Decode-Compress and Forward Relay: AWGN and Constellation Constrained Channels

Ahmed Attia Abotabl and Aria Nosratinia
Department of Electrical Engineering, University of Texas at Dallas, Richardson, TX 75083-0688 USA,
E-mail: ahmed.abotabl@utdallas.edu;aria@utdallas.edu.

***Abstract*—Even though there exists a lot of different transmission techniques, the capacity of the general relay channel is still unknown. This paper shows that combining decode-and-forward (DF) with compress-and-forward (CF) can have a performance advantage over each of the two techniques individually. Combining the two techniques goes back to the celebrated work of Cover and El-Gamal [1], more specifically, they gave an achievable rate for the DF-CF combination. In this paper, we re-derive the achievable rate of combining DF with CF in the full-duplex discrete memoryless relay channel. We derive the achievable rate in the AWGN relay channel as well as the constrained constellation AWGN relay channel. We show that even though in the AWGN channel, DF-CF combination does not provide any advantage, under constrained constellation, combining the two techniques has an advantage.**

I. Introduction

Shortly after the introduction of the relay channel by Van Der Meulen [2], Cover and El-Gamal [1] introduced the main transmission techniques of cooperation, namely, decode-and-forward (DF) and compress-and-forward (CF). More details about the transmission over the relay channel can be found in [3], [4]

Each of the two techniques, DF and CF has its own advantage and imposed constraints. For example, the decode-and-forward achieves the capacity of the degraded relay channel but it enforces the source to transmit with a rate that is decodable at the relay, a restriction that causes DF to fail when the source-relay channel is very weak. Compress-and-forward performs very-well when the relay-destination channel is very strong [5]. However, it only sends a description of the received signal using Wyner-Ziv coding [6] but the relay is not allowed to attempt noise cancellation.

The constraints of the two techniques act in an opposite manner. In order to overcome the restriction of each technique, the combination of the two of them in decode-compress-and-forward (DCF) transmission is studied. We first re-derive the achievable rate of DCF which was first presented by Cover and El-Gamal in [1, Theorem 7]. Second, we derive the achievable rate in the AWGN relay channel and show that DCF does not provide any performance gain over the best of DF and CF. Finally, we present the achievable rate under a constrained constellation AWGN relay channel, showing that DCF provides performance gain over DF and CF.

This work was supported in part by the NSF grants 1546969 and 1711689.

978-1-5386-3531-5/17/$31.00

II. Decode, Compress-and-Forward Technique

In this Section, we re-derive Theorem 7 [1] for the discrete memoryless relay channel and obtain the achievable rate for the AWGN relay channel as well as the constellation constrained AWGN relay channel.

A. Discrete Memoryless Full-Duplex Relay

For the discrete memory-less relay channel, the transmission technique is summarized in Fig. 1 over four transmission blocks. In each transmission block, the relay sends a compress-and-forward component superimposed on a decode-and-forward component, whereas the source sends a compress-and-forward component that is superimposed on new information to be sent to the relay which is superimposed on the assistance that the source provides to the relay transmission.

Codebook generation at the relay node: $2^{nR_{2d}}$ codewords are generated at the relay node independently and identically distributed (i.i.d.) according to a distribution $P_{U_{2d}}(u_{2d}) = \prod_{i=1}^{n} P_{U_{2di}}(u_{2di})$ where U_{2di} is the symbol number i in the codeword U_{2d}. For every codeword U_{2d}, a codebook of $2^{nR_{2c}}$ codewords are generated i.i.d. according to a distribution $P_{U_{2c}|U_{2d}}(u_{2c}|u_{2d}) = \prod_{i=1}^{n} P_{U_{2ci}|U_{2d}}(u_{2ci}|u_{2d})$.

Codebook generation at the source node: For every codeword U_{2d}, a codebook of size $2^{nR_{2d}}$ is generated i.i.d. with a distribution $P_{U_{1d}|U_{2d}}(u_{1d}|u_{2d}) = \prod_{i=1}^{n} P_{U_{1di}|U_{2d}}(u_{1di}|u_{2d})$. For every U_{1d}, a codebook of size $2^{nR_{1c}}$ codewords is i.i.d. generated with a distribution $P_{U_{1c}|U_{1d},U_{2d}}(u_{1c}|,u_{1d},u_{2d})$.

