

# A Two-Round Interactive Receiver Cooperation Scheme for Multicast Channels

Victor Exposito, Sheng Yang, Nicolas Gresset

► **To cite this version:**

Victor Exposito, Sheng Yang, Nicolas Gresset. A Two-Round Interactive Receiver Cooperation Scheme for Multicast Channels. 54th Annual Allerton Conference on Communication, Control, and Computing, IEEE, Sep 2016, Monticello, United States. pp.695-702, 10.1109/ALLERTON.2016.7852299 . hal-01435515

**HAL Id: hal-01435515**

**<https://hal-centralesupelec.archives-ouvertes.fr/hal-01435515>**

Submitted on 10 Jun 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A Two-Round Interactive Receiver Cooperation Scheme for Multicast Channels

Victor Exposito<sup>\*†</sup>, Sheng Yang<sup>\*</sup>, Nicolas Gresset<sup>†</sup>

<sup>\*</sup>L2S (UMR 8506), CentraleSupélec, CNRS, Université Paris-Sud, Université Paris-Saclay, Gif-sur-Yvette, France

<sup>†</sup>Mitsubishi Electric R&D Centre Europe, Rennes, France

Email: {victor.exposito,sheng.yang}@centralesupelec.fr, {v.exposito,n.gresset}@fr.merce.mee.com

**Abstract**—We consider the transmission of a common message from a transmitter to two receivers over a broadcast channel, also called a multicast channel in this case. The two receivers are allowed to cooperate with each other in full-duplex over non-orthogonal cooperation links. We investigate the information-theoretic upper and lower bounds on the broadcast rate. In particular, we propose a two-round cooperation scheme in which the receivers interactively perform compress-forward (CF) and then decode-forward (DF) to improve the achievable rate. Numerical results compare the proposed scheme to existing schemes and the cutset upper bound in the Gaussian case. We show that the proposed scheme outperforms the non-interactive DF and CF schemes as well as the noisy network coding scheme. The gain over the DF scheme becomes larger when the main channel becomes symmetric, while the gain over the CF scheme becomes larger when the main channel becomes asymmetric.

## I. INTRODUCTION

The increase of multimedia content delivery in wireless communication systems is leading, among many, to a spectrum crunch or interference intensification. In traditional broadcasting systems such as digital video broadcasting through satellite (DVBS), the same content is intended to a group of users. Multilayer strategies are often used so that the transmission rate is not limited by the worst user in terms of channel quality. However, if all users wish to obtain the same content quality, the worst user would set the rate and thus impact the whole group. With the recent study of device-to-device mechanisms in standards, user cooperation in close proximity becomes possible and would benefit all the users.

In this work, we investigate the broadcast channel (BC) with one transmitter sending a common message to two receivers, also called a multicast channel in this case. The receivers can cooperate through a cooperation link. The goal is to characterize the benefit of cooperation in terms of achievable broadcast rate through an information-theoretic analysis. Since the receivers also transmit signals through the cooperation link, our channel is a mixture of the BC and the relay channel (RC).

BCs were introduced in [1]. The capacity of BCs remains unknown except for special cases such as the stochastically degraded BC [2] for which superposition coding is optimal. The largest achievable rate-region known to date for the general case is based on binning [3]. RCs were introduced in [4]. The compress-forward (CF) and decode-forward (DF), fundamental relaying strategies, were proposed in [5]. Relaying strategies based on CF and DF were proposed for larger

networks [6]–[9]. The capacity of the BC with cooperation is unknown in general, except for special cases such as the physically degraded main channel. The setup of the current work has been partially studied in [10], referred to as BCs with cooperative decoders, and in [11], [12], referred to as relay BCs. A BC with orthogonal cooperation links was considered in [10]. In [11], [12], although the cooperation links are not restricted to be orthogonal, the authors assumed that either the main channel is degraded or the cooperation link is uni-directional. It is worth noting that achievable rate regions of both common and private messages were provided in [10]–[12]. Although distributed decode-forward [13] seems promising, its bound is hard to evaluate due to the high number of parameters to optimize. For this reason, it is not considered in our paper.

Unlike previous works, we study a general multicast channel with a bi-directional non-orthogonal cooperation link. First, we present for our channel the cutset upper bound and four lower bounds derived from existing results in the literature. Two of the non-interactive schemes are, 1) the one-sided cooperation scheme where the “strongest” receiver decodes the message first and then serves as DF relay to help the “weakest” receiver decode, and 2) the one-round cooperation scheme where the users serve as CF relay to help each other decode. Then a new two-round interactive cooperation scheme is proposed as the main result of our paper. In this scheme, one receiver first relays the received signal using CF to help the other one decode during the first round, which in turn relays back side information on the decoded message during the second round to help the first one decode. The proposed scheme is based on block Markov superposition coding, contrary to [10]–[12].

Sec. II introduces the system model and the Single-Input Single-Output (SISO) Gaussian BC as a special case. Sec. III presents existing upper and lower bounds. In Sec. IV, we detail the proposed scheme and show that it surpasses existing schemes regarding the achievable rate. Numerical results are provided in Sec. V for the SISO Gaussian BC, and the Multiple-Input Single-Output (MISO) Gaussian BC introduced in Sec. IV. We show that the gain of the proposed scheme over the one-sided scheme is larger when the main channel becomes symmetric while the gain over the one-round scheme (as well as the noisy network coding scheme [7]) is larger when the main channel becomes asymmetric. In the symmetric MISO case the proposed scheme strictly outperforms the one-round

scheme while they were equal in the SISO counterpart.

## II. SYSTEM MODEL

We consider a simple multicast network where one transmitter sends the same information to two receivers through the main channel. The two receivers can cooperate with each other in full-duplex, i.e., they can transmit and receive simultaneously, through a bi-directional non-orthogonal cooperation link. This setup includes, 1) the cooperation links orthogonal to the main channel, orthogonal links being either physically separated medium, e.g., using different transmission technologies over different resources, or created with artificial orthogonalization, e.g., in time or frequency and, 2) the half-duplex mode if the receivers transmit and receive at a different time. To derive upper and lower bounds on the achievable rate, we resort to information theory. The current channel belongs to a class of stationary memoryless channels shown in Fig. 1,  $(\mathcal{X} \times \mathcal{X}_1 \times \mathcal{X}_2, p(y_1, y_2|x, x_1, x_2), \mathcal{Y}_1 \times \mathcal{Y}_2)$ , defined as  $p(y_1^n, y_2^n|x^n, x_1^n, x_2^n) = \prod_{i=1}^n p(y_{1i}, y_{2i}|x_i, x_{1i}, x_{2i})$  where  $x^n \in \mathcal{X}^n$ ,  $x_1^n \in \mathcal{X}_1^n$ , and  $x_2^n \in \mathcal{X}_2^n$  are the sequences of transmitted signals from the transmitter, receiver 1 and 2, respectively, and  $y_1^n \in \mathcal{Y}_1^n$  and  $y_2^n \in \mathcal{Y}_2^n$  are the sequences of received signals at the receivers 1 and 2, respectively. The probability distribution of the channel is known at every node (perfect channel state information (CSI) at the transmitter and receivers) by assumption. The common message  $M$  is assumed to be uniformly distributed in  $\mathcal{M} \triangleq [1 : 2^{nR}]$  where  $R$  is the number of bits per channel use. An encoder at the transmitter side is a map  $f_i^{(n)}$  from the message  $M$  to the sequence of input symbols  $x^n$ , an encoder at the receiver  $k$ ,  $k = 1, 2$ , is a sequence of maps  $\{f_{k,i}^{(n)}\}_i$  from the past received symbols  $y_k^{i-1}$  to the transmitted symbol  $x_{k,i}$ . A decoder at the receiver  $k$  is a map  $\{g_{k,i}^{(n)}\}_i$  from the received sequence  $y_k^n$  to  $\hat{M}^{(k)} \in \mathcal{M}$ . The probability of error is defined as  $P_e^{(n)} \triangleq \Pr(M \neq \hat{M}' \text{ or } M \neq \hat{M}'')$ . Finally, a rate  $R$  is achievable if there exist a sequence of encoders/decoders  $(f_i^{(n)}, \{f_{1,i}^{(n)}\}_i, \{f_{2,i}^{(n)}\}_i, \{g_{1,i}^{(n)}\}_i, \{g_{2,i}^{(n)}\}_i)$  such that  $P_e^{(n)} \rightarrow 0$  when  $n \rightarrow \infty$ . Note that we obtain an orthogonal channel if, 1) we split  $\mathcal{Y}_k = \mathcal{Y}_k^m \times \mathcal{Y}_k^c$ , 2) we split  $Y_1 = (Y_1^m, Y_1^c)$ ,  $Y_2 = (Y_2^m, Y_2^c)$  with  $Y_k^m \in \mathcal{Y}_k^m$ ,  $Y_k^c \in \mathcal{Y}_k^c$ ,  $k = 1, 2$ , and 3) we have

