

Message and State Cooperation in Multiple Access Channels

Haim H. Permuter, *Member, IEEE*, Shlom (Shitz) Shamai, *Fellow, IEEE*, and Anelia Somekh-Baruch, *Member, IEEE*

Abstract—We investigate the capacity of a multiple access channel with cooperating encoders where partial state information is known to each encoder and full state information is known to the decoder. The cooperation between the encoders has a two-fold purpose: to generate empirical state coordination between the encoders, and to share information about the private messages that each encoder has. For two-way cooperation, this two-fold purpose is achieved by double-binning, where the first layer of binning is used to generate the state coordination similarly to the two-way source coding, and the second layer of binning is used to transmit information about the private messages. The complete result provides the framework and perspective for addressing a complex level of cooperation that mixes states and messages in an optimal way.

Index Terms—Channel state information, cooperating encoders, coordination, double-binning, message-state cooperation, multiple access channel, superbin.

I. INTRODUCTION

STATE-DEPENDENT channels describe a rich variety of communication models spanning the cases where the states are governed by physical phenomena (such as fading), and accounting also for situations where the states model effects of interfering transmissions. Their wide applicability, theoretical importance, and practical implications, led to intensive information theoretic studies. We focus here on a multiple-access channel (MAC), where the channel is affected by the state (S_1, S_2) known partly at the transmitters. That is, state S_1 is available at Transmitter 1, while S_2 is known at Transmitter 2. This can be associated with local cognition, that is, Transmitter 1 learns beforehand about the sequence S_1 , while Transmitter 2 learns about S_2 . We further assume that the states, which can be viewed as channel-affecting parameters, are known at the receiving point, or alternatively are retrieved accurately by the receiver. This is a standard problem, which falls within

the class of decentralized processing at the transmitters. The focus of this work is the implications of transmitter cooperation facilitated by an orthogonal finite capacity link. This link can be used both to share state information, as to facilitate a more coordinated operation, up to a degree of central coordination, achieved when both transmitters know accurately (S_1, S_2) . The cooperation link can also be employed to share messages, to the extreme of full message cooperation, turning the problem into a single two-elements (antennas) transmitter. The interplay among these types of cooperation is at the center of our paper, and here the optimal approach, given in terms of the associated capacity region, is found. Evidently the derivation of this general result is extending previous important cases as it is detailed in the following.

Willems [1], [2] introduced and derived the capacity region of the multiple access channel (MAC) with cooperating encoders. He showed that to achieve the capacity region the encoders should use the cooperation link in order to share parts of their private messages and then use a coding scheme for the ordinary MAC, which was found earlier by Slepian and Wolf [3]. Willems' model allows interactive communication between the encoders, however Willems showed [1], [2] that the optimality is achieved even in a single round of communication between the encoders.

In this paper, we consider the problem of MAC with cooperating encoder, where different partial state information is known at each encoder and perfect state information is known at the decoder. The setting of the problem is depicted in Fig. 1. The state of the channel is given by the pair (S_1, S_2) , where Encoder 1 knows S_1 , Encoder 2 knows S_2 , and the decoder knows the pair (S_1, S_2) . The cooperation links C_{12} and C_{21} may increase the capacity region by transmission of the state information that is missing to the encoders and by sharing parts of the private messages (m_1, m_2) . Here the transmission of the state information is done by achieving an empirical coordination [4] of the state information, namely, generating sequences of action that are functions of the cooperation and are jointly typical with the state information. Simultaneously, these sequences of action are designed in such a way that they allow the encoders to share parts of their private messages. We illustrate the idea of simultaneity generating coordination and sharing a message by a simple toy-problem that is presented in Section II. In the two way cooperation we use double-binning to generate coordination and share message. The double-binning is a technique that was used previously by Liu *et al.* [5], [6] for achieving secrecy capacity in the broadcast channel.

The problem of cooperating encoders with partial state information combines two kinds of settings that are widely treated in the literature; the first is limited-rate noise-free cooperation

Manuscript received June 10, 2010; revised March 07, 2011; accepted April 12, 2011. Date of current version October 07, 2011. This work was supported by the CORNET Consortium sponsored by the Chief Scientist of the Israel Ministry for Industry and Commerce.

H. H. Permuter is with the Department of Electrical and Computer Engineering, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel (e-mail: haimp@bgu.ac.il).

S. S. Shamai is with the Department of Electrical Engineering, Technion-Israel Institute of Technology, Haifa 32000, Israel (e-mail: sshlomo@ee.technion.ac.il).

A. Somekh-Baruch is with the School of Engineering, Bar-Ilan University, Ramat-Gan 52900, Israel (e-mail: anelia.somekhbaruch@gmail.com).

Communicated by E. Erkip, Associate Editor for Shannon Theory.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIT.2011.2165804

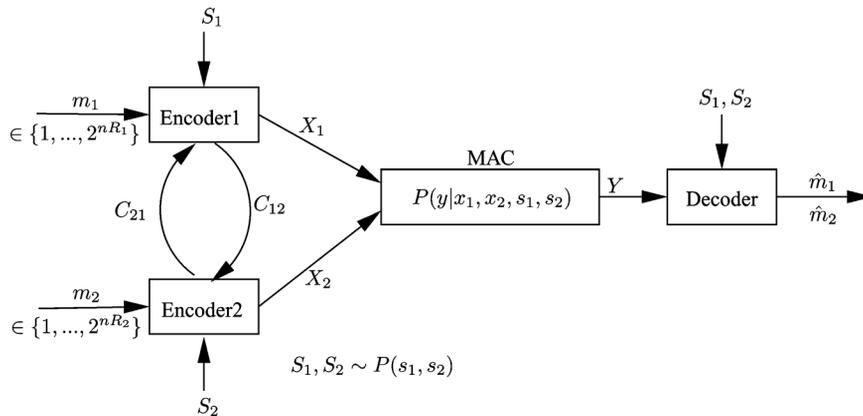


Fig. 1. MAC with cooperation where different partial state information is known to each encoder, and full state information is known to the decoder.

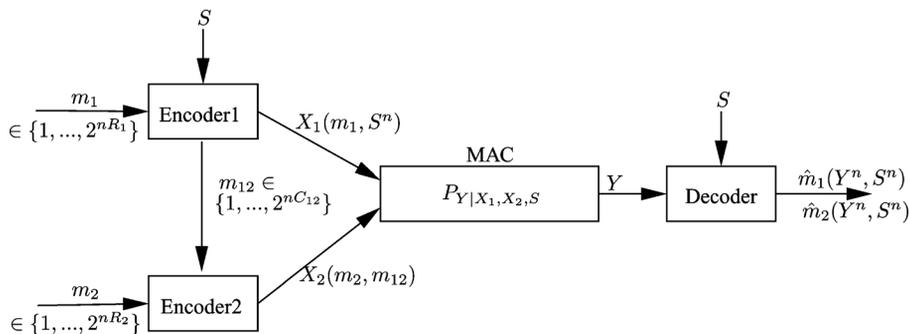


Fig. 2. MAC with one-way conferencing and state information at one encoder and the decoder.

between users and the second is limited-rate noise-free state information that is available to encoders/decoders.

Cooperation between users through a noise-free limited-rate link has been investigated in various of multi-user settings such as in MAC [1], [2], [7], [8], interference channel [9]–[15], broadcast channel [16], relay channels [17]–[19], and cellular networks [20]. A comprehensive survey of cooperation and its role in communication is given in [21]. Recently, cooperation between encoders where state information is available was considered in [22], [23] where it is assumed that the cooperation is allowed only before the state information is available at the encoders. In this paper, we take a different approach, assuming that the cooperation occurs after the state information becomes available, the cooperation may include parts of the private message and the state information as well.

The second setting, that is, limited-rate state information at encoders/decoders, was first treated by Heegard and El-Gamal [24]. The case, most related to the setting in this paper, where full state information is available at the decoder and limited-rate state information is known at the encoder was solved by Cemal and Steinberg for the point-to-point channel [25] and for the MAC [26]. The main difference between the setting here and the setting in [26] is that here the limited-rate encoder knows the state and the private message rather than just the state; therefore, a scheme which combines message information and state information is needed.

The remainder of the paper is organized as follows. In Section II, we derive the capacity region where only one coop-

eration link from Encoder 1 to Encoder 2 exists. This setting helps us to gain the intuition necessary for solving the extended problem of two-way cooperation, which is solved in Section III. In Section IV, we solve a specific example and compare the capacity region to two different cooperation settings given in [22] and in [26]. In addition, in Section IV, we check the strategy of splitting the cooperation link into message-only link and state-only link, and we show that this naive strategy is strictly suboptimal.

II. ONE-WAY COOPERATION

In this section, we consider a special case, in which there is only one-way cooperation from Encoder 1 to Encoder 2. In addition, we assume that Encoder 1 and the decoder have full non-causal state information. This setting captures the idea of, simultaneously, sharing a part of the private message m_1 and sharing the information on channel state S . The setting is depicted in Fig. 2. We start by defining the notation and the code for this setting, then we state the capacity region, explain the intuition and provide its proof.

The MAC setting consists of two transmitters (encoders) and one receiver (decoder). Each sender $l \in \{1, 2\}$ chooses an index m_l uniformly from the set $\{1, \dots, 2^{nR_l}\}$ and independently of the other sender. The input to the channel from encoder $l \in \{1, 2\}$ is denoted by $\{X_{l,1}, X_{l,2}, X_{l,3}, \dots\}$, and the output of the channel is denoted by $\{Y_1, Y_2, Y_3, \dots\}$. The state at time i , i.e., $S_i \in \mathcal{S}$, takes values in a finite set of possible states \mathcal{S} . The channel is characterized by a conditional prob-

ability $P(y_i|x_{1,i}, x_{2,i}, s_i)$ and by the state probability $P(s_i)$. Both probabilities do not depend on the time index i and satisfy

$$P(y_i, s_{i+1}|x_1^i, x_2^i, s^i, y^{i-1}) = P(y_i|x_{1,i}, x_{2,i}, s_i)P(s_{i+1}), \quad (1)$$

where the superscripts denote sequences in the following way: $x_l^i = (x_{l,1}, x_{l,2}, \dots, x_{l,i})$, $l \in \{1, 2\}$.

Definition 1: A $(2^{nR_1}, 2^{nR_2}, 2^{nC_{12}}, n)$ code with one-way cooperating encoder as shown in Fig. 2 consists of three encoding functions

$$\begin{aligned} f_1 &: \{1, \dots, 2^{nR_1}\} \times \mathcal{S}^n \mapsto \mathcal{X}_1^n, \\ f_{12} &: \{1, \dots, 2^{nR_1}\} \times \mathcal{S}^n \mapsto \{1, \dots, 2^{nC_{12}}\}, \\ f_2 &: \{1, \dots, 2^{nR_2}\} \times \{1, \dots, 2^{nC_{12}}\} \mapsto \mathcal{X}_2^n, \end{aligned} \quad (2)$$

and a decoding function,

$$g : \mathcal{Y}^n \times \mathcal{S}^n \mapsto \{1, \dots, 2^{nR_1}\} \times \{1, \dots, 2^{nR_2}\}. \quad (3)$$

The average probability of error for $(2^{nR_1}, 2^{nR_2}, 2^{nC_{12}}, n)$ code is defined in (4) at the bottom of the page. A rate (R_1, R_2) is said to be *achievable* for the one-way cooperating MAC with cooperation link C_{12} , if there exists a sequence of $(2^{nR_1}, 2^{nR_2}, 2^{nC_{12}}, n)$ codes with $P_e^{(n)} \rightarrow 0$. The *capacity region* of MAC is the closure of all achievable rates. The following theorem describes the capacity region of one-way cooperating MAC.

Theorem 1: The capacity region of the MAC with a cooperating encoder that has state information as shown in Fig. 2 is the closure of the set that contains all rates that satisfy

$$C_{12} \geq I(U; S) \quad (5)$$

$$R_1 \leq I(X_1; Y|X_2, S, U) + C_{12} - I(U; S) \quad (6)$$

$$R_2 \leq I(X_2; Y|X_1, S, U) \quad (7)$$

$$R_1 + R_2 \leq \min \left\{ \begin{array}{l} I(X_1, X_2; Y|S, U) + C_{12} - I(U; S), \\ I(X_1, X_2; Y|S) \end{array} \right\}, \quad (8)$$

for some joint distribution of the form

$$P(s)P(u, x_1|s)P(x_2|u)P(y|x_1, x_2, s). \quad (9)$$

Lemma 2:

- 1) The capacity region described in Theorem 1, given in (5)–(9), is convex.
- 2) It is enough to restrict the alphabet of the auxiliary random variable U in Theorem 1 to satisfy

$$|\mathcal{U}| \leq \min(|\mathcal{X}_1||\mathcal{X}_2||\mathcal{S}| + 3, |\mathcal{Y}||\mathcal{S}| + 4). \quad (10)$$

Before proving the theorem and the lemma let us investigate the role of the auxiliary random variable U in Theorem 1. The random variable U plays a double role: first, it generates an empirical coordination between the two encoders regarding the state of the channel, i.e., with high probability the sequence U^n would be jointly typical with S^n ; second, it generates a common message between the two encoders. Let us look at two special cases which emphasize these two roles.

