NON-LINEAR PRECODING APPROACHES FOR NON-REGENERATIVE MULTIUSER MIMO RELAY SYSTEMS

I. Jimenez, M. Barrenechea, M. Mendicute, and E. Arruti

Signal Theory and Communications Area, Department of Electronics and Computer Science
University of Mondragón, Mondragón, Spain
Email: ijimenez,mbarrenetxea,mmendikute,earruti@mondragon.edu

ABSTRACT

In this paper we consider the design of minimum mean square error (MMSE) transceivers for non-regenerative multiuser multiple-input multiple-output (MIMO) relay systems, where the main challenge resides in the joint optimal design of the precoder to be used at the base station and the relaying algorithms. We propose the use of non-linear precoding techniques, such as Tomlinson-Harashima precoding (THP) and vector precoding (VP), at the downlink relayed transmission scenario with the aim of outperforming the more common linear strategies. In order to reduce the computational cost of the proposed schemes, we propose a novel suboptimal matrix design approach for VP transmission. Provided simulation results show that the proposed non-linear precoding schemes outperform the best MIMO-AF relaying architectures in the literature, even when a reduced-complexity suboptimal strategy is adopted, considering both BER performance and mean square error minimization.

***Index Terms*—** Multiuser MIMO, non-linear precoding, vector precoding (VP), cooperative communications

1. INTRODUCTION

Relaying or multi-hop systems have drawn a considerably research interest due to their increased coverage and low cost. Furthermore, the combination of relays and MIMO technology [1][2][3] has been used to improve channel capacity, reliability and network coverage.

In multiuser relaying schemes, the relay is used to connect a base station (BS) and multiple mobile stations (MS), or vice versa. Relays are generally classified into two categories depending on their processing capabilities: amplify-and-forward (AF) [4] and decode-and-forward (DF) [5], also known as non-regenerative and regenerative relaying techniques respectively. While the first only amplifies the received signal before retransmission, the second decode and encode it again. When AF relaying scenarios are considered, which is the case for this reseach work, the design of the precoding and relaying algorithms plays an important role.

Most of the work in the area of multiuser MIMO relaying is based on the joint design of precoding and relaying matrices. In [6], a joint linear optimization problem is considered for both uplink and downlink systems based on the MMSE criterion. Apart from giving the optimal solution, a suboptimal approach is proposed based on the singular value decomposition (SVD) of the wireless channels. A similar solution is also applied in [7] to a slightly different multiuser MIMO relaying scenario, where each user is equipped with an MMSE equalizer.

It is well known for single-hop multiuser MIMO systems that linearly prefiltering at the transmitter can be outperformed by non-linear precoding techniques, such as Tomlinson-Harashima precoding [8] or vector precoding [9]. Regarding their application to multiuser AF relaying schemes, suboptimal THP solutions are given in [10] and [11]. MMSE and zero forcing (ZF) THP solutions are also given in [12] for a multiuser MIMO relaying scenario with a fixed channel-independent relaying matrix.

Two non-iterative suboptimal approaches are proposed in [13] for VP-precoded transmission. The first minimizes the mean square error (MSE) of each hop independently, while the second, based on block diagonal geometric mean decomposition (BD-GMD), applies GMD and BD-GMD for the first and the second hops respectively, outperforming the precoding strategies in the literature at the cost of a higher processing complexity.

Besides their better performance in comparison to linear precoding techniques, the optimal transceiver designs based on THP and VP have not been derived yet for multiuser MIMO AF relaying schemes. In this paper, we consider the problem of MMSE design of precoder and relaying algorithms for the aforementioned system. The proposed iterative design schemes greatly outperform the iterative linear design approaches in both performance and iteration efficiency. Furthermore, we propose a suboptimal non-iterative design strategy for VP, which simplifies the MSE optimization problem by dividing it into a master problem and subproblem.

The main contributions of this paper are the following:

1. We propose and derive novel optimal and suboptimal design strategies for non-linear precoding in multiuser MIMO AF relaying downlink systems.

2. We compare the BER performance and the convergence features of the proposed novel iterative algorithms with the optimal linear approach presented in [6].

3. Simulation results are provided showing that the proposed approaches clearly outperform the optimal linear solution. Moreover, we confirm that the proposed iterative approaches minimize the MSE faster, requiring a much lower number of iterations than the optimal linear schemes for similar results.

