The State-Dependent Multiple-Access Channel with States Available at a Cribbing Encoder

Shraga I. Bross
School of Engineering
Bar-Ilan University
Ramat-Gan 52900, Israel
brosss@macs.biu.ac.il

Amos Lapidoth
ETH Zurich
CH-8092 Zurich, Switzerland
lapidoth@isi.ee.ethz.ch

***Abstract*—The two-user discrete memoryless state-dependent multiple-access channel (MAC) models a scenario in which two encoders transmit independent messages to a single receiver via a MAC whose channel law is governed by an i.i.d. state random variable. In the cooperative state-dependent MAC model it is further assumed that Message 1 is shared by both encoders whereas Message 2 is known only to Encoder 2 – the cognitive transmitter. The capacity of the cooperative state-dependent MAC where the realization of the state sequence is known non-causally to the cognitive encoder was derived by Somekh-Baruch *et. al.*.**

In this work we dispense with the assumption that Message 1 is shared by both encoders. Instead, we study the case in which Encoder 2 cribs causally from Encoder 1. We determine the capacity region for the case where the cribbing is strictly causal.

Classification: Information theory, communications.

I. Channel Model and Main Results

A discrete memoryless state-dependent multiple-access channel (MAC) is a triple $(\mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{S}, p(y|x_1, x_2, s), \mathcal{Y})$ where $\mathcal{X}_1$ and $\mathcal{X}_2$ are finite sets corresponding to the input alphabets of Encoder 1 and Encoder 2 respectively, $\mathcal{S}$ is a finite set corresponding to the alphabet of the state governing the channel law, the finite set $\mathcal{Y}$ is the output alphabet at the receiver, and $p(\cdot|x_1, x_2, s)$ is a collection of probability laws on $\mathcal{Y}$ indexed by the input symbols $x_1 \in \mathcal{X}_1$, $x_2 \in \mathcal{X}_2$ and $s \in \mathcal{S}$. The channel's law extends to n-tuples according to a memoryless law.

Encoder 1 sends a message W_1, which is drawn uniformly over the set $\{1, \ldots, e^{nR_1}\} \triangleq \mathcal{W}_1$, while Encoder 2 sends a message W_2 which is independent of W_1 and is drawn uniformly over the set $\{1, \ldots, e^{nR_2}\} \triangleq \mathcal{W}_2$. The channel state sequence S^n, which is drawn i.i.d. according to the law p_S, is available non-causally to Encoder 2. It is further assumed that Encoder 2 "cribs" causally and learns the sequence of channel inputs emitted by Encoder 1 in all past transmissions before generating its next channel input. The model is depicted in Figure 1.

An (e^{nR_1}, e^{nR_2}, n) code for the state-dependent MAC with a *strictly causal* cribbing encoder consists of:

1) Encoder 1 defined by a mapping $f_1 \colon \mathcal{W}_1 \to \mathcal{X}_1^n$.
2) Encoder 2 defined by a collection of encoding functions

$$f_{2,k} \colon \mathcal{W}_2 \times \mathcal{S}^n \times \mathcal{X}_1^{k-1} \to \mathcal{X}_2 \quad k = 1, 2, \ldots, n.$$

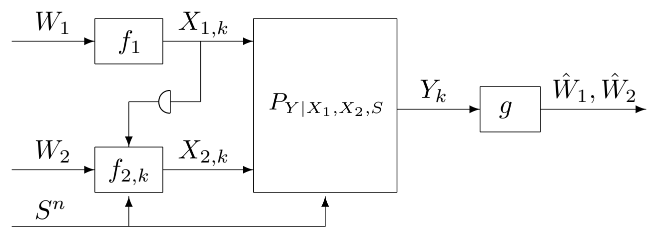

Fig. 1. State-dependent MAC with a cribbing encoder.

3) The receiver defined by a mapping $g \colon \mathcal{Y}^n \to \mathcal{W}_1 \times \mathcal{W}_2$.

For a given code, the block average probability of error is $P_e^{(n)} = e^{-n(R_1+R_2)} \sum_{w_1=1}^{e^{nR_1}} \sum_{w_2=1}^{e^{nR_2}} P_e^{(n)}(w_1, w_2)$, where

$$P_e^{(n)}(w_1, w_2) = \Pr\left\{(\hat{w}_1, \hat{w}_2) \neq (w_1, w_2) | (w_1, w_2) \text{ sent}\right\}.$$

Here $\hat{w}_1$ and $\hat{w}_2$ are the receiver's guesses of the transmitted messages.

A rate-pair (R_1, R_2) is said to be achievable if there exists a sequence of (e^{nR_1}, e^{nR_2}, n) codes with $\lim_{n\to\infty} P_e^{(n)} = 0$. The capacity region of the state-dependent MAC with a cribbing encoder is the closure of the set of achievable rate-pairs.

