A Note on Multiple-Access Channels with Strictly-Causal State Information

Amos Lapidoth
ETH Zurich
Switzerland
Email: lapidoth@isi.ee.ethz.ch

Yossef Steinberg
Technion—Israel Institute of Technology
Israel
Email: ysteinbe@ee.technion.ac.il

***Abstract*—We propose a new inner bound on the capacity region of a memoryless multiple-access channel that is governed by a memoryless state that is known strictly causally to the encoders. The new inner bound contains the previous bounds, and we provide an example demonstrating that the inclusion can be strict.**

A variation on this example is then applied to the case where the channel is governed by two independent state sequences, where each transmitter knows one of the states strictly causally. The example proves that, as conjectured by Li *et al.*, an inner bound that they derived for this scenario can indeed by strictly better than previous bounds.

I. Introduction

If a memoryless single-user channel is governed by an independent and identically distributed (IID) state sequence, then its capacity is not increased if the state is made available to the encoder in a strictly-causal way. The picture changes dramatically on the multiple-access channel (MAC) [1], [2]: In the "single-state scenario," where the channel is governed by a single state sequence, the capacity region typically increases if the state is revealed to both transmitters in a strictly causal way [1]. Some of the gains can be attributed to the ability of the two encoders to compress the state information and to cooperate in sending the compressed version to the receiver. But strictly-causal side information (SI) is beneficial even in the "double-state scenario," where the channel is governed by two independent states, with each transmitter knowing one of the sequences strictly causally. In this case too, the side information can be helpful even though the transmitters cannot cooperate in compressing the states or in sending them [2].

The present note deals with both the single-state and the double-state scenarios. For the single-state scenario, we present a new inner bound on the capacity region. This bound contains the inner bound of [1] (which was extended to the many-transmitters scenario in [3]). We also provide an example showing that the inclusion can be strict.

By adapting this example to the double-state scenario, we provide an example showing that—as conjectured in [3]—the inner bound proposed by Li *et al.* in [3] can be strictly larger than that in [2].

To keep the contribution focused, we do not consider causal side information in this note, although our results can be carried over to that setting as in [1], [2].

We next describe the two scenarios more explicitly. Our descriptions are identical to those in [1], [2] except that, for simplicity, we do not consider cost constraints and we assume throughout that all the alphabets are finite.

A. The Single-State Scenario

In the single-state scenario we are given a discrete memoryless state-dependent MAC of law $P_{Y|W,X_1,X_2}$ with state alphabet $\mathcal{W}$, state probability mass function (PMF) P_W, input alphabets $\mathcal{X}_1$ and $\mathcal{X}_2$, and output alphabet $\mathcal{Y}$. Sequences of letters from $\mathcal{W}$ are denoted $w^n = (w_1, w_2, \ldots, w_n)$ and $w_i^j = (wi, w_{i+1} \ldots, w_j)$. Similar notation holds for all alphabets, e.g. $x_1^n = (x_{1,1}, x_{1,2}, \ldots, x_{1,n})$, $x_{2,i}^j = (x_{2,i}, x_{2,i+1}, \ldots, x_{2,j})$. When there is no risk of ambiguity, n-sequences will sometimes be denoted by boldface letters, $\mathbf{y}$, $\mathbf{x}_1$, $\mathbf{w}$, etc. The laws governing n-sequences of output letters and states are

$$P^n_{Y|W,X_1,X_2}(\mathbf{y}|\mathbf{w},\mathbf{x}_1,\mathbf{x}_2) = \prod_{i=1}^{n} P_{Y|W,X_1,X_2}(y_i|w_i,x_{1,i},x_{2,i}),$$

$$P^n_W(\mathbf{w}) = \prod_{i=1}^{n} P_W(w_i).$$

For notational convenience, we henceforth omit the superscript n, and we denote the channel by P.

Definition 1: Given positive integers ν_1, ν_2, let $\mathcal{M}_1$ denote the set $\{1, 2, \ldots, \nu_1\}$, and let $\mathcal{M}_2$ denote the set $\{1, 2, \ldots, \nu_2\}$. An $(n, \nu_1, \nu_2, \epsilon)$ code with strictly-causal side information (SI) at the encoders is a pair of sequences of encoder mappings

$$f_{k,i} \colon \mathcal{W}^{i-1} \times \mathcal{M}_k \to \mathcal{X}_k, \quad k = 1, 2, \quad i = 1, \ldots, n \tag{1}$$

and a decoding mapping

$$g \colon \mathcal{Y}^n \to \mathcal{M}_1 \times \mathcal{M}_2$$

such that the average probability of error P_e does now exceed ϵ. Here P_e is $1 - P_\mathrm{c}$;

$$P_\mathrm{c} = \frac{1}{\nu_1 \nu_2} \sum_{m_1=1}^{\nu_1} \sum_{m_2=1}^{\nu_2} \Pr(\text{correct}|m_1, m_2);$$

and

$$\Pr(\text{correct}|m_1, m_2) = \\ \sum_{\mathbf{w}} P_W(\mathbf{w}) P\left(g^{-1}(m_1, m_2)|\mathbf{w}, \mathbf{f}_1(\mathbf{w}, m_1), \mathbf{f}_2(\mathbf{w}, m_2)\right),$$

where $g^{-1}(m_1, m_2) \subset \mathcal{Y}^n$ is the decoding set of the pair of messages (m_1, m_2), and

$$\mathbf{f}_k(\mathbf{w}, m_k) = (f_{k,1}(m_k), f_{k,2}(w_1, m_k), \ldots, f_{k,n}(w^{n-1}, m_k)).$$

