SRC TR 89-29

State Constraints for the Multiple-Access Arbitrarily Varying Channel

By

J. A. Gubner

IEEE Transactions on Information Theory, in review

State Constraints for the Multiple-Access Arbitrarily Varying Channel

JOHN A. GUBNER

Abstract — It is known ~~was shown in [15]~~ that if a multiple-access arbitrarily varying channel is symmetrizable, then the capacity region has an empty interior. In this paper, we show that if a suitable constraint is placed on the channel state sequences, then the capacity region can contain certain open rectangles and thereby possess a nonempty interior. We also prove a new weak converse under a state constraint. We use our results to establish the capacity region of the two-user adder channel under state constraint $\frac{1}{2}$.

I. Introduction

This paper is a continuation of our work in [15]. We assume that the reader is somewhat familiar with the notation and the results found there. To understand the results of this paper it is not necessary to be familar with the proofs found in [15].

Recall that a two-user multiple-access arbitrarily varying channel (or AVC for brevity) is a transition probability W from $\mathcal{X} \times \mathcal{Y} \times \mathcal{S}$ into $\mathcal{Z}$, where $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{S}$, and $\mathcal{Z}$ are finite sets, each containing at least two elements. We interpret $W(z|x, y, s)$ as the conditional probability that the channel output is $z \in \mathcal{Z}$ given that the channel input symbol from user 1 is $x \in \mathcal{X}$, the channel input symbol from user 2 is $y \in \mathcal{Y}$, and that the channel *state* is $s \in \mathcal{S}$. The channel operation on n-tuples $\mathbf{x} = (x_1, \ldots, x_n) \in \mathcal{X}^n$, $\mathbf{y} \in \mathcal{Y}^n$, $\mathbf{s} \in \mathcal{S}^n$, and $\mathbf{z} \in \mathcal{Z}^n$ is given by

$$W^n(\mathbf{z}|\mathbf{x}, \mathbf{y}, \mathbf{s}) \triangleq \prod_{k=1}^{n} W(z_k|x_k, y_k, s_k).$$

This research was conducted while the author was supported by an IEEE Frank A. Cowan Scholarship and by fellowships from the Minta Martin Fund, the University of Maryland Graduate School, and the University of Maryland Systems Research Center under NSF Grant OIR-85-00108.

The author was with the Electrical Engineering Department and Systems Research Center, University of Maryland, College Park. He is now with the Department of Electrical and Computer Engineering, University of Wisconsin, Madison, WI 53706.

II. State Constraints

Fix any function $\ell: \mathcal{S} \to [0, \infty)$ such that $\min_{s \in \mathcal{S}} \ell(s) = 0$. Set $\ell_{\max} \triangleq \max_{s \in \mathcal{S}} \ell(s)$. Next, for any $\mathbf{s} = (s_1, \ldots, s_n) \in \mathcal{S}^n$, let

$$\ell_n(\mathbf{s}) \triangleq \frac{1}{n} \sum_{k=1}^{n} \ell(s_k).$$

Example: If $\mathcal{S} = \{0, 1\}$ and $\ell(s) = s$, then $\ell_n(\mathbf{s})$ is the normalized Hamming weight of $\mathbf{s}$, i.e., the fraction of 1's in $\mathbf{s}$.

Definition 2.1: For any *state constraint*, $L \geq 0$, let

$$\mathcal{S}^n(L) \triangleq \{\mathbf{s} \in \mathcal{S}^n : \ell_n(\mathbf{s}) \leq L\}.$$

Of course, if $L \geq \ell_{\max}$, then $\mathcal{S}^n(L) = \mathcal{S}^n$.

Consider the following modification of [15, Definition 2.2].

Definition 2.2: A pair of nonnegative real numbers, (R_1, R_2), is said to be *achievable under state constraint* L for the AVC W if:

> For every $0 < \lambda < 1$, and every $\Delta R > 0$, there exists a positive integer n_0 such that for all $n \geq n_0$, there exist positive integers N and M such that
>
> $$\frac{\log N}{n} > R_1 - \Delta R \quad \text{and} \quad \frac{\log M}{n} > R_2 - \Delta R,$$
>
> and such that there exists a code (f, g, φ) with
>
> $$\frac{1}{NM} \sum_{i=1}^{N} \sum_{j=1}^{M} W^n(\{\mathbf{z} \in \mathcal{Z}^n : \varphi(\mathbf{z}) \neq (i,j)\} | \mathbf{x}_i, \mathbf{y}_j, \mathbf{s}) \leq \lambda, \qquad \forall \mathbf{s} \in \mathcal{S}^n(L).$$

