Transmit Beamforming Aided Amplify-and-Forward MIMO Full-Duplex Relaying with Limited Dynamic Range[☆]

Omid Taghizadeh[1]

Institute for Theoretical Information Technology, RWTH Aachen University, D-52074 Aachen, Germany (email: taghizadeh@ti.rwth-aachen.de).

Jianshu Zhang, Martin Haardt[2]

Communications Research Laboratory, Ilmenau University of Technology, D-98684 Ilmenau, Germany (email: martin.haardt@tu-ilmenau.de, jianshu.zhang@tu-ilmenau.de).

Abstract

In this paper we address the design of a relay-assisted communication system where two half-duplex (HD) users communicate with each other via the help of an amplify-and-forward (AF) and full-duplex (FD) relay node. We design joint user beamforming and relay transmit strategies to deal with the self-interference signal at the FD relay node. Our approach is to maximize the system sum rate via linear transmit strategies by exploiting multiple antennas at all the involved nodes. Convex optimization based sub-optimal and optimal solutions are developed. More specifically, an alternating optimization for self-interference aware FD relaying (AO) is devised when the transmission of multiple streams is considered. Moreover, a unified approach via gradient projections is proposed as a benchmark. It can be applied regardless of the number of antennas at each node but has a significantly higher computational complexity than the AO algorithm. Simulation results show that the proposed algorithms can achieve a good performance compared to the benchmark algorithm. Moreover, a significant FD gain is achieved in terms of the sum rate when the residual self-interference power is smaller than the noise.

Keywords: MIMO, beamforming, full-duplex, one-way relay,

 Part of this work has been presented at the Tenth International Symposium on Wireless Communication Systems, Ilmenau, Germany (ISWCS 13) [1].

[1]Corresponding author. Currently, O. Taghizadeh is with the Institute for Theoretical Information Technology, RWTH Aachen University, D-52074 Aachen, Germany. This work was performed while O. Taghizadeh was with the Communications Research Laboratory, Ilmenau University of Technology, D-98684 Ilmenau, Germany.

[2]EURASIP Member.

amplify-and-forward.

1. Introduction

Full-duplex (FD) operation is defined as a transceiver's capability to transmit and receive at the same time and frequency. While FD-enabled communication schemes are beneficial in many desired aspects (e.g., lower delay, higher efficiency and security, improved access layer function, ... [2]), they were long considered to be practically infeasible due to the inherent self-interference. Recently, exploiting specialized cancellation techniques [3–8], FD communications have been introduced as a possibility for short range applications. In theory, the loopback self-interference signal is known to the receiver and hence it can be successfully subtracted from the received signal if i) we have a sufficiently high dynamic range at the receiver and the transmitter, and ii) we have a perfect knowledge of the self-interference path, i.e., the channel between the transmit and the receive antennas of the same node. It is clear that in practice none of the aforementioned requirements are perfectly satisfied, considering the erroneous behavior of the transceiver and more importantly the strength of the self-interference path compared to the desired channel [9].

In order to address this challenge, more sophisticated methods have been introduced to deal with the interference in different stages. First, an effective physical isolation between the transmit and the receive ends must be ensured for any FD transceiver. This can be achieved via proper separation of the antennas or via exploiting the antennas' directivity [10]. Furthermore, we can exploit the knowledge of the transmit signal as well as the knowledge of the self-interference channel such that the self-interference is estimated and canceled in the RF domain. Thereby the main part of interference is suppressed at the receiver prior to down-conversion, in order to avoid the destructive effects of the limited dynamic range of the analog-to-digital convertor (ADC), cf. [11, 12]. As shown in [7], digital cancellation methods can also be used to estimate the residual self-interference components in the receiver-end in order to improve the cancellation quality.

Other than self-interference cancellation techniques, many papers have focused on potential applications of FD operation in future communication systems [13]. This includes identifying appropriate use cases as well as system optimization in the presence of FD devices. To this end, performance optimization of FD-enabled communication systems has been studied for a point-to-point communication system in [9]. Performance optimization involving a FD base station has been considered in [14], and the issues regarding physical layer security have been addressed in [15]. In particular, FD relay-assisted communication has been investigated to enhance the efficiency of classic half-duplex relaying schemes [16].

A FD relay is capable of receiving the signal from the source, while simultaneously communicating to the destination. This capability not only reduces the required time slots in order to accomplish an end-to-end communication,

but also reduces the latency compared to traditionally Time Division Duplex (TDD)-based half-duplex (HD) relays [16]. Therefore, a FD decode-and-forward (DF) relaying scheme is studied in [17, 18], regarding the methodologies for relay selection and power allocation, and in [19, 20] in order to maximize the system sum rate for relays with multiple antennas. Amplify-and-forward (AF) relays are specially attractive due to their performance and operation simplicity [21, 22]. Therefore, algorithms for relay selection and power allocation for single-antenna AF-FD relays have been proposed in [23–27] for various single-antenna relaying setups, taking into account the effects of residual self-interference at the relay. Regarding the relaying systems with multiple antenna nodes, schemes to mitigate the self-interference intensity at the relay via linear transmit strategies are presented in [28, 29], while algorithms to maximize the end-to-end performance are studied in [30–36]. In [32], a minimum-mean-squared-error (MMSE) approach is derived to design an AF-FD relaying system where perfect self-interference cancellation is assumed. In [33] a setup with multiple single antenna users and a relay with multiple transmit antennas and a single receive antenna is studied. A setup with single antenna users and a multiple antenna AF-FD relay is studied in [35] from the perspective of joint relay/antenna selection and performance analysis, and in [37] regarding the joint transmit/receive filter design, assuming a perfect hardware operation. The work in [34], proposes low complexity designs for signal-to-interference-plus-noise (SINR) maximization in an AF-FD relaying system. Nevertheless, a joint user and relay optimization for the users and a relay with multiple antennas, where digital and analog domain self-interference cancellation methods are simultaneously applied, is still an open problem.

In this paper, we investigate the sum rate[3] maximization problem for a relay assisted communication system with a FD relay and two HD users. Our main contributions are as follows: In Section 2, we define a system where digital domain and analog domain cancellation techniques are simultaneously exploited to benefit from FD transmission. As a result, the limitation of the RF domain suppression techniques is taken into account as one of our design constraints in the digital domain. Afterwards, advanced linear transmit strategies for the relay as well as the end users are derived to maximize the system sum rate in Sections 3 and 4. More precisely, when all the nodes have single antennas, a closed-form solution is derived. When only the relay has multiple antennas, a rank-one optimal solution is derived by using semi-definite relaxation. When all the nodes have multiple antennas, an alternating optimization for self-interference aware FD relaying (AO) is devised. A reduced-complexity AO (RC-AO) algorithm is proposed based on the optimal transmit strategy for a point-to-point MIMO system, i.e., singular value decomposition (SVD) based precoding together with water-filling based power allocation. To benchmark the proposed algorithms, a

[3]Please note that the term *sum rate* may refer to the sum of the communication rates in different links, or to the sum of communications rates corresponding to different data streams in a single link. The latter case is the intention of this paper.

unified design is proposed via gradient projections, which is applicable regardless of the number of antennas at each node. Numerical simulations are presented in Section 5.

Throughout this paper, column vectors and matrices are denoted as lower-case and upper-case bold letters, respectively. The rank of a matrix, expectation, trace, determinant, transpose, conjugate and Hermitian transpose are denoted by $\mathrm{rank}(\cdot)$, $\mathbb{E}(\cdot)$, $\mathrm{Tr}(\cdot)$, $|\cdot|$, $(\cdot)^{\mathrm{T}}$, $(\cdot)^*$ and $(\cdot)^{\mathrm{H}}$, respectively. The identity matrix with dimension K is denoted as $\boldsymbol{I}_K$ and $\mathrm{vec}(\cdot)$ operator stacks the elements of a matrix into a vector. Moreover, $(\cdot)^{-1}$ represents the inverse of a matrix and $||\cdot||_2$ represents the Euclidean norm of a vector. The set of all positive semi-definite matrices with Hermitian symmetry is denoted by $\mathcal{H}$.

2. System model

In this work we investigate the case where a single pair of HD, multiple-antenna users communicate with the help of a one-way FD relay as depicted in Fig. 1. The relay has $M_{\mathrm{t}}^{(\mathrm{R})}$ transmit and $M_{\mathrm{r}}^{(\mathrm{R})}$ receive antennas and uses the AF strategy. Due to the simultaneous transmission and reception of the FD relay, it only takes a single communication phase to accomplish an end-to-end communication between a source and a destination node. The source and the destination nodes are equipped with M_{t} and M_{r} antennas, respectively. Our channels are full-rank, quasi-stationary[4] flat-fading, and perfect synchronization is assumed among all transmit and receive chains. The channel between the source and the relay is denoted as $\boldsymbol{H}_{\mathrm{SR}} \in \mathbb{C}^{M_{\mathrm{r}}^{(\mathrm{R})} \times M_{\mathrm{t}}}$, the channel between the relay transmit and receiver ends (self-interference channel) is denoted as $\boldsymbol{H}_{\mathrm{RR}} \in \mathbb{C}^{M_{\mathrm{r}}^{(\mathrm{R})} \times M_{\mathrm{t}}^{(\mathrm{R})}}$, and the channel between the relay and the destination is denoted as $\boldsymbol{H}_{\mathrm{RD}} \in \mathbb{C}^{M_{\mathrm{r}} \times M_{\mathrm{t}}^{(\mathrm{R})}}$. We assume that perfect channel state information is available and the direct channel between the source and the destination is ignored. We may now formulate the transmitted signal from the source as

$$\boldsymbol{x} = \boldsymbol{F}\boldsymbol{s}, \tag{1}$$

where $\boldsymbol{F} \in \mathbb{C}^{M_{\mathrm{t}} \times M_{\mathrm{t}}}$ is the source precoder and $\boldsymbol{s} \in \mathbb{C}^{M_{\mathrm{t}}}$ is our transmitted data vector ($\mathbb{E}\{\boldsymbol{s}\boldsymbol{s}^{\mathrm{H}}\} = \boldsymbol{I}_{M_t}$). The transmit power constraint at the source has to fulfill $\mathbb{E}\{\|\boldsymbol{x}\|_2^2\} \leq P_{\max}$, where $P_{\max}$ represents the maximum transmit power. The received signal at the relay can be written as

$$\boldsymbol{y}_{\mathrm{R}} = \boldsymbol{H}_{\mathrm{SR}}\boldsymbol{x} + \boldsymbol{n}_{\mathrm{R}} + \boldsymbol{H}_{\mathrm{RR}}\boldsymbol{x}_{\mathrm{R}}, \tag{2}$$

where $\boldsymbol{n}_{\mathrm{R}} \in \mathbb{C}^{M_{\mathrm{r}}^{(\mathrm{R})}}$ denotes the zero-mean complex Gaussian (ZMCG) noise, where we have $\mathbb{E}\{\boldsymbol{n}_{\mathrm{R}}\boldsymbol{n}_{\mathrm{R}}{}^{\mathrm{H}}\} = \sigma_{\mathrm{nr}}^2 \boldsymbol{I}_{M_{\mathrm{r}}^{(\mathrm{R})}}$. The last term in (2) denotes the received self-interference where $\boldsymbol{x}_{\mathrm{R}} \in \mathbb{C}^{M_{\mathrm{t}}^{(\mathrm{R})}}$ represents the transmit signal from the relay.

[4]It implies that the channel is constant within one frame but may vary from frame to frame.

Although recent cancellation methods demonstrate suppression of the self-interference signal down to the receiver noise floor for short range communications, e.g., [7] for WiFi 802.11ac, such designs are not yet suitable for high-power communication schemes. This stems from the fact that while higher self-interference power requires higher levels of suppression, the involved hardware components tend to become less accurate and closer to saturation as the power increases. Furthermore, various real-world imperfections, e.g., aging of the analog circuit elements and limited channel coherence time, or limited estimation accuracy in the digital domain, are inevitable. These effects degrade the performance of self-interference cancellation techniques [38]. In [38], it has been observed that different available self-interference cancellation schemes result in different cancellation capabilities.

In order to take into account the aforementioned limits we assume a residual self-interference signal with zero mean and a Gaussian distribution, which increases the receiver noise floor [39, 40]. Furthermore, a constraint on the suppressible self-interference power is applied according to [41, 42] in order to define the functional dynamic range of the relay: $\mathbb{E}\{\|\boldsymbol{H}_{\mathrm{RR}}\boldsymbol{x}_{\mathrm{R}}\|_2^2\} \leq P_{\mathrm{th}}^{(\mathrm{R})}$, where $P_{\mathrm{th}}^{(\mathrm{R})}$ is the maximum tolerable self-interference power.[5] Note that the role of the defined threshold is to ensure that the involved hardware components in the self-interference cancellation process operate in their accurate functional range, and prevent saturation and nonlinear/undesired behavior. As a result, it is highly dependent on the implemented self-interference cancellation. In our work, since the P_{th} is defined as the maximum average interference power, it should be chosen by considering a peak-to-average-power-ratio (PAPR) margin, regarding the hardware components which are sensitive to the instantaneous power. Thereby, the signal transmitted from the relay can be formulated as

$$\begin{aligned} \boldsymbol{x}_{\mathrm{R}} &= \boldsymbol{G}(\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{x} + \boldsymbol{n}_{\mathrm{R}} + \boldsymbol{i}_{\mathrm{R}}), \\ \mathbb{E}\{\|\boldsymbol{H}_{\mathrm{RR}}\boldsymbol{x}_{\mathrm{R}}\|_2^2\} &\leq P_{\mathrm{th}}^{(\mathrm{R})},\ \mathbb{E}\{\|\boldsymbol{x}_{\mathrm{R}}\|_2^2\} \leq P_{\mathrm{max}}^{(\mathrm{R})}, \end{aligned} \tag{3}$$

where $\boldsymbol{i}_{\mathrm{R}}$ is a ZMCG residual self-interference with $\mathbb{E}\{\boldsymbol{i}_{\mathrm{R}}\boldsymbol{i}_{\mathrm{R}}{}^{\mathrm{H}}\} = \sigma_{\mathrm{si}}^2\boldsymbol{I}_{M_{\mathrm{r}}^{(\mathrm{R})}}$, and $\sigma_{\mathrm{r}}^2 := \sigma_{\mathrm{nr}}^2 + \sigma_{\mathrm{si}}^2$ is the variance of the noise plus residual self-interference.[6][7] The relay amplification matrix is denoted as $\boldsymbol{G} \in \mathbb{C}^{M_{\mathrm{t}}^{(\mathrm{R})} \times M_{\mathrm{r}}^{(\mathrm{R})}}$ and $P_{\mathrm{max}}^{(\mathrm{R})}$ represents the maximum transmit power at the relay. Finally, the received signal at the

[5] Other than the total tolerable self-interference power, we may define the tolerable self-interference power on each individual receiver chain, on the collective self-interference which originates from each transmit chain, or on a mixture of the aforementioned cases. The respective choice depends on the implemented self-interference cancellation scheme and the affordable design complexity.

[6] In this work a constant σ_{si}^2 is assumed within the relay self-interference power range defined by $P_{\mathrm{th}}^{(\mathrm{R})}$, see (3). More accurate models are also available which take into account the dependence of the residual interference covariance matrix on the instantaneous relay power, e.g., [24, 25], at the expense of a higher complexity.

[7] Note that for a given FD transceiver, the appropriate values of P_{th} and σ_{si}^2 can be obtained by conducting experiments and measurements.

destination is given as

$$\boldsymbol{y}_{\mathrm{D}} = \boldsymbol{H}_{\mathrm{RD}}\boldsymbol{x}_{\mathrm{R}} + \boldsymbol{n}_{\mathrm{D}}, \tag{4}$$

where $\boldsymbol{n}_{\mathrm{D}}$ is the ZMCG noise with $\mathbb{E}\{\boldsymbol{n}_{\mathrm{D}}\boldsymbol{n}_{\mathrm{D}}{}^{\mathrm{H}}\} = \sigma_{\mathrm{nd}}^2\boldsymbol{I}_{M_{\mathrm{r}}}$. In the following we study the rate maximization problem for the defined scenario.

