# New Bounds for the Capacity Region of the Finite-State Multiple Access Channel

Haim Permuter*, Tsachy Weissman*† and Jun Chen‡,
* Department of Electrical Engineering, Stanford University, Stanford, CA, USA, {haim1, tsachy}@stanford.edu
† Department of Electrical Engineering, Technion, Haifa, Israel
‡Department of Electrical and Computer Engineering, McMaster University, Hamilton, Ontario, Canada, junchen@ece.mcmaster.ca

***Abstract*— The capacity region of the Finite-State Multiple Access Channel (FS-MAC) with feedback that may be an arbitrary time-invariant function of the channel output samples is considered. We provided a sequence of inner and outer bounds for this region. These bounds are shown to coincide, and hence yield the capacity region, of FS-MACs where the state process is stationary and ergodic and not affected by the inputs, and for indecomposable FS-MAC when feedback is not allowed. Though the capacity region is 'multi-letter' in general, our results yield explicit conclusions when applied to specific scenarios of interest.**

## I. Introduction

The Multiple Access Channel (MAC) has received much attention in the literature. To put our contributions in context, we begin by briefly describing some of the key results in the area. The capacity region for the memoryless MAC was derived by Ahlswede in [1]. Cover and Leung derived an achievable region for a memoryless MAC with feedback in [2]. Ozarow derived the capacity of a memoryless Gaussian MAC with feedback in [3]

In [4] [5], Kramer derived several capacity results for discrete memoryless networks with feedback. By using the idea of code-trees instead of code-words, Kramer derived a 'mulit-letter' expression for the capacity of the discrete memoryless MAC. One of the main results we develop in the present paper extends Kramer's capacity result to the case of a stationary and ergodic Markov Finite-State MAC (FS-MAC), to be formally defined below.

In [6], Han used the information-spectrum method in order to derive the capacity of a general MAC without feedback. Han also considered the additive mod-2 MAC, which we shall use here to illustrate the way in which our general results characterize special cases of interest. In particular, our results will imply that feedback does not increase the capacity region of the additive mod-2 MAC.

In this work, we consider the capacity region of the Finite-State Multiple Access Channel (FS-MAC), with feedback that may be an arbitrary time-invariant function of the channel output samples. We characterize a sequence of inner and outer bounds for this region and show that it yields the capacity region, for the important subfamily of FS-MACs.

Our derivation of the capacity region is rooted in the derivation of the capacity of finite-state channels in Gallager's book [7, ch 4,5]. More recently, Lapidoth and Telatar [8] have used it in order to derive the capacity of a compound channel without feedback, where the compound channel consists of a family of finite-state channels. In particular, they have introduced into Gallager's proof the idea of concatenating codewords, which we extend here to concatenating code-trees.

The paper is organized as follows. We concretely describe the communication model in Section II. In Section III, we introduce the causal conditioning, directed information and an important idea of sup/sub-additivity of regions. We state our main capacity results in Sections IV and V, and we present a few applications of the capacity results in Section VI. Because of space limitation we do not provide the proofs. The proofs, with the exceptions of Lemmas 3 and 5, and Theorem 4, can be found in the preprint [9].

## II. Channel Model

In this paper, we consider an FS-MAC (Finite State MAC) with a time invariant feedback as illustrated in Fig. 1. The

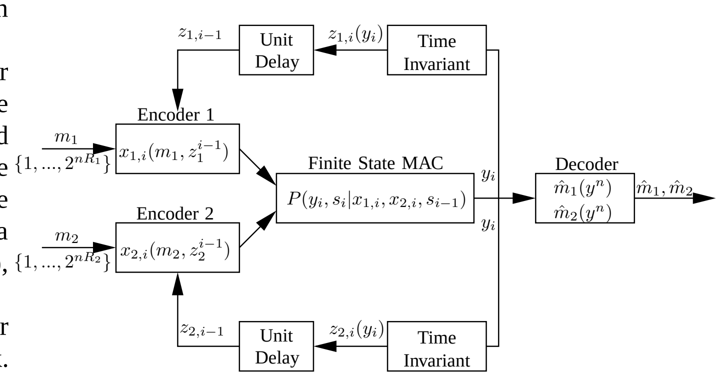


Fig. 1. Channel with feedback that is a time invariant deterministic function of the output.