$$p(y_1, y_2|x, x_1, x_2) = p(y_1^m, y_2^m|x)p(y_1^c|x_2)p(y_2^c|x_1), \quad (1)$$

i.e., the received signals from the main channel are independent of the received signals from the cooperation links. The information-theoretic bounds derived under those general classes of channels can be specialized for any stationary memoryless channels compliant to the corresponding requirements. We decided to derive the bounds for the Gaussian BC (Gaussian inputs and noises) as a special case of the class of non-orthogonal stationary memoryless channels. The received signal at an instant  $i$  is written as

$$y_1[i] = h_{11}x[i] + h_{21}x_2[i] + z_1[i], \quad (2)$$

$$y_2[i] = h_{22}x[i] + h_{12}x_1[i] + z_2[i] \quad (3)$$

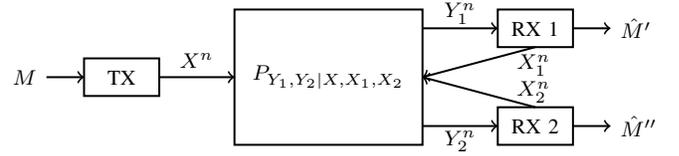


Fig. 1: The BC with receiver cooperation.

where the transmitted signals are subject to the same average power constraint  $\sum_{i=1}^n |x[i]|^2 \leq nP$ ,  $\sum_{i=1}^n |x_k[i]|^2 \leq nP_k$ ,  $k = 1, 2$ , i.e.,  $P_1 = P_2 = P$ , and the additive white Gaussian noises (AWGN)  $z_k[i] \sim \mathcal{CN}(0, \sigma^2)$  are independent across resources and receivers by assumption. The signal-to-noise ratios of the main channels are  $\text{SNR}_k = |h_k|^2 \frac{P}{\sigma^2}$ ,  $k = 1, 2$ , and those of the cooperative links are  $\text{SNR}_{kl} = |h_{kl}|^2 \frac{P}{\sigma^2}$ ,  $k \neq l \in \{1, 2\}$ . Note that in the above model, self-interference is not considered at the receivers, i.e., it can be removed using the perfect CSI assumption. In practice, self-interference could be dealt with by data processing or resource decoupling.

## III. EXISTING UPPER AND LOWER BOUNDS

In the following, we present the cutset upper bound, and four lower bounds readily derived from existing schemes in the literature. Proofs are omitted and the readers are referred to [14], [15] for details.

### A. Cutset Upper Bound

According to the cutset upper bound [15, Th. 18.1], the capacity of the channel is upper bounded by

$$C \leq R_{\text{CS}} \triangleq \max_{p(x, x_1, x_2)} \min\{I(X; Y_1, Y_2 | X_1, X_2), I(X, X_2; Y_1 | X_1), I(X, X_1; Y_2 | X_2)\}. \quad (4)$$

The capacity cannot be larger than any of those three cases, 1) both receivers cooperate perfectly to decode the source message, i.e.,  $I(X; Y_1, Y_2 | X_1, X_2)$ , 2) the transmitter and receiver 2 cooperate perfectly to send the message to receiver 1, i.e.,  $I(X, X_2; Y_1 | X_1)$ , and 3) same for receivers 1 and 2 exchanging roles. The maximization is performed over all possible input distributions  $p(x, x_1, x_2)$ , since the cooperation is assumed to be perfect in the upper bound.

### B. No Cooperation

A natural lower bound is obtained by ignoring the cooperation link, it is in fact the capacity of a compound channel

$$C \geq R_{\text{NC}} \triangleq \max_{p(x)p(x_1)p(x_2)} \min\{I(X; Y_1 | X_1, X_2), I(X; Y_2 | X_1, X_2)\}. \quad (5)$$

### C. One-Sided Cooperation

We can obtain another lower bound when one of the cooperation links is ignored, e.g., receiver 1 does not exploit the presence of the other and always decodes the source message first. After decoding each message, receiver 1 and the source

can cooperatively send the message to receiver 2, through the main channel and the cooperation link. The achievable rate in this case is similar to the one of a conventional DF scheme [5], [15] when receiver 1 is considered as relay, i.e.,  $\min\{I(X; Y_1|X_1), I(X, X_1; Y_2|X_2)\}$ . Note that the cooperation between the transmitter and receiver 1 is possible after the latter decodes the message, thus the rate can be maximized over all possible  $p(x, x_1)p(x_2)$ . Same for receivers 1 and 2 exchanging roles. Finally, we maximize the following lower bound by choosing the decoding order that gives a higher rate.

$$C \geq R_{1SC} \triangleq \max_{k \neq l} \max_{\mathcal{P}_{l \rightarrow k}} \min\{I(X; Y_l|X_k, X_l), I(X, X_l; Y_k|X_k)\} \quad (6)$$

where  $\mathcal{P}_{l \rightarrow k}$  is the set of distributions  $p(x, x_l)p(x_k)$ . Note that the short message noisy network coding with DF option (SNNC-DF) [8, Th. 2] is designed for the multiple-relay channel (unicast over a cooperative network) in which relays can use either quantize-forward (QF) or DF. Since we consider a multicast channel, the DF option of the SNNC-DF has to be forced otherwise one receiver would not be able to decode the message in general, giving an identical bound.

#### D. Two-Sided One-Round Cooperation

To exploit both cooperation links, a simple scheme based on CF scheme [5], [15] works as follows. Both receivers compress simultaneously and independently their own observations and then send the compression index to the other receiver using the cooperation links. In other words, each receiver acts as a CF relay to the other receiver.

$$C \geq R_{1RC} \triangleq \max_{\mathcal{P}} \min_{k \neq l} \min\left\{I(X; Y_l, \hat{Y}_k|X_k, X_l), I(X, X_k; Y_l|X_l) - I(Y_k; \hat{Y}_k|X, X_k, X_l, Y_l)\right\} \quad (7)$$

where  $\mathcal{P}$  is the set of conditional distribution  $p(x)p(x_1)p(x_2)p(\hat{y}_1|x_1, y_1)p(\hat{y}_2|x_2, y_2)$  with  $|\hat{\mathcal{Y}}_k| \leq |\mathcal{X}_k||\mathcal{Y}_k| + 1$ ,  $k = 1, 2$ . For brevity, we also refer to this scheme as one-round cooperation scheme.