3. Enhanced transmit strategies at user and relay nodes

In this section, our goal is to maximize the mutual information (MI) among the end users via a joint design of our transmit precoder ($\boldsymbol{F}$) and the relay amplification matrix ($\boldsymbol{G}$). The optimization problem is formulated into

$$\max_{\boldsymbol{F},\boldsymbol{G}} \quad \mathrm{MI}\left(\boldsymbol{x};\boldsymbol{y}_{\mathbf{D}}\right) \tag{5a}$$

$$\text{s.t.} \quad \mathbb{E}\left\{\|\boldsymbol{x}\|_2^2\right\} \leq P_{\max}, \tag{5b}$$

$$\mathbb{E}\{\|\boldsymbol{x}_{\mathrm{R}}\|_2^2\} \leq P_{\max}^{(\mathrm{R})}, \tag{5c}$$

$$\mathbb{E}\{\|\boldsymbol{H}_{\mathrm{RR}}\boldsymbol{x}_{\mathrm{R}}\|_2^2\} \leq P_{\mathrm{th}}^{(\mathrm{R})}, \tag{5d}$$

where $\mathrm{MI}\left(\boldsymbol{x};\boldsymbol{y}_{\mathbf{D}}\right)$ represents the mutual information between the input arguments given the current system parameters ($\boldsymbol{H}_{\mathrm{RD}}, \boldsymbol{G}, \boldsymbol{F}, \boldsymbol{H}_{\mathrm{SR}}$) and (5b-d) represent the user, relay and the self-interference power constraints, respectively. The mutual information, given $\boldsymbol{F}$ and $\boldsymbol{G}$, is calculated as

$$\mathrm{MI}\left(\boldsymbol{x};\boldsymbol{y}_{\mathbf{D}}\right) = \log_2\Big|\boldsymbol{I}_{M_{\mathrm{r}}} + (\boldsymbol{H}_{\mathrm{RD}}\boldsymbol{G}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{F})(\boldsymbol{H}_{\mathrm{RD}}\boldsymbol{G}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{F})^{\mathrm{H}} \left(\sigma_{\mathrm{r}}^2\boldsymbol{H}_{\mathrm{RD}}\boldsymbol{G}\boldsymbol{G}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{RD}}^{\mathrm{H}} + \sigma_{\mathrm{nd}}^2\boldsymbol{I}_{M_r}\right)^{-1}\Big|. \tag{6}$$

Note that (6) is only achievable for sources with a Gaussian distribution and thus can be viewed as an achievable upper-bound for the end-to-end mutual information. Benefiting from the results of majorization theory [43] and using the Karush-Kuhn-Tucker (KKT) conditions for optimality [44], the optimal solution to (5) with a relaxed interference power constraint (5d) has been provided in [21, 30, 45]. Although the additional constraint (5d) changes the nature of our problem, the aforementioned results are still useful for our investigation. In the following, we solve the non-convex problem (5) under different system settings.

3.1. Special case: $M_{\mathrm{t}} = M_{\mathrm{r}} = 1,\ M_t^{(R)} = M_r^{(R)} = 1$

In this case all of the desired and interference channels simplify to SISO channels. Hence, the design of optimal transmit strategies simplifies to a power adjustment at the relay and at the source node

$$\max_{P_{\mathrm{T}},g} \quad \log_2\left(1+\frac{P_{\mathrm{T}}|h_{\mathrm{SR}}|^2|g|^2|h_{\mathrm{RD}}|^2}{\sigma_{\mathrm{r}}^2|g|^2|h_{\mathrm{RD}}|^2+\sigma_{\mathrm{nd}}^2}\right) \tag{7a}$$

$$\text{s.t.} \quad P_{\mathrm{T}} \leq P_{\max}, \tag{7b}$$

$$P_{\mathrm{T}}|h_{\mathrm{SR}}|^2|g|^2+\sigma_{\mathrm{r}}^2|g|^2 \leq P_{\max}^{(\mathrm{R})}, \tag{7c}$$

$$|h_{\mathrm{RR}}|^2\left(P_{\mathrm{T}}|h_{\mathrm{SR}}|^2|g|^2+\sigma_{\mathrm{r}}^2|g|^2\right) \leq P_{\mathrm{th}}^{(\mathrm{R})}, \tag{7d}$$

where P_{T} and g are the user's transmit power and the relay's amplification coefficient, and $h_{\mathrm{SR}}, h_{\mathrm{RR}}, h_{\mathrm{RD}}$ are scalar (SISO) notations of the corresponding channel matrices. Furthermore, we observe that the common component $(P_{\mathrm{T}}|h_{\mathrm{SR}}|^2|g|^2+\sigma_{\mathrm{r}}^2|g|^2)$ from (7c), (7d), can be substituted with a single equivalent power constraint

$$\max_{P_{\mathrm{T}},g} \quad \log_2\left(1+\frac{P_{\mathrm{T}}|h_{\mathrm{SR}}|^2|g|^2|h_{\mathrm{RD}}|^2}{\sigma_{\mathrm{r}}^2|g|^2|h_{\mathrm{RD}}|^2+\sigma_{\mathrm{nd}}^2}\right) \tag{8a}$$

$$\text{s.t.} \quad P_{\mathrm{T}} \leq P_{\max}, \tag{8b}$$

$$P_{\mathrm{T}}|h_{\mathrm{SR}}|^2|g|^2+\sigma_{\mathrm{r}}^2|g|^2 \leq \min\left\{P_{\max}^{(\mathrm{R})}, \frac{P_{\mathrm{th}}^{(\mathrm{R})}}{|h_{\mathrm{RR}}|^2}\right\}. \tag{8c}$$

At this point we observe that in the optimality of (8), the transmit power constraint is always satisfied with equality for the source node (see Proposition 1 for a detailed discussion). This leaves us with the following single parameter optimization problem:

$$\max_{g} \quad \frac{P_{\max}|h_{\mathrm{SR}}|^2|g|^2|h_{\mathrm{RD}}|^2}{\sigma_{\mathrm{r}}^2|g|^2|h_{\mathrm{RD}}|^2+\sigma_{\mathrm{nd}}^2} \tag{9a}$$

$$\text{s.t.} \quad |g|^2 \leq \frac{\min\left\{P_{\max}^{(\mathrm{R})}, \frac{P_{\mathrm{th}}^{(\mathrm{R})}}{|h_{\mathrm{RR}}|^2}\right\}}{P_{\max}|h_{\mathrm{SR}}|^2+\sigma_{\mathrm{r}}^2} =: \zeta. \tag{9b}$$

Due to the fact that (9b) bounds $|g|^2$ from above and our cost function is monotonically increasing with respect to $|g|^2$, we achieve an optimal analytic solution for g, P_{T} as

$$|g| = \sqrt{\zeta}, \quad P_{\mathrm{T}} = P_{\max}. \tag{10}$$

Note that as the objective value (9a) and the constraint (9b) are both invariant to the phase of g. Therefore, the phase can be chosen arbitrarily without loss of optimality. It is worth mentioning that this setup corresponds to the single user scenario studied in [33] with a single antenna base station, as well as the investigated power allocation strategies for single antenna FD relays in [24, 25].

3.2. Special case: $M_{\mathrm{t}} = M_{\mathrm{r}} = 1,\ M_t^{(R)}, M_r^{(R)} > 1$

Under this setting, our original problem in (5) can be reformulated as

$$\max_{P_{\mathrm{T}}, \boldsymbol{G}} \quad \log_2 \left(1 + \frac{P_{\mathrm{T}} \boldsymbol{h}_{\mathrm{RD}}^{\mathrm{T}} \boldsymbol{G} \boldsymbol{h}_{\mathrm{SR}} \boldsymbol{h}_{\mathrm{SR}}^{\mathrm{H}} \boldsymbol{G}^{\mathrm{H}} \boldsymbol{h}_{\mathrm{RD}}^{*}}{\sigma_{\mathrm{r}}^2 \boldsymbol{h}_{\mathrm{RD}}^{\mathrm{T}} \boldsymbol{G} \boldsymbol{G}^{\mathrm{H}} \boldsymbol{h}_{\mathrm{RD}}^{*} + \sigma_{\mathrm{nd}}^2}\right) \tag{11a}$$

$$\text{s.t.} \quad P_{\mathrm{T}} \leq P_{\max}, \tag{11b}$$

$$\mathrm{Tr}\left(\boldsymbol{G}\left[P_{\mathrm{T}} \boldsymbol{h}_{\mathrm{SR}} \boldsymbol{h}_{\mathrm{SR}}^{\mathrm{H}} + \sigma_{\mathrm{r}}^2 \boldsymbol{I}_{M_{\mathrm{r}}^{(\mathrm{R})}}\right] \boldsymbol{G}^{\mathrm{H}}\right) \leq P_{\max}^{(\mathrm{R})}, \tag{11c}$$

$$\mathrm{Tr}\left(\boldsymbol{H}_{\mathrm{RR}} \boldsymbol{G}\left[P_{\mathrm{T}} \boldsymbol{h}_{\mathrm{SR}} \boldsymbol{h}_{\mathrm{SR}}^{\mathrm{H}} + \sigma_{\mathrm{r}}^2 \boldsymbol{I}_{M_{\mathrm{r}}^{(\mathrm{R})}}\right] \boldsymbol{G}^{\mathrm{H}} \boldsymbol{H}_{\mathrm{RR}}^{\mathrm{H}}\right) \leq P_{\mathrm{th}}^{(\mathrm{R})}, \tag{11d}$$

where $\boldsymbol{h}_{\mathrm{SR}}, \boldsymbol{h}_{\mathrm{RD}}^{\mathrm{T}}$ are the vector representations of $\boldsymbol{H}_{\mathrm{SR}}, \boldsymbol{H}_{\mathrm{RD}}$, respectively.

Proposition 1. *At the optimality of (11) the transmit power constraint (11b) is active ($P_{\mathrm{T}} = P_{\max}$).*

Proof It is observed that the following variable update: $\forall \kappa > 1,\ P_{\mathrm{T}} \leftarrow \kappa P_{\mathrm{T}}$ and $\boldsymbol{G} \leftarrow \frac{\boldsymbol{G}}{\sqrt{\kappa}}$ increases the value of the objective (11a) while preserving the feasibility (the relay's transmit and the interference powers are not increased). Thus we have $P_{\mathrm{T}} = P_{\max}$ at the optimality.

Proposition 1 simplifies our problem as

$$\max_{\boldsymbol{G}} \quad \frac{P_{\max} \boldsymbol{h}_{\mathrm{RD}}^{\mathrm{T}} \boldsymbol{G} \boldsymbol{h}_{\mathrm{SR}} \boldsymbol{h}_{\mathrm{SR}}^{\mathrm{H}} \boldsymbol{G}^{\mathrm{H}} \boldsymbol{h}_{\mathrm{RD}}^{*}}{\sigma_{\mathrm{r}}^2 \boldsymbol{h}_{\mathrm{RD}}^{\mathrm{T}} \boldsymbol{G} \boldsymbol{G}^{\mathrm{H}} \boldsymbol{h}_{\mathrm{RD}}^{*} + \sigma_{\mathrm{nd}}^2} \tag{12a}$$

$$\text{s.t.} \quad \mathrm{Tr}\left(\boldsymbol{G}\left[P_{\max} \boldsymbol{h}_{\mathrm{SR}} \boldsymbol{h}_{\mathrm{SR}}^{\mathrm{H}} + \sigma_{\mathrm{r}}^2 \boldsymbol{I}_{M_{\mathrm{r}}^{(\mathrm{R})}}\right] \boldsymbol{G}^{\mathrm{H}}\right) \leq P_{\max}^{(\mathrm{R})}, \tag{12b}$$

$$\mathrm{Tr}\left(\boldsymbol{H}_{\mathrm{RR}} \boldsymbol{G}\left[P_{\max} \boldsymbol{h}_{\mathrm{SR}} \boldsymbol{h}_{\mathrm{SR}}^{\mathrm{H}} + \sigma_{\mathrm{r}}^2 \boldsymbol{I}_{M_{\mathrm{r}}^{(\mathrm{R})}}\right] \boldsymbol{G}^{\mathrm{H}} \boldsymbol{H}_{\mathrm{RR}}^{\mathrm{H}}\right) \leq P_{\mathrm{th}}^{(\mathrm{R})}. \tag{12c}$$

The above problem is neither convex nor can be solved in closed-form. This motivates us to resort to a sub-optimal solution. Inspired by the low-rank nature of our communication scheme (single stream, single antenna end users), in the following we solve (12) by imposing a rank-1 constraint on $\boldsymbol{G}$. In this way, although we lose optimality due to eliminating parts of the feasible set, as it is presented in Section 5, the resulting performance is very close to our baseline performance which is defined in Section 4. We incorporate the rank-1 constraint by assuming without loss of generality ($\boldsymbol{w}_r \in \mathbb{C}^{M_{\mathrm{r}}^{(\mathrm{R})}}$, $\boldsymbol{w}_t \in \mathbb{C}^{M_{\mathrm{t}}^{(\mathrm{R})}}$):

$$\boldsymbol{G} = \boldsymbol{w}_t \boldsymbol{w}_r^{\mathrm{H}},\ \|\boldsymbol{w}_r\|_2 = 1, \tag{13}$$

which redefines our optimization problem as:

$$\max_{\boldsymbol{w}_r, \boldsymbol{w}_t} \quad \frac{P_{\max} \boldsymbol{h}_{\mathrm{RD}}^{\mathrm{T}} \boldsymbol{w}_t \boldsymbol{w}_r^{\mathrm{H}} \boldsymbol{h}_{\mathrm{SR}} \boldsymbol{h}_{\mathrm{SR}}^{\mathrm{H}} \boldsymbol{w}_r \boldsymbol{w}_t^{\mathrm{H}} \boldsymbol{h}_{\mathrm{RD}}^{*}}{\sigma_{\mathrm{r}}^2 \boldsymbol{h}_{\mathrm{RD}}^{\mathrm{T}} \boldsymbol{w}_t \boldsymbol{w}_t^{\mathrm{H}} \boldsymbol{h}_{\mathrm{RD}}^{*} + \sigma_{\mathrm{nd}}^2} \tag{14a}$$

$$\text{s.t.} \quad ||\boldsymbol{w}_r||_2 = 1, \tag{14b}$$

$$\mathrm{Tr}\left(\boldsymbol{w}_t \boldsymbol{w}_r^{\mathrm{H}}\left[P_{\max} \boldsymbol{h}_{\mathrm{SR}} \boldsymbol{h}_{\mathrm{SR}}^{\mathrm{H}} + \sigma_{\mathrm{r}}^2 \boldsymbol{I}_{M_{\mathrm{r}}^{(\mathrm{R})}}\right] \boldsymbol{w}_r \boldsymbol{w}_t^{\mathrm{H}}\right) \leq P_{\max}^{(\mathrm{R})}, \tag{14c}$$

$$\mathrm{Tr}\left(\boldsymbol{H}_{\mathrm{RR}} \boldsymbol{w}_t \boldsymbol{w}_r^{\mathrm{H}}\left[P_{\max} \boldsymbol{h}_{\mathrm{SR}} \boldsymbol{h}_{\mathrm{SR}}^{\mathrm{H}} + \sigma_{\mathrm{r}}^2 \boldsymbol{I}_{M_{\mathrm{r}}^{(\mathrm{R})}}\right] \boldsymbol{w}_r \boldsymbol{w}_t^{\mathrm{H}} \boldsymbol{H}_{\mathrm{RR}}^{\mathrm{H}}\right) \leq P_{\mathrm{th}}^{(\mathrm{R})}. \tag{14d}$$

Proposition 2. *The optimal $\boldsymbol{w}_r$ in (14) is invariant to the choices of $\boldsymbol{w}_t$ and $\boldsymbol{h}_{\mathrm{RD}}$, and obtained as*

$$\boldsymbol{w}_r^{(\mathrm{opt})} = \frac{\boldsymbol{h}_{\mathrm{SR}}}{||\boldsymbol{h}_{\mathrm{SR}}||_2}. \tag{15}$$

Proof See Appendix 1.

