Asymmetric Cooperative Multiple Access Channels with Delayed CSI

Abdellatif Zaidi † Shlomo Shamai (Shitz) ‡

† Université Paris-Est Marne La Vallée, Champs-sur-Marne 77454, France
‡ Department of EE, Technion-Israel Institute of Technology, Haifa, Israel
abdellatif.zaidi@univ-mlv.fr, sshlomo@ee.technion.ac.il

Abstract— **We consider a two-user multiaccess channel with degraded messages sets in which the channel state information (CSI) is revealed, strictly causally or with one-unit delay, to only the encoder that sends the common message. We study the capacity region of this model. We establish inner and outer bounds on the capacity region. We also identify some special cases in which the bounds meet, thereby characterizing the capacity region in these cases. The outer bound is non-trivial and has a relatively simple convenient expression (it incorporates *only one auxiliary random variable*). The coding scheme that we use for the inner bound utilizes rate-splitting to resolve a tension at the informed encoder among exploiting the knowledge of the (delayed) CSI (through a noisy network coding or quantize-map-and-forward state compression) and sending information cooperatively with the other encoder. Together with some previous results on closely related models, the results in this paper shed more light on the utility of delayed CSI for increasing the capacity region of multiaccess channels; and tie with some recent progress in this framework.**

I. Introduction

In this work, we study a two-user multiaccess channel (MAC) in which both encoders transmit a common message and one of the encoders also transmits an individual message. That is, a MAC with degraded messages sets or cooperative MAC. We assume that the channel state information (CSI) is known, strictly causally or with one-unit delay, to only the encoder that transmits only the common message. More precisely, let W_c and W_1 denote the common message and the individual message to be transmitted in, say, n uses of the channel; and $S^n = (S_1, \ldots, S_n)$ denote the CSI sequence during the transmission. At time i, Encoder 2 knows the channel states only up to time $i-1$, i.e., the sequence $S^{i-1} = (S_1, \ldots, S_{i-1})$, and both Encoder 1 and the decoder do not the CSI at all. We refer to this model as "asymmetric cooperative MAC with delayed CSI". From a practical viewpoint, the state sequence may model an interfering signal that is overheard, and estimated with high precision, by only Encoder 2, for example due to proximity.

We study the capacity region of this network model. We establish inner and outer bounds on the capacity region. The outer bound is non trivial, and has the advantage of having a relatively simple form that incorporates directly the channel inputs X_1 and X_2 and *only one auxiliary random variable*.

The inner bound is based on a coding scheme in which the encoder that knows the delayed CSI sends a compressed version of the CSI to the receiver, in addition to the cooperative information that it sends collaboratively with the other encoder. For this model, because the encoder that transmits both messages does not know the channel states, there is a *dilemma* among 1) exploiting the knowledge of these states by the other encoder, and 2) sending information cooperatively (i.e., the common message). The coding scheme of the inner bound resolves this tension by splitting the common message into two independent parts, one that is sent cooperatively by the two encoders, and one that is sent only by the encoder that knows the delayed CSI. The CSI compression is performed à-la noisy network coding by Lim, Kim, El Gamal and Chung [1] or the quantize-map-and-forward by Avestimeher, Diggavi and Tse [2], i.e., with no binning.

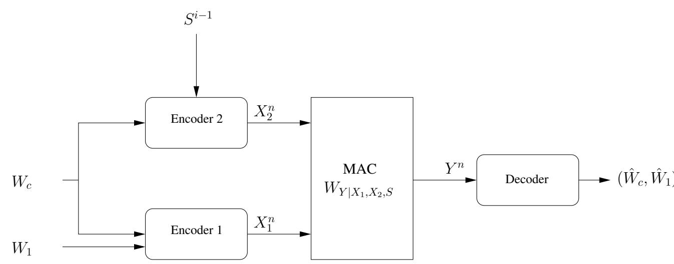

Fig. 1. Asymmetric cooperative Multiple Access Channel with delayed CSI.

Furthermore, we also identify some special cases in which the inner and outer bounds meet, and so characterize the capacity region in these cases. Throughout, the results are illustrated through some insightful examples.

This work tie with recent and very recent efforts [3]–[11] aiming at better understanding the utility of delayed CSI in increasing the capacity region of multiaccess channels. Other related models are treated in [12]–[15].

II. Problem Setup

We consider a stationary memoryless two-user MAC $W_{Y|X_1,X_2,S}$ whose output $Y \in \mathcal{Y}$ is controlled by the channel inputs $X_1 \in \mathcal{X}_1$ and $X_2 \in \mathcal{X}_2$ from the encoders and a channel state $S \in \mathcal{S}$ which is drawn according to a memoryless probability law Q_S. Both encoders transmit a common message W_c; and one of the encoders, Encoder 1, also transmits an independent individual message W_1. The channel state is revealed, strictly causally, to only the encoder that sends only the common message. That is, at time i Encoder 2 knows the values of the state sequence up to time $i-1$, i.e.,

$S^{i-1} = (S_1, \ldots, S_{i-1})$; and both Encoder 1 and the decoder do not know these states at all.