#### E. Noisy Network Coding

Finally, we can apply the noisy network coding (NNC) scheme [7, Th. 1] and obtain the following lower bound.

$$C \geq R_{NNC} \triangleq \max_{\mathcal{P}} \min_{k \neq l} \min\left\{I(X; Y_l, \hat{Y}_k, \hat{Y}_l|X_k, X_l), I(X, X_k; Y_l, \hat{Y}_l|X_l) - I(Y_k; \hat{Y}_k|X, X_k, X_l, Y_l, \hat{Y}_l)\right\} \quad (8)$$

where  $\mathcal{P}$  is the set of conditional distributions  $p(x)p(x_1)p(x_2)p(\hat{y}_1|x_1, y_1)p(\hat{y}_2|x_2, y_2)$  with  $|\hat{\mathcal{Y}}_k| \leq |\mathcal{X}_k||\mathcal{Y}_k| + 1$ ,  $k = 1, 2$ . In this scheme, the same long message (high rate) is sent using independent codebooks in each block, and then each receiver uses QF, i.e., quantizes the observation and relays the description in the next block. The message is decoded only at the end of the whole transmission at both receivers. Note that the short message noisy network coding (SNNC) [8, Th. 1], in which the long message is

cut into small independent ones each sent using independent codebooks in each block gives an identical bound.

**Remark 1.** Note that the NNC and the one-round cooperation lower bounds coincide by applying the Markov chain  $(X, X_k) \leftrightarrow (X_l, Y_l) \leftrightarrow \hat{Y}_l$ ,  $k \neq l$  on the right-hand side of (8) to obtain (7). A similar result arises in the three-node RC where the NNC reduces to the CF [15, Rmk. 18.5]. Intuitively, the multi-round cooperation introduced by the NNC scheme cannot improve over the one-round cooperation scheme in a two-receiver channel, since the receivers do not benefit from the multi-round compression which is not better than one-round compression. The same conclusion does not hold when the network has more than two receivers.

Due to the same achievable rate, we do not distinguish between the one-round cooperation and NNC schemes hereafter.

#### IV. PROPOSED COOPERATION SCHEME

In this section, we propose a two-round interactive cooperation scheme. In a nutshell, the first round corresponds to a CF cooperation, for receiver  $k$  to help receiver  $l$  decode, and the second round corresponds to a DF cooperation, for receiver  $l$  to help receiver  $k$  decode. The main result of this work is the following proposition.

**Proposition 1.** With the two-round interactive cooperation scheme, we achieve the following lower bound

$$C \geq R_{2RC} \triangleq \max_{k \neq l} \max_{\mathcal{P}_{k \leftrightarrow l}} \min\left\{I(X, X_l; Y_k|X_k), I(X; Y_l, \hat{Y}_k|X_k, X_l), I(X, X_k; Y_l|X_l) - I(Y_k; \hat{Y}_k|X, X_k, X_l, Y_l)\right\} \quad (9)$$

where  $\mathcal{P}_{k \leftrightarrow l}$  is the set of conditional distributions  $p(x, x_l)p(x_k)p(\hat{y}_k|x_k, y_k)$  with  $|\hat{\mathcal{Y}}_k| \leq |\mathcal{X}_k||\mathcal{Y}_k| + 1$ .

**Remark 2.** Note that a time-sharing random variable can be added to improve further the rate in all the information theoretic bounds presented in this paper.

We explain the proposed scheme with the help of Fig. 2. Due to the space restrictions, the formal proof will be provided in the full version of the paper, and the formal setup is in the Appx. Our scheme is based on block Markov superposition coding. Without loss of generality, we assume that receiver 1 performs CF to help receiver 2 decode first. In each block  $b$ , the transmitter sends a codeword as a function of the current message  $m_b$  and the past message  $m_{b-2}$  (with two blocks of delay). Upon the reception of  $Y_{1,b}$ , receiver 1 performs a compression and obtains a description  $\hat{Y}_1(k_b|l_{b-1})$  using a source codebook indexed by  $l_{b-1}$ . The index of the description  $k_b$  is further compressed into  $l_b$  using a binning procedure. The compressed index  $l_b$  is then sent during the next block  $b+1$  through the cooperation link to receiver 2, with receiver 1's codebook. At the end of block  $b+1$ , receiver 2 jointly decodes the compression index  $l_b$  from  $Y_{2,b+1}$  to get side information about  $Y_{1,b}$  and its own observation  $Y_{2,b}$  to decode  $m_b$  as  $\hat{m}_b''$ . The first round cooperation is complete. It is almost

the same as the CF scheme with receiver 1 as relay and receiver 2 as destination, except that in this case receiver 2 is also transmitting constantly. One can show that  $\hat{m}_b'' = m_b$  with high probability if the rate satisfies

$$R \leq \min \left\{ I(X; Y_2, \hat{Y}_1 | X_1, X_2), \right. \\ \left. I(X, X_1; Y_2 | X_2) - I(Y_1; \hat{Y}_1 | X, X_1, X_2, Y_2) \right\} \quad (10)$$

as in the CF scheme [15, Th. 16.4], except for the conditioning on the transmitted signals  $X_1$  and  $X_2$  of receivers 1 and 2.

The second round cooperation related to  $m_b$  starts once receiver 2 recovers  $m_b$  at the end of block  $b + 1$ . The side information on  $m_b$  is sent during the next block  $b + 2$  through the cooperation link to receiver 1, with receiver 2's codebook. Meanwhile, the source sends the new  $m_{b+2}$  using a codebook indexed by  $m_b$ , i.e., the codewords sent by the source and receiver 2 are correlated. At the end of block  $b + 2$ , receiver 1 simultaneously decodes the two observations that are related to  $m_b$ , namely,  $Y_{1,b}$  and  $Y_{1,b+2}$  to decode  $m_b$  as  $\hat{m}_b'$ . One can show that  $\hat{m}_b' = m_b$  with high probability if the rate satisfies

$$R \leq I(X, X_2; Y_1 | X_1) \quad (11)$$

as if the source and receiver 2 form a virtual transmitter to send the message to receiver 1. Therefore, for both receivers to decode  $m_b$  successfully with high probability, both (10) and (11) should be satisfied. Same for receivers 1 and 2 exchanging roles. Finally, we maximize the following lower bound by choosing the decoding order that gives a higher rate from which we obtain the rate given in Prop. 1.

**Remark 3.** Note that the proposed scheme has low latency (two blocks for the weakest user) since receiver 1 uses sliding window decoding, instead of high latency (wait until the last block for the weakest user) if it used backward decoding, and that both decoding procedures give identical performances.