Exploiting the result of Proposition 2, we reformulate (14) as

$$\max_{\boldsymbol{w}_t} \quad \frac{P_{\max}||\boldsymbol{h}_{\mathrm{SR}}||_2^2 \boldsymbol{h}_{\mathrm{RD}}^{\mathrm{T}} \boldsymbol{w}_t \boldsymbol{w}_t^{\mathrm{H}} \boldsymbol{h}_{\mathrm{RD}}^*}{\sigma_{\mathrm{r}}^2 \boldsymbol{h}_{\mathrm{RD}}^{\mathrm{T}} \boldsymbol{w}_t \boldsymbol{w}_t^{\mathrm{H}} \boldsymbol{h}_{\mathrm{RD}}^* + \sigma_{\mathrm{nd}}^2} \tag{16a}$$

$$\text{s.t.} \quad \mathrm{Tr}\left(\boldsymbol{w}_t \left[P_{\max}||\boldsymbol{h}_{\mathrm{SR}}||_2^2 + \sigma_{\mathrm{r}}^2\right] \boldsymbol{w}_t^{\mathrm{H}}\right) \leq P_{\max}^{(\mathrm{R})}, \tag{16b}$$

$$\mathrm{Tr}\left(\boldsymbol{H}_{\mathrm{RR}} \boldsymbol{w}_t \left[P_{\max}||\boldsymbol{h}_{\mathrm{SR}}||_2^2 + \sigma_{\mathrm{r}}^2\right] \boldsymbol{w}_t^{\mathrm{H}} \boldsymbol{H}_{\mathrm{RR}}^{\mathrm{H}}\right) \leq P_{\mathrm{th}}^{(\mathrm{R})}. \tag{16c}$$

At this point, by observing the fact that our cost function in (16a) is monotonically increasing with respect to the term $\boldsymbol{h}_{\mathrm{RD}}^{\mathrm{T}} \boldsymbol{w}_t \boldsymbol{w}_t^{\mathrm{H}} \boldsymbol{h}_{\mathrm{RD}}^*$, we write our problem as

$$\max_{\boldsymbol{w}_t} \quad \mathrm{Tr}\left(\boldsymbol{w}_t \boldsymbol{w}_t^{\mathrm{H}} \boldsymbol{h}_{\mathrm{RD}}^* \boldsymbol{h}_{\mathrm{RD}}^{\mathrm{T}}\right) \tag{17a}$$

$$\text{s.t.} \quad \mathrm{Tr}\left(\boldsymbol{w}_t \boldsymbol{w}_t^{\mathrm{H}}\right) \leq P_{\max}^{(\mathrm{R})}/c_0, \tag{17b}$$

$$\mathrm{Tr}\left(\boldsymbol{w}_t \boldsymbol{w}_t^{\mathrm{H}} \boldsymbol{H}_{\mathrm{RR}}^{\mathrm{H}} \boldsymbol{H}_{\mathrm{RR}}\right) \leq P_{\mathrm{th}}^{(\mathrm{R})}/c_0, \tag{17c}$$

where $c_0 := P_{\max}||\boldsymbol{h}_{\mathrm{SR}}||_2^2 + \sigma_{\mathrm{r}}^2$. In order to modify the above problem into an efficient convex structure, we define $\hat{\boldsymbol{Q}} := \boldsymbol{w}_t \boldsymbol{w}_t^{\mathrm{H}}$ and reformulate (17) into

$$\max_{\hat{\boldsymbol{Q}} \in \mathcal{H}} \quad \mathrm{Tr}\left(\bar{\boldsymbol{H}}_{\mathrm{RD}} \hat{\boldsymbol{Q}}\right) \tag{18a}$$

$$\text{s.t.} \quad \mathrm{Tr}\left(\hat{\boldsymbol{Q}}\right) \leq P_{\max}^{(\mathrm{R})}/c_0, \;\; \mathrm{Tr}\left(\bar{\boldsymbol{H}}_{\mathrm{RR}} \hat{\boldsymbol{Q}}\right) \leq P_{\mathrm{th}}^{(\mathrm{R})}/c_0, \tag{18b}$$

where $\bar{\boldsymbol{H}}_{\mathrm{RD}} := \boldsymbol{h}_{\mathrm{RD}}^* \boldsymbol{h}_{\mathrm{RD}}^{\mathrm{T}}$, $\bar{\boldsymbol{H}}_{\mathrm{RR}} := \boldsymbol{H}_{\mathrm{RR}}^{\mathrm{H}} \boldsymbol{H}_{\mathrm{RR}}$. It is clear that any rank-1 optimal solution for $\hat{\boldsymbol{Q}}$ results in an optimal $\boldsymbol{w}_t$. Following Corollary 3.4 in [46], we can always achieve an optimal rank-1 solution for $\hat{\boldsymbol{Q}} \in \mathcal{H}$ for (18), which results in an optimal $\boldsymbol{w}_{\mathrm{t}}$. Consequently, the optimal rank-1 $\boldsymbol{G}$ is obtained as

$$\boldsymbol{w}_t = \hat{\boldsymbol{Q}}^{\frac{1}{2}}, \;\; \boldsymbol{G} = \boldsymbol{w}_{\mathrm{t}} \boldsymbol{w}_r^{\mathrm{H}}. \tag{19}$$

In Section 5 it is illustrated that the performance of the proposed rank-1 solution is very close to the achieved baseline performance.

3.3. General case: $M_{\mathrm{t}} = M_{\mathrm{r}}$, $M_r^{(R)} = M_t^{(R)} =: M_{\mathrm{R}} > 1$

The general optimization problem is formulated as:

$$\max_{\boldsymbol{F}, \boldsymbol{G}} \quad \mathrm{MI}\left(\boldsymbol{x}; \boldsymbol{y}_{\mathbf{D}}\right) \tag{20a}$$

$$\text{s.t.} \quad \mathrm{Tr}\left(\boldsymbol{F}\boldsymbol{F}^{\mathrm{H}}\right) \leq P_{\max}, \tag{20b}$$

$$\mathrm{Tr}\left(\boldsymbol{G}\left[\boldsymbol{H}_{\mathrm{SR}} \boldsymbol{F}\boldsymbol{F}^{\mathrm{H}} \boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}} + \sigma_{\mathrm{r}}^2 \boldsymbol{I}_{M_{\mathrm{R}}}\right] \boldsymbol{G}^{\mathrm{H}}\right) \leq P_{\max}^{(\mathrm{R})}, \tag{20c}$$

$$\mathrm{Tr}\left(\boldsymbol{H}_{\mathrm{RR}} \boldsymbol{G}\left[\boldsymbol{H}_{\mathrm{SR}} \boldsymbol{F}\boldsymbol{F}^{\mathrm{H}} \boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}} + \sigma_{\mathrm{r}}^2 \boldsymbol{I}_{M_{\mathrm{R}}}\right] \boldsymbol{G}^{\mathrm{H}} \boldsymbol{H}_{\mathrm{RR}}^{\mathrm{H}}\right) \leq P_{\mathrm{th}}^{(\mathrm{R})}, \tag{20d}$$

where $\mathrm{MI}\left(\boldsymbol{x};\boldsymbol{y_D}\right)$ is defined in (6). Once again, we are dealing with a non-convex optimization problem. Concerning computational complexity, we propose two sub-optimal solutions in the following.

3.3.1. Alternating Optimization for self-interference aware FD relaying (AO)

In order to deal with the non-convex problem (20), we choose an iterative optimization over the precoder ($\boldsymbol{F}$) and the relay amplification matrix ($\boldsymbol{G}$). In each step, one of the $\boldsymbol{F}$ or $\boldsymbol{G}$ is optimized assuming that the other matrix is fixed. This process is continued until a stable pair of $\boldsymbol{F},\boldsymbol{G}$ is achieved. As the first iteration we solve our optimization problem over $\boldsymbol{F}$ assuming a fixed $\boldsymbol{G}$:

$$\max_{\boldsymbol{F}} \quad \mathrm{MI}\left(\boldsymbol{x};\boldsymbol{y_D}\right) \tag{21a}$$

$$\text{s.t.} \quad \mathrm{Tr}\left(\boldsymbol{F}\boldsymbol{F}^{\mathrm{H}}\right) \leq P_{\max}, \tag{21b}$$

$$\mathrm{Tr}\left(\boldsymbol{G}\left[\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{F}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}+\sigma_{\mathrm{r}}^{2}\boldsymbol{I}_{M_{\mathrm{R}}}\right]\boldsymbol{G}^{\mathrm{H}}\right) \leq P_{\max}^{(\mathrm{R})}, \tag{21c}$$

$$\mathrm{Tr}\left(\boldsymbol{H}_{\mathrm{RR}}\boldsymbol{G}\left[\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{F}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}+\sigma_{\mathrm{r}}^{2}\boldsymbol{I}_{M_{\mathrm{R}}}\right]\boldsymbol{G}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{RR}}^{\mathrm{H}}\right) \leq P_{\mathrm{th}}^{(\mathrm{R})}. \tag{21d}$$

This problem can be equivalently formulated as a convex optimization problem by defining $\boldsymbol{Q} := \boldsymbol{F}\boldsymbol{F}^{\mathrm{H}}$ as

$$\max_{\boldsymbol{Q}\in\mathcal{H}} \quad \log_2\left|\boldsymbol{B}_1\boldsymbol{Q}\boldsymbol{B}_1^{\mathrm{H}}+\boldsymbol{B}_{\mathrm{n}}\right| \tag{22a}$$

$$\text{s.t.} \quad \mathrm{Tr}\left(\boldsymbol{Q}\right) \leq P_{\max}, \tag{22b}$$

$$\mathrm{Tr}\left(\boldsymbol{B}_2\boldsymbol{Q}\boldsymbol{B}_2^{\mathrm{H}}\right) \leq \xi_1,\ \mathrm{Tr}\left(\boldsymbol{B}_3\boldsymbol{Q}\boldsymbol{B}_3^{\mathrm{H}}\right) \leq \xi_2, \tag{22c}$$

where:

$$\boldsymbol{B}_1 := \boldsymbol{H}_{\mathrm{RD}}\boldsymbol{G}\boldsymbol{H}_{\mathrm{SR}},\ \boldsymbol{B}_2 := \boldsymbol{G}\boldsymbol{H}_{\mathrm{SR}}, \tag{22d}$$

$$\boldsymbol{B}_3 := \boldsymbol{H}_{\mathrm{RR}}\boldsymbol{G}\boldsymbol{H}_{\mathrm{SR}},\ \boldsymbol{B}_{\mathrm{n}} := \sigma_{\mathrm{r}}^{2}\boldsymbol{H}_{\mathrm{RD}}\boldsymbol{G}\boldsymbol{G}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{RD}}^{\mathrm{H}}+\sigma_{\mathrm{nd}}^{2}\boldsymbol{I}_{M_r}, \tag{22e}$$

$$\xi_1 := P_{\max}^{(\mathrm{R})}-\mathrm{Tr}\left(\sigma_{\mathrm{r}}^{2}\boldsymbol{G}\boldsymbol{G}^{\mathrm{H}}\right),\ \xi_2 := P_{\mathrm{th}}^{(\mathrm{R})}-\mathrm{Tr}\left(\sigma_{\mathrm{r}}^{2}\boldsymbol{H}_{\mathrm{RR}}\boldsymbol{G}\boldsymbol{G}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{RR}}^{\mathrm{H}}\right). \tag{22f}$$

The above problem possesses a convex structure. Therefore, we can get an optimal $\boldsymbol{Q}\in\mathcal{H}$, and consequently an optimal $\boldsymbol{F}$ as $\boldsymbol{F}=\boldsymbol{Q}^{\frac{1}{2}}$. The next step is to compute an optimal $\boldsymbol{G}$, given the precoder $\boldsymbol{F}$. The corresponding problem is formulated as

$$\max_{\boldsymbol{G}} \quad \mathrm{MI}\left(\boldsymbol{x};\boldsymbol{y_D}\right) \tag{23a}$$

$$\text{s.t.} \quad \mathrm{Tr}\left(\boldsymbol{G}\left[\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{F}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}+\sigma_{\mathrm{r}}^{2}\boldsymbol{I}_{M_{\mathrm{R}}}\right]\boldsymbol{G}^{\mathrm{H}}\right) \leq P_{\max}^{(\mathrm{R})}, \tag{23b}$$

$$\mathrm{Tr}\left(\boldsymbol{H}_{\mathrm{RR}}\boldsymbol{G}\left[\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{F}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}+\sigma_{\mathrm{r}}^{2}\boldsymbol{I}_{M_{\mathrm{R}}}\right]\boldsymbol{G}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{RR}}^{\mathrm{H}}\right) \leq P_{\mathrm{th}}^{(\mathrm{R})}. \tag{23c}$$

The transmit power constraint is dropped because it is not a function of $\boldsymbol{F}$. Note that at the optimality of problem (23a-c) at least one of the constraints has to be active, i.e., the constraint is satisfied with equality. Otherwise, we can scale $\boldsymbol{G}$ up such that the optimal value increases. This contradicts the optimality. This motives us to solve problem (23a-c) by solving three sub-problems, i.e.,

one of the constraints ((23b) or (23c)) is active, or both constraints are active. At this point we recall that the KKT conditions have been studied in [30] for the problem defined by (23a-b). Inspired by [30], we first consider the special case that our system is dominated by the level of allowed transmit power and thus the interference power constraint can be dropped:

$$\max_{\boldsymbol{G}} \quad \mathrm{MI}\left(\boldsymbol{x};\boldsymbol{y_D}\right) \tag{24a}$$

$$\text{s.t.} \quad \mathrm{Tr}\left(\boldsymbol{G}\left[\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{F}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}+\sigma_{\mathrm{r}}^{2}\boldsymbol{I}_{M_{\mathrm{R}}}\right]\boldsymbol{G}^{\mathrm{H}}\right)\leq P_{\max}^{(\mathrm{R})}, \tag{24b}$$

which results in an optimal $\boldsymbol{G}$ via the solution in [30]. If the resulting $\boldsymbol{G}$ does not violate the self-interference constraint, we have already obtained an optimal relay amplification matrix for (23). If the resulting $\boldsymbol{G}$ is not feasible, we then investigate the other extreme case in which our problem is dominated by the allowed interference power. Our optimization problem turns into

$$\max_{\boldsymbol{G}} \quad \mathrm{MI}\left(\boldsymbol{x};\boldsymbol{y_D}\right) \tag{25a}$$

$$\text{s.t.} \quad \mathrm{Tr}\left(\boldsymbol{H}_{\mathrm{RR}}\boldsymbol{G}\left[\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{F}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}+\sigma_{\mathrm{r}}^{2}\boldsymbol{I}_{M_{\mathrm{R}}}\right]\boldsymbol{G}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{RR}}^{\mathrm{H}}\right)\leq P_{\mathrm{th}}^{(\mathrm{R})}, \tag{25b}$$

where the self-interference power constraint (25b) is active at the optimality[8]. By defining the following auxiliary variables:

$$\boldsymbol{G}^{'}:=\boldsymbol{H}_{\mathrm{RR}}\boldsymbol{G},\ \ \boldsymbol{H}_{\mathrm{RD}}^{'}:=\boldsymbol{H}_{\mathrm{RD}}\boldsymbol{H}_{\mathrm{RR}}^{-1}, \tag{26}$$

we turn (25) into

$$\max_{\boldsymbol{G}^{'}} \quad \log_2\left|\boldsymbol{H}_{\mathrm{RD}}^{'}\boldsymbol{G}^{'}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{F}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\boldsymbol{G}^{'\mathrm{H}}\boldsymbol{H}_{\mathrm{RD}}^{'\mathrm{H}}+\sigma_{\mathrm{r}}^{2}\boldsymbol{H}_{\mathrm{RD}}^{'}\boldsymbol{G}^{'}\boldsymbol{G}^{'\mathrm{H}}\boldsymbol{H}_{\mathrm{RD}}^{'\mathrm{H}}+\sigma_{\mathrm{nd}}^{2}\boldsymbol{I}_{M_r}\right|$$
$$-\log_2\left|\sigma_{\mathrm{r}}^{2}\boldsymbol{H}_{\mathrm{RD}}^{'}\boldsymbol{G}^{'}\boldsymbol{G}^{'\mathrm{H}}\boldsymbol{H}_{\mathrm{RD}}^{'\mathrm{H}}+\sigma_{\mathrm{nd}}^{2}\boldsymbol{I}_{M_r}\right|=\mathrm{MI}^{'}\left(\boldsymbol{x};\boldsymbol{y_D}\right) \tag{27a}$$

$$\text{s.t.} \quad \mathrm{Tr}\left(\boldsymbol{G}^{'}\left[\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{F}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}+\sigma_{\mathrm{r}}^{2}\boldsymbol{I}_{M_{\mathrm{R}}}\right]\boldsymbol{G}^{'\mathrm{H}}\right)\leq P_{\mathrm{th}}^{(\mathrm{R})}, \tag{27b}$$

where $\mathrm{MI}^{'}\left(\boldsymbol{x};\boldsymbol{y_D}\right)$ represents the mutual information between $\boldsymbol{x},\boldsymbol{y}_{\mathrm{D}}$ given $\boldsymbol{H}_{\mathrm{RD}}^{'}$, $\boldsymbol{G}^{'},\boldsymbol{F},\boldsymbol{H}_{\mathrm{SR}}$. The above problem follows the same structure as problem (24) and an optimal $\boldsymbol{G}^{'}$ can be obtained using the iterative method in [30]. The optimum $\boldsymbol{G}$ can be then computed as

$$\boldsymbol{G}=\boldsymbol{H}_{\mathrm{RR}}^{-1}\boldsymbol{G}^{'}, \tag{28}$$

which is the inverse representation of (26). Once again, we check if the solution is feasible. If the transmit power constraint is not violated, the achieved $\boldsymbol{G}$

[8]Please note that the constraint (24b) is active at the optimality of (24a-b). Otherwise, we can scale $\boldsymbol{G}$ up such that the optimal value increases, which contradicts with optimality. The same argument holds for the constraint (25b) in (25a-b), which represents the case where the self-interference power constraint is active.

is optimal for (23). If the resulting $\boldsymbol{G}$ is not feasible, we resort to a suboptimal solution for $\boldsymbol{G}$, where both constraints (23b) and (23c) are satisfied with equality[9].