Case 1: The Point-to-Point Case [25], i.e., $R_1 = 0$ and $P(y|x_1, x_2, s) = P(y|x_2, s)$: For this case the rate region of Theorem 1 becomes

$$C_{12} \geq I(U; S) \quad (11)$$

$$R_2 \leq I(X_2; Y|S, U) \quad (12)$$

$$R_2 \leq \min \left\{ \begin{array}{l} I(X_2; Y|S, U) + C_{12} - I(U; S) \\ I(X_2; Y|S) \end{array} \right\}, \quad (13)$$

which is simply

$$C_{12} \geq I(U; S) \quad (14)$$

$$R_2 \leq I(X_2; Y|S, U) \quad (15)$$

for a joint distribution of the form

$$P(s)P(u|s)P(x_2|u)P(y|x_2, s).$$

Case 2: $|\mathcal{S}| = 1$, the Memoryless Case [2]: In this case $I(U; S) = 0$, hence we obtain a special case of MAC with cooperation and the rate region of Theorem 1 becomes

$$R_1 \leq I(X_1; Y|X_2, U) + C_{12} \quad (16)$$

$$R_2 \leq I(X_2; Y|X_1, U) \quad (17)$$

$$R_1 + R_2 \leq \min \left\{ \begin{array}{l} I(X_1, X_2; Y|U) + C_{12} \\ I(X_1, X_2; Y) \end{array} \right\}, \quad (18)$$

for a joint distribution of the form

$$P(u)P(x_1|u)P(x_2|u)P(y|x_2, x_1).$$

$$P_e^{(n)} = \frac{1}{2^{n(R_1+R_2)}} \sum_{m_1, m_2} \Pr\{g(Y^n, S^n) \neq (m_1, m_2) | (m_1, m_2) \text{ sent}\}. \quad (4)$$

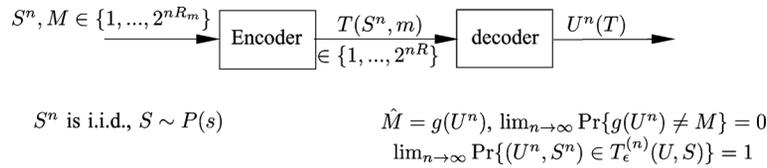


Fig. 3. A problem that illustrates the double role of cooperation. On one hand, the sequence U^n needs to be jointly typical with S^n , and on the other hand, one should be able to reconstruct the message m with high probability.

Note that in the first case the role of the auxiliary random variable U is to generate an empirical coordination $P_{U|S}$, and then use the sequence U^n as common side information at the encoder and decoder. In the second case, the auxiliary random variable represents the common message m_0 between the two encoders, and the decoder needs to decode it. In Theorem 1, these two roles are combined. Namely, the sequence U^n needs to be coordinated with S^n and simultaneously represents a common message. Fig. 3 illustrates the role of cooperation. The cooperation is generated through a transmission of a limited-rate message $T \in \{1, 2, \dots, 2^{nR}\}$ which is a function of another message $M \in \{1, 2, \dots, 2^{nR_m}\}$ and of a sequence S^n . On one hand, the cooperation needs to generate a sequence U^n that is jointly typical with S^n , i.e., $\lim_{n \rightarrow \infty} \Pr\{(U^n, S^n) \in T_\epsilon^{(n)}(U, S)\} = 1$, and on the other hand, there should be a function $g(U^n)$ such that one can estimate the message M with high probability, i.e., $\lim_{n \rightarrow \infty} \Pr\{g(U^n) \neq M\} = 0$. If $R > R_m + I(U; S)$ and $H(U|S) \geq R_m$, this goal can be achieved. Combining these two roles (generating empirical coordination and transmitting a message) is done by binning,¹ where the bin number represents the common message and in each bin there will be enough codewords U^n such that at least one codeword is jointly typical with S^n . This is similar to the role of the auxiliary random variable in Gelfand-Pinsker [28], where the sequence of the auxiliary random variables that is generated needs to represent a message that is transmitted via the channel and needs to be jointly typical with the sequence of the channel states.

We would like to mention two comments about the importance of using the binning technique in the achievability proof. First, a naive scheme that splits the rate into two parts, one for generating coordination with the state sequence S^n and one for sharing part of the message, is not optimal as it is shown in an example in Section IV. Second, we should note that since the auxiliary random variable U^n is a function of the sequence S^n and the message, the receiver of the MAC which knows S^n does not know U^n but needs to decode it, which is different from Case 1 (the point-to-point case [25]), discussed previously, where the auxiliary random variable U^n is a function of only the sequence S^n .

Next we present a formal proof of Theorem 1. Throughout the achievability proofs in the paper we use the definition of a strong typical set. The set $T_\epsilon^{(n)}(X, Y, Z)$ of ϵ -typical n -sequences is defined by $\{(x^n, y^n, z^n) : \frac{1}{n} N(x, y, z | x^n, y^n, z^n) - p(x, y, z) \leq \epsilon p(x, y, z) \forall (x, y, z) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}\}$, where $N(x, y, z | x^n, y^n, z^n)$ is the number of appearances of (x, y, z) in the n -sequence (x^n, y^n, z^n) . Fur-

¹In channel coding problems, the collection of codewords in a bin, are also called a subcodebook [26] and whole codebook is called multicoding.

thermore, we will use the following well-known lemma [27], [29]–[31],

Lemma 3 (Joint Typicality Lemma): Consider a joint distribution $P_{X,Y,Z}$ and suppose $(x^n, y^n) \in T_\epsilon^{(n)}(X, Y)$. Let \tilde{Z}^n be distributed according to $\prod_{i=1}^n P_{Z|X}(\tilde{z}_i | x_i)$. Then,

$$\Pr\{(x^n, y^n, \tilde{Z}^n) \in T_\epsilon^{(n)}(X, Y, Z)\} \leq 2^{-n(I(Y;Z|X) - \delta(\epsilon))}, \quad (19)$$

where $\lim_{\epsilon \rightarrow 0} \delta(\epsilon) = 0$.

Proof of Theorem 1: Achievability Part: Fix a joint distribution $P(s)P(u, x_1|s)P(x_2|u)P(y|x_1, x_2, s)$ where $P(s)$ and $P(y|x_1, x_2, s)$ are given by the channel.

Code Construction: Generate $2^{nC_{12}}$ codewords U^n independently using i.i.d. $\sim P(u)$, and assign them into $2^{n(C_{12} - I(U;S) - \epsilon)}$ bins. Hence, in each bin there are $2^{n(I(U;S) + \epsilon)}$ codewords. The role of the auxiliary random variable here is the same as in the example depicted in Fig. 3, namely to transmit a message at rate $C_{12} - I(U;S) - \epsilon$ that is associated with a bin number and generate coordination via the fact that in each bin there are $2^{n(I(U;S) + \epsilon)}$ codewords. For each codeword $u^n(j)$, where $j = 1, 2, \dots, 2^{nC_{12}}$ and for each $s^n \in \mathcal{S}^n$ generate $2^{n(R_1 - (C_{12} - I(U;S) - \epsilon))}$ codewords X_1^n according to i.i.d. $\sim P(x_1|u, s)$ and for each $u^n(j)$, where $j = 1, 2, \dots, 2^{nC_{12}}$, generate $2^{n(R_2 - (C_{12} - I(U;S)))}$ codewords X_2^n according to i.i.d. $\sim P(x_2|u)$.

Encoder: Split message $m_1 \in [1, \dots, 2^{nR_1}]$ into two messages $m_{1,a} \in [1, \dots, 2^{n(C_{12} - I(U;S) - \epsilon)}]$ and $m_{1,b} \in [1, \dots, 2^{n(R_1 - (C_{12} - I(U;S) - \epsilon))}]$. Now, associate each message $m_{1,a} \in [1, \dots, 2^{n(C_{12} - I(U;S) - \epsilon)}]$ with a bin, where in each bin there are $2^{n(I(U;S) + \epsilon)}$ codewords u^n , indexed by $l \in [1, \dots, 2^{n(I(U;S) + \epsilon)}]$. Find in the chosen bin a codeword, denoted by $u^n(m_{1,a}, s^n)$, with the smallest lexicographical order that is jointly typical with s^n and send its index $[1, \dots, 2^{nC_{12}}]$ to Encoder 2. If such a codeword u^n does not exist, namely, among the codewords in the bin none is jointly typical with s^n , choose an arbitrary u^n from the bin (in such a case the decoder will declare an error). Now, Encoder 1 transmits $x_1^n(s^n, u^n(m_{1,a}, s^n), m_{1,b})$, and Encoder 2 transmits $x_2^n(u^n(m_{1,a}, s^n), m_2)$.

Decoder: The decoder knows s^n and y^n and looks for the indices $\hat{m}_{1,a} \in [1, \dots, 2^{n(C_{12} - I(U;S) - \epsilon)}]$, $\hat{m}_{1,b} \in [1, \dots, 2^{n(R_1 - (C_{12} - I(U;S) - \epsilon))}]$, $\hat{m}_2 \in [1, \dots, 2^{nR_2}]$ such that (see (20) at the bottom of the next page), where throughout the proof, the joint distribution that specifies the typical set $T_\epsilon^{(n)}(U, X_1, X_2, S, Y)$ is the one we have fixed at the beginning of the proof, i.e., $P(s)P(u, x_1|s)P(x_2|u)P(y|x_1, x_2, s)$. If none or more than one such triplet is found, an error is declared. The estimated message sent from Encoder 1 is $(\hat{m}_{1,a}, \hat{m}_{1,b})$, and the estimated message transmitted from Encoder 2 is \hat{m}_2 .

Error Analysis: Assume $(m_{1,a}, m_{1,b}, m_2) = (1, 1, 1)$. Let us define the event as shown in (21) at the bottom of the page. An error occurs if either the correct codewords are not jointly typical with the received sequences, i.e., $E_{1,1,1}^c$, or there exists a different $(i, j, k) \neq (1, 1, 1)$ such that $E_{i,j,k}$ occurs. From the union of bounds we obtain that

$$\begin{aligned} P_e^{(n)} &\leq \Pr(E_{1,1,1}^c) + \sum_{i=1, j=1, k>1} \Pr(E_{i,j,k}) \\ &+ \sum_{i=1, j>1, k=1} \Pr(E_{i,j,k}) + \sum_{i=1, j>1, k>1} \Pr(E_{i,j,k}) \\ &+ \sum_{i>1, j\geq 1, k\geq 1} \Pr(E_{i,j,k}). \end{aligned} \quad (22)$$

Now let us show that each term in (22) goes to zero as the block-length of the code n goes to infinity.

- Upper-bounding $\Pr(E_{1,1,1}^c)$: Since the number of codewords in each bin is larger than $2^{nI(U;S)}$, and since the codewords were generated i.i.d., with high probability there will be at least one codeword that is jointly typical with s^n . We denote this sequence as $u^n(1, s^n)$. Furthermore, given that $(u^n(1, s^n), s^n) \in T_\epsilon^{(n)}(U, S)$, it follows from the law of large numbers that $\Pr(E_{1,1,1}^c) \rightarrow 0$ as n goes to infinity.
- Upper-bounding $\sum_{i=1, j=1, k>1} \Pr(E_{i,j,k})$: The probability that Y^n , which is generated according to $P(y|x_1, s, u)$, is jointly typical with x_2^n , which was generated according to $P(x_2|u) = P(x_2|u, s, x_1)$, where $(x_1^n, s^n, u^n) \in T_\epsilon^{(n)}(X_1, S, U)$ is bounded by (Lemma 3)

$$\Pr\{(x_1^n, X_2^n, u^n, s^n, Y^n) \in T_\epsilon^{(n)} | (x_1^n, u^n, s^n) \in T_\epsilon^{(n)}\} \leq 2^{-n(I(X_2;Y|X_1, S, U) - \delta(\epsilon))}. \quad (23)$$

Hence, we obtain

$$\sum_{i=1, j=1, k>1} \Pr(E_{i,j,k}) \leq 2^{nR_2} 2^{-n(I(X_2;Y|X_1, S, U) - \delta(\epsilon))} \quad (24)$$

- Upper-bounding $\sum_{i=1, j>1, k=1} \Pr(E_{i,j,k})$: The probability that Y^n which is generated according to $P(y|x_2, s, u)$ is jointly typical with x_1^n which was generated according to $P(x_1|u, s) = P(x_1|u, s, x_2)$, where $(x_2^n, s^n, u^n) \in T_\epsilon^{(n)}(X_2, S, U)$ is upper bounded by $2^{-n(I(X_1;Y|X_2, S, U) - \delta(\epsilon))}$, hence

$$\sum_{i=1, j>1, k=1} \Pr(E_{i,j,k}) \leq 2^{n(R_1 - (C_{12} - I(U;S) - \epsilon))} 2^{-n(I(X_1;Y|X_2, S, U) - \delta(\epsilon))}. \quad (25)$$

- Upper-bounding $\sum_{i=1, j>1, k>1} \Pr(E_{i,j,k})$

$$\sum_{i=1, j>1, k>1} \Pr(E_{i,j,k}) \leq 2^{n(R_2 + R_1 - (C_{12} - I(U;S) - \epsilon))} 2^{-n(I(X_2, X_1;Y|S, U) - \delta(\epsilon))}. \quad (26)$$

- Upper-bounding $\sum_{i>1, j\geq 1, k\geq 1} \Pr(E_{i,j,k})$: see (27) at the bottom of the page.