The rate pair (R_1, R_2) of the code is defined as

$$R_1 = \frac{1}{n} \log \nu_1, \quad R_2 = \frac{1}{n} \log \nu_2.$$

A rate-pair (R_1, R_2) is said to be achievable if for every positive ϵ and sufficiently large n there exists an $(n, 2^{nR_1}, 2^{nR_2}, \epsilon)$ code with strictly-causal SI for the channel $P_{Y|W,X_1,X_2}$. The capacity region of the channel with strictly-causal SI is the closure of the set of all achievable pairs (R_1, R_2), and is denoted $\mathcal{C}^{\text{com}}_{\text{s-c}}$. The subscript "s-c" stands for strictly-causal.

B. The Double-State Scenario

In the double-state scenario we are given a discrete memoryless state-dependent MAC $P_{Y|S_1,S_2,X_1,X_2}$ with state alphabets $\mathcal{S}_1$ and $\mathcal{S}_2$, state probability mass functions (PMFs) P_{S_1} and P_{S_2}, input alphabets $\mathcal{X}_1$ and $\mathcal{X}_2$, and output alphabet $\mathcal{Y}$. The laws governing n sequences of output letters and states are

$$P^n_{Y|S_1,S_2,X_1,X_2}(\mathbf{y}|\mathbf{s}_1\mathbf{s}_2, \mathbf{x}_1, \mathbf{x}_2) \\ = \prod_{i=1}^{n} P_{Y|S_1,S_2,X_1,X_2}(y_i|s_{1,i}, s_{2,i}, x_{1,i}, x_{2,i}),$$

$$P^n_{S_1,S_2}(\mathbf{s}_1, \mathbf{s}_2) = \prod_{i=1}^{n} P_{S_1}(s_{1,i}) P_{S_2}(s_{2,i}).$$

For notational convenience, we henceforth omit the superscript n, and we denote the channel by P.

Given positive integers ν_1, ν_2, let $\mathcal{M}_1$ be the set $\{1, 2, \ldots, \nu_1\}$ and $\mathcal{M}_2$ the set $\{1, 2, \ldots, \nu_2\}$. An $(n, \nu_1, \nu_2, \epsilon)$ code with strictly causal independent SI at the encoders is a pair of sequences of encoder mappings

$$f_{k,i} \colon \mathcal{S}_k^{i-1} \times \mathcal{M}_k \to \mathcal{X}_k, \quad k = 1, 2, \quad i = 1, \ldots, n \tag{3}$$

and a decoding mapping

$$g \colon \mathcal{Y}^n \to \mathcal{M}_1 \times \mathcal{M}_2$$

such that the average probability of error P_e is bounded by ϵ, where $P_{\text{e}} = 1 - P_{\text{c}}$ and

$$P_{\text{c}} = \frac{1}{\nu_1 \nu_2} \sum_{m_1=1}^{\nu_1} \sum_{m_2=1}^{\nu_2} \sum_{\mathbf{s}_1, \mathbf{s}_2} \\ P_{S_1}(\mathbf{s}_1) P_{S_2}(\mathbf{s}_2) P\big(g^{-1}(m_1, m_2)|\mathbf{s}_1, \mathbf{s}_2, \mathbf{f}_1(\mathbf{s}_1, m_1), \mathbf{f}_2(\mathbf{s}_2, m_2)\big)$$

where $g^{-1}(m_1, m_2) \subset \mathcal{Y}^n$ is the decoding set of the pair of messages (m_1, m_2), and

$$\mathbf{f}_k(\mathbf{s}_k, m_k) = \big(f_{k,1}(m_k), f_{k,2}(s_{k,1}, m_k), \ldots, f_{k,n}(s_k^{n-1}, m_k)\big).$$

The rate pair (R_1, R_2) of the code is defined as

$$R_1 = \frac{1}{n} \log \nu_1, \quad R_2 = \frac{1}{n} \log \nu_2.$$

A rate-pair $(R_1, R_2,)$ is said to be achievable if for every $\epsilon > 0$ and sufficiently large n there exists an $(n, 2^{nR_1}, 2^{nR_2}, \epsilon)$ code with strictly-causal SI for the channel $P_{Y|S,X_1,X_2}$. The capacity region of the channel with strictly-causal independent SI is the closure of the set of all achievable pairs (R_1, R_2), and is denoted $\mathcal{C}^{\text{ind}}_{\text{s-c}}$. The superscript "ind" indicates that the two states are independent.

