

Low-complexity beamforming design for full-duplex relay-assisted cooperative NOMA

Hanlin Liao¹ and Teruyuki Miyajima^{1, a)}

¹ Graduate School of Science and Engineering, Ibaraki University,
4–12–1 Nakanarusawa, Hitachi-shi, Ibaraki, 316–8511, Japan

a) teruyuki.miyajima.spc@vc.ibaraki.ac.jp

Abstract: In this paper, we propose a beamforming design method for a cooperative NOMA downlink system wherein a relay node assisting a far user operates in full-duplex mode to achieve high spectral efficiency. The proposed method, based on the zero-forcing (ZF) principle, suppresses both self-interference and inter-user interference. Simulation results show that the proposed method is considerably superior to the conventional transmit ZF method, achieving performance comparable to that of the optimum method while reducing complexity.

Keywords: non-orthogonal multiple-access, full-duplex mode, cooperative relaying, beamforming, zero-forcing criterion

Classification: Wireless Communication Technologies

References

- [1] M. Zeng, W. Hao, O.A. Dobre, and Z. Ding, “Cooperative NOMA: State of the art, key techniques and open challenges,” *IEEE Network*, vol. 34, no. 5, pp. 205–211, July 2020. DOI: [10.1109/MNET.011.1900601](https://doi.org/10.1109/MNET.011.1900601)
- [2] M. Mohammadi, X. Shi, B.K. Chalise, Z. Ding, H.A. Suraweera, C. Zhong, and J.S. Thompson, “Full-duplex non-orthogonal multiple access for next generation wireless systems,” *IEEE Commun. Mag.*, vol. 57, no. 5, pp. 110–116, May 2019. DOI: [10.1109/MCOM.2019.1800578](https://doi.org/10.1109/MCOM.2019.1800578)
- [3] C. Zhong and Z. Zhang, “Non-orthogonal multiple access with cooperative full-duplex relaying,” *IEEE Commun. Lett.*, vol. 20, no. 12, pp. 2478–2481, Dec. 2016. DOI: [10.1109/LCOMM.2016.2611500](https://doi.org/10.1109/LCOMM.2016.2611500)
- [4] V. Aswathi and A.V. Babu, “Full/Half duplex cooperative NOMA under imperfect successive interference cancellation and channel state estimation errors,” *IEEE Access*, vol. 7, pp. 179961–179984, Dec. 2019. DOI: [10.1109/ACCESS.2019.2959001](https://doi.org/10.1109/ACCESS.2019.2959001)
- [5] X. Pei, H. Yu, M. Wen, S. Mumtaz, S.A. Otaibi, and M. Guizani, “NOMA-based coordinated direct and relay transmission with a half-duplex/full-duplex relay,” *IEEE Trans. Commun.*, vol. 68, no. 11, pp. 6750–6760, Nov. 2020. DOI: [10.1109/TCOMM.2020.3017002](https://doi.org/10.1109/TCOMM.2020.3017002)
- [6] Z. Mobini, M. Mohammadi, B.K. Chalise, H.A. Suraweera, and Z. Ding, “Beamforming design and performance analysis of full-duplex cooperative NOMA systems,” *IEEE Trans. Wireless Commun.*, vol. 18, no. 6, pp. 3295–3311, May 2019. DOI: [10.1109/TWC.2019.2913425](https://doi.org/10.1109/TWC.2019.2913425)

1 Introduction

Various promising technologies such as non-orthogonal multiple access (NOMA) and cooperative relaying, have been extensively explored to fulfill the diverse demands of future wireless systems. NOMA is a practical solution to improve the spectral efficiency of a downlink multiuser system by accommodating multiple users on the same frequency band and at the same time. On the contrary, cooperative relaying effectively improves reception reliability and expands the communication area, especially when there is no direct link between a base station (BS) and users owing to obstacles or heavy shadowing. Combining NOMA and cooperative relaying can take advantage of the benefits of both [1]. The spectral efficiency of relay-assisted cooperative NOMA (RA-CNOMA) can be further improved by employing full-duplex (FD) relaying [2]. FD relaying has the potential to double the spectral efficiency compared with half-duplex (HD) relaying.

