MMSE Filter Design for Full-duplex Filter-and-forward MIMO Relays under Limited Dynamic Range

Emilio Antonio-Rodríguez*, Stefan Werner†*, Roberto López-Valcarce‡, and Risto Wichman*
*Department of Signal Processing and Acoustics, Aalto University, Helsinki, Finland
†Department of Electronic Systems, Norwegian University of Science and Technology, Trondheim, Norway
‡Department of Signal Theory and Communications, University of Vigo, Vigo, Spain

Abstract

We study the problem of optimizing the end-to-end performance of a full-duplex filter-and-forward MIMO relay link, consisting of a source, a relay, and a destination node, by employing linear filtering at each node. The system model accounts for multipath propagation and self-interference at the relay, as well as transmitter impairments and limited dynamic range at every node. The design accommodates signals with arbitrary spectra and includes the direct link between the source and destination nodes. Under the minimum mean square error criterion, the resulting non-convex problem is approximated by a sequence of convex problems and solved by means of an alternating minimization method. Linear constraints allocate some of the degrees of freedom in the relay to guarantee a sufficiently small residual self-interference. Simulations quantify the impact of degrees of freedom, the dynamic range, and the balance between direct and relay paths on the link performance.

I. Introduction

Relays expand the source-destination architecture by extending coverage area and improving end-to-end performance [1], [2]. Among the different relay protocols, filter-and-forward (FF) relaying constitutes an attractive alternative, in terms of complexity and performance, to other relaying techniques such as amplify-and-forward (AF), decode-and-forward (DF) or compress-and-forward (CF) [3]–[7]. In contrast to AF relaying, where the signal is forwarded after a spatial transformation, FF relaying forwards the signal after passing it through a linear filter, usually a linear finite impulse response (FIR) filter [8]–[12]. The spectrum shaping capabilities of FF relaying bring about a performance edge over its AF counterpart, as reported in [13]. Additionally, FF relaying allows for a scalable design in comparison to the implementation of DF and CF relaying (normally comprising, among others, timing recovery, frame alignment and/or signal regeneration operations), because the number of parameters grows linearly with the number of antennas and the number of filter taps. By avoiding the decoding and re-encoding of the signal, AF and FF relays are able to comply with the strict latency requirements of 5G systems.

In combination with a full-duplex (FD) protocol, any efficient relay design must deal with the presence of self-interference (SI) distortion which is a consequence of simultaneous transmission and reception in the same frequency [14]–[20]. If not properly mitigated, SI severely impacts the performance, as it has been reported in [3], [21]. To cope with SI distortion, different mitigation methods, both in the analog and digital domains, have been developed, from which we highlight those based on interference suppression (exploitation of spatial diversity), e.g., [14], [22]–[25], and those based on interference cancellation (generation of an SI replica), e.g., [26]–[30]. Due to insufficient mitigation and high power transmission of the relay, residual SI may still be strong enough to limit performance [14], and, therefore, the relay design must account for the presence of residual SI.

Any relay implementing a linear filtering based FF protocol, under the presence of residual SI, is subjected to an impulse response of infinite duration, in which the relay operation is contained inside the feedback loop caused by the SI [29]. As a consequence, data signal and SI are correlated, which turns any optimal design strategy into an intractable problem and possibly leading to an unstable system. This is in contrast to DF and CF relays, where the processing delay, defined as the relay input-to-output time delay, is long enough to decorrelate the data signal and the SI [30].

Decorrelation between data and SI can be achieved by introducing additional processing delay, i.e., the relay waits a sufficiently long number of samples before transmission. This decorrelation property applies in relays operating in frequency domain, where the processing delay is at least one symbol or several time samples [14]. Under those conditions, the relay design results in a tractable problem. Processing delay is an important parameter, and in the case of an orthogonal frequency-division multiplexing (OFDM) system it should not exceed the cyclic prefix duration to avoid inter-symbol interference due to the direct link. We propose a time-domain design in which no additional delay is introduced in the processing path. Still, the method ensures decorrelation between data signal and SI by transmitting in the nullspace of the SI channel.

SI distortion is not the only performance-limiting factor in an FD architecture. Another important cause of performance degradation is the limited dynamic range at reception and transmission [31]–[33]. Limited dynamic range is due to imperfections in analog/digital conversion and nonlinear effects in the modulation/demodulation process [31], and translates into additional

distortion in the system. Our design method incorporates limited dynamic range at every node of the link as additional parameters of the system model.

A. Contributions of the paper

We consider a MIMO FD FF relay link with limited dynamic range and residual SI distortion after antenna isolation/analog cancellation at the relay [17]. We introduce linear filters at each node, which are designed by minimizing the mean square error (MSE) at the destination node. The resulting optimization problem is non-convex and the IIR nature of the relay makes the problem intractable. We transform the original problem into a tractable one by imposing interference-free constraints on the relay. This modified problem is solved by employing an alternating optimization technique. In particular, node filters are obtained by solving a sequence of decoupled convex optimization problems.

Existing literature in FF design [13], [34]–[36] only considers half-duplex protocol and excludes direct link between source and destination [34], whereas we model a FD system with direct link between source and destination. Although [37] contemplates the presence of a direct link between source and destination, it does not account for a FD relay with SI or limited dynamic range at each node of the link. Our design incorporates additional noise sources that model limited dynamic range at both transmission and reception.

In [38], the authors study the problem of an FF FD relay network. In contrast to our approach, the system model considers filters in the frequency domain, where there is no correlation between data signal and SI, and dynamic range is infinite. Furthermore, frequency-domain processing introduces an unavoidable of at least one OFDM symbol. While [39], [40] consider source-destination link, only the case of single antenna with flat-frequency channel response and infinite dynamic range is treated.

In [41], the authors consider an FF FD relay network in the frequency domain with limited dynamic range and direct link between source and destination. A gradient-projection based approach is used to maximize the signal-to-noise-plus-interference ratio. Our design method works in the time-domain and is, therefore, modulation independent, and able to deal with multipath distortion as well as reducing the end-to-end delay because it does not require synchronization at the relay.

Finally, [31] considers the problem of limited dynamic range for a DF relay link assuming uncorrelated input and output signals. Concretely, we consider the correlation between data signal and self-interference at the FF relay, which has a major impact on the filter design, particularly in the assignment of its degrees of freedom.

B. Organization of the paper

The paper is organized as follows. Section II describes the system model of the relay link, with special attention to the limited dynamic range of the nodes and the equivalent impulse response of the link. Section III formulates the filter design problem based on the minimum MSE (MMSE) criterion and details the required approximations in order to obtain a tractable optimization problem. Section IV solves the optimization problem by means of an alternating optimization approach, in which every filter is the solution to a convex optimization problem. Section V illustrates the performance of the design algorithm for an OFDM relay system. Finally Section VI draws the conclusions.

C. Notation

Let $\{\mathbf{A}[k]\}_{k=0}^{L_A}$, with $\mathbf{A}[k] \in \mathbb{C}^{M\times N}$, denote the impulse response of a complex-valued, L_Ath-order causal FIR MIMO filter. The *row-expanded matrix* of $\mathbf{A}[n]$, denoted by $\boldsymbol{\mathcal{A}}$, is defined as $\boldsymbol{\mathcal{A}} = [\mathbf{A}[0] \dots \mathbf{A}[L_A]] \in \mathbb{C}^{M\times N(L_A+1)}$. We define the squared Frobenius filter norm as $\|\mathbf{A}\|^2 = \mathrm{tr}\{\boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{A}}^H\} = \sum_{k=0}^{L_A} \mathrm{tr}\{\mathbf{A}[k]\mathbf{A}^H[k]\}$.

The convolution between filters $\mathbf{A}[n]$ and $\mathbf{B}[n]$, of respective sizes $M \times N$ and $N \times P$ and orders L_A and L_B, yields filter $\mathbf{C}[n] = \mathbf{A}[n] \star \mathbf{B}[n] = \sum_{k=0}^{L_C} \mathbf{A}[k]\mathbf{B}[n-k]$, of size $M \times P$ and order $L_C = L_A + L_B$. In its row-expanded matrix, $\mathbf{C}[n]$ is expressed as

$$\boldsymbol{\mathcal{C}} = \boldsymbol{\mathcal{A}}\mathcal{R}(\boldsymbol{\mathcal{B}}) \tag{1}$$

where $\mathcal{R}(\boldsymbol{\mathcal{B}}) \in \mathbb{C}^{N(L_A+1)\times P(L_C+1)}$ is the *row-diagonal-expanded matrix* of order L_A of $\mathbf{B}[n]$ [42, Sec.7.5.],

$$\mathcal{R}(\boldsymbol{\mathcal{B}}) = \underbrace{\begin{Bmatrix} \\ \\ \\ \\ \end{Bmatrix}}_{L_\mathrm{A}\text{ times}} \begin{bmatrix} \overbrace{\mathbf{B}[0] & \dots & \mathbf{B}[L_B]}^{\boldsymbol{\mathcal{B}}} & \overbrace{\mathbf{0}_{N\times P} & \dots & \mathbf{0}_{N\times P}}^{L_\mathrm{A}\text{ times}} \\ \mathbf{0}_{N\times P} & \mathbf{B}[0] & \dots & \mathbf{B}[L_B] & \dots & \mathbf{0}_{N\times P} \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \mathbf{0}_{N\times P} & \dots & \mathbf{0}_{N\times P} & \mathbf{B}[0] & \dots & \mathbf{B}[L_B] \end{bmatrix} \tag{2}$$

where $\mathbf{0}_{N\times P}$ denotes a null matrix of size $N \times P$. Note that $\mathcal{R}(\boldsymbol{\mathcal{C}}) = \mathcal{R}(\boldsymbol{\mathcal{A}})\mathcal{R}(\boldsymbol{\mathcal{B}})$ for a proper size of the expanded matrices. The *column-diagonal-expanded matrix* of order L_A of $\mathbf{B}[n]$, denoted by $\mathcal{C}(\boldsymbol{\mathcal{B}}) \in \mathbb{C}^{N(L_C+1)\times P(L_A+1)}$, is defined as [42,

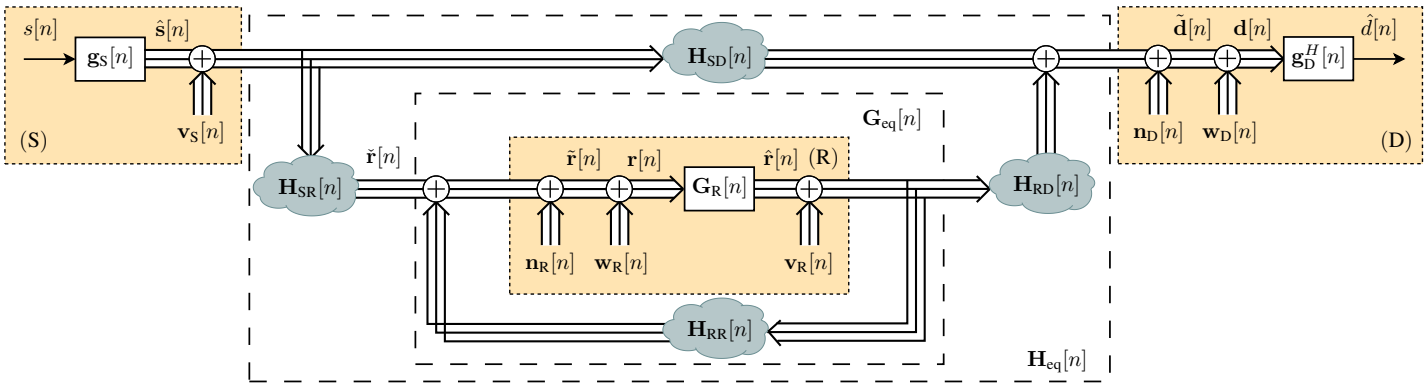

Fig. 1. System model of an FF relay link incorporating the proposed design.

Sec.7.5.]

$$\mathcal{C}(\boldsymbol{\mathcal{B}}) = \underbrace{\begin{Bmatrix} \\ \\ \\ \\ \\ \\ \end{Bmatrix}}_{L_{\mathrm{A}}\text{ times}} \overbrace{\begin{bmatrix} \mathbf{B}[0] & \mathbf{0}_{N\times P} & \dots & \mathbf{0}_{N\times P} \\ \vdots & \mathbf{B}[0] & \ddots & \vdots \\ \mathbf{B}[L_B] & \vdots & \ddots & \mathbf{0}_{N\times P} \\ \mathbf{0}_{N\times P} & \mathbf{B}[L_B] & \ddots & \mathbf{B}[0] \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{N\times P} & \mathbf{0}_{N\times P} & \dots & \mathbf{B}[L_B] \end{bmatrix}}^{L_{\mathrm{A}}\text{ times}} \tag{3}$$

The convolution between signal vector $\mathbf{x}[n]$ and filter $\mathbf{A}[n]$ yields signal $\mathbf{y}[n] = \mathbf{A}[n] \star \mathbf{x}[n] = \sum_{k=0}^{L_A} \mathbf{A}[k]\mathbf{x}[n-k]$, which can be equivalently expressed as

$$\boldsymbol{y} = \mathcal{R}(\boldsymbol{\mathcal{A}})\boldsymbol{x} \tag{4}$$

with $\boldsymbol{y} = [\mathbf{y}^T[n] \dots \mathbf{y}^T[n-L]]^T \in \mathbb{C}^{M(L+1)}$ and $\boldsymbol{x} = [\mathbf{x}^T[n] \dots \mathbf{x}^T[n-L-L_A]]^T \in \mathbb{C}^{N(L+L_A+1)}$ being the *column-expansion* (of certain order) of $\mathbf{y}[n]$ and $\mathbf{x}[n]$, respectively. The Lth-order autocorrelation matrix of $\mathbf{y}[n]$ is denoted as $\mathbf{R}_{\mathbf{y}} = \mathbb{E}\{\boldsymbol{y}\boldsymbol{y}^H\}$. Operator $\langle \mathbf{A}[n] \rangle^{-1}$ denotes the impulse response of the inverse system of the causal and stable filter $\mathbf{A}[n]$ with coefficients $\{\mathbf{A}[k]\}_{k=0}^{L_A} \in \mathbb{C}^{M\times M}$, i.e., $\langle \mathbf{A}[n] \rangle^{-1} \star \mathbf{A}[n] = \delta[n]\mathbf{I}_M$.

Let $\mathbf{x} \sim \mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Gamma})$ denote a vector $\mathbf{x}$ that follows a circularly-symmetric complex normal distribution with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Gamma}$. Matrix $\mathbf{I}_M$ is the identity matrix of size $M \times M$, whereas $\mathbf{I}_M[L] \in \mathbb{C}^{(L+1)M\times M}$ denotes the block matrix given by

$$\mathbf{I}_M[L] = \mathbf{1}_{L+1}^T \otimes \mathbf{I}_M = [\underbrace{\mathbf{I}_M \ \mathbf{I}_M \ \dots \ \mathbf{I}_M}_{L+1\text{ times}}] \tag{5}$$

where $\mathbf{1}_L$ is the all-ones column vector of size L and $\otimes$ denotes the Kronecker product. Operator $\mathrm{diag}\{\mathbf{A}\}$ creates a diagonal matrix with the principal diagonal of $\mathbf{A}$.

II. System Model

The single-stream full-duplex MIMO relay link under consideration is depicted in Fig. 1 and consists of a source node (S) equipped with M_{S} antennas that transmits the signal $\hat{\mathbf{s}}[n] \in \mathbb{C}^{M_{\mathrm{S}}}$, a destination node (D) equipped with N_{D} antennas that receives the signal $\mathbf{d}[n] \in \mathbb{C}^{N_{\mathrm{D}}}$, and a relay node (R) equipped with N_{R} receive and M_{R} transmit antennas that receives the signal $\mathbf{r}[n] \in \mathbb{C}^{N_{\mathrm{R}}}$ while it simultaneously transmits the signal $\hat{\mathbf{r}}[n] \in \mathbb{C}^{M_{\mathrm{R}}}$. Assuming a block fading model, the L_{ij}th-order channel between node i and node j, where $i \in \{\mathrm{S}, \mathrm{R}\}$ and $j \in \{\mathrm{R}, \mathrm{D}\}$, is denoted by $\mathbf{H}_{ij}[n] \in \mathbb{C}^{N_j \times M_i}$. Note that the strictly causal $\mathbf{H}_{\mathrm{RR}}[n]$ with coefficients $\{\mathbf{H}_{\mathrm{RR}}[k]\}_{k=1}^{L_{\mathrm{RR}}}$ is the residual SI channel after antenna isolation/analog cancellation [6], [17], [43]. Mitigation in the analog domain eliminates a significant part of the self-interference. The mitigation level is a function of the frequency and normally insufficient for a wideband signal. Impulse responses $\mathbf{H}_{ij}[n]$ account for the analog front-end distortion and the propagation effects between nodes, as well as their respective delays. Noise sources are represented by vectors $\mathbf{v}_i[n] \in \mathbb{C}^{M_i}$, $\mathbf{w}_j[n] \in \mathbb{C}^{N_j}$, and $\mathbf{n}_j[n] \in \mathbb{C}^{N_j}$, while aggregated noise sources containing all noise terms at a node are represented by vectors $\mathbf{z}_j[n] \in \mathbb{C}^{N_j}$, where $i \in \{\mathrm{S}, \mathrm{R}\}$ and $j \in \{\mathrm{R}, \mathrm{D}\}$. As seen from Fig. 1, each node filters the locally observed signal. The design criterion for $\mathbf{g}_{\mathrm{S}}[n]$, $\mathbf{g}_{\mathrm{D}}[n]$, and $\mathbf{G}_{\mathrm{R}}[n]$ is to minimize the MMSE given by $\mathbb{E}\{|\hat{d}[n] - s[n-\tau]|^2\}$,

with $\tau \geq 0$ being a design parameter. The joint design of these filters under the MSE criterion is the topic of subsequent sections of this paper.