MAC setting consists of two senders and one receiver. Each sender $l \in \{1, 2\}$ chooses an index $m_l$ uniformly from the set $\{1, ..., 2^{nR_l}\}$ and independently of the other sender. The input to the channel from encoder $l$ is denoted by $\{X_{l1}, X_{l2}, X_{l3}, ...\}$, and the output of the channel is denoted by $\{Y_1, Y_2, Y_3, ...\}$. The state at time $i$, i.e., $S_i \in \mathcal{S}$, takes values in a finite set of possible states. The channel is stationary and is characterized by a conditional probability $P(y_i, s_i|x_{1i}, x_{2i}, s_{i-1})$ that satisfies

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

where the superscripts denote sequences in the following way: $x_l^i = (x_{l1}, x_{l2}, ..., x_{li}), \quad l \in \{1, 2\}$. We assume a communication with feedback $z_l^i$, where the element $z_{li}$ is a time-invariant function of the output $y_i$. For example, $z_{li}$ could equal $y_i$ (perfect feedback), or a quantized version of $y_i$, or null (no feedback). The encoders receive the feedback samples with one unit delay.

A code with feedback consists of two encoding functions $g_l : \{1, ..., 2^{nR_1}\} \times \mathcal{Z}_l^{n-1} \to \mathcal{X}_l^n,\ l = 1, 2$, where the $k$th coordinate of $x_l^n \in \mathcal{X}_l^n$ is given by the function

$$x_{lk} = g_{lk}(m_l, z_l^{k-1}), \qquad k = 1, 2, \dots, n, \quad l = 1, 2 \tag{2}$$

and a decoding function,

$$g : \mathcal{Y}^n \to \{1, ..., 2^{nR_1}\} \times \{1, ..., 2^{nR_2}\}. \tag{3}$$

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

$$P_e^{(n)} = \frac{1}{2^{n(R_1+R_2)}} \sum_{w_1, w_2} \Pr\{g(Y^n) \neq (w_1, w_2) | (w_1, w_2) \text{ sent}\}. \tag{4}$$

A rate $(R_1, R_2)$ is said to be *achievable* for the MAC if there exists a sequence of $((2^{nR_1}, 2^{nR_2}), n)$ codes with $P_e^{(n)} \to 0$. The *capacity region* of MAC is the closure of the set of achievable $(R_1, R_2)$ rates.

## III. Preliminaries

### A. Causal conditioning and directed information

Throughout this paper we use the *causal conditioning* notation $(\cdot||\cdot)$. We denote the probability mass function (pmf) of $Y^N$ causally conditioned on $X^{N-d}$, for some integer $d \geq 0$, as $P(y^N || x^{N-d})$ which is defined as

$$P(y^N || x^{N-d}) \triangleq \prod_{i=1}^{N} P(y_i | y^{i-1}, x^{i-d}), \tag{5}$$

(if $i - d \leq 0$ then $x^{i-d}$ is set to null). In particular, we extensively use the cases where $d = 0, 1$:

$$P(y^N || x^N) \triangleq \prod_{i=1}^{N} P(y_i | y^{i-1}, x^i) \tag{6}$$

$$Q(x^N || y^{N-1}) \triangleq \prod_{i=1}^{N} Q(x_i | x^{i-1}, y^{i-1}), \tag{7}$$

where the letters $Q$ and $P$ are both used for denoting pmfs.

The *Directed information* was defined by Massey in [10] as

$$I(X^N \to Y^N) \triangleq \sum_{i=1}^{N} I(X^i; Y_i | Y^{i-1}), \tag{8}$$

and in [4], Kramer introduced the notation

$$I(X_1^N \to Y^N || X_2^N) \triangleq \sum_{i=1}^{N} I(X_1^i; Y_i | Y^{i-1}, X_2^i). \tag{9}$$

Directed inofrmation has been widely used in the characterization of capacity of channels [4], [11]–[16], and rate distortion function [17], [18]. Throughout the proofs, we are using several properties of causal conditioning and directed information. We summarize them in the following lemma.

