On the capacity of memoryless finite-state multiple access channels with asymmetric noisy state information at the encoders

Sen, Nevroz; Como, Giacomo; Yuksel, Serdar; Alajaji, Fady

On the Capacity of Memoryless Finite-State Multiple Access Channels with Asymmetric Noisy State Information at the Encoders

Nevroz Şen
Queen's University
nsen@mast.queensu.ca

Giacomo Como
M.I.T.
giacomo@mit.edu

Serdar Yüksel
Queen's University
yuksel@mast.queensu.ca

Fady Alajaji
Queen's University
fady@mast.queensu.ca

Abstract—We consider the capacity of memoryless finite-state multiple access channels (FS-MACs) with causal asymmetric noisy state information available at both transmitters and complete state information available at the receiver. Single letter inner and outer bounds are provided for the capacity of such channels when the state process is independent and identically distributed. The outer bound is attained by observing that the proposed inner bound is tight for the sum-rate capacity.

I. INTRODUCTION AND LITERATURE REVIEW

Modeling communication channels with a state process fits well for many physical scenarios. For single-user channels, the characterization of the capacity with various degrees of channel state information at the transmitter (CSIT) and at the receiver (CSIR) is well understood. Among them, Shannon [1] determined the capacity formula when causal noiseless state information is available at the transmitter, where state is independent identically distributed (i.i.d.). The same problem with non-causal side information is considered in [2]. In [3], Shannon's result is extended to the case where noisy state observation is available at both the transmitter and the receiver. Later, in [4] this result has been shown to be a special case of Shannon's model and the authors also determined that when CSIT is a deterministic function of CSIR optimal codes can be constructed directly on the input alphabet.

In the multi-user setting, [5] provides a multi-letter characterization of the capacity region of time-varying multiple access channels (MACs) with various degrees of CSIT and CSIR. In [6], a general framework for the capacity region of MACs with causal and non-causal CSI is presented. In particular, an achievable rate region is presented for memoryless FS-MAC with correlated CSI and the sum-rate capacity is determined under the condition that the state information available to each encoders are independent [6, Theorem 4]. In a related work, MACs where the encoders have degraded information on the channel state, which is coded to the encoders, is considered [7]. In [8], memoryless FS-MACs with two independent states (see also [9] for the single state case), each known causally and strictly causally to one encoder, are considered and an achievable rate region, which is shown to contain an achievable region where each user applies Shannon

strategies, is proposed. In [8] and [9] it is also shown that strictly casual state information does not increase the sum-rate capacity. More recently, in [10] finite-state Markovian MACs with asymmetric delayed state information at the transmitters are studied and their capacity region is determined.

The most relevant work to our paper is [11], which obtained a single letter characterization of the capacity region for memoryless FS-MAC in which transmitters have asymmetric partial quantized state observations and the receiver has full state information. In this work, the authors were inspired from team decision theory [12], [13]. We herein mainly adopt the converse technique presented in [11] and partially extend it to a noisy setup. The present paper, thus, studies the FS-MAC in which each of the transmitters have an asymmetric state information which is corrupted by i.i.d. noise processes and the receiver has complete state information. We provide a single letter inner bound to the capacity region, in terms of Shannon strategies [1]. By observing that this inner bound is tight for the sum-rate capacity, we also provide an outer bound to the channel's capacity region. We modify the approach in [11] to account for the fact that the decoder does not have access to the state information at the encoders, and that the past state information does not lead to a tractable recursion.

The rest of the paper is organized as follows. In Section II we formally state the problem, present inner and outer bounds to the capacity region with the achievability and converse proofs and in Section III we present concluding remarks.

Throughout the paper we will use the following notations. A random variable will be denoted by an upper case letter X and its particular realization by a lower case letter x. For a vector v, and a positive integer i, v_i will denote the i-th entry of v, while $v_{[i]} = (v_1, \cdots, v_i)$ will denote the vector of the first i entries of v. For a finite set $\mathcal{A}$, $\mathcal{P}(\mathcal{A})$ will denote the simplex of probability distributions over $\mathcal{A}$. Probability distributions are denoted by $P(\cdot)$ and subscripted by the name of the random variables and conditioning, e.g., $P_{U,T|V,S}(u,t|v,s)$ is the conditional probability of $(U = u, T = t)$ given $(V = v, S = s)$. Finally, for a positive integer n, we shall denote by $\mathcal{A}^{(n)} := \bigcup_{0 < s < n} \mathcal{A}^s$ the set of $\mathcal{A}$-strings of length smaller than n. We denote the indicator function of an event by $1_{\{E\}}$. All sets considered hereafter are finite.
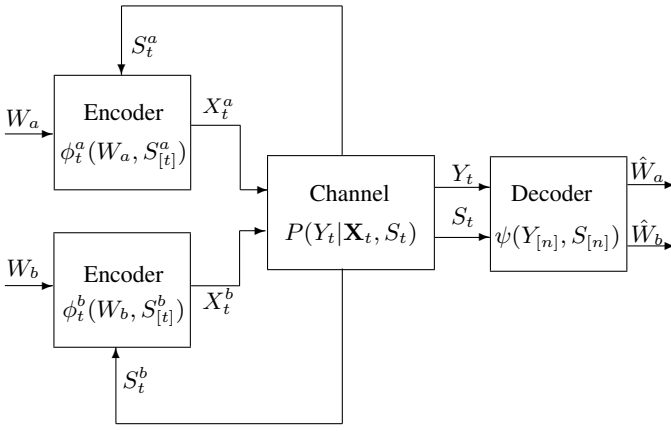
Fig. 1. The multiple-access channel with noisy state feedback.