II. The Single-State Scenario

For the single-state scenario, an inner bound on $\mathcal{C}^{\text{com}}_{\text{s-c}}$ was derived in [1] and later extended to many-transmitters in [3]. In the absence of cost constraints this bound can be described as follows: Let $\mathcal{P}^{\text{com}}_{\text{s-c}}$ be the collection of all random variables (U, V, X_1, X_2, W, Y) whose joint distribution satisfies

$$P_{U,V,X_1,X_2,W,Y} = P_W P_{X_1|U} P_{X_2|U} P_U P_{V|W} P_{Y|W,X_1,X_2}. \tag{4}$$

Note that (4) implies the Markov relations $X_1 \multimap U \multimap X_2$ and $V \multimap W \multimap Y$, and that the triplet (X_1, U, X_2) is independent of (V, W). Let $\mathcal{R}^{\text{com}}_{\text{s-c}}$ be the convex hull of the collection of all (R_1, R_2) satisfying

$$R_1 \le I(X_1; Y|X_2, U, V) \tag{5a}$$
$$R_2 \le I(X_2; Y|X_1, U, V) \tag{5b}$$
$$R_1 + R_2 \le I(X_1, X_2; Y|U, V) \tag{5c}$$
$$R_1 + R_2 \le I(X_1, X_2, V; Y) - I(V; W) \tag{5d}$$

for some $(U, V, X_1, X_2, W, Y) \in \mathcal{P}^{\text{com}}_{\text{s-c}}$.

Theorem 1 ([1]): $\mathcal{R}^{\text{com}}_{\text{s-c}} \subseteq \mathcal{C}^{\text{com}}_{\text{s-c}}$.

The achievability of this region is based on a Block-Markov scheme where at Block $\nu + 1$ the transmitters send fresh private messages as well as a common message that is used to send a compressed version of the state sequence of Block ν. The compression is of the Wyner-Ziv type with the side information being the channel outputs at Block ν.

We next present a tighter inner bound. At Block $\nu + 1$ we still use the MAC by sending private messages and a common message. The common message is still a compressed version of the state information from the previous block. The twist, however, is that the private messages need not be entirely composed of fresh information. The private message of Transmitter 1 has two parts. The first, of rate R_1, is indeed fresh information. But the second, of rate $R_0^{(1)}$, is a compressed version of the pair of sequences $(\mathbf{x}_1, \mathbf{w})$ from Block ν (again with the side information being the received symbols in the previous block). Since Transmitter 1 knows which symbols it sent in the previous block, and since it knows the state of the channel in the previous block, it can compress the pair $(\mathbf{x}_1, \mathbf{w})$. Likewise Transmitter 2. Using Gastpar's results on the compression of correlated sources with side information [4] we obtain the following bound:

Theorem 2: The rate-pair (R_1, R_2) is achievable if for some joint distribution of the form

$$P_{U,V,V_1,V_2,X_1,X_2,W,Y} = P_W P_{X_1|U} P_{X_2|U} P_U P_{V|W} P_{V_1|W,X_1} P_{V_2|W,X_2} P_{Y|W,X_1,X_2} \tag{6}$$

there exist nonnegative numbers $R_0^{(1)}$ and $R_0^{(2)}$ such that

$$R_1 + R_0^{(1)} \leq I(X_1; Y, V_1, V_2, V | X_2, U) \tag{7}$$
$$R_2 + R_0^{(2)} \leq I(X_2; Y, V_1, V_2, V | X_1, U) \tag{8}$$
$$R_1 + R_2 + R_0^{(1)} + R_0^{(2)} \leq I(X_1, X_2; Y, V_1, V_2, V | U) \tag{9}$$
$$R_0 + R_1 + R_2 + R_0^{(1)} + R_0^{(2)} \leq I(X_1, X_2; Y, V_1, V_2, V) \tag{10}$$

and

$$R_0^{(1)} \geq I(X_1, W; V_1 | V, V_2, Y) \tag{11a}$$
$$R_0^{(2)} \geq I(X_2, W; V_2 | V, V_1, Y) \tag{11b}$$
$$R_0 \geq I(W; V | V_1, V_2, Y) \tag{11c}$$
$$R_0^{(1)} + R_0^{(2)} \geq I(X_1, X_2, W; V_1, V_2 | V, Y) \tag{11d}$$
$$R_0^{(1)} + R_0 \geq I(X_1, W; V_1, V | V_2, Y) \tag{11e}$$
$$R_0^{(2)} + R_0 \geq I(X_2, W; V_2, V | V_1, Y) \tag{11f}$$
$$R_0^{(1)} + R_0^{(2)} + R_0 \geq I(X_1, X_2, W; V_1, V_2, V | Y). \tag{11g}$$

If we only consider joint distributions where V_1 and V_2 are deterministic, and if we set $R_0^{(1)}$, $R_0^{(2)}$ to zero, we obtain the inner bound of [1]. Thus,

Remark 1: The proposed inner bound contains the inner bound of [1]

The following example shows that the inclusion can be strict.

