Massive MIMO Full-Duplex Relaying with Optimal Power Allocation for Independent Multipairs

João S. Lemos*, Francisco Rosário*, Francisco A. Monteiro†, João Xavier‡, António Rodrigues*

*Instituto de Telecomunicações, and Instituto Superior Técnico, Universidade de Lisboa, Portugal
†Instituto de Telecomunicações, and ISCTE - Instituto Universitário de Lisboa, Portugal
‡Instituto de Sistemas e Robótica, and Instituto Superior Técnico, Universidade de Lisboa, Portugal
{joao.sande.lemos, francisco.rosario}@tecnico.ulisboa.pt , francisco.monteiro@lx.it.pt

Abstract—With the help of an in-band full-duplex relay station, it is possible to simultaneously transmit and receive signals from multiple users. The performance of such system can be greatly increased when the relay station is equipped with a large number of antennas on both transmitter and receiver sides. In this paper, we exploit the use of massive arrays to effectively suppress the loopback interference (LI) of a decode-and-forward relay (DF) and evaluate the performance of the end-to-end (e2e) transmission. This paper assumes imperfect channel state information is available at the relay and designs a minimum mean-square error (MMSE) filter to mitigate the interference. Subsequently, we adopt zero-forcing (ZF) filters for both detection and beamforming. The performance of such system is evaluated in terms of bit error rate (BER) at both relay and destinations, and an optimal choice for the transmission power at the relay is shown. We then propose a complexity efficient optimal power allocation (OPA) algorithm that, using the channel statistics, computes the minimum power that satisfies the rate constraints of each pair. The results obtained via simulation show that when both MMSE filtering and OPA method are used, better values for the energy efficiency are attained.

Index terms - Decode-and-forward relay, in-band full-duplex, massive multiple-input multiple-output (MIMO), MMSE, ZF, loopback interference cancellation, optimal power allocation.

I. Introduction

In-band full-duplex systems in wireless communications have attracted a lot of research interest in the past few years. This feature is expected to be implemented in future generation networks and gives the possibility of transmitting and receiving data at the same time instant and frequency band, which means that with the same amount of energy as in the half-duplex (HD) counterpart, it is possible to double spectral efficiency [1]. However, the inherent complexity of such systems has been hindering the deployment of full-duplex enabled equipment. The main obstacle to overcome is the difficulty in canceling the signal leakage from the relay output to its input [2], [3]. Moreover, hardware constraints, algorithms' inefficiency and difficulty in modeling the channel have also posed problems, hence the need for further analysis.

Alongside the developments in full-duplex, a lot of research has been conducted recently on the capabilities of multiple-input multiple-output (MIMO) systems and its extension to very large arrays (dubbed as massive MIMO). The idea in future generation systems is to use hundreds of antennas at the base station in the hope of serving more users and devices. This study involves various signal processing techniques, such as channel modeling and estimation, precoding and detection algorithms, which must be efficient in terms of complexity and perform as close to optimal as possible.

In this work, we consider a massive MIMO full-duplex relay architecture, which uses both of the aforementioned state-of-the-art technologies. We assume a group of K sources communicating with a group of K destinations. This transmission is supported by a decode-and-forward (DF) relay station.

In the first part of this paper, we are interested in loopback interference (LI) mitigation techniques for the considered setup. We assume that both source to relay and LI channels are known (apart from an estimation error) to design a minimum-mean square error filter. This is done in [4] for a small number of antennas. We show that by exploiting the massive MIMO effects and using simple linear processing techniques, LI can be effectively attenuated. The results are evaluated in terms of end-to-end (e2e) bit error rate (BER) performance. Furthermore, an algorithm to find the optimal power allocation (OPA) that fulfils the needs of each communication pair in terms of achievable rate and peak power is presented.

II. System Model

Let us consider K user pairs establish a wireless connection through a relay station, which operates in in-band full-duplex mode (sharing the same time-frequency resources). For this purpose, each user is equipped with a single antenna, while the relay has N_{rx} receiving antennas and N_{tx} transmitting antennas. The number of antennas at the relay may go up to a few hundred and the number of served users is in general considered to be much lower ($K << N_{rx}, N_{tx}$). The nonexistence of a direct link between the sources and destinations is assumed, hence all links are established via the relay station. A simplified representation, as well as the used notation to be defined in this section, can be seen in figure 1.

A. Channel Model

For the evaluation of the performance of such setup, we are interested in the vectors containing the transmitted symbols from the sources to the relay and the beamformed symbols from the relay to the destinations, denoted by $\mathbf{x} \in \mathbb{C}^{K \times 1}$ and $\mathbf{t} \in \mathbb{C}^{N_{tx} \times 1}$, respectively. These vectors are assumed to be taken from an M-QAM constellation and the output signals to be normalized and uncorrelated, such that $\mathbb{E}(\mathbf{x}\mathbf{x}^H) = \mathbf{I}$ and
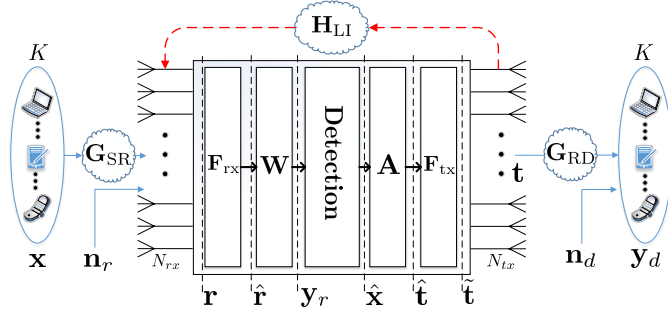
Figure 1. Full-duplex relay system.