II. ON THE CAPACITY OF FS-MAC WITH NOISY CSIT AND COMPLETE CSIR

Consider a two-user memoryless FS-MAC, with two encoders, a, b, and two independent message sources W_a and W_b which are uniformly distributed in the sets $W_a \in \{1, 2, \cdots, M_a\}$ and $W_b \in \{1, 2, \cdots, M_b\}$, respectively. The channel inputs of the encoders are X^a and X^b, respectively. The channel state process is modeled as a sequence $\{S_t\}_{t=1}^{\infty}$ of i.i.d. random variables in some space $\mathcal{S}$. The two encoders have access to causal noisy version of the state information S_t at each time $t \geq 1$ modeled by $S_t^a \in \mathcal{S}^a$, $S_t^b \in \mathcal{S}^b$, respectively and the joint distribution of (S_t, S_t^a, S_t^b) satisfies

$$P_{S_t^a, S_t^b, S_t}(s_t^a, s_t^b, s_t) = P_{S_t^a|S_t}(s_t^a|s_t) P_{S_t^b|S_t}(s_t^b|s_t) P_{S_t}(s_t). \quad (1)$$

We also assume that S_t is fully available at the receiver (see Fig. 1) and that (S_t, S_t^a, S_t^b) are independent of (W_a, W_b) $\forall t \geq 1$. The channel inputs at time t, i.e., X_t^a and X_t^b, are functions of the locally available information $(W_a, S_{[t]}^a)$ and $(W_b, S_{[t]}^b)$. Let $\mathbf{W} := (W_a, W_b)$ and $\mathbf{X} := (X^a, X^b)$. Then, the laws governing n-sequences of state, input and output letters are given by

$$P_{Y_{[n]}|\mathbf{W}, \mathbf{X}_{[n]}, S_{[n]}, S_{[n]}^a, S_{[n]}^b}(y_{[n]}|\mathbf{w}, \mathbf{x}_{[n]}, s_{[n]}, s_{[n]}^a, s_{[n]}^b)$$
$$= \prod_{t=1}^{n} P_{Y_t|X_t^a, X_t^b, S_t}(y_t|x_t^a, x_t^b, s_t), \quad (2)$$

where the channel's transition probability distribution, $P_{Y_t|X_t^a, X_t^b, S_t}(y_t|x_t^a, x_t^b, s_t)$, is given a priori.

Definition 1: An $(n, 2^{nR_a}, 2^{nR_b})$ code with block length n and rates (R_a, R_b) for an FS-MAC with noisy state feedback consists of

(1) A sequence of mappings for each encoder
$$\phi_t^{(a)} : (\mathcal{S}^a)^t \times \mathcal{W}_a \to \mathcal{X}_a, \quad t = 1, 2, \dots n;$$
$$\phi_t^{(b)} : (\mathcal{S}^b)^t \times \mathcal{W}_b \to \mathcal{X}_b, \quad t = 1, 2, \dots n.$$

2) An associated decoding function
$$\psi : (\mathcal{S})^n \times \mathcal{Y}^n \to \mathcal{W}_a \times \mathcal{W}_b.$$

The system's probability of error, $P_e^{(n)}$, is given by

$$\frac{1}{2^{n(R_a+R_b)}} \sum_{w_a=1}^{2^{nR_a}} \sum_{w_b=1}^{2^{nR_b}} P\left(\psi(Y_{[n]}, S_{[n]}) \neq (w_a, w_b)|\mathbf{W} = \mathbf{w}\right).$$

A rate pair (R_a, R_b) is achievable if for any $\epsilon > 0$ there exists, for all n sufficiently large, an $(n, 2^{nR_a}, 2^{nR_b})$ code such that $\frac{1}{N} \log M_a \geq R_a \geq 0$, $\frac{1}{N} \log M_b \geq R_b \geq 0$ and $P_e^{(n)} \leq \epsilon$. The capacity region of the FS-MAC, $\mathcal{C}_{FS}$, is the closure of the set of all achievable rate pairs (R_a, R_b) and the sum-rate capacity is defined as $\mathcal{C}_{\Sigma}^{FS} := \max_{(R_a, R_b) \in \mathcal{C}_{FS}} (R_a + R_b)$.

Before proceeding with the main result, we introduce *memoryless stationary team policies* [11] and their associated rate regions. We first define Shannon strategies.

Definition 2: Let the set of all possible functions from $\mathcal{S}^a$ to $\mathcal{X}^a$ and $\mathcal{S}^b$ to $\mathcal{X}^b$ be denoted by $\mathcal{T}^a$ and $\mathcal{T}^b$, respectively, where $\mathcal{T}^a = \mathcal{X}^{a|\mathcal{S}^a|}$ and $\mathcal{T}^b = \mathcal{X}^{b|\mathcal{S}^b|}$. Let $T^a \in \mathcal{T}^a$ and $T^b \in \mathcal{T}^b$ be two $\mathcal{T}^a$-valued and $\mathcal{T}^b$-valued random vectors, respectively, referred to as Shannon strategies.

Definition 3: [11] A memoryless stationary (in time) team policy is a family

$$\Pi = \left\{\pi = (\pi_{T^a}(\cdot), \pi_{T^b}(\cdot)) \in \mathcal{P}(\mathcal{T}^a) \times \mathcal{P}(\mathcal{T}^b)\right\} \quad (3)$$

of probability distributions on the two sets of random functions. For every memoryless stationary team policy π, $\mathcal{R}(\pi)$ denotes the region of all rate pairs $R = (R_a, R_b)$ satisfying

$$R_a < I(T^a; Y|T^b, S) \quad (4)$$
$$R_b < I(T^b; Y|T^a, S) \quad (5)$$
$$R_a + R_b < I(T^a, T^b; Y|S) \quad (6)$$

where S, T^a, T^b and Y are random variables taking values in $\mathcal{S}$, $\mathcal{T}^a$, $\mathcal{T}^b$ and $\mathcal{Y}$, respectively and whose joint probability distribution factorizes as

$$P_{S, T^a, T^b, Y}(s, t^a, t^b, y)$$
$$= P_S(s) P_{Y|T^a, T^b, S}(y|t^a, t^b, s) \pi_{T^a}(t^a) \pi_{T^b}(t^b). \quad (7)$$

We can now state the inner bound to the capacity region. Let $\mathcal{C}_{IN} := \overline{co}\left(\bigcup_{\pi} \mathcal{R}(\pi)\right)$ denotes the closure of the convex hull of the rate regions $\mathcal{R}(\pi)$ given by (4)-(6) associated to all possible memoryless stationary team polices as defined in (3).

