SUM-RATE MAXIMIZATION OF TWO-WAY AMPLIFY-AND-FORWARD RELAY NETWORKS WITH IMPERFECT CHANNEL STATE INFORMATION

Yupeng Jia and Azadeh Vosoughi

Department of Electrical and Computer Engineering, University of Rochester, NY14627, USA

ABSTRACT

Considering a two-way amplify-and-forward (AF) relay network and aiming to simultaneously maximize the two users’ mutual information lower bounds in the presence of channel estimation errors, we study the Pareto-front of users’ mutual information lower bounds. Based on the Pareto-front we investigate the optimal power allocation among the two users and the relay, as well as the optimal power allotment between training and data symbols that maximize the average sum-rate lower bound, and explore the variations of these optimal power allocation as the relay position changes. We also show that the mean square error (MSE) of channel estimation is minimized when training vectors transmitted by the two users are orthogonal.

Index Terms— Two-way relay channel, channel estimation, bi-objective optimization, Pareto-front, power allocation, sum-rate lower bound maximization, training design

1. INTRODUCTION

Two-way relay networks have gained a great attention due to its ability to reduce the spectral loss caused by half-duplex relaying [1]. Current literature has studied several problems concerning with two-way amplify-and-forward (AF) relaying, including sum-rate maximization [1][2][3], power consumption minimization and network lifetime maximization [4], and beamforming design for multi-antenna two-way relaying [5][6][7]. These works have mainly assumed perfect channel state information (CSI) at relay and users’ terminals.

The literature on two-way relaying system with imperfect CSI is sparse. We note that the channel estimation problem and the negative impact of channel estimation errors on the information rate in two-way relaying systems are different from those of point-to-point and one-way relaying communications, since in the former systems channel estimation errors lead to self-interference and thus further reduce information rate. For DF two-way relaying systems [8] proposes an iterative directed self-interference aided channel estimation method, that does not require training symbols. The advantage of this method, however, is offset by the increasing complexity required for decoding and iterative self-interference suppression. In [9], the authors propose optimal training sequence design for AF two-way relaying via minimizing the channel mean square error (MSE). [10] considers the optimal training sequence design for AF two-way relaying and analyzes the received signal-to-noise ratio (SNR) in the presence of imperfect CSI. In our previous work [11], we considered the effect of channel estimation errors on the sum-rate in an AF two-relaying network, where the relay appends an extra training symbol to the transmit packet, to facilitate the channel estimation at the terminals.

Considering a different and more bandwidth efficient training transmission strategy from [11], we study the simultaneous max-

Fig. 1. System model of two-way AF relaying systemimization of the two users’ mutual information lower bounds in the presence of channel estimation errors and find the Pareto-front. Based on the Pareto-front we investigate the optimal power allocation among the three nodes, as well as the optimal power allotment between training and data symbols that maximize the average sum-rate lower bound, and explore the variations of these optimal power allocation as the relay position changes.

2. SYSTEM MODEL

We consider a two-way relay network, consisting of terminals \( T_A \) and \( T_B \) and a half duplex relay node \( R \), where each node is equipped with a single antenna (Fig.1). The complex baseband channel between \( T_A \) and \( R \) is denoted as \( h_a \) and the one between \( T_B \) and \( R \) is denoted as \( h_b \). To capture the effect of pathloss and Rayleigh flat fading we assume \( h_a \sim \mathcal{CN}(0, \sigma_{h_a}^2) \) and \( h_b \sim \mathcal{CN}(0, \sigma_{h_b}^2) \). The variances \( \sigma_{h_a}^2 = G_i/d_{i,R}^2 \) and \( \sigma_{h_b}^2 = G_R/d_{R,i}^2 \) where \( d_{i,R} \) denotes the distance between \( T_i \) and \( R \), \( \epsilon \) is the pathloss exponent, and \( G \) is a constant that depends on antenna gains and the wavelength. Although \( \epsilon \) and \( G \) may vary for each channel link, throughout this paper it is assumed that \( \epsilon \) and \( G \) are identical for all links. We assume block fading model, in which \( h_a \), \( h_b \) are constant within one transmission block, after which they change to different independent values that hold for another block.