Our main result is.

Theorem 1. *Consider the discrete memoryless state-dependent MAC $(\mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{S}, p(y|x_1, x_2, s), \mathcal{Y})$ with state-sequence available non-causally at a strictly causal cribbing encoder and finite alphabets $\mathcal{S}, \mathcal{X}_1, \mathcal{X}_2$. The capacity region of this channel is*

$$\begin{aligned}\mathcal{C} = \bigcup_{p_{VSUX_1X_2Y}} \Big\{ &(R_1, R_2) : \\ &0 \leq R_1 \leq H(X_1|V) \\ &0 \leq R_2 \leq I(U; Y|VX_1) - I(U; S|V) \\ &0 \leq R_1 + R_2 \leq I(VUX_1; Y) - I(U; S|V) \Big\}, \end{aligned} \qquad (1)$$

where the union in (1) is over all laws of the form

$$p_{VSUX_1X_2Y}(v,s,u,x_1,x_2,y)$$
$$= p_V(v)p_S(s)p_{X_1|V}(x_1|v)p_{UX_2|SV}(u,x_2|s,v)p(y|x_1,x_2,s). \tag{2}$$

The cardinalities of the auxiliary r.v.'s V and U are bounded by $|\mathcal{V}| \leq |\mathcal{X}_1||\mathcal{X}_2||\mathcal{S}|+5$, and $|\mathcal{U}| \leq |\mathcal{X}_1||\mathcal{X}_2||\mathcal{S}||\mathcal{V}|+2$.

II. Proofs

A. Proof of the achievability part in Theorem 1

We propose a coding scheme that is based on Block-Markov superposition encoding and which combines the coding technique of [1] with that of [2]. The decoder uses backward decoding.

1) Coding Scheme: We consider B blocks, each of n symbols. A sequence of $B-1$ message pairs $(W_1^{(b)}, W_2^{(b)})$, for $b=1,\ldots,B-1$, will be transmitted during B transmission blocks. Here the sequence $\{W_1^{(b)}\}$ is an i.i.d. sequence of uniform random variables over $\{1,\ldots,e^{nR_1}\}$, and, independent thereof, $\{W_2^{(b)}\}$ is an i.i.d. sequence of uniform random variables over $\{1,\ldots,e^{nR_2}\}$. As $B\to\infty$, for fixed n, the rate pair of the message (W_1,W_2), $(\tilde{R}_1,\tilde{R}_2) = (R_1(B-1)/B, R_2(B-1)/B)$, is arbitrarily close to (R_1,R_2).

We assume a tuple of random variables $V\in\mathcal{V}, S\in\mathcal{S}, U\in\mathcal{U}, X_1\in\mathcal{X}_1, X_2\in\mathcal{X}_2, Y\in\mathcal{Y}$, of joint law (2).

Random coding and partitioning: In each block $b, b=1,2,\ldots,B$, we use the following construction.

- Generate e^{nR_1} sequences $\boldsymbol{v}=(v_1,\ldots,v_n)$, each with probability $\Pr(\boldsymbol{v})=\prod_{k=1}^n p_V(v_k)$. Label them $\boldsymbol{v}(\omega_0)$ where $\omega_0\in\{1,\ldots,e^{nR_1}\}$.
- For each $\boldsymbol{v}(\omega_0)$ generate e^{nR_1} sequences $\boldsymbol{x}_1=(x_{1,1},x_{1,2},\ldots,x_{1,n})$, each with probability $\Pr(\boldsymbol{x}_1|\boldsymbol{v}(\omega_0))=\prod_{k=1}^n p_{X_1|V}(x_{1,k}|v_k(\omega_0))$. Label them $\boldsymbol{x}_1(i,\omega_0), i\in\{1,\ldots,e^{nR_1}\}$.
- For each $\boldsymbol{v}(\omega_0)$ generate $e^{n(R_2+R')}$ sequences $\boldsymbol{u}=(u_1,u_2,\ldots,u_n)$, each with probability $\Pr(\boldsymbol{u}|\boldsymbol{v}(\omega_0))=\prod_{k=1}^n p_{U|V}(u_k|v_k(\omega_0))$. Randomly partition the set $\{\boldsymbol{u}\}$ into e^{nR_2} bins, each consisting of $e^{nR'}$ codewords. Now label the codewords by $\boldsymbol{u}(j,\jmath,\omega_0), j\in\{1,\ldots,e^{nR_2}\}, \jmath\in\{1,\ldots,e^{nR'}\}$ where j identifies the bin and $\jmath$ the index within the bin.