**Lemma 1** The following four properties ,(10)-(13), hold for any discrete random vectors $(X_1^N, X_2^N, Y^N)$,

$$P(x_1^N, y^N || x_2^N) = P(x_1^N || y^{N-1}, x_2^N) P(y^N || x_1^N, x_2^N). \tag{10}$$

$$|I(X_1^N \to Y^N || X_2^N) - I(X_1^N \to Y^N || X_2^N, S)| \leq H(S). \tag{11}$$

$$\mathcal{I}(Q(x_1^N || y^{N-1}), Q(x_2^N || y^{N-1}), P(y^N || x_1^N, x_2^N)) = I(X_1^N \to Y^N || X_2^N), \tag{12}$$

where $\mathcal{I}(Q(x_1), Q(x_2), P(y|x_1, x_2))$ denotes the functional $I(X_1; Y | X_2)$, i.e.,

$$\mathcal{I}(Q(x_1), Q(x_2), P(y|x_1, x_2)) \triangleq \sum_{y, x_1, x_2} Q(x_1) Q(x_2) P(y|x_1, x_2) \log \frac{P(y|x_1, x_2)}{\sum_{x_1'} Q(x_1') P(y|x_1', x_2)}.$$

If there is no feedback, i.e., $Q(x_1^N, x_2^N || y^{N-1}) = Q(x_1^N) Q(x_2^N)$, then

$$I(X_1^N; Y^N | X_2^N) = I(X_1^N \to Y^N || X_2^N). \tag{13}$$

### B. Sup/Sub-additivity, and Convergence of 2D regions

In this subsection we define basic operations (summation and multiplication by scalar), convergence, sup-additivity and sub-additivity of 2D regions. Furthermore we show that the limit of a sup-additive sequence of regions converges to the union of all the regions, and the limit of a sub-additive and convex sequence 2D regions converges to the intersection of all the regions.

Let $\mathcal{A}, \mathcal{B}$ be sets in $\mathbb{R}^2$, i.e., $\mathcal{A}$ and $\mathcal{B}$ are sets of 2D vectors. The sum of two regions is denoted as $\mathcal{A} + \mathcal{B}$ and defined as

$$\mathcal{A} + \mathcal{B} = \{\mathbf{a} + \mathbf{b} :\ \mathbf{a} \in \mathcal{A}, \mathbf{b} \in \mathcal{B}\}, \tag{14}$$

and multiplication of a set $A$ with a scalar $c$ is defined as

$$c\mathcal{A} = \{c\mathbf{a} :\ \mathbf{a} \in \mathcal{A}\}. \tag{15}$$

A sequence $\{\mathcal{A}_n\},\ n = 1, 2, 3, ...$, of 2D regions is said to *converge* to a region $\mathcal{A}$, written $\mathcal{A} = \lim \mathcal{A}_n$ if

$$\limsup \mathcal{A}_n = \liminf \mathcal{A}_n = \mathcal{A}, \tag{16}$$

where

$$\begin{aligned} \liminf \mathcal{A}_n &= \{\mathbf{a} : \mathbf{a} = \lim \mathbf{a}_n, \mathbf{a}_n \in \mathcal{A}_n\}, \\ \limsup \mathcal{A}_n &= \{\mathbf{a} : \mathbf{a} = \lim \mathbf{a}_k, \mathbf{a}_k \in \mathcal{A}_{n_k}\}, \end{aligned} \tag{17}$$

and $n_k$ denotes an arbitrary increasing subsequence of the integers. Let us denote $\overline{\mathcal{A}} = \text{cl}\left(\bigcup_{n\geq 1} \mathcal{A}_n\right)$ and $\underline{\mathcal{A}} = \text{cl}\left(\bigcap_{n\geq 1} \mathcal{A}_n\right)$.

We say that a sequence $\{\mathcal{A}_n\}_{n\geq 1}$ is *bounded* if $\sup\{||\mathbf{a}|| : \mathbf{a} \in \overline{\mathcal{A}}\} < \infty$ where $||\cdot||$ denotes a norm in $\mathbb{R}^2$.