**Corollary 1.** The proposed scheme outperforms both the one-sided and one-round cooperation schemes,

$$R_{CS} \geq R_{2RC} \geq R_{ISC} \geq R_{NC}, \quad \text{and} \quad (12)$$

$$R_{CS} \geq R_{2RC} \geq R_{IRC} = R_{NNC} \geq R_{NC}. \quad (13)$$

*Proof.* We show that the proposed scheme is at least as good as the one-sided scheme by restricting  $\hat{Y}_k$ ,  $k = 1, 2$ , to be independent of  $(X_k, Y_k)$  in  $\mathcal{P}_{k \leftrightarrow l}$ , thus  $R_{2RC}$  in (9) reduces to  $R_{1SC}$  in (6). By rewriting (7) for the one-round cooperation as

$$\max_{\mathcal{P}} \min \left\{ I(X; Y_l, \hat{Y}_k | X_k, X_l), I(X; Y_k, \hat{Y}_l | X_k, X_l), \right. \\ \left. I(X, X_k; Y_l | X_l) - I(Y_k; \hat{Y}_k | X, X_k, X_l, Y_l), \right. \\ \left. I(X, X_l; Y_k | X_k) - I(Y_l; \hat{Y}_l | X, X_k, X_l, Y_k) \right\}, \quad (14)$$

it is readily shown that each term in the minimization on the right-hand side of (9) is larger than at least one term in (14). By restricting  $p(x, x_l) = p(x)p(x_l)$  in  $\mathcal{P}_{k \leftrightarrow l}$ , we finally show that a lower bound of (9) is larger than (7).  $\square$

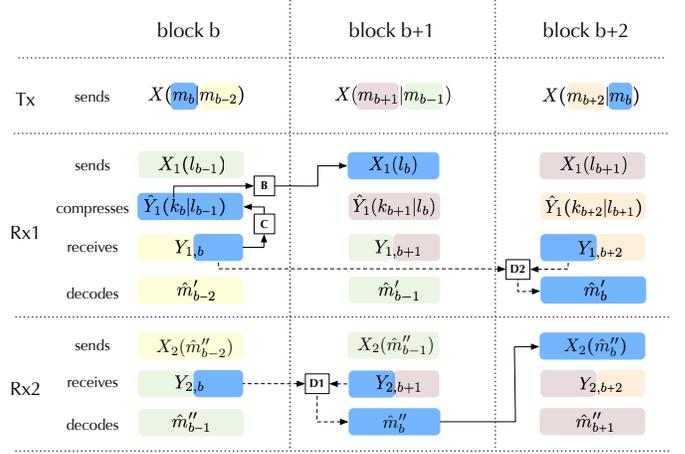


Fig. 2: Illustration of the proposed scheme in strategy  $STG_{1 \rightarrow 2 \rightarrow 1}$ . The encoding and decoding related to the message  $m_b$  are in blue. Solid arrows represent encoding and dashed arrows decoding. Square blocks B, C, D1, D2 stand for “binning”, “compression”, “decoder 1”, and “decoder 2”.

From Prop. 1, we obtain the following rate when the cooperation links are orthogonal to the main channel as in (1).

**Corollary 2** (Orthogonal cooperation links). *The proposed scheme achieves the following lower bound in a multicast channel with orthogonal cooperation links*

$$R_{2RC}^{ortho} \triangleq \max_{k \neq l} \max_{\mathcal{P}_{k \leftrightarrow l}^o} \min \left\{ I(X; Y_k^m) + C_{l,k}, I(X; Y_l^m, \hat{Y}_k), \right. \\ \left. I(X; Y_l^m) - I(Y_k^m; \hat{Y}_k | X, Y_l^m) + C_{k,l} \right\} \quad (15)$$

where  $\mathcal{P}_{k \leftrightarrow l}^o$  is the set of distributions  $p(x)p(\hat{y}_k|y_k^m)$  with  $|\mathcal{Y}_k^m| \leq |\mathcal{Y}_k^m| + 1$ , and  $C_{l,k} \triangleq \max_{p(x_l)} I(X_l; Y_k^c)$  is the capacity of the cooperation link from receiver  $l$  to receiver  $k$ .

*Proof.* From the definition in (1), we have

$$I(X, X_l; Y_k | X_k) = I(X; Y_k^m) + I(X_l; Y_k^c) \quad (16)$$

$$I(X; Y_l, \hat{Y}_k | X_k, X_l) = I(X; Y_l^m, \hat{Y}_k) \quad (17)$$

$$I(X, X_k; Y_l | X_l) - I(Y_k; \hat{Y}_k | X, X_k, X_l, Y_l) \\ = I(X; Y_l^m) + I(X_k; Y_l^c) - I(Y_k^m; \hat{Y}_k | X, Y_l^m) \quad (18)$$

which follows directly from the independence between  $(X, Y_1^m, Y_2^m, \hat{Y}_1, \hat{Y}_2)$  and  $(X_1, X_2, Y_1^c, Y_2^c)$ . We obtain (15) by, 1) maximizing separately  $I(X_k; Y_l^c)$  and  $I(X_l; Y_k^c)$ , which is possible because each term only depends on its own input distribution, and 2) maximizing the rest of the terms by letting  $X$  be independent of  $X_l$  in  $\mathcal{P}_{k \leftrightarrow l}^o$ , which is without loss of optimality since these terms do not depend on  $X_l$  given  $X$ .  $\square$

**Remark 4.** Note that under the same setting, the rate (15) coincides with the one derived in [10]. It is worth mentioning that the coding techniques in both works are conceptually different. Our results are based on the block Markov superposition coding aiming for the general non-orthogonal cooperation links whereas the schemes in [10] are based on

separated coding only suitable for orthogonal cooperation links.

**Corollary 3** (Gaussian channel). *The proposed scheme achieves the following lower bound expressed explicitly from Prop.1 in a Gaussian multicast channel with non-orthogonal cooperation links*

$$R_{2RC}^{Gauss} \triangleq \max_{k \neq l} \max_{\rho \in [0,1]} \min \left\{ \log(1 + \text{SNR}_k + \text{SNR}_{lk} + 2\sqrt{\text{SNR}_k \text{SNR}_{lk} \rho}), \right. \\ \left. \log \left( 1 + \left( \text{SNR}_l + \frac{\text{SNR}_k}{1 + \Delta_k^*(\rho)} \right) (1 - \rho^2) \right) \right\} \quad (19)$$

where  $\Delta_k^*(\rho) \triangleq \frac{1 + (\text{SNR}_k + \text{SNR}_l)(1 - \rho^2)}{\text{SNR}_{kl}}$ .

*Proof.* We let  $(X, X_l) \sim \mathcal{CN}(0, \mathbf{Q})$  and  $X_k \sim \mathcal{CN}(0, P)$  be independent, with

$$\mathbf{Q} \triangleq P \begin{bmatrix} 1 & \rho e^{j\theta} \\ \rho e^{-j\theta} & 1 \end{bmatrix}, \quad \rho \in [0, 1], \quad \theta \in [0, 2\pi), \quad (20)$$

$\hat{Y}_k = Y_k + N_k$  with  $N_k \sim \mathcal{CN}(0, \Delta_k)$ ,  $k = 1, 2$ , being the compression noise. We follow three steps to obtain the expression in (19). First, we apply (9) with the above distributions. Then, we optimize over  $\theta$  according to the phases of the channel coefficients. Finally, we maximize over  $\Delta_k \geq 0$  for a given  $\rho$  and obtain the optimal value of  $\Delta_k$  as  $\Delta_k^*(\rho)$ .  $\square$