We assume that the common message W_c and the individual message W_1 are independent random variables drawn uniformly from the sets $\mathcal{W}_c = \{1, \cdots, M_c\}$ and $\mathcal{W}_1 = \{1, \cdots, M_1\}$, respectively. The sequences X_1^n and X_2^n from the encoders are sent across the state-dependent MAC modeled as a memoryless conditional probability distribution $W_{Y|X_1,X_2,S}$. The laws governing the state sequence and the output letters are given by

$$W^n_{Y|X_1,X_2,S}(y^n|x_1^n, x_2^n, s^n) = \prod_{i=1}^{n} W_{Y|X_1,X_2,S}(y_i|x_{1i}, x_{2i}, s_i) \quad (1)$$

$$Q_S^n(s^n) = \prod_{i=1}^{n} Q_S(s_i). \quad (2)$$

The receiver guesses the pair $(\hat{W}_c, \hat{W}_1)$ from the channel output Y^n.

***Definition** 1:* For positive integers n, M_c and M_1, an (M_c, M_1, n, ϵ) code for the multiple access channel with asymmetric delayed states consists of a mapping

$$\phi_1^n : \mathcal{W}_c \times \mathcal{W}_1 \longrightarrow \mathcal{X}_1^n, \quad (3)$$

at Encoder 1, a sequence of mappings

$$\phi_{2,i} : \mathcal{W}_c \times \mathcal{S}^{i-1} \longrightarrow \mathcal{X}_2, \quad i = 1, \ldots, n \quad (4)$$

at Encoder 2, and a decoder map

$$\psi : \mathcal{Y}^n \longrightarrow \mathcal{W}_c \times \mathcal{W}_1 \quad (5)$$

such that the average probability of error is bounded by ϵ,

$$P_e^n = \mathbb{E}_S\Big[\Pr\Big(\psi(Y^n) \neq (W_c, W_1)|S^n = s^n\Big)\Big] \leq \epsilon. \quad (6)$$

The rate of the common message and the rate of the individual message are defined as

$$R_c = \frac{1}{n}\log M_c \qquad \text{and} \qquad R_1 = \frac{1}{n}\log M_1,$$

respectively. A rate pair (R_c, R_1) is said to be achievable if for every $\epsilon > 0$ there exists an $(2^{nR_c}, 2^{nR_1}, n, \epsilon)$ code for the channel $W_{Y|X_1,X_2,S}$. The capacity region $\mathcal{C}_{\text{asym,s-c}}$ of the state-dependent MAC with asymmetric delayed states is defined as the closure of the set of achievable rate pairs.

Due to space limitation, the results of this paper are either outlined only or mentioned without proofs. Detailed proofs can be found in [4].

III. Bounds on the Capacity Region

In this section, it is assumed that the alphabets $\mathcal{S}, \mathcal{X}_1, \mathcal{X}_2$ are finite.

A. Outer Bound on the Capacity Region

Let $\mathcal{P}^{\text{out}}_{\text{asym,s-c}}$ stand for the collection of all random variables (S, V, X_1, X_2, Y) such that V, X_1 and X_2 take values in finite alphabets $\mathcal{V}$, $\mathcal{X}_1$ and $\mathcal{X}_2$, respectively, and satisfy

$$\begin{aligned} &P_{S,V,X_1,X_2,Y}(s, v, x_1, x_2, y) \\ &\quad = P_{S,V,X_1X_2}(s, v, x_1, x_2) W_{Y|X_1,X_2,S}(y|x_1, x_2, s) &(7a)\\ &P_{S,V,X_1,X_2}(s, v, x_1, x_2) = Q_S(s) P_{X_2}(x_2) P_{X_1|X_2}(x_1|x_2) \\ &\quad \cdot P_{V|S,X_1,X_2}(v|s, x_1, x_2) &(7b)\\ &\sum_{v,x_1,x_2} P_{S,V,X_1,X_2}(s, v, x_1, x_2) = Q_S(s) &(7c)\end{aligned}$$

and

$$0 \leq I(V, X_2; Y) - I(V, X_2; S). \quad (8)$$

The relations in (7) imply that $V \leftrightarrow (S, X_1, X_2) \leftrightarrow Y$ is a Markov chain, and X_1 and X_2 are independent of S.

