Mathematics

https://www.aimspress.com/journal/Math

AIMS Mathematics, 10(7): 15821–15840.
DOI: 10.3934/math.2025709
Received: 14 February 2025
Revised: 10 June 2025
Accepted: 13 June 2025
Published: 11 July 2025

Research article

Causal state quantization with or without cribbing for the MAC with a state-cognizant receiver

Amos Lapidoth* **and Baohua Ni**

Department of Information Technology and Electrical Engineering, ETH Zurich, 8092 Zurich, Switzerland

* **Correspondence:** Email: lapidoth@isi.ee.ethz.ch.

Abstract: A two-to-one memoryless state-dependent multiple-access channel is studied in a setting where helpers that observe the state sequence causally provide rate-limited (lossy) causal descriptions of it to the encoders. It is shown that, when the receiver is cognizant of the channel state, the optimal causal descriptions take the form of scalar symbol-by-symbol quantizers whose descriptions of the current state do not depend on the past states. This holds irrespective of whether the description is provided to both encoders (the "common description" architecture), to one of the encoders (the "one-sided description" architecture), or whether different descriptions are provided to the different encoders (the "general" architecture). In fact, in the common description architecture, it also holds when the encoders can crib. Thus, the prevalent assumption that the receiver is cognizant of the channel state greatly simplifies the design of the state quantizer.

Keywords: capacity region; causal state information; helper; MAC; multiple-access channel; rate-limited help; state-dependent; state information; state quantization
Mathematics Subject Classification: 94A15, 94A34

1. Introduction

How to causally describe the state of a channel succinctly to an encoder and how the latter should utilize the description to maximize throughput, is a challenging problem even for single-user channels [1], let alone for multiple-access channels (MACs). Indeed, as shown in [1, Claim 10], even on memoryless single-user channels with memoryless states, symbol-by-symbol (scalar) state quantizers need not be optimal. Moreover, on the MAC it is not even known how the encoders should utilize the state description when the latter is provided to them perfectly: Shannon strategies—memoryless mappings of states to channel inputs—are optimal for single-user channels [2] but not for MACs [3–5], and the question of how the encoders of a MAC should instead utilize perfect causal state information is open. (For a survey of the literature on state-dependent channels with unquantized state information, see [6].)

However, there is a ray of hope. On the single-user channel, the problem of state quantization is greatly simplified when the receiver—as in many wireless applications—is cognizant of the (unquantized) channel state [1, Theorem 6]. Could this also be the case on the MAC? Here, we answer this question in the affirmative and demonstrate that, in this setting, scalar quantizers are optimal and the descriptions are optimally utilized using Shannon strategies. To some extent, this is also the case when the encoders can crib à la Willems and Van der Meulen [7–9].

When it comes to the MAC, one can consider a number of different helper architectures depending on which encoder is provided with which description of which aspect of the state sequence. The "common description" architecture provides the *same description* of the state sequence to *both encoders*; see Figure 1 (without the dashed lines). The "one-sided description" architecture provides only one of the encoders with a description of the state [10–13]. And the "general" architecture allows for the different encoders to be provided with different descriptions of different aspects of the state; see Figure 2. Specifically, in the general architecture, we envision that the time-i channel state S_i is a triple

$$(S_{0,i}, S_{1,i}, S_{2,i}) \tag{1.1}$$

whose components are, in general, dependent. The description of $\{S_{0,i}\}_{i=1}^n$ is provided to both encoders; the description of $\{S_{1,i}\}_{i=1}^n$ is only provided to Encoder 1; and the description of $\{S_{2,i}\}_{i=1}^n$ is only provided to Encoder 2. We denote the time-i descriptions

$$T_i = (T_{0,i}, T_{1,i}, T_{2,i}) \tag{1.2}$$

with the understanding that $T_{0,i}$ describes $\{S_{0,j}\}_{j=1}^i$ and is provided at time-i to both encoders, $T_{1,i}$ describes $\{S_{1,j}\}_{j=1}^i$ and is provided at time-i to Encoder 1, and $T_{2,i}$ describes $\{S_{2,j}\}_{j=1}^i$ and is provided at time-i to Encoder 2. Note that all three descriptions are causal:

$$T_{1,i} = t_{1,i}(S_1^i), \quad T_{2,i} = t_{2,i}(S_2^i), \quad T_{0,i} = t_{0,i}(S_0^i). \tag{1.3}$$

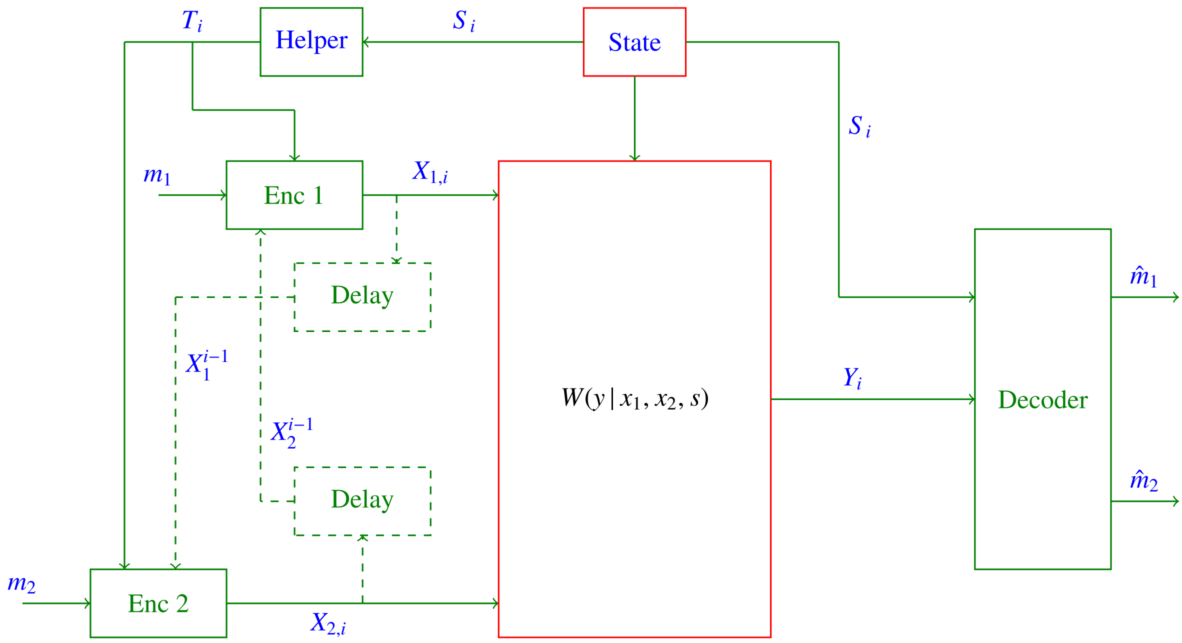

Figure 1. State-dependent MAC with the "common description" architecture.

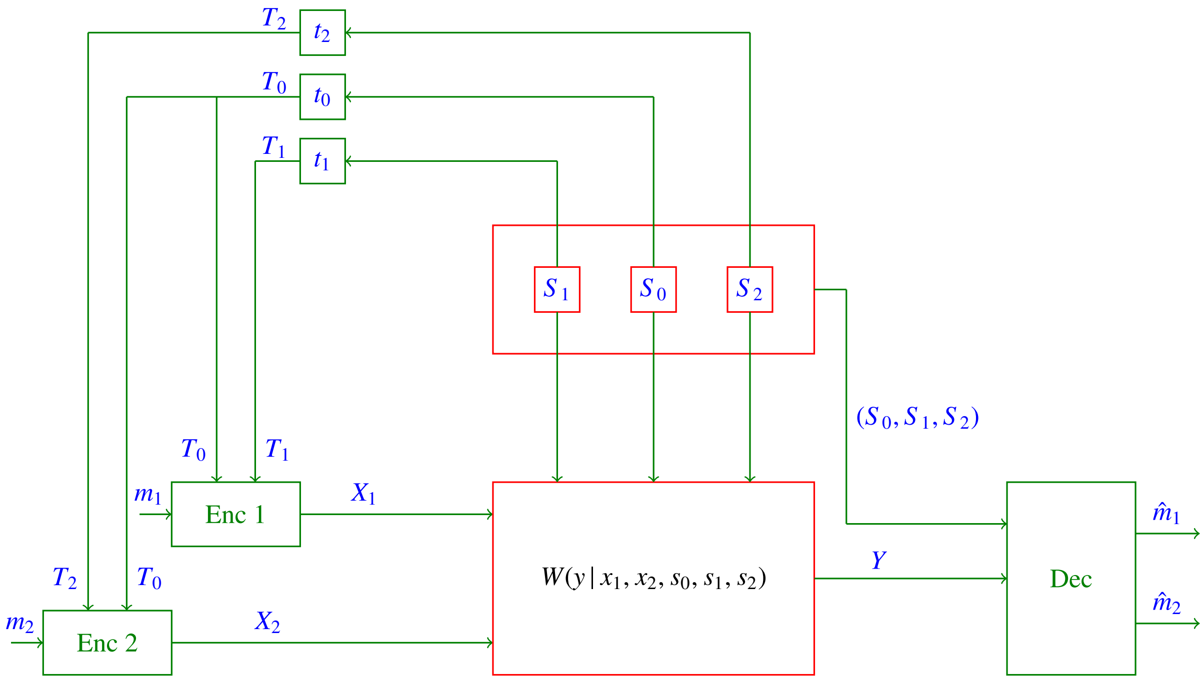

Figure 2. State-dependent MAC with the "general" architecture.

Here and throughout, a length-i sequence of random variables $A_1, \ldots, A_i$ is denoted A^i, and this convention is extended to double indices, so $X_{1,1}, X_{1,2}, \ldots, X_{1,i}$ is denoted X_1^i.

To account for rate constraints, we require that the descriptions take values in alphabets of given cardinalities: $T_{0,i}$ must take values in the (possibly empty) set $\mathcal{T}_0$ whose cardinality is denoted $|\mathcal{T}_0|$, while $T_{1,i}$ and $T_{2,i}$ must take values in the sets $\mathcal{T}_1$ and $\mathcal{T}_2$ whose cardinalities are denoted $|\mathcal{T}_1|$ and $|\mathcal{T}_2|$, respectively. Thus, the mappings in (1.3) have the following forms:

$$t_{0,i}\colon \mathcal{S}_0^{\times i} \to \mathcal{T}_0 \tag{1.4a}$$
$$t_{1,i}\colon \mathcal{S}_1^{\times i} \to \mathcal{T}_1 \tag{1.4b}$$
$$t_{2,i}\colon \mathcal{S}_2^{\times i} \to \mathcal{T}_2 \tag{1.4c}$$

where $\mathcal{S}_0$, $\mathcal{S}_1$, and $\mathcal{S}_2$ are the sets in which $S_{0,i}$, $S_{1,i}$, and $S_{2,i}$ take values, respectively, and $\mathcal{A}^{\times i}$ denotes the i-fold Cartesian product $\mathcal{A} \times \cdots \times \mathcal{A}$.

The common description architecture corresponds to the special case of the general architecture where S_1 and S_2 are null (or $|\mathcal{T}_1| = |\mathcal{T}_2| = 1$). The one-sided architecture (to Encoder 1) corresponds to the case where S_0 and S_2 are null (or $|\mathcal{T}_0| = |\mathcal{T}_2| = 1$).

The encoders produce channel inputs that are based on the message they wish to transmit and on the state descriptions they have received thus far. Denoting the messages that Encoder 1 and Encoder 2 wish to transmit M_1 and M_2, respectively , we denote the time-i channel symbols they produce

$$X_{1,i} = X_{1,i}(M_1, T_0^i, T_1^i), \qquad X_{2,i} = X_{2,i}(M_2, T_0^i, T_2^i). \tag{1.5}$$

Our main result is that when the state is known to the decoder, there is no loss of optimality in restricting the mappings in (1.3) to "time-dependent scalar state quantizers," i.e., to having the form

$$T_{0,i} = t_{0,i}(S_{0,i}), \quad T_{1,i} = t_{1,i}(S_{1,i}), \quad T_{2,i} = t_{2,i}(S_{2,i}), \tag{1.6}$$

so that the time-i descriptions only depend on the time-i state and not on the past states. Moreover, this quantized state information is optimally utilized using Shannon strategies. This allows us to solve for the capacity region.

We also consider a scenario where the encoders can crib in the sense introduced by Willems and Van der Meulen [7]. This is not to be confused with the setting considered in [6], where it is the helper that does the cribbing. For this scenario, we only consider the common description setting; see Figure 1 with the dashed lines. Thus, we assume that the encoders have the form

$$X_{1,i} = X_{1,i}(M_1, T_0^i, X_2^{i-1}), \qquad X_{2,i} = X_{2,i}(M_2, T_0^i, X_1^{i-1}) \tag{1.7}$$

because the channel input produced by each encoder at time-i may now also depend on the past channel inputs produced by the other encoder. Here, X_2^{i-1} stands for the past symbols produced by Encoder 2, i.e., $X_{2,1}, X_{2,2}, \ldots, X_{2,i-1}$ and likewise X_1^{i-1}.

Our main result for this setting is that, here too, if the decoder is cognizant of the channel state, then there is no loss in capacity in restricting the common description to correspond to a time-dependent scalar state quantizer, i.e., to having the form

$$T_{0,i}(S_0^i) = t_{0,i}(S_{0,i}). \tag{1.8}$$

Once this restriction is made, we can solve for the capacity region. (Even with this restriction, the results do not follow directly from Asnani and Permuter [8], because the cribbing encoders do not see a known fixed function of the strategy employed by the other user: what they see is the result of applying said strategy to the assistance.)

For typographical reasons, when discussing cribbing with the common description architecture, we shall drop the subscript 0 from S_0 and T_0 and use S and T instead. Thus, we denote the time-i state S_i, and the description of S^i that is provided to both encoders T_i.

The rest of the paper is organized as follows: We conclude this introductory section with some notations; Section 2 defines the capacity regions we seek; the results on noncribbing encoders with general architectures are presented in Section 3; and those on cribbing encoders and common descriptions in Section 4.

1.1. Notations

When k is a positive integer, $[1 : k]$ denotes the set $\{1, \ldots, k\}$. More generally, if a is any positive number, then $[1 : a]$ denotes the (possibly empty) set $\{1, 2, \ldots, \lfloor a \rfloor\}$. Upper-case calligraphic fonts are used to denote sets (e.g., $\mathcal{S}$ for the set of possible channel states). All the sets that are denoted using calligraphic fonts are assumed to be finite. The cardinality of a set $\mathcal{A}$ is denoted $|\mathcal{A}|$ (so $|\mathcal{S}|$ denotes the number of channel states). A random variable that takes values in the set $\mathcal{A}$ is usually denoted A and its Probability Mass Function (PMF) P_A (so the random channel state that takes values in $\mathcal{S}$ is denoted S and its PMF P_S). A generic realization of such a random variable is usually denoted a (so $P_S(s)$ is the probability that the state S equals s). Given a PMF P_X and a conditional PMF $P_{Y|X}$, we write $P_X\, P_{Y|X}$ for the joint PMF that assigns to the pair (x, y) the probability $P_X(x)\, P_{Y|X}(y|x)$. The n-fold Cartesian products of a set $\mathcal{A}$ with itself is denoted $\mathcal{X}^{\times n}$ and comprises all n-tuples with entries from $\mathcal{A}$. The indicator function of an event $\mathcal{E}$ is denoted $\mathbb{I}\{\mathcal{E}\}$ and equals 1 if $\mathcal{E}$ holds and equals 0 otherwise. For example, $\mathbb{I}\{x = y\}$ equals 1 if $x = y$, and equals 0 otherwise.

2. Problem setup

We are given a discrete memoryless two-to-one state-dependent multiple-access channel (SD-MAC) whose input alphabets $\mathcal{X}_1$, $\mathcal{X}_2$; its output alphabet $\mathcal{Y}$; and its state alphabet $\mathcal{S}$ are all finite. The time-i state is denoted S_i, and it is assumed throughout that the state sequence $\{S_i\}$ comprises independent and identically distributed (IID) random variables of some given PMF P_S on $\mathcal{S}$. Conditional on the state sequence, the channel behaves like a MAC of inputs X_1 and X_2 and of channel law $P_{Y|X_1X_2S}$. We refer to the pair $(P_{Y|X_1,X_2,S}, P_S)$ as "the channel".

For a fixed rate pair (R_1, R_2), the message sets for blocklength-n communications are denoted $\mathcal{M}_1 = [1 : 2^{nR_1}]$ and $\mathcal{M}_2 = [1 : 2^{nR_2}]$. We assume that the message pair to be transmitted (M_1, M_2) is drawn equiprobably from $\mathcal{M}_1 \times \mathcal{M}_2$, and that M_1 and M_2 are thus a fortiori statistically independent. We also assume throughout that the decoder observes the channel output sequence Y^n and is cognizant of the state sequence S^n. Thus, it is a mapping of the form

$$\phi\colon \mathcal{Y}^{\times n} \times \mathcal{S}^{\times n} \to \mathcal{M}_1 \times \mathcal{M}_2. \tag{2.1}$$

We say that an error occurs if the tuple produced by the decoder differs from the transmitted tuple (M_1, M_2). Thus, the probability of error $P_{\mathrm{e}}^{(n)}$ is

$$P_{\mathrm{e}}^{(n)} = \Pr\Big(\phi(Y^n, S^n) \neq (M_1, M_2)\Big) \tag{2.2}$$

with the understanding that, in the above expression, Y^n are the channel outputs produced when the message pair (M_1, M_2) is presented to the encoders. The structure of the encoders depends on the setting and will be discussed next.

2.1. *No cribbing and general architecture*

When discussing general helping architectures, we assume that the state alphabet $\mathcal{S}$ has the form $\mathcal{S}_0 \times \mathcal{S}_1 \times \mathcal{S}_2$ and denote the time-i state S_i or $(S_{0,i}, S_{1,i}, S_{2,i})$. For the description of the state, we are given three finite description alphabets: $\mathcal{T}_0$, $\mathcal{T}_1$, and $\mathcal{T}_2$. A blocklength-n helping scheme comprises n helping triples

$$\{(t_{0,i}, t_{1,i}, t_{2,i})\}_{i=1}^{n} \tag{2.3}$$

where, for each $i \in [1:n]$,

$$t_{0,i} \colon \mathcal{S}_0^{\times i} \to \mathcal{T}_0 \tag{2.4a}$$

$$t_{1,i} \colon \mathcal{S}_1^{\times i} \to \mathcal{T}_1 \tag{2.4b}$$

$$t_{2,i} \colon \mathcal{S}_2^{\times i} \to \mathcal{T}_2. \tag{2.4c}$$

The time-i assistance T_i is the triple $(T_{0,i}, T_{1,i}, T_{2,i})$, which is given by

$$T_i = \Big(t_{0,i}(S_0^i), t_{1,i}(S_1^i), t_{2,i}(S_2^i)\Big) \in \mathcal{T}_0 \times \mathcal{T}_1 \times \mathcal{T}_2. \tag{2.5}$$

We sometimes write T_i as $T_i(S^i)$ and $\mathcal{T}_0 \times \mathcal{T}_1 \times \mathcal{T}_2$ as $\mathcal{T}$.

In defining the capacity region, we have restricted ourselves to deterministic helpers. More generally, we could consider stochastic helpers whose time-i description is determined not only by the state up-to time-i, but also on some source of randomization Θ that is shared by the helpers and over which the probability of error $P_e^{(n)}$ is averaged. Allowing for such helpers does not increase the capacity region, because, given such stochastic helpers, we can consider the realization $\theta^\star$ that minimizes the conditional probability of error given $\Theta = \theta$ and replace the stochastic helpers with the deterministic helpers that correspond to Θ being deterministically equal to $\theta^\star$. Thus:

Remark 1. *Allowing for stochastic helpers does not increase the capacity region.*

The help is *(time-dependent) symbol-by-symbol* if $T_i(S^i)$ is determined by i and S_i (i.e., if the time-i assistance can be determined from the time-i state, with the past states being irrelevant). Symbol-by-symbol help thus takes the form

$$t_{0,i} \colon \mathcal{S}_0 \to \mathcal{T}_0 \tag{2.6a}$$

$$t_{1,i} \colon \mathcal{S}_1 \to \mathcal{T}_1 \tag{2.6b}$$

$$t_{2,i} \colon \mathcal{S}_2 \to \mathcal{T}_2, \tag{2.6c}$$

with

$$T_i = \Big(t_{0,i}(S_{0,i}), t_{1,i}(S_{1,i}), t_{2,i}(S_{2,i})\Big) \in \mathcal{T}_0 \times \mathcal{T}_1 \times \mathcal{T}_2. \tag{2.7}$$

We now turn to the encoders. The symbol each encoder produces at time-i is determined by the message it wishes to convey and by the assistance it has received thus far. Encoder 1 is thus specified by n functions $\{f_{1,i}\}_{i\in[1:n]}$, where, for each $i \in [1:n]$, the function $f_{1,i}$ has the form

$$f_{1,i} : \mathcal{M}_1 \times \mathcal{T}_0^{\times i} \times \mathcal{T}_1^{\times i} \to \mathcal{X}_1. \tag{2.8a}$$

Likewise, Encoder 2 is specified by functions $\{f_{2,i}\}_{i\in[1:n]}$, with

$$f_{2,i} : \mathcal{M}_2 \times \mathcal{T}_0^{\times i} \times \mathcal{T}_2^{\times i} \to \mathcal{X}_2. \tag{2.8b}$$

The time-i symbol produced by Encoder 1 in order to convey Message M_1 after having observed the present-and-past assistance T_0^i, T_1^i is

$$X_{1,i} = f_{1,i}(M_1, T_0^i, T_1^i) \tag{2.9}$$

and likewise

$$X_{2,i} = f_{2,i}(M_2, T_0^i, T_2^i). \tag{2.10}$$

Given a channel $(P_{Y|X_1,X_2,S}, P_S)$ and description alphabets $\mathcal{T}_0, \mathcal{T}_1, \mathcal{T}_2$, we say that the pair (R_1, R_2) is achievable (without cribbing) if, for every $\epsilon > 0$, there exist, for every blocklength n, helping functions as in (2.4), encoders $\{f_{1,i}, f_{2,i}\}_{i=1}^n$ of rates exceeding $(R_1 - \epsilon, R_2 - \epsilon)$ as in (2.8), and a decoding function as in (2.1) such that

$$\lim_{n\to\infty} P_{\mathrm{e}}^{(n)} = 0. \tag{2.11}$$

The capacity region comprises the achievable rate pairs.

2.2. *Cribbing and common description*

For the case where the encoders can crib, we only consider the common description architecture. The state S (which now need not have the form (1.1)) is presented to the helper who describes it using the finite description alphabet $\mathcal{T}$. To that end, the helper uses n functions

$$\left\{t_i : \mathcal{S}^{\times i} \to \mathcal{T}\right\}_{i=1}^n \tag{2.12}$$

to produce the descriptions

$$T_i = t_i(S^i) \in \mathcal{T}. \tag{2.13}$$

The help is *(time-dependent) symbol-by-symbol* if $T_i(S^i)$ is determined by i and S_i and thus has the form

$$t_i : \mathcal{S} \to \mathcal{T} \tag{2.14}$$

and

$$T_i = t_i(S_i). \tag{2.15}$$

The time-i symbol produced by each cribbing encoder may depend not only on the message it wishes to transmit and the help it has received thus far, but also on the past symbols produced by the other encoder. Thus, for every $i \in [1:n]$, the cribbing encoders have the form

$$f_{1,i} : \mathcal{M}_1 \times \mathcal{T}^{\times i} \times \mathcal{X}_2^{\times(i-1)} \to \mathcal{X}_1 \tag{2.16a}$$
$$f_{2,i} : \mathcal{M}_2 \times \mathcal{T}^{\times i} \times \mathcal{X}_1^{\times(i-1)} \to \mathcal{X}_2 \tag{2.16b}$$

and

$$X_{1,i} = f_{1,i}(M_1, T^i, X_2^{i-1}) \tag{2.17a}$$
$$X_{2,i} = f_{2,i}(M_2, T^i, X_1^{i-1}). \tag{2.17b}$$

The definition of an achievable rate pair and of the capacity region is analogous to that in the absence of cribbing; the only difference is that the encoding functions now have the form (2.16).

3. No cribbing and general architecture

3.1. The capacity region

Let $\mathcal{C}^{(I)}$ denote the union over all the joint PMFs $P_{QST_0T_1T_2X_1X_2Y}$ of the form

$$P_S\, P_Q\, P_{T_0|S_0,Q}\, P_{T_1|S_1,Q}\, P_{T_2|S_2,Q}\, P_{X_1|T_0,T_1,Q}\, P_{X_2|T_0,T_2,Q}\, P_{Y|X_1,X_2,S} \tag{3.1}$$

of the pentagons

$$\left\{ (R_1,R_2)\in\mathbb{R}_{\geq 0}^{\times 2} \;\middle|\; \begin{aligned} R_1 &\leq I(X_1;Y\,|\,X_2,T,S,Q) \\ R_2 &\leq I(X_2;Y\,|\,X_1,T,S,Q) \\ R_1+R_2 &\leq I(X_1,X_2;Y\,|\,T,S,Q) \end{aligned} \right\}, \tag{3.2}$$

where Q is an auxiliary random variable of finite support. Since

$$\begin{pmatrix} I(X_1;Y\,|\,X_2,T,S,Q) \\ I(X_2;Y\,|\,X_1,T,S,Q) \\ I(X_1,X_2;Y\,|\,T,S,Q) \end{pmatrix} = \sum_{q\in\mathcal{Q}} P_Q(q)\cdot \begin{pmatrix} I(X_1;Y_1\,|\,X_2,T,S,Q=q) \\ I(X_2;Y\,|\,X_1,T,S,Q=q) \\ I(X_1,X_2;Y\,|\,T,S,Q=q) \end{pmatrix}, \tag{3.3}$$

we can infer from Carathéodory Theorem for connected subsets of $\mathbb{R}^3$ (e.g., [14, Appendix 4A]):

Remark 2. *Without altering the union, we may impose the constraint*

$$|\mathcal{Q}| \leq 3. \tag{3.4}$$

Moreover, by imposing this constraint, we can use a compactness-and-continuity argument to infer:

Remark 3. *The mapping*

$$P_{QSTX_1X_2Y} \mapsto \Big(I(X_1;Y_1\,|\,X_2,T,S,Q),\; I(X_2;Y\,|\,X_1,T,S,Q),\; I(X_1,X_2;Y\,|\,T,S,Q) \Big) \tag{3.5}$$

maps the family of PMFs of the form (3.1) to a compact convex subset of $\mathbb{R}^3$.

Theorem 1. *The capacity region of the SD-MAC with causal helpers and non-cribbing encoders is the set $\mathcal{C}^{(I)}$. Moreover, all rate pairs in this region can be achieved using (time-dependent) symbol-by-symbol helpers.*

3.2. Converse

First, we prove the converse. Given any coding/helping scheme with $P_e^{(n)}$ tending to zero, we consider its behavior when the messages are independent and equiprobably distributed to obtain:

$$\begin{aligned} n\,R_1 &= H(M_1) && (3.6a) \\ &= H(M_1\,|\,M_2) - H(M_1\,|\,Y^n,S^n,M_2) + H(M_1\,|\,Y^n,S^n,M_2) && (3.6b) \\ &\leq H(M_1\,|\,M_2) - H(M_1\,|\,Y^n,S^n,M_2) + 1 + n\,R_1\,P_e^{(n)} && (3.6c) \\ &\leq I(M_1;Y^n,S^n\,|\,M_2) + n\delta_n && (3.6d) \end{aligned}$$

$$= \sum_{i=1}^{n} I(M_1; Y_i, S_i \,|\, M_2, Y^{i-1}, S^{i-1}) + n\delta_n \tag{3.6e}$$

$$= \sum_{i=1}^{n} I(M_1; S_i \,|\, M_2, Y^{i-1}, S^{i-1}) + I(M_1; Y_i \,|\, M_2, Y^{i-1}, S^{i}) + n\delta_n \tag{3.6f}$$

$$= \sum_{i=1}^{n} I(M_1; Y_i \,|\, M_2, Y^{i-1}, S^{i}) + n\delta_n \tag{3.6g}$$

$$= \sum_{i=1}^{n} I(M_1; Y_i \,|\, M_2, Y^{i-1}, S^{i}, X_{2,i}) + n\delta_n \tag{3.6h}$$

$$= \sum_{i=1}^{n} I(M_1, X_{1,i}; Y_i \,|\, M_2, X_{2,i}, Y^{i-1}, S^{i}) + n\delta_n \tag{3.6i}$$

$$\leq \sum_{i=1}^{n} I(M_1, M_2, X_{1,i}, Y^{i-1}; Y_i \,|\, X_{2,i}, S^{i}) + n\delta_n \tag{3.6j}$$

$$= \sum_{i=1}^{n} I(M_1, M_2, Y^{i-1}; Y_i \,|\, X_{1,i}, X_{2,i}, S^{i}) + I(X_{1,i}; Y_i \,|\, X_{2,i}, S_i, S^{i-1}) + n\delta_n \tag{3.6k}$$

$$= \sum_{i=1}^{n} I(X_{1,i}; Y_i \,|\, X_{2,i}, S_i, S^{i-1}) + n\delta_n \tag{3.6l}$$

$$= \sum_{i=1}^{n} I(X_{1,i}; Y_i \,|\, X_{2,i}, T_i, S_i, S^{i-1}) + n\delta_n \tag{3.6m}$$

$$= \sum_{i=1}^{n} I(X_{1,i}; Y_i \,|\, X_{2,i}, T_i, S_i, Q_i) + n\delta_n, \tag{3.6n}$$

where in (3.6n) we define

$$Q_i \triangleq S^{i-1}. \tag{3.7}$$

Here, (3.6c) follows from Fano's inequality (and the fact that the probability of incorrectly guessing M_1 is upper bounded by the probability of incorrectly guessing the pair (M_1, M_2)), in (3.6d) we define $\delta_n \triangleq n^{-1}(1 + P_e^{(n)}\, n\,(R_1 + R_2))$, (3.6g) follows from the Markov chain $M_1 \multimap\!\!- (M_2, Y^{i-1}, S^{i-1}) \multimap\!\!- S_i$, (3.6h) holds because $X_{2,i}$ can be determined from (M_2, S^i), (3.6i) holds because $X_{1,i}$ is determined by (M_1, S^i), (3.6l) holds because $I(M_1, M_2, Y^{i-1}; Y_i \,|\, X_{1,i}, X_{2,i}, S^i) = 0$, as can be seen from the channel law, and (3.6m) holds because T_i is a deterministic function of S_i, S^{i-1}. Thus,

$$R_1 - \delta_n \leq \frac{1}{n} \sum_{i=1}^{n} I(X_{1,i}; Y_i \,|\, X_{2,i}, T_i, S_i, Q_i) \tag{3.8a}$$

and, by symmetry,

$$R_2 - \delta_n \leq \frac{1}{n} \sum_{i=1}^{n} I(X_{2,i}; Y_i \,|\, X_{1,i}, T_i, S_i, Q_i). \tag{3.8b}$$

We next show that

$$R_1 + R_2 - \delta_n \leq \frac{1}{n} \sum_{i=1}^{n} I(X_{1,i}, X_{2,i}; Y_i \,|\, T_i, S_i, Q_i) \tag{3.8c}$$

from which the converse will follow from Remark 3 upon letting $n \to \infty$ (with δ_n thus tending to zero).

The proof of (3.8c) is similar to the proof of [1, Lemma 2], except for the presence of $Q_i = S^{i-1}$. For completeness, we provide the following proof:

$$
\begin{aligned}
n(R_1 + R_2) &= H(M_1, M_2) && (3.9a)\\
&\leq H(M_1, M_2) - H(M_1, M_2 \,|\, Y^n, S^n) + 1 + P_e^{(n)}\, n\,(R_1 + R_2) && (3.9b)\\
&= \sum_{i=1}^{n} I(M_1, M_2; Y_i, S_i \,|\, Y^{i-1}, S^{i-1}) + n\delta_n && (3.9c)\\
&= \sum_{i=1}^{n} I(M_1, M_2; S_i \,|\, Y^{i-1}, S^{i-1}) + I(M_1, M_2; Y_i \,|\, Y^{i-1}, S^{i}) + n\delta_n && (3.9d)\\
&= \sum_{i=1}^{n} I(M_1, M_2; Y_i \,|\, Y^{i-1}, S^{i}) + n\delta_n && (3.9e)\\
&= \sum_{i=1}^{n} I(X_{1,i}, X_{2,i}, M_1, M_2; Y_i \,|\, Y^{i-1}, S^{i}) + n\delta_n && (3.9f)\\
&\leq \sum_{i=1}^{n} I(X_{1,i}, X_{2,i}, M_1, M_2, Y^{i-1}; Y_i \,|\, S^{i}) + n\delta_n && (3.9g)\\
&= \sum_{i=1}^{n} I(X_{1,i}, X_{2,i}; Y_i \,|\, S^{i}) + n\delta_n && (3.9h)\\
&= \sum_{i=1}^{n} I(X_{1,i}, X_{2,i}; Y_i \,|\, T_i, S^{i}) + n\delta_n && (3.9i)\\
&= \sum_{i=1}^{n} I(X_{1,i}, X_{2,i}; Y_i \,|\, T_i, S_i, Q_i) + n\delta_n, && (3.9j)
\end{aligned}
$$

where (3.9f) holds because $X_{1,i}, X_{2,i}$ are determined by (S^i, M_1, M_2), and (3.9h) follows from the channel law.

Having established (3.8), we next verify that the joint PMF of the time-i random variables is as in (3.1). This can be verified by noting that in any valid helping and encoding scheme, $T_{0,i}$ is determined by $S_{0,i}$ and Q_i; $T_{1,i}$ is determined by $S_{1,i}$ and Q_i; $T_{2,i}$ is determined by $S_{2,i}$ and Q_i; $X_{1,i}$ is a function of $(M_1, T_{0,i}, T_{1,i}, Q_i)$; and $X_{2,i}$ is a function of $(M_2, T_{0,i}, T_{2,i}, Q_i)$ (with M_1 and M_2 being independent also conditionally on the past states and assistances).

The converse now follows from (3.8) and the desired form of the joint PMF by letting n tend to infinity (with δ_n hence tending to 0) and using Remark 3.

3.3. Achievability

Given some joint PMF of the form (3.1), it follows from the Functional Representation Lemma [14, Appendix B] that any (not necessarily zero-one valued) $P_{T_1|S_1,Q}$ can be implemented by having T_1 be of the form $T_1 = t_1(S_1, Q, \tilde{Q}_1)$ for some deterministic function $t_1(\cdot)$, where $\tilde{Q}_1$ is independent of all the random variables and has a bounded cardinality $|\tilde{Q}_1| \leq |\mathcal{T}_1|\,(|\mathcal{S}_1| - 1)$. Repeating this argument for T_0 and T_2 with $\tilde{Q}_0$ and $\tilde{Q}_2$ and merging these auxiliary random variables yields $\tilde{Q} = (\tilde{Q}_0, \tilde{Q}_1, \tilde{Q}_2)$ with a bounded cardinality $|\tilde{Q}| \leq |\mathcal{T}_0|\,(|\mathcal{S}_0| - 1)\,|\mathcal{T}_1|\,(|\mathcal{S}_1| - 1)\,|\mathcal{T}_2|\,(|\mathcal{S}_2| - 1)$.

Generate a time-sharing sequence $Q_1, Q_2, \ldots, Q_n$ IID $\sim P_Q$ and reveal its realization $q_1, q_2, \ldots, q_n$ to all parties. Independently of that sequence, generate a functional-representing sequence $\tilde{Q}_1, \tilde{Q}_2, \ldots, \tilde{Q}_n$ and reveal its realization $\tilde{q}_1, \tilde{q}_2, \ldots, \tilde{q}_n$ to the helpers and to the receiver but not to the encoders.

At each time instance $i \in [1 : n]$, let the assistance be

$$T_{0,i} = t_0(S_{0,i}, q_i, \tilde{q}_i) \tag{3.10a}$$
$$T_{1,i} = t_1(S_{1,i}, q_i, \tilde{q}_i) \tag{3.10b}$$
$$T_{2,i} = t_2(S_{2,i}, q_i, \tilde{q}_i). \tag{3.10c}$$

The values of $T_{0,i}$ and $T_{1,i}$ are provided to Encoder 1, those of $T_{0,i}$ and $T_{2,i}$ to Encoder 2, and all three can be computed by the receiver who is cognizant of S, Q, and $\tilde{Q}$.

Let $\mathcal{W}_1$ be a finite set, P_{W_1} a PMF on it, and $f_1 \colon \mathcal{W}_1 \times \mathcal{T}_0 \times \mathcal{T}_1 \times \mathcal{Q} \to \mathcal{X}_1$ a mapping for which the PMF of $f_1(W_1, \tau_0, \tau_1, q)$ is the given $P_{X_1|T_0,T_1,Q}(x_1|\tau_0, \tau_1, q)$ for all $(\tau_0, \tau_1, q) \in \mathcal{T}_0 \times \mathcal{T}_1 \times \mathcal{Q}$. Generate 2^{nR_1} independent n-tuples $\{(W_{1,1}(m_1), \ldots, W_{1,n}(m_1))\}_{m_1 \in \mathcal{M}_1}$ each with IID $\sim P_{W_1}$ components, and define $X_{1,i}(m_1)$ for every $i \in [1 : n]$ to be $f_1(W_{1,i}(m_1), T_{0,i}, T_{1,i}, Q_i)$. Similarly, choose P_{W_2} and $f_2(\cdot)$ so that $f_2(W_2, \tau_0, \tau_2, q)$ is the given $P_{X_2|T_0,T_2,Q}$; independently of $\{(W_{1,1}(m_1), \ldots, W_{1,n}(m_1))\}_{m_1 \in \mathcal{M}_1}$, generate 2^{nR_2} independent n-tuples $\{(W_{2,1}(m_2), \ldots, W_{2,n}(m_2))\}_{m_2 \in \mathcal{M}_2}$ each with IID $\sim P_{W_2}$ components; and set $X_{2,i}(m_2)$ to be $f_2(W_{2,i}(m_2), T_{0,i}, T_{2,i}, Q_i)$. To convey the messages m_1 and m_2, the encoders produce the n-tuples $(X_{1,1}(m_1), \ldots, X_{1,n}(m_1))$ and $(X_{2,1}(m_2), \ldots, X_{2,n}(m_2))$, respectively.

This coding scheme can be analyzed as a time-sharing scheme for a MAC whose inputs are W_1 and W_2 and whose output is (Y, S), where the decoder searches for the unique pair $(\hat{m}_1, \hat{m}_2)$ such that $(q^n, \tilde{q}^n, s^n, y^n, W_1^n(\hat{m}_1), W_2^n(\hat{m}_2))$ is jointly typical with respect to the corresponding marginal of

$$\begin{aligned} P_S\, P_Q\, P_{\tilde{Q}}\, P_{W_1}\, P_{W_2}\, P_{T_0|S_0,Q,\tilde{Q}}\, P_{T_1|S_1,Q,\tilde{Q}}\, P_{T_2|S_2,Q,\tilde{Q}}\, \mathbb{I}\{X_1 = f_1(W_1, T_0, T_1, Q)\} \\ \mathbb{I}\{X_2 = f_2(W_2, T_0, T_2, Q)\}\, P_{Y|X_1,X_2,S}. \end{aligned} \tag{3.11}$$

In this way, it can be shown that the scheme's probability of error decays to zero (as the blocklength n tends to infinity) whenever the rate tuple is in the pentagon

$$\left\{ (R_1, R_2) \in \mathbb{R}_{\geq 0}^{\times 2} \;\middle|\; \begin{aligned} R_1 &\leq I(W_1; Y, S \mid W_2, Q, \tilde{Q}) \\ R_2 &\leq I(W_2; Y, S \mid W_1, Q, \tilde{Q}) \\ R_1 + R_2 &\leq I(W_1, W_2; Y, S \mid Q, \tilde{Q}) \end{aligned} \right\}, \tag{3.12}$$

where the mutual information terms are computed for the joint distribution (3.11).

We next show that the set in (3.12) contains the pentagon (3.2) and thus conclude the achievability proof. To this end, it is helpful to consider the factor graph of the channel, which is depicted in Figure 3 (without Q, and assuming *symbol-by-symbol* assistance).

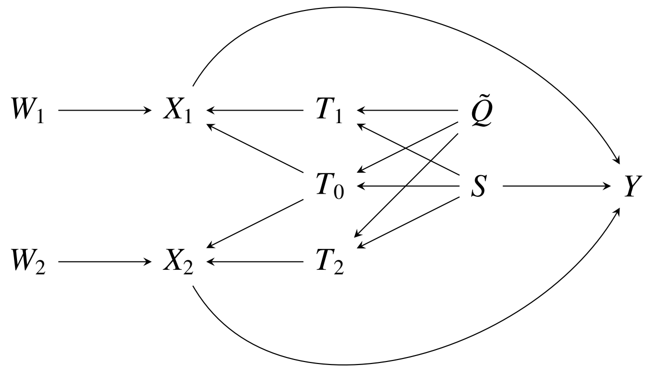

Figure 3. Factor graph of the state-dependent MAC with symbol-by-symbol assistance.

With S standing for (S_0, S_1, S_2) and T standing for (T_0, T_1, T_2),

$$\begin{aligned} I(W_1; Y, S \mid W_2, Q, \tilde{Q}) &= I(W_1; Y, S, T, W_2, \tilde{Q} \mid Q) && (3.13a) \\ &\geq I(W_1; Y, S, T, W_2 \mid Q) && (3.13b) \\ &= I(W_1; Y, S \mid W_2, T, Q) + I(W_1; W_2, T \mid Q) && (3.13c) \\ &= I(W_1; Y, S \mid W_2, T, Q) && (3.13d) \\ &= I(W_1; Y, S \mid W_2, X_2, T, Q) && (3.13e) \\ &= I(W_1; Y, S, W_2 \mid X_2, T, Q) - I(W_1; W_2 \mid X_2, T, Q) && (3.13f) \\ &= I(W_1; Y, S, W_2 \mid X_2, T, Q) && (3.13g) \\ &\geq I(W_1; Y, S \mid X_2, T, Q) && (3.13h) \\ &= I(W_1, X_1; Y, S \mid X_2, T, Q) && (3.13i) \\ &\geq I(X_1; Y \mid X_2, T, S, Q), && (3.13j) \end{aligned}$$

where (3.13a) holds because T is determined by $(S, Q, \tilde{Q})$, (3.13d) follows from the Markovity $W_1 \multimap\!\!- Q \multimap\!\!- (W_2, T)$, (3.13e) holds because X_2 is computable from (W_2, T), (3.13g) follows from the Markovity $W_1 \multimap\!\!- (X_2, T, Q) \multimap\!\!- W_2$, and (3.13i) holds because X_1 is determined by (W_1, T). By symmetry,

$$I(W_2; Y, S \mid W_1, Q, \tilde{Q}) \geq I(X_2; Y \mid X_1, T, S, Q). \tag{3.14}$$

As to the sum-rate constraint,

$$\begin{aligned} I(W_1, W_2; Y, S \mid Q, \tilde{Q}) &= I(W_1, W_2; Y, T, S, \tilde{Q} \mid Q) && (3.15a) \\ &\geq I(W_1, W_2; Y, T, S \mid Q) && (3.15b) \\ &\geq I(W_1, W_2; Y \mid T, S, Q) && (3.15c) \\ &= I(W_1, W_2, X_1, X_2; Y \mid T, S, Q) && (3.15d) \\ &\geq I(X_1, X_2; Y \mid T, S, Q) && (3.15e) \end{aligned}$$

where (3.15a) holds because T is determined by $(S, Q, \tilde{Q})$, and (3.15d) holds because (X_1, X_2) are determined by (W_1, W_2, T, Q). □

Parallel single-user channels: Next, we analyze a special case of Theorem 1 corresponding to a MAC that has the structure of two non-interfering state-dependent single-user channels. We show that, as expected, the capacity region corresponds to a rectangle, and we recover the single-user results of [1, Section II C].

The output of the MAC of parallel single-user channels is a tuple $Y = (Y_1, Y_2)$; the time-i state S_i is a tuple $(S_{1,i}, S_{2,i})$ with the state sequence $\{(S_{1,i}, S_{2,i})\}$ drawn IID according to some product distribution $P_S = P_{S_1} P_{S_2}$; and the channel law factorizes as

$$P_{Y_1,Y_2|X_1,X_2,S_1,S_2} = P_{Y_1|X_1,S_1} P_{Y_2|X_2,S_2}. \tag{3.16}$$

The state sequence S_1^n is described to Encoder 1, and the state sequence S_2^n to Encoder 2; there is no common description, i.e., $|\mathcal{T}_0| = 1$.

Remark 4. *The capacity region of the MAC with parallel single-user channels is the rectangle*

$$\left\{(R_1, R_2) \in \mathbb{R}_{\geq 0}^{\times 2} \middle| \begin{aligned} R_1 &\leq \max I(X_1; Y \,|\, S_1) \\ R_2 &\leq \max I(X_2; Y \,|\, S_2) \end{aligned} \right\}, \tag{3.17}$$

where the maxima are over all joint distributions having the forms $P_{S_1} P_{T_1|S_1} P_{X_1|T_1} P_{Y_1|X_1,S_1}$ and $P_{S_2} P_{T_2|S_2} P_{X_2|T_2} P_{Y_2|X_2,S_2}$, respectively, and where both $P_{T_1|S_1}$ and $P_{T_2|S_2}$ are zero-one valued.

Proof. The conditional independence of (X_1, S_1, T_1, Y_1) and (X_2, S_2, T_2, Y_2) given Q implies that the terms in (3.2) can be expressed as:

$$I(X_1; Y \,|\, X_2, S, Q) = I(X_1; Y_1 \,|\, T_1, S_1, Q) \tag{3.18}$$

$$I(X_2; Y \,|\, X_1, S, Q) = I(X_2; Y_2 \,|\, T_2, S_2, Q) \tag{3.19}$$

$$I(X_1, X_2; Y \,|\, S, Q) = I(X_1; Y_1 \,|\, T_1, S_1, Q) + I(X_2; Y_2 \,|\, T_2, S_2, Q). \tag{3.20}$$

Since the sum-rate bound is the sum of the marginal bounds, it is superfluous. This reduces the pentagon to the rectangle

$$\left\{(R_1, R_2) \in \mathbb{R}_{\geq 0}^{\times 2} \middle| \begin{aligned} R_1 &\leq I(X_1; Y_1 \,|\, T_1, S_1, Q) \\ R_2 &\leq I(X_2; Y_2 \,|\, T_2, S_2, Q) \end{aligned} \right\}. \tag{3.21}$$

The expectations over Q on the right can be replaced with maximizations by choosing Q to be deterministically equal to $(q_1^\star, q_2^\star)$, where $q_1^\star$ maximizes $I(X_1; Y_1 \,|\, T_1, S_1, Q = q)$, and where $q_2^\star$ maximizes $I(X_2; Y_2 \,|\, T_2, S_2, Q = q)$ (and by choosing the conditional PMFs given $Q = (q_1^\star, q_2^\star)$ accordingly). This allows us to eliminate Q from (3.21).

We next argue that we may restrict $P_{T_1|S_1}$ to be zero-one valued and may hence eliminate T_1 from (3.21). To this end, we express $I(X_1; Y_1 \,|\, T_1, S_1)$ (where Q has already been eliminated) as

$$I(X_1; Y_1 \,|\, T_1, S_1) = \sum_{s_1} P_{S_1}(s_1) \sum_{t_1} P_{T_1|S_1}(t_1|s_1)\, I(X_1; Y_1 \,|\, T_1 = t_1, S_1 = s_1). \tag{3.22}$$

If we fix $P_{X_1|T_1}$, then, for any given s_1, the term $I(X_1; Y_1 \,|\, T_1 = t_1, S_1 = s_1)$ can be viewed as a function of t_1. Thus, we can maximize the expectation over T_1 given S_1 by choosing T_1 to be a deterministic function of S_1, i.e., by having $P_{T_1|S_1}$ be zero-one valued. Repeating the argument for R_2 concludes the proof. □

An altruistic encoder: We now consider the special case of Theorem 1 when R_2 is zero. It was pointed out in [5] that—even when Encoder 2 has no message to send, i.e., when $R_2 = 0$—it can behave altruistically and aid the link to the receiver from Encoder 1 by using its input to convey to the decoder some information about the state. This point was further elaborated in [15]. However, in the present setting, the receiver is cognizant of the state, so this reasoning is inapplicable. Nevertheless, as we next show, the altruistic encoder still has an important role to play, and it should not send a constant symbol: its time-i inputs $X_{2,i}$ should depend on the state information at its disposal and should mitigate some of the detrimental effects the state may have on the link to the receiver from Encoder 1. For example, if the SD-MAC is such that its output Y is equal to X_1 (deterministically) whenever X_2 coincides with the state, and its output Y is otherwise independent of X_1, then the altruistic Encoder 2 should strive to transmit a symbol that matches the state and not a constant symbol. It is, however, true that, in order to be most helpful, the altruistic encoder should use a constant Shannon strategy, i.e., that X_2 be a deterministic function of (t_0, t_2). Moreover, the PMFs $P_{T_0|S_0}$, $P_{T_1|S_1}$, and $P_{T_2|S_2}$ can be chosen to be zero-one valued.

Corollary 1. In the setting of Theorem 1, let C_1 be the supremum over all rates R_1 for which $(R_1, 0)$ is achievable. Then,

$$C_1 = \max I(X_1; Y \mid S), \tag{3.23}$$

where the maximum is over the same family of PMFs as in the definition of $C^{(I)}$ of Theorem 1 but with Q deterministic, with X_2 being a deterministic function of (T_0, T_2), and with the conditional PMFs $P_{T_0|S_0}$, $P_{T_1|S_1}$, and $P_{T_2|S_2}$ all being zero-one valued.

Proof. Since $R_2 = 0$, the pinching constraint in (3.2) is the constraint on R_1 (and not on $R_1 + R_2$), so

$$C_1 = \max I(X_1; Y \mid X_2, T, S, Q) \tag{3.24}$$

where the maximum is over the same family of PMFs as in the definition of $C^{(I)}$. By replacing the expectation over Q on the right hand side (RHS) of (3.24) with a maximization, we conclude that Q may be chosen deterministic, so

$$C_1 = \max I(X_1; Y \mid X_2, T, S). \tag{3.25}$$

We next claim that, given $P_{X_1|T_0,T_1}$ and $P_{X_2|T_0,T_2}$, we may restrict $P_{T_0|S_0}$, $P_{T_1|S_1}$, and $P_{T_2|S_2}$ (formerly $P_{T_0|S_0,Q}$, $P_{T_1|S_1,Q}$, and $P_{T_2|S_2,Q}$) in the maximization in (3.25) to all be zero-one valued. To this end, we rewrite the RHS of (3.25) as

$$\begin{aligned} C_1 &= \max \sum_{s\in\mathcal{S}} \sum_{t_0\in\mathcal{T}_0} \sum_{t_1\in\mathcal{T}_1} \sum_{t_2\in\mathcal{T}_2} P_S(s)\, P_{T_0|S_0}(t_0 \mid s_0)\, P_{T_1|S_1}(t_1 \mid s_1)\, P_{T_2|S_2}(t_2 \mid s_2) \\ &\quad I(X_1; Y \mid X_2, T_0 = t_0, T_1 = t_1, T_2 = t_2, S = s) \qquad (3.26) \\ &= \max \sum_{s\in\mathcal{S}} \sum_{t_0\in\mathcal{T}_0} \sum_{t_1\in\mathcal{T}_1} \sum_{t_2\in\mathcal{T}_2} P_S(s)\, P_{T_0|S_0}(t_0 \mid s_0)\, P_{T_1|S_1}(t_1 \mid s_1)\, P_{T_2|S_2}(t_2 \mid s_2)\, h(t_0, t_1, t_2, s_0, s_1, s_2) \qquad (3.27) \end{aligned}$$

where in the last equation we have noted that, since $P_{X_1|T_0,T_1}$ and $P_{X_2|T_0,T_2}$ are fixed, the conditional mutual information in (3.26) is a function of $(t_0, t_1, t_2, s_0, s_1, s_2)$, which we have denoted $h(\cdot)$.

Since the RHS of (3.27) is linear and, *a fortiori*, convex in $P_{T_0|S_0}$, we may restrict the latter to be zero-one valued (because any conditional PMF can be expressed as a convex combination of zero-one valued conditional PMFs). An analogous argument shows that also $P_{T_1|S_1}$ and $P_{T_2|S_2}$ can be restricted to being zero-one valued. Thus, we can write T_0, T_1, and T_2 as $\tau_0(S_0)$, $\tau_1(S_1)$, and $\tau_2(S_2)$ respectively, where τ_0, τ_1, and τ_2 are deterministic functions.

We now fix these functions as well as $P_{X_1|T_0,T_1}$ and consider the maximization over $P_{X_2|T_0,T_2}$. We shall argue that the maximum is achieved by a zero-one PMF, i.e., by having X_2 be a deterministic function of (T_0, T_2). To this end, we express the objective function as

$$\sum_{t_0,t_2} \sum_{t_1} \sum_{\substack{s \in \mathcal{S}: \\ \tau_0(s_0)=t_0 \\ \tau_1(s_1)=t_1 \\ \tau_2(s_2)=t_2}} P_S(s) \sum_{x_2} P_{X_2|T_0,T_2}(x_2 \,|\, t_0, t_2)\, I(X_1; Y \,|\, X_2 = x_2, T = t, S = s) \tag{3.28}$$

or

$$\sum_{t_0,t_2} \sum_{x_2} P_{X_2|T_0,T_2}(x_2 \,|\, t_0, t_2) \Bigg(\sum_{t_1} \sum_{\substack{s \in \mathcal{S}: \\ \tau_0(s_0)=t_0 \\ \tau_1(s_1)=t_1 \\ \tau_2(s_2)=t_2}} P_S(s) I(X_1; Y \,|\, X_2 = x_2, T = t, S = s) \Bigg), \tag{3.29}$$

where, for each pair (t_0, t_2), the expression in the parentheses is only a function of x_2. The sum over x_2 thus has the form of an average, which is maximized by the maximum, i.e., by choosing X_2 to be a deterministic function of (t_0, t_2).

□

4. Cribbing and common description

4.1. The capacity region

We begin by recalling the result of Willems and Van der Meulen [7] on the capacity of the discrete memoryless MAC with cribbing encoders.

Theorem 2 (Willems and Van der Meulen'85)**.** *The capacity region of the MAC $P_{Y|X_1,X_2}$ with strictly-causal cribbing encoders is the union over all the joint PMFs $P_{UX_1X_2Y}$ of the form $P_U\, P_{X_1|U}\, P_{X_2|U}\, P_{Y|X_1,X_2}$ of the pentagons*

$$\left\{ (R_1, R_2) \in \mathbb{R}_{\geq 0}^{\times 2} \,\middle|\, \begin{aligned} R_1 &\leq H(X_1 \,|\, U) \\ R_2 &\leq H(X_2 \,|\, U) \\ R_1 + R_2 &\leq I(X_1, X_2; Y) \end{aligned} \right\}, \tag{4.1}$$

where the cardinality of the support of U may be restricted to satisfy $|\mathcal{U}| \leq \min\{|\mathcal{X}_1| \cdot |\mathcal{X}_2| + 1, |\mathcal{Y}| + 2\}$.

We now turn to the SD-MAC with a helper. Let $C_{\text{crib}}^{(I)}$ denote the union over all the joint PMFs $P_{QSTUX_1X_2Y}$ of the form

$$P_S\, P_Q\, P_{T|S,Q}\, P_{U|T}\, P_{X_1|U,T,Q}\, P_{X_2|U,T,Q}\, P_{Y|X_1,X_2,S} \tag{4.2}$$

of the pentagons

$$\left\{(R_1,R_2)\in\mathbb{R}_{\geq 0}^{\times 2}\left|\begin{aligned} R_1 &\leq H(X_1\,|\,U,T,S,Q)\\ R_2 &\leq H(X_2\,|\,U,T,S,Q)\\ R_1+R_2 &\leq I(X_1,X_2;Y\,|\,T,S,Q)\end{aligned}\right.\right\}, \tag{4.3}$$

where Q, U are auxiliary random variables of finite supports. Analogously to Remark 2 (and with a similar proof), we have:

Remark 5. *Without altering the union defining $C_{\text{crib}}^{(I)}$, we may impose the constraint*

$$|\mathcal{Q}| \leq 3. \tag{4.4}$$

Likewise, in analogy to Remark 3 (and with a similar proof), we have:

Remark 6. *The mapping*

$$P_{QSTUX_1X_2Y} \mapsto \Big(H(X_1\,|\,U,T,S,Q),\ H(X_2\,|\,U,T,S,Q),\ I(X_1,X_2;Y\,|\,T,S,Q)\Big) \tag{4.5}$$

maps the subset of PMFs of the form (4.2) to a compact convex subset of $\mathbb{R}^3$.

Theorem 3. *The capacity region of the SD-MAC with a common state description and cribbing encoders is the set $C_{\text{crib}}^{(I)}$. Moreover, all achievable rate pairs are achievable by a time-dependent symbol-by-symbol helper, and the cardinality of the auxiliary random variable U may be upper-bounded by $|\mathcal{U}| \leq \min\{|\mathcal{X}_1|\cdot|\mathcal{X}_2|+1, |\mathcal{Y}|+2\}$.*

4.2. Converse

We begin with the converse part. Given any coding/helping scheme with $P_{\text{e}}^{(n)}$ tending to zero, we consider its behavior when the messages are independent and drawn equiprobably from the respective message sets to obtain,

$$\begin{aligned}
nR_1 &= H(M_1) && \text{(4.6a)}\\
&\leq H(M_1\,|\,M_2) - H(M_1\,|\,Y^n,S^n,M_2) + 1 + nP_{\text{e}}^{(n)}R_1 && \text{(4.6b)}\\
&= I(M_1;Y^n,S^n\,|\,M_2) + n\delta_n && \text{(4.6c)}\\
&= I(M_1;Y^n\,|\,S^n,M_2) + n\delta_n && \text{(4.6d)}\\
&\leq I(M_1,X_1^n;Y^n\,|\,S^n,M_2) + n\delta_n && \text{(4.6e)}\\
&= H(Y^n\,|\,S^n,M_2) - H(Y^n\,|\,S^n,M_2,M_1,X_1^n) + n\delta_n && \text{(4.6f)}\\
&= H(Y^n\,|\,S^n,M_2) - H(Y^n\,|\,S^n,M_2,M_1,X_1^n,X_2^n) + n\delta_n && \text{(4.6g)}\\
&= H(Y^n\,|\,S^n,M_2) - H(Y^n\,|\,S^n,M_2,X_1^n,X_2^n) + n\delta_n && \text{(4.6h)}\\
&= H(Y^n\,|\,S^n,M_2) - H(Y^n\,|\,S^n,M_2,X_1^n) + n\delta_n && \text{(4.6i)}\\
&= I(X_1^n;Y^n\,|\,S^n,M_2) + n\delta_n && \text{(4.6j)}\\
&\leq H(X_1^n\,|\,S^n,M_2) + n\delta_n && \text{(4.6k)}\\
&= \sum_{i=1}^{n} H(X_{1,i}\,|\,S^n,M_2,X_1^{i-1}) + n\delta_n && \text{(4.6l)}
\end{aligned}$$

$$= \sum_{i=1}^{n} H(X_{1,i} \,|\, S^n, M_2, X_2^{i-1}, X_1^{i-1}) + n\delta_n \tag{4.6m}$$

$$= \sum_{i=1}^{n} H(X_{1,i} \,|\, S^{i-1}, S_i, X_2^{i-1}, X_1^{i-1}) + n\delta_n \tag{4.6n}$$

$$= \sum_{i=1}^{n} H(X_{1,i} \,|\, U_i, T_i, S_i, Q_i) + n\delta_n, \tag{4.6o}$$

where

$$U_i \triangleq (X_1^{i-1}, X_2^{i-1}), \qquad Q_i \triangleq S^{i-1}. \tag{4.7}$$

Here, (4.6d) holds because M_1, M_2, and S^n are independent; (4.6i) holds because X_2^n is determined by (S^n, M_2, X_1^n); (4.6j) holds because $M_1 \multimap X_1^n \multimap (S^n, X_2^n)$; and (4.6m) holds because X_2^{i-1} is determined by $(X_1^{i-2}, S^{i-1}, M_2)$.

Using symmetry, we can obtain a similar bound for $n R_2$. As for the sum rate, we can follow the steps we took in the proof in Theorem 1, specifically in (3.9), to conclude that

$$R_1 - \delta_n \leq \frac{1}{n} \sum_{i=1}^{n} H(X_{1,i} \,|\, U_i, T_i, S_i, Q_i) \tag{4.8a}$$

$$R_2 - \delta_n \leq \frac{1}{n} \sum_{i=1}^{n} H(X_{2,i} \,|\, U_i, T_i, S_i, Q_i) \tag{4.8b}$$

$$R_1 + R_2 - \delta_n \leq \frac{1}{n} \sum_{i=1}^{n} I(X_{1,i}, X_{2,i}; Y_i \,|\, T_i, S_i, Q_i). \tag{4.8c}$$

The condition $S_i \multimap (T_i, S^{i-1}, X_1^{i-1}, X_2^{i-1}) \multimap (X_{1,i}, X_{2,i})$ follows from the problem setup, and the converse thus follows from Remark 6.

4.3. Achievability

To prove achievability, we consider symbol-by-symbol helpers. Given a general $P_{T|S,Q}$ (not necessarily zero-one valued), we can use the Functional Representation Lemma to infer the existence of a random variable $\tilde{Q} \sim P_{\tilde{Q}}$ such that $T = t(S, Q, \tilde{Q})$. We draw a time-sharing sequence Q^n IID $\sim P_Q$ and a functional-representing sequence $\tilde{Q}^n \sim P_{\tilde{Q}}$ and reveal them to all parties. Then, we employ the symbol-by-symbol helper that maps $(S_i, Q_i, \tilde{Q}_i)$ to the T_i determined by $t(\cdot)$. All parties are then cognizant of T_i: the encoders thanks to the helper, and the receiver—being cognizant of the state sequence—can compute T_i from $(S_i, Q_i, \tilde{Q}_i)$. Thus, we can view $(T_i, Q_i, \tilde{Q}_i)$ as a new channel state that is known to all parties. Conditioning on $(T_i, Q_i, \tilde{Q}_i)$, applying Theorem 2, and averaging over $(T_i, Q_i, \tilde{Q})$ demonstrates the achievability of all the set of the form

$$\left\{ (R_1, R_2) \in \mathbb{R}_{\geq 0}^{\times 2} \,\middle|\, \begin{aligned} R_1 &\leq H(X_1 \,|\, U, T, Q, \tilde{Q}) \\ R_2 &\leq H(X_2 \,|\, U, T, Q, \tilde{Q}) \\ R_1 + R_2 &\leq I(X_1, X_2; Y, S \,|\, T, Q, \tilde{Q}) \end{aligned} \right\}, \tag{4.9}$$

where the joint distribution is of the form

$$P_Q\, P_{\tilde{Q}}\, P_{S,T|Q,\tilde{Q}}\, P_{U|T,Q}\, P_{X_1|U,T,Q} P_{X_2|U,T,Q}\, P_{Y|X_1,X_2,S} \tag{4.10}$$

i.e., of the form (4.2) (with $P_{T|S\,Q\tilde{Q}}$ zero-one valued).

It remains to show that these sets coincide with the pentagons in (4.3). We begin with the bound on R_1:

$$\begin{aligned} H(X_1 \mid U, T, Q, \tilde{Q}) &= H(X_1 \mid U, T, S, Q, \tilde{Q}) + I(X_1; S \mid U, T, Q, \tilde{Q}) && (4.11a)\\ &= H(X_1 \mid U, T, S, Q, \tilde{Q}) && (4.11b)\\ &= H(X_1 \mid U, T, S, Q) - I(X_1; \tilde{Q} \mid U, T, S, Q) && (4.11c)\\ &= H(X_1 \mid U, T, S, Q), && (4.11d) \end{aligned}$$

where (4.11a) holds because $X_1 \multimap (U, T, Q, \tilde{Q}) \multimap S$ forms a Markov chain, and (4.11d) holds because $X_1 \multimap (U, T, S, Q) \multimap \tilde{Q}$ forms a Markov chain. Likewise, for R_2

$$H(X_2 \mid U, T, S, Q) = H(X_2 \mid U, T, Q, \tilde{Q}). \tag{4.12a}$$

As for the sum-rate:

$$\begin{aligned} I(X_1, X_2; Y, S \mid T, Q, \tilde{Q}) &= I(X_1, X_2; Y, S, \tilde{Q} \mid T, Q) - I(X_1, X_2; \tilde{Q} \mid T, Q) && (4.13a)\\ &= I(X_1, X_2; Y, S, \tilde{Q} \mid T, Q) && (4.13b)\\ &= I(X_1, X_2; Y, S \mid T, Q) + I(X_1, X_2; \tilde{Q} \mid T, Q, Y, S) && (4.13c)\\ &= I(X_1, X_2; Y, S \mid T, Q) && (4.13d)\\ &= I(X_1, X_2; Y \mid T, S, Q) + I(X_1, X_2; S \mid T, Q) && (4.13e)\\ &= I(X_1, X_2; Y \mid T, S, Q). && (4.13f) \end{aligned}$$

The pentagons in (4.3) and the pentagons in (4.9) thus coincide, and the proof of the direct part is concluded.

Author contributions

Amos Lapidoth and Baohua Ni: Writing—original draft, writing—review and editing. All authors have read and approved the final version of the manuscript for publication.

Use of Generative-AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

The authors gratefully acknowledge discussions with Yossef Steinberg and Ligong Wang.

This work was supported by the Swiss National Science Foundation (SNSF) under Grant 200021-215090.

Conflict of interest

All authors declare no conflicts of interest in this paper.

References

1. A. Lapidoth, L. Wang, State-dependent DMC with a causal helper, *IEEE T. Inform. Theory*, **70** (2023), 3162–3174. https://doi.org/10.1109/TIT.2023.3305539
2. C. E. Shannon, Channels with side information at the transmitter, *IBM J. Res. Dev.*, **2** (1958), 289–293. https://doi.org/10.1147/rd.24.0289
3. A. Lapidoth, Y. Steinberg, *The multiple access channel with causal and strictly causal side information at the encoders*, In: International Zurich Seminar on Communications (IZS 2010), Eidgenössische Technische Hochschule, 2010, 13–16. https://doi.org/10.3929/ethz-a-005997126
4. A. Lapidoth, Y. Steinberg, The multiple-access channel with causal side information: Common state, *IEEE T. Inform. Theory*, **59** (2012), 32–50. https://doi.org/10.1109/TIT.2012.2216096
5. A. Lapidoth, Y. Steinberg, The multiple-access channel with causal side information: Double state, *IEEE T. Inform. Theory*, **59** (2012), 1379–1393. https://doi.org/10.1109/TIT.2012.2230214
6. A. Lapidoth, Y. Steinberg, The state-dependent channel with a rate-limited cribbing helper, *Entropy*, **26** (2024), 570. https://doi.org/10.3390/e26070570
7. F. Willems, E. Van der Meulen, The discrete memoryless multiple-access channel with cribbing encoders, *IEEE T. Inform. Theory*, **31** (1985), 313–327. https://doi.org/10.1109/TIT.1985.1057042
8. H. Asnani, H. H. Permuter, Multiple-access channel with partial and controlled cribbing encoders, *IEEE T. Inform. Theory*, **59** (2013), 2252–2266. https://doi.org/10.1109/TIT.2012.2231462
9. S. I. Bross, A. Lapidoth, *The state-dependent multiple-access channel with states available at a cribbing encoder*, In: 2010 IEEE 26-th Convention of Electrical and Electronics Engineers in Israel, 2010, 000665–000669. https://doi.org/10.1109/EEEI.2010.5662131
10. A. Somekh-Baruch, S. Shamai, S. Verdú, *Cooperative multiple access encoding with states available at one transmitter*, In: 2007 IEEE International Symposium on Information Theory, 2007, 1556–1560. https://doi.org/10.1109/ISIT.2007.4557443
11. W. Yang, Y. Liang, S. Shamai, H. V. Poor, State-dependent Gaussian multiple access channels: New outer bounds and capacity results, *IEEE T. Inform. Theory*, **64** (2018), 7866–7882. https://doi.org/10.1109/TIT.2018.2839643
12. L. Dikstein, H. H. Permuter, S. Shamai, MAC with action-dependent state information at one encoder, *IEEE T. Inform. Theory*, **61** (2015), 173–188. https://doi.org/10.1109/TIT.2014.2372769
13. S. P. Kotagiri, J. N. Laneman, Multiaccess channels with state known to some encoders and independent messages, *EURASIP J. Wirel. Comm.*, **2008** (2008), 1–14. https://doi.org/10.1155/2008/450680
14. A. El Gamal, Y.-H. Kim, *Network information theory*, Cambridge University Press, 2011. https://doi.org/10.1017/CBO9781139030687

15. A. Zaidi, S. Shamai, *On cooperative multiple access channels with delayed CSI*, In: 2013 IEEE International Symposium on Information Theory, 2013, 982–986. https://doi.org/10.1109/ISIT.2013.6620373