Generalized Backward Decoding Strategies for the Relay Channel

Hon-Fah Chong, Mehul Motani, *Member, IEEE*, and Hari Krishna Garg

***Abstract*—This correspondence studies coding strategies for a three-node relay channel. We start with the basic coding strategies of Cover and El Gamal: the relay decodes the source message and forward it to the destination (*cooperation*); the relay transmits its compressed channel outputs to the destination (*facilitation*); or the relay superimposes both cooperation and facilitation (generalized). In this paper, two new generalized strategies superimposing cooperation and facilitation are introduced and investigated on the general relay channel. The first strategy makes use of sequential backward (*SeqBack*) decoding while the second strategy makes use of simultaneous backward (*SimBack*) decoding. The achievable rate for the second strategy is shown to include that of the generalized strategy of Cover and El Gamal. Assuming zero-mean, jointly Gaussian random variables, the two new strategies give higher achievable rates than the generalized strategy of Cover and El Gamal for certain parameters on the Gaussian relay channel.**

***Index Terms*—Backward decoding, compress-and-forward, decode-and-forward, relay channel.**

I. Introduction

The relay channel was first introduced and studied by van der Meulen in [1], [2]. Two important coding theorems for the single relay channel were established in a fundamental paper by Cover and El Gamal [3]. In the first strategy, termed *cooperation*, the relay decodes the source message and forward it to the destination. In the second strategy, termed *facilitation*, the relay transmits a compressed version of its channel outputs to the destination. These strategies were generalized in [3, Th. 7] to achieve a tighter lower bound for the maximum achievable rate. To date, the capacity of the general single relay channel remains unsolved.

The purpose of this correspondence is to investigate other generalizations of the two basic coding strategies for the single relay channel. In the next section, we introduce some mathematical definitions and notations. Next, we look at some of the fundamental results in [3] and apply them to the Gaussian relay channel. We then discuss and derive achievable rates for two alternative strategies that superimpose cooperation and facilitation. The first strategy performs sequential backward (*SeqBack*) decoding at the receiver while the second strategy performs simultaneous backward (*SimBack*) decoding. In [4], it was shown that simultaneous decoding results in superior performance compared to sequential decoding for the interference channel. We also show that the achievable rate for the SimBack decoding strategy will always contain the rate achieved by the generalized strategy of Cover and El Gamal [3, Th. 7]. Finally, we compute the achievable rates for these strategies in a Gaussian relay channel. As it may be formidable to compute the maximum achievable rate over all input distributions, we impose the customary restriction to the class of jointly Gaussian input distributions.

Manuscript received March 29, 2005; revised March 17, 2006. This work was supported in part by the National University of Singapore by Grants R-263-000-261-112 and R-263-000-293-112. The material in this correspondence was presented in part at the IEEE International Symposium on Information Theory, Adelaide, Australia, September 2005.

The authors are with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore 117576, Singapore.

Communicated by R. W. Yeung, Associate Editor for Shannon Theory.

Digital Object Identifier 10.1109/TIT.2006.887067

Backward decoding for the relay channel was also discussed in [5]. It was shown that for decode-and-forward strategies for multiple relays, backward decoding is not needed for optimal performance. However, in this paper, we show that assuming zero-mean, jointly Gaussian random variables, the two new strategies give higher achievable rates than the generalized strategy of Cover and El Gamal for certain parameters of the Gaussian relay channel.

II. Mathematical Model

We closely follow the formulation and notation of [3]. A discrete memoryless relay channel consists of four finite sets $\mathcal{X}_1$, $\mathcal{X}_2$, $\mathcal{Y}_2$, $\mathcal{Y}_3$, and a collection of probability distributions $p(.,.|x_1x_2)$ on $\mathcal{Y}_2, \mathcal{Y}_3$. x_1 is the source input, x_2 is the relay input, y_2 is the relay output, and y_3 is the destination output.

A $(2^{nR}, n)$ code for the relay channel is composed of a set of integers $\mathcal{W} = \{1, 2, \ldots, 2^{nR}\}$, an encoding function

$$X : \{1, 2, \ldots, 2^{nR}\} \to \mathcal{X}_1^n$$

a set of relay functions $\{f_i\}_{i=1}^n$ such that

$$x_{2i} = f_i(Y_{21}, Y_{22}, \ldots, Y_{2i-1}), 1 \le i \le n$$

and a decoding function

$$g : \mathcal{Y}_3^n \to \{1, 2, \ldots, 2^{nR}\}.$$

The relay is causal in nature. Hence, the relay channel input is allowed to depend only on the past observations $y_{21}, y_{22}, \ldots, y_{2i-1}$. The channel is assumed to be memoryless. Hence, for any choice $p(w)$, $w \in \mathcal{W}$, any code choice $X : \{1, 2, \ldots, 2^{nR}\} \to \mathcal{X}_1^n$, and relay functions $\{f_i\}_{i=1}^n$, the joint probability distribution function on $\mathcal{W} \times \mathcal{X}_1^n \times \mathcal{X}_2^n \times \mathcal{Y}_2^n \times \mathcal{Y}_3^n$ is given by

$$p\left(w, \mathbf{x}_1^n, \mathbf{x}_2^n, \mathbf{y}_2^n, \mathbf{y}_3^n\right) = p(w) \prod_{i=1}^n p(x_{1i}|w) p\left(x_{2i}|y_2^{i-1}\right) \cdot p(y_{2i}, y_{3i}|x_{1i}, x_{2i}).$$

If the message $w \in \mathcal{W}$ is sent, let

$$\lambda(w) = \Pr\left\{g\left(\mathbf{Y}_3^n\right) \neq w | w \text{ sent}\right\}$$

be the conditional probability of error. The average probability of error is defined by

$$P_e^{(n)} = \frac{1}{2^{nR}} \sum_w \lambda(w).$$

The probability of error is calculated under the uniform distribution over the codewords $w \in \mathcal{W}$. The rate R is said to be achievable by the relay channel if there exists a sequence of $(2^{nR}, n)$ codes with $P_e^{(n)} \to 0$. The capacity C_{R} is the supremum of the set of achievable rates.

A. Model for the Gaussian Relay Channel

Consider the Gaussian relay channel of Fig. 1, in which the source node intends to transmit information to the destination node by using the direct link between source and destination as well as with the help of another relay node.

The dependency of the outputs on the inputs are as follows. The relay output is given by

$$Y_2 = h_0 X_1 + N_2 \tag{1}$$

and the destination output is given by

$$Y_3 = h_1 X_1 + h_2 X_2 + N_3. \tag{2}$$

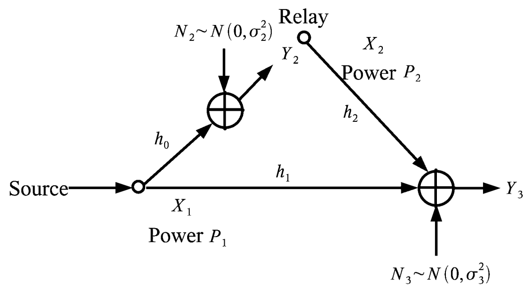

Fig. 1. Gaussian Relay Channel.

The constants h_0, h_1, and h_2 are channel losses and are assumed to be constant. $N_2 \sim \mathcal{N}(0, \sigma_2^2)$ and $N_3 \sim \mathcal{N}(0, \sigma_3^2)$ are independent Gaussian noises. The input power constraints are given by $\mathbb{E}|X_1^2| \le P_1$ and $\mathbb{E}|X_2^2| \le P_2$.