Define $\mathcal{R}^{\text{out}}_{\text{asym,s-c}}$ to be the set of all rate pairs (R_c, R_1) such that

$$\begin{aligned} R_1 &\leq I(X_1; Y|V, X_2) \\ R_c + R_1 &\leq I(V, X_1, X_2; Y) - I(V, X_1, X_2; S) \\ &\qquad \text{for some } (S, V, X_1, X_2, Y) \in \mathcal{P}^{\text{out}}_{\text{asym,s-c}}. \quad (9)\end{aligned}$$

As stated in the following theorem, the set $\mathcal{R}^{\text{out}}_{\text{asym,s-c}}$ is an outer bound on the capacity region of the cooperative state-dependent discrete memoryless MAC with delayed CSI states at the encoder that sends only the common message.

***Theorem** 1:* The capacity region of the cooperative multiple access channel with delayed CSI known only at the encoder that sends only the common message satisfies

$$\mathcal{C}_{\text{asym,s-c}} \subseteq \mathcal{R}^{\text{out}}_{\text{asym,s-c}}. \quad (10)$$

We now state a proposition that provides an alternative outer bound on the capacity region of the cooperative multiaccess channel with delayed CSI that we study. This proposition will turn out to be useful in Section IV.

Let $\breve{\mathcal{R}}^{\text{out}}_{\text{asym,s-c}}$ be the set of all rate pairs (R_c, R_1) satisfying

$$\begin{aligned} R_1 &\leq I(X_1; Y|X_2, S) \\ R_c + R_1 &\leq I(X_1, X_2; Y) \end{aligned} \quad (11)$$

for some measure

$$P_{S,X_1,X_2,Y} = Q_S P_{X_1,X_2} W_{Y|S,X_1,X_2}. \quad (12)$$

***Proposition** 1:* The capacity region $\mathcal{C}_{\text{asym,s-c}}$ of the cooperative multiple access channel with delayed CSI known only at the encoder that sends only the common message satisfies

$$\mathcal{C}_{\text{asym,s-c}} \subseteq \breve{\mathcal{R}}^{\text{out}}_{\text{asym,s-c}}. \quad (13)$$

The bound on the sum rate of Theorem 1 is at least at tight as that of Proposition 1. This can be seen through the following inequalities.

$$\begin{aligned} &I(V, X_1, X_2; Y) - I(V, X_1, X_2; S) \\ &= I(X_1, X_2; Y) + I(V; Y|X_1, X_2) - I(V; S|X_1, X_2) &(14)\\ &= I(X_1, X_2; Y) + I(V; Y|S, X_1, X_2) - I(V; S|X_1, X_2, Y) &(15)\\ &= I(X_1, X_2; Y) - I(V; S|X_1, X_2, Y) + H(Y|S, X_1, X_2) \\ &\quad - H(Y|V, S, X_1, X_2) &(16)\\ &= I(X_1, X_2; Y) - I(V; S|X_1, X_2, Y) + H(Y|S, X_1, X_2) \\ &\quad - H(Y|S, X_1, X_2) &(17)\\ &= I(X_1, X_2; Y) - I(V; S|X_1, X_2, Y) &(18)\\ &\leq I(X_1, X_2; Y) &(19)\end{aligned}$$

where: (14) follows since X_1 and X_2 are independent of the state S; (17) follows since for all random variables A, B and C, we have $I(A; B) - I(A; C) = I(A; B|C) - I(A; C|B)$; and (17) follows since $V \leftrightarrow (S, X_1, X_2) \leftrightarrow Y$ is a Markov chain.

For some channels, the outer bound of Theorem 1 is strictly contained in the outer bound of Proposition 1, i.e.,

$$\mathcal{R}^{\text{out}}_{\text{asym,s-c}} \subsetneq \breve{\mathcal{R}}^{\text{out}}_{\text{asym,s-c}}. \tag{20}$$

The following example, shows this.

Example 1: Consider the following discrete memoryless channel, considered initially in [7],

$$Y = X_S \tag{21}$$

where $\mathcal{X}_1 = \mathcal{X}_2 = \mathcal{Y} = \{0, 1\}$, and the state S is uniformly distributed over the set $\mathcal{S} = \{1, 2\}$ and acts as a random switch that connects a randomly chosen transmitter to the output. For this channel, the rate-pair $(R_c, R_1) = (1/2, 1/2)$ is in the outer bound of Proposition 1, but not in that of Theorem 1, i.e., $(1/2, 1/2) \in \breve{\mathcal{R}}^{\text{out}}_{\text{asym,s-c}}$ and $(1/2, 1/2) \notin \mathcal{R}^{\text{out}}_{\text{asym,s-c}}$.