The source node: In block t, the source node splits the message $W^{(t)}$ into messages $W_d^{(t)}$ for decode-and-forward transmission and $W_c^{(t)}$ for compress-and-forward transmission. Subsequently, the source sends the codeword $X_1(W_d^{(t)}, W_d^{(t-1)}, W_c^{(t)})$ where $W_d^{(t-1)}$ is the decode-and-forward which is assumed to be perfectly decoded at the relay during transmission block $t-1$. In the first transmission block, the source sends $X_1(W_d^{(1)}, 1, W_c^{(t)})$.

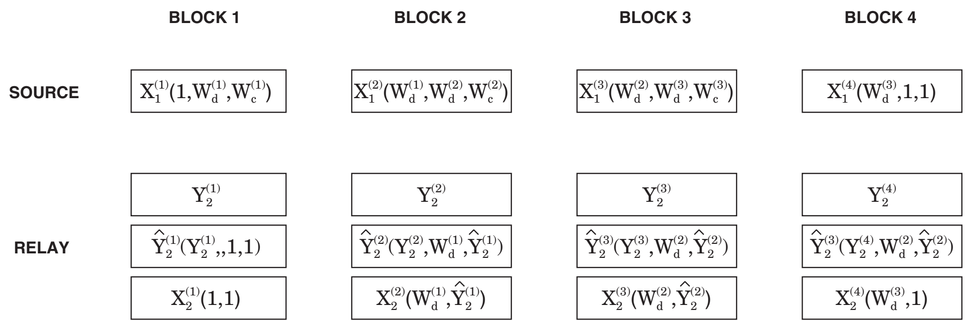

Fig. 1. Decode-Compress and Forward transmission over four transmission blocks.

The relay node: In transmission block $t-1$, the relay receives the signal $Y_2^{(t-1)}$. Subsequently, it decodes the message $W_d^{(t-1)}$ using its received signal $Y_2^{(t-1)}$ and $W_d^{(t-2)}$ which is assumed to be successfully decoded during transmission block $t-2$. Please note that this decoding is performed while treating $W_c^{(t-1)}$ as noise. After decoding $W_d^{(t)}$, the relay removes the effect of $W_d^{(t-1)}$ and $W_d^{(t-2)}$ from $Y_2^{(t-1)}$ and quantizes the resulting signal to generate $\hat{Y}_2^{(t)}$. In transmission block t, the relay receives $Y_2^{(t)}$ and similarly decodes W_d^t while treating $W_c^{(t)}$ as noise. In the transmission phase, in block t, the relay generates a codeword $U_{2d}(W_d^{(t)})$ and conditionally on U_{2d}, generates a codeword $X_2^{(t)}(W_d^{(t)}, \hat{Y}_2^{(t)})$ to be transmitted.

The destination node: In every block t, the destination receives $Y_3^{(t)}$. The destination uses backward decoding where it waits until the last transmission block B. In the last transmission block, destination receives $y_3^{(B)}$ which depends only on $W_d^{(B)}$ and $W_c^{(B)}$ and free of interference from any other message. The destination uses $y_3^{(B)}$ to decode $W_d^{(B)}$ while treating $W_c^{(B)}$ as noise. Subsequently, the destination removes the effect of $W_3^{(B)}$ from $Y_3^{(B)}$ and $y_3^{(B-1)}$ and use them jointly to decode $W_c^{(B)}$ and then decode $W_d^{(B-1)}$. This process continues until the destination decodes all the messages.