We further consider the MISO  $2 \times 1$  Gaussian BC case with the following signal model at an instant  $i$

$$y_1[i] = \mathbf{h}_1^T \mathbf{x}[i] + h_{21}x_2[i] + z_1[i], \quad (21)$$

$$y_2[i] = \mathbf{h}_2^T \mathbf{x}[i] + h_{12}x_1[i] + z_2[i] \quad (22)$$

where  $\mathbf{h}_1, \mathbf{h}_2 \sim \mathbb{C}^{2 \times 1}$  are the channel vectors from the transmitter to, respectively, the receiver 1 and 2,  $h_{kl} \sim \mathbb{C}$ ,  $k \neq l \in \{1, 2\}$  is the channel gain from receiver  $k$  to receiver  $l$ , and  $z_k \sim \mathcal{CN}(0, \sigma^2)$ ,  $k = 1, 2$  are the AWGNs independent across resources and receivers by assumption. Note that as in the SISO case, all channel coefficients are assumed to be known and remain constant during the transmission, and that self-interference is not considered at the receivers. Assume that the transmitted signals are subject to the same average power constraint  $\sum_{i=1}^n |\mathbf{x}[i]|^2 \leq nP$ ,  $\sum_{i=1}^n |x_k[i]|^2 \leq nP$ ,  $k = 1, 2$ . For simplicity, we fix the channel coefficients

$$\mathbf{h}_1^T = \left[ \sqrt{\frac{\text{SNR}_1 \sigma^2}{2P}} \quad \sqrt{\frac{\text{SNR}_1 \sigma^2}{2P}} \right] \quad (23)$$

$$\mathbf{h}_2^T = \left[ \sqrt{\frac{\text{SNR}_2 \sigma^2}{2P}} \quad -\sqrt{\frac{\text{SNR}_2 \sigma^2}{2P}} \right] \quad (24)$$

$$h_{kl} = \sqrt{\frac{\text{SNR}_{kl} \sigma^2}{P}}, \quad (25)$$

where  $\sigma^2$  is discarded afterward in the channel coefficients since it appears in the SNRs' denominator, and by considering

normalized SNRs. For the proposed scheme, we let  $(\mathbf{X}, X_l) \sim \mathcal{CN}(0, \mathbf{Q})$  and  $X_k \sim \mathcal{CN}(0, P)$  be independent, with

$$\mathbf{Q} \triangleq P \begin{bmatrix} \mathbf{Q}_{\mathbf{X}} & \mathbf{Q}_{\mathbf{X}, X_l} \\ \mathbf{Q}_{\mathbf{X}, X_l}^H & P \end{bmatrix} \succeq 0 \quad (26)$$

$$\mathbf{Q}_{\mathbf{X}|[X_l, X_k]} = \mathbf{Q}_{\mathbf{X}|X_l} = \mathbf{Q}_{\mathbf{X}} - \mathbf{Q}_{\mathbf{X}, X_l} P^{-1} \mathbf{Q}_{\mathbf{X}, X_l}^H \succeq 0. \quad (27)$$

Note that  $\mathbf{Q} \succeq 0$  is equivalent to  $\mathbf{Q}_{\mathbf{X}|X_l} \succeq 0$  and  $P \geq 0$  since  $\mathbf{Q}_{\mathbf{X}|X_l}$  is the Schur complement of  $P$  in  $\mathbf{Q}$ . It follows that  $\mathbf{Q}$  is uniquely determined by  $\mathbf{Q}_{\mathbf{X}|X_l}$ ,  $\mathbf{Q}_{\mathbf{X}, X_l}$ , and  $P$ . The matrices are parametrized as follows without loss of generality

$$\mathbf{Q}_{\mathbf{X}|X_l} = \begin{bmatrix} A & \sqrt{C} e^{j\theta_1} \\ \sqrt{C} e^{-j\theta_1} & B \end{bmatrix} \succeq 0 \quad (28)$$

$$\mathbf{Q}_{\mathbf{X}, X_l} = \sqrt{D} \begin{bmatrix} u e^{j\theta_2} \\ \sqrt{1 - u^2} \end{bmatrix}, \quad (29)$$

where  $A, B, C, D \geq 0$ ,  $u \in [0, 1]$ , and  $\theta_1, \theta_2 \in [0, 2\pi)$ . Note that  $\mathbf{Q}_{\mathbf{X}, X_l} / \sqrt{D}$  is a unit vector, and that the eigenvalues of  $\mathbf{Q}_{\mathbf{X}|X_l}$  are non-negative if and only if  $AB \geq C$ . Then

$$\mathbf{Q}_{\mathbf{X}} = \mathbf{Q}_{\mathbf{X}|X_l} + \mathbf{Q}_{\mathbf{X}, X_l} P^{-1} \mathbf{Q}_{\mathbf{X}, X_l}^H \succeq 0, \quad (30)$$

and satisfy the power constraint  $\text{Tr}(\mathbf{Q}_{\mathbf{X}}) \leq P$  which implies

$$A + B + \frac{D}{P} \leq P. \quad (31)$$

It can be further shown that assuming the equality in (31) does not bring loss of optimality. The compression noise is  $\hat{Y}_k = Y_k + N_k$  with  $N_k \sim \mathcal{CN}(0, \Delta_k)$ ,  $k = 1, 2$ .

**Corollary 4** (MISO Gaussian channel). *The proposed scheme achieves the following lower bound expressed explicitly from Prop. 1 in a MISO  $2 \times 1$  Gaussian multicast channel with non-orthogonal cooperation links*

$$R_{2RC}^{MISO \text{ Gauss}} \triangleq \max_{k \neq l} \max_{\mathbf{Q} \succeq 0, \text{Tr}(\mathbf{Q}) \leq 2P} \min \left\{ \log(1 + \kappa_k), \log \left( 1 + \gamma_l + \frac{\gamma_k + \gamma_m}{1 + \Delta_k^*} \right) \right\} \quad (32)$$

where  $\Delta_k^*(A, B, C, \theta_1) \triangleq \frac{1 + \gamma_k + \gamma_l + \gamma_m}{\kappa_l - \gamma_l}$ , with

$$\kappa_l = (A + B \pm 2\sqrt{C} \cos(\theta_1)) \frac{\text{SNR}_l}{2P} + \text{SNR}_{kl} \quad (33)$$

$$\kappa_k = (P \mp 2\sqrt{C} \cos(\theta_1) \mp 2(P - (A + B)u\sqrt{1 - u^2} \cos(\theta_2))) \frac{\text{SNR}_k}{2P} + \text{SNR}_{lk} + 2\sqrt{P - (A + B)(u \cos(\theta_2) \mp \sqrt{1 - u^2})} \sqrt{\frac{\text{SNR}_k \text{SNR}_{lk}}{2P}} \quad (34)$$

$$\gamma_l = (A + B \pm 2\sqrt{C} \cos(\theta_1)) \frac{\text{SNR}_l}{2P} \quad (35)$$

$$\gamma_k = (A + B \mp 2\sqrt{C} \cos(\theta_1)) \frac{\text{SNR}_k}{2P} \quad (36)$$

$$\gamma_m = 4(AB - C) \frac{\text{SNR}_l \text{SNR}_k}{4P^2}. \quad (37)$$

*Proof.* Under the considered setup of the MISO  $2 \times 1$ , we follow two steps to obtain the expression in (32). First,

we apply (9) with the corresponding distributions. Then, we maximize over  $\Delta_k \geq 0$  for a given tuple  $(A, B, C, \theta_1)$  and obtain the optimal value of  $\Delta_k$  as  $\Delta_k^*(A, B, C, \theta_1)$ .  $\square$

