Arbitrarily Varying Multiple-Access Channels Part I—Ericson's Symmetrizability Is Adequate, Gubner's Conjecture Is True

Rudolf Ahlswede and Ning Cai

Abstract— **In 1981 Jahn used the elimination technique of the first author to determine the average errors capacity region of an arbitrarily varying multiple-access channel (AVMAC), when this region has nonempty interior. *Here we remove this restriction.* In his thesis (1990), Gubner missed this result because he used the first author's first approach to the MAC, which is based on conditional decoding, and not the first author's second approach, which is based on maximum likelihood decoding. This second approach was originally needed for a kind of compound MAC. For the AVMAC the difference between the approaches is naturally even more essential.**

Index Terms— **Ahlswede's dichotomy, arbitrarily varying multiple-access channels, capacity region, jamming.**

I. Introduction

The discovery of [3] was the following *dichotomy*: the capacity $C(\mathcal{W})$ of an arbitrarily varying channel (AVC) $\mathcal{W} = \{W(\cdot|\cdot, s)\colon \mathcal{X} \to \mathcal{Z}, s \in \mathcal{S}\}$ with input alphabet $\mathcal{X}$ and output alphabet $\mathcal{Z}$ under the average error criterion equals either zero or else equals the random code capacity

$$C_R(\mathcal{W}) = \max_{P \in \mathcal{P}(\mathcal{X})} \min_{W \in \overline{\mathcal{W}}} I(P, W)$$

where $I(P, W)$ denotes the mutual information for input distribution P and channel W, $\overline{\mathcal{W}}$ is the convex hull of $\mathcal{W}$, and $\mathcal{P}(\mathcal{X})$ is the set of all probability distributions on $\mathcal{X}$.

The problem of positivity of $C(\mathcal{W})$ was also addressed in [3]. It was shown that the separability of two random words is necessary and sufficient for $C(\mathcal{W}) > 0$. Moreover, if $\overline{\mathcal{W}}$ equals the row-convex hull

$$\bar{\bar{\mathcal{W}}} = \{W\colon W(\cdot|x) \in \overline{\mathcal{W}}(x) = \text{conv}\{W(\cdot|x) \in \mathcal{W}\}\}$$

then already separability of two random letters, that is, for some $P, P' \in \mathcal{P}(\mathcal{X})$

$$\left\{\sum_x P(x)W(\cdot|x)\colon W \in \bar{\bar{\mathcal{W}}}\right\} \cap \left\{\sum_x P'(x)W(\cdot|x)\colon W \in \bar{\bar{\mathcal{W}}}\right\} = \emptyset \tag{1}$$

is necessary and sufficient for $C(\mathcal{W}) > 0$. Arbitrarily varying channels are designed as a robust channel model—in jamming, for instance. If the jammer can respond to every individual message, then the communicators are forced to use the maximal error criterion (see discussion in [3]). Then one readily verifies (see again [3]) that the coding problems for $\mathcal{W}$, $\overline{\mathcal{W}}$, and $\bar{\bar{\mathcal{W}}}$ are all equivalent; that is, we can assume $\mathcal{W} = \bar{\bar{\mathcal{W}}}$. Using the average error criterion it is of course also more realistic to assume that $\overline{\mathcal{W}} = \bar{\bar{\mathcal{W}}}$ and, therefore, to confine oneself (as in [3]) to the "single letter" condition (1).

Still there remained the mathematical problem of giving a single-letter characterization for positivity of a general $\mathcal{W}$ with $\overline{\mathcal{W}} \neq \bar{\bar{\mathcal{W}}}$.

The key step was made by T. Ericson [5] by introducing an adequate concept. He called $\mathcal{W}$ symmetrizable iff for a stochastic $E\colon \mathcal{X} \to \mathcal{S}$

$$\sum_s W(z|x, s)E(s|x') = \sum_s W(z|x', s)E(s|x, \qquad \text{for all } x, x' \in \mathcal{X} \text{ and } z \in \mathcal{Z}. \tag{2}$$

After having verified that for positivity of $C(\mathcal{W})$ it is necessary that $\mathcal{W}$ be nonsymmetrizable Ericson conjectured that this is also sufficient.