Definition 2.3: The *capacity region under state constraint* L, denoted $C(W, L)$, is defined by

$$C(W, L) \triangleq \{(R_1, R_2) : (R_1, R_2) \text{ is achievable under state constraint } L\}.$$

If a pair (R_1, R_2) is achievable in the sense of [15, Definition 2.2], it is achievable in the sense Definition 2.2. Thus, we always have

$$C(W) \subset C(W, L).$$

Proof: First, we clearly have $\bigcap_{0<\delta<L} \mathcal{R}^{L-\delta}(W) \supset \mathcal{R}^L(W)$. It remains to prove the reverse inclusion. Lemmas A.1 and A.2 in Appendix A will establish that for every $\varepsilon > 0$, there exists a $\delta > 0$ such that for all p and q,

$$\begin{aligned} R^{L-\delta}(p,q,W) \subset \{(R_1,R_2): \quad & 0 \le R_1 < I^L_{\mathcal{X}\wedge\mathcal{Z}|\mathcal{Y}}(p,q,W)+\varepsilon, \\ & 0 \le R_2 < I^L_{\mathcal{Y}\wedge\mathcal{Z}|\mathcal{X}}(p,q,W)+\varepsilon, \\ & 0 \le R_1+R_2 < I^L_{\mathcal{X}\mathcal{Y}\wedge\mathcal{Z}}(p,q,W)+\varepsilon\}. \end{aligned} \tag{3.2}$$

Let $\mathcal{R}^L_\varepsilon(p,q,W)$ denote the set on the right-hand side of (3.2). Let $\mathcal{R}^L_\varepsilon(W)$ denote the closed convex hull of

$$\bigcup_{p\in\mathcal{D}(\mathcal{X}),\, q\in\mathcal{D}(\mathcal{Y})} \mathcal{R}^L_\varepsilon(p,q,W).$$

Clearly, for every $\varepsilon > 0$, there exists a $0 < \delta < L$ with $\mathcal{R}^{L-\delta}(W) \subset \mathcal{R}^L_\varepsilon(W)$. It follows that

$$\bigcap_{0<\delta<L} \mathcal{R}^{L-\delta}(W) \subset \bigcap_{\varepsilon>0} \mathcal{R}^L_\varepsilon(W).$$

Now, it is easy to see that every point in $\mathcal{R}^L_\varepsilon(W)$ is within distance ε of $\mathcal{R}^L(W)$. Since $\mathcal{R}^L(W)$ is closed set, $\bigcap_{\varepsilon>0} R^L_\varepsilon(W) \subset \mathcal{R}^L(W)$, and so

$$\bigcap_{0<\delta<L} \mathcal{R}^{L-\delta}(W) \subset \bigcap_{\varepsilon>0} \mathcal{R}^L_\varepsilon(W) \subset \mathcal{R}^L(W).$$

□

Having established (3.1), we can now prove the following result.

Theorem 3.2 (Weak Converse Under State Constraint L):

$$C(W,L) \subset \mathcal{R}^L(W).$$

Proof: It suffices to prove that for every $0 < \delta < L$,

$$C(W,L) \subset \mathcal{R}^{L-\delta}(W).$$

Fix $0 < \delta < L$. Let $0 < \lambda < 1$ and $\Delta R > 0$ be arbitrary. Suppose that $(R_1, R_2) \in C(W,L)$. Then by Definition 2.2, for all $n \ge n_0$, there exist positive integers N and M such that

$$\frac{\log N}{n} > R_1 - \Delta R \qquad \text{and} \qquad \frac{\log M}{n} > R_2 - \Delta R, \tag{3.3}$$

$$
\begin{aligned}
\mathsf{P}(\varphi(\mathbf{Z}) \neq (A,B)) &= \sum_{i,j} \mathsf{P}(\varphi(\mathbf{Z}) \neq (i,j), A = i, B = j) \\
&= \sum_{\mathbf{s} \in \mathcal{S}^n} r(\mathbf{s}) \Big(\frac{1}{NM} \sum_{i=1}^{N} \sum_{j=1}^{M} W^n(\{\mathbf{z} : \varphi(\mathbf{z}) \neq (i,j)\} | \mathbf{x}_i, \mathbf{y}_j, \mathbf{s}) \Big) \\
&= \mathsf{E}[\, e(\mathbf{S})] \\
&\leq \max_{\mathbf{s} \in \mathcal{S}^n(L)} e(\mathbf{s}) + \mathsf{P}(\ell_n(\mathbf{S}) > L) \\
&\leq \lambda/2 + \lambda/2 \;=\; \lambda. \qquad (3.7)
\end{aligned}
$$