The performance of full-duplex RA-CNOMA (FD-RA-CNOMA) is mainly bounded by two types of interference: self-interference (SI) at the relay and inter-user interference (IUI) at the near user. In [3], they showed that FD-RA-CNOMA is superior to half-duplex RA-CNOMA (HD-RA-CNOMA) if both SI and IUI can be canceled perfectly. It was shown in [4] that FD-RA-CNOMA outperforms HD-RA-CNOMA at low SNRs and that IUI has a greater impact on performance than SI in the case of incomplete interference cancellation. In [5], they studied the effects of imperfect channel state information and power allocation at the BS. All the above-mentioned studies considered FD-RA-CNOMA with a single-antenna relay. If we employ a relay with multiple antennas, we can expect interference suppression in spatial domain by beamforming (BF). In [6], the authors considered a suboptimum BF design method, called transmit zero-forcing (TZF) that cancels the SI by projecting the transmit signal to the null space of the received signal at the relay. Although this method is computationally efficient, its performance degrades owing to IUI. In addition, the authors in [6] proposed an optimum BF design method that can suppress both SI and IUI by maximizing the received SINR at the near user while satisfying the target SINR at the far user. However, the drawback of this method is its high computational costs.

In this study, we propose a new BF design method. The basic idea of the proposed method is to suppress IUI by the transmit BF and SI by the receive BF. The proposed method is based on the ZF criterion and is thus computationally efficient. Through simulations, we demonstrate that the performance of the proposed method is considerably superior to that of TZF and is comparable to that of the optimum BF.

2 Full-duplex relay-assisted CNOMA system model

Figure 1 shows the FD-RA-CNOMA system model using a relay with multiple antennas. We consider a downlink system that supports two users, where the near user U1 directly communicates with the BS, while the far user U2 requires the assistance of an FD relay employing a decode-and-forward (DF) protocol. Successive interference cancellation (SIC) is adopted at U1. We assume that BS, U1, and U2 are equipped with a single antenna, and the relay is equipped with N_T transmit antennas and N_R receive antennas.

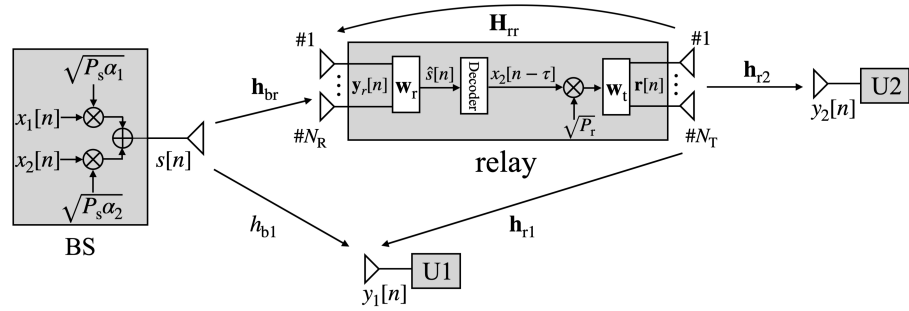


Fig. 1. System model.

BS transmits a sum of information symbols to both users at time n , which is given by $s[n] = \sqrt{P_s\alpha_1}x_1[n] + \sqrt{P_s\alpha_2}x_2[n]$, where P_s is the BS transmit power, and $x_i[n], i \in \{1, 2\}$ is the information symbol intended for U_i , and α_i is the power allocation coefficient that satisfies $\alpha_1 + \alpha_2 = 1$.