In order to design our sub-optimal approach, we recall that the solutions to (24) and (25) exploit the allowed transmit and interference power in an optimal fashion, respectively. Hence, when both of the constraints are active, the choice of $\boldsymbol{G}$ must be a trade-off between the results from (24) and (25). Therefore we define

$$\boldsymbol{\Gamma} := \boldsymbol{G}\left[\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{F}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}} + \sigma_{\mathrm{r}}^{2}\boldsymbol{I}_{M_{\mathrm{R}}}\right]\boldsymbol{G}^{\mathrm{H}}, \tag{29}$$

where $\boldsymbol{\Gamma}$ represents the resulting relay's transmit covariance matrix. The matrix $\boldsymbol{G}$ can then be calculated as

$$\boldsymbol{G} = \boldsymbol{\Gamma}^{\frac{1}{2}}\left[\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{F}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}} + \sigma_{\mathrm{r}}^{2}\boldsymbol{I}_{M_{\mathrm{R}}}\right]^{-\frac{1}{2}}. \tag{30}$$

Let the relay transmit covariance matrix obtained by solving (24) and (25) be $\boldsymbol{\Gamma}_1$ and $\boldsymbol{\Gamma}_2$, respectively. We look for a linear combination of them ($\boldsymbol{\Gamma}^{\star}$) in a way that both constraints in (23b) and (23c) are satisfied with equality:

$$\boldsymbol{\Gamma}^{\star} = \alpha_1\boldsymbol{\Gamma}_1 + \alpha_2\boldsymbol{\Gamma}_2,\ \alpha_1, \alpha_2 \in \mathbb{R}, \tag{31}$$

where α_1 and α_2 are the corresponding weights which should be determined. Then we have

$$\begin{aligned}
&\mathrm{Tr}\left(\boldsymbol{\Gamma}^{\star}\right) = \mathrm{Tr}\left(\alpha_1\boldsymbol{\Gamma}_1 + \alpha_2\boldsymbol{\Gamma}_2\right) = P_{\max}^{(\mathrm{R})} \\
&\Rightarrow \alpha_1\mathrm{Tr}\left(\boldsymbol{\Gamma}_1\right) + \alpha_2\mathrm{Tr}\left(\boldsymbol{\Gamma}_2\right) = P_{\max}^{(\mathrm{R})}, &\text{(32a)}\\
&\mathrm{Tr}\left(\bar{\boldsymbol{H}}_{\mathrm{RR}}\boldsymbol{\Gamma}^{\star}\right) = \mathrm{Tr}\left(\alpha_1\bar{\boldsymbol{H}}_{\mathrm{RR}}\boldsymbol{\Gamma}_1 + \alpha_2\bar{\boldsymbol{H}}_{\mathrm{RR}}\boldsymbol{\Gamma}_2\right) \\
&\Rightarrow \alpha_1\mathrm{Tr}\left(\bar{\boldsymbol{H}}_{\mathrm{RR}}\boldsymbol{\Gamma}_1\right) + \alpha_2\mathrm{Tr}\left(\bar{\boldsymbol{H}}_{\mathrm{RR}}\boldsymbol{\Gamma}_2\right) = P_{\mathrm{th}}^{(\mathrm{R})}. &\text{(32b)}
\end{aligned}$$

where $\bar{\boldsymbol{H}}_{\mathrm{RR}} := \boldsymbol{H}_{\mathrm{RR}}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{RR}}$. By defining the auxiliary constants ($\beta_1, \beta_2 > 1$) as

$$\beta_1 := \frac{\mathrm{Tr}\left(\boldsymbol{\Gamma}_2\right)}{P_{\max}^{(\mathrm{R})}},\ \ \beta_2 := \frac{\mathrm{Tr}\left(\bar{\boldsymbol{H}}_{\mathrm{RR}}\boldsymbol{\Gamma}_1\right)}{P_{\mathrm{th}}^{(\mathrm{R})}}, \tag{33}$$

and applying (33) to (32) we have:

$$\begin{aligned}
&\alpha_1 + \alpha_2\beta_1 = 1,\ \ \alpha_2 + \alpha_1\beta_2 = 1, &\text{(34a)}\\
&\Rightarrow \alpha_1 = \frac{\beta_1 - 1}{\beta_1\beta_2 - 1},\ \ \alpha_2 = \frac{\beta_2 - 1}{\beta_2\beta_1 - 1}. &\text{(34b)}
\end{aligned}$$

[9]Other primitive sub-optimal solutions would be the scaled versions of the solutions to (24) and (25), where the relay's transmit power constraint or the self-interference constraint is active, respectively. The case with two active constraints is studied in (29)-(34) using a low complexity solution.

Then by using (31) and (30) we obtain the desired $\boldsymbol{\Gamma}^{\star}$ and $\boldsymbol{G}$. Note that the resulting $\boldsymbol{G}$ is sub-optimal and satisfies both the relay's transmit power constraint and the interference power constraint. This procedure, i.e., the alternating design of $\boldsymbol{F}$ and $\boldsymbol{G}$, is continued until no further improvement is observed in the objective value (20a). The proposed AO method is summarized in Algorithm 1.

The proposed AO algorithm can be computationally inefficient, especially when none of the basic power constraints dominates. This motivates us to calculate $\boldsymbol{F}$ and $\boldsymbol{G}$ sequentially. The proposed sequential method is also used as the initialization step for the AO algorithm. Fig. 7 illustrates the convergence speed of the proposed AO algorithm under different system parameters.

3.3.2. The Reduced-Complexity AO (RC-AO) design of $\boldsymbol{F}$ and $\boldsymbol{G}$

This method is composed of two separated designs for the precoder and the relay amplification matrix. In the first stage, the precoder is designed to maximize the mutual information between the transmitter and the relay node. This is performed via optimal power allocation on the singular modes of $\boldsymbol{H}_{\mathrm{SR}}$, i.e., by using the water-filling algorithm in [44]. In the second stage, given the resulting $\boldsymbol{F}$, we design $\boldsymbol{G}$ by applying the proposed steps in (23)-(34). In Section 5, we observe that the performance of this method is close to that of the alternating AO solution when the system dynamic range is very low or very high.

4. A unified design via gradient projections

As we have observed in the previous sections, due to the non-convex nature of our problem, we have to resort to sub-optimal methods in parts of our solutions. In order to provide a baseline for the performance evaluation, we introduce an iterative gradient projection based optimization to find a jointly stable point for $\boldsymbol{F}$ and $\boldsymbol{G}$. Although we may not guarantee the global optimality via a gradient projection-based approach, we will ensure that the optimum performance is not far from the achieved one by repeating the procedure with multiple initialization points. Please note that the provided design in this part is applicable to a general setup, regardless of the number of antennas at the different nodes. We begin our method by calculating the gradient of the cost function with respect to $\boldsymbol{F}^{*}, \boldsymbol{G}^{*}$ (see Appendix 3):

$$\frac{\partial \mathrm{MI}\left(\boldsymbol{x};\boldsymbol{y}_{\mathbf{D}}\right)}{\partial \boldsymbol{G}^{*}} = \frac{1}{\ln(2)}\Big\{\boldsymbol{H}_{\mathrm{RD}}^{\mathrm{H}}\boldsymbol{A}_{1}^{-1}\boldsymbol{H}_{\mathrm{RD}}\boldsymbol{G}\left[\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{F}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}} + \sigma_{\mathrm{r}}^{2}\boldsymbol{I}_{M_{\mathrm{r}}^{(\mathrm{R})}}\right] - \sigma_{\mathrm{r}}^{2}\boldsymbol{H}_{\mathrm{RD}}^{\mathrm{H}}\boldsymbol{A}_{2}^{-1}\boldsymbol{H}_{\mathrm{RD}}\boldsymbol{G}\Big\}, \tag{35a}$$

$$\frac{\partial \mathrm{MI}\left(\boldsymbol{x};\boldsymbol{y}_{\mathbf{D}}\right)}{\partial \boldsymbol{F}^{*}} = \frac{1}{\ln(2)}\Big\{\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{RD}}^{\mathrm{H}}\boldsymbol{A}_{1}^{-1}\boldsymbol{H}_{\mathrm{RD}}\boldsymbol{G}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{F}\Big\}, \tag{35b}$$

where $\boldsymbol{A}_{2} := \sigma_{\mathrm{r}}^{2}\boldsymbol{H}_{\mathrm{RD}}\boldsymbol{G}\boldsymbol{G}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{RD}}^{\mathrm{H}} + \sigma_{\mathrm{nd}}^{2}\boldsymbol{I}_{M_{r}}$ and $\boldsymbol{A}_{1} := \boldsymbol{H}_{\mathrm{RD}}\boldsymbol{G}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{F}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{RD}}^{\mathrm{H}} + \boldsymbol{A}_{2}$. The natural logarithm is represented by $\ln(\cdot)$. Using the above gradients, we

define the update equations as follows (in each iteration only one of the matrices is updated):

$$\Delta \boldsymbol{F} = \frac{\partial \mathrm{MI}\left(\boldsymbol{x};\boldsymbol{y}_{\mathbf{D}}\right)}{\partial \boldsymbol{F}^{*}} \cdot \delta_{f} \ , \ \boldsymbol{F}_{\mathrm{new}} = \boldsymbol{F}_{\mathrm{old}} + \Delta \boldsymbol{F}, \tag{36a}$$

$$\Delta \boldsymbol{G} = \frac{\partial \mathrm{MI}\left(\boldsymbol{x};\boldsymbol{y}_{\mathbf{D}}\right)}{\partial \boldsymbol{G}^{*}} \cdot \delta_{g} \ , \ \boldsymbol{G}_{\mathrm{new}} = \boldsymbol{G}_{\mathrm{old}} + \Delta \boldsymbol{G}, \tag{36b}$$

where $\delta_f, \delta_g \in \mathbb{R}$ represent the step-size values which are determined in each iteration by a line search based on Armijo's step-size rule [47].

For the defined update matrices, we must guarantee that the corresponding update does not violate the power constraints (20b-d). In order to fulfill this requirement, we calculate the projection matrix to the non-power-increasing region (for both $\Delta \boldsymbol{F}$ and $\Delta \boldsymbol{G}$). In this way, once an update matrix defined by (36) violates one of our power constraints, we project it to the feasible region. In order to derive the projection matrices we first formulate the user and relay's transmit and interference power values as a function of $\boldsymbol{F}$ and $\boldsymbol{G}$:

$$P^{(\mathrm{Relay})} = \mathrm{Tr}\left\{\boldsymbol{G}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{F}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}} + \sigma_{\mathrm{r}}^{2}\boldsymbol{G}\boldsymbol{G}^{\mathrm{H}}\right\}, \tag{37a}$$

$$P^{(\mathrm{Tx})} = \mathrm{Tr}\left\{\boldsymbol{F}\boldsymbol{F}^{\mathrm{H}}\right\}, \tag{37b}$$

$$P^{(\mathrm{Int})} = \mathrm{Tr}\left\{\boldsymbol{H}_{\mathrm{RR}}\boldsymbol{G}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{F}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{RR}}^{\mathrm{H}} + \sigma_{\mathrm{r}}^{2}\boldsymbol{H}_{\mathrm{RR}}\boldsymbol{G}\boldsymbol{G}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{RR}}^{\mathrm{H}}\right\}, \tag{37c}$$

where $P^{(\mathrm{Relay})}, P^{(\mathrm{Tx})}, P^{(\mathrm{Int})}$ represent the relay's actual transmit power, the transmit power from the source, and the interference power, respectively. Using simple matrix calculations, it is possible to calculate the gradients of the formulated powers with respect to $\boldsymbol{F}$ and $\boldsymbol{G}$ as follows: (see Appendix 3)

$$\frac{\partial P^{(\mathrm{Tx})}}{\partial \boldsymbol{F}^{*}} = \boldsymbol{F}, \tag{38a}$$

$$\frac{\partial P^{(\mathrm{Tx})}}{\partial \boldsymbol{G}^{*}} = \boldsymbol{0}, \tag{38b}$$

$$\frac{\partial P^{(\mathrm{Relay})}}{\partial \boldsymbol{F}^{*}} = \boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}}\boldsymbol{G}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{F}, \tag{38c}$$

$$\frac{\partial P^{(\mathrm{Relay})}}{\partial \boldsymbol{G}^{*}} = \boldsymbol{G}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{F}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}} + \sigma_{\mathrm{r}}^{2}\boldsymbol{G}, \tag{38d}$$

$$\frac{\partial P^{(\mathrm{Int})}}{\partial \boldsymbol{F}^{*}} = \boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{RR}}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{RR}}\boldsymbol{G}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{F}, \tag{38e}$$

$$\frac{\partial P^{(\mathrm{Int})}}{\partial \boldsymbol{G}^{*}} = \boldsymbol{H}_{\mathrm{RR}}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{RR}}\boldsymbol{G}\left[\sigma_{\mathrm{r}}^{2}\boldsymbol{I}_{M_{\mathrm{r}}^{(\mathrm{R})}} + \boldsymbol{H}_{\mathrm{SR}}\boldsymbol{F}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\right]. \tag{38f}$$

Exploiting the calculated gradients, we calculate the desired projection matrices

and the corresponding projections as follows:

$$\boldsymbol{\Pi}_g^{(C)} := \boldsymbol{I}_{M_{\mathrm{r}}^{(\mathrm{R})} M_{\mathrm{t}}^{(\mathrm{R})}} - \frac{\bar{\boldsymbol{p}}_g^{(C)} \bar{\boldsymbol{p}}_g^{(C)\mathrm{H}}}{\bar{\boldsymbol{p}}_g^{(C)\mathrm{H}} \bar{\boldsymbol{p}}_g^{(C)}}, \tag{39a}$$

$$\boldsymbol{\Pi}_f^{(C)} := \boldsymbol{I}_{M_t M_t} - \frac{\bar{\boldsymbol{p}}_f^{(C)} \bar{\boldsymbol{p}}_f^{(C)\mathrm{H}}}{\bar{\boldsymbol{p}}_f^{(C)\mathrm{H}} \bar{\boldsymbol{p}}_f^{(C)}}, \tag{39b}$$

$$\mathrm{vec}\left(\Delta \boldsymbol{F}^{(\mathrm{Feasible})}\right) = \boldsymbol{\Pi}_f^{(C)} \mathrm{vec}\left(\Delta \boldsymbol{F}^{(\mathrm{Infeasible})}\right), \tag{39c}$$

$$\mathrm{vec}\left(\Delta \boldsymbol{G}^{(\mathrm{Feasible})}\right) = \boldsymbol{\Pi}_g^{(C)} \mathrm{vec}\left(\Delta \boldsymbol{G}^{(\mathrm{Infeasible})}\right), \tag{39d}$$

where $\bar{\boldsymbol{p}}_g^{(C)} := \mathrm{vec}\left(\frac{\partial P^{(C)}}{\partial \boldsymbol{G}^*}\right)$, $\bar{\boldsymbol{p}}_f^{(C)} := \mathrm{vec}\left(\frac{\partial P^{(C)}}{\partial \boldsymbol{F}^*}\right)$. The projection matrices are denoted as $\boldsymbol{\Pi}$, and $C \in \{\mathrm{Relay}, \mathrm{Int}, \mathrm{Tx}\}$ refers to the power constraints corresponding to (37a), (37b), (37c), respectively. This update procedure (35-39) is continued until a stable point is achieved and repeated for multiple initial choices of $\boldsymbol{F}, \boldsymbol{G}$ to avoid (with higher confidence) the convergence into local stationary points.