Example 1: Consider a MAC with two binary inputs $\mathcal{X}_1 = \mathcal{X}_2 = \{0, 1\}$; a common state $W = (W_0, W_1) \in \{0, 1\}^2$, where W_0, W_1 are IID with entropy

$$H(W_0) = H(W_1) = 1/2; \tag{12}$$

and an output $Y = (Y_1, Y_2) \in \{0, 1\}^2$ with

$$Y_1 = X_1 \oplus W_{X_2} \tag{13a}$$
$$Y_2 = X_2. \tag{13b}$$

Thus, if X_2 is equal to zero, then Y_1 is the mod-2 sum of X_1 and W_0, and otherwise it is the mod-2 sum of X_1 and W_1. We study the highest rate at which User 2 can communicate when User 1 transmits at rate 1. We show that for this channel

$$\max\{R_2 : (1, R_2) \in \mathcal{R}_{\text{s-c}}^{\text{com}}\} = 0 \tag{14}$$

but

$$\max\{R_2 : (1, R_2) \in \mathcal{C}_{\text{s-c}}^{\text{com}}\} = 1/2, \tag{15}$$

and that the rate-pair $(1, 1/2)$ is in the new inner bound.

Proof: We first prove (15). To this end we note that if $(1, R_2)$ is achievable, then R_2 cannot exceed $1/2$. This can be shown using the full-cooperation outer-bound [1], which implies that (R_1, R_2) can only be achievable if $R_1 + R_2 \leq 3/2$. Of more interest to us is the fact that the rate-pair $(1, 1/2)$ is achievable. We demonstrate this using the new inner bound. Indeed, it is straightforward to verify that setting

$$R_0^{(1)} = R_0 = 0, \quad R_0^{(2)} = 1/2, \tag{16a}$$
$$V = V_1 = 0, \quad V_2 = W_{X_2}, \tag{16b}$$
$$U = 0, \tag{16c}$$
$$X_1, X_2 \sim \text{IID Bernoulli } 1/2, \tag{16d}$$

and

$$(R_1, R_2) = (1, 1/2) \tag{16e}$$

satisfies all the required inequalities. This choice corresponds to the following Block-Markov scheme: In the Block-Markov scheme Transmitter 1 sends its data uncoded. At Block $b+1$ Transmitter 2 sends n bits, half of which are fresh data bits and half of which are used to describe the n-length sequence $\mathbf{w}_{\mathbf{x}_2}$ of the previous block. Note that Transmitter 2 does not describe the entire state sequence $\mathbf{w}$ of the previous block but only $\mathbf{w}_{\mathbf{x}_2}$. This latter sequence is known to Transmitter 2 at the beginning of Block $b+1$ thanks to the strictly-causal state information and because it knows the sequence $\mathbf{x}_2$ it transmitted in the previous block. And $n/2$ bits suffice to describe this sequence because W_{X_2} is of entropy $1/2$.

We now turn to proving (14). We fix some distribution $P_{U,V,X_1,X_2,W,Y}$ of the form (4), we assume that $(R_1 = 1, R_2)$ satisfy Inequalities (5), and we then prove that R_2 must be zero. Since $R_1 = 1$ and since $\mathcal{X}_1$ is binary, Inequality (5a) must hold with equality, and X_1 must be independent of (X_2, U, V). By (4), this implies that

$$X_1 \text{ is independent of } (X_2, U, V, W). \tag{17a}$$

From (5a) (that we know holds with equality) and the fact that $R_1 = 1$ we also infer that

$$\begin{aligned} 1 &= H(Y|X_2, U, V) - H(Y|X_1, X_2, U, V) \\ &= H(Y_1|X_2, U, V) - H(Y_1|X_1, X_2, U, V) \end{aligned} \tag{17b}$$

where the second equality holds because Y_2 is a deterministic function of X_2. Since Y_1 is binary, $H(Y_1|X_2, U, V)$ is upper-bounded by 1, and we conclude from (17b) that

$$\begin{aligned} 0 &= H(Y_1|X_1, X_2, U, V) \\ &= H(Y_1 \oplus X_1|X_1, X_2, U, V) \\ &= H(W_{X_2}|X_1, X_2, U, V) \\ &= H(W_{X_2}|X_2, U, V) \end{aligned} \tag{17c}$$

where the last equality follows from (17a).

