Multiaccess Channels with State Known to One Encoder: A Case of Degraded Message Sets

Shivaprasad Kotagiri and J. Nicholas Laneman
Department of Electrical Engineering
University of Notre Dame
Notre Dame, IN 46556, USA
Email: {skotagir,jnl}@nd.edu

***Abstract*— We consider a state-dependent multiple access channel $p(y|x_1, x_2, s)$ whose output Y is controlled by the channel inputs X_1 and X_2 from two encoders and the channel state S. It is assumed that the channel state is known non-causally at one encoder, called the informed encoder. We derive the capacity region for the case of degraded messages in which the informed encoder knows the message of the uninformed encoder.**

I. INTRODUCTION

We consider a multiple access channel (MAC) controlled by a random parameter or channel state generated from a memoryless source as shown in Figure 1. We assume that channel state information (CSI) is known non-causally at one encoder, called the *informed encoder*; the other encoder is called *uninformed*. The informed encoder knows both messages W_1 with rate R_1 and W_2 with rate R_2, and the uninformed encoder knows only message W_2, i.e., there is a *degraded message set*. These two encoders want to communicate their messages to a decoder through a state-dependent memoryless MAC $p(y|x_1, x_2, s)$ controlled by the channel inputs X_1 and X_2 from the informed encoder and the uninformed encoder, respectively, and the channel state S. The decoder forms estimates $(\hat{W}_1, \hat{W}_2)$ of (W_1, W_2) from the channel output Y^n. We study the capacity region of this model.

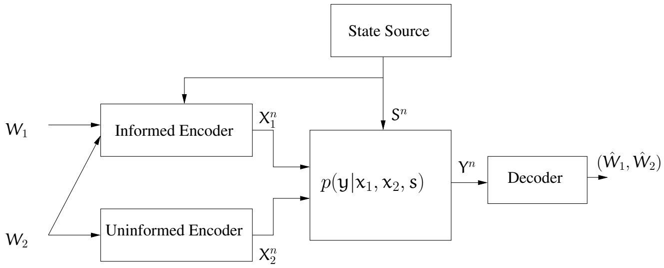

Fig. 1. Multiple access channel (MAC) with state information known to one encoder and a degraded message set.

Shannon initiated study of single-user channel models with random parameters, and there have been several recent applications of his and related work in, e.g., watermarking and information embedding, precoding in vector broadcast channels, and so forth. The capacity of discrete memoryless (DM) channels $p(y|x, s)$ is given by [1], where X and Y are the channel input and output, respectively, and memoryless CSI $S \sim p(s)$ is known *causally* at the encoder. The case of CSI known *non-causally* at the encoder is introduced in [2], [3] in the context of computer memories with defects. In [4], Gel'fand and Pinsker derive the capacity of such models, which is given by

$$C = \max_{p(u|s),\ X=f(U,S)} \mathbb{I}(U;Y) - \mathbb{I}(U;S) \tag{1}$$

where U is an auxiliary random variable, and X is a deterministic function of (U, S). In [5], Costa extends this result to the state-dependent additive Gaussian model with encoder CSI, and inspires the terminology of *dirty paper coding*.

The model in Figure 1 is one of many possible extensions of state-dependent channels to *multiuser* problems, e.g., [6], [7], [8], [9], [10], [11]. In particular, it is a variation on the model of *asymmetric* encoder CSI initiated in [12], [13], and related to the MAC with cribbing encoders but without state [14] . In contrast to the symmetric Gaussian case in which both encoders have CSI, and for which the capacity region is derived in [6], [7], only inner and outer bounds on the capacity region are obtained in [13] for both the discrete memoryless and Gaussian cases.

However, as we show formally in Theorem 1, if there is a degraded message set as in Figure 1, rates (R_1, R_2) are achievable only if

$$R_1 \le \mathbb{I}(U;Y|X_2,Q) - \mathbb{I}(U;S|X_2,Q)$$
$$R_1 + R_2 \le \mathbb{I}(U,X_2;Y|Q) - \mathbb{I}(U;S|X_2,Q)$$

for some joint distribution of the form

$$p(\mathsf{q})p(\mathsf{s})p(\mathsf{x}_2|\mathsf{q})p(\mathsf{u},\mathsf{x}_1|\mathsf{s},\mathsf{x}_2,\mathsf{q})p(\mathsf{y}|\mathsf{s},\mathsf{x}_1,\mathsf{x}_2),$$

and achievable if both inequalities are strict. A coding scheme that achieves the capacity region is a combination of superposition coding and Gel'fand-Pinsker coding at the informed encoder. This problem is also studied in the discrete memoryless and Gaussian cases in [15], building on the special case of only a common message i.e., $W_1 = 0$, studied in [16].