Therefore, combining the upper bounds (24)–(27) into (22), we obtain that if rate-pair (R_1, R_2) is inside the rate region given by (5)–(9), then there exists a sequence of codes $(2^{nR_1}, 2^{nR_2}, 2^{nC_{12}}, n)$ such that $P_e^{(n)}$ goes to zero as $n \rightarrow \infty$.

Converse Part: The converse is deferred to Appendix A. ■

Proof of Lemma 2: First we prove that the capacity region described in Theorem 1, (5)–(9), is convex and therefore there is no need to convexify it. Let $P_i, i = 1, 2, 3$ be three distributions of the form

$$P(s)P(u, x_1|s)P(x_2|u)P(y|x_1, x_2, s), \quad (28)$$

which induce the quantities shown in (29) at the bottom of the next page. for $i = 1, 2, 3$, respectively. In addition, let $P_3 = \alpha P_1 + \bar{\alpha} P_2$, where $0 \leq \alpha \leq 1$ and $\bar{\alpha} = 1 - \alpha$, furthermore when $q = 1$ the distribution of U, X_1, X_2 is according to P_1 and when $q = 2$ it is according to P_2 . Let Q be a binary random variable with $P(q = 1) = \alpha$ and $P(q = 2) = 1 - \alpha$ independent of S . Let us denote $\tilde{U} = (U, Q)$, and note that

$$(u^n(\hat{m}_{1,a}, s^n), x_1^n(s^n, u^n(\hat{m}_{1,a}, s^n), \hat{m}_{1,b}), x_2^n(u^n(\hat{m}_{1,a}, s^n), \hat{m}_2), s^n, y^n) \in T_\epsilon^{(n)}(U, X_1, X_2, S, Y), \quad (20)$$

$$E_{i,j,k} \triangleq \left\{ (u^n(i, s^n), x_1^n(s^n, u^n(i, s^n), j), x_2^n(u^n(i, s^n), k), s^n, y^n) \in T_\epsilon^{(n)}(U, X_1, X_2, S, Y) \right\}. \quad (21)$$

$$\begin{aligned} \sum_{i>1, j\geq 1, k\geq 1} \Pr(E_{i,j,k}) &\leq 2^{n(C_{12} - I(U;S) - \epsilon)} 2^{n(R_1 - (C_{12} - I(U;S) - \epsilon))} 2^{nR_2} 2^{-n(I(X_2, X_1, U;Y|S) - \delta(\epsilon))} \\ &= 2^{n(R_1 + R_2 - I(X_2, X_1, U;Y|S) - \delta(\epsilon))} \end{aligned} \quad (27)$$

P_3 is of the form of (28) and is obtained when we marginalize $P(s)P(\tilde{u}, x_1|s, q)P(x_2|u)P(y|x_1, x_2, s)$ which equals to $P(s)P(q)P(u, x_1|s, q)P(x_2|u)P(y|x_1, x_2, s)$ over Q . Finally, the convexity of the region in (5)–(9) follows from the equalities $\alpha I_1(U; S) + \bar{\alpha} I_2(U; S) = I_3(\bar{U}; S)$, and similar equalities for the other terms in (29), and from the inequality

$$\begin{aligned} \alpha I_1(X_1, X_2; Y|S) + \bar{\alpha} I_2(X_1, X_2; Y|S) &= I_3(X_1, X_2; Y|S, Q) \\ &\leq I_3(X_1, X_2; Y|S). \end{aligned} \quad (30)$$

Now, to prove the cardinality bound on U , we invoke the support lemma [29, p. 310]. The auxiliary random variable U needs to have $|\mathcal{X}_1||\mathcal{X}_2||S| - 1$ letters to preserve $p(x_1, x_2, s)$ plus four more to preserve the expressions $H(S|U)$, $I(X_1; Y|X_2, S, U)$, $I_i(X_2; Y|X_1, S, U)$, and $I(X_1, X_2; Y|S, U)$. Note that the joint distribution $p(x_1, x_2, s, y)$ is preserved because of the Markov form $U - (X_1, X_2, S) - Y$. Alternatively, the external random variable U needs to have $|\mathcal{Y}||S| - 1$ letters to preserve $P(y, s)$ plus five more to preserve the expressions $H(S|U)$, $I(X_1; Y|X_2, S, U)$, $I_i(X_2; Y|X_1, S, U)$, $I(X_1, X_2; Y|S, U)$, and $H(Y|X_1, X_2, S, U)$. ■

III. TWO-WAY COOPERATION

Here we extend the setting from the previous section to a MAC with two-way cooperation where different state information is available at each encoder and full state information is available at the receiver, as depicted in Fig. 1.

Definition 2: A $(2^{nR_1}, 2^{nR_2}, 2^{nC_{12}}, 2^{nC_{21}}, n)$ code with two-way cooperating encoders, where each encoder has partial state information, consists of four encoding functions

$$\begin{aligned} f_{12} &: \{1, \dots, 2^{nR_1}\} \times \mathcal{S}_1^n \mapsto \{1, \dots, 2^{nC_{12}}\}, \\ f_{21} &: \{1, \dots, 2^{nR_2}\} \times \{1, \dots, 2^{nC_{12}}\} \times \mathcal{S}_2^n \mapsto \{1, \dots, 2^{nC_{21}}\}, \end{aligned}$$

$$\begin{aligned} f_1 &: \{1, \dots, 2^{nR_1}\} \times \{1, \dots, 2^{nC_{21}}\} \times \mathcal{S}_1^n \mapsto \mathcal{X}_1^n, \\ f_2 &: \{1, \dots, 2^{nR_2}\} \times \{1, \dots, 2^{nC_{12}}\} \times \mathcal{S}_2^n \mapsto \mathcal{X}_2^n, \end{aligned} \quad (31)$$

and a decoding function,

$$g : \mathcal{Y}^n \times \mathcal{S}_1^n \times \mathcal{S}_2^n \mapsto \{1, \dots, 2^{nR_1}\} \times \{1, \dots, 2^{nR_2}\}. \quad (32)$$

The probability of error, achievable rates, and the capacity region are defined similarly to Definition 1. The next theorem states the capacity region of the two-way cooperating encoders with partial state information.

Theorem 4: The capacity region of the MAC with two-way cooperating encoders and with partial state information as shown in Fig. 1 is the closure of the set of rates that satisfy (33)–(37) at the bottom of the page. for some joint distribution of the form shown in (38) at the bottom of the page, where U and V are auxiliary random variables with bounded cardinality.

In the achievability proof of the theorem we use double-binning, which was introduced by Liu *et al.* [5], [6] to achieve secrecy capacity in the broadcast channel. Here the double-binning is needed since one layer of binning will be used for transmitting a common message between the encoders and an additional layer of binning is needed for choosing a specific typical sequence using side information as done in the Wyner-Ziv problem [32] and two-way source coding [33]. In a double-binning coding scheme we have special bins that contain other bins rather than codewords, and we call such a special bin a *superbin*, as depicted in Fig. 4.

We would like to emphasize that in the two way cooperation problem that we discuss here the ordering of operations on the conferencing links (first transmitter 1, then transmitter 2) is imposed in the problem definition in (31). This is different from the definition of conferencing that is in Willems' work [2] which allows for any ordering and for multiple rounds of conferencing. In order to deal with multiple rounds of conferencing one should use techniques of two-way source coding developed by Kaspi

$$(I_i(U; S), I_i(X_1; Y|X_2, S, U), I_i(X_2; Y|X_1, S, U), I_i(X_1, X_2; Y|S, U), I_i(X_1, X_2; Y|S)), \quad (29)$$

$$C_{12} \geq I(U; S_1|S_2) \quad (33)$$

$$C_{21} \geq I(V; S_2|S_1, U) \quad (34)$$

$$R_1 \leq I(X_1; Y|X_2, S_1, S_2, U, V) + C_{12} - I(U; S_1|S_2) \quad (35)$$

$$R_2 \leq I(X_2; Y|X_1, S_1, S_2, U, V) + C_{21} - I(V; S_2|S_1, U) \quad (36)$$

$$R_1 + R_2 \leq \min \left\{ \begin{array}{l} I(X_1, X_2; Y|S_1, S_2, U, V) + C_{12} + C_{21} - I(U; S_1|S_2) - I(V; S_2|S_1, U) \\ I(X_1, X_2; Y|S_1, S_2) \end{array} \right\}, \quad (37)$$

$$P(s_1, s_2)P(u|s_1)P(v|s_2, u)P(x_1|s_1, u, v)P(x_2|s_2, u, v)P(y|x_1, x_2, s_1, s_2), \quad (38)$$

Now, Encoder 1 transmits $x_1^n(s_1^n, u^n, v^n, m_{1,b})$, and Encoder 2 transmits $x_2^n(s_2^n, u^n, v^n, m_{2,b})$.

Decoder: The decoder knows (s_1^n, s_2^n, y^n) and looks for the indices $\hat{m}_{1,a}$, $\hat{m}_{1,b}$, $\hat{m}_{2,a}$ and $\hat{m}_{2,b}$ such that (see (39) at the bottom of the page), where u^n and v^n which are not indexed denote $u^n(\hat{m}_{1,a}, s_1^n)$ and $v^n(\hat{m}_{2,a}, s_2^n, u^n(\hat{m}_{1,a}, s_1^n))$, respectively. If none or more than one such quadruplet is found, an error is declared. The estimated message sent from Encoder 1 is $(\hat{m}_{1,a}, \hat{m}_{1,b})$, and the estimated message transmitted from Encoder 2 is $(\hat{m}_{2,a}, \hat{m}_{2,b})$.

Error Analysis: Assume $(m_{1,a}, m_{2,a}, m_{1,b}, m_{2,b}) = (1, 1, 1, 1)$. Let us define the event shown in (40) at the bottom of the page, where u^n and v^n which are not indexed denotes $u^n(i, s_1^n)$ and $v^n(j, s_2^n, u^n(i, s_1^n))$, respectively.

We have an error if either the correct codewords are not jointly typical with the received sequences, i.e., $E_{1,1,1,1}^c$, or there exists a different $(i, j, k, l) \neq (1, 1, 1, 1)$ such that $E_{i,j,k,l}$ occurs. From the union of bounds we obtain that

$$\begin{aligned} P_e^{(n)} &\leq \Pr(E_{1,1,1,1}^c) + \sum_{i=1, j=1, k=1, l>1} \Pr(E_{i,j,k,l}) \\ &+ \sum_{i=1, j=1, k>1, l=1} \Pr(E_{i,j,k,l}) + \sum_{i=1, j=1, k>1, l>1} \Pr(E_{i,j,k,l}) \\ &+ \sum_{(i,j) \neq 1, k \geq 1, l \geq 1} \Pr(E_{i,j,k,l}). \end{aligned} \quad (41)$$

Now let us show that each term in (41) goes to zero as the block-length of the code n goes to infinity.

- Upper-bounding $\Pr(E_{1,1,1,1}^c)$: Since the total number of codewords in each superbin associated with i (or $m_{1,a}$) is larger than $I(U; S_1)$, and since the codewords were generated i.i.d. $\sim P(u)$, with high probability there will be at least one codeword that is jointly typical with s_1^n . Let us denote this codeword by $u^n(1, s_1^n)$. Since the Markov form $U - S_1 - S_2$ holds, from the Markov lemma [34] with high probability $u^n(1, s_1^n)$ would be jointly typical with S_2^n . Furthermore, since each bin in the superbin that

is associated with i contains $2^{n(I(U; S_2) - \epsilon)}$ codewords, with high probability, there will not be any additional codeword that is jointly typical with s_2^n , hence, Encoder 2 would identify $u^n(1, s_1^n)$ from the received bin.

Similarly, for a given $u^n \in T_\epsilon^{(n)}(U|s_1^n, s_2^n)$, which is known to Encoder 2, the total number of codewords in each superbin associated with j (or $m_{2,a}$) is larger than $I(V; S_2|U)$, and since the codewords were generated i.i.d. according to $P(v|u)$, with high probability there will be at least one codeword that is jointly typical with (s_2^n, u^n) . Let us denote this codeword by $v^n(1, s_2^n, u^n)$. Since the Markov form $V - (S_2, U) - S_1$ holds, it follows from the Markov lemma that with high probability $v^n(1, s_2^n, u^n)$ would be jointly typical with (s_1^n, u^n) . Furthermore, since each bin in the superbin that is associated with j contains $2^{n(I(V; S_1|U) - \epsilon)}$ codewords, with high probability, there would not be any additional codeword that is jointly typical with (s_1^n, u^n) , hence, Encoder 2, would identify $v^n(1, s_2^n, u^n)$ from the bin.