We next show that

$$U \multimap\!\!- (X_2, V) \multimap\!\!- W_{X_2}. \tag{17d}$$

To this end we note that, by (4), the pair (V, W) is independent of (U, X_2) and hence

$$U \multimap\!\!- (X_2, V) \multimap\!\!- W. \tag{17e}$$

Since W_{X_2} is a deterministic function of (X_2, V, W), this implies (17d), because if $A \multimap\!\!- B \multimap\!\!- C$ forms a Markov chain then $A \multimap B \multimap f(B,C)$. Having established (17d), we now obtain from (17c)

$$H(W_{X_2}|X_2,V)=0. \tag{17f}$$

We now focus on the case where X_2 is not deterministic

$$\Pr\bigl[X_2=\eta\bigr]>0, \quad \eta\in\{0,1\}, \tag{17g}$$

because if X_2 is deterministic then R_2 must be zero by (5b). We also assume that the PMF of V is strictly positive

$$\Pr\bigl[V=v\bigr]>0, \quad v\in\mathcal{V}, \tag{17h}$$

because outcomes of the auxiliary random variable that have zero probability can be removed from $\mathcal{V}$ without affecting the inner bound. Since , by (4), V is independent of X_2, it follows from (17g) and (17h) that

$$\Pr\bigl[X_2=\eta, V=v\bigr]>0, \quad \eta\in\{0,1\},\ v\in\mathcal{V}. \tag{17i}$$

This and (17f) imply that

$$H(W_\eta|X_2=\eta, V=v)=0, \quad \eta\in\{0,1\},\ v\in\mathcal{V}. \tag{17j}$$

Since, by (4), X_2 is independent of (V,W) and, *a fortiori*, of (V,W_η), it follows from (17j) that

$$H(W_\eta|V=v)=0, \quad \eta\in\{0,1\},\ v\in\mathcal{V}. \tag{17k}$$

Thus, $H(W_\eta|V)=0$, and since $W=(W_0,W_1)$,

$$H(W|V)=0. \tag{17l}$$

Consequently,

$$\begin{aligned} I(V;W) &= H(W) \\ &= 1, \end{aligned} \tag{17m}$$

where the second equality follows from (12) and the independence of W_0 and W_1. From (17m), (5d), and the fact that $\mathcal{Y}$ has four elements we then conclude that $R_1+R_2\le 1$. This combines with $R_1=1$ to establish that R_2 must be zero. ∎

Terminating the Block-Markov scheme: To conclude the sketch of the achievability of the new inner bound, we still need to describe how the Block-Markov scheme is terminated. We thus assume that B blocks have been transmitted, and we proceed to describe Blocks $B+1$, $B+2$, and $B+3$. We think about these blocks as "overhead," because they contain no fresh information. Fortunately, this overhead does not affect the throughput because we can choose B very large.

The next lemma shows that if the full-cooperation capacity of the MAC without SI is zero, then the new inner bound contains only the rate-pair $(0,0)$ and is thus trivially an inner bound.

Lemma 1: If the capacity of the MAC without any side information but with full cooperation is zero, i.e., if

$$\max_{P_{X_1,X_2}} I(X_1,X_2;Y)=0, \tag{18}$$

then the proposed new inner bound contains only the all-zero rate tuple.

Proof: By (11g) and (10), we conclude that if R_1,R_2 is in the new inner bound, then for some joint distribution of the form (6)

$$\begin{aligned} &R_1+R_2 \\ &\quad\le I(X_1,X_2;Y,V_1,V_2,V)-I(X_1,X_2,W;V_1,V_2,V|Y) \\ &\quad= I(X_1,X_2;Y)+I(X_1,X_2;V_1,V_2,V|Y) \\ &\qquad -I(X_1,X_2,W;V_1,V_2,V|Y). \end{aligned}$$

Consequently, if (18) holds and hence $I(X_1,X_2;Y)$ is zero, then R_1+R_2 must be upper-bounded by $I(X_1,X_2;V_1,V_2,V|Y)-I(X_1,X_2,W;V_1,V_2,V|Y)$, which is nonpositive. ∎

In view of Lemma 1, it only remains to prove the achievability of the new inner bound when the full-cooperation capacity without SI is positive. The next lemma shows that we can also assume that the channel between Transmitter 1 (uninformed) and the receiver (informed) is of positive capacity and likewise from Transmitter 2.

Lemma 2: If the channel between Transmitter 1 (uninformed) to the receiver (informed) is of zero capacity, i.e.,

$$\max_{x_2\in\mathcal{X}_2}\max_{P_{X_1}} I(X_1;Y,W|X_2=x_2)=0, \tag{19}$$

then the new inner bound contains only rate pairs (R_1,R_2) with $R_1=0$ and $R_2\le\max I(X_2;Y)$. An analogous result holds if

$$\max_{x_1\in\mathcal{X}_1}\max_{P_{X_2}} I(X_2;Y,W|X_1=x_1)=0, \tag{20}$$

Proof: We first prove that if a rate pair (R_1,R_2) is in the new inner bound, and if (19) holds, then R_1 must be zero. Fix some joint distribution of the form (6) and let (R_1,R_2) satisfy the inequalities of Theorem 2. We next argue that Hypothesis (19) implies