$\mathbb{E}(\mathbf{t}^H \mathbf{t}) = 1$. In this manner, the power transmitted by each source is independent of K and the total transmitted power at the relay dissociated of N_{tx}. The received vectors at the relay and at the destinations are given by (1) and (2), respectively.

$$\mathbf{r} = \mathbf{G}_{\text{SR}}\mathbf{D}_{ps}^{1/2}\mathbf{x} + \sqrt{p_{\text{R}}}\mathbf{H}_{\text{LI}}\mathbf{t} + \mathbf{n}_r. \tag{1}$$

$$\mathbf{y}_d = \sqrt{p_{\text{R}}}\mathbf{G}_{\text{RD}}\mathbf{t} + \mathbf{n}_d. \tag{2}$$

The uplink symbols in $\mathbf{x}$ pass through a MIMO channel from the sources to the relay denoted by $\mathbf{G}_{\text{SR}} \in \mathbb{C}^{N_{rx} \times K}$, whereas the symbols from the relay to the destinations pass through $\mathbf{G}_{\text{RD}} \in \mathbb{C}^{K \times N_{tx}}$. Both matrices are expressed as $\mathbf{G}_{\text{SR}} = \mathbf{H}_{\text{SR}}\mathbf{D}_{\text{SR}}^{1/2}$ and $\mathbf{G}_{\text{RD}} = \mathbf{D}_{\text{RD}}^{1/2}\mathbf{H}_{\text{RD}}$, where $\mathbf{D}_{\text{SR}}$ and $\mathbf{D}_{\text{RD}}$ are diagonal matrices with entries $\beta_{\text{SR},k}$ and $\beta_{\text{RD},k}$ taken from a log-normal distribution, which account for large-scale channel effects. The fast-fading channel components are present in $\mathbf{H}_{\text{SR}}$ and $\mathbf{H}_{\text{RD}}$, both with independent entries taken from a $\mathcal{CN}(0,1)$ distribution. The LI channel is represented by $\mathbf{H}_{\text{LI}}$ with distribution $\mathcal{CN}(0,\sigma_{\text{LI}}^2)$, where σ_{LI}^2 accounts for the residual LI power after self-interference suppression imposed in the analog-domain. The vectors $\mathbf{n}_r \sim \mathcal{CN}(0,\sigma_{\mathbf{n}_r}^2)$ and $\mathbf{n}_d \sim \mathcal{CN}(0,\sigma_{\mathbf{n}_d}^2)$ account for additive white Gaussian noise at the relay and destinations, respectively. Finally, the diagonal matrix $\mathbf{D}_{ps}$ with entries $p_{\text{S},k}$ regulates each source transmit power, while p_{R} denotes the relay's average transmit power.

B. Channel Estimation

In order to efficiently apply detection, precoding and LI mitigation techniques, channel estimates $\tilde{\mathbf{H}}_{\text{SR}}$, $\tilde{\mathbf{H}}_{\text{LI}}$ and $\tilde{\mathbf{H}}_{\text{RD}}$ of the true $\mathbf{H}_{\text{SR}}$, $\mathbf{H}_{\text{LI}}$ and $\mathbf{H}_{\text{RD}}$ are required. Applying any of the estimation algorithms proposed in the literature [5], we assume an error in the channel matrices modelled by $\mathcal{E}_{\mathbf{H}_{\text{SR}}}, \mathcal{E}_{\mathbf{H}_{\text{RD}}}, \mathcal{E}_{\mathbf{H}_{\text{LI}}} \sim \mathcal{CN}(0,\epsilon_{\mathbf{H}}^2)$. The difference between the non-ideal estimates and the true channel values are

$$\mathbf{H}_{\text{SR}} = \tilde{\mathbf{H}}_{\text{SR}} + \mathcal{E}_{\mathbf{H}_{\text{SR}}}; \mathbf{H}_{\text{LI}} = \tilde{\mathbf{H}}_{\text{LI}} + \mathcal{E}_{\mathbf{H}_{\text{LI}}}; \mathbf{H}_{\text{RD}} = \tilde{\mathbf{H}}_{\text{RD}} + \mathcal{E}_{\mathbf{H}_{\text{RD}}}. \tag{3}$$

No errors in the large-scale fading matrices are assumed. The impact of any hardware imperfection at the relay is modeled by means of an additive error [4] in the transmitted vector given by $\mathbf{t} = \tilde{\mathbf{t}} + \mathcal{E}_{\mathbf{t}}$, where $\tilde{\mathbf{t}}$ is the vector to be transmitted after baseband filtering and where all the elements of $\mathcal{E}_{\mathbf{t}} \sim \mathcal{CN}(0,\epsilon_{\mathbf{t}}^2)$ are assumed to be uncorrelated with $\tilde{\mathbf{t}}$. It is

worth mentioning that the covariance of the LI term $\sqrt{p_{\text{R}}}\mathbf{H}_{\text{LI}}\mathbf{t}$ in (1) is controlled by the parameters p_{R}, σ_{LI}^2, $\epsilon_{\mathbf{t}}^2$ and $\epsilon_{\mathbf{H}}^2$.

