

# Compound Relay Channel with Informed Relay and Destination

O. Simeone, D. Gündüz and S. Shamai (Shitz)

**Abstract**—A two-state compound relay channel is considered where the relay and the destination are informed about the channel state while the source is not. Achievable rates and upper bounds are derived for discrete memoryless and Gaussian models, and specialized to a scenario with orthogonal components. It is shown that, apart from some special cases, optimality conditions valid for decode-and-forward (DF)-based solutions on a standard relay channel do not carry over to a compound setting, and more flexible transmission strategies are generally advantageous. For instance, partial decode-and-forward (PDF) that superimposes transmission of three layers and uses joint decoding at the destination performs better than the standard two-layer PDF with successive decoding, even when the latter is optimal for the regular relay channel. Moreover, the capacity is derived in the special case in which the relay is not active in one state. Extension to the broadcast coding approach, as an alternative to the compound model, is also discussed.

## I. INTRODUCTION

The relay channel introduced in [1] represents one of the most fundamental building blocks of communication systems in which *cooperative transmission* is enabled. In its basic form, a relay channel consists of a three-node network including a source, a destination and a relay, the sole purpose of the latter being that of improving the source-destination communication. Initial work on the relay channel has focused on discrete memoryless and Gaussian models [2]. In these models, an underlying assumption is that the source, relay and destination are all informed about the channel, i.e., about the conditional distribution of the outputs (of the relay and the destination) given the inputs (of the source and the relay). Given the interest of cooperative models for wireless systems, relaying over fading channels has later been studied, under both the assumption of ergodic fading variations (for fast-varying channels) and quasi-static fading (for slow-varying channels), see, e.g., [3] and references therein. In both cases, channel state information at the receiver side (for both the relay and the destination) is typically assumed, whereas the transmitters (the source and the relay) may or may not be aware of the current channel realizations. The latter condition is generally assumed for quasi-static fading channels, where the concept of outage probability is used to define the probability that a certain fixed rate cannot be

supported by the current channel realization, see, e.g., [3]-[6].

Consider a relay channel where the channel state remains constant for the entire duration of a codeword, and assume that the source, and possibly the relay, are unaware of the current channel realization. The scenario just described can be seen as a *compound* relay channel. In a compound channel, a rate is said to be achievable if a code can be found that guarantees vanishing probability of error for any channel in the uncertainty set [8]. In this sense, the quasi-static fading relay model can be studied as a compound relay channel where the channel uncertainty set is selected so that the measure of its complement provides the desired outage probability. This standpoint is taken in [4] to derive an achievable rate for half-duplex quasi-static fading relay channels.

In this paper, we revisit the model of a compound relay channel. Unlike [4], where the goal was that of deriving an achievable scheme with optimal outage properties, here we tackle the analysis of discrete memoryless and Gaussian compound relay channels *per se*. We focus on a two-state model and assume that while the source is unaware of the current channel conditions, the relay and the destination are informed. Achievable rates and upper bounds are derived for discrete memoryless and Gaussian models, and specialized to a scenario with orthogonal components, extending the analysis in [7]. Generalization to the broadcast coding approach [9] [10], as an alternative to the compound model, is also discussed. After completing this paper, the authors became aware of the concurrent work [12] that considers the same model studied here by deriving achievable rates based on Decode-and-Forward (DF) and the coding scheme of [13].

We start by describing the system model in Sec. II, and then analyze the general two-state discrete and Gaussian compound relay channel in Sec. III. The orthogonal-component model is studied in Sec. IV, and the broadcast coding approach scenario is discussed in Sec. V. Finally, the "occasional relay" scenario is studied in Sec. VI.

*Notation:* We define  $x^i = (x_1, \dots, x_i)$ ; Probability distributions are identified by their subscripts, e.g.,  $P_X(x)$  or  $P_X$  equals  $P_X(x) = \Pr[X = x]$  for a random variable  $X$ ; For a set  $\mathcal{X}$ ,  $|\mathcal{X}|$  denotes the cardinality;  $[1, M] = \{1, \dots, M\}$  with  $M$  integer.

## II. SYSTEM MODEL

Consider a discrete memoryless compound relay (CR) channel  $(\mathcal{X}, \mathcal{T}, P_{Y,Z|X,T}^{(\theta)}, \mathcal{Y}, \mathcal{Z})$  parameterized by  $\theta \in \Theta$ , with  $\Theta$  being the channel uncertainty set, as shown in Fig. 1-(a). The input symbol of the source is taken from alphabet  $\mathcal{X}$ ,

O. Simeone is with CWCSRP, ECE Dept. NJIT {osvaldo.simeone@njit.edu}.

D. Gündüz is with the Dept. of EE, Princeton Univ., and with the Dept. of EE, Stanford Univ. {dgunduz@princeton.edu}.

S. Shamai (Shitz) is with the Department of EE, Technion {sshlo@ee.technion.ac.il}.

This work was supported by the NEWCOM++ network of excellence within the 7th European Community Framework Programmes, and the ISRC consortium for wireless communication, and by the U.S. National Science Foundation under grant CCF-0914899.

while the input alphabet for the relay is  $\mathcal{T}$ . Output alphabets for the relay and the destination are denoted as  $\mathcal{Y}$  and  $\mathcal{Z}$ , respectively. The channel at state  $\theta \in \Theta$  is characterized by the distribution  $P_{Y,Z|X,T}^{(\theta)}$ . The encoder is assumed to be unaware of the current channel realization  $\theta \in \Theta$ , while both the relay and the destination are informed about  $\theta$ .

*Remark 1:* The analysis of this paper carries over to more general CR models  $(\mathcal{X}, \mathcal{T}^{(\theta)}, P_{Y,Z|X,T}^{(\theta)}, \mathcal{Y}^{(\theta)}, \mathcal{Z}^{(\theta)})$  where the input and output alphabets of the relay and the output alphabet of the destination depend on the current value of  $\theta$ .

*Definition 1:* An  $(M, n)$  code for the CR channel of Fig. 1-(a) consists of a message set  $[1, M]$ , a source encoder

$$f: [1, M] \rightarrow \mathcal{X}^n, \quad (1)$$

that maps a message  $W \in [1, M]$  into a codeword  $x^n(W) \in \mathcal{X}^n$ ; a set of  $n|\Theta|$  relaying functions

$$f_i^{(\theta)}: \mathcal{Y}^{i-1} \rightarrow \mathcal{T}, \quad (2)$$

that map the received sequence  $y^{i-1}$  into an input symbol  $t_i = f_i^{(\theta)}(y^{i-1}) \in \mathcal{T}$ , for  $i = 1, \dots, n$  and  $\theta \in \Theta$ ; and  $|\Theta|$  decoding functions

$$g^{(\theta)}: \mathcal{Z}^n \rightarrow [1, M], \quad (3)$$

that provide the decoded message as  $\hat{W}^{(\theta)} = g^{(\theta)}(z^n) \in [1, M]$  for  $\theta \in \Theta$ .