Achievable Rate The source-relay transmission should be with rates that allow the relay node to decode the message $W_d^{(t)}$ assuming that $W_d^{(t-1)}$ is decoded correctly at the relay during block $t-1$. In order for this to be satisfied, the rate R_{1d} should be

$$R_{1d} \leq I(U_{1d}; Y_2 | U_{2d}) \tag{1}$$

The message W_d is to be sent using decode-and-forward transmission, and hence, using the received signal at the destination Y_3, the destination should be able to decode W_d. Therefore, similar to the decode-and-forward analysis, the rate R_d should satisfy the following inequality

$$R_d \leq I(U_{1d}; Y_3 | U_{2d}) + I(U_{2d}; Y_3)$$
$$= I(U_{1d}, U_{2d}; Y_3) \tag{2}$$

At this point, we consider that the destination knows the message that corresponds to the decode-and-forward part. For the compress-and-forward part, the relay quantizes the received signal Y_2 after decoding U_{1d} to obtain a description $\hat{Y}_2$ using a Wyner-Ziv coding approach [6]. According to Shannon's rate-distortion theory [7, Chapter 13], the rate of the quantization codebook $\hat{Y}_2$ should be upper bounded by

$$R_Q(D) \leq I(Y_2; \hat{Y}_2 | X_2, U_{1d}) \tag{3}$$

The relay sends this quantized signal superimposed on U_{2d} to the destination. The destination should be able to decode $\hat{Y}_2$, which requires the quantization rate to satisfy

$$R_Q(D) \leq I(X_2; Y_3 | U_{2d}) \tag{4}$$

The destination now decodes the message corresponding to the compress-and-forward transmission W_c in a regular manner, and hence, for successful decoding, the rate R_{1c} should satisfy the following inequality

$$R_{1c} \leq I(X_1; \hat{Y}_2, Y_3 | X_2, U_{1d}) \tag{5}$$

The total transmission rate is then $R = R_d + R_c$. This rate is a function of the distributions of the variables involved in the bounds of R_d and R_c. Therefore, by combining these bounds we reach to the following theorem

The achievable rate is then given by the following Theorem:

Theorem 1: The achievable rate of the decode-compress and forward is given by

$$R \leq \min\left\{ I(U_{1d}; Y_2 | U_{2d}), I(U_{1d}, U_{2d}; Y_3) \right\} + I(X_1; \hat{Y}_2, Y_3 | X_2, U_{1d}) \tag{6}$$

subject to

$$I(Y_2; \hat{Y}_2 | X_2, U_{1d}) \leq I(X_2; Y_3) - I(U_{2d}; Y_3) \tag{7}$$

where

$$\begin{aligned} &P(y_3, y_2, \hat{y}_2, u_1, u_2, x_1, x_2) \\ =&P(u_2)P(u_{1d}|u_{2d})P(x_1|u_1)P(x_2|u_2)P(y_2|x_1). \\ &P(\hat{y}_2|u_1, x_2, y_2)p(y_3|x_1, x_2) \end{aligned} \tag{8}$$

B. AWGN Full-Duplex Relay

Assume that all the variables in the previous Section are Gaussian variables [1] while the source and relay have an average power constrained by P_1 and P_2 respectively. The DCF transmission in the AWGN relay channel is described in Fig. 2 where the source and relay signals are given by

$$X_1 = U_{1d} + \beta U_{2d} + U_{1c} \tag{9}$$

$$X_2 = U_{2d} + U_{2c} \tag{10}$$

respectively.

Each of the codewords U_{1d}, U_{2d}, U_{1c} and U_{2c} are normally distributed. The term βU_{2d} represents the assistance that the source provides to the relay destination transmission. This assistance depends on the correlation between $U_{1d} + \beta U_{2d}$ and U_{2d} which is denoted by ρ.

Remark 1: In decode-and-forward AWGN relay channel, the design variable that should be optimized to maximize the transmission rate is the correlation between the source and the relay transmissions. Whereas in compress-and-forward, once the source and relay signals are set to be normally distributed with variances equal to the source power and the relay power respectively, the transmission rate is obtained. Sitting the codewords to be normally distributed leaves the room for a lot of parameters to be optimized. These parameters are basically the power allocation of each component codeword at the source and the relay. For this reason, we fix the power of two signals, U_{1d} and U_{1c} which are the decode-and-forward the compress-and-forward components at the source respectively. In this way, the power of all the other signals can be obtain as a function of the power of U_{1d} and U_{1c} and the power constraint at the source and relay nodes. The design variable of the rate maximization problem becomes the power of U_{1d}, U_{1c} and the correlation ρ, and hence, the maximum rate can be obtained using exhaustive search over all possible power values of U_{1d} and U_{1c} and the correlation ρ.