$$I(X_1;Y,V_2,V|X_2,U)=0. \tag{21}$$

Indeed,

$$\begin{aligned} &I(X_1;Y,V_2,V|X_2,U) \\ &\quad\le I(X_1;Y,V_2,V|X_2,U,W) && \text{(22a)} \\ &\quad= I(X_1;Y|X_2,U,W,V_2,V) && \text{(22b)} \\ &\quad= I(X_1;Y|X_2,U,W), && \text{(22c)} \\ &\quad= I(X_1;Y,W|X_2,U), && \text{(22d)} \\ &\quad\le \max_{u\in\mathcal{U}}\max_{x_2\in\mathcal{X}_2} I(X_1;Y,W|X_2=x_2,U=u) && \text{(22e)} \\ &\quad\le \max_{u\in\mathcal{U}}\max_{x_2\in\mathcal{X}_2}\max_{P_{X_1|U=u}} I(X_1;Y,W|X_2=x_2,U=u) && \text{(22f)} \\ &\quad= \max_{x_2\in\mathcal{X}_2}\max_{P_{X_1}} I(X_1;Y,W|X_2=x_2) && \text{(22g)} \end{aligned}$$

where the first line follows from

$$X_1 \multimap (X_2,U) \multimap W; \tag{23}$$

the second from the chain rule and because

$$X_1 \multimap (X_2,U,W) \multimap (V_2,V) \tag{24}$$

so $I(X_1;V_2,V|X_2,U,W)$ is zero; the third from

$$(X_1,Y) -\!\!\circ\!\!- (X_2,U,W) -\!\!\circ\!\!- (V_2,V); \tag{25}$$

the fourth again by (23); the fifth by upper bounding the average by the maximal value; the sixth by maximizing over the conditional distribution of X_1 given $U=u$; and the last because the maximization over u on the RHS of (22f) is unnecessary.

Continuing our proof that R_1 must be zero, we note that (7) and (11a) imply

$$\begin{aligned} R_1 &\le I(X_1;Y,V_1,V_2,V|X_2,U) - I(X_1,W;V_1|V,V_2,Y) \\ &= I(X_1;Y,V_2,V|X_2,U) + I(X_1;V_1|X_2,U,Y,V_2,V) \\ &\qquad - I(X_1,W;V_1|V,V_2,Y) \\ &= I(X_1;V_1|X_2,U,Y,V_2,V) - I(X_1,W;V_1|V,V_2,Y) \\ &= H(V_1|X_2,U,Y,V_2,V) - H(V_1|X_1,X_2,U,Y,V_2,V) \\ &\qquad + H(V_1|X_1,W,V,V_2,Y) - H(V_1|V,V_2,Y) \\ &= H(V_1|X_2,U,Y,V_2,V) - H(V_1|X_1,X_2,U,Y,V_2,V) \\ &\qquad + H(V_1|X_1,W) - H(V_1|V,V_2,Y) \\ &\le 0, \end{aligned}$$

where the second equality (third line) follows from (21), and where in the last inequality we have used

$$H(V_1|X_2,U,Y,V_2,V) \le H(V_1|V,V_2,Y)$$

(conditioning reduces entropy) and

$$H(V_1|X_1,X_2,U,Y,V_2,V) \ge H(V_1|X_1,W),$$

which can be argued as follows:

$$\begin{aligned} H(V_1|X_1,X_2,U,Y,V_2,V) &\ge H(V_1|X_1,W,X_2,U,Y,V_2,V) \\ &= H(V_1|X_1,W), \end{aligned}$$

where the first inequality is because conditioning cannot increase entropy, and the second by (6), which implies that, conditional on (X_1,W), the auxiliary random variable V_1 is independent of (X_2,U,Y,V_2,V).

Having established that R_1 is zero, we now conclude from (11g) and (10)

$$\begin{aligned} R_2 &= R_1 + R_2 \\ &\le I(X_1,X_2;Y,V_1,V_2,V) - I(X_1,X_2,W;V_1,V_2,V|Y) \\ &= I(X_1,X_2;Y) + \\ &\quad I(X_1,X_2;V_1,V_2,V|Y) - I(X_1,X_2,W;V_1,V_2,V|Y) \\ &\le I(X_1,X_2;Y) \\ &= I(X_2;Y) + I(X_1;Y|X_2) \\ &= I(X_2;Y). \end{aligned}$$

■

Lemma 2 shows that if either (19) or (20) holds, then the new inner bound is achievable. It thus only remains to prove its achievability when

$$\max_{x_2\in\mathcal{X}_2} \max_{P_{X_1}} I(X_1;Y,W|X_2=x_2) > 0 \tag{26}$$

and

$$\max_{x_1\in\mathcal{X}_1} \max_{P_{X_2}} I(X_2;Y,W|X_1=x_1) > 0, \tag{27}$$

both of which we now assume.

We are now ready to describe the termination of the Block-Markov scheme. Block $B+1$ is split into two parts. In the first, Transmitter 1 sends the $\mathbf{v}_1$-sequence of Block B assuming that the receiver knows the state sequence $\mathbf{w}$ of Block $B+1$. This can be done (under this assumption) by (26). In the second, Transmitter 2 sends the $\mathbf{v}_2$-sequence of Block B assuming that the receiver knows the state of Block $B+1$. This is possible by (27). In Block $B+2$ the transmitters cooperate to send the sequence $\mathbf{w}$ of Block $B+1$, and in Block $B+3$ they cooperate to send the $\mathbf{v}$ sequence of Block B.