Before proceeding to a precise statement of the problem in Section II and derivation of the capacity region in Section III, we say a few words about our notation. Throughout the paper, random variables and sample values are denoted in a special font, e.g., the random variable X and sample value x. Alphabets are denoted in calligraphic font, e.g., $\mathcal{X}$, and are all discrete. The shorthand X_1^n represents the sequence $\mathsf{X}_{1,1}, \mathsf{X}_{1,2}, \ldots, \mathsf{X}_{1,n}$, and $\mathsf{X}_{1,i}^n$ represents the sequence $\mathsf{X}_{1,i}, \mathsf{X}_{1,i+1}, \ldots, \mathsf{X}_{1,n}$. Finally, $\mathbb{H}(\cdot)$ and $\mathbb{I}(\cdot;\cdot)$ denote the standard information theoretic quantities of (ensemble average) entropy and mutual information.

II. CHANNEL MODEL AND DEFINITIONS

As shown in Figure 1, the MAC is embedded in some environment in which CSI is non-causally known at one encoder. A memoryless state-dependent MAC with CSI, denoted $p(\mathsf{y}|\mathsf{x}_1,\mathsf{x}_2,\mathsf{s})$, is controlled by the output $\mathsf{S} \in \mathcal{S}$ of a memoryless state source $p(\mathsf{s})$ and the channel input pair $(\mathsf{X}_1, \mathsf{X}_2) \in (\mathcal{X}_1, \mathcal{X}_2)$, and generates the channel output $\mathsf{Y} \in \mathcal{Y}$. We assume that S_i's are independent and identically distributed (i.i.d.) random variables drawn according to $p(\mathsf{s})$, $i = 1, 2, \ldots, n$.

The message sources produce random integers $W_1 \in \{1, 2, \ldots, M_1\}$ and $W_2 \in \{1, 2, \ldots, M_2\}$ at the beginning of each block of integer length n channel uses. We assume that the messages are independent, and the probability of each pair of messages $(W_1 = w_1, W_2 = w_2)$ is given by $\frac{1}{M_1 M_2}$. The rate of message W_i is $(1/n)\log M_i$, for $i = 1, 2$. We also assume that both messages W_1 and W_2 are provided to the informed encoder and only the message W_2 is provided to the uninformed encoder.

Definition 1: For a positive integer n and a pair (R_1, R_2) of real, non-negative rates, a $(\lceil 2^{nR_1}\rceil, \lceil 2^{nR_2}\rceil, n)$ code consists of encoding functions

$$f_1^n : \mathcal{S}^n \times \mathcal{W}_1 \times \mathcal{W}_2 \to \mathcal{X}_1^n \text{ and } f_2^n : \mathcal{W}_2 \to \mathcal{X}_2^n$$

at the informed encoder and the uninformed encoder, respectively, and a decoding function

$$g^n : \mathcal{Y}^n \to \mathcal{W}_1 \times \mathcal{W}_2.$$

From a $(\lceil 2^{nR_1}\rceil, \lceil 2^{nR_2}\rceil, n)$ code, the sequences X_1^n and X_2^n from the informed encoder and the uninformed encoder, respectively, are transmitted across a memoryless state-dependent MAC with CSI and without feedback modeled as a discrete memoryless conditional probability distribution, so that

$$p(\mathsf{y}^n|\mathsf{s}^n,\mathsf{x}_1^n,\mathsf{x}_2^n) = \prod_{j=1}^{n} p(\mathsf{y}_j|\mathsf{s}_j,\mathsf{x}_{1,j},\mathsf{x}_{2,j}). \tag{2}$$

The decoder reconstructs the messages from the channel output Y^n. The average probability of error is defined as $P_e^n := \mathbb{P}[g(\mathsf{Y}^n) \neq (W_1, W_2)]$.