B. Inner Bound on the Capacity Region

Let $\mathcal{P}^{\text{in}}_{\text{asym,s-c}}$ stand for the collection of all random variables (S, U, V, X_1, X_2, Y) such that U, V, X_1 and X_2 take values in finite alphabets $\mathcal{U}$, $\mathcal{V}$, $\mathcal{X}_1$ and $\mathcal{X}_2$, respectively, and satisfy

$$\begin{aligned} &P_{S,U,V,X_1,X_2,Y}(s, u, v, x_1, x_2, y) \\ &\quad = P_{S,U,V,X_1,X_2}(s, u, v, x_1, x_2)W_{Y|X_1,X_2,S}(y|x_1, x_2, s) && \text{(22a)} \\ &P_{S,U,V,X_1,X_2}(s, u, v, x_1, x_2) \\ &\quad = Q_S(s)P_U P_{X_2|U}(x_2|u)P_{X_1|U}(x_1|u)P_{V|S,U,X_2}(v|s, u, x_2) && \text{(22b)} \\ &\sum_{u,v,x_1,x_2} P_{S,U,V,X_1,X_2}(s, u, v, x_1, x_2) = Q_S(s) && \text{(22c)} \end{aligned}$$

The relations in (22) imply that $(U, V) \leftrightarrow (S, X_1, X_2) \leftrightarrow Y$, $X_1 \leftrightarrow U \leftrightarrow X_2$ and $(U, X_1) \leftrightarrow (V, X_2) \leftrightarrow S$ are Markov chains; and X_1 and X_2 are independent of S.

Define $\mathcal{R}^{\text{in}}_{\text{asym,s-c}}$ to be the set of all rate pairs (R_c, R_1) such that

$$\begin{aligned} R_1 &\leq I(X_1; Y|U, V, X_2) \\ R_1 &\leq I(V, X_1, X_2; Y|U) - I(V; S|U, X_2) \\ R_c + R_1 &\leq I(U, V, X_1, X_2; Y) - I(V; S|U, X_2) \\ &\qquad \text{for some } (S, U, V, X_1, X_2, Y) \in \mathcal{P}^{\text{in}}_{\text{asym,s-c}}. \end{aligned} \tag{23}$$

As stated in the following theorem, the set $\mathcal{R}^{\text{in}}_{\text{asym,s-c}}$ is an inner bound on the capacity region of the state-dependent discrete memoryless cooperative MAC with delayed CSI known only at the encoder that sends only the common message.

Theorem 2: The capacity region of the cooperative multiple access channel with delayed CSI revealed only to the encoder that sends only the common message satisfies

$$\mathcal{R}^{\text{in}}_{\text{asym,s-c}} \subseteq \mathcal{C}_{\text{asym,s-c}}. \tag{24}$$

Proof: A description of the coding scheme that we use for the proof of Theorem 2 is given in Section V-A.

The following remark helps better understanding the coding scheme that we use for the proof of Theorem 2.

Remark 1: For the model that we study, a good codebook at the encoder that sends only the common message should resolve a *dilemma* among 1) exploiting the knowledge of the state that is available at this encoder and 2) sending information cooperatively with the other encoder (i.e., the common message). The coding scheme of Theorem 2 resolves this tension by splitting the common rate R_c into two parts. More specifically, the common message W_c is divided into two parts, $W = (W_{c1}, W_{c2})$. The part W_{c1} is sent cooperatively by the two encoders, at rate R_{c1}; and the part W_{c2} is sent only by the encoder that exploits the available state, at rate R_{c2}. The total rate for the common message is $R_c = R_{c1} + R_{c2}$. In Theorem 2, the random variable U stands for the information that is sent cooperatively by the two encoders, and the random variable V stands for the compression of the state by the encoder that sends only the common message, in a manner that is described in more details in Section V-A. ∎

IV. Capacity Results

Consider the following class of discrete memoryless channels, which we denote as $\mathcal{D}_{\text{IH}}$. Encoder 1 does not know the state sequence at all, and transmits an individual message $W_1 \in [1, 2^{nR_1}]$. Encoder 2 knows the state sequence strictly causally, and does not transmits any message. In this model, Encoder 2 plays the role of a helper that is informed of the channel state sequence only strictly causally. This network may model one in which there is an external node that interferes with the transmission from Encoder 1 to the destination, and that is overheard only by Encoder 2 which then assists the destination by providing some information about the interference. Furthermore, we assume that the state S an be obtained as a deterministic function of the inputs X_1, X_2 and the channel output Y, as

$$S = f(X_1, X_2, Y). \tag{25}$$

For channels with a helper that knows the states strictly causally, the class of channels $\mathcal{D}_{\text{IH}}$ is larger than that considered in [9], as the channel output needs not be a deterministic function of the channel inputs and the state. The following theorem characterizes the capacity region for the class of channels $\mathcal{D}_{\text{IH}}$.

The capacity of the class of channels $\mathcal{D}_{\text{IH}}$ can be characterized as follows.