Starting with the signal in S, from Fig. 1, the L_{S}th-order filter $\mathbf{g}_{\mathrm{S}}[n] \in \mathbb{C}^{M_{\mathrm{S}}}$ precodes the data signal $s[n]$

$$\hat{\mathbf{s}}[n] = \mathbf{g}_{\mathrm{S}}[n] \star s[n] \tag{6}$$

The relay implements an FF protocol, modeled as the L_{R}th-order filter $\mathbf{G}_{\mathrm{R}}[n] \in \mathbb{C}^{M_{\mathrm{R}} \times N_{\mathrm{R}}}$ and transmits the signal

$$\hat{\mathbf{r}}[n] = \mathbf{G}_{\mathrm{R}}[n] \star \mathbf{r}[n] \tag{7}$$

where the input signal to the relay filter $\mathbf{r}[n]$ can be decomposed into

$$\begin{aligned} \mathbf{r}[n] =& \check{\mathbf{r}}[n] + \underbrace{\mathbf{H}_{\mathrm{RR}}[n] \star \hat{\mathbf{r}}[n]}_{\text{self-interference } \mathbf{i}[n]} \\ &+ \underbrace{\mathbf{n}_{\mathrm{R}}[n] + \mathbf{w}_{\mathrm{R}}[n] + \mathbf{H}_{\mathrm{RR}}[n] \star \mathbf{v}_{\mathrm{R}}[n]}_{\text{noise terms}} \end{aligned} \tag{8}$$

with $\check{\mathbf{r}}[n] \in \mathbb{C}^{N_{\mathrm{R}}}$ being the incoming source signal, i.e.,

$$\check{\mathbf{r}}[n] = \mathbf{H}_{\mathrm{SR}}[n] \star (\hat{\mathbf{s}}[n] + \mathbf{v}_{\mathrm{S}}[n]) \tag{9}$$

and $\mathbf{i}[n]$ denoting the self-interference. The presence of the self-interference path introduces a feedback loop, due to which the relay becomes an infinite impulse response (IIR) linear MIMO filter. The impulse response of the relay $\mathbf{G}_{\mathrm{eq}}[n] \in \mathbb{C}^{M_{\mathrm{R}} \times N_{\mathrm{R}}}$ is given by

$$\begin{aligned} \mathbf{G}_{\mathrm{eq}}[n] =& \langle \mathbf{I}_{M_{\mathrm{R}}} \delta[n] - \mathbf{G}_{\mathrm{R}}[n] \star \mathbf{H}_{\mathrm{RR}}[n] \rangle^{-1} \star \mathbf{G}_{\mathrm{R}}[n] \\ =& \mathbf{G}_{\mathrm{R}}[n] \star \langle \mathbf{I}_{N_{\mathrm{R}}} \delta[n] - \mathbf{H}_{\mathrm{RR}}[n] \star \mathbf{G}_{\mathrm{R}}[n] \rangle^{-1} \end{aligned} \tag{10}$$

The L_{D}th-order filter $\mathbf{g}_{\mathrm{D}}[n] \in \mathbb{C}^{N_{\mathrm{D}}}$ linearly combines the received signal $\mathbf{d}[n]$ at D, and produces the output

$$\hat{d}[n] = \mathbf{g}_{\mathrm{D}}^{H}[n] \star \mathbf{d}[n] \tag{11}$$

from which source signal $s[n]$ is recovered. We may now rewrite (11), in terms of the various channels and noise sources, as follows

$$\hat{d}[n] = \mathbf{g}_{\mathrm{D}}^{H}[n] \star (\mathbf{H}_{\mathrm{eq}}[n] \star \mathbf{g}_{\mathrm{S}}[n] \star s[n] + \mathbf{z}_{\mathrm{D}}[n]) \tag{12}$$

with

$$\mathbf{H}_{\mathrm{eq}}[n] = \mathbf{H}_{\mathrm{SD}}[n] + \mathbf{H}_{\mathrm{RD}}[n] \star \mathbf{G}_{\mathrm{eq}}[n] \star \mathbf{H}_{\mathrm{SR}}[n] \tag{13}$$

$$\mathbf{z}_{\mathrm{D}}[n] = \mathbf{H}_{\mathrm{SD}}[n] \star \mathbf{v}_{\mathrm{S}}[n] + \mathbf{w}_{\mathrm{D}}[n] + \mathbf{n}_{\mathrm{D}}[n] + \mathbf{H}_{\mathrm{RD}}[n] \star \mathbf{z}_{\mathrm{R}}[n] \tag{14}$$

The L_{eq}th-order filter $\mathbf{H}_{\mathrm{eq}}[n] \in \mathbb{C}^{N_{\mathrm{D}} \times M_{\mathrm{S}}}$ represents the overall channel from S to D including direct (S-D) and relay (S-R-D) paths. From (12)-(14) we see that synchronization between direct-link and source-relay-destination signals is not needed as long as the delay spread of $\mathbf{H}_{\mathrm{eq}}[n]$ is smaller than the cyclic prefix length.

From (10), $L_{\mathrm{eq}} \to \infty$, as the equivalent channel response has an infinite duration in principle. Vector $\mathbf{z}_{\mathrm{D}}[n] \in \mathbb{C}^{N_{\mathrm{D}}}$ in (12) comprises all noise terms at D and $\mathbf{z}_{\mathrm{R}}[n] \in \mathbb{C}^{M_{\mathrm{R}}}$ denotes the aggregated noise sources at the relay output:

$$\begin{aligned} \mathbf{z}_{\mathrm{R}}[n] =& \mathbf{G}_{\mathrm{eq}}[n] \star (\mathbf{H}_{\mathrm{SR}}[n] \star \mathbf{v}_{\mathrm{S}}[n] + \mathbf{n}_{\mathrm{R}}[n] + \mathbf{w}_{\mathrm{R}}[n]) \\ &+ \langle \mathbf{I}_{M_{\mathrm{R}}} \delta[n] - \mathbf{G}_{\mathrm{R}}[n] \star \mathbf{H}_{\mathrm{RR}}[n] \rangle^{-1} \star \mathbf{v}_{\mathrm{R}}[n] \end{aligned} \tag{15}$$

where vectors $\mathbf{n}_{\mathrm{R}}[n] \in \mathbb{C}^{N_{\mathrm{R}}}$ and $\mathbf{n}_{\mathrm{D}}[n] \in \mathbb{C}^{N_{\mathrm{D}}}$ are the thermal noise sources in R and D, respectively. Their respective distributions are

$$\mathbf{n}_{\mathrm{R}}[n] \sim \mathcal{CN}\left(\mathbf{0}, \sigma_{\mathrm{R}}^2 \mathbf{I}\right) \tag{16}$$

$$\mathbf{n}_{\mathrm{D}}[n] \sim \mathcal{CN}\left(\mathbf{0}, \sigma_{\mathrm{D}}^2 \mathbf{I}\right) \tag{17}$$

Vectors $\mathbf{v}_{\mathrm{S}}[n] \in \mathbb{C}^{M_{\mathrm{S}}}$ and $\mathbf{v}_{\mathrm{R}}[n] \in \mathbb{C}^{M_{\mathrm{R}}}$ model imperfections at the transmit sides of S and R, and are assumed normally distributed and statistically independent of the transmit signals $\hat{\mathbf{s}}[n]$ and $\hat{\mathbf{r}}[n]$. Following [31], they are modeled as

$$\mathbf{v}_{\mathrm{S}}[n] \sim \mathcal{CN}\left(\mathbf{0}, \delta_{\mathrm{S}} \operatorname{diag}\left\{\mathbb{E}\{\hat{\mathbf{s}}[n]\hat{\mathbf{s}}^{H}[n]\}\right\}\right) \tag{18}$$

$$\mathbf{v}_{\mathrm{R}}[n] \sim \mathcal{CN}\left(\mathbf{0}, \delta_{\mathrm{R}} \operatorname{diag}\left\{\mathbb{E}\{\hat{\mathbf{r}}[n]\hat{\mathbf{r}}^{H}[n]\}\right\}\right) \tag{19}$$

with $0 \leq \{\delta_{\mathrm{S}}, \delta_{\mathrm{R}}\} \ll 1$. Vectors $\mathbf{w}_{\mathrm{R}}[n] \in \mathbb{C}^{N_{\mathrm{R}}}$ and $\mathbf{w}_{\mathrm{D}}[n] \in \mathbb{C}^{N_{\mathrm{D}}}$ model limited dynamic range distortion at R and D, and are

assumed normally distributed and statistically independent of received signals at R and D, with

$$\mathbf{w}_{\mathrm{R}}[n] \sim \mathcal{CN}\left(\mathbf{0}, \epsilon_{\mathrm{R}} \operatorname{diag}\left\{\mathbb{E}\{\tilde{\mathbf{r}}[n]\tilde{\mathbf{r}}^H[n]\}\right\}\right) \tag{20}$$

$$\mathbf{w}_{\mathrm{D}}[n] \sim \mathcal{CN}\left(\mathbf{0}, \epsilon_{\mathrm{D}} \operatorname{diag}\left\{\mathbb{E}\{\tilde{\mathbf{d}}[n]\tilde{\mathbf{d}}^H[n]\}\right\}\right) \tag{21}$$

where $\tilde{\mathbf{r}}[n]$ and $\tilde{\mathbf{d}}[n]$ are the received signal prior to digital conversion at R and D, and $0 \leq \{\epsilon_{\mathrm{R}}, \epsilon_{\mathrm{D}}\} \ll 1$. The independent Gaussian distortion models (18)-(19) and (20)-(21) accurately captures the combined effect of DAC and ADC nonlinearities and practical hardware impairments, see [31] and the references therein.

III. MMSE Design

The end-to-end performance depends primarily on the ability of the system to reconstruct signal $s[n]$ at destination. Therefore, a reasonable approach for designing filters $\mathbf{g}_{\mathrm{S}}[n]$, $\mathbf{G}_{\mathrm{R}}[n]$ and $\mathbf{g}_{\mathrm{D}}[n]$ is to minimize the MSE at node D, i.e.,

$$\begin{aligned} \min_{\{\mathbf{g}_{\mathrm{S}}[n], \mathbf{G}_{\mathrm{R}}[n], \mathbf{g}_{\mathrm{D}}[n]\}} \quad & \mathbb{E}\{|\hat{d}[n] - s[n-\tau]|^2\} \\ \text{subject to} \quad & \mathbb{E}\{\|\hat{\mathbf{s}}[n]\|^2\} \leq P_{\mathrm{S}} \\ & \mathbb{E}\{\|\hat{\mathbf{r}}[n]\|^2\} \leq P_{\mathrm{R}} \end{aligned} \tag{22}$$

where the end-to-end delay $\tau \geq 0$ is a design parameter, and constants $P_{\mathrm{S}} > 0$ and $P_{\mathrm{R}} > 0$ denote the maximum transmit power at S and R, respectively. In view of (10), the recursive nature of the SI-affected relay makes problem (22) intractable. To overcome this, we introduce additional linear constraints. Since SI is the dominant source of distortion in a FD relay link, see, e.g., [44], we modify problem (22) to incorporate an explicit SI suppression constraint, i.e.,

$$\begin{aligned} \min_{\{\mathbf{g}_{\mathrm{S}}[n], \mathbf{G}_{\mathrm{R}}[n], \mathbf{g}_{\mathrm{D}}[n]\}} \quad & \mathbb{E}\{|\hat{d}[n] - s[n-\tau]|^2\} \\ \text{subject to} \quad & \mathbb{E}\{\|\hat{\mathbf{s}}[n]\|^2\} \leq P_{\mathrm{S}} \\ & \mathbb{E}\{\|\hat{\mathbf{r}}[n]\|^2\} \leq P_{\mathrm{R}} \\ & \mathbf{H}_{\mathrm{RR}}[n] \star \mathbf{G}_{\mathrm{R}}[n] = \mathbf{0} \end{aligned} \tag{23}$$

The immediate effect of introducing the linear constraints $\mathbf{H}_{\mathrm{RR}}[n] \star \mathbf{G}_{\mathrm{R}}[n] = \mathbf{0}$ in (23) is summarized as follows.

Remark 1. *From* (10) *we see that the equivalent relay impulse response has now finite order, i.e.,* $\mathbf{G}_{eq}[n] = \mathbf{G}_{R}[n]$ *and* $L_{eq} = \max\{L_{SD}, L_{SR} + L_R + L_{RD}\}$.

Remark 2. *The SI is suppressed, i.e.,* $\mathbf{i}[n] = \mathbf{0}$*, making problem* (23) *tractable.*

Remark 3. *Since the SI suppression constraints constitute a linear subspace, as will be shown in Sec. IV, the available degrees of freedom in* $\mathbf{G}_{R}[n]$ *are reduced. Therefore, problems* (23) *and* (22) *are not equivalent and their solutions may perform differently.*

An important consequence of the SI suppression constraint in (23) is exposed by the next result.

Lemma 1. *When* $\mathbf{H}_{RR}[n] \star \mathbf{G}_{R}[n] = \mathbf{0}$*, the inverse system* $\langle \mathbf{I}_{M_R}\delta[n] - \mathbf{G}_{R}[n] \star \mathbf{H}_{RR}[n]\rangle^{-1} = \mathbf{I}_{M_R}\delta[n] + \mathbf{G}_{R}[n] \star \mathbf{H}_{RR}[n]$ *and is of finite order* $L_R + L_{RR}$.

Proof. By taking the Fourier transform of the inverse system, denoted by $\mathcal{F}\{\langle \mathbf{I}_{M_{\mathrm{R}}}\delta[n] - \mathbf{G}_{\mathrm{R}}[n] \star \mathbf{H}_{\mathrm{RR}}[n]\rangle^{-1}\}$, we obtain

$$\begin{aligned} &\mathcal{F}\{\langle \mathbf{I}_{M_{\mathrm{R}}}\delta[n] - \mathbf{G}_{\mathrm{R}}[n] \star \mathbf{H}_{\mathrm{RR}}[n]\rangle^{-1}\} \\ &\qquad = (\mathbf{I}_{M_{\mathrm{R}}} - \mathbf{G}_{\mathrm{R}}(e^{j\omega})\mathbf{H}_{\mathrm{RR}}(e^{j\omega}))^{-1} \end{aligned} \tag{24}$$

Using the Woodbury matrix identity, the right-hand side of (24) is expanded as

$$\begin{aligned} &(\mathbf{I}_{M_{\mathrm{R}}} - \mathbf{G}_{\mathrm{R}}(e^{j\omega})\mathbf{H}_{\mathrm{RR}}(e^{j\omega}))^{-1} \\ &\qquad = \mathbf{I}_{M_{\mathrm{R}}} + \mathbf{G}_{\mathrm{R}}(e^{j\omega})(\mathbf{I}_{N_{\mathrm{R}}} + \mathbf{H}_{\mathrm{RR}}(e^{j\omega})\mathbf{G}_{\mathrm{R}}(e^{j\omega}))^{-1}\mathbf{H}_{\mathrm{RR}}(e^{j\omega}) \\ &\qquad = \mathbf{I}_{M_{\mathrm{R}}} + \mathbf{G}_{\mathrm{R}}(e^{j\omega})\mathbf{H}_{\mathrm{RR}}(e^{j\omega}) \end{aligned} \tag{25}$$

where we have used $\mathbf{H}_{\mathrm{RR}}(e^{j\omega})\mathbf{G}_{\mathrm{R}}(e^{j\omega}) = \mathbf{0}$. By taking the inverse Fourier transform of (25), the inverse system impulse response is given by

$$\mathcal{F}^{-1}\{\mathbf{I}_{M_{\mathrm{R}}} + \mathbf{G}_{\mathrm{R}}(e^{j\omega})\mathbf{H}_{\mathrm{RR}}(e^{j\omega})\} = \mathbf{I}_{M_{\mathrm{R}}}\delta[n] + \mathbf{G}_{\mathrm{R}}[n] \star \mathbf{H}_{\mathrm{RR}}[n] \tag{26}$$

whose order is $L_{\mathrm{R}} + L_{\mathrm{RR}}$.

■

$$\mathbb{E}\{\|\mathbf{G}_{\mathrm{R}}[n]\star\mathbf{H}_{\mathrm{RR}}[n]\star\mathbf{v}_{\mathrm{R}}[n]\|^2\} = \delta_{\mathrm{R}}\,\mathrm{tr}\left\{\boldsymbol{\mathcal{G}}_{\mathrm{R}}\mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{RR}})\left(\mathbf{I}_{(L_{\mathrm{R}}+L_{\mathrm{RR}}+1)}\otimes\mathrm{diag}\{\boldsymbol{\mathcal{G}}_{\mathrm{R}}\mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{SR}})\mathcal{R}(\boldsymbol{g}_{\mathrm{S}})\mathbf{R}_s\mathcal{R}^H(\boldsymbol{g}_{\mathrm{S}})\mathcal{R}^H(\boldsymbol{\mathcal{H}}_{\mathrm{SR}})\boldsymbol{\mathcal{G}}_{\mathrm{R}}^H\}\right)\mathcal{R}^H(\boldsymbol{\mathcal{H}}_{\mathrm{RR}})\boldsymbol{\mathcal{G}}_{\mathrm{R}}^H\right\} \quad (40)$$

Algorithm 1 Alternating MMSE linear filter design procedure

1: **Initialization point**: $\mathbf{g}_{\mathrm{S}}^{(0)}[n]$, $\mathbf{G}_{\mathrm{R}}^{(0)}[n]$ and $\mathbf{g}_{\mathrm{D}}^{(0)}[n]$.
2: **repeat** for each iteration $k = 1, 2, 3, \ldots$
3: Solve problem (23) with respect to $\mathbf{g}_{\mathrm{D}}^{(k)}[n]$ for fixed $\mathbf{g}_{\mathrm{S}}^{(k-1)}[n]$ and $\mathbf{G}_{\mathrm{R}}^{(k-1)}[n]$.
4: Solve problem (23) with respect to $\mathbf{G}_{\mathrm{R}}^{(k)}[n]$ for fixed $\mathbf{g}_{\mathrm{S}}^{(k-1)}[n]$ and $\mathbf{g}_{\mathrm{D}}^{(k)}[n]$.
5: Solve problem (23) with respect to $\mathbf{g}_{\mathrm{S}}^{(k)}[n]$ for fixed $\mathbf{G}_{\mathrm{R}}^{(k)}[n]$ and $\mathbf{g}_{\mathrm{D}}^{(k)}[n]$.
6: **until** the convergence criterion is met.