Notice that the relay encoding function (2) and the decoding function (3) depend on the channel state  $\theta$  since the relay and the destination are assumed to know the current parameter  $\theta \in \Theta$ . The channel is assumed to be memoryless so that given the above definitions, the joint distribution of the involved random variables  $P_{WX^n T^n Y^n Z^n \hat{W}}^{(\theta)}$  for a given  $\theta$  factorizes as

$$\frac{1}{M} P_{X^n|W}(x^n|w) \left( \prod_{i=1}^n P_{T_i|Y^{i-1}}^{(\theta)}(t_i|y^{i-1}) \right) \cdot P_{Y Z|X T}^{(\theta)}(y_i, z_i|x_i, t_i) P_{\hat{W}|Z^n}^{(\theta)}(\hat{w}|z^n), \quad (4)$$

where  $P_{X^n|W}$ ,  $P_{T_i|Y^{i-1}}^{(\theta)}$  and  $P_{\hat{W}|Z^n}^{(\theta)}$  are defined by (1), (2) and (3), respectively, and are thus deterministic (i.e., they take values 0,1), and the message  $W$  is assumed to be uniformly distributed in  $[1, M]$ . For a given  $(M, n)$  code the probability of error is defined as

$$P_e^n = \sup_{\theta \in \Theta} \Pr_{\theta}[\hat{W} \neq W], \quad (5)$$

where  $\Pr_{\theta}[\cdot]$  is taken with respect to the measure (4).

*Definition 2:* A rate  $R$  is said to be *achievable* for the CR channel if for any  $\epsilon > 0$  and  $\delta > 0$  and for all sufficiently large  $n$  there exist  $(M, n)$  codes such that  $P_e^n \leq \epsilon$  and  $\log M/n \geq R - \delta$ . The *capacity* of the CR channel is the supremum of all achievable rates.

#### A. Equivalent System with Two States ( $\Theta = \{1, 2\}$ )

From now on, we focus for simplicity on the case with only two states, i.e.,  $\Theta = \{1, 2\}$ . Under the stated assumption, it can be seen that the CR channel is equivalent to a multicasting system with two relays and two

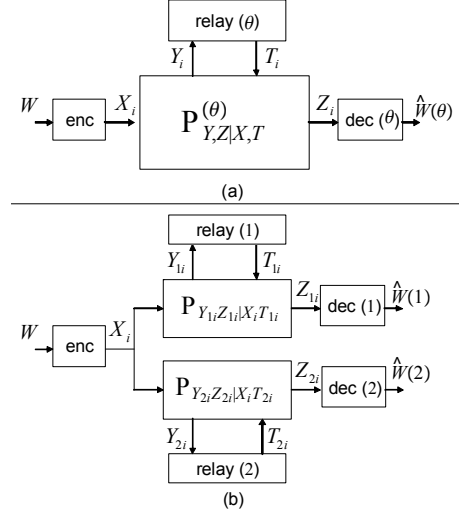


Fig. 1. (a) A Compound Relay (CR) channel; (b) Equivalent model for uncertainty set  $\Theta = \{1, 2\}$ .

receivers as shown in Fig. 1-(b). Denoting as  $T_1, T_2$  the relays' inputs, as  $Y_1, Y_2$  the relays' outputs and as  $Z_1, Z_2$  the destinations' outputs, we will use the definitions  $P_{Y_\theta Z_\theta|X T_\theta} \triangleq P_{Y Z|X T}^{(\theta)}$ ,  $P_{T_{\theta,i}|Y_\theta^{i-1}} \triangleq P_{T_i|Y^{i-1}}^{(\theta)}$  and  $P_{\hat{W}_\theta|Z_\theta^n} \triangleq P_{\hat{W}|Z^n}^{(\theta)}$  for  $\theta \in \{1, 2\}$  to clarify the relationship between the equivalent models in Fig. 1-(a) and Fig. 1-(b). Note that the achievable rate as defined above depends on the joint distribution  $P_{WX^n T_1^n T_2^n Y_1^n Y_2^n Z_1^n Z_2^n \hat{W}_1 \hat{W}_2}$  only through the marginals  $P_{WX^n T_1^n Y_1^n Z_1^n \hat{W}_1} = P_{WX^n T^n Y^n Z^n \hat{W}}^{(1)}$  and  $P_{WX^n T_2^n Y_2^n Z_2^n \hat{W}_2} = P_{WX^n T^n Y^n Z^n \hat{W}}^{(2)}$  (recall (4)).

#### B. Gaussian Model with Power Constraints

Beside the discrete memoryless CR discussed above, the corresponding Gaussian CR model is considered, which is given by, with the notation of the equivalent two-relay model of Fig. 1-(b),

$$Y_{\theta,i} = \gamma_\theta X_i + N_{Y_{\theta,i}} \quad (6a)$$

$$Z_{\theta,i} = \alpha_\theta X_i + \beta_\theta T_{\theta,i} + N_{Z_{\theta,i}}, \quad (6b)$$

for  $\theta \in \{1, 2\}$ . We assume all channel gains  $\gamma_\theta, \alpha_\theta, \beta_\theta \geq 0$ , and white, unit-power, independent Gaussian noises  $N_{Y_{\theta,i}}$  and  $N_{Z_{\theta,i}}$ . Finally, we assume per-block power constraints

$$\frac{1}{n} \sum_{i=1}^n E[X_i^2] \leq P \text{ and } \frac{1}{n} \sum_{i=1}^n E[T_{\theta,i}^2] \leq P_\theta, \quad (7)$$

$\theta = 1, 2$ , where the average  $E[\cdot]$  is taken with respect to the message set and the channel statistics. Note here that we consider separate average power constraints at the relay for different channel states.

### III. SYSTEM ANALYSIS

In this section, we derive upper and lower bounds on the capacity of the CR channel.

### A. Upper Bound

*Proposition 1:* The following is an upper bound on the capacity of the discrete memoryless CR channel:

$$R \leq \max_{P_X} \min \left\{ I_1^{(ub)}(P_X), I_2^{(ub)}(P_X) \right\} \quad (8a)$$

$$I_\theta^{(ub)}(P_X) = \max_{P_{T_\theta|X}} \min \left\{ I(X; Y_\theta Z_\theta | T_\theta), I(X T_\theta; Z_\theta) \right\}, \quad (8b)$$

whereas for the Gaussian CR channel, we have the upper bound

$$R \leq \min \left\{ I_1^{(ub)}, I_2^{(ub)} \right\} \quad (9a)$$

$$I_\theta^{(ub)} = \max_{0 \leq \rho_\theta \leq 1} \min \left\{ \begin{array}{l} \frac{1}{2} \log_2 \left( 1 + (\gamma_\theta^2 + \alpha_\theta^2) P (1 - \rho_\theta^2) \right) \\ \frac{1}{2} \log_2 \left( \frac{1 + \alpha_\theta^2 P + \beta_\theta^2 P_\theta}{+ 2 \rho_\theta \alpha_\theta \beta_\theta \sqrt{P P_\theta}} \right) \end{array} \right\}. \quad (9b)$$

*Proof:* The bound (8) follows from the standard cut-set bound, and uses the fact that  $I_\theta^{(ub)}(P_X)$  depends only on the joint distribution  $P_{X T_\theta}$ . The bound for the Gaussian case follows similarly by further accounting for the power constraints and using the maximum conditional entropy theorem.

### B. Achievable Rates

Achievable rates via Compress-and-Forward (CF) for both discrete-memoryless and Gaussian CF, and Amplify-and-Forward (AF) for the Gaussian CR channel are straightforward generalizations of the corresponding results for the standard relay channel model of [2] (see also Remark 7). Therefore, here we limit the scope to DF protocols, whose analysis does not follow from the standard relay channel model.

1) *Decode-and-Forward (DF):* In DF the message is decoded by both of the relays, which then collaborate with the source to forward the message to the destinations.