Remark 1: Throughout the paper, we make use of strong typicality in order to invoke Berger's Markov lemma [6, Lemma 14.8.1]. However, since strong typicality does not apply to continuous random variables, the coding theorems can also be proven using weak typicality by making modifications along the lines of Oohama [7]. Our focus, in this paper, is to look at new generalized backward decoding strategies for the relay channel.

B. Notations

We denote a discrete random variable with capital letter X and its realization with lower case letter x. We denote vectors with boldface letters, e.g., $\mathbf{X}^n$ denote a random vector and $\mathbf{x}^n$ denote a realization of the random vector. We denote the entropy of a (discrete or continuous) random variable X by $H(X)$. For the sake of simplicity, we define the following two functions which will be used throughout the correspondence:

$$C(x) \triangleq \frac{1}{2}\log_2(1+x) \tag{3}$$

$$J(x) \triangleq \frac{1}{2}\log_2(2\pi e x). \tag{4}$$

III. Coding Strategies for the Relay Channel

In this section, we review the cut-set upper bound on the capacity of the relay channel. We also review some achievable coding strategies of [3] and then derive two new generalized backward decoding strategies. For all the strategies, we compute rates for the Gaussian relay channel shown in Fig. 1.

A. Capacity Upper Bound

The capacity of the relay channel satisfies

$$R_{\mathrm{U}} \le \sup_{p(x_1,x_2)} \min\{I(X_1X_2;Y_3), I(X_1;Y_2Y_3|X_2)\}. \tag{5}$$

This capacity upper bound can be achieved under certain conditions. The source and the relay could transmit to the destination with rate $I(X_1X_2;Y_3)$ if the relay had complete knowledge of the source message. The rate $I(X_1;Y_2Y_3|X_2)$ could be achieved if the source had knowledge of X_2 and the destination had knowledge of Y_2.

A conditional maximum entropy theorem of [8] ensures that the capacity upper bound can be maximized by making $p(x_1,x_2)$ zero-mean Gaussian. Hence, for the Gaussian relay channel, let $X_1 = aX_2 + W$, where a is a constant. In Appendix A, we compute the cut-set bound to be

$$R_{\mathrm{U}} \le \sup_{0\le\alpha\le1} \min \left\{ \begin{array}{c} C\left(\frac{h_1^2P_1+h_2^2P_2+2h_1h_2\sqrt{(1-\alpha)P_1P_2}}{\sigma_3^2}\right) \\ C\left(\alpha P_1\left(\frac{h_0^2}{\sigma_2^2}+\frac{h_1^2}{\sigma_3^2}\right)\right) \end{array} \right\}. \tag{6}$$

B. Cooperation via Decode-And-Forward

For the first strategy of Cover and El Gamal [3, Th. 1], the relay decodes and knows all the information transmitted to the receiver. Hence, the authors in [5] interpret this as a decode-and-forward strategy. This strategy can achieve any rate up to

$$R_1 = \sup_{p(x_1,x_2)} \{\min\{I(X_1X_2;Y_3), I(X_1;Y_2|X_2)\}\}. \tag{7}$$

In the literature, several different strategies have been suggested to achieve rate R_1. In [3], Cover and El Gamal use irregular block Markov superposition encoding and successive decoding. In [9], Willems suggests regular block Markov superposition encoding and backward decoding. In [10], Carleial uses regular block Markov superposition encoding and sliding window decoding. The advantage of the third strategy by Carleial is that both the source and the relay employ an equal number of codewords. Moreover, a delay of only one block length is necessary for the receiver to perform decoding.

The conditional maximum entropy theorem of [8] again ensures that R_1 is maximized by choosing X_1 and X_2 to be zero-mean Gaussian. Similar to the computation of the cut-set upper bound, we let $X_1 = aX_2 + W$, where a is a constant. Rate R_1 is then given by

$$R_1 = \sup_{0\le\alpha\le1} \min \left\{ \begin{array}{c} C\left(\frac{h_1^2P_1+h_2^2P_2+2h_1h_2\sqrt{(1-\alpha)P_1P_2}}{\sigma_3^2}\right) \\ C\left(\frac{\alpha h_0^2 P_1}{\sigma_2^2}\right) \end{array} \right\}. \tag{8}$$

C. Facilitation via Compress-and-Forward

For the strategy of [3, Th. 6], the relay forward a compressed version of its channel outputs to the destination. For any relay channel, the following rate is achievable:

$$R_2 = \sup_{p(x_1)p(x_2)} I\left(X_1;\hat{Y}_2Y_3|X_2\right) \tag{9}$$

where the supremum is taken over all joint probability density functions of the form

$$p(x_1,x_2,\hat{y}_2,y_2,y_3) = p(x_1)p(x_2) \cdot p(y_2,y_3|x_1,x_2)p(\hat{y}_2|y_2,x_2) \tag{10}$$

subject to the constraint

$$I(X_2;Y_3) \ge I(Y_2;\hat{Y}_2|X_2Y_3). \tag{11}$$

To compute an achievable rate for the Gaussian relay channel, let $\hat{Y}_2 = Y_2 + N_W$, where $N_W \sim \mathcal{N}(0,\sigma_W^2)$. We also assume that X_1 and X_2 are independent, zero-mean Gaussian random variables. We compute the rate in Appendix B and obtain

$$R_2 = C\left(P_1\left(\frac{h_1^2}{\sigma_3^2}+\frac{h_0^2}{\sigma_2^2+\sigma_W^2}\right)\right) \tag{12}$$

subject to the constraint

$$\sigma_W^2 \geq \frac{h_1^2 P_1 \sigma_2^2 + h_0^2 P_1 \sigma_3^2 + \sigma_2^2 \sigma_3^2}{h_2^2 P_2}. \tag{13}$$

Remark 2: The optimal distribution on $(X_1, X_2, \hat{Y}_2)$ is currently still unknown. However, restricting $(X_1, X_2, \hat{Y}_2)$ to the Gaussian distribution allows one to compute an achievable rate for the Gaussian relay channel using compress-and-forward strategy. Throughout the rest of the paper, we restrict our attention to the class of Gaussian input distributions, which may not necessarily be optimal.

D. Generalized Lower Bound of Cover and El Gamal

The strategy of [3, Th. 7] is a combination of the decode-and-forward strategy with the compress-and-forward strategy. For any relay channel, the following rate is achievable:

$$R_3 = \sup \left\{ \min \left\{ \begin{array}{c} I(X_1; \hat{Y}_2 Y_3 | U X_2) + I(U; Y_2 | V X_2), \\ I(X_1 X_2; Y_3) - I(Y_2; \hat{Y}_2 | U X_1 X_2 Y_3) \end{array} \right\} \right\} \tag{14}$$

where the supremum is taken over all joint probability density functions of the form

$$p(u, v, x_1, x_2, y_2, \hat{y}_2, y_3) = p(v)p(u|v) \cdot p(x_1|u)p(x_2|v)p(y_2, y_3|x_1, x_2)p(\hat{y}_2|x_2, y_2, u) \tag{15}$$

subject to the constraint

$$I(X_2; Y_3|V) \geq I(\hat{Y}_2; Y_2|U X_2 Y_3). \tag{16}$$

Remark 3: Cooperation via decode-and-forward strategy is attained by setting $V = X_2, U = X_1$ and $\hat{Y}_2 = \phi$. Facilitation via compress-and-forward strategy is attained by setting $V = \phi$ and $U = \phi$.