By virtue of Lemma 1, the noise term $\mathbf{z}_{\mathrm{R}}[n]$ at the relay output (15) can be written as

$$\begin{aligned}\mathbf{z}_{\mathrm{R}}[n] =& \mathbf{G}_{\mathrm{R}}[n]\star(\mathbf{H}_{\mathrm{SR}}[n]\star\mathbf{v}_{\mathrm{S}}[n]+\mathbf{n}_{\mathrm{R}}[n]+\mathbf{w}_{\mathrm{R}}[n])\\ &+(\mathbf{I}_{M_{\mathrm{R}}}\delta[n]+\mathbf{G}_{\mathrm{R}}[n]\star\mathbf{H}_{\mathrm{RR}}[n])\star\mathbf{v}_{\mathrm{R}}[n]\end{aligned} \quad (27)$$

We make the following assumption about the limited dynamic range

Assumption 1. *The statistics of $\mathbf{w}_R[n]$, $\mathbf{w}_D[n]$ and $\mathbf{v}_R[n]$ depend only on their respective data signal, which follows from $\delta_i\delta_j \approx 0$, $\delta_i\epsilon_j \approx 0$, $\delta_i\sigma_j^2 \approx 0$, $\epsilon_i\delta_j \approx 0$, $\epsilon_i\sigma_j^2 \approx 0$ and $\epsilon_i\epsilon_j \approx 0$. Let $\mathbf{s}_R[n] = \mathbf{H}_{SR}[n]\star\mathbf{g}_S[n]\star s[n]$ and $\mathbf{s}_D[n] = \mathbf{H}_{eq}[n]\star\mathbf{g}_S[n]\star s[n]$ denote the data signal arriving at S and D, respectively. Therefore,*

$$\mathbf{w}_R[n] \sim \mathcal{CN}\left(\mathbf{0}, \epsilon_R\,\mathrm{diag}\left\{\mathbb{E}\{\mathbf{s}_R[n]\mathbf{s}_R^H[n]\}\right\}\right) \quad (28)$$

$$\mathbf{w}_D[n] \sim \mathcal{CN}\left(\mathbf{0}, \epsilon_D\,\mathrm{diag}\left\{\mathbb{E}\{\mathbf{s}_D[n]\mathbf{s}_D^H[n]\}\right\}\right) \quad (29)$$

$$\mathbf{v}_R[n] \sim \mathcal{CN}\left(\mathbf{0}, \delta_R\,\mathrm{diag}\left\{\mathbb{E}\left\{\left(\mathbf{G}_R[n]\star\mathbf{s}_R[n]\right)\left(\mathbf{G}_R[n]\star\mathbf{s}_R[n]\right)^H\right\}\right\}\right) \quad (30)$$

Problem (23) is non-convex due to the coupling between $\mathbf{g}_{\mathrm{S}}[n]$, $\mathbf{G}_{\mathrm{R}}[n]$ and $\mathbf{g}_{\mathrm{D}}[n]$. Furthermore, the covariance of $\mathbf{v}_{\mathrm{S}}[n]$, $\mathbf{w}_{\mathrm{R}}[n]$, $\mathbf{v}_{\mathrm{R}}[n]$, $\mathbf{w}_{\mathrm{D}}[n]$ are functions of filters $\mathbf{g}_{\mathrm{S}}[n]$, $\mathbf{G}_{\mathrm{R}}[n]$ and $\mathbf{g}_{\mathrm{D}}[n]$, see (18)−(21). In the next section we make use of an alternating procedure to solve (23).

IV. Alternating Filter Design

A sub-optimal solution to the non-convex problem (23) can be iteratively obtained by means of alternating (or cyclic) minimization, in which the design of each individual filter is decoupled from the others by fixing them at each iteration. The steps of such procedure are summarized in Algorithm 1. The destination node is in charge of the computation and broadcasts the filter coefficients through the feedback channel.

Symbols $\mathbf{g}_{\mathrm{S}}^{(k)}[n]$, $\mathbf{G}_{\mathrm{R}}^{(k)}[n]$ and $\mathbf{g}_{\mathrm{D}}^{(k)}[n]$ denote the node filters at iteration k. In summary, Algorithm 1 attempts to solve (23) by iterating over a sequence of simpler optimization problems, which, as explained in the following, are convex in their respective variables and can be solved semi-analytically. Although a global optimizer of (23) is not guaranteed, the convergence of Algorithm 1 is ensured under the conditions described in [45]. Specifically, since the sequence of MSE values obtained by the algorithm is nonincreasing and bounded below by 0, it must eventually converge. In [45], it is shown that algorithms of the same form as Algorithm 1 have linear convergence, i.e., the ratio of the norm of the difference between the current iteration and a solution and the norm of the difference between the previous iteration and a solution is less than or equal to a constant. The algorithm can be sensitive to the initialization point due to existence of several local minima. Therefore, running the algorithm with different initialization points can robustify the solution in case the algorithm falls in a local minima.

From now on, unless otherwise stated, we will simply denote $\mathbf{g}_{\mathrm{S}}[n] = \mathbf{g}_{\mathrm{S}}^{(k)}[n]$, $\mathbf{G}_{\mathrm{R}}[n] = \mathbf{G}_{\mathrm{R}}^{(k)}[n]$ and $\mathbf{g}_{\mathrm{D}}[n] = \mathbf{g}_{\mathrm{D}}^{(k)}[n]$.

A. Solution of problem (23) with respect to $\mathbf{g}_D[n]$

We first solve the problem (23) with respect to $\mathbf{g}_{\mathrm{D}}[n]$, while keeping both $\mathbf{g}_{\mathrm{S}}[n]$ and $\mathbf{G}_{\mathrm{R}}[n]$ fixed, i.e.,

$$\mathbf{g}_{\mathrm{D}}^{\star}[n] = \arg\min_{\mathbf{g}_{\mathrm{D}}[n]} \quad \mathbb{E}\{|\hat{d}[n]-s[n-\tau]|^2\} \quad (31)$$

Note that, since $\mathbf{g}_{\mathrm{D}}[n]$ is at the receive side of D, its optimal expression will correspond to that of the MIMO normal equations. Let $\boldsymbol{g}_{\mathrm{D}}$ and $\boldsymbol{g}_{\mathrm{S}}$ be the column-expansions of $\mathbf{g}_{\mathrm{D}}[n]$ and $\mathbf{g}_{\mathrm{S}}[n]$, respectively. From (12), we can now express $\hat{d}[n]$ as

$$\hat{d}[n] = \boldsymbol{g}_{\mathrm{D}}^H \left(\mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{eq}})\mathcal{R}(\boldsymbol{g}_{\mathrm{S}})\boldsymbol{s} + \boldsymbol{z}_{\mathrm{D}}\right) \tag{32}$$

where vectors $\boldsymbol{s}$ and $\boldsymbol{z}_{\mathrm{D}}$ are defined as

$$\boldsymbol{s} = [s[n]\ s[n-1]\ \ldots\ s[n-L_{\mathrm{S}}-L_{\mathrm{eq}}-L_{\mathrm{D}}]]^T \tag{33}$$

$$\boldsymbol{z}_{\mathrm{D}} = \left[\mathbf{z}_{\mathrm{D}}^T[n]\ \mathbf{z}_{\mathrm{D}}^T[n-1]\ \ldots\ \mathbf{z}_{\mathrm{D}}^T[n-L_{\mathrm{D}}]\right]^T \tag{34}$$

Let $\mathbf{e}_\tau$ denote the τth canonical basis vector with $\mathbf{e}_\tau[\tau] = 1$ and 0 elsewhere. Using (32), the MSE can be expressed as

$$\begin{aligned}\mathbb{E}\{|\hat{d}[n] - s[n-\tau]|^2\} = \\ \left(\boldsymbol{g}_{\mathrm{D}}^H\mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{eq}})\mathcal{R}(\boldsymbol{g}_{\mathrm{S}}) - \mathbf{e}_\tau^H\right)\mathbf{R}_s\left(\mathcal{R}^H(\boldsymbol{g}_{\mathrm{S}})\mathcal{R}^H(\boldsymbol{\mathcal{H}}_{\mathrm{eq}})\boldsymbol{g}_{\mathrm{D}} - \mathbf{e}_\tau\right) \\ + \boldsymbol{g}_{\mathrm{D}}^H\mathbf{R}_{\boldsymbol{z}_{\mathrm{D}}}\boldsymbol{g}_{\mathrm{D}}\end{aligned} \tag{35}$$

with $\mathbf{R}_s = \mathbb{E}\{\boldsymbol{s}\boldsymbol{s}^H\}$ and $\mathbf{R}_{\boldsymbol{z}_{\mathrm{D}}} = \mathbb{E}\{\boldsymbol{z}_{\mathrm{D}}\boldsymbol{z}_{\mathrm{D}}^H\}$ (see Table I for a full expression) denoting the autocorrelation matrices of $s[n]$ and $\mathbf{z}_{\mathrm{D}}[n]$, respectively. Combining (35) and (31), the optimal filter $\boldsymbol{g}_{\mathrm{D}}^\star$ is given by

$$\boldsymbol{g}_{\mathrm{D}}^\star = \mathbf{R}_{\mathrm{D}}^{-1}\mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{eq}})\mathcal{R}(\boldsymbol{g}_{\mathrm{S}})\mathbf{R}_s\mathbf{e}_\tau \tag{36}$$

where

$$\mathbf{R}_{\mathrm{D}} = \mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{eq}})\mathcal{R}(\boldsymbol{g}_{\mathrm{S}})\mathbf{R}_s\mathcal{R}^H(\boldsymbol{g}_{\mathrm{S}})\mathcal{R}^H(\boldsymbol{\mathcal{H}}_{eq}) + \mathbf{R}_{\boldsymbol{z}_{\mathrm{D}}} \tag{37}$$

is the autocorrelation matrix of the input signal to $\mathbf{g}_{\mathrm{D}}[n]$. From (36) we see that τ, with $0 \le \tau \le (L_{\mathrm{eq}} + L_D + L_S)$, selects one of the columns of $\mathbf{R}_{\mathrm{D}}^{-1}\mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{eq}})\mathcal{R}(\boldsymbol{g}_{\mathrm{S}})\mathbf{R}_s$, and can be interpreted as a filter selector.

B. Solution of problem (23) *with respect to* $\mathbf{G}_R[n]$

The next step of Algorithm 1 is to design $\mathbf{G}_{\mathrm{R}}[n]$ as the solution to (23) assuming both $\mathbf{g}_{\mathrm{S}}[n]$ and $\mathbf{g}_{\mathrm{D}}[n]$ fixed, i.e.,

$$\begin{aligned}\mathbf{G}_{\mathrm{R}}^\star[n] = \arg\min_{\mathbf{G}_{\mathrm{R}}[n]} \quad & \mathbb{E}\{|\hat{d}[n] - s[n-\tau]|^2\} \\ \text{subject to} \quad & \mathbb{E}\{\|\hat{\mathbf{r}}[n]\|^2\} \le P_{\mathrm{R}} \\ & \mathbf{H}_{\mathrm{RR}}[n] \star \mathbf{G}_{\mathrm{R}}[n] = \mathbf{0}\end{aligned} \tag{38}$$

Even under the SI suppression constraints, solving (38) requires the use of certain approximations. Concretely, from (27), the power of $\mathbf{z}_{\mathrm{R}}[n]$ has the following expression

$$\begin{aligned}\mathbb{E}\{\|\mathbf{z}_{\mathrm{R}}[n]\|^2\} = \mathbb{E}\{\|\mathbf{G}_{\mathrm{R}}[n] \star \mathbf{n}_{\mathrm{R}}[n]\|^2\} + \mathbb{E}\{\|\mathbf{v}_{\mathrm{R}}[n]\|^2\} \\ + \mathbb{E}\{\|\mathbf{G}_{\mathrm{R}}[n] \star \mathbf{H}_{\mathrm{SR}}[n] \star \mathbf{v}_{\mathrm{S}}[n]\|^2\} \\ + \mathbb{E}\{\|\mathbf{G}_{\mathrm{R}}[n] \star \mathbf{w}_{\mathrm{R}}[n]\|^2\} \\ + \mathbb{E}\{\|\mathbf{G}_{\mathrm{R}}[n] \star \mathbf{H}_{\mathrm{RR}}[n] \star \mathbf{v}_{\mathrm{R}}[n]\|^2\}\end{aligned} \tag{39}$$

From (7) and (19) we see that the covariance of $\mathbf{v}_{\mathrm{R}}[n]$ depends on $\mathbf{G}_{\mathrm{R}}[n]$. Consequently, the last term in (39), $\mathbb{E}\{\|\mathbf{G}_{\mathrm{R}}[n] \star \mathbf{H}_{\mathrm{RR}}[n] \star \mathbf{v}_{\mathrm{R}}[n]\|^2\}$, exhibits a "*fourth-order*" relation with $\mathbf{G}_{\mathrm{R}}[n]$, as seen in (40).

To avoid such higher-order relation, we approximate $\mathbf{G}_{\mathrm{R}}^{(k)}[n] \star \mathbf{H}_{\mathrm{RR}}[n]$ as $\mathbf{G}_{\mathrm{R}}^{(k-1)}[n] \star \mathbf{H}_{\mathrm{RR}}[n]$. This transforms (39) into

$$\begin{aligned}\mathbb{E}\{\|\mathbf{z}_{\mathrm{R}}[n]\|^2\} \approx \mathbb{E}\{\|\mathbf{G}_{\mathrm{R}}^{(k)}[n] \star \mathbf{n}_{\mathrm{R}}[n]\|^2\} + \mathbb{E}\{\|\mathbf{v}_{\mathrm{R}}[n]\|^2\} \\ + \mathbb{E}\{\|\mathbf{G}_{\mathrm{R}}^{(k)}[n] \star \mathbf{H}_{\mathrm{SR}}[n] \star \mathbf{v}_{\mathrm{S}}[n]\|^2\} \\ + \mathbb{E}\{\|\mathbf{G}_{\mathrm{R}}^{(k)}[n] \star \mathbf{w}_{\mathrm{R}}[n]\|^2\} \\ + \mathbb{E}\{\|\mathbf{G}_{\mathrm{R}}^{(k-1)}[n] \star \mathbf{H}_{\mathrm{RR}}[n] \star \mathbf{v}_{\mathrm{R}}[n]\|^2\}\end{aligned} \tag{41}$$

with

$$\mathbb{E}\{\|\mathbf{v}_{\mathrm{R}}[n]\|^2\} = \delta_{\mathrm{R}}\,\mathrm{tr}\{\mathrm{diag}\{\boldsymbol{\mathcal{G}}_{\mathrm{R}}^{(k)}\mathbf{R}_{\mathbf{s}_{\mathrm{R}}}(\boldsymbol{\mathcal{G}}_{\mathrm{R}}^{(k)})^H\}\} \tag{42}$$

Matrix $\mathbf{R}_{\mathbf{s}_{\mathrm{R}}} = \mathbb{E}\{\boldsymbol{s}_{\mathrm{R}}\boldsymbol{s}_{\mathrm{R}}{}^H\}$ is the autocorrelation of $\mathbf{s}_{\mathrm{R}}[n] = \mathbf{H}_{\mathrm{SR}}[n] \star \mathbf{g}_{\mathrm{S}}[n] \star s[n]$, see (30). Vector $\boldsymbol{s}_{\mathrm{R}}$ is defined as

$$\boldsymbol{s}_{\mathrm{R}} = \left[\mathbf{s}_{\mathrm{R}}^T[n]\ \mathbf{s}_{\mathrm{R}}^T[n-1]\ \ldots\ \mathbf{s}_{\mathrm{R}}^T[n-L_{\mathrm{R}}]\right]^T \tag{43}$$

By using the approximation in (41), $\mathbb{E}\{\|\mathbf{z}_{\mathrm{R}}[n]\|^2\}$ has a second-order relation with $\mathbf{G}_{\mathrm{R}}^{(k)}[n]$, and (38) can be cast as a constrained least-squares problem. The difference between (39) and (41) is roughly proportional to δ_{R} and the approximation

Algorithm 2 Design steps for $\mathbf{G}_{\mathrm{R}}^{(k)}[n]$

1: Calculate the nullspace basis $\mathbf{N}$ of $(\mathbf{I}_{N_{\mathrm{R}}} \otimes \mathcal{C}(\boldsymbol{\mathcal{H}}_{\mathrm{RR}}))$.
2: Approximate $\mathbb{E}\{\|\mathbf{z}_{\mathrm{R}}[n]\|^2\}$ according to expression (41).
3: Calculate $\mathbf{w}^{\star}$ by solving (49).
4: Obtain $\mathbf{G}_{\mathrm{R}}^{(k)}[n]$ through expression $\mathbf{g}_{\mathrm{R}}^{\star} = \mathbf{N}\mathbf{w}^{\star}$.

in (41) will be tight whenever δ_{R} is not the dominant noise source. Typically, noise due to limited dynamic range is lower than thermal noise [31].