The remainder of the proof is omitted since it is almost identical to Jahn's proof [16] of the weak converse (without state constraint L, of course). The complete remainder of this proof can be found in [14, pp. 79-82]. □

IV. The Additive AVC

Having established a weak converse in Theorem 3.2, we now compute $\mathcal{R}^L(W)$ for the two-user *additive AVC*. We also consider the special case of the two-user *group adder AVC*. We begin with a few preliminaries. Let $\mathcal{G}$ denote a finite commutative group under $+$, and set $\mathcal{X} = \mathcal{Y} = \mathcal{Z} = \mathcal{G}$. Let $\mathcal{S}$ denote an arbitrary finite set.

Definition 4.1: We say W is an *additive AVC* if

$$W(z|x, y, s) \;=\; V(z - x - y|s) \qquad (4.1)$$

for some transition probability V from $\mathcal{S}$ into $\mathcal{G}$. We also write $(rW)(z|x,y) = (rV)(z - x - y)$, where $(rV)(t) \triangleq \sum_s r(s)V(t|s)$.

To simplify matters, we need the following definiton.

Definition 4.2: For $p, q \in \mathcal{D}(\mathcal{G})$, the *convolution* of p and q is given by

$$(p * q)(t) \;\triangleq\; \sum_{b \in \mathcal{G}} p(t - b)q(b).$$

for some $V_0 \in \mathcal{D}(\mathcal{G})$. For the group adder channel,

$$(rV)(t) \;=\; (r * V_0)(t).$$

We see immediately that if r or V_0 is uniform, $r * V_0 = u$, and $H(r * V_0) = \log|\mathcal{G}|$. Thus, if $u \in \mathcal{D}^L(\mathcal{S})$, or if $V_0 = u$, then (4.3) reduces to $\mathcal{R}^L(W) = \{(0,0)\}$. Note that $u \in \mathcal{D}^L(\mathcal{S})$ if and only if

$$\frac{1}{|\mathcal{G}|}\sum_{s\in\mathcal{G}} \ell(s) \;\leq\; L. \tag{4.4}$$

We can further specialize the group adder AVC to a *noiseless* group adder channel by setting $V_0(t) = \delta(t)$, where $\delta(t) = 1$ if $t = 0$, and $\delta(t) = 0$ otherwise. In this case, $r * V_0 = r * \delta = r$, and

$$\mathcal{R}^L(W) \;=\; \{(R_1, R_2) : 0 \leq R_1 + R_2 \leq \log|\mathcal{G}| - \max_{r\in\mathcal{D}^L(\mathcal{S})} H(r)\}.$$

Thus, for the *noiseless* group adder AVC, $\mathcal{R}^L(W) = \{(0,0)\}$ if and only if (4.4) holds.

V. Forward Theorems

In this section, we prove forward theorems which provide inner bounds on the capacity region under state constraint L. We first recall from [15] some important definitions and results.

Definition 5.1: The AVC W is said to be *symmetrizable-$\mathcal{XY}$* if there exists a transition probability U from $\mathcal{X} \times \mathcal{Y}$ into $\mathcal{S}$ such that

$$\sum_s W(z|x,y,s)U(s|x',y') = \sum_s W(z|x',y',s)U(s|x,y), \qquad \forall x,\, x',\, y,\, y',\, z. \tag{5.1}$$

If no such U exists, we say that W is *nonsymmetrizable-$\mathcal{XY}$*.

Definition 5.2: The AVC W is said to be *symmetrizable-$\mathcal{X}$* if there exists a transition probability U from $\mathcal{X}$ into $\mathcal{S}$ such that

$$\sum_s W(z|x,y,s)U(s|x') = \sum_s W(z|x',y,s)U(s|x), \qquad \forall x,\, x',\, y,\, z. \tag{5.2}$$

If no such U exists, we say that W is *nonsymmetrizable-$\mathcal{X}$*.

Remark: If W is symmetrizable-$\mathcal{XY}$ and U satisfies (5.1), and if q is any element of $\mathcal{D}(\mathcal{Y})$, then multiplying both sides by $q(y)q(y')$ and summing over all y, y' shows that qW is symmetrizable-$\mathcal{X}$. Similarly, if W is symmetrizable-$\mathcal{X}$ and U satisfies (5.2), multiplying both sides by $q(y)$ and summing over all y shows that qW is symmetrizable-$\mathcal{X}$ for every $q \in \mathcal{D}(\mathcal{Y})$.