C. Detection and Precoding

Let us consider for the moment an equivalent received signal $\hat{\mathbf{r}}$, whose LI component has been minimized. Hence, we are in the presence of an uplink channel where $K << N_{rx}$. Under this assumption, linear processing techniques are proven to perform close to optimal [2]. We consider the usage of zero-forcing (ZF) filtering for both detection and beamforming [6]. Using the estimation of $\mathbf{G}_{\text{SR}}$, denoted by $\tilde{\mathbf{G}}_{\text{SR}} = \hat{\mathbf{H}}_{\text{SR}}\mathbf{D}_{\text{SR}}^{1/2}$, the estimated symbols after a ZF filter $\mathbf{W}_{\text{zf}}$ are given by

$$\hat{\mathbf{x}} = \mathcal{Q}(\mathbf{W}_{\text{zf}}\hat{\mathbf{r}}) = \mathcal{Q}((\tilde{\mathbf{G}}_{\text{SR}}^H\tilde{\mathbf{G}}_{\text{SR}})^{-1}\tilde{\mathbf{G}}_{\text{SR}}^H\hat{\mathbf{r}}), \tag{4}$$

where $\mathcal{Q}(\cdot)$ is a symbol-wise quantizer to the constellation set. Upon detection based on $\hat{\mathbf{x}}$, the estimated symbols are forwarded to the destinations $\mathbf{y}_d$, which we assume to have very limited processing capabilities. Thus, a ZF precoder $\mathbf{A}_{\text{zf}}$ is used. The corresponding expression is

$$\hat{\mathbf{t}} = \mathbf{A}_{\text{zf}}\hat{\mathbf{x}} = \alpha_{\text{zf}}\tilde{\mathbf{G}}_{\text{RD}}^H(\tilde{\mathbf{G}}_{\text{RD}}\tilde{\mathbf{G}}_{\text{RD}}^H)^{-1}\hat{\mathbf{x}}, \tag{5}$$

where $\tilde{\mathbf{G}}_{\text{RD}} = \mathbf{D}_{\text{RD}}^{1/2}\tilde{\mathbf{H}}_{\text{RD}}$ is the estimation of the true $\mathbf{G}_{\text{RD}}$ and $\alpha_{\text{zf}} = \left(\mathbb{E}(tr\{(\tilde{\mathbf{G}}_{\text{RD}}\tilde{\mathbf{G}}_{\text{RD}}^H)^{-1}\})\right)^{-\frac{1}{2}}$ is a scalar chosen to normalize the power of $\hat{\mathbf{t}}$, i.e., $\mathbb{E}(\hat{\mathbf{t}}^H\hat{\mathbf{t}}) = 1$. From the definition of $\tilde{\mathbf{H}}_{\text{RD}} \sim \mathcal{CN}(0, 1+\epsilon_{\mathbf{H}}^2)$, α_{zf} comes as (cf. appendix A)

$$\alpha_{\text{zf}} = \sqrt{\frac{(N_{tx} - K)}{\sum_{k=1}^{K}\left(\beta_{\text{RD},k}(1+\epsilon_{\mathbf{H}}^2)\right)^{-1}}}. \tag{6}$$

III. LOOPBACK INTERFERENCE MITIGATION

A. Linear Filtering

We now evaluate the problem of suppressing LI, using linear processing. Under the model given by (1), we are interested in minimizing the LI term $\sqrt{p_{\text{R}}}\mathbf{H}_{\text{LI}}\mathbf{t}$, while preserving the signal $\mathbf{G}_{\text{SR}}\mathbf{D}_{ps}^{1/2}\mathbf{x}$ and taking into account the covariance of the noise vector $\mathbf{R}_{\mathbf{n}_r}$ at the relay. Hence, the mean square error between the desired signal and the one received by the relay station is $\mathbb{E}((\mathbf{G}_{\text{SR}}\mathbf{D}_{ps}^{1/2}\mathbf{x} - \hat{\mathbf{r}})(\mathbf{G}_{\text{SR}}\mathbf{D}_{ps}^{1/2}\mathbf{x} - \hat{\mathbf{r}})^H)$. Using the method proposed in [4], we consider a linear prefilter $\mathbf{F}_{\text{tx}}$ and postfilter $\mathbf{F}_{\text{rx}}$, such that $\tilde{\mathbf{t}} = \mathbf{F}_{\text{tx}}\hat{\mathbf{t}}$ and $\hat{\mathbf{r}} = \mathbf{F}_{\text{rx}}\mathbf{r}$. The filter that finds the MMSE when $\mathbf{F}_{\text{rx}}$ is fixed satisfies (cf. appendix B)

$$\mathbf{F}_{\text{rx}}\tilde{\mathbf{H}}_{\text{LI}}\mathbf{F}_{\text{tx}} = \mathbf{0}, \tag{7}$$

and when $\mathbf{F}_{\text{tx}}$ is fixed becomes

$$\mathbf{F}_{\text{rx}} = \tilde{\mathbf{G}}_{\text{SR}}\mathbf{D}_{ps}\tilde{\mathbf{G}}_{\text{SR}}^H(\tilde{\mathbf{G}}_{\text{SR}}\mathbf{D}_{ps}\tilde{\mathbf{G}}_{\text{SR}}^H + p_{\text{R}}\tilde{\mathbf{H}}_{\text{LI}}\mathbf{R}_{\mathbf{t}}\tilde{\mathbf{H}}_{\text{LI}}^H + \mathbf{R}_{\mathbf{n}_r})^{-1}, \tag{8}$$

