Leveraging Strictly Causal State Information at the Encoders for Multiple Access Channels

Min Li[1], Osvaldo Simeone[2] and Aylin Yener[1]

[1]Dept. of Electrical Engineering, The Pennsylvania State University, University Park, PA 16802
[2]Dept. of Electrical and Computer Engineering, New Jersey Institute of Technology, University Heights, NJ 07102
mxl971@psu.edu, osvaldo.simeone@njit.edu, yener@ee.psu.edu

Abstract—The state-dependent multiple access channel (MAC) is considered where the state sequences are known strictly causally to the encoders. First, a two-user MAC with two independent states each known strictly causally to one encoder is revisited, and a new achievable scheme inspired by the recently proposed noisy network coding is presented. This scheme is shown to achieve a rate region that is potentially larger than that provided by recent work for the same model. Next, capacity results are presented for a class of channels that include modulo-additive state-dependent MACs. It is shown that the proposed scheme can be easily extended to an arbitrary number of users. Finally, a similar scheme is proposed for a MAC with common state known strictly causally to all encoders. The corresponding achievable rate region is shown to reduce to the one given in the previous work as a special case for two users.

I. INTRODUCTION

State-dependent channels model relevant phenomena in wireless communication links, such as fading and interference. It is usually assumed that there exists a state sequence s^n, with each component s_i denoting the state value affecting the channel at time instant i. Understanding the merit of state information (SI), i.e., information about the sequence s^n, for reliable communication is a key problem of both theoretical and practical interest. State-dependent channels are mainly classified into three groups with respect to the availability of SI at the encoders, including non-causal, causal and strictly causal SI [1]–[5]. Here we shall focus on the most realistic one, namely, strictly causal SI where at channel use i the encoders know the states up to the previous time instant. Furthermore, our focus is on models where the decoders have no SI as done in [1]–[4].

In a point-to-point discrete memoryless (DM) channel without decoder side SI, strictly causal SI at the encoder cannot increase the capacity for the channel with an independently and identically distributed (i.i.d.) state sequence s^n [6]. On the contrary, causal and non-causal SI turns out to be useful for the same channel and capacity-achieving strategies have been derived for the former [1] and the latter [2]. State-dependent multi-user channels have also been widely investigated, see [5] and references therein. In particular, on state-dependent multiple access channels (MAC), reference [7] derived a genie-aided bound to assess the capacity advantage of non-causal SI over causal SI for the case with independent state sequences available at the two encoders. Reference [8] explicitly characterized the capacity region of a cooperative MAC with SI non-causally available at one encoder, while reference [9] proposed several achievable rates and upper bounds for a MAC with non-causal SI at some encoders.

While the work mentioned above mostly focused on non-causal or causal SI, recently, references [3] and [4] have demonstrated that, even with strictly causal SI at the encoders of a state-dependent MAC channel, an improvement in the capacity region is possible. In [3], a common state sequence is known either strictly causally or causally at both of encoders in a two-user MAC, while in [4], two independent state sequences are present with each strictly causally or causally known to one encoder. An achievable rate region is derived in both papers and the capacity region is identified for some special cases including Gaussian models. The main idea in the achievability proofs is to use a block Markov coding scheme in which the two users cooperatively [3] or non-cooperatively [4] transmit compressed state information to the decoder, which in turn uses such information to perform coherent decoding. The results show that an increase in the capacity region can be obtained by devoting part of the transmission resources to the transmission of the compressed state.

In this work, we start with the MAC with independent states each strictly causally known to one transmitter of [4]. Combining the same idea of letting transmitters convey a compressed version of the state sequence to the destination, we propose a new achievable scheme with simple encoding and a decoding strategy similar to noisy network coding [10]. The resulting achievable rate region is shown to be potentially larger than the original one of [4]. Moreover, the capacity result for Gaussian channels of [4] for the case of a single state sequence is generalized to a larger class of channels that includes modulo-additive state-dependent MACs. The proposed scheme is then extended to arbitrary number of users. Finally, an achievable rate region is presented for the MAC of arbitrary number of users with common state. As a special case, the region for the two-user MAC with common state given in [3] is recovered. Due to space limitation, most of the proofs are omitted and can be found in [11].