Definition 2: A rate pair of real, non-negative rates (R_1, R_2) is said to be achievable if there exists a sequence of $(\lceil 2^{nR_1}\rceil, \lceil 2^{nR_2}\rceil, n)$ codes (f_1^n, f_2^n, g^n) with $\lim_{n\to\infty} P_e^n = 0$.

Definition 3: The capacity region $\mathcal{C} \subseteq \mathbb{R}^2$ is the closure of the set of achievable rate pairs (R_1, R_2).

III. THE CAPACITY REGION

The following theorem presents the capacity region of the model shown in Figure 1.

Theorem 1: The capacity region $\mathcal{C}$ of the model shown in Figure 1 is the closure of all rate pairs (R_1, R_2) satisfying

$$R_1 < \mathbb{I}(\mathsf{U};\mathsf{Y}|\mathsf{X}_2,\mathsf{Q}) - \mathbb{I}(\mathsf{U};\mathsf{S}|\mathsf{X}_2,\mathsf{Q}) \tag{3a}$$

$$R_1 + R_2 < \mathbb{I}(\mathsf{U},\mathsf{X}_2;\mathsf{Y}|\mathsf{Q}) - \mathbb{I}(\mathsf{U};\mathsf{S}|\mathsf{X}_2,\mathsf{Q}) \tag{3b}$$

for some random variables $(\mathsf{Q},\mathsf{S},\mathsf{U},\mathsf{X}_1,\mathsf{X}_2,\mathsf{Y})$ whose distribution is of the form

$$p(\mathsf{q})p(\mathsf{s})p(\mathsf{x}_2|\mathsf{q})p(\mathsf{u},\mathsf{x}_1|\mathsf{s},\mathsf{x}_2,\mathsf{q})p(\mathsf{y}|\mathsf{s},\mathsf{x}_1,\mathsf{x}_2),$$

where, U is an auxiliary random variable, and Q is a time-sharing auxiliary random variable.

Remarks:

- The region in (3) without the time-sharing random variable is not convex. The time-sharing auxiliary random variable Q is introduced in the above theorem to take the convex closure.
- To compute the region in the above theorem, it is sufficient to consider the auxiliary random variables U and Q with $|\mathcal{U}| \leq |\mathcal{X}_1||\mathcal{X}_2||\mathcal{S}|$ and $|\mathcal{Q}| \leq 3$, respectively.

A. Converse

First, we prove that, for any sequence of $(\lceil 2^{nR_1}\rceil, \lceil 2^{nR_2}\rceil, n)$ codes with $P_e^n \to 0$ as $n \to \infty$, the rate pair $(R_1, R_2) \in \mathcal{C}$. Fix n and consider a given code of block length n. The joint distribution of random variables W_1, W_2, S^n, X_1^n, and X_2^n is given by

$$\begin{aligned} & p(\mathsf{s}^n, w_1, w_2, \mathsf{x}_1^n, \mathsf{x}_2^n, \mathsf{y}^n) = \\ & \quad \frac{1}{\lceil 2^{nR_1}\rceil \lceil 2^{nR_2}\rceil} p(\mathsf{s}^n) p(\mathsf{x}_1^n | w_1, w_2, \mathsf{s}^n) p(\mathsf{x}_2^n | w_2) \\ & \quad \times p(\mathsf{y}^n | \mathsf{s}^n, \mathsf{x}_1^n, \mathsf{x}_2^n) \end{aligned}$$

where, $p(\mathsf{x}_1^n|w_1, w_2, \mathsf{s}^n)$ is 1 if $\mathsf{x}_1^n = f_1^n(w_1, w_2, \mathsf{s}^n)$ and 0 otherwise, and $p(\mathsf{x}_2^n|w_2)$ is 1 if $\mathsf{x}_2^n = f_2^n(w_2)$ and 0 otherwise.