Using this result, and assuming $(U, V, X_1, X_2, \hat{Y}_2)$ to be jointly Gaussian, zero-mean random variables, we can compute an achievable rate for the Gaussian relay channel. Let U, X_1, and X_2 be zero-mean Gaussian random variables of the following form:

$$\begin{aligned} U &= aV + W_0 \\ X_1 &= bU + W_1 = abV + bW_0 + W_1 \\ X_2 &= cV + W_2 \end{aligned} \tag{17}$$

where a, b, and c are constants, and V, W_0, W_1, and W_2 are independent, zero-mean Gaussian random variables. For $\alpha \in [0,1], \beta \in [0,1]$ and $\gamma \in [0,1]$, define

$$\begin{aligned} \alpha &= \frac{\mathbb{E}\left[W_1^2\right]}{P_1} \\ \beta &= \frac{\mathbb{E}\left[b^2 W_0^2\right]}{(1-\alpha)P_1} \\ \gamma &= \frac{\mathbb{E}\left[W_2^2\right]}{P_2}. \end{aligned} \tag{18}$$

We also define the following:

$$\overline{\alpha} = 1 - \alpha, \quad \overline{\beta} = 1 - \beta, \quad \text{and } \overline{\gamma} = 1 - \gamma. \tag{19}$$

The received signals Y_2 and Y_3, and the compressed signal $\hat{Y}_2$ can then be written as

$$\begin{aligned} Y_2 &= abh_0 V + bh_0 W_0 + h_0 W_1 + N_2 \\ Y_3 &= (abh_1 + ch_2)V + bh_1 W_0 + h_1 W_1 \\ &\quad + h_2 W_2 + N_3 \\ \hat{Y}_2 &= Y_2 + N_W \end{aligned} \tag{20}$$

where $N_W \sim \mathcal{N}(0, \sigma_W^2)$. We derive the following in Appendix C

$$I(X_1; \hat{Y}_2 Y_3 | U X_2) = C\left(\alpha P_1 \left(\frac{h_1^2}{\sigma_3^2} + \frac{h_0^2}{\sigma_2^2 + \sigma_W^2}\right)\right) \tag{21}$$

$$I(U; Y_2 | V X_2) = C\left(\frac{h_0^2 \beta \overline{\alpha} P_1}{h_0^2 \alpha P_1 + \sigma_2^2}\right) \tag{22}$$

$$I(X_1 X_2; Y_3) = C\left(\frac{h_1^2 P_1 + h_2^2 P_2 + 2h_1 h_2 \sqrt{\overline{\alpha}\overline{\beta}\overline{\gamma} P_1 P_2}}{\sigma_3^2}\right) \tag{23}$$

$$I(\hat{Y}_2; Y_2 | U X_1 X_2 Y_3) = C\left(\frac{\sigma_2^2}{\sigma_W^2}\right) \tag{24}$$

subject to the constraint as shown in

$$\sigma_W^2 \geq \frac{\left[\alpha h_1^2 P_1 \sigma_2^2 + \sigma_3^2 \left(\alpha h_0^2 P_1 + \sigma_2^2\right)\right] \left[(\beta - \alpha\beta + \alpha) h_1^2 P_1 + \sigma_3^2\right]}{\gamma h_2^2 P_2 \left(\alpha h_1^2 P_1 + \sigma_3^2\right)}. \tag{25}$$

E. Seqback Decoding Strategy

In this section, we derive a new achievable rate region for the discrete memoryless relay channel. Similar to the derivation of [3, Th. 7], we superimpose *cooperation* and the transmission of $\hat{Y}_2$, making use of regular block Markov superposition encoding and backward decoding. The following theorem establishes an achievable rate for this strategy.

Theorem 1: For any relay channel $(\mathcal{X}_1 \times \mathcal{X}_2, p(y_2 y_3 | x_1 x_2), \mathcal{Y}_2 \times \mathcal{Y}_3)$, the following rate is achievable:

$$R_4 = \sup \left\{ \min \left\{ \begin{array}{c} I(U; Y_2 | V X_2) + I(X_1; \hat{Y}_2 Y_3 | U X_2), \\ I(UV; Y_3) + I(X_1; \hat{Y}_2 Y_3 | U X_2) \end{array} \right\} \right\} \tag{26}$$

where the supremum is taken over all joint probability density functions of the form (15) and subject to the constraint

$$I(X_2; Y_3 | UV) \geq I(\hat{Y}_2; Y_2 | U X_2 Y_3). \tag{27}$$

Proof: The encoding and decoding methods differ from those of Cover and El Gamal. For encoding, we make use of regular block Markov superposition encoding and for decoding we make use of backward decoding [9]. We consider $B + 2$ blocks, each of n symbols. A sequence of B messages $w_{1i} \times w_{2i} = [1, 2^{nR'}] \times [1, 2^{nR''}], i = 1, 2, \ldots, B$ will be sent over the channel in $n(B+2)$ transmissions. In each n-block $b = 1, 2, \ldots, B+2$, we shall use the same set of codewords. We consider only the probability of error in each block as the total average probability of error can be upper bounded by the sum of the decoding error probabilities at each step, under the assumption that no error propagation from the previous steps has occurred [11].

Codebook Generation:

The random codewords to be used in each block are generated as follows:

a) Generate at random $2^{nR'}$ i.i.d. n-sequences, $\mathbf{v}^n$, each drawn according to

$$p(\mathbf{v}^n) = \prod_{i=1}^{n} p(v_i).$$

Label these $\mathbf{v}^n(w_p), w_p = [1, 2^{nR'}]$.

b) For each codeword $\mathbf{v}^n(w_p)$, generate $2^{n\hat{R}}$ conditionally independent n-sequences, $\mathbf{x}_2^n$, each drawn according to

$$p\left(\mathbf{x}_2^n|\mathbf{v}^n(w_p)\right) = \prod_{i=1}^{n} p(x_{2i}|v_i(w_p)).$$

Label these $\mathbf{x}_2^n(w_p, z_p), z_p = [1, 2^{n\hat{R}}]$.

c) For each codeword $\mathbf{v}^n(w_p)$, generate $2^{nR'}$ conditionally independent n-sequences, $\mathbf{u}^n$, each drawn according to

$$p(\mathbf{u}^n|\mathbf{v}^n(w_p)) = \prod_{i=1}^{n} p(u_i|v_i(w_p)).$$

Label these $\mathbf{u}^n(w_p, w), w = [1, 2^{nR'}]$.

d) For each codeword $\mathbf{u}^n(w_p, w)$ and for each codeword $\mathbf{x}_2^n(w_p, z_p)$, generate $2^{n\hat{R}}$ conditionally independent n-sequences, $\hat{\mathbf{y}}_2^n$, each drawn according to

$$p\left(\hat{\mathbf{y}}_2^n|\mathbf{x}_2^n(w_p, z_p), \mathbf{u}^n(w_p, w)\right) = \prod_{i=1}^{n} p(\hat{y}_{2i}|x_{2i}(w_p, z_p), u_i(w_p, w)).$$

Label these $\hat{\mathbf{y}}_2^n(w_p, w, z_p, z), z = [1, 2^{n\hat{R}}]$.

e) For each codeword $\mathbf{u}^n(w_p, w)$, generate $2^{nR''}$ conditionally independent n-sequences, $\mathbf{x}_1^n$, each drawn according to

$$p\left(\mathbf{x}_1^n|\mathbf{u}^n(w_p, w)\right) = \prod_{i=1}^{n} p(x_{1i}|u_i(w_p, w)).$$

Label these $\mathbf{x}_1^n(w_p, w, w_n), w_n = [1, 2^{nR''}]$.