Assuming that the power of U_{1d} is P_{1d} and the power of U_{1c} is P_{1c}, the power of U_{2d} is then given by

$$P_{2d} = \frac{P_1 - P_{1d} - P_{1c}}{\beta^2} \tag{11}$$

and the power of U_{2c} is

$$P_{2c} = P_2 - P_{2d} = P_2 - \frac{P_1 - P_{1d} - P_{1c}}{\beta^2} \tag{12}$$

where

$$\rho = \frac{E[(U_{1d} + \beta U_{2d})U_{2d}]}{\sqrt{(P_{1d} + \beta^2 P_{2d})P_{2d}}} \tag{13}$$

$$\beta = \sqrt{\frac{\rho^2 P_{1d} P_{2d}}{P_{2d}^2(1-\rho^2)}} \tag{14}$$

[1]Gaussian random variables are not necessarily optimal, therefore, the achievable rate is only a lower bound

Now, the signals $Y_2, \hat{Y}_2$ and Y_3 are given by

$$Y_2 = H_{12}X_1 + n_2 \tag{15}$$

$$\hat{Y}_2 = Y_2 + \hat{n} \tag{16}$$

$$Y_3 = H_{13}X_1 + H_{23}X_2 + n_3 \tag{17}$$

where n_2, n_3 and $\hat{n}$ are zero mean Gaussian noise with variance σ_2^2, σ_3^2 and $\hat{N}$ respectively.

Based on the previous characterization for each of the distributions of the variables involved in calculating the transmission rate, the achievable rate for the AWGN relay channel is given by the following theorem.

Theorem 2: The achievable rate of decode-compress and forward for the AWGN relay channel with all codewords normally distributed is given by

$$\begin{aligned} R \le \min\Bigg\{ &\frac{1}{2}\log\left(1 + \frac{|H_{12}|^2 P_{1d}}{|H_{12}|^2 P_{1c} + \sigma_2^2}\right), \\ &\frac{1}{2}\log\Bigg(1 + \frac{(P_{1d} + \beta^2 P_{2c})|H_{13}|^2}{\sigma_3^2} + \frac{P_{2d}|H_{23}|^2}{\sigma_3^2} \\ &+ 2\rho\sqrt{\frac{(P_{1d} + \beta^2 P_2)P_{2d}|H_{13}|^2|H_{23}|^2}{\sigma_3^4}}\Bigg)\Bigg\} \\ &+ \frac{1}{2}\log\Bigg[\Big((|H_{12}|^2 P_{1c} + \sigma_3^2 + \hat{N})(|H_{13}|^2 P_{1c} + \sigma_3^2) \\ &- (|H_{12}|^2|H_{13}|^2)P_{1c}^2\Big)/(\sigma_3^2 + \hat{N})\sigma_3^2\Bigg] \end{aligned} \tag{18}$$

where

$$\hat{N} = \frac{(|H_{12}|^2 P_{1c} + \sigma_2^2)(|H_{13}|^2(P_{1d} + P_{1c}) + \sigma_2^2)}{|H_{23}|^2 P_{2c}} \tag{19}$$

Proof: See Appendix A. ∎

In order to see the performance of DCF approach, we consider the same model that is considered in [8], namely, the source, relay and destination are all considered to be on a line. The source and destination locations are fixed while the relay is moving. For simplicity, only a path-loss model is considered at which the channel coefficient $H_{ij} = 1/d_{ij}^{\alpha}$ where d_{ij} is the distance between node i and node j and α is the path-loss coefficient. The distance between the source and the destination is fixed to $d_{13} = 1$ while the distances d_{12} and d_{23} depend on the relay location where $d_{23} = 1 - d_{12}$. In Fig. 3, we draw the achievable rate of different transmission technique as a function of the distance between the source and the relay d_{12} where negative values of d_{12} means that the relay is in the side of the source that is far from the destination and positive values means that the relay is between the source and destination.