Furthermore, given that $(u^n, v^n, s_1^n, s_2^n) \in T_\epsilon^{(n)}(U, V, S_1, S_2)$, it follows from the law of large numbers that $\Pr(E_{1,1,1,1}^c) \rightarrow 0$ as n goes to infinity. Note that the event $E_{1,1,1,1}^c$ includes the error that may occur at the encoders side.

- Upper-bounding $\sum_{i=1, j=1, k=1, l>1} \Pr(E_{i,j,k,l})$: The probability that Y^n , which is generated according to $P(y|x_1, s, u, v)$, is jointly typical with x_2^n , which was generated according to $P(x_2|u, v, s_2) = P(x_2|u, v, s_2, s_1, x_1)$, where $(x_1^n, s_1^n, s_2^n, u^n, v^n) \in T_\epsilon^{(n)}(X_1, S_1, S_2, U, V)$ is upper bounded according to Lemma 3 by (42) at the bottom of the page. Hence, we obtain (43) at the bottom of the next page.
- Upper-bounding $\sum_{i=1, j=1, k>1, l=1} \Pr(E_{i,j,k,l})$: The probability that Y^n , which is generated according to $P(y|x_2, s, u, v)$, is jointly typical with x_1^n , which was generated according to $P(x_1|u, v, s_1) = P(x_1|u, v, s_2, s_1, x_2)$, where $(x_2^n, s_1^n, s_2^n, u^n, v^n) \in$

$$(u^n(\hat{m}_{1,a}, s_1^n), v^n(\hat{m}_{2,a}, s_2^n, u^n), x_1^n(s_1^n, u^n, v^n, \hat{m}_{1,b}), x_2^n(s_2^n, u^n, v^n, \hat{m}_{2,b}), s_1^n, s_2^n, y^n) \in T_\epsilon^{(n)}(U, V, X_1, X_2, S_1, S_2, Y), \quad (39)$$

$$E_{i,j,k,l} \triangleq \left\{ (u^n(i, s_1^n), v^n(j, s_2^n, u^n), x_1^n(s_1^n, u^n, v^n, k), x_2^n(s_2^n, u^n, v^n, l), s_1^n, s_2^n, y^n) \in T_\epsilon^{(n)}(U, V, X_1, X_2, S_1, S_2, Y) \right\}, \quad (40)$$

$$\Pr\{(x_1^n, X_2^n, u^n, v^n, s_1^n, s_2^n, Y^n) \in T_\epsilon^{(n)} | (x_1^n, u^n, v^n, s_1^n, s_2^n) \in T_\epsilon^{(n)}\} \leq 2^{-n(I(X_2; Y|X_1, S_1, S_2, U, V) - \delta(\epsilon))}. \quad (42)$$

$T_\epsilon^{(n)}(X_2, S_1, S_2, U, V)$ is upper bounded according to Lemma 3 by (44) at the bottom of the page. Hence, we obtain

$$\sum_{i=1, j=1, k>1, l=1} \Pr(E_{i,j,k,l}) \leq 2^{n(R_1 - (C_{12} - I(U; S_1) + I(U; S_2) - 2\epsilon))} 2^{-n(I(X_1; Y | X_2, S_1, S_2, U, V) - \delta(\epsilon))} \quad (45)$$

- Upper-bounding $\sum_{i=1, j=1, k>1, l>1} \Pr(E_{i,j,k,l})$: see (46) at the bottom of the page.
- Upper-bounding $\sum_{i=1, j=1, k>1, l>1} \Pr(E_{i,j,k,l})$

$$\sum_{(i,j) \neq 1, k \geq 1, l \geq 1} \Pr(E_{i,j,k,l}) \leq 2^{n(R_1 + R_2)} 2^{-n(I(U, V, X_2, X_1; Y | S_1, S_2) - \delta(\epsilon))} \quad (47)$$

Finally, we note that if the rate-pair (R_1, R_2) is in the rate region that is given by (33)–(38), then each term in (41) goes to zero as $n \rightarrow \infty$; hence there exists a sequence of codes $(2^{nR_1}, 2^{nR_2}, 2^{nC_{12}}, 2^{nC_{21}}, n)$ such that $P_\epsilon^{(n)}$ goes to zero as $n \rightarrow \infty$.

Converse Part: The converse part combines techniques from cooperation in a MAC [2] and two-way source coding [33]. Assume that we have a $(2^{nR_1}, 2^{nR_2}, 2^{nC_{12}}, 2^{nC_{21}}, n)$ code as in Definition 2. We will show the existence of a joint distribution $P(s_1, s_2)P(u|s_1)P(v|s_2, u)P(x_1|s_1, u, v)P(x_2|s_2, u, v)P(y|x_1, x_2, s_1, s_2)$ that satisfies (33)–(37) within some ϵ_n , where ϵ_n goes to zero as $n \rightarrow \infty$. Consider

$$\begin{aligned} nC_{12} &\geq H(M_{12}) \\ &= H(M_{12}|S_2^n) \\ &\geq I(M_{12}; S_1^n | S_2^n) \\ &= \sum_{i=1}^n H(S_{1,i}|S_{2,i}) - H(S_{1,i}|M_{12}, S_1^{i-1}, S_2^n) \end{aligned}$$

$$\begin{aligned} &\stackrel{(a)}{=} \sum_{i=1}^n H(S_{1,i}|S_{2,i}) - H(S_{1,i}|M_{12}, S_1^{i-1}, S_{2,i}^n, S_{2,i+1}^n) \\ &= \sum_{i=1}^n I(S_{1,i}; M_{12}, S_1^{i-1}, S_{2,i+1}^n | S_{2,i}) \\ &\stackrel{(b)}{=} \sum_{i=1}^n I(S_{1,i}; U_i | S_{2,i}), \end{aligned} \quad (48)$$

where (a) follows from the Markov chain $S_{1,i} - (M_{12}, S_1^{i-1}, S_{2,i}^n, S_{2,i+1}^n) - S_2^{i-1}$ and (b) follows from the definition

$$U_i \triangleq (M_{12}, S_1^{i-1}, S_{2,i+1}^n). \quad (49)$$

Now, consider

$$\begin{aligned} nC_{21} &\geq H(M_{21}) \\ &\geq H(M_{21}|M_{12}, S_1^n) \\ &\geq I(M_{21}; S_2^n | M_{12}, S_1^n) \\ &= \sum_{i=1}^n H(S_{2,i}|S_{2,i+1}^n, S_1^n, M_{12}) \\ &\quad - H(S_{2,i}|S_{2,i+1}^n, S_1^n, M_{12}, M_{21}) \\ &\stackrel{(a)}{=} \sum_{i=1}^n H(S_{2,i}|S_{2,i+1}^n, S_1^n, M_{12}) \\ &\quad - H(S_{2,i}|S_{2,i+1}^n, S_1^i, M_{12}, M_{21}) \\ &\stackrel{(b)}{=} \sum_{i=1}^n H(S_{2,i}|U_i, S_{1,i}) - H(S_{2,i}|U_i, S_{1,i}, V_i) \\ &= \sum_{i=1}^n I(S_{2,i}; V_i | U_i, S_{1,i}), \end{aligned} \quad (50)$$

where (a) follows from the Markov chain $S_{2,i} - (S_{2,i+1}^n, S_1^i, M_{12}, M_{21}) - S_{1,i+1}^n$ and (b) follows from the definitions of U_i given in (49) and V_i which is given by

$$V_i \triangleq M_{21}. \quad (51)$$

$$\sum_{i=1, j=1, k=1, l>1} \Pr(E_{i,j,k,l}) \leq 2^{n(R_2 - (C_{21} - I(V; S_2|U) + I(V; S_1|U) - 2\epsilon))} 2^{-n(I(X_2; Y | X_1, S_1, S_2, U, V) - \delta(\epsilon))} \quad (43)$$

$$\Pr\{X_1^n, x_2^n, u^n, v^n, s_1^n, s_2^n, Y^n \in T_\epsilon^{(n)} | x_2^n, u^n, v^n, s_1^n, s_2^n \in T_\epsilon^{(n)}\} \leq 2^{-n(I(X_1; Y | X_2, S_1, S_2, U, V) - \delta(\epsilon))}. \quad (44)$$

$$\sum_{i=1, j=1, k>1, l>1} \Pr(E_{i,j,k,l}) \leq 2^{n(R_1 - (C_{12} - I(U; S_1) + I(U; S_2) - 2\epsilon) + R_2 - (C_{21} - I(V; S_2|U) + I(V; S_1|U) - 2\epsilon))} 2^{-n(I(X_2, X_1; Y | S_1, S_2, U, V) - \delta(\epsilon))} \quad (46)$$

Now, consider

$$\begin{aligned}
nR_1 &= H(M_1) \\
&= H(M_1|S_1^n, S_2^n, M_2) \\
&= H(M_1, M_{12}|S_1^n, S_2^n, M_2) \\
&= H(M_{12}|S_1^n, S_2^n, M_2) + H(M_1|S_1^n, S_2^n, M_2, M_{12}) \\
&\leq H(M_{12}|S_1^n, S_2^n) + H(M_1|S_1^n, S_2^n, M_2, M_{12}). \quad (52)
\end{aligned}$$

Now, let us consider the terms $H(M_{12}|S_1^n, S_2^n)$ and $H(M_1|S_1^n, S_2^n, M_2, M_{12})$ separately.

$$\begin{aligned}
H(M_{12}|S_1^n, S_2^n) &= H(M_{12}|S_1^n, S_2^n) - H(M_{12}) + H(M_{12}) \\
&\leq nC_{12} - I(S_1^n, S_2^n; M_{12}) \\
&\stackrel{(a)}{=} nC_{12} - I(S_1^n; M_{12}|S_2^n) \\
&\stackrel{(b)}{=} nC_{12} - \sum_{i=1}^n I(S_{1,i}; U_i|S_{2,i}), \quad (53)
\end{aligned}$$

where (a) follows from the fact that M_{12} is independent of S_2^n , and (b) follows from (48), where it is shown that $I(S_1^n; M_{12}|S_2^n) = \sum_{i=1}^n I(S_{1,i}; U_i|S_{2,i})$. Now consider the second term,

$$\begin{aligned}
H(M_1|S_1^n, S_2^n, M_2, M_{12}) &= H(M_1|S_1^n, S_2^n, M_2, M_{12}, M_{21}) \\
&\stackrel{(a)}{=} I(M_1; Y^n|S_1^n, S_2^n, M_2, M_{12}, M_{21}) + n\epsilon_n \\
&= H(Y^n|S_1^n, S_2^n, M_2, M_{12}, M_{21}) \\
&\quad - H(Y^n|S_1^n, S_2^n, M_2, M_{12}, M_{21}, M_1) + n\epsilon_n \\
&= H(Y^n|S_1^n, S_2^n, X_2^n, M_2, M_{12}, M_{21}) \\
&\quad - H(Y^n|S_1^n, S_2^n, X_2^n, X_1^n, M_2, M_{12}, M_{21}, M_1) + n\epsilon_n \\
&= \sum_{i=1}^n H(Y_i|Y^{i-1}, S_1^n, S_2^n, X_2^n, M_2, M_{12}, M_{21}) \\
&\quad - H(Y_i|S_1^n, S_2^n, X_2^n, X_1^n, M_2, M_{12}, M_{21}, M_1, Y^{i-1}) \\
&\quad + n\epsilon_n \\
&\stackrel{(b)}{\leq} \sum_{i=1}^n H(Y_i|S_{1,i}, S_{2,i}, X_{2,i}, M_{12}, S_1^{i-1}, S_{2,i+1}^n, M_{21}) \\
&\quad - H(Y_i|S_{1,i}, S_{2,i}, X_{2,i}, X_{1,i}, M_{12}, S_1^{i-1}, S_{2,i+1}^n, M_{21}) \\
&\quad + n\epsilon_n \\
&= \sum_{i=1}^n I(X_{1,i}; Y_i|S_{1,i}, S_{2,i}, X_{2,i}, U_i, V_i) \\
&\quad + n\epsilon_n, \quad (54)
\end{aligned}$$

where (a) follows from Fano's inequality and $\epsilon_n \triangleq R_1 P_e^{(n)}$ and (b) from that fact that conditioning reduces entropy. Substituting Inequalities (53) and (54) into (52), we obtain

$$\begin{aligned}
nR_1 &\leq \sum_{i=1}^n I(X_{1,i}; Y_i|S_{1,i}, S_{2,i}, X_{2,i}, U_i, V_i) \\
&\quad - I(S_{1,i}; U_i|S_{2,i}) + nC_{12}. \quad (55)
\end{aligned}$$

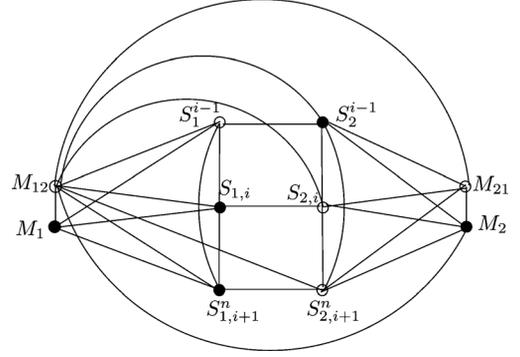


Fig. 5. Proof of the Markov chain $(M_2, S_2^n) - (S_{2,i}, M_{12}, S_1^{i-1}, S_{2,i+1}, M_{21}) - (M_1, S_1^n)$ using an undirected graphical technique [35]. The undirected graph corresponds to the joint distribution $P(S_1^{i-1}, S_2^{i-1})P(S_{1,i}, S_{2,i})P(S_{1,i+1}, S_{2,i+1})P(M_1)P(M_2)P(M_{12}|m_1, s_1^n)P(M_{21}|m_2, m_{12}, s_2^n)$. The Markov chain follows from the fact that all the paths from (M_1, S_1^n) to (M_2, S_2^n) go through the nodes $(S_{2,i}, M_{12}, S_1^{i-1}, S_{2,i+1}, M_{21})$.