Definition 5.10: For any $p \in \mathcal{D}(\mathcal{X})$, set $(pW)(z|y,s) \triangleq \sum_x p(x)W(z|x,y,s)$. We say that pW is *symmetrizable-$\mathcal{Y}$* if there exists a transition probability U from $\mathcal{Y}$ into $\mathcal{S}$ such that

$$\sum_s (pW)(z|y,s)U(s|y') \;=\; \sum_s (pW)(z|y',s)U(s|y), \qquad \forall y,\, y',\, z. \tag{5.5}$$

If no such U exists, we say that pW is *nonsymmetrizable-$\mathcal{Y}$*.

A. Nonsymmetrizable Channels

The following theorem is an obvious analog of [15, Theorem 5.1] when the permissible state sequences are constrained to lie in $\mathcal{S}^n(L)$. We prove the existence of nonempty, open rectangles of achievable rate pairs, provided that certain nonsymmetrizability conditions are satisfied.

Theorem 5.11: Suppose W is nonsymmetrizable-$\mathcal{Y}$. Fix any $p \in \mathcal{D}(\mathcal{X})$ and $q \in \mathcal{D}(\mathcal{Y})$. Further, suppose qW is nonsymmetrizable-$\mathcal{X}$. If

$$0 \;<\; R_1 \;<\; I^L_{\mathcal{X}\wedge\mathcal{Z}}(p,q,W) \qquad \text{and} \qquad 0 \;<\; R_2 \;<\; I^L_{\mathcal{Y}\wedge\mathcal{Z}|\mathcal{X}}(p,q,W), \tag{5.6}$$

then (R_1, R_2) is achievable under state constraint L (cf. Definition 2.2).

Remark: It trivially follows from [15, Remark 5.2] that if $p \in \mathcal{D}(\mathcal{X})$ and $q \in \mathcal{D}(\mathcal{Y})$ are strictly positive, then the mutual information quantities in (5.6) are strictly positive under the preceding nonsymmetrizability assumptions.

Proof: The proof of this result is easily obtained by repeating the proof of [15, Theorem 5.1], provided that every occurrence of $\mathcal{D}(\mathcal{S})$ is changed to $\mathcal{D}^L(\mathcal{S})$, and every occurrence of $\mathcal{S}^n$ is changed to $\mathcal{S}^n(L)$. □

Analogous modifications can be made to [15, Lemma D.1] and to [15, Theorem 5.5].

Lemma 5.16: Assume that $U_{\mathcal{X}}(q, W)$ does not depend on q. Fix any $p \in \mathcal{D}(\mathcal{X})$ and $q \in \mathcal{D}(\mathcal{Y})$. If

$$L < \ell_{\mathcal{X}}^{W}(p,q) \qquad \text{and} \qquad L < \ell_{\mathcal{Y}}^{W}(q), \tag{5.7}$$

and if

$$0 < R_1 < I_{\mathcal{X}\wedge\mathcal{Z}}^{L}(p,q,W) \qquad \text{and} \qquad 0 < R_2 < I_{\mathcal{Y}\wedge\mathcal{Z}|\mathcal{X}}^{L}(p,q,W), \tag{5.8}$$

then (R_1, R_2) is achievable under state constraint L (cf. Definition 2.2).

Proof: See Appendix B.

Definition 5.17: Let

$$\mathcal{R}_{\mathcal{X}}^{L}(p,q,W) \triangleq \{(R_1,R_2) : 0 < R_1 < I_{\mathcal{X}\wedge\mathcal{Z}|\mathcal{Y}}^{L}(p,q,W),\ 0 < R_2 < I_{\mathcal{Y}\wedge\mathcal{Z}}^{L}(p,q,W)\},$$

and

$$\mathcal{R}_{\mathcal{Y}}^{L}(p,q,W) \triangleq \{(R_1,R_2) : 0 < R_1 < I_{\mathcal{X}\wedge\mathcal{Z}}^{L}(p,q,W),\ 0 < R_2 < I_{\mathcal{Y}\wedge\mathcal{Z}|\mathcal{X}}^{L}(p,q,W)\}.$$

Observe that if $\mathcal{U}_{\mathcal{X}}(q,W) = \mathcal{U}_{\mathcal{X}}(W)$, then $\ell_{\mathcal{X}}^{W}(p,q) = \ell_{\mathcal{X}}^{W}(p)$. Hence, we have the following theorem.

