Adaptive Self-interference Cancellation in Wideband Full-Duplex Decode-and-Forward MIMO Relays

Emilio Antonio-Rodríguez*†, Roberto López-Valcarce*, Taneli Riihonen†, Stefan Werner†, and Risto Wichman†
*Department of Signal Theory and Communications, University of Vigo, Vigo, Spain
†Department of Signal Processing and Acoustics, Aalto University, Helsinki, Finland

***Abstract*—Full-duplex relays suffer from self-interference, due to simultaneous reception and transmission in the same frequency. This self-induced interference signal of the relay can be of significant power and considerably distort the information-bearing signal, leading to unacceptable performance degradation. This paper presents an adaptive gradient-based method for full-duplex decode-and-forward multiple-input multiple-output (MIMO) relays that mitigates the self-interference signal at the relay. The scheme makes use of the cross-correlation between signals available at the relay to adaptively estimate the self-interference channel in order to cancel it. We analyze the behaviour of the algorithm in terms of stationary points and mean convergence, and the performed experiments show that the proposed method can mitigate the self-interference signal by over 30 dB in a low SNR scenario.**

Keywords*—Full-duplex, regenerative relaying, self-interference, MIMO, adaptive filtering.

I. Introduction

Relays are devices widely used for coverage extension and cooperative communications, especially in the context of single frequency networks (SFNs). At the same time, modern communication techniques like spatial multiplexing demand the use of several antennas in order to avoid the key-hole effect [1] while allowing the use of precoding schemes, justifying the need to use several antennas in both transmission and reception in the relay [2]. Based on the transmission mode, relays can be classified into half-duplex [3] and full-duplex [4]. While half-duplex relays receive and transmit in orthogonal time slots, full-duplex relays receive and transmit at the same time, and consequently they can achieve a higher spectral efficiency than the half-duplex counterpart. However, full-duplex relays suffer from self-interference, due to simultaneous transmission and reception in the same frequency. This self-interference signal may severely impact the performance of the relay in the case of a high power imbalance between the self-interference signal and the information-bearing signal arriving at the relay. Although some physical isolation between antennas on the receive and transmit side of the relay is usually provided during the design stage [5]–[7], large power imbalance can make the physical isolation insufficient, and therefore hampering the relay performance. Consequently, in order to exploit the full potential of full-duplex relaying, efficient self-interference mitigation techniques are needed, like for example [2], [5], [8]–[10].

This work was partially supported by Fundación Pedro Barrié de la Maza under the program Becas de Postgrado en el Extranjero 2012, the Centre of Excellence in Smart Radios and Wireless Research (SMARAD) and by the Academy of Finland.

Among the different types of full-duplex relaying protocols, decode-and-forward relays decode and re-encode the message from the original source [11]. This additional complexity renders better system performance compared to amplify-and-forward relays, which only forward messages without being aware of their content. Therefore, decode-and-forward relays contain a decoder device, whose mission is to regenerate and re-encode the source data streams. Furthermore, Decode-and-forward relays can be classified according to the decoder characteristics of the relay: Repetition coding-based relays use the same codebook as in the transmitter, but the relay could also use a different codebook than in the transmitter, which translates into a multiple access channel at the final destination.

Nevertheless, to ensure reliable decoding of the information signal in the relay, it is necessary to mitigate any interference signal arriving at the relay. Otherwise the decoded signal may contain errors which will propagate to the final destination. In fact, this imposes a maximum interference power level that can be present in the relay for a desired performance at the final destination. In contrast to the decode-and-forward case, in the amplify-and-forward case the signal does not contain errors but only a distortion (though it can be highly selective in frequency [9]). As explained before, in a full-duplex relay the major source of interference is usually the relay itself, and consequently, the use of efficient self-interference mitigation schemes in a decode-and-forward relay is extremely important, since the overall link quality is bounded by the performance achieved in the relay.

In this work, we propose a method to mitigate the self-interference signal for full-duplex decode-and-forward MIMO relays. The characteristics of the decode-and-forward relay allows to identify the self-interference signal using a gradient-descent approach. Furthermore, the use of an adaptive solution makes possible to track the temporal variations of the self-interference channel at the relay.

Notation: Throughout the paper, discrete time-domain signals are functions of the discrete time index n, e.g., $\mathbf{x}(n)$. Similarly, frequency-domain systems are denoted by their z-transform, e.g., $\mathbf{A}(z) = \sum_k \mathbf{A}[k]\, z^{-k}$, and, z, when applied to signals, represents the delay operator, i.e., $z^{-k}\{\mathbf{x}(n)\} = \mathbf{x}(n-k)$. Therefore, the output of a generic linear time-invariant system can be expressed as $\mathbf{y}(n) = \mathbf{H}_{ab}(z)\,\mathbf{x}(n)$. The matrix $\mathbf{H}_{ab}(z)$ represents a causal channel of order L_{ab}, i.e., $\mathbf{H}_{ab}(z) = \sum_{k=0}^{L_{ab}} \mathbf{H}_{ab}[k] z^{-k}$. Also, the power spectral density of $\mathbf{x}(n)$ is represented by the matrix $\mathbf{R}_{xx}(z)$, whereas $P_x = 1/2\pi \,\mathrm{tr} \int_{-\pi}^{\pi} \mathbf{R}_{xx}(e^{j\omega}) e^{j\omega} d\omega$ denotes its power.