We can bound the rate R_1 as

$$\begin{aligned} nR_1 \leq & \mathbb{H}(W_1) \\ = & \mathbb{H}(W_1|W_2) \\ = & \mathbb{I}(W_1; \mathsf{Y}^n|W_2) + \mathbb{H}(W_1|W_2, \mathsf{Y}^n) \\ \overset{(a)}{\leq} & \mathbb{I}(W_1; \mathsf{Y}^n|W_2) + n\epsilon_n \\ \overset{(b)}{=} & \sum_{i=1}^{n} [\mathbb{I}(W_1, \mathsf{S}_{i+1}^n; \mathsf{Y}^i|W_2) - \mathbb{I}(W_1, \mathsf{S}_i^n; \mathsf{Y}^{i-1}|W_2)] + n\epsilon_n \\ \overset{(c)}{=} & \sum_{i=1}^{n} [\mathbb{I}(W_1, \mathsf{S}_{i+1}^n; \mathsf{Y}_i|W_2, \mathsf{Y}^{i-1}) \\ & \quad - \mathbb{I}(\mathsf{S}_i; \mathsf{Y}^{i-1}|W_1, \mathsf{S}_{i+1}^n, W_2)] + n\epsilon_n \\ \overset{(d)}{=} & \sum_{i=1}^{n} [\mathbb{I}(W_1, \mathsf{S}_{i+1}^n; \mathsf{Y}_i|W_2, \mathsf{Y}^{i-1}, \mathsf{X}_{2,i}) \\ & \quad - \mathbb{I}(\mathsf{S}_i; \mathsf{Y}^{i-1}|W_1, \mathsf{S}_{i+1}^n, W_2, \mathsf{X}_{2,i})] + n\epsilon_n \\ \overset{(e)}{\leq} & \sum_{i=1}^{n} [\mathbb{H}(\mathsf{Y}_i|\mathsf{X}_{2,i}) - \mathbb{H}(\mathsf{Y}_i|W_2, \mathsf{Y}^{i-1}, W_1, \mathsf{S}_{i+1}^n, \mathsf{X}_{2,i}) \\ & \quad + \mathbb{H}(\mathsf{S}_i|W_2, \mathsf{Y}^{i-1}, W_1, \mathsf{S}_{i+1}^n, \mathsf{X}_{2,i}) \\ & \quad - \mathbb{H}(\mathsf{S}_i|\mathsf{X}_{2,i})] + n\epsilon_n \\ = & \sum_{i=1}^{n} [\mathbb{I}(W_1, W_2, \mathsf{S}_{i+1}^n, \mathsf{Y}^{i-1}; \mathsf{Y}_i|\mathsf{X}_{2,i}) \\ & \quad - \mathbb{I}(W_1, W_2, \mathsf{S}_{i+1}^n, \mathsf{Y}^{i-1}; \mathsf{S}_i|\mathsf{X}_{2,i})] + n\epsilon_n \end{aligned} \tag{4}$$

where:

(a) follows from Fano's inequality with $\epsilon_n \to 0$ as $P_e^n \to 0$,

(b) follows from $\mathbb{I}(W_1, \mathsf{S}_{i+1}^n; \mathsf{Y}^i|W_2) = \mathbb{I}(W_1; \mathsf{Y}^n|W_2)$ for $i = n$; $\mathbb{I}(W_1, \mathsf{S}_i^n; \mathsf{Y}^{i-1}|W_2) = 0$ for $i = 1$; and the sum of the remaining terms equals to zero,

(c) follows from applying the chain rule for mutual information to $(\mathsf{Y}^{i-1}, \mathsf{Y}_i)$ in the first term and to $(\{W_1, \mathsf{S}_{i+1}^n\}, \mathsf{S}_i)$ in the second term,

(d) follows from $\mathsf{X}_{2,i}$ is a deterministic function of W_2 for $i = \{1, 2, \ldots, n\}$,

(e) follows from $\mathbb{H}(\mathsf{Y}_i|\mathsf{X}_{2,i}) \geq \mathbb{H}(\mathsf{Y}_i|W_2, \mathsf{Y}^{i-1}, \mathsf{X}_{2,i})$ and S_i being independent of $(\mathsf{S}_{i+1}^n, W_1, W_2)$.