Encoding : We denote the realizations of the sequences $\{W_1^{(b)}\}$ and $\{W_2^{(b)}\}$ by $\{w_1^{(b)}\}$ and $\{w_2^{(b)}\}$, and the realization of the state sequence $(S_1^{(b)},S_2^{(b)},\ldots,S_n^{(b)})$ by $\boldsymbol{s}^{(b)}$. The code builds upon a Block-Markov structure in which the message $(w_1^{(b)},w_2^{(b)})$ is encoded over the successive blocks b and $(b+1)$ such that, $\omega_0^{(b+1)}=w_1^{(b)}$, for $b=1,\ldots,B-1$.

The messages $\{w_1^{(b)}\}$ and $\{w_2^{(b)}\}$, $b=1,2,\ldots,B-1$ are encoded as follows:

In Block 1 the encoders send

$$\boldsymbol{x}_1^{(1)}=\boldsymbol{x}_1(w_1^{(1)},1)$$
$$\boldsymbol{x}_2^{(1)}=\boldsymbol{x}_2(\boldsymbol{s}^{(1)},w_2^{(1)},1),$$

where $\boldsymbol{x}_2(\boldsymbol{s}^{(b)},w_2^{(b)},\omega_0^{(b)})$ is defined in two steps:

1) Find the typical $\boldsymbol{u}(w_2^{(b)},\jmath_0,\omega_0^{(b)})$: Search within the bin $\boldsymbol{u}(w_2^{(b)},\cdot,\omega_0^{(b)})$ for the lowest $\jmath_0\in\{1,\ldots,e^{nR'}\}$ such that $\boldsymbol{u}(w_2^{(b)},\jmath_0,\omega_0^{(b)})$ is jointly typical with the pair $(\boldsymbol{v}(\omega_0^{(b)}),\boldsymbol{s}^{(b)})$; denote this $\jmath_0$ as $\jmath_0(\boldsymbol{s}^{(b)},w_2^{(b)},\omega_0^{(b)})$. If such $\jmath_0$ is not found or if the state sequence $\boldsymbol{s}^{(b)}$ is non-typical an error is declared and $\jmath_0(\boldsymbol{s}^{(b)},w_2^{(b)},\omega_0^{(b)})=1$.
2) Generate the codeword $\boldsymbol{x}_2(\boldsymbol{s}^{(b)},w_2^{(b)},\omega_0^{(b)})$ by drawing its components i.i.d. conditionally on the triple $(\boldsymbol{s}^{(b)},\boldsymbol{u}(w_2^{(b)},\jmath_0,\omega_0^{(b)}),\boldsymbol{v}(\omega_0^{(b)}))$, where the conditional law is induced by (2).

Suppose that, as a result of cribbing from Encoder 1, before the beginning of Block $b=2,3,\ldots,B$, Encoder 2 has an estimate $\hat{\hat{w}}_1^{(b-1)}$ for $w_1^{(b-1)}$. Then, in Block $b=2,3,\ldots,B-1$, the encoders send

$$\boldsymbol{x}_1^{(b)}=\boldsymbol{x}_1(w_1^{(b)},w_1^{(b-1)})$$
$$\boldsymbol{x}_2^{(b)}=\boldsymbol{x}_2(\boldsymbol{s}^{(b)},w_2^{(b)},\hat{\hat{w}}_1^{(b-1)}),$$

and in Block B

$$\boldsymbol{x}_1^{(B)}=\boldsymbol{x}_1(1,w_1^{(B-1)})$$
$$\boldsymbol{x}_2^{(B)}=\boldsymbol{x}_2(\boldsymbol{s}^{(B)},1,\hat{\hat{w}}_1^{(B-1)}).$$

Decoding at the receiver: After the reception of Block B the receiver uses backward decoding starting from Block B to Block 1 and decodes the messages as follows.

In Block B the receiver looks for $\hat{w}_1^{(B-1)}$ such that

$$\Big(\boldsymbol{v}(\hat{w}_1^{(B-1)}),\boldsymbol{x}_1(1,\hat{w}_1^{(B-1)}),\boldsymbol{u}(1,\jmath_0,\hat{w}_1^{(B-1)}),$$
$$\boldsymbol{x}_2(\boldsymbol{s}^{(B)},1,\hat{w}_1^{(B-1)}),\boldsymbol{y}^{(B)}\Big)\in\mathcal{A}_\epsilon(V,X_1,U,X_2,Y),$$

where $\jmath_0=\jmath_0(\boldsymbol{s}^{(B)},1,\hat{w}_1^{(B-1)})$.