Notation: Probability distributions are identified by their

arguments, e.g., $P_X(x) = \Pr[X = x]$. x_k^i denotes vector $[x_{k1}, ..., x_{ki}]$. $\mathbb{E}[X]$ denotes the expectation of random variable X. $\mathbb{R}_i^+$ denotes the set of non-negative real vectors in i dimension. Notation co refers to the convex hull operation of a set. $\mathcal{N}(0, \sigma^2)$ denotes a zero-mean Gaussian distribution with variance σ^2. $C(x)$ is defined as $C(x) = \frac{1}{2}\log_2(1+x)$.

II. Channel Model and Preliminaries

In this section, we describe our channel models, formulate the problem and revisit related results derived in reference [4].

A. MAC with Independent States

We first investigate an M-user DM MAC with M state sequences, which is denoted by:

$$(\mathcal{X}_1 \times ... \times \mathcal{X}_M, \mathcal{S}_1 \times ... \times \mathcal{S}_M, \mathcal{Y},$$
$$P(s_1)...P(s_M), P(y|x_1, ..., x_M, s_1, ..., s_M)) \quad (1)$$

with input alphabets $(\mathcal{X}_1, ..., \mathcal{X}_M)$, output alphabet $\mathcal{Y}$ and state alphabets $(\mathcal{S}_1, ..., \mathcal{S}_M)$. The state sequences are assumed to be i.i.d. and are mutually independent, i.e., $P(s_1^n...s_M^n) = \prod_{i=1}^n \prod_{k=1}^M P(s_{ki})$. The channel is memoryless in the sense that at any discrete time $i = 1, ..., n$:

$$P(y_i|x_1^n, ..., x_M^n, s_1^n, ..., s_M^n, y^{i-1})$$
$$= P(y_i|x_{1i}, ..., x_{Mi}, s_{1i}, ..., s_{Mi}). \quad (2)$$

Each state realization is available to its corresponding encoder in a *strictly causal* manner. Encoder k's signal x_k^n is subject to an average input cost constraint:

$$\frac{1}{n}\sum_{i=1}^n \mathbb{E}[c_k(X_{ki})] \le \Gamma_k, \quad k = 1, ..., M, \quad (3)$$

where $c_k : \mathcal{X}_k \to \mathbb{R}^+$ is a single-letter input cost function for encoder k and the expectation is taken with respect to all the messages and states. We now define the following code.

Definition 1: Let W_k, uniformly distributed over the set $\mathcal{W}_k = [1 : 2^{nR_k}]$, be the message sent by transmitter k. A $(2^{nR_1}, ..., 2^{nR_M}, n, \Gamma_1, ..., \Gamma_M)$ code with strictly causal and independent SI at the encoders consists of sequences of encoder mappings:

$$f_{k,i} : \mathcal{W}_k \times \mathcal{S}_k^{i-1} \to \mathcal{X}_{ki}, \ i = 1, ..., n, \ k = 1, ..., M, \quad (4)$$

each of which generates a channel input sequence such that (3) is satisfied, and a decoder mapping $g : \mathcal{Y}^n \to \mathcal{W}_1 \times ... \times \mathcal{W}_M$.

The average probability of error, $\Pr(E)$, is defined by:

$$\Pr(E) = \frac{1}{\prod_{k=1}^M 2^{nR_k}} \sum_{w_1=1}^{2^{nR_1}} ... \sum_{w_M=1}^{2^{nR_M}}$$
$$\left(\Pr(g(y^n) \ne (w_1, ..., w_M)|(w_1, ..., w_M)\ sent)\right) \quad (5)$$

Given a cost tuple $\Gamma = (\Gamma_1, ..., \Gamma_M)$, a rate tuple $(R_1, ..., R_M)$ is said to be Γ-achievable if there exists a sequence of codes $(2^{nR_1}, ..., 2^{nR_M}, n, \Gamma_1, ..., \Gamma_M)$ defined above such that $\Pr(E) \to 0$ as $n \to \infty$. The capacity region $\mathcal{C}(\Gamma)$ is the closure of all the Γ-achievable rate tuples.

B. MAC with Common State

We also investigate a general M-user DM MAC with common state. It is defined similarly to the first model by replacing states $s_1 = ... = s_M = s$, i.e., a common state. Each state realization is assumed to be *strictly causally* known to *all* encoders. Each transmitted sequence x_k^n is subject to an average input cost constraint defined as (3). Then the code for the channel remains the same as the one given in Definition 1 with the exception that each $s_1 = ... = s_M = s$. The average probability of error, the Γ-achievable rates and capacity region are also defined as was done in Section II-A.