The received signal at the relay can be written as $\mathbf{y}_r[n] = \sqrt{L(d_{br})}\mathbf{h}_{br}s[n] + \sqrt{\mu\rho}\mathbf{H}_{rr}\mathbf{r}[n] + \mathbf{n}_r[n]$, where $L(d)$ is the path loss at distance d , μ is the isolation of the transmit and receive antennas at the relay, ρ is the SI cancellation level including analog and digital cancellations, $\mathbf{r}[n]$ is the transmit signal of the relay, $\mathbf{n}_r[n]$ is the additive white Gaussian noise (AWGN) at the relay with $\mathbb{E}\{\mathbf{n}_r[n]\mathbf{n}_r^H[n]\} = \sigma_r^2\mathbf{I}_{N_R}$, $\mathbf{H}_{rr} \in \mathbb{C}^{N_R \times N_T}$ is the SI channel at the relay, and $\mathbf{h}_{br} \in \mathbb{C}^{N_R \times 1}$ and d_{br} are the channel coefficient and distance between BS and the relay, respectively. The relay employing the DF protocol first estimates $s[n]$ by applying a receive BF vector $\mathbf{w}_r \in \mathbb{C}^{N_R \times 1}$ with $\|\mathbf{w}_r\|^2 = 1$ on $\mathbf{y}_r[n]$ as follows

$$\hat{s}[n] = \mathbf{w}_r^H \mathbf{y}_r[n] = \sqrt{L(d_{br})}\mathbf{w}_r^H \mathbf{h}_{br}s[n] + \sqrt{\mu\rho}\mathbf{w}_r^H \mathbf{H}_{rr}\mathbf{r}[n] + \mathbf{w}_r^H \mathbf{n}_r[n]. \quad (1)$$

The second term on the right-hand side of (1) corresponds to SI. Next, the relay decodes the information symbol $x_2[n]$ intended for U_2 . After decoding, the relay forwards $x_2[n]$ to U_2 by applying the transmit BF vector $\mathbf{w}_t \in \mathbb{C}^{N_T \times 1}$ with $\|\mathbf{w}_t\|^2 = 1$ as $\mathbf{r}[n] = \sqrt{P_r}\mathbf{w}_t x_2[n - \tau]$, where P_r is the transmit power of the relay and τ is the time delay due to relay processing.

The received signal at U_1 is given by

$$y_1[n] = \sqrt{L(d_{b1})}h_{b1}s[n] + \sqrt{\beta P_r L(d_{r1})}\mathbf{h}_{r1}^T \mathbf{w}_t x_2[n - \tau] + n_1[n], \quad (2)$$

where β is the IUI cancellation level including analog and digital cancellations, $n_1[n]$ is the AWGN at U_1 with $\mathbb{E}\{|n_1[n]|^2\} = \sigma_1^2$, h_{b1} and d_{b1} are the channel and distance between the BS and U_1 , $\mathbf{h}_{r1} \in \mathbb{C}^{N_T \times 1}$ and d_{r1} are the channel and distance between the relay and U_1 , respectively. The second term on the right-hand side of (2) corresponds to IUI. U_1 carries out SIC, that is, U_1 first decodes the symbol $x_2[n]$ of U_2 and then subtracts it from the received signal to detect its symbol $x_1[n]$. Finally, the received signal at U_2 can be written as

$$y_2[n] = \sqrt{P_r L(d_{r2})}\mathbf{h}_{r2}^T \mathbf{w}_t x_2[n - \tau] + n_2[n], \quad (3)$$

where $n_2[n]$ is the AWGN at U_2 with $\mathbb{E}\{|n_2[n]|^2\} = \sigma_2^2$, $\mathbf{h}_{r2} \in \mathbb{C}^{N_T \times 1}$ and d_{r2} are the channel and distance between the relay and U_2 , respectively. U_2 decodes delayed symbol $x_2[n - \tau]$ from $y_2[n]$.