Theorem 1 (Inner Bound): $\mathcal{C}_{IN} \subseteq \mathcal{C}_{FS}$.
The achievability proof follows the standard arguments of joint ϵ-typical n-sequences [14, Section 15.2].

Definition 4: [14, Section 15.2] The set $\mathcal{A}_{\epsilon}^n$ of ϵ-typical n-sequences $\{(x_{[n]}^1, \cdots, x_{[n]}^k)\}$ with respect to the distribution $P_{X^1, \cdots, X^k}(x^1, \cdots, x^k)$ is defined by

$$\mathcal{A}_{\epsilon}^n = \Big\{(x_{[n]}^1, \cdots, x_{[n]}^k) \in \mathcal{X}^1 \times \cdots \mathcal{X}^k :$$
$$\Big| -\frac{1}{n} \log(P(\mathbf{s})) - H(S)\Big| < \epsilon, \forall S \subseteq \{X^1, \cdots, X^k\}\Big\}$$

where $\mathbf{s}$ denotes ordered set of sequences in $x_{[n]}^1, \cdots, x_{[n]}^k$ corresponding to S.

Proof of Theorem 1: Fix $(R_a, R_b) \in \mathcal{R}(\pi)$.

Codebook Generation Fix $\pi_{T^a}(t^a)$ and $\pi_{T^b}(t^b)$. For each $w_a \in \{1, \cdots, 2^{nR_a}\}$ randomly generate its corresponding n-tuple $t^a_{[n],w_a}$, each according to $\prod_{i=1}^n \pi_{T_i^a}(t^a_{i,w_a})$. Similarly, For each $w_b \in \{1, \cdots, 2^{nR_b}\}$ randomly generate its corresponding n-tuple $t^b_{[n],w_b}$, each according to $\prod_{i=1}^n \pi_{T_i^b}(t^b_{i,w_b})$. These codeword pairs form the codebook, which is revealed to the decoder while codewords t^l_{i,w_l} is revealed to encoder l, $l = \{a, b\}$.

Encoding Define the encoding functions as follows: $x_i^a(w_a) = \phi_i^a(w_a, s^a_{[i]}) = t^a_{i,w_a}(s_i^a)$ and $x_i^b(w_b) = \phi_i^b(w_b, s^b_{[i]}) = t^b_{i,w_b}(s_i^b)$ where t^a_{i,w_a} and t^b_{i,w_b} denote the ith component of $t^a_{[n],w_a}$ and $t^b_{[n],w_b}$, respectively, and s_i^a and s_i^b denote the last component $s^a_{[i]}$ and $s^b_{[i]}$, respectively, $i = 1, \cdots, n$. Therefore, to send the messages w_a and w_b, we simply transmit the corresponding $t^a_{[n],w_a}$ and $t^b_{[n],w_b}$, respectively.

Decoding After receiving $(y_{[n]}, s_{[n]})$, the decoder looks for the only (w_a, w_b) pair such that $(t^a_{[n],w_a}, t^b_{[n],w_b}, y_{[n]}, s_{[n]})$ are jointly $\epsilon-$typical and declares this pair as its estimate $(\hat{w}_a, \hat{w}_b)$.

Error Analysis Without loss of generality, we can assume that $(w_a, w_b) = (1, 1)$ was sent. An error occurs, if the correct codewords are not typical with the received sequence or there is a pair of incorrect codewords that are typical with the received sequence. Define the events $E_{\alpha,\beta} \overset{\triangle}{=} \{(T^a_{[n],\alpha}, T^b_{[n],\beta}, Y_{[n]}, S_{[n]}) \in A_\epsilon^n\}$, $\alpha \in \{1, \cdots, 2^{nR_a}\}$ and $\beta \in \{1, \cdots, 2^{nR_b}\}$. Then by the union bound we get

$$P_e^n = P(E_{1,1}^c \bigcup_{(\alpha,\beta) \neq (1,1)} E_{\alpha,\beta}) \leq P(E_{1,1}^c)$$
$$+ \sum_{\alpha=1,\beta \neq 1} P(E_{\alpha,\beta}) + \sum_{\alpha \neq 1,\beta=1} P(E_{\alpha,\beta}) + \sum_{\alpha \neq 1,\beta \neq 1} P(E_{\alpha,\beta})$$
$$(8)$$

where $P(E_{1,1}^c)$ denotes the probability that no message pair is jointly typical. It can easily be verified that $\{Y_i, S_i, T_i^a, T_i^b\}_{i=1}^\infty$ is an i.i.d. sequence and by [14, Theorem 15.1.2], $P(E_{1,1}^c) \to 0$ for sufficiently large n. Next, let us consider the second term

$$\sum_{\alpha=1,\beta \neq 1} P(E_{\alpha=1,\beta \neq 1})$$
$$= \sum_{\alpha=1,\beta \neq 1} P((T^a_{[n],1}, T^b_{[n],\beta}, Y_{[n]}, S_{[n]}) \in A_\epsilon^n)$$
$$\overset{(i)}{=} \sum_{\alpha=1,\beta \neq 1} \sum_{(t^a_{[n]}, t^b_{[n]}, y_{[n]}, s_{[n]}) \in A_\epsilon^n} P_{T^b_{[n]}}(t^b_{[n]})$$
$$P_{T^a_{[n]}, Y_{[n]}, S_{[n]}}(t^a_{[n]}, y_{[n]}, s_{[n]})$$
$$\leq \sum_{\alpha=1,\beta \neq 1} |A_\epsilon^n| 2^{-n[H(T^b)-\epsilon]} 2^{-n[H(T^a,Y,S)-\epsilon]}$$
$$\leq 2^{nR_b} 2^{-n[H(T^b)+H(T^a,Y,S)-H(T^a,T^b,Y,S)-3\epsilon]}$$
$$\overset{(ii)}{=} 2^{n[R_b - I(T^b;Y|S,T^a)-3\epsilon]}$$
$$(9)$$