f) For each codeword $\mathbf{x}_2^n(1, z_p), z_p = [1, 2^{n\hat{R}}]$, generate a n-sequence, $\mathbf{y}_2^n$, with probability

$$p\left(\mathbf{y}_2^n|\mathbf{x}_1^n(1,1,1), \mathbf{x}_2^n(1, z_p)\right) = \prod_{i=1}^{n} p(y_{2i}|x_{1i}(1,1,1), x_{2i}(1, z_p)).$$

Next, choose a z^*, where $z^* = [1, 2^{n\hat{R}}]$, such that $(\mathbf{u}^n(1,1), \mathbf{x}_2^n(1, z_p), \hat{\mathbf{y}}_2^n(1,1,z_p,z^*), \mathbf{y}_2^n)$ are jointly ϵ-typical. Swap the labeling for $\hat{\mathbf{y}}_2^n(1,1,z_p,z^*)$ and $\hat{\mathbf{y}}_2^n(1,1,z_p,1)$ if one such z^* can be found. Nothing is done if no such z^* is found. This is to facilitate backward decoding starting from the last block. For sufficiently large n, such a z^* will exist with high probability if

$$\hat{R} > I(Y_2; \hat{Y}_2|UX_2). \tag{28}$$

The index w_p represents the index w of the previous block while the index z_p represents the index z of the previous block. The index w_n represents the index of the current block which the relay cannot decode. The encoding and decoding procedure for each block i is as follows:

$$\begin{aligned}
&\mathbf{x}_1^n(1, w_{11}, w_{21}), \mathbf{x}_2^n(1,1)\\
&\quad \hat{\mathbf{y}}_2^n(1, \hat{w}_{11}, 1, z_1), \quad i = 1\\
&\mathbf{x}_1^n(w_{1i-1}, w_{1i}, w_{2i}), \mathbf{x}_2^n(\hat{w}_{1i-1}, z_{i-1})\\
&\quad \hat{\mathbf{y}}_2^n(\hat{w}_{1i-1}, \hat{w}_{1i}, z_{i-1}, z_i), \quad i = 2, \dots, B\\
&\mathbf{x}_1^n(w_{1B}, 1, 1), \mathbf{x}_2^n(\hat{w}_{1B}, z_B)\\
&\quad \hat{\mathbf{y}}_2^n(\hat{w}_{1B}, 1, z_B, z_{B+1}), \quad i = B+1\\
&\mathbf{x}_1^n(1,1,1), \mathbf{x}_2^n(1, z_{B+1})\\
&\quad \hat{\mathbf{y}}_2^n(1, 1, z_{B+1}, 1), \quad i = B+2
\end{aligned}$$

Encoding and Decoding at the Relay for block i:

Suppose the relay has decoded accurately w_{1i-1} from the previous block. The relay then determines $\hat{w}_{1i}$ such that $(\mathbf{u}^n(w_{1i-1}, \hat{w}_{1i}), \mathbf{v}^n(w_{1i-1}), \mathbf{x}_2^n(w_{1i-1}, z_{i-1}), \mathbf{y}_2^n(i))$ are jointly ϵ-typical. For sufficiently large n, $\hat{w}_{1i} = w_{1i}$ with high probability if

$$R' < I(U; Y_2|VX_2). \tag{29}$$

Next, the relay estimates z_i such that

$$(\mathbf{u}^n(w_{1i-1}, w_{1i}), \mathbf{x}_2^n(w_{1i-1}, z_{i-1}), \hat{\mathbf{y}}_2^n(w_{1i-1}, w_{1i}, z_{i-1}, z_i), \mathbf{y}_2^n(i))$$

are jointly ϵ-typical. For sufficiently large n, such a z_i will exist with high probability if (28) is satisfied.

Decoding at the Receiver for block i}:

The receiver starts decoding only after receiving the last block $\mathbf{y}_3^n(B+2)$. Assume that w_{1i} and z_i have been decoded accurately from block $i+1$. The receiver determines the unique $\hat{w}_{1i-1}$ such that $(\mathbf{v}^n(\hat{w}_{1i-1}), \mathbf{u}^n(\hat{w}_{1i-1}, w_{1i}), \mathbf{y}_3^n(i))$ are jointly ϵ-typical. For sufficiently large n, $\hat{w}_{1i-1} = w_{1i-1}$ with high probability if

$$R' < I(UV; Y_3). \tag{30}$$

Next, it searches for the unique $\hat{z}_{i-1}$ such that

$$\begin{aligned}(\mathbf{v}^n(w_{1i-1}), \mathbf{u}^n(w_{1i-1}, w_{1i}), \hat{\mathbf{y}}_2^n(w_{1i-1}, w_{1i}, \hat{z}_{i-1}, z_i),\\ \mathbf{x}_2^n(w_{1i-1}, \hat{z}_{i-1}), \mathbf{y}_3^n(i))\end{aligned}$$

are jointly ϵ-typical. For sufficiently large n, $\hat{z}_{i-1} = z_{i-1}$ with high probability if

$$\hat{R} < I(X_2\hat{Y}_2; Y_3|UV). \tag{31}$$

Finally, the receiver searches the list for the unique $\hat{w}_{2i}$ such that

$$\begin{aligned}(\mathbf{v}^n(w_{1i-1}), \mathbf{u}^n(w_{1i-1}, w_{1i}), \mathbf{x}_1^n(w_{1i-1}, w_{1i}, \hat{w}_{2i}), \mathbf{x}_2^n(w_{1i-1}, z_{i-1}),\\ \hat{\mathbf{y}}_2^n(w_{1i-1}, w_{1i}, z_{i-1}, z_i), \mathbf{y}_3^n(i))\end{aligned}$$

are jointly ϵ-typical. For sufficiently large n, $\hat{w}_{2i} = w_{2i}$ with high probability if

$$R'' < I(X_1; \hat{Y}_2Y_3|UX_2). \tag{32}$$

We consider the following

$$\begin{aligned}I(\hat{Y}_2; Y_2Y_3|UX_2) &= I(\hat{Y}_2; Y_2|UX_2) + I(\hat{Y}_2; Y_3|UX_2Y_2)\\ &= I(\hat{Y}_2; Y_2|UX_2).\end{aligned}$$

From (29) and (32), we obtain the first term of (26). From (30) and (32), we obtain the second term of (26). From (28) and (31), we obtain the constraint as follows:

$$\begin{aligned}&I(Y_2; \hat{Y}_2|UX_2) \le I(X_2\hat{Y}_2; Y_3|UV)\\ \Rightarrow\ &I(Y_2; \hat{Y}_2|UX_2) \le I(X_2; Y_3|UV) + I(\hat{Y}_2; Y_3|UX_2)\\ \Rightarrow\ &I(\hat{Y}_2; Y_2|UX_2Y_3) \le I(X_2; Y_3|UV).\end{aligned}$$

□

1) Achievable Rate for the Gaussian Relay Channel: We assume $(U, V, X_1, X_2, \hat{Y}_2)$ to be zero-mean, jointly Gaussian random variables of the same form as (17) and (20). The parameters α, β and γ are as defined in (18). We can show that

$$\begin{aligned}I(UV; Y_3) &= H(Y_3) - H(Y_3|UV)\\ &= C\left(\frac{h_1^2\overline{\alpha}P_1 + h_2^2\overline{\gamma}P_2 + 2h_1h_2\sqrt{\overline{\alpha}\overline{\beta}\overline{\gamma}P_1P_2}}{h_1^2\alpha P_1 + h_2^2\gamma P_2 + \sigma_3^2}\right).\end{aligned} \tag{33}$$