1.1. Notation

Throughout the paper, we denote vectors and matrices by lower and upper case bold letters, respectively. We use $E\left[\bullet\right]$, $\left(\bullet\right)^T$, $\left(\bullet\right)^H$, $tr\left(\bullet\right)$, $\Re\left(\bullet\right)$ and $\Im\left(\bullet\right)$ for expectation, transpose, Hermitian, trace of a matrix, real and imaginary part respectively. $\lfloor\bullet\rfloor$ represents the floor operator which returns the largest integer that is smaller than or equal to the argument while $\mathbf{I}_N$ stands for an $N \times N$ identity matrix.

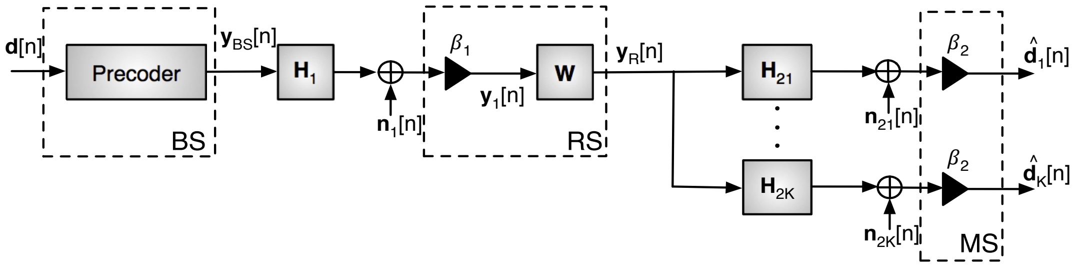

Fig. 1. Multiuser MIMO downlink AF relay system.

2. SYSTEM MODEL

We consider the multiuser MIMO AF relaying system depicted in Fig. 1, which divides the communication between a base station and *K* single antenna mobile stations in two time slots: in the first one, the data streams are transmitted trough the channel $\mathbf{H}_1 \in \mathbb{C}^{R\times M}$, where *R* and *M* correspond to the relay and BS antennas respectively. Once the data streams are received and amplified, they are transmitted across the second channel $\mathbf{H}_2 = \left[\mathbf{h}_{21}^T, \ldots, \mathbf{h}_{2k}^T, \ldots, \mathbf{h}_{2K}^T\right]^T \in \mathbb{C}^{K\times M}$ in the second time slot, where $\mathbf{h}_{2k}$ stands for the channel created between the relay and user *k*.

The transmitted data symbols are represented by vector $\mathbf{d}[n] = \left[d_1^T[n] \ldots d_K^T[n]\right]^T \in \mathbb{C}^K$, where $d_k[n]$ are the symbols transmitted for user *k*. The number of transmitted symbols per frame is set to N_B and the covariance of the transmitted symbols is $E\left[\mathbf{d}[n]\mathbf{d}^H[n]\right] = \mathbf{I}_K$. At the receiver side, the recovered data matrix will be denoted by $\hat{\mathbf{d}}[n] \in \mathbb{C}^K$.

In order to separate the streams of different users, the precoded symbol vector $\mathbf{y}_{BS}[n] \in \mathbb{C}^M$ is generated at the BS. To limit the total transmitted power to $\frac{1}{N_B}\sum_{n=1}^{N_B}\|\mathbf{y}_{BS}[n]\|_2^2 = P_S$, a scaling factor of $1/\beta_1$ is applied at the precoder.

After re-scaling the received symbols, the relay station (RS) applies a non-regenerative process and generates the transmitted signal vector $\mathbf{y}_R[n]$, which will be transmitted in the second time slot:

$$\mathbf{y}_R[n] = \beta_1\mathbf{W}\mathbf{H}_1\mathbf{y}_{BS}[n] + \beta_1\mathbf{W}\mathbf{n}_1[n] \in \mathbb{C}^R,$$

where $\mathbf{W} \in \mathbb{C}^{R\times R}$ is the processing matrix at the relay and $\mathbf{n}_1[n]\mathbb{C}^R$ is the white Gaussian noise (AWGN) vector with covariance matrix $\mathbf{R}_{n_1} = \sigma_1^2\mathbf{I}_R$. In the same way as at the precoder, the relaying matrix $\mathbf{W}$ includes a scaling factor $1/\beta_2$ in order to constraint the total transmitted power at the relay to $\frac{1}{N_B}\sum_{n=1}^{N_B}\|\mathbf{y}_R[n]\|_2^2 = P_R$.

Once the signal is received at the final user terminals, symbols are re-scaled by β_2 yielding:

$$\hat{\mathbf{d}}[n] = \beta_2\beta_1\mathbf{H}_2\mathbf{W}\mathbf{H}_1\mathbf{y}_{BS}[n] + \beta_2\beta_1\mathbf{H}_2\mathbf{W}\mathbf{n}_1[n] + \beta_2\mathbf{n}_2[n],$$

$\mathbf{n}_2[n]$ being the AWGN noise with covariance $\mathbf{R}_{n_2} = \sigma_2^2\mathbf{I}_K$ at the user terminals.