Finally, the sum rate $R_1 + R_2$ can be upper bounded as

$$\begin{aligned} n(R_1 + R_2) = & \mathbb{H}(W_1, W_2) \\ = & \mathbb{I}(W_1, W_2; \mathsf{Y}^n) + \mathbb{H}(W_1, W_2|\mathsf{Y}^n) \\ \overset{(a)}{\leq} & \mathbb{I}(W_1, W_2; \mathsf{Y}^n) + n\epsilon_n \\ \overset{(b)}{=} & \sum_{i=1}^{n} \mathbb{I}(W_1, W_2, \mathsf{S}_{i+1}^n; \mathsf{Y}^i) \\ & \quad - \mathbb{I}(W_1, W_2, \mathsf{S}_i^n; \mathsf{Y}^{i-1}) + n\epsilon_n \\ \overset{(c)}{=} & \sum_{i=1}^{n} \mathbb{I}(W_1, W_2, \mathsf{S}_{i+1}^n; \mathsf{Y}_i|\mathsf{Y}^{i-1}) \\ & \quad - \mathbb{I}(\mathsf{S}_i; \mathsf{Y}^{i-1}|W_1, \mathsf{S}_{i+1}^n, W_2) + n\epsilon_n \\ = & \sum_{i=1}^{n} \mathbb{I}(W_1, W_2, \mathsf{S}_{i+1}^n; \mathsf{Y}_i|\mathsf{Y}^{i-1}) \\ & \quad - \mathbb{H}(\mathsf{S}_i|W_1, \mathsf{S}_{i+1}^n, W_2) \\ & \quad + \mathbb{H}(\mathsf{S}_i|W_1, \mathsf{S}_{i+1}^n, \mathsf{Y}^{i-1}, W_2) + n\epsilon_n \\ \leq & \sum_{i=1}^{n} \mathbb{I}(W_1, W_2, \mathsf{S}_{i+1}^n, \mathsf{Y}^{i-1}; \mathsf{Y}_i) \\ & \quad - \mathbb{H}(\mathsf{S}_i) + \mathbb{H}(\mathsf{S}_i|W_1, \mathsf{S}_{i+1}^n, \mathsf{Y}^{i-1}, W_2) + n\epsilon_n, \\ \overset{(d)}{=} & \sum_{i=1}^{n} \mathbb{I}(W_1, W_2, \mathsf{S}_{i+1}^n, \mathsf{Y}^{i-1}, \mathsf{X}_{2,i}; \mathsf{Y}_i) \\ & \quad - \mathbb{I}(W_1, W_2, \mathsf{S}_{i+1}^n, \mathsf{Y}^{i-1}, \mathsf{X}_{2,i}; \mathsf{S}_i), \\ \overset{(e)}{=} & \sum_{i=1}^{n} \mathbb{I}(W_1, W_2, \mathsf{S}_{i+1}^n, \mathsf{Y}^{i-1}, \mathsf{X}_{2,i}; \mathsf{Y}_i) \\ & \quad - \mathbb{I}(W_1, W_2, \mathsf{S}_{i+1}^n, \mathsf{Y}^{i-1}; \mathsf{S}_i|\mathsf{X}_{2,i}), \end{aligned} \tag{5}$$

where:

(a) follows from Fano's inequality with $\epsilon_n \to 0$ as $P_e^n \to 0$,

(b) follows from $\mathbb{I}(W_1, W_2, \mathsf{S}_{i+1}^n; \mathsf{Y}^i) = \mathbb{I}(W_1, W_2; \mathsf{Y}^n)$ for $i = n$; $\mathbb{I}(W_1, W_2, \mathsf{S}_i^n; \mathsf{Y}^{i-1}) = 0$ for $i = 1$; and the sum of remaining terms equals zero,

(c) follows from applying the chain rule for mutual information to $(\mathsf{Y}^{i-1}, \mathsf{Y}_i)$ in the first term and to $(\{W_1, \mathsf{S}_{i+1}^n, W_2\}, S_i)$ in the second term,

(d) follows from $\mathsf{X}_{2,i}$ being a deterministic function of W_2 for $i = \{1, 2, \ldots, n\}$.