Although the SI suppression constraints in (38), $\mathbf{H}_{\mathrm{RR}}[n] \star \mathbf{G}_{\mathrm{R}}[n] = \mathbf{0}$, force the possible solutions to lie within a subspace of reduced dimension, (38) cannot be analytically solved by means of the normal equations due to the second-order power inequality constraints $\mathbb{E}\{\|\hat{\mathbf{r}}[n]\|^2\} \leq P_{\mathrm{R}}$. To characterize the interference-free subspace, we arrange constraints $\mathbf{H}_{\mathrm{RR}}[n] \star \mathbf{G}_{\mathrm{R}}[n] = \mathbf{0}$ in the following matrix form

$$\mathcal{C}(\boldsymbol{\mathcal{H}}_{\mathrm{RR}}) \left[\mathbf{G}_{\mathrm{R}}^T[0] \ \ldots \ \mathbf{G}_{\mathrm{R}}^T[L_{\mathrm{R}}]\right]^T = \left[\mathbf{0} \ \ldots \ \mathbf{0}\right]^T \tag{44}$$

Applying the *vec* operator, (44) can be written in a compact form

$$(\mathbf{I}_{N_{\mathrm{R}}} \otimes \mathcal{C}(\boldsymbol{\mathcal{H}}_{\mathrm{RR}}))\mathbf{g}_{\mathrm{R}} = \mathbf{0} \tag{45}$$

with $\mathbf{g}_{\mathrm{R}} = \mathrm{vec}\{\left[\mathbf{G}_{\mathrm{R}}^T[0] \ \cdots \ \mathbf{G}_{\mathrm{R}}^T[L_{\mathrm{R}}]\right]^T\} \in \mathbb{C}^{N_{\mathrm{R}}M_{\mathrm{R}}(L_{\mathrm{R}}+1)\times 1}$. For a nontrivial solution to (44) to exist, the rank of $\mathcal{C}(\boldsymbol{\mathcal{H}}_{\mathrm{RR}})$ must be strictly less than the number of its columns, which is $M_{\mathrm{R}}(L_{\mathrm{R}}+1)$. Assuming a full-row rank $\mathcal{C}(\boldsymbol{\mathcal{H}}_{\mathrm{RR}})$, this condition reads as $N_{\mathrm{R}}(L_{\mathrm{R}}+L_{\mathrm{RR}}) < M_{\mathrm{R}}(L_{\mathrm{R}}+1)$. For this to hold, the number of transmit antennas must exceed the number of receive antennas ($M_{\mathrm{R}} > N_{\mathrm{R}}$), and in addition, the order of the relay filter must satisfy $L_{\mathrm{R}} \geq N_{\mathrm{R}}(L_{\mathrm{RR}}-1)/(M_{\mathrm{R}}-N_{\mathrm{R}})$. Consequently, the subspace of feasible $\mathbf{G}_{\mathrm{R}}[n]$ is given by

$$\mathbf{g}_{\mathrm{R}} = \mathbf{N}\mathbf{w} \tag{46}$$

where the columns of $\mathbf{N}$ constitute a basis of the nullspace of $(\mathbf{I}_{N_{\mathrm{R}}} \otimes \mathcal{C}(\boldsymbol{\mathcal{H}}_{\mathrm{RR}}))$ and $\mathbf{w}$ is an arbitrary vector. Due to the constraints (45), the number of degrees of freedom (DoF) in $\mathbf{G}_{\mathrm{R}}[n]$ is equal to the dimension of this null space, which is given by

$$\begin{aligned} \mathrm{DoF} &= N_{\mathrm{R}}\left(M_{\mathrm{R}}(L_{\mathrm{R}}+1) - \mathrm{rank}\{\mathcal{C}(\boldsymbol{\mathcal{H}}_{\mathrm{RR}})\}\right) \\ &= N_{\mathrm{R}}M_{\mathrm{R}}(L_{\mathrm{R}}+1)\rho \end{aligned} \tag{47}$$

where ρ defines the fraction of DoF

$$\rho = 1 - \frac{\mathrm{rank}\{\mathcal{C}(\boldsymbol{\mathcal{H}}_{\mathrm{RR}})\}}{M_{\mathrm{R}}(L_{\mathrm{R}}+1)} \in [0,1] \tag{48}$$

When $\rho \to 1$ either $\mathbf{H}_{\mathrm{RR}}[n] \to \mathbf{0}$ or $M_{\mathrm{R}}(L_{\mathrm{R}}+1)$ is very large, whereas when $\rho \to 0$ the degrees of freedom in the relay that are used to suppress the self-interference is maximized. For a given relay configuration, ρ is minimized when $\mathcal{C}(\boldsymbol{\mathcal{H}}_{\mathrm{RR}})$ is of full-rank. Generally speaking, $\mathcal{C}(\boldsymbol{\mathcal{H}}_{\mathrm{RR}})$ is of full-rank when $\mathbf{H}_{\mathrm{RR}}[n]$ follows a Rayleigh fading channel model, setting a worst case scenario benchmark in terms of self-interference.

Combining (46) and (38) yields the equivalent optimization problem

$$\begin{aligned} \mathbf{w}^{\star} = \arg\min_{\mathbf{w}} \quad & \mathbb{E}\{|\hat{d}[n] - s[n-\tau]|^2\} \\ \text{subject to} \quad & \mathbb{E}\{\|\hat{\mathbf{r}}[n]\|^2\} \leq P_{\mathrm{R}} \end{aligned} \tag{49}$$

Problem (49) is a Least Squares minimization problem with a Quadratic Inequality constraint (LSQI) [46]. Appendix A provides the details of the computation and a semi-analytical expression of the solution. Finally, the optimal $\mathbf{G}_{\mathrm{R}}^{\star}[n]$ is recovered directly from $\mathbf{g}_{\mathrm{R}}^{\star} = \mathbf{N}\mathbf{w}^{\star}$ and reordering of $\mathbf{g}_{\mathrm{R}} = \mathrm{vec}\{\left[\mathbf{G}_{\mathrm{R}}^T[0] \ \cdots \ \mathbf{G}_{\mathrm{R}}^T[L_{\mathrm{R}}]\right]^T\}$. Table 2 summarizes the steps needed to design filter $\mathbf{G}_{\mathrm{R}}[n]$ at each iteration.

C. Solution of problem (23) *with respect to* $\mathbf{g}_S[n]$

In order to complete an iteration, and as a final step, we must solve (23) with respect to $\mathbf{g}_{\mathrm{S}}[n]$ for both $\mathbf{G}_{\mathrm{R}}[n]$ and $\mathbf{g}_{\mathrm{D}}[n]$ fixed, i.e.,

$$\begin{aligned} \mathbf{g}_{\mathrm{S}}^{\star}[n] = \arg\min_{\mathbf{g}_{\mathrm{S}}[n]} \quad & \mathbb{E}\{|\hat{d}[n] - s[n-\tau]|^2\} \\ \text{subject to} \quad & \mathbb{E}\{\|\hat{\mathbf{s}}[n]\|^2\} \leq P_{\mathrm{S}} \end{aligned} \tag{50}$$

Since $\mathbf{g}_{\mathrm{S}}[n]$ is located at node S, its optimal expression will be that of a precoding filter tailored to the combined channel $\mathbf{g}_{\mathrm{D}}[n] \star \mathbf{H}_{\mathrm{eq}}[n]$. Similarly to previous step, problem (50) can be cast as an LSQI problem and solved using the theory in [46]. We refer the reader to Appendix B for details of the computation and a semi-analytical expression of the solution.

V. Simulation results and discussion

We respectively define the signal-to-noise ratio of the SR hop, the SD hop and the RD hop as

$$\text{SNR}_{\text{SR}} = \frac{\mathbb{E}\left\{\|\mathbf{H}_{\text{SR}}[n] \star \mathbf{g}_{\text{S}}[n] \star s[n]\|^2\right\}}{\mathbb{E}\left\{\|\mathbf{n}_{\text{R}}[n]\|^2\right\}}$$

$$\text{SNR}_{\text{SD}} = \frac{\mathbb{E}\left\{\|\mathbf{H}_{\text{SD}}[n] \star \mathbf{g}_{\text{S}}[n] \star s[n]\|^2\right\}}{\mathbb{E}\left\{\|\mathbf{n}_{\text{D}}[n]\|^2\right\}}$$

$$\text{SNR}_{\text{RD}} = \frac{\mathbb{E}\left\{\|\mathbf{H}_{\text{RD}}[n] \star \mathbf{G}_{\text{R}}[n] \star \mathbf{H}_{\text{SR}}[n] \star \mathbf{g}_{\text{S}}[n] \star s[n]\|^2\right\}}{\mathbb{E}\left\{\|\mathbf{n}_{\text{D}}[n]\|^2\right\}}$$

The final signal-to-noise ratio at the destination takes into account the limited dynamic range parameters δ_{S}, δ_{R}, ϵ_{R} and ϵ_{D} and is defined as

$$\text{SNR}_{\text{D}} = \frac{\mathbb{E}\left\{\|\mathbf{g}_{\text{D}}^H[n] \star \mathbf{H}_{\text{eq}}[n] \star \mathbf{g}_{\text{S}}[n] \star s[n]\|^2\right\}}{\mathbb{E}\left\{\|\mathbf{g}_{\text{D}}^H[n] \star \mathbf{z}_{\text{D}}[n]\|^2\right\}}$$

We consider a source node with $M_{\text{S}} = 2$ antennas transmitting a 64-QAM OFDM modulated signal. The number of subcarriers is $n_c = 1024$ and the cyclic prefix length is $n_p = 32$. The sampling frequency equals the Nyquist frequency and $\mathbb{E}\{|s[n]|^2\} = 1$. Destination node has $N_{\text{D}} = 2$ receive antennas, while the relay node has $N_{\text{R}} = 2$ receive antennas and $M_{\text{R}} = 5$ transmit antennas. The maximum transmit powers at the source and the relay are set to $P_{\text{S}} = 0$ dB and $P_{\text{R}} = 0$ dB, respectively. The channels follow a Rayleigh fading model, so each tap of $\mathbf{H}_{ij}[k]$ is independently drawn from a circularly-symmetric complex normal distribution, i.e., $\text{vec}\{\mathbf{H}_{ij}[k]\} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_j M_i})$ and then a scaling is performed in order to set $\|\mathbf{H}_{ij}\|^2 = \gamma_{ij}$. Parameter γ_{ij} modifies the SNR for a fixed noise power σ_i^2. In this work, we only consider perfect channel state information, threrefore, the following results can be seen as a benchmark case. The direct link is weaker than the relay hop and subjected to stronger multipath components, and channel orders are $L_{\text{SD}} = 5$, $L_{\text{SR}} = 2$, $L_{\text{RD}} = 2$ and $L_{\text{RR}} = 2$ and normalization constants are $\gamma_{\text{SD}} = 0.1$, $\gamma_{\text{SR}} = 1$, $\gamma_{\text{RD}} = 1$ and $\gamma_{\text{RR}} = 1$, respectively. Noise levels are $\delta_{\text{S}} = \delta_{\text{R}} = -30$ dB, $\epsilon_{\text{R}} = \epsilon_{\text{D}} = -20$ dB, and $\sigma_{\text{R}}^2 = \sigma_{\text{D}}^2 = -20$ dB. Filters $\mathbf{g}_{\text{S}}[n]$, $\mathbf{G}_{\text{R}}[n]$ and $\mathbf{g}_{\text{D}}[n]$ all have orders $L_{\text{S}} = L_{\text{R}} = L_{\text{D}} = 3$. Consequently, from (48) we have that a fraction DoF $\rho = 0.5$ is used by $\mathbf{G}_{\text{R}}[n]$ for suppressing the self-interference. The end-to-end delay is set to $\tau = 3$ by using the rule of thumb recommending a value half the length of the equivalent channel, $L_{\text{eq}} = 7$. A method to optimize τ can be found in [47], [48], and a study about its influence on the performance in [49].

Algorithm 1 is assumed to have converged when the weighted filter norm between two consecutive iterations falls below a certain threshold $\varepsilon^2 \geq 0$, i.e.,

$$\|\mathbf{g}_{\text{S}}^{(k)} - \mathbf{g}_{\text{S}}^{(k-1)}\|^2 / \|\mathbf{g}_{\text{S}}^{(k)}\|^2 \leq \varepsilon^2$$
$$\|\mathbf{G}_{\text{R}}^{(k)} - \mathbf{G}_{\text{R}}^{(k-1)}\|^2 / \|\mathbf{G}_{\text{R}}^{(k)}\|^2 \leq \varepsilon^2$$
$$\|\mathbf{g}_{\text{D}}^{(k)} - \mathbf{g}_{\text{D}}^{(k-1)}\|^2 / \|\mathbf{g}_{\text{D}}^{(k)}\|^2 \leq \varepsilon^2$$

where the sensitivity threshold ε is set to $\varepsilon^2 = 0.003$. The number of iterations to reach convergence varies as function of the system parameters and channel impulse reponses. From the observed results, the algorithm may take from a few dozens up to several thousands iterations.

Initialization points are $\mathbf{g}_{\text{S}}^{(0)}[n] = \mathbf{1}\delta[n]$, $\mathbf{g}_{\text{D}}^{(0)}[n] = \mathbf{1}\delta[n]$ and $\mathbf{w}^{(0)} = \mathbf{1}$, with $\mathbf{1}$ being the all-ones vector of appropriate dimension. The results are obtained by averaging over 400 independent realizations. The parameters listed above are fixed throughout simulations unless otherwise stated.

Figure 2 depicts the MSE as a function of SNR_{SD} and SNR_{RD}, when $\gamma_{\text{SR}} = 1$, $\gamma_{\text{RD}} \in [0.1, 1]$ and $\gamma_{\text{SD}} \in [0, 1]$. We see that for low values of SNR_{RD} the presence of the direct link, i.e., when $\text{SNR}_{\text{SD}} > -\infty$, improves the end-to-end performance by approximately 5 dB. On the other hand, when SNR_{RD} is large, the improvement due to the direct link is roughly 1 dB. We can conclude that the contribution of the direct link to the overall performance is a nonlinear function of the SNR of the RD hop. When the RD hop supports a high SNR, the contribution of the direct link to the end-to-end performance is marginal.

Figure 3 shows the MSE as a function of SNR_{SR} and SNR_{RD}, when $\text{SNR}_{\text{RD}} = 0$, i.e., no direct link, $\gamma_{\text{SR}} \in [0.1, 1]$ and $\gamma_{\text{RD}} \in [0.1, 1]$. We see that the impact of SNR_{SR} on the MSE is more significant than that of SNR_{RD}, due to the noise from S being propagated to D. For low SNR_{SR}, increasing SNR_{RD} by 10 dB results in an MSE improvement of 3 dB, whereas, when SNR_{SR} is large, the same operation results in an MSE improvement of 5.5 dB. Therefore, a system where the SR hop supports higher SNR values than the RD hop leads to better performance.

Figure 4 depicts the MSE as a function of SNR_{SR} and SNR_{SD}, when $\gamma_{\text{RD}} = 1$, $\gamma_{\text{SR}} \in [0.1, 1]$ and $\gamma_{\text{SD}} \in [0.1, 1]$. Similarly to the results in Fig. 2, the MSE gain decreases as SNR_{SR} increases.

Figure 5 shows the contour lines of the MSE as a function of the distortion at the destination, ϵ_{D}, and the distortion at the relay, ϵ_{R}. From the obtained results, although the MSE depends on both ϵ_{D} and ϵ_{R}, ϵ_{D} has a stronger impact on the end-to-end

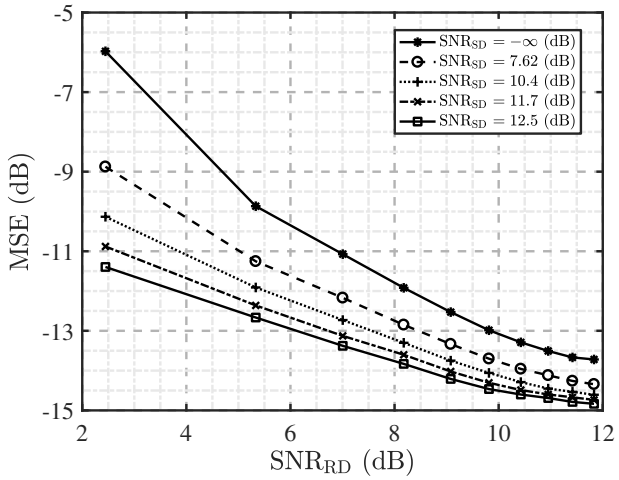

Fig. 2. MSE versus SNR_{RD} for various SNR_{SD}.

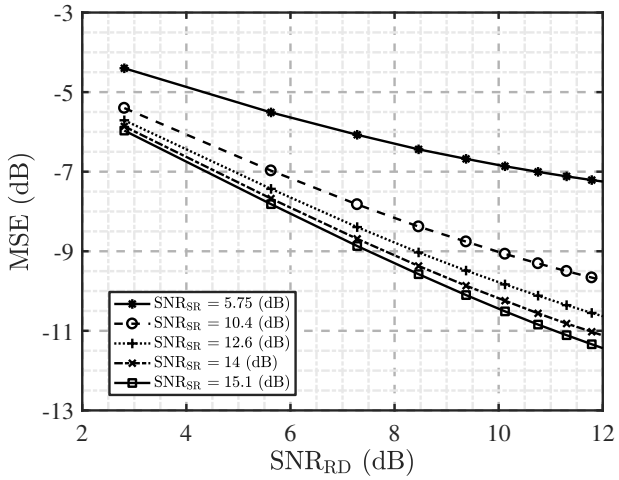

Fig. 3. MSE versus SNR_{RD} for various SNR_{SR}. No S-D link.

performance than ϵ_{R}, e.g., compare the point $(-10,-30)$ with the point $(-30,-10)$ in Fig. 5. This is a consequence of the cumulative effect of $\mathbf{w}_{\text{D}}[n]$ at D, whose power depends on other noise sources, such as $\mathbf{w}_{\text{R}}[n]$.