In the following, we first restrict our attention to a two-user MAC with two independent states, and then generalize to an arbitrary M-user MAC with independent states, common state in Section V and Section VI respectively.

C. Preliminaries

We first summarize and slightly extend a key result of [4].

Theorem 1: [4] Let $\Gamma = (\Gamma_1, \Gamma_2)$ be given. For the two-user MAC with strictly causal SI, a Γ-achievable rate region $\bar{\mathcal{R}}_{in1}(\Gamma_1, \Gamma_2)$ is given by:

$$\left\{ \begin{array}{c} (R_1, R_2) : ((R_1, R_2), (\Gamma_1, \Gamma_2)) \in \\ \text{co} \bigcup_{\beta_1 > 0, \beta_2 > 0} (\mathcal{R}_{in1}(\beta_1, \beta_2), (\beta_1, \beta_2)) \end{array} \right\} \quad (6)$$

where $\mathcal{R}_{in1}(\beta_1, \beta_2)$ is the union of the sets:

$$\bigcup \left\{ \begin{array}{l} (R_1, R_2) \in \mathbb{R}_2^+ : \\ R_1 \le I(X_1; Y|X_2, V_1, V_2) - I(V_1; S_1|Y, V_2) \\ R_2 \le I(X_2; Y|X_1, V_1, V_2) - I(V_2; S_2|Y, V_1) \\ R_1 + R_2 \le I(X_1, X_2; Y|V_1, V_2) \\ \qquad\qquad - I(V_1, V_2; S_1, S_2|Y) \end{array} \right\} \quad (7)$$

with the union taken over the distributions in the set of

$$\mathcal{P}_{sc} = \{P_{V_1, V_2, S_1, S_2, X_1, X_2, Y} :$$
$$P_{V_1|S_1} P_{V_2|S_2} P_{S_1} P_{S_2} P_{X_1} P_{X_2} P_{Y|S_1, S_2, X_1, X_2}\} \quad (8)$$

satisfying $\mathbb{E}[c_k(X_k)] \le \beta_k, \ k = 1, 2$.

Remark 1: The Γ-achievable region given here is a straight-forward extension from the original one in [4] where time-sharing among different cost constraints is allowed, see, e.g., [12]. The basic idea of the achievable scheme is to let transmitters convey a compressed version of the state, namely V_1 for S_1 and V_2 for S_2, to the receiver. The decoder can then use this partial information about the state to improve decoding. The proof of the theorem, though not available in detail in [4], is there indicated to be based on a scheme that exploits distributed Wyner-Ziv compression [13] and block Markov encoding in [4]. □

III. A New Achievable Rate Region

In this section, for the two-user MAC with two independent states, we propose a new achievable scheme that is based on simple encoding, whereby no block-Markov operation is necessary, and on a decoding strategy that is similar to that of noisy network coding in [10]. The scheme achieves a

potentially larger achievable rate region and can be easily generalized to arbitrarily M-user MACs as discussed in Section V.

Theorem 2: Let $\Gamma = (\Gamma_1, \Gamma_2)$ be given. For the two-user MAC with strictly causal SI, a Γ-achievable rate region $\mathcal{R}_{in2}(\Gamma_1, \Gamma_2)$ is given by

$$\left\{ \begin{array}{c} (R_1, R_2) : ((R_1, R_2), (\Gamma_1, \Gamma_2)) \in \\ \text{co} \bigcup_{\beta_1 > 0, \beta_2 > 0} (\mathcal{R}_{in2}(\beta_1, \beta_2), (\beta_1, \beta_2)) \end{array} \right\} \quad (9)$$

where $\mathcal{R}_{in2}(\beta_1, \beta_2)$ is the union of the sets:

$$\bigcup \left\{ \begin{array}{l} (R_1, R_2) \in \mathbb{R}_2^+ : \\ R_1 < I(X_1, V_1; Y | X_2, V_2) - I(V_1; S_1 | X_1) \\ R_2 < I(X_2, V_2; Y | X_1, V_1) - I(V_2; S_2 | X_2) \\ R_1 + R_2 < I(X_1, X_2, V_1, V_2; Y) \\ \qquad - I(V_1; S_1 | X_1) - I(V_2; S_2 | X_2) \end{array} \right\} \quad (10)$$

with the union taken over the distributions in the set of

$$\mathcal{P}_{sc}^* = \big\{ P_{V_1, V_2, S_1, S_2, X_1, X_2, Y} :$$
$$P_{V_1 | S_1, X_1} P_{V_2 | S_2, X_2} P_{S_1} P_{S_2} P_{X_1} P_{X_2} P_{Y | S_1, S_2, X_1, X_2} \big\} \quad (11)$$

satisfying $\mathbb{E}[c_k(X_k)] \le \beta_k$, $k = 1, 2$.