Similarly, we obtain

$$\begin{aligned}
nR_2 &\leq \sum_{i=1}^n I(X_{2,i}; Y_i|S_{1,i}, S_{2,i}, X_{1,i}, U_i, V_i) \\
&\quad - I(S_{2,i}; V_i|S_{1,i}, U_i) + nC_{21}. \quad (56)
\end{aligned}$$

Regarding the sum-rate we have

$$\begin{aligned}
nR_1 + nR_2 &= H(M_1, M_2) \\
&= H(M_1, M_2|S_1^n, S_2^n) \\
&= I(M_1, M_2; Y^n|S_1^n, S_2^n) + n\epsilon_n \\
&\leq I(X_{1,i}, X_{2,i}; Y_i|S_{1,i}, S_{2,i}) + n\epsilon_n, \quad (57)
\end{aligned}$$

and

$$\begin{aligned}
nR_1 + nR_2 &= H(M_1, M_2) \\
&= H(M_1, M_2, M_{12}, M_{21}|S_1^n, S_2^n) \\
&= H(M_{12}|S_1^n, S_2^n) + H(M_{21}|S_1^n, S_2^n, M_{12}) \\
&\quad + H(M_1, M_2|S_1^n, S_2^n, M_{12}, M_{21}) + \epsilon_n, \quad (58)
\end{aligned}$$

and now using (53) we bound

$$H(M_{12}|S_1^n, S_2^n) \leq nC_{12} - \sum_{i=1}^n I(S_{1,i}; U_i|S_{2,i}), \quad (59)$$

and similarly

$$H(M_{21}|S_1^n, S_2^n, M_{12}) \leq nC_{21} - \sum_{i=1}^n I(S_{2,i}; V_i|S_{1,i}, U_i). \quad (60)$$

Using similar steps as in (54) we bound

$$\begin{aligned}
H(M_1, M_2|S_1^n, S_2^n, M_{12}, M_{21}) &\leq \sum_{i=1}^n I(X_{2,i}, X_{1,i}; Y_i|S_{1,i}, S_{2,i}, U_i, V_i). \quad (61)
\end{aligned}$$

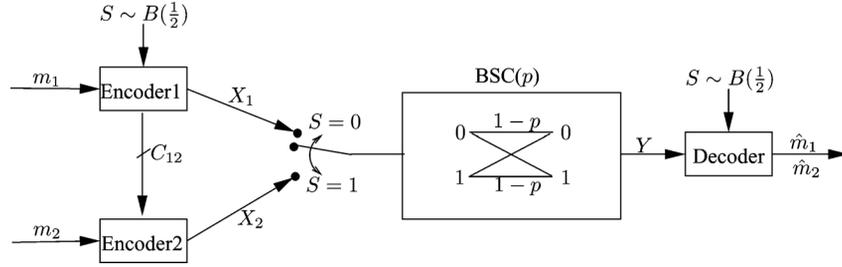


Fig. 6. An example of a MAC with one-way cooperation and state information at one encoder and the decoder. The state S controls the switch. When $S = 0$, $Y = X_1 + Z$ and when $S = 1$, $Y = X_2 + Z$, and $Z \sim B(p)$.

Hence we obtain

$$nR_1 + nR_2 \leq \sum_{i=1}^n I(X_{2,i}, X_{1,i}; Y_i | S_{1,i}, S_{2,i}, U_i, V_i) - I(S_{1,i}; U_i | S_{2,i}) - I(S_{2,i}; V_i | S_{1,i}, U_i) + nC_{12} + nC_{21}.$$

Now we need to verify that the following Markov chains hold:

$$(M_{12}, S_1^{i-1}, S_{2,i+1}^n) - S_{1,i} - S_{2,i}, \quad (62)$$

$$M_{21} - (S_{2,i}, M_{12}, S_1^{i-1}, S_{2,i+1}^n) - S_{1,i}, \quad (63)$$

$$X_{1,i}(M_1, S_1^n, M_{21}) - (S_{1,i}, M_{12}, S_1^{i-1}, S_{2,i+1}^n, M_{21}) - S_{2,i}, \quad (64)$$

$$X_{2,i}(M_2, S_2^n, M_{12}) - (S_{2,i}, M_{12}, S_1^{i-1}, S_{2,i+1}^n, M_{21}) - (S_{1,i}, X_{1,i}(M_1, S_1^n, M_{21})). \quad (65)$$

Proving the Markov chains (62)–(64) is straightforward and therefore omitted. To prove the Markov chain in (65), we use the undirected graphical method from [35, Section II], (the technique is similar in spirit to techniques in [36, Ch 1.2], [37], [38, pp. 231]). Fig. 5 proves the Markov chain $(M_2, S_2^n) - (S_{2,i}, M_{12}, S_1^{i-1}, S_{2,i+1}^n, M_{21}) - (M_1, S_1^n)$, and as a consequence the Markov chain in (65) holds too.

Finally, let Q be a random variable independent of (X^n, S_1^n, S_2^n, Y^n) , and uniformly distributed over the set $\{1, 2, 3, \dots, n\}$. Define the random variables $U \triangleq (Q, U_Q)$, $V \triangleq (Q, V_Q)$, and we obtain that the region given by (33)–(38) is an outer bound to the set of all achievable rate-pairs.

To show that the cardinalities of the random variables U and V are bounded we follow similar steps as in Lemma 2, first for U and then for V . We note that the cardinality of auxiliary random variables U and V may be bounded by $|\mathcal{U}| \leq \min(|\mathcal{X}_1||\mathcal{X}_2||\mathcal{S}_1||\mathcal{S}_2| + 4, |\mathcal{Y}||\mathcal{S}_1||\mathcal{S}_2| + 5)$, and $|\mathcal{V}| \leq \min(|\mathcal{X}_1||\mathcal{X}_2||\mathcal{S}_1||\mathcal{S}_2||\mathcal{U}| + 3, |\mathcal{Y}||\mathcal{S}_1||\mathcal{S}_2||\mathcal{U}| + 4)$. ■

IV. EXAMPLE AND COMPARISON TO MESSAGE-ONLY AND STATE-ONLY COOPERATION

Consider the example given in Fig. 6, where the state of the channel controls the switch that determines which input goes through a binary symmetric channel (BSC) with parameter p . When $S = 0$, the binary input X_1 goes through and when $S =$

1 the binary input X_2 goes through, hence the output of the channel Y is given by

$$Y = \bar{S}X_1 \oplus SX_2 \oplus Z, \quad (66)$$

where $Z \sim \text{Bernoulli}(\frac{1}{2})$ and is independent of S , the symbol \oplus denotes XOR, and \bar{S} denotes $1 - S$. We also have the constraint on the portion of '1's at the encoders, namely for any pair of codeword (x_1^n, x_2^n) , $\frac{1}{n} \sum_{i=1}^n x_{1,i} \leq p_1$ and $\frac{1}{n} \sum_{i=1}^n x_{2,i} \leq p_2$. The capacity region is given by Theorem 1 where additional constraints $\mathbb{E}[X_1] \leq p_1$ and $\mathbb{E}[X_2] \leq p_2$ are needed ($\mathbb{E}[\cdot]$ denotes expectation). The achievability proof of the Theorem 1 with the input constraint is the same since by the law of large numbers the constraints are satisfied with high probability for each codeword. The Converse is also similar just adding a step $p_1 \geq \frac{1}{n} \mathbb{E}[\sum_{i=1}^n X_{1,i}] = \mathbb{E}[\mathbb{E}[X_{1,Q}|Q]] = \mathbb{E}[X_{1,Q}] = \mathbb{E}[X_1]$ where the last equality follows from the definition of the random variable X_1 to be $X_{1,Q}$. The cardinality bound in Lemma 2 when there are two input constraints increases to $|\mathcal{U}| \leq \min(|\mathcal{X}_1||\mathcal{X}_2||\mathcal{S}| + 3, |\mathcal{Y}||\mathcal{S}| + 6)$ since additional terms $\mathbb{E}[X_1] = \mathbb{E}[\mathbb{E}[X_1|U]]$ and $\mathbb{E}[X_2] = \mathbb{E}[\mathbb{E}[X_2|U]]$ need to be preserved.

Invoking the following identities

$$\begin{aligned} I(X_2; Y | X_1, S, U) &= \frac{1}{2} I(X_2; Y | X_1, S = 1, U) \\ &= \frac{1}{2} I(X_2; Y | S = 1, U) \\ &= \frac{1}{2} H(X_2 \oplus Z | S = 1, U) - \frac{1}{2} H(Z) \\ I(X_1; Y | X_2, S, U) &= \frac{1}{2} H(X_1 \oplus Z | S = 0) - \frac{1}{2} H(Z) \\ I(X_1, X_2; Y | S, U) &= \frac{1}{2} H(X_1 \oplus Z | S = 0) \\ &\quad + \frac{1}{2} H(X_2 \oplus Z | S = 1, U) - H(Z) \\ I(X_1, X_2; Y | S) &= \frac{1}{2} H(X_1 \oplus Z | S = 0) \\ &\quad + \frac{1}{2} H(X_2 \oplus Z | S = 1) - H(Z), \end{aligned} \quad (67)$$

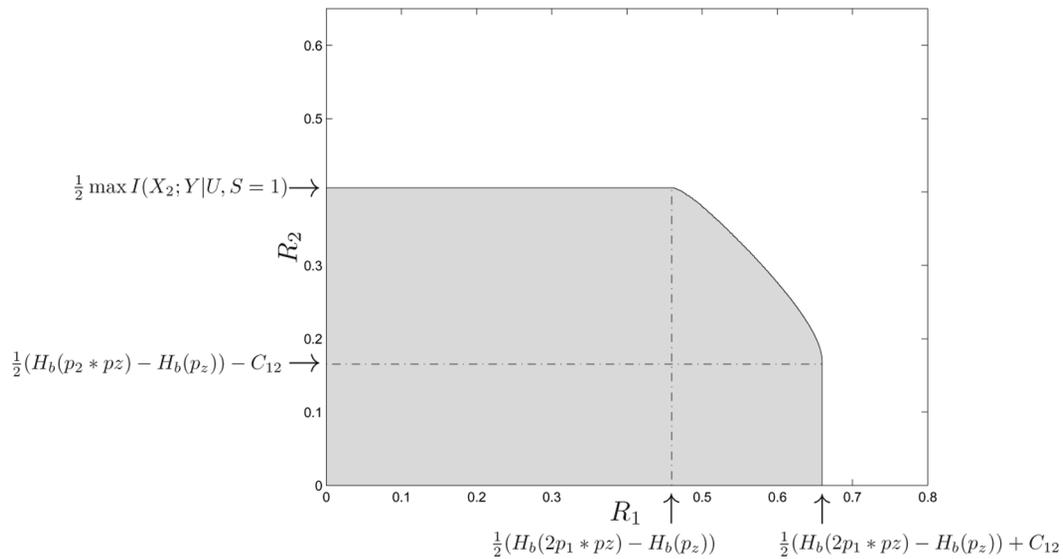


Fig. 7. Capacity region of the example depicted in Fig. 6 where $C_{12} = 0.2$, $p_z = 0.01$, $p_1 = p_2 = 0.25$.

we obtain from Theorem 1 that the capacity region is the set of all rate-pairs (R_1, R_2) that satisfy

$$\begin{aligned} R_1 &\leq \frac{1}{2}H_b(p_1 * p_z) - \frac{1}{2}H_b(p_z) + C_{12} - I(U; S) \\ R_2 &\leq \frac{1}{2}H(X_2 \oplus Z|S = 1, U) - \frac{1}{2}H_b(p_z) \\ R_1 + R_2 &\leq \frac{1}{2}H_b(p_1 * p_z) + \frac{1}{2}H(X_2 \oplus Z|S = 1) - H_b(p_z) \end{aligned} \quad (68)$$

for some conditional distributions $P(u|s)$ and $P(x_2|u)$ where $I(U; S) \leq C_{12}$. The term $H_b(p)$ denotes the binary entropy function, which is defined for $0 \leq p \leq 1$ as $H_b(p) = -p \log p - (1-p) \log(1-p)$. The term $p * q$ denotes the parameter of a Bernoulli distribution that results from convolving mod-2 two Bernoulli distributions with parameters p and q , i.e., $p * q = (1-p)q + (1-q)p$.