C. Constellation-Constrained Full-Duplex relay

For the discrete input AWGN relay channel, the rate in Theorem 1 can be obtained via numerical integrations. The optimizing distribution may require an exhaustive

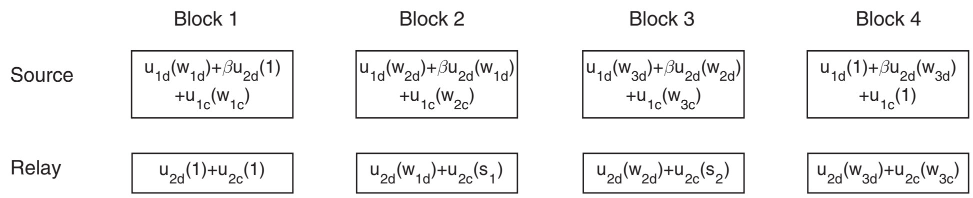

Fig. 2. Decode-compress and forward transmission for the AWGN full-duplex relay channel over four blocks.

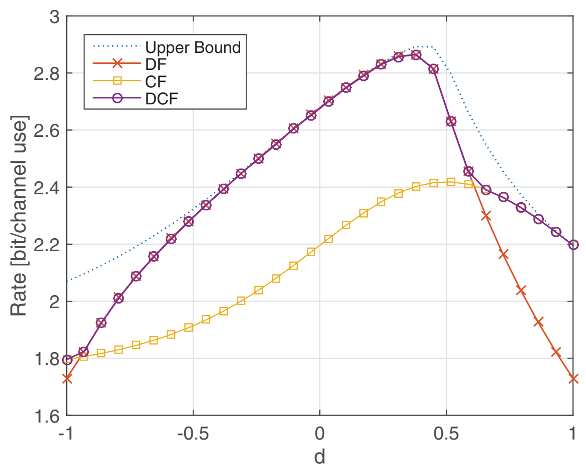

Fig. 3. The achievable rates in the AWGN full-duplex relay channel.

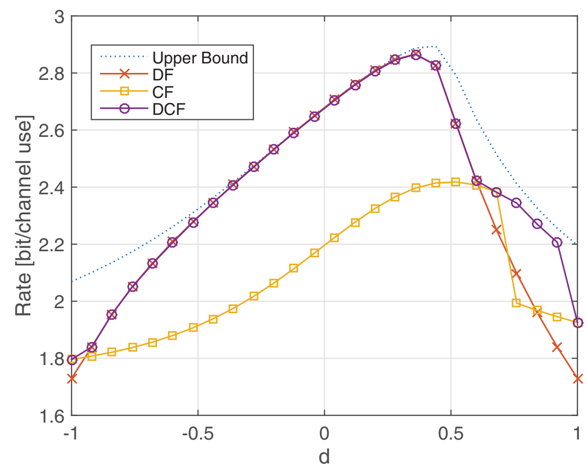

Fig. 4. The achievable rates in the AWGN relay channel under a 16-PAM at the source and 4-PAM at the relay.

search. As shown in many point-to-point and multi-user scenarios [9], [10], when the constellation becomes large enough, the achievable rate under a constrained constellation becomes very close to the Gaussian input rate. However, the main difficulty comes from the restriction in (7) which is now even harder to satisfy since the mutual information $I(X_2;Y_3)$ is no longer equal to

$$\frac{1}{2}\log\left(1+\frac{|H_{23}|^2P_2}{N_3}\right)$$

and is limited by the cardinality of the input size $|X_2|$. The exact value of the constellation constrained capacity [11] is

$$I(X_2;Y_3)=\max_{P_{X_2}}\sum_{X_2}P_{X_2}(x_2)\int_{y_3} P_{Y_3|X_2}(y_3|x_2)\log\left(\frac{P_{Y_3|X_2}(y_3|x_2)}{P_{Y_3}(y_3)}\right)dy_3 \tag{20}$$

A very accurate approximation for $I(X_2;Y_3)$ under constrained constellation can be obtained using the Blahut-Arimoto algorithm [12], [13].