Theorem 3: For any channel in the class $\mathcal{D}_{\text{IH}}$ defined above, the capacity $C_{\text{s-c}}$ is given by

$$C_{\text{s-c}} = \max \Big\{ I(X_1; Y|S, X_2),\ I(X_1, X_2; Y) \Big\} \tag{26}$$

where the maximization is over measures of the form

$$P_{S,X_1,X_2,Y} = Q_S P_{X_1} P_{X_2} W_{Y|S,X_1,X_2}. \tag{27}$$

Remark 2: The class $\mathcal{D}_{\text{IH}}$ includes the Gaussian model $Y = X_1 + X_2 + S$ where the state $S \sim \mathcal{N}(0, Q)$ comprises the channel noise, and the inputs are subjected to the input power constraints $(1/n)\sum_{i=1}^n \mathbb{E}[X_{k,i}^2] \leq P_k$, $k = 1, 2$. Encoder 1 does not know the state sequence and transmits message W_1. Encoder 2 knows the state sequence strictly causally, and does not transmit any message. The capacity of this model is given by

$$C^{\text{G}}_{\text{s-c}} = \frac{1}{2}\log(1 + \frac{P_1 + P_2}{Q}). \tag{28}$$

The capacity (31) can be obtained from Theorem 3 by maximizing the two terms of the minimization utilizing the *Maximum Differential Entropy Lemma* [16, Section 2.2]. Observe that the first term of the minimization in (26) is redundant in this case. Also, we note that the capacity (31) of this example can also be obtained as a special case of that of the Gaussian example considered in [9, Remark 4]. ∎

In the following example the channel output can *not* be obtained as a deterministic function of the channel inputs and the channel state, and yet, its capacity can be characterized using Theorem 3.

Example 2: Consider the following Gaussian example with $Y = (Y_1, Y_2)$, and

$$Y_1 = X_1 + X_2 + S \tag{29}$$

$$Y_2 = X_2 + Z \tag{30}$$

where the state process is memoryless Gaussian, with $S \sim \mathcal{N}(0, Q)$, and the noise process is memoryless Gaussian independent of all other processes, $Z \sim \mathcal{N}(0, N)$. Encoder 1 does not know the state sequence, and transmits message $W_1 \in [1, 2^{nR_1}]$. Encoder 2 knows the state strictly causally, and does not transmit any message. The inputs are subjected to the input power constraints $\sum_{i=1}^{n} \mathbb{E}[X_{1,i}^2] \leq nP_1$ and $\sum_{i=1}^{n} \mathbb{E}[X_{2,i}^2] \leq nP_2$. The capacity of this model can be computed easily using Theorem 3, as

$$C_{\text{s-c}}^{\text{G}} = \frac{1}{2}\log\Big(1 + \frac{P_1}{Q} + \frac{P_2}{Q}\frac{N}{P_2+N}\Big) + \frac{1}{2}\log\Big(1 + \frac{P_2}{N}\Big). \tag{31}$$

Note that the knowledge of the states strictly causally at Encoder 2 makes it possible to send at positive rates by Encoder 1 even if the allowed average power P_1 is zero.

Example 3: Consider the following binary example in which the state models fading. The channel output has two components, i.e., $Y = (Y_1, Y_2)$, with

$$Y_1 = S{\cdot}X_1 \tag{32a}$$

$$Y_2 = X_2 + Z \tag{32b}$$

where $\mathcal{X}_1 = \mathcal{X}_2 = \mathcal{S} = \mathcal{Z} = \{+1, -1\}$, and the noise Z is independent of (S, X_1, X_2) with $\Pr\{Z = 1\} = p$ and $\Pr\{Z = -1\} = 1 - p$, $0 \leq p \leq 1$, and the state S, known strictly causally to only Encoder 2, is such that $\Pr\{S = 1\} = \Pr\{S = -1\} = 1/2$. Using Theorem 3, it is easy to compute the capacity of this example, as

$$C_{\text{s-c}}^{\text{B}} = \max_{0 \leq q_1, q_2 \leq 1} \min \Big\{h_2(q_1), g(p, q_2) - h_2(p)\Big\} \tag{33}$$

where

$$\begin{aligned} g(p, q_2) = &-pq_2 \log(pq_2) - (1-p)(1-q_2)\log((1-p)(1-q_2)) \\ &- p * q_2 \log(p * q_2). \end{aligned} \tag{34}$$

Observe that $C_{\text{s-c}}^{\text{B}} \geq 1 - \frac{1}{2}h_2(p) \geq 0.5$.