Fig. 7 depicts the capacity region for the case where $C_{12} = 0.2$, $p_z = 0.01$ and $p_1 = p_2 = 0.25$. The capacity region was numerically evaluated using (68), where the cardinality of the auxiliary random variable U was assumed to be $|\mathcal{U}| = 2$; changing the cardinality to 3, 4, 5 did not increase the numerical capacity region.

Fig. 8 illustrates the influence of the cooperation rate on the capacity region. It shows the capacity regions for several rates of cooperation $C_{12} = [0, 0.2, 0.5, 1]$ where $p_z = 0.01$, $p_1 = p_2 = 0.25$. One can see that when the cooperation rate is small an increase in the cooperation rate significantly influences the capacity region; however, for a large cooperation rate, such as $C_{12} > 0.5$, an increase in the cooperation rate hardly influences the capacity region.

Comparison to Two Different Kinds of Cooperation: In the setting analyzed in this paper, we assumed a cooperation link that may use both the message and the state information. Recent works assumed similar settings where the cooperation depends

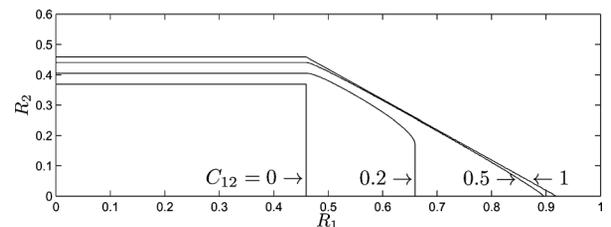


Fig. 8. Capacity region of the example depicted in Fig. 6 for several values of C_{12} , i.e., $C_{12} = [0, 0.2, 0.5, 1]$ and $p_z = 0.01$, $p_1 = p_2 = 0.25$.

only on the state [26], as depicted in Fig. 9, or on the message only [22], [23] as depicted in Fig. 10. For the first case where the cooperation may use only the state information (Fig. 9), the capacity region was derived in [26] and may be written as

$$C_{12} = I(U; S) \quad (69)$$

$$R_1 \leq I(X_1; Y|X_2, S, U) \quad (70)$$

$$R_2 \leq I(X_2; Y|X_1, S, U) \quad (71)$$

$$R_1 + R_2 \leq I(X_1, X_2; Y|S, U), \quad (72)$$

for some joint distribution of the form

$$P(s, u)P(x_1|s, u)P(x_2|u)P(y|x_1, x_2, s). \quad (73)$$

For the second case where the cooperation may use only the message (Fig. 10) the capacity region was considered in [22], [23] and may be written as

$$R_1 \leq I(X_1; Y|X_2, S, U) + C_{12} \quad (74)$$

$$R_2 \leq I(X_2; Y|X_1, S, U) \quad (75)$$

$$R_1 + R_2 \leq \min \left\{ \begin{array}{l} I(X_1, X_2; Y|S, U) + C_{12}, \\ I(X_1, X_2; Y|S) \end{array} \right\}, \quad (76)$$

for some joint distribution of the form

$$P(s)P(u)P(x_1|s, u)P(x_2|u)P(y|x_1, x_2, s), \quad (77)$$

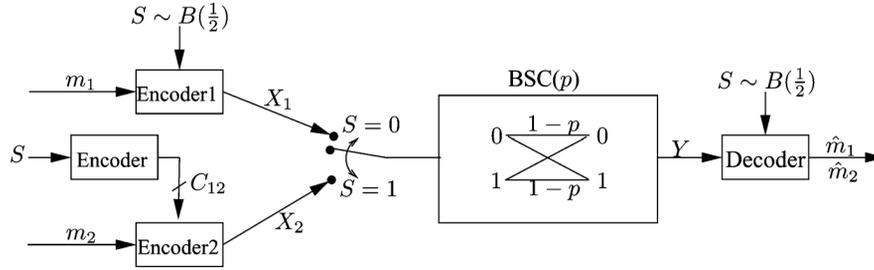


Fig. 9. State cooperation. An example inspired by the setting in [26], where the cooperation is a limited rate state information and is independent of the message.

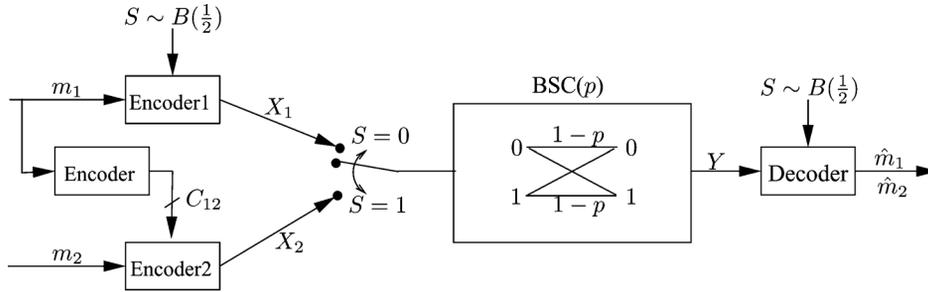


Fig. 10. Message cooperation. An example inspired by the setting in [22], where the cooperation is a function of the message only, and then after the cooperation stage the channel state is available to Encoder 1.

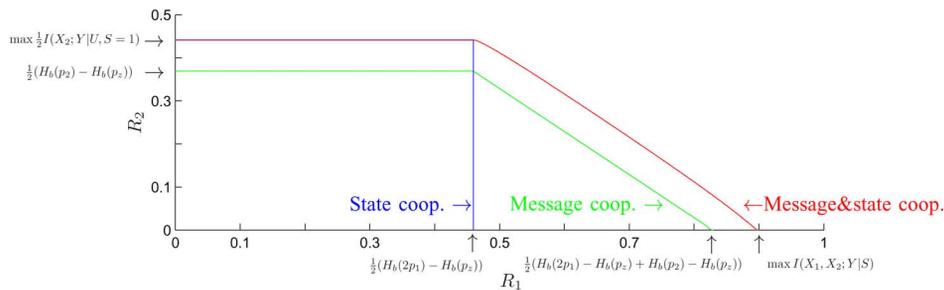


Fig. 11. The regions of three settings with a cooperation link $C_{12} = 0.5$. The blue region corresponds to the case where the cooperation is based only on the state information as depicted in Fig. 9. The green region corresponds to the case where the cooperation is based only on the message and not on the state as depicted in Fig. 10. Finally, the red region is the one that corresponds to the setting of this paper where the cooperation may use both the state and the message as depicted in Fig. 6.

where U and V are auxiliary random variables with bounded cardinality.

Both regions, the one in (72)–(73) and the one in (76)–(77), are contained in the region of Theorem 1 where the cooperation may use both the message and the state. It is interesting to note that one can obtain the regions (72)–(73), and (76)–(77) by adding only an additional constraint to the region of Theorem 1. More precisely, to obtain the regions (72)–(73) add the constraint $C_{12} = I(U; S)$, and to obtain the region (76)–(77) add the constraint $I(U; S) = 0$ to the region (5)–(9) of Theorem 1.

Fig. 11 depicts the capacity regions obtained for a cooperation link $C_{12} = 0.5$ for the three settings:

- 1) state-cooperation, where the cooperation is based only on the state information (Fig. 9),
- 2) message-cooperation, where the cooperation is based only on the message (Fig. 10),
- 3) message-state cooperation, where the cooperation may use both the state and the message (Fig. 6).

In this example, one can note from Fig. 11 that state cooperation and message cooperation have different gains that are com-

bined in the message-state cooperation. The state-cooperation increases the capacity region only in the direction of R_2 since only Encoder 2 receives the state information via the cooperation link. Message-cooperation increases the capacity region only in the direction of R_1 since only the message from Encoder 1 is transmitted via the cooperation link. However the message-state cooperation increases the capacity region in the direction of both R_1 and R_2 since it combines the advantages of both message and state cooperation.

From the comparison above, it is interesting to note that there are special cases where the state-only cooperation or the message-only cooperation performs as well as the combined state-message cooperation.

Equal rates, i.e., $R_1 = R_2$: consider the example of one-way cooperation depicted in Fig. 6, where we are interested in equal-rates working-point, i.e., $R_1 = R_2$. Since on the boundary region $R_2 \leq R_1$, the best equal-rate working point is achieved by maximizing R_2 . To maximize R_2 in one-way cooperation, there is no need for message cooperation and therefore the state-only cooperation achieves the maximum equal rate point.

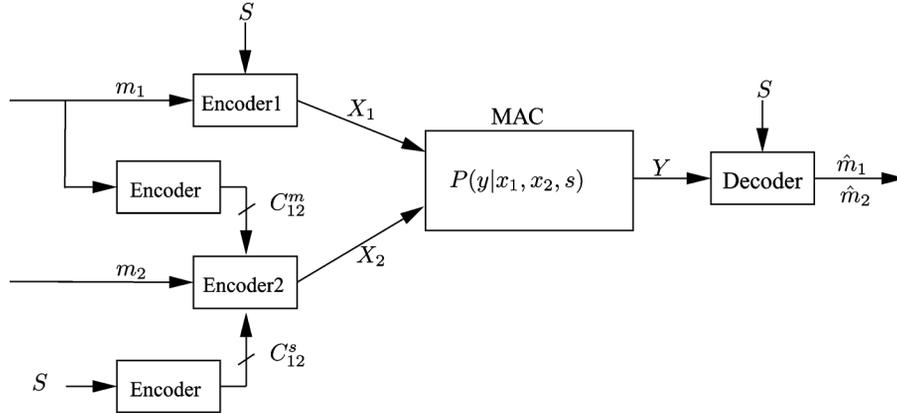


Fig. 12. Separate message and state cooperation. There are two cooperation links at rates C_{12}^m and C_{12}^s . The first link uses only the message information m_1 , the second link uses only the state information S^n .

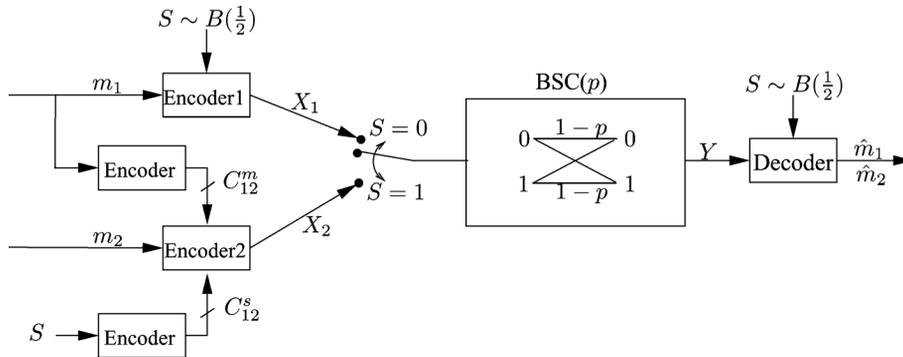


Fig. 13. An example with separate message and state cooperation.

Effectively, no power constraint, $p_1 = p_2 = 0.5$: consider the one-way cooperation as depicted in Fig. 6, where, effectively, there is no power constraint; this means that p_1, p_2 may be equal to or larger than 0.5. For this case, the state information at the transmitter does not enlarge the rate region, hence the message-only cooperation as introduced by Willems [2] is optimal.

A. Splitting the Cooperation Link in Message-Only and State-Only Links

In this subsection we investigate what happens if we split the cooperation link into two links: one link for message-only cooperation at rate C_{12}^m and the other link for state-only cooperation at rate C_{12}^s as shown in Fig. 12. We derive the capacity region for this setting for a general MAC with state $P(y|x_1, x_2, s)$ and show that the split is strictly suboptimal.