To prove this is by no means easy. However, it was easy for an expert familiar with the new coding methods of [4] (maximum probability (not likelihood) decoding in conjunction with large deviational ideas). The actual results of [4] are for the harder case of *maximal* errors. The proof has been recast in another language in [9]. It also is the germ of the proof of Csiszár/Narayan in [6] where Ericson's conjecture has been established. Finally, in the theory of identification the cycle closes: both, the nonsingle letter separability by random words and symmetrizability are used and needed!

Another development started with the earlier method of proof in [3], the so called "elimination technique." It is a forerunner of what is called now *derandomization* in computer science. It converts random codes into deterministic codes. Jahn [8] used it to extend the results of [3] to multiuser channels, in particular to the multiple-access channel (MAC). For an arbitrarily varying MAC (AVMAC), defined by a set $\mathcal{W} = \{W(\cdot|\cdot, \cdot, s)\colon s \in \mathcal{S}\}$ of stochastic $(\mathcal{X} \times \mathcal{Y} \to \mathcal{Z})$-matrices, he characterized the region of achievable rates $\mathcal{R}(\mathcal{W})$ under the average error criterion as follows.

For a pair of RV's (X, Y) with joint distribution $P_{XY} = P_X \cdot P_Y$ define the set $\mathcal{R}(X, Y)$ of pairs $(\mathcal{R}_1, \mathcal{R}_2)$ satisfying

$$\begin{aligned} 0 \le R_1 &\le \inf I(X \wedge Z|Y), \\ 0 \le R_2 &\le \inf I(Y \wedge Z|X) \\ R_1 + R_2 &\le \inf I(XY \wedge Z) \end{aligned} \tag{3}$$

where all infima range over the quadruples (X, Y, Z, S) with

$$P_{XYZS}(x, y, z, s) = P_X(x)P_Y(y)P_S(s)W(z|x, y, z).$$

Furthermore, write $\mathcal{R}_R(\mathcal{W})$ for the closed convex hull of $\bigcup_{(X,Y)} \mathcal{R}(X, Y)$ and write "int(A)" for the topological interior of a set $A \subset \mathbb{R}^2$.

Theorem (Jahn [8]):

$$\mathcal{R}(\mathcal{W}) = \mathcal{R}_R(\mathcal{W}), \qquad \text{when int}(\mathcal{R}(\mathcal{W})) \neq \emptyset.$$

After the result of Csiszár/Narayan existed, it was natural to try to characterize int$(\mathcal{R}(\mathcal{W})) \neq \emptyset$ in terms of symmetrizability. This was the topic of Gubner's thesis [7], where he introduced (among others) the following symmetrizability conditions for the AVMAC:

Manuscript received September 17, 1996; revised May 25, 1998. Presented at the IEEE International Symposium on Information Theory, Ulm, Germany, June 29–July 4, 1997.

The authors are with the Fakultät für Mathematik, Universität Bielefeld, Bielefeld, 33501 Germany.

Communicated by S. Shamai, Associate Editor for Shannon Theory.

Publisher Item Identifier S 0018-9448(99)01407-8.

i) $\mathcal{W}$ is $(\mathcal{X}, \mathcal{Y})$ symmetrizable iff for a stochastic $\sigma: \mathcal{X} \times \mathcal{Y} \to \mathcal{S}$

$$\sum_s W(z|x, y, s)\sigma(s|x', y') = \sum_s W(z|x', y', s)\sigma(s|x, y),$$

for all $x, x' \in \mathcal{X}$, $y, y' \in \mathcal{Y}$ and $z \in \mathcal{Z}$.

ii) $\mathcal{W}$ is $\mathcal{X}$ symmetrizable iff for a stochastic $\sigma_1: \mathcal{X} \to \mathcal{S}$

$$\sum_s W(z|x, y, s)\sigma_1(s|x') = \sum_s W(z|x', y, s)\sigma_1(s|x),$$

for all $x, x' \in \mathcal{X}$, $y \in \mathcal{Y}$ and $z \in \mathcal{Z}$.