We obtain from (21) and (22) the following:

$$I(X_1;\hat{Y}_2Y_3|UX_2) = C\left(\alpha P_1\left(\frac{h_1^2}{\sigma_3^2}+\frac{h_0^2}{\sigma_2^2+\sigma_W^2}\right)\right)$$
$$I(U;Y_2|VX_2) = C\left(\frac{h_0^2\beta\overline{\alpha}P_1}{h_0^2\alpha P_1+\sigma_2^2}\right).$$

Next, let us consider

$$\begin{aligned}I(Y_2;\hat{Y}_2|UX_2Y_3) &= H(\hat{Y}_2|UX_2Y_3) - H(\hat{Y}_2|UX_2Y_2Y_3)\\ &= H(h_0W_1+N_2+N_W|h_1W_1+N_3) - H(N_W)\\ &= C\left(\frac{\alpha P_1\left(h_1^2\sigma_2^2+h_0^2\sigma_3^2\right)+\sigma_2^2\sigma_3^2}{\sigma_W^2\left(\alpha h_1^2P_1+\sigma_3^2\right)}\right).\end{aligned} \quad (34)$$

We also consider

$$\begin{aligned}I(X_2;Y_3|UV) &= H(Y_3|UV) - H(Y_3|UVX_2)\\ &= H(h_1W_1+h_2W_2+N_3) - H(h_1W_1+N_3)\\ &= C\left(\frac{h_2^2\gamma P_2}{h_1^2\alpha P_1+\sigma_3^2}\right).\end{aligned} \quad (35)$$

Finally, from (34) and (35), we obtain

$$\sigma_W^2 \geq \frac{\alpha h_1^2P_1\sigma_2^2+\alpha h_0^2P_1\sigma_3^2+\sigma_2^2\sigma_3^2}{\gamma h_2^2P_2}. \quad (36)$$

F. SimBack Decoding Strategy

In this section, we exploit the use of *simultaneous decoding* to obtain a new achievable region for the discrete memoryless relay channel. The codebook generation is similar to that in the proof of Theorem 1. However, instead of performing sequential decoding at the receiver, we perform simultaneous decoding. In [4], it was shown that the use of simultaneous decoding results in superior performance compared to sequential decoding for the interference channel. The following theorem establishes an achievable rate for this strategy.

Theorem 2: For any relay channel $(\mathcal{X}_1\times\mathcal{X}_2, p(y_2y_3|x_1x_2), \mathcal{Y}_2\times\mathcal{Y}_3)$, the following rate is achievable:

$$R_5 = \sup\left\{\min\left\{\begin{array}{l}I(X_1;\hat{Y}_2Y_3|UX_2)+I(U;Y_2|VX_2),\\ I(X_1X_2;Y_3)-I(Y_2;\hat{Y}_2|UX_1X_2Y_3)\end{array}\right\}\right\} \quad (37)$$

where the supremum is taken over all joint probability density functions of the form (15) and subject to the constraint

$$I(X_2;Y_3|UV) \geq I(\hat{Y}_2;Y_2|UX_2Y_3). \quad (38)$$

Remark 4: Comparing the SimBack decoding strategy with the generalized lower bound of Cover and El Gamal [3, Th. 7], we see that the rate region of the SimBack decoding strategy will always include that defined by the generalized strategy of Cover and El Gamal. (14) and (37) are the same but the SimBack decoding strategy is subjected to a more relaxed constraint (compare (16) and (38)). This can be seen from

$$\begin{aligned}I(X_2;Y_3|UV) &= H(X_2|UV) - H(X_2|UVY_3)\\ &= H(X_2|V) - H(X_2|UVY_3)\\ &\geq H(X_2|V) - H(X_2|VY_3)\\ &= I(X_2;Y_3|V).\end{aligned}$$

Proof: Similar to the proof of Theorem 1, a sequence of B messages $w_{1i}\times w_{2i} = [1,2^{nR'}]\times[1,2^{nR''}], i=1,2,\ldots,B$ will be sent over the channel in $n(B+2)$ transmissions. In each n-block $b=1,2,\ldots,B+2$, we shall use the same set of codewords. Again, we consider only the probability of error in each decoding step as the total average probability of error can be upper bounded by the sum of the decoding error probabilities at each step, under the assumption that no error propagation from the previous steps has occurred.

Codebook Generation:

The codebook to be used in each block are generated exactly as the codebook generation in the proof of Theorem 1.

Encoding and Decoding at the Relay for block i:

The encoding and decoding at the relay for block i is carried out in exactly the same manner as the SeqBack decoding strategy.

Decoding at the Receiver for block i:

Again, the receiver starts decoding only after receiving the last block $\mathbf{y}_3^n(B+2)$. Assume that w_{1i} and z_i have been decoded accurately from block $i+1$. The receiver determines the unique $\hat{w}_{1i-1}$ and $\hat{z}_{i-1}$ such that

$$(\mathbf{v}^n(\hat{w}_{1i-1}), \mathbf{u}^n(\hat{w}_{1i-1},w_{1i}), \mathbf{x}_2^n(\hat{w}_{1i-1},\hat{z}_{i-1}), \hat{\mathbf{y}}_2^n(\hat{w}_{1i-1},w_{1i},\hat{z}_{i-1},z_i), \mathbf{y}_3^n(i))$$

are jointly ϵ-typical. For sufficiently large n, $\hat{w}_{1i-1}=w_{1i-1}$ and $\hat{z}_{i-1}=z_{i-1}$ with high probability if

$$R'+\hat{R} < I(UX_2\hat{Y}_2;Y_3) \quad (39)$$
$$\hat{R} < I(X_2\hat{Y}_2;Y_3|UV). \quad (40)$$

Finally, the receiver searches the list for the unique $\hat{w}_{2i}$ such that

$$(\mathbf{v}^n(w_{1i-1}), \mathbf{u}^n(w_{1i-1},w_{1i}), \mathbf{x}_1^n(w_{1i-1},w_{1i},\hat{w}_{2i}), \mathbf{x}_2^n(w_{1i-1},z_{i-1}), \hat{\mathbf{y}}_2^n(w_{1i-1},w_{1i},z_{i-1},z_i), \mathbf{y}_3^n(i))$$

are jointly ϵ-typical. For sufficiently large n, $\hat{w}_{2i}=w_{2i}$ with high probability if

$$R'' < I(X_1;\hat{Y}_2Y_3|UX_2). \quad (41)$$

In [4], the authors show that simultaneous decoding performs better than sequential decoding for the interference channel. Hence, in the second strategy, instead of decoding $\hat{w}_{1i-1}$ before before decoding $\hat{z}_{i-1}$, we decode $\hat{z}_{i-1}$ and $\hat{w}_{1i-1}$ simultaneously. From (28), (39) and (41), we obtain the following:

$$\begin{aligned}R_5 &< I(UX_2\hat{Y}_2;Y_3) - I(Y_2;\hat{Y}_2|UX_2)\\ &\quad + I(X_1;\hat{Y}_2Y_3|UX_2)\\ &= I(X_1X_2;Y_3) - I(\hat{Y}_2;Y_2|UX_1X_2Y_3).\end{aligned}$$

From (29) and (41), we obtain

$$R_5 < I(U;Y_2|VX_2) + I(X_1;\hat{Y}_2Y_3|UX_2).$$

From (28) and (40), we obtain the constraint

$$I(X_2;Y_3|UV) \geq I(\hat{Y}_2;Y_2|UX_2Y_3). \qquad \square$$

1) Achievable Rate for the Gaussian Relay Channel: We again assume $(U,V,X_1,X_2,\hat{Y}_2)$ to be zero-mean, jointly Gaussian random