Sketch of Proof: Similar to [4], at each block, each transmitter sends out a codeword which carries both message information and information regarding the state sequence in the previous block. However, unlike [4], the encoders do not use block Markov coding but instead encode the entire message over all the blocks, inspired by noisy network coding. Specifically, the same message w_k, $w_k \in [1 : 2^{nbR_k}]$, is sent at each transmitter $k = 1, 2$, over b blocks of transmission with each consisting of n channel uses. At the end of each block, each transmitter k compresses the state sequence s_k^n over the block without distributed Wyner-Ziv coding, that is, without binning as in [10]. In next block, transmitter k sends out a codeword about both message w_k and the compression index to the receiver. After b blocks of transmission, the decoder performs jointly decoding over all receptions, without explicitly decoding the compressed state information correctly. ■

Remark 2: The resulting region (10) is potentially larger than (7) since it allows a more general input distribution. In fact, it is admissible to generate V_k dependent on both S_k and X_k. Moreover, the cost in terms of rate to be paid for the transmission of compressed state information in (10) can be smaller than that of (7). This is shown in Theorem 3. □

Theorem 3: $\mathcal{R}_{in2}(\Gamma_1, \Gamma_2) \supseteq \mathcal{R}_{in1}(\Gamma_1, \Gamma_2)$ for all cost constraint pairs (Γ_1, Γ_2).

Sketch of Proof: Given any cost constraint pair (β_1, β_2), by setting $P_{V_1 | S_1, X_1} = P_{V_1 | S_1}$ and $P_{V_2 | S_2, X_2} = P_{V_2 | S_2}$ in $\mathcal{R}_{in2}(\beta_1, \beta_2)$ and exploiting facts that $(V_1, V_2) \leftrightarrow (S_1, S_2) \leftrightarrow Y$ forms a Markov chain and (V_1, S_1) is independent of (V_2, S_2), we can show that the sum rate bound in (10) is equal to the one in (7), while the individual bounds on R_1 and R_2 of (10) are larger or equal than the ones in (7). Based on this, and applying time-sharing among different cost constraints, it is possible to conclude the proof. Details can be found in [11]. ■

IV. Capacity Result

In this section, we generalize the capacity result for Gaussian channels of [4] for the special case of a single state sequence to a larger class of channels. We also provide an example that demonstrates the benefits of strictly causal SI.

Consider a class of two-user DM deterministic MACs denoted by $\mathcal{D}_{MAC}$, in which the output Y is a deterministic function of the inputs X_1, X_2 and the channel state S:

$$Y = f(X_1, X_2, S), \quad (12)$$

and where the channel state S, strictly causally known to encoder 1, is another deterministic function of the inputs X_1, X_2 and the output Y:

$$S = g(X_1, X_2, Y), \quad (13)$$

over all product distributions $P_{X_1} P_{X_2}$ satisfying input cost constraints (3) for $k = 1, 2$.

Then the capacity region for $\mathcal{D}_{MAC}$ is identified as follows.

Theorem 4: Let $\Gamma = (\Gamma_1, \Gamma_2)$ be given. For any MAC in $\mathcal{D}_{MAC}$ defined above, the capacity region $\mathcal{C}(\Gamma)$ is given by:

$$\mathcal{C}(\Gamma) \triangleq \bigcup \left\{ \begin{array}{l} (R_1, R_2) \in \mathbb{R}_2^+ : \\ R_1 \le H(Y | X_2, Q) - H(S) \\ R_2 \le H(Y | X_1, S, Q) \\ R_1 + R_2 \le H(Y | Q) - H(S) \end{array} \right\} \quad (14)$$

where the union is taken over all input distributions in form of $P_{X_1 | Q} P_{X_2 | Q} P_Q$ satisfying $\mathbb{E}[c_k(X_k)] \le \Gamma_k$, $k = 1, 2$, and Q is an auxiliary random variable with cardinality $|\mathcal{Q}| \le 5$.