5. Simulation Results

In this section we evaluate the derived transmit strategies using Monte Carlo simulations. For a FD device with a very high self-interference cancellation capability, the traditional HD transmit strategies are optimal in the FD case. In the following we study how a FD-specific design helps our system when the self-interference cancellation quality is limited. An un-correlated Rayleigh flat-fading channel model is used, where $\rho_{\mathrm{SR}}, \rho_{\mathrm{RR}}, \rho_{\mathrm{RD}} \in \mathbb{R}^+$ represent the variance of the source-to-relay, relay-to-relay (self-interference), and relay-to-destination channel coefficients, respectively. We average the simulation results over 1000 channel realizations. We evaluate the effects of different noise and residual interference levels ($\sigma_{\mathrm{n}}^2 := \sigma_{\mathrm{nr}}^2 = \sigma_{\mathrm{nd}}^2$, $P_{\max} = 0$ dBW, and $\mathrm{SNR} := \frac{P_{\max}}{\sigma_{\mathrm{n}}^2}$), the effect of the strength of the interference channel (ρ_{RR}), the maximum allowed self-interference power that can be handled by the implemented cancellation scheme ($\epsilon := \frac{P_{\mathrm{th}}^{(\mathrm{R})}}{P_{\max}}$), the level of inevitable residual self-interference after self-interference cancellation ($\mathrm{SIR} := \frac{P_{\max}}{\sigma_{\mathrm{si}}^2}$), and the numbers of antenna elements at the users ($M_{\mathrm{t}}, M_{\mathrm{r}}$) and at the relay node (M_{R}). In each case, we compare the resulting performance with that of the traditional HD setup (with the same number of transmit and receive chains as in the FD scenario) to illustrate the FD gain in different scenarios.

5.1. Single antenna users, multiple antenna relay

For the case with single antenna users, our simulations include the proposed method with an optimal rank-1 relay amplification matrix (FD-SDR), the gradient based optimization in Section 4 (FD-GP), and the HD precoding method

which is a scaled version of the traditional optimal HD transmit strategy such that the self-interference constraint is satisfied with equality (FD-ScaledHD).

Fig. 2 and Fig. 3 demonstrate the achievable sum rate as a function of the SNR. The proposed rank-1 solution (FD-SDR) performs very close to the general rank solution based on gradient projection (FD-GP) for all noise levels. This convergence is particularly observed in Fig. 3 for different number of antennas at the relay. Furthermore, a significant FD gain in terms of sum rate can be achieved in the high SNR regime or when the number of antennas at the relay is large. Moreover, all the proposed methods outperform the scaled version of the optimal HD transmit strategy.

In Fig. 4 and Fig. 5 we show the achievable sum rate of different algorithms as a function of dynamic range and the interference channel intensity, respectively. Fig. 4 illustrates that a two-fold gain can be obtained if the dynamic range is large. Moreover, as the number of antennas at the relay increases, the difference between the proposed solutions and the scaled HD solution increases. It is clear that as the interference channel intensity increases, a higher dynamic range is needed at the relay to provide the same performance. The same performance trend is observable in Fig. 5. The achievable FD gain increases as the interference channel intensity becomes small. If the number of the antennas at the relay is large, a larger FD gain is obtained. In both figures the rank-1 beamformer matches the gradient projection based method.

Fig. 6 demonstrates the achievable sum rate as a function of the SIR. Similar to Figs. 2, 4 and 5, the proposed rank-1 solution (FD-SDR) performs very close to the general rank solution based on gradient projection (FD-GP) for all SIR levels. Furthermore, it is observable that a FD gain in terms of sum rate can be only achieved if the residual interference power is much smaller than the noise. For large values of the SIR, the FD system performance is dominated by the noise and no significant change is observable. Since the HD setup does not suffer from residual self-interference, its performance remains constant for all SIR values.

5.2. Multiple antenna users, multiple antenna relay

When both the users and the relay have multiple antennas, we compare the performance of the proposed AO algorithm in Section 3 (FD-AO), the sequential method RC-AO in Section 3 (FD-RC-AO), the gradient based algorithm (FD-GP) in Section 4, and the scaled version of optimal HD design (FD-ScaledHD). The performance of the baseline scenario, i.e., the performance of an equivalent HD system (HD) is also included.

In Fig. 7 the required number of iterations for the AO algorithm to obtain a stable pair of the user precoder and the relay amplification matrix is plotted. It is observed that the algorithm provides a faster convergence speed when the limited dynamic ranges are very high or very low. These correspond to the cases where one of the constraints dominates, i.e., only (24) or (25) needs to be solved. In this case, only a few iterations are needed. Furthermore, a larger number of antennas results in a slightly slower convergence as the dimension of the solution space increases.

Fig. 8 demonstrates the achievable sum rate of different algorithms when the SNR values vary. The gradient projection based method provides the best performance. The AO has a performance close to the gradient projection algorithm especially when the number of the antennas at the relay is small. The low complexity method RC-AO performs almost the same as the AO method regardless of the number of antennas at the relay. Again, the proposed algorithms outperform the scaled optimal HD design. Interestingly, compared to the HD baseline scenario, a significant FD gain is always obtained. This suggests that when all the nodes have multiple antennas the system gains more from a FD operation at the relay.

Fig. 9 and Fig. 10 show the system sum rate as a function of the limited dynamic range and the number of antennas at the relay, respectively. As seen in Fig. 9, when the number of antennas at the relay is large, the scaled optimal HD design suffers more from a low dynamic range value compared to the other algorithms. Moreover, the low-complexity RC-AO algorithm provides a good trade-off between the performance and the computational complexity. However, it suffers from a little loss when the limited dynamic range is large. Fig. 10 illustrates that the gradient projection based method provides the best performance among the proposed algorithms. The difference between the performance of the gradient projection algorithm and the other algorithms increases as M_{R} increases. However, the gradient projection based solution has the worst computational complexity as illustrated in the next section. Again, a two-fold FD gain is obtained when the limited dynamic range is high or when the number of antennas at the relay is large.

Fig. 11 demonstrates the achievable sum rate as a function of the SIR. Similar to Fig. 6, it is observable that a FD gain in terms of sum rate can only be achieved if the residual interference power is much smaller than the noise. Furthermore, as the number of antennas increases, the FD setup outperforms the equivalent HD setup with a smaller SIR. This is expected as the signal space can be better separated from the residual interference as the signal dimension increases. For large values of the SIR, the FD system performance is dominated by the noise and no significant change is observable. The performance of the HD setup remains constant for every SIR as it does not suffer from the residual self-interference.

5.3. Computational complexity

In this part we discuss the computational complexity of the proposed methods. As the overall computational complexity of the iterative algorithms depends also on their convergence speed which is not known in general, we use the following parameters so that the computational burden of all the proposed algorithms are visualized in a better way. We denote the required number of iterations for the gradient projection based algorithm (FD-GP) to achieve a stable point as ζ_1, while ζ_2 represents the number of initial points for which the gradient search is repeated. The required number of iterations for the AO method is denoted as ζ_3 and the number of required line search steps in (36a) and (36b) is denoted as ζ_4. Furthermore, the required number of iterations for

the convex solver to achieve the desired numerical accuracy for problems (18) and (22) are denoted as L_1 and L_2, respectively. Note that the values of these parameters, i.e., $\zeta_1, \zeta_3, L_1, L_2$, are highly dependent on the desired numerical accuracy. A more detailed analysis of the relationship between the numerical accuracy and the number of required iterations is found in [44, 47]. But it is out of the scope of this paper. Tables 1-4 present the computational complexity of the proposed algorithms using both the approximate number of required floating point operations and the required CPU time[10]. Clearly, among the proposed algorithms, the gradient projection algorithm has the highest computational complexity. When all the nodes have multiple antennas, the RC-AO algorithm requires the lowest CPU time.

6. Conclusion

In this work we focus on FD enhancements for a relay-assisted communication environment which allows the HD operation of the end users (HD-compatible FD relaying). Enhanced transmit strategies for the users as well as the relay node are derived to maximize the system sum rate. For the case that all nodes have a single antenna, the optimal power levels for the transmitter and the relay nodes are derived. For the scenario with a single antenna user and a multi-antenna relay, we have obtained an optimal rank-1 relay amplification matrix which performs very close to the optimal solution using gradient projection techniques. When all the nodes are equipped with multiple antennas, we have developed the alternating optimization for FD relaying with self-interference (AO) algorithm to iteratively design the relay amplification matrix and the beamformers at the user terminals. Alternatively, we have also proposed a low complexity and independent design via the reduced-complexity AO (RC-AO) algorithm. Moreover, a unified approach via gradient projections is proposed as a benchmark. It can be applied regardless of the number of antennas at each node at the expense of significantly higher computational complexity than the AO algorithm. The performance of all the proposed methods and the corresponding computational complexities are then compared to the proposed gradient-based optimization via numerical simulations. It is observed that while the system performance suffers as the self-interference suppression quality decreases, the proposed designs provide a decent robustness against this degradation. As expected, we observe a gain in terms of the system sum rate by a factor of two, if the relay has a high self-interference cancellation capability.

[10]Please note that the values for the required CPU time enable a comparison between the computational complexity for different algorithms. Nevertheless they may vary depending on the used software/hardware simulation platforms. Our numerical results are obtained using an Intel Core i5-3320M processor with the clock rate of 2.6 GHz and 8 GB of random-access memory (RAM). As our software platform we have used MATLAB 2013a, on a 64-bit operating system.

Appendices

Appendix 1: Proof to Proposition 2

Let $\bar{\boldsymbol{w}}_t, \bar{\boldsymbol{w}}_r$ be an optimal solution for (14). We define the following variable update

$$\boldsymbol{w}_r^\star = \frac{\boldsymbol{h}_{\mathrm{SR}}}{||\boldsymbol{h}_{\mathrm{SR}}||_2}, \quad \boldsymbol{w}_t^\star = \bar{\boldsymbol{w}}_t \cdot \frac{1}{\sqrt{\psi}}, \tag{41a}$$

$$\psi := \frac{\boldsymbol{w}_r^{\star\mathrm{H}} \boldsymbol{h}_{\mathrm{SR}} \boldsymbol{h}_{\mathrm{SR}}^{\mathrm{H}} \boldsymbol{w}_r^\star}{\bar{\boldsymbol{w}}_r{}^{\mathrm{H}} \boldsymbol{h}_{\mathrm{SR}} \boldsymbol{h}_{\mathrm{SR}}^{\mathrm{H}} \bar{\boldsymbol{w}}_r} = \frac{||\boldsymbol{h}_{\mathrm{SR}}||_2^2}{\bar{\boldsymbol{w}}_r{}^{\mathrm{H}} \boldsymbol{h}_{\mathrm{SR}} \boldsymbol{h}_{\mathrm{SR}}^{\mathrm{H}} \bar{\boldsymbol{w}}_r}, \tag{41b}$$

where $\psi \geq 1$ is observable from (41b) based on the fact that $\|\bar{\boldsymbol{w}}_r\|_2 = 1$, and hence the denominator is upper-bounded by $||\boldsymbol{h}_{\mathrm{SR}}||_2^2$. In the following we study the effect of the defined update on the objective (14a) and on the problem constraints (14c), (14d). By substituting (41) into (14a) we have

$$\text{updated (14a)} = \frac{\psi \cdot \frac{1}{\psi} \cdot P_{\max} \cdot \boldsymbol{h}_{\mathrm{RD}}^{\mathrm{T}} \bar{\boldsymbol{w}}_t \bar{\boldsymbol{w}}_r{}^{\mathrm{H}} \boldsymbol{h}_{\mathrm{SR}} \boldsymbol{h}_{\mathrm{SR}}^{\mathrm{H}} \bar{\boldsymbol{w}}_r \bar{\boldsymbol{w}}_t{}^{\mathrm{H}} \boldsymbol{h}_{\mathrm{RD}}^*}{\frac{1}{\psi} \cdot \sigma_{\mathrm{r}}^2 \boldsymbol{h}_{\mathrm{RD}}^{\mathrm{T}} \bar{\boldsymbol{w}}_t \bar{\boldsymbol{w}}_t{}^{\mathrm{H}} \boldsymbol{h}_{\mathrm{RD}}^* + \sigma_{\mathrm{nd}}^2} \overset{(\psi \geq 1)}{\geq} \text{Old (14a)}. \tag{41}$$

On the other hand we need to verify the feasibility of our new variable set. By substituting (41) into (14c,d) we have

$$\begin{aligned} \text{updated } P^{(\mathrm{Tx})} &= \mathrm{Tr}\left(\frac{1}{\psi} \bar{\boldsymbol{w}}_t \bar{\boldsymbol{w}}_r^{\mathrm{H}} [\psi P_{\max} \boldsymbol{h}_{\mathrm{SR}} \boldsymbol{h}_{\mathrm{SR}}^{\mathrm{H}} + \sigma_{\mathrm{r}}^2 \boldsymbol{I}_{M_{\mathrm{r}}^{(\mathrm{R})}}] \bar{\boldsymbol{w}}_r \bar{\boldsymbol{w}}_t^{\mathrm{H}} \right) \\ &= \mathrm{Tr}\left(\bar{\boldsymbol{w}}_t \bar{\boldsymbol{w}}_r^{\mathrm{H}} [P_{\max} \boldsymbol{h}_{\mathrm{SR}} \boldsymbol{h}_{\mathrm{SR}}^{\mathrm{H}} + \frac{1}{\psi} \cdot \sigma_{\mathrm{r}}^2 \boldsymbol{I}_{M_{\mathrm{r}}^{(\mathrm{R})}}] \bar{\boldsymbol{w}}_r \bar{\boldsymbol{w}}_t^{\mathrm{H}} \right) \leq \text{old } P^{(\mathrm{Tx})} \leq P_{\max}^{(\mathrm{R})}, \end{aligned} \tag{43a}$$

$$\begin{aligned} \text{updated } P^{(\mathrm{Int})} &= \mathrm{Tr}\left(\boldsymbol{H}_{\mathrm{RR}} \frac{1}{\psi} \bar{\boldsymbol{w}}_t \bar{\boldsymbol{w}}_r^{\mathrm{H}} [\psi P_{\max} \boldsymbol{h}_{\mathrm{SR}} \boldsymbol{h}_{\mathrm{SR}}^{\mathrm{H}} + \sigma_{\mathrm{r}}^2 \boldsymbol{I}_{M_{\mathrm{r}}^{(\mathrm{R})}}] \bar{\boldsymbol{w}}_r \bar{\boldsymbol{w}}_t^{\mathrm{H}} \boldsymbol{H}_{\mathrm{RR}}^{\mathrm{H}} \right) \\ &= \mathrm{Tr}\left(\boldsymbol{H}_{\mathrm{RR}} \bar{\boldsymbol{w}}_t \bar{\boldsymbol{w}}_r^{\mathrm{H}} [P_{\max} \boldsymbol{h}_{\mathrm{SR}} \boldsymbol{h}_{\mathrm{SR}}^{\mathrm{H}} + \frac{1}{\psi} \cdot \sigma_{\mathrm{r}}^2 \boldsymbol{I}_{M_{\mathrm{r}}^{(\mathrm{R})}}] \bar{\boldsymbol{w}}_r \bar{\boldsymbol{w}}_t^{\mathrm{H}} \boldsymbol{H}_{\mathrm{RR}}^{\mathrm{H}} \right) \leq \text{old } P^{(\mathrm{Int})} \leq P_{\max}^{(\mathrm{R})}. \end{aligned} \tag{43b}$$

The improvement in (42) as well as the feasibility of (43) suggest that the choice of $\boldsymbol{w}_r = \frac{\boldsymbol{h}_{\mathrm{SR}}}{||\boldsymbol{h}_{\mathrm{SR}}||_2}$ is optimal for (14).