**Lemma 2** Let $\mathcal{A}_n,\ n = 1, 2, ...$, be a bounded sequence of sets in $\mathbb{R}^2$ that includes the origin, i.e., $(0, 0)$. If $n\mathcal{A}_n$ is sup-additive, i.e., for all $n \geq 1$ and all $N > n$

$$N\mathcal{A}_N \supseteq n\mathcal{A}_n + (N - n)\mathcal{A}_{N-n} \tag{18}$$

then

$$\lim_{n\to\infty} \mathcal{A}_n = \overline{\mathcal{A}}. \tag{19}$$

**Lemma 3** Let $\mathcal{A}_n,\ n = 1, 2, ...,$ be a sequence of convex, closed and bounded sets in $\mathbb{R}^2$. If $n\mathcal{A}_n$ is sub-additive, i.e., for all $n \geq 1$ and all $N > n$

$$N\mathcal{A}_N \subseteq n\mathcal{A}_n + (N-n)\mathcal{A}_{N-n} \tag{20}$$

then

$$\lim_{n\to\infty} \mathcal{A}_n = \underline{\mathcal{A}}. \tag{21}$$

## IV. FS-MAC WITH TIME-INVARIANT FEEDBACK

### A. Inner Bound

Let $\underline{\mathcal{R}}_n$ denote the following region in $\mathbb{R}_+^2$ (2D set of nonnegative real numbers):

$$\underline{\mathcal{R}}_n = \bigcup_Q \begin{cases} R_1 \leq \min_{s_0} \frac{1}{n} I(X_1^n \to Y^n || X_2^n, W, s_0) - \frac{\log|\mathcal{S}|}{n}, \\ R_2 \leq \min_{s_0} \frac{1}{n} I(X_2^n \to Y^n || X_1^n, W, s_0) - \frac{\log|\mathcal{S}|}{n}, \\ R_1 + R_2 \leq \min_{s_0} \frac{1}{n} I((X_1,X_2)^n \to Y^n | W, s_0) - \frac{\log|\mathcal{S}|}{n}, \end{cases} \tag{22}$$

where the union is over the set of all input distributions of the form $Q(w)Q(x_1^n||z_1^{n-1}, w)Q(x_2^n||z_2^{n-1}, w)$. Having the auxiliary random variable $W$ is equivalent to taking the convex hull of the region. Furthermore, the set of three-inequalities is equivalent to an intersection of three regions, and $\min_{s_0}$ is equivalent to $\cap_{s_0}$. Hence, an equivalent region is

$$\underline{\mathcal{R}}_n = \text{conv} \bigcup_Q \bigcap_{s_0} \begin{cases} R_1 \leq \frac{1}{n} I(X_1^n \to Y^n || X_2^n, s_0) - \frac{\log|\mathcal{S}|}{n}, \\ R_2 \leq \frac{1}{n} I(X_2^n \to Y^n || X_1^n, s_0) - \frac{\log|\mathcal{S}|}{n}, \\ R_1 + R_2 \leq \frac{1}{n} I((X_1,X_2)^n \to Y^n | s_0) - \frac{\log|\mathcal{S}|}{n}, \end{cases} \tag{23}$$

where conv denotes the convex hull, and the input distribution is of the form $Q(x_1^n||z_1^{n-1})Q(x_2^n||z_2^{n-1})$. In general, the right hand side (RHS) of each of the three inequalities that define $\mathcal{R}_n$ can be negative. In such a case, we assume that the RHS is zero.

**Theorem 1** (*Inner bound.*) For any FS-MAC with time invariant feedback as shown in Fig. 1, and for any integer $n \geq 1$, the region $\underline{\mathcal{R}}_n$ is achievable.

The proof is similar to the point-to-point FSC with time-invariant feedback, given in [14, Sec. V]. In the proof we use Gallager's techniques to analyze the error probability of a ML decoder of a randomly-generated code. There are two main differences compared to the point-to-point FSC:

1) In the case of FSC, only one message is sent, and in the case of FS-MAC, two independent messages are sent. This requires that we analyze three different types of errors, and they yield three inequalities in the achievable region.
2) For the FS-MAC case, we need to prove the achivebility for a set of input distributions while for the point-to-point channel it was enough to prove it only for the input distribution that achieves the maximum. Because of this difference, we introduce the idea of concatenating code-trees (see Fig. 2). This difference influences the encoding scheme and the analysis.

The following lemma establishes the sub-additivity of $\{\underline{\mathcal{R}}_n\}$.

