbf{P}_1)^{-1}]$$
$$\text{s.to} \quad \begin{cases} \mathrm{tr}(\mathbf{P}_1\,\mathbf{P}_1^{\mathrm{H}}) \le \mathcal{P}_{\mathrm{T},1} \\ \mathrm{tr}(\mathbf{B}_2\mathbf{B}_2^{\mathrm{H}}) \le \widetilde{\mathcal{P}}_{\mathrm{R},2} \end{cases} . \tag{A1}$$

We note that under (a4) and (a5), one has $\mathrm{rank}(\mathbf{B}_2\mathbf{H}_1\mathbf{P}_1) = N_{\mathrm{S},1} \le N_{\mathrm{T},1}$. Let $\mathbf{U}_{\mathrm{a}}\boldsymbol{\Lambda}_{\mathrm{a}}\mathbf{U}_{\mathrm{a}}^{\mathrm{H}}$ be the eigenvalue decomposition (EVD) of $\mathbf{A} \triangleq \mathbf{H}_1^{\mathrm{H}}\mathbf{B}_2^{\mathrm{H}}\mathbf{B}_2\mathbf{H}_1 \in \mathbb{C}^{N_{\mathrm{T},1}\times N_{\mathrm{T},1}}$, where the diagonal matrix $\boldsymbol{\Lambda}_{\mathrm{a}} \in \mathbb{R}^{N_{\mathrm{T},1}\times N_{\mathrm{T},1}}$ and the unitary matrix $\mathbf{U}_{\mathrm{a}} \in \mathbb{C}^{N_{\mathrm{T},1}\times N_{\mathrm{T},1}}$ collect the eigenvalues, arranged in increasing order, and the eigenvectors of $\mathbf{A}$, respectively. The objective function in (A1) is a Schur-concave function of the diagonal elements of $(\mathbf{I}_{N_{S,1}} + \sigma_{n,2}^{-2}\,\mathbf{P}_1^{\mathrm{H}}\mathbf{A}\mathbf{P}_1)^{-1}$. In this case, it can be shown [23] that there is an optimal $\mathbf{P}_1$ such that $\mathbf{P}_1^{\mathrm{H}}\mathbf{A}\mathbf{P}_1$ is diagonal, whose diagonal elements are assumed to be arranged in increasing order, and such an optimal matrix, which also minimizes $\mathrm{tr}(\mathbf{P}_1\mathbf{P}_1^{\mathrm{H}})$, is given by

$$\mathbf{P}_1 = \mathbf{U}_{\mathrm{a,right}}\boldsymbol{\Omega}_1, \tag{A2}$$

where $\mathbf{U}_{\mathrm{a,right}} \in \mathbb{C}^{N_{\mathrm{T},1}\times N_{\mathrm{S},1}}$ contains the $N_{\mathrm{S},1} \leq N_{\mathrm{T},1}$ rightmost columns from $\mathbf{U}_{\mathrm{a}}$, and $\boldsymbol{\Omega}_1 \in \mathbb{C}^{N_{\mathrm{S},1}\times N_{\mathrm{S},1}}$ is a diagonal matrix. Let $\mathbf{Q}_2 \in \mathbb{C}^{N_{\mathrm{T},2}\times N_{\mathrm{T},1}}$ be an arbitrary semi-unitary matrix, i.e., $\mathbf{Q}_2^{\mathrm{H}}\mathbf{Q}_2 = \mathbf{I}_{N_{\mathrm{T},1}}$, it follows from the EVD of the matrix $\mathbf{A}$ that $\mathbf{B}_2\mathbf{H}_1 = \mathbf{Q}_2\boldsymbol{\Lambda}_{\mathrm{a}}^{1/2}\mathbf{U}_{\mathrm{a}}^{\mathrm{H}}$. Noting that $\mathrm{rank}(\mathbf{H}_1) = N_{\mathrm{T},1}$, by substituting the ordered SVD of $\mathbf{H}_1 = \mathbf{U}_{\mathrm{h},1}\boldsymbol{\Lambda}_{\mathrm{h},1}\mathbf{V}_{\mathrm{h},1}^{\mathrm{H}}$ in this equation, after some algebraic manipulations, one has that the minimum-norm solution [24] of the matrix equation $\mathbf{B}_2\mathbf{U}_{\mathrm{h},1}\boldsymbol{\Lambda}_{\mathrm{h},1} = \mathbf{Q}_2\boldsymbol{\Lambda}_{\mathrm{a}}^{1/2}\mathbf{U}_{\mathrm{a}}^{\mathrm{H}}\mathbf{V}_{\mathrm{h},1}$ is

$$\mathbf{B}_2 = \mathbf{Q}_2\boldsymbol{\Lambda}_{\mathrm{a}}^{1/2}\widetilde{\mathbf{U}}_{\mathrm{a}}\boldsymbol{\Lambda}_{\mathrm{h},1,\mathrm{right}}^{-1}\mathbf{U}_{\mathrm{h},1,\mathrm{right}} \tag{A3}$$