Proof: Using (32), we have $S = Y_1/X_1$, and, so, S is a deterministic function of (X_1, X_2, Y). Thus, the capacity of this channel can be computed using Theorem 3. Let $0 \leq q_1 \leq 1$ such that $\Pr\{X_1 = 1\} = q_1$ and $\Pr\{X_1 = -1\} = 1 - q_1$. Also, let $0 \leq q_2 \leq 1$ such that $\Pr\{X_2 = 1\} = q_2$ and $\Pr\{X_2 = -1\} = 1 - q_2$. Then, considering the first term on the RHS of (26), we get

$$\begin{aligned} I(X_1; Y|S, X_2) &= H(Y|S, X_2) - H(Y|S, X_1, X_2) && (35)\\ &= H(SX_1, X_2 + Z|S, X_2) - H(Z|S, X_1, X_2) && (36)\\ &= H(X_1, Z|S, X_2) - H(Z) && (37)\\ &= H(X_1, Z) - H(Z) && (38)\\ &= H(X_1) && (39)\\ &= h_2(q_1) && (40) \end{aligned}$$

where (37) holds since Z is independent of (S, X_1, X_2), (38) holds since (X_1, Z) is independent of (S, X_2), and (39) holds since X_1 and Z are independent.

Similarly, considering the second term on the RHS of (26), we get

$$\begin{aligned} I(X_1, X_2; Y) &= H(Y) - H(Y|X_1, X_2) && (41)\\ &= H(Y) - (SX_1, Z|X_1, X_2) && (42)\\ &= H(Y) - H(Z) - H(S) && (43)\\ &= H(SX_1) + H(X_2 + Z) - H(Z) - H(S) && (44)\\ &= H(X_2 + Z) - H(Z) && (45)\\ &= g(p, q_2) - h_2(p) && (46) \end{aligned}$$

where (43) holds since S and Z are independent of (X_1, X_2) and independent of each other, (44) holds since $Y_1 = SX_1$ and $Y_2 = X_2 + Z$ are independent, (45) follows because

$$\Pr\{SX_1 = 1\} = \Pr\{SX_1 = -1\} = \frac{1}{2} \tag{47}$$

and, so, $H(SX_1) = 1 = H(S)$, and (46) follows because

$$\begin{aligned} &\Pr\{X_2 + Z = 0\} = p * q_2, \qquad \Pr\{X_2 + Z = 2\} = pq_2 \\ &\Pr\{X_2 + Z = -2\} = (1-p)(1-q_2) \end{aligned} \tag{48}$$

and, so, $H(X_2 + Z) = g(p, q_2)$ as given by (34). ∎

Remark 3: The result of Theorem 3 can be extended to the case in which the encoders send separate messages and each observes (strictly causally) an independent state. In this case, denoting by S_1 the state that is observed by Encoder 1 and by S_2 the state that is observed by Encoder 2, it can be shown that, if both S_1 and S_2 can be obtained as deterministic functions of the inputs X_1 and X_2 and the channel output Y, then the capacity region is given by the set of all rates satisfying

$$R_1 \leq I(X_1; Y|X_2, S_2) \tag{49a}$$

$$R_2 \leq I(X_2; Y|X_1, S_1) \tag{49b}$$

$$R_1 + R_2 \leq I(X_1, X_2; Y) \tag{49c}$$

for some measure of the form $Q_{S_1,S_2,X_1,X_2} = Q_{S_1}Q_{S_2}P_{X_1}P_{X_2}$. This result can also be obtained by noticing that, if both S_1 and S_2 are deterministic functions of (X_1, X_2, Y), then the inner bound of [9, Theorem 2] reduces to (49), which is also an outer bound as stated in [8, Proposition 3].

V. Proofs

A. Proof of Theorem 3

The transmission takes place in B blocks. The common message W_c and the individual message W_1 are sent over *all* blocks. We thus have $B_{W_c} = nBR_c$, $B_{W_1} = nBR_1$, $N = nB$, $R_{W_c} = B_{W_c}/N = R_c$ and $R_{W_1} = B_{W_1}/N = R_1$, where B_{W_c} is the number of common message bits, B_{W_1} is the number of individual message bits, N is the number of channel uses and R_{W_c} and R_{W_1} are the overall rates of the common and individual messages, respectively.

Codebook Generation: Fix a measure $P_{S,U,V,X_1,X_2,Y} \in \mathcal{P}^{\text{in}}_{\text{asym,s-c}}$. Fix $\epsilon > 0$, $\eta_c > 0$, $\eta_1 > 0$, $\hat{\eta} > 0$, $\delta > 1$ and denote $M_c = 2^{nB[R_c - \eta_c\epsilon]}$, $M_1 = 2^{nB[R_1 - \eta_1\epsilon]}$, and $\hat{M} = 2^{n[\hat{R} + \hat{\eta}\epsilon]}$. Also, let $\eta_{c1} > 0$, $\eta_{c2} > 0$ and $M_{c1} = 2^{nB[R_{c1} - \eta_{c1}\epsilon]}$ and $M_{c2} = 2^{nB[R_{c2} - \eta_{c2}\epsilon]}$ such that $R_c = R_{c1} + R_{c2}$.