Let \( s_i = [s_{N_i}^1, s_{N_i}^2, ..., s_{N_i}^{N_i}] \) be the block transmitted by \( T_i \) for \( i \in \{ A, B \} \). Also, let \( N_A \) and \( N_B \) denote the number of training and data symbols, respectively, in a transmission block of \( N \) symbols. We assume that each transmission block is augmented as \( s_i = [s_i^a, s_i^b] \) in which \( s_i^a = [s_{N_a}^1, s_{N_a}^2, ..., s_{N_a}^a] \) and \( s_i^b = [s_{N_a}^1, s_{N_a}^2, ..., s_{N_a}^b] \) denote the \( 1 \times N_A \) and \( 1 \times N_B \) row vectors including training and data symbols, respectively, and \( s_i^a \sim \mathcal{CN}(0, \rho_i^a I_{N_A}) \) with \( \rho_i^a \) defined in (1). Suppose \( s_i^a \) and \( s_i^b \) are independent. We note that Gaussian input has been adopted before [12][13] to study the effect of channel estimation on capacity, although it may not be the optimal distribution in terms of maximizing mutual information given imperfect CSI.

We consider a transmission protocol consisting of two phases (Fig.1). During phase I, \( T_A \) and \( T_B \) transmit, respectively, vectors \( s_A \) and \( s_B \) to \( R \) and \( R \) receive \( y_R = h_A s_A + h_B s_B + n_R \). During phase II, \( R \) amplifies \( y_R \) and forwards \( x = f y_R \) to \( T_A \) and \( T_B \), where \( f \) is the amplifying factor. Assuming channel reciprocity for \( h_a \) and \( h_b \) terminals \( T_A \) and \( T_B \) receive

\[
\begin{align*}
y_A &= h_a x + n_A = f g s_B + f h_A s_A + f h_b n_R + n_A \\
y_B &= h_b x + n_B = f g s_A + f h_B s_B + f h_a n_R + n_B
\end{align*}
\]
in which $g = h_a h_b$ is the overall complex relay channel between $T_A$ and $T_B$, $h_a = h_a^2$ and $h_B = h_B^2$. The noise vectors are $n_i \sim CN(0, \sigma_i^2 I_N)$ for $i \in \{A, B\}$. The noises and fading channels are mutually independent. To keep the complexity of training assisted AF two-way relay system low, we assume that only the terminals are equipped with channel estimator. Terminal $T_A$ first uses received training symbols to estimate the unknown channels $h_A, g$ at $h_B, g$. It uses these channel estimates to suppress self-interference, caused by back-propagation, and then recovers the data symbols $s_i^R, (s_i^D)$. Due to channel estimation errors self-interference at the terminals cannot be completely canceled and the residual self-interference reduces the rate, compared to the case when self-interference is perfectly canceled.

Let $\mathcal{P}_A, \mathcal{P}_B, \mathcal{P}_R$ be the average power constraints per transmission block at $T_A, T_B$, and $T_R$. Suppose the total transmit power of the system $P = \mathcal{P}_A + \mathcal{P}_B + \mathcal{P}_R$ is fixed. Let $\rho_A, \rho_B, \rho$ for $i \in \{A, B\}$ denote the average transmit power per training symbol, per data symbol, and per symbol in a transmission block, and $\mathcal{P}_i, \mathcal{P}_D$ be the training power and the average power per transmission block. Conservation of time and energy yields $N = N_t + N_d$ and $\mathcal{P}_i = \mathcal{P}_D$ for $i \in \{A, B\}$ where $\mathcal{P}_D = \rho N = E[|s|^2]$. Let $\alpha \in \{0 < \alpha < 1\}$ be the fraction of the total transmission power that $T_i$ assigns to its $N_d$ data symbols. We have

$$\mathcal{P}_i = \alpha \mathcal{P}_i, \quad \mathcal{P}_D = (1 - \alpha) \mathcal{P}_i$$

$$\rho_A = \frac{\mathcal{P}_A}{N}, \quad \rho_B = \frac{\mathcal{P}_B}{N}, \quad \rho = \frac{\mathcal{P}_D}{N}$$

Considering the average power constraint $\mathcal{P}_R = E[\|xx^H\|] = f^T (\mathcal{P}_A \sigma_{h_A}^2 + \mathcal{P}_B \sigma_{h_B}^2 + \sigma_{h_R}^2 N) \mathcal{R}$ we find

$$f = \sqrt{\mathcal{P}_A \sigma_{h_A}^2 + \mathcal{P}_B \sigma_{h_B}^2 + \sigma_{h_R}^2 N}$$