where $\mathbf{U}_{\mathrm{h},1,\mathrm{right}}$ collects the $N_{\mathrm{T},1}$ rightmost columns of $\mathbf{U}_{\mathrm{h},1}$, whereas $\widetilde{\mathbf{U}}_{\mathrm{a}} \triangleq \mathbf{U}_{\mathrm{a}}^{\mathrm{H}}\mathbf{V}_{\mathrm{h},1} \in \mathbb{C}^{N_{\mathrm{T},1}\times N_{\mathrm{T},1}}$ and the diagonal a $\boldsymbol{\Lambda}_{\mathrm{h},1,\mathrm{right}} \in \mathbb{R}^{N_{\mathrm{T},1}\times N_{\mathrm{T},1}}$ gathers the $N_{\mathrm{T},1}$ nonzero singular values of $\mathbf{H}_1$ in increasing order. The aim is now to further determine (A3) by properly choosing $\widetilde{\mathbf{U}}_{\mathrm{a}}$ such that $\mathrm{tr}(\mathbf{B}_2\mathbf{B}_2^{\mathrm{H}}) = \mathrm{tr}[(\widetilde{\mathbf{U}}_{\mathrm{a}}\boldsymbol{\Lambda}_{\mathrm{h},1,\mathrm{right}}^{-2}\widetilde{\mathbf{U}}_{\mathrm{a}})\boldsymbol{\Lambda}_{\mathrm{a}}]$ has the smallest value[4]. By observing that $\widetilde{\mathbf{U}}_{\mathrm{a}}^{\mathrm{H}}\widetilde{\mathbf{U}}_{\mathrm{a}} = \mathbf{I}_{N_{\mathrm{T},1}}$ and using a known trace inequality, one has

$$\mathrm{tr}[(\widetilde{\mathbf{U}}_{\mathrm{a}}\boldsymbol{\Lambda}_{\mathrm{h},1,\mathrm{right}}^{-2}\widetilde{\mathbf{U}}_{\mathrm{a}})\boldsymbol{\Lambda}_{\mathrm{a}}] \geq \sum_{\ell=1}^{N_{\mathrm{T},1}} \lambda_{\mathrm{h},1,\ell}^{-2}\lambda_{\mathrm{a},\ell} \tag{A4}$$

where $\lambda_{\mathrm{h},1,\ell}$ and $\lambda_{\mathrm{a},\ell}$ denote the ℓth diagonal entry of $\boldsymbol{\Lambda}_{\mathrm{h},1,\mathrm{right}}$ and $\boldsymbol{\Lambda}_{\mathrm{a}}$, respectively. The equality in (A4) holds when

$$\widetilde{\mathbf{U}}_{\mathrm{a}} = \mathbf{U}_{\mathrm{a}}^{\mathrm{H}}\mathbf{V}_{\mathrm{h},1} = \mathbf{I}_{N_{\mathrm{T},1}}. \tag{A5}$$

Substituting (A5) in (A3), after some algebraic manipulations, one obtains $\mathbf{B}_2 = \mathbf{Q}_2\,\boldsymbol{\Delta}_2\mathbf{U}_{\mathrm{h},1,\mathrm{right}}^{\mathrm{H}}$, with $\boldsymbol{\Delta}_2 \triangleq \boldsymbol{\Lambda}_{\mathrm{a}}^{1/2}\boldsymbol{\Lambda}_{\mathrm{h},1,\mathrm{right}}^{-1} \in \mathbb{R}^{N_{\mathrm{T},1}\times N_{\mathrm{T},1}}$. Solution (15) comes from substituting in (A2) the minimum-norm solution [24] of (A5), i.e., $\mathbf{U}_{\mathrm{a}} = \mathbf{V}_{\mathrm{h},1,\mathrm{right}}$.

Author Contributions: Conceptualization, D.D., G.G., I.I. and F.V.; methodology, D.D., G.G., I.I. and F.V.; writing–original draft preparation, G.G. and F.V.; writing–review and editing, I.I. and F.V.; supervision, F.V..

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Yang, Y.; Hu, H.; Xu, J.; Mao, G. Relay technologies for WiMax and LTE-advanced mobile systems. *IEEE Communications Magazine* **2009**, *47*, 100–105.
2. Loa, K.; Wu, C.; Sheu, S.; Yuan, Y.; Chion, M.; Huo, D.; Xu, L. IMT-advanced relay standards [WiMAX/LTE Update]. *IEEE Communications Magazine* **2010**, *48*, 40–48.
3. Bhat, P.; Nagata, S.; Campoy, L.; Berberana, I.; Derham, T.; Liu, G.; Shen, X.; Zong, P.; Yang, J. LTE-advanced: an operator perspective. *IEEE Communications Magazine* **2012**, *50*, 104–114.
4. Omoniwa, B.; Hussain, R.; Adil, M.; Shakeel, A.; Tahir, A.K.; Hasan, Q.U.; Malik, S.A. An Optimal Relay Scheme for Outage Minimization in Fog-Based Internet-of-Things (IoT) Networks. *IEEE Internet of Things Journal* **2019**, *6*, 3044–3054.
5. Chen, Z.; Wang, L.; Han, D.; Zeng, H. A Unified Performance Analysis of Relaying Communication System for IoT Application With Hybrid Fading. *IEEE Internet of Things Journal* **2020**, *7*, 570–583.