*Proposition 2:* The following rate is achievable via DF for the discrete memoryless CR channel:

$$R < \max_{P_{X T_1 T_2}} \min \left\{ I_1^{(DF)}(P_{X T_1 T_2}), I_2^{(DF)}(P_{X T_1 T_2}) \right\}, \quad (10)$$

with

$$I_\theta^{(DF)}(P_{X T_1 T_2}) = \min \left\{ I(X; Y_\theta | T_1 T_2), I(X T_\theta; Z_\theta) \right\}, \quad (11)$$

whereas for the Gaussian CR channel we have

$$R < \max_{0 \leq \rho \leq 1} \min \left\{ I_1^{(DF)}, I_2^{(DF)} \right\} \quad (12)$$

$$I_\theta^{(DF)} = \min \left\{ \begin{array}{l} \frac{1}{2} \log_2 \left( 1 + \gamma_\theta^2 P (1 - \rho^2) \right), \\ \frac{1}{2} \log_2 \left( 1 + \alpha_\theta^2 P + \beta_\theta^2 P_\theta + 2 \rho \alpha_\theta \beta_\theta \sqrt{P P_\theta} \right) \end{array} \right\}.$$

*Proof:* See Appendix-A.

*Remark 2:* Consider the case where both component relay channels are physically degraded, i.e., the symbol-wise Markov chain conditions  $X - T_\theta Y_\theta - Z_\theta$  hold for  $\theta = 1, 2$ . In the standard relay channel case, it is well known that DF achieves the cut-set bound and is thus optimal under this assumption. In contrast, for the CR channel, this is generally

not the case. Intuitively, this is because the source cannot choose the optimal resource allocation between the signal that carries the newly generated information in each block (to be decoded at the relays) and the signal correlated with the relay input  $T_\theta$  (to boost reception at the destination), simultaneously for both of the relay channels. For instance, if a relay is closer to the destination than the other, the source would have to target the farther destination when allocating resources to the transmission of the new message. From an analytical standpoint, notice that in (11) the conditioning on both  $T_1$  and  $T_2$  in the first term accounts for the fact that the part of the source input  $X$  that is correlated with *either* of the relay inputs does not carry new information in a given block. It follows that DF for the CR is optimal when the two channel components are degraded and, additionally, one of the two conditions hold: (i) Among the pairs of distributions  $P_{X T_1}$  and  $P_{X T_2}$  maximizing the cut-set bound (8), there exist one such that  $P_{X T_1} = P_{X T_2} \triangleq P_{X T}^* = P_T^* P_{X|T}^*$ . Under this assumption, we can take  $P_{X T_1 T_2}(x, t_1, t_2) = P_T^*(t_1) P_{X|T}^*(x, t_1) \delta(t_1 - t_2)$  in (10), which coincides with the optimized cut-set bound (This is the case, e.g., when one channel component is "better" than the other so that  $P_{X T}^*$  can be chosen to be the optimal distribution for the worse component); (ii) The distributions  $P_{X T_1}$  and  $P_{X T_2}$  maximizing the cut-set bound (8) are of the form  $P_X^* P_{T_\theta}^*$  so that we can set  $P_{X T_1 T_2} = P_X^* P_{T_1}^* P_{T_2}^*$  in (10) and obtain the cut-set bound (8) (This case occurs, e.g., in a "multihop" scenario where we have  $P_{Y_\theta Z_\theta | X T_\theta} = P_{Y_\theta | X} P_{Z_\theta | T_\theta}$ ).

Further elaboration on the performance of DF is provided by focusing on the Gaussian model via some numerical results. Fig. 2 shows the upper bound (9) along with the achievable rate (12) via DF (upper figure) and the optimal correlation coefficients for  $I_1^{(ub)}, I_2^{(ub)}$  and for the achievable rate (12) (lower figure). Parameters are selected as  $P = P_1 = P_2 = 3$ ,  $\alpha_1^2 = \alpha_2^2 = 1$ ,  $\gamma_\theta^2 = 1/d_\theta^2$ ,  $\beta_\theta^2 = 1/(1 - d_\theta)^2$ , where  $d_\theta$  denotes the source-to-relay normalized distance for the two relays. It is assumed that  $d_2 = d_1 + x/2$  so that as  $x$  increase the second relay moves away from the source. Moreover, we set  $d_1 = 0.01$ . For  $x = 0$ , both relays are in the same location and close to the source, and thus DF is approximately optimal (i.e., it coincides with the upper bound). However, as  $x$  increases, one of the two relays moves further away and DF is no longer optimal. As for the optimal values of the correlation coefficient, it can be seen that in the upper bound (9) as far as the first relay is concerned, the optimal correlation coefficient is  $\rho_1^* \simeq 1$ , while  $\rho_2^*$  decreases with  $x$ . The achievable DF scheme selects a single correlation factor, whose optimal value falls somewhere in between  $\rho_1^*$  and  $\rho_2^*$ .

2) *Partial Decode and Forward (PDF):* As discussed above, DF is limited by the fact that the source cannot tailor the transmission strategy similarly for both relays. Partial decode-and-forward (PDF) based on a three-layer coding strategy is investigated in this section as a solution to this limitation.

*Proposition 3:* The following rate is achievable via PDF

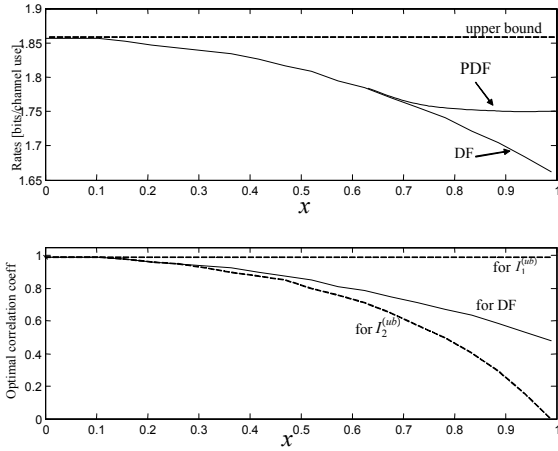


Fig. 2. Upper bound (9) and achievable rates via DF (12) and PDF (13) (upper figure), and the optimal correlation coefficients for  $I_1^{(ub)}$ ,  $I_2^{(ub)}$  and for the achievable rate (12) (lower figure) versus  $x$  with relay distances  $d_1 = 0.01$  and  $d_2 = d_1 + x/2$  ( $P = P_1 = P_2 = 3$ ,  $\alpha_1^2 = \alpha_2^2 = 1$ ,  $\gamma_\theta^2 = 1/d_\theta^2$ ,  $\beta_\theta^2 = 1/(1 - d_\theta)^2$ ).

for the discrete memoryless CR channel:

$$\begin{aligned}
 R < \min\{ & \min_{\theta} I(XT_\theta; Z_\theta), \\
 & \min\{I(XT_1; Z_1|VT_2U_1), I(X; Z_2|VT_2U_1)\} \\
 & + \min_{\theta} I(U_1; Y_\theta|VT_2), \\
 & \min\{I(X; Z_1|U_1U_2VT_1T_2), I(X; Z_2|U_1U_2VT_1T_2)\} \\
 & + \min_{\theta} I(U_1; Y_\theta|VT_2) + I(U_2; Y_1|U_1VT_1T_2)\} \quad (13)
 \end{aligned}$$

for any joint distribution

$$P_{U_1U_2VT_1T_2X} = P_{VU_1T_2}P_{T_1|VT_2}P_{U_2X|U_1VT_1T_2}.$$

For the Gaussian model rate (13) is achievable via PDF with the mutual information terms calculated in (31)-(32) for any parameter set  $0 \leq \eta, \eta_1, \eta_2, \eta'_3, \eta''_3 \leq 1$  with  $\eta'_3 + \eta''_3 \leq 1$  (and definitions  $\bar{\eta}_j = 1 - \eta_j$  and  $\bar{\eta}_3 = 1 - \eta'_3 - \eta''_3$ ). The rate defined above is also achievable with swapped subscripts 1 and 2.