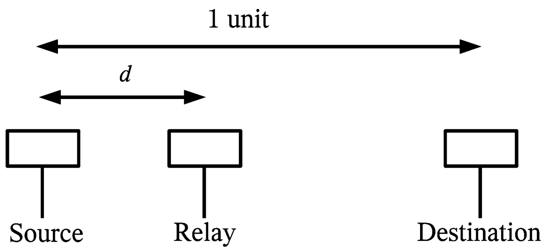

Fig. 2. Linear configuration for the Gaussian relay channel.

variables of the same form as (17) and (20). The parameters α, β, and γ are as defined in (18). We obtain from Appendix C the following

$$I(X_1;\hat{Y}_2Y_3|UX_2) = C\left(\alpha P_1\left(\frac{h_1^2}{\sigma_3^2}+\frac{h_0^2}{\sigma_2^2+\sigma_W^2}\right)\right)$$
$$I(U;Y_2|VX_2) = C\left(\frac{h_0^2\beta(1-\alpha)P_1}{h_0^2\alpha P_1+\sigma_2^2}\right)$$
$$I(X_1X_2;Y_3) = C\left(\frac{h_1^2P_1+h_2^2P_2+2h_1h_2\sqrt{\bar{\alpha}\bar{\beta}\bar{\gamma}P_1P_2}}{\sigma_3^2}\right)$$
$$I(\hat{Y}_2;Y_2|UX_1X_2Y_3) = C\left(\frac{\sigma_2^2}{\sigma_W^2}\right).$$

We also obtain from (36) the following constraint

$$\sigma_W^2 \geq \frac{\alpha h_1^2P_1\sigma_2^2+\alpha h_0^2P_1\sigma_3^2+\sigma_2^2\sigma_3^2}{\gamma h_2^2P_2}.$$

IV. Numerical Computations

In this section, we numerically compute the rates for the various strategies described in the previous section, i.e., cut-set upper bound $(R_{\rm U})$, decode-and-forward (R_1), compress-and-forward (R_2), the generalized lower bound of Cover and El Gamal (R_3), the SeqBack decoding strategy (R_4), and the SimBack decoding strategy (R_5). The physical setup is the Gaussian relay channel shown in Fig. 2. Here, the nodes are collinear, the distance between the source and the destination is 1 unit, and the distance between the source and the relay is d. The quantities h_0, h_1 and h_2 are given by

$$h_0=\frac{1}{d},\quad h_1=1,\quad h_2=\frac{1}{|1-d|}. \tag{42}$$

In all our computations, we have assumed zero-mean, jointly Gaussian random variables. Even though this may not necessarily be optimal, it allows us to compare the rates achieved by the various strategies for this restricted class of probability distributions.

Remark 5: From Tables I and II, we note that as the relay node gets closer to the source, i.e., as d decreases, the rates of all the generalized strategies coincide with that of the decode-and-forward strategy. Conversely, as the relay node gets closer to the destination, i.e., as d increases, the rates of all the generalized strategies coincide with that of the compress-and-forward strategy. In [5], the authors observed that decode-and-forward performs better as the relay moves toward the source while compress-and-forward performs better as the relay moves toward the destination. In fact, decode-and-forward achieves the capacity when the relay is at the source, while compress-and-forward achieves the capacity when the relay is at the destination. The generalized strategies will offer no improvement over decode-and-forward or compress-and-forward when either one of the two strategies dominates.

Remark 6: We also observe that for certain values of h_0, h_1 and h_2, our strategies outperform the decode-and-forward strategy, the compress-and-forward strategy and the generalized strategy of Cover and El Gamal. In general, our strategies outperform the generalized strategy of Cover and El Gamal in regions where the decode-and-forward strategy performs almost as well as the compress-and-forward strategy, i.e., neither decode-and-forward strategy nor compress-and-forward strategy dominates. On the basis of the numerical simulations for the restricted class of zero-mean, jointly Gaussian random variables and the fact that $R_5 \geq R_3$, we conjecture that in general $R_5 > R_3$.

TABLE I
Comparison of Achievable Rates for the Relay Channel for Various Coding Strategies

$P_1=5,\ P_2=1,\ \sigma_2^2=1,\ \sigma_3^2=1$						
d	$R_{\rm U}$	R_1	R_2	R_3	R_4	R_5
0.72	1.9838	1.7061	1.6845	1.7061	1.7061	1.7061
0.73	1.9716	1.6881	1.6908	1.6927	1.6984	1.6984
0.74	1.9597	1.6703	1.6971	1.6971	1.7012	1.7012
0.75	1.9481	1.6529	1.7033	1.7033	1.7052	1.7052
0.76	1.9367	1.6358	1.7094	1.7094	1.7099	1.7099
0.78	1.9148	1.6022	1.7210	1.7210	1.7210	1.7210

TABLE II
Comparison of Achievable Rates for the Relay Channel for Various Coding Strategies

$P_1=10,\ P_2=1,\ \sigma_2^2=1,\ \sigma_3^2=1$						
d	$R_{\rm U}$	R_1	R_2	R_3	R_4	R_5
0.76	2.4117	2.0974	2.0680	2.0974	2.0974	2.0974
0.77	2.4002	2.0796	2.0776	2.0816	2.0911	2.0911
0.78	2.3890	2.0620	2.0874	2.0874	2.0967	2.0967
0.79	2.3781	2.0447	2.0973	2.0973	2.1033	2.1033
0.80	2.3674	2.0276	2.1073	2.1073	2.1108	2.1108
0.82	2.3467	1.9942	2.1274	2.1274	2.1279	2.1279
0.84	2.3269	1.9617	2.1471	2.1471	2.1471	2.1471

Remark 7: We also observe that the SeqBack decoding strategy and the SimBack decoding strategy perform equally well for the given parameters. In fact, both strategies perform equally well when the constraint $I(\hat{Y}_2;Y_2|UX_2Y_3) \leq I(X_2;Y_3|UV)$ holds with no slack. We can see this from the following:

$$\begin{aligned} I(X_1;\hat{Y}_2Y_3|UX_2)&+I(UV;Y_3)\\ &= I(X_1;Y_3|UX_2)+I(UV;Y_3)+I(X_1;\hat{Y}_2|UX_2Y_3)\\ &= I(X_1X_2;Y_3)-I(X_2;Y_3|UV)+I(X_1;\hat{Y}_2|UX_2Y_3)\\ &= I(X_1X_2;Y_3)-I(\hat{Y}_2;Y_2|UX_2Y_3)+I(X_1;\hat{Y}_2|UX_2Y_3)\\ &= I(X_1X_2;Y_3)-I(Y_2;\hat{Y}_2|UX_1X_2Y_3). \end{aligned}$$

For the Gaussian relay channel assuming zero-mean, jointly Gaussian random variables, the rate maximizing distribution is always such that the constraint holds with no slack for both the SeqBack decoding strategy and the SimBack decoding strategy. Hence, both our strategies perform equally well. In general, we are not sure whether the two rate regions are similar since the rate maximizing probability distribution may have slack in the constraint. However, the rate region of the SimBack decoding strategy will definitely include that of the SeqBack decoding strategy.

V. Conclusion

In this correspondence, we have derived two coding strategies for the general relay channel. The strategies make use of regular block

Markov superposition encoding and Willems' backward decoding. In the SeqBack decoding strategy, the receiver performs sequential backward decoding in each block. In the SimBack decoding strategy, the receiver performs simultaneous backward decoding in each block. We also show that the achievable rate for the SimBack decoding strategy includes that of the generalized lower bound of Cover and El Gamal. For a Gaussian relay channel, under the assumption of zero-mean, jointly Gaussian distributions, we see that our strategies outperform the generalized strategy of Cover and El Gamal.