where (i) holds since for $\beta \neq 1$, $T^b_{[n],\beta}$ is independent of $(T^a_{[n],1}, Y_{[n]}, S_{[n]})$ and (ii) follows since T^b and (T^a, S)

are independent and $I(T^b; Y, T^a, S) = I(T^b; T^a, S) + I(T^b; Y|T^a, S) = I(T^b; Y|T^a, S)$, where $I(T^b; T^a, S) = 0$. Following the same steps for $(\alpha \neq 1, \beta = 1)$ and $(\alpha \neq 1, \beta \neq 1)$ we get

$$\sum_{\alpha \neq 1,\beta=1} P(E_{\alpha,\beta}) \leq 2^{n[R_a - I(T^a;Y|T^b,S)-3\epsilon]},$$
$$\sum_{\alpha \neq 1,\beta \neq 1} P(E_{\alpha,\beta}) \leq 2^{n[R_a+R_b-I(T^a,T^b;Y|S)-3\epsilon]}, (10)$$

and the rate conditions of the $\mathcal{R}(\pi)$ imply that each term tends in (8) tends to zero as $n \to \infty$. This shows the achievability of a rate pair $(R_a, R_b) \in \mathcal{R}(\pi)$. Achievability of any rate pair in $\mathcal{C}_{IN}$ follows from a standard time-sharing argument. ∎

We now present an outer bound to $\mathcal{C}_{FS}$, which is obtained by providing a tight converse to the sum-rate capacity. Let

$$\mathcal{C}_{OUT} := \Big\{ (R_a, R_b) \in \mathbb{R}^+ \times \mathbb{R}^+ :$$
$$R_a + R_b \leq \sup_{\pi_{T^a}(t^a)\pi_{T^b}(t^b)} I(T^a, T^b; Y|S) \Big\}.$$

Theorem 2 (Outer Bound): $\mathcal{C}_{FS} \subseteq \mathcal{C}_{OUT}$.

Proof: We need to show that all achievable rates satisfy

$$R_a + R_b \leq \sup_{\pi_{T^a}(t^a)\pi_{T^b}(t^b)} I(T^a, T^b; Y|S),$$

i.e., a converse for the sum-rate capacity. We use the converse technique of [11] and extend it to a noisy setup. Therefore, following [11] let $\alpha_\sigma := \frac{1}{n} P_{S_{[t-1]}}(\sigma)$, $\eta(\epsilon) := \frac{\epsilon}{1-\epsilon} \log |\mathcal{Y}| + \frac{H(\epsilon)}{1-\epsilon}$. Observe that $\lim_{\epsilon \to 0} \eta(\epsilon) = 0$ and

$$\sum_{\sigma \in \mathcal{S}^{(n)}} \alpha_\sigma = \frac{1}{n} \sum_{1 \leq t \leq n} \sum_{\sigma \in \mathcal{S}^{(t-1)}} P_{S_{[t-1]}}(\sigma) = 1,$$

where $\mathcal{S}^{(n)}$ and $\mathcal{S}^{(t-1)}$ are the sets of all $\mathcal{S}$-strings of length n and $(t-1)$, respectively. First recall that, since $X_t^a = \phi_t^{(a)}(W_a, S^a_{[t-1]}, S_t^a)$ and $X_t^b = \phi_t^{(b)}(W_b, S^b_{[t-1]}, S_t^b)$, we have

$$T_t^a = \phi_t^{(a)}(W_a, S^a_{[t-1]}) \in \mathcal{X}^{a|\mathcal{S}^a|},$$
$$T_t^b = \phi_t^{(b)}(W_b, S^b_{[t-1]}) \in \mathcal{X}^{b|\mathcal{S}^b|}. \quad (11)$$

We now show that the sum of any achievable rate pair can be written as the convex combination of conditional mutual information terms which are indexed by the realization of past complete state information.

Lemma 1: Let $T_t^a \in \mathcal{T}^a$ and $T_t^b \in \mathcal{T}^b$ be the Shannon strategies induced by ϕ_t^a and ϕ_t^b, respectively, as shown in (11). Assume that a rate pair $R = (R_a, R_b)$, with block length $n \geq 1$ and a constant $\epsilon \in (0, 1/2)$, is achievable. Then,

$$R_a + R_b \leq \sum_{\sigma \in \mathcal{S}^{(n)}} \alpha_\sigma I(T_t^a, T_t^b; Y_t|S_t, S_{[t-1]} = \sigma) + \eta(\epsilon). \quad (12)$$

Proof: Let $\mathbf{T}_t := (T_t^a, T_t^b)$. By Fano's inequality, we get

$$H(\mathbf{W}|Y_{[n]}, S_{[n]}) \leq H(\epsilon) + \epsilon \log(|\mathcal{W}_a||\mathcal{W}_b|). \quad (13)$$

Observing that

$$
\begin{aligned}
I(\mathbf{W}; Y_{[n]}, S_{[n]}) &= H(\mathbf{W}) - H(\mathbf{W}|Y_{[n]}, S_{[n]}) \\
&= \log(|\mathcal{W}_a||\mathcal{W}_b|) - H(\mathbf{W}|Y_{[n]}, S_{[n]}).
\end{aligned}
\tag{14}
$$