Figure 6 shows the contour lines of the MSE as a function of the source transmit noise, δ_{S}, and the relay transmit noise, δ_{R}. Note the similar behavior of Fig. 6 and Fig. 5. It is observed that δ_{S}, or, δ_{R} has a stronger impact on the MSE than δ_{S}. For instance, comparing the point ($\delta_{\text{R}} = -10$, $\delta_{\text{S}} = -30$) and the point ($\delta_{\text{R}} = -30$, $\delta_{\text{S}} = -10$) of Fig. 6, it is seen that there is approximately a 1 dB gap between both. This is a consequence of the relay transmit noise being filtered by $\mathbf{H}_{\text{RR}}[n]$.

Figure 7 shows the final SNR at destination SNR_{D} for different values of SNR_{RD} and SNR_{SD}. The case $\text{SNR}_{\text{RD}} \to -\infty$ indicates the absence of a relay. Therefore, Fig. 7 shows the performance gain resulting from the relay in the link. Note the saturation effect when SNR_{SD} is large, where performance gain is $3-4$ dB. On the other hand, when SNR_{SD} is low, the relay can boost performance in $10-15$ dB.

Figure 8 compares the final SNR at destination SNR_{D} for different values of SNR_{SR} and SNR_{SD}. Solid lines depict the unlimited dynamic range (u.d.r.) cases, $\delta_{\text{S}} = \epsilon_{\text{R}} = \delta_{\text{R}} = \epsilon_{\text{D}} = 0$, whereas dashed lines depict the limited dynamic range (l.d.r.) cases, $\delta_{\text{S}} = \delta_{\text{R}} = -30$ dB and $\epsilon_{\text{R}} = \epsilon_{\text{D}} = -20$ dB. The gap between l.d.r and u.d.r. cases is of approximately $1-2$ dB, and it increases alongside SNR_{SD}.

Figure 9 shows the MSE as a function of the fraction of DoF ρ in (48) for different values of M_{R} and L_{R}. Dashed lines show ρ for the same values of M_{R} and L_{R}. Note the clear relation between ρ and the MSE. Figure 9 highlights that different values of M_{R} and L_{R} may lead to the same ρ or MSE. This is relevant if additional coefficients of $\mathbf{G}_{\text{R}}[n]$ are affordable, as performance may reach that of a system with more antennas, for example, cases $(M_{\text{R}} = 7, L_{\text{R}} = 10)$ and $(M_{\text{R}} = 9, L_{\text{R}} = 1)$ results in the same MSE. However, M_{R} has a bigger impact on the performance than L_{R}.

Figure 9 also highlights the difference in performance between an AF relay ($L_{\text{R}} = 0$) and an FF relay. In all the tested cases, the FF relay outperforms the AF relay by $2-3$ dB. In fact, an FF relay with less antennas may outperform an AF relay

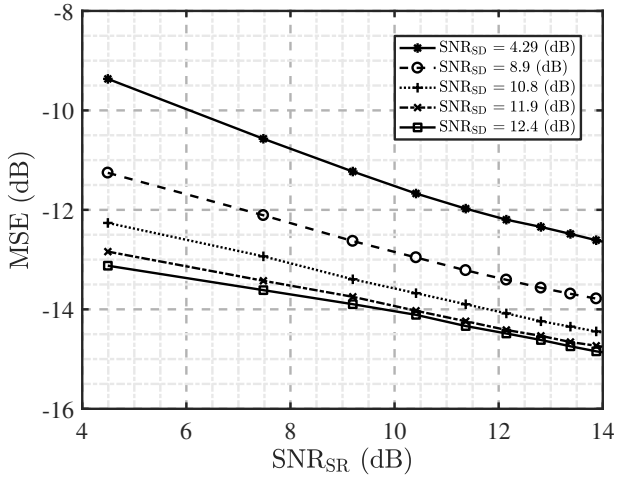

Fig. 4. MSE versus SNR_{SR} for various SNR_{SD}.

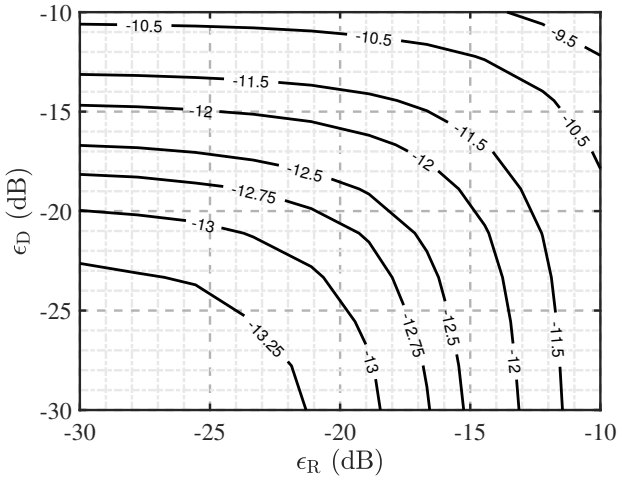

Fig. 5. MSE contour lines (in dB) versus destination distortion ϵ_{D} and relay distortion ϵ_{R}.

equipped with more antennas. This is, because for an AF relay, ρ is low even if M_{R} is large. Consequently, an FF protocol has clearly the edge in performance over an AF protocol, particularly in the case where the number of transmit antennas M_{R} is fixed and L_{R} is set on demand.

Figures 10 and 11 show the histogram of the number of iterations and the MSE for randomized initialization points when simulations parameters are fixed. The number of randomized initialization points is 1000, whereas simulations parameters are set to default values. From Fig. 10 the number of iterations to reach convergence is of the order of several dozens. Concretely, most of the times the algorithm takes than 25 iterations before obtaining a solution. Note that this number depends on the initialization point and the value of ε^2, where lower values of ε^2 result, on average, on a higher number of iterations. Similarly, because the stop criterion is based on the normalized change per iteration, the resulting MSE follows the histogram in Fig. 11. Note that the MSE is most likely to lie within an interval of $1.5 - 2$ dB around its mean value.

VI. Conclusions

We presented a method for MMSE design of node filters in a filter-and-forward full-duplex MIMO relay link subjected to limited dynamic range. The original non-convex optimization problem is approximated by an alternating optimization algorithm, where each node's filter is designed individually at each iteration. Simulations show that the balance between direct path (source-destination) and relay path (source-relay-destination) has a strong influence on the end-to-end performance, particularly when the source-to-relay path supports low SNR. A filter-and-forward protocol outperforms an amplify-and-forward protocol in most of the cases, even for fewer antennas. Limited dynamic range decreases the performance by a factor that depends on the individual hop SNR.

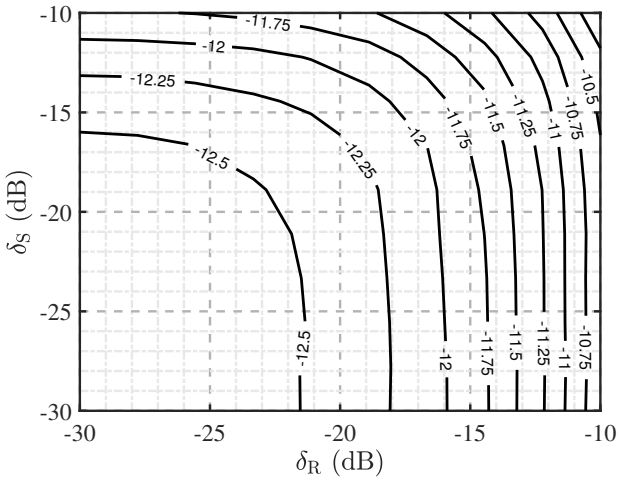

Fig. 6. MSE contour lines (in dB) versus source transmit noise δ_{S} and relay transmit noise δ_{R}.

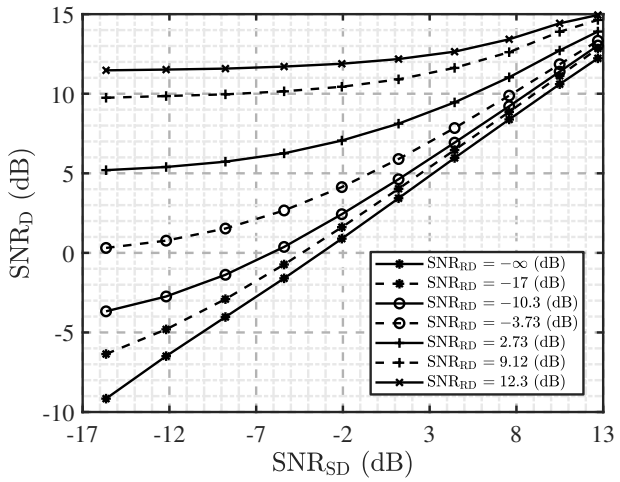

Fig. 7. $\mathrm{SNR_D}$ versus $\mathrm{SNR_{SD}}$ for various $\mathrm{SNR_{RD}}$.

APPENDIX A

Solution of (23) *in terms of* $\mathbf{G}_R[n]$ *for fixed* $\mathbf{g}_S[n]$ *and* $\mathbf{g}_D[n]$. Let first write the MSE as a function of $\mathbf{w}$. Using (32), $\hat{d}[n]$ has the following expression

$$\begin{aligned}\hat{d}[n] =& \boldsymbol{g}_{\mathrm{D}}^H(\mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{RD}})\mathcal{R}(\boldsymbol{\mathcal{G}}_{\mathrm{R}})\mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{SR}}) + \mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{SD}}))\mathcal{R}(\boldsymbol{g}_{\mathrm{S}})\boldsymbol{s} \\ &+ \boldsymbol{g}_{\mathrm{D}}^H \boldsymbol{z}_{\mathrm{D}}\end{aligned} \tag{51}$$

where $\boldsymbol{s}$ denotes a vector of appropriate size that contains current and L past samples of $s[n]$, i.e.,

$$\boldsymbol{s} = [s[n]\ s[n-1]\ \ldots\ s[n-L]]^T \tag{52}$$

By vectorizing both sides of (51), we obtain

$$\begin{aligned}\hat{d}[n] =& \boldsymbol{g}_{\mathrm{D}}^H(\mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{SD}})\mathcal{R}(\boldsymbol{g}_{\mathrm{S}})\boldsymbol{s} + \boldsymbol{z}_{\mathrm{D}}) \\ &+ \ \boldsymbol{g}_{\mathrm{D}}^H((\mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{SR}})\mathcal{R}(\boldsymbol{g}_{\mathrm{S}})\boldsymbol{s})^T \otimes \mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{RD}}))\mathbf{P}\mathbf{w}\end{aligned} \tag{53}$$

where $\mathbf{P} = \mathbf{V}\mathbf{N}$, and $\mathbf{V}$ being a reordering matrix satisfying $\mathrm{vec}\{\mathcal{R}(\boldsymbol{\mathcal{G}}_{\mathrm{R}})\} = \mathbf{V}\,\mathrm{vec}\{\boldsymbol{\mathcal{G}}_{\mathrm{R}}\}$, see Appendix C. From (53),

$$\begin{aligned}\mathbb{E}\{|\hat{d}[n] - s[n-\tau]|^2\} =& \mathbf{w}^H\mathbf{Q}_s\mathbf{w} - 2\,\mathrm{Re}\{\mathbf{w}^H\mathbf{q}_\tau\} \\ &+ \boldsymbol{g}_{\mathrm{D}}^H(\mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{RD}})\mathbf{R}_{\mathbf{z}_{\mathrm{R}}}\mathcal{R}^H(\boldsymbol{\mathcal{H}}_{\mathrm{SD}}) + \mathbf{R}_{\mathbf{w}_{\mathrm{D}}})\boldsymbol{g}_{\mathrm{D}} + q\end{aligned} \tag{54}$$

where matrix $\mathbf{Q}_s$, vector $\mathbf{q}_\tau$ and scalar q do not depend on $\mathbf{w}$ and are defined in Table II. Matrix $\mathbf{R}_{\mathbf{w}_{\mathrm{D}}}$ denotes the autocorrelation matrix of $\mathbf{w}_{\mathrm{D}}[n]$ that depends on $\mathbf{G}_{\mathrm{R}}[n]$.

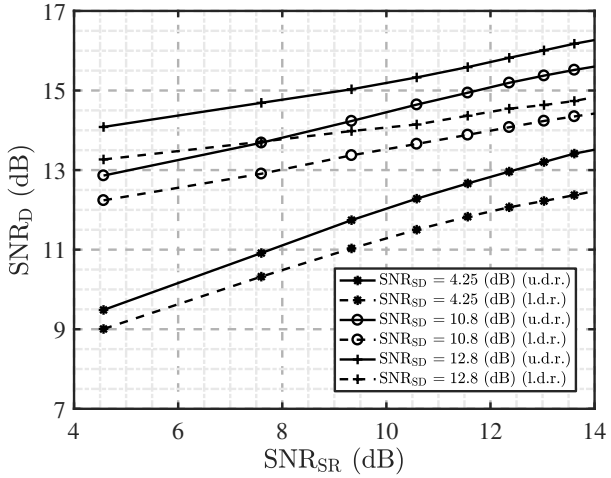

Fig. 8. $\mathrm{SNR_D}$ versus $\mathrm{SNR_{SR}}$ for cases of *limited dynamic range* (l.d.r.) and *unlimited dynamic range* (u.d.r.).

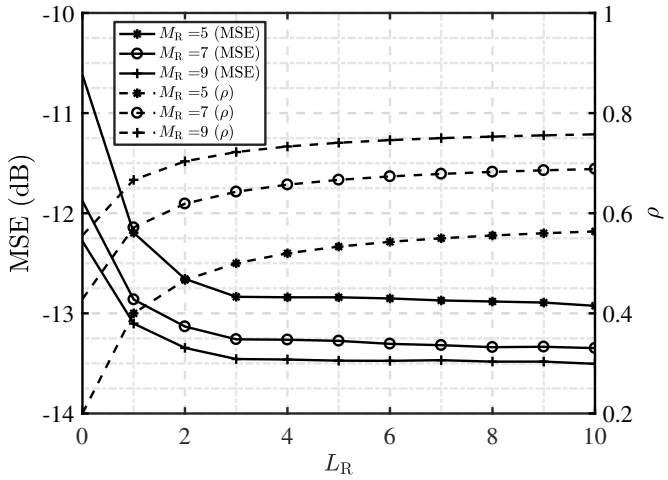

Fig. 9. MSE (solid) and fraction of DoF ρ (dashed) versus relay filter order L_{R} and number of transmit antennas of the relay M_{R}.

Equation (55) expresses $\boldsymbol{g}_{\mathrm{D}}^H \mathbf{R}_{\mathbf{w}_{\mathrm{D}}} \boldsymbol{g}_{\mathrm{D}}$ as a convex function of $\mathbf{w}$, where in Equality 1 we used the linearity of expectation and the properties of Kronecker product, whereas in Equality 2 we used Lemmas 2 and 3, presented below.

Lemma 2. *For any constant matrices* $\mathbf{A} \in \mathbb{C}^{M\times N}$ *and* $\mathbf{C} \in \mathbb{C}^{P\times R}$ *and any random matrix* $\mathbf{B} \in \mathbb{C}^{R\times M}$*, the following property holds*

$$\mathbb{E}\{\mathrm{tr}\{\mathbf{C}\,\mathrm{diag}\{\mathbf{B}\mathbf{A}\mathbf{A}^H\mathbf{B}^H\}\mathbf{C}^H\}\} = \mathrm{tr}\{\mathbf{A}^H\mathbb{E}\{\mathbf{B}^H\,\mathrm{diag}\{\mathbf{C}^H\mathbf{C}\}\mathbf{B}\}\mathbf{A}\} \quad (56)$$

Proof. Let $\mathbf{e}_i \in \mathbb{C}^R$ denote the ith canonical basis vector. Matrix $\mathbf{J}_i = \mathbf{e}_i\mathbf{e}_i^T \in \mathbb{C}^{R\times R}$ denotes the single-entry matrix whose ith diagonal element is one and zero elsewhere. The diag operation can be written as

$$\mathrm{diag}\{\mathbf{B}\mathbf{A}\mathbf{A}^H\mathbf{B}^H\} = \sum_{i=1}^{R} \mathbf{J}_i\mathbf{B}\mathbf{A}\mathbf{A}^H\mathbf{B}^H\mathbf{J}_i^H \quad (57)$$

Combining (57) and (56) yields

$$\mathbb{E}\{\mathrm{tr}\{\mathbf{C}\,\mathrm{diag}\{\mathbf{B}\mathbf{A}\mathbf{A}^H\mathbf{B}^H\}\mathbf{C}^H\}\} = \mathbb{E}\{\mathrm{tr}\{\mathbf{C}\sum_{p=1}^{R}\mathbf{J}_p\mathbf{B}\mathbf{A}\mathbf{A}^H\mathbf{B}^H\mathbf{J}_p^H\mathbf{C}^H\}\}$$

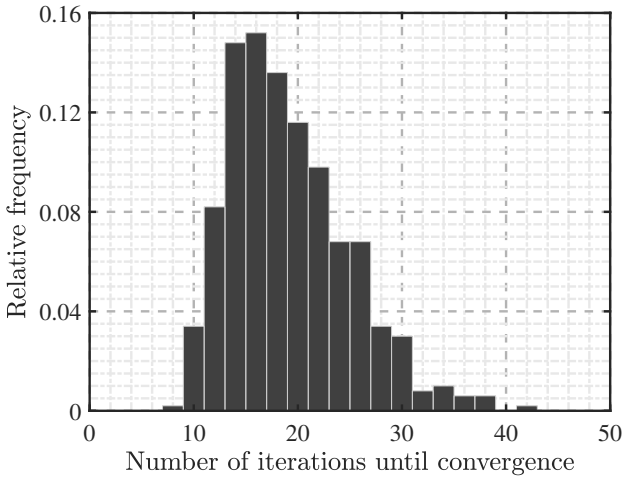

Fig. 10. Histogram of the number of iterations until convergence for randomized initialization points.