3. OPTIMAL JOINT ITERATIVE APPROACHES

In this section we derive and analyze the optimal transmission and relaying strategies with linear and non-linear precoding at the BS. We propose two novel joint optimal iterative approaches that use THP and VP at the BS, which outperform the joint optimal iterative linear approach introduced in [6], which will be briefly described in the following subsection.

3.1. Joint linear MMSE design (Lin-MMSE-opt)

The optimal linear iterative approach was derived in [6] with the aim of minimizing the mean square error between the BS and the end users. In order to obtain the closed-form expressions of the matrices, the Karush-Kuhn-Tucker (KKT) conditions are applied to the optimization problem, which is subject to two power constraints. The solution, which is obtained as a result of an iterative process, provides a local optimal solution.

The sum-MSE between the BS and the *K* users is defined as

$$\xi = \frac{1}{N_B}\sum_{k=1}^{K}\sum_{n=1}^{N_B}\mathrm{E}\left[\|\hat{d}_k[n] - d_k[n]\|_2^2\right]. \qquad (1)$$

The optimization problem can be formulated as

$$\{\mathbf{F}, \mathbf{W}, \beta_1, \beta_2\} = \underset{\{\mathbf{F},\mathbf{W},\beta_1,\beta_2\}}{\operatorname{argmin}}\ \xi$$

$$\text{s.t} \quad \frac{1}{N_B}\sum_{n=1}^{N_B}\|\mathbf{y}_{BS}[n]\|_2^2 = P_S$$

$$\frac{1}{N_B}\sum_{n=1}^{N_B}\|\mathbf{y}_R[n]\|_2^2 = P_R,$$

where $\mathbf{y}_{BS}[n] = \mathbf{F}\mathbf{d}[n] \in \mathbb{C}^M$ are the precoded symbols and $\mathbf{F} \in \mathbb{C}^{M\times K}$ is the linear precoding matrix. The interdependence of the solutions obtained for matrices $\mathbf{F}$ and $\mathbf{W}$ leads to the use of an iterative process [6], which updates $\mathbf{F}$ for a fixed $\mathbf{W}$ in each iteration and vice versa.

3.2. Joint non-linear MMSE designs

Non-linear precoding will be analysed in this section due to its better performance in comparison to linear precoders in single-hop multiuser MIMO downlink systems [8][9]. We propose and derive two novel designs, based on THP and VP non linear techniques for the multiuser AF relaying scenario under consideration.

3.2.1. Tomlinson-Harashima precoding (THP-MMSE-opt)

Tomlinson-Harashima precoding was first proposed in [14] for inter-symbol interference mitigation and has been extended to multiuser schemes for interuser interference cancellation.

At the BS, two filters are included at the precoder: the feedback filter $\mathbf{B} \in \mathbb{C}^{K\times K}$, which is a lower triangular matrix with zero elements at the main diagonal for successive interference cancellation, and the feedforward filter $\mathbf{F} \in \mathbb{C}^{M\times K}$. In this scenario the transmitted signal $\mathbf{d}[n]$ fulfils $\mathbf{d}[n] = (\mathbf{I}_K - \mathbf{B})\mathbf{v}[n] \in \mathbb{C}^K$, where $\mathbf{v}[n] \in \mathbb{C}^K$ is the input signal at the feedforward filter. The following modulo operator is applied at both transmitter and receiver side:

$$\mathbb{M}(\bullet) = x - \tau\lfloor\frac{\Re(x)}{\tau} + \frac{1}{2}\rfloor + j\tau\lfloor\frac{\Im(x)}{\tau} + \frac{1}{2}\rfloor, \quad (2)$$

where *x* is the symbol to be mapped and τ denotes the modulo constant, which is a scalar that provides a symmetric decoding region around every signal constellation point and depends on the modulation constellation [15]. Put in other words, the modulo operator, which is applied to reduce the power increments due to $\mathbf{B}$, maps both the real and imaginary parts of its input to the interval $\left[\frac{-\tau}{2}, \frac{\tau}{2}\right)$, where the final symbol power is lower.