Combining (13) and (14) gives

$$
(1 - \epsilon)\log(|\mathcal{W}_a||\mathcal{W}_b|) \leq I(\mathbf{W}; Y_{[n]}, S_{[n]}) + H(\epsilon)
$$

and

$$
\begin{aligned}
R_a + R_b &= \frac{1}{n}\log(|\mathcal{W}_a||\mathcal{W}_b|) \\
&\leq \frac{1}{1-\epsilon}\frac{1}{n}\left(I(\mathbf{W}; Y_{[n]}, S_{[n]}) + H(\epsilon)\right).
\end{aligned}
\tag{15}
$$

Furthermore, $I(\mathbf{W}; Y_{[n]}, S_{[n]})$ can be written as

$$
\begin{aligned}
&\sum_{t=1}^{n}\left[H(Y_t, S_t|S_{[t-1]}, Y_{[t-1]}) - H(Y_t, S_t|\mathbf{W}, S_{[t-1]}, Y_{[t-1]})\right] \\
&\overset{(i)}{=} \sum_{t=1}^{n}\left[H(Y_t|S_{[t]}, Y_{[t-1]}) - H(Y_t|\mathbf{W}, S_{[t]}, Y_{[t-1]})\right] \\
&\overset{(ii)}{\leq} \sum_{t=1}^{n}\left[H(Y_t|S_{[t]}) - H(Y_t|\mathbf{W}, S_{[t]}, Y_{[t-1]}, \mathbf{T}_t)\right] \\
&\overset{(iii)}{=} \sum_{t=1}^{n}\left[H(Y_t|S_{[t]}) - H(Y_t|S_{[t]}, \mathbf{T}_t)\right] \\
&= \sum_{t=1}^{n}I(\mathbf{T}_t; Y_t|S_{[t]})
\end{aligned}
\tag{16}
$$

where (i) follows from the fact that S_t is i.i.d. and independent of $\mathbf{W}$, in (ii), $\mathbf{T}_t := (T_t^a, T_t^b)$ are Shannon strategies whose realizations are mappings $t_t^i : S_t^i \to X_t^i$ for $i = \{a, b\}$ and thus (ii) holds since conditioning reduces entropy. Finally, (iii) follows since

$$
\begin{aligned}
&P_{Y_t|\mathbf{W}, S_t, S_{[t-1]}, Y_{[t-1]}, T_t^a, T_t^b}(y_t|\mathbf{w}, s_t, s_{[t-1]}, y_{[t-1]}, t_t^a, t_t^b) \\
&= \sum_{s_t^a, s_t^b} P_{Y_t|S_t, S_t^a, S_t^b, T_t^a, T_t^b}(y_t|s_t, s_t^a, s_t^b, t_t^a, t_t^b) \\
&\qquad\qquad \times P_{S_t^a, S_t^b|S_t}(s_t^a, s_t^b|s_t) \\
&= P_{Y_t|S_t, T_t^a, T_t^b}(y_t|s_t, t_t^a, t_t^b)
\end{aligned}
\tag{17}
$$

where the first equality is verified by (2), where $x_t^i = t_t^i(s_t^i)$ for $i = \{a, b\}$, and by $\{S_t\}$ being i.i.d. and independent of

$\mathbf{W}$. Now, let $\chi(\epsilon) := \frac{H(\epsilon)}{n(1-\epsilon)}$ and combining (15)-(16) gives

$$
\begin{aligned}
R_a + R_b &= \frac{1}{n}\log(|\mathcal{W}_a||\mathcal{W}_b|) \\
&\leq \frac{1}{1-\epsilon}\frac{1}{n}\sum_{t=1}^{n}I(T_t^a, T_t^b; Y_t|S_{[t]}) + \chi(\epsilon) + (n-1)\chi(\epsilon) \\
&\overset{(a)}{\leq} \frac{1}{1-\epsilon}\frac{1}{n}\sum_{t=1}^{n}I(T_t^a, T_t^b; Y_t|S_{[t]}) + \eta(\epsilon) \\
&\qquad\qquad - \frac{\epsilon}{1-\epsilon}\frac{1}{n}\sum_{t=1}^{n}I(T_t^a, T_t^b; Y_t|S_{[t]}) \\
&= \frac{1}{n}\sum_{t=1}^{n}I(T_t^a, T_t^b; Y_t|S_{[t]}) + \eta(\epsilon)
\end{aligned}
\tag{18}
$$

where (a) is valid since $I(T_t^a, T_t^b; Y_t|S_{[t]}) \leq \log|\mathcal{Y}|$. Furthermore,

$$
\begin{aligned}
&I(T_t^a, T_t^b; Y_t|S_{[t]}) \\
&\qquad = n\sum_{\sigma \in \mathcal{S}^{(t-1)}}\alpha_\sigma I(T_t^a, T_t^b; Y_t|S_t, S_{[t-1]} = \sigma),
\end{aligned}
\tag{19}
$$

and substituting the above into (18) yields (12). ∎

Observe now that, for any $t \geq 1$, $I(T_t^a, T_t^b; Y_t|S_t, S_{[t-1]} = \sigma)$ is a function of the joint conditional distribution of channel state S_t, inputs T_t^a, T_t^b and output Y_t given the past realization $(S_{[t-1]} = \sigma)$. Hence, to complete the proof of the outer bound, we need to show that $P_{T_t^a, T_t^b, Y_t, S_t|S_{[t-1]}}(t_t^a, t_t^b, y_t, s_t|\sigma)$ factorizes as in (7). This is done in the lemma below. In particular, it is crucial to observe that the complete state observation at the decoder is enough to provide a product form on T^a and T^b. Before stating the lemma, let us introduce some more notations. Let $\sigma_{\mathbf{a}}$ and $\sigma_{\mathbf{b}}$ denote particular realizations of $S_{[t-1]}^a$ and $S_{[t-1]}^b$, respectively. Let