**Lemma 4** (*sup-additivity of $\underline{\mathcal{R}}_n$.* ) For any FS-MAC, the sequence $\{\underline{\mathcal{R}}_n\}$ is sup-additive. Therefore, $\lim_{n\to\infty} \underline{\mathcal{R}}_n$ exists, it is an achievable region, and it equals to $\overline{\mathcal{R}}$.

### B. Outer Bound

The following outer bound is proved using Fano's inequality.

**Theorem 2** (*Outer bound.*) Let $(R_1, R_2)$ be an achievable pair for a FS-MAC with time invariant feedback, as shown in Fig. 1. Then, for any $n$ there exists a distribution $Q(x_1^n||z_1^{n-1})Q(x_2^n||z_2^{n-1})$ such that the following inequalities hold:

$$\begin{aligned} R_1 &\leq \frac{1}{n} I(X_1^n \to Y^n || X_2^n) + \epsilon_n, \\ R_2 &\leq \frac{1}{n} I(X_2^n \to Y^n || X_1^n) + \epsilon_n, \\ R_1 + R_2 &\leq \frac{1}{n} I((X_1,X_2)^n \to Y^n) + \epsilon_n, \end{aligned} \tag{24}$$

where $\epsilon_n$ goes to zero as $n$ goes to infinity.

This theorem implies that $\liminf \mathcal{R}_n$ is an outer bound, where $\mathcal{R}_n$ is defined as

$$\mathcal{R}_n = \bigcup_Q \begin{cases} R_1 \leq \frac{1}{n} I(X_1^n \to Y^n || X_2^n), \\ R_1 \leq \frac{1}{n} I(X_2^n \to Y^n || X_1^n), \\ R_1 + R_2 \leq \frac{1}{n} I((X_1,X_2)^n \to Y^n), \end{cases} \tag{25}$$

and the union is over input distributions of the form $Q(x_1^n||z_1^{n-1})Q(x_2^n||z_2^{n-1})$.

### C. Capacity

Based on the bounds above, we have the following capacity result.

**Theorem 3** For any FS-MAC of the form

$$P(y_i, s_i | x_{1,i}, x_{2,i}, s_{i-1}) = P(s_i|s_{i-1}) P(y_i | x_{1,i}, x_{2,i}, s_{i-1}), \tag{26}$$

where the state process $S_i$ is stationary and ergodic, the achievable region is $\lim_{n\to\infty} \mathcal{R}_n$, and the limit exists.

## V. FS-MAC WITHOUT FEEDBACK

The case where there is no feedback is a special case of deterministic time-invariant feedback in which $z_i$ is null, and therefore the theorems in the previous section hold for the case of no feedback. Here we show additional results, which apply only for the case without feedback. The results include a sequence of upper bounds for all FS-MACs, and a capacity formula for indecomposable FS-MACs.

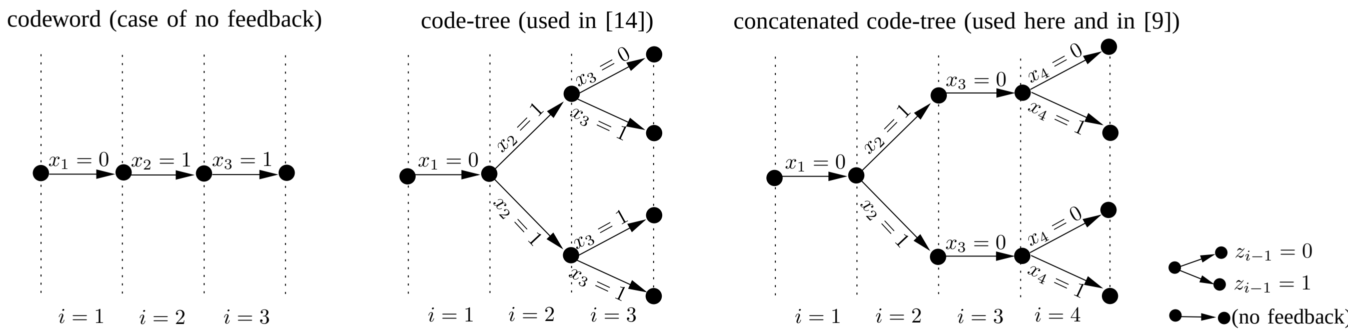


Fig. 2. Illustration of coding scheme for setting without feedback, setting with feedback as used for point-to-point channel [14] and a code-tree that was created by concatenating smaller code-trees. In the case of no feedback each message is mapped to a codeword, and in the case of feedback each message is mapped to a code-tree. The third scheme is a code-tree of depth 4 created by concatenating two trees of depth 2.