*Proof:* See Appendix-B.

*Remark 3 (sketch of the proof):* The proposed PDF scheme works by using block-Markov encoding (with number of blocks  $B \rightarrow \infty$ ), along with superposition coding, and the basic idea is as follows. The source splits the message to be sent in each block (of rate  $R$ ) into three parts as  $R = R_1 + R_2 + R_3$ . The first part, of rate  $R_1$ , is decoded by both of the relays and is "carried" by the auxiliary random variable  $U_1$ . The second part is decoded only by the first relay (a symmetric result holds by considering the second relay instead) and is conveyed via the auxiliary random variable  $U_2$ . The third part is not decoded by any relay and is transmitted directly to the destinations. Relay 1 and relay 2 cooperate with the source, in the slot following decoding, for transmission of the first message, while relay 1 also cooperates for the transmission of the second message. Specifically, cooperation on the first message at relay 1 takes place through the auxiliary variable  $V$ . The two destinations

perform *joint decoding of the three messages* using either backward or sliding window decoding based on two consecutive received blocks.

*Remark 4:* Setting  $V = T_1$ ,  $U_1 = X$  and  $U_2 = U_1$  in (13), we recover the rate achievable with DF. The same holds in the Gaussian case by choosing  $\eta_1 = 1$ ,  $\eta = 0$ , and  $\eta'_3 = 1$ .

*Remark 5:* In the standard relay channel, PDF transmits a single message to the relay [2]. Moreover, there is no loss in performance by using successive decoding at the destination, whereby the message that is forwarded with the help of the relay is decoded first, followed by the message sent directly to the destination. On the contrary, in the CR channel, a more general scheme that transmits two messages to the relays, as explained above, may be more advantageous. Furthermore, the performance with successive decoding can be improved upon by joint decoding. To illustrate this latter fact, consider the special case of (13) where the rate of the second message (see description above) is set to zero, i.e.,  $R_2 = 0$  and  $V = T_1$ ,  $U_2 = \text{constant}$ , so that a single message is decoded by both relays. Rate (13) reduces to

$$\begin{aligned}
 R < \max_{P_{XU_1T_1T_2}} \min\{ & \min_{\theta} I(XT_\theta; Z_\theta), \min_{\theta} I(U_1; Y_\theta|T_1T_2) \\
 & + \min_{\theta} I(X; Z_\theta|U_1T_1T_2)\}. \quad (14)
 \end{aligned}$$

By employing successive decoding as in the standard relay channel case, we instead achieve:

$$\begin{aligned}
 R < \max_{P_{XU_1T_1T_2}} \min\{ & \min(I(U_1; Y_\theta|T_1T_2), I(U_1T_1T_2; Z_\theta))\} \\
 & + \min_{\theta} \{I(X; Z_\theta|U_1T_1T_2)\}, \quad (15)
 \end{aligned}$$

which is generally smaller than (14).

*Remark 6:* For the numerical example of Fig. 2, it is seen that, as expected, PDF improves the achievable rate when the distance of the two relays from the source is sufficiently different, i.e., when  $x$  is large enough.

*Remark 7:* It is interesting to compare the performance of the DF and the PDF schemes discussed above with the CF strategy. Focusing on the Gaussian CR (same considerations apply for the discrete memoryless CR), it can be seen that CF achieves the rate

$$R = \frac{1}{2} \log_2 \left( 1 + \min_{\theta} \left\{ \alpha_\theta^2 P + \frac{\gamma_\theta^2 P}{1 + \sigma_\theta^2} \right\} \right) \quad (16)$$

with  $\sigma_\theta^2 = \frac{1 + P(\gamma_\theta^2 + \alpha_\theta^2)}{\beta_\theta^2 P_\theta}$ . This is simply the minimum of the CF rates for the two channels (see, e.g., [11]). For the example of Fig. 2, the rate of CF is approximately equal to 1.4 [bits/channel use] for any  $x$  and is thus largely outperformed by DF. This is due to the fact that the performance of CF is limited by the "worst" relay channel, which in our example is the first relay, being close to the source.

#### IV. ORTHOGONAL-COMPONENTS COMPOUND RELAY MODEL

In this section, we focus on the orthogonal-components (OC) CR model, which generalizes the scenario of [7]. In this model, the channel from the source to the relay is orthogonal to that from the source-relay pair to the destination, as

detailed below. Reference [7] shows that in a standard OC relay channel, PDF is optimal, in that it achieves the cut-set upper bound.

#### A. Orthogonal-Components CR channel

Similar to [7], the OC-CR channel is defined as a special case of the general CR model of Fig. 1 by setting  $X = [X_R X_D]$  and

$$P_{Y_\theta, Z_\theta | X, T_\theta} = P_{Y_\theta | X_R, T_\theta} P_{Z_\theta | X_D, T_\theta}, \quad (17)$$

for  $\theta \in \{1, 2\}$ , where  $X_R$  represents the symbol sent on the channel component from the source to the two relays, i.e.,  $P_{Y_\theta | X_R, T_\theta}$ , and  $X_D$  represents the symbol sent on the channel component towards the destination, i.e.,  $P_{Z_\theta | X_D, T_\theta}$ . We also consider the Gaussian version of the OC-CR channel, which is defined, extending [7] to the compound model at hand, as:

$$Y_{\theta,i} = \gamma_\theta X_{R,i} + N_{Y_{\theta,i}} \quad (18a)$$

$$Z_\theta = \alpha_\theta X_{D,i} + \beta_\theta T_{\theta,i} + N_{Z_{\theta,i}}, \quad (18b)$$

with white, unit-power, independent Gaussian noises  $N_{Y_{\theta,i}}$  and  $N_{Z_{\theta,i}}$ , and power constraints  $\frac{1}{n} \sum_{i=1}^n E[X_{R,i}^2] + E[X_{D,i}^2] \leq P$  and  $\frac{1}{n} \sum_{i=1}^n E[T_{\theta,i}^2] \leq P_\theta$ .

#### B. Upper Bound

*Proposition 4:* The following is an upper bound on the capacity of the discrete memoryless OC-CR channel:

$$R \leq \max_{P_{X_D X_R}} \min \left\{ I_1^{(ub)}(P_{X_D X_R}), I_2^{(ub)}(P_{X_D X_R}) \right\} \quad (19a)$$

$$I_\theta^{(ub)}(P_X) = \max_{P_{T_\theta | X_D X_R}} \min \{ I(X_R; Y_\theta | T_\theta) + I(X_D; Z_\theta | T_\theta), I(X_D T_\theta; Z_\theta) \}, \quad (19b)$$

while for the Gaussian OC-CR channel, we have:

$$R \leq \max_{0 \leq \eta \leq 1} \min \left\{ I_1^{(ub)}(\eta), I_2^{(ub)}(\eta) \right\} \quad (20a)$$

$$I_\theta^{(ub)}(\eta) = \max_{0 \leq \rho_\theta \leq 1} \min \left\{ \frac{1}{2} \log_2 (1 + (1 - \eta) \gamma_\theta^2 P) + \frac{1}{2} \log_2 (1 + \eta \alpha_\theta^2 P (1 - \rho_\theta^2)), \frac{1}{2} \log_2 (1 + \eta \alpha_\theta^2 P + \beta_\theta^2 P_\theta + 2\rho_\theta \alpha_\theta \beta_\theta \sqrt{\eta P P_\theta}) \right\}. \quad (20b)$$

*Proof:* The bound (19) follows from Proposition 1 by setting  $X = [X_D X_R]$  and then proceeding as in Sec. II of [7]. The bound for the Gaussian case also follows similarly to the discussion in Sec. III of [7].