$$
\begin{aligned}
\Upsilon_{\sigma_{\mathbf{a}}}^a(t^a) &:= \{w_a : \phi_t^{(a)}(w_a, s_{[t-1]}^a = \sigma_{\mathbf{a}}) = t^a\}, \\
\Upsilon_{\sigma_{\mathbf{b}}}^b(t^b) &:= \{w_b : \phi_t^{(b)}(w_b, s_{[t-1]}^b = \sigma_{\mathbf{b}}) = t^b\}
\end{aligned}
\tag{20}
$$

and

$$
\begin{aligned}
\pi_{T^a}^{\sigma_{\mathbf{a}}}(t^a) &:= \sum_{w_a \in \Upsilon_{\sigma_{\mathbf{a}}}^a(t^a)}\frac{1}{|\mathcal{W}_a|}, \\
\pi_{T^b}^{\sigma_{\mathbf{b}}}(t^b) &:= \sum_{w_b \in \Upsilon_{\sigma_{\mathbf{b}}}^b(t^b)}\frac{1}{|\mathcal{W}_b|}, \\
\pi_{T^a}^{\sigma}(t^a) &:= \sum_{\sigma_{\mathbf{a}}}\pi_{T^a}^{\sigma_{\mathbf{a}}}(t^a)P_{S_{[t-1]}^a|S_{[t-1]}}(\sigma_{\mathbf{a}}|\sigma), \\
\pi_{T^b}^{\sigma}(t^b) &:= \sum_{\sigma_{\mathbf{b}}}\pi_{T^b}^{\sigma_{\mathbf{b}}}(t^b)P_{S_{[t-1]}^b|S_{[t-1]}}(\sigma_{\mathbf{b}}|\sigma).
\end{aligned}
\tag{21}
$$

Lemma 2: For every $1 \leq t \leq n$ and $\sigma \in (\mathcal{S})^{t-1}$, the following holds

$$
\begin{aligned}
&P_{T_t^a, T_t^b, Y_t, S_t|S_{[t-1]}}(t^a, t^b, y, s|\sigma) \\
&\qquad = P_S(s)P_{Y|S, T^a, T^b}(y|s, t^a, t^b)\pi_{T^a}^{\sigma}(t^a)\pi_{T^b}^{\sigma}(t^b).
\end{aligned}
\tag{22}
$$

Proof: Let $\mathbf{S} := (S, S_t^a, S_t^b)$ and $\mathbf{s} := (s, s_t^a, s_t^b)$. Observe that

$$P_{T_t^a, T_t^b, Y_t, S_t | S_{[t-1]}}(t^a, t^b, y, s | \sigma)$$
$$= \sum_{s_t^a \in \mathcal{S}^a} \sum_{s_t^b \in \mathcal{S}^b} P_{\mathbf{S}, T_t^a, T_t^b, Y_t | S_{[t-1]}}(\mathbf{s}, t^a, t^b, y | \sigma)$$
$$= \sum_{s_t^a \in \mathcal{S}^a} \sum_{s_t^b \in \mathcal{S}^b} P_{Y | \mathbf{S}, T_t^a, T_t^b}(y | \mathbf{s}, t^a, t^b)$$
$$\times P_{\mathbf{S}, T_t^a, T_t^b | S_{[t-1]}}(\mathbf{s}, t^a, t^b | \sigma) \quad (23)$$

where the second equality verified by (2) since $x_t^i = t_t^i(s_t^i)$ for $i = \{a, b\}$. Let us now consider the term $P_{\mathbf{S}, T_t^a, T_t^b | S_{[t-1]}}(\mathbf{s}, t^a, t^b | \sigma)$ above. We have the following