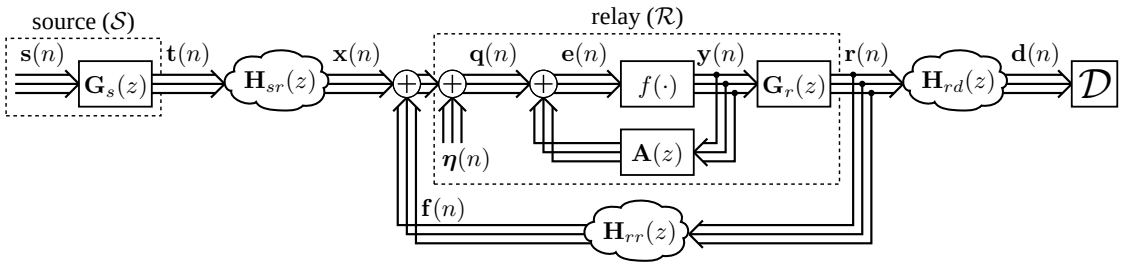

Fig. 1. System model of a single-frequency decode-and-forward relay with self-interference mitigation.

II. System Model

Figure 1 depicts the considered single-frequency MIMO relay link, which is composed of a source node ($\mathcal{S}$) equipped with M_t transmit antennas, and a relay ($\mathcal{R}$) with N_r and M_r receive and transmit antennas, respectively. Finally, the destination receiver ($\mathcal{D}$) has N_d antennas.

At time instant n, $\mathcal{S}$ transmits the signal $\mathbf{t}(n)$, $\mathcal{R}$ receives $\mathbf{q}(n)$ while transmitting $\mathbf{r}(n)$, and $\mathcal{D}$ receives the signal $\mathbf{d}(n)$. The expressions for both $\mathbf{q}(n)$ and $\mathbf{d}(n)$ are

$$\mathbf{q}(n) = \mathbf{H}_{sr}(z)\mathbf{t}(n) + \mathbf{H}_{rr}(z)\mathbf{r}(n) + \boldsymbol{\eta}(n) \quad (1)$$

$$\mathbf{d}(n) = \mathbf{H}_{rd}(z)\mathbf{t}(n) + \boldsymbol{\nu}(n) \quad (2)$$

where $\mathbf{H}_{sr}(z)$, with size $N_r \times M_t$ and order L_{sr}, denotes the channel from $\mathcal{S}$ to $\mathcal{R}$, $\mathbf{H}_{rd}(z)$, with size $N_d \times M_r$ and order L_{rd}, is the channel from $\mathcal{R}$ to $\mathcal{D}$ and $\mathbf{H}_{rr}(z)$, with size $N_r \times M_r$ and order L_{rr}, is the self-interference channel at the relay. Finally, $\boldsymbol{\eta}(n)$ and $\boldsymbol{\nu}(n)$ are the additive noise sources at the relay and destination, respectively.

We assume that in both nodes, $\mathcal{S}$ and $\mathcal{R}$, linear precoding is applied before transmission. Consequently,

$$\mathbf{t}(n) = \mathbf{G}_s(z)\mathbf{s}(n) \quad (3)$$

$$\mathbf{r}(n) = \mathbf{G}_r(z)\mathbf{y}(n) \quad (4)$$

where $\mathbf{s}(n)$ represents the original $M_s \leq M_t$ data streams that are precoded by $\mathbf{G}_s(z)$, of size $M_t \times M_s$ and order L_s. In the same way, $\mathbf{y}(n)$ represents the M_s retrieved data streams at the relay, and $\mathbf{G}_r(z)$, of size $M_r \times M_s$ and order L_r, the precoding matrix for transmitting $\mathbf{r}(n)$.

Every decode-and-forward relay contains a decoder device, which regenerates the original M_s source data streams by using the received signal at $\mathcal{R}$. In general, we consider the decoder to be a nonlinear function of past samples of its input, $\mathbf{e}(n)$, i.e., the decoder output $\mathbf{y}(n)$ can be expressed as $\mathbf{y}(n) = f(\mathbf{e}(n-\tau), \mathbf{e}(n-\tau-1), \ldots)$, being τ the decoding delay. Therefore, in the ideal case, $\mathbf{y}(n) = \mathbf{s}(n-(\tau+\delta_{sr}))$ (where δ_{sr} accounts for propagation and transmission delay between the source and the relay) if the same coding scheme is applied at both hops [12].

Since the decoder regenerates the data streams, the associated decoding delay is expected to be relatively high, in comparison to the amplify-and-forward case, where such decoding process is not present. Throughout this paper, we assume that τ is long enough so that $\mathbf{t}(n)$ and $\mathbf{y}(n)$ are uncorrelated in the sense that $\mathbb{E}\{\mathbf{t}(n)\mathbf{y}^H(n-k)\} \approx \mathbf{0}$, for any $k \geq 0$.