Appendix 2: Gradient details of (35a), (35b)

The goal is to calculate the derivatives of mutual information function (6)

$$\mathrm{MI}\left(\boldsymbol{x}; \boldsymbol{y}_{\mathbf{D}}\right) = \log_2 |\boldsymbol{A}_1| - \log_2 |\boldsymbol{A}_2|, \tag{43}$$

with defined matrices $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$ in (35). Now, we may formulate the differentiation and consequently the derivative of our mutual information function

with respect to $\boldsymbol{F}^*$ as following (applying Wirtingers calculus [48–50] and some famous matrix equalities [51]): The differentiation of the mutual information function can be written as:

$$\begin{aligned}
\partial \mathrm{MI}\left(\boldsymbol{x};\boldsymbol{y_D}\right) &= \partial \log_2|\boldsymbol{A}_1| - \partial \log_2|\boldsymbol{A}_2| \\
&= \frac{1}{\ln(2)\partial \boldsymbol{F}^*}\cdot \mathrm{Tr}\left(\boldsymbol{A}_1^{-1}\cdot \partial \boldsymbol{A}_1\right) - \frac{1}{\ln(2)}\cdot \mathrm{Tr}\left(\boldsymbol{A}_2^{-1}\cdot \partial \boldsymbol{A}_2\right) \Rightarrow \\
\frac{\partial \mathrm{MI}\left(\boldsymbol{x};\boldsymbol{y_D}\right)}{\partial \boldsymbol{F}^*} &= \frac{1}{\ln(2)\partial \boldsymbol{F}^*}\cdot \left[\mathrm{Tr}\left(\boldsymbol{A}_1^{-1}\boldsymbol{H}_{\mathrm{RD}}\boldsymbol{G}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{F}\cdot \partial \boldsymbol{F}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{RD}}^{\mathrm{H}}\right) + \mathrm{Tr}(\boldsymbol{A}_2^{-1}\cdot \boldsymbol{0})\right] \\
&= \frac{1}{\ln(2)\partial \boldsymbol{F}^*}\cdot \mathrm{Tr}\left(\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{RD}}^{\mathrm{H}}\boldsymbol{A}_1^{-1}\boldsymbol{H}_{\mathrm{RD}}\boldsymbol{G}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{F}\cdot \partial \boldsymbol{F}^{\mathrm{H}}\right) \\
\frac{\partial \mathrm{MI}\left(\boldsymbol{x};\boldsymbol{y_D}\right)}{\partial \boldsymbol{G}^*} &= \frac{1}{\ln(2)\partial \boldsymbol{G}^*}\cdot \Big[\mathrm{Tr}\left(\boldsymbol{A}_1^{-1}\boldsymbol{H}_{\mathrm{RD}}\boldsymbol{G}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{F}\cdot \boldsymbol{F}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\partial \boldsymbol{G}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{RD}}^{\mathrm{H}}\right) \\
&\quad + \mathrm{Tr}\left(\boldsymbol{A}_1^{-1}\sigma_{\mathrm{r}}^2\boldsymbol{H}_{\mathrm{RD}}\boldsymbol{G}\partial \boldsymbol{G}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{RD}}^{\mathrm{H}}\right) - \mathrm{Tr}\left(\boldsymbol{A}_2^{-1}\sigma_{\mathrm{r}}^2\boldsymbol{H}_{\mathrm{RD}}\boldsymbol{G}\cdot \partial \boldsymbol{G}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{RD}}^{\mathrm{H}}\right)\Big] \\
&= \frac{1}{\ln(2)\partial \boldsymbol{G}^*}\mathrm{Tr}\Big[\big(\boldsymbol{H}_{\mathrm{RD}}^{\mathrm{H}}\boldsymbol{A}_1^{-1}\boldsymbol{H}_{\mathrm{RD}}\boldsymbol{G}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{F}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}} \\
&\quad + \boldsymbol{H}_{\mathrm{RD}}^{\mathrm{H}}\boldsymbol{A}_1^{-1}\sigma_{\mathrm{r}}^2\boldsymbol{H}_{\mathrm{RD}}\boldsymbol{G} - \boldsymbol{H}_{\mathrm{RD}}^{\mathrm{H}}\boldsymbol{A}_2^{-1}\sigma_{\mathrm{r}}^2\boldsymbol{H}_{\mathrm{RD}}\boldsymbol{G}\big)\partial \boldsymbol{G}^{\mathrm{H}}\Big], \qquad (44)
\end{aligned}$$

which consequently result in the equations (35a) and (35b).

Appendix 3: Gradient details of (38)

Recalling (37a-c), the differentiation of the system power constraints with respect to $\boldsymbol{F}^*$ can be written as

$$\partial P^{(\mathrm{Relay})} = \mathrm{Tr}\left\{\boldsymbol{G}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{F}\cdot \partial \boldsymbol{F}^{\mathrm{H}}\cdot \boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}}\right\} = \mathrm{Tr}\left\{\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}}\boldsymbol{G}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{F}\cdot \partial \boldsymbol{F}^{\mathrm{H}}\right\}, \qquad (46a)$$

$$\partial P^{(\mathrm{Tx})} = \mathrm{Tr}\{\boldsymbol{F}\partial \boldsymbol{F}^{\mathrm{H}}\}, \qquad (46b)$$

$$\begin{aligned}
\partial P^{(\mathrm{Int})} &= \mathrm{Tr}\left\{\boldsymbol{H}_{\mathrm{RR}}\boldsymbol{G}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{F}\cdot \partial \boldsymbol{F}^{\mathrm{H}}\cdot \boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{RR}}^{\mathrm{H}}\right\} \\
&= \mathrm{Tr}\left\{\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\boldsymbol{G}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{RR}}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{RR}}\boldsymbol{G}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{F}\cdot \partial \boldsymbol{F}^{\mathrm{H}}\right\}, \qquad (46c)
\end{aligned}$$

which results in identities (38a), (38c) and (38e). Similar differentiations can be achieved with respect to $\boldsymbol{G}^*$ as

$$\begin{aligned}
\partial P^{(\mathrm{Relay})} &= \mathrm{Tr}\left\{\boldsymbol{G}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{F}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\partial \boldsymbol{G}^{\mathrm{H}} + \sigma_{\mathrm{r}}^2\boldsymbol{G}\cdot \partial \boldsymbol{G}^{\mathrm{H}}\right\} \\
&= \mathrm{Tr}\left\{[\boldsymbol{G}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{F}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}} + \sigma_{\mathrm{r}}^2\boldsymbol{G}]\cdot \partial \boldsymbol{G}^{\mathrm{H}}\right\}, \qquad (47a) \\
\partial P^{(\mathrm{Tx})} &= \mathrm{Tr}\{0\}, \qquad (47b) \\
\partial P^{(\mathrm{Int})} &= \mathrm{Tr}\left\{\boldsymbol{H}_{\mathrm{RR}}\boldsymbol{G}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{F}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}}\cdot \partial \boldsymbol{G}^{\mathrm{H}}\cdot \boldsymbol{H}_{\mathrm{RR}}^{\mathrm{H}} + \sigma_{\mathrm{r}}^2\boldsymbol{H}_{\mathrm{RR}}\boldsymbol{G}\cdot \partial \boldsymbol{G}^{\mathrm{H}}\cdot \boldsymbol{H}_{\mathrm{RR}}^{\mathrm{H}}\right\} \\
&= \mathrm{Tr}\left\{[\boldsymbol{H}_{\mathrm{RR}}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{RR}}\boldsymbol{G}\boldsymbol{H}_{\mathrm{SR}}\boldsymbol{F}\boldsymbol{F}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{SR}}^{\mathrm{H}} + \sigma_{\mathrm{r}}^2\cdot \boldsymbol{H}_{\mathrm{RR}}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{RR}}\boldsymbol{G}]\cdot \partial \boldsymbol{G}^{\mathrm{H}}\right\}, \qquad (47c)
\end{aligned}$$

which result in identities (38b), (38d) and (38f).

7. Acknowledgements

The authors acknowledge the financial support by the Carl-Zeiss-Foundation (http://carl-zeiss-stiftung.de/).

References

[1] J. Zhang, O. Taghizadeh, M. Haardt, Joint source and relay precoding design for one-way full-duplex MIMO relaying systems, Proceedings of the Tenth International Symposium on Wireless Communication Systems (ISWCS), Ilmenau, Germany, 2013.

[2] J. I. Choi, S. Hong, M. Jain, S. Katti, P. Levis, J. Mehlman, Beyond full duplex wireless, in: Forty Sixth Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, 2012.

[3] M. Duarte, A. Sabharwal, Full-duplex wireless communications using off-the-shelf radios: Feasibility and first results, in: Proceedings of 44th Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, 2010.

[4] R. Askar, T. Kaiser, B. Schubert, T. Haustein, W. Keusgen, Active self-interference cancellation mechanism for full-duplex wireless transceivers, in: Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM), 2014 9th International Conference on, 2014.

[5] D. Bharadia, S. Katti, Full duplex mimo radios, in: Proceedings of the 11th USENIX Conference on Networked Systems Design and Implementation, NSDI'14, Berkeley, CA, USA, 2014, pp. 359–372.

[6] A. Tang, X. Wang, Balanced RF-circuit based self-interference cancellation for full duplex communications, Elsevier Ad Hoc Networks 24 (2015) 214–227.

[7] D. Bharadia, E. McMilin, S. Katti, Full duplex radios, in: Proceedings of the ACM SIGCOMM, Vol. 43, 2013.

[8] Y. Hua, Y. Ma, A. Gholian, Y. Li, A. C. Cirik, P. Liang, Radio self-interference cancellation by transmit beamforming, all-analog cancellation and blind digital tuning, Elsevier Signal Processing 108 (2015) 322–340.

[9] B. Day, A. Margetts, D. Bliss, P. Schniter, Full-duplex bidirectional MIMO: Achievable rates under limited dynamic range, IEEE Transactions on Signal Processing 60 (7) (2012) 3702–3713.

[10] A. S. E. Everett, A. Sahai, Passive self-interference suppression for full-duplex infrastructure nodes, IEEE Transactions on Wireless Communication 13 (2) (2014) 680–694.

[11] T. Riihonen, R. Wichman, Analog and digital self-interference cancellation in full-duplex MIMO-OFDM transceivers with limited resolution in A/D conversion, in: Signals, Systems and Computers (ASILOMAR), 2012 Conference Record of the Forty Sixth Asilomar Conference on, 2012, pp. 45–49.

[12] O. Taghizadeh, M. Rothe, A. C. Cirik, R. Mathar, Distortion-Loop analysis for Full-Duplex Amplify-and-Forward relaying in cooperative multicast scenarios, in: 2015 9th International Conference on Signal Processing and Communication Systems (ICSPCS), Cairns, Australia, 2015, pp. 107–115.

[13] S. Hong, J. Brand, J. Choi, M. Jain, J. Mehlman, S. Katti, P. Levis, Applications of self-interference cancellation in 5G and beyond, IEEE Communications Magazine 52 (2) (2014) 114–121.

[14] D. Nguyen, L.-N. Tran, P. Pirinen, M. Latva-aho, On the spectral efficiency of full-duplex small cell wireless systems, IEEE Transactions on Wireless Communications 13 (9) (2014) 4896–4910.

[15] F. Zhu, F. Gao, M. Yao, H. Zou, Joint information- and jamming-beamforming for physical layer security with full duplex base station, IEEE Transactions on Signal Processing 62 (24) (2014) 6391–6401.

[16] G. Liu, F. Yu, H. Ji, V. Leung, X. Li, In-band full-duplex relaying: A survey, research issues and challenges, IEEE Communications Surveys Tutorials 32 (9) (2014) 1637–1652.

[17] B. Zhong, D. Zhang, Z. Zhang, Z. Pan, K. Long, A. Vasilakos, Opportunistic full-duplex relay selection for decode-and-forward cooperative networks over rayleigh fading channels, in: IEEE International Conference on Communications (ICC), Sydney Australia, 2014.

[18] L. Chen, S. Han, W. Meng, C. Li, Optimal power allocation for dual-hop full-duplex decode-and-forward relay, IEEE Communications Letters 19 (3) (2015) 471–474.

[19] B. Day, A. Margetts, D. Bliss, P. Schniter, Full-duplex MIMO relaying: Achievable rates under limited dynamic range, IEEE Journal on Selected Areas in Communications 30 (8) (2012) 1541–1553.

[20] O. Taghizadeh, R. Mathar, Robust multi-user decode-and-forward relaying with full-duplex operation, in: The Eleventh International Symposium on Wireless Communication Systems (ISWCS 2014), Barcelona, Spain, 2014.

[21] L. Sanguinetti, A. D. Amico, Y. Rong, A tutorial on the optimization of amplify-and-forward MIMO relay systems, IEEE Journal on Selected Areas in Communications 30 (8) (2012) 1331–1346.

[22] C. Xing, S. Ma, M. Xia, Y.-C. Wu, Cooperative beamforming for dual-hop amplify-and-forward multi-antenna relaying cellular networks, Elsevier Signal processing 92 (11) (2012) 2689–2699.

[23] T. Riihonen, S. Werner, R. Wichman, Optimized gain control for single-frequency relaying with loop interference, IEEE Transactions on Wireless Communications, 8 (6) (2009) 2801–2806.

[24] L. Jimenez Rodriguez, N. Tran, T. Le-Ngoc, Optimal power allocation and capacity of full-duplex AF relaying under residual self-interference, IEEE Wireless Communications Letters 3 (2) (2014) 233–236.

[25] O. Taghizadeh, R. Mathar, Cooperative strategies for distributed full-duplex relay networks with limited dynamic range, in: International Conference on Wireless for Space and Extreme Environments (WiSEE), 2014 IEEE, Noordwijk, Netherlands.

[26] C. Dang, L. Jimenez Rodriguez, N. Tran, S. Shelly, S. Sastry, Secrecy capacity of the full-duplex AF relay wire-tap channel under residual self-interference, in: IEEE Wireless Communications and Networking Conference (WCNC), 2015.

[27] X. Cheng, B. Yu, X. Cheng, L. Yang, Two-way full-duplex amplify-and-forward relaying, in: IEEE Military Communications Conference, MILCOM, 2013.

[28] T. Riihonen, S. Werner, R. Wichman, Mitigation of loopback self-interference in full-duplex MIMO relays, IEEE Transactions on Signal Processing 59 (12) (2011) 5983–5993.

[29] T. Riihonen, A. Balakrishnan, K. Haneda, S. Wyne, S. Werner, R. Wichman, Optimal eigen-beamforming for suppressing self-interference in full-duplex MIMO relays, in: 45th Annual Conference on Information Sciences and Systems (CISS), 2011, pp. 1–6.

[30] Y. Y. Kang, J. H. Cho, Capacity of MIMO wireless channel with full-duplex amplify-and-forward relay, in: Proceedings of IEEE 20th International Symposium on Personal, Indoor and Mobile Radio Communications, Pohang, South Korea, 2009.

[31] J. Sangiamwong, T. Asai, J. Hagiwara, Y. Okumura, T. Ohya, Joint multi-filter design for full-duplex MU-MIMO relaying, in: IEEE Vehicular Technology Conference, 2009.

[32] K. Lee, H. Kwon, M. Jo, H. Park, Y. Lee, MMSE-based optimal design of full-duplex relay system, in: IEEE Vehicular Technology Conference (VTC Fall), Quebec, Canada, 2012.

[33] Y. Y. Kang, B.-J. Kwak, J. H. Cho, An optimal full-duplex AF relay for joint analog and digital domain self-interference cancellation, IEEE Transactions on Communications 13 (2) (2014) 913–927.

[34] H. Suraweera, I. Krikidis, G. Zheng, C. Yuen, P. Smith, Low-complexity end-to-end performance optimization in MIMO full-duplex relay systems, IEEE Transactions on Wireless Communications, 13 (2) (2014) 913–927.

[35] K. Yang, H. Cui, L. Song, Y. Li, Efficient full-duplex relaying with joint antenna-relay selection and self-interference suppression, IEEE Transactions on Wireless Communications 14 (7) (2015) 3991–4005.

[36] G. Zheng, Joint beamforming optimization and power control for full-duplex MIMO two-way relay channel, Signal Processing, IEEE Transactions on 63 (3) (2015) 555–566.

[37] U. Ugurlu, T. Riihonen, R. Wichman, Optimized in-band full-duplex MIMO relay under single-stream transmission, Vehicular Technology, IEEE Transactions on PP (99) (2015) 1–1.

[38] M. Duarte, C. Dick, A. Sabharwal, Experiment-driven characterization of full-duplex wireless systems, IEEE Transactions on Wireless Communications 11 (12) (2012) 4296–4307.

[39] I. Krikidis, H. Suraweera, P. Smith, C. Yuen, Full-duplex relay selection for amplify-and-forward cooperative networks, IEEE Transactions on Wireless Communications 11 (12) (2012) 4381–4393.

[40] Z. Liu, Y. Liu, F. Liu, Joint resource scheduling for full-duplex cellular system, in: 22nd International Conference on Telecommunications (ICT), Sydney, Australia, 2015.

[41] S. Huberman, T. Le-Ngoc, Self-interference threshold-based MIMO full-duplex precoding, IEEE Transactions on Vehicular Technology 64 (8) (2015) 3803–3807.

[42] J. Zhang, O. Taghizadeh, M. Haardt, Robust transmit beamforming design for full-duplex point-to-point MIMO systems, Proceedings of the Tenth International Symposium on Wireless Communication Systems (ISWCS), Ilmenau, Germany, 2013.

[43] A. W. Marshall, I. Olkin, B. C. Arnold, Inequalities: Theory of Majorization and Its Applications, Springer, 2010.

[44] S. P. Boyd, L. Vandenberghe, Convex optimization, Cambridge University Press, 2004.