iii) $\mathcal{W}$ is $\mathcal{Y}$ symmetrizable iff for a stochastic $\sigma_2: \mathcal{Y} \to \mathcal{S}$

$$\sum_s W(z|x, y, s)\sigma_2(s|y') = \sum_s W(z|x, y', s)\sigma_2(s|y),$$

for all $x \in \mathcal{X}$, $y, y' \in \mathcal{Y}$ and $z \in \mathcal{Z}$.

Their connections are clearer in light of the following:

Example 1: $\mathcal{W}$ can be $(\mathcal{X}, \mathcal{Y})$ symmetrizable without being $\mathcal{X}$ symmetrizable or $\mathcal{Y}$ symmetrizable. Choose $\mathcal{X} = \mathcal{Y} = \mathcal{Z} = \mathcal{S} = \{0, 1\}$ and choose as distributions on $\mathcal{Z}$ for $s = 0$

$$W(\cdot|0, 0, 0) = W(\cdot|1, 1, 0) = (1, 0)$$
$$W(\cdot|0, 1, 0) = W(\cdot|0, 1, 0) = (\tfrac{1}{2}, \tfrac{1}{2})$$

and for $s = 1$

$$W(\cdot|0, 0, 1) = W(\cdot|1, 1, 1) = (\tfrac{1}{2}, \tfrac{1}{2})$$
$$W(\cdot|0, 1, 1) = W(\cdot|1, 0, 1) = (0, 1).$$

Now, $\mathcal{W}$ is $(\mathcal{X}, \mathcal{Y})$ symmetrizable, because for σ defined by

$$\sigma(s|x, y) = \begin{cases} 0, & \text{if } x = y \text{ and } s = 1, \text{ or } x \neq y \text{ and } s = 0 \\ 1, & \text{if } x = y \text{ and } s = 0, \text{ or } x \neq y \text{ and } s = 1, \end{cases}$$

we have for all x, x', y, y'

$$\sum_s W(\cdot|x, y, s)\sigma(s|x', y') = \sum_s W(\cdot|x', y', s)\sigma(s|x, y).$$

On the other hand, $\mathcal{W}$ is not $\mathcal{X}$ symmetrizable (and similarly not $\mathcal{Y}$ symmetrizable). Indeed, for $y = 0$, $x = 0$, $x' = 1$ a σ_1 satisfying

$$\sum_s W(\cdot|0, 0, s)\sigma_1(s|1) = \sum_s W(\cdot|1, 0, s)\sigma_1(s|0)$$

must be of the form $\sigma_1(0|0) = \sigma_1(1|1) = 1$. However, for this σ_1, for $y = 1$, $x = 0$, $x' = 1$

$$\sum_s W(\cdot|0, 1, s)\sigma_1(s|1) = (0, 1) \neq (1, 0) = \sum_s W(\cdot|1, 1, s)\sigma_1(s|0).$$

An example for the opposite relation was given in [7].

Gubner introduces two further concepts.

1) $\mathcal{W}$ is called $(\mathcal{X}, Q)$ symmetrizable iff for some $\sigma_1': \mathcal{X} \to \mathcal{S}$

$$\sum_s \left(\sum_y Q(y)W(z|x, y, s)\right)\sigma_1'(s|x') = \sum_s \left(\sum_y Q(y)W(z|x', y, s)\right)\sigma_1'(s|x)$$

for all x, x' and z.

2) $\mathcal{W}$ is called $(\mathcal{Y}, P)$ symmetrizable for $P \in \mathcal{P}(\mathcal{X})$ iff relations analogous to those in 1) hold.

Following the argument of Ericson he obtains his first result.