For simplicity, let $\mathbf{x}(n) = \mathbf{H}_{sr}(z)\mathbf{t}(n)$ and $\mathbf{f}(n) = \mathbf{H}_{rr}(z)\mathbf{r}(n)$, so (1) can be expressed as

$$\mathbf{q}(n) = \mathbf{x}(n) + \mathbf{f}(n) + \boldsymbol{\eta}(n) \quad (5)$$

Since the ratio P_f/P_x can be very high (even tens of dBs), the normally neglected transmit side noise, which is due to distortion effects in practical implementations, can be of significant power at the relay receive side. Therefore, $\boldsymbol{\eta}(n)$ is decomposed into two different noise sources, i.e, $\boldsymbol{\eta}(n) = \boldsymbol{\gamma}(n) + \boldsymbol{\sigma}(n)$, where $\boldsymbol{\gamma}(n)$ represents the additive noise at reception and $\boldsymbol{\sigma}(n) = \mathbf{H}_{rr}(z)\boldsymbol{\varepsilon}(n)$ represents the additive noise propagated from the relay transmit side. We regard the noise sources as independent from each other and from the information-bearing signal.

III. Self-interference Mitigation

In this section, we detail the proposed method for self-interference mitigation. First we describe the architecture used and later we outline an adaptation rule for it.

A. Proposed Architecture

In order to mitigate the self-interference signal $\mathbf{f}(n)$, we propose the cancellation architecture depicted in Figure 1. It consists of the L_a-th order matrix filter $\mathbf{A}(z)$, with $N_r \times M_r \times (L_a+1)$ adjustable parameters. When compared to suppression schemes, where the relay input is filtered to achieve mitigation, a cancellation scheme does not increase the overall delay in the relay, which can be important for some applications. Besides, suppression schemes that filter the signal at the relay transmit side (for nullspace transmission in $\mathbf{H}_{rr}(z)$) may disturb the link between the relay and the destination.

The cancellation scheme proposed here does not modify the actual architecture of the relay or its normal operation. In the best case, the relay can be designed as it was free of self-interference. The decoder input signal, $\mathbf{e}(n)$, has the expression

$$\mathbf{e}(n) = \mathbf{x}(n) + \mathbf{i}(n) + \boldsymbol{\eta}(n) \quad (6)$$

where $\mathbf{i}(n) = \mathbf{f}(n) + \mathbf{A}(z)\mathbf{y}(n)$ is the residual self-interference after mitigation. Equation (6) clearly shows that in order to cancel self-interference in the relay we must select $\mathbf{A}(z) = -\mathbf{H}_{rr}(z)\mathbf{G}_r(z)$, yielding $\mathbf{e}(n) = \mathbf{x}(n) + \boldsymbol{\eta}(n)$. Consequently, to achieve perfect cancellation $L_a \geq L_r + L_{rr}$. To evaluate the performance of the self-interference mitigation scheme we define the signal-to-interference ratio before mitigation and after mitigation as $\text{SIR}_{pre} = P_x/P_f$ and $\text{SIR}_{post} = P_x/P_i$, respectively. Additionally, we define the signal-to-noise ratio as $\text{SNR} = P_x/P_\eta$.

B. Adaptation Rule

Once we have set the architecture of the mitigation scheme, we need a rule for adapting the coefficients of $\mathbf{A}(z)$. In the first place, note that the power spectral density of $\mathbf{e}(n)$ has the expression

$$\mathbf{R}_{ee}(z) = \mathbf{R}_{xx}(z) + \mathbf{R}_{ii}(z) + \mathbf{R}_{\eta\eta}(z) \tag{7}$$

where $\mathbf{R}_{ii}(z) = (\mathbf{A}(z) + \mathbf{H}_{rr}(z)\mathbf{G}_r(z))\mathbf{R}_{yy}(z)(\mathbf{A}^H(1/z^*) + \mathbf{G}_r^H(1/z^*)\mathbf{H}_{rr}^H(1/z^*))$. We see from (7) that P_e has a minimum with respect to $\mathbf{A}(z)$ when $\mathbf{A}(z) = -\mathbf{H}_{rr}(z)\mathbf{G}_r(z)$ and consequently, a gradient-descent approach for the adaptation rule is possible. Note that this minimum is unique if $\mathbf{R}_{yy}(z)$ is full rank. The proposed adaptation rule for $\mathbf{A}(z)$ is then

$$\mathbf{A}[k](n+1) = \mathbf{A}[k](n) + \mu_a \left(\mathbf{R}_\star[k](n) - \mathbf{e}(n)\mathbf{y}^H(n-k)\right) \tag{8}$$

for $k = 0, \ldots, L_a$, where μ_a is the step-size of the algorithm and $\mathbf{R}_\star[k](n)$ are L_a+1 user-defined matrix sequences. When $\mathbf{R}_\star[k](n) = \mathbf{0}$, then (8) is a gradient descent of P_e with respect to $\mathbf{A}(z)$. Next section details how to select $\mathbf{R}_\star[k](n)$ for several cases.

It is clear from (8) that the algorithm only makes use of the input and output decoder signal in the adaptation, which are easily accessible inside the relay architecture. As a consequence, the mitigation scheme and the relay can be designed independently from each other. Moreover, the mitigation scheme does not need information about $\mathbf{G}_r(z)$, since it can be assimilated into the self-interference channel.