From (1), the received SINR at the relay to decode $x_2[n]$ is given by

$$\gamma_R^{x_2} = \frac{P_s \alpha_2 L(d_{br}) |\mathbf{w}_r^H \mathbf{h}_{br}|^2}{P_s \alpha_1 L(d_{br}) |\mathbf{w}_r^H \mathbf{h}_{br}|^2 + \mu \rho P_r |\mathbf{w}_r^H \mathbf{H}_{rr} \mathbf{w}_t|^2 + \sigma_r^2}. \quad (4)$$

From (2), the received SINR at U1 to decode $x_2[n]$ is given by

$$\gamma_1^{x_2} = \frac{P_s \alpha_2 L(d_{b1}) |h_{b1}|^2}{P_s \alpha_1 L(d_{b1}) |h_{b1}|^2 + \beta P_r L(d_{r1}) |\mathbf{h}_{r1}^T \mathbf{w}_t|^2 + \sigma_1^2}. \quad (5)$$

After U1 successfully adopts SIC, the received SINR at U1 to decode x_1 is given by

$$\gamma_1^{x_1} = \frac{P_s \alpha_1 L(d_{b1}) |h_{b1}|^2}{\beta P_r L(d_{r1}) |\mathbf{h}_{r1}^T \mathbf{w}_t|^2 + \sigma_1^2}. \quad (6)$$

From (3), the received SNR at U2 is given by $\gamma_2^{x_2} = P_r L(d_{r2}) |\mathbf{h}_{r2}^T \mathbf{w}_t|^2 / \sigma_2^2$. The achievable rate at U1 is then given by $R_1 = \log_2(1 + \gamma_1^{x_1})$. In dual-hop DF relaying, the weakest link determines the achievable rate. In addition, because $x_2[n]$ must be decoded at U1 for SIC, we can write the achievable rate for U2 as $R_2 = \log_2(1 + \min(\gamma_R^{x_2}, \gamma_2^{x_2}, \gamma_1^{x_2}))$.

3 Proposed beamforming design method

The purpose of the BF design is to suppress both SI in (1) and IUI in (2). To this end, we impose the following constraints.

$$\mathbf{h}_{r1}^T \mathbf{w}_t = 0 \quad \text{and} \quad \mathbf{w}_r^H \mathbf{H}_{rr} \mathbf{w}_t = 0. \quad (7)$$

Under these constraints, $\gamma_1^{x_1}$ in (6), hence R_1 , is independent of the choice of BF vectors. Now, we consider the following BF design problem.

$$\max_{\mathbf{w}_t, \mathbf{w}_r} R_2, \quad \text{s.t.} \quad \mathbf{h}_{r1}^T \mathbf{w}_t = 0, \quad \mathbf{w}_r^H \mathbf{H}_{rr} \mathbf{w}_t = 0, \quad \|\mathbf{w}_t\| = \|\mathbf{w}_r\| = 1. \quad (8)$$

Recall that $\gamma_1^{x_2}$ in (5) is also independent of the BF vectors and that $\gamma_2^{x_2}$ and $\gamma_R^{x_2}$ depend only on \mathbf{w}_t and \mathbf{w}_r , respectively. We propose to decompose problem (8) into the following two subproblems:

$$\max_{\mathbf{w}_t} |\mathbf{h}_{r2}^T \mathbf{w}_t|^2, \quad \text{s.t.} \quad \mathbf{h}_{r1}^T \mathbf{w}_t = 0, \quad \|\mathbf{w}_t\| = 1. \quad (9)$$

$$\max_{\mathbf{w}_r} |\mathbf{w}_r^H \mathbf{h}_{br}|^2, \quad \text{s.t.} \quad \mathbf{w}_r^H \mathbf{H}_{rr} \mathbf{w}_t = 0, \quad \|\mathbf{w}_r\| = 1. \quad (10)$$

We obtain \mathbf{w}_t and \mathbf{w}_r by solving these problems in order.