Theorem G1: If the AVMAC $\mathcal{W}$ is $(\mathcal{X}, \mathcal{Y})$ symmetrizable or $\mathcal{X}$ symmetrizable or $\mathcal{Y}$ symmetrizable, then $\text{int}(\mathcal{R}) = \emptyset$.

His second result goes in the other direction.

Theorem G2: If the AVMAC $\mathcal{W}$ is not $\mathcal{X}$-symmetrizable and for some $P \in \mathcal{P}(\mathcal{X})$ not $(\mathcal{Y}, P)$ symmetrizable [resp. not $\mathcal{Y}$ symmetrizable and for some $Q \in \mathcal{P}(\mathcal{Y})$ not $(\mathcal{X}, Q)$ symmetrizable] then $\text{int}(\mathcal{R}(\mathcal{W})) \neq \emptyset$.

This sufficient condition is different from the necessary condition in Theorem G1 and Gubner conjectured the condition in Theorem G1 to be exact. Why did he or anybody else not settle the problem?

After some initial excitement about the concept of symmetrizability is this a new puzzle?

Fortunately a simple explanation can be given. Gubner extends the decoding rule of [6] to the multiple-access situation by following Ahlswede's [1] approach of "conditional decoding" (one message, say of the $\mathcal{X}$ encoder, is decoded against the average over the messages of the $\mathcal{Y}$ encoder; then the receiver, after knowing this message, uses this knowledge is in decoding the message of the $\mathcal{Y}$ encoder—and vice versa).

This is a suboptimal decoding rule, but suited for the discovery of the capacity theorem for the MAC. Its drawback for systems of channels was soon realized by Ahlswede [2]. For the derivation of the capacity theorem for the *compound* MAC he introduced, therefore, maximum likelihood decoding, which also could be analyzed.

For the AVMAC suboptimality of the conditional decoding rule becomes even more significant. In fact, Gubner does not even get by his approach the region established by Jahn if $\text{int}(\mathcal{R}(\mathcal{W})) \neq \emptyset$.

Here is our new result.

Theorem 1 (Proof of Gubner's conjecture): For the AVMAC $\mathcal{W}$ $\text{int}(\mathcal{R}(\mathcal{W})) \neq \emptyset$ iff $\mathcal{W}$ is not $(\mathcal{X}, \mathcal{Y})$ symmetrizable, not $\mathcal{X}$ symmetrizable, and not $\mathcal{Y}$ symmetrizable.

Proof: Having read [2] (also a good literature for experts on "turbo codes") and understood our previous reasoning, this should be a good exercise for people in "Shannon Theory."

Proof: Instead of proving the whole capacity theorem all over again, we only show how to get positive rates. (Combining this with Jahn's result, the capacity theorem follows.)

This makes calculations and formalisms much simpler. We use two auxiliary results. The Lemma 1 in Section II concerns the coding rule and the Lemma 2 in Section III large deviational properties to cope with the $|\mathcal{S}^n|$; that is, exponentially many individual channels. Finally, we prove the Theorem in Section IV.

II. Decoding

We begin with the description of a decoding rule for the AVMAC $\mathcal{W}$. Fix two RV's X and Y such that for their distributions $P_X \in \mathcal{P}(n, \mathcal{X})$ and $P_Y \in \mathcal{P}(n, \mathcal{Y})$; that is, they are n-empirical distributions. $\mathcal{T}_X^n$ and $\mathcal{T}_Y^n$ are the sets of typical sequences with relative frequencies specified by P_X, respectively, P_Y.

For sets $\mathcal{U} \subset \mathcal{T}_X^n$ and $\mathcal{V} \subset \mathcal{T}_Y^n$ (of codewords) and (small) positive numbers ξ, ζ_1, ζ_2, and ζ we define decoding sets $D_{uv}(u \in \mathcal{U}, v \in \mathcal{V})$ as follows: $z^n \in D_{uv}$ iff there is an $s^n \in \mathcal{S}^n$ and a quadruple (X, Y, S, Z) of RV's with $(u, v, s^n, z^n) \in \mathcal{T}_{XYSZ}^n$ satisfying simultaneously the conditions