IV. Algorithm Analysis

In this section, we analyze the behaviour of the algorithm. We provide an expression for the stationary points, and we study the mean convergence properties of the algorithm.

A. Stationary Points

In this section, we show that the adaptation rule in (8) leads to an estimate of the self-interference multipath channel when the algorithm has converged. In fact, upon convergence of the algorithm, any stationary point (denoted by $\mathbf{A}_\star(z)$) will fulfill

$$\mathbb{E}\{\mathbf{e}(n)\mathbf{y}^H(n-k)\} = \mathbb{E}\{\mathbf{R}_\star[k](n)\} \tag{9}$$

for $k = 0, \ldots, L_a$. Using (6), the cross-correlation between $\mathbf{e}(n)$ and $\mathbf{y}(n)$ can be expressed as

$$\mathbf{R}_{ey}(z) = (\mathbf{A}_\star(z) + \mathbf{H}_{rr}(z)\mathbf{G}_r(z))\mathbf{R}_{yy}(z) \tag{10}$$

If we define $\mathbf{H}_{eq}(z) \triangleq \mathbf{H}_{rr}(z)\mathbf{G}_r(z)$ with order $L_{eq} = L_{rr} + L_r$, then (9) is equivalent to

$$\mathbf{A}_\star \mathbf{R}_Y^{(L_a, L_a)} + \mathbf{H}\mathbf{R}_Y^{(L_{eq}, L_a)} = \mathbf{R}_\star \tag{11}$$

where $\mathbf{A}_\star = [\mathbf{A}_\star[0] \ldots \mathbf{A}_\star[L_a]]$ contains the coefficients of $\mathbf{A}(z)$. In the same way, $\mathbf{H} = [\mathbf{H}_{eq}[0] \ldots \mathbf{H}_{eq}[L_{eq}]]$, and $\mathbf{R}_\star = \mathbb{E}\{[\mathbf{R}_\star[0] \ldots \mathbf{R}_\star[L_a]]\}$. The matrix $\mathbf{R}_Y^{(\alpha,\beta)}$ refers to

$$\mathbf{R}_Y^{(\alpha,\beta)} = \begin{pmatrix} \mathbf{R}_{yy}[0] & \mathbf{R}_{yy}[1] & \ldots & \mathbf{R}_{yy}[\beta] \\ \mathbf{R}_{yy}[-1] & \mathbf{R}_{yy}[0] & \ldots & \mathbf{R}_{yy}[\beta-1] \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{R}_{yy}[-\alpha] & \mathbf{R}_{yy}[1-\alpha] & \ldots & \mathbf{R}_{yy}[\beta-\alpha] \end{pmatrix} \tag{12}$$

where α and β are positive integers. In most of the cases, all the M_s streams of $\mathbf{y}(n)$ will be independent and share the same spectrum, which renders $\mathbf{R}_{yy}(z) = \rho(z)\rho(1/z^*)\mathbf{I}$ (where $\rho(z)$ has no zeroes on the unit circle). Consequently, it is reasonable to assume that $\mathbf{R}_Y^{(\alpha,\alpha)}$ is full rank for any choice of α. Under these conditions, (11) has the unique solution

$$\mathbf{A}_\star = \left(\mathbf{R}_\star - \mathbf{H}\mathbf{R}_Y^{(L_{eq}, L_a)}\right)\left(\mathbf{R}_Y^{(L_a, L_a)}\right)^{-1} \tag{13}$$

From (13) we can differentiate the three following cases:

Undermodelled case ($L_a < L_{eq}$): The algorithm has not enough degrees of freedom to fully mitigate the self-interference signal, leading to a biased estimate of the self-interference path. This is seen by dividing $\mathbf{H}$ into $\mathbf{H} = [\mathbf{H}_L \ \mathbf{H}_U]$, where the matrix $\mathbf{H}_L = [\mathbf{H}[0] \ldots \mathbf{H}[L_a]]$ and, equivalently, the matrix $\mathbf{H}_U = [\mathbf{H}[L_a+1] \ldots \mathbf{H}[L_{eq}]]$. Additionally, $\mathbf{R}_Y^{(L_{eq}, L_a)} = \left[\left(\mathbf{R}_Y^{(L_a, L_a)}\right)^T \ \left(\mathbf{D}_Y^{(L_{eq}, L_a)}\right)^T\right]^T$, with $\mathbf{D}_Y^{(L_{eq}, L_a)}$ being the last $L_h - L_a$ rows of $\mathbf{R}_Y^{(L_{eq}, L_a)}$, as can be seen from (12). Therefore, (13) has the following expression for the undermodelled case

$$\mathbf{A}_\star = -\mathbf{H}_L + \left(\mathbf{R}_\star - \mathbf{H}_U\mathbf{D}_Y^{(L_{eq}, L_a)}\right)\left(\mathbf{R}_Y^{(L_a, L_a)}\right)^{-1} \tag{14}$$

We see from (14) that the algorithm is able to estimate the first L_a coefficients of $\mathbf{H}$ up to a bias term which is the consequence of the higher order of the self-interference path. This bias can be still mitigated by selecting $\mathbf{R}_\star[k](n)$ such that $\mathbf{R}_\star = \mathbf{H}_U\mathbf{D}_Y^{(L_{eq}, L_a)}$, if the information is available.