The MMSE solution is obtained minimizing the sum-MSE defined in (1) subject to two power constraints. The optimization problem can be stated as:

$$\begin{aligned}
&\{\mathbf{F}, \mathbf{W}, \mathbf{B}, \beta_1, \beta_2\} = \underset{\{\mathbf{F},\mathbf{W},\mathbf{B},\beta_1,\beta_2\}}{\text{argmin}} \ \xi \\
&\text{s.t} \quad \text{Tr}\left(\mathbf{F}\mathbf{R}_{vv}\mathbf{F}^H\right) = P_S \\
&\quad \beta_1^2 \text{Tr}\left(\mathbf{W}\left(\mathbf{H}_1\mathbf{F}\mathbf{R}_{vv}\mathbf{F}^H\mathbf{H}_1^H + \sigma_1^2\mathbf{I}_R\right)\mathbf{W}^H\right) = P_R \\
&\quad \mathbf{S}_k\mathbf{B}\mathbf{e}_k = 0 \quad k = 1\ldots K, \qquad (3)
\end{aligned}$$

where (3) is the constraint established for the triangularization of $\mathbf{B}$. $\mathbf{S}_k = [\mathbf{I}_k, \mathbf{0}_{k\times K-k}] \in \{0,1\}^{k\times K}$ is the selection matrix which selects the first *k* elements from a *K*-dimensional vector, while $\mathbf{e}_k$ corresponds to the k^{th} column of $\mathbf{I}_K$. $\mathbf{R}_{vv} \in \mathbb{C}^{R\times R}$ corresponds to the covariance matrix of the precoded symbols as defined in [8].

After we solve the optimization problem, the following expressions are obtained:

$$\overline{\mathbf{F}} = \left(\mathbf{H}_1^H\overline{\mathbf{W}}^H\left(\mathbf{H}_2^H\mathbf{H}_2 + \alpha_2\right)\overline{\mathbf{W}}\mathbf{H}_1 + \alpha_1\right)^{-1}\mathbf{H}_1^H\overline{\mathbf{W}}^H\mathbf{H}_2^H\overline{\mathbf{B}} \quad (4)$$

$$\overline{\mathbf{B}} = \mathbf{I}_K - \sum_{k=1}^{K}\left(\mathbf{S}_k^T\mathbf{S}_k - \mathbf{I}_K\right)\mathbf{H}_2\overline{\mathbf{W}}\mathbf{H}_1\overline{\mathbf{F}}\mathbf{e}_k\mathbf{e}_k^T$$

$$\overline{\mathbf{W}} = \left(\mathbf{H}_2^H\mathbf{H}_2 + \alpha_2\right)^{-1}\overline{\mathbf{B}}\mathbf{R}_{vv}\overline{\mathbf{F}}^H\mathbf{H}_1^H\left(\mathbf{H}_1\overline{\mathbf{F}}\mathbf{R}_{vv}\overline{\mathbf{F}}\mathbf{H}_1^H + \beta_1^2\sigma_1^2\mathbf{I}\right)^{-1},$$

where $\overline{\mathbf{F}} = \beta_1\mathbf{F}$ and $\overline{\mathbf{W}} = \beta_2\mathbf{W}$ are the unnormalized precoding and relaying matrices. The expressions for α_1 and α_2 are the following: $\alpha_1 = \frac{\sigma_1^2}{P_S}\text{Tr}\left(\overline{\mathbf{W}}^H\left(\mathbf{H}_2^H\mathbf{H}_2 + \alpha_2\right)\overline{\mathbf{W}}\right)\mathbf{I}_M$ and $\alpha_2 = \frac{\sigma_2^2 K}{P_R}\mathbf{I}_R$, respectively.

As it happens for the linear solution, due to the interdependence of the solutions provided for matrices $\mathbf{F}$ and $\mathbf{W}$, an iterative process has to be applied to get to the optimal solution.

3.2.2. *Vector precoding approach (VP-MMSE-opt)*

The vector perturbation approach or vector precoding [9] was inspired by the idea of THP. Instead of applying an iterative process for the interference cancellation, VP adds a perturbation vector $\mathbf{a}[n]$, removing the need for a modulo operation and a feedback filter at the transmitter side. The novel MMSE-VP solution proposed here for multuser AF-relaying scenarios computes the perturbation vector $\mathbf{a}[n]$, the precoding matrix $\mathbf{F}$ and the relaying matrix $\mathbf{W}$ that minimize the MSE of the whole transmission.