Next, assume that, decoding backwards up to (and including) Block $b+1$, the receiver decoded $\hat{w}_1^{(B-1)},(\hat{w}_2^{(B-1)},\hat{w}_1^{(B-2)}),\ldots,(\hat{w}_2^{(b+1)},\hat{w}_1^{(b)})$. To decode Block b, the receiver looks for $(\hat{w}_2^{(b)},\hat{w}_1^{(b-1)})$ such that

$$\Big(\boldsymbol{v}(\hat{w}_1^{(b-1)}),\boldsymbol{x}_1(\hat{w}_1^{(b)},\hat{w}_1^{(b-1)}),\boldsymbol{u}(\hat{w}_2^{(b)},\jmath_0,\hat{w}_1^{(b-1)}),$$
$$\boldsymbol{x}_2(\boldsymbol{s}^{(b)},\hat{w}_2^{(b)},\hat{w}_1^{(b-1)}),\boldsymbol{y}^{(b)}\Big)\in\mathcal{A}_\epsilon(V,X_1,U,X_2,Y),$$

where $\jmath_0=\jmath_0(\boldsymbol{s}^{(b)},\hat{w}_2^{(b)},\hat{w}_1^{(b-1)})$.

Decoding at Encoder 2: To obtain cooperation, after Block $b=1,2,\ldots,B-1$, Encoder 2 chooses $\tilde{w}_1^{(b)}$ such that

$$\Big(\boldsymbol{v}(\tilde{\omega}_0^{(b)}),\boldsymbol{x}_1(\tilde{w}_1^{(b)},\tilde{\omega}_0^{(b)}),\boldsymbol{x}_1^{(b)}\Big)\in\mathcal{A}_\epsilon(V,X_1,X_1),$$

where $\tilde{\omega}_0^{(b)}=\tilde{w}_1^{(b-1)}$ was determined at the end of Block $b-1$ and $\tilde{\omega}_0^{(1)}=1$.

When a decoding step either fails to recover a unique index (or index pair) which satisfies the decoding rule, or there is more than one index (or index pair), then an index (or an index pair) is chosen at random.

2) Bounding the Probability of Error: Genie-aided arguments as in [3] and [4] can be used to show that the probability that either Endoder 2 makes an encoding error or the receiver makes a decoding error after Block b in the above scheme is upper bounded by the probability that at least one of the following events $E_0^{(b)} - E_5^{(b)}$ happens.

Error events:

- $E_0^{(b)}$:
$$\left(\boldsymbol{v}(\omega_0^{(b)}), \boldsymbol{u}(w_2^{(b)}, \jmath_0, \omega_0^{(b)}), \boldsymbol{x}_1(w_1^{(b)}, \omega_0^{(b)})\right) \notin \mathcal{A}_\epsilon(V, U, X_1).$$
- $E_1^{(b)}$: There exists $\tilde{w}_1 \neq w_1^{(b)}$ such that
$$\left(\boldsymbol{v}(\omega_0^{(b)}), \boldsymbol{x}_1(\tilde{w}_1, \omega_0^{(b)}), \boldsymbol{x}_1^{(b)}\right) \in \mathcal{A}_\epsilon(V, X_1, X_1).$$
- $E_2^{(b)}$: There doesn't exist $\jmath_0 \in \{1, \ldots, e^{nR'}\}$ such that
$$\left(\boldsymbol{v}(\omega_0^{(b)}), \boldsymbol{u}(w_2^{(b)}, \jmath_0, \omega_0^{(b)}), \boldsymbol{s}^{(b)}\right) \in \mathcal{A}_\epsilon(V, U, S).$$
- $E_3^{(b)}$:
$$\Big(\boldsymbol{v}(\omega_0^{(b)}), \boldsymbol{u}(w_2^{(b)}, \jmath_0, \omega_0^{(b)}), \boldsymbol{x}_1(w_1^{(b)}, \omega_0^{(b)}), \boldsymbol{x}_2(\boldsymbol{s}^{(b)}, w_2^{(b)}, w_1^{(b-1)}), \boldsymbol{y}^{(b)}\Big) \notin \mathcal{A}_\epsilon(V, U, X_1, X_2, Y).$$
- $E_4^{(b)}$: There exists $\tilde{\omega}_0 \neq \omega_0^{(b)}$ such that
$$\Big(\boldsymbol{v}(\tilde{\omega}_0^{(b)}), \boldsymbol{x}_1(w_1^{(b)}, \tilde{\omega}_0^{(b)}), \boldsymbol{u}(j, \jmath_0, \tilde{\omega}_0^{(b)}), \boldsymbol{x}_2(\boldsymbol{s}^{(b)}, j, \tilde{\omega}_0^{(b)}), \boldsymbol{y}^{(b)}\Big) \in \mathcal{A}_\epsilon(V, U, X_1, X_2, Y),$$
for some pair $(j, \jmath_0)$, $j \in \mathcal{W}_2$, $\jmath_0 \in \{1, \ldots, e^{nR'}\}$.
- $E_5^{(b)}$: There exists $\tilde{w}_2 \neq w_2^{(b)}$ such that
$$\Big(\boldsymbol{v}(\omega_0^{(b)}), \boldsymbol{x}_1(w_1^{(b)}, \omega_0^{(b)}), \boldsymbol{u}(\tilde{w}_2, \jmath_0, \omega_0^{(b)}), \boldsymbol{x}_2(\boldsymbol{s}^{(b)}, \tilde{w}_2, \omega_0^{(b)}), \boldsymbol{y}^{(b)}\Big) \in \mathcal{A}_\epsilon(V, U, X_1, X_2, Y),$$
for some index $\jmath_0 \in \{1, \ldots, e^{nR'}\}$.

We define the event
$$F_1^{(b)} \triangleq \bigcup_{j=4}^{5} E_j^{(b)}, \qquad b = 1, \ldots, B,$$
the event
$$F_2 \triangleq \bigcup_{j=1}^{B} E_0^{(b)},$$
the event
$$F_3 \triangleq \bigcup_{j=1}^{B} \left(E_0^{(b)} \cup E_1^{(b)}\right),$$
the event
$$F_4 \triangleq \bigcup_{j=1}^{B} \left(E_0^{(b)} \cup E_1^{(b)} \cup E_2^{(b)}\right),$$
and the event
$$F_4 \triangleq \bigcup_{j=1}^{B} \left(E_0^{(b)} \cup E_1^{(b)} \cup E_2^{(b)} \cup E_3^{(b)}\right).$$

We can upper bound the average probability of error $\bar{P}_e$ averaged over all codebooks and all random partitions by
$$\begin{aligned}\bar{P}_e \leq & \sum_{b=1}^{B} \left\{\Pr\left[E_0^{(b)}\right] + \Pr\left[E_1^{(b)} | {F_2}^c, {E_1^{(1\ldots b-1)}}^c\right]\right\} \\ & + \sum_{b=1}^{B} \left\{\Pr\left[E_2^{(b)} | F_3^c\right] + \Pr\left[E_3^{(b)} | {F_4}^c, {E_3^{(1\ldots b-1)}}^c\right]\right\} \\ & + \sum_{b=1}^{B} \Pr\left[F_1^{(b)} | {F_4}^c, {F_1^{(b+1\ldots B)}}^c\right],\end{aligned}$$
where ${F^{(1\ldots b-1)}}^c$ denotes the complement of the event $F^{(1)} \cup \ldots \cup F^{(b-1)}$.

Furthermore, we can upper bound each of the summands in the last sum as
$$\begin{aligned}& \Pr\left(F_1^{(b)} | {F_4}^c, {F_1^{(b+1\ldots B)}}^c\right) \\ & = \Pr\left(\bigcup_{j=4}^{5} E_j^{(b)} | {F_4}^c, {F_1^{(b+1\ldots B)}}^c\right) \\ & \leq \Pr\left(E_4^{(b)} | {F_4}^c, {F_1^{(b+1\ldots B)}}^c\right) \\ & \quad + \Pr\left(E_5^{(b)} | {F_4}^c, {F_1^{(b+1\ldots B)}}^c\right).\end{aligned}$$

In the following we separately examine each of the above summands.

By the AEP $\Pr\left[E_3^{(b)} | {F_4}^c, {E_3^{(1\ldots b-1)}}^c\right]$ and $\Pr\left[E_0^{(b)}\right]$ can be made arbitrarily small for sufficiently large n.

Also,

- If
$$R_1 < H(X_1|V), \tag{3}$$
then $\Pr\left[E_1^{(b)} | {F_2}^c, {E_1^{(1\ldots b-1)}}^c\right]$ can be made arbitrarily small, provided that n is sufficiently large;
- If
$$R_1 + R_2 + R' < I(VUX_1; Y), \tag{4}$$
then $\Pr\left(E_4^{(b)} | {F_4}^c, {F_1^{(b+1\ldots B)}}^c\right)$ can be made arbitrarily small, provided that n is sufficiently large;

- If

$$R_2 + R' < I(U;Y|VX_1) \quad (5)$$

then $\Pr\left(E_5^{(b)} \middle| {F_4}^c, {F_1^{(b+1\ldots B)}}^c\right)$ can be made arbitrarily small, provided that n is sufficiently large;

Finally, by the covering lemma (See [6], [7], [8] or [9, Chapter 13]), if

$$R' > I(U;S|V) \quad (6)$$

then $\Pr\left[E_2^{(b)} \middle| {F_3}^c\right]$ can be made arbitrarily small, provided that n is sufficiently large.

The combination of (3), (4), (5), and (6) establishes the achievability of the rate region (1) for a law of the form (2).

B. Proof of the converse in Theorem 1

Consider an (e^{nR_1}, e^{nR_2}, n) code with average block error probability $P_e^{(n)}$, and a law on $\mathcal{W}_1 \times \mathcal{W}_2 \times \mathcal{X}_1^n \times \mathcal{X}_2^n \times \mathcal{Y}^n \times \mathcal{S}^n$ given by

$$\begin{aligned} &p_{W_1W_2X_1^nX_2^nS^nY^n} = p_{W_1}p_{W_2}p_{S^n}I_{\{X_1^n=f_1(W_1)\}} \\ &\cdot p_{X_2^n|W_1W_2S^n}\prod_{k=1}^{n} p_{Y_k|X_{1,k}X_{2,k}S_k}\ , \end{aligned} \quad (7)$$

where

$$p_{X_2^n|W_1W_2S^n} = \prod_{k=1}^{n} I_{\{X_{2,k}=f_{2,k}(W_2,X_1^{k-1},S^n)\}}. \quad (8)$$

Let V_k and U_k be the random variables defined by

$$V_k \triangleq X_1^{k-1}, \quad (9)$$

$$U_k \triangleq W_2Y^{k-1}S_{k+1}^n. \quad (10)$$

We start with an upper bound on R_1 by following similar steps as in [2, Section V—Converse for situation 2].

$$\begin{aligned} nR_1 &= H(W_1|W_2) \\ &= I(W_1;Y^n|W_2) + H(W_1|W_2Y^n) \\ &\leq I(W_1;Y^n|W_2) + n\delta(P_e) \\ &\overset{(a)}{=} I(X_1^n;Y^n|W_2) + n\delta(P_e) \\ &= \sum_{k=1}^{n} I(X_{1,k};Y^n|W_2X_1^{k-1}) + n\delta(P_e) \\ &\leq \sum_{k=1}^{n} H(X_{1,k}|X_1^{k-1}) + n\delta(P_e) \\ &= \sum_{k=1}^{n} H(X_{1,k}|V_k) + n\delta(P_e). \end{aligned} \quad (11)$$

where (a) follows from the encoding rule at Encoder 1.

Next, consider R_2:

$$\begin{aligned} nR_2 &= H(W_2|W_1) \\ &\leq I(W_2;Y^n|W_1) + n\delta(P_e) \\ &= \sum_{k=1}^{n} I(W_2;Y_k|W_1Y^{k-1}) + n\delta(P_e) \\ &\leq \sum_{k=1}^{n} I(W_2Y^{k-1};Y_k|W_1) + n\delta(P_e) \\ &= \sum_{k=1}^{n} \big[I(W_2Y^{k-1}S_{k+1}^n;Y_k|W_1) \\ &\quad -I(S_{k+1}^n;Y_k|W_1W_2Y^{k-1})\big] + n\delta(P_e) \\ &\overset{(b)}{=} \sum_{k=1}^{n} \big[I(W_2Y^{k-1}S_{k+1}^n;Y_k|W_1) \\ &\quad -I(Y^{k-1};S_k|W_1W_2S_{k+1}^n)\big] + n\delta(P_e) \\ &\overset{(c)}{=} \sum_{k=1}^{n} \big[I(W_2Y^{k-1}S_{k+1}^n;Y_k|W_1) \\ &\quad -I(W_2Y^{k-1}S_{k+1}^n;S_k|W_1)\big] + n\delta(P_e) \\ &\overset{(d)}{=} \sum_{k=1}^{n} \big[I(W_2Y^{k-1}S_{k+1}^n;Y_k|W_1X_1^{k-1}X_{1,k}) \\ &\quad -I(W_2Y^{k-1}S_{k+1}^n;S_k|W_1X_1^{k-1})\big] + n\delta(P_e) \\ &\overset{(e)}{=} \sum_{k=1}^{n} \big[I(W_2Y^{k-1}S_{k+1}^n;Y_k|X_1^{k-1}X_{1,k}) \\ &\quad -I(W_2Y^{k-1}S_{k+1}^n;S_k|W_1X_1^{k-1})\big] + n\delta(P_e) \\ &\overset{(f)}{=} \sum_{k=1}^{n} \big[I(W_2Y^{k-1}S_{k+1}^n;Y_k|X_1^{k-1}X_{1,k}) \\ &\quad -I(W_2Y^{k-1}S_{k+1}^n;S_k|X_1^{k-1})\big] + n\delta(P_e) \\ &= \sum_{k=1}^{n} \left[I(U_k;Y_k|V_kX_{1,k}) - I(U_k;S_k|V_k)\right]. \end{aligned} \quad (12)$$

Here,

- (b) follows by the Csiszár-Körner's identity [5, Lemma 7];
- (c) follows since $(W_2S_{k+1}^n)$ is independent of S_k given W_1;
- (d) follows by the encoding rule at Encoder 1.
- (e) follows since $W_1 \multimap X_{1,k}X_1^{k-1} \multimap W_2Y_kY^{k-1}S_{k+1}^n$ and $W_1 \multimap X_{1,k}X_1^{k-1} \multimap Y_k$ are Markov strings; and
- (f) follows since $W_1 \multimap X_1^{k-1} \multimap W_2S_kY^{k-1}S_{k+1}^n$ is a Markov string.

Finally, we consider the sum-rate $R_1 + R_2$:

$$\begin{aligned} n(R_1+R_2) &= H(W_1W_2) \\ &\leq I(W_1W_2;Y^n) + n\delta(P_e) \\ &= \sum_{k=1}^{n} I(W_1W_2;Y_k|Y^{k-1}) + n\delta(P_e) \\ &\overset{(g)}{\leq} \sum_{k=1}^{n} \big[I(W_1W_2Y^{k-1}S_{k+1}^n;Y_k) \\ &\quad -I(W_1W_2Y^{k-1}S_{k+1}^n;S_k)\big] + n\delta(P_e) \end{aligned}$$

$$= \sum_{k=1}^{n} \big[I(W_1 X_1^{k-1} X_{1,k} W_2 Y^{k-1} S_{k+1}^n; Y_k) - I(W_1 W_2 Y^{k-1} S_{k+1}^n; S_k)\big] + n\delta(P_e)$$

$$\stackrel{(h)}{=} \sum_{k=1}^{n} \big[I(X_1^{k-1} X_{1,k} W_2 Y^{k-1} S_{k+1}^n; Y_k) - I(W_1 W_2 Y^{k-1} S_{k+1}^n; S_k)\big] + n\delta(P_e)$$

$$= \sum_{k=1}^{n} [I(V_k U_k X_{1,k}; Y_k) - I(W_1; S_k) - I(W_2 Y^{k-1} S_{k+1}^n; S_k | W_1)] + n\delta(P_e)$$

$$= \sum_{k=1}^{n} [I(V_k U_k X_{1,k}; Y_k) - I(W_1; S_k) - I(W_2 Y^{k-1} S_{k+1}^n; S_k | W_1 X_1^{k-1})] + n\delta(P_e)$$

$$\stackrel{(i)}{=} \sum_{k=1}^{n} [I(V_k U_k X_{1,k}; Y_k) - I(U_k; S_k | V_k)]. \quad (13)$$

Here,

- (g) follows by the same procedure as (b) and (c);
- (h) follows by the encoding rule at Encoder 1 and since $W_1 \multimap X_{1,k} X_1^{k-1} W_2 Y^{k-1} S_{k+1}^n \multimap Y_k$ is a Markov string; and
- (i) follows since W_1 is independent of S_k and since $W_1 \multimap X_1^{k-1} W_2 Y^{k-1} S_{k+1}^n \multimap S_k$ and $W_1 \multimap X_1^{k-1} \multimap S_k$ are Markov strings.

Next we verify the joint law of the auxiliary random variables.

By (7) and the encoding rule at Encoder 1 we may write

$$p_{W_1 W_2 X_1^{k-1} X_{1,k} S^{k-1} S_k S_{k+1}^n X_2^k Y^{k-1}} = p_{W_1} p_{X_1^{k-1}|W_1} P_{X_{1,k}|W_1 X_1^{k-1}} p_{S^{k-1}} p_{S_k} p_{S_{k+1}^n} \cdot p_{W_2} p_{X_2^k | W_2 X_1^{k-1} S^n} p_{Y^{k-1}|X_1^{k-1} X_2^{k-1} S^{k-1}}$$

Summing this joint law over w_1 we obtain

$$\sum_{w_1} p_{W_1 W_2 X_1^{k-1} X_{1,k} S^{k-1} S_k S_{k+1}^n X_2^k Y^{k-1}} = p_{W_2 X_1^{k-1} X_{1,k} S^{k-1} S_k S_{k+1}^n X_2^k Y^{k-1}} = p_{X_1^{k-1}} P_{X_{1,k}|X_1^{k-1}} p_{S^{k-1}} p_{S_k} p_{S_{k+1}^n} \cdot p_{W_2} p_{X_2^k | W_2 X_1^{k-1} S^n} p_{Y^{k-1}|X_1^{k-1} X_2^{k-1} S^{k-1}}$$

Summing this joint law over all possible sub-sequences $(s_1, s_2, \ldots, s_{k-1})$ we obtain

$$\sum_{(s_1, s_2, \ldots, s_{k-1})} p_{W_2 X_1^{k-1} X_{1,k} S^{k-1} S_k S_{k+1}^n X_2^k Y^{k-1}} = p_{W_2 X_1^{k-1} X_{1,k} S_k S_{k+1}^n X_2^k Y^{k-1}} = p_{X_1^{k-1}} P_{X_{1,k}|X_1^{k-1}} p_{S_k} p_{S_{k+1}^n} \cdot p_{W_2} p_{X_2^k | W_2 X_1^{k-1} S_k S_{k+1}^n} p_{Y^{k-1}|X_1^{k-1} X_2^{k-1}}$$

This establishes the Markov relation

$$X_{2,k} U_k \multimap S_k V_k \multimap X_{1,k}. \quad (14)$$

Next, let J be a r.v. uniformly distributed over $\{1, \ldots, n\}$ and independent of $(X_{1,k}, X_{2,k}, V_k, U_k, S_k, Y_k)$, $k = 1, \ldots, n$, and define

$$(S, X_1, X_2, V, U, Y) = (S_J, X_{1,J}, X_{2,J}, V_J, U_J, Y_J).$$

We may express (11) as follows

$$R_1 \le \frac{1}{n} \sum_{k=1}^{n} H(X_{1,k}|V_k) = H(X_1|V, J) = H(X_1|\bar{V}). \quad (15)$$

where in the last step we defined $\bar{V} \triangleq (V, J)$.
Similarly, we may express (12) as follows

$$R_2 \le \frac{1}{n} \sum_{k=1}^{n} [I(U_k; Y_k | V_k X_{1,k}) - I(U_k; S_k | V_k)] = I(U; Y | V, X_1, J) - I(U; S | V, J) = I(U; Y | \bar{V}, X_1) - I(U; S | \bar{V}), \quad (16)$$

Finally, we may express (13) as follows

$$R_1 + R_2 \le \frac{1}{n} \sum_{k=1}^{n} [I(V_k U_k X_{1,k}; Y_k) - I(U_k; S_k | V_k)] = I(V, U, X_1; Y | J) - I(U; S | V, J) = I(V, J, U, X_1; Y) - I(J; Y) - I(U; S | V, J) \le I(V, J, U, X_1; Y) - I(U; S | V, J) = I(\bar{V}, U, X_1; Y) - I(U; S | \bar{V}). \quad (17)$$

This establishes the single letter expression for the achievable rate region (1). The convexity of the rate region (1) can be shown in a similar way.

Inequalities (11), (12), (13) combined with their respective single-letter expressions and the Markov relation (14) establish the converse part of Theorem 1.

Acknowledgment

The work of S. Bross was supported by the Israel Science Foundation, grant no. 497/09.

References

[1] A. Somekh-Baruch, S. Shamai (Shitz) and S. Verdú, "Cooperative multiple-access encoding with states available at one transmitter," *IEEE Trans. Inform. Theory,* vol. IT-54, no. 10, pp. 4448-4469, Oct. 2008.

[2] F.M.J. Willems and E.C. van der Meulen, "The discrete memoryless multiple-access channel with cribbing encoders", *IEEE Trans. Inform. Theory,* vol. IT-31, no. 3, pp. 313-327, May 1985.

[3] B. Rimoldi and R. Urbanke, "A rate-splitting approach to the Gaussian multiple-access channel", *IEEE Trans. Inform. Theory,* vol. IT-42, no. 2, pp. 364 – 375, Mar 1996.

[4] J. M. Wozencraft and I. M. Jacobs, *Principles of Communication Engineering.* John Wiley & Sons, 1965.

[5] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Inform. Theory,* vol. IT-24, No. 3 pp. 339-348, May 1978.

[6] R. Ahlswede and J. Körner, "Source coding with side information and a converse for degraded broadcast channels," *IEEE Trans. Inform. Theory,* vol. IT-21, No. 6, pp. 629-637, Nov. 1975.

[7] A. D. Wyner, "On source coding with side information at the decoder," *IEEE Trans. Inform. Theory,* vol. IT-21, No. 6, pp. 294-300, May 1975.

[8] T. Berger, "Multiterminal source coding." Lecture notes presented at the 1977 CISM Summer School, Udine, Italy, July 18-20, 1977, Springer-Verlag.

[9] T. M. Cover and J. A. Thomas, "*Elements of Information Theory,*" Wiley, 1991.