Theorem 5: The capacity region of the MAC with separated links, for message and state cooperation, as shown in Fig. 12 is the closure of the set that contains all rates that satisfy

$$C_{12}^s \geq I(U; S) \quad (78)$$

$$R_1 \leq I(X_1; Y|X_2, S, U, V) + C_{12}^m \quad (79)$$

$$R_2 \leq I(X_2; Y|X_1, S, U, V) \quad (80)$$

$$R_1 + R_2 \leq \min \left\{ \begin{array}{l} I(X_1, X_2; Y|S, U, V) + C_{12}^m \\ I(X_1, X_2; Y|S, U) \end{array} \right\}, \quad (81)$$

for some joint distribution of the form

$$P(s, u)P(v)P(x_1|s, u, v)P(x_2|u, v)P(y|x_1, x_2, s). \quad (82)$$

The proof of the theorem for the case where the MAC is of the general form $P(y|x_1, x_2, s)$ is given in Appendix B. The converse is based on the identification of the auxiliary random variable U being a function of the state sequence only, and the identification of the auxiliary random variable V being a function of the message M_1 only; hence the pair (U, S^n) is independent of V , since M_1 is independent of S^n . The achievability part is based on generating the coordination $P_{S,U}$ and then multiplexing the cooperation MAC codebooks according to U . There is no need for binning in the achievability part where the cooperation link is split.

Fig. 14 depicts the capacity region of the example shown in Fig. 13 where separate state cooperation and message cooperation exists and $C_{12}^s = C_{12}^m = 0.25$. From Fig. 14 we learn that using the naive strategy of splitting the cooperation link into message-only cooperation and state-only cooperation is strictly suboptimal.

APPENDIX A CONVERSE PROOF OF THEOREM 1

Converse Part: Assume that we have a $(2^{nR_1}, 2^{nR_2}, 2^{nC_{12}}, n)$ code as in Definition 1.

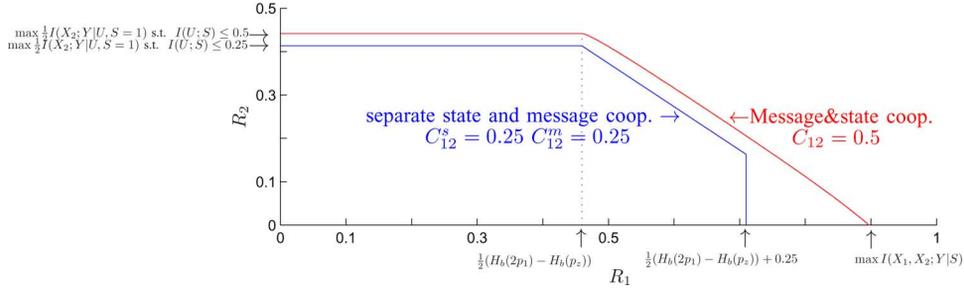


Fig. 14. Comparison between the capacity regions of separate state and message cooperation and joint state-message cooperation. The blue line corresponds to the capacity region of the setting in Fig. 14 where $C_{12}^s = C_{12}^m = 0.25$. The red line corresponds to the capacity region of the setting in Fig. 10 where $C_{12} = 0.5$.

We will show the existence of a joint distribution $P(s)P(u|s)P(x_1|s, u)P(x_2|u)P(y|x_1, x_2)$ that satisfies (5)–(8) within some ϵ_n , where ϵ_n goes to zero as $n \rightarrow \infty$. Denote $M_{12} = f_{12}(M_1, S^n)$. Then,

$$\begin{aligned} nC_{12} &\geq H(M_{12}) \\ &\geq I(M_{12}; S^n) \\ &\stackrel{(a)}{=} \sum_{i=1}^n I(S_i; M_{12}, S^{i-1}) \\ &\stackrel{(b)}{=} \sum_{i=1}^n I(S_i; U_i), \end{aligned} \quad (83)$$

where (a) follows from the fact that S_i is i.i.d. and (b) follows from the definition of U_i , which is

$$U_i \triangleq (M_{12}, S^{i-1}). \quad (84)$$

Next, consider

$$\begin{aligned} nR_1 &= H(M_1) \\ &= H(M_1|S^n, M_2) \\ &\stackrel{(a)}{=} H(M_1, M_{12}|S^n, M_2) \\ &= H(M_{12}|S^n, M_2) + H(M_1|S^n, M_2, M_{12}) \\ &\leq H(M_{12}|S^n) + H(M_1|S^n, M_2, M_{12}), \end{aligned} \quad (85)$$

where step (a) follows from the fact that M_{12} is a function of M_1 and S^n . Now, let us consider the terms $H(M_{12}|S^n)$ and $H(M_1|S^n, M_2, M_{12})$ separately.

$$\begin{aligned} H(M_{12}|S^n) &= H(M_{12}|S^n) - H(M_{12}) + H(M_{12}) \\ &\leq nC_{12} - I(S^n; M_{12}) \\ &= nC_{12} - \sum_{i=1}^n I(S_i; U_i), \end{aligned} \quad (86)$$

where the last equality follows from (83) where it is shown that $I(M_{12}; S^n) = \sum_{i=1}^n I(S_i; U_i)$. Further,

$$\begin{aligned} H(M_1|S^n, M_2, M_{12}) &\stackrel{(a)}{=} I(M_1; Y^n|S^n, M_2, M_{12}) + n\epsilon_n \\ &= H(Y^n|S^n, M_2, M_{12}) \\ &\quad - H(Y^n|S^n, M_2, M_{12}, M_1) + n\epsilon_n \\ &\stackrel{(b)}{=} H(Y^n|S^n, X_2^n, M_2, M_{12}) \end{aligned}$$

$$\begin{aligned} &- H(Y^n|S^n, X_2^n, X_1^n, M_2, M_{12}, M_1) + n\epsilon_n \\ &= \sum_{i=1}^n H(Y_i|Y^{i-1}, S^n, X_2^n, M_2, M_{12}) \\ &\quad - H(Y_i|S^n, X_2^n, X_1^n, M_2, M_{12}, M_1, Y^{i-1}) + n\epsilon_n \\ &\stackrel{(c)}{\leq} \sum_{i=1}^n H(Y_i|S_i, X_{2,i}, M_{12}, S^{i-1}) \\ &\quad - H(Y_i|S_i, X_{2,i}, X_{1,i}, M_{12}, S^{i-1}) + n\epsilon_n, \end{aligned} \quad (87)$$

where (a) follows from Fano's inequality which state that $H(M_1|Y^n, S^n) \leq 1 + nR_1P_e^{(n)}$ and from the definition $\epsilon_n \triangleq \frac{1}{n} + R_1P_e^{(n)}$, (b) follows from the fact that X_1^n is a deterministic function of (S^n, M_1) and X_2^n is a deterministic function of (M_2, M_{12}) , and (c) from the fact that conditioning reduces entropy and from the Markov chain $Y_i - (S_i, X_{2,i}, X_{1,i}) - (S^n, X_2^n, X_1^n, M_2, M_{12}, M_1)$. Substituting Inequalities (86) and (87) into (85), we obtain

$$nR_1 \leq \sum_{i=1}^n I(Y_i; X_{1,i}|S_i, X_{2,i}, U_i) - I(S_i; U_i) + nC_{12}. \quad (88)$$

Similarly, we have

$$\begin{aligned} nR_2 &= H(M_2) \\ &= H(M_2|S^n, M_1, M_{12}) \\ &\leq \sum_{i=1}^n I(Y_i; X_{2,i}|S_i, X_{1,i}, U_i), \end{aligned} \quad (89)$$

where the last inequality follows from similar steps as in (87). Regarding the sum-rate we have

$$\begin{aligned} nR_1 + nR_2 &= H(M_1, M_2) \\ &= H(M_1, M_2|S^n) \\ &= I(M_1, M_2; Y^n|S^n) + n\epsilon_n \\ &= \sum_{i=1}^n H(Y_i|Y^{i-1}, S^n) \\ &\quad - H(Y_i|M_1, M_2, S^n, X_1^n, X_2^n) + n\epsilon_n \\ &\leq \sum_{i=1}^n H(Y_i|S_i) - H(Y_i|S_i, X_{1,i}, X_{2,i}) + n\epsilon_n \\ &\leq \sum_{i=1}^n I(X_{1,i}, X_{2,i}; Y_i|S_i) + n\epsilon_n \end{aligned} \quad (90)$$

and

$$\begin{aligned}
nR_1 + nR_2 &= H(M_1, M_2) \\
&= H(M_1, M_2, M_{12}|S^n) \\
&= H(M_{12}|S^n) + H(M_1, M_2|M_{12}, S^n) + n\epsilon_n,
\end{aligned} \tag{91}$$

and now using (86) and similar steps as in (87) we obtain

$$nR_1 + nR_2 \leq \sum_{i=1}^n I(Y_i; X_{1,i}, X_{2,i}|S_i, U_i) - I(S_i; U_i) + nC_{12}. \tag{92}$$

Now we verify that the Markov chain $X_{2,i} - U_i - (X_{1,i}, S_i)$ holds (this is due to the Markov chain $M_2 - (M_{12}, S^{i-1}) - (M_1, S^n)$). Finally, let Q be a random variable independent of (X^n, S^n, Y^n) , and uniformly distributed over the set $\{1, 2, 3, \dots, n\}$. Define the random variables $U \triangleq (Q, U_Q)$. Using the simple observation that $I(X_1, X_2; Y|S, Q) \leq I(X_1, X_2; Y|S)$, we obtain that the region given in (5)–(9) is an outer bound to any achievable rate.

APPENDIX B PROOF OF THEOREM 5

Here we present the proof of Theorem 5, where MAC is of the general form $P(y|x_1, x_2, s)$.

Proof of Theorem 5:

Converse Part: Assume that we have a $(2^{nR_1}, 2^{nR_2}, 2^{nC_{12}^m}, 2^{nC_{12}^s}, n)$ code. We will show the existence of a joint distribution $P(s)P(u|s)P(v)P(x_1|s, u, v)P(x_2|u, v)P(y|x_1, x_2, s)$ that satisfies (78)–(102) within some ϵ_n , where ϵ_n goes to zero as $n \rightarrow \infty$. Let $M_{12}^s \in \{1, 2, \dots, 2^{nC_{12}^s}\}$ and $M_{12}^m \in \{1, 2, \dots, 2^{nC_{12}^m}\}$ be the message sent on the state cooperation link and the message cooperation link, respectively. Consider

$$\begin{aligned}
nC_{12}^s &\geq H(M_{12}^s) \\
&\geq I(M_{12}^s; S^n) \\
&\stackrel{(a)}{=} \sum_{i=1}^n I(S_i; M_{12}^s, S^{i-1}) \\
&\stackrel{(b)}{=} \sum_{i=1}^n I(S_i; U_i),
\end{aligned} \tag{93}$$

where (a) follows from the fact that S_i is i.i.d. and (b) follows from the definition of U_i , which is

$$U_i \triangleq (M_{12}^s, S^{i-1}). \tag{94}$$

Now, consider

$$\begin{aligned}
nR_1 &= H(M_1) \\
&\stackrel{(a)}{=} H(M_1, M_{12}^m) \\
&= H(M_{12}^m) + H(M_1|M_{12}^m) \\
&\stackrel{(b)}{\leq} nC_{12}^m + H(M_1|M_{12}^m, M_2, S^n, M_{12}^s) \\
&\stackrel{(c)}{\leq} nC_{12}^m + I(M_1; Y^n|M_{12}^m, M_2, S^n, M_{12}^s) + n\epsilon_n
\end{aligned}$$

$$\begin{aligned}
&= nC_{12}^m + I(M_1, X_1^n; Y^n|M_{12}^m, M_2, S^n, X_2^n) + n\epsilon_n \\
&\leq nC_{12}^m + \sum_{i=1}^n I(X_{1,i}; Y_i|M_{12}^m, M_{12}^s, S^i, X_{2,i}) + n\epsilon_n \\
&\stackrel{(d)}{=} nC_{12}^m + \sum_{i=1}^n I(X_{1,i}; Y_i|V_i, U_i, S_i, X_{2,i}) + n\epsilon_n,
\end{aligned} \tag{95}$$

where (a) follows from the fact that M_{12}^m is a deterministic function of M_1 , (b) from the fact that S^n is independent of M_1 and M_{12}^s is a deterministic function of S^n , (c) from Fano's inequality and the definition of $\epsilon_n \triangleq (R_1 + R_2)P_e^{(n)}$. Step (d) follows from the definition of the auxiliary random variable

$$V_i \triangleq M_{12}^m. \tag{96}$$

Now using similar steps as above we obtain the following additional upper bounds

$$\begin{aligned}
nR_2 &= H(M_2) \\
&= H(M_2|M_1, M_{12}^m, S^n, M_{12}^s) \\
&\leq I(M_2; Y^n|M_1, M_{12}^m, S^n, M_{12}^s) + n\epsilon_n \\
&\leq \sum_{i=1}^n I(X_{2,i}; Y_i|X_{1,i}, V_i, U_i, S_i) + n\epsilon_n
\end{aligned} \tag{97}$$

and

$$\begin{aligned}
nR_1 + nR_2 &= H(M_1, M_2, M_{12}^m) \\
&= H(M_1, M_2, M_{12}^m|S^n) \\
&= H(M_{12}^m|S^n) + H(M_1, M_2|M_{12}^m, S^n, M_{12}^s) \\
&\leq nC_{12}^m + I(M_1, M_2; Y^n|M_{12}^m, S^n, M_{12}^s) + n\epsilon_n \\
&\leq nC_{12}^m + \sum_{i=1}^n I(X_{1,i}, X_{2,i}; Y_i|V_i, U_i, S_i) + n\epsilon_n,
\end{aligned} \tag{98}$$

and

$$\begin{aligned}
nR_1 + nR_2 &= H(M_1, M_2, M_{12}^m) \\
&= H(M_1, M_2, M_{12}^m|S^n) \\
&\leq I(M_1, M_2; Y^n|S^n, M_{12}^s) + n\epsilon_n \\
&\leq \sum_{i=1}^n I(X_{1,i}, X_{2,i}; Y_i|U_i, S_i) + n\epsilon_n
\end{aligned} \tag{99}$$

Now, we note that V_i is independent of (U_i, S_i) since M_1 is independent of S^n , and $X_2 - (U, V) - (X_1, S_1)$ is a Markov chain since $(M_2, M_{12}^m, M_{12}^s) - (S^{i-1}, M_{12}^m, M_{12}^s) - (M_1, S^n)$ holds. Finally, let Q be a random variable independent of (X^n, S^n, Y^n) , and uniformly distributed over the set $\{1, 2, 3, \dots, n\}$. Define the random variables $U \triangleq (Q, U_Q)$, $V \triangleq (Q, V_Q)$, and we obtain that the region given by (78)–(82) is an outer bound to the achievable region.