By finding the value of $I(X_2;Y_3)$, one can find the achievable rate of compress-and-forward and DCF. In the following we give an upper bound on the achievable rate of DCF under discrete relay-destination input $\mathcal{X}_2$. The upper bound is based on

$$I(X_2;Y_3)\le\min\left\{|X_2|,\log\left(1+\frac{|H_{23}|^2P_2}{\sigma_3^2}\right)\right\} \tag{21}$$

Therefore, the constraint in (7) becomes

$$I(Y_2;\hat{Y}_2|X_2,U_{1d})\le\min\left\{|X_2|,\log_2\left(1+\frac{|H_{23}|^2P_2}{\sigma_3^2}\right)\right\} - I(U_{2d};Y_3) \tag{22}$$

Via an exhaustive search for the optimal input distributions that satisfy (22), the achievable rate can be obtained. In a similar manner to the previous Section, we show the achievable rate of different strategies under a constrained constellation in Fig. 4.

An intuitive explanation for this observation is as follows: For a 4-PAM constellation and a weak source-relay channel, DF enforces the source to transmit with a small enough rate for the relay to be able to decode both of the bits. CF sends a description of the received signal after quantizing it using only two bits. DCF on the other hand, decode the most reliable bit and sends a description of the least reliable bit.

Remark 2: Compress-and-forward works well when the source relay channel is weak but the relay destination channel is very strong that it can send a very precise quantized version of Y_2. However, when the relay is constellation constrained, even if the quality of the relay-destination link is extremely good, the relay cannot send a precise description of Y_2.

APPENDIX A
ACHIEVABLE RATE OF DCF IN THE AWGN RELAY

First we start by obtaining the optimal value of $\hat{N}$. In the AWGN relay channel with all codewords normally distributed, the compress-and-forward constraint in (7) becomes

$$\frac{1}{2}\log\Big(1+\frac{|H_{12}|^2P_{1c}+\sigma_2^2}{\hat{N}}\Big) \le \frac{1}{2}\log\Big(1+\frac{|H_{23}|^2P_{2c}}{|H_{13}|^2(P_{1d}+P_{1c})+\sigma_3^2}\Big) \tag{23}$$

The following value $\hat{N}$ is the value that satisfies (23) with equality [14], [15], and hence,

$$\hat{N}=\frac{(|H_{12}|^2P_{1c}+\sigma_2^2)(|H_{13}|^2(P_{1d}+P_{1c})+\sigma_2^2)}{|H_{23}|^2P_{2c}} \tag{24}$$

Now, we calculate the two terms that represent the decode-and-forward bound. Given that $Y_2 = H_{12}X_1 + n_2$,

$$\begin{aligned} I(U_{1d};Y_2|U_{2d}) &= h(Y_2|U_{2d}) - h(Y_2|U_{1d},U_{2d}) \\ &= h(U_{1d}+U_{1c}) - h(U_{1c}) \\ &= \frac{1}{2}\log\Big(1+\frac{|H_{12}|^2P_{1d}}{|H_{12}|^2P_{1c}+\sigma_2^2}\Big) \end{aligned} \tag{25}$$

The other term in the decode-and-forward bound is $I(U_{1d}+\beta u_{2d},U_{2d};Y_3)$. Given that $Y_3 = H_{13}X_1 + H_{23}X_2 + n_3$, this term can be obtained by:

$$\begin{aligned} I(U_{1d}+\beta U_{2d},U_{2d};Y_3) &= h(Y_3) - h(Y_3|U_{1d},U_{2d}) \\ &= \frac{1}{2}\log\left(\frac{|H_{12}|^2P_1+|H_{23}|^2P_2+\sigma_3^2}{|H_{13}|^2P_{1c}+|H_{23}|^2P_{2c}+\sigma_3^2}\right) \end{aligned} \tag{26}$$