O)

$$D(P_{XYSZ}||P_X \times P_Y \times P_S \times W) < \xi \tag{4}$$

I) If there are $u' \neq u$, $v' \neq v$, $s'^n \in \mathcal{S}^n$, and RV's X', Y', and S' such that $(u, u', v, v', s^n, s'^n, z^n) \in \mathcal{T}^n_{XX'YY'SS'Z}$ and

$$D(P_{X'Y'S'Z} \| P_{X'} \times P_{Y'} \times P_{S'} \times W) < \xi \tag{5}$$

then

$$I(XYZ \wedge X'Y'|S) < \zeta. \tag{6}$$

II) If there are $u' \neq u$, $s'^n \in \mathcal{S}^n$, and RV's X' and S' such that

$$(u, u', v, s^n, s'^n, z^n) \in \mathcal{T}^n_{XX'YSS'Z}$$

and

$$D(P_{X'YS'Z} \| P_{X'} \times P_Y \times P_{S'} \times W) < \xi \tag{7}$$

then

$$I(XYZ \wedge X'|S) < \zeta_1 \tag{8}$$

and [symmetrically to II)]

III) if there are $v' \neq v$, $s'^n \in \mathcal{S}^n$, and RV's Y' and S' such that

$$(u, v, v', s^n, s'^n, z^n) \in \mathcal{T}^n_{XYY'SS'Z}$$

and

$$D(P_{XY'S'Z} \| P_X \times P_{Y'} \times P_{S'} \times W) < \xi \tag{9}$$

then

$$I(XYZ \wedge Y'|S) < \zeta_2. \tag{10}$$

Of course, we have to ensure that the D_{uv}'s are disjoint. Here is where the three nonsymmetrizabilities come in.

Lemma 1: We assume that $\mathcal{W}$ is not $(\mathcal{X}, \mathcal{Y})$, $\mathcal{X}$, and not $\mathcal{Y}$ symmetrizable. For (small) $\alpha, \beta > 0$ consider distributions $P_X \in \mathcal{P}(n, \mathcal{X})$, $P_Y \in \mathcal{P}(n, \mathcal{Y})$ with

$$\min_x P_X(x) \geq \alpha \text{ and } \min_y P_Y(y) \geq \beta.$$

One can choose positive ξ, ζ_1, ζ_2, and ζ (depending on α, β, and $\mathcal{W}$) such that for any sets of codewords $\mathcal{U} \subset \mathcal{T}^n_X$ and $\mathcal{V} \subset \mathcal{T}^n_Y$ the decoding sets defined above are disjoint

$$D_{uv} \cap D_{u'v'} = \varnothing \text{ for } (u, v) \neq (u', v'). \tag{11}$$

More specifically,

a) the condition non-i) and rules O) and I) imply

$$D_{uv} \cap D_{u'v'} = \varnothing, \text{ if } u \neq u' \textbf{ and } v' \neq v'.$$

b) the condition non-ii) [resp. non-iii)] and rules O) and II) [resp. III)] imply

$$D_{uv} \cap D_{u'v} = \varnothing, \text{ if } u \neq u'$$

(respectively, $D_{uv} \cap D_{uv'} = \varnothing$, if $v \neq v'$).

Proof: See Appendix.

III. Large Deviational Methods

We present here our second auxiliary result, Lemma 2 below. It is analogous in formulation to Lemma 3 of [6] and Theorem C1 of [7]. However, the underlying idea based on large deviations for sums of RV's is solely due to [4]. In fact, we derive first from [4, Lemma 1(b)]

Lemma [6, A1]: Let $Z_1, \cdots, Z_N$ be arbitrary RV's and $f_i(Z_1, \cdots, Z_i)$ be arbitrary with $0 \leq f_i \leq 1$ $(1 \leq i \leq N)$. Then the condition

$$E[f_i(Z_1, \cdots, Z_i)|Z_1, \cdots, Z_{i-1}] \leq a \text{ a.s. } 1 \leq i \leq N$$

implies that

$$\Pr\left\{\frac{1}{N}\sum_{i=1}^N f_i(Z_1, \cdots, Z_i) > t\right\} \leq \exp\{-N(t - a \log e)\}.$$