Proof:

i) **Achievability**: We provide the proof of achievability for $Q = q$ constant and drop the conditioning on q for simplicity. The region (14) then follows by using coded time-sharing [14]. Fix $P_{X_1} P_{X_2} P_S P_{Y | X_1, X_2, S}$ and $\mathbb{E}[c_k(X_k)] \le \Gamma_k$, $k = 1, 2$, for a given $Q = q$. By setting $V_2 = S_2 = \emptyset$ and $V_1 = S_1 = S$ in the region (10) we derived, and exploiting the deterministic property of channels, rate pair (R_1, R_2) is achievable where

$$R_1 < H(Y | X_2) - H(S), R_2 < H(Y | X_1, S), \quad (15)$$
$$R_1 + R_2 < H(Y) - H(S), \quad (16)$$
$$\mathbb{E}[c_k(X_k)] \le \Gamma_k, \quad k = 1, 2. \quad (17)$$

ii) **Converse**: From Proposition 1 and 2 in [4], we have that

$$R_1 \le \frac{1}{n} \sum_{i=1}^n I(X_{1i}; Y_i | X_{2i}) + \epsilon_n \quad (18)$$

$$R_1 + R_2 \le \frac{1}{n} \sum_{i=1}^n I(X_{1i}, X_{2i}; Y_i) + \epsilon_n, \quad (19)$$

where $\epsilon_n \to 0$ as $n \to \infty$. Introducing a random variable Q uniformly distributed in the set $[1 : n]$, we have

$$R_1 \le \frac{1}{n} \sum_{i=1}^n I(X_{1q}; Y_q | X_{2q}, Q = q = i) + \epsilon_n \quad (20)$$

$$= I(X_{1Q}; Y_Q | X_{2Q}, Q) + \epsilon_n \quad (21)$$

$$= I(X_1; Y | X_2, Q) + \epsilon_n \quad (22)$$

$$= H\left(Y \left|X_2, Q\right.\right) - H\left(S\right) + \epsilon_n. \tag{23}$$

Along similar lines to (20)−(22), one can show that

$$R_1 + R_2 \le I\left(X_1, X_2; Y \left|Q\right.\right) + \epsilon_n \tag{24}$$
$$= H\left(Y \left|Q\right.\right) - H\left(S\right) + \epsilon_n. \tag{25}$$

Moreover, by providing perfect state information to the receiver, one can easily prove the following bound:

$$R_2 \le \frac{1}{n} \sum_{i=1}^{n} I\left(X_{2i}; Y_i \left|X_{1i}, S_i\right.\right) + \epsilon_n \tag{26}$$

$$= I\left(X_2; Y \left|X_1, S, Q\right.\right) + \epsilon_n \tag{27}$$
$$= H\left(Y \left|X_1, S, Q\right.\right) + \epsilon_n. \tag{28}$$

It can be seen that the distribution on (Q, S, X_1, X_2, Y) from a given code is of the form $P_{Q,S,X_1,X_2,Y} = P_Q P_{X_1|Q} P_{X_2|Q} P_S P_{Y|X_1,X_2,S}$. Notice that both (23) and (25) exploit property (13) and the fact that S is independent of Q.

For the cost constraints, starting from (3), we have

$$\Gamma_k \ge \frac{1}{n} \sum_{i=1}^{n} \sum_{x_{ki}} \Pr\left(x_{ki}\right) c_k\left(x_{ki}\right) \tag{29}$$

$$= \sum_{q=1}^{n} \frac{1}{n} \sum_{x_k} \Pr\left(x_k \left|q\right.\right) c_k\left(x_k\right) = \mathbb{E}\left[c_k\left(X_k\right)\right]. \tag{30}$$

Finally, the cardinality bound $|\mathcal{Q}| \le 5$ follows from the Lemma in [14, Appendix C]. ∎

Remark 3: In the achievability proof, we set $V_1 = S_1 = S$, which implies that V_1 is independent of X_1. Hence the achievable scheme proposed in [4] is also optimal for the class of channels considered here. □