After some mathematical manipulations,

$$\begin{aligned} &I(U_{1d}+\beta U_{2d},U_{2d};y_3) = \\ &\frac{1}{2}\log\Big(1+\frac{(P_{1d}+\beta^2P_{2c})|H_{13}|^2}{\sigma_3^2}+\frac{P_{2d}|H_{23}|^2}{\sigma_3^2} \\ &+2\rho\sqrt{\frac{(P_{1d}+\beta^2P_2)P_{2d}|H_{13}|^2|H_{23}|^2}{\sigma_3^4}}\Big)\Big\} \end{aligned} \tag{27}$$

The rate of the compress-and-forward part is given by

$$\begin{aligned} I(X_1;\hat{Y}_2,Y_3|X_2,U_{1d}) &= h(\hat{Y}_2,Y_3|X_2,U_{1d}) \\ &- h(\hat{Y}_2,Y_3|X_2,U_{1d},X_1) \end{aligned} \tag{28}$$

where U_{1d} is in the given expression because the assumption of decoding the decode-and-forward part first. By treating $[\hat{Y}_2Y_3]$ as a random vector, from the covariance matrix, the entropies in (28) can be calculating and give

$$\begin{aligned} &I(X_1;\hat{Y}_2,Y_3|X_2,U_{1d}) = \\ &\frac{1}{2}\log\Big(\frac{(|H_{12}|^2P_{1c}+\sigma_3^2+\hat{N})(|H_{13}|^2P_{1c}+\sigma_3^2)}{(\sigma_3^2+\hat{N})\sigma_3^2} \\ &\qquad+\frac{(|H_{12}|^2|H_{13}|^2)P_{1c}^2}{(\sigma_3^2+\hat{N})\sigma_3^2}\Big) \end{aligned} \tag{29}$$

By combining these mutual informations and substituting in the rate described in (6), we obtain the rate in Theorem 2.

REFERENCES

[1] T. Cover and A. Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inform. Theory*, vol. 25, no. 5, pp. 572–584, Sept. 1979.

[2] E. C. Van Der Meulen, "Three-terminal communication channels," *Advances in applied Probability*, pp. 120–154, 1971.

[3] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. Inform. Theory*, vol. 51, no. 9, pp. 3037–3063, Sept. 2005.

[4] J. N. Laneman, "Cooperative diversity in wireless networks: Algorithms and architectures," Ph.D. dissertation, Massachusetts Institute of Technology, 2002.

[5] S. H. Lee and S. Y. Chung, "When is compress-and-forward optimal?" *Information Theory and Applications Workshop (ITA)*, pp. 1–3, Jan. 2010.

[6] A. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inform. Theory*, vol. 22, no. 1, pp. 1–10, Jan. 1976.

[7] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley and Sons, 1991.

[8] G. Kramer, I. Marić, and R. D. Yates, "Cooperative communications," *Foundations and Trends® in Networking*, vol. 1, no. 3–4, pp. 271–425, 2007.

[9] Z. Mheich, F. Alberge, and P. Duhamel, "Achievable rates optimization for broadcast channels using finite size constellations under transmission constraints," *Journal on Wireless Communications and Networking EURASIP*, no. 1, p. 254, 2013.

[10] A. A. Abotabl and A. Nosratinia, "Broadcast coded modulation: Multilevel and bit-interleaved construction," *IEEE Trans. Commun.*, vol. 65, no. 3, pp. 969–980, March 2017.

[11] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. Inform. Theory*, vol. 28, no. 1, pp. 55–67, 1982.

[12] R. Blahut, "Computation of channel capacity and rate-distortion functions," *IEEE Trans. Inf. Theory*, vol. 18, no. 4, pp. 460–473, 1972.

[13] S. Arimoto, "An algorithm for computing the capacity of arbitrary discrete memoryless channels," *IEEE Trans. Inf. Theory*, vol. 18, no. 1, pp. 14–20, 1972.

[14] A. Host-Madsen and J. Zhang, "Capacity bounds and power allocation for wireless relay channels," *IEEE Trans. Inform. Theory*, vol. 51, no. 6, pp. 2020–2040, June 2005.

[15] M. Gastpar, G. Kramer, and P. Gupta, "The multiple-relay channel: Coding and antenna-clustering capacity," in *IEEE International Symposium on Information Theory. Proceedings.* IEEE, 2002, p. 136.