We assume that $\mathcal{R}$ does not estimate the channels, hence it cannot utilize $h_a, h_b$ in choosing $f$. Given the knowledge of the channel variances, however, $\mathcal{R}$ can incorporate $\sigma_{h_A}^2$ and $\sigma_{h_B}^2$ in designing $f$. The received signal vectors at $T_A$ and $T_B$ can be partitioned as $y_A = [y_A^t, y_A^d]$ and $y_B = [y_B^t, y_B^d]$, where $y_A^t, y_B^t, i \in \{A, B\}$ are the received training and the received data row vectors

$$y_A^t = f q_A^t s_i^R + f h_a n_i^R + n_A^t$$
$$y_A^d = f q_A^d s_i^D + f h_a n_i^D + n_A^d$$
$$y_B^t = f q_B^t s_i^R + f h_b n_i^R + n_B^t$$
$$y_B^d = f q_B^d s_i^D + f h_b n_i^D + n_B^d$$

where we define the $1 \times 2$ effective channel vectors $q_i = [h_i, g_i]$, the $2 \times N_t$ training matrix $S_i = [s_i^t, s_i^d]$ and the $2 \times N_d$ data matrix $S_i = [s_i^d, s_i^d]$. The vectors $n_i^t, n_i^d$ are noise terms during training and data transmission, and are mutually independent. Terminal $T_i$ uses $y_i^t$ to estimate the effective channel vector $\hat{q}_i$ and suppress self-interference caused by the term $f h_a n_i^d$. Using the estimate $\hat{q}_i$ and the self-interference suppressed received vector $z_i^d = y_i^d - f h_a n_i^d$, terminal $T_i$ recovers $s_i^R$ for $i, j \in \{A, B\}, i \neq j$.

### 3. CHANNEL ESTIMATION

We model the channel vectors at terminals $T_i$ as [14]

$$q_i = q_i^t + q_i^d, \quad \hat{q}_i = [\hat{h}_i, \hat{g}_i], \quad i \in \{A, B\}$$

where $q_i$ is the estimation vector and $\hat{q}_i$ is the zero mean error vector with covariance matrix $R_{\hat{q}_i} = E[\hat{q}_i^H \hat{q}_i]$. Let $R_{\\hat{q}_i} = E[\hat{q}_i^H q_i]$ be the channel covariance matrix, where $R_{\\hat{q}_i}$ is a diagonal matrix with diagonal elements of $\sigma_{h_A}^2$ and $\sigma_{h_B}^2$. Given $y_i^t$ at $T_A$, the LMMSE estimate $\hat{q}_i$ is [14]

$$\hat{q}_i = y_i^t E[y_i^H y_i^t]^{-1} E[y_i^H q_i]$$

$$= y_i^t (S_i^H R_{\\hat{q}_i} S_i + (f^2 \sigma_{h_A}^2 \sigma_{h_B}^2 + \sigma_{h_R}^2 I_N))^{-1} S_i^H R_{\\hat{q}_i}$$

and the estimation error covariance matrix is [14]

$$\mathcal{R}_{\hat{q}_i} = \left( \frac{f^2 S_i^H S_i^H}{(f^2 \sigma_{h_A}^2 \sigma_{h_B}^2 + \sigma_{h_R}^2 I_N)} \right)^{-1}$$

We can find $\hat{q}_i$ and $\mathcal{R}_{\hat{q}_i}$ by substituting $y_i^t, A, \hat{q}_i, \sigma_{h_A}^2, \sigma_{h_B}^2$ in (3) and (4) with $y_i^t, R_{\hat{q}_i}, \sigma_{h_A}^2, \sigma_{h_B}^2$. Under the power constraint on the training matrix $S_i$ we wish to design $S_i$ such that $tr(R_{\hat{q}_i})$ is minimized. Schwartz’s inequality states that for a positive definite matrix $A$ we have $|A|_{ii} \geq \frac{\sigma_{A}}{\sigma_{A}^2}$ in which the equality holds when $A$ is diagonal [13]. Noting that $R_{\hat{q}_i}$ is diagonal, we realize that $tr(R_{\hat{q}_i})$ is minimized if and only if $S_i^H S_i$ is diagonal. However

$$S_i^H S_i = \left[ \begin{array}{cc} 1 & 0 \\ 0 & S_i^H S_i \\ \end{array} \right]$$