where $\mathbf{R}_{\mathbf{t}} = \mathbf{F}_{\text{tx}}\mathbf{A}_{\text{zf}}\mathbf{R}_{\hat{\mathbf{x}}}\mathbf{A}_{\text{zf}}^H\mathbf{F}_{\text{tx}}^H + \epsilon_{\mathbf{t}}^2\mathbf{I}$. The solution provided by (7) corresponds to the null-space projection already presented in [4], and in general finding a reliable solution for large dimensions is hard. On the contrary, (8) is a closed-form expression containing the channel estimates and covariance matrices of transmitted vectors and noise, that are assumed to be known. Under the considered model and for a given set of parameters, a trade-off in the e2e BER is expected: on the one

hand, a lower transmitted power at the relay p_R reduces the LI power and diminishes any impact of the estimation errors and radio-frequency impairments; on the other hand, a higher p_R leads to improved signal-to-noise ratios (SNR) levels at the destinations and hence a lower BER in the forward link channel. Thus, for a fixed source transmit power, an optimal choice for p_R that minimizes the e2e BER can be anticipated.

B. Optimal Power Allocation

The level of interference suffered by the relay station depends critically on its transmission power and directly perturbs the e2e performance. Thus, finding the optimal power that meets the requirements of the system is desired. Let us consider the achievable rate for each individual link. The procedure done in [7] allows us to obtain an expression for the transmission link rate between each source and destination

$$R_k = \min\{R_{\text{SR},k}, R_{\text{RD},k}\}, \tag{9}$$

where $R_{\text{SR},k}$ and $R_{\text{RD},k}$ denote the achievable rates between the sources and the relay and between the relay and the destinations, respectively. To obtain R_k, we start by considering the received signal at the relay before detection given by

$$
\begin{aligned}
y_{\text{r},k} = & \sqrt{p_{\text{S},k}}(\mathbf{W}_{\text{zf}}\mathbf{F}_{\text{rx}})_k^T \mathbf{g}_{\text{SR},k} x_k \\
& + \sum_{j\neq k}^{K} \sqrt{p_{\text{S},j}}(\mathbf{W}_{\text{zf}}\mathbf{F}_{\text{rx}})_k^T \mathbf{g}_{\text{SR},j} x_j \\
& + \sqrt{p_{\text{R}}}(\mathbf{W}_{\text{zf}}\mathbf{F}_{\text{rx}})_k^T \mathbf{H}_{\text{LI}}\mathbf{t} + (\mathbf{W}_{\text{zf}}\mathbf{F}_{\text{rx}})_k^T \mathbf{n}_r,
\end{aligned}
\tag{10}
$$

where $\mathbf{g}_{\text{SR},k}$ denotes the kth column of $\mathbf{G}_{\text{SR}}$. The received signal at each destination link before detection is given by

$$
\begin{aligned}
y_{\text{d},k} = & \sqrt{p_{\text{R}}}\mathbf{g}_{\text{RD},k}^T (\mathbf{F}_{\text{tx}}\mathbf{A}_{\text{zf}})_k \hat{x}_k \\
& + \sqrt{p_{\text{R}}} \sum_{j\neq k}^{K} \mathbf{g}_{\text{RD},k}^T (\mathbf{F}_{\text{tx}}\mathbf{A}_{\text{zf}})_j \hat{x}_j + n_{d,k}.
\end{aligned}
\tag{11}
$$

Both expressions may be seen as a known mean gain times the desired signal plus an uncorrelated effective noise term that includes channel impairment effects, interpair and loopback interference, and Gaussian noise. A technique commonly used in large MIMO systems [8] is to approximate the effective noise term by a Gaussian noise term under the central limit theorem, and which gives a good approximation for $R_{\text{SR},k}$ as in (13) and for $R_{\text{RD},k}$ as in (14).

The goal is to find the system power allocation, i.e., compute the required power transmitted by both sources and relay, such that the desired rate for each communication pair k is guaranteed. Moreover, peak power constraints should be satisfied and overall power consumption taken into consideration; in other words, the system energy efficiency, defined as $\text{EE} = \sum_{k=1}^{K} R_k / (p_{\text{R}} + \sum_{k=1}^{K} p_{\text{S},k})$, should be maximized. This can be described as the following optimization problem

$$
\begin{aligned}
\min_{p_{\text{R}}, p_{\text{S},1}\cdots p_{\text{S},K}} \quad & \sum_{k=1}^{K} p_{\text{S},k} + p_{\text{R}}, \\
\text{s.t.} \quad & R_k \geq R_{0,k}, \ k=1,...,K; \\
& 0 \leq p_{\text{S},k} \leq p_{\text{S}_{0,k}}, \ k=1,...,K; \\
& 0 \leq p_{\text{R}} \leq p_{\text{R}_0},
\end{aligned}
\tag{12}
$$

where $R_{0,k}$ and $p_{\text{S}_{0,k}}$ are the required rate and peak power for pair k respectively, and p_{R_0} is the relay station peak power. The solution to problem (12) involves the computation of the channel statistics which are the output of involved expressions due to the non-linear dependence of the optimization variables in $\mathbf{F}_{\text{rx}}$. For that reason, we propose a simple and computationally efficient iterative algorithm, based on linear programming, capable of obtaining the referred OPA. The algorithm computes the channel statistics coefficients in (13) and (14), the loopback mitigation filters and the power vectors iteratively until the optimal power vector is reached. The procedure is summarized in algorithm 1.