#### C. Achievable Rates

Here, we study the rate achievable via PDF. It is recalled that PDF is optimal for the standard OC relay channel [7].

*Proposition 5:* The following rate is achievable via PDF for the discrete memoryless OC-CR channel:

$$R < \max_{P_{X_D X_R U_1 T_2}} \min \left\{ \min_\theta I(X_R; Y_\theta | T_1 T_2) + \min_\theta I(X_D; Z_\theta | X_R T_1 T_2), \min_\theta I(X_D T_\theta; Z_\theta) \right\}, \quad (21)$$

while for the Gaussian OC-CR channel we have:

$$R < \max_{\substack{0 \leq \eta \leq 1 \\ 0 \leq \rho \leq 1}} \min \left\{ \min_\theta \frac{1}{2} \log (1 + \gamma_\theta^2 \eta P) + \min_\theta \frac{1}{2} \log (1 + \alpha_\theta^2 (1 - \eta) (1 - \rho^2) P), \min_\theta \frac{1}{2} \log \left( \frac{1 + \alpha_\theta^2 (1 - \eta) P + \beta_\theta^2 P_\theta}{+ 2\alpha_\theta \beta_\theta \rho \sqrt{(1 - \eta) P P_\theta}} \right) \right\}. \quad (22)$$

*Proof:* Rate (21) follows from (14) by setting  $U_1 = X_R$ . For the Gaussian model, we evaluate the involved mutual informations with  $T_\theta = \sqrt{P_\theta} \tilde{T}$ ,  $X_R = \sqrt{\eta} P \tilde{U}$  and  $X_D = \sqrt{(1 - \eta) P} (\rho \tilde{T} + \sqrt{1 - \rho^2} \tilde{X})$ , where  $\tilde{T}, \tilde{X}, \tilde{U} \sim \mathcal{N}(0, 1)$  are independent.

*Remark 8:* While the scheme of Proposition 5 employs a two-layer PDF scheme, a generally better rate can be obtained by setting  $U_1 = X_R$  in (13), rather than in (14), thereby using a three-layer transmission. This is not further pursued here.

*Remark 9:* Similar to the discussion in Remark 5, rate (21) is obtained via joint decoding at the destination. Using the successive decoding scheme that would be optimal in a standard OC relay channel, we instead obtain from (15) the achievable rate:

$$R < \max_{P_{X_D X_R T_1 T_2}} \min_\theta \left\{ \begin{array}{l} \min(I(X_R; Y_\theta | T_1 T_2), \\ I(X_R T_1 T_2; Z_\theta)) \end{array} \right\} + \min_\theta \{ I(X_D; Z_\theta | X_R T_1 T_2) \}, \quad (23)$$

and for the Gaussian OC-CR:

$$R < \max_{\substack{0 \leq \eta \leq 1 \\ 0 \leq \rho \leq 1}} \min_\theta \left\{ \min \left( \frac{1}{2} \log (1 + \gamma_\theta^2 \eta P), \frac{1}{2} \log \left( \frac{1 + \alpha_\theta^2 (1 - \eta) P + \beta_\theta^2 P_\theta + 2\alpha_\theta \beta_\theta \rho \sqrt{(1 - \eta) P P_\theta}}{1 + \alpha_\theta^2 (1 - \eta) (1 - \rho^2) P} \right) \right) + \min_\theta \left\{ \frac{1}{2} \log (1 + \alpha_\theta^2 (1 - \eta) (1 - \rho^2) P) \right\} \right\}. \quad (24)$$

Fig. 3 shows the achievable rate for a Gaussian OC-CR model via PDF with joint decoding (JD) as in (22), successive decoding (SD) as in (24), and the upper bound (20) for  $P = P_1 = P_3 = 3$ ,  $\gamma_1^2 = \beta_1^2 = a > 1$ ,  $\gamma_2^2 = \beta_2^2 = 1$ ,  $\alpha_2^2 = a$  versus  $\alpha_1^2$ . In other words, the first channel has the best gains to/from the relay by a factor of  $a > 1$ , while the direct channel to the destination of the second channel can be potentially better for channel 2 (since  $\alpha_2^2 = a$ ). Two major differences are observed with respect to the standard non-compound case [7]: (i) Here, PDF is not necessarily optimal: In particular, it is optimal if one of the two channels dominates the other, that is, all the channel gains are larger for one of the two channels (here channel 1 dominates if  $\alpha_1^2 \geq a$ , see also Remark 2); (ii) Successive interference cancellation at the receiver may fail to achieve the upper bound (when  $\alpha_1^2$  is sufficiently large), even when PDF is optimal: in this case, joint decoding at the receiver is necessary.

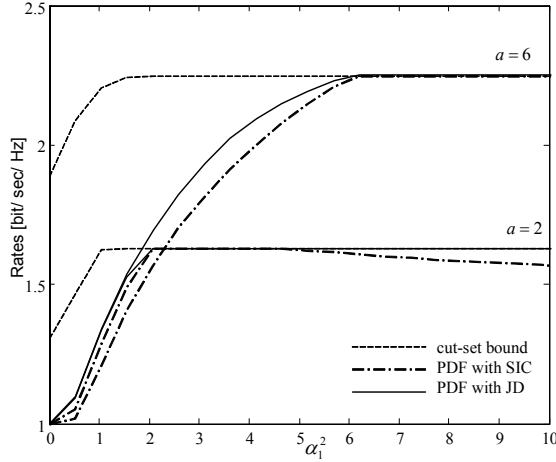


Fig. 3. Achievable rate on a Gaussian OC-CR model via PDF with joint decoding (JD), (22), and successive decoding (SD), (24), as compared to the upper bound (20) versus  $\alpha_1^2$  ( $P = P_1 = P_3 = 3$ ,  $\gamma_1^2 = \beta_1^2 = a > 1$ ,  $\gamma_2^2 = \beta_2^2 = 1$ ,  $\alpha_2^2 = a$ ).

## V. BROADCAST CODING APPROACH

In this section, we consider an alternative definition of the achievable rate that allows delivery of a variable number of bits depending on the network state [9] [10]. This strategy is typically referred to as the broadcast coding approach and entails the following definitions of coding and achievable rates. Notice that we focus on the general CR but the results are specialized easily to the OC-CR model.