To solve (9), we adopt the ZF approach [6] that cancels IUI by projecting the transmit signal from the relay on the null space of the received signal at U1. Similarly, problem (10) can be solved by nulling the SI at the receive antenna of the relay. The resulting transmit and receive BF vectors can be represented as

$$\mathbf{w}_t = \frac{\mathbf{D} \mathbf{h}_{r2}^*}{\|\mathbf{D} \mathbf{h}_{r2}^*\|}, \quad \mathbf{D} = \mathbf{I}_{N_T} - \frac{\mathbf{h}_{r1}^* \mathbf{h}_{r1}^T}{\|\mathbf{h}_{r1}\|^2}, \quad (11)$$

$$\mathbf{w}_r = \frac{\mathbf{E} \mathbf{h}_{br}}{\|\mathbf{E} \mathbf{h}_{br}\|}, \quad \mathbf{E} = \mathbf{I}_{N_R} - \frac{\mathbf{H}_{rr} \mathbf{D} \mathbf{h}_{r2}^* \mathbf{h}_{r2}^T \mathbf{D}^H \mathbf{H}_{rr}^H}{\|\mathbf{H}_{rr} \mathbf{D} \mathbf{h}_{r2}^*\|^2}, \quad (12)$$

respectively.

The computational complexity of the proposed method is evaluated. The numbers of multiplications required to obtain \mathbf{w}_t and \mathbf{w}_r are $5N_T + 1$ and $N_T^2 + N_T N_R + 5N_R + 1$, respectively. This is slightly larger but comparable to the TZF method [6] whose complexities are $N_T N_R + 5N_T + 1$ and $2N_R$ for \mathbf{w}_t and \mathbf{w}_r , respectively. By contrast, the optimum BF design method [6] requires $N_T^3 + N_T^2 N_R + N_R^2 N_T + 5N_R^2 + 6N_T^2 + 3N_T N_R + 5N_R + 22$ multiplications and two $N_R \times N_R$ matrix inversions of complexity $O(N_R^3)$. Moreover, the optimum method is based on the semidefinite relaxation (SDR) approach, where SDR is solved with a worst-case complexity of $O(N_T^{4.5} \log \epsilon^{-1})$ given a solution accuracy $\epsilon > 0$, and the principal eigenvector of the SDR solution is computed with a complexity of $O(N_T^2)$.

4 Simulation results

We evaluated the FD-RA-CNOMA system using the proposed method by comparing it with other methods through computer simulations. The transmission bandwidth was 1.0 MHz, carrier frequency was 2.0 GHz, and noise power spectral density was -169 dBm/Hz. The path loss model $-128.1 - 37.6 \log_{10}(d_{ij})$ dB was used. The distances d_{b1}, d_{br}, d_{r1} , and d_{r2} were 0.1, 0.15, 0.18, and 0.15 km, respectively. We modeled the channel coefficients as $CN(0, 1)$ random variables. We set $\alpha_1 = 0.1$, $\alpha_2 = 0.9$, $N_T = N_R = 2$, $\mu = -30$ dB, and $\rho = -80$ dB. Quality-of-service (QoS) requirements for U1 and U2 were both 1.0 bps/Hz. We ran 10^5 simulations.

Figure 2 shows the outage probability (OP) and average achievable sum rate (SR) of different BF methods as a function of IUI cancellation level β . The OP is the probability that the rate of U1 or U2 falls below the QoS requirement. The transmit power at BS and relay were $P_s = P_r = 30$ dBm. The random BF method, where BF vectors are randomly chosen, is the worst owing to the influence of SI and IUI. The performance of the TZF method [6] worsens as the IUI cancellation level β increases because TZF does not suppress IUI. In contrast, the performance of the proposed BF method was unaffected by the IUI cancellation level. The proposed method achieved almost the same performance as the optimum SDR method [6].