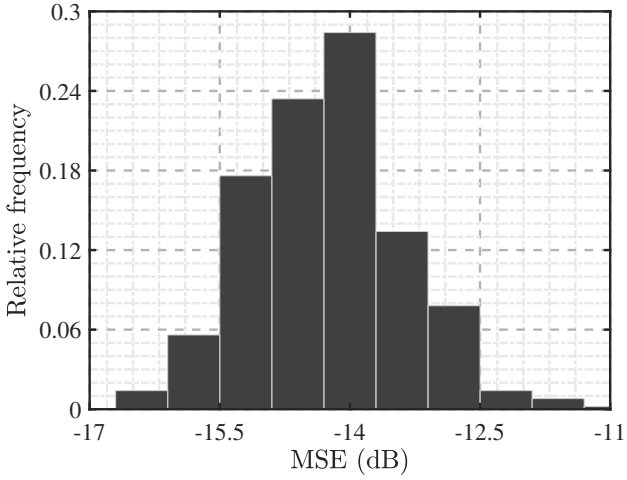

Fig. 11. Histogram of the MSE in dB for randomized initialization points.

$$\begin{aligned}
\boldsymbol{g}_{\mathrm{D}}^{H}\mathbf{R}_{\mathbf{w}_{\mathrm{D}}}\boldsymbol{g}_{\mathrm{D}} &= \boldsymbol{g}_{\mathrm{D}}^{H}\left(\mathbf{I}_{(L_{\mathrm{D}}+1)}\otimes\epsilon_{\mathrm{D}}\,\mathrm{diag}\left\{\left(\mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{RD}})\mathcal{R}(\boldsymbol{\mathcal{G}}_{\mathrm{R}})\mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{SR}})\mathcal{R}(\boldsymbol{g}_{\mathrm{S}})\right)\mathbb{E}\{\boldsymbol{s}\boldsymbol{s}^{H}\}\left(\mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{RD}})\mathcal{R}(\boldsymbol{\mathcal{G}}_{\mathrm{R}})\mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{SR}})\mathcal{R}(\boldsymbol{g}_{\mathrm{S}})\right)^{H}\right\}\right)\boldsymbol{g}_{\mathrm{D}} \\
&\overset{1}{=} \epsilon_{\mathrm{D}}\sum_{i=0}^{L_{\mathrm{D}}}\mathbb{E}\left\{\mathbf{g}_{\mathrm{D}}^{H}[i]\,\mathrm{diag}\left\{\left(\left(\mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{SR}})\mathcal{R}(\boldsymbol{g}_{\mathrm{S}})\boldsymbol{s}\right)^{T}\otimes\mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{RD}})\right)\mathbf{P}\mathbf{w}\mathbf{w}^{H}\mathbf{P}^{H}\left(\left(\mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{SR}})\mathcal{R}(\boldsymbol{g}_{\mathrm{S}})\boldsymbol{s}\right)^{T}\otimes\mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{RD}})\right)^{H}\right\}\mathbf{g}_{\mathrm{D}}[i]\right\} \\
&\overset{2}{=} \epsilon_{\mathrm{D}}\mathbf{w}^{H}\underbrace{\mathbb{E}\left\{\mathbf{P}^{H}\left(\left(\mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{SR}})\mathcal{R}(\boldsymbol{g}_{\mathrm{S}})\boldsymbol{s}\right)^{T}\otimes\mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{RD}})\right)^{H}\mathbf{I}_{N_{\mathrm{D}}}[L_{\mathrm{D}}]\,\mathrm{diag}\left\{\boldsymbol{g}_{\mathrm{D}}\boldsymbol{g}_{\mathrm{D}}^{H}\right\}\mathbf{I}_{N_{\mathrm{D}}}^{T}[L_{\mathrm{D}}]\left(\left(\mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{SR}})\mathcal{R}(\boldsymbol{g}_{\mathrm{S}})\boldsymbol{s}\right)^{T}\otimes\mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{RD}})\right)\mathbf{P}\right\}}_{\mathbf{Q}_{\epsilon_{\mathrm{D}}}}\mathbf{w} \qquad (55)
\end{aligned}$$

$$\begin{aligned}
&= \sum_{p=1}^{R}\mathbb{E}\{\mathrm{tr}\{\mathbf{A}^{H}\mathbf{B}^{H}\mathbf{J}_{p}^{H}\mathbf{C}^{H}\mathbf{C}\mathbf{J}_{p}\mathbf{B}\mathbf{A}\}\} \\
&= \mathrm{tr}\{\mathbf{A}^{H}\mathbb{E}\{\mathbf{B}^{H}\sum_{p=1}^{R}\left(\mathbf{J}_{p}\mathbf{C}^{H}\mathbf{C}\mathbf{J}_{p}^{H}\right)\mathbf{B}\}\mathbf{A}\} \\
&= \mathrm{tr}\{\mathbf{A}^{H}\mathbb{E}\{\mathbf{B}^{H}\,\mathrm{diag}\{\mathbf{C}^{H}\mathbf{C}\}\mathbf{B}\}\mathbf{A}\} \qquad (58)
\end{aligned}$$

where we have used that $\mathrm{tr}\{\mathbf{A}\mathbf{A}^{H}\} = \mathrm{tr}\{\mathbf{A}^{H}\mathbf{A}\}$ and that $\mathbf{J}_i$ is a Hermitian matrix. ■

Lemma 3. *For any filter* $\mathbf{A}[n] \in \mathbb{C}^{M\times N}$ *of order* L_A *the following property holds*

$$\sum_{k=0}^{L_A} \operatorname{diag}\left\{\mathbf{A}^H[k]\mathbf{A}[k]\right\} = \mathbf{I}_N[L_A] \operatorname{diag}\left\{\boldsymbol{\mathcal{A}}^H\boldsymbol{\mathcal{A}}\right\} \mathbf{I}_N^T[L_A] \tag{59}$$

Proof. Gramian matrix $\boldsymbol{\mathcal{A}}^H\boldsymbol{\mathcal{A}}$ has the following expression

$$\boldsymbol{\mathcal{A}}^H\boldsymbol{\mathcal{A}} = \begin{pmatrix} \mathbf{A}^H[0]\mathbf{A}[0] & \dots & \mathbf{A}^H[0]\mathbf{A}[L_A] \\ \vdots & \ddots & \vdots \\ \mathbf{A}^H[L_A]\mathbf{A}[0] & \dots & \mathbf{A}^H[L_A]\mathbf{A}[L_A] \end{pmatrix} \tag{60}$$

which, after applying the diag operation, results in

$$\operatorname{diag}\left\{\boldsymbol{\mathcal{A}}^H\boldsymbol{\mathcal{A}}\right\} = \begin{pmatrix} \operatorname{diag}\left\{\mathbf{A}^H[0]\mathbf{A}[0]\right\} & & \\ & \ddots & \\ & & \operatorname{diag}\left\{\mathbf{A}^H[L_A]\mathbf{A}[L_A]\right\} \end{pmatrix} \tag{61}$$

from where the result of the lemma follows immediately. ■

Proceeding in a similar fashion with the additional terms in $\boldsymbol{g}_{\mathrm{D}}^H\mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{RD}})\mathbf{R}_{\mathbf{z}_{\mathrm{R}}}\mathcal{R}^H(\boldsymbol{\mathcal{H}}_{\mathrm{SD}})\boldsymbol{g}_{\mathrm{D}}$, the MSE can be written as

$$\mathbb{E}\{|\hat{d}[n] - s[n-\tau]|^2\} = \mathbf{w}^H\mathbf{Q}\mathbf{w} - 2\operatorname{Re}\{\mathbf{w}^H\mathbf{q}_\tau\} + q \tag{62}$$

where $\mathbf{Q} = \mathbf{Q}_s + \mathbf{Q}_{\epsilon_{\mathrm{R}}} + \mathbf{Q}_{\epsilon_{\mathrm{D}}} + \mathbf{Q}_{\delta_{\mathrm{S}}} + \mathbf{Q}_{\delta_{\mathrm{R}}} + \mathbf{Q}_{\sigma_{\mathrm{R}}^2}$. Each matrix $\mathbf{Q}_i$ is a result of noise source with parameter i and signal $s[n]$ propagating to D and is defined in Table II.

The inequality constraint, $\mathbb{E}\{\|\hat{\mathbf{r}}[n]\|^2\} \leq P_{\mathrm{R}}$, can be written as

$$\mathbb{E}\{\|\hat{\mathbf{r}}[n]\|^2\} = \mathbf{w}^H\mathbf{Q}_{\mathbf{r}}\mathbf{w} \leq P_{\mathrm{R}} \tag{63}$$

with $\mathbf{Q}_{\mathbf{r}} = \mathbf{N}^H(\mathbf{R}_{\mathbf{r}}^T \otimes \mathbf{I}_{M_{\mathrm{R}}})^H\mathbf{N}$ and $\mathbf{R}_{\mathbf{r}}$ the autocorrelation of $\mathbf{r}[n]$, see Table II. Finally, the optimization problem can be expressed as:

$$\begin{aligned} \min_{\mathbf{w}} \quad & \mathbf{w}^H\mathbf{Q}\mathbf{w} - 2\operatorname{Re}\{\mathbf{w}^H\mathbf{q}_\tau\} + q \\ \text{subject to} \quad & \mathbf{w}^H\mathbf{Q}_{\mathbf{r}}\mathbf{w} \leq P_{\mathrm{R}} \end{aligned} \tag{64}$$

Problem (64), as explained in [46], is an LSQI problem, whose unique solution is obtained as follows:

1) If the unconstrained solution $\mathbf{w}^\star = \mathbf{Q}^{-1}\mathbf{q}_\tau$ satisfies $(\mathbf{w}^\star)^H\mathbf{Q}_{\mathbf{r}}\mathbf{w}^\star \leq P_{\mathrm{R}}$ then $\mathbf{w}^\star$ is a solution of (64).
2) If not, the optimal solution is obtained using the Lagrange function and is given by $\mathbf{w}^\star = (\mathbf{Q} + \lambda_{\mathrm{R}}^\star\mathbf{Q}_{\mathbf{r}})^{-1}\mathbf{q}_\tau$ where $\lambda_{\mathrm{R}}^\star$ is a Lagrange multiplier satisfying $(\mathbf{w}^\star)^H\mathbf{Q}_{\mathbf{r}}\mathbf{w}^\star = P_{\mathrm{R}}$.

The Lagrange multiplier $\lambda_{\mathrm{R}}^\star$ can be obtained by means of any standard root finding technique.

APPENDIX B

Solution of (23) *in terms of* $\mathbf{g}_S[n]$ *for fixed* $\mathbf{G}_R[n]$ *and* $\mathbf{G}_R[n]$. Let $\mathbf{f}_{\mathrm{S}} = \operatorname{vec}\{[\mathbf{g}_{\mathrm{S}}[0]\dots\mathbf{g}_{\mathrm{S}}[L_{\mathrm{S}}]]\} \in \mathbb{C}^{M_{\mathrm{S}}(L_{\mathrm{S}}+1)}$. By vectorizing both sides of (32), we obtain

$$\hat{d}[n] = \left(\boldsymbol{s}^T \otimes \boldsymbol{g}_{\mathrm{D}}^H\mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{eq}})\right)\mathbf{V}\mathbf{f}_{\mathrm{S}} + \boldsymbol{g}_{\mathrm{D}}^H\boldsymbol{z}_{\mathrm{D}} \tag{65}$$

with $\mathbf{V}$ being matrix satisfying $\operatorname{vec}\{\mathcal{R}(\boldsymbol{g}_{\mathrm{S}})\} = \mathbf{V}\mathbf{f}_{\mathrm{S}}$, see Appendix C. From (65),

$$\begin{aligned} &\mathbb{E}\{|\hat{d}[n] - s[n-\tau]|^2\} \\ &\qquad = \mathbf{f}_{\mathrm{S}}^H\mathbf{T}_s\mathbf{f}_{\mathrm{S}} - 2\operatorname{Re}\{\mathbf{f}_{\mathrm{S}}^H\mathbf{t}_\tau\} + \boldsymbol{g}_{\mathrm{D}}^H\mathbf{R}_{\mathbf{z}_{\mathrm{D}}}\boldsymbol{g}_{\mathrm{D}} \end{aligned} \tag{66}$$

where matrix $\mathbf{T}_s$ and vector $\mathbf{t}_\tau$ are defined in Table III. Proceeding in a similar fashion as in Appendix A, term $\boldsymbol{g}_{\mathrm{D}}^H\mathbf{R}_{\mathbf{z}_{\mathrm{D}}}\boldsymbol{g}_{\mathrm{D}}$ can be expressed as a convex function of $\mathbf{f}_{\mathrm{S}}$, resulting in

$$\mathbb{E}\{|\hat{d}[n] - s[n-\tau]|^2\} = \mathbf{f}_{\mathrm{S}}^H\mathbf{T}\mathbf{f}_{\mathrm{S}} - 2\operatorname{Re}\{\mathbf{f}_{\mathrm{S}}^H\mathbf{t}_\tau\} + t \tag{67}$$

where $\mathbf{T} = \mathbf{T}_s + \mathbf{T}_{\epsilon_{\mathrm{R}}} + \mathbf{T}_{\epsilon_{\mathrm{D}}} + \mathbf{T}_{\delta_{\mathrm{S}}} + \mathbf{T}_{\delta_{\mathrm{R}}}$ and scalar t collects all the terms independent of $\mathbf{g}_{\mathrm{S}}$. Each matrix $\mathbf{T}_i$ is a result of noise source i and signal $s[n]$ propagating to D and is defined in Table III.

The inequality constraint, $\mathbb{E}\{\|\hat{\mathbf{s}}[n]\|^2\} \leq P_{\mathrm{S}}$, can be written in terms of $\mathbf{g}_{\mathrm{S}}$ as

$$\mathbb{E}\{\|\hat{\mathbf{s}}[n]\|^2\} = \mathbf{f}_{\mathrm{S}}^H\mathbf{T}_{\hat{\mathbf{s}}}\mathbf{f}_{\mathrm{S}} \leq P_{\mathrm{S}} \tag{68}$$

with $\mathbf{T}_{\hat{\mathbf{s}}} = (\mathbf{R}_s^T \otimes \mathbf{I}_{M_\mathrm{S}})^H$ and $\mathbf{R}_s$ the autocorrelation of $s[n]$. Finally, $\mathbf{f}_\mathrm{S}$ is the solution to the following optimization problem:

$$\begin{aligned} \min_{\mathbf{f}_\mathrm{S}} \quad & \mathbf{f}_\mathrm{S}^H \mathbf{T} \mathbf{f}_\mathrm{S} - 2\,\mathrm{Re}\{\mathbf{f}_\mathrm{S}^H \mathbf{t}_\tau\} + t \\ \text{subject to} \quad & \mathbf{f}_\mathrm{S}^H \mathbf{T}_{\hat{\mathbf{s}}} \mathbf{f}_\mathrm{S} \leq P_\mathrm{S} \end{aligned} \tag{69}$$

From Appendix A, (69) is seen to be an LSQI problem, whose solution $\mathbf{f}_\mathrm{S}^\star$ is given by

$$\mathbf{f}_\mathrm{S}^\star = (\mathbf{T} + \lambda_\mathrm{S}^\star \mathbf{T}_{\hat{\mathbf{s}}})^{-1} \mathbf{t}_\tau \tag{70}$$

where Lagrange multiplier $\lambda_\mathrm{S}^\star \geq 0$ ensures $(\mathbf{f}_\mathrm{S}^\star)^H \mathbf{T}_{\hat{\mathbf{s}}} \mathbf{f}_\mathrm{S}^\star = P_\mathrm{S}$ when the unconstrained solution $\mathbf{f}_\mathrm{S}^\star = \mathbf{T}^{-1}\mathbf{t}_\tau$ yields $(\mathbf{f}_\mathrm{S}^\star)^H \mathbf{T}_{\hat{\mathbf{s}}} \mathbf{f}_\mathrm{S}^\star > P_\mathrm{S}$.