[45] Y. Rong, X. Tang, Y. Hua, A unified framework for optimizing linear nonregenerative multicarrier MIMO relay communication systems, IEEE Transactions on Signal Processing 57 (12) (2009) 4837–4851.

[46] Y. Huang, D. P. Palomar, Rank-constrained separable semidefinite programming with applications to optimal beamforming, IEEE Transactions on Signal Processing 58 (2010) 664–678.

[47] A. Cohen, Stepsize analysis for descent methods, in: Proceedings of IEEE 12th International Symposium on Adaptive Processes, 1973.

[48] W. Wirtinger, Zur formalen Theorie der Funktionen von mehreren komplexen Veränderlichen, Mathematische Annalen 97 (1) (1927) 357–375.

[49] Y. Huang, D. P. Palomar, Complex-valued matrix differentiation: Techniques and key results, IEEE Transactions on Signal Processing 55 (2007) 2740–2746.

[50] T. Abrudan, J. Eriksson, V. Koivunen, Conjugate gradient algorithm for optimization under unitary matrix constraint, Elsevier Signal Processing 89 (9) (2009) 1704–1714.

[51] S. Petersen, M. Pedersen, Matrix Cookbook, 2008.

Algorithm 1 The alternating optimization for FD relaying with self-interference (AO) design of $\boldsymbol{F}, \boldsymbol{G}$.

$\boldsymbol{U}_{\rm sr}\boldsymbol{\Sigma}_{\rm sr}\boldsymbol{V}_{\rm sr}^{\rm H} \leftarrow \text{SVD}\left(\boldsymbol{H}_{\rm SR}\right)$
$\boldsymbol{F} \leftarrow \boldsymbol{V}_{\rm sr}\boldsymbol{\Sigma}^{\star}$, $\boldsymbol{\Sigma}^{\star} \leftarrow$ Water-Filling [44]
$k \leftarrow 1$
while $|res(k) - res(k-1)| > 1e-3$ **do**
  $k \leftarrow k+1$
  $\boldsymbol{G} \leftarrow$ solution to (24);
  if ($\boldsymbol{G}$ is *not* feasible) **then**
    $\boldsymbol{G}_0 \leftarrow \boldsymbol{G}/s_0$; (scaling down to satisfy (23c))
    $\boldsymbol{G} \leftarrow$ solution to (25);
    if ($\boldsymbol{G}$ is *not* feasible) **then**
      $\boldsymbol{G}_1 \leftarrow \boldsymbol{G}/s_1$; (scaling down to satisfy (23b))
      $\boldsymbol{G}_2 \leftarrow$ calculated from (30)-(34);
      $\boldsymbol{G} \leftarrow \max\limits_{\boldsymbol{G}\in\{\boldsymbol{G}_0,\boldsymbol{G}_1,\boldsymbol{G}_2\}} \text{MI}\left(\boldsymbol{x};\boldsymbol{y}_{\mathbf{D}}\big|_{\boldsymbol{H}_{\rm RD},\boldsymbol{G},\boldsymbol{F},\boldsymbol{H}_{\rm SR}}\right)$
    end if
  end if
  $\boldsymbol{F} \leftarrow$ calculate via convex optimization (22);
  $res(k) \leftarrow$ calculate the objective via (6);
end while

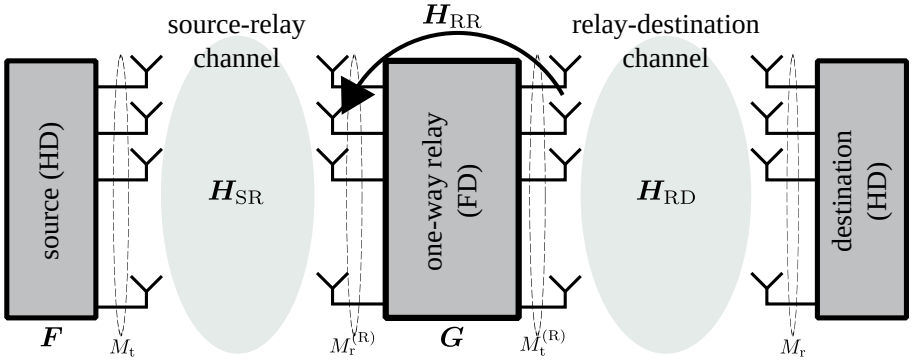

Figure 1: A one-way full-duplex (FD) relaying system. A source node communicates with a destination node with a help of a FD relay. The arrow ($\boldsymbol{H}_{\mathrm{RR}}$) represents the loopback self-interference path between the relay transmitter and receiver ends.

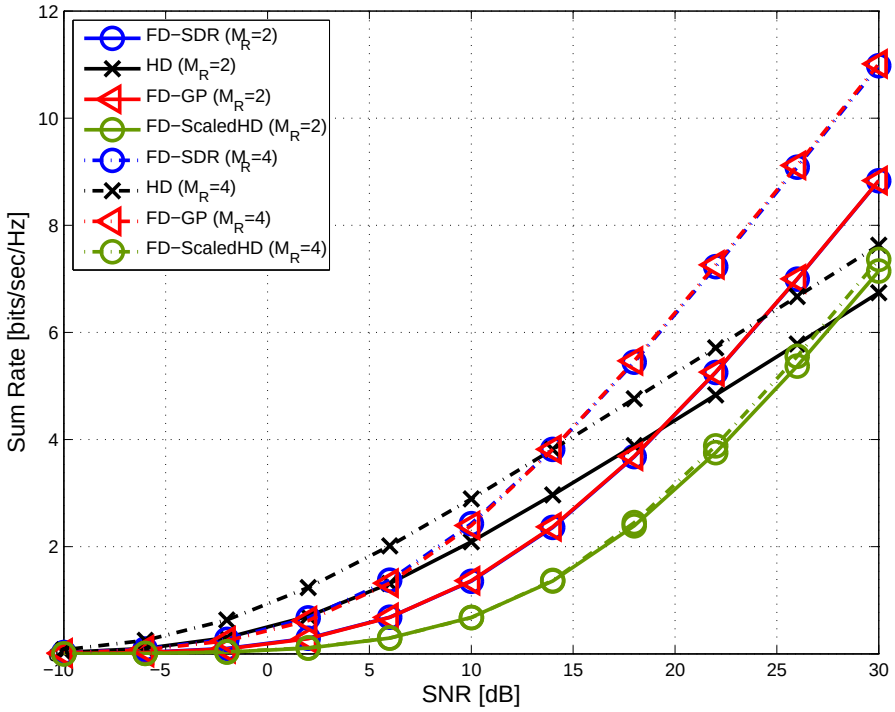

Figure 2: Sum Rate [bits/sec/Hz] vs. SNR [dB]. $\rho_{\mathrm{SR}} = \rho_{\mathrm{RD}} = 0$ dB, $\rho_{\mathrm{RR}} = 20$ dB, $M_{\mathrm{r}} = M_{\mathrm{t}} = 1, M_{\mathrm{r}}^{(\mathrm{R})} = M_{\mathrm{t}}^{(\mathrm{R})} = M_{\mathrm{R}}$, $\epsilon = 5$ dB, $\sigma_{\mathrm{si}}^2 = 0$. Similar performance is observed for the rank-1 design, and the gradient projection based optimization.

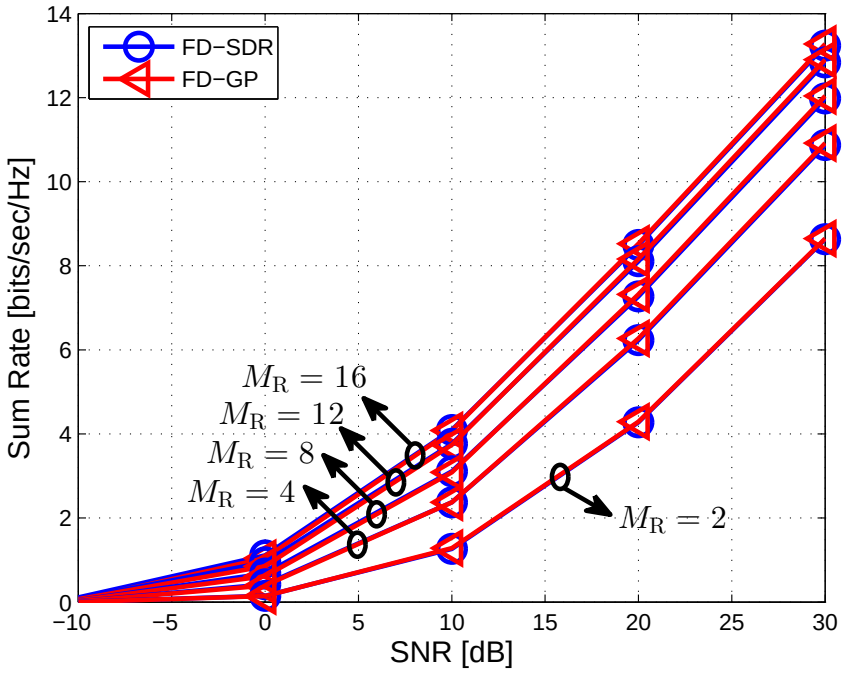

Figure 3: Sum Rate [bits/sec/Hz] vs. SNR [dB]. $\rho_{\mathrm{SR}} = \rho_{\mathrm{RD}} = 0$ dB, $\rho_{\mathrm{RR}} = 20$ dB, $M_{\mathrm{r}} = M_{\mathrm{t}} = 1, M_{\mathrm{r}}^{(\mathrm{R})} = M_{\mathrm{t}}^{(\mathrm{R})} = M_{\mathrm{R}}$, $\epsilon = 5$ dB, $\sigma_{\mathrm{si}}^2 = 0$. Similar performance is observed for the rank-1 design and the gradient projection based optimization, with larger antenna array at the relay.

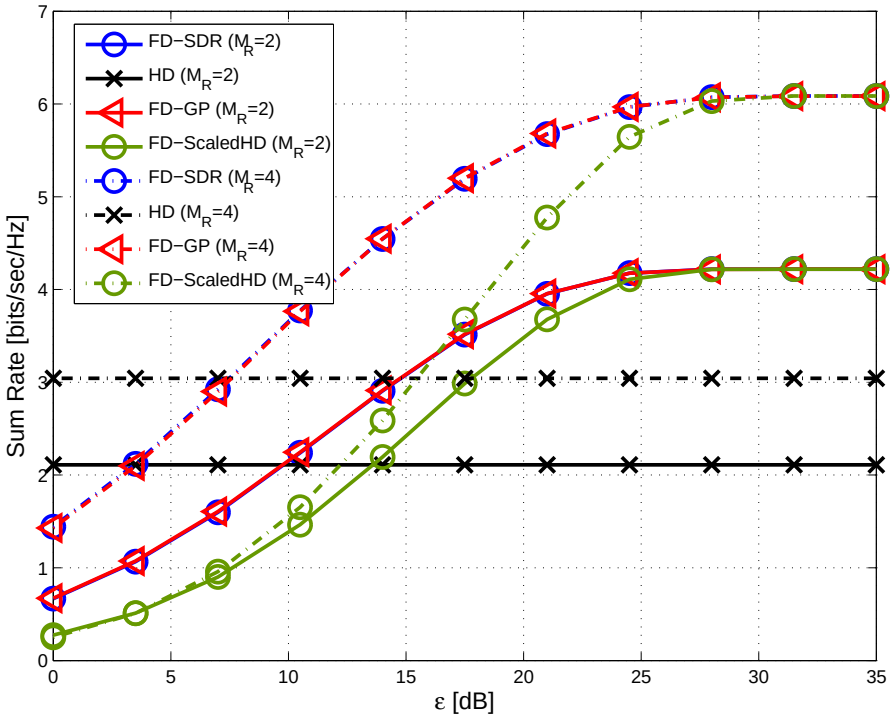

Figure 4: Sum Rate [bits/sec/Hz] vs. ϵ [dB]. $\rho_{\mathrm{SR}} = \rho_{\mathrm{RD}} = 0$ dB, $\rho_{\mathrm{RR}} = 20$ dB, $M_{\mathrm{r}} = M_{\mathrm{t}} = 1, M_{\mathrm{r}}^{(\mathrm{R})} = M_{\mathrm{t}}^{(\mathrm{R})} = M_{\mathrm{R}}$, SNR $= 10$ dB, $\sigma_{\mathrm{si}}^2 = 0$. The FD gain decreases as the dynamic range decreases. For a system with a high dynamic range, the two-fold gain is observable with FD operation.

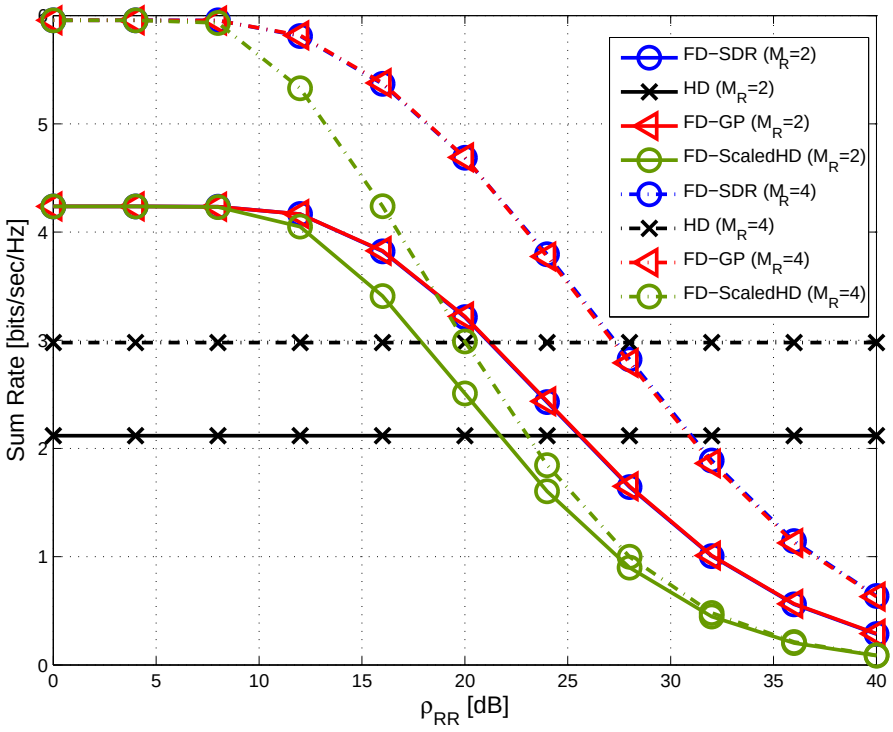

Figure 5: Sum Rate [bits/sec/Hz] vs. ρ_{RR} [dB]. $\rho_{\mathrm{SR}} = \rho_{\mathrm{RD}} = 0$ dB, $M_{\mathrm{r}} = M_{\mathrm{t}} = 1, M_{\mathrm{r}}^{(\mathrm{R})} = M_{\mathrm{t}}^{(\mathrm{R})} = M_{\mathrm{R}}$, SNR $= 10$ dB, $\epsilon = 15$ dB, $\sigma_{\mathrm{si}}^2 = 0$. The FD gain decreases as the self-interference channel gets stronger. Similar performance is observed for the rank-1 design, and the gradient projection based optimization.

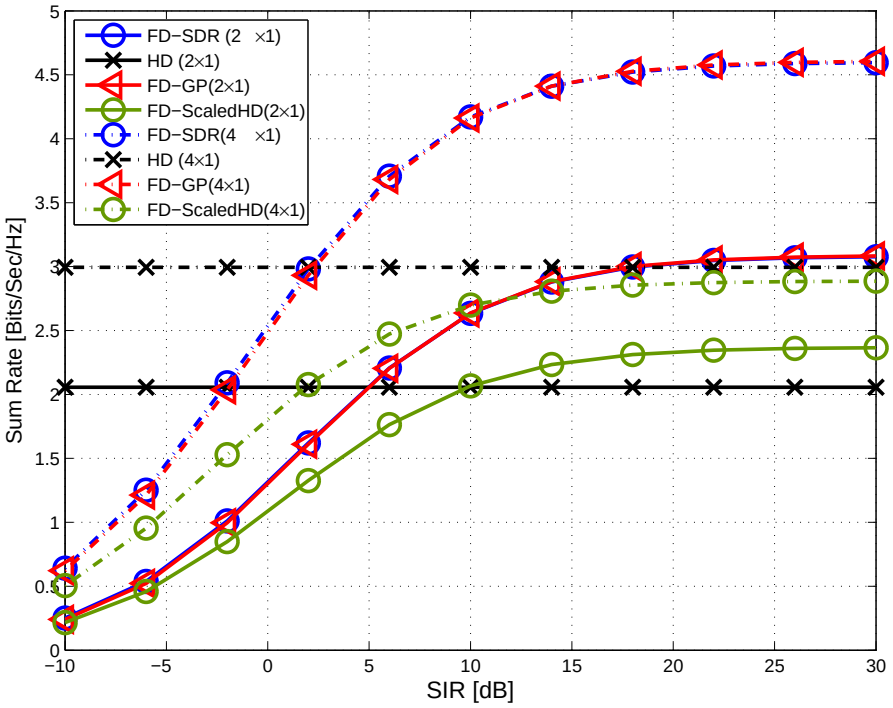

Figure 6: Sum Rate [bits/sec/Hz] vs. SIR [dB]. $\rho_{\mathrm{SR}} = \rho_{\mathrm{RD}} = 0$ dB, $M_{\mathrm{r}} = M_{\mathrm{t}} = 1, M_{\mathrm{r}}^{(\mathrm{R})} = M_{\mathrm{t}}^{(\mathrm{R})} = M_{\mathrm{R}}$, SNR $= 10$ dB, $\epsilon = 15$ dB. The FD gain decreases as the residual self-interference gets stronger.