In Cor. 4, the upper sign of the plus-minus signs correspond to  $\text{STG}_{2 \rightarrow 1 \rightarrow 2}$  and the lower sign to  $\text{STG}_{1 \rightarrow 2 \rightarrow 1}$  due to (23) and (24), and the constraints are summarized as follows

$$A, B, C \geq 0, AB \geq C, A + B \leq P \quad (38)$$

$$D = P(P - (A + B)) \quad (39)$$

$$u \in [0, 1], \theta_1, \theta_2 \in [0, 2\pi). \quad (40)$$

## V. NUMERICAL RESULTS

In this section, we first focus on the SISO Gaussian BC as defined in (2) and (3), and evaluate through numerical simulations the achievable rate of the proposed scheme given in (19), as well as the rate bounds provided in Sec. III. Note that to provide a fair comparison the parameters such as input correlation and compression noise variance are optimized for each bound. We study the impact of the cooperation link on the throughput of the channel. We assume that the SNR of the cooperation links is symmetric, i.e.,  $\text{SNR}_{12} = \text{SNR}_{21} = \text{SNR}_{\text{coop}}$ . In Fig. 3, we fix the SNR of the main channel, and plot the throughput in terms of spectral efficiency (bit/s/Hz) by varying  $\text{SNR}_{\text{coop}}$  from  $-40$  dB to  $40$  dB. In Fig. 3a, the main channel is symmetric with a SNR of  $10$  dB at each receiver. In this case, the proposed scheme has the same performance as the one-round scheme, both of which go from the “no cooperation” lower bound  $R_{\text{NC}}$ , when the cooperation link is weak, to the cutset upper bound  $R_{\text{CS}}$ , when the cooperation link is strong, while the one-sided scheme has the same performance as  $R_{\text{NC}}$ . When the channel becomes asymmetric, as shown in Fig. 3b and Fig. 3c, both the one-round and the proposed schemes still bridge  $R_{\text{NC}}$  and  $R_{\text{CS}}$  as the strength of the cooperation link increases. Note that the one-sided scheme outperforms the one-round scheme with weak cooperation, and conversely with strong cooperation, and that as demonstrated in the information theoretic setting and its analysis in Cor. 1, the proposed scheme outperforms both the one-sided and the one-round schemes in all configurations. We insist on the fact that the proposed scheme surpasses a simple selection scheme between the one-sided and the one-round schemes. This results from the interaction, the first receiver (using CF) brings knowledge to the second receiver (using DF) that improves the performances of the DF, thus the latter one can improve its help toward the first one. Note that in Fig. 3a the proposed scheme has the same performance as the one-round scheme for completely symmetric channels, which can be explained as follows, 1) with the one-round scheme, both receivers decode simultaneously, using the other one as a CF relay, 2) with the two-round scheme, one of the receivers decodes first, using the other one as a CF relay. When the channel is completely symmetric, both receivers in the one-round scheme achieve the same performance as the first decoder of the two-round scheme, which proves that the one-round scheme cannot achieve a lower rate than the two-round

scheme for completely symmetric channels. However, from Cor. 1, we know that the two-round scheme cannot be worse than the one-round scheme, thus both schemes should have the same performance for completely symmetric channels.

We focus in a second time on the MISO  $2 \times 1$  Gaussian BC as defined in (21) and (22), and evaluate through numerical simulations the achievable rate of the proposed scheme given in (32), as well as the rate bounds provided in Sec. III for their corresponding optimal parameters. In Fig. 4a, the main channel is symmetric. Contrary to the SISO case, in the MISO case the proposed scheme strictly outperforms the one-round scheme due to the beamforming optimization. In the one-round scheme, the equipartition of the power allocation enhances the achievable rate of both receivers, while in the proposed scheme, the beam in cooperation with the receiver using DF enhances the link towards the receiver using CF, virtually establishing an asymmetric channel leading to an improvement of the achievable rate. Same for the one-sided scheme no longer equal to the “no cooperation” lower bound. When the channel becomes asymmetric, as shown in Fig. 4b and Fig. 4c, the comments for the SISO case also hold for the MISO case.

## VI. CONCLUSION

In this paper, we investigated the impact of receiver cooperation on the throughput of a two-receiver multicast channel. We proposed an interactive cooperation scheme and proved that it outperforms existing schemes in which no interaction is exploited. The numerical results validated the theoretical analysis focusing on the Gaussian BC. Our results revealed that interaction is particularly helpful when the main channel is asymmetric, and that the MISO BC enhances further the interaction properties. Our ongoing work includes the extension of the current results to a larger network.

## APPENDIX SETUP OF PROP. 1

We consider  $\text{STG}_{2 \rightarrow 1 \rightarrow 2}$ . A sequence of  $(b - 2)$  messages  $M_j, j \in [1 : b - 2]$ , are selected independently and uniformly over  $[1 : 2^{nR}]$  and are separately encoded and transmitted over  $b$  blocks. The average rate  $R \frac{b-2}{b}$  tends to  $R$  as  $b$  tends to infinity.

*Codebook generation.* Fix the probability distribution

$$p(x, x_1, x_2, y_1, y_2, \hat{y}_2) = p(x, x_1)p(x_2)p(y_1, y_2|x, x_1, x_2)p(\hat{y}_2|x_2, y_2). \quad (41)$$

Generate at random an independent codebook for each block (only three such independent codebooks used for every consecutive triple-block are required, so that joint decoding over any three adjacent blocks result in independent error events). For  $j \in [1 : b]$ , randomly and independently generate  $2^{nR}$  sequences  $x_1^n(m_{j-2}), m_{j-2} \in [1 : 2^{nR}]$ , each according to  $\prod_{i=1}^n p_{X_1}(x_{1i})$ . For each  $m_{j-2} \in [1 : 2^{nR}]$ , randomly and conditionally independently generate  $2^{nR}$  sequences  $x^n(m_j|m_{j-2}), m_j \in [1 : 2^{nR}]$ , each according to  $\prod_{i=1}^n p_{X|X_1}(x_i|x_{1i}(m_{j-2}))$ . Randomly and independently

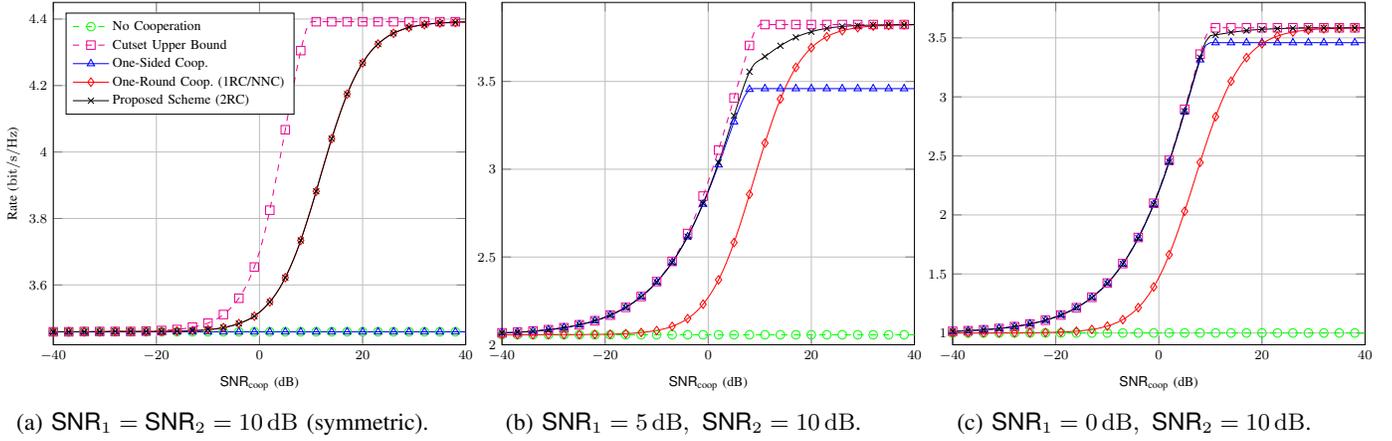


Fig. 3: Comparison of the upper and lower bounds on the achievable rate for the Gaussian BC with receiver cooperation.