Algorithm 1

---

1. Initialization: Set $i = 1$; initialize powers $p_{\text{S},k,1} = p_{\text{S}_{0,k}}$ and $p_{\text{R},0} = p_{\text{R}_0}$; define L as the total number of iterations and set N_{it} as the number of channel realizations per iteration.

2. Iteration i:

  1) Compute channel statistics:

  for N_{it} **do**

    i) Generate $\mathbf{G}_{\text{SR}}$, $\mathbf{G}_{\text{RD}}$, $\mathbf{H}_{\text{LI}}$, $\mathbf{W}_{\text{zf}}$ and $\mathbf{A}_{\text{zf}}$

    ii) Compute filter $\mathbf{F}_{\text{rx}}$ with $p_{\text{S},k,i}$ and $p_{\text{R},i}$

    iii) Compute instantaneous rate coefficients for all k pairs:

    $\text{MV}_{\text{SR},k}$, $\text{V}_{\text{SR},k}$, $\text{MP}_{\text{SR},(k,j)}$, $\text{LI}_{\text{SR},k}$ and $\text{AN}_{\text{SR},k}$ (as in (13))

    $\text{MV}_{\text{RD},k}$, $\text{V}_{\text{RD},k}$, $\text{MP}_{\text{RD},k}$ and $\text{AN}_{\text{RD},k}$ (as in (14))

  end for

  2) Average to obtain channel statistics.

  3) Solve the linear program (12) with the coefficients found in step 2) to obtain the new $p_{\text{S},k,i}$ and $p_{\text{R},i}$.

  4) Set $p_{\text{S},k,i+1} = p_{\text{S},k,i}$ and $p_{\text{R},i+1} = p_{\text{R},i}$.

3. Check: If $i = L$ end algorithm, else set $i = i + 1$.

IV. NUMERICAL RESULTS

In this section, we compare the performance of the proposed filters with special emphasis on the e2e link reliability and on the efficiency of the power allocation scheme. All results are obtained via Monte Carlo simulation, using uncoded MIMO.

A. System Parameters

We assume a symmetric system with the same number of users K on both sides with one antenna each and $N = N_{tx} = N_{rx}$ antennas at the relay. Transmission and reception at the relay are done in the same time-slot and frequency band, and an arbitrary processing delay $d \geq 1$ is assumed such that at time instant i: $\hat{\mathbf{x}}[i] = f(\mathbf{x}[i-d])$. We consider $\mathbf{F}_{\text{rx}}$ as in (8) and $\mathbf{F}_{\text{tx}} = \mathbf{I}$. Without loss of generality, we also consider $\sigma_{\text{LI}}^2 = 1$. The SNR at the relay is defined as $\text{SNR}_{\text{R}} = \frac{\sum_{k=1}^{K} \beta_{\text{SR},k} p_{\text{S},k}}{\sigma_{\mathbf{n}_r}^2}$.

B. Transmitted Power vs. BER

We start by evaluating the performance of the system in terms of BER at both the relay and destinations, considering

$$R_{\text{SR},k} = \log_2 \left(1 + \frac{p_{\text{S},k} \text{MV}_{\text{SR},k}}{p_{\text{S},k} \text{V}_{\text{SR},k} + \sum_{j\neq k}^{K} p_{\text{S},j} \text{MP}_{\text{SR},(k,j)} + p_{\text{R}} \text{LI}_{\text{SR},k} + \text{AN}_{\text{SR},k}} \right), \tag{13}$$

where $\text{MV}_{\text{SR},k} = |\mathbb{E}\{(\mathbf{W}_{\text{zf}}\mathbf{F}_{\text{rx}})_k^T \mathbf{g}_{\text{SR},k}\}|^2$, $\text{V}_{\text{SR},k} = \mathbb{V}\text{ar}\{(\mathbf{W}_{\text{zf}}\mathbf{F}_{\text{rx}})_k^T \mathbf{g}_{\text{SR},k}\}$, $\text{MP}_{\text{SR},(k,j)} = \mathbb{E}\{|(\mathbf{W}_{\text{zf}}\mathbf{F}_{\text{rx}})_k^T \mathbf{g}_{\text{SR},j}|^2\}$, $\text{LI}_{\text{SR},k} = \mathbb{E}\{\| \mathbf{w}_{\text{zf},k}^T \mathbf{F}_{\text{rx}} \mathbf{H}_{\text{LI}} \mathbf{F}_{\text{tx}} \mathbf{A}_{\text{zf}} \|^2\}$ and $\text{AN}_{\text{SR},k} = \sigma_{n_r}^2 \mathbb{E}\{\| (\mathbf{W}_{\text{zf}}\mathbf{F}_{\text{rx}})_k \|^2\}$.

$$R_{\text{RD},k} = \log_2 \left(1 + \frac{p_{\text{R}} \text{MV}_{\text{RD},k}}{p_{\text{R}} \text{V}_{\text{RD},k} + p_{\text{R}} \text{MP}_{\text{RD},k} + \text{AN}_{\text{RD},k}} \right), \tag{14}$$

where $\text{MV}_{\text{RD},k} = |\mathbb{E}\{\mathbf{g}_{\text{RD},k}^T (\mathbf{F}_{\text{tx}}\mathbf{A}_{\text{zf}})_k\}|^2$, $\text{V}_{\text{RD},k} = \mathbb{V}\text{ar}\{\mathbf{g}_{\text{RD},k}^T (\mathbf{F}_{\text{tx}}\mathbf{A}_{\text{zf}})_k\}$, $\text{MP}_{\text{RD},k} = \sum_{j\neq k}^{K} \mathbb{E}\{|\mathbf{g}_{\text{RD},k}^T (\mathbf{F}_{\text{tx}}\mathbf{A}_{\text{zf}})_j|^2\}$ and $\text{AN}_{\text{RD},k} = \sigma_{n_d}^2$.