### A. Outer bound

Let us denote,

$$\overline{\mathcal{R}}_n = \text{conv}\bigcup_Q \begin{cases} R_1 \le \max_P \frac{1}{n} I(X_1^n \to Y^n || X_2^n, S_0) + \frac{H(S_0)}{n}, \\ R_2 \le \max_P \frac{1}{n} I(X_2^n \to Y^n || X_1^n, S_0) + \frac{H(S_0)}{n}, \\ R_1 + R_2 \le \max_P \frac{1}{n} I((X_1, X_2)^n \to Y^n | S_0) + \frac{H(S_0)}{n} \end{cases}$$

where the union is over all input distributions of the form $Q(x_1^n)Q(x_2^n)$, and $\max_P$ denote a maximization over distribution of the form $P(s_0|x_1^n, x_2^n)$. The sup-additivity of $\{\overline{\mathcal{R}}_n\}$ is the key property for establishing the outer bound.

**Lemma 5** (*sub-additivity of* $\overline{\mathcal{R}}_n$. ) For any FS-MAC, the sequence $\{\overline{\mathcal{R}}_n\}$ is sub-additive, i.e.,

$$(n+l)\overline{\mathcal{R}}_{n+l} \subseteq n\overline{\mathcal{R}}_n + l\overline{\mathcal{R}}_l. \tag{27}$$

**Theorem 4** (*Outer bound*) For any FS-MAC and all $n \ge 1$, $\overline{\mathcal{R}}_n$ contains the capacity region.

### B. Capacity

**Theorem 5** (*Capacity of FS-MAC without feedback.*) For any indecomposable FS-MAC without feedback,

$$\lim_{n\to\infty} \underline{\mathcal{R}}_n = \lim_{n\to\infty} \overline{\mathcal{R}}_n, \tag{28}$$

and therefore its capacity region is $\lim_{n\to\infty} \mathcal{R}_n$, and the limit exists.

## VI. Applications

In this section we use the capacity results in order to derive the following conclusions:

1) For a stationary ergodic Markovian channels, the capacity region is zero, if and only if the capacity region with feedback is zero.
2) For the additive mod-$|\mathcal{X}|$ MAC, where the noise may have memory:
    a) feedback does not enlarge the capacity;
    b) source-channel coding separation holds for lossless reconstruction.

### A. Zero capacity

The first conclusion is given in Theorem 6.

**Theorem 6** For the channel described in (26), where the state process $S_i$ is stationary and ergodic, if the capacity without feedback is zero, then it is also zero in the case that there is feedback.

The proof of Theorem 6 is based on the fact that for any MAC

$$\max_{Q(x_1^n||y^{n-1})Q(x_2^n||y^{n-1})} I(X_1^n, X_2^n \to Y^n) = 0 \tag{29}$$

if and only if

$$\max_{Q(x_1^n)Q(x_2^n)} I(X_1^n, X_2^n \to Y^n) = 0, \tag{30}$$

and on the fact that for the family of channels that is mentioned in the theorem, the sequence $\mathcal{R}_n$ is sup-additive.

For the case of additive Gaussian MAC, one can deduce the result from the fact that feedback can at most double its capacity region [19]. Clearly, Theorem 6 also holds for the case of a stationary and ergodic FS-Markov point-to-point channel because a MAC is an extension of a point-to-point channel. However, it does not hold for the case of a broadcast channel.

### B. Additive mod-$|\mathcal{X}|$ MAC

In this section we consider the additive mod-$|\mathcal{X}|$ MAC with and without feedback. The channel is described in Fig. 3. In the binary case, the channel is simply $Y = X_{1,i} \oplus X_{2,i} \oplus V_i$, where $V_i$ is the binary noise, possibly with memory, and $\oplus$ denotes addition mod-2.

The following theorem is an extension of of Alajaj's result [20] to the additive MAC .