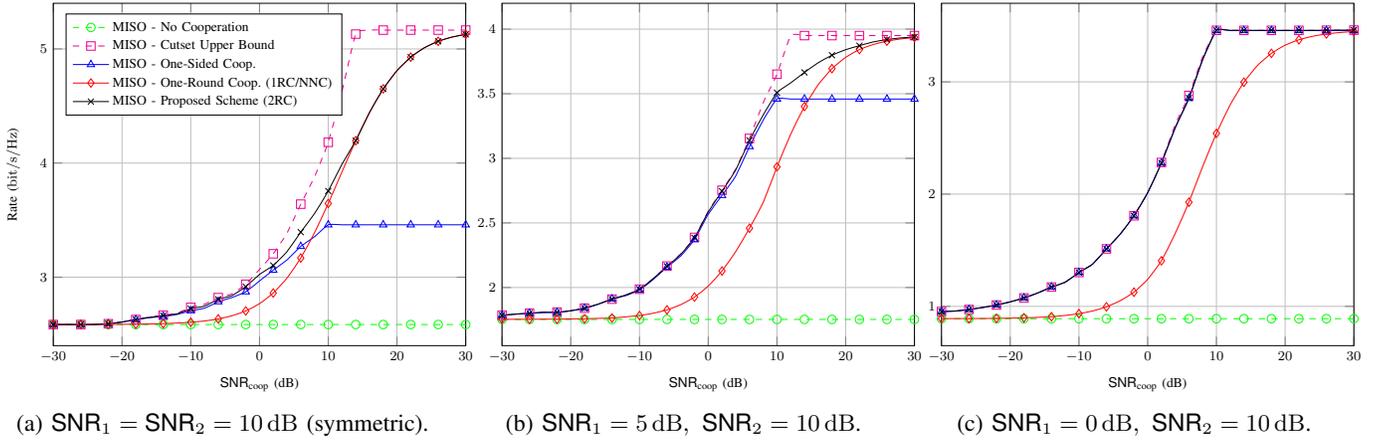


Fig. 4: Comparison of the upper and lower bounds on the achievable rate for the MISO  $2 \times 1$  Gaussian BC with receiver cooperation.

generate  $2^{nR_2}$  sequences  $x_2^n(l_{j-1}), l_{j-1} \in [1 : 2^{nR_2}]$ , each according to  $\prod_{i=1}^n p_{X_2}(x_{2i})$ . For each  $l_{j-1} \in [1 : 2^{nR_2}]$ , randomly and conditionally independently generate  $2^{n\hat{R}_2}$  sequences  $\hat{y}_2^n(k_j|l_{j-1}), k_j \in [1 : 2^{n\hat{R}_2}]$ , each according to  $\prod_{i=1}^n p_{\hat{Y}_2|X_2}(\hat{y}_{2i}|x_{2i}(l_{j-1}))$ .

The codebooks are defined as,

$$\mathcal{C}_j = \left\{ \left( x^n(m_j|m_{j-2}), x_1^n(m_{j-2}), x_2^n(l_{j-1}), \hat{y}_2^n(k_j|l_{j-1}) \right) \middle| m_j, m_{j-2} \in [1 : 2^{nR}], l_{j-1} \in [1 : 2^{nR_2}], k_j \in [1 : 2^{n\hat{R}_2}] \right\}, \quad (42)$$

for  $j \in [1 : b]$ . Partition the set  $[1 : 2^{n\hat{R}_2}]$  into  $2^{nR_2}$  equal size bins  $\mathcal{B}(l_j) = [(l_j - 1)2^{n(\hat{R}_2 - R_2)} + 1 : l_j 2^{n(\hat{R}_2 - R_2)}], l_j \in [1 : 2^{nR_2}], \hat{R}_2 \geq R_2$ . The codebooks and the bin assignments are revealed to all parties.

*Encoding.* Let  $m_j \in [1 : 2^{nR}]$  be the message to be sent over the block  $j$ . The encoder transmits  $x^n(m_j|m_{j-2})$  from the codebook  $\mathcal{C}_j$ , where  $m_{-1} = m_0 = m_{b-1} = m_b = 1$  by convention.

*Relay encoding at receiver 2.* Let  $l_0 = l_{b-1} = 1$  by convention. At the end of block  $j$ , the relay receiver 2 finds an index  $k_j$  s.t.  $(y_2^n(j), \hat{y}_2^n(k_j|l_{j-1}), x_2^n(l_{j-1})) \in \mathcal{T}_\epsilon^{(n)}$ . If there is more than one such index, it selects one of them uniformly at random. If there is no such index, it selects an index from  $[1 : 2^{n\hat{R}_2}]$  uniformly at random. In block  $j + 1$  the relay receiver 2 transmits  $x_2^n(l_j)$  from codebook  $\mathcal{C}_{j+1}$ , where  $k_j \in \mathcal{B}(l_j)$ .

*Decoding at receiver 1.* Let  $\epsilon > \epsilon'$ . At the end of block  $j + 1$ , the decoder receiver 1 finds the unique index  $\hat{l}_j$  s.t.  $(x_2^n(\hat{l}_j), x_1^n(\hat{m}'_{j-1}), y_1^n(j + 1)) \in \mathcal{T}_\epsilon^{(n)}$ . It then finds the unique message  $\hat{m}'_j$  s.t.,  $(x^n(\hat{m}'_j|\hat{m}'_{j-2}), x_2^n(\hat{l}_{j-1}), \hat{y}_2^n(\hat{k}_j|\hat{l}_{j-1}), x_1^n(\hat{m}'_{j-2}), y_1^n(j)) \in \mathcal{T}_\epsilon^{(n)}$ , for some  $\hat{k}_j \in \mathcal{B}(\hat{l}_j)$ .

*Relay encoding at receiver 1.* Let  $\hat{m}_{-1} = \hat{m}_0 = 1$  by convention. In block  $j + 2$  the relay receiver 1 transmits  $x_1^n(\hat{m}'_j)$  from the codebook  $\mathcal{C}_{j+2}$ .

*Sliding window decoding at receiver 2.* At the end of block  $j + 2$ , the decoder receiver 2 finds the unique message  $\hat{m}''_j$  s.t.  $(x^n(\hat{m}''_j|\hat{m}''_{j-2}), x_1^n(\hat{m}''_{j-2}), x_2^n(l_{j-1}), y_2^n(j)) \in \mathcal{T}_\epsilon^{(n)}$

and  $(x_1^n(\hat{m}_j''), x_2^n(l_{j+1}), y_2^n(j+2)) \in \mathcal{T}_\epsilon^{(n)}$  simultaneously.