We randomly and independently generate a codebook for each block.

1) For each block i, $i = 1, \ldots, B$, we generate M_{c1} independent and identically distributed (i.i.d.) codewords $\mathbf{u}_i(w_{c1})$ indexed by $w_{c1} = 1, \ldots, M_{c1}$, each with i.i.d. components drawn according to P_U.
2) For each block i, for each codeword $\mathbf{u}_i(w_{c1})$, we generate $M_{c2}\hat{M}$ independent and identically distributed (i.i.d.) codewords $\mathbf{x}_{2,i}(w_{c1}, w_{c2}, t_i')$ indexed by $w_{c2} = 1, \ldots, M_{c2}$, $t_i' = 1, \ldots, \hat{M}$, each with i.i.d. components drawn according to $P_{X_2|U}$.
3) For each block i, for each pair of codewords $(\mathbf{u}_i(w_{c1}), \mathbf{x}_{2,i}(w_{c1}, w_{c2}, t_i'))$, we generate $\hat{M}$ i.i.d. codewords $\mathbf{v}_i(w_{c1}, w_{c2}, t_i', t_i)$ indexed by $t_i = 1, \ldots, \hat{M}$, each with i.i.d. components drawn according to $P_{V|U,X_2}$.
4) For each block i, for each codeword $\mathbf{u}_i(w_{c1})$, we generate M_1 independent and identically distributed (i.i.d.) codewords $\mathbf{x}_{1,i}(w_{c1}, w_1)$ indexed by $w_1 = 1, \ldots, M_1$, each with i.i.d. components drawn according to $P_{X_1|U}$.

Encoding: Suppose that a common message $W_c = w_c = (w_{c1}, w_{c2})$ and an individual message $W_1 = w_1$ are to be transmitted. As we mentioned previously, w_c and w_1 will be sent over *all* blocks. We denote by $\mathbf{s}[i]$ the state affecting the channel in block i, $i = 1, \ldots, B$. For convenience, we let $\mathbf{s}[0] = \emptyset$ and $t_{-1} = t_0 = 1$ (a default value). The encoding at the beginning of block i, $i = 1, \ldots, B$, is as follows.

Encoder 2, which has learned the state sequence $\mathbf{s}[i-1]$, knows t_{i-2} and looks for a compression index $t_{i-1} \in [1 : \hat{M}]$ such that $\mathbf{v}_{i-1}(w_{c1}, w_{c2}, t_{i-2}, t_{i-1})$ is strongly jointly typical with $\mathbf{s}[i-1]$ and $\mathbf{x}_{2,i-1}(w_{c1}, w_{c2}, t_{i-2})$. If there is no such index or the observed state $\mathbf{s}[i-1]$ is not typical, t_{i-1} is set to 1 and an error is declared. If there is more than one such index t_{i-1}, choose the smallest. It can be shown that the error in this step has vanishing probability as long as n and B are large and

$$\hat{R} > I(V; S|U, X_2). \tag{50}$$

Encoder 2 then transmits the vector $\mathbf{x}_{2,i}(w_{c1}, w_{c2}, t_{i-1})$. Encoder 1 transmits the vector $\mathbf{x}_{1,i}(w_{c1}, w_1)$.

Decoding: At the end of the transmission, the decoder has collected all the blocks of channel outputs $\mathbf{y}[1], \ldots, \mathbf{y}[B]$.

Step (a): The decoder estimates message $w_c = (w_{c1}, w_{c2})$ using all blocks $i = 1, \ldots, B$, i.e., simultaneous decoding. It declares that $\hat{w}_c = (\hat{w}_{c1}, \hat{w}_{c2})$ is sent if there exist $t^B = (t_1, \ldots, t_B) \in [1 : \hat{M}]^B$ and $w_1 \in [1 : M_1]$ such that $\mathbf{u}_i(\hat{w}_{c1})$, $\mathbf{x}_{2,i}(\hat{w}_{c1}, \hat{w}_{c2}, t_{i-1})$, $\mathbf{v}_i(\hat{w}_{c1}, \hat{w}_{c2}, t_{i-1}, t_i)$, $\mathbf{x}_{1,i}(\hat{w}_{c1}, w_1)$ and $\mathbf{y}[i]$ are jointly typical for all $i = 1, \ldots, B$. One can show that the decoder obtains the correct $w_c = (w_{c1}, w_{c2})$ as long as n and B are large and

$$R_{c2} + R_1 \leq I(V, X_1, X_2; Y|U) - \hat{R} \tag{51a}$$
$$R_{c1} + R_{c2} + R_1 \leq I(U, V, X_1, X_2; Y) - \hat{R}. \tag{51b}$$