Theorem 5.18: If for every $q \in \mathcal{D}(\mathcal{Y})$, $\mathcal{U}_{\mathcal{X}}(q,W) = \mathcal{U}_{\mathcal{X}}(W)$, and if for every $p \in \mathcal{D}(\mathcal{X})$, $\mathcal{U}_{\mathcal{Y}}(p,W) = \mathcal{U}_{\mathcal{Y}}(W)$, then $C(W,L)$ contains the closed convex hull of

$$\bigcup_{p \in \mathcal{D}(\mathcal{X}) : L \le \ell_{\mathcal{X}}^{W}(p),\ q \in \mathcal{D}(\mathcal{Y}) : L \le \ell_{\mathcal{Y}}^{W}(q)} [\mathcal{R}_{\mathcal{X}}^{L}(p,q,W) \cup \mathcal{R}_{\mathcal{Y}}^{L}(p,q,W)].$$

An Example: The Adder Channel

To conclude our discussion of state constraints, we return to the adder channel defined in Example 5.4. We take $\ell(s) = s$ so that $\ell_n(\mathbf{s})$ is the average number of 1's in the sequence $\mathbf{s}$. We claim that

$$C(W_a, \tfrac{1}{2}) = \{(R_1,R_2) : 0 \le R_1 \le \tfrac{1}{2},\ 0 \le R_2 \le \tfrac{1}{2},\ 0 \le R_1 + R_2 \le 2 - \tfrac{3}{4}\log 3\}. \tag{5.9}$$

To establish our claim, we proceed as follows. Let $p^*(0) = p^*(1) = \frac{1}{2}$ and $q^*(0) = q^*(1) = \frac{1}{2}$. In Appendix C, we prove that for all $p \in \mathcal{D}(\mathcal{X})$ and all $q \in \mathcal{D}(\mathcal{Y})$,

$$\mathcal{R}^{\frac{1}{2}}(p,q,W_a) \subset \mathcal{R}^{\frac{1}{2}}(p^*,q^*,W_a). \tag{5.10}$$

The key point is to observe that since

$$2 - \frac{3}{4}\log 3 = \left(\frac{3}{2} - \frac{3}{4}\log 3\right) + \frac{1}{2},$$

we can write

$$I^{\frac{1}{2}}_{\mathcal{X}\mathcal{Y}\wedge\mathcal{Z}}(p^*, q^*, W_a) = I^{\frac{1}{2}}_{\mathcal{X}\wedge\mathcal{Z}}(p^*, q^*, W_a) + I^{\frac{1}{2}}_{\mathcal{Y}\wedge\mathcal{Z}|\mathcal{X}}(p^*, q^*, W_a) \tag{5.14}$$

and

$$I^{\frac{1}{2}}_{\mathcal{X}\mathcal{Y}\wedge\mathcal{Z}}(p^*, q^*, W_a) = I^{\frac{1}{2}}_{\mathcal{Y}\wedge\mathcal{Z}}(p^*, q^*, W_a) + I^{\frac{1}{2}}_{\mathcal{X}\wedge\mathcal{Z}|\mathcal{Y}}(p^*, q^*, W_a). \tag{5.15}$$

Remark: Equations (5.14) and (5.15) are nontrivial since in general, "the infimum of a sum is greater than the sum of the infima."

From (5.14) and (5.15), it now follows that $\mathcal{R}^{\frac{1}{2}}(W_a)$, that is, the set on the right in equation (5.12), is equal to the closed convex hull of the union of the two open rectangles (cf. Definition 5.17)

$$\begin{aligned}
&\mathcal{R}^{\frac{1}{2}}_{\mathcal{X}}(p^*, q^*, W_a) \\
&\quad = \{(R_1, R_2) : 0 < R_1 < I^{\frac{1}{2}}_{\mathcal{X}\wedge\mathcal{Z}|\mathcal{Y}}(p^*, q^*, W_a),\ 0 < R_2 < I^{\frac{1}{2}}_{\mathcal{Y}\wedge\mathcal{Z}}(p^*, q^*, W_a)\} \\
&\quad = \{(R_1, R_2) : 0 < R_1 < \tfrac{1}{2},\ 0 < R_2 < \tfrac{3}{2} - \tfrac{3}{4}\log 3\}
\end{aligned}$$

and

$$\begin{aligned}
&\mathcal{R}^{\frac{1}{2}}_{\mathcal{Y}}(p^*, q^*, W_a) \\
&\quad = \{(R_1, R_2) : 0 < R_1 < I^{\frac{1}{2}}_{\mathcal{X}\wedge\mathcal{Z}}(p^*, q^*, W_a),\ 0 < R_2 < I^{\frac{1}{2}}_{\mathcal{Y}\wedge\mathcal{Z}|\mathcal{X}}(p^*, q^*, W_a)\} \\
&\quad = \{(R_1, R_2) : 0 < R_1 < \tfrac{3}{2} - \tfrac{3}{4}\log 3,\ 0 < R_2 < \tfrac{1}{2}\}.
\end{aligned}$$

By Theorem 5.18, it follows that $\mathcal{R}^{\frac{1}{2}}(W_a) \subset C(W_a, \frac{1}{2})$. This proves our claim.