$$P_{\mathbf{S}, T_t^a, T_t^b | S_{[t-1]}}(\mathbf{s}, t^a, t^b | \sigma)$$
$$= \sum_{w_a \in \mathcal{W}_a} \sum_{w_b \in \mathcal{W}_b} \sum_{\sigma_\mathbf{a}} \sum_{\sigma_\mathbf{b}}$$
$$P_{\mathbf{W}, S_{[t-1]}^a, S_{[t-1]}^b, \mathbf{S}, T_t^a, T_t^b | S_{[t-1]}}(\mathbf{w}, \sigma_\mathbf{a}, \sigma_\mathbf{b}, \mathbf{s}, t^a, t^b | \sigma)$$
$$\overset{(i)}{=} P_{\mathbf{S}}(\mathbf{s}) \sum_{w_a \in \mathcal{W}_a} \sum_{w_b \in \mathcal{W}_b} \sum_{\sigma_\mathbf{a}} \sum_{\sigma_\mathbf{b}}$$
$$P_{\mathbf{W}, S_{[t-1]}^a, S_{[t-1]}^b, T_t^a, T_t^b | S_{[t-1]}}(\mathbf{w}, \sigma_\mathbf{a}, \sigma_\mathbf{b}, t^a, t^b | \sigma)$$
$$\overset{(ii)}{=} P_{\mathbf{S}}(\mathbf{s}) \sum_{w_a \in \mathcal{W}_a} \sum_{w_b \in \mathcal{W}_b} \sum_{\sigma_\mathbf{a}} \sum_{\sigma_\mathbf{b}} 1_{\{t^l = \phi_t^{(l)}(w_l, \sigma_l), \; l=a,b\}}$$
$$\times P_{\mathbf{W}, S_{[t-1]}^a, S_{[t-1]}^b | S_{[t-1]}}(\mathbf{w}, \sigma_\mathbf{a}, \sigma_\mathbf{b} | \sigma)$$
$$\overset{(iii)}{=} P_{\mathbf{S}}(\mathbf{s}) \sum_{w_a \in \mathcal{W}_a} \sum_{w_b \in \mathcal{W}_b} \sum_{\sigma_\mathbf{a}} \sum_{\sigma_\mathbf{b}} 1_{\{t^l = \phi_t^{(l)}(w_l, \sigma_l), \; l=a,b\}}$$
$$\times \frac{1}{|\mathcal{W}_a|} \frac{1}{|\mathcal{W}_b|} P_{S_{[t-1]}^a, S_{[t-1]}^b | S_{[t-1]}}(\sigma_\mathbf{a}, \sigma_\mathbf{b} | \sigma)$$
$$\overset{(iv)}{=} P_{\mathbf{S}}(\mathbf{s}) \sum_{\sigma_\mathbf{a}} P_{S_{[t-1]}^a | S_{[t-1]}}(\sigma_\mathbf{a} | \sigma) \sum_{\sigma_\mathbf{b}} P_{S_{[t-1]}^b | S_{[t-1]}}(\sigma_\mathbf{b} | \sigma)$$
$$\sum_{w_a \in \mathcal{W}_a} \frac{1}{|\mathcal{W}_a|} 1_{\{t^a = \phi_t^{(a)}(w_a, \sigma_\mathbf{a})\}} \sum_{w_b \in \mathcal{W}_b} \frac{1}{|\mathcal{W}_b|} 1_{\{t^b = \phi_t^{(b)}(w_b, \sigma_\mathbf{b})\}}$$
$$\overset{(v)}{=} P_{\mathbf{S}}(\mathbf{s}) \sum_{\sigma_\mathbf{a}} P_{S_{[t-1]}^a | S_{[t-1]}}(\sigma_\mathbf{a} | \sigma) \sum_{w_a \in \Upsilon_{\sigma_\mathbf{a}}^a(t^a)} \frac{1}{|\mathcal{W}_a|}$$
$$\sum_{\sigma_\mathbf{b}} P_{S_{[t-1]}^b | S_{[t-1]}}(\sigma_\mathbf{b} | \sigma) \sum_{w_b \in \Upsilon_{\sigma_\mathbf{b}}^b(t^b)} \frac{1}{|\mathcal{W}_b|}$$
$$\overset{(vi)}{=} P_{\mathbf{S}}(\mathbf{s}) \sum_{\sigma_\mathbf{a}} P_{S_{[t-1]}^a | S_{[t-1]}}(\sigma_\mathbf{a} | \sigma) \pi_{T^a}^{\sigma_\mathbf{a}}(t^a)$$
$$\sum_{\sigma_\mathbf{b}} P_{S_{[t-1]}^b | S_{[t-1]}}(\sigma_\mathbf{b} | \sigma) \pi_{T^b}^{\sigma_\mathbf{b}}(t^b)$$
$$\overset{(vii)}{=} P_{\mathbf{S}}(\mathbf{s}) \pi_{T^a}^\sigma(t^a) \pi_{T^b}^\sigma(t^b) \quad (24)$$

where (i) is valid since the current state is independent of $\mathbf{W}$ and (T^a, T^b), (ii) is valid by (11), (iii) is valid since $\mathbf{W}$ is independent from the state processes, (iv) is valid by (1) and (11), (v) is valid due to (20) and $(vi) - (vii)$ is valid due to (21). Substituting (24) into (23) proves the lemma. ■

We can now complete the proof of Theorem 2. With Lemma 1 it is shown that the sum of any achievable rate pair can be approximated by the convex combinations of rate conditions given in (6) which are indexed by $\sigma \in \mathcal{S}^{(n)}$ and satisfy (7) for joint state-input-output distributions. More explicitly, we have

$$R_a + R_b \leq \sum_{\sigma \in \mathcal{S}^{(n)}} \alpha_\sigma I(T_t^a, T_t^b; Y_t | S_t, S_{[t-1]} = \sigma) + \eta(\epsilon)$$
$$= \sum_{\sigma \in \mathcal{S}^{(n)}} \alpha_\sigma I(T_t^a, T_t^b; Y_t | S_t)_{\pi_{T^a}^\sigma(t^a) \pi_{T^b}^\sigma(t^b)} + \eta(\epsilon)$$
$$\leq \sup_{\left(\pi_{T^a}^\sigma(t^a) \pi_{T^b}^\sigma(t^b), \; \sigma\right)} I(T_t^a, T_t^b; Y_t | S_t) + \eta(\epsilon)$$
$$\leq \sup_{\left(\pi_{T^a}(t^a) \pi_{T^b}(t^b) \in \Pi\right)} I(T_t^a, T_t^b; Y_t | S_t) + \eta(\epsilon)$$

where the second step is valid since $I(T_t^a, T_t^b; Y_t | S_t, S_{[t-1]} = \sigma)$ is a function of the joint conditional distribution of channel state S_t, inputs T_t^a, T_t^b and output Y_t given the past realization $(S_{[t-1]} = \sigma)$. Hence, since $\lim_{n \to \infty} \eta(\epsilon) = 0$, any achievable pair satisfies $R_a + R_b \leq \sup_{\pi_{T^a}(t^a) \pi_{T^b}(t^b)} I(T^a, T^b; Y | S)$. ■ As a direct consequence of Theorems 1 and 2, we have the following corollary.