[4] It is readily seen that $\mathrm{tr}(\mathbf{B}_2\mathbf{B}_2^{\mathrm{H}})$ is invariant to the choice of $\mathbf{Q}_2$.

6. Ji, B.; Chen, Z.; Chen, S.; Zhou, B.; Li, C.; Wen, H. Joint optimization for ambient backscatter communication system with energy harvesting for IoT. *Mechanical Systems and Signal Processing* **2020**, *135*, 1–10.
7. Parkvall, S.; Dahlman, E.; Furuskar, A.; Frenne, M. NR: The New 5G Radio Access Technology. *IEEE Communications Standards Magazine* **2017**, *1*, 24–30.
8. Rappaport, T.S.; Xing, Y.; Kanhere, O.; Ju, S.; Madanayake, A.; Mandal, S.; Alkhateeb, A.; Trichopoulos, G.C. Wireless Communications and Applications Above 100 GHz: Opportunities and Challenges for 6G and Beyond. *IEEE Access* **2019**, *7*, 78729–78757.
9. Wu, S.; Atat, R.; Mastronarde, N.; Liu, L. Improving the Coverage and Spectral Efficiency of Millimeter-Wave Cellular Networks Using Device-to-Device Relays. *IEEE Transactions on Communications* **2018**, *66*, 2251–2265.
10. Cui, S.; Goldsmith, A.J.; Bahai, A. Energy-efficiency of MIMO and cooperative MIMO techniques in sensor networks. *IEEE Journal on Selected Areas in Communications* **2004**, *22*, 1089–1098. doi:10.1109/JSAC.2004.830916.
11. Jayaweera, S.K. Virtual MIMO-based cooperative communication for energy-constrained wireless sensor networks. *IEEE Trans. Wireless Communications* **2006**, *5*, 984–989. doi:10.1109/TWC.2006.1633350.
12. Behbahani, A.S.; Merched, R.; Eltawil, A.M. Optimizations of a MIMO Relay Network. *IEEE Transactions on Signal Processing* **2008**, *56*, 5062–5073. doi:10.1109/TSP.2008.929120.
13. Darsena, D.; Gelli, G.; Verde, F. Design and performance analysis of multiple-relay cooperative MIMO networks. *Journal of Communications and Networks* **2019**, *21*, 25–32. doi:10.1109/JCN.2019.000003.
14. Katti, S.; Gollakota, S.; Katabi, D. Embracing Wireless Interference: Analog Network Coding. *ACM SIGCOMM Computer Communication Review* **2007**, *9*. doi:10.1145/1282427.1282425.
15. Lee, K.; Sung, H.; Park, E.; Lee, I. Joint Optimization for One and Two-Way MIMO AF Multiple-Relay Systems. *IEEE Transactions on Wireless Communications* **2010**, *9*, 3671–3681. doi:10.1109/TWC.2010.102210.091021.
16. Rong, Y. Joint Source and Relay Optimization for Two-Way MIMO Multi-Relay Networks. *IEEE Communications Letters* **2011**, *15*, 1329–1331. doi:10.1109/LCOMM.2011.092911.111584.
17. Hu, C.; Chou, Y. Precoding Design of MIMO AF Two-Way Multiple-Relay Systems. *IEEE Signal Processing Letters* **2013**, *20*, 623–626. doi:10.1109/LSP.2012.2231069.
18. Hu, C.; Liu, G.; Chen, B. Joint Relay/Antenna Selection and Precoding Design for Two-Way MIMO Amplify-and-Forward Relaying Systems. *IEEE Transactions on Vehicular Technology* **2016**, *65*, 4854–4864. doi:10.1109/TVT.2015.2469727.
19. Hu, C.; Chen, B.; Liu, G. Precoding design for two-way MIMO relaying with antenna selection. Proc. IEEE Int. Conf. Consumer Electronics (ICCE), 2018, pp. 1–5. doi:10.1109/ICCE.2018.8326316.
20. Rong, Y. Joint Source and Relay Optimization for Two-Way Linear Non-Regenerative MIMO Relay Communications. *IEEE Transactions on Signal Processing* **2012**, *60*, 6533–6546. doi:10.1109/TSP.2012.2213081.
21. Wu, T.; Rappaport, T.S.; Collins, C.M. Safe for Generations to Come: Considerations of Safety for Millimeter Waves in Wireless Communications. *IEEE Microwave Magazine* **2015**, *16*, 65–84.
22. Gastpar, M. On Capacity Under Receive and Spatial Spectrum-Sharing Constraints. *IEEE Transactions on Information Theory* **2007**, *53*, 471–487.
23. Palomar, D.P.; Cioffi, J.M.; Lagunas, M.A. Joint Tx-Rx beamforming design for multicarrier MIMO channels: a unified framework for convex optimization. *IEEE Transactions on Signal Processing* **2003**, *51*, 2381–2401. doi:10.1109/TSP.2003.815393.
24. Ben-Israel, A.; Greville, T.N. *Generalized inverses: theory and applications*; Vol. 15, Springer Science & Business Media, 2003.