Sufficient order case ($L_a \geq L_{eq}$): We can extend $\mathbf{H}$ by performing zero padding, and obtain an equivalent L_a-th order matrix $\mathbf{H}_A = [\mathbf{H} \ \mathbf{0}]$. After this zero padding, (13) becomes

$$\mathbf{A}_\star = \mathbf{R}_\star - \mathbf{H}_A \tag{15}$$

which clearly indicates that we must select $\mathbf{R}_\star[k](n) = \mathbf{0}$, to obtain an unbiased estimate of the self-interference channel, i.e., $\mathbf{A}_\star(z) = -\mathbf{H}_{rr}(z)\mathbf{G}_r(z)$. Consequently, after convergence $\mathbf{e}(n) = \mathbf{x}(n) + \boldsymbol{\eta}(n)$, and the self-interference signal has been completely mitigated. In conclusion, for the sufficient order case the gradient-descent algorithm in (8) with $\mathbf{R}_\star[k](n) = \mathbf{0}$ provides cancellation of the self-interference signal.

Singular case ($\mathrm{rank}\{\mathbf{R}_{yy}(z)\} < M_s$): Now $\mathbf{R}_Y^{(L_a, L_a)}$ is singular. As opposed to the previous cases, there exist several stationary points fulfilling

$$\mathbf{A}_\star \mathbf{R}_Y^{(L_a, L_a)} = \left(\mathbf{R}_\star - \mathbf{H}\mathbf{R}_Y^{(L_{eq}, L_a)}\right) \tag{16}$$

which shows that, due to the existence of local minima, the estimation of the self-interference channel may be biased. In fact, all the possible solutions to (16) conform an $M_s \times (L_a + 1) - \mathrm{rank}\{\mathbf{R}_Y^{(L_a, L_a)}\}$ dimensional subspace, meaning that $\mathbf{A}_\star + \boldsymbol{\Delta}$ is also a solution, with $\mathbf{A}_\star$ satisfying (16) and $\boldsymbol{\Delta}$ being in the nullspace of $\mathbf{R}_Y^{(L_a, L_a)}$. Since the solutions of (16) are not bounded, this can result, in the worst case, in the algorithm amplifying the self-interference signal, i.e., $\mathrm{SIR}_{pre} < \mathrm{SIR}_{post}$, which is, of course, detrimental for the relay. Obviously, the desired solution will be the one that globally minimizes P_e with respect to $\mathbf{A}$. In order to alleviate

this unwanted bias, we can use a regularization technique. By selecting $\mathbf{R}_\star[k](n) = \sum_{i=0}^{L_a} \mathbf{A}[i](n)\mathbf{\Theta}_{ik}$, (16) is transformed to

$$\mathbf{A}_\star(\mathbf{R}_Y^{(L_a,L_a)} - \mathbf{\Theta}) = -\mathbf{H}\mathbf{R}_Y^{(L_{eq},L_a)} \tag{17}$$

where $\mathbf{\Theta}$ is of size $(L_a+1)M_s \times (L_a+1)M_s$ and has elements $\mathbf{\Theta}_{ik}$. If we select $\mathbf{\Theta}$ such that $\mathbf{R}_Y^{(L_a,L_a)} - \mathbf{\Theta}$ is full rank, then (17) will have the unique solution

$$\mathbf{A}_\star = -\mathbf{H}\mathbf{R}_Y^{(L_{eq},L_a)}(\mathbf{R}_Y^{(L_a,L_a)} - \mathbf{\Theta})^{-1} \tag{18}$$

and the equivalent self-interference channel after mitigation has the expression

$$\begin{aligned}\mathbf{A}_\star + \mathbf{H} = [&\mathbf{H}_L(\mathbf{I} - \mathbf{R}_Y^{(L_a,L_a)}(\mathbf{R}_Y^{(L_a,L_a)} - \mathbf{\Theta})^{-1}) \\ &+ \mathbf{H}_U\mathbf{D}_Y^{(L_{eq},L_a)}(\mathbf{R}_Y^{(L_a,L_a)} - \mathbf{\Theta})^{-1})\ \mathbf{H}_U]\end{aligned} \tag{19}$$

If we consider the case of sufficient order, when $\mathbf{H}_U = \mathbf{0}$, we can select $\mathbf{\Theta}$ during the design stage to minimize the residual self-interference power after mitigation, i.e., using the notation $\mathbf{H}_{res} = (\mathbf{I} - \mathbf{R}_Y^{(L_a,L_a)}(\mathbf{R}_Y^{(L_a,L_a)} - \mathbf{\Theta})^{-1})$, $\mathbf{\Theta}$ is the solution to the problem

$$\begin{aligned}\underset{\mathbf{\Theta}}{\text{minimize}} \quad & \mathrm{tr}\{\mathbf{H}_{res}\mathbf{R}_Y^{(L_a,L_a)}\mathbf{H}_{res}^H\} \\ \text{subject to} \quad & \mathrm{rank}\{\mathbf{R}_Y^{(L_a,L_a)} - \mathbf{\Theta}\} = M_s(L_a+1)\end{aligned}$$

which may be sufficient depending on the particular scenario and application. If isolation provided at this point is not enough, the scheme needs additional information about $\mathbf{\Delta}$ in order to provide further attenuation of the self-interference. This information could be provided by the decoder device of the relay, although a joint design of the decoder and the mitigation scheme will be required in that case.