Corollary 1:

$$\mathcal{C}_\Sigma^{FS} = \sup_{\pi_{T^a}(t^a) \pi_{T^b}(t^b)} I(T^a, T^b; Y | S). \quad (25)$$

Proof: We need to show that $\exists \; (R_a, R_b) \in \mathcal{C}_{IN}$ achieving (25). We follows steps akin to [14, p.535] where discrete memoryless MACs are considered. Let us fix $\pi_{T^a}(t^a) \pi_{T^b}(t^b)$ and consider the rate constraints given in $\mathcal{C}_{IN}$

$$I(T^a; Y | T^b, S) = H(T^a | T^b, S) - H(T^a | T^b, Y, S)$$
$$= H(T^a) - H(T^a | T^b, Y, S) \quad (26)$$
$$I(T^b; Y | T^a, S) = H(T^b | T^a, S) - H(T^b | T^a, Y, S)$$
$$= H(T^b) - H(T^b | T^a, Y, S) \quad (27)$$

and

$$I(T^a, T^b; Y | S)$$
$$= H(T^a, T^b) - H(T^a, T^b | Y, S)$$
$$= H(T^a) + H(T^b) - H(T^a | T^b, Y, S) - H(T^b | Y, S), \quad (28)$$

where (26), (27) and (28) are valid since T^a and T^b are independent of each other and independent of S. Observe now that for any $\pi_{T^a}(t^a) \pi_{T^b}(t^b)$, $I(T^a; Y | T^b, S) + I(T^b; Y | T^a, S) \geq I(T^a, T^b; Y | S)$ since $H(T^b | Y, S) \geq H(T^b | T^a, Y, S)$. Therefore, the sum-rate constraint in $\mathcal{C}_{IN}$ is always active and hence, there exists $(R_a, R_b) \in \mathcal{C}_{IN}$ achieving (25). ■

Remark 1: One main observation about the proof of Theorem 2 is the fact that, once we have the complete state information, conditioning on which allows a product form on T^a and T^b, there is no loss of optimality (for the sum-rate capacity) in using associated memoryless team policies instead of using all the past information at the receiver. This fact is observed in [11] when the information at the encoders are asymmetric quantized version of the information at the decoder.

Remark 2: For the validity of Corollary 1, it is crucial to have the product form on (T^a, T^b). If this

is not the case, we would get that $I(T^a; Y|T^b, S) + I(T^b; Y|T^a, S) = H(T^a|T^b) + H(T^b|T^a) - H(T^a|T^b, Y, S) - H(T^b|T^a, Y, S)$ and $I(T^a, T^b; Y|S) = H(T^a|T^b) + H(T^b) - H(T^a|T^b, Y, S) - H(T^b|Y, S)$. Therefore, it is possible to get an obsolete sum-rate constraint in $\mathcal{C}_{IN}$ and hence, achievability of $\mathcal{C}_{\sum}^{FS}$ is not guaranteed.

Remark 3: It should be noted that the main difference between the problem that we consider here and the one considered in [11] is the information at the decoder about the information at the encoders. More explicitly, in [11], the information at the encoders are available at the decoder and as such, as the authors explicitly mention in their paper, the decoder does not need to estimate the coding policies used in a decentralized time-sharing. From this perspective, the main contribution of our work can be thought as showing that when this is not the case, by enlarging the input space, there is no loss of optimality (for the sum-rate capacity) if the optimization is performed by ignoring the past information at the encoders given that the decoder has complete CSI.

III. Conclusion and Remarks

The present paper has investigated the memoryless FS-MAC with asymmetric noisy CSI at the encoders and complete CSI at the decoder. Single letter inner and outer bounds are presented when the channel state is a sequence of i.i.d. random variables. The main contribution of the paper, i.e., the tight converse for the sum-rate capacity and hence, an outer bound to the capacity region, is realized by observing that the information available at the decoder is enough to attain a product form on the channel input functions and hence, there is no loss of optimality if we ignore the past noisy state information at the encoders.

References

[1] C. E. Shannon, "Channels with side information at the transmitter," *IBM J. Res. Develop,* vol. 2, pp. 289-293, 1958.
[2] S. I. Gelfand and M. S. Pinsker, "Coding for channels with random parameters" *Prob. of Cont. and Inform.,* vol. 9, pp. 19-31, 1980.
[3] M. Salehi, "Capacity and coding for memories with real-time noisy defect information at encoder and decoder," *IEE-Proceedings-I* vol. 139, pp. 113-117, Apr. 1992.
[4] G. Caire and S. Shamai, "On the capacity of some channels with channel state information," *IEEE Trans. Inform. Theory,* vol. 45, no. 6, pp. 2007-2019, Sep. 1999.
[5] A. Das and P. Narayan, "Capacities of time-varying multiple access channels with side information," *IEEE Trans. Inform. Theory,* vol. 48, no. 1, pp. 4-25, Jan. 2002.
[6] S. A. Jafar, "Capacity with casual and non-casual side information-a unified view," *IEEE Trans. Inform. Theory,* vol. 52 no. 12, pp. 5468-5474, Dec. 2006.
[7] Y. Cemal and Y. Steinberg, "The multiple-access channel with partial state information at the encoders," *IEEE Trans. Inform. Theory,* vol. 51, no. 11, pp. 3992-4003, Nov. 2005.
[8] A. Lapidoth and Y. Steinberg, "The multiple-access channel with two independent states each known causally to one encoder," *Proc. IEEE Int. Symp. Information Theory* Austin, Texas, U.S.A, June 2010.
[9] A. Lapidoth and Y. Steinberg, "The multiple-access channel with causal and strictly causal side information at the encoders," *Proc. Int. Zurich Sem. on Communications* Zurich, March 2010.
[10] U. Basher, A. Shirazi and H. Permuter, "Capacity region of finite state multiple-access channel with delayed state information at the transmitters," available at [arXiv:1101.2389].
[11] G. Como and S. Yüksel, "On the capacity of memoryless finite state multiple access channels with asymmetric state information at the encoders," *IEEE Trans. Inform. Theory,* vol. 57, no. 3, pp. 1267-1273, March 2011.
[12] S. Yüksel, "Stochastic nestedness and the belief sharing information pattern," *IEEE Trans. Auto. Control,* vol. 55, pp. 2773-2786, Dec. 2009.
[13] H. S. Witsenhausen, "Equivalent stochastic control problems," *Mathematics of Control, Signal and Systems,* vol. 1, pp. 3-11, Springer-Verlag, 1988.
[14] T. M. Cover and J. Thomas, *Elements of Information Theory,* New York: Wiley, 2nd edition, 2006.