APPENDIX A
DERIVATION OF (6)

The proof for (8) follows exactly along the same lines. Hence, we only show the explicit derivation of (6). Let $\alpha = \frac{\mathbb{E}[W^2]}{P_1}$. The relay output Y_2 is given by

$$Y_2 = h_0 X_1 + N_2 = a h_0 X_2 + h_0 W + N_2. \tag{A43}$$

The destination output Y_3 is given by

$$Y_3 = h_1 X_1 + h_2 X_2 + N_3 = (ah_1 + h_2) X_2 + h_1 W + N_3. \tag{A44}$$

Since Y_3 is a zero-mean Gaussian random variable, the variance of Y_3 is given by

$$\begin{aligned}\mathbb{E}\left(Y_3^2\right) &= h_1^2 \mathbb{E}\left[X_1^2\right] + h_2^2 \mathbb{E}\left[X_2^2\right] + 2h_1h_2\mathbb{E}\left[X_1X_2\right] + \mathbb{E}\left[N_3^2\right] \\ &= h_1^2P_1 + h_2^2P_2 + 2h_1h_2\sqrt{(1-\alpha)P_1P_2} + \sigma_3^2. \end{aligned} \tag{A45}$$

Hence, we have

$$H(Y_3) = J\left(h_1^2P_1 + h_2^2P_2 + 2h_1h_2\sqrt{(1-\alpha)P_1P_2} + \sigma_3^2\right). \tag{A46}$$

We may compute the first term of the cut-set upper bound as follows:

$$\begin{aligned} I(X_1X_2;Y_3) &= H(Y_3) - H(Y_3|X_1X_2) \\ &= H(Y_3) - H(N_3) \\ &= C\left(\frac{h_1^2P_1 + h_2^2P_2 + 2h_1h_2\sqrt{(1-\alpha)P_1P_2}}{\sigma_3^2}\right). \end{aligned}$$

Next, let us consider

$$\begin{aligned} H(Y_2Y_3|X_2) &= H(h_0W + N_2, h_1W + N_3) \\ &= J\left(2\pi e \begin{vmatrix} h_0^2\alpha P_1 + \sigma_2^2 & h_0h_1\alpha P_1 \\ h_0h_1\alpha P_1 & h_1^2\alpha P_1 + \sigma_3^2 \end{vmatrix}\right) \\ &= J\left((2\pi e)\left(\alpha P_1\left(h_0^2\sigma_3^2 + h_1^2\sigma_2^2\right) + \sigma_2^2\sigma_3^2\right)\right). \end{aligned} \tag{A47}$$

Finally, we may compute the second term of the cut-set upper bound as follows:

$$\begin{aligned} I(X_1;Y_2Y_3|X_2) &= H(Y_2Y_3|X_2) - H(Y_2Y_3|X_1X_2) \\ &= H(Y_2Y_3|X_2) - H(N_2N_3) \\ &= C\left(\alpha P_1\left(\frac{h_0^2}{\sigma_2^2} + \frac{h_1^2}{\sigma_3^2}\right)\right). \end{aligned}$$

APPENDIX B
DERIVATION OF (12) AND (2)

The relay output and the destination output is given by (1) and (2), respectively. We have

$$\begin{aligned} I(X_2;Y_3) &= H(Y_3) - H(Y_3|X_2) = H(Y_3) - H(h_1X_1 + N_3) \\ &= C\left(\frac{h_2^2P_2}{h_1^2P_1 + \sigma_3^2}\right). \end{aligned} \tag{B48}$$

Since $I(Y_2;\hat{Y}_2|X_2Y_3) = H(Y_2|X_2Y_3) - H(Y_2|X_2\hat{Y}_2Y_3)$, let us consider

$$\begin{aligned} H(Y_2|X_2Y_3) &= H(h_0X_1 + N_2|h_1X_1 + N_3) \\ &= H(h_0X_1 + N_2, h_1X_1 + N_3) - H(h_1X_1 + N_3) \\ &= J\left(\frac{\begin{vmatrix} h_0^2P_1 + \sigma_2^2 & h_0h_1P_1 \\ h_0h_1P_1 & h_1^2P_1 + \sigma_3^2 \end{vmatrix}}{h_1^2P_1 + \sigma_3^2}\right) \\ &= J\left(\frac{h_1^2P_1\sigma_2^2 + h_0^2P_1\sigma_3^2 + \sigma_2^2\sigma_3^2}{h_1^2P_1 + \sigma_3^2}\right). \end{aligned} \tag{B49}$$

We also consider

$$\begin{aligned} &H(Y_2|X_2\hat{Y}_2Y_3) \\ &= H(h_0X_1 + N_2|h_1X_1 + N_3, h_0X_1 + N_2 + N_W) \\ &= J\left(\frac{\begin{vmatrix} h_0^2P_1 + \sigma_2^2 & h_0h_1P_1 & h_0^2P_1 + \sigma_2^2 \\ h_0h_1P_1 & h_1^2P_1 + \sigma_3^2 & h_0h_1P_1 \\ h_0^2P_1 + \sigma_2^2 & h_0h_1P_1 & h_0^2P_1 + \sigma_2^2 + \sigma_W^2 \end{vmatrix}}{\begin{vmatrix} h_1^2P_1 + \sigma_3^2 & h_0h_1P_1 \\ h_0h_1P_1 & h_0^2P_1 + \sigma_2^2 + \sigma_W^2 \end{vmatrix}}\right) \\ &= J\left(\frac{h_1^2P_1\sigma_2^2\sigma_W^2 + h_0^2P_1\sigma_3^2\sigma_W^2 + \sigma_2^2\sigma_3^2\sigma_W^2}{h_1^2P_1\sigma_2^2 + h_0^2P_1\sigma_3^2 + \sigma_2^2\sigma_3^2 + h_1^2P_1\sigma_W^2 + \sigma_3^2\sigma_W^2}\right). \end{aligned} \tag{B50}$$

From (B48), (B49), and (B50), we can show that the constraint (11) is satisfied if

$$\sigma_W^2 \geq \frac{h_1^2P_1\sigma_2^2 + h_0^2P_1\sigma_3^2 + \sigma_2^2\sigma_3^2}{h_2^2P_2}.$$

Finally, we can compute the rate as follows:

$$\begin{aligned} I(X_1;\hat{Y}_2Y_3|X_2) &= H(\hat{Y}_2Y_3|X_2) - H(\hat{Y}_2Y_3|X_1X_2) \\ &= H(h_0X_1 + N_2 + N_W, h_1X_1 + N_3) \\ &\quad - H(N_2 + N_W, N_3) \\ &= \frac{1}{2}\log_2\left(\frac{\begin{vmatrix} h_0^2P_1 + \sigma_2^2 + \sigma_W^2 & h_0h_1P_1 \\ h_0h_1P_1 & h_1^2P_1 + \sigma_3^2 \end{vmatrix}}{\begin{vmatrix} \sigma_2^2 + \sigma_W^2 & 0 \\ 0 & \sigma_3^2 \end{vmatrix}}\right) \\ &= C\left(P_1\left(\frac{h_1^2}{\sigma_3^2} + \frac{h_0^2}{\sigma_2^2 + \sigma_W^2}\right)\right). \end{aligned}$$

APPENDIX C
DERIVATION OF (21)–(25)