Let us define $\mathsf{U}(i) := (W_1, W_2, \mathsf{S}_{i+1}^n, \mathsf{Y}^{i-1})$ and Q to take values uniformly in the set $\mathcal{Q} = \{1, 2, \ldots, n\}$. As $n \to \infty$, we obtain $(R_1, R_2) \in \mathcal{C}$ from (4) and (5) where the distribution on $(\mathsf{Q}, \mathsf{S}, \mathsf{X}_2, \mathsf{X}_1, \mathsf{U}, \mathsf{Y})$ is $p(\mathsf{s})p(\mathsf{q})p(\mathsf{x}_{2,\mathsf{q}}|\mathsf{q})p(\mathsf{u}_\mathsf{q}, \mathsf{x}_{1,\mathsf{q}}|\mathsf{q})p(\mathsf{y}|\mathsf{s}, \mathsf{x}_{1,\mathsf{q}}, \mathsf{x}_{2,\mathsf{q}})$.

B. Achievability

We denote by $T_\epsilon^n[X,Y]$ the set of jointly strongly typical sequences [17], [18] with distribution $p(x,y)$. Let $T_\epsilon^n[X,Y|x^n] := \{y^n : (x^n,y^n) \in T_\epsilon^n[X,Y]\}$. In this section, we construct a sequence of $(\lceil 2^{nR_1}\rceil, \lceil 2^{nR_2}\rceil, n)$ codes with $P_e^n \to 0$ as $n \to \infty$ if (R_1,R_2) satisfies (3). Fix $\epsilon > 0$ and a distribution $p(q)p(s)p(x_2|q)p(u,x_1|q,s,x_2)$ on (Q,S,U,X_1,X_2). In this case, Q is a time-sharing random variable. Generate the time-sharing sequence $Q^n = q^n$ according to $\prod p(q_i)$. Without loss generality, it is assumed that the time-sharing sequence is non-causally known at the encoders and the decoder.

1) Encoding: The encoding strategy at the two encoders is as follows. Let $M_1 = 2^{n(R_1-4\epsilon)}$, $M_2 = 2^{n(R_2-2\epsilon)}$, and $J = 2^{n(\mathbb{I}(U;S|Q,X_2)+2\epsilon)}$. At the uninformed encoder, generate the sequences X_2^n, according to $\prod p(x_{2,i}|q_i)$, where $1 \le m_2 \le M_2$. The uninformed encoder sends the codeword $X_2^n(q^n,w_2)$ to send the message $W_2 \in \{1,2,\dots,M_2\}$ for a given time-sharing sequence $Q^n = q^n$.

At the informed encoder, generate JM_1 sequences $U^n(q^n,m_1,m_2,j)$, with $\prod p(u_i|q_i,x_{2,i}(q,m_2))$, where $1 \le m_1 \le M_1$, $1 \le m_2 \le M_2$ and $1 \le j \le J$. Here, (m_1,m_2) indexes bins and j indexes sequences within a particular bin (m_1,m_2). To encode the message $(W_1,W_2) \in \{1,2,\dots,M_1\} \times \{1,2,\dots,M_2\}$ given $S^n = s^n$ and $Q^n = q^n$, the informed encoder looks in bin (W_1,W_2) for a sequence $U^n(q^n,w_1,w_2,j)$, $1 \le j \le J$, such that $U^n(q^n,w_1,w_2,j) \in T_\epsilon^n[Q,U,S,X_2|q^n,s^n,X_2^n(w_2)]$. Then, the informed encoder generates X_1^n with $\prod p(x_{1,i}|q_i,s_i,u_i(q^n,w_1,w_2,j),x_{2,i}(q^n,w_2))$.

Given (S^n,X_1^n,X_2^n), the channel generates the output Y^n according to conditional probability distribution $\prod_i p(y_i|s_i,x_{1,i},x_{2,i})$.

2) Decoding: The decoder chooses a pair $(U^n(q^n,m_1,m_2,j),X_2^n(m_2))$, $1 \le m_1 \le M_1$, $1 \le j \le J$, and $1 \le m_2 \le M_2$ such that $(U^n(q^n,m_1,m_2,j),X_2^n(q^n,m_2)) \in T_\epsilon^n[Q,U,X_2,Y|q^n,Y^n]$. If such a pair exists and is unique, the decoder declares that $(\hat{W}_1,\hat{W}_2) = (m_1,m_2)$. Otherwise, the decoder declares an error.