*Definition 3:* An  $(M_1, M_2, n)$  broadcast code for the CR model of Fig. 1-(a) with  $\Theta = \{1, 2\}$  consists of message sets  $[1, M_1]$  and  $[1, M_2]$ , an encoder

$$f: [1, M_1] \times [1, M_2] \rightarrow \mathcal{X}^n, \quad (25)$$

that maps messages  $W_i \in [1, M_i]$ ,  $i = 1, 2$ , into a codeword  $x^n(W_1, W_2) \in \mathcal{X}^n$ ; a set of  $2n$  relaying functions (2) and two decoding functions

$$g^{(1)}: \mathcal{Z}^n \rightarrow [1, M_1] \times [1, M_2] \quad (26)$$

$$g^{(2)}: \mathcal{Z}^n \rightarrow [1, M_1], \quad (27)$$

that provide the decoded messages as  $(\hat{W}_1^{(1)}, \hat{W}_2) = g^{(1)}(z^n) \in [1, M_1] \times [1, M_2]$  and  $\hat{W}_1^{(2)} = g^{(2)}(z^n)$ .

The definition above implies that when in state  $\theta = 1$  (the "good" state) the destination is required to decode both  $W_1$  and  $W_2$ , thus receiving  $\log_2 M_1 + \log_2 M_2$  bits, whereas in state  $\theta = 2$  (the "bad" state) only the message  $W_1$  is decoded. As explained in [10], this approach to coding is especially of interest for multimedia applications such as video streaming, where multilayer (successive refinement) coding enable progressively better "reception quality" as more layers (messages) are correctly decoded. Notice that one could clearly switch the indices of channels 1 and 2.

*Definition 4:* A pair of rates  $(R_1, R_2)$  is said to be achievable for the CR channel with broadcast coding if for any  $\epsilon > 0$  and  $\delta > 0$  and for all sufficiently large  $n$  there exist  $(M_1, M_2, n)$  broadcast codes such that  $P_e^n = \max\{\Pr_{\theta=1}[(\hat{W}_1^{(1)}, \hat{W}_2) \neq (W_1, W_2)], \Pr_{\theta=2}[\hat{W}_1^{(2)} \neq$

$W_1]\} \leq \epsilon$  and  $\log M_1/n \geq R_1 - \delta$ ,  $\log M_2/n \geq R_2 - \delta$ . The capacity region of the CR channel is the closure of the set of all achievable rate pairs.

The proposition below provides an achievable rate region based on DF at the two relays and superposition coding. It is noted that superposition coding is known to be optimal for a broadcast channel with degraded message sets, which coincides with the CR channel under study for  $Y_1 = Y_2 = \emptyset$ . The use of DF implies that message  $W_1$  is decoded at both relays while  $W_2$  is only decoded at relay 1 (i.e., in the "good" channel state).

*Proposition 6:* The closure of the union of all rates  $(R_1, R_2)$  satisfying the inequalities

$$R_1 < \min \left\{ \min_{\theta} I(U_1; Y_{\theta} | VT_2), I(U_1 T_2; Z_2) \right\}, \quad (28a)$$

$$R_2 < \min \{ I(X; Y_1 | U_1 VT_1 T_2), I(X T_1; Z_1 | VT_2 U_1) \} \quad (28b)$$

$$R_1 + R_2 < I(X T_1; Z_1) \quad (28c)$$

for some joint distribution

$$P_{U_1 V T_1 T_2 X} = P_{V U_1 T_2} P_{T_1 | V T_2} P_{X | U_1 V T_1 T_2},$$

is achievable via DF for the CR with broadcast coding. In the Gaussian model, region (28) is achievable via DF with the mutual information terms calculated in (31)-(32) and (33) for any parameter set  $0 \leq \eta_1, \eta_2, \eta'_3, \eta''_3 \leq 1$  with  $\eta'_3 + \eta''_3 \leq 1$  and  $\eta = 0$ .

*Proof:* The encoding scheme is a special case of the one used for PDF in Proposition 3 in which we set  $R_3 = 0$  and  $U_2 = X$  (i.e., no message is sent directly to the destinations without being decoded at the relays). However, decoding at the destinations takes place differently in that decoder 1 is interested in both  $W_1$  and  $W_2$ , while decoder 2 only recovers  $W_1$ . This modifies the error events as briefly discussed in Appendix-C.

*Remark 10:* Setting  $R_2 = 0$  and  $X = U_1$ , one recovers the rate  $R = R_1$  achievable with DF of Proposition 3. From this, and noticing that the latter choice maximizes the achievable rate  $R_1$  in Proposition 6, one can conclude that the broadcast coding approach allows to transmit a larger rate  $R_1 + R_2$  in the "good state"  $\theta = 1$  at the expense of a reduction in the maximum rate  $R_1$  achievable in the "bad state"  $\theta = 2$ .

## VI. THE OCCASIONAL RELAY CHANNEL SCENARIO

In this section, we consider a special case of the CR channel in which the relay may be present (say, in state  $\theta = 1$ ) or not (in state  $\theta = 2$ ). We refer to this model as the Occasional Relay (OR) channel. This amounts to setting  $P_{Y_2 Z_2 | X T_2} = P_{Z_2 | X}$  or equivalently  $Y_2$  and  $T_2$  to constants (for the Gaussian model,  $\gamma_2 = 0$  and/ or  $\beta_2 = 0$ ). We treat this channel separately here for its relevance and to point out two results specific to this model. We start with the following capacity result that is the counterpart of the conclusions in [6] for a quasi-static fading relay channel.

*Proposition 7:* If channel 1 is physically degraded, i.e., the Markov chain  $X - T_1 Y_1 - Z_1$  holds symbol-wise, the

capacity of the OR channel is given by

$$C = \max_{P_{T_1 X}} \min \{I(X; Y_1|T_1), I(X; Z_2), I(XT_1; Z_1)\}.$$

*Proof:* The converse follows from Proposition 1, while achievability is a special case of the PDF scheme of Proposition 3 by setting  $U_2 = X$  and  $V, U_1$  as constants.

Consider now the broadcast coding approach for the OR channel. The region of Proposition 6 reduces to  $R_1 = 0$  and  $R_2 \leq \min\{I(X; Y_1|T_1), I(XT_1; Z_1)\}$ . The next proposition provides a better achievable scheme tailored to the OR.

*Proposition 8:* The closure of the union of all rates  $(R_1, R_2)$  satisfying the inequalities

$$\begin{aligned} R_1 &< \min\{I(U; Y_1|T_1), I(UT_1; Z_2)\} \\ R_2 &< I(X; Z_1|UT_1) \\ R_1 + R_2 &< I(XT_1; Z_1) \end{aligned}$$

for some joint distribution  $P_{UT_1 X}$  is achievable for the OR channel with broadcast coding.

*Proof:* Message  $W_1$  is decoded at relay 1 and at both destinations 1 and 2, while  $W_2$  decoded only at destination 1. Encoding takes place in each block  $b$  by transmitting codewords  $t_1^n(w_{1b-1})$ ,  $u^n(w_{1b}, w_{1b-1})$  and  $x^n(w_{2b}, w_{1b}, w_{1b-1})$ , jointly generated and decoded similarly to Appendix A and B.

*Remark 11:* Similar considerations as Remark 10 apply here.

## VII. CONCLUDING REMARKS

In the presence of inherent uncertainty regarding the channel state at the transmitter, design of robust coding schemes becomes a critical issue. This is especially true for cooperative networks where more links and nodes participate in the ongoing communication. Rather than considering a standard outage formulation, this paper has taken a compound view of the problem for a single-relay channel with two possible states. Under the assumption that the relay and the destination are informed about the current channel state but the source is not, achievable rates based on DF and PDF as well as an upper bound are derived for discrete memoryless and Gaussian models, and specialized to a scenario with orthogonal components. It is shown that, apart from some special cases (like the "occasional relay" scenario), the optimality of DF for a physically degraded relay and the optimality of PDF with orthogonal components for a standard relay channel do not carry over to a compound setting, and more general transmission strategies may be more advantageous. For instance, partial decode-and-forward (PDF) that superimposes transmission of three layers and uses joint decoding at the destination performs better than the standard two-layer PDF with successive decoding, even when the latter is optimal for the regular relay channel.