APPENDIX C

Proof of $\mathrm{vec}\{\mathcal{R}(\boldsymbol{\mathcal{B}})\} = \mathbf{V}\,\mathrm{vec}\{\boldsymbol{\mathcal{B}}\}$. In terms of $\boldsymbol{\mathcal{B}}$, (2) can be rewritten as

$$\mathcal{R}(\boldsymbol{\mathcal{B}}) = \sum_{j=0}^{L_\mathrm{A}} \mathbf{U}_j \boldsymbol{\mathcal{B}} \mathbf{W}_j \tag{71}$$

where $\mathbf{U}_j \in \mathbb{C}^{N(L_\mathrm{A}+1)\times N}$ and $\mathbf{W}_j \in \mathbb{C}^{P(L_\mathrm{B}+1)\times P(L_\mathrm{C}+1)}$ are given by

$$\mathbf{U}_j = \begin{bmatrix} \mathbf{0}_{jN\times N}^T & \mathbf{I}_N^T & \mathbf{0}_{(L_\mathrm{A}-j)N\times N}^T \end{bmatrix}^T \tag{72}$$

$$\mathbf{W}_j = \begin{bmatrix} \mathbf{0}_{P(L_\mathrm{B}+1)\times jP} & \mathbf{I}_{P(L_\mathrm{B}+1)} & \mathbf{0}_{P(L_\mathrm{B}+1)\times(L_\mathrm{A}-j)P} \end{bmatrix} \tag{73}$$

Using (71) and the properties of the vectorization operator,

$$\begin{aligned} \mathrm{vec}\{\mathcal{R}(\boldsymbol{\mathcal{B}})\} &= \mathrm{vec}\left\{ \sum_{j=0}^{L_\mathrm{A}} \mathbf{U}_j \boldsymbol{\mathcal{B}} \mathbf{W}_j \right\} \\ &= \sum_{j=0}^{L_\mathrm{A}} \mathrm{vec}\left\{ \mathbf{U}_j \boldsymbol{\mathcal{B}} \mathbf{W}_j \right\} \\ &= \sum_{j=0}^{L_\mathrm{A}} \left(\mathbf{W}_j^T \otimes \mathbf{U}_j \right) \mathrm{vec}\left\{ \boldsymbol{\mathcal{B}} \right\} \\ &= \mathbf{V}\,\mathrm{vec}\left\{ \boldsymbol{\mathcal{B}} \right\} \end{aligned} \tag{74}$$

where the resulting vector $\mathrm{vec}\left\{\boldsymbol{\mathcal{B}}\right\} \in \mathbb{C}^{NP(L_\mathrm{B}+1)}$ and matrix $\mathbf{V} = \sum_{j=0}^{L_\mathrm{A}} \left(\mathbf{W}_j^T \otimes \mathbf{U}_j \right) \in \mathbb{C}^{NP(L_\mathrm{C}+1)(L_\mathrm{A}+1)\times NP(L_\mathrm{B}+1)}$.

TABLE I
AUTOCORRELATION MATRIX OF $\mathbf{z}_\mathrm{D}[n]$, $\mathbf{R}_{\mathbf{z}_\mathrm{D}}$.

$\left\{L_{\mathrm{RD}'} = L_\mathrm{R} + L_\mathrm{RD} + L_\mathrm{D}\right\}$, $\left\{L_{\mathrm{RD}''} = L_{\mathrm{RD}'} + L_\mathrm{RR}\right\}$
$\mathbf{R}_{\mathbf{z}_\mathrm{D}} = \delta_\mathrm{R} \mathcal{R}(\boldsymbol{\mathcal{H}}_\mathrm{RD}) \mathcal{R}(\boldsymbol{\mathcal{G}}_\mathrm{R}) \mathcal{R}(\boldsymbol{\mathcal{H}}_\mathrm{RR}) \left(\mathbf{I}_{\left(L_{\mathrm{RD}''}+1\right)} \otimes \mathrm{diag}\left\{ \boldsymbol{\mathcal{G}}_\mathrm{R} \mathcal{R}(\boldsymbol{\mathcal{H}}_\mathrm{SR}) \mathcal{R}(\boldsymbol{g}_\mathrm{S}) \mathbf{R}_s \mathcal{R}^H(\boldsymbol{g}_\mathrm{S}) \mathcal{R}^H(\boldsymbol{\mathcal{H}}_\mathrm{SR}) \boldsymbol{\mathcal{G}}_\mathrm{R}^H \right\} \right) \mathcal{R}^H(\boldsymbol{\mathcal{H}}_\mathrm{RR}) \mathcal{R}^H(\boldsymbol{\mathcal{G}}_\mathrm{R}) \mathcal{R}^H(\boldsymbol{\mathcal{H}}_\mathrm{RD})$
$+\delta_\mathrm{R} \mathcal{R}(\boldsymbol{\mathcal{H}}_\mathrm{RD}) \left(\mathbf{I}_{(L_\mathrm{D}+L_\mathrm{RD}+1)} \otimes \mathrm{diag}\left\{ \boldsymbol{\mathcal{G}}_\mathrm{R} \mathcal{R}(\boldsymbol{\mathcal{H}}_\mathrm{SR}) \mathcal{R}(\boldsymbol{g}_\mathrm{S}) \mathbf{R}_s \mathcal{R}^H(\boldsymbol{g}_\mathrm{S}) \mathcal{R}^H(\boldsymbol{\mathcal{H}}_\mathrm{SR}) \boldsymbol{\mathcal{G}}_\mathrm{R}^H \right\} \right) \mathcal{R}^H(\boldsymbol{\mathcal{H}}_\mathrm{RD}) + \sigma_\mathrm{R}^2 \mathcal{R}(\boldsymbol{\mathcal{H}}_\mathrm{RD}) \mathcal{R}(\boldsymbol{\mathcal{G}}_\mathrm{R}) \mathcal{R}^H(\boldsymbol{\mathcal{G}}_\mathrm{R}) \mathcal{R}^H(\boldsymbol{\mathcal{H}}_\mathrm{RD})$
$+\epsilon_\mathrm{R} \mathcal{R}(\boldsymbol{\mathcal{H}}_\mathrm{RD}) \mathcal{R}(\boldsymbol{\mathcal{G}}_\mathrm{R}) \left(\mathbf{I}_{\left(L_{\mathrm{RD}'}+1\right)} \otimes \mathrm{diag}\left\{ \boldsymbol{\mathcal{H}}_\mathrm{SR} \mathcal{R}(\boldsymbol{g}_\mathrm{S}) \mathbf{R}_s \mathcal{R}^H(\boldsymbol{g}_\mathrm{S}) \boldsymbol{\mathcal{H}}_\mathrm{SR}^H \right\} \right) \mathcal{R}^H(\boldsymbol{\mathcal{G}}_\mathrm{R}) \mathcal{R}^H(\boldsymbol{\mathcal{H}}_\mathrm{RD}) + \sigma_\mathrm{D}^2 \mathbf{I}_{N_\mathrm{D}(L_\mathrm{D}+1)}$
$+\epsilon_\mathrm{D} \left(\mathbf{I}_{(L_\mathrm{D}+1)} \otimes \mathrm{diag}\left\{ \boldsymbol{\mathcal{H}}_{eq} \mathcal{R}(\boldsymbol{g}_\mathrm{S}) \mathbf{R}_s \mathcal{R}^H(\boldsymbol{g}_\mathrm{S}) \boldsymbol{\mathcal{H}}_{eq}^H \right\} \right) + \delta_\mathrm{S} \mathcal{R}(\boldsymbol{\mathcal{H}}_\mathrm{eq}) \left(\mathbf{I}_{(L_\mathrm{D}+L_\mathrm{eq}+1)} \otimes \mathrm{diag}\left\{ \boldsymbol{g}_\mathrm{S} \mathbf{R}_s \boldsymbol{g}_\mathrm{S}^H \right\} \right) \mathcal{R}^H(\boldsymbol{\mathcal{H}}_\mathrm{eq})$

ACKNOWLEDGMENT

This work was partially supported by the Academy of Finland under Grants 296849 and 288249, by the Research Council of Norway, the Agencia Estatal de Investigación (Spain) and the European Regional Development Fund (ERDF) under project WINTER (TEC2016-76409-C2-2-R), and by the Xunta de Galicia (Agrupación Estratéxica Consolidada de Galicia accreditation 2016-2019; Red Temática RedTEIC 2017-2018).

TABLE II
MSE IN TERMS OF $\mathbf{G}_{\mathrm{R}}[n]$ FOR FIXED $\mathbf{g}_{\mathrm{S}}[n]$ AND $\mathbf{g}_{\mathrm{D}}[n]$. OPERATOR * DENOTES CONJUGATION.

DEFINITIONS
$\left\{L_{\mathrm{SD}'} = L_{\mathrm{SR}} + L_{\mathrm{R}} + L_{\mathrm{RD}} + L_{\mathrm{D}}\right\}$, $\left\{L_{\mathrm{RD}''} = L_{\mathrm{R}} + L_{\mathrm{RR}} + L_{\mathrm{RD}} + L_{\mathrm{D}}\right\}$, $\left\{\mathbf{D}_{\mathrm{R}}[n] = \mathbf{I}_{M_{\mathrm{R}}}\delta[n] + \mathbf{G}_{\mathrm{R}}^{(k-1)}[n] \star \mathbf{H}_{\mathrm{RR}}[n]\right\}$
$\mathbf{Q}_s = \mathbf{P}^H \mathbb{E}\left\{\left((\mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{SR}})\mathcal{R}(\boldsymbol{g}_{\mathrm{S}})\boldsymbol{s})^* \otimes \mathcal{R}^H(\boldsymbol{\mathcal{H}}_{\mathrm{RD}})\boldsymbol{g}_{\mathrm{D}}\right)\left((\mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{SR}})\mathcal{R}(\boldsymbol{g}_{\mathrm{S}})\boldsymbol{s})^T \otimes \boldsymbol{g}_{\mathrm{D}}^H\mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{RD}})\right)\right\}\mathbf{P}$
$\mathbf{q}_\tau = \mathbf{P}^H\left((\mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{SR}})\mathcal{R}(\boldsymbol{g}_{\mathrm{S}})\mathbb{E}\left\{\boldsymbol{s}s^*[n-\tau]\right\})^* \otimes \mathcal{R}^H(\boldsymbol{\mathcal{H}}_{\mathrm{RD}})\boldsymbol{g}_{\mathrm{D}}\right)$
$\mathbf{Q}_{\delta_{\mathrm{S}}} = \delta_{\mathrm{S}}\mathbf{P}^H\left(\left(\mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{SR}})\left(\mathbf{I}_{(L_{\mathrm{SD}'}+1)} \otimes \mathrm{diag}\left\{\boldsymbol{g}_{\mathrm{S}}\mathbf{R}_s\boldsymbol{g}_{\mathrm{S}}^H\right\}\right)\mathcal{R}^H(\boldsymbol{\mathcal{H}}_{\mathrm{SR}})\right)^* \otimes \mathcal{R}^H(\boldsymbol{\mathcal{H}}_{\mathrm{RD}})\boldsymbol{g}_{\mathrm{D}}\boldsymbol{g}_{\mathrm{D}}^H\mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{RD}})\right)\mathbf{P}$
$\mathbf{Q}_{\epsilon_{\mathrm{D}}} = \epsilon_{\mathrm{D}}\mathbf{P}^H\mathbb{E}\left\{\left((\mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{SR}})\mathcal{R}(\boldsymbol{g}_{\mathrm{S}})\boldsymbol{s})^* \otimes \mathcal{R}^H(\boldsymbol{\mathcal{H}}_{\mathrm{RD}})\right)\mathbf{I}_{N_{\mathrm{D}}}[L_{\mathrm{D}}]\,\mathrm{diag}\left\{\boldsymbol{g}_{\mathrm{D}}\boldsymbol{g}_{\mathrm{D}}^H\right\}\mathbf{I}_{N_{\mathrm{D}}}^T[L_{\mathrm{D}}]\left((\mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{SR}})\mathcal{R}(\boldsymbol{g}_{\mathrm{S}})\boldsymbol{s})^T \otimes \mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{RD}})\right)\right\}\mathbf{P}$
$\mathbf{Q}_{\delta_{\mathrm{R}}} = \delta_{\mathrm{R}}\mathbf{P}^H\mathbb{E}\left\{\left((\mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{SR}})\mathcal{R}(\boldsymbol{g}_{\mathrm{S}})\boldsymbol{s})^* \otimes \mathbf{I}_{M_{\mathrm{R}}}\right)\mathbf{I}_{M_{\mathrm{R}}}[L_{\mathrm{RD}''}]\,\mathrm{diag}\left\{\mathcal{R}^H(\boldsymbol{\mathcal{D}}_{\mathrm{R}})\mathcal{R}^H(\boldsymbol{\mathcal{H}}_{\mathrm{RD}})\boldsymbol{g}_{\mathrm{D}}\boldsymbol{g}_{\mathrm{D}}^H\mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{RD}})\mathcal{R}(\boldsymbol{\mathcal{D}}_{\mathrm{R}})\right\}\mathbf{I}_{M_{\mathrm{R}}}^T[L_{\mathrm{RD}''}]\left((\mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{SR}})\mathcal{R}(\boldsymbol{g}_{\mathrm{S}})\boldsymbol{s})^T \otimes \mathbf{I}_{M_{\mathrm{R}}}\right)\right\}\mathbf{P}$
$\mathbf{Q}_{\epsilon_{\mathrm{R}}} = \epsilon_{\mathrm{R}}\mathbf{P}^H\left(\left(\mathbf{I}_{(L_{\mathrm{R}}+L_{\mathrm{RD}}+L_{\mathrm{D}}+1)} \otimes \boldsymbol{\mathcal{H}}_{\mathrm{SR}}\mathcal{R}(\boldsymbol{g}_{\mathrm{S}})\mathbf{R}_s\mathcal{R}^H(\boldsymbol{g}_{\mathrm{S}})\boldsymbol{\mathcal{H}}_{\mathrm{SR}}^H\right)^* \otimes \mathcal{R}^H(\boldsymbol{\mathcal{H}}_{\mathrm{RD}})\boldsymbol{g}_{\mathrm{D}}\boldsymbol{g}_{\mathrm{D}}^H\mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{RD}})\right)\mathbf{P}$
$\mathbf{Q}_{\sigma_{\mathrm{R}}^2} = \sigma_{\mathrm{R}}^2\mathbf{P}^H\left(\mathbf{I}_{(N_{\mathrm{R}}(L_{\mathrm{R}}+L_{\mathrm{RD}}+L_{\mathrm{D}}+1))} \otimes \mathcal{R}^H(\boldsymbol{\mathcal{H}}_{\mathrm{RD}})\boldsymbol{g}_{\mathrm{D}}\boldsymbol{g}_{\mathrm{D}}^H\mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{RD}})\right)\mathbf{P}$
$q = \boldsymbol{g}_{\mathrm{D}}^H\left(\sigma_{\mathrm{D}}^2\mathbf{I}_{(N_{\mathrm{D}}(L_{\mathrm{D}}+1))} + \mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{SD}})\mathcal{R}(\boldsymbol{g}_{\mathrm{S}})\mathbf{R}_s\mathcal{R}^H(\boldsymbol{g}_{\mathrm{S}})\mathcal{R}^H(\boldsymbol{\mathcal{H}}_{\mathrm{SD}})\right)\boldsymbol{g}_{\mathrm{D}} + \mathbb{E}\left\{\|s[n]\|^2\right\} - 2\,\mathrm{Re}\left\{\mathbf{g}_{\mathrm{D}}^H\mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{SD}})\mathcal{R}(\boldsymbol{g}_{\mathrm{S}})\mathbb{E}\left\{\boldsymbol{s}s^*[n-\tau]\right\}\right\}$ $+\boldsymbol{g}_{\mathrm{D}}^H\left(\delta_{\mathrm{S}}\mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{SD}})\,\mathrm{diag}\left\{\mathbf{I}_{(L_{\mathrm{SD}}+L_{\mathrm{D}}+1)} \otimes \boldsymbol{g}_{\mathrm{S}}\mathbf{R}_s\boldsymbol{g}_{\mathrm{S}}^H\right\}\mathcal{R}^H(\boldsymbol{\mathcal{H}}_{\mathrm{SD}}) + \epsilon_{\mathrm{R}}\,\mathrm{diag}\left\{\mathbf{I}_{(L_{\mathrm{D}}+1)} \otimes \boldsymbol{\mathcal{H}}_{\mathrm{SD}}\mathcal{R}(\boldsymbol{g}_{\mathrm{S}})\mathbf{R}_s\mathcal{R}^H(\boldsymbol{g}_{\mathrm{S}})\boldsymbol{\mathcal{H}}_{\mathrm{SD}}^H\right\}\right)\boldsymbol{g}_{\mathrm{D}}$
$\mathbf{Q}_{\mathbf{r}} = \mathbf{N}^H\left(\mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{SR}})\mathcal{R}(\boldsymbol{g}_{\mathrm{S}})\mathbf{R}_s\mathcal{R}^H(\boldsymbol{g}_{\mathrm{S}})\mathcal{R}^H(\boldsymbol{\mathcal{H}}_{\mathrm{SR}}) \otimes \mathbf{I}_{M_{\mathrm{R}}}\right)^*\mathbf{N}$
MSE as a function of $\mathbf{w}$, $\mathbf{g}_{\mathrm{R}} = \mathbf{N}\mathbf{w} = \mathrm{vec}\{\left[\mathbf{G}_{\mathrm{R}}^T[0] \;\cdots\; \mathbf{G}_{\mathrm{R}}^T[L_{\mathrm{R}}]\right]^T\}$
$\mathbb{E}\left\{\left\|\hat{d}[n] - s[n-\tau]\right\|^2\right\} = \mathbf{w}^H\left(\underbrace{\left(\mathbf{Q}_s + \mathbf{Q}_{\delta_{\mathrm{S}}} + \mathbf{Q}_{\delta_{\mathrm{R}}} + \mathbf{Q}_{\epsilon_{\mathrm{R}}} + \mathbf{Q}_{\sigma_{\mathrm{R}}} + \mathbf{Q}_{\epsilon_{\mathrm{D}}}\right)}_{\mathbf{Q}}\right)\mathbf{w} - 2\,\mathrm{Re}\left\{\mathbf{w}^H\mathbf{q}_\tau\right\} + q$