In order to get the complete MMSE solution, we define the optimization problem as

$$\begin{aligned}
&\{\mathbf{a}[n], \mathbf{F}, \mathbf{W}, \beta_1, \beta_2\} = \underset{\{\mathbf{a},\mathbf{F},\mathbf{W},\beta_1,\beta_2\}}{\text{argmin}} \ \xi \\
&\text{s.t} \quad \frac{1}{N_B}\sum_{n=1}^{N_B}\|\mathbf{y}_{BS}[n]\|_2^2 = P_S \\
&\quad \frac{1}{N_B}\sum_{n=1}^{N_B}\|\mathbf{y}_R[n]\|_2^2 = P_R,
\end{aligned}$$

where ξ is the sum-MSE defined in (1). For VP approaches, $\mathbf{d}[n] = \mathbf{s}[n] + \mathbf{a}[n]$ is considered, $\mathbf{a}[n]$ being the perturbation vector, while $\hat{\mathbf{d}}[n]$ stands for the received symbols before the modulo operation.

This optimization problem leads us to the next unnormalized precoding matrix

$$\overline{\mathbf{F}} = \left(\mathbf{H}_1^H\overline{\mathbf{W}}^H\left(\mathbf{H}_2^H\mathbf{H}_2 + \alpha_2\right)\overline{\mathbf{W}}\mathbf{H}_1 + \alpha_1\right)^{-1}\mathbf{H}_1^H\overline{\mathbf{W}}^H\mathbf{H}_2^H,$$

where α_1 and α_2 are described in the previous section. In the same way, the unnormalized processing matrix stands for

$$\overline{\mathbf{W}} = \left(\mathbf{H}_2^H\mathbf{H}_2 + \alpha_2\right)^{-1}\mathbf{H}_2^H\mathbf{D}\overline{\mathbf{Y}}_{BS}^H\mathbf{H}_1^H\left(\mathbf{H}_1\overline{\mathbf{Y}}_{BS}\overline{\mathbf{Y}}_{BS}^H\mathbf{H}_1^H + \gamma\right)^{-1},$$

where $\gamma = \left(\frac{\sigma_1^2}{P_S}\right)\text{Tr}\left(\overline{\mathbf{Y}}_{BS}^H\overline{\mathbf{Y}}_{BS}\right)\mathbf{I}_M$. Matrix $\overline{\mathbf{Y}}_{BS} = \frac{1}{\beta_1}\mathbf{F}\mathbf{D}$ and $\mathbf{D}$ contain the unnormalized precoded symbols and perturbed symbols respectively.

In the same way, the expressions for the precoding factors are the following

$$\beta_i = \sqrt{\frac{\sum_{n=1}^{N_B}\mathbf{y}_i^H[n]\,\mathbf{y}_i[n]}{N_B P_i}} \quad i = 1, 2 \quad (5)$$

where for the first hop (i=1), $\mathbf{y}_i[n] = \mathbf{y}_{BS}[n]$ and $P_i = P_S$, while for the second one (i=2), $\mathbf{y}_i[n] = \mathbf{y}_R[n]$ and $P_i = P_R$.

Once $\mathbf{W}$ and $\mathbf{F}$ are found, the optimal perturbation vector which minimizes the overall MSE between the source and the end users can be obtained. After some manipulations we get:

$$\xi = \frac{1}{N_B}\sum_{n=1}^{N_B}\mathbf{d}^H[n]\left(\mathbf{I}_K - \mathbf{H}_2\overline{\mathbf{W}}\mathbf{H}_1\overline{\mathbf{F}}\right)\mathbf{d}[n]. \quad (6)$$

Using the Cholesky factorization of the inner term in (6)

$$\left(\mathbf{I}_K - \mathbf{H}_2\overline{\mathbf{W}}\mathbf{H}_1\overline{\mathbf{F}}\right) = \mathbf{L}^H\mathbf{L},$$

the computation of $\mathbf{a}$ can be simplified to [9]

$$\mathbf{a}[n] = \underset{\mathbf{a}'[n]\in\tau\mathbb{Z}^K + j\tau\mathbb{Z}^K}{\text{argmin}}\|\mathbf{L}\left(\mathbf{s}[n] + \mathbf{a}'[n]\right)\|_2^2. \quad (7)$$

As it happens with the aforementioned *Lin-MMSE-opt* and *THP-MMSE-opt*, an iterative algorithm has to be implemented. This time, it must include the search of the optimal precoding vector in each iteration. This process is executed until a reference convergence error is achieved. As it will be shown in the simulation results, the addition of the perturbation vector improves considerably the performance of the joint designs, but the search of the perturbation vector in each iteration increases the computational complexity. In order to reduce it, we propose and derive the suboptimal solution that will be shown in the next section.

3.3. Suboptimal vector precoding approach (VP-MMSE-sub)

In this section we propose a novel suboptimal system for the design of the precoding and relaying matrices with VP non-linear precoding.