## VIII. APPENDIX

### A. Appendix-A: Proof of Proposition 2

We use random coding arguments as follows. Generate codewords  $t_1^n(w')$  and  $t_2^n(w')$  with  $w' \in [1, 2^{nR}]$  for each

block independently, where each letter  $(t_{1,i}, t_{2,i})$  is drawn independent and identically distributed (i.i.d.) according to the given distribution  $P_{T_1 T_2}$ . Then, for each  $w'$ , generate  $2^{nR}$  codewords  $x^n(w, w')$  with each letter i.i.d. with respect to  $P_{X|T_1 T_2}(\cdot|t_{1,i}(w'), t_{2,i}(w'))$ . Coding takes place via a standard block-Markov strategy [2] with the mapping between messages and codewords for block  $b$  given by  $t_1^n(w_{b-1})$ ,  $t_2^n(w_{b-1})$  and  $x^n(w_b, w_{b-1})$ , where  $w_b$  denotes the message sent over block  $b$  (initial and final blocks are dealt as in [2]). Decoding at the relays is performed block-by-block via joint typicality. Namely, assume that at block  $b$  decoding at the previous block  $b - 1$  has been successful so that  $w_{b-1}$  is known by the relays (this can be assumed without loss of generality by reasoning from the first block and imposing the conditions discussed below [2]). Now, a message  $w_b \in [1, 2^{nR}]$  is sought by relay  $(\theta)$  such that  $(x^n(w_b, w_{b-1}), t_1^n(w_{b-1}), t_2^n(w_{b-1}), y_\theta^n(b))$ , where  $y_\theta^n(b)$  is the received signal at block  $b$ , are jointly typical with respect to distribution  $P_{X T_1 T_2 Y_\theta}$ . The relay declares an error if none or more than one such  $w_b$  is found. Following standard arguments, the probability of error in this step can be made arbitrarily small if  $R < I(X; Y_\theta|T_1 T_2)$ . At the receiver, sliding window decoding (or backward decoding) can be employed. Here we consider sliding window. Assume again without loss of generality that message  $w_{b-1}$  is known at the end of block  $b$  by the destinations. The destinations look for a message  $w_b$  such that the tuples  $(x^n(w_b, w_{b-1}), t_1^n(w_{b-1}), t_2^n(w_{b-1}), z_\theta^n(b))$  and  $(x^n(w_{b+1}, w_b), t_1^n(w_b), t_2^n(w_b), z_\theta^n(b+1))$  are both jointly typical each with respect to distribution  $P_{X T_1 T_2 Z_\theta}$ . The probability of error can be made arbitrarily small if  $R < I(X; Z_\theta|T_1 T_2) + I(T_1 T_2; Z_\theta) = I(X T_1 T_2; Z_\theta) = I(X T_\theta; Z_\theta)$ .

For the Gaussian model, it can be seen that it is sufficient to consider Gaussian inputs of the type  $T_\theta = \sqrt{P_\theta} \tilde{T}$  and  $X = \rho \sqrt{P \tilde{T}} + \sqrt{(1 - \rho^2) P} \tilde{X}$  with  $\tilde{T}, \tilde{X} \sim \mathcal{N}(0, 1)$  independent.

### B. Appendix-B: Proof of Proposition 3

For each block independently, we generate the following set of codewords (recall also sketch of proof in Sec. III-B.2): Two sets of  $2^{nR_1}$  codewords denoted by  $v^n(w_1)$  and  $t_2^n(w_1)$  i.i.d. with distributions  $P_{V T_2}$  with  $w_1 \in [1, 2^{nR_1}]$ ; For each  $w_1$ ,  $2^{nR_1}$  codewords  $u_1^n(w'_1, w_1)$  according to  $\prod_{i=1}^n P_{U_1|V T_2}(u_{1,i}|v_i(w_1), v_i(w_1))$ ; For each  $w_1$ ,  $2^{nR_2}$  codewords  $t_1^n(w_2, w_1)$  with  $w_2 \in [1, 2^{nR_2}]$  according to  $\prod_{i=1}^n P_{T_1|V T_2}(t_{1,i}|v_i(w_1), v_i(w_1))$ ; for each  $w_2$ ,  $2^{nR_2}$  codewords  $u_2^n(w'_2, w_2, w'_1, w_1)$  with distribution  $\prod_{i=1}^n P_{U_2|T_1 U_1 V T_2}(u_{2,i}|t_{1,i}(w_2, w'_1, w_1), u_{1,i}(w'_1, w_1), v_i(w_1), t_{2,i}(w_2))$ ; Finally, for each  $(w'_2, w_2, w'_1, w_1)$  generate  $2^{nR_3}$  codewords  $x^n(w_3, w'_2, w_2, w'_1, w_1)$  with distribution  $\prod_{i=1}^n P_{X|U_2 T_1 U_1 V T_2}(x_i|u_{2,i}(w'_2, w_2, w'_1, w_1), t_{1,i}(w_2, w'_1, w_1), u_{1,i}(w'_1, w_1), v_i(w_1), t_{2,i}(w_2))$ . Coding takes place following a standard block-Markov coding strategy with mapping in the  $b$ th block given by  $v^n(w_{1b-1})$ ,  $t_2^n(w_{1b-1})$ ,  $u_1^n(w_{1b}, w_{1b-1})$ ,  $t_1^n(w_{2b-1}, w_{1b-1})$ ,  $u_2^n(w_{2b}, w_{2b-1}, w_{1b}, w_{1b-1})$ ,  $x^n(w_{3b}, w_{2b}, w_{2b-1}, w_{1b}, w_{1b-1})$ .

Relays decode block by block. Assume without loss of generality that at block  $b$ , messages  $w_{1b-1}$  and  $w_{2b-1}$  are known at relay 1, while relay 2 knows  $w_{1b-1}$ . Relay 1 and decode  $w_{1b}$  by looking for a tuple  $(u_1^n(w_{1b}, w_{1b-1}), v^n(w_{1b-1}), t_2^n(w_{1b-1}), y_\theta^n(b))$  that is jointly typical with respect to  $P_{UVT_2Y_\theta}$ . Imposing vanishing error probability leads to the condition  $R_1 < \min_\theta I(U_1; Y_\theta | VT_2)$ . Relay 1 successively decodes also message  $w_{2b}$  by looking for a jointly typical tuple  $(u_1^n(w_{1b}, w_{1b-1}), v^n(w_{1b-1}), t_1^n(w_{2b-1}, w_{1b-1}), t_2^n(w_{1b-1}), u_2^n(w_{2b}, w_{2b-1}, w_{1b}, w_{1b-1}), y_2^n(b))$  which yields the condition  $R_2 < I(U_2; Y_1 | U_1 VT_1 T_2)$ . Decoding at the destinations may take place either via backward decoding or sliding window. Consider sliding window and assume that at block  $b$  messages  $w_{1b-1}, w_{2b-1}, w_{3b-1}$  have been already decoded at the destinations. Now, at both destinations, *joint decoding* of  $w_{1b}, w_{2b}, w_{3b}$ , is carried out by looking for one such triple so that the pairs of all involved sequences in blocks  $b$  and  $b+1$  are jointly typical with respect to the distribution  $P_{U_1 U_2 V T_1 T_2 X}$ . It can be seen that there are three dominating error events due to the nested coded structure. These events lead to the following three conditions

$$R_3 < I(X; Z_1 | U_1 U_2 V T_1 T_2) \quad (29a)$$

$$R_2 + R_3 < I(X T_1; Z_1 | V T_2 U_1) \quad (29b)$$

$$R_1 + R_2 + R_3 < I(X T_1; Z_1) \quad (29c)$$

for destination 1, while for destination 2 we have

$$R_3 < I(X; Z_2 | U_1 U_2 V T_1 T_2) \quad (30a)$$

$$R_2 + R_3 < I(X; Z_2 | V T_2 U_1) \quad (30b)$$

$$R_1 + R_2 + R_3 < I(X T_2; Z_2), \quad (30c)$$

This leads to the rate claimed in the Proposition.