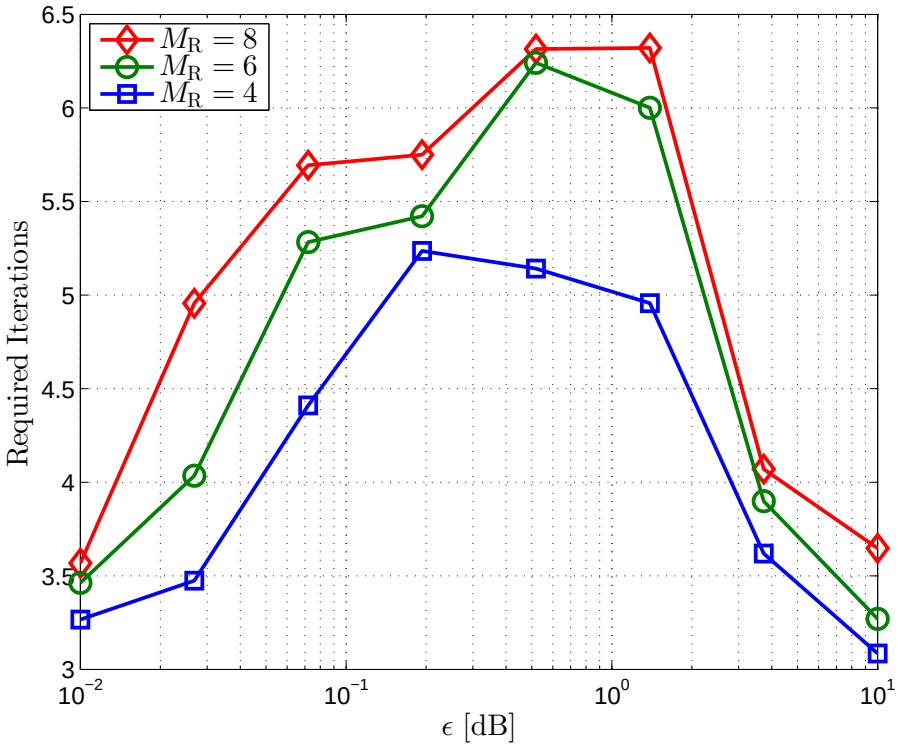

Figure 7: Average number of required iterations for the proposed AO algorithm. $\rho_{\mathrm{SR}} = \rho_{\mathrm{RD}} = \rho_{\mathrm{RR}} = 0$ dB, $M_{\mathrm{r}} = M_{\mathrm{t}} = M_{\mathrm{r}}^{(\mathrm{R})} = M_{\mathrm{t}}^{(\mathrm{R})} = M_{\mathrm{R}}$, SNR $= 10$ dB, $\sigma_{\mathrm{si}}^2 = 0$. A faster convergence is observed for very high or very low system dynamic ranges.

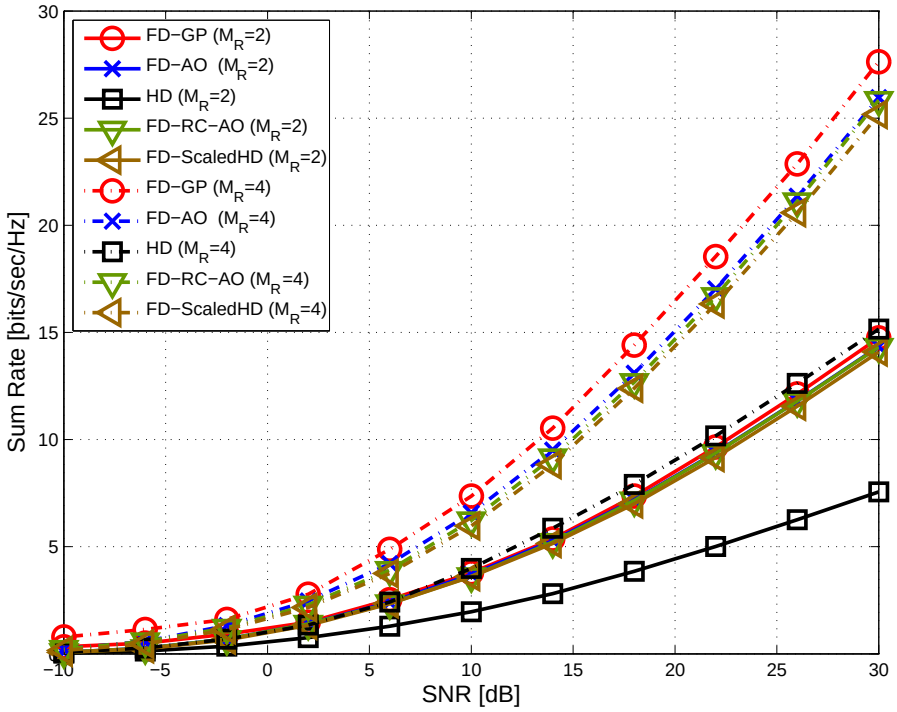

Figure 8: Sum Rate [bits/sec/Hz] vs. SNR [dB]. $\rho_{\mathrm{SR}} = \rho_{\mathrm{RD}} = 0$ dB, $\rho_{\mathrm{RR}} = 20$ dB, $M_{\mathrm{r}} = M_{\mathrm{t}} = M_{\mathrm{r}}^{(\mathrm{R})} = M_{\mathrm{t}}^{(\mathrm{R})} = M_{\mathrm{R}}$, $\epsilon = 20$ dB, $\sigma_{\mathrm{si}}^2 = 0$. A FD gain is observable for a wide range of SNR values.

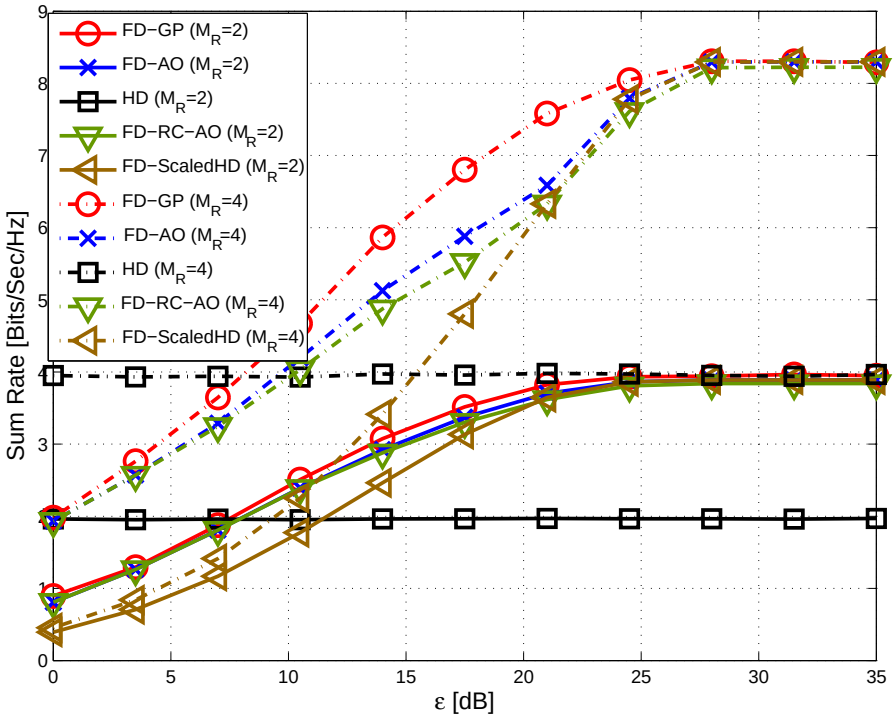

Figure 9: Sum Rate [bits/sec/Hz] vs. ϵ [dB]. $\rho_{\mathrm{SR}} = \rho_{\mathrm{RD}} = 0$ dB, $\rho_{\mathrm{RR}} = 20$ dB, $M_{\mathrm{r}} = M_{\mathrm{t}} = M_{\mathrm{r}}^{(\mathrm{R})} = M_{\mathrm{t}}^{(\mathrm{R})} = M_{\mathrm{R}}$, SNR $= 10$ dB, $\sigma_{\mathrm{si}}^2 = 0$.

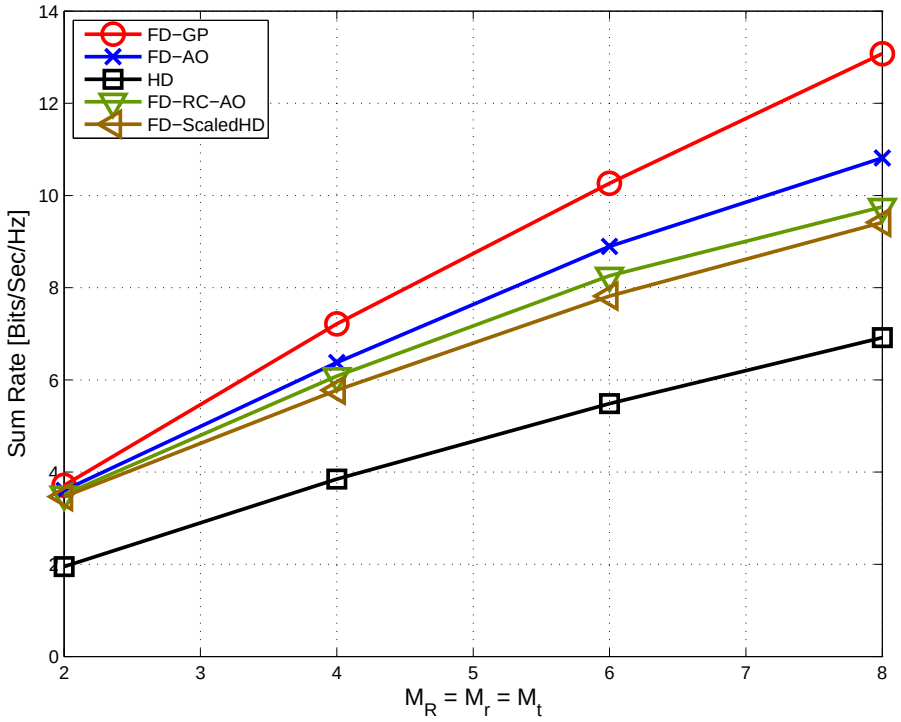

Figure 10: Sum Rate [bits/sec/Hz] vs. M_{R}. $\rho_{\mathrm{SR}} = \rho_{\mathrm{RD}} = 0$ dB, $\rho_{\mathrm{RR}} = 20$ dB, $M_{\mathrm{r}} = M_{\mathrm{t}} = M_{\mathrm{r}}^{(\mathrm{R})} = M_{\mathrm{t}}^{(\mathrm{R})} = M_{\mathrm{R}}$, SNR $= 10$ dB, $\epsilon = 20$ dB, $\sigma_{\mathrm{si}}^2 = 0$.

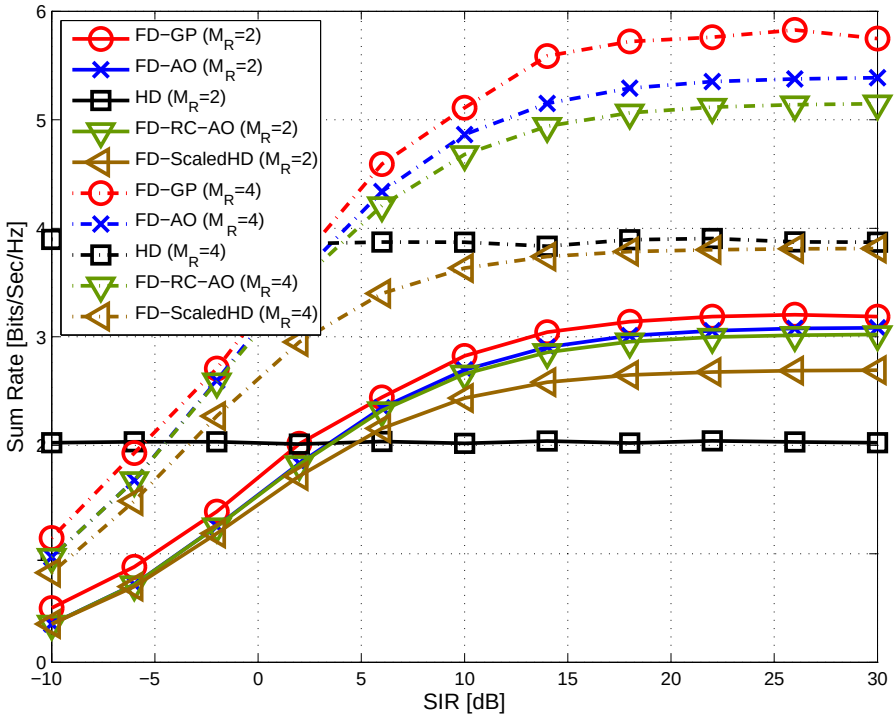

Figure 11: Sum Rate [bits/sec/Hz] vs. SIR [dB]. $\rho_{\mathrm{SR}} = \rho_{\mathrm{RD}} = 0$ dB, $M_{\mathrm{r}} = M_{\mathrm{t}} = M_{\mathrm{r}}^{(\mathrm{R})} = M_{\mathrm{t}}^{(\mathrm{R})} = M_{\mathrm{R}}$, SNR = 10 dB, $\epsilon = 15$ dB. The FD gain decreases as the residual self-interference gets stronger.

Table 1: Required number of floating point operations for different design algorithms.

	FD-GP	FD-SDR
Floating Point Operations	$\mathcal{O}\left(2\zeta_1\zeta_2{M_\mathrm{R}}^3 + 39\zeta_1\zeta_2{M_\mathrm{R}}^2 + 13\zeta_1\zeta_2 M_\mathrm{R}\right)$	$\mathcal{O}\left(\left(21 + L_1\right){M_\mathrm{R}}^3 + 10{M_\mathrm{R}}^2 + 9M_\mathrm{R}\right)$

Table 2: CPU Time [sec.] for different M_{R}. $\zeta_2 = 10$.

	FD-GP	FD-SDR
$M_{\mathrm{R}} = 2$	46.07	0.324
$M_{\mathrm{R}} = 4$	114.95	0.684

Table 3: Required number of floating point operations for different design algorithms.

	FD-GP	AO	RC-AO
Floating Point Operations	$\mathcal{O}\big(4\zeta_1\zeta_2 {M_\mathrm{R}}^4 + (33\zeta_1\zeta_2\zeta_4 + 165\zeta_1\zeta_2)\, {M_\mathrm{R}}^3 + 54\zeta_1\zeta_2 {M_\mathrm{R}}^2 + 31\zeta_1\zeta_2 M_\mathrm{R}\big)$	$\mathcal{O}\big(\zeta_3\,(78 + 3L_2)\, {M_\mathrm{R}}^3 + \zeta_3\,(29 + 9L_2)\, {M_\mathrm{R}}^2 + 26\zeta_3 M_\mathrm{R}\big)$	$\mathcal{O}\big(56 {M_\mathrm{R}}^3 + 2{M_\mathrm{R}}^2 + 3M_\mathrm{R}\big)$

Table 4: CPU Time [sec.] for different M_{R}. $\zeta_2 = 10$.

	FD-GP	AO	RC-AO
$M_{\mathrm{R}} = 2$	271.31	2.09	0.35
$M_{\mathrm{R}} = 4$	803.09	4.60	0.89