**Theorem 7** Feedback does not enlarge the capacity region of a discrete additive (mod-$|\mathcal{X}|$) noise MAC. Moreover, the capacity region is given by

$$\sum_{m=1}^{M} R_m \le \log q - H(\mathcal{V}), \tag{31}$$

where $H(\mathcal{V})$ is the entropy rate of the additive noise.

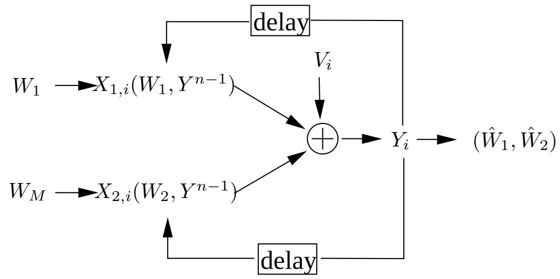


Fig. 3. Additive noise MAC with feedback. The random variables $X_{1,i}, X_{2,i}, Y_i, V_i,\ i \in \mathbb{Z}^+$, are from a common alphabet, and they denote the input from sender 1,2, the output and the noise at time $i$, respectively. The output satisfies $y_i = x_{1,i} \oplus x_{2,i} \oplus v_i$ where $\oplus$ denotes addition mod-$|\mathcal{X}|$. The noise $V_i$, possibly with memory, is independent of the messages $W_1, W_2$.

The theorem can be shown to hold for a larger family of MACs. The family includes all the MACs that can be represented as multiplexer followed by a point-to-point channel. The main idea of the proof is that all three inequalities that defines the region $\mathcal{R}_n$ are maximized by uniform and i.i.d distribution, even if feedback is allowed.

### *C. Source-channel coding separation*

Cover, El-Gamal and Salehi [21] showed that, in general, the source channel coding separation does not hold for MACs even for a memoryless channel without feedback. However, for the case where the MAC is a additive mod-$|\mathcal{X}|$, and the goal is to reconstruct the sources losslessly, then it does hold.

**Theorem 8** (*Source-channel coding theorem for a additive mod-$|\mathcal{X}|$ MAC.*) Let $(U_1, U_2)_{n\geq 1}$ be a finite alphabet, jointly stationary and ergodic pair of processes, and let the MAC channel be an additive mod-$|\mathcal{X}|$ MAC with stationary and ergodic noise. Define $P_e^{(n)} \triangleq \Pr((\hat{U}_1^n, \hat{U}_2^n) \neq (U_1^n, U_2^n))$, where $\hat{U}_1^n, \hat{U}_2^n$ are the reconstructed sources at the decoder.

(*direct part.)* There exists a sequence of source-channel codes with $P_e^{(n)} \to 0$, if $H(\mathcal{U}_1, \mathcal{U}_2) < \log|\mathcal{X}| - H(\mathcal{V})$, where $H(\mathcal{U}_1, \mathcal{U}_2)$ is the entropy rate of the sources and $H(\mathcal{V})$ is the entropy rate of the noise.

(*converse part.)* If $H(\mathcal{U}_1, \mathcal{U}_2) > \log|\mathcal{X}| - H(\mathcal{V})$, then the probability of error is bounded away from zero, independent of the blocklength.

It is interesting to notice that, even though the source-channel coding separation theorem holds when the reconstruction of the sources has to be lossless, the theorem does not hold when distortion is allowed. Such an example was shown by Nazar and Gastpar [22].

## VII. Conclusions and Future Directions

In this paper we have shown that directed information and causal conditioning emerge naturally in characterizing the capacity region of FS-MACs in the presence of a time-invariant feedback. We provided a sequence of inner and outer bounds, and for some large families of channels we characterize the capacity region in terms of a 'multi-letter' expression, which is a first step toward deriving useful concepts in communication. For instance, we use this characterization to show that for a stationary and ergodic Markovian channel, the capacity is zero if and only if the capacity with feedback is zero. Further, we identify FS-MACs for which feedback does not enlarge the capacity region and for which source-channel separation holds.

One future direction is to use the characterizations developed in this paper to explicitly compute the capacity regions of classes of MACs with memory and feedback (other than the additive mod-$|\mathcal{X}|$ channel), and to find optimal coding schemes.