B. Global Convergence

In this section, we prove that regardless of the initialization point of the algorithm, it will ultimately converge to an stationary point. Under the assumption of a sufficiently small μ_a (although not necessarily vanishing), the mean convergence properties of the algorithm are characterized by the associated ordinary differential equation [13], which is given by

$$\frac{d\mathbf{A}[k](t)}{dt} = \mathbb{E}\{\mathbf{R}_\star[k] - \mathbf{e}(n)\mathbf{y}^H(n-k)\} \tag{20}$$

for $k = 0, \ldots, L_a$. In the same way as we do in Section IV-A, (20) can be expressed as a linear system

$$\frac{d\mathbf{A}(t)}{dt} = \mathbf{R}_\star - \mathbf{A}(t)\mathbf{R}_Y^{(L_a,L_a)} - \mathbf{H}\mathbf{R}_Y^{(L_{eq},L_a)} \tag{21}$$

where $\mathbf{A}(t) = [\mathbf{A}[0](t) \ldots \mathbf{A}[L_a](t)]$ and the remaining elements are defined in Section IV-A. If we split (22) into real and imaginary parts and concatenate each part into an extended matrix, we obtain the linear system

$$\frac{d\tilde{\mathbf{A}}(t)}{dt} = \tilde{\mathbf{R}}_\star - \tilde{\mathbf{A}}(t)\tilde{\mathbf{R}}_Y^{(L_a,L_a)} - \tilde{\mathbf{H}}\tilde{\mathbf{R}}_Y^{(L_{eq},L_a)} \tag{22}$$

with $\tilde{\mathbf{R}} = [\Re\mathbf{R}_\star\ \Im\mathbf{R}_\star]$, $\tilde{\mathbf{A}}(t) = [\Re\mathbf{A}(t)\ \Im\mathbf{A}(t)]$ and $\tilde{\mathbf{H}} = [\Re\mathbf{H}\ \Im\mathbf{H}]$. On the other hand, $\tilde{\mathbf{R}}_Y^{(L_a,L_a)}$ is given by

$$\tilde{\mathbf{R}}_Y^{(\alpha,\beta)} = \begin{pmatrix} \Re\mathbf{R}_Y^{(\alpha,\beta)} & \Im\mathbf{R}_Y^{(\alpha,\beta)} \\ -\Im\mathbf{R}_Y^{(\alpha,\beta)} & \Re\mathbf{R}_Y^{(\alpha,\beta)} \end{pmatrix} \tag{23}$$

For the convergence analysis, it is convenient to express $\tilde{\mathbf{A}}(t)$ as a $2N_r \times M_s \times (L_a+1)$ vector. Vectorizing both sides of (22) results in

$$\frac{d\mathbf{a}(t)}{dt} = \mathbf{r}_\star - \mathbf{\Sigma}_{L_a}^{L_a}\mathbf{a}(t) - \mathbf{\Sigma}_{L_a}^{L_{eq}}\mathbf{h} \tag{24}$$

where $\mathbf{r}_\star = \mathrm{vec}\{\tilde{\mathbf{R}}_\star\}$, $\mathbf{a}(t) = \mathrm{vec}\{\tilde{\mathbf{A}}(t)\}$ and $\mathbf{h} = \mathrm{vec}\{\tilde{\mathbf{H}}\}$. Additionally, we define $\mathbf{\Sigma}_\beta^\alpha \triangleq \left(\tilde{\mathbf{R}}_Y^{(\alpha,\beta)}\right)^T \otimes \mathbf{I}$, with $\otimes$ denoting the Kronecker product. The spectral factorization of $\mathbf{\Sigma}_\alpha^\alpha$ is $\mathbf{\Gamma}_\alpha$, i.e., $\mathbf{\Sigma}_\alpha^\alpha = \mathbf{\Gamma}_\alpha^H\mathbf{\Gamma}_\alpha$.

In order to prove that the algorithm is globally convergent, we propose the following Lyapunov function [14]

$$V(\mathbf{a}(t)) = \frac{1}{2}\|\mathbf{a}(t) - \mathbf{a}_\star\|^2 \tag{25}$$

with time derivative

$$\frac{dV(\mathbf{a}(t))}{dt} = \Re\left\{(\mathbf{a}(t) - \mathbf{a}_\star)^H\frac{d\mathbf{a}(t)}{dt}\right\} \tag{26}$$

where $\mathbf{a}_\star = \mathrm{vec}\{\tilde{\mathbf{A}}_\star\} = \mathrm{vec}\{[\Re\mathbf{A}_\star\ \Im\mathbf{A}_\star]\}$. If we vectorize (16) we obtain that $\mathbf{\Sigma}_{L_a}^{L_a}\mathbf{a}_\star = (\mathbf{r}_\star - \mathbf{\Sigma}_{L_a}^{L_{eq}}\mathbf{h})$. Using this in (26) gives the following result

$$\frac{dV(\mathbf{a}(t))}{dt} = -\|\mathbf{\Gamma}_\alpha(\mathbf{a}(t) - \mathbf{a}_\star)\|^2 \tag{27}$$