Decoding is performed as follows. The decoder first decodes Block $B+3$ without any side-information and thus learns the sequence $\mathbf{v}$ of Block B. It then decodes Block $B+2$ (again without any side information) and learns the state sequence $\mathbf{w}$ of Block $B+1$. Now that it knows the state sequence of Block-$B+1$, it can decode that block and learn the $\mathbf{v}_1$-sequence and the $\mathbf{v}_2$-sequence of Block B. From here on, it can proceed with the regular backward decoding: in decoding Block b it knows the sequences $\mathbf{v}$, $\mathbf{v}_1$, and $\mathbf{v}_2$ of Block b and it can therefore decode the common message and the messages transmitted by each of the transmitters in Block b. From this decoding it learns the private messages of Block b, and the sequences $\mathbf{v}$, $\mathbf{v}_1$, and $\mathbf{v}_2$ of Block $b-1$.

III. The Double-State Scenario

For the double-state scenario, an inner bound on $\mathcal{C}^{\text{ind}}_{\text{s-c}}$ was proposed in [2]. In the absence of cost constraints this bound can be described as follows: Let $\mathcal{P}^{\text{ind}}_{\text{s-c}}$ be the collection of all random variables $(V_1,V_2,S_1,S_2,X_1,X_2,Y)$ whose joint distribution satisfies

$$\begin{aligned} &P_{V_1,V_2,S_1,S_2,X_1,X_2,Y} = \\ &\qquad P_{V_1|S_1} P_{V_2|S_2} P_{S_1} P_{S_2} P_{X_1} P_{X_2} P_{Y|S_1,S_2,X_1,X_2}. \end{aligned} \tag{28}$$

Note that (28) implies the Markov relations

$$\begin{aligned} &V_1 -\!\!\circ\!\!- S_1 -\!\!\circ\!\!- (V_2,Y,S_2) \\ &V_2 -\!\!\circ\!\!- S_2 -\!\!\circ\!\!- (V_1,Y,S_1) \\ &(V_1,V_2) -\!\!\circ\!\!- (S_1,S_2) -\!\!\circ\!\!- Y \end{aligned} \tag{29}$$

and that X_1,X_2 are independent of each other and of the quadruple (V_1,V_2,S_1,S_2). Let $\mathcal{R}^{\text{ind}}_{\text{s-c}}$ be the convex hull of the collection of all rate-pairs (R_1,R_2) satisfying

$$0 \le R_1 \le I(X_1;Y|X_2,V_1,V_2) - I(V_1;S_1|Y,V_2) \tag{30}$$

$$0 \le R_2 \le I(X_2;Y|X_1,V_1,V_2) - I(V_2;S_2|Y,V_1) \tag{31}$$

$$R_1 + R_2 \le I(X_1,X_2;Y|V_1,V_2) - I(V_1,V_2;S_1,S_2|Y) \tag{32}$$

for some $(V_1,V_2,S_1,S_2,X_1,X_2,Y) \in \mathcal{P}^{\text{ind}}_{\text{s-c}}$.

Theorem 3 ([2]): $\mathcal{R}^{\text{ind}}_{\text{s-c}} \subseteq \mathcal{C}^{\text{ind}}_{\text{s-c}}$.

The proof is based on a scheme where lossy versions of the state sequences are conveyed to the decoder using distributed Wyner-Ziv compression [4] and Block-Markov encoding for the MAC, to transmit the messages and the Wyner-Ziv codewords. The channel output serves as the decoder's SI in the distributed Wyner-Ziv code. Since the two components of the source are independent, there is no direct cooperation between the encoders via a common message as in single-state scenario. Instead, each user spends part of its private rate on the transmission of its Wyner-Ziv codeword.

An improved inner bound was proposed by Li *et al.* in [3]. There it was shown that the improved inner bound always contains the inner bound of [2], and it was conjectured that there are cases where the inclusion is strict. We next present the inner bound of Li *et al.* and then show that the inclusion can, indeed, be strict.

Li *et al.* consider all joint distributions of the form

$$P_{V_1,V_2,S_1,S_2,X_1,X_2,Y} = P_{V_1|S_1,X_1} P_{V_2|S_2,X_2} P_{S_1} P_{S_2} P_{X_1} P_{X_2} P_{Y|S_1,S_2,X_1,X_2} \tag{33}$$

and prove the achievability of rate pairs (R_1, R_2) satisfying

$$R_1 \le I(X_1, V_1; Y|X_2, V_2) - I(V_1; S_1|X_1) \tag{34a}$$

$$R_2 \le I(X_2, V_2; Y|X_1, V_1) - I(V_2; S_2|X_2) \tag{34b}$$

$$R_1 + R_2 \le I(X_1, X_2, V_1, V_2; Y) - I(V_1; S_1|X_1) - I(V_2; S_2|X_2). \tag{34c}$$

Roughly speaking, the improvement in the inner bound is the result of Transmitter 1 compressing the pair $(\mathbf{s}_1, \mathbf{x}_1)$ from the previous block (with the outputs from the previous block serving as side information) and not just $\mathbf{s}_1$ and likewise for Transmitter 2. We next show, by example, that the bound of Li *et al.* can, indeed, be tighter than that of Theorem 3

The example is very similar to Example 1. In fact, the channel is as in Example 1, but with the state S_1 being null (deterministic) and the state S_2 consisting of the pair (W_0, W_1) of Example 1:

$$S_1 = 0 \quad S_2 = (W_0, W_1), \tag{35a}$$

where W_0, W_1 are IID binary random variables, each of entropy $1/2$.