only small-scale fading, that is $\mathbf{D}_{\text{SR}}^{1/2} = \mathbf{D}_{\text{RD}}^{1/2} = \mathbf{I}$. Firstly, for a given SNR_{R} and fixed uniform transmitted power $p_{\text{S,k}} = 1$ for all k, we are interested in studying different allocated powers at the relay p_{R} and for an increasing number of antennas N. For comparison purposes, we implemented natural isolation (NI) (when $\mathbf{F}_{\text{rx}} = \mathbf{F}_{\text{tx}} = \mathbf{I}$, which corresponds to ignoring the LI component) and the HD counterpart (when $\mathbf{t} = 0$ and $\mathbf{F}_{\text{rx}} = \mathbf{F}_{\text{tx}} = \mathbf{I}$). Setting the variance in the errors to be $\epsilon_{\mathbf{t}}^2 = \epsilon_{\mathbf{H}}^2 = 10^{-3}$, curves of BER for different numbers of antennas and relay power p_{R} are depicted in figure 2. It can be seen that for an increasing number of antennas, the system using the MMSE filter becomes more robust to the LI, showing a BER performance closer to the HD for higher values of power p_{R} than when compared with NI. We then evaluate the e2e BER in terms of the allocated power p_{R} and using the same setup as before. For this purpose, we set the variance of the noise at the destinations $\sigma_{\mathbf{n}_d}^2$, quantize the received signals in (2) and compare them with the symbols transmitted by the sources. The results are shown in figure 3. One can confirm that for a given configuration there is an optimal choice for the power at the relay that minimizes the e2e BER. Moreover, it is concluded that a larger number of antennas attains the minimum BER with less power.

C. Optimal Power Allocation Algorithm

Finally, we evaluate the results of the proposed algorithm in section III-B, finding the OPA that meets the rate constraints of each link, while minimizing the overall power of the system. Large-scale fading effects are now taken into consideration, more precisely, $\beta_{\text{SR},k}$ and $\beta_{\text{RD},k}$ are independent variables generated from a log-normal distribution with mean value $m = 1$ and standard deviation $\sigma = 6$ dB. Additionally, we set the normalized peak power of the sources and relay as $p_{\text{S}_{0,k}} = 3$ dB and $p_{\text{R}_0} = 10$ dB, respectively. In addition, parameters $N_{it} = 10^3$ and $L = 5$ are empirically defined, such that both channel statistics and output powers in algorithm 1 are good approximations of their real value. The performance of algorithm 1 is determined in terms of EE for the case where interference effects are disregarded (OPA-NI) and when the MMSE filter is used (OPA-MMSE). This method is also compared with the situation where an optimal uniform power allocation (OUPA) is used, which corresponds to consider $p_{\text{S},k} = p_{\text{S}}$, for all k, in (12). Figure 4 shows the curves of average EE for different values of desired e2e sum-rate,
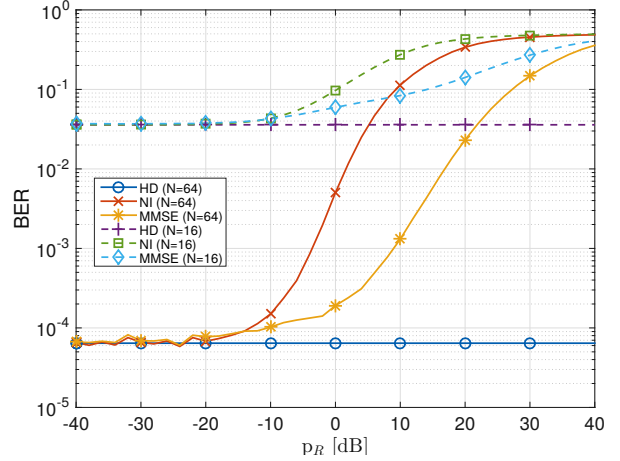
Figure 2. BER performance at the relay for different numbers of antennas N, $K = 5$ pairs, $\text{SNR}_{\text{R}} = 8$ dB and 16-QAM.
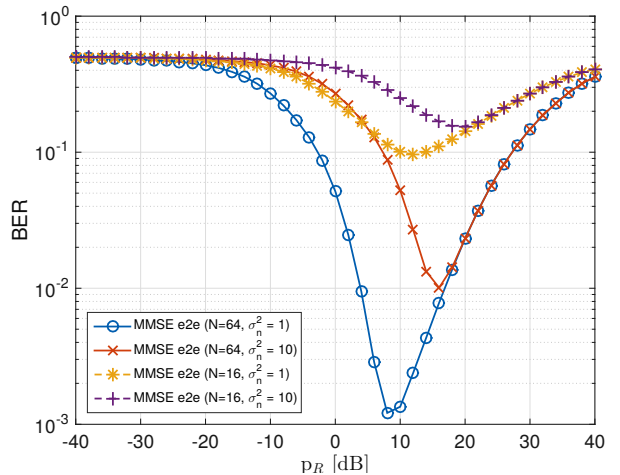
Figure 3. End-to-end BER performance for different N and $\sigma_{\mathbf{n}_d}^2$, $K = 5$ pairs, $\text{SNR}_{\text{R}} = 8$ dB and 16-QAM.

defined as $\sum_{k=1}^{K} R_{0,k}$, where the individual required rates $R_{0,k}$ are taken from a discrete uniform distribution.