After manipulating (5) and substituting it in the sum-MSE definition in (1), the MSE is equivalent to

$$\xi = \frac{1}{N_B}\sum_{n=1}^{N_B} \mathbf{d}^H[n]\,\mathbf{X}\mathbf{d}[n], \tag{8}$$

where $\mathbf{X} = \overline{\mathbf{F}}^H \mathbf{H}_1^H \overline{\mathbf{W}}^H \left(\mathbf{H}_2^H \mathbf{H}_2 + \frac{\sigma_2^2 K}{P_R}\mathbf{I}_R \overline{\mathbf{W}}\mathbf{H}_1\overline{\mathbf{F}}\right) + \frac{\sigma_1^2}{P_S}\overline{\mathbf{F}}^H\overline{\mathbf{F}}$ $\mathrm{Tr}\left(\overline{\mathbf{W}}^H\left(\mathbf{H}_2^H\mathbf{H}_2 + \frac{\sigma_2^2 K}{P_R}\mathbf{I}_R\right)\overline{\mathbf{W}}\right)$. It can be seen in (8) that $\mathbf{F}$ and $\mathbf{W}$ are constant over the perturbed symbols.

As a design strategy and in order to avoid the search of the precoding vector in each iteration, we divide the optimization problem in a subproblem and a master problem. While the first searches the optimal precoding and relaying matrices for a fixed perturbation vector taking into account the power constraints, the second finds the vector $\mathbf{a}[n]$ that minimizes the MSE with the previously selected locally optimal $\mathbf{F}$ and $\mathbf{W}$. The optimization problem can be then defined as follows:

$$\{\mathbf{a}[n], \overline{\mathbf{F}}, \overline{\mathbf{W}},\} = \overbrace{\min_{\mathbf{a}[n]}\left(\frac{1}{N_B}\sum_{n=1}^{N_B}\mathbf{d}[n]\left(\underbrace{\min_{\overline{\mathbf{F}},\overline{\mathbf{W}}}\left(\mathrm{Tr}\left(\mathbf{X}\right)\right)}_{\text{sub-problem}}\right)\mathbf{d}[n]\right)}^{\text{master problem}}$$

$$\text{st.} \quad \frac{1}{N_B}\sum_{n=1}^{N_B}\|\mathbf{y}_{BS}[n]\|_2^2 = P_S$$

$$\frac{1}{N_B}\sum_{n=1}^{N_B}\|\mathbf{y}_{R}[n]\|_2^2 = P_R.$$

Solving the subproblem,we get the expressions:

$$\overline{\mathbf{F}} = (\mathbf{H}_1^H\overline{\mathbf{W}}^H\left(\mathbf{H}_2^H\mathbf{H}_2 + \alpha_2\right)\overline{\mathbf{W}}\mathbf{H}_1 + \alpha_1)^{-1}\mathbf{H}_1^H\overline{\mathbf{W}}^H\mathbf{H}_2^H$$

$$\overline{\mathbf{W}} = \left(\mathbf{H}_2^H\mathbf{H}_2 + \alpha_2\right)^{-1}\mathbf{H}_2^H\overline{\mathbf{F}}^H\mathbf{H}_1^H(\mathbf{H}_1\overline{\mathbf{F}}\overline{\mathbf{F}}^H\mathbf{H}_1^H + \phi_W\mathbf{I}_R)^{-1},$$

where α_1 and α_2 were introduced before and $\phi_W = \frac{\sigma_1^2}{P_S}\mathrm{Tr}\left(\overline{\mathbf{F}}\overline{\mathbf{F}}^H\right)$. It turns out that these expressions lead to a unique solution for a fixed perturbation vector. Due to the interdependence of the matrices, an iterative process is applied, which does not now depend on the perturbation vector $\mathbf{a}[n]$. Once $\mathbf{F}$ and $\mathbf{W}$ are derived, the master problem is solved under the assumption that the optimal precoding and relaying matrices are employed.

Carrying out the same procedure realized for the *VP-MMSE-opt* system, we have to find the perturbation vector that minimizes the overall MSE between the BS and the end users. Taking into account the MSE defined in (8) and using the Cholesky factorization for the inner term as $\mathbf{X} = \mathbf{L}^H\mathbf{L}$, the search of $\mathbf{a}[n]$ can be established as defined in (7).

As it will be seen in the provided simulation results, this suboptimal design strategy outperforms both optimal linear and THP designs. Furthermore, the cost due to the inclusion of VP is greatly reduced due to the fact that the perturbation vector is not searched in each iteration.