For the Gaussian case, we fix (without claim of optimality) the following random variables  $T_2 = \sqrt{P_2} \tilde{T}_2$ ,  $T_1 = \sqrt{\eta_1 P_1} \tilde{T}_2 + \sqrt{\bar{\eta}_1 P_1} \tilde{T}_1$ ,  $V = \tilde{T}_2$ ,  $U_1 = \sqrt{\eta_2} \tilde{T}_2 + \sqrt{\bar{\eta}_2} \tilde{U}_1$ ,  $U_2 = \sqrt{\eta_3} U_1 + \sqrt{\bar{\eta}_3} \tilde{U}_2 + \sqrt{\bar{\eta}_3} \tilde{T}_1$  and  $X = \sqrt{\eta} P \tilde{X} + \sqrt{\bar{\eta}} P U_2$  with parameters defined as in Proposition 3 and  $\tilde{T}_1, \tilde{T}_2, \tilde{U}_1, \tilde{U}_2$  independent  $\mathcal{N}(0, 1)$ . This leads to the following mutual information terms:

$$\min_\theta I(X T_\theta; Z_\theta) \quad (31a)$$

$$= \min \left\{ \frac{1}{2} \log \left( \frac{1 + \alpha_1^2 P + \beta_1^2 P_1 + \sqrt{\bar{\eta} \eta_1 \eta_2 \eta_3' P P_1}}{2 \alpha_1 \beta_1 \left( \sqrt{\bar{\eta} \eta_1 \eta_2 \eta_3' P P_1} \right)} \right) \right\}, \quad (31b)$$

$$\frac{1}{2} \log \left( 1 + \alpha_2^2 P + \beta_2^2 P_2 + 2 \alpha_2 \beta_2 \sqrt{\bar{\eta} \eta_3' \eta_2 P P_2} \right) \quad (31c)$$

$$I(X T_1; Z_1 | V T_2 U_1) \quad (31d)$$

$$= \frac{1}{2} \log \left( \frac{1 + \alpha_1^2 (\eta + \bar{\eta} \eta_3') P}{\beta_1^2 \bar{\eta}_1 P_1 + 2 \alpha_1 \beta_1 \sqrt{\bar{\eta} \eta_3' \bar{\eta}_1 P_1 P}} \right) \quad (31e)$$

and

$$I(X; Z_2 | V T_2 U_1) = \frac{1}{2} \log \left( 1 + \alpha_2^2 (\eta + \bar{\eta} \eta_3') P \right) \quad (32a)$$

$$I(U_1; Y_\theta | V T_2) = \frac{1}{2} \log \left( 1 + \frac{\gamma_\theta^2 \bar{\eta} \eta_2 \eta_3' P}{1 + \gamma_\theta^2 (\eta + \bar{\eta} \eta_3') P} \right) \quad (32b)$$

$$I(X; Z_1 | U_1 U_2 V T_1 T_2) = \frac{1}{2} \log \left( 1 + \alpha_1^2 \eta P \right) \quad (32c)$$

$$I(X; Z_2 | U_1 U_2 V T_1 T_2) = \frac{1}{2} \log \left( 1 + \alpha_2^2 \eta P \right) \quad (32d)$$

$$I(U_2; Y_1 | U_1 V T_1 T_2) = \frac{1}{2} \log \left( 1 + \frac{\gamma_1^2 \bar{\eta} \eta_3'' P}{1 + \gamma_1^2 \eta P} \right). \quad (32e)$$

### C. Appendix-C: Proof of Proposition 6

As explained in Sec. V, the encoding scheme is the same as in Appendix-B but with  $R_3 = 0$  and thus  $U_2 = X$ . Moreover, decoding at the relay and destination 1 takes place in the same way so that the same rate conditions derived above remain valid here. However, since destination 2 only decodes message  $W_1$ , we need to further impose the condition  $R_1 < I(U_1 T_2; Z_2)$ , which concludes the proof. Finally, for the Gaussian model, the mutual information above reads

$$I(U_1 T_2; Z_2) = \frac{1}{2} \log \left( 1 + \frac{\alpha_2^2 \eta_3' P + \beta_2^2 P_2 + 2 \alpha_2 \beta_2 \sqrt{\eta_3' \eta_2 P P_2}}{1 + \bar{\eta}_3' P} \right). \quad (33)$$

### REFERENCES

- [1] E. C. van der Meulen, "Three-terminal communication channels," *Adv. Appl. Prob.*, vol. 3, pp. 120-154, 1971.
- [2] T. Cover and A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inform. Theory*, vol. 25, no. 5, pp. 572-584, Sep 1979.
- [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] P. Mitran, H. Ochiari, and V. Tarokh, "Space-time diversity enhancements using collaborative communications," *IEEE Trans. Information Theory*, vol. 51, no. 6, pp. 2041-2057, June 2005.
- [5] M. Katz, and S. Shamai (Shitz), "Transmitting to colocated users in wireless ad hoc and sensor networks," *IEEE Trans. Inform. Theory*, vol. 51, no. 10, pp. 3540-3563, Oct. 2005.
- [6] M. Katz and S. Shamai (Shitz), "Cooperative schemes for a source and an occasional nearby relay in wireless networks," to appear in *IEEE Trans. Inform. Theory*.
- [7] A. El Gamal and S. Zahedi, "Capacity of a class of relay channels with orthogonal components," *IEEE Trans. Information Theory*, vol. 51, no. 5, pp. 1815-1817, May 2005.
- [8] E. Csiszar and J. Korner, *Information theory: Coding theorems for discrete memoryless systems*, Academic Press, New York, 1981.
- [9] T. Cover, "Broadcast channels," *IEEE Trans. Inform. Theory*, vol. 18, no. 1, pp. 2-14, Jan. 1972.
- [10] S. Shamai and A. Steiner, "A broadcast approach for a single-user slowly fading MIMO channel," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2617-2635, Oct. 2003.
- [11] G. Kramer, *Topics in multiuser information theory*, Foundations and Trends in Communications and Information Theory, vol. 4, no. 4-5, pp. 265-444, Now Publishers, 2008.
- [12] A. Behboodi and P. Piantanida, "On the simultaneous relay channel with informed receivers," in *Proc. IEEE Intern. Symposium on Inform. Theory (ISIT 2009)*, June 28-July 3, Seoul, Korea, 2009.
- [13] K. Marton, "A coding theorem for the discrete memoryless broadcast channel," *IEEE Trans. Inform. Theory*, vol. 25, no. 3, pp. 306-311, May 1979.