In Fig. 3(a), the performance of the FD-RA-CNOMA system using the proposed method is compared with that of HD-RA-CNOMA using the maximum ratio combining (MRC) BF at the receiver and maximal ratio transmission (MRT) BF at the

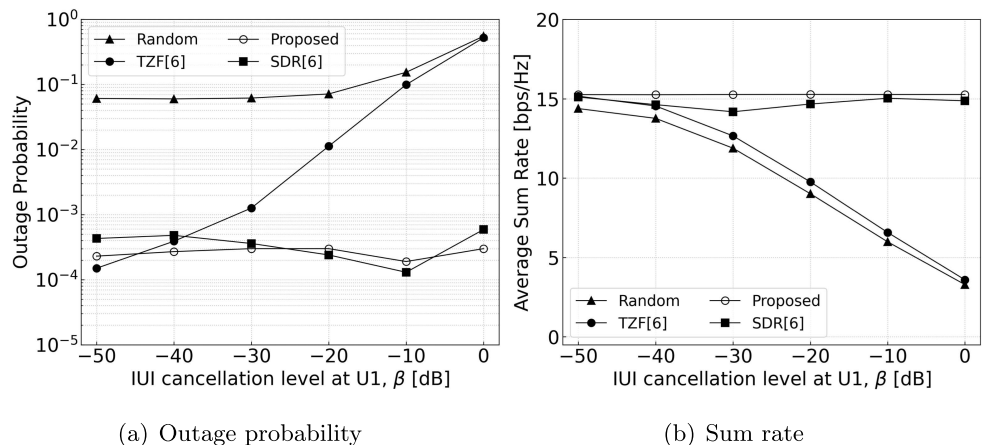
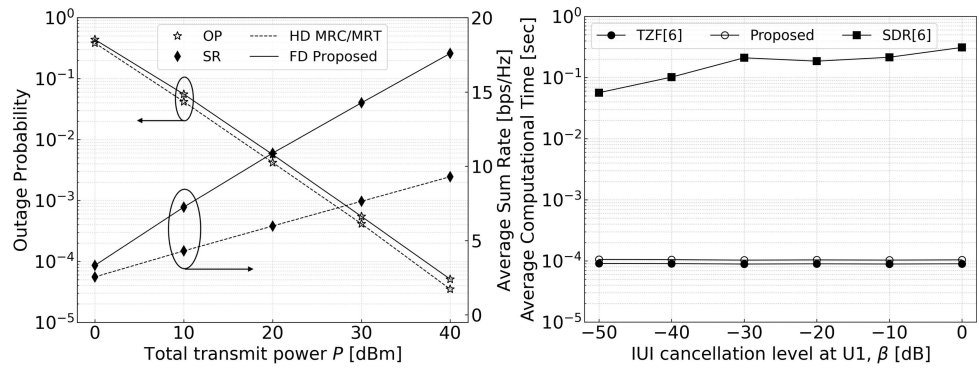


Fig. 2. Performance comparison with different methods.



(a) Comparison with HD

(b) Computational time

Fig. 3. Performance evaluation.

transmitter, where IUI cancellation level was $\beta = -30$ dB. The total transmit power P satisfied that $P_s = P_r = P/2$ in FD and $P_s = P_r = P$ in HD. In Fig. 3(a), as with HD, the proposed method shows that the OP improves as the total transmission power increases. FD with the proposed method can achieve approximately twice the rate of HD. Figure 3(b) shows a computational time comparison between the proposed method and the conventional methods [6]. Computer simulations were performed using Python 3.8 on a computer with a 128 GB memory and an AMD Ryzen Threadripper 3960X processor. We used CVXPY 1.2.2 as the solver for the optimum SDR method. The computational time required by the proposed method was almost the same as that of TZF and considerably reduced compared with that of the optimum SDR method.

5 Conclusion

In this paper, we proposed a BF design method for FD-RA-CNOMA, where both SI and IUI are suppressed. Simulation results show that the proposed method is superior to the conventional method and achieves almost the same performance as the optimum method while reducing complexity.

Acknowledgments

This study was supported by JSPS KAKENHI Grant Number JP20K04479.