Lemma [4, 1(b)]: Let $T_1, \cdots, T_k$ be a sequence of RV's then

$$\Pr\left\{\frac{1}{K}\sum_{i=1}^K T_i \geq a\right\} \leq e^{-aK/2b} \prod_{i=1}^K \max_{(t_1, \cdots, t_{i-1})} E(1 + b^{-1}T_i | T_1 = t_1, \cdots, T_{i-1} = t_{i-1})$$

if $T_1, \cdots, T_k$ take values in $[0, b]$.

Let $f_i(Z_1, \cdots, Z_i) = T_i$. Then $E(1 + T_i|T^{i-1}) = E(E(1 + T_i|Z^{i-1})|T^{i-1})$, since T^{i-1} is a function of Z^{i-1}. Therefore,

$$\max_{(t_1, \cdots, t_{i-1})} E(1 + T_i | T_1 = t_1, \cdots, T_{i-1} = t_{i-1}) \leq \max_{(z_1, \cdots, z_{i-1})} E(1 + T_i | Z_1 = z_1, \cdots, Z_{i-1} = z_{i-1})$$

which with Lemma 1(b) of [4] and the well-known inequality $\ell n(1 + x) \leq x$ together implies that

$$\Pr\left\{\frac{1}{N}\sum_{i=1}^N f_i(Z_1, \cdots, Z_i) > t\right\} \leq \exp_e\left\{-\frac{Nt}{2} + Na\right\} = \exp\left\{-N\left(\frac{\log e}{2}t - a \log e\right)\right\}$$

under the present assumption. The constant $\log e/2$ obviously makes no difference. Next we derive from [6, *Lemma A1*].

Proposition: For RV's $\tilde{U}_0, \cdots, \tilde{U}_m$ and functions $g_i(\tilde{U}_0, \tilde{U}_1, \cdots, \tilde{U}_i)$ with $0 \leq g_i \leq 1$ $(1 \leq i \leq m)$ the condition

$$\mathbb{E}(g_i(\tilde{U}_0, \tilde{U}_1, \cdots, \tilde{U}_i)|\tilde{U}_0, \tilde{U}_1, \cdots, \tilde{U}_{i-1}) \leq a, \quad \text{for } 1 \leq i \leq m \text{ (a.s.)} \tag{12}$$

implies

$$\Pr\left\{\sum_{i=1}^m g_i(\tilde{U}_0, \tilde{U}_1, \cdots, \tilde{U}_i) > mb\right\} \leq \left(\frac{e}{2}\right)^a \exp\{-m(b - a \log e)\}. \tag{13}$$

Proof: Take $N = m + 1$, $Z_1 = \tilde{U}_0$, $f_1 = a$, $Z_i = \tilde{U}_{i-1}$, $f_i = g_{i-1}$ (for $2 \leq i \leq m$) and $t = ((mb + a)/(m + 1))$ in *Lemma A1*. Then (13) follows. ∎

For sets of codewords $\mathcal{U} \subset \mathcal{T}^n_X$ and $\mathcal{V} \subset \mathcal{T}^n_Y$ of cardinality

$$|\mathcal{U}| = |\mathcal{V}| = M \tag{14}$$

the numbers $r = \frac{1}{n}\log M$, and $\varepsilon > 0$ define now for $s^n \in \mathcal{S}^n$ three sets

$$\mathcal{A}_\varepsilon(s^n) = \{(u, v) \in \mathcal{U} \times \mathcal{V}: (u, v, s^n) \in \mathcal{T}^n_{XYS} \text{ implies } D(P_{XYS} \| P_X \times P_Y \times P_S) \leq \varepsilon\}, \tag{15}$$

$$\mathcal{B}_\varepsilon(s^n) = \{u \in \mathcal{U}: \text{for some } u' \neq u, v \neq v' \text{ implies } I(X \wedge X'YY'S) \leq 3r + \varepsilon\} \tag{16}$$

and

$$\mathcal{C}_\varepsilon(s^n) = \{v \in \mathcal{V}: \text{for some } u \neq u' \text{ and } v' \neq v \cdot (u, u', v, v', s^n) \in \mathcal{T}^n_{XX'YY'S} \text{ implies } I(Y \wedge XX'Y'S) \leq 3r + \varepsilon\}. \tag{17}$$