implying that orthogonal training vectors for which $\rho = 0$ minimize $tr(R_{\hat{q}_i})$. When $\rho = 0$, $R_{\hat{q}_i}$ is diagonal and $tr(R_{\hat{q}_i})$ attains its minimum value of

$$\text{tr}(R_{\hat{q}_i}) = \sigma_{h_A}^2 + \sigma_{h_B}^2 \left( 1 + \frac{f^2 \rho_A \sigma_{h_A}^2}{f^2 \sigma_{h_A}^2 \sigma_{h_B}^2 + \sigma_{h_R}^2} \right)^{-1}$$

Therefore, we conclude that the design that training vectors sent from $T_A$ and $T_B$ are orthogonal is the optimal, in the sense of minimizing MSE.

### 4. MUTUAL INFORMATION LOWER BOUND

In this section, we obtain the mutual information lower bound at each terminal given imperfect CSI. Substituting (2) into $y_i^t$ and $y_i^d$ and subtracting self-interference at each terminal, we obtain

$$z_i^t = y_i^t - f h_a s_i^d = f g_A s_i^d + f h_a s_i^d + f h_b n_i^R + n_i^d$$

$$z_i^d = y_i^d - f h_a s_i^d = f g_B s_i^d + f h_b n_i^R + f h_b s_i^d + n_i^d$$

where $w_i^d$ are zero mean noise vectors with diagonal covariance matrices $R_{w_i^d} = E[\|w_i^d\|^H] = \sigma_{w_i^d}^2 I_N$ and the variances $\sigma_{w_i^d}$ and $\sigma_{w_i^d}$ are given below

$$\sigma_{w_i^d}^2 = f^2 \sigma_{h_A}^2 + f^2 \sigma_{h_B}^2 + f^2 \sigma_{h_A}^2 \sigma_{h_B}^2 + \sigma_{h_R}^2$$

$$\sigma_{w_i^d}^2 = f^2 \sigma_{h_B}^2 + f^2 \sigma_{h_B}^2 + f^2 \sigma_{h_A}^2 \sigma_{h_B}^2 + \sigma_{h_R}^2$$
\[ \rho_A^{\text{eff}} = \frac{P_A^2 \rho_A^2 \sigma_A^2}{\sigma_w^2} = \frac{P_R P_B^s \sigma_A^2 + P_R \sigma_A^2 \sigma_h^2 + P_R \sigma_h^2 \sigma_n^2 N_A + P_A \sigma_h^2 \sigma_n^2 N_d + \sigma_A^2 \sigma_n^2 N_d}{\sigma_w^2} \]
\[ \rho_B^{\text{eff}} = \frac{P_B^2 \rho_B^2 \sigma_B^2}{\sigma_w^2} = \frac{P_R P_A^s \sigma_B^2 + P_R \sigma_B^2 \sigma_h^2 + P_R \sigma_h^2 \sigma_n^2 N_B + P_A \sigma_h^2 \sigma_n^2 N_d + \sigma_B^2 \sigma_n^2 N_d}{\sigma_w^2} \]

Note that due to channel estimation error perfect self-interference cancelation is not feasible. One can show that for LMMSE channel estimates where \( E\{q_A|q_A\} = 0 \) [14], the additive noises \( w_A^d \) and \( w_B^d \) are conditionally uncorrelated. Furthermore, we have \( E\{s_B^d \mid w_A^d|q_A\} = 0 \) and \( E\{s_A^d \mid w_B^d|q_A\} = 0 \), since
\[
E\{s_B^d \mid w_A^d|q_A\} = f E\{s_B^d E\{q_A\}\}
\]
\[
+ f E\{s_B^d \mid w_A^d\} E\{\hat{h}_A\mid q_A\} + E\{s_B^d \mid f h_s n_i + n_d\} = 0
\]

Given \( q_1 \) at terminal \( T_j \), we denote the conditional mutual information between data vector \( x_i^j \) transmitted by terminal \( T_i \) and the corresponding received data vector \( x_i^j \) at terminal \( T_j \) as \( I(s_i^j; z_i^j|q_1) \) for \( i, j \in \{A, B\} \) and \( j \neq i \). The quantity \( I(x_i^j; z_i^j|q_1) \) is the mutual information of a known channel system \( q_1 \), subject to the additive noise vector \( w^d \) that is uncorrelated with \( s_i^j \). Lemma 1 provides mutual information lower bounds, given the LMMSE channel estimates at the terminals.