Step (b): Next, the decoder estimates message w_1 using again all blocks $i = 1, \ldots, B$, i.e., simultaneous decoding. It declares that $\hat{w}_1$ is sent if there exist $t^B = (t_1, \ldots, t_B) \in [1 : \hat{M}]^B$ such that $\mathbf{u}_i(\hat{w}_{c1})$, $\mathbf{x}_{2,i}(\hat{w}_{c1}, \hat{w}_{c2}, t_{i-1})$, $\mathbf{v}_i(\hat{w}_{c1}, \hat{w}_{c2}, t_{i-1}, t_i)$, $\mathbf{x}_{1,i}(\hat{w}_{c1}, w_1)$ and $\mathbf{y}[i]$ are jointly typical for all $i = 1, \ldots, B$. One can show that the decoder obtains the correct $w_c = (w_{c1}, w_{c2})$ as long as n and B are large and

$$R_1 \leq I(X_1; Y|U, V, X_2) \tag{52a}$$
$$R_1 \leq I(V, X_1, X_2; Y|U) - \hat{R}. \tag{52b}$$

The rest of the proof follows by applying Fourier-Motzkin Elimination to successively eliminate R_{c2} and $\hat{R}$ from the rate constraints defined by (50), (51) and (52).

Acknowledgement

This work has been supported by the European Commission in the framework of the FP7 Network of Excellence in Wireless Communications (NEWCOM#).

References

[1] S. H. Lim, Y.-H. Kim, A. E. Gamal, and S.-Y. Chung, "Noisy network coding," *IEEE Trans. Inf. Theory*, vol. 57, pp. 3132–3152, May 2011.

[2] A. Avestimeher, S. Diggavi, and D. Tse, "Wireless network information flow: a determenistic approach," *IEEE Trans. Inf. Theory*, vol. 57, pp. 1872–1905, April 2011.

[3] A. Zaidi and S. Shamai (Shitz), "On multiple access channels with delayed CSI," in *Proc. IEEE Int. Symp. Information Theory*, Istanbul, Turkey, Jun.-Jul. 2013.

[4] ——, "On cooperative multiple access channels with delayed CSI at transmitters," *IEEE Trans. Inf. Theory, submitted for publication, available at http://www-syscom.univ-mlv.fr/~zaidi/publications/proofs-isit2014.pdf*, 2014.

[5] A. Zaidi, P. Piantanida, and S. Shamai (Shitz), "Wyner-ziv type versus noisy network coding for a state-dependent mac," in *Proc. IEEE Int. Symp. Information Theory*, Cambridge, USA, Jul. 2012, pp. 1682–1686.

[6] ——, "Capacity region of cooperative multiple access channel with states," *IEEE Trans. Inf. Theory*, vol. 59, pp. 6153–6174, Oct. 2013.

[7] A. Lapidoth and Y. Steinberg, "The multiple access channel with causal side information: common state," *IEEE Trans. Inf. Theory*, vol. 59, pp. 32–50, Jan. 2013.

[8] ——, "The multiple access channel with causal side information: double state," *IEEE Trans. Inf. Theory*, vol. 99, pp. 1379–1393, 2013.

[9] M. Li, O. Simeone, and A. Yener, "Multiple access channels with states causally known at the transmitters," *IEEE Trans. Inf. Theory*, vol. 99, pp. 1394–1404, 2013.

[10] G. Como and S. Yüksel, "On the capacity of memoryless finite state multiple-access channels with asymmetric state information at the encoders," *IEEE Trans. Inf. Theory*, vol. 57, pp. 1267–1273, Oct. 2011.

[11] Z. Goldfeld, H. Permuter, and B. M. Zaidel, "Capacity region of the finite state mac with cooperative encoders and delayed CSI," in *Proc. IEEE Int. Symp. Information Theory*, Cambridge, MA, Jun. 2012, pp. 1692–1696.

[12] C. Choudhuri, Y.-H. Kim, and U. Mitra, "Causal state amplification," *Available in http://arxiv.org/abs/1203.6027*, 2012.

[13] R. Khosravi-Farsani and F. Marvasti, "Capacity bounds for multiuser channels with non-causal channel state information at the transmitters," in *Proc. IEEE Information Theory Workshop*, Paraty, Brasil, Oct. 2011, pp. 195 – 199.

[14] A. Vahid, M. A. Maddah-Ali, and A. S. Avestimehr, "Capacity results for binary fading interference channels with delayed CSIT," *IEEE Trans. Inf. Theory, submitted for publication. Available at urlhttp://arxiv.org/abs/1301.5309*, 2013.

[15] T. J. Oechtering and M. Skoglund, "Bidirectional broadcast channel with random states noncausally known at the encoder," *IEEE Trans. Inf. Theory*, vol. 59, pp. 64–75, 2013.

[16] A. E. Gamal and Y.-H. Kim, *Network information theory.* Cambridge University Press, 2011.