To show that the cardinality of the random variables U and V is bounded we follow similar steps as in Lemma 2, first for U and then for V . We note that the cardinality of auxiliary random variables U may be bounded by $|U| \leq |S| + 4$ and for auxiliary random variables V , we have $|V| \leq \min(|X_1||X_2||S||U| + 3, |Y||S||U| + 4)$.

Outline of Achievability Part: The achievability part is straightforward once we observe that we can generate a coordination $P_{U,S}$ with a rate $C_{12}^s > I(U; S)$ and then use a multiplexer where U^n , which is known to all encoders and to the decoder, is the control sequence of the multiplexer. For a given $U = u$ we obtain a MAC with cooperation C_{12}^m where the state is known to one encoder and to the decoder, hence the region

$$R_1 \leq I(X_1; Y|X_2, S, U = u, V) + C_{12}^m \quad (100)$$

$$R_2 \leq I(X_2; Y|X_1, S, U = u, V) \quad (101)$$

$$R_1 + R_2 \leq \min \left\{ \begin{array}{l} I(X_1, X_2; Y|S, U = u, V) + C_{12}^m \\ I(X_1, X_2; Y|S, U = u) \end{array} \right\}, \quad (102)$$

is achievable. Averaging over $P(u)$ results in the region given by (78)–(82). ■

ACKNOWLEDGMENT

The authors would like to thank the associate editor and the anonymous referees for their suggestions which helped improving the content of this paper. In addition, the authors would like to thank Yossef Steinberg for an enlightening discussion on the difference between deterministic and random binning which is mentioned in the paper.

REFERENCES

- [1] F. M. J. Willems, "Information-Theoretical Results for the Discrete Memoryless Multiple Access Channel," Ph.D., Katholieke Universiteit Leuven, Haverlee, Belgium, 1982.
- [2] F. M. J. Willems, "The discrete memoryless multiple channel with partially cooperating encoders," *IEEE Trans. Inf. Theory*, vol. 29, no. 6, pp. 441–445, 1983.
- [3] D. Slepian and J. K. Wolf, "A coding theorem for multiple-access channel with correlated sources," *Bell Syst. Tech. J.*, vol. 51, pp. 1037–1076, 1973.
- [4] P. Cuff, H. H. Permuter, and T. M. Cover, "Coordination capacity," *IEEE Trans. Inf. Theory*, Sep. 2009, submitted for publication.
- [5] R. Liu, I. Maric, P. Spasojevic, and R. D. Yates, "Discrete memoryless interference and broadcast channels with confidential messages: Secrecy capacity regions," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2493–2507, 2008.
- [6] R. Liu and V. Poor, "Secrecy capacity region of a multi-antenna Gaussian broadcast channel with confidential messages," *IEEE Trans. Inf. Theory*, vol. 55, no. 3, pp. 1235–1249, 2009.
- [7] O. Simeone, D. Gündüz, V. H. Poor, A. J. Goldsmith, and S. Shamai, "Compound multiple access channels with partial cooperation," *IEEE Trans. Inf. Theory*, vol. 55, no. 10, pp. 2425–2441, 2008.
- [8] S. I. Bross, A. Lapidoth, and M. A. Wigger, "The Gaussian MAC with conferencing encoders," in *Arxiv*, 2008, abs/0805.0516.
- [9] M. Prabhakaran, V. and P. Viswanath, "Interference channels with destination cooperation," *IEEE Trans. Inf. Theory*, 2009, submitted for publication.
- [10] V. M. Prabhakaran and P. Viswanath, "Interference channels with source cooperation," *IEEE Trans. Inf. Theory*, 2009, submitted for publication.
- [11] I.-H. Wang and D. N. C. Tse, "Interference mitigation through limited receiver cooperation," *IEEE Trans. Inf. Theory*, 2009, submitted for publication.
- [12] H. Bagheri, A. S. Motahari, and A. K. Khandani, "On the symmetric Gaussian interference channel with partial unidirectional cooperation," *IEEE Trans. on Wireless Communication*, 2009, submitted for publication.
- [13] C. T. K. Ng, N. Jindal, A. J. Goldsmith, and U. Mitra, "Capacity gain from two-transmitter and two-receiver cooperation," *IEEE Tran. on Wireless Communications*, vol. 53, no. 10, pp. 3822–3827, 2007.
- [14] I. Maric, R. D. Yates, and G. Kramer, "Capacity of interference channels with partial transmitter cooperation," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3536–3548, 2007.

- [15] I. Maric, A. J. Goldsmith, G. Kramer, and S. Shamai, "On the capacity of interference channels with one cooperating transmitter," *European Transactions on Telecommunications*, vol. 19, pp. 405–420, 2008.
- [16] R. Dabora and S. D. Servetto, "On the role of estimate-and-forward with time sharing in cooperative communication," *IEEE Trans. Inf. Theory*, vol. 54, no. 10, pp. 4409–4431, 2008.
- [17] D. Gündüz and E. Erkip, "Source and channel coding for cooperative relaying," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3454–3475, 2007.
- [18] L. Sankar, G. Kramer, and N. B. Mandayam, "Relay vs. user cooperation in time-duplexed multiaccess networks," *IEEE Trans. on Wireless Communication*, 2008, submitted for publication.
- [19] C. T. K. Ng and A. J. Goldsmith, "The impact of csi and power allocation on relay channel capacity and cooperation strategies," *IEEE Tran. on Wireless Communications*, vol. 7, no. 12–2, pp. 5380–5389, 2008.
- [20] O. Simeone, O. Somekh, H. V. Poor, and S. Shamai, "Local base station cooperation via finite-capacity links for the uplink of linear cellular networks," *IEEE Trans. Inf. Theory*, vol. 55, no. 1, pp. 190–204, 2009.
- [21] G. Kramer, I. Maric, and R. D. Yates, "Cooperative communications," *Foundations and Trends in Networking*, vol. 1, no. 3/4, 2006.
- [22] A. Haghi, R. Khosravi-Farsani, M. R. Aref, and F. Marvasti, "The capacity region of fading multiple access channels with cooperative encoders and partial csit," *ISIT10*, 2010, submitted for publication.
- [23] A. Haghi, R. Khosravi-Farsani, M. R. Aref, and F. Marvasti, "The capacity region of p -transmitter/ q -receiver multiple access channels with common information," *IEEE Trans. on Wireless Communication*, 2010, submitted for publication.
- [24] C. Heegard and A. A. El Gamal, "On the capacity of computer memory with defects," *IEEE Trans. Inf. Theory*, vol. 29, no. 5, pp. 731–739, 1983.
- [25] Y. Cermal and Y. Steinberg, "Coding problems for channels with partial state information at the transmitter," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4521–4536, 2007.
- [26] Y. Cermal and Y. Steinberg, "The multiple-access channel with partial state information at the encoders," *IEEE Trans. Inf. Theory*, vol. 51, no. 11, pp. 3992–4003, 2005.
- [27] A. El Gamal and Y.-H. Kim, "Lecture notes on network information theory," *Arxiv*, 2010.
- [28] S. I. Gel'fand and M. S. Pinsker, "Coding for channel with random parameters," *Probl. Contr. and Inf. Theory*, vol. 9, no. 1, pp. 19–31, 1980.
- [29] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. New York: Academic, 1981.
- [30] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New-York: Wiley, 2006.
- [31] G. Kramer, "Topics in multi-user information theory," *Foundations and Trends in Communications and Information Theory*, vol. 4, no. 4/5, pp. 265–444, 2007.
- [32] A. D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inf. Theory*, vol. 22, no. 1, pp. 1–10, 1976.
- [33] A. H. Kaspi, "Two-way source coding with a fidelity criterion," *IEEE Trans. Inf. Theory*, vol. 31, no. 6, pp. 735–740, 1985.
- [34] T. Berger, "Multiterminal source coding," in *Information Theory Approach to Communications*, G. Longo, Ed. : CISM Course and Lecture, 1978, pp. 171–231.
- [35] H. H. Permuter, Y. Steinberg, and T. Weissman, "Two-way source coding with a helper," *IEEE Trans. Inf. Theory*, Apr. 2009.
- [36] J. Pearl, *Causality: Models, Reasoning and Inference*. : Cambridge Univ. Press, 2000.
- [37] G. Kramer, "Capacity results for the discrete memoryless network," *IEEE Trans. Inf. Theory*, vol. IT-49, pp. 4–21, 2003.
- [38] M. I. Jordan, *Learning in Graphical Models*. Cambridge, MA: MIT Press, 1999.

Haim H. Permuter (M'08) received his B.Sc. (*summa cum laude*) and M.Sc. (*summa cum laude*) degree in Electrical and Computer Engineering from the Ben-Gurion University, Israel, in 1997 and 2003, respectively, and Ph.D. degrees in Electrical Engineering from Stanford University, California in 2008.

Between 1997 and 2004, he was an officer at a research and development unit of the Israeli Defense Forces. He is currently a senior lecturer at Ben-Gurion University of the Israel.

Dr. Permuter is a recipient of the Fullbright Fellowship, the Stanford Graduate Fellowship (SGF), Allon Fellowship, and the 2009 U.S.-Israel Binational Science Foundation Bergmann Memorial Award.

Shlom (Shitz) Shamai received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from the Technion—Israel Institute of Technology, in 1975, 1981, and 1986 respectively.

During 1975–1985 he was with the Communications Research Labs in the capacity of a Senior Research Engineer. Since 1986 he is with the Department of Electrical Engineering, Technion—Israel Institute of Technology, where he is now the William Fondiller Professor of Telecommunications. His research interests encompasses a wide spectrum of topics in information theory and statistical communications.

Dr. Shamai is an IEEE Fellow, and the recipient of the 2011 Claude E. Shannon Award. He is the recipient of the 1999 van der Pol Gold Medal of the Union Radio Scientifique Internationale (URSI), and a co-recipient of the 2000 IEEE Donald G. Fink Prize Paper Award, the 2003, and the 2004 joint IT/COM societies paper award, the 2007 IEEE Information Theory Society Paper Award, the 2009 European Commission FP7, Network of Excellence in Wireless COMMunications (NEWCOM++) Best Paper Award, and the 2010 Thomson Reuters Award for International Excellence in Scientific Research. He is also the recipient of 1985 Alon Grant for distinguished young scientists and the 2000 Technion Henry Taub Prize for Excellence in Research. He has served as Associate Editor for the Shannon Theory of the IEEE TRANSACTIONS ON INFORMATION THEORY, and has also served on the Board of Governors of the Information Theory Society.

Anelia Somekh-Baruch (S'01–M'03) received the B.Sc. degree from Tel-Aviv University, Tel-Aviv, Israel, in 1996 and the M.Sc. and Ph.D. degrees from the Technion-Israel Institute of Technology, Haifa, Israel, in 1999 and 2003, respectively, all in electrical engineering. During 2003–2004, she was with the Technion Electrical Engineering Department. During 2005–2008, she was a Visiting Research Associate at the Electrical Engineering Department, Princeton University, Princeton, NJ. From 2008 to 2009 she was a researcher at the Electrical Engineering Department, Technion, and from 2009 she has been with the Bar-Ilan University School of Engineering, Ramat-Gan, Israel. Her research interests include topics in information theory and communication theory. Dr. Somekh-Baruch received the Tel-Aviv University program for outstanding B.Sc. students scholarship, the Viterbi scholarship, the Rothschild foundation scholarship for postdoctoral studies, and the Marie Curie Outgoing International Fellowship.