VI. Conclusions

In our prior paper [15], we showed that if an AVC W is symmetrizable-$\mathcal{X}$, symmetrizable-$\mathcal{Y}$, or symmetrizable-$\mathcal{X}\mathcal{Y}$, then the capacity region $C(W)$ has an empty interior. In this paper, we showed that if a suitable constraint is placed on the state sequences, then the

Appendix A

Lemmas A.1 and A.2

Lemma A.1: For every $\eta > 0$, there exists a δ, $0 < \delta < L$, such that for all $r \in \mathcal{D}^L(\mathcal{S})$, there exists an $\hat{r} \in \mathcal{D}^{L-\delta}(\mathcal{S})$ with $d(r, \hat{r}) < \eta$.

Proof: Recall our assumption that $\min_s \ell(s) = 0$. Hence, there is some $s_0 \in \mathcal{S}$ with $\ell(s_0) = 0$. Let $\eta > 0$ be given. Choose $0 < \delta < L$ such that

$$2(1 - \frac{L-\delta}{L}) \; < \; \eta.$$

For $s \neq s_0$, set $\hat{r}(s) = r(s) \cdot (L-\delta)/L$. Since

$$\sum_{s \neq s_0} \hat{r}(s) = (1 - r(s_0))\frac{L-\delta}{L} < 1,$$

we can set $\hat{r}(s_0) = 1 - \sum_{s \neq s_0} \hat{r}(s)$. Observe that since $\ell(s_0) = 0$,

$$\sum_s \ell(s)\hat{r}(s) = \frac{L-\delta}{L} \sum_s \ell(s) r(s) \leq L - \delta.$$

Finally, note that

$$d(r, \hat{r}) \leq 2 \sum_{s \neq s_0} |r(s) - \hat{r}(s)| = 2(1 - \frac{L-\delta}{L}) \sum_{s \neq s_0} r(s) < \eta.$$

$\square$

Lemma A.2: For every $\varepsilon > 0$, there exists a $\delta > 0$ such that for all $p \in \mathcal{D}(\mathcal{X})$ and all $q \in \mathcal{D}(\mathcal{Y})$,

$$\begin{aligned} I^{L-\delta}_{\mathcal{X}\wedge\mathcal{Z}|\mathcal{Y}}(p, q, W) &< I^{L}_{\mathcal{X}\wedge\mathcal{Z}|\mathcal{Y}}(p, q, W) + \varepsilon, \\ I^{L-\delta}_{\mathcal{Y}\wedge\mathcal{Z}|\mathcal{X}}(p, q, W) &< I^{L}_{\mathcal{Y}\wedge\mathcal{Z}|\mathcal{X}}(p, q, W) + \varepsilon, \\ I^{L-\delta}_{\mathcal{X}\mathcal{Y}\wedge\mathcal{Z}}(p, q, W) &< I^{L}_{\mathcal{X}\mathcal{Y}\wedge\mathcal{Z}}(p, q, W) + \varepsilon. \end{aligned}$$

Proof: It suffices to prove the first inequality. Let $\varepsilon > 0$ be given. Choose $\eta > 0$ such that $d(r, \hat{r}) < \eta$ implies

$$|I_{\mathcal{X}\wedge\mathcal{Z}|\mathcal{Y}}(p \times q \times \hat{r}W) - I_{\mathcal{X}\wedge\mathcal{Z}|\mathcal{Y}}(p \times q \times rW)| \; < \; \varepsilon.$$

Choose $\delta > 0$ so small that

$$\begin{aligned} 0 &< 2\delta < \min\{\xi_{\mathcal{X}}^{\alpha}(q, W),\, \xi_{\mathcal{Y}}^{\alpha}(W)\}, \qquad \text{(B.1)} \\ 0 &< R_1 < I_{\mathcal{X}\wedge\mathcal{Z}}^{L}(p, q, W) - 2\delta, \\ 0 &< R_2 < I_{\mathcal{Y}\wedge\mathcal{Z}|\mathcal{X}}^{L}(p, q, W) - 2\delta. \end{aligned}$$