4. SIMULATION RESULTS

We consider a system composed of a BS with *M=4* antennas, which wants to communicate with *K=4* single-antenna users through a relay with *R=4* antennas. An ensemble of 10^4 channel realizations have been simulated for matrices $\mathbf{H}_i, i = \{1, 2\}$, whose coefficients are drawn from independent and identically distributed complex Gaussian processes. $N_B = 100$ quadrature phase shift keying (QPSK) symbols are transmitted per user and channel realization. The signal-to-noise ratios (SNR) are independent for each hop and are defined as $\mathrm{SNR}_1 = \frac{P_S}{\sigma_1^2}$ and $\mathrm{SNR}_2 = \frac{P_R}{\sigma_2^2}$ for the first and the second hops, respectively. A maximum convergence error, i.e. variation from iteration to iteration, of $\varepsilon = 0.0025$ has been set as stop criterion for the iterative algorithms.

Four different schemes have been evaluated: the *Lin-MMSE-opt* scheme introduced in [6], the joint optimal approaches with non-linear transmission (*THP-MMSE-opt* and *VP-MMSE-opt*) and the novel suboptimal solution named *VP-MMSE-sub*.

Fig. 2 shows the BER performance curves for a fixed SNR of 15 dB in the second hop. Clearly, all the non-linear precoding schemes outperform the linear optimal approach. *VP-MMSE-opt* exhibits the best performance at expense of a higher computational cost. The suboptimal scheme designed for the reduction of the computational cost improves the BER of the joint optimal linear and THP approaches. It can also be seen in Fig. 2 that the schemes proposed with vector precoding (*VP-MMSE-opt* and *VP-MMSE-sub*) outperform the joint THP precoder, specially at high SNRs.

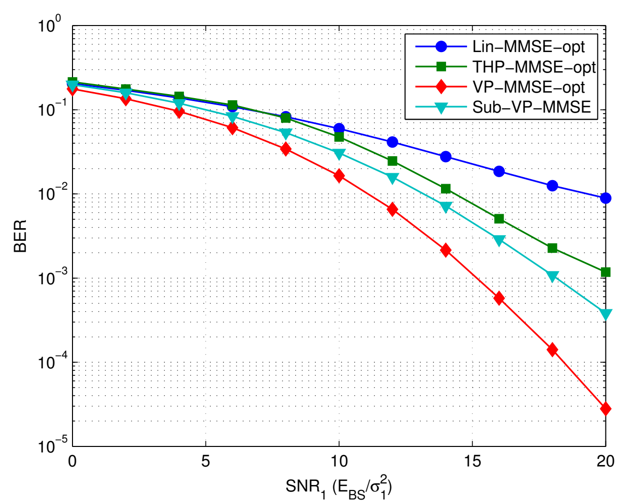

Fig. 2. BER performance for a fixed $\mathrm{SNR}_2 = 15$ dB.

The same conclusions hold for Fig. 3, where the SNR at the first hop has been set to 15 dB. The optimal non-linear proposals *THP-MMSE-opt* and *VP-MMSE-opt* outperform the linear optimal approach, being *VP-MMSE-opt* once again the one with the best performance. Apart from that, the suboptimal solution presents a slight gain over *THP-MMSE-opt*, reaching *THP-MMSE-opt*'s performance at high ranges of SNR_2.

Finally Fig. 4 shows the convergence properties of the proposed approaches. We compare the sum-MSE versus the number of iterations for a scenario with fixed values of $\mathrm{SNR}_1 = 10$ dB and $\mathrm{SNR}_2 = 15$ dB. The figure shows that all the algorithms converge in a few number of iterations. We firmly assert that the mean square

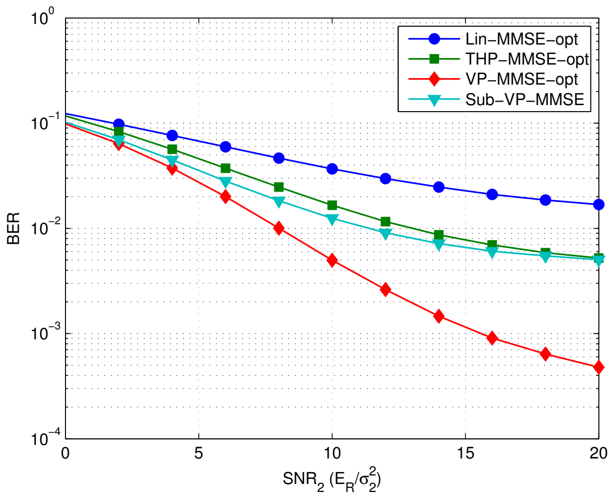

Fig. 3. BER performance for a fixed $\mathrm{SNR}_1 = 15$ dB.

error is reduced applying non-linear precoding techniques, being the VP approaches the ones which reduce the error faster. Apart from that, simulation results prove that the non-linear precoding designs decrease the error faster than the linear one and outperform the linear error after a few number of iterations.