The rate pair $(R_1, R_2) = (1, 1/2)$ is in the inner bound of Li *et al.*. To see this we set $V_1 = 0$ and $V_2 = W_{X_2}$ with X_1, X_2 IID random bits. However, as we next prove, the pair $(1, 1/2)$ is not in $\mathcal{R}^{\text{ind}}_{\text{s-c}}$.

We prove this by showing that if $(1, R_2)$ is in $\mathcal{R}^{\text{ind}}_{\text{s-c}}$, then R_2 must be zero. Suppose then that $(1, R_2) \in \mathcal{R}^{\text{ind}}_{\text{s-c}}$. Since S_1 is null, it follows from the structure (28) of the joint distribution, that V_1 must be independent of all the other random variables. Consequently, we can strike it out from (30), (31), and (32). Since $R_1 = 1$, it follows from (30) that X_1 must be Bernoulli(1/2) and that $H(X_1|X_2, V_2, Y)$ must be zero. This implies that $H(W_{X_2}|X_2, V_2, Y)$ must also be zero (because $X_1 = Y_1 \oplus W_{X_2}$). Consequently, $H(W_{X_2}|X_2, V_2, Y_1)$ must also be zero (because $Y_2 = X_2$). This implies that

$$H(W_{X_2}|X_2, V_2) = 0 \tag{35b}$$

because X_1 is Bernoulli(1/2) and independent of (X_2, V_2, W), so Y_1, which is equal to $X_1 \oplus W_{X_2}$, must also be independent of (X_2, V_2, W). Equation (35b) is reminiscent of (17f) (with V_2 replacing V).

As in Example 1, we now distinguish between two cases depending on whether X_2 is deterministic or not. If it is deterministic, then the rate R_2 must be zero by (31). Consider now the case when it is not. In this case $\Pr[X_2 = \eta]$ is positive for all $\eta \in \{0, 1\}$. Since V_2 is independent of X_2 (by (28)), and since without changing the inner bound we can assume that $\Pr[V_2 = v_2]$ is positive for all $v_2 \in \mathcal{V}_2$, it follows that in this case

$$\Pr[X_2 = \eta,\, V_2 = v_2] > 0, \quad \eta \in \{0, 1\},\ v_2 \in \mathcal{V}_2. \tag{35c}$$

This combines with (35b) to imply that

$$H(W_\eta|X_2 = \eta, V_2 = v_2) = 0, \quad \eta \in \{0, 1\},\ v_2 \in \mathcal{V}_2. \tag{35d}$$

This implies that

$$H(W_\eta|V_2 = v_2) = 0, \quad \eta \in \{0, 1\},\ v_2 \in \mathcal{V}_2, \tag{35e}$$

because, by (28), X_2 is independent of (V_2, S_2) and hence *a fortiori* of (V_2, W_η). Thus, $H(W_\eta|V_2) = 0$, and since $S_2 = (W_0, W_1)$,

$$H(S_2|V_2) = 0. \tag{35f}$$

Consequently,

$$I(V_2; S_2) = H(S_2) = 1. \tag{35g}$$

This implies that also

$$I(V_2; S_2|Y) = 1, \tag{35h}$$

because Y is independent of (V_2, S_2). It now follows from (35h), the fact that V_1 is deterministic, and from (31) that R_2 must be zero.

IV. Summary

We have presented an improved inner bound on the capacity region of the memoryless multiple-access channel that is controlled by an IID state that is known strictly causally to the two encoders. This bound contains the bound of [1], and we have provided an example showing that the inclusion can be strict.

We also adapted this example to a memoryless multiple-access channel that is governed by two independent states, where each transmitter knows one of the states strictly causally. The resulting example demonstrates that—as conjecture by Li *et al.* [3]—the inner bound of Li *et al.* can be strictly tighter than that of [2].

References

[1] A. Lapidoth and Y. Steinberg, "The multiple-access channel with causal and strictly causal side information at the encoders," in *Proceedings International Zurich Seminar*, March 2010.

[2] A. Lapidoth and Y. Steinberg, "The multiple access channel with two independent states each known causally to one encoder," in *Proceedings IEEE International Symposium on Information Theory (ISIT)*, June 2010.

[3] M. Li, O. Simeone, and A. Yener, "Multiple access channels with states causally known at transmitters," *Preprint*. arXiv:1011.6639v1 30 November 2010.

[4] M. Gastpar, "On Wyner-Ziv networks," in *Proc. 37th Asilomar Conf. Signals, Systems, and Computers*, Asilomar, CA, Nov. 2003.