Choose $\hat{p} \in \mathcal{D}(\mathcal{X})$ and $\hat{q} \in \mathcal{D}(\mathcal{Y})$, both strictly positive with $d(p, \hat{p})$ and $d(q, \hat{q})$ both so small that

$$\begin{aligned} \ell_{\mathcal{X}}^{W}(p, q) &\leq \ell_{\mathcal{X}}^{W}(\hat{p}, \hat{q}) + \alpha/2, \\ \ell_{\mathcal{Y}}^{W}(q) &\leq \ell_{\mathcal{Y}}^{W}(\hat{q}) + \alpha/2, \\ \xi_{\mathcal{X}}^{\alpha}(q, W) &\leq \xi_{\mathcal{X}}^{\alpha}(\hat{q}, W) + \delta/2, \\ I_{\mathcal{X}\wedge\mathcal{Z}}^{L}(p, q, W) &\leq I_{\mathcal{X}\wedge\mathcal{Z}}^{L}(\hat{p}, \hat{q}, W) + \delta/2, \\ I_{\mathcal{Y}\wedge\mathcal{Z}|\mathcal{X}}^{L}(p, q, W) &\leq I_{\mathcal{Y}\wedge\mathcal{Z}|\mathcal{X}}^{L}(\hat{p}, \hat{q}, W) + \delta/2. \end{aligned}$$

(Since we have assumed that $U_{\mathcal{X}}(q, W)$ does not depend on q, it is easy to see that $\ell_{\mathcal{X}}^{W}(p, q)$ is a continuous function of p, and does not depend on q.) The remainder of the proof is almost identical to the proof of [15, Theorem 5.1] except as detailed below.

The first step is to replace $\mathcal{S}^n$ by $\mathcal{S}^n(L)$ and $\mathcal{D}(\mathcal{S})$ by $\mathcal{D}^L(\mathcal{S})$. In particular, [15, equation (4.7)] becomes

$$G_{ij} \triangleq \bigcup_{\mathbf{s}' \in \mathcal{S}^n(L)} [\, K_{ij}^{0}(\mathbf{s}') \cap K_{ij}^{1}(\mathbf{s}')\,].$$

To show that for each i, $G_{i1}, \ldots, G_{iM}$ are pairwise disjoint, we proceed as in [15, Appendix B]. Set $P_{XX'YY'SS'} = P_{\mathbf{x}_i, \mathbf{x}_{i'}, \mathbf{y}_j, \mathbf{y}_{j'}, \mathbf{s}, \mathbf{s}'}$. Write

$$\begin{aligned} \ell_n(\mathbf{s}) &= \frac{1}{n} \sum_{k=1}^{n} \ell(s_k) \\ &= \sum_{s} \ell(s) P_S(s) \\ &= \mathsf{E}[\,\ell(S)\,] \\ &= \mathsf{E}[\,\mathsf{E}[\,\ell(S) \mid Y'\,]\,] \\ &= \sum_{y,s} \ell(s) P_{S|Y'}(s|y) Q(y) \\ &= \ell(Q P_{S|Y'}). \end{aligned}$$

where $H(t_1,\ldots,t_m) \triangleq -\sum_{k=1}^{m} t_k \log t_k$, and $h(r) \triangleq H(r, 1-r)$. We also point out that (C.1) – (C.3) can be used to simplify

$$I_{\mathcal{X}\wedge\mathcal{Z}}(p \times q \times rW_a) = I_{\mathcal{X}\mathcal{Y}\wedge\mathcal{Z}}(p \times q \times rW_a) - I_{\mathcal{Y}\wedge\mathcal{Z}|\mathcal{X}}(p \times q \times rW_a)$$

and

$$I_{\mathcal{Y}\wedge\mathcal{Z}}(p \times q \times rW_a) = I_{\mathcal{X}\mathcal{Y}\wedge\mathcal{Z}}(p \times q \times rW_a) - I_{\mathcal{X}\wedge\mathcal{Z}|\mathcal{Y}}(p \times q \times rW_a)$$

for use in verifying (5.13). Now, (5.10) will be established if we can show that

$$\sup_{(p,q)\in[0,1]\times[0,1]} I^{\frac{1}{2}}_{\mathcal{X}\wedge\mathcal{Z}|\mathcal{Y}}(p,q,W_a) = I^{\frac{1}{2}}_{\mathcal{X}\wedge\mathcal{Z}|\mathcal{Y}}(p^*,q^*,W_a), \tag{C.4}$$

$$\sup_{(p,q)\in[0,1]\times[0,1]} I^{\frac{1}{2}}_{\mathcal{Y}\wedge\mathcal{Z}|\mathcal{X}}(p,q,W_a) = I^{\frac{1}{2}}_{\mathcal{Y}\wedge\mathcal{Z}|\mathcal{X}}(p^*,q^*,W_a), \tag{C.5}$$

and

$$\sup_{(p,q)\in[0,1]\times[0,1]} I^{\frac{1}{2}}_{\mathcal{X}\mathcal{Y}\wedge\mathcal{Z}}(p,q,W_a) = I^{\frac{1}{2}}_{\mathcal{X}\mathcal{Y}\wedge\mathcal{Z}}(p^*,q^*,W_a). \tag{C.6}$$