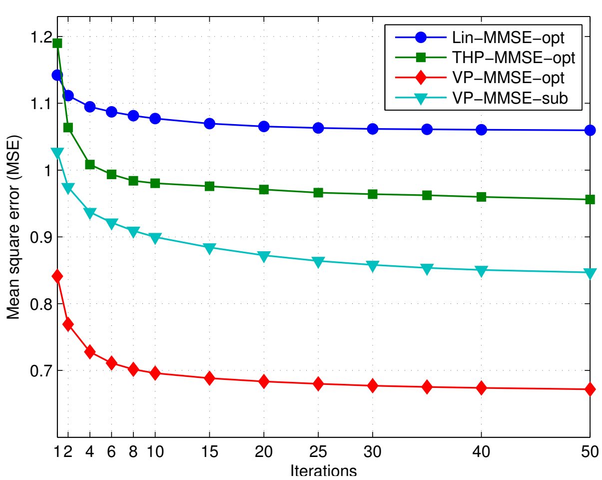

Fig. 4. Convergence error for a fixed $\mathrm{SNR}_2 = 15$ dB.

5. CONCLUSIONS

The design of precoding and relaying schemes is analysed in this paper for the downlink of multiuser MIMO AF relay systems. We have proposed three novel strategies for the design of the precoding and relaying matrices. In the first one, the optimal MMSE solution is derived for the joint design of the THP precoder at the BS and the processing matrices at the relay. In the second one, vector perturbation is applied and a complete MMSE optimization criterion is established to get the joint design of the aforementioned matrices and the perturbation vector. Finally and in order to show the effectiveness of VP, we have proposed a suboptimal MMSE design based on this technique with a lower computational complexity, which is achieved by decomposing the overall optimization problem in a subproblem and master problem. Provided BER simulation results show that these non-linear design strategies clearly outperform the optimal linear iterative approach, considered the best proposed in the literature for AF relaying systems. Furthermore, we show that our non-linear proposals minimize the error faster and require a lower number of iterations to converge.

6. REFERENCES

[1] G.J. Foschini and M.J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wir. per. com.*, 1998.

[2] E. Telatar, "Capacity of multi-antenna gaussian channels," *Eur. trans. on telecom.*, 1999.

[3] J. Winters, "On the capacity of radio communication systems with diversity in a rayleigh fading environment," *IEEE Jour. on Selec. Areas in Com.*, 1987.

[4] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. on Inform. Theory*, 2005.

[5] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity. part i. system description," *IEEE Trans. on Com.*, 2003.

[6] G. Li, Y. Wang, T. Wu, and J. Huang, "Joint linear filter design in multi-user cooperative non-regenerative mimo relay systems," *EURASIP Journal on Wir. Com. and Networking*, 2009.

[7] S. Jang, J. Yang, and D.K. Kim, "Minimum mse design for multiuser mimo relay," *IEEE Com. Let.*, 2010.

[8] K. Kusume, M. Joham, W. Utschick, and G. Bauch, "Efficient tomlinson-harashima precoding for spatial multiplexing on flat mimo channel," in *IEEE Intern. Conf. on Com.*, 2005.

[9] D.A. Schmidt, M. Joham, and W. Utschick, "Minimum mean square error vector precoding," *European Trans. on Telecom.*, 2008.

[10] A.P. Millar, S. Weiss, and R.W. Stewart, "Tomlinson harashima precoding design for non-regenerative mimo relay networks," in *IEEE 73rd Vehic.r Tech. Conf.*, 2011.

[11] C.B. Chae, T. Tang, R.W. Heath, and S. Cho, "Mimo relaying with linear processing for multiuser transmission in fixed relay networks," *IEEE Trans. on Signal Proc.*, 2008.

[12] H. Kim, S. Lee, K. Kwak, H. Min, and D. Hong, "On the design of zf and mmse tomlinson-harashima precoding in multiuser mimo amplify-and-forward relay system," in *Intern. Symp. on Pers., Indoor and Mob. Radio Com.*, 2009.

[13] I. Jimenez, S. Weiss, M. Mendicute, and E. Arruti, "Multiuser mimo amplify-and-forward relaying schemes with vector precoding," *IEEE Intern. Symp. on Sig. Proc. and Inform. Tech.*, 2011.

[14] M. Tomlinson, "New automatic equaliser employing modulo arithmetic," *Electronics letters*, , no. 5, 1971.

[15] F.A. Dietrich, *Robust signal processing for wireless communications*, Springer Verlag, 2008.