We can compute $I(X_1;\hat{Y}_2Y_3|UX_2)$ as follows:

$$\begin{aligned} &I(X_1;\hat{Y}_2Y_3|UX_2) \\ &= H(\hat{Y}_2Y_3|UX_2) - H(\hat{Y}_2Y_3|UX_1X_2) \\ &= H(h_0W_1 + N_2 + N_W, h_1W_1 + N_3) \\ &\quad - H(N_2 + N_W, N_3) \\ &= \frac{1}{2}\log_2\left(\frac{\begin{vmatrix} h_0^2\alpha P_1 + \sigma_2^2 + \sigma_W^2 & h_0h_1\alpha P_1 \\ h_0h_1\alpha P_1 & h_1^2\alpha P_1 + \sigma_3^2 \end{vmatrix}}{\begin{vmatrix} \sigma_2^2 + \sigma_W^2 & 0 \\ 0 & \sigma_3^2 \end{vmatrix}}\right) \\ &= C\left(\alpha P_1\left(\frac{h_1^2}{\sigma_3^2} + \frac{h_0^2}{\sigma_2^2 + \sigma_W^2}\right)\right). \end{aligned}$$

We can compute $I(U;Y_2|VX_2)$ as follows:

$$\begin{aligned} &I(U;Y_2|VX_2) \\ &= H(Y_2|VX_2) - H(Y_2|UVX_2) \\ &= H(bh_0W_0 + h_0W_1 + N_2) - H(h_0W_1 + N_2) \\ &= C\left(\frac{h_0^2\beta\overline{\alpha}P_1}{h_0^2\alpha P_1 + \sigma_2^2}\right). \end{aligned}$$

We can compute $I(X_1X_2;Y_3)$ as follows:

$$\begin{aligned} I(X_1X_2;Y_3) &= H(Y_3) - H(Y_3|X_1X_2) \\ &= H(Y_3) - H(N_3) \\ &= C\left(\frac{h_1^2P_1 + h_2^2P_2 + 2h_1h_2\sqrt{\overline{\alpha}\overline{\beta}\overline{\gamma}P_1P_2}}{\sigma_3^2}\right). \end{aligned}$$

We can compute $I(Y_2;\hat{Y}_2|UX_1X_2Y_3)$ as follows:

$$I(Y_2;\hat{Y}_2|UX_1X_2Y_3) = I(N_2+N_W;N_2) = C\left(\frac{\sigma_2^2}{\sigma_W^2}\right).$$

Next, let us consider

$$H(h_0W_1+N_2+N_W \mid h_1W_1+N_3) = J\left(\frac{\begin{vmatrix} h_0^2\alpha P_1+\sigma_2^2+\sigma_W^2 & h_0h_1\alpha P_1 \\ h_0h_1\alpha P_1 & h_1^2\alpha P_1+\sigma_3^2 \end{vmatrix}}{h_1^2\alpha P_1+\sigma_3^2}\right) \quad \text{(C51)}$$

and

$$H(h_0W_1+N_2+N_W|h_0W_1+N_2,h_1W_1+N_3) = J\left(\frac{\begin{vmatrix} h_0^2\alpha P_1+\sigma_2^2+\sigma_W^2 & h_0^2\alpha P_1+\sigma_2^2 & h_0h_1\alpha P_1 \\ h_0^2\alpha P_1+\sigma_2^2 & h_0^2\alpha P_1+\sigma_2^2 & h_0h_1\alpha P_1 \\ h_0h_1\alpha P_1 & h_0h_1\alpha P_1 & h_1^2\alpha P_1+\sigma_3^2 \end{vmatrix}}{\begin{vmatrix} h_0^2\alpha P_1+\sigma_2^2 & h_0h_1\alpha P_1 \\ h_0h_1\alpha P_1 & h_1^2\alpha P_1+\sigma_3^2 \end{vmatrix}}\right). \quad \text{(C52)}$$

From (C51) and (C52), we have

$$\begin{aligned} &I(\hat{Y}_2;Y_2|UX_2Y_3) \\ &= I(h_0W_1+N_2+N_W;h_0W_1+N_2|h_1W_1+N_3) \\ &= H(h_0W_1+N_2+N_W|h_1W_1+N_3) \\ &\quad - H(h_0W_1+N_2+N_W|h_0W_1+N_2,h_1W_1+N_3) \\ &= C\left(\frac{\alpha P_1\left(h_1^2\sigma_2^2+h_0^2\sigma_3^2\right)+\sigma_2^2\sigma_3^2}{\sigma_W^2\left(\alpha h_1^2P_1+\sigma_3^2\right)}\right). \end{aligned} \quad \text{(C53)}$$

We also compute $I(X_2;Y_3|V)$ as follows:

$$\begin{aligned} I(X_2;Y_3|V) &= H(Y_3|V) - H(Y_3|V,X_2) \\ &= H(bh_1W_0+h_1W_1+h_2W_2+N_3) \\ &\quad - H(bh_1W_0+h_1W_1+N_3) \\ &= C\left(\frac{h_2^2\gamma P_2}{h_1^2\overline{\alpha}\beta P_1+h_1^2\alpha P_1+\sigma_3^2}\right). \end{aligned} \quad \text{(C54)}$$

Finally, from (C53) and (C.54), the constraint (16) is satisfied if

$$\sigma_W^2 \geq \frac{\left[\alpha h_1^2P_1\sigma_2^2+\sigma_3^2\left(\alpha h_0^2P_1+\sigma_2^2\right)\right]\left[(\beta-\alpha\beta+\alpha)h_1^2P_1+\sigma_3^2\right]}{\gamma h_2^2P_2\left(\alpha h_1^2P_1+\sigma_3^2\right)}.$$

ACKNOWLEDGMENT

The authors are grateful to the anonymous reviewers for their constructive comments and suggestions.

REFERENCES

[1] E. C. van der Meulen, “Transmission of information in a T-terminal discrete memoryless channel,” Ph.D. dissertation, Univ. California, Berkeley, CA, Jun. 1968.
[2] E. C. van der Meulen, “Three-terminal communication channels,” *Adv. Appl. Prob.*, vol. 3, pp. 120–154, 1971.
[3] T. M. Cover and A. A. E. Gamal, “Capacity theorems for the relay channel,” *IEEE Trans. Inf. Theory*, vol. 25, no. 5, pp. 572–584, Sep. 1979.
[4] T. S. Han and K. Kobayashi, “A new achievable rate region for the interference channel,” *IEEE Trans. Inf. Theory*, vol. 27, no. 1, pp. 49–60, Jan. 1981.
[5] G. Kramer, M. Gastpar, and P. Gupta, “Cooperative strategies and capacity theorems for relay networks,” *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3037–3063, Sep. 2005.
[6] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley Interscience, 1991.
[7] Y. Oohama, “Gaussian multiterminal source coding,” *IEEE Trans. Inf. Theory*, vol. 43, no. 6, pp. 1912–1923, Nov. 1997.
[8] J. A. Thomas, “Feedback can at most double Gaussian multiple access channel capacity,” *IEEE Trans. Inf. Theory*, vol. 33, no. 5, pp. 711–716, Sep. 1987.
[9] F. M. J. Willems, “Information theoretical results for the discrete memoryless multiple access channel,” Ph.D. dissertation, Katholieke Univ. Leuven, Leuven, Belgium, Oct. 1982.
[10] A. B. Carleial, “Multiple-access channels with different generalized feedback signals,” *IEEE Trans. Inf. Theory*, vol. 28, no. 6, pp. 841–850, Nov. 1982.
[11] C. M. Zeng, F. Kuhlmann, and A. Buzo, “Achievability proof of some multiuser channel coding theorems using backward decoding,” *IEEE Trans. Inf. Theory*, vol. 35, no. 6, pp. 1160–1165, Nov. 1989.