To establish (C.4) – (C.6), we first note that

$$\sup_{(p,q)\in[0,1]\times[0,1]} I^{\frac{1}{2}}_{\ldots}(p,q,W_a) = \sup_{(p,q)\in[0,1]\times[0,1]} \inf_{r\in\mathcal{D}^{\frac{1}{2}}(\mathcal{S})} I_{\ldots}(p \times q \times rW_a).$$

Hence, (C.4) – (C.6) will be established if we can show that $((p^*,q^*),r^*)$ is a *saddle point* in each case. Recall that if F is a real-valued function of two variables, say u and v, then (u^*,v^*) is a saddle point for F if for all u and v,

$$F(u,v^*) \le F(u^*,v^*) \le F(u^*,v). \tag{C.7}$$

If (C.7) holds, it is trivial to show that

$$\sup_u \inf_v F(u,v) = F(u^*,v^*).$$

To establish (C.7), one shows that v^* is a *global* minimum of $F(u^*,v)$ regarded as a function of v, and that u^* is a *global* maximum of $F(u,v^*)$ regarded as a function of u. Now, it is not too difficult to establish (C.4) and (C.5) since (C.1) implies $I_{\mathcal{X}\wedge\mathcal{Z}|\mathcal{Y}}(p \times q \times rW_a)$ does not depend on q, and (C.2) implies $I_{\mathcal{Y}\wedge\mathcal{Z}|\mathcal{X}}(p \times q \times rW_a)$ does not depend on p. To establish (C.6) is extemely tedious, and we only sketch the derivation:

Clearly, $m(\frac{1}{2}) = 0$. To show that $p = \frac{1}{2}$ is the only possible solution, it is sufficient to show that m is a strictly increasing function on $(0, 1)$. This can be accomplished by showing that $m' > 0$ on $(0, 1)$. In fact, we show that m' has a unique minimum at $p = \frac{1}{2}$, and that $m'(\frac{1}{2}) > 0$. To show that m' has a unique minimum at $p = \frac{1}{2}$, we show that the only solution of $m''(p) = 0$ is $p = \frac{1}{2}$, and that $m'''(\frac{1}{2}) > 0$. Showing that $m''(p) = 0$ has the unique solution $p = \frac{1}{2}$ is the tedious part of the task. Having done all of the above, it follows that f has a *unique* maximum on $(0, 1) \times (0, 1)$. With only a little more work, it is easy to verify that $(\frac{1}{2}, \frac{1}{2})$ maximizes f on $[0, 1] \times [0, 1]$. $\square$

REFERENCES

[1] R. Ahlswede, "Multi-way communication channels," in *Proc. 2nd Int. Symp. Inform. Theory* (Tsahkadsor, Armenian S.S.R.), pp. 23-52, 1971. (Publishing House of the Hungarian Academy of Sciences, 1973.)

[2] R. Ahlswede, "An elementary proof of the strong converse theorem for the multiple-access channel," *J. Comb., Inform. Syst. Sci.*, vol. 7, no. 3, pp. 216-230, 1982.

[3] P. Billingsley, *Probability and Measure*. New York: Wiley, 1979.

[4] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. New York: Academic, 1981.

[5] I. Csiszár and J. Körner, "On the capacity of the arbitrarily varying channel for the maximum probability of error," *Z. Wahrscheinlichkeitstheorie verw. Geb.*, vol. 57, pp. 87-101, 1981.

[6] I. Csiszár, J. Körner, and K. Marton, "A new look at the error exponent of discrete memoryless channels," preprint, presented at the IEEE Int. Symp. Information Theory, Cornell University, Ithaca, NY, 1977.

[7] I. Csiszár and P. Narayan, "Arbitrarily varying channels with constrained inputs and states," *IEEE Trans. Inform. Theory*, vol. IT-34, pp. 27-34, Jan. 1988.

[20] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inform. Theory*, vol. IT-19, pp. 471-480, July 1973.

[21] E. C. van der Meulen, "Recent coding theorems and converses for multi-way channels. Part II: The multiple-access channel (1976-1985)," preprint.

[22] J. van Tiel, *Convex Analysis, An Introductory Text.* Chichester: Wiley, 1984.