Note that (27) is still valid for the singular case in section IV-A by defining $\mathbf{\Gamma}_k$ as the spectral factorization of the full-rank matrix $\left(\tilde{\mathbf{R}}_Y^{(\alpha,\beta)} - \tilde{\mathbf{\Theta}}\right)^T \otimes \mathbf{I}$, with

$$\tilde{\mathbf{\Theta}} = \begin{pmatrix} \Re\mathbf{\Theta} & \Im\mathbf{\Theta} \\ -\Im\mathbf{\Theta} & \Re\mathbf{\Theta} \end{pmatrix} \tag{28}$$

We can verify from (27) that $dV(\mathbf{a}(t))/dt < 0$ for any value of $\mathbf{a}(t)$, except for the equilibrium point $\mathbf{a}_\star$. Therefore, in this case, we can conclude that the stationary point $\mathbf{a}_\star$ is globally asymptotically stable, and the algorithm will converge to it regardless of the initial point, i.e., $\mathbf{a}(t) \to \mathbf{a}_\star$ as $t \to \infty$ for any initialization $\mathbf{a}(0)$.

V. Simulations and Results

In this section, we study the performance and behaviour of the algorithm under two practical scenarios. In particular, we simulate a system in which the source signal $\mathbf{s}(n)$ consists of $M_s = 2$ independent OFDM-modulated streams with $N_{sub} = 8192$ subcarriers and a cyclix prefix of length $N_{pre} = 1/4N_{sub}$ (in samples). The transmitter has $M_t = 3$ antennas, while the relay has $N_r = 4$ receive antennas and $M_r = 3$ transmit antennas. Additionally, the channel $\mathcal{S}$-$\mathcal{R}$ has order $L_{sr} = 2$, while $L_{rr} = 1$, and the precoders have orders $L_r = 1$ and $L_s = 0$, so $L_{eq} = 2$. Finally, $L_a = 2$ and therefore $\mathbf{R}_\star[k](n) = \mathbf{0}$. The relay uses an oversampling factor of 2 in both $\mathbf{e}(n)$ and $\mathbf{y}(n)$. The incoming signal $\mathbf{x}(n)$ has normalized power, i.e., $P_x = 0$ dB. Finally, $\mathbf{R}_{\eta\eta}(z) = P_\gamma\mathbf{I} + P_\sigma\mathbf{H}_{rr}(z)\mathbf{H}_{rr}^H(1/z^*)$, considering that $\mathbf{H}_{rr}(z)$ has normalized gain.

In the first case, we analyze how $\mathbf{H}_{rr}(z)$ affects the convergence time of the algorithm. The convergence time is defined as the number of samples needed, from the initialization point $\tilde{\mathbf{A}}(0) = \mathbf{0}$, to reach a point where $\|\tilde{\mathbf{A}}(\tau)+\tilde{\mathbf{H}}\|_F^2/\|\tilde{\mathbf{H}}\|_F^2 < -25$

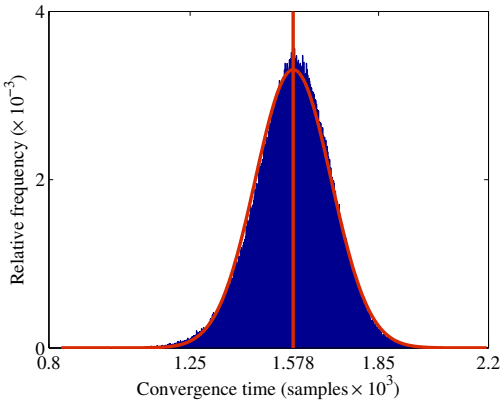

Fig. 2. Convergence time distribution (the vertical line indicates the mean convergence time).

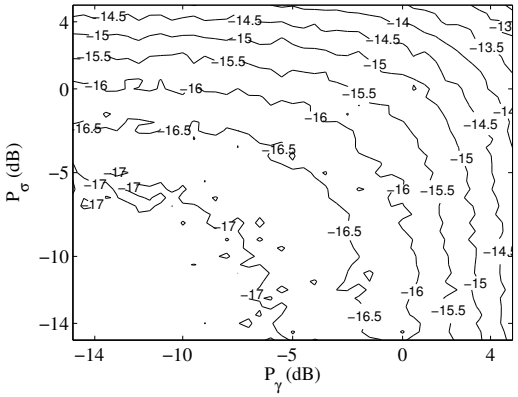

Fig. 3. Self-interference power in dB after mitigation.

dB is fulfilled. For this case, $\mathbf{H}_{sr}(z)\mathbf{G}_s(z)$ is fixed across all simulations and selected as a snapshot of a multipath Rayleigh fading channel. Additionally, $\text{SIR}_{pre} = -20$ dB and $\text{SNR} = 3.8$ dB (with $P_\sigma = -5$ dB and $P_\gamma = -10$ dB). In each simulation, $\tilde{\mathbf{H}} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ up to power normalization. Finally, $\mu_a = 0.01$ and the number of simulations is 500000.