*Analysis of the probability of error at receiver 1 and 2.* The probability of decoding error is analyzed at the decoder receiver 1 and 2, for the message  $M_j$  averaged over codebooks. Assume without loss of generality that  $M_{j-2} = M_j = 1$  and let  $L_{j-1}, L_j, K_j$  denote the indices chosen by the relay receiver 2 in block  $j$ . Then, the decoder receiver 1 makes an error only if one or more of the following events occur,

$$\mathcal{E}_{(1)2 \rightarrow 1 \rightarrow 2}(j-2) = \left\{ \hat{M}'_{j-2} \neq 1 \right\} \quad (43)$$

$$\hat{\mathcal{E}}_{(2)}(j) = \left\{ (Y_2^n(j), \hat{Y}_2^n(k_j | L_{j-1}), X_2^n(L_{j-1})) \notin \mathcal{T}_\epsilon^{(n)}, \forall k_j \in [1 : 2^{\hat{R}_2}] \right\} \quad (44)$$

$$\mathcal{E}_{(1)1}(j) = \left\{ \hat{L}_j \neq L_j \right\} \text{ and } \mathcal{E}_{(1)1}(j-1) \quad (45)$$

$$\mathcal{E}_{(1)2}(j) = \left\{ (X^n(1 | \hat{M}'_{j-2}), Y_1^n(j), \hat{Y}_2^n(K_j | \hat{L}_{j-1}), X_1^n(\hat{M}'_{j-2}), X_2^n(\hat{L}_{j-1})) \notin \mathcal{T}_\epsilon^{(n)} \right\} \quad (46)$$

$$\mathcal{E}_{(1)3}(j) = \left\{ (X^n(m'_j | \hat{M}'_{j-2}), Y_1^n(j), \hat{Y}_2^n(K_j | \hat{L}_{j-1}), X_1^n(\hat{M}'_{j-2}), X_2^n(\hat{L}_{j-1})) \in \mathcal{T}_\epsilon^{(n)} \text{ for some } m'_j \neq 1 \right\} \quad (47)$$

$$\mathcal{E}_{(1)4}(j) = \left\{ (X^n(m'_j | \hat{M}'_{j-2}), X_1^n(\hat{M}'_{j-2}), Y_1^n(j), \hat{Y}_2^n(\hat{k}_j | \hat{L}_{j-1}), X_2^n(\hat{L}_{j-1})) \in \mathcal{T}_\epsilon^{(n)} \text{ for some } \hat{k}_j \in \mathcal{B}(\hat{L}_j), \hat{k}_j \neq K_j, m'_j \neq 1 \right\}. \quad (48)$$

Applying the law of large numbers, the conditional typicality lemma [15, Sec. 2.5], the joint typicality lemma [15, Sec. 2.5.1], the packing lemma [15, Lem. 3.1, Sec 3.2], the covering lemma [15, Lem. 3.3, Sec 3.7], the lemma 11.1 [15, Sec. 11.3.1], the chain rule, and the Fourier-Motzkin elimination procedure [15, Appx. D] gives the bounds

$$R \leq I(X; Y_1, \hat{Y}_2 | X_1, X_2) - \delta(\epsilon) \quad (49)$$

$$R \leq I(X, X_2; Y_1 | X_1) - I(Y_2; \hat{Y}_2 | X, X_1, X_2, Y_1) - (2\delta(\epsilon) + \delta(\epsilon')). \quad (50)$$

Assume without loss of generality that  $M_{j-2} = M_j = 1$ . Then, the decoder receiver 2 makes an error only if one or

more of the following events occur,

$$\mathcal{E}_{(1)2 \rightarrow 1 \rightarrow 2}(j-2) = \left\{ \hat{M}'_{j-2} \neq 1 \right\} \text{ and } \mathcal{E}_{(1)2 \rightarrow 1 \rightarrow 2}(j) \quad (51)$$

$$\mathcal{E}_{(2)2 \rightarrow 1 \rightarrow 2}(j-2) = \left\{ \hat{M}''_{j-2} \neq 1 \right\} \quad (52)$$

$$\mathcal{E}_{(2)1}(j) = \left\{ (X^n(\hat{M}_j'' | \hat{M}''_{j-2}), Y_2^n(j), X_1^n(\hat{M}''_{j-2}), X_2^n(L_{j-1})) \notin \mathcal{T}_\epsilon^{(n)} \text{ or } (X_1^n(\hat{M}_j''), Y_2^n(j+2), X_2^n(L_{j+1})) \notin \mathcal{T}_\epsilon^{(n)} \right\} \quad (53)$$

$$\mathcal{E}_{(2)2}(j) = \left\{ (X^n(m_j'' | \hat{M}''_{j-2}), Y_2^n(j), X_1^n(\hat{M}''_{j-2}), X_2^n(L_{j-1})) \in \mathcal{T}_\epsilon^{(n)} \text{ and } (X_1^n(m_j''), Y_2^n(j+2), X_2^n(L_{j+1})) \in \mathcal{T}_\epsilon^{(n)} \text{ for some } m_j'' \neq \hat{M}_j' \right\}. \quad (54)$$

Applying the law of large numbers, the joint typicality lemma and the chain rule gives the bound

$$R \leq I(X, X_1; Y_2 | X_2) - 2\delta(\epsilon). \quad (55)$$

Combining the bounds (49), (50), and (55), taking the limit over  $n$  and maximizing over the two strategies gives the result in Prop. 1.

## REFERENCES

- [1] T. Cover, "Broadcast channels," *IEEE Trans. Inf. Theory*, vol. 18, no. 1, pp. 2–14, Jan. 1972.
- [2] P. Bergmans, "Random coding theorem for broadcast channels with degraded components," *IEEE Trans. Inf. Theory*, vol. 19, no. 2, pp. 197–207, Mar. 1973.
- [3] K. Marton, "A coding theorem for the discrete memoryless broadcast channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 3, pp. 306–311, May 1979.
- [4] E. van der Meulen, "Three-terminal communication channels," *Adv. Appl. Probab.*, vol. 3, no. 1, pp. 120–154, Spring 1971.
- [5] T. Cover and A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 5, pp. 572–584, Sep. 1979.
- [6] 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.
- [7] S. H. Lim, Y.-H. Kim, A. El Gamal, and S.-Y. Chung, "Noisy network coding," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 3132–3152, May 2011.
- [8] J. Hou and G. Kramer, "Short message noisy network coding with a decode-forward option," *IEEE Trans. Inf. Theory*, vol. 62, no. 1, pp. 89–107, Jan. 2016.
- [9] A. Behboodi and P. Piantanida, "Cooperative strategies for simultaneous and broadcast relay channels," *IEEE Trans. Inf. Theory*, vol. 59, no. 3, pp. 1417–1443, Mar. 2013.
- [10] R. Dabora and S. Servetto, "Broadcast channels with cooperating decoders," *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5438–5454, Dec. 2006.
- [11] Y. Liang and V. Veeravalli, "Cooperative relay broadcast channels," *IEEE Trans. Inf. Theory*, vol. 53, no. 3, pp. 900–928, Mar. 2007.
- [12] Y. Liang and G. Kramer, "Rate regions for relay broadcast channels," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3517–3535, Oct. 2007.
- [13] S. H. Lim, K. T. Kim, and Y.-H. Kim, "Distributed decode-forward for multicast," in *Proc. IEEE Int. Symp. Inf. Theory*, 2014, pp. 636–640.
- [14] T. Cover and J. Thomas, *Elements of Information Theory, Second Edition*. Wiley-Interscience, 2006.
- [15] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge Univ. Press, 2012.