REFERENCES

[1] M. Iwamura, H. Takahashi, and S. Nagata, "Relay technology in LTE-advanced," *NTT DoCoMo Technical Journal*, vol. 12, no. 2, pp. 29–36, Sep. 2010.
[2] 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.
[3] A. Mattsson, "Single frequency networks in DTV," *IEEE Trans. Broadcast.*, vol. 51, no. 4, pp. 413–422, Dec. 2005.
[4] M. Yu and J. Li, "Is amplify-and-forward practically better than decode-and-forward or vice versa?" in *Proc. IEEE Int. Conf. Acoust., Speech and Signal Process. (ICASSP)*, vol. 3, Mar. 2005, pp. 365–368.
[5] N. Bornhorst and M. Pesavento, "Filter-and-forward beamforming with adaptive decoding delays in asynchronous multi-user relay networks," *Signal Processing*, vol. 109, pp. 132–147, Apr. 2015.
[6] T. Riihonen, S. Werner, and R. Wichman, "Hybrid full-duplex/half-duplex relaying with transmit power adaptation," *IEEE Trans. Wireless Commun.*, vol. 10, no. 9, pp. 3074–3085, Sep. 2011.
[7] M. Souryal and B. Vojcic, "Performance of amplify-and-forward and decode-and-forward relaying in Rayleigh fading with turbo codes," in *Proc. IEEE Int. Conf. Acoust., Speech and Signal Process. (ICASSP)*, vol. 4, May 2006.
[8] S. Mohammadkhani and G. K. Karagiannidis, "Filter-and-forward relaying in cognitive networks with blind channel estimation," *IET Commun.*, vol. 10, no. 18, pp. 2678–2686, Dec. 2016.
[9] M. Agrawal, D. J. Love, and V. Balakrishnan, "Communicating over filter-and-forward relay networks with channel output feedback," *IEEE Trans. Signal Process.*, vol. 64, no. 5, pp. 1117–1131, March 2016.
[10] K. P. Chou, J. C. Lin, and H. V. Poor, "Disintegrated channel estimation in filter-and-forward relay networks," *IEEE Trans. Commun.*, vol. 64, no. 7, pp. 2835–2847, July 2016.
[11] D. Kim, Y. Sung, and J. Chung, "Filter-and-forward relay design for mimo-ofdm systems," *IEEE Trans. Commun.*, vol. 62, no. 7, pp. 2329–2339, July 2014.
[12] Y. w. Liang, A. Ikhlef, W. Gerstacker, and R. Schober, "Two-way filter-and-forward beamforming for frequency-selective channels," *IEEE Trans. Wireless Commun.*, vol. 10, no. 12, pp. 4172–4183, Dec. 2011.
[13] H. Chen, A. B. Gershman, and S. Shahbazpanahi, "Filter-and-forward distributed beamforming in relay networks with frequency selective fading," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1251–1262, Mar. 2010.
[14] 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.
[15] M. Duarte and A. Sabharwal, "Full-duplex wireless communications using off-the-shelf radios: Feasibility and first results," in *Proc. Conf. Signals, Syst. and Comput. (ASILOMAR)*, Nov. 2010, pp. 1558–1562.
[16] M. Duarte, C. Dick, and A. Sabharwal, "Experiment-driven characterization of full-duplex wireless systems," *IEEE Trans. Wireless Commun.*, vol. 11, no. 12, pp. 4296–4307, Dec. 2012.
[17] M. Heino, D. Korpi, T. Huusari, E. Antonio-Rodriguez, S. Venkatasubramanian, T. Riihonen, L. Anttila, C. Icheln, K. Haneda, R. Wichman, and M. Valkama, "Recent advances in antenna design and interference cancellation algorithms for in-band full duplex relays," *IEEE Commun. Mag.*, vol. 53, no. 5, pp. 91–101, May 2015.
[18] A. Sabharwal, P. Schniter, D. Guo, D. W. Bliss, S. Rangarajan, and R. Wichman, "In-band full-duplex wireless: Challenges and opportunities," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 9, pp. 1637–1652, Sept. 2014.

TABLE III
MSE IN TERMS OF $\mathbf{g}_{\mathrm{S}}[n]$ FOR FIXED $\mathbf{G}_{\mathrm{R}}[n]$ AND $\mathbf{g}_{\mathrm{D}}[n]$. OPERATOR * DENOTES CONJUGATION.

DEFINITIONS
$\left\{L_{\mathrm{RD}'} = L_{\mathrm{R}} + L_{\mathrm{RD}} + L_{\mathrm{D}}\right\}, \left\{L_{\mathrm{RD}''} = L_{\mathrm{RD}'} + L_{\mathrm{RR}}\right\}, \left\{\mathbf{D}_{\mathrm{R}}[n] = \mathbf{I}_{M_{\mathrm{R}}}\delta[n] + \mathbf{G}_{\mathrm{R}}[n] \star \mathbf{H}_{\mathrm{RR}}[n]\right\}$
$\mathbf{T}_s = \mathbf{V}^H \left(\mathbf{R}_s^* \otimes \mathcal{R}^H(\boldsymbol{\mathcal{H}}_{\mathrm{eq}}) \boldsymbol{g}_{\mathrm{D}} \boldsymbol{g}_{\mathrm{D}}^H \mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{eq}})\right) \mathbf{V}$
$\mathbf{t}_\tau = \mathbf{V}^H \left(\mathbb{E}\{\boldsymbol{s}^* s[n-\tau]\} \otimes \mathcal{R}^H(\boldsymbol{\mathcal{H}}_{\mathrm{eq}}) \boldsymbol{g}_{\mathrm{D}}\right)$
$\mathbf{T}_{\delta_{\mathrm{S}}} = \delta_{\mathrm{S}} \mathbb{E}\left\{(\boldsymbol{s}^* \otimes \mathbf{I}_{M_{\mathrm{S}}}) \mathbf{I}_{M_{\mathrm{S}}}[L_{\mathrm{eq}} + L_{\mathrm{D}}] \operatorname{diag}\left\{\mathcal{R}^H(\boldsymbol{\mathcal{H}}_{\mathrm{eq}}) \boldsymbol{g}_{\mathrm{D}} \boldsymbol{g}_{\mathrm{D}}^H \mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{eq}})\right\} \mathbf{I}_{M_{\mathrm{S}}}^T[L_{\mathrm{eq}} + L_{\mathrm{D}}] \left(\boldsymbol{s}^T \otimes \mathbf{I}_{M_{\mathrm{S}}}\right)\right\}$
$\mathbf{T}_{\epsilon_{\mathrm{D}}} = \epsilon_{\mathrm{D}} \mathbf{V}^H \mathbb{E}\left\{\left(\boldsymbol{s}^* \otimes \boldsymbol{\mathcal{H}}_{\mathrm{eq}}^H\right) \mathbf{I}_{N_{\mathrm{D}}}[L_{\mathrm{D}}] \operatorname{diag}\{\boldsymbol{g}_{\mathrm{D}} \boldsymbol{g}_{\mathrm{D}}^H\} \mathbf{I}_{N_{\mathrm{D}}}^T[L_{\mathrm{D}}] \left(\boldsymbol{s}^T \otimes \boldsymbol{\mathcal{H}}_{\mathrm{eq}}\right)\right\} \mathbf{V}$
$\mathbf{T}_{\epsilon_{\mathrm{R}}} = \epsilon_{\mathrm{R}} \mathbf{V}^H \mathbb{E}\left\{\left(\boldsymbol{s}^* \otimes \boldsymbol{\mathcal{H}}_{\mathrm{SR}}^H\right) \mathbf{I}_{N_{\mathrm{R}}}[L_{\mathrm{RD}'}] \operatorname{diag}\left\{\mathcal{R}^H(\boldsymbol{\mathcal{G}}_{\mathrm{R}}) \mathcal{R}^H(\boldsymbol{\mathcal{H}}_{\mathrm{RD}}) \boldsymbol{g}_{\mathrm{D}} \boldsymbol{g}_{\mathrm{D}}^H \mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{RD}}) \mathcal{R}(\boldsymbol{\mathcal{G}}_{\mathrm{R}})\right\} \mathbf{I}_{N_{\mathrm{R}}}^T[L_{\mathrm{RD}'}] \left(\boldsymbol{s}^T \otimes \boldsymbol{\mathcal{H}}_{\mathrm{SR}}\right)\right\} \mathbf{V}$
$\mathbf{T}_{\delta_{\mathrm{R}}} = \delta_{\mathrm{R}} \mathbf{V}^H \mathbb{E}\left\{\left(\boldsymbol{s}^* \otimes \mathcal{R}^H(\boldsymbol{\mathcal{H}}_{\mathrm{SR}}) \boldsymbol{\mathcal{G}}_{\mathrm{R}}^H\right) \mathbf{I}_{M_{\mathrm{R}}}[L_{\mathrm{RD}''}] \operatorname{diag}\left\{\mathcal{R}^H(\boldsymbol{\mathcal{D}}_{\mathrm{R}}) \mathcal{R}^H(\boldsymbol{\mathcal{H}}_{\mathrm{RD}}) \boldsymbol{g}_{\mathrm{D}} \boldsymbol{g}_{\mathrm{D}}^H \mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{RD}}) \mathcal{R}(\boldsymbol{\mathcal{D}}_{\mathrm{R}})\right\} \mathbf{I}_{M_{\mathrm{R}}}^T[L_{\mathrm{RD}''}] \left(\boldsymbol{s}^T \otimes \boldsymbol{\mathcal{G}}_{\mathrm{R}} \mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{SR}})\right)\right\} \mathbf{V}$
$t = \mathbb{E}\left\{\|s[n]\|^2\right\} + \boldsymbol{g}_{\mathrm{D}}^H \left(\sigma_{\mathrm{D}}^2 \mathbf{I}_{(N_{\mathrm{D}}(L_{\mathrm{D}}+1))} + \sigma_{\mathrm{R}}^2 \mathcal{R}(\boldsymbol{\mathcal{H}}_{\mathrm{RD}}) \mathcal{R}(\boldsymbol{\mathcal{G}}_{\mathrm{R}}) \mathcal{R}^H(\boldsymbol{\mathcal{G}}_{\mathrm{R}}) \mathcal{R}^H(\boldsymbol{\mathcal{H}}_{\mathrm{RD}})\right) \boldsymbol{g}_{\mathrm{D}}$
MSE as a function of $\mathbf{f}_{\mathrm{S}} = \mathrm{vec}\{[\mathbf{g}_{\mathrm{S}}[0] \dots \mathbf{g}_{\mathrm{S}}[L_{\mathrm{S}}]]\}$
$\mathbb{E}\left\{\lvert\hat{d}[n] - s[n-\tau]\rvert^2\right\} = \mathbf{f}_{\mathrm{S}}^H \Bigg(\underbrace{\mathbf{T}_s + \mathbf{T}_{\delta_{\mathrm{S}}} + \mathbf{T}_{\delta_{\mathrm{R}}} + \mathbf{T}_{\epsilon_{\mathrm{R}}} + \mathbf{T}_{\epsilon_{\mathrm{D}}}}_{\mathbf{T}}\Bigg) \mathbf{f}_{\mathrm{S}} - 2\,\mathrm{Re}\left\{\mathbf{f}_{\mathrm{S}}^H \mathbf{t}_\tau\right\} + t$

[19] Z. Zhang, X. Chai, K. Long, A. V. Vasilakos, and L. Hanzo, "Full duplex techniques for 5G networks: self-interference cancellation, protocol design, and relay selection," *IEEE Commun. Mag.*, vol. 53, no. 5, pp. 128–137, May 2015.

[20] G. Liu, F. R. Yu, H. Ji, V. C. M. Leung, and X. Li, "In-band full-duplex relaying: A survey, research issues and challenges," *IEEE Commun. Surveys & Tutorials*, vol. 17, no. 2, pp. 500–524, 2015.

[21] H. Hamazumi, K. Imamura, N. Iai, K. Shibuya, and M. Sasaki, "A study of a loop interference canceller for the relay stations in an SFN for digital terrestrial broadcasting," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, vol. 1, 2000, pp. 167–171.

[22] 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. Annu. Conf. Inf. Sci. and Syst. (CISS)*, Mar. 2011.

[23] P. Lioliou, M. Viberg, M. Coldrey, and F. Athley, "Self-interference suppression in full-duplex MIMO relays," in *Proc. Conf. Signals, Syst. and Comput. (ASILOMAR)*, Nov. 2010, pp. 658–662.

[24] P. Larsson and M. Prytz, "MIMO on-frequency repeater with self-interference cancellation and mitigation," in *Proc. IEEE Veh. Technol. Conf. (VTC)*, Apr. 2009.

[25] E. Everett, A. Sahai, and A. Sabharwal, "Passive self-interference suppression for full-duplex infrastructure nodes," *IEEE Trans. Wireless Commun.*, vol. 13, no. 2, pp. 680–694, Feb. 2014.

[26] N. Li, W. Zhu, and H. Han, "Digital interference cancellation in single channel, full duplex wireless communication," in *Proc. Int. Conf. Wireless Commun., Networking and Mobile Computing (WiCOM)*, Sep. 2012.

[27] E. Antonio-Rodríguez, R. López-Valcarce, T. Riihonen, S. Werner, and R. Wichman, "Autocorrelation-based adaptation rule for feedback equalization in wideband full-duplex amplify-and-forward MIMO relays," in *Proc. IEEE Int. Conf. Acoust., Speech and Signal Process. (ICASSP)*, May 2013.

[28] E. Antonio-Rodríguez, R. López-Valcarce, T. Riihonen, S. Werner, and R. Wichman, "Adaptive self-interference cancellation in wideband full-duplex decode-and-forward MIMO relays," in *Proc. IEEE Int. Workshop on Signal Process. Advances in Wireless Commun. (SPAWC)*, Jun. 2013, pp. 370–374.

[29] 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.

[30] E. Antonio-Rodríguez, S. Werner, R. López-Valcarce, T. Riihonen, and R. Wichman, "Wideband full-duplex mimo relays with blind adaptive self-interference cancellation," *Signal Processing*, vol. 130, pp. 74–85, Jan. 2017.

[31] B. Day, A. Margetts, D. Bliss, and P. Schniter, "Full-duplex MIMO relaying: Achievable rates under limited dynamic range," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 8, pp. 1541–1553, Sep. 2012.

[32] D. Korpi, T. Riihonen, V. Syrjl, L. Anttila, M. Valkama, and R. Wichman, "Full-duplex transceiver system calculations: Analysis of ADC and linearity challenges," *IEEE Trans. Wireless Commun.*, vol. 13, no. 7, pp. 3821–3836, July 2014.

[33] O. Taghizadeh, J. Zhang, and M. Haardt, "Transmit beamforming aided amplify-and-forward mimo full-duplex relaying with limited dynamic range," *Signal Processing*, no. 127, pp. 266–281, Oct. 2016.

[34] D. Kim, J. Seo, and Y. Sung, "Filter-and-forward transparent relay design for OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 62, no. 9, pp. 4392–4407, Nov. 2013.

[35] H. Chen, S. Shahbazpanahi, and A. B. Gershman, "Filter-and-forward distributed beamforming for two-way relay networks with frequency selective channels," *IEEE Trans. Signal Process.*, vol. 60, no. 4, pp. 1927–1941, Apr. 2012.

[36] M. Maleki and V. T. Vakili, "Filter-and-forward transceiver design for cognitive two-way relay networks," *IET Commun.*, vol. 9, no. 17, pp. 2061–2069, 2015.

[37] Y. w. Liang, A. Ikhlef, W. Gerstacker, and R. Schober, "Cooperative filter-and-forward beamforming for frequency-selective channels with equalization," *IEEE Trans. Wireless Commun.*, vol. 10, no. 1, pp. 228–239, Jan. 2011.

[38] S. Koyanagi and T. Miyajima, "Filter-and-forward based full-duplex relay networks with cooperative beamforming," *Proc. IEEE Int. Workshop on Signal Process. Advances in Wireless Commun. (SPAWC)*, pp. 1–5, July 2017.

[39] B. C. Nguyen, X. N. Tran, and D. T. Tran, "Performance analysis of in-band full-duplex amplify-and-forward relay system with direct link," *Proc. Int. Conf. on Recent Advances in Signal Process. Telecomm. Computing (SigTelCom)*, pp. 192–197, Jan. 2018.

[40] D. P. M. Osorio, E. E. B. Olivo, H. Alves, J. C. S. S. Filho, and M. Latva-aho, "Exploiting the direct link in full-duplex amplify-and-forward relaying networks," *IEEE Signal Process. Letters*, vol. 22, no. 10, pp. 1766–1770, Oct. 2015.

[41] O. Taghizadeh, A. C. Cirik, and R. Mathar, "Hardware impairments aware transceiver design for full-duplex amplify-and-forward mimo relaying," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 1644–1659, March 2018.
[42] A. Hjørungnes, *Complex-valued matrix derivatives: With Applications in Signal Processing and Communications*. Cambridge University Press, Mar. 2011.
[43] K. E. Kolodziej, J. G. McMichael, and B. T. Perry, "Multitap RF canceller for in-band full-duplex wireless communications," *IEEE Trans. Wireless Commun.*, vol. 15, no. 6, pp. 4321–4334, June 2016.
[44] T. Riihonen, S. Werner, and R. Wichman, "Comparison of full-duplex and half-duplex modes with a fixed amplify-and-forward relay," in *Proc. IEEE Wireless Commu. and Networking Conf. (WCNC)*, Apr. 2009.
[45] J. C. Bezdek and R. J. Hathaway, "Convergence of alternating optimization," *Neural, Parallel Sci. Comput.*, vol. 11, no. 4, pp. 351–368, 2003.
[46] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 4th ed., ser. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, Dec. 2012.
[47] Y. Gong and C. F. N. Cowan, "Optimum decision delay of the finite-length DFE," *IEEE Signal Process. Letters*, no. 11, pp. 858–861, Nov. 2004.
[48] W. Zhou and S. Zhang, "The decision delay in finite-length MMSE/DFE systems," *Wireless Personal Communications*, no. 1, pp. 175–189, Jul. 2005.
[49] P. A. Voois, I. Lee, and J. M. Cioffi, "The effect of decision delay in finite-length decision feedback equalization," *IEEE Trans. Information Theory*, no. 2, pp. 618–621, March 1996.