**Lemma 1**: The conditional mutual information per transmission block at terminals \( T_A \) and \( T_B \) can be lower bounded as \( I(s_B^d; x_A^d|q_A) \geq I_{\text{lower}}^A \) and \( I(s_A^d; x_B^d|q_B) \geq I_{\text{lower}}^B \)
\[
I_{\text{lower}}^A = \frac{N_B}{2N} \log (1 + \rho_A^{\text{eff}} |\hat{g}_A|^2) \]
\[
I_{\text{lower}}^B = \frac{N_A}{2N} \log (1 + \rho_B^{\text{eff}} |\hat{g}_B|^2)
\]

where \( \hat{g}_A = \frac{\hat{h}_A}{\sigma_h^2} \), \( \hat{g}_B = \frac{\hat{h}_B}{\sigma_h^2} \) are the normalized channel estimates, \( \rho_A^{\text{eff}} \) and \( \rho_B^{\text{eff}} \) are shown in (7) and (8) and the variances of the LMMSE channel estimates can be found using orthogonality principle [14] \( \sigma_h^2 = \sigma_A^2 \sigma_h^2 = \sigma_B^2 \sigma_h^2 = \sigma_h^2 \).

Proof of Lemma 1 is omitted due to space limitation, we refer interested readers to [11] [12] [15].

### 5. PARETO-FRONT AND SUM-RATE MAXIMIZATION

We aim to maximize the mutual information lower bounds simultaneously under total transmit power constraint. This task can be accomplished via scalarizing the bi-objective optimization problem [16] and finding the Pareto-front. Using the Pareto-front we can solve the average sum-rate lower bound maximization problem and find the effect of channel estimation errors upon the solution. The optimization problem is

\[
\max_{(\alpha, P_A, P_B, P_R)} \{ I_{\text{lower}}^A, I_{\text{lower}}^B \}
\]
subject to \( P_A + P_B + P_R \leq P, \quad \alpha \in (0, 1) \)

We aim to find the optimal trade-off surface (Pareto-front). To simplify the optimization, we relax the constraints and let \( \alpha \) be any given constant between (0,1). We can rewrite (9) as the following constrained weighted-sum maximization problem [16]

\[
\max_{(P_A, P_B, P_R)} \lambda_1 I_{\text{lower}}^A + \lambda_2 I_{\text{lower}}^B
\]
subject to \( P_A + P_B + P_R \leq P \)
\[
\lambda_1 + \lambda_2 = 1, \quad \alpha = \text{constant} \in (0, 1)
\]

### 6. NUMERICAL RESULTS

In this section, we present numerical results to evaluate the effects of channel estimation errors upon Pareto-front and average sum-rate lower bound maximization. We consider a linear network model and define \( d_{A, R}/d_{A, B} = \eta \) and \( d_{B, R}/d_{A, B} = 1 - \eta \), where \( d_{A, B} \) is the distance between \( T_A \) and \( T_B \). The noise variances \( \sigma_{h_A}^2 = \sigma_{h_B}^2 = \sigma_{n_A}^2 \) and the system SNR is defined as \( SNR = \bar{P} = 15dB \). The block length \( N \) = 20 and the training length \( N_t = 2 \) [15]. The pathloss exponent \( \varepsilon = 3 \) and \( G = 1 \). We run 100 channel realizations to measure the average performance.

Fig. 3 shows the Pareto-front with three different relay locations \( \eta = 0.3, 0.5, 0.7 \) and fixed \( 1 - \alpha = 0.2 \). The Pareto-front for two-way relaying system with perfect CSI is plotted to serve as a bench-
7. CONCLUSIONS

Considering a two-way AF relay network and aiming to simultaneously maximize the two users’ mutual information lower bounds in the presence of channel estimation errors, we studied the Pareto-front of users’ mutual information lower bounds. Based on the Pareto-front we investigated the optimal power allocation among the two users and the relay, as well as the optimal power allotment between training and data symbols that maximize the average sum-rate lower bound. Our simulations show that the optimal training power fraction is about 0.2 and this value stays constant as the relay position changes, while the optimal relay location is at the midpoint between $T_A$ and $T_B$.

8. REFERENCES