## References

[1] R. Ahlswede, "Multi-way communication channels," in *Proceedings of 2nd International Symposium on Information Theory (Thakadsor, Armenian SSR, Sept. 1971)*, Publishing House of the Hungarian Academy of Science, Budapest, 1973, pp. 23–52.

[2] T. M. Cover and C. S. K. Leung, "An achievable rate region for the multiple-access channel with feedback," *IEEE Trans. on Info. Theory*, vol. 27, no. 3, pp. 292–298, 1981.

[3] L. H. Ozarow, "The capacity of the white Gaussian multiple access channel with feedback," *IEEE Transactions on Information Theory*, vol. 30, no. 4, pp. 623–628, 1984.

[4] G. Kramer, "Directed information for channels with feedback," Ph.D. Dissertation, Swiss Federal Institute of Technology Zurich, 1998.

[5] ——, "Capacity results for the discrete memoryless network," *IEEE Trans. Inform. Theory*, vol. 49, pp. 4–21, 2003.

[6] T. S. Han, "An information-spectrum approach to capacity theorems for the general multiple-access channel," *IEEE Trans. Inform. Theory*, vol. 44, no. 7, pp. 2773–2795, 1998.

[7] R. G. Gallager, *Information theory and reliable communication*. New York: Wiley, 1968.

[8] A. Lapidoth and I. Telatar, "The compound channel capacity of a class of finite-state channels," *IEEE Trans. Inform. Theory*, vol. 44, pp. 973–983, 1998.

[9] H. H. Permuter and T. Weissman, "Capacity region of the finite-state multiple access channel with and without feedback," Aug. 2007, submitted to IEEE Trans. Inform. Theory. Availble at arxiv.org/pdf/cs.IT/0608070.

[10] J. Massey, "Causality, feedback and directed information," *Proc. Int. Symp. Information Theory Application (ISITA-90)*, pp. 303–305, 1990.

[11] S. Tatikonda, "Control under communication constraints," *Ph.D. disertation, MIT, Cambridge, MA*, 2000.

[12] J. Chen and T. Berger, "The capacity of finite-state Markov channels with feedback," *IEEE Trans. on Information theory*, vol. 51, pp. 780–789, 2005.

[13] S. Yang, A. Kavcic, and S. Tatikonda, "Feedback capacity of finite-state machine channels," *IEEE Trans. Inform. Theory*, pp. 799–810, 2005.

[14] H. H. Permuter, T. Weissman, and A. J. Goldsmith, "Finite state channels with time-invariant deterministic feedback," Sep 2006, submitted to IEEE Trans. Inform. Theory. Availble at arxiv.org/pdf/cs.IT/0608070.

[15] Y. Kim, "A coding theorem for a class of stationary channels with feedback," Jan 2007, submitted to IEEE Trans. Inform. Theory. Availble at arxiv.org/cs.IT/0701041.

[16] B. Shrader and H. H. Permuter, "On the compound finite state channel with feedback," in *ISIT 2007*, Nice, France, 2007.

[17] R. Venkataramanan and S. S. Pradhan, "Source coding with feedforward: Rate-distortion function for general sources," in *IEEE Information theory workshop (ITW)*, 2004.

[18] R. Zamir, Y. Kochman, and U. Erez, "Achieving the gaussian rate-distortion function by prediction," July 2006, submitted for publication in "IEEE Trans. Inform. Theory". [Online]. Available: http://www.eng.tau.ac.il/~zamir/papers/dpcm.pdf

[19] J. A. Thomas, "Feedback can at most double gaussian multiple access channel capacity," *IEEE Trans. Inform. Theory*, vol. 33, pp. 711–716, 1987.

[20] F. Alajaji, "Feedback does not increase the capacity of discrete channels with additive noise," *IEEE Trans. Inform. Theory*, vol. 41, pp. 546–549, 1995.

[21] T. M. Cover, A. E. Gamal, and M. Salehi, "Multiple access channels with arbitrarily correlated sources," *IEEE Trans. Inform. Theory*, vol. 26, pp. 648–657, 1980.

[22] B. Nazer and M. Gastpar, "The case for structured random codes in network communication theorems," in *2007 nformation Theory Workshop (ITW), Lake Tahoe, CA,*, Sep. 2007.