Figure 2 shows the resulting histogram of convergence time with a superimposed Gaussian distribution with the same mean and variance. As indicated by the vertical line, the mean convergence time is approximately 1578 samples, which translates to much less than one OFDM symbol, under the current parameters. Also, note that due to the frequency dependence of $\mathbf{H}_{rr}(z)$ and the stopping criterion used, SIR_{post} is also a random variable with $\mathbb{E}\{\text{SIR}_{post}\} = 8.6$ dB, or a mean attenuation of the self-interference signal of 28.6 dB.

In our second case, we analyze the effect of the SNR on the algorithm performance. As opposed to the first case, $\tilde{\mathbf{H}}$ is fixed across all simulations and $\mathbf{H} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ up to a power normalization. Like before, $\text{SIR}_{pre} = -20$ dB and we modify the values of both P_σ and P_γ. Each simulation spans over one OFDM symbol (or 20480 samples) and the adaptation rate used is $\mu_a = 0.001$. The number of simulations is 300 while the remaining parameters are unchanged from the previous case.

Figure 3 depicts the power of signal $\mathbf{i}(n)$ and how the performance of the algorithm is affected by the noise, after the algorithm has converged. There we can see that the performance loss is independent of the noise power spectral density (note that $\boldsymbol{\gamma}(n)$ is a frequency-selective noise process), and only depends on its overall power, $P_\sigma + P_\gamma$, while the minimum attenuation from mitigation is above 30 dB.

VI. Conclusions

In this paper, we presented a method for self-interference mitigation in full-duplex decode-and-forward MIMO relays. We showed that a gradient descent approach is possible in order to mitigate the self-interference without modifying the relay design. Also, for the case of sufficient order, the algorithm has a unique stationary point which matches with the self-interference channel. We also addressed the singular case by introducing a regularization term, and we showed that the algorithm is globally convergent. Finally, according to simulations the algorithm is able to attenuate the self-interference by over 30 dB, when SNR ≈ 6 dB.

References

[1] P. Almers, F. Tufvesson, and A. F. Molisch, "Keyhole effect in MIMO wireless channels: Measurements and theory." *IEEE Trans. Wireless Comm.*, vol. 5, no. 12, pp. 3596–3604, Dec. 2006.

[2] T. Riihonen, S. Werner, and R. Wichman, "Mitigation of loopback self-interference in full-duplex MIMO relays," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 5983–5993, Dec. 2011.

[3] B. Rankov and A. Wittneben, "Spectral efficient signaling for half-duplex relay channels," in *Proc. ACSSC*, Nov. 2005, pp. 1066–1071.

[4] D. W. Bliss, P. A. Parker, and A. R. Margetts, "Simultaneous transmission and reception for improved wireless network performance," in *Proc. SSP*, Aug. 2007, pp. 478–482.

[5] T. Riihonen, A. Balakrishnan, K. Haneda, S. Wyne, S. Werner, and R. Wichman, "Optimal eigenbeamforming for suppressing self-interference in full-duplex MIMO relays," in *Proc. CISS*, Mar. 2011, pp. 1–6.

[6] P. Persson, M. Coldrey, A. Wolfgang, and P. Bohlin, "Design and evaluation of a 2 x 2 MIMO repeater," in *Proc. EuCAP*, Mar. 2009, pp. 1509–1512.

[7] K. Haneda, E. Kahra, S. Wyne, C. Icheln, and P. Vainikainen, "Measurement of loop-back interference channels for outdoor-to-indoor full-duplex radio relays," in *Proc. EuCAP*, Apr. 2010, pp. 1–5.

[8] P. Larsson and M. Prytz, "MIMO on-frequency repeater with self-interference cancellation and mitigation," in *Proc. VTS*, Apr. 2009, pp. 1–5.

[9] R. Lopez-Valcarce, E. Antonio-Rodriguez, C. Mosquera, and F. Perez-Gonzalez, "An adaptive feedback canceller for full-duplex relays based on spectrum shaping," *IEEE J. Sel. Areas Commun*, vol. 30, no. 8, pp. 1566–1577, Sep. 2012.

[10] E. Antonio-Rodríguez and R. López-Valcarce, "Adaptive self-interference suppression for full-duplex relays with multiple receive antennas," in *Proc. SPAWC*, Jun. 2012, pp. 454–458.

[11] T. Riihonen, S. Werner, and R. Wichman, "Transmit power optimization for multiantenna decode-and-forward relays with loopback self-interference from full-duplex operation," in *Proc. ACSSC*, Nov. 2011, pp. 1408–1412.

[12] J. Laneman, D. Tse, and G. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[13] H. Fan, "Application of Benveniste's convergence results in the study of adaptive IIR filtering algorithms," *IEEE Trans. Inf. Theory*, vol. 34, no. 4, pp. 692–709, Jul. 1988.

[14] P. Regalia, *Adaptive IIR Filtering in Signal Processing and Control*. Marcel Dekker, 1995.