Example 1: The class $\mathcal{D}_{MAC}$ includes the Gaussian model considered in [4], which is defined as $Y_i = X_{1i} + X_{2i} + S_i$ with input power constraints $\frac{1}{n} \sum_{i=1}^{n} \mathbb{E}\left[X_{ki}^2\right] \le P_k$ and $S_i \sim \mathcal{N}\left(0, \sigma_s^2\right)$ strictly causally known to encoder 1. The capacity region $\mathcal{C}$ can be identified from Theorem 4 by simple extension to continuous alphabets [6], and is given by:

$$\left\{ \begin{array}{l} (R_1, R_2) \in \mathbb{R}_2^+ : \\ R_1 \le C\left(\frac{P_1}{\sigma_s^2}\right), R_1 + R_2 \le C\left(\frac{P_1 + P_2}{\sigma_s^2}\right) \end{array} \right\}. \tag{31}$$

Example 2: The class $\mathcal{D}_{MAC}$ contains more channels beside the Gaussian model discussed in Example 1. In particular, consider a class of binary modulo-additive state-dependent MAC channels, e.g., $Y = X_1 \oplus X_2 \oplus S$, where $S \sim Bernoulli\left(p_s\right)$, with input cost constraints $\frac{1}{n} \sum_{i=1}^{n} \mathbb{E}[X_{1i}] \le p_1$ and $\frac{1}{n} \sum_{i=1}^{n} \mathbb{E}[X_{2i}] \le p_2$, $0 < p_1, p_2, p_s \le \frac{1}{2}$. The capacity region $\mathcal{C}_{bin}^s$ for this class of channels is given by:

$$\left\{ \begin{array}{l} (R_1, R_2) \in \mathbb{R}_2^+ : \\ R_1 \le H_b\left(p_1 * p_s\right) - H_b\left(p_s\right), R_2 \le H_b\left(p_2\right) \\ R_1 + R_2 \le H_b\left(p_1 * p_2 * p_s\right) - H_b\left(p_s\right) \end{array} \right\}, \tag{32}$$

where $p_1 * p_2 = p_1\left(1 - p_2\right) + p_2\left(1 - p_1\right)$, and $H_b\left(p\right) = -p \log_2 p - \left(1 - p\right) \log_2\left(1 - p\right)$.
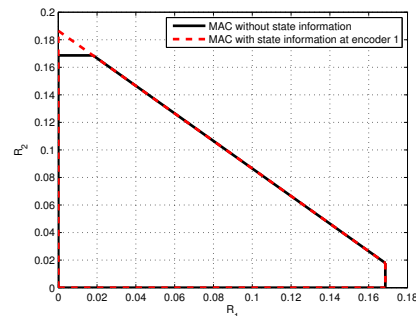

Fig. 1. Capacity region for an example of modulo-additive state-dependent MAC with input constraints $(p_1 = p_2 = \frac{1}{3}, p_s = \frac{1}{4})$

The capacity region (32) can be identified from Theorem 4 as follows. Given a fixed $Q = q$ and input constraints $\mathbb{E}\left[c_k\left(X_k\right) \left|Q = q\right.\right] \le p_{kq}$, for some p_{kq} such that $\sum_{\mathcal{Q}} p_{kq} \Pr\left(Q = q\right) \le p_k$, each bound in (14) is maximized by the same input distributions $Bernoulli\left(p_{kq}\right)$, $k = 1, 2$. For example, on bound R_1

$$H\left(Y \left|X_2, Q = q\right.\right) - H\left(S\right) \tag{33}$$
$$= H\left(X_1 \oplus S \left|Q = q\right.\right) - H_b\left(p_s\right) \tag{34}$$
$$\le H_b\left(p_{1q} * p_s\right) - H_b\left(p_s\right), \tag{35}$$

where (35) follows from the fact that entropy $H\left(X_1 \oplus S \left|Q = q\right.\right)$ monotonically increases with the probability of $X_1 = 1$. Hence,

$$H\left(Y \left|X_2, Q\right.\right) - H\left(S\right) \tag{36}$$
$$= \sum_{q \in \mathcal{Q}} H\left(Y \left|X_2, Q = q\right.\right) \Pr\left(Q = q\right) - H\left(S\right) \tag{37}$$
$$\le \sum_{q \in \mathcal{Q}} H_b\left(p_{1q} * p_s\right) \Pr\left(Q = q\right) - H_b\left(p_s\right) \tag{38}$$
$$\le H_b\left(p_1 * p_s\right) - H_b\left(p_s\right), \tag{39}$$

where (39) follows by the concavity of entropy. Thus there is no need for time-sharing among different cost constraints. By similar arguments on the other two bounds, we claim the capacity region in (32).

Remark 4: It is known from [6] that, without SI, the capacity region $\mathcal{C}_{bin}^{ns}$ for this MAC channel is given by:

$$\left\{ \begin{array}{l} (R_1, R_2) \in \mathbb{R}_2^+ : \\ R_1 \le H_b\left(p_1 * p_s\right) - H_b\left(p_s\right), \\ R_2 \le H_b\left(p_2 * p_s\right) - H_b\left(p_s\right), \\ R_1 + R_2 \le H_b\left(p_1 * p_2 * p_s\right) - H_b\left(p_s\right) \end{array} \right\}. \tag{40}$$

Hence $\mathcal{C}_{bin}^{ns} \subseteq \mathcal{C}_{bin}^s$. For a numerical example, we set $p_1 = p_2 = \frac{1}{3}$ and $p_s = \frac{1}{4}$. The corresponding regions (32) and (40) are depicted and compared in Fig. 1. It is evident that the presence of strictly causal SI at encoder 1 improves the maximum rate of user 2. □

V. M-USER MAC WITH INDEPENDENT STATES

In this section, we generalize the proposed scheme to an arbitrary number of users with independent states. Due to

limited space, from now on, we only present the achievable rate region given a cost tuple $(\beta_1, ..., \beta_M)$, from which a Γ-achievable rate region can be obtained by convex hull operation and time-sharing among different cost constraints as similarly characterized in (9) for the two-user case.

Let $\mathcal{S}$ denote any subset of $[1 : M]$, i.e., $\mathcal{S} \subseteq [1 : M]$ and $\mathcal{S}^c$ be the complement of $\mathcal{S}$. Any random vector $\mathbf{X}(\mathcal{S})$ is defined by $\mathbf{X}(\mathcal{S}) \triangleq (X_1, X_2, ..., X_l)$, $l = |\mathcal{S}|$, the cardinality of $\mathcal{S}$.

Theorem 5: Given a cost tuple $(\beta_1, ..., \beta_M)$, for the M-user MAC with strictly causal and independent SI, an achievable rate region $\mathcal{R}_{in}^M(\beta_1, ..., \beta_M)$ is given by the union of the sets:

$$\bigcup \left\{ \begin{array}{l} (R_1, ..., R_M) \in \mathbb{R}_M^+ : \\ \sum_{k \in \mathcal{T}} R_k < \\ \min_{\substack{\mathcal{S} \subseteq [1:M] \\ \mathcal{T} \subseteq \mathcal{S}}} \left(\begin{array}{c} I(\mathbf{X}(\mathcal{S}), \mathbf{V}(\mathcal{S}); Y | \mathbf{X}(\mathcal{S}^c), \mathbf{V}(\mathcal{S}^c)) \\ - \sum_{l \in \mathcal{S}} I(V_l; S_l | X_l) \end{array} \right), \\ \forall \ \mathcal{T} \subseteq [1:M] \end{array} \right\}$$

with the union taken over the distributions in the set of

$$\mathcal{P}_{sc}^M = \Big\{ P_{V_1,...,V_M,S_1,...,S_M,X_1,...,X_M,Y} : $$
$$\prod_{k=1}^M \left(P_{V_k|S_k,X_k} P_{S_k} P_{X_k} \right) P_{Y|S_1,...,S_M,X_1,...,X_M} \Big\} \quad (41)$$

satisfying $\mathbb{E}[c_k(X_k)] \leq \beta_k$, $k = 1, ..., M$.

Proof: The coding scheme follows from that of Theorem 2 for the two-user DM MAC. Details can be found in [11]. ∎

VI. M-USER MAC WITH COMMON STATE

In this section, we present an achievable rate region for the M-user MAC with common SI by applying a similar scheme to that for the case with independent SI.

Theorem 6: Given a cost tuple $(\beta_1, ..., \beta_M)$, for the M-user MAC with strictly causal common SI, an achievable rate region $\mathcal{R}_{in}^{M,c}(\beta_1, ..., \beta_M)$ is given by the union of the sets:

$$\bigcup \left\{ \begin{array}{l} (R_1, ..., R_M) \in \mathbb{R}_M^+ : \\ \sum_{k \in [1:M]} R_k < I(\mathbf{X}([1:M]), V; Y) - I(V; S), \\ \sum_{k \in \mathcal{S}} R_k < I(\mathbf{X}(\mathcal{S}); Y | U, V, \mathbf{X}(\mathcal{S}^c)), \\ \forall \ \mathcal{S} \subseteq [1:M] \end{array} \right\}$$

with the union taken over the distributions in the set of

$$\mathcal{P}_{sc}^{M,c} = \Big\{ P_{V,S,U,X_1,...,X_M,Y} : $$
$$P_S P_{V|S} P_U \left(\prod_{k=1}^M P_{X_k|U} \right) P_{Y|S,X_1,...,X_M} \Big\} \quad (42)$$

satisfying $\mathbb{E}[c_k(X_k)] \leq \beta_k$, $k = 1, ..., M$.

Proof: The proof follows a similar idea of forwarding compressed state based on noisy network coding while coupled with a cooperative transmission strategy for the DM MAC with common message [15]. We refer for details to [11]. ∎

Remark 5: Given any cost pair (β_1, β_2), setting $M = 2$ in $\mathcal{R}_{in}^{M,c}$ derived above, we recover the achievable region for the two-user DM MAC with common state in [3]. Moreover, as

shown in [3], this region is the capacity region for the two-user Gaussian MAC, $Y = X_1 + X_2 + S$, where $S \sim \mathcal{N}(0, \sigma_s^2)$ and is known strictly causally at both encoders. ∎

VII. CONCLUDING REMARKS

In this work, we have investigated whether it is beneficial to have strictly causal state information at encoders only in a MAC. To answer this question, we first revisited a two-user MAC in [4] and derived a new achievable rate region based on simple encoding at the transmitter and a decoding strategy inspired by noisy network coding. The new region turns out to be potentially larger than that in [4]. The capacity result for the Gaussian model of [4] for the special case of a single state sequence is generalized to a larger class of channels that includes modulo-additive state-dependent MACs. Moreover, we generalized the proposed scheme to arbitrarily M-user MACs and established the achievable rate regions for the case with independent states and common state respectively. In general, our results point to the advantages of state information at encoders in the MACs considered, despite it being known only strictly causally.

REFERENCES

[1] C. E. Shannon, "Channels with side information at the transmitter," *IBM Journal of Research and Development*, vol. 2, no. 4, pp. 289–293, October 1958.

[2] S. I. Gel'fand and M. S. Pinsker, "Coding for channel with random parameters," *Problems of Control and Information Theory*, vol. 9, no. 1, pp. 19–31, 1980.

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

[4] ——, "The multiple access channel with two independent states each known causally to one encoder," in *Proceedings of IEEE International Symposium on Information Theory*, June 2010.

[5] G. Keshet, Y. Steinberg, and N. Merhav, "Channel coding in the presence of side information," *Foundations and Trends in Communications and Information Theory*, vol. 4, no. 6, pp. 445–586, 2007.

[6] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley-Interscience, July 2006.

[7] S. Jafar, "Capacity with causal and noncausal side information: A unified view," *IEEE Transactions on Information Theory*, vol. 52, no. 12, pp. 5468–5474, December 2006.

[8] A. Somekh-Baruch, S. Shamai, and S. Verdu, "Cooperative multiple-access encoding with states available at one transmitter," *IEEE Transactions on Information Theory*, vol. 54, no. 10, pp. 4448–4469, October 2008.

[9] S. P. Kotagiri and J. N. Laneman, "Multiaccess channels with state known to some encoders and independent messages," *EURASIP Journal on Wireless Communications and Networking*, vol. 2008, pp. 1–14, January 2008.

[10] S. H. Lim, Y. H. Kim, A. El Gamal, and S. Y. Chung, "Noisy network coding," 2010, available online at http://arxiv.org/abs/1002.3188v2.

[11] M. Li, O. Simeone, and A. Yener, "Multiple access channels with states causally known at transmitters," November 2010, submitted to *IEEE Transactions on Information Theory*, available online at http://arxiv.org/abs/1011.6639.

[12] S. Verdu, "On channel capacity per unit cost," *IEEE Transactions on Information Theory*, vol. 36, no. 5, pp. 1019–1030, September 2002.

[13] M. Gastpar, "The Wyner-Ziv problem with multiple sources," *IEEE Transactions on Information Theory*, vol. 50, no. 11, pp. 2762–2768, November 2004.

[14] A. El Gamal and Y. H. Kim, "Lecture notes on network information theory," 2010, available online at http://arxiv.org/abs/1001.3404.

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