3) Analysis of Probability of Error: The average probability of error is given by

$$\begin{aligned} P_e^n &= \sum_{s^n\in\mathcal{S}^n,\,q^n\in\mathcal{Q}^n} p(s^n)p(q^n)\Pr[\text{error}|s^n,q^n] \\ &\le \sum_{s^n\notin T_\epsilon^n[S]} p(s^n) + \sum_{q^n\notin T_\epsilon^n[Q]} p(q^n) \\ &\quad + \sum_{s^n\in T_\epsilon^n[S],\,q^n\in T_\epsilon^n[Q]} p(s^n)p(q^n)\Pr[\text{error}|s^n,q^n]. \end{aligned} \quad (6)$$

The first term, $\Pr[s^n \notin T_\epsilon^n[S]]$, and the second term, $\Pr[q^n \notin T_\epsilon^n[Q]]$, in the right hand side of (6) go to zero as $n \to \infty$ by the strong asymptotic equipartition property (AEP) [17].

Without loss of generality, we can assume that $(W_1,W_2) = (1,1)$ is sent, the time sharing sequence is $Q^n = q^n$, and the channel state realization is $S^n = s^n$. Then $\Pr[\text{error}|s^n,q^n]$ is the conditional probability of error given $(W_1,W_2) = (1,1)$, $Q^n = q^n \in T_\epsilon^n[Q]$, and $S^n = s^n \in T_\epsilon^n[S]$.

- Let E_1 be the event that there is no sequence $U^n(q^n,1,1,j)$ such that $U^n(q^n,1,1,j) \in T_\epsilon^n[Q,U,S,X_2|q^n,s^n,X_2^n(q^n,1)]$. Since $U^n(q^n,1,1,j)$ and $S^n = s^n$ are generated independently according to $\prod p(u_i|q_i,x_{2,i}(q^n,1))$ and $\prod p(s_i)$, respectively; and there are J sequences in each bin, the probability of event E_1 goes to zero as $n \to \infty$.
  Under the event E_1^c, we can also assume that a particular sequence $U^n(q^n,1,1,1)$ in bin $(1,1)$ is jointly strongly typical with $(s^n,q^n,X_2^n(q^n,1))$. Thus, X_1^n corresponding to $(U^n(q^n,1,1,1),s^n,q^n,X_2^n(q^n,1))$ and $X_2^n(q^n,1)$ are sent from the informed and the uninformed encoders, respectively.
- Let E_2 be the event that $(U^n(q^n,1,1,1),X_2^n(q^n,1),Y^n) \notin T_\epsilon^n[Q,U,X_2,Y|q^n]$. According to the Markov lemma [17], $\Pr[E_2|E_1^c] \to 0$ as $n \to \infty$. Let E_3 be the event that $U^n(q^n,m_1,1,j) \in T_\epsilon^n[Q,U,X_2,Y|q^n,Y^n,X_2^n(q^n,1)]$ for $m_1 \neq 1$ and $j \neq 1$. Using properties of strongly typical sequences [17], it can be easily shown that $\Pr[E_3|E_1^c,E_2^c] \to 0$ as $n \to \infty$ if $R_1 < \mathbb{I}(U;Y|Q,X_2) - \mathbb{I}(U;S|Q,X_2)$.
- Finally, let E_4 be the event that $(U_1^n(q^n,m_1,m_2,j),X_2^n(q^n,m_2)) \in T_\epsilon^n[Q,U,X_2,Y|q^n,Y^n]$ for $m_1 \neq 1$, $1 \le j \le J$, and $m_2 \neq 1$. Using the properties of strongly typical sequences, it can be shown that the $\Pr[E_5|E_1^c,E_2^c] \to 0$ as $n \to \infty$ if $R_2 + R_2 < \mathbb{I}(U,X_2;Y|Q) - \mathbb{I}(U;S|Q,X_2)$.

In terms of these events, $\Pr[\text{error}|s^n,q^n]$ in (6) can be upper-bounded via the union bound, and the fact that probabilities are less than one, as

$$\begin{aligned} \Pr[\text{error}|s^n,q^n] \le &\Pr[E_1] + \Pr[E_2|E_1^c] + \Pr[E_3|E_1^c,E_2^c] \\ &+ \Pr[E_4|E_1^c,E_2^c]. \end{aligned} \quad (7)$$

From (7) and the above limiting arguments, we have $\Pr[\text{error}|s^n,q^n] \to 0$ as $n \to \infty$. Therefore, the average probability of error $P_e^n \to 0$ as $n \to \infty$, completing the proof.