Lemma 2: For any $0 < \varepsilon < \delta$ and all $n \geq n_0(\varepsilon, \delta)$, suitable for $M = 2^{nr}$ with $r \geq \delta$ there exist for $P_X \in \mathcal{P}(n, \mathcal{X})$, $P_Y \in \mathcal{P}(n, \mathcal{Y})$ sets of codewords $\mathcal{U}$, $\mathcal{V}$ as in (14) such that for all $s^n \in \mathcal{S}^n$

$$|\mathcal{A}^c_\varepsilon(s^n)| \leq 2^{-(n\varepsilon/4)}M^2 \tag{18}$$

and

$$|\mathcal{B}^c_\varepsilon(s^n)|, |\mathcal{C}^c_\varepsilon(s^n)| \leq 2^{-(n\varepsilon/4)}M. \tag{19}$$

Proof: Let $U_i, V_j (1 \le i, j \le M)$ be independent, uniformly distributed RV's taking values in $\mathcal{T}_X^n$ and $\mathcal{T}_Y^n$, respectively.

To obtain (19), for any quintuple $(XX'YY'S)$ of RV's with $I(X \wedge X'YY'S) > 3r + \varepsilon$ and $s^n \in \mathcal{S}^n$ define

$$g_i(V^M, U_1, \cdots, U_i) = \begin{cases} 1, & \text{if exist } i' < i \text{ and } j \neq j' \\ & \text{with } (U_i, U_i', V_j, V_j', s^n) \in \mathcal{T}_{XX'YY'S}^n \\ 0, & \text{otherwise.} \end{cases} \tag{20}$$

Since $\mathcal{P}(n, \mathcal{X} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Y} \times \mathcal{S})$ grows polynomially in n by the symmetry in i and i' it suffices to show that the event

$$\sum_{i=1}^{M} g_i(V^M, U_1, \cdots, U_i) > 2^{-(n\varepsilon/2)} M \tag{21}$$

has double exponentially small probability.

This follows from the Proposition with the choices $\tilde{U}_0 = V^M$, $\tilde{U}_i = U_i$, $m = M$, $a = M^3 \exp\{-n(I(X \wedge X'YY'S) - \frac{\varepsilon}{4})\}$, and $b = 2^{-(n\varepsilon/2)}$. Analogously $|\mathcal{C}_\varepsilon(s^n)|$ can be bounded and (19) is established. To obtain (18), for a fixed $s^n \in \mathcal{S}^n$ and triple of RV's (X, Y, S) with

$$D(P_{XYS} \| P_X \times P_Y \times P_S) > \varepsilon \tag{22}$$

we define

$$f(V_j) = \begin{cases} 1, & \text{if } (V_j, s^n) \in \mathcal{T}_{YS}^n \\ 0, & \text{otherwise} \end{cases} \tag{23}$$

and for $y^n \in \mathcal{T}_{Y|S}^n(s^n)$

$$f^{(y^n)}(U_i) = \begin{cases} 1, & \text{if } (U_i, y^n, s^n) \in \mathcal{T}_{XYS}^n \\ 0, & \text{otherwise} \end{cases} \tag{24}$$

$(1 \le i \le M)$.

We observe that the event

$$'' |\{(U_i, V_j): (U_i, V_i, s^n) \in \mathcal{T}_{XYS}^n\}| > 2^{-(n\varepsilon/2)} M^2 '' \tag{25}$$

is contained in the union of the event

$$'' \sum_{j=1}^{M} f_j(V_j) > \exