Multiple Access Channels with Combined Cooperation and Partial Cribbing

Tal Kopetz, Haim Permuter and Shlomo Shamai (Shitz)

arXiv:1404.0832v1 [cs.IT] 3 Apr 2014

Abstract

In this paper we study the multiple access channel (MAC) with combined cooperation and partial cribbing and characterize its capacity region. Cooperation means that the two encoders send a message to one another via a rate-limited link prior to transmission, while partial cribbing means that each of the two encoders obtains a deterministic function of the other encoder's output with or without delay. Prior work in this field dealt separately with cooperation and partial cribbing. However, by combining these two methods we can achieve significantly higher rates. Remarkably, the capacity region does not require an additional auxiliary random variable (RV) since the purpose of both cooperation and partial cribbing is to generate a common message between the encoders. In the proof we combine methods of block Markov coding, backward decoding, double rate-splitting, and joint typicality decoding. Furthermore, we present the Gaussian MAC with combined one-sided cooperation and quantized cribbing. For this model, we give an achievability scheme that shows how many cooperation or quantization bits are required in order to achieve a Gaussian MAC with full cooperation/cribbing capacity region. After establishing our main results, we consider two cases where only one auxiliary RV is needed. The first is a rate distortion dual setting for the MAC with a common message, a private message and combined cooperation and cribbing. The second is a state-dependent MAC with cooperation, where the state is known at a partially cribbing encoder and at the decoder. However, there are cases where more than one auxiliary RV is needed, e.g., when the cooperation and cribbing are not used for the same purposes. We present a MAC with an action-dependent state, where the action is based on the cooperation but not on the cribbing. Therefore, in this case more than one auxiliary RV is needed. We deduce a general rule for this result.

Index Terms

Action, Block Markov coding, Cooperation, Duality, Double rate splitting, Gaussian MAC, Gelfand-Pinsker coding, Multiple access channels, Partial cribbing, State.

I. INTRODUCTION

The MAC with cooperating encoders was first studied by Willems [1]-[3]. Willems introduced two separate approaches to cooperating encoders; in the first, using a rate-limited cooperation link between the two encoders, the

two encoders cooperate and share as much of their private messages as possible, while in the second, each encoder "listens" to the other encoder and obtains its output. The second approach was named cribbing. Capacity regions for the two approches, separately, were established by Willems. Furthermore, the cribbing setting was generalized in [4] to partial cribbing which means that each of the two encoders obtains a deterministic function of the other encoders output. The partial cribbing is especially important in the continuous alphabet, such as the Gaussian MAC, since in a continuous alphabet perfect cribbing means full cooperation between the encoders regardless of the cribbing delay.

In this paper, we combine cooperation and partial cribbing and use them simultaneously, thus obtaining better performance and a larger capacity region. A MAC with combined cooperation and partial cribbing is depicted in Fig. 1. Encoder 1 and Encoder 2 obtain messages M_{21} and M_{12} prior to transmission. For the cribbing part, we address two cases. In Case A, the cribbing is done strictly causally by both encoders, i.e., $X_{1,i}$ is a function of (M_{21}, Z_2^{i-1}) and $X_{2,i}$ is a function of (M_{12}, Z_1^{i-1}). In Case B, the cribbing is done strictly causally by Encoder 1 and causally by Encoder 2, i.e., $X_{1,i}$ is a function of (M_{21}, Z_2^{i-1}) and $X_{2,i}$ is a function of (M_{12}, Z_1^i). The idea is that this deterministic function, Z_1, is on a sliding scale where one end is $Z_{1,i} = X_{1,i}$ (the actual output) and the other end is when $Z_{1,i}$ is a constant, which does not give any information about $X_{1,i}$. The same applies for Z_2. In this research, it was our goal to obtain a generic capacity region for a scheme with both cooperation and partial cribbing.
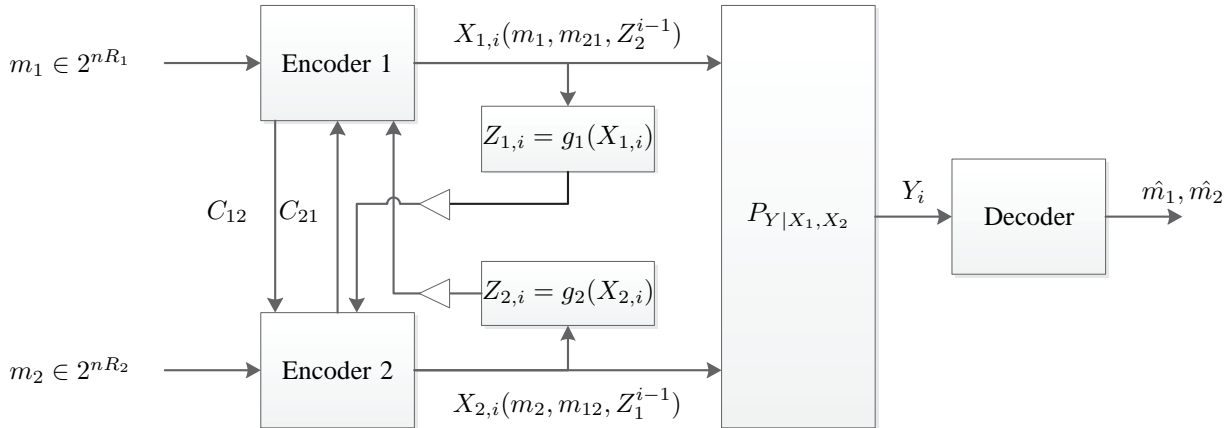

Fig. 1. MAC with combined cooperation and partial cribbing. Encoder 1 and Encoder 2 obtain messages M_{21} and M_{12} prior to transmission. The cribbing is done strictly causally by both encoders. This setting corresponds to Case A.

Cooperation and cribbing carry practical implications. In [5, Chapter 8], Simone et. al. considered cooperative wireless cellular systems and analyzed their performance with separate cooperation and cribbing (referred to as Out-of-Band cooperation and In-Band cooperation, respectively). The results show how cooperation and cribbing separately increase capacity in wireless cellular systems. In the expected 3GPP Release 12, a standard called Proximity Services (ProSE) will be added to the LTE-Advanced "grab bag" of technologies [6]. The ProSE protocol

will address issues of spectrum utilization, overall throughput, and energy consumption, while enabling new peer to peer and location based applications and services, all of which will be applied using cooperation between "nearby" users in the network. The communication between the users can be attained by using mobile ad hoc networks (Out-of-Band/Cooperation) or by using the same band as the cell sites (In-Band/Cribbing). Settings of combined cooperation and cribbing considered in this paper give the fundamental limits and insights on how to design optimal coding for communication systems where the users have cognition capabilities and, therefore, "listen" to each other's signals and, in addition, cooperate with each other via dedicated links. We show that combining cribbing and cooperation is straightforward since it does not require any additional auxiliary RV compared with only cribbing or only cooperation. Therefore, the combination of cooperation and cribbing should be considered in future cooperative wireless cellular systems such as ProSE.

In this paper, we solve the general model that incorporates both cooperation and partial cribbing. The capacity regions that were found for cooperation and partial cribbing, separately, in [1] and [4] were constructed using an auxiliary RV, U. That RV signified the information that both encoders share. In [9], Slepian and Wolf discovered that the capacity region for the MAC is larger if the encoders share a common message. Therefore, we can refer to the information obtained via cooperation and cribbing as common information shared by both encoders. One of the results in our work is that the combination of the models does not require an additional auxiliary RV; it is possible to use only one auxiliary RV that represents the common information. This implies that if for the MAC with partial cribbing we have a "good code", namely, a code that achieves the capacity region, then by performing minor modifications, namely, increasing the common message rate, we can construct a "good code" for the MAC with combined cooperation and partial cribbing. The coding techniques we use in this paper include block Markov coding (introduced by Willems), joint typicality decoding, backward decoding, and double rate splitting. Double rate splitting is necessary since we need to split the original message twice; one part will be obtained through the cooperation link and the other part will be obtained using partial cribbing.

Combining cooperation and cribbing was first considered by Bracher and Lapidoth [10] in the context of feedback and state information. However, only strictly-causal perfect cribbing was considered and in our paper we consider partial cribbing both causal and strictly-causal.

After establishing our main results, we present the Gaussian MAC with combined one-sided cooperation and partial cribbing. One can see that an outer bound for the capacity region of this setting is when Encoder 2 knows the message of Encoder 1. Inspired by the work of Asnani et al. [4] and Bross et al. [11], we describe an achievability scheme that coincides with this outer bound in some cases.

Additionaly, we provide a duality between a MAC with a common message, a private message and combined cooperation and cribbing and the rate distortion model known as "Successive Refinement (SR) With Decoder Cooperation" presented in [12]. The decoder cooperation is through a dedicated link and partial cribbing. In this paper we combine both cooperation and partial cribbing in the SR problem and obtain a rate region with only one auxiliary RV.

We go on to study the impact of cooperation and cribbing on state-dependent MACs where the state may provide

a refined characterization of the channel, as state-dependent channels are widely studied in the literature. We address two different state-dependent MACs with cooperation and cribbing (see [10], [13] for further reading). The first is a MAC with cooperation and channel state known non-causally at a partially cribbing encoder and at the decoder. In this case we use our results to find a solution with a lone auxiliary RV. Only one auxiliary RV is needed since the purpose of both cooperation and partial cribbing is to generate a common message between the encoders. The second is a MAC where action-dependent state is known non-causally at a cribbing encoder. Additionally, a one-sided cooperation link is attained at the cribbing encoder. Action-dependent states were introduced by Weissman in [14]. The action is based on the private message of the cribbing encoder and the message from the cooperation link. In this case, a lone auxiliary RV will not suffice since the purpose of the cooperation is not only to generate a common message but also to contribute to the action and affect the channel state.

The remainder of the paper is organized as follows: In Section II, we define the MAC with combined cooperation and partial cribbing and provide its capacity region for two cases. The first is for strictly causal partial cribbing (Case A) and the second is for mixed causal and strictly causal partial cribbing (Case B). Thereafter, the proof for both cases is provided. In Section III, we give an achievability scheme for the Gaussian MAC with combined one-sided cooperation and partial cribbing. In Section IV, we establish the duality between the MAC with combined cooperation and partial cribbing at the encoders and the SR problem with combined cooperation and partial cribbing at the decoders. We show that a lone RV is needed to characterize the rate region of the SR problem. In Section V, we give an example of a state-dependent MAC with combined cooperation and partial cribbing where only one auxiliary RV is needed. In Section VI, we study the case of the MAC with an action-dependent state where more than one auxiliary RV is needed and consider its implications. In Section VII we conclude the paper and suggest some research directions that have not yet been solved such as noncausal partial cribbing and combined cooperation and cribbing in the interference channel.

II. THE MAC WITH COMBINED COOPERATION AND PARTIAL CRIBBING

A. Definitions and Main Results

Let us consider the MAC with combined cooperation and partial cribbing depicted in Fig. 1. The MAC setting consists of two transmitters (encoders) and one receiver (decoder). Each transmitter $l \in \{1, 2\}$ chooses an index m_l uniformly from the set $\{1, \ldots, 2^{nR_l}\}$ and independently of the other transmitter. The input to the channel from Encoder $l \in \{1, 2\}$ is denoted by $\{X_{l,1}, X_{l,2}, X_{l,3}, \ldots\}$. Encoder 1 and Encoder 2 obtain deterministic functions of the form $Z_{2,i} = g_2(X_{2,i})$ and $Z_{1,i} = g_1(X_{1,i})$, respectively. We address two cases in this setting:

- Case A : Both Encoder 1 and Encoder 2 obtain $Z_{2,i}$ and $Z_{1,i}$, respectively, with unit delay.
- Case B : Encoder 1 obtains $Z_{2,i}$ with unit delay and Encoder 2 obtains $Z_{1,i}$ without delay.

Additionally, Encoder 1 obtains a message $m_{21} \in \{1, \ldots, 2^{nC_{21}}\}$ from Encoder 2 and Encoder 2 obtains a message $m_{12} \in \{1, \ldots, 2^{nC_{12}}\}$ from Encoder 1. Both messages are obtained prior to the transmission of (X_1^n, X_2^n) through the channel. The output of the channel is denoted by $\{Y_1, Y_2, Y_3, \ldots\}$. The channel is characterized by a conditional

probability $P(y_i|x_{1,i}, x_{2,i})$. The channel probability does not depend on the time index i and is memoryless, i.e.,

$$P(y_i|x_1^i, x_2^i, y^{i-1}) = P(y_i|x_{1,i}, x_{2,i}), \tag{1}$$

where the superscripts denote sequences in the following way: $x_l^i = (x_{l,1}, x_{l,2}, \ldots, x_{l,i}), l \in \{1, 2\}$. Since the settings in this paper do not include feedback from the receiver to the transmitters, i.e., $P(x_{1,i}, x_{2,i}|x_1^{i-1}, x_2^{i-1}, y^{i-1}) = P(x_{1,i}, x_{2,i}|x_1^{i-1}, x_2^{i-1})$, equation (1) implies that

$$P(y_i|x_1^n, x_2^n, y^{i-1}) = P(y_i|x_{1,i}, x_{2,i}). \tag{2}$$

Definition 1 A $(2^{nR_1}, 2^{nR_2}, 2^{nC_{12}}, 2^{nC_{21}}, n)$ *code* for the MAC with combined cooperation and partial cribbing, as shown in Fig. 1, consists at time i of encoding functions at Encoder 1 and Encoder 2

$$f_{12} \quad : \quad \{1, \ldots, 2^{nR_1}\} \mapsto \{1, \ldots, 2^{nC_{12}}\}, \tag{3}$$

$$f_{21} \quad : \quad \{1, \ldots, 2^{nR_2}\} \mapsto \{1, \ldots, 2^{nC_{21}}\}, \tag{4}$$

$$f_{1,i} \quad : \quad \{1, \ldots, 2^{nR_1}\} \times \{1, \ldots, 2^{nC_{21}}\} \times \mathcal{Z}_2^{i-1} \mapsto \mathcal{X}_{1,i}, \tag{5}$$

$$f_{2,i}^A \quad : \quad \{1, \ldots, 2^{nR_2}\} \times \{1, \ldots, 2^{nC_{12}}\} \times \mathcal{Z}_1^{i-1} \mapsto \mathcal{X}_{2,i}, \tag{6}$$

$$f_{2,i}^B \quad : \quad \{1, \ldots, 2^{nR_2}\} \times \{1, \ldots, 2^{nC_{12}}\} \times \mathcal{Z}_1^i \mapsto \mathcal{X}_{2,i}, \tag{7}$$

and a decoding function

$$g : \mathcal{Y}^n \mapsto \{1, \ldots, 2^{nR_1}\} \times \{1, \ldots, 2^{nR_2}\}. \tag{8}$$

The average probability of error for a $(2^{nR_1}, 2^{nR_2}, 2^{nC_{12}}, 2^{nC_{21}}, n)$ code is defined as

$$P_e^{(n)} = \frac{1}{2^{n(R_1+R_2)}} \sum_{m_1, m_2} \Pr\{g(Y^n) \neq (m_1, m_2)|(m_1, m_2) \text{ sent}\}. \tag{9}$$

A rate (R_1, R_2) is said to be *achievable* for the MAC with combined cooperation and partial cribbing if there exists a sequence of $(2^{nR_1}, 2^{nR_2}, 2^{nC_{12}}, 2^{nC_{21}}, n)$ codes s.t. $P_e^{(n)} \to 0$. The *capacity region* of the MAC is the closure of all achievable rates. The following theorem describes the capacity region of a MAC with combined cooperation and partial cribbing.

Let us define the following regions, $\mathcal{R}^A$ and $\mathcal{R}^B$, that are contained in $\mathbb{R}_+^2$, namely, contained in the set of nonnegative two-dimensional real numbers.

$$\mathcal{R}^A = \left\{ \begin{array}{c} R_1 \leq I(X_1; Y|X_2, Z_1, U) + H(Z_1|U) + C_{12}, \\ R_2 \leq I(X_2; Y|X_1, Z_2, U) + H(Z_2|U) + C_{21}, \\ R_1 + R_2 \leq I(X_1, X_2; Y|U, Z_1, Z_2) + H(Z_1, Z_2|U) + C_{12} + C_{21}, \\ R_1 + R_2 \leq I(X_1, X_2; Y), \text{ for} \\ P(u)P(x_1|u)\mathbb{1}_{z_1=f(x_1)}P(x_2|u)\mathbb{1}_{z_2=f(x_2)}P(y|x_1, x_2). \end{array} \right\}. \tag{10}$$

The region $\mathcal{R}^B$ is defined with the same set of inequalities as in (10), but the joint distribution is of the form

$$P(u)P(x_1|u)\mathbb{1}_{z_1=f(x_1)}P(x_2|u,z_1)\mathbb{1}_{z_2=f(x_2)}P(y|x_1,x_2). \tag{11}$$

Theorem 1 *(Capacity Region of the MAC with Combined Cooperation and Partial Cribbing)* The capacity regions of the MAC with combined cooperation and strictly causal (Case A) and mixed strictly causal and causal (Case B) partial cribbing, as described in Def. 1, are $\mathcal{R}^A$ and $\mathcal{R}^B$, respectively.

We note that $H(Z_1|U) = I(Z_1; X_1|U)$; thus the cribbing, $I(Z_1; X_1|U)$, plays the same role (in a quantitative sense) to the cooperation link, C_{12}. Similarly, the role of $I(Z_2; X_2|U)$ to C_{21} and of $I(Z_1, Z_2; X_1, X_2|U)$ to $C_{12} + C_{21}$. Hence, the important feature is the mutual information of the cooperation, whether the cooperation is done by cribbing or by dedicated links, and they both act in a similar way.

A straightforward result from Theorem 1 is the capacity region for the compound MAC [15] with combined cooperation and partial cribbing. The region and proof for the compound MAC are omitted for brevity.

B. Proof of Theorem 1

1) Converse: We will start with the converse of Case A.

Converse for Case A: Given an achievable rate (R_1, R_2) we need to show that there exists a joint distribution of the form $P(u)P(x_1|u)\mathbb{1}_{z_1=f(x_1)}P(x_2|u)\mathbb{1}_{z_2=f(x_2)}P(y|x_1,x_2)$ such that the inequalities (10) are satisfied. Since (R_1, R_2) is an achievable rate-pair, there exists a $(2^{nR_1}, 2^{nR_2}, 2^{nC_{12}}, 2^{nC_{21}}, n)$ code with an arbitrarily small error probability $P_e^{(n)}$. By Fano's inequality,

$$H(M_1, M_2|Y^n) \le n(R_1 + R_2)P_e^{(n)} + H(P_e^{(n)}). \tag{12}$$

We set

$$(R_1 + R_2)P_e^{(n)} + \frac{1}{n}H(P_e^{(n)}) \triangleq \epsilon_n, \tag{13}$$

where $\epsilon_n \to 0$ as $P_e^{(n)} \to 0$. Hence,

$$H(M_1|Y^n, M_2) \le H(M_1, M_2|Y^n) \le n\epsilon_n, \tag{14}$$

$$H(M_2|Y^n, M_1) \le H(M_1, M_2|Y^n) \le n\epsilon_n. \tag{15}$$

For R_1 we have the following:

$$nR_1 = H(M_1) \tag{16}$$

$$\overset{(a)}{=} H(M_1, M_{12}, Z_1^n|M_2) \tag{17}$$

$$\overset{(b)}{=} H(M_{12}|M_2) + H(Z_1^n|M_{12}, M_2) + H(M_1|Z_1^n, M_{12}, M_2) \tag{18}$$

$$= H(M_{12}) + H(Z_1^n|M_{12}, M_{21}, M_2) + H(M_1|Z_1^n, M_{12}, M_2)$$

$$+ H(M_1|Y^n, Z_1^n, M_{12}, M_2) - H(M_1|Y^n, Z_1^n, M_{12}, M_2) \tag{19}$$

$$\stackrel{(c)}{\leq} H(M_{12}) + H(Z_1^n|M_{12}, M_{21}, M_2) + I(M_1; Y^n|Z_1^n, M_{12}, M_2, M_{21}) + n\epsilon_n \tag{20}$$

$$\stackrel{(d)}{=} H(M_{12}) + \sum_{i=1}^{n}[H(Z_{1,i}|Z_1^{i-1}, M_{12}, M_{21}, M_2)$$
$$+ I(M_1; Y_i|Y^{i-1}, Z_1^n, M_{12}, M_2, M_{21})] + n\epsilon_n \tag{21}$$

$$\stackrel{(e)}{=} H(M_{12}) + \sum_{i=1}^{n}[H(Z_{1,i}|Z_1^{i-1}, Z_2^{i-1}, M_{12}, M_{21}, M_2)$$
$$+ I(M_1, X_{1,i}; Y_i|Y^{i-1}, Z_1^n, Z_2^{i-1}, M_{12}, M_2, M_{21})] + n\epsilon_n \tag{22}$$

$$\stackrel{(f)}{\leq} H(M_{12}) + \sum_{i=1}^{n}[H(Z_{1,i}|Z_1^{i-1}, Z_2^{i-1}, M_{12}, M_{21})$$
$$+ I(X_{1,i}; Y_i|X_{2,i}, Z_1^i, Z_2^{i-1}, M_{12}, M_{21})] + n\epsilon_n, \tag{23}$$

where (a) follows since messages M_1 and M_2 are independent and since $(M_{12}, Z_1^n) = f(M_1, M_2)$, (b) and (d) follow from the chain rule, (c) follows from Fano's inequality and because M_{21} is a function of M_2, (e) follows since Z_2^{i-1} is a function of (M_{12}, M_2) and $X_{1,i}$ is a function of (M_1, M_{21}), and step (f) follows since conditioning reduces entropy and from the Markov chain $Y_i - (X_{1,i}, X_{2,i}, M_{12}, M_{21}, Z_1^i, Z_2^{i-1}) - (M_1, M_2, Y^{i-1})$. From the definition of a RV

$$U_i \triangleq (Z_1^{i-1}, Z_2^{i-1}, M_{12}, M_{21}), \tag{24}$$

we obtain

$$R_1 \leq C_{12} + \frac{1}{n}\sum_{i=1}^{n}[H(Z_{1,i}|U_i) + I(X_{1,i}; Y_i|X_{2,i}, Z_{1,i}, U_i)] + \epsilon_n. \tag{25}$$

Similarly to (25), we obtain

$$R_2 \leq C_{21} + \frac{1}{n}\sum_{i=1}^{n}[H(Z_{2,i}|U_i) + I(X_{2,i}; Y_i|X_{1,i}, Z_{2,i}, U_i)] + \epsilon_n. \tag{26}$$

Now, consider

$$n(R_1 + R_2) = H(M_1, M_2) \tag{27}$$

$$\stackrel{(a)}{=} H(M_1, M_2, Z_1^n, Z_2^n, M_{12}, M_{21}) \tag{28}$$

$$\stackrel{(b)}{=} H(M_{12}) + H(M_{21}|M_{12}) + H(Z_1^n, Z_2^n|M_{12}, M_{21})$$
$$+ H(M_1, M_2|Z_1^n, Z_2^n, M_{12}, M_{21}) \tag{29}$$

$$\stackrel{(c)}{\leq} H(M_{12}) + H(M_{21}) + H(Z_1^n, Z_2^n|M_{12}, M_{21})$$
$$+ I(M_1, M_2; Y^n|Z_1^n, Z_2^n, M_{12}, M_{21}) + n\epsilon_n \tag{30}$$

$$\stackrel{(d)}{\leq} nC_{12} + nC_{21} + \sum_{i=1}^{n}[H(Z_{1,i}, Z_{2,i}|Z_1^{i-1}, Z_2^{i-1}, M_{12}, M_{21})$$
$$+ I(M_1, M_2; Y_i|Y^{i-1}, Z_1^n, Z_2^n, M_{12}, M_{21})] + n\epsilon_n \tag{31}$$

$$\stackrel{(e)}{=} nC_{12} + nC_{21} + \sum_{i=1}^{n}[H(Z_{1,i}, Z_{2,i}|Z_1^{i-1}, Z_2^{i-1}, M_{12}, M_{21})$$

$$+I(M_1, X_{1,i}, M_2, X_{2,i}; Y_i | Y^{i-1}, Z_1^n, Z_2^n, M_{12}, M_{21})] + n\epsilon_n \tag{32}$$

$$\overset{(f)}{=} \ nC_{12} + nC_{21} + \sum_{i=1}^{n}[H(Z_{1,i}, Z_{2,i} | Z_1^{i-1}, Z_2^{i-1}, M_{12}, M_{21})$$

$$+I(X_{1,i}, X_{2,i}; Y_i | Z_1^i, Z_2^i, M_{12}, M_{21})] + n\epsilon_n, \tag{33}$$

where (a) follows from the fact that $(M_{12}, M_{21}, Z_1^n, Z_2^n) = f(M_1, M_2)$, (b) and (d) follow from the chain rule, (c) follows from Fano's inequality and because M_{21} is independent of M_{12}, (e) follows from the fact that $(X_{1,i}, X_{2,i}) = f(M_1, M_2)$, and step (f) follows from the Markov chain $Y_i - (X_{1,i}, X_{2,i}, Z_1^i, Z_2^i, M_{12}, M_{21}) - (M_1, M_2, Y^{i-1})$. From the definition of the RV U, we obtain

$$R_1 + R_2 \ \le \ C_{12} + C_{21} + \frac{1}{n}\sum_{i=1}^{n}[H(Z_{1,i}, Z_{2,i} | U_i) + I(X_{1,i}, X_{2,i}; Y_i | Z_{1,i}, Z_{2,i}, U_i)] + \epsilon_n. \tag{34}$$

Furthermore, consider

$$n(R_1 + R_2) \ = \ H(M_1, M_2) \tag{35}$$

$$= \ H(M_1, M_2) + H(M_1, M_2 | Y^n) - H(M_1, M_2 | Y^n) \tag{36}$$

$$\overset{(a)}{\le} \ I(M_1, M_2; Y^n) + n\epsilon_n \tag{37}$$

$$\overset{(b)}{=} \ I(X_1^n, X_2^n; Y^n) + n\epsilon_n \tag{38}$$

$$\overset{(c)}{=} \ \sum_{i=1}^{n} I(X_1^n, X_2^n; Y_i | Y^{i-1}) + n\epsilon_n \tag{39}$$

$$\overset{(d)}{=} \ \sum_{i=1}^{n} I(X_{1,i}, X_{2,i}; Y_i) + n\epsilon_n, \tag{40}$$

where (a) follows from Fano's inequality, (b) follows from the fact that (X_1^n, X_2^n) is a deterministic function of (M_1, M_2) and from the Markov chain $Y^n - (X_1^n, X_2^n) - (M_1, M_2)$, (c) follows from the chain rule, and step (d) follows from the memoryless property of the channel. Thus we obtain

$$R_1 + R_2 \le \frac{1}{n}\sum_{i=1}^{n} I(X_{1,i}, X_{2,i}; Y_i) + \epsilon_n. \tag{41}$$

Finally, we will prove the following Markov chains:

- $Z_{2,i} - U_i - Z_{1,i}$ - We will prove this graphically as in [16, Section II]. Using the undirected graph in Fig. 2, we can see that the Markov Chain $Z_{2,i} - (M_{12}, M_{21}, Z_1^{i-1}, Z_2^{i-1}) - Z_{1,i}$ holds since we cannot get from node $Z_{2,i}$ to node $Z_{1,i}$ without going through nodes $(M_{12}, M_{21}, Z_1^{i-1}, Z_2^{i-1})$.

- $X_{1,i} - (U_i, Z_{1,i}) - Z_{2,i}$ - Using the undirected graph in Fig. 2, we can see that the Markov Chain $X_{1,i} - (M_{12}, M_{21}, Z_1^i, Z_2^{i-1}) - Z_{2,i}$ holds since we cannot get from node $X_{1,i}$ to node $Z_{2,i}$ without going through nodes $(M_{12}, M_{21}, Z_1^i, Z_2^{i-1})$.

- $X_{2,i} - (U_i, Z_{2,i}) - X_{1,i}$ - Using the undirected graph in Fig. 2, we can see that the Markov Chain $X_{2,i} - (M_{12}, M_{21}, Z_1^{i-1}, Z_2^i) - X_{1,i}$ holds since we cannot get from node $X_{2,i}$ to node $X_{1,i}$ without going through nodes $(M_{12}, M_{21}, Z_1^{i-1}, Z_2^i)$.
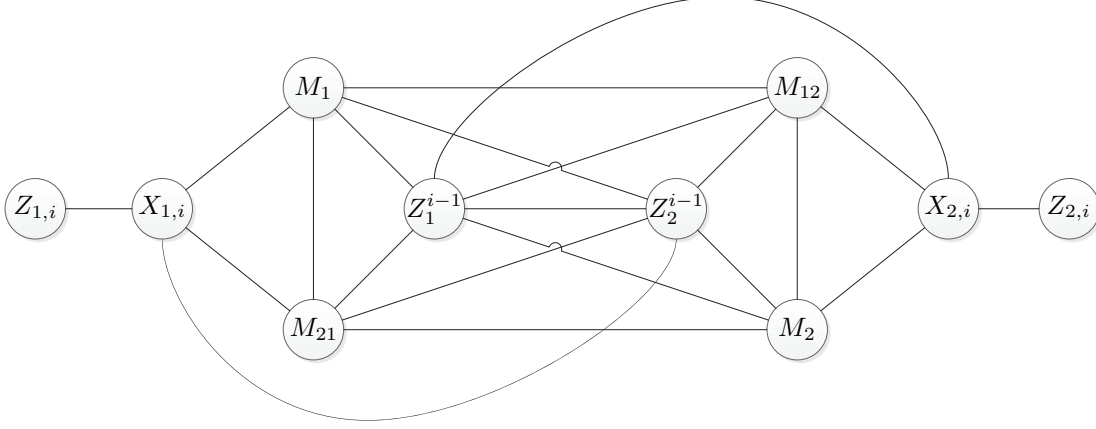
Fig. 2. Proof of the Markov Chain $X_{2,i} - (M_{12}, M_{21}, Z_1^{i-1}, Z_2^{i-1}) - X_{1,i}$ using the undirected graphical technique [16, Section II]. This graph corresponds to the joint distribution $P(m_1)P(m_2)P(m_{12}|m_1)P(m_{21}|m_2) \prod_{k=1}^{i-1} P(z_{1,k}|m_1, m_{21}, z_2^{k-1})P(z_{2,k}|m_2, m_{12}, z_1^{k-1})P(x_{1,i}|m_{21}, m_1, z_2^{i-1})$ $P(x_{2,i}|m_{12}, m_2, z_1^{i-1})P(z_{1,i}|x_{1,i})P(z_{2,i}|x_{2,i})$.

- $Y_i - (X_{1,i}, X_{2,i}) - (Z_{1,i}, Z_{2,i}, U_i)$ - Follows since the channel output at time i depends on the history (X_1^i, X_2^i) only through $(X_{1,i}, X_{2,i})$.

Finally, let Q be an RV independent of (X_1^n, X_2^n, Y^n) and uniformly distributed over the set $\{1, 2, 3, \ldots, n\}$. We define the RVs $U \triangleq (Q, U_Q), X_1 \triangleq X_{1,Q}, X_2 \triangleq X_{2,Q}$, and $Y \triangleq Y_Q$ to obtain the region given in (10). This completes the converse for Case A. ∎

Converse for Case B: We repeat the same approach as for Case A, except that in the final step we need to show the Markov chain $X_{2,i} - (U_i, Z_{1,i}, Z_{2,i}) - X_{1,i}$ rather than $X_{2,i} - (U_i, Z_{2,i}) - X_{1,i}$ as in Case A. Since for Case A $X_{2,i} - (M_{12}, M_{21}, Z_1^{i-1}, Z_2^i) - X_{1,i}$ holds, then $X_{2,i} - (M_{12}, M_{21}, Z_1^i, Z_2^i) - X_{1,i}$ also holds. ∎

2) Achievability: Achievability for Case A: To prove the achievability of the capacity region, we need to show that for a fixed distribution of the form $P(u)P(x_1|u)\mathbb{1}_{z_1=f(x_1)}P(x_2|u)\mathbb{1}_{z_2=f(x_2)}P(y|x_1, x_2)$ and for (R_1, R_2) that satisfy the inequalities in (10), there exists a sequence of $(2^{nR_1}, 2^{nR_2}, 2^{nC_{12}}, 2^{nC_{21}}, n)$ codes for which $P_e^{(n)} \to 0$ as $n \to \infty$.

The idea behind this proof is to convert the cooperation problem into a setting that corresponds to the MAC with a common message and partially cribbing encoders considered in [4] and rely on its capacity region to show that the cooperation capacity region is indeed achievable. This is done by sharing as much as possible of the original private messages, (m_1, m_2), through the communication links in order to create a common message; the unshared parts of the original messages serve as the new private messages. By doing so, the coding scheme of the setting with a common message can be employed. The capacity region found in [4] for the MAC with a common message and partially cribbing encoders is

$$\tilde{R}_1 \leq H(Z_1|U) + I(X_1; Y|X_2, Z_1, U),$$

$$\tilde{R}_2 \leq H(Z_2|U) + I(X_2; Y|X_1, Z_2, U),$$

$$\tilde{R}_1 + \tilde{R}_2 \leq I(X_1, X_2; Y|U, Z_1, Z_2) + H(Z_1, Z_2|U),$$

$$\tilde{R}_0 + \tilde{R}_1 + \tilde{R}_2 \leq I(X_1, X_2; Y). \tag{42}$$

The achievability proof for the MAC with a common message and partially cribbing encoders is available in Appendix A. Let us define the following rates

$$\tilde{R}_0 = C_{12} + C_{21}, \tag{43}$$

$$\tilde{R}_1 = R_1 - C_{12}, \tag{44}$$

$$\tilde{R}_2 = R_2 - C_{21}, \tag{45}$$

i.e., we defined the common message as the messages that are transmitted through the cooperation links. With respect to these definitions, the inequalities in (42) can be rewritten as

$$R_1 - C_{12} \leq H(Z_1|U) + I(X_1; Y|X_2, Z_1, U),$$

$$R_2 - C_{21} \leq H(Z_2|U) + I(X_2; Y|X_1, Z_2, U),$$

$$(R_1 - C_{12}) + (R_2 - C_{21}) \leq I(X_1, X_2; Y|U, Z_1, Z_2) + H(Z_1, Z_2|U),$$

$$(C_{12} + C_{21}) + (R_1 - C_{21}) + (R_2 - C_{21}) \leq I(X_1, X_2; Y), \tag{46}$$

which is equivalent to the region in (10). ∎

Achievability for Case B: The achievability of case B is very similar to that of case A, only the codewords of X_2 need to be generated according to Shannon's strategy (or a code-tree) rather than codewords. This is due to the fact that $Z_{1,i}$ is known causally and X_2 is generated according to a distribution $P(x_2|u, z_1, z_2)$. ∎

III. GAUSSIAN MAC WITH COMBINED COOPERATION AND QUANTIZED CRIBBING

We now consider a Gaussian MAC, i.e., $Y = X_1 + X_2 + W$ where $W \sim N(0, N)$, depicted in Fig. 3.

We assume that the power constraints over the outputs of Encoder 1 and Encoder 2 are P_1 and P_2, respectively. Prior to transmission, Encoder 1 sends a message M_{12} to Encoder 2. In addition, Encoder 2 cribs causally from Encoder 1 and obtains Z_i, which is a scalar quantization of the signal $X_{1,i}$. First, we look at an inner bound to the capacity region, which is the Gaussian MAC without cooperation and cribbing. The capacity region in this case is

$$R_1 \leq \frac{1}{2} \log(1 + \frac{P_1}{N}),$$

$$R_2 \leq \frac{1}{2} \log(1 + \frac{P_2}{N}),$$

$$R_1 + R_2 \leq \frac{1}{2} \log(1 + \frac{P_1 + P_2}{N}). \tag{47}$$

On the other hand, an outer bound is obtained when there is full cooperation or perfect cribbing, i.e., Encoder 2 obtains the message m_1 before sending X_2. The capacity region in this case is

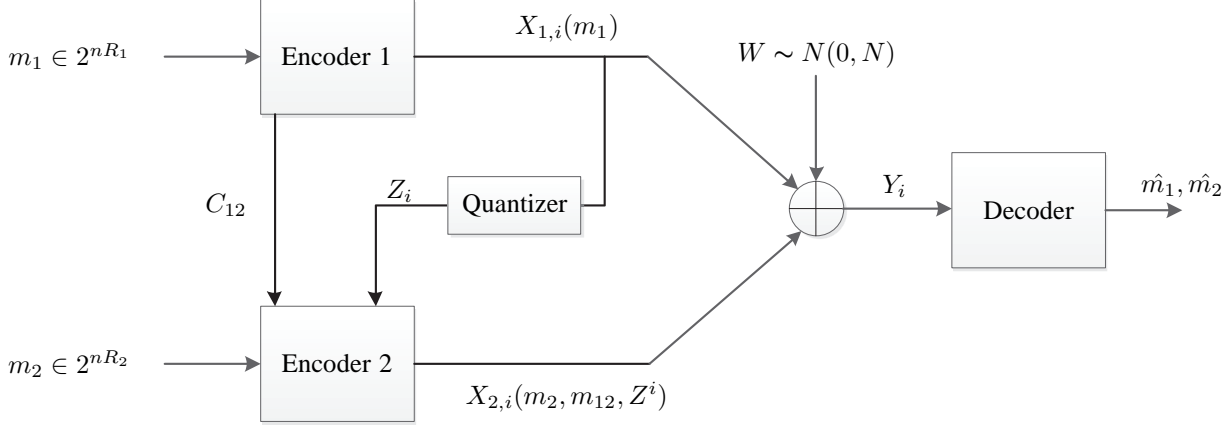$$R_2 \leq \frac{1}{2} \log(1 + \frac{P_2}{N}(1 - \rho^2)),$$

Fig. 3. Gaussian MAC with one-sided combined cooperation and quantized cribbing. Message M_{12} is sent prior to transmission and Z_i is known causally at Encoder 2.

$$R_1 + R_2 \quad \leq \quad \frac{1}{2}\log(1 + \frac{P_1 + 2\rho\sqrt{P_1 P_2} + P_2}{N}). \tag{48}$$

We now present an achievability scheme inspired by the work of Asnani et al. [4] and Bross et al. [11]. In [4], an achievable region for the Gaussian MAC with quantized cribbing has been described, whereas in [11], an achievable region for the Gaussian MAC with a common message was provided. In our work, we combine the two achievability schemes. We set the following distributions:

$$X_1 \quad = \quad \lambda U + X_1', \tag{49}$$

$$X_2 \quad = \quad \bar{\lambda} U + X_2', \tag{50}$$

where

$$U \sim N(0, P_0) \quad , \quad P_0 = \left(\sqrt{\bar{\beta}_1 P_1} + \sqrt{\bar{\beta}_2 P_2} \right)^2,$$

$$P_{X_2'|Z,U}(x_2'|z,u) \quad = \quad \bar{\rho} P_{X_2''}(x_2') + \rho P_{X_1'|Z,U}(x_2'|z,u),$$

$$X_1' \quad \sim \quad N(0, \beta_1 P_1),$$

$$X_2'' \quad \sim \quad N(0, \beta_2 P_2),$$

$$\lambda = \sqrt{\frac{\bar{\beta}_1 P_1}{P_0}} \quad , \quad \bar{\lambda} = 1 - \lambda,$$

$$\beta_1, \beta_2, \rho \quad \in \quad [0, 1]. \tag{51}$$

The intuition behind the choice of these distributions is as follows. The common message, signified as U, is obtained via the rate-limited link and the two encoders cooperate to send that common message. Since the cooperation and cribbing are one-sided, only Encoder 2 can help Encoder 1 send his private message. The idea behind the choice of $P_{X_2'|Z,U}(x_2'|z,u)$ is that Encoder 2 will send $\bar{\rho}$ of the time his private message and ρ of the time the estimation
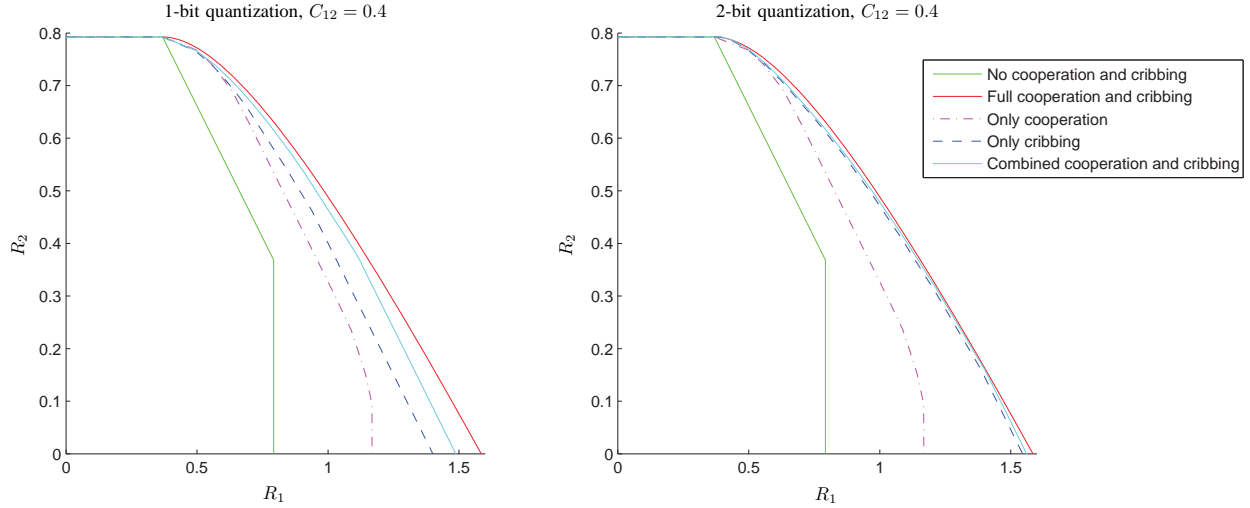
Fig. 4. Achievable regions for the Gaussian MAC with combined cooperation and quantized cribbing.

of Encoder 1's private message, X_1', conditioned on the cribbing Z and the cooperation U. Notice that under these definitions, by setting the power constraints as $P_1 = P_2 = 1$, the power constraints on both encoders hold. Evaluation of region $\mathcal{R}_B$ with Z_2 constant and $N = \frac{1}{2}$ is depicted in Fig. 4; achievable regions for 1-bit and 2-bit quantizations are illustrated where $C_{12} = 0.4$. When only one bit of quantization is available (LHS of Fig. 4), the region of combined cooperation and cribbing encloses special cases of cribbing [4] and cooperation [11]. However, when two bits of quantization are available (RHS of Fig. 4), combining cooperation and cribbing does not significantly increase the region. This is because the difference between the achievable region with a 2-bit quantizer ($C_{12} = 0$) and full cooperation is negligible.

IV. DUAL RATE DISTORTION SETTING

The information-theoretic duality between rate distortion and channel coding was first introduced by Shannon in [17]. An important duality between the Wyner-Ziv rate distortion problem [18] and the Gelfand-Pinsker channel coding problem [19] was pointed out by Cover and Chiang in [20] (see [21] and [22] for further reading). In some cases, the corner points of a rate distortion region and its dual channel coding capacity region are the same. This property can help one find a region based on its dual region. In general, there is no solution for the dual setting of the MAC. However, the rate distortion dual of the MAC with a common message has been solved. In [12], Asnani et. al. considered the SR problem with decoder cooperation and its channel coding duals. In this section we show how our methods of combined cooperation and cribbing can be implemented in the rate distortion dual. We establish the duality between the MAC with a common message, a private message, and combined cooperation and partial cribbing and the SR problem with combined cooperation and partial cribbing at the decoder. As expected, the rate region for the rate distortion dual consists of a single RV. Table I describes the principles of duality between channel coding and source coding. We start by defining the channel coding problem and state its capacity region. We continue by solving its rate distortion dual, i.e., the SR problem with combined cooperation and partial cribbing

TABLE I

PRINCIPLES OF DUALITY BETWEEN CHANNEL CODING AND SOURCE CODING

Channel coding	Source coding
Channel decoder	Source encoder
Encoder 1 input $(M_0, M_1) \in \{1, \ldots, 2^{n(R_0+R_1)}\}$	Decoder 1 input $(T_0, T_1) \in \{1, \ldots, 2^{n(R_0+R_1)}\}$
Encoder 1 output $X_1 \in \mathcal{X}_1$	Decoder 1 output $\hat{X}_1 \in \hat{\mathcal{X}}_1$
Encoder 2 input $M_0 \in \{1, \ldots, 2^{nR_0}\}$, $Z_i(X_{1,i}), M_{12}(M_0, M_1)$	Decoder 2 input $T_0 \in \{1, \ldots, 2^{nR_0}\}$, $Z_i(\hat{X}_{1,i}), T_{12}(T_0, T_1)$
Encoder 2 output $X_2 \in \mathcal{X}_2$	Decoder 2 output $\hat{X}_2 \in \hat{\mathcal{X}}_2$
Decoder input $Y \in \mathcal{Y}$	Encoder input $X \in \mathcal{X}$
Decoder output $(\hat{M}_0, \hat{M}_1) \in \{1, \ldots, 2^{n(R_0+R_1)}\}$	Encoder output $(T_0, T_1) \in \{1, \ldots, 2^{n(R_0+R_1)}\}$
Encoding function $f_1 : \mathcal{M}_0 \times \mathcal{M}_1 \mapsto \mathcal{X}_1^n$	Decoding function $g_1 : \mathcal{T}_0 \times \mathcal{T}_1 \mapsto \hat{\mathcal{X}}^n$
Causal cribbing encoding function $f_2 : \mathcal{M}_0 \times \mathcal{M}_{12} \times \mathcal{Z}^i \mapsto \mathcal{X}_{2,i}$	Causal cribbing decoding function $g : \mathcal{T}_0 \times \mathcal{T}_{12} \times \mathcal{Z}^i \mapsto \hat{\mathcal{X}}_{2,i}$
Decoding function $g : \mathcal{Y}^n \mapsto \mathcal{M}_0 \times \mathcal{M}_1$	Encoding function $f_0 : \mathcal{X}^n \mapsto \mathcal{T}_0, f_1 : \mathcal{X}^n \mapsto \mathcal{T}_1$
Auxiliary RV U	Auxiliary RV U
Joint distribution $p(u, x_1, x_2, y)$	Joint distribution $p(u, \hat{x}_1, \hat{x}_2, x)$
Constraint: $p(y|x_1, x_2)$ is fixed	Constraint: $p(x)$ is fixed

at the decoder. We end this section by pointing out the dualities between these two settings and show how the

PSfrag replacementspoints of the two regions are the same.

A. The MAC with a Common Message, a Private Message, and Combined Cooperation and Partial Cribbing

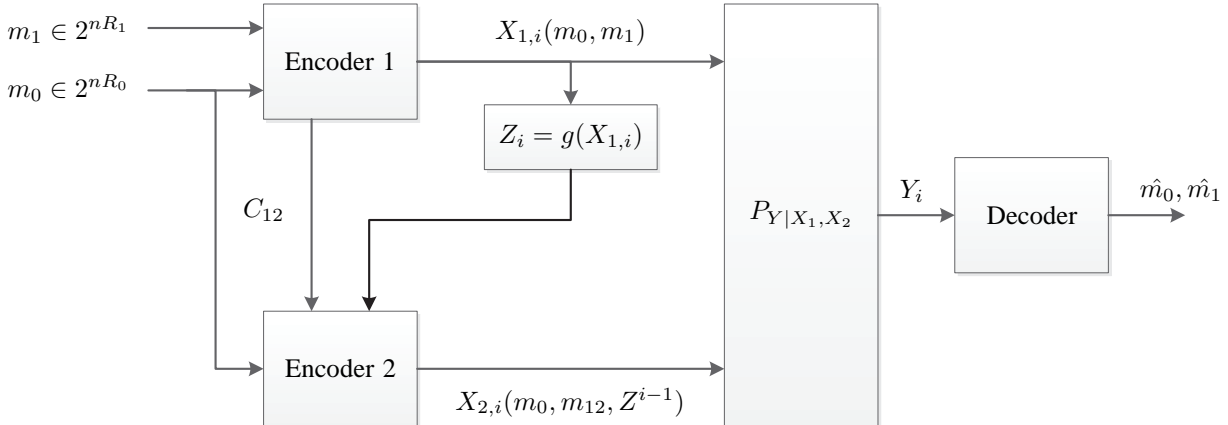Let us define the setting depicted in Fig. 5.

Fig. 5. MAC with common message, private message, and combined cooperation and cribbing. Encoder 2 obtains message M_{12} prior to transmission. The cribbing is done causally.

Definition 2 A $(2^{nR_0}, 2^{nR_1}, 2^{nC_{12}}, n)$ *code* for the MAC with a common message, a private message, and combined cooperation and partial cribbing, as shown in Fig. 5, consists at time i of encoding functions at Encoder 1 and Encoder 2

$$f_{12} \quad : \quad \{1, \ldots, 2^{nR_0}\} \times \{1, \ldots, 2^{nR_1}\} \mapsto \{1, \ldots, 2^{nC_{12}}\}, \tag{52}$$

$$f_1 \quad : \quad \{1, \ldots, 2^{nR_0}\} \times \{1, \ldots, 2^{nR_1}\} \mapsto \mathcal{X}_1^n, \tag{53}$$

$$f_{2,i} \quad : \quad \{1, \ldots, 2^{nR_0}\} \times \{1, \ldots, 2^{nC_{12}}\} \times \mathcal{Z}^i \mapsto \mathcal{X}_{2,i}, \tag{54}$$

and a decoding function

$$g : \mathcal{Y}^n \mapsto \{1, \ldots, 2^{nR_0}\} \times \{1, \ldots, 2^{nR_1}\}. \tag{55}$$

The average probability of error for a $(2^{nR_0}, 2^{nR_1}, 2^{nC_{12}}, n)$ code is defined as

$$P_e^{(n)} = \frac{1}{2^{n(R_0+R_1)}} \sum_{m_0, m_1} \Pr\{g(Y^n) \neq (m_0, m_1) | (m_0, m_1) \text{ sent}\}. \tag{56}$$

Let us define the following region and $\mathcal{R}_{MAC}$ that is contained in $\mathbb{R}_+^2$, namely, contained in the set of nonnegative two-dimensional real numbers.

$$\mathcal{R}_{MAC} = \left\{ \begin{array}{c} R_1 \leq I(X_1; Y | Z, U) + H(Z|U) + C_{12}, \\ R_0 + R_1 \leq I(X_1, U; Y), \text{ for} \\ P(u)P(x_1|u)\mathbb{1}_{z=f(x_1)}P(x_2|u, z)P(y|x_1, x_2). \end{array} \right\}. \tag{57}$$

Theorem 2 *(Capacity Region of the MAC with Combined Cooperation and Partial Cribbing)* The capacity region of the MAC with common message, private message, and combined cooperation and causal partial cribbing, as described in Def. 2, is $\mathcal{R}_{MAC}$.

Since the proof for Theorem 2 can be obtained by using the same methods described in Subsection II-B, it is omitted for brevity. We go on to define the SR setting with combined cooperation and partial cribbing at the decoders.

B. The Successive Refinement with Combined Cooperation and Partial Cribbing at the Decoders

We address the rate distortion setting depicted in Fig. 6.

The source sequence $X_i \in \mathcal{X}, i = 1, 2, \ldots$ is drawn i.i.d. $\sim p(x)$. Let $\hat{\mathcal{X}}_1$ and $\hat{\mathcal{X}}_2$ denote the reconstruction alphabets, and $d_i : \mathcal{X} \times \hat{\mathcal{X}}_i \mapsto [0, \infty)$, for $i = 1, 2$ denote single letter distortion measures. Distortion between sequences is defined in the usual way;

$$d_i(x^n, \hat{x}_i^n) = \frac{1}{n} \sum_{j=1}^n d_i(x_j, \hat{x}_{i,j}), \text{ for } i = 1, 2. \tag{58}$$

Definition 3 A $(2^{nR_0}, 2^{nR_1}, 2^{nC_{12}}, n)$ *rate-distortion code* for the SR with combined cooperation and partial cribbing at the decoders, as shown in Fig. 6, consists at time i of encoding functions at Encoder 1 and Encoder 2

$$f_0 \quad : \quad \mathcal{X}^n \mapsto \{1, \ldots, 2^{nR_0}\}, \tag{59}$$
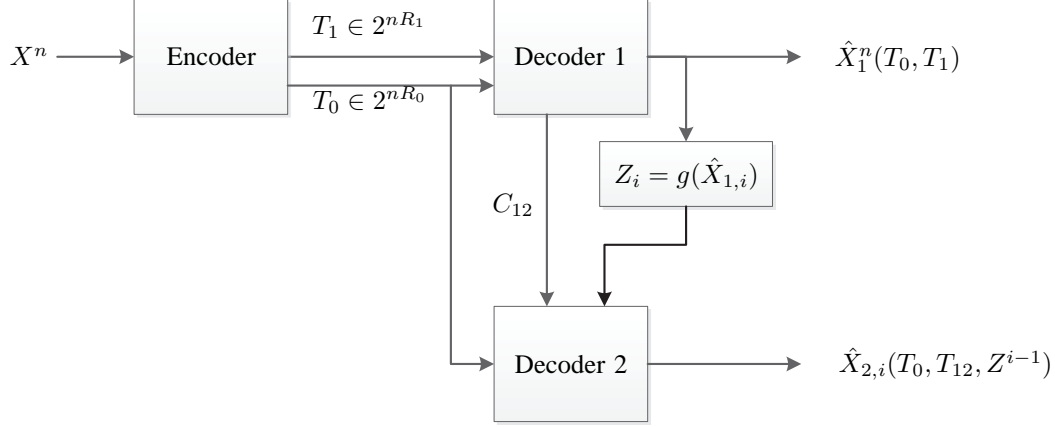
Fig. 6. SR with combined cooperation and partial cribbing at the decoders. The cribbing is done causally.

$$
f_1 \quad : \quad \mathcal{X}^n \mapsto \{1, \ldots, 2^{nR_1}\}, \tag{60}
$$

$$
f_{12} \quad : \quad \{1, \ldots, 2^{nR_0}\} \times \{1, \ldots, 2^{nR_1}\} \mapsto \{1, \ldots, 2^{nC_{12}}\}, \tag{61}
$$

$$
\tag{62}
$$

and a decoding function

$$
g_1 \quad : \quad \{1, \ldots, 2^{nR_0}\} \times \{1, \ldots, 2^{nR_1}\} \mapsto \hat{\mathcal{X}}_1^n, \tag{63}
$$

$$
g_{2,i} \quad : \quad \{1, \ldots, 2^{nR_0}\} \times \{1, \ldots, 2^{nC_{12}}\} \times \mathcal{Z}^i \mapsto \mathcal{X}_{2,i}. \tag{64}
$$

A rate (R_0, R_1, D_1, D_2) is said to be *achievable* for the SR with combined cooperation and partial cribbing at the decoders if $\forall \epsilon > 0$ and a $(2^{nR_0}, 2^{nR_1}, 2^{nC_{12}}, n)$ rate-distortion code the expected distortion for the decoders is bounded as,

$$
E\left[d_i(X^n, \hat{X}_i^n)\right] \le D_i + \epsilon, \text{ for } i = 1, 2. \tag{65}
$$

The *rate-distortion region* $\mathcal{R}(D_1, D_2)$ is defined as the closure of the set of all achievable rate-distortion tuples (R_0, R_1, D_1, D_2).

Let us define the following region $\mathcal{R}_{SR}(D_1, D_2)$ that is contained in $\mathbb{R}_+^2$, namely, contained in the set of nonnegative two-dimensional real numbers.

$$
\mathcal{R}_{SR}(D_1, D_2) = \left\{ \begin{array}{c} R_0 \ge I(X; Z, U) - H(Z|U) - C_{12}, \\ R_0 + R_1 \ge I(\hat{X}_1, U; X), \text{ for} \\ P(x, x_1, u)\mathbb{1}_{z=f(x_1), x_2=f(u, z_1)} \text{ s.t.} \\ E\left[d_i(X^n, \hat{X}_i^n)\right] \le D_i + \epsilon, \text{ for } i = 1, 2. \end{array} \right\}. \tag{66}
$$

Theorem 3 *(Rate Distortion Region of the Successive Refinement with Combined Cooperation and Partial Cribbing Decoders)* The rate-distortion region for the SR with combined cooperation and partial cribbing, as defined in Def.

3, is $\mathcal{R}_{SR}(D_1, D_2)$.

Proof:

Achievability: The achievability for this model is the same as in [12] where the achievable region was

$$
\begin{align}
\tilde{R}_0 &\geq I(X; Z, U) - H(Z|U), \\
\tilde{R}_0 + \tilde{R}_1 &\geq I(\hat{X}_1, U; X). \tag{67}
\end{align}
$$

In our case, we use rate splitting and set the following rates

$$
\begin{align}
\tilde{R}_0 &= R_0 + C_{12}, \tag{68} \\
\tilde{R}_1 &= R_0 - C_{12}. \tag{69}
\end{align}
$$

By setting these rates we obtain the region in (66).

Converse: Assume we have a $(2^{nR_0}, 2^{nR_1}, 2^{nC_{12}}, n)$ rate distortion code s.t. a (R_0, R_1, D_1, D_2) tuple is feasible. For the first inequality

$$
\begin{align}
nR_0 \quad &\geq \quad H(T_0) \tag{70} \\
&\overset{(a)}{=} \quad H(Z^n, T_0, T_{12}) - H(Z^n|T_{12}, T_0) - H(T_{12}|T_0) \tag{71} \\
&\overset{(b)}{\geq} \quad I(X^n; Z^n, T_0, T_{12}) - H(Z^n|T_{12}, T_0) - H(T_{12}) \tag{72} \\
&\overset{(c)}{\geq} \quad \sum_{i=1}^{n}[I(X_i; Z^n, T_0, T_{12}|X^{i-1}) - H(Z_i|T_{12}, T_0, Z^{i-1})] - nC_{12} \tag{73} \\
&= \quad \sum_{i=1}^{n}[I(X_i; Z^n, T_0, T_{12}, X^{i-1}) - H(Z_i|T_{12}, T_0, Z^{i-1})] - nC_{12} \tag{74} \\
&\overset{(d)}{\geq} \quad \sum_{i=1}^{n}[I(X_i; Z^i, T_0, T_{12}) - H(Z_i|T_{12}, T_0, Z^{i-1})] - nC_{12} \tag{75} \\
&\overset{(e)}{=} \quad \sum_{i=1}^{n}[I(X_i; Z_i, U_i) - H(Z_i|U_i)] - nC_{12} \tag{76} \\
&= \quad n\sum_{i=1}^{n}\frac{1}{n}[I(X_i; Z_i, U_i) - H(Z_i|U_i)] - nC_{12} \tag{77} \\
&\overset{(f)}{=} \quad n[I(X_Q; Z_Q, U_Q|Q) - H(Z_Q, U_Q|Q) - C_{12}] \tag{78} \\
&= \quad n[I(X_Q; Z_Q, U_Q, Q) - H(Z_Q, U_Q|Q) - C_{12}] \tag{79} \\
&\geq \quad n[I(X_Q; Z_Q, U_Q) - H(Z_Q, U_Q) - C_{12}], \tag{80}
\end{align}
$$

where (a) and (c) follow from the chain rule, (b) follows since conditionality reduces entropy, (d) follows since X_i is independent of X^{i-1}, (e) follows by setting the random variable $U_i = (Z^{i-1}, T_0, T_{12})$, and (f) follows by defining the RV Q independent of X^n and uniformly distributed over the set $\{1, 2, 3, \ldots, n\}$. For the second inequality

$$
\begin{align}
n(R_0 + R_1) \quad &\geq \quad H(T_0, T_1) \tag{81} \\
&\overset{(a)}{=} \quad I(X^n; T_0, T_1) \tag{82}
\end{align}
$$

$$= \sum_{i=1}^{n} I(X_i; T_0, T_1 | X^{i-1}) \tag{83}$$

$$\overset{(b)}{=} \sum_{i=1}^{n} I(X_i; T_0, T_1, X^{i-1}) \tag{84}$$

$$\overset{(c)}{=} \sum_{i=1}^{n} I(X_i; T_0, T_1, \hat{X}_{1,i}, Z^{i-1}, T_{12}, X^{i-1}) \tag{85}$$

$$\geq \sum_{i=1}^{n} I(X_i; \hat{X}_{1,i}, Z^{i-1}, T_0, T_{12}) \tag{86}$$

$$= \sum_{i=1}^{n} I(X_i; \hat{X}_{1,i}, U_i) \tag{87}$$

$$= nI(X_Q; \hat{X}_{1,Q}, U_Q), \tag{88}$$

where (a) follows since (T_0, T_1) is a function of X^n, (b) follows since since $X_{1,i}$ is independent of X_1^{i-1}, and (c) follows since $(\hat{X}_{1,i}, Z^{i-1}, T_{12})$ is a function of (T_0, T_1). We complete the proof by noting that the joint distribution of $(X_Q, \hat{X}_{1,Q}, Z_Q, U_Q)$ is the same as that of $(X, \hat{X}_1, Z, U)$. ∎

C. Duality Results Between the MAC and the Successive Refinement settings with combined cooperation and partial cribbing

After establishing Theorems 2 and 3, we now point out the dualities between the two settings. The similarity between the rate regions of the two settings is evident. Let us consider the corner points depicted in Table II and Fig. 7. One can see that the corner points are the same if we apply the duality rules $\hat{X}_1 \leftrightarrow X_1$, $\hat{X}_2 \leftrightarrow X_2$, $X \leftrightarrow Y$

TABLE II

CORNER POINTS OF MAC AND SR

	(R_0, R_1)
MAC	$(I(Y; Z, U) - H(Z|U) - C_{12}, I(Y; X_1|Z, U) + H(Z|U) + C_{12})$
(Theorem 2)	$(I(Y; X_1, U), 0)$
SR	$(I(X; Z, U) - H(Z|U) - C_{12}, I(X; \hat{X}_1|Z, U) + H(Z|U) + C_{12})$
(Theorem 3)	$(I(X; \hat{X}_1, U), 0)$

and $\geq \leftrightarrow \leq$. We notice that only one RV was used to describe the common message in both settings. This means that our methods of combining cooperation and cribbing can also be implemented in source coding problems. In the next section we address another case where only one RV is needed to describe both cooperation and cribbing.

V. STATE-DEPENDENT MAC WITH COMBINED COOPERATION AND PARTIAL CRIBBING

Following our results from Section II, we now show that our methods can also be implemented for a state-dependent channel where still only one auxiliary RV is needed. Let us consider the MAC with cooperation and non-causal state known at a partially cribbing encoder and at the decoder, depicted in Fig. 8.

We note that message M_{12} is sent prior to message M_{21}. For this model we address two different cases:

- The strictly causal case (sc) : Encoder 2 obtains Z_i with unit delay.

Fig. 7. Capacity region of the MAC and rate-distortion region of SR with combined cooperation and cribbing where A is $I(Y;X_1|Z,U)+H(Z|U)+C_{12}$ and B is $I(X;\hat{X}_1|Z,U)+H(Z|U)+C_{12}$.

Fig. 8. The MAC with cooperation and state known at a partially cribbing encoder and at the decoder. Encoder 1 and Encoder 2 obtain messages M_{21} and M_{12} prior to transmission. The partial cribbing is done strictly causally only by Encoder 2. This setting corresponds to the strictly causal case.

- The causal case (c) : Encoder 2 obtains Z_i without delay.

The channel probability does not depend on the time index i and is memoryless, i.e.,

$$P(y_i|x_1^i,x_2^i,s^i,y^{i-1})=P(y_i|x_{1,i},x_{2,i},s_i) \tag{89}$$

Definition 4 A $\left(2^{nR_1},2^{nR_2},2^{nC_{12}},2^{nC_{21}},n\right)$ *code* for the MAC with cooperation and non-causal state known at a partially cribbing encoder and at the decoder, as shown in Fig. 8, consists at time i of encoding functions at

Encoder 1 and Encoder 2.

$$f_{12} \quad : \quad \{1, \ldots, 2^{nR_1}\} \mapsto \{1, \ldots, 2^{nC_{12}}\}, \tag{90}$$

$$f_{21} \quad : \quad \{1, \ldots, 2^{nR_2}\} \times \mathcal{S}^n \times \{1, \ldots, 2^{nC_{12}}\} \mapsto \{1, \ldots, 2^{nC_{21}}\}, \tag{91}$$

$$f_1 \quad : \quad \{1, \ldots, 2^{nC_{21}}\} \times \{1, \ldots, 2^{nR_1}\} \mapsto \mathcal{X}_1^n, \tag{92}$$

$$f_{2,i}^{sc} \quad : \quad \{1, \ldots, 2^{nC_{12}}\} \times \{1, \ldots, 2^{nR_2}\} \times \mathcal{S}^n \times \mathcal{Z}^{i-1} \mapsto \mathcal{X}_{2,i}, \tag{93}$$

$$f_{2,i}^{c} \quad : \quad \{1, \ldots, 2^{nC_{12}}\} \times \{1, \ldots, 2^{nR_2}\} \times \mathcal{S}^n \times \mathcal{Z}^{i} \mapsto \mathcal{X}_{2,i}, \tag{94}$$

and a decoding function

$$g \quad : \quad \mathcal{S}^n \times \mathcal{Y}^n \mapsto \{1, \ldots, 2^{nR_1}\} \times \{1, \ldots, 2^{nR_2}\}. \tag{95}$$

The average probability of error for a $(2^{nR_1}, 2^{nR_2}, 2^{nC_{12}}, 2^{nC_{21}}, n)$ code is defined as

$$P_e^{(n)} = \frac{1}{2^{n(R_1+R_2)}} \sum_{m_1, m_2} \Pr\{g(Y^n, S^n) \neq (m_1, m_2) | (m_1, m_2) \text{ sent}\}. \tag{96}$$

Let us define the following regions, $\mathcal{R}_{State}^{sc}$ and $\mathcal{R}_{State}^{c}$, that are contained in $\mathbb{R}_+^2$, namely, contained in the set of nonnegative two-dimensional real numbers.

$$\mathcal{R}_{State}^{sc} = \left\{ \begin{array}{c} C_{21} \geq I(U; S), \\ R_1 \leq H(Z|U) + I(X_1; Y|S, U, X_2, Z) + C_{12}, \\ R_2 \leq I(X_2; Y|X_1, S, U) + C_{21} - I(U; S), \\ R_1 + R_2 \leq I(X_1, X_2; Y|S), \\ R_1 + R_2 \leq I(X_1, X_2; Y|U, Z, S) + H(Z|U) + C_{12} + C_{21} - I(U; S), \text{ for} \\ P(s)P(u|s)P(x_1|u)\mathbb{1}_{z=f(x_1)}P(x_2|s, u)P(y|x_1, x_2, s). \end{array} \right\} \tag{97}$$

The region $\mathcal{R}_{State}^{c}$ is defined with the same set of inequalities as in (97), but the joint distribution is of the form

$$P(s)P(u|s)P(x_1|u)\mathbb{1}_{z=f(x_1)}P(x_2|s, u, z)P(y|x_1, x_2, s). \tag{98}$$

Theorem 4 *(Capacity Region of the MAC with Cooperation and State Known at a Partial Cribbing Encoder)* The capacity regions of the MAC with cooperation and non-causal state known at a partially cribbing encoder and at the decoder for the strictly causal case and the causal case, as described in Def. 4, are $\mathcal{R}_{State}^{sc}$ and $\mathcal{R}_{State}^{c}$, respectively.

The role of the RV U is to generate an empirical coordination between the two encoders regarding the state channel and to generate a common message between the two encoders by combining the cooperation links and the partial cribbing. We now examine two special cases of this capacity region.

   Case 1: The One-Sided Cooperation and No Cribbing Case, i.e., $|\mathcal{Z}| = 1$ and $C_{12} = 0$: In this case $H(Z|U) = 0$ and hence the region $\mathcal{R}_{State}^{sc}$ coincides with the region in [23, Theorem 1].

Case 2: $|\mathcal{S}| = 1$, *The Memoryless Case:* Notice that in this case $I(U; S) = 0$ and the region $\mathcal{R}^{sc}_{State}$ reduces to

$$
\mathcal{R}^2_{State} = \left\{ \begin{array}{c} R_1 \leq H(Z|U) + I(X_1; Y|U, X_2, Z) + C_{12}, \\ R_2 \leq I(X_2; Y|X_1, U) + C_{21}, \\ R_1 + R_2 \leq I(X_1, X_2; Y), \\ R_1 + R_2 \leq I(X_1, X_2; Y|U, Z) + H(Z|U) + C_{12} + C_{21}, \text{ for} \\ P(u)P(x_1|u)\mathbb{1}_{z=f(x_1)}P(x_2|u)P(y|x_1, x_2). \end{array} \right\} \tag{99}
$$

which is the region in Theorem 1 where $Z_1 = Z$ and only Encoder 2 cribs from Encoder 1, i.e., $|\mathcal{Z}_2| = 1$.

The proof of Theorem 4 is given in Appendix B.

Although we have shown that for combined cooperation and cribbing only one auxiliary RV is needed to describe the capacity region, in some cases this is not possible. For instance, if the role of the cribbing and cooperation in the communication setting is different, then more then one auxiliary RV is needed. In the next section, we introduce a MAC with cooperation and action-dependent state known at a cribbing encoder. Because of the nature of actions and of non-causal states, the actions depend only on the cooperation and, therefore, two auxiliary RVs are needed, one for the cooperation and one for the cribbing.

VI. MAC WITH COOPERATION AND ACTION-DEPENDENT STATE KNOWN AT A CRIBBING ENCODER

We now address a MAC where two auxiliary RVs are needed in order to combine cooperation and cribbing. Consider the MAC with one-way cooperation and action-dependent state known at a cribbing encoder, depicted in Fig. 9. Notice that the action A^n is taken from (m_2, m_{12}).

We address two cases for this setting:

- The strictly causal case (sc) : Encoder 2 obtains $X_{1,i}$ with unit delay.
- The causal case (c) : Encoder 2 obtains $X_{1,i}$ without delay.

The channel probability is defined as in (89).

Definition 5 A $(2^{nR_1}, 2^{nR_2}, 2^{nC_{12}}, n)$ *code* for the MAC with one-way cooperation and action-dependent state known at a cribbing encoder, as shown in Fig. 9, consists at time i of encoding functions at Encoder 1 and Encoder 2

$$
f_{12} \quad : \quad \{1, \ldots, 2^{nR_1}\} \mapsto \{1, \ldots, 2^{nC_{12}}\}, \tag{100}
$$

$$
f_1 \quad : \quad \{1, \ldots, 2^{nR_1}\} \mapsto \mathcal{X}_1^n, \tag{101}
$$

$$
f_{Action} \quad : \quad \{1, \ldots, 2^{nR_2}\} \times \{1, \ldots, 2^{nC_{12}}\} \mapsto \mathcal{A}^n, \tag{102}
$$

$$
f^{sc}_{2,i} \quad : \quad \{1, \ldots, 2^{nR_2}\} \times \{1, \ldots, 2^{nC_{12}}\} \times \mathcal{S}^n \times \mathcal{X}_1^{i-1} \mapsto \mathcal{X}_{2,i}, \tag{103}
$$

$$
f^c_{2,i} \quad : \quad \{1, \ldots, 2^{nR_2}\} \times \{1, \ldots, 2^{nC_{12}}\} \times \mathcal{S}^n \times \mathcal{X}_1^i \mapsto \mathcal{X}_{2,i}, \tag{104}
$$

and a decoding function

$$
g \quad : \quad \mathcal{Y}^n \mapsto \{1, \ldots, 2^{nR_1}\} \times \{1, \ldots, 2^{nR_2}\}. \tag{105}
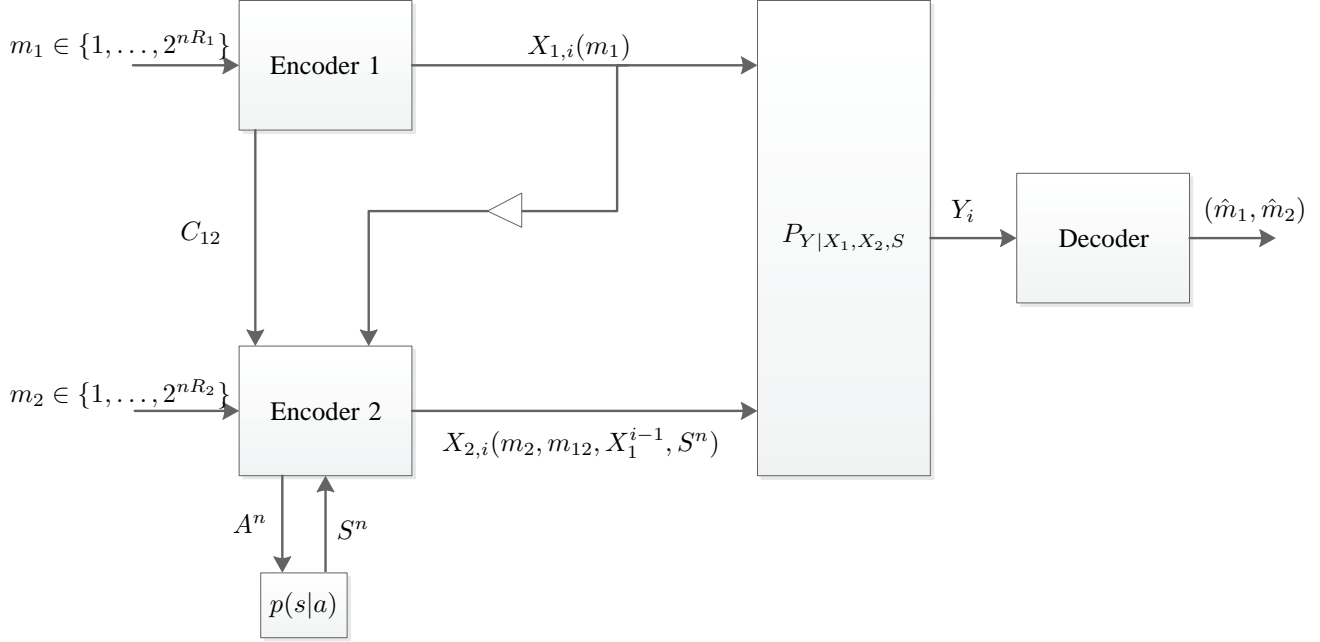$$

Fig. 9. The MAC with one-way cooperation and action-dependent state known at a cribbing encoder. Encoder 2 obtains messages M_{12} prior to transmission. The cribbing is done strictly causally only by Encoder 2. This setting corresponds to the strictly causal case.

The average probability of error for a $(2^{nR_1}, 2^{nR_2}, 2^{nC_{12}}, n)$ code is defined as

$$P_e^{(n)} = \frac{1}{2^{n(R_1+R_2)}} \sum_{m_1, m_2} \Pr\{g(Y^n) \neq (m_1, m_2)|(m_1, m_2) \text{ sent}\}. \tag{106}$$

Let us define the following regions $\mathcal{R}_{Action}^{sc}$ and $\mathcal{R}_{Action}^{c}$ that are contained in $\mathbb{R}_+^2$, namely, contained in the set of nonnegative two-dimensional real numbers.

$$\mathcal{R}_{Action}^{sc} = \left\{ \begin{array}{c} R_1 \leq \min\{H(X_1|V,W), I(Y;V,X_1,U|W,A) - I(S;U|W,V,A)\} + C_{12}, \\ R_2 \leq I(U,A;Y|X_1,V,W) - I(U;S|W,V,A), \\ R_1 + R_2 \leq I(X_1,V,U,A;Y|W) - I(U;S|W,V,A) + C_{12}, \\ R_1 + R_2 \leq I(X_1,V,U,A,W;Y) - I(U;S|W,V,A), \text{ for} \\ P(w)P(v|w)p(a|w)P(s|a)P(x_1|v,w)P(u,x_2|s,v,a,w)P(y|x_1,x_2,s). \end{array} \right\} \tag{107}$$

The region $\mathcal{R}_{Action}^{c}$ is defined with the same set of inequalities as in (107), but the joint distribution is of the form

$$P(w)P(v|w)p(a|w)P(s|a)P(x_1|v,w)P(u|s,v,a,w)P(x_2|v,u,s,a,w,x_1)P(y|x_1,x_2,s). \tag{108}$$

Theorem 5 *(Capacity Region of the MAC with Cooperation and Action-Dependent State Known at a Cribbing Encoder)* The capacity regions of the MAC with one-way cooperation and action-dependent state known at a strictly causal and causal cribbing encoder, as described in Def. 4, are $\mathcal{R}_{Action}^{sc}$ and $\mathcal{R}_{Action}^{c}$, respectively.

In this case, U is a Gelfand-Pinsker coding RV [19]. The role of the RV W is to generate a common message based on the cooperation link, whereas the RV V generates a common message based on the cribbing. The reason why in this case we cannot combine the cooperation and cribbing is that only part of the common information of both encoders is being used to generate the action sequence A^n. This example shows that in cases where only part of the common information that the encoders share is being used for arbitrary purposes, cooperation and cribbing cannot be combined into one RV. We now address two previous results in this field and show that they are special cases of our result.

Case 1: The Action-Dependent MAC where $C_{12} = R_1$: In this case the region reduces to

$$
\mathcal{R}_{Action}^1 = \left\{
\begin{array}{c}
R_2 \leq I(U, A; Y | X_1, V, W) - I(U; S | W, V, A), \\
R_1 + R_2 \leq I(X_1, V, U, A, W; Y) - I(U; S | W, V, A), \text{ for} \\
P(w)P(v|w)p(a|w)P(s|a)P(x_1|v,w)P(u,x_2|s,v,a,w)P(y|x_1,x_2,s).
\end{array}
\right\}
\tag{109}
$$

First, we notice that the cribbing in this case is redundant. Second, since the action is now taken from (M_1, M_2) we can set the RV $W = X_1$ and V as a constant and the region coincides with the capacity region in [24].

Case 2: The State-Dependent MAC with State Known at a Cribbing Encoder, i.e., $|\mathcal{A}| = 1$ and $C_{12} = 0$: Notice that in this case the state is not action-dependent and the region reduces to

$$
\mathcal{R}_{Action}^2 = \left\{
\begin{array}{c}
R_1 \leq H(X_1 | V, W), \\
R_2 \leq I(U; Y | X_1, V, W) - I(U; S | W, V), \\
R_1 + R_2 \leq I(X_1, V, U; Y | W) - I(U; S | W, V), \\
R_1 + R_2 \leq I(X_1, V, U, W; Y) - I(U; S | W, V), \text{ for} \\
P(w)P(v|w)P(s)P(x_1|v,w)P(u,x_2|s,v,w)P(y|x_1,x_2,s).
\end{array}
\right\}
\tag{110}
$$

If we set W as constant, the region coincides with the capacity region in [25]. Since these regions are equal, this shows that the capacity region in [25] is a special case of the region in Theorem 5.

The proof of Theorem 5 is given in Appendix C.

VII. CONCLUSIONS AND FUTURE WORK

In this paper, we have presented the capacity region for the MAC with combined cooperation and partial cribbing. Remarkably, the solution necessitates the use of only one auxiliary RV. Additionally, we have shown an achievability scheme for the Gaussian MAC with combined one-sided cooperation and causal partial cribbing. In this case, partial cribbing is a scalar quantization of Encoder 1's output obtained by Encoder 2. Graphs of achievability regions were presented for various number of quantization bits and capacity links. Using these results, it is possible to find under which conditions the outer bound is achieved. Thereafter, we considered a dual setting for the MAC with a common message, a private message, and combined cooperation and cribbing. We successfully characterized the rate-distortion region for the dual model using a single auxiliary RV. We applied our methods in order to find the capacity region for a MAC with cooperation and state known at a cribbing encoder and at the decoder. Again, the capacity region consisted of only one auxiliary RV. Finally, we addressed a MAC with one-way cooperation

and cribbing and action-dependent state, where the action was based on the cooperation between the encoders. In this case two auxiliary RVs were needed. We stated that if only part of the common information that the encoders share is being used for arbitrary purposes, then cooperation and cribbing cannot be combined into one RV. We suggest, for future work, considering the non-causal partial cribbing case and the interference channel with combined cooperation and cribbing. An additional case to consider is where the state or action is known at the weak encoder (the non-cognitive encoder).

APPENDIX A

ACHIEVABILITY FOR THE MAC WITH A COMMON MESSAGE AND PARTIALLY CRIBBING ENCODERS

Fix a joint distribution $P(u)P(x_1|u)\mathbb{1}_{z_1=f(x_1)}P(x_2|u)\mathbb{1}_{z_2=f(x_2)}P(y|x_1,x_2)$. In the following achievability scheme we use Block Markov Coding and Rate-Splitting.

Coding Scheme: We consider B blocks, each consisting of n symbols and thus we transmit nB symbols. We transmit $B-1$ message-pairs (M_1, M_2) in B blocks of information. Here, $M_i \in \{1, \ldots, 2^{nR_i}\}$ for $i \in \{1,2\}$; thus, asymptotically, for a large enough n, our transmission rate would be $\frac{nR_i(B-1)}{nB} \overset{n\to\infty}{\longrightarrow} R_i$ for $i \in \{1,2\}$. In each block we split messages M_1 and M_2 into (M_1', M_1'') and (M_2', M_2''), respectively, s.t. $R_1 = R_1' + R_1''$ and $R_2 = R_2' + R_2''$.

Code Design: Generate $2^{n(R_0+R_1'+R_2')}$ codewords u^n i.i.d. using $P(u^n) = \Pi_{i=1}^n P(u_i)$. For each u^n, generate $2^{nR_1'}$ codewords z_1^n i.i.d. using $P(z_1^n|u^n) = \Pi_{i=1}^n P(z_{1,i}|u_i)$ and $2^{nR_1''}$ codewords x_1^n i.i.d. using $P(x_1^n|u^n, z_1^n) = \Pi_{i=1}^n P(x_{1,i}|u_i, z_{1,i})$. Additionally, for each u^n, generate $2^{nR_2'}$ codewords z_2^n i.i.d. using $P(z_2^n|u^n) = \Pi_{i=1}^n P(z_{2,i}|u_i)$ and $2^{nR_2''}$ codewords x_2^n i.i.d. using $P(x_2^n|u^n, z_2^n) = \Pi_{i=1}^n P(x_{2,i}|u_i, z_{2,i})$.

Encoding: We denote the realizations of the sequences (M_0, M_1, M_2) at block b as $(m_{0,b}, m_{1,b}, m_{2,b})$. Since we use block Markov coding, we set $(m_{1,B}', m_{1,B}') = (1,1)$. In block $b \in \{1, \ldots, B\}$, encode message $(m_{0,b}, m_{1,b-1}', m_{2,b-1}')$ using $u^n(m_{0,b}, m_{1,b-1}', m_{2,b-1}')$. Encode message $m_{1,b}'$ conditioned on $(m_{0,b}, m_{1,b-1}', m_{2,b-1}')$ using $z_1^n(m_{1,b}', u^n)$ and message $m_{1,b}''$ conditioned on $(m_{0,b}, m_{1,b-1}', m_{2,b-1}', m_{1,b}')$ using $x_1^n(m_{1,b}'', u^n, z_1^n)$. Additionally, encode message $m_{2,b}'$ conditioned on $(m_{0,b}, m_{1,b-1}', m_{2,b-1}')$ using $z_2^n(m_{2,b}', u^n)$ and message $m_{2,b}''$ conditioned on $(m_{0,b}, m_{1,b-1}', m_{2,b-1}', m_{2,b}')$ using $x_2^n(m_{2,b}'', u^n, z_2^n)$. Send $x_1^n(m_{1,b}'', u^n, z_1^n)$ and $x_2^n(m_{2,b}'', u^n, z_2^n)$ over the channel.

Decoding at Encoder 1: At the end of block b, Encoder 1 tries to decode message $m_{2,b}'$. Given $(m_{0,b}, m_{1,b-1}')$ and assuming that message $m_{2,b-1}'$ was decoded correctly at the end of block $b-1$, Encoder 1 looks for $\hat{m}_{2,b}'$ s.t.

$$(u^n(m_{0,b}, m_{1,b-1}', m_{2,b-1}'), z_2^n(\hat{m}_{2,b}', u^n)) \in T_\epsilon^{(n)}(U, Z_2). \tag{111}$$

If no such $\hat{m}_{2,b}'$, or more than one such $\hat{m}_{2,b}'$, was found, an error is declared at block b and therefore in the whole super-block nB.

Decoding at Encoder 2: Similarly for Encoder 2; at the end of block b, Encoder 2 tries to decode message $m_{1,b}'$. Given $(m_{0,b}, m_{2,b-1}')$ and assuming that message $m_{1,b-1}'$ was decoded correctly at the end of block $b-1$, Encoder 2 looks for $\hat{m}_{1,b}'$ s.t.

$$(u^n(m_{0,b}, m_{1,b-1}', m_{2,b-1}'), z_1^n(\hat{m}_{1,b}', u^n)) \in T_\epsilon^{(n)}(U, Z_1). \tag{112}$$

If no such $\hat{m}'_{1,b}$, or more than one such $\hat{m}'_{1,b}$, was found, an error is declared at block b and therefore in the whole super-block nB.

Decoding at the receiver: At the end of block B, the decoding is done backwards. At block b, the decoder looks for the triplet $(\hat{m}_{0,b}, \hat{m}'_{1,b-1}, \hat{m}''_{1,b}, \hat{m}'_{2,b-1}, \hat{m}''_{2,b})$ s.t.

$$(u^n(\hat{m}_{0,b}, \hat{m}'_{1,b-1}, \hat{m}'_{2,b-1}), z_1^n(\hat{m}'_{1,b}, u^n), z_2^n(\hat{m}'_{2,b}, u^n), x_1^n(\hat{m}''_{1,b}, u^n, z_1^n), x_2^n(\hat{m}''_{2,b}, u^n, z_2^n), y^n)$$
$$\in T_\epsilon^{(n)}(U, Z_1, Z_2, X_1, X_2, Y). \tag{113}$$

If no such tuple, or more than one such tuple, was found, an error is declared at block b and therefore in the whole super-block nB.

Error Analysis: The probability that $z_1^n(1, u^n) = z_1^n(i, u^n)$ where $i > 1$ and where $(u^n(i), z_1^n(1, u^n)) \in T_\epsilon^{(n)}(U, Z_1)$ is bounded by $2^{-n(H(Z_1|U)-\delta(\epsilon))}$, where $\delta(\epsilon)$ goes to zero as ϵ goes to zero. Hence, if

$$R'_1 < H(Z_1|U), \tag{114}$$

then the probability that an incorrect message $m'_{1,b}$ was decoded goes to zero for a large enough n.

From symmetry, we can see that if

$$R'_2 < H(Z_2|U), \tag{115}$$

then the probability that an incorrect message $m'_{2,b}$ was decoded goes to zero for a large enough n. We define the following event at block b:

$$E_{i,j,k,b} \triangleq (u^n(i), z_1^n(\hat{m}'_{1,b}, u^n), z_2^n(\hat{m}'_{2,b}, u^n), x_1^n(j, u^n, z_1^n), x_2^n(k, u^n, z_2^n), y^n) \in T_\epsilon^{(n)}(U, Z_1, Z_2, X_1, X_2, Y). \tag{116}$$

We can bound the probability of error as follows:

$$P_{e,b}^{(n)} \le \Pr(E_{1,1,1,b}^c) \quad + \sum_{i=1,j>1,k=1} \Pr(E_{1,j,1,b}) + \sum_{i=1,j=1,k>1} \Pr(E_{1,1,k,b})$$
$$+ \sum_{i=1,j>1,k>1} \Pr(E_{1,j,k,b}) + \sum_{i>1,j>1,k>1} \Pr(E_{i,j,k,b}). \tag{117}$$

We now show that each term in (117) goes to zero for a large enough n.

- Upper-bounding $\Pr(E_{1,1,1,b}^c)$: Since we assume that Transmitters 1 and 2 encode the correct message triplet $(m_{0,b}, m_{1,b-1}, m_{2,b-1})$ at block b and that the receiver decoded the right triplet $(m_{0,b+1}, m_{1,b}, m_{2,b})$ at block $b+1$, by the law of large numbers (LLN), $\Pr(E_{1,1,1,b}^c) \to 0$ when $n \to \infty$.

- Upper-bounding $\sum_{i=1,j>1,k=1} \Pr(E_{1,j,1,b})$: Assuming that $(m'_{1,b}, m'_{2,b})$ were decoded correctly at block $b+1$, the probability for this event is bounded by

$$\sum_{i=1,j>1,k=1} \Pr(E_{1,j,1,b}) \le 2^{nR''_1} 2^{n(I(X_1;Y|U,Z_1,X_2)-\delta(\epsilon))}. \tag{118}$$

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

$$\sum_{i=1,j=1,k>1} \Pr(E_{1,1,k,b}) \le 2^{nR''_2} 2^{n(I(X_2;Y|U,Z_2,X_1)-\delta(\epsilon))}. \tag{119}$$

- Upper-bounding $\sum_{i=1,j>1,k>1} \Pr(E_{1,j,k,b})$: Again we assume that $(m'_{1,b}, m'_{2,b})$ were decoded correctly at block $b+1$; the probability for this event is bounded by

$$\sum_{i=1,j>1,k>1} \Pr(E_{1,j,k,b}) \leq 2^{n(R''_1 + R''_2)} 2^{n(I(X_1,X_2;Y|U,Z_1,Z_2) - \delta(\epsilon))}. \tag{120}$$

- Upper-bounding $\sum_{i>1,j>1,k>1} \Pr(E_{i,j,k,b})$: We assume that $(m'_{1,b}, m'_{2,b})$ were decoded correctly at block $b+1$; the probability for this event is bounded by

$$\sum_{i>1,j>1,k>1} \Pr(E_{i,j,k,b}) \leq 2^{n(R_0 + R_1 + R_2)} 2^{n(I(X_1,X_2;Y) - \delta(\epsilon))}. \tag{121}$$

Using the Fourier-Motzkin Elimination on equations (114), (115), (118), (119), (120), and (121) yields the achievable region in (42), thus completing the proof. ∎

APPENDIX B
PROOF OF THEOREM 4

A. Converse

Converse for the strictly causal case: Given an achievable rate-pair (R_1, R_2) we need to show that there exists a joint distribution of the form $P(s)P(u|s)P(v|u)P(z, x_1|v, u)P(x_2|s, v, u)P(y|x_1, x_2, s)$ such that the inequalities in (97) are satisfied. Since (R_1, R_2) is an achievable rate-pair, there exists a $(2^{nR_1}, 2^{nR_2}, 2^{nC_{12}}, 2^{nC_{21}}, n)$ code with an arbitrarily small error probability $P_e^{(n)}$. By Fano's inequality,

$$H(M_1, M_2 | Y^n, S^n) \leq n(R_1 + R_2)P_e^{(n)} + H(P_e^{(n)}). \tag{122}$$

We set

$$(R_1 + R_2)P_e^{(n)} + \frac{1}{n}H(P_e^{(n)}) \triangleq \epsilon_n, \tag{123}$$

where $\epsilon_n \to 0$ as $P_e^{(n)} \to 0$. Hence,

$$H(M_1 | Y^n, M_2, S^n) \leq H(M_1, M_2 | Y^n, S^n) \leq n\epsilon_n, \tag{124}$$

$$H(M_2 | Y^n, M_1, S^n) \leq H(M_1, M_2 | Y^n, S^n) \leq n\epsilon_n. \tag{125}$$

For R_1 we have the following:

$$nR_1 = H(M_1) \tag{126}$$

$$= H(M_1 | M_{12}) + H(M_{12}) \tag{127}$$

$$\overset{(a)}{=} H(M_1 | M_{12}, M_2, S^n) + H(M_{12}) \tag{128}$$

$$= I(M_1; Y^n | M_{12}, M_2, S^n) + H(M_1 | Y^n, M_{12}, M_2, S^n) + H(M_{12}) \tag{129}$$

$$\overset{(b)}{\leq} I(M_1; Y^n | M_{12}, M_2, S^n) + nC_{12} + n\epsilon_n \tag{130}$$

$$\overset{(c)}{=} I(X_1^n, Z^n; Y^n | M_{12}, M_2, S^n) + nC_{12} + n\epsilon_n \tag{131}$$

$$\overset{(d)}{=} I(Z^n; Y^n | M_{12}, M_2, S^n) + I(X_1^n; Y^n | M_{12}, M_2, S^n, Z^n) + nC_{12} + n\epsilon_n \tag{132}$$

$$\overset{(e)}{=} \sum_{i=1}^{n}[I(Z_i;Y^n|M_{12},M_{21},M_2,Z^{i-1},S^n) + I(X_1^n;Y_i|Y^{i-1},M_{12},M_{21},M_2,S^n,Z^n)]$$
$$+nC_{12} + n\epsilon_n \tag{133}$$

$$\overset{(f)}{\leq} \sum_{i=1}^{n}[H(Z_i|M_{21},Z^{i-1},M_{12},S^{i-1}) + I(X_1^n;Y_i|Y^{i-1},M_{12},M_{21},M_2,S^n,Z^n,X_{2,i})]$$
$$+nC_{12} + n\epsilon_n \tag{134}$$

$$\overset{(g)}{\leq} \sum_{i=1}^{n}[H(Z_i|M_{21},Z^{i-1},M_{12},S^{i-1}) + I(X_{1,i};Y_i|M_{21},S^i,Z^{i-1},M_{12},X_{2,i},Z_i)]$$
$$+nC_{12} + n\epsilon_n \tag{135}$$

$$\overset{(h)}{=} \sum_{i=1}^{n}[H(Z_i|U_i) + I(X_{1,i};Y_i|U_i,X_{2,i},S_i,Z_i)] + nC_{12} + n\epsilon_n, \tag{136}$$

where (a) follows from the fact that the messages M_1 and (M_2,S^n) are independent, (b) follows from Fano's inequality, (c) follows from the Markov chain $M_1 - (X_1^n,Z^n,M_{12},M_2,S^n) - Y^n$, (d) and (e) follow from the chain rule and since $M_{21} = f(S^n,M_2,M_{12})$, (f) follows since conditioning reduces entropy and since $X_{2,i} = f(S^n,Z^{i-1},M_{12},M_2)$, (g) follows from the Markov Chain $Y_i - (X_{1,i},X_{2,i},S^i,M_{12},M_{21},Z^i) - (Y^{i-1},M_2,S_{i+1}^n,Z_{i+1}^n)$, and (h) follows by setting the RV

$$U_i \triangleq (M_{12},M_{21},Z^{i-1},S^{i-1}). \tag{137}$$

Thus, we obtained

$$R_1 \leq \frac{1}{n}\sum_{i=1}^{n}[H(Z_i|U_i) + I(X_{1,i};Y_i|U_i,X_{2,i},S_i,Z_i)] + C_{12} + \epsilon_n. \tag{138}$$

Next, we consider R_2;

$$nR_2 = H(M_2) \tag{139}$$

$$\overset{(a)}{=} H(M_2|S^n,M_1) \tag{140}$$

$$\overset{(b)}{=} H(M_{21},M_2|S^n,M_1) \tag{141}$$

$$= H(M_{21}|S^n,M_1) + H(M_2|S^n,M_{21},M_1) \tag{142}$$

$$\overset{(c)}{\leq} H(M_{21}|M_1) - I(M_{21};S^n|M_1) + I(M_2;Y^n|S^n,M_1,M_{21}) + n\epsilon_n \tag{143}$$

$$\overset{(d)}{\leq} nC_{21} + \sum_{i=1}^{n}[I(M_2;Y_i|Y^{i-1},S^n,M_1,M_{21}) - I(S_i;M_{21}|S^{i-1},M_1)] + n\epsilon_n \tag{144}$$

$$\overset{(e)}{=} nC_{21} + \sum_{i=1}^{n}[I(M_2,X_{2,i};Y_i|Y^{i-1},M_1,M_{12},M_{21},S^n,X_{1,i},Z^{i-1})$$
$$-I(S_i;M_{21},S^{i-1},M_1,M_{12},Z^{i-1})] + n\epsilon_n \tag{145}$$

$$\overset{(f)}{\leq} nC_{21} + \sum_{i=1}^{n}[I(X_{2,i};Y_i|M_{21},M_{12},S^i,Z^{i-1},X_{1,i})$$
$$-I(S_i;M_{21},S^{i-1},M_{12},Z^{i-1})] + n\epsilon_n \tag{146}$$

$$= nC_{21} + \sum_{i=1}^{n}[I(X_{2,i};Y_i|U_i,S_i,V_i,X_{1,i}) - I(S_i;U_i)] + n\epsilon_n, \tag{147}$$

where (a) follows since M_2 is independent of S^n and M_1, (b) follows since $M_{21} = f(S^n, M_2, M_1)$, (c) follows from Fano's inequality, (d) follows from the chain rule, (e) follows since S_i is independent of (S^{i-1}, M_1) and since $(M_{12}, Z^{i-1}, X_{1,i}) = f(M_{21}, M_1)$, and (f) follows from the same argument as in (135) and since conditioning reduces entropy. Thus, we obtained

$$R_2 \leq C_{21} + \frac{1}{n}\sum_{i=1}^{n}[I(X_{2,i}; Y_i | U_i, S_i, X_{1,i}) - I(S_i; U_i)] + \epsilon_n. \tag{148}$$

Now, consider

$$n(R_1 + R_2) = H(M_1, M_2, M_{12}) \tag{149}$$

$$\overset{(a)}{=} H(M_1, M_2 | S^n, M_{12}) + H(M_{12}) \tag{150}$$

$$\leq H(M_1, M_2 | M_{21}, S^n, M_{12}) + H(M_{21} | S^n, M_{12}) + nC_{12} \tag{151}$$

$$\overset{(c)}{\leq} nC_{12} + I(M_1, M_2, Z^n; Y^n | S^n, M_{12}, M_{21}) + H(M_{21} | S^n, M_1) + n\epsilon_n \tag{152}$$

$$\leq nC_{12} + I(M_1, M_2; Y^n | S^n, M_{12}, M_{21}, Z^n) + I(Z^n | S^n, M_{12}, M_{21})$$

$$+ H(M_{21} | S^n, M_1) + n\epsilon_n \tag{153}$$

$$\overset{(d)}{\leq} nC_{12} + I(X_1^n, X_2^n; Y^n | S^n, M_{12}, M_{21}, Z^n) + nC_{21}$$

$$+ \sum_{i=1}^{n}[H(Z_i | U_i) - I(S_i; U_i)] + n\epsilon_n \tag{154}$$

$$\overset{(e)}{=} nC_{12} + nC_{21} + \sum_{i=1}^{n}[I(X_1^n, X_2^n; Y_i | S^n, Y^{i-1}, M_{12}, M_{21}, Z^n)$$

$$+ H(Z_i | U_i) - I(S_i; U_i)] + n\epsilon_n \tag{155}$$

$$\overset{(f)}{=} nC_{12} + nC_{21} + \sum_{i=1}^{n}[I(X_{1,i}, X_{2,i}; Y_i | S_i, S^{i-1}, M_{21}, M_{12}, Z^i)$$

$$+ H(Z_i | U_i) - I(S_i; U_i)] + n\epsilon_n \tag{156}$$

$$\leq nC_{12} + nC_{21} + \sum_{i=1}^{n}[I(X_{1,i}, X_{2,i}; Y_i | S_i, U_i, Z_i)$$

$$+ H(Z_i | U_i) - I(S_i; U_i)] + n\epsilon_n, \tag{157}$$

where (a) follows since (M_1, M_2) is independent of S^n, (b) follows since $Z^n = f(M_1, M_{21})$, (c) follows from Fano's inequality and since M_{21} is independent of M_{12}, (d) follows from the same arguments as given in (143)-(146) and from the Markov chain $(M_1, M_2) - (X_1^n, X_2^n, M_{12}, M_{21}, Z^n, S^n) - Y^n$, (e) follows from the chain rule, and (f) follows from the same argument as given in (135). Thus, we obtained

$$R_1 + R_2 \leq C_{12} + C_{21} + \frac{1}{n}\sum_{i=1}^{n}[I(X_{1,i}, X_{2,i}; Y_i | S_i, U_i, Z_i) + H(Z_i | U_i) - I(S_i; U_i)] + \epsilon_n. \tag{158}$$

Additionally,

$$n(R_1 + R_2) \leq H(M_1, M_2) \tag{159}$$

$$\overset{(a)}{=} H(M_1, M_2 | S^n) \tag{160}$$

$$\overset{(b)}{\leq} \quad I(M_1, M_2; Y^n | S^n) + n\epsilon_n \tag{161}$$

$$\overset{(c)}{\leq} \quad I(X_1^n, X_2^n; Y^n | S^n) + n\epsilon_n \tag{162}$$

$$\overset{(d)}{=} \quad \sum_{i=1}^{n} I(X_1^n, X_2^n; Y_i | S^n, Y^{i-1}) + n\epsilon_n \tag{163}$$

$$\overset{(e)}{\leq} \quad \sum_{i=1}^{n} I(X_{1,i}, X_{2,i}; Y_i | S_i) + n\epsilon_n, \tag{164}$$

where (a) follows since (M_1, M_2) is independent of S^n, (b) follows from Fano's inequality, (c) follows from encoding relations (90)-(94), (d) follows from the chain rule, and step (e) follows from the Markov Chain $Y_i - X_{1,i}, X_{2,i}, S_i - Y^{i-1}$ and since conditioning reduces entropy. Thus we obtained

$$R_1 + R_2 \quad \leq \quad \frac{1}{n} \sum_{i=1}^{n} I(X_{1,i}, X_{2,i}; Y_i | S_i) + \epsilon_n. \tag{165}$$

Finally,

$$nC_{21} \quad \geq \quad H(M_{21}) \tag{166}$$

$$\geq \quad H(M_{21} | M_1) \tag{167}$$

$$\geq \quad I(M_{21}; S^n | M_1) \tag{168}$$

$$= \quad \sum_{i=1}^{n} I(S_i; M_{21} | S^{i-1}, M_1) \tag{169}$$

$$\overset{(a)}{=} \quad \sum_{i=1}^{n} I(S_i; M_{21}, S^{i-1}, M_1) \tag{170}$$

$$\geq \quad \sum_{i=1}^{n} I(S_i; M_{21}, S^{i-1}, Z^{i-1}, M_{12}) \tag{171}$$

$$= \quad \sum_{i=1}^{n} I(S_i; U_i), \tag{172}$$

where (a) follows since S_i is independent of (S^{i-1}, M_1). Finally, let Q be an RV independent of (X_1^n, X_2^n, Y^n) and uniformly distributed over the set $\{1, 2, 3, \ldots, n\}$. We define the RV $U \triangleq (Q, U_Q)$ and obtain the region given in (97).

To complete the converse, we need to show the following Markov relations:

- $Z_i - U_i - S_i$, $X_{1,i} - (U_i, Z_i) - S_i$, and $X_{2,i} - (M_{12}, M_{21}, Z^{i-1}, S^{i-1}) - X_{1,i}$ - These Markov relations can be proven by using the undirected graph method in Fig. 10. For the first Markov chain, see that it is impossible to get from node Z_i to node S_i without going through nodes $(S^{i-1}, Z^{i-1}, M_{12}, M_{21})$. For the second Markov chain, it is impossible to get from node $X_{1,i}$ to node S_i without going through nodes $(S^{i-1}, Z^i, M_{12}, M_{21})$. Finally, for the third Markov chain, we can see that it is impossible to get from node $X_{1,i}$ to node $X_{2,i}$ without going through nodes $(S^i, Z^{i-1}, M_{12}, M_{21})$.
- $Y_i - (X_{1,i}, X_{2,i}) - (Z_{1,i}, U_i)$ - Follows from the fact that the channel output at any time i is assumed to depend only on the channel inputs and state at time i.
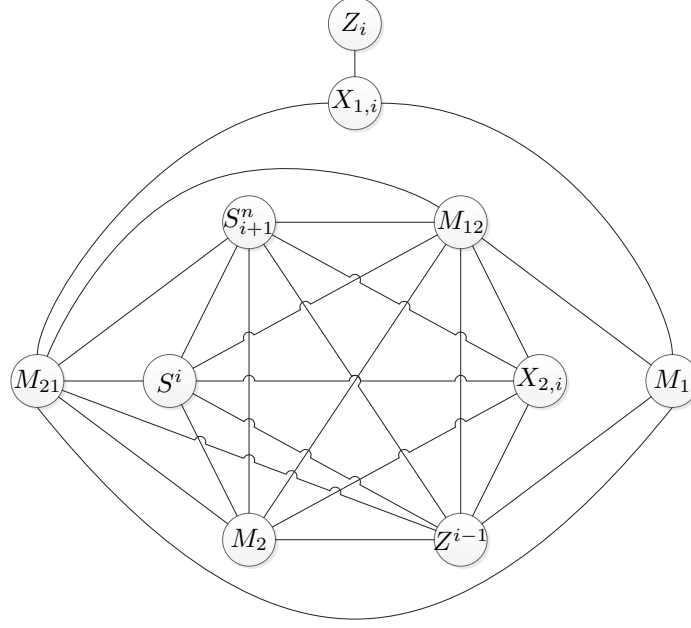
Fig. 10.   Proof of the Markov chains $Z_i - U_i - S_i$, $X_{1,i} - (U_i, Z_i) - S_i$, and $X_{2,i} - (M_{12}, M_{21}, Z^{i-1}, S^{i-1}) - X_{1,i}$ using the undirected graphical technique [16, Section II]. This graph corresponds to the joint distribution $P(s^n)P(m_1)P(m_2)P(m_{12}|m_1)P(m_{21}|m_2, s^n, m_{21})P(z^{i-1}|m_1, m_{21})P(x_{1,i}|m_{21}, m_1)P(z_i|x_{1,i})P(x_{2,i}|m_{12}, m_2, s^n, z^{i-1})$.

This completes the converse part.   ∎

Converse for the causal case: For the causal case we repeat the same converse as for the strictly causal case, except that in the final step we need to show the Markov chain $X_{2,i} - (U_i, Z_i, S_i) - X_{1,i}$, rather than $X_{2,i} - (U_i, S_i) - X_{1,i}$, as in the strictly causal case. If we change node Z^{i-1} to Z^i in Fig. 10, we can see that the Markov chain $X_{2,i} - (M_{12}, M_{21}, Z^i, S^i) - X_{1,i}$ holds since we cannot get from node $X_{2,i}$ to node $X_{1,i}$ without going through nodes $(M_{12}, M_{21}, Z^i, S^i)$.   ∎

B. Achievability

In order to prove the achievability, we will consider a similar setting and then, by doing minor modifications, we will prove our setting. We first prove the achievability for the strictly causal case.

Achievability for the strictly causal case: Let us look at a similar model depicted in Fig. 11.

First, we will solve the achievability for this model. Fix a joint distribution $P(s)P(u|s)P(z, x_1|u)P(x_2|s, u)P(y|x_1, x_2, s)$ where $P(s)$ and $P(y|x_1, x_2, s)$ are given by the channel. In the following achievability scheme we use block Markov coding, rate splitting, and double binning.

Coding Scheme: We consider B blocks, each consisting of n symbols; thus we transmit nB symbols. We transmit $B-1$ messages M_1 in B blocks of information. Here, $M_1 \in \{1, \ldots, 2^{nR_1}\}$; thus asymptotically, for a large enough n, our transmission rate would be $\frac{nR_1(B-1)}{nB} \xrightarrow{n \to \infty} R_1$. At each block we split messages M_1 and M_2 into (M_1', M_1'') and (M_2', M_2'') at rates (R_1', R_1'') and (R_2', R_2''), respectively. We note that $R_1' + R_1'' = R_1$ and $R_2' + R_2'' = R_2$.
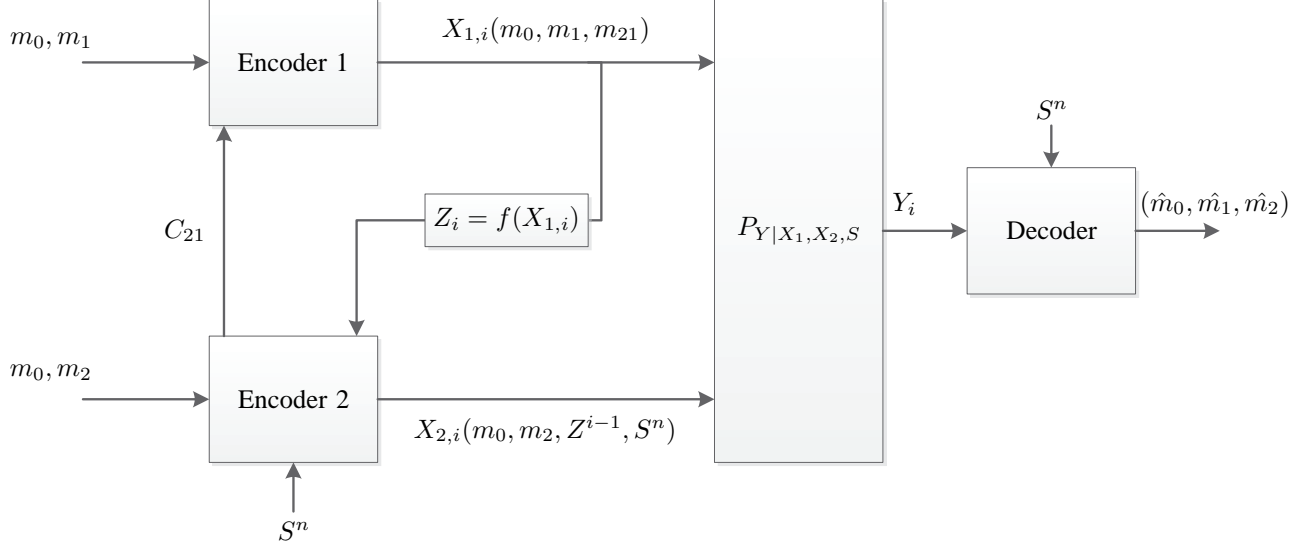
Fig. 11.   MAC with a common message and state known at a partially cribbing Encoder.

Code Design: The following binning process is depicted in Fig. 12. Generate $2^{n(R_0+R'_1+C_{21})}$ codewords u^n i.i.d. using $P(u^n) = \Pi_{i=1}^n P(u_i)$. Bin all u^ns into $2^{n(R_0+R'_1)}$ super-bins. In each super-bin, bin all u^ns into $2^{nR'_2}$ bins. Thus we have $2^{n(R_0+R'_1)}$ super-bins, each consisting of $2^{nR'_2}$ bins, where in each bin we have $2^{n(C_{21}-R'_2)}$ u^n codewords. For each u^n, generate $2^{nR'_1}$ codewords z^n i.i.d. using $P(z^n|u^n) = \Pi_{i=1}^n P(z_i|u_i)$. For each pair (u^n, z^n), generate $2^{nR''_1}$ codewords x_1^n i.i.d. using $P(x_1^n|u^n, z^n) = \Pi_{i=1}^n P(x_{1,i}|u_i, z_i)$. Additionally, for each pair (u^n, s^n), generate $2^{nR''_2}$ codewords x_2^n i.i.d. using $P(x_2^n|u^n, s^n) = \Pi_{i=1}^n P(x_{2,i}|u_i, s_i)$.
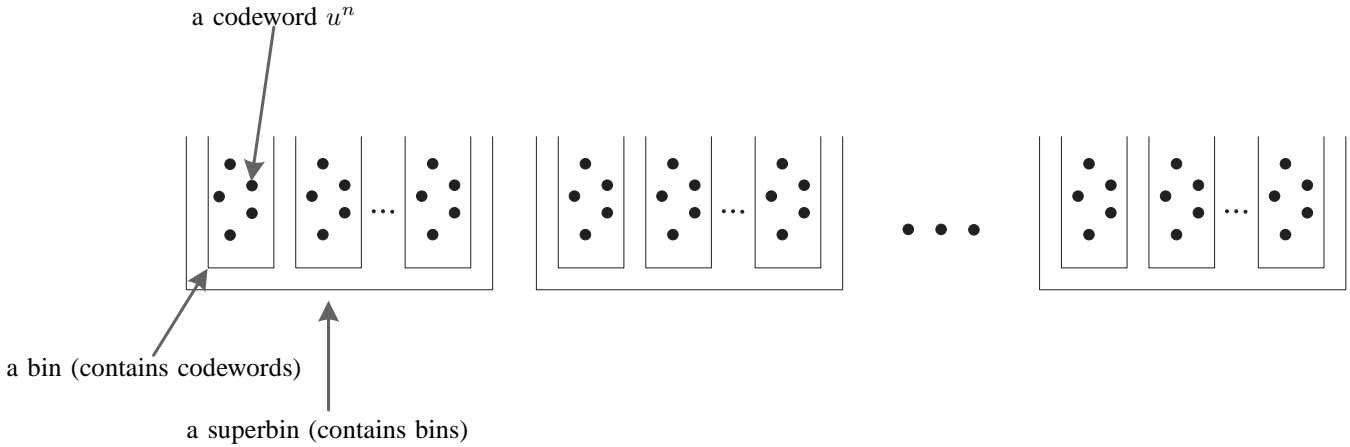

Fig. 12.   The binning process as explained in the code design. There are $2^{n(R_0+R'_1)}$ super-bins and $2^{nR'_2}$ bins in each super-bin. The number of codewords in each bin must be greater than $I(U; S)$ in order to find u^n such that $(u^n, s^n) \in T_\epsilon^{(n)}(U, S)$.

Encoding: We denote the realizations of the sequences $(M_0, M'_1, M''_1, M'_2, M''_2)$ at block b as $(m_{0,b}, m'_{1,b}, m''_{1,b}, m'_{2,b}, m''_{2,b})$. Since we use block Markov coding, we set $m'_{1,B} = 1$. In block $b \in \{1, \dots, B\}$,

Encoder 2 looks in super-bin $(m_{0,b}, m'_{1,b})$ and bin $m'_{2,b}$ for u^n such that $(u^n, s^n) \in T_\epsilon^{(n)}(U, S)$ and sends its index l inside the super-bin over the rate-limited cooperation link to Encoder 1, where $l \in \{1, \ldots, 2^{nC_{21}}\}$. If such a codeword u^n does not exist, namely, among the codewords in the bin none is jointly typical with s^n, choose an arbitrary u^n from the bin $m'_{2,b}$ (in such a case the decoder will declare an error). Encoder 1 looks in super-bin $(m_{0,b}, m'_{1,b})$ for the bin that $u^n(l)$ lies in. That bin's index is $m'_{2,b}$. Then, Encoder 1 encodes message $m'_{1,b}$ conditioned on $(m_{0,b}, m_{1,b-1}, m'_{2,b})$ using $z^n(m'_{1,b}, u^n)$ and encodes message $m''_{1,b}$ conditioned on $(m_{0,b}, m_{1,b-1}, m'_{2,b}, m'_{1,b})$ using $x_1^n(m''_{1,b}, u^n, z^n)$. Encoder 2 encodes message $m''_{2,b}$ conditioned on $(m_{0,b}, m_{1,b-1}, m'_{2,b})$ and s^n using $x_2^n(m_{2,b}, u^n, s^n)$. Send $x_1^n(m''_{1,b}, u^n, z^n)$ and $x_2^n(m''_{2,b}, u^n, s^n)$ over the channel.

Decoding at Encoder 2: At the end of block b, Encoder 2 tries to decode message $m'_{1,b}$. Given $(m_{0,b}, m'_{2,b})$ and assuming that message $m'_{1,b-1}$ was decoded correctly at the end of block $b-1$, Encoder 2 looks for $\hat{m}'_{1,b}$ s.t.

$$(u^n(m_{0,b}, m'_{1,b-1}, m'_{2,b}), z^n(\hat{m}'_{1,b}, u^n)) \in T_\epsilon^{(n)}(U, Z). \tag{173}$$

If no such $\hat{m}'_{1,b}$, or more than one such $\hat{m}'_{1,b}$, was found, an error is declared at block b and therefore in the whole super-block nB.

Decoding at the receiver: At the end of block B, the decoding is done backwards. At block b, assuming that $(m_{0,b+1}, m_{1,b}, m'_{2,b+1})$ was decoded correctly in block $b+1$, the decoder looks for the set $(\hat{m}_{0,b}, \hat{m}'_{1,b-1}, \hat{m}''_{1,b}, \hat{m}'_{2,b}, \hat{m}''_{2,b})$ s.t.

$$(u^n(\hat{m}_{0,b}, \hat{m}'_{1,b-1}, \hat{m}'_{2,b}, s^n), z^n(\hat{m}'_{1,b}, u^n), x_1^n(m''_{1,b}, u^n, z^n), x_2^n(\hat{m}''_{2,b}, u^n, s^n), s^n, y^n) \in T_\epsilon^{(n)}(U, Z, X_1, X_2, S, Y).$$

If no such tuple, or more than one such tuple, was found, an error is declared in block b and therefore at the whole super-block nB.

Error Analysis: The probability that $z^n(1, u^n) = z^n(i, u^n)$, where $i > 1$ and where $(u^n, z^n(1, u^n)) \in T_\epsilon^{(n)}(U, Z)$ is bounded by $2^{-n(H(Z|U)-\delta(\epsilon))}$, where $\delta(\epsilon)$ goes to zero as ϵ goes to zero. Hence, if

$$R'_1 < H(Z|U), \tag{174}$$

then the probability that an incorrect message $m'_{1,b}$ was decoded goes to zero for a large enough n. In order to find in super-bin $(\hat{m}_{0,b}, \hat{m}'_{1,b-1})$ and in bin $m'_{2,b}$ a codeword u^n that is jointly typical with s^n, we need to have more than $I(U; S)$ codewords in each bin; thus if

$$C_{21} - R'_2 \geq I(U; S), \tag{175}$$

$$R'_2 \leq C_{21} - I(U; S), \tag{176}$$

then the probability of finding a codeword u^n such that $(u^n, s^n) \in T_\epsilon^{(n)}(U, S)$ goes to 1 for a large enough n. We define the following event at block b:

$$E_{i,j,k,b} \triangleq (u^n(i, s^n), z^n(\hat{m}'_{1,b}, u^n), x_1^n(j, u^n, z^n), x_2^n(k, s^n), s^n, y^n) \in T_\epsilon^{(n)}(U, Z, X_1, X_2, S, Y). \tag{177}$$

We can bound the probability of error as follows:

$$P_{e,b}^{(n)} \leq \Pr(E_{1,1,1,b}^c) + \sum_{i=1,j=1,k>1} \Pr(E_{1,1,k,b}) + \sum_{i=1,j>1,k=1} \Pr(E_{1,j,1,b})$$

$$+ \sum_{i=1,j>1,k>1} \Pr(E_{1,j,k,b}) + \sum_{i>1,j>1,k>1} \Pr(E_{i,j,k,b}). \tag{178}$$

We now show that each term in (178) goes to zero for a large enough n.

- Upper-bounding $\Pr(E^c_{1,1,1,b})$: Since we assume that Encoders 1 and 2 encode the correct message-tuple $(m_{0,b}, m'_{1,b-1}, m''_{1,b}, m'_{2,b}, m''_{2,b})$ at block b and that the decoder decoded the right $(m_{0,b+1}, m'_{1,b}, m''_{1,b+1}, m'_{2,b+1}, m''_{2,b+1})$ at block $b+1$, by the LLN, $\Pr(E^c_{1,1,1,b}) \to 0$.

- Upper-bounding $\sum_{i=1,j=1,k>1} \Pr(E_{1,1,k,b})$: Assuming that $m'_{1,b}$ was decoded correctly at block $b+1$, the probability for this event is bounded by

$$\sum_{i=1,j=1,k>1} \Pr(E_{1,1,k,b}) \leq 2^{nR''_2} 2^{-n(I(X_2;Y|S,U,Z,X_1)-\delta(\epsilon)} \tag{179}$$

$$= 2^{nR''_2} 2^{-n(I(X_2;Y|S,U,X_1)-\delta(\epsilon))}. \tag{180}$$

- Upper-bounding $\sum_{i=1,j>1,k=1} \Pr(E_{1,j,1,b})$: Assuming that $m'_{1,b}$ was decoded correctly at block $b+1$, the probability for this event is bounded by

$$\sum_{i=1,j>1,k=1} \Pr(E_{1,j,1,b}) \leq 2^{n(R''_1)} 2^{-n(I(X_1;Y|S,U,Z,X_2)-\delta(\epsilon))}. \tag{181}$$

- Upper-bounding $\sum_{i=1,j>1,k>1} \Pr(E_{1,j,k,b})$: Assuming that $m'_{1,b}$ was decoded correctly at block $b+1$, the probability for this event is bounded by

$$\sum_{i=1,j>1,k>1} \Pr(E_{1,j,k,b}) \leq 2^{n(R''_1+R''_2)} 2^{-n(I(X_1,X_2;Y|S,U,Z)-\delta(\epsilon))}. \tag{182}$$

- Upper-bounding $\sum_{i>1,j>1,k>1} \Pr(E_{i,j,k,b})$: Assuming that $m'_{1,b}$ was decoded correctly at block $b+1$, the probability for this event is bounded by

$$\sum_{i>1,j>1,k>1} \Pr(E_{1,j,k,b}) \leq 2^{n(R_0+R'_1+R''_1+R'_2+R''_2)} 2^{-n(I(U,V,Z,X_1,X_2;Y|S)-\delta(\epsilon))} \tag{183}$$

$$\leq 2^{n(R_0+R'_1+R''_1+R'_2+R''_2)} 2^{-n(I(X_1,X_2;Y|S)-\delta(\epsilon))}. \tag{184}$$

To summarize, we note that $R'_1 = R_1 - R''_1$ and $R'_2 = R_2 - R''_2$ and thus we obtained that if (R''_1, R''_2, R_1, R_2) satisfy

$$R_1 - R''_1 \leq H(Z|U), \tag{185}$$

$$R_2 - R''_2 \leq C_{21} - I(U;S), \tag{186}$$

$$R''_2 \leq I(X_2;Y|S,U,X_1), \tag{187}$$

$$R''_1 \leq I(X_1;Y|S,U,Z,X_2), \tag{188}$$

$$R''_1 + R''_2 \leq I(X_1,X_2;Y|S,U,Z), \tag{189}$$

$$R_0 + R''_1 + R''_2 \leq I(X_1,X_2;Y|S), \tag{190}$$

then there exists a code with a probability of error that goes to zero as the block length goes to infinity. Using the Fourier-Motzkin elimination and by setting $R_1 = \tilde{R}_1, R_0 = \tilde{R}_0$, we obtain the following region

$$
\begin{aligned}
\tilde{R}_1 &\leq H(Z|V,U) + I(X_1;Y|S,U,X_2,Z), \\
R_2 &\leq I(X_2;Y|X_1,S,U) + C_{21} - I(U;S), \\
\tilde{R}_1 + R_2 &\leq I(X_1,X_2;Y|U,Z,S) + H(Z|U) + C_{21} - I(U;S), \\
\tilde{R}_0 + \tilde{R}_1 + R_2 &\leq I(X_1,X_2;Y|S).
\end{aligned}
\tag{191}
$$

Now we can easily see that if we set

$$
\tilde{R}_0 = C_{12}, \tag{192}
$$

$$
\tilde{R}_1 = R_1 - C_{12}, \tag{193}
$$

then the inequalities can be rewritten as

$$
\begin{aligned}
R_1 - C_{12} &\leq H(Z|U) + I(X_1;Y|S,U,X_2,Z), \\
R_2 &\leq I(X_2;Y|X_1,V,S,U) + C_{21} - I(U;S), \\
R_1 - C_{12} + R_2 &\leq I(X_1,X_2;Y|U,Z,S) + H(Z|U) + C_{21} - I(U;S), \\
C_{12} + (R_1 - C_{12}) + R_2 &\leq I(X_1,X_2;Y|S),
\end{aligned}
\tag{194}
$$

and thus we obtain the region in (97). ∎

Achievability for the causal case: The achievability part follows similarly to that of the strictly causal case, only now the generation of X_2^n is done i.i.d. according to the conditional distribution of $p(x_2|u,s,z)$ induced by (98). ∎

APPENDIX C

PROOF OF THEOREM 5

A. Converse

Converse for the strictly causal case: Given an achievable rate-pair (R_1, R_2), we need to show that there exists a joint distribution of the form $P(w)P(v|w)p(a|w)P(s|a)P(x_1|v,w)P(u,x_2|s,v,a,w)P(y|x_1,x_2,s)$ such that the inequalities in (107) are satisfied. Since (R_1, R_2) is an achievable rate-pair, there exists a $(2^{nR_1}, 2^{nR_2}, 2^{nC_{12}}, n)$ code with an arbitrarily small error probability $P_e^{(n)}$. By Fano's inequality,

$$
H(M_1, M_2|Y^n) \leq n(R_1 + R_2)P_e^{(n)} + H(P_e^{(n)}). \tag{195}
$$

We set

$$
(R_1 + R_2)P_e^{(n)} + \frac{1}{n}H(P_e^{(n)}) \triangleq \epsilon_n, \tag{196}
$$

where $\epsilon_n \to 0$ as $P_e^{(n)} \to 0$. Hence,

$$
H(M_1|Y^n, M_2) \leq H(M_1, M_2|Y^n) \leq n\epsilon_n, \tag{197}
$$

$$H(M_2|Y^n, M_1) \leq H(M_1, M_2|Y^n) \leq n\epsilon_n. \tag{198}$$

For R_1 we have the following:

$$
\begin{align}
nR_1 &= H(M_1) \tag{199} \\
&= H(M_1, M_{12}) \tag{200} \\
&\overset{(a)}{=} H(M_1|M_2, M_{12}) + H(M_{12}) \tag{201} \\
&\leq nC_{12} + I(M_1; Y^n|M_2, M_{12}) + H(M_1|Y^n, M_2, M_{12}) \tag{202} \\
&\overset{(b)}{\leq} nC_{12} + I(M_1; Y^n|M_2, M_{12}) + n\epsilon_n \tag{203} \\
&\overset{(c)}{=} nC_{12} + I(X_1^n; Y^n|M_2, M_{12}) + n\epsilon_n \tag{204} \\
&\overset{(d)}{=} nC_{12} + \sum_{i=1}^n I(X_{1,i}; Y^n|M_2, X_1^{i-1}, M_{12}) + n\epsilon_n \tag{205} \\
&\leq nC_{12} + \sum_{i=1}^n H(X_{1,i}|M_2, X_1^{i-1}, M_{12}) + n\epsilon_n \tag{206} \\
&\overset{(e)}{\leq} nC_{12} + \sum_{i=1}^n H(X_{1,i}|V_i, W_i) + n\epsilon_n, \tag{207}
\end{align}
$$

where (a) follows from the fact that the messages M_1 and M_2 are independent, (b) follows from Fano's inequality, (c) follows from the encoding relation in (101), (d) follows from the chain rule, and step (e) follows since conditioning reduces entropy and by setting the RVs

$$
\begin{align}
V_i &\triangleq X_1^{i-1}, \tag{208} \\
W_i &\triangleq M_{12}. \tag{209}
\end{align}
$$

Thus, we obtained

$$R_1 \leq C_{12} + \frac{1}{n}\sum_{i=1}^n H(X_{1,i}|V_i, W_i) + \epsilon_n. \tag{210}$$

Additionally,

$$
\begin{align}
nR_1 &= H(M_1) \tag{211} \\
&= H(M_1|M_2, M_{12}) + H(M_{12}) \tag{212} \\
&\overset{(a)}{\leq} nC_{12} + I(M_1; Y^n|M_2, M_{12}) + n\epsilon_n \tag{213} \\
&\overset{(b)}{=} nC_{12} + \sum_{i=1}^n I(M_1; Y_i|Y^{i-1}, M_2, M_{12}) + n\epsilon_n \tag{214} \\
&\leq nC_{12} + \sum_{i=1}^n I(Y^{i-1}, M_1, M_2; Y_i|M_{12}) + n\epsilon_n \tag{215} \\
&= nC_{12} + \sum_{i=1}^n [I(Y^{i-1}, M_1, M_2, S_{i+1}^n; Y_i|M_{12}) \\
&\qquad - I(S_{i+1}^n; Y_i|M_1, M_2, Y^{i-1}, M_{12})] + n\epsilon_n \tag{216}
\end{align}
$$

$$\overset{(c)}{=} nC_{12} + \sum_{i=1}^{n}[I(Y^{i-1}, M_1, M_2, S_{i+1}^n, X_1^{i-1}, X_{1,i}; Y_i | M_{12})$$
$$-I(S_i; Y^{i-1} | M_1, M_2, S_{i+1}^n, M_{12})] + n\epsilon_n \tag{217}$$

$$\overset{(d)}{=} nC_{12} + \sum_{i=1}^{n}[I(Y^{i-1}, M_2, S_{i+1}^n, X_1^{i-1}, X_{1,i}; Y_i | A_i, M_{12})$$
$$-I(S_i; Y^{i-1} | M_1, M_2, A_i, S_{i+1}^n, M_{12})] + n\epsilon_n \tag{218}$$

$$\overset{(e)}{\leq} nC_{12} + \sum_{i=1}^{n}[I(Y^{i-1}, M_2, S_{i+1}^n, X_1^{i-1}, X_{1,i}; Y_i | A_i, M_{12})$$
$$-I(S_i; Y^{i-1}, M_2, S_{i+1}^n | M_1, A_i, M_{12})] + n\epsilon_n \tag{219}$$

$$= nC_{12} + \sum_{i=1}^{n}[I(Y^{i-1}, M_2, S_{i+1}^n, X_1^{i-1}, X_{1,i}; Y_i | A_i, M_{12})$$
$$-I(S_i; Y^{i-1}, M_2, S_{i+1}^n | M_1, A_i, X_1^{i-1}, M_{12})] + n\epsilon_n \tag{220}$$

$$\overset{(f)}{=} nC_{12} + \sum_{i=1}^{n}[I(Y^{i-1}, M_2, S_{i+1}^n, X_1^{i-1}, X_{1,i}; Y_i | A_i, M_{12})$$
$$-I(S_i; Y^{i-1}, M_2, S_{i+1}^n | A_i, X_1^{i-1}, M_{12})] + n\epsilon_n \tag{221}$$

$$\overset{(g)}{=} nC_{12} + \sum_{i=1}^{n}[I(V_i, U_i, X_{1,i}; Y_i | A_i, W_i) - I(S_i; U_i | V_i, A_i, W_i)] + n\epsilon_n, \tag{222}$$

where (a) follows from Fano's inequality, (b) follows from the chain rule, (c) follows since $X_1^i = f(M_1)$ and by using the Csiszar Sum Equality, (d) follows since $A_i = f(M_{12}, M_2)$ and from the Markov Chain $M_1 - (M_{12}, X_{1,i}, X_1^{i-1}, Y^{i-1}, M_2, S_{i+1}^n, A_i, M_{12}) - Y_i$, (e) follows since S_i is independent of (M_2, S_{i+1}^n) given (M_1, A_i), (f) follows from the Markov Chain $M_1 - (M_{12}, X_1^{i-1}, A_i) - (Y^{i-1}, M_2, S_{i+1}^n)$, and (g) follows by setting the RVs W, V and

$$U_i \triangleq (Y^{i-1}, M_2, S_{i+1}^n). \tag{223}$$

Thus, we obtained

$$R_1 \leq C_{12} + \frac{1}{n}\sum_{i=1}^{n}[I(V_i, U_i, X_{1,i}; Y_i | A_i, W_i) - I(S_i; U_i | V_i, A_i, W_i)] + \epsilon_n. \tag{224}$$

Next, we consider R_2

$$nR_2 = H(M_2) \tag{225}$$

$$= H(M_2 | M_1) \tag{226}$$

$$\overset{(a)}{\leq} I(M_2; Y^n | M_1) + n\epsilon_n \tag{227}$$

$$= \sum_{i=1}^{n} I(M_2; Y_i | Y^{i-1}, M_1) + n\epsilon_n \tag{228}$$

$$\leq \sum_{i=1}^{n} I(Y^{i-1}, M_2; Y_i | M_1) + n\epsilon_n \tag{229}$$

$$\overset{(b)}{=} \sum_{i=1}^{n}[I(Y^{i-1}, M_2, S_{i+1}^n; Y_i | M_1) - I(S_{i+1}^n; Y_i | M_1, M_2, Y^{i-1})] + n\epsilon_n \tag{230}$$

$$\overset{(c)}{=} \sum_{i=1}^{n}[I(Y^{i-1}, M_2, S_{i+1}^n; Y_i|M_1, M_{12}, X_{1,i}, X_1^{i-1})$$

$$-I(S_i; Y^{i-1}, M_2, S_{i+1}^n|A_i, M_{12}, X_1^{i-1})] + n\epsilon_n \tag{231}$$

$$\overset{(d)}{=} \sum_{i=1}^{n}[I(Y^{i-1}, M_2, S_{i+1}^n, A_i; Y_i|M_1, M_{12}, X_{1,i}, X_1^{i-1})$$

$$-I(S_i; Y^{i-1}, M_2, S_{i+1}^n|A_i, M_{12}, X_1^{i-1})] + n\epsilon_n \tag{232}$$

$$\overset{(e)}{\leq} \sum_{i=1}^{n}[I(Y^{i-1}, M_2, S_{i+1}^n, A_i; Y_i|M_{12}, X_{1,i}, X_1^{i-1})$$

$$-I(S_i; Y^{i-1}, M_2, S_{i+1}^n|A_i, M_{12}, X_1^{i-1})] + n\epsilon_n \tag{233}$$

$$\overset{(f)}{=} \sum_{i=1}^{n}[I(U_i, A_i; Y_i|W_i, X_{1,i}, V_i) - I(S_i; U_i|W_i, V_i, A_i)] + n\epsilon_n, \tag{234}$$

where (a) follows from Fano's inequality, (b) follows from the chain rule, (c) follows since $(M_{12}, X_1^i) = f(M_1)$ and from the same arguments as given in (217) - (222), (d) follows since $A_i = f(M_{12}, M_2)$, (e) follows from the same arguments as given in (218), and (f) follows by setting the RVs U, V and W. Thus, we obtained

$$R_2 \leq \frac{1}{n}\sum_{i=1}^{n}[I(U_i, A_i; Y_i|W_i, X_{1,i}, V_i) - I(S_i; U_i|W_i, V_i, A_i)] + \epsilon_n. \tag{235}$$

Now, consider

$$n(R_1 + R_2) = H(M_1, M_2) \tag{236}$$

$$= H(M_1, M_2|M_{12}) + H(M_{12}) \tag{237}$$

$$\overset{(a)}{\leq} nC_{12} + I(M_1, M_2; Y^n|M_{12}) + n\epsilon_n \tag{238}$$

$$\overset{(b)}{=} nC_{12} + \sum_{i=1}^{n} I(M_1, M_2; Y_i|Y^{i-1}, M_{12}) + n\epsilon_n \tag{239}$$

$$\overset{(c)}{\leq} nC_{12} + \sum_{i=1}^{n}[I(M_1, Y^{i-1}, M_2, S_{i+1}^n; Y_i|M_{12})$$

$$-I(Y^{i-1}, M_2, S_{i+1}^n; S_i|M_{12}, A_i, X_1^{i-1})] + n\epsilon_n \tag{240}$$

$$\overset{(d)}{=} nC_{12} + \sum_{i=1}^{n}[I(M_1, X_{1,i}, X_1^{i-1}, Y^{i-1}, M_2, S_{i+1}^n; Y_i|M_{12})$$

$$-I(Y^{i-1}, M_2, S_{i+1}^n; S_i|M_{12}, A_i, X_1^{i-1})] + n\epsilon_n \tag{241}$$

$$\overset{(e)}{=} nC_{12} + \sum_{i=1}^{n}[I(X_{1,i}, X_1^{i-1}, Y^{i-1}, M_2, S_{i+1}^n, A_i; Y_i|M_{12})$$

$$-I(Y^{i-1}, M_2, S_{i+1}^n; S_i|M_{12}, A_i, X_1^{i-1})] + n\epsilon_n \tag{242}$$

$$= nC_{12} + \sum_{i=1}^{n}[I(U_i, V_i, X_{1,i}, A_i; Y_i|W_i) - I(U_i; S_i|V_i, A_i, W_i)] + n\epsilon_n, \tag{243}$$

where (a) follows from Fano's inequality, (b) follows from the chain rule, (c) follows from the same arguments as given in (217)-(222), (d) follows since $X_1^i = f(M_1)$ and $A_i = f(M_{12}, M_2)$, and (e) follows from the same

arguments as given in (218). Thus we obtained

$$R_1 + R_2 \quad \leq \quad C_{12} + \frac{1}{n} \sum_{i=1}^{n} [I(U_i, V_i, X_{1,i}, A_i; Y_i | W_i) - I(U_i; S_i | V_i, A_i | W_i)] + \epsilon_n. \tag{244}$$

Again,

$$n(R_1 + R_2) \quad = \quad H(M_1, M_2) \tag{245}$$

$$\overset{(a)}{\leq} \quad I(M_1, M_2; Y^n) + n\epsilon_n \tag{246}$$

$$\overset{(b)}{=} \quad \sum_{i=1}^{n} I(M_1, M_2; Y_i | Y^{i-1}) + n\epsilon_n \tag{247}$$

$$\overset{(c)}{\leq} \quad \sum_{i=1}^{n} [I(M_1, Y^{i-1}, M_2, S_{i+1}^n; Y_i)$$
$$\qquad - I(Y^{i-1}, M_2, S_{i+1}^n; S_i | M_{12}, A_i, X_1^{i-1})] + n\epsilon_n \tag{248}$$

$$\overset{(d)}{=} \quad \sum_{i=1}^{n} [I(M_1, M_{12}, X_{1,i}, X_1^{i-1}, Y^{i-1}, M_2, S_{i+1}^n; Y_i)$$
$$\qquad - I(Y^{i-1}, M_2, S_{i+1}^n; S_i | M_{12}, A_i, X_1^{i-1})] + n\epsilon_n \tag{249}$$

$$\overset{(e)}{=} \quad \sum_{i=1}^{n} [I(M_{12}, X_{1,i}, X_1^{i-1}, Y^{i-1}, M_2, S_{i+1}^n, A_i; Y_i)$$
$$\qquad - I(Y^{i-1}, M_2, S_{i+1}^n; S_i | M_{12}, A_i, X_1^{i-1})] + n\epsilon_n \tag{250}$$

$$\leq \quad \sum_{i=1}^{n} [I(W_i, U_i, V_i, X_{1,i}, A_i; Y_i) - I(U_i; S_i | W_i, V_i, A_i)] + n\epsilon_n, \tag{251}$$

where (a) follows from Fano's inequality, (b) follows from the chain rule, (c) follows from the same arguments as given in (217)-(222), (d) follows since $X_1^i = f(M_1)$ and $A_i = f(M_{12}, M_2)$, and (e) follows from the same arguments as given in (218). Thus, we obtained

$$R_1 + R_2 \quad \leq \quad \frac{1}{n} \sum_{i=1}^{n} [I(W_i, U_i, V_i, X_{1,i}, A_i; Y_i) - I(U_i; S_i | W_i, V_i, A_i)] + \epsilon_n. \tag{252}$$

Finally, we need to prove the following Markov chains:

- $A_i - W_i - V_i$ -

$$p(a_i | m_{12}, x_1^{i-1}) \quad = \quad \sum_{m_2 \in \mathcal{M}_2} p(m_2 | m_{12}, x_1^{i-1}) p(a_i | m_{12}, m_2, x_1^{i-1}) \tag{253}$$

$$\overset{(a)}{=} \quad \sum_{m_2 \in \mathcal{M}_2} p(m_2 | m_{12}) p(a_i | m_{12}, m_2) \tag{254}$$

$$= \quad p(a_i | m_{12}), \tag{255}$$

where (a) follows since m_2 is independent of m_1 and since $a_i = f(m_2, m_{12})$.

- $S_i - A_i - (W_i, V_i)$ - Follows from the fact that the channel state at any time i is assumed to depend only on the action at time i.

- $X_{1,i} - (V_i, W_i) - (A_i, S_i)$ -

$$p(x_{1,i} | m_{12}, x_1^{i-1}, a_i, s_i) \quad = \quad \sum_{m_1 \in \mathcal{M}_1} p(m_1 | m_{12}, x_1^{i-1}, a_i, s_i) p(x_{1,i} | m_{12}, m_1, x_1^{i-1}, a_i, s_i) \tag{256}$$

$$\overset{(a)}{=} \sum_{m_1 \in \mathcal{M}_1} p(m_1|m_{12}, x_1^{i-1}) p(x_{1,i}|m_1, m_{12}, x_1^{i-1}) \tag{257}$$

$$= p(x_{1,i}|m_{12}, x_1^{i-1}), \tag{258}$$

where (a) follows since m_1 is independent of (a_i, s_i) given (m_{12}, x_1^{i-1}) and since $x_{1,i} = f(m_1)$.

- $(U_i, X_{2,i}) - S_i, A_i, W_i, V_i - X_{1,i}$ -

$$p(x_{1,i}|m_{12}, x_1^{i-1}, a_i, s_i^n, y^{i-1}, m_2, x_{2,i}) = \sum_{m_1 \in \mathcal{M}_1} p(m_1|m_{12}, x_1^{i-1}, a_i, s_i^n, y^{i-1}, m_2, x_{2,i})$$
$$p(x_{1,i}|m_1, m_{12}, x_1^{i-1}, a_i, s_i^n, y^{i-1}, m_2, x_{2,i}) \tag{259}$$

$$\overset{(a)}{=} \sum_{m_1 \in \mathcal{M}_1} p(m_1|m_{12}, x_1^{i-1}, a_i, s_i)$$
$$p(x_{1,i}|m_1, m_{12}, x_1^{i-1}, a_i, s_i) \tag{260}$$

$$= p(x_{1,i}|m_{12}, x_1^{i-1}, a_i, s_i), \tag{261}$$

where (a) follows since m_1 is independent of $(s_{i+1}^n, y^{i-1}, m_2, x_{2,i})$ given $(m_{12}, x_1^{i-1}, a_i, s_i)$ and since $x_{1,i} = f(m_1)$.

- $Y_i - (X_{1,i}, X_{2,i}, S_i) - (W_i, V_i, U_i, A_i)$ - Follows from the fact that the channel output at any time i is assumed to depend only on the channel inputs and state at time i.

Finally, let Q be an RV independent of (X_1^n, X_2^n, Y^n) and uniformly distributed over the set $\{1, 2, 3, \ldots, n\}$. We define the RV $W \triangleq (Q, W_Q)$ and obtain the region given in (107). ∎

Converse for the causal case: For the causal case we repeat the same approach as for the strictly causal case, except that in the final step we need to show the Markov chain $U_i - (S_i, A_i, W_i, V_i) - X_{1,i}$. We can see from the following derivations that this Markov chain holds

$$p(x_{1,i}|m_{12}, x_1^{i-1}, a_i, s_i^n, y^{i-1}, m_2) = \sum_{m_1 \in \mathcal{M}_1} p(m_1|m_{12}, x_1^{i-1}, a_i, s_i^n, y^{i-1}, m_2)$$
$$p(x_{1,i}|m_1, m_{12}, x_1^{i-1}, a_i, s_i^n, y^{i-1}, m_2) \tag{262}$$

$$\overset{(a)}{=} \sum_{m_1 \in \mathcal{M}_1} p(m_1|m_{12}, x_1^{i-1}, a_i, s_i)$$
$$p(x_{1,i}|m_1, m_{12}, x_1^{i-1}, a_i, s_i) \tag{263}$$

$$= p(x_{1,i}|m_{12}, x_1^{i-1}, a_i, s_i), \tag{264}$$

where (a) follows since m_1 is independent of $(s_{i+1}^n, y^{i-1}, m_2)$ given $(m_{12}, x_1^{i-1}, a_i, s_i)$ and since $x_{1,i} = f(m_1)$. ∎

B. Achievability

Achievability for the strictly causal case: Fix a joint distribution $P(w)P(v|w)P(a|w)P(s|a)P(x_1|v, w)P(u|s, w, v, a)$ $p(x_2|w, a, v, u, s)P(y|x_1, x_2, s)$ where $P(s|a)$ and $P(y|x_1, x_2, s)$ are given by the channel. In the following achievability scheme we use block Markov coding, rate splitting, and Gelfand-Pinsker coding.

Coding Scheme: We consider B blocks, each consisting of n symbols; thus we transmit nB symbols. We transmit $B - 1$ messages M_1 in B blocks of information. Here, $M_1 \in \{1, \dots, 2^{nR_1}\}$; thus, asymptotically, for a large enough n, our transmission rate would be $\frac{nR_1(B-1)}{nB} \overset{n \to \infty}{\longrightarrow} R_1$. We also split message M_1 into (M_1', M_1'') such that $(R_1', R_1'') = (C_{12}, R_1 - C_{12})$.

Code Design: Generate $2^{nR_1'}$ codewords w^n i.i.d. using $P(w^n) = \Pi_{i=1}^n P(w_i)$. For each w^n, generate $2^{nR_1''}$ codewords v^n i.i.d. using $P(v^n|w^n) = \Pi_{i=1}^n P(v_i|w_i)$. For each w^n, generate 2^{nR_2} codewords a^n i.i.d. using $P(a^n|w^n) = \Pi_{i=1}^n P(a_i|w_i)$. For each pair (w^n, v^n), generate $2^{nR_1''}$ codewords x_1^n i.i.d. using $P(x_1^n|v^n, w^n) = \Pi_{i=1}^n P(x_{1,i}|v_i, w_i)$. Additionally, for each triplet (w^n, v^n, a^n), generate $2^{n(R_2 + \tilde{R})}$ codewords u^n i.i.d. using $P(u^n|a^n, v^n, w^n) = \Pi_{i=1}^n P(u_i|a_i, v_i, w_i)$. Randomly bin all u^n codewords into 2^{nR_2} bins where each bin contains $2^{n\tilde{R}}$ codewords.

Encoding: We denote the realizations of the messages (M_1', M_1'', M_2) at block b as $(m_{1,b}', m_{1,b}'', m_{2,b})$. Since we use block Markov coding, we set $m_{1,B}' = 1$. In block $b \in \{1, \dots, B\}$, send $m_{1,b}'$ from Encoder 1 to Encoder 2 via the rate-limited cooperation link. Encode message $m_{1,b}'$ using $w^n(m_{1,b}')$. Encode message $m_{1,b-1}''$ conditioned on $m_{1,b}'$ using $v^n(m_{1,b-1}'', w^n)$ and encode message $m_{1,b}''$ conditioned on $(m_{1,b-1}'', m_{1,b}')$ using $x_1^n(m_{1,b}'', v^n, w^n)$. Given $(m_{1,b}', m_{2,b})$, Encoder 2 chooses an action sequence a^n. Given (s^n, w^n, v^n, a^n), look in bin $m_{2,b}$ for a codeword $u^n(w^n, v^n, a^n, m_{2,b}, l)$ that is jointly typical with $(w^n(m_{1,b}'), v^n(m_{1,b-1}''), s^n, a^n(m_{2,b}))$, where $l \in \{1, \dots, 2^{n\tilde{R}}\}$. Send $x_1^n(m_{1,b}'', w^n, v^n)$ and x_2^n according to $p(x_2|w, v, u, s)$ i.i.d. over the channel.

Decoding at Encoder 2: At the end of block b, Encoder 2 tries to decode message $m_{1,b}''$. Given $m_{1,b}'$ and assuming that message $m_{1,b-1}''$ was decoded correctly at the end of block $b - 1$, Encoder 2 looks for $\hat{m}_{1,b}''$ s.t.

$$(w^n(m_{1,b}'), v^n(m_{1,b-1}'', w^n), x_1^n(\hat{m}_{1,b}'', w^n, v^n)) \in T_\epsilon^{(n)}(W, V, X_1). \tag{265}$$

If no such $\hat{m}_{1,b}'$, or more than one such $\hat{m}_{1,b}'$, was found, an error is declared at block b and therefore in the whole super-block nB.

Decoding at the receiver: At the end of block B, the decoding is done backwards. At block b, assuming that $m_{1,b}$ was decoded correctly in block $b + 1$, the decoder looks for the triplet $(m_{1,b}', m_{1,b-1}'', \hat{m}_{2,b})$ s.t.

$$(w^n(\hat{m}_{1,b}'), v^n(\hat{m}_{1,b-1}'', w^n), x_1^n(m_{1,b}'', w^n, v^n), a^n(\hat{m}_{2,b}, w^n), u^n(\hat{m}_{2,b}, w^n, v^n, s^n, a^n, l), y^n) \in T_\epsilon^{(n)}(W, V, X_1, A, U, Y). \tag{266}$$

If no such pair, or more than one such pair, was found, an error is declared at block b and therefore in the whole super-block nB.

Error Analysis: Without loss of generality, we assume that $(m_{1,b}', m_{1,b-1}'', m_{2,b}) = (1, 1, 1)$. The probability that $x_1^n(1, w^n, v^n) = x_1^n(i, w^n, v^n)$ where $i > 1$ and where $(w^n(1), v^n(1, w^n), x_1^n(1, w^n, v^n)) \in T_\epsilon^{(n)}(W, V, X_1)$ is bounded by $2^{-n(H(X_1|V,W) - \delta(\epsilon))}$, where $\delta(\epsilon)$ goes to zero as ϵ goes to zero. Hence, if

$$R_1 - C_{12} < H(X_1|V, W), \tag{267}$$

then the probability that an incorrect message $m_{1,b}$ was decoded goes to zero for a large enough n. We define the

following event at block b:

$$E_{i,j,k,l,b} \triangleq (w^n(i), v^n(j, w^n), x_1^n(\hat{m}_{1,b}'', v^n, w^n), a^n(k, w^n), u^n(k, v^n, s^n, a^n, w^n, l), y^n) \in T_\epsilon^{(n)}(W, V, X_1, A, U, Y).$$

(268)

We can bound the probability of error as follows:

$$
\begin{aligned}
P_{e,b}^{(n)} \quad \leq \quad & \Pr(E_{1,1,1,1,b}^c) + \sum_{\substack{i=1,j=1 \\ k>1,l>1}} \Pr(E_{1,1,k,l,b}) + \sum_{\substack{i=1,j>1 \\ k=1,l>1}} \Pr(E_{1,j,1,l,b}) \\
& + \sum_{\substack{i=1,j>1 \\ k>1,l>1}} \Pr(E_{1,j,k,l,b}) + \sum_{\substack{i>1,j>1 \\ k>1,l>1}} \Pr(E_{i,j,k,l,b}).
\end{aligned}
$$

(269)

We now show that each term in (269) goes to zero for a large enough n.

- Upper-bounding $\Pr(E_{1,1,1,1,b}^c)$: Since we assume that Transmitters 1 and 2 encode the correct message triplet $(m_{1,b}', m_{1,b-1}'', m_{2,b})$ at block b and that the receiver decoded the right $(m_{1,b+1}', m_{1,b}'', m_{2,b+1})$ at block $b+1$, by the LLN, $\Pr(E_{1,1,1,1,b}^c) \to 0$.

- Upper-bounding $\sum_{\substack{i=1,j=1 \\ k>1,l>1}} \Pr(E_{1,1,k,l,b})$: Assuming that $m_{1,b}''$ was decoded correctly at block $b+1$, the probability for this event is bounded by

$$\sum_{\substack{i=1,j=1 \\ k>1,l>1}} \Pr(E_{1,1,k,l,b}) \leq 2^{n(R_2+\tilde{R})} 2^{-n(I(U,A;Y|W,V,X_1)-\delta(\epsilon))}.$$

(270)

- Upper-bounding $\sum_{\substack{i=1,j>1 \\ k=1,l>1}} \Pr(E_{1,j,1,l,b})$: Similarly to (270) we obtain

$$\sum_{\substack{i=1,j>1 \\ k=1,l>1}} \Pr(E_{1,j,1,l,b}) \leq 2^{n(R_1-C_{12}+\tilde{R})} 2^{-n(I(V,X_1,U;Y|W,A)-\delta(\epsilon))}.$$

(271)

- Upper-bounding $\sum_{\substack{i=1,j>1 \\ k>1,l>1}} \Pr(E_{1,j,k,l,b})$: Similarly to (270) we obtain

$$\sum_{\substack{i=1,j>1 \\ k>1,l>1}} \Pr(E_{1,j,k,l,b}) \leq 2^{n(R_1-C_{12}+R_2+\tilde{R})} 2^{-n(I(U,A,V,X_1;Y|W)-\delta(\epsilon))}.$$

(272)

- Upper-bounding $\sum_{\substack{i>1,j>1 \\ k>1,l>1}} \Pr(E_{1,j,k,l,b})$: Similarly to (270) we obtain

$$\sum_{\substack{i>1,j>1 \\ k>1,l>1}} \Pr(E_{1,j,k,l,b}) \leq 2^{n(R_1+R_2+\tilde{R})} 2^{-n(I(U,A,W,V,X_1;Y)-\delta(\epsilon))}.$$

(273)

Finally, we analyze the probability of error for finding u^n at Encoder 2. By the covering lemma, if

$$\tilde{R} > I(U; S|W, V, A)$$

(274)

then with high probability, in block b we can find a codeword u^n that is jointly typical with s^n in bin number $m_{2,b}$. The combination of (267), (270), (271), (272), (273), and (274) yields the capacity region in (107), thus completing the proof. ∎

Achievability for the causal case: The achievability part follows similarly to that of the strictly causal case, only now the generation of X_2^n is done i.i.d. according to the conditional distribution of $p(x_2|w, v, u, s, x_1)$ induced by (108). ∎

REFERENCES

[1] F. M. J. Willems, "Information-theoretical results for the discrete memoryless multiple access channel," Ph.D. dissertation, Katholieke Universiteit Leuven, Haverlee, Belgium, 1982.

[2] ——, "The discrete memoryless multiple access channel with partially cooperating encoders," *IEEE Trans. Inf. Theory*, vol. 29, no. 3, pp. 441–445, 1983.

[3] F. M. J. Willems and E. Van der Meulen, "The discrete memoryless multiple-access channel with cribbing encoders," *IEEE Trans. Inf. Theory*, vol. 31, no. 3, pp. 313–327, 1985.

[4] H. Asnani and H. H. Permuter, "Multiple-access channel with partial and controlled cribbing encoders," *IEEE Trans. Inf. Theory*, vol. 59, no. 4, pp. 2252–2266, 2013.

[5] O. Simeone, N. Levy, A. Sanderovich, O. Somekh, B. M. Zaidel, H. V. Poor, S. Shamai *et al.*, "Cooperative wireless cellular systems: An information-theoretic view," *Foundations and Trends in Communications and Information Theory*, vol. 8, no. 1-2, pp. 1–177, 2012.

[6] X. Lin, J. G. Andrews, A. Ghosh, and R. Ratasuk, "An overview on 3GPP device-to-device proximity services," *Submitted to IEEE Communications Magazine available at arXiv preprint arXiv:1310.0116*, 2013.

[7] Y. Cao and B. Chen, "The cognitive radio channel: From spectrum sensing to message cribbing," *Ad Hoc Networks*, pp. 868–883, 2010.

[8] J. Shimonovich, A. S. Baruch, and S. S. Shitz, "Cognitive cooperative communications on the multiple access channel," in *Proc. IEEE Inf. Theory Workshop*. IEEE, 2013, pp. 1–5.

[9] D. Slepian and J. K. Wolf, "A coding theorem for multiple access channels with correlated sources," *Bell Syst. Tech. J.*, vol. 52, no. 7, pp. 1037–1076, 1973.

[10] A. Bracher and A. Lapidoth, "Feedback, cribbing, and causal state-information on the multiple-access channel," *Submitted to IEEE Trans. Inf. Theory*, 2013.

[11] S. I. Bross, A. Lapidoth, and M. A. Wigger, "The gaussian mac with conferencing encoders," in *Proc. IEEE Int. Symp. Inf. Theory*. IEEE, 2008, pp. 2702–2706.

[12] H. Asnani, H. H. Permuter, and T. Weissman, "Successive refinement with decoder cooperation and its channel coding duals," *IEEE Trans. Inf. Theory*, vol. 59, no. 9, pp. 5511–5533, 2013.

[13] Y. Steinberg, "Cooperative coding for channels with uncertainty in the cooperation links," in *Proc. Inf. Theory and App. Workshop*, UCSD, San Diego, CA, USA, 2014.

[14] T. Weissman, "Capacity of channels with action-dependent states," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5396–5411, 2010.

[15] R. Ahlswede, "The capacity region of a channel with two senders and two receivers," *Ann. Probab.*, vol. 2, pp. 805–814, 1974.

[16] H. H. Permuter, Y. Steinberg, and T. Weissman, "Two-way source coding with a helper," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2905–2919, 2010.

[17] C. E. Shannon, "Coding theorems for a discrete source with a fidelity criterion," *IRE Nat. Conv. Rec*, vol. 4, no. 142–163, p. 1, 1959.

[18] A. D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inf. Theory*, vol. 22, no. 1, pp. 1–10, 1976.

[19] S. I. Gel'fand and M. S. Pinsker, "Coding for channel with random parameters," *Probl. Contr. and Inf. Theory*, vol. 9, no. 1, pp. 19–31, 1980.

[20] T. M. Cover and M. Chiang, "Duality between channel capacity and rate distortion with two-sided state information," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1629–1638, 2002.

[21] S. S. Pradhan, J. Chou, and K. Ramchandran, "Duality between source coding and channel coding and its extension to the side information case," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1181–1203, 2003.

[22] A. Gupta and S. Verdú, "Operational duality between lossy compression and channel coding," *IEEE Trans. Inf. Theory*, vol. 57, no. 6, pp. 3171–3179, 2011.

[23] H. H. Permuter, S. Shamai, and A. Somekh-Baruch, "Message and state cooperation in multiple access channels," *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6379–6396, 2011.

[24] L. Dikstein, H. H. Permuter, and S. Shamai, "MAC with action-dependent state information at one encoder," *Submitted to IEEE Trans. Inf. Theory available at arXiv preprint arXiv:1212.4626*, 2012.

[25] S. I. Bross and A. Lapidoth, "The state-dependent multiple-access channel with states available at a cribbing encoder," in *IEEE 26th Convention of Electrical and Electronics Engineers in Israel (IEEEI). Longer version available at arXiv:1007.3896v1.* IEEE, 2010, pp. 665–669.

[26] A. Somekh-Baruch, S. Shamai, and S. Verdú, "Cooperative multiple-access encoding with states available at one transmitter," *IEEE Trans. Inf. Theory*, vol. 54, no. 10, pp. 4448–4469, 2008.

[27] T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley & Sons, 2012.

[28] Z. Goldfeld, H. H. Permuter, and B. M. Zaidel, "The finite state MAC with cooperative encoders and delayed CSI," *Submitted to IEEE Trans. Inf. Theory available at arXiv preprint arXiv:1303.7083*, 2013.