IV. CONCLUSIONS

We consider a state-dependent MAC with state known at one encoder. In general the capacity region is not known for this model when the informed encoder and the uninformed encoder are provided with independent messages. In this paper, we assume that the informed encoder is aware of the message of the uninformed encoder, i.e., a case of degraded message sets, and obtain the capacity region.

ACKNOWLEDGMENTS

This work has been supported in part by the National Science Foundation through grant CCF05-46618 and by a Graduate Fellowship from the Center for Applied Mathematics, University of Notre Dame.

REFERENCES

[1] C. Shannon, “Channels with Side Information at the transmitter,” *IBM J. Res. Devel.*, vol. vol.2, pp. 289–293, 1958.

[2] A. V. Kusnetsov and B. S. Tsybakov, “Coding in a Memory with Defective Cells,” *Probl. Peredach. Inform.*, vol. vol.10, no. 2, pp. 52–60, April/June 1974.

[3] C. Heegard and A. E. Gamal, “On the Capacities of Computer Memories with Defects,” *IEEE Trans. Inform. Theory*, vol. vol.IT-29, pp. 731–739, September 1983.

[4] S. I. Gel’fand and M. S. Pinsker, “Coding for Channel with Random Parameters,” *Probl.Contr. and Information Theory*, vol. 9, no. 1, pp. pp.19–31, 1980.

[5] M. H. M. Costa, “Writing on Dirty Paper,” *IEEE Trans. Inform. Theory*, vol. vol.IT-29, pp. 439–441, May 1983.

[6] S. I. Gel’fand and M. S. Pinsker, “On Gaussian Channels with Random Parameters,” in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, 1983.

[7] Y. H. Kim, A. Sutivong, and S. Sigurjónsson, “Multiple User Writing on Dirty Paper,” in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, June 27 - July 2 2004.

[8] A. Khisti, U. Erez, and G. W. Wornell, “Writing on Many Pieces of Dirty Paper at Once: The Binary Case,” in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, June 27 – July 2 2004.

[9] Y. Steinberg, “Coding for the Degraded Broadcast Channel with Random parameters, with Causal and Noncausal Side Information,” *IEEE Trans. Inform. Theory*, vol. vol.51, pp. 2867–2877, August 2005.

[10] Y. Cemal and Y. Steinberg, “Multiple Access Channel with Partial State Information at the Encoders,” *IEEE Trans. Inform. Theory*, vol. vol.IT-51, pp. 3992–4003, November 2005.

[11] S. A. Jafar, “Capacity with Causal and Non-Causal Side Information - A Unified View,” *IEEE Trans. Inform. Theory*, vol. 52, no. 12, pp. 5468–5475, Dec. 2006.

[12] S. Kotagiri and J. N. Laneman, “Achievable Rates for Multiple Access Channels with State Information Known at One Encoder,” in *Proc. Allerton Conf. Communications, Control, and Computing*, 2004.

[13] ——, “Multiple Access Channels with State Information Known at Some Encoders,” July 2006, submitted to ”*IEEE Trans. Inform. Theory*”.

[14] F. M. J. Willems and E. C. D. Meulen, “The Discrete Memoryless Multiple Access Channel with Cribbing Encoders,” *IEEE Trans. Inform. Theory*, vol. 31, pp. 313–327, May 1985.

[15] A. Somekh-Baruch, S. Shamai, and S. Verdu, “Cooperative Multiple Access Encoding with States Available at One Transmitter,” in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, 2007.

[16] ——, “Cooperative Encoding with Asymmetric State Information at the Transmitters,” in *Proc. Allerton Conf. Communications, Control, and Computing*, 2006.

[17] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: John Wiley & Sons, Inc., 1991.

[18] I. Csiszár and J. Korner, Eds., *Information Theory: Coding Thoerems for Discrete Memoryless Systems*. New York: Academic Press Inc., 1981.