As one may see, for the same desired sum-rate, the EE of OPA-MMSE is improved significantly when compared to OUPA, while guaranteeing that no link is in outage ($R_k \geq R_{0,k}$ for all k). Furthermore, since MMSE filter effectively reduces the LI, a lower amount of energy is spent to achieve the same sum-rate when compared with OPA-NI, an
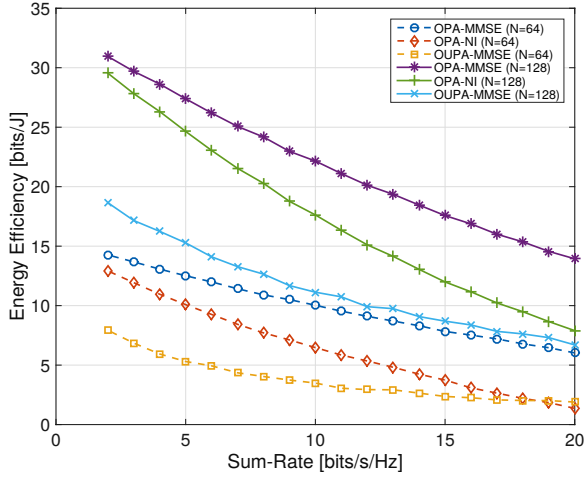
Figure 4. Energy efficiency for different power allocation schemes, filters and number of antennas. The remaining parameters are $\sigma_{\mathbf{n}_d}^2 = 1$, $K = 10$ pairs and $SNR_R = 16$ dB.

effect that becomes more preponderant in higher rates regimes.

V. Conclusions

This paper studies an independent single-antenna multipair communication system using an in-band full-duplex relay station equipped with a large number of antennas. Considering imperfect channel state information, the relay applies two stages of filtering. Firstly, inner ZF filters are used in order to perform both detection and beamforming. Secondly, an MMSE outer filter is designed to mitigate LI, while preserving the e2e channel. We assess the effectiveness of the MMSE filter via simulation in terms of BER at the relay and comparing it with NI and the HD counterpart. Moreover, in the e2e BER analysis, an optimal choice for the power at the relay was identified. We further confirmed the low complexity, high performance capabilities of linear processing in systems employing massive MIMO. We provide the expressions for the e2e achievable rate of the studied setup and propose a low complexity iterative algorithm to perform OPA. The latter takes into consideration large-scale fading effects, peak power constraints and the rate requirements for each individual pair. The results were evaluated in terms of EE, and the system employing MMSE filtering outperformed both NI and OUPA, specially for high sum-rates.

Appendix

A. Normalization factor α_{zf}

The objective is to design α_{zf} such that $\mathbb{E}\{\mathbf{t}^H\mathbf{t}\} = 1$. Using its definition $\mathbf{t} = \alpha_{zf}\tilde{\mathbf{G}}_{RD}^H(\tilde{\mathbf{G}}_{RD}\tilde{\mathbf{G}}_{RD}^H)^{-1}\hat{\mathbf{x}} = \alpha_{zf}\mathbf{P}\hat{\mathbf{x}}$, and the fact that $\mathbf{C} = \mathbb{E}\{\hat{\mathbf{x}}\hat{\mathbf{x}}^H\} = \mathbf{I}$, it follows that

$$\mathbb{E}\{\mathbf{t}^H\mathbf{t}\} = \alpha_{zf}^2\mathbb{E}\{(\mathbf{P}\hat{\mathbf{x}})^H(\mathbf{P}\hat{\mathbf{x}})\} = \alpha_{zf}^2 tr\{\mathbf{P}\mathbf{C}\mathbf{P}^H\} = 1. \quad (15)$$

Exploiting $\mathbf{C} = \mathbf{I}$, it follows that $tr\{\mathbf{P}\mathbf{C}\mathbf{P}^H\}$ is given by $tr\{(\tilde{\mathbf{G}}_{RD}\tilde{\mathbf{G}}_{RD}^H)^{-1}\} = tr\{(\mathbf{D}_{RD}\tilde{\mathbf{H}}_{RD}\tilde{\mathbf{H}}_{RD}^H)^{-1}\}$. Now noting that $\mathbf{W} = \tilde{\mathbf{H}}_{RD}\tilde{\mathbf{H}}_{RD}^H$ is a central Wishart matrix, where the columns of $\tilde{\mathbf{H}}_{RD} \in \mathbb{C}^{K \times N_{tx}}$ are zero-mean complex Gaussian vectors with covariance matrix $(1 + \epsilon_{\mathbf{H}}^2)\mathbf{I}$, it follows from [9,

Lemma 2.10] that

$$\mathbb{E}\{(tr\{\mathbf{D}_{RD}\mathbf{W}\})^{-1}\} = \frac{\sum_{k=1}^{K}\left(\beta_{RD,k}(1 + \epsilon_{\mathbf{H}}^2)\right)^{-1}}{N_{tx} - K}. \quad (16)$$

Merging (15) and (16), the factor α_{zf} in (6) follows.

B. $\mathbf{F}_{tx}$ and $\mathbf{F}_{rx}$ expressions

The error covariance matrix $\mathbf{Q}$ of the relay input signal is given by $\mathbb{E}((\mathbf{G}_{SR}\mathbf{x} - \hat{\mathbf{r}})(\mathbf{G}_{SR}\mathbf{x} - \hat{\mathbf{r}})^H)$, yielding [4]

$$\begin{aligned}\mathbf{Q} =&(\mathbf{I} - \mathbf{F}_{rx})\tilde{\mathbf{G}}_{SR}\mathbf{D}_{ps}\tilde{\mathbf{G}}_{SR}^H(\mathbf{I} - \mathbf{F}_{rx})^H \\ &+ p_R\mathbf{F}_{rx}\tilde{\mathbf{H}}_{LI}\mathbf{R}_t\tilde{\mathbf{H}}_{LI}^H\mathbf{F}_{rx}^H + \mathbf{F}_{rx}\mathbf{R}_{n_r}\mathbf{F}_{rx}^H,\end{aligned} \quad (17)$$

where $\mathbf{R_t} = \mathbf{R_{\tilde{t}}} + \mathbf{R_{\mathcal{E}_t}} = \mathbb{E}(\tilde{\mathbf{t}}\tilde{\mathbf{t}}^H) + \mathbb{E}(\mathcal{E}_t\mathcal{E}_t^H) = \mathbf{F}_{tx}\mathbf{A}_{zf}\mathbf{R}_{\hat{\mathbf{x}}}\mathbf{A}_{zf}^H\mathbf{F}_{tx}^H + \epsilon_{\mathbf{t}}^2\mathbf{I}$. Using the fact that the function $f = tr\{\mathbf{Z}\mathbf{A}_0\mathbf{Z}^H\mathbf{A}_1\}$ has derivative given by $\frac{d}{d\mathbf{Z}^*}f = \mathbf{A}_1\mathbf{Z}\mathbf{A}_0$ [10, Table 4], one may easily find $\frac{\partial}{\partial\mathbf{F}_{tx}^*}tr\{\mathbf{Q}\} = \mathbf{0}$ and $\frac{\partial}{\partial\mathbf{F}_{rx}^*}tr\{\mathbf{Q}\} = \mathbf{0}$. Fixing $\mathbf{F}_{tx}$ comes

$$\begin{aligned}\frac{\partial}{\partial\mathbf{F}_{rx}^*}tr\{\mathbf{Q}\} =&\mathbf{F}_{rx}(p_R\tilde{\mathbf{H}}_{LI}\mathbf{R}_t\tilde{\mathbf{H}}_{LI}^H + \mathbf{R}_{n_r}) \\ &- (\mathbf{I} - \mathbf{F}_{rx})\tilde{\mathbf{G}}_{SR}\mathbf{D}_{ps}\tilde{\mathbf{G}}_{SR}^H = \mathbf{0},\end{aligned} \quad (18)$$

which, when solved, gives (8). Fixing $\mathbf{F}_{rx}$ leads to

$$\begin{aligned}\frac{\partial}{\partial\mathbf{F}_{tx}^*}tr\{\mathbf{Q}\} =&\mathbf{D}_{ps}\mathbf{F}_{rx}\tilde{\mathbf{H}}_{LI}(p_R\mathbf{F}_{rx}^H\tilde{\mathbf{H}}_{LI}^H\mathbf{F}_{tx}\mathbf{A}_{zf}\mathbf{R}_{\hat{\mathbf{x}}}\mathbf{A}_{zf}^H) \\ =&\mathbf{D}_{ps}\mathbf{F}_{rx}\tilde{\mathbf{H}}_{LI}(\mathbf{F}_{rx}\tilde{\mathbf{H}}_{LI}\mathbf{F}_{tx})\mathbf{A}_{zf}\mathbf{R}_{\hat{\mathbf{x}}}\mathbf{A}_{zf}^H = \mathbf{0},\end{aligned} \quad (19)$$

whose solution is given by (7).

References

[1] S. Hong *et al.*, "Applications of self-interference cancellation in 5G and beyond," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 114–121, 2014.
[2] H. Q. Ngo, H. A. Suraweera, M. Matthaiou, and E. G. Larsson, "Multipair full-duplex relaying with massive arrays and linear processing," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 9, pp. 1721–1737, 2014.
[3] J. S. Lemos, F. A. Monteiro, I. Sousa, and A. Rodrigues, "Recursive least squares self-interference mitigation in broadband MIMO full-duplex relays," 2015. [Online]. Available: http://arxiv.org/abs/1505.00850
[4] T. Riihonen, S. Werner, and R. Wichman, "Mitigation of loopback self-interference in full-duplex MIMO relays," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 5983–5993, 2011.
[5] M. Duarte, C. Dick, and A. Sabharwal, "Experiment-driven characterization of full-duplex wireless systems," *IEEE Trans. Wireless Commun.*, vol. 11, no. 2, pp. 4296–4307, 2012.
[6] H. Yang and T. L. Marzetta, "Performance of conjugate and zero-forcing beamforming in large-scale antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 31, pp. 172–179, 2013.
[7] T. Riihonen, S. Werner, and R. Wichman, "Transmit power optimization for multiantenna decode-and-forward relays with loopback self-interference from full-duplex operation," *in Proc. Annu. Asilomar Conf. Signals, Syst., Comput.*, pp. 1408–1412, 2011.
[8] B. Hassibi and B. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 951–963, 2003.
[9] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communications," *Found. Trends Commun. Inf. Theory*, vol. 1, no. 1, pp. 1–182, 2004.
[10] A. Hjorungnes and D. Gesbert, "Complex-valued matrix differentiation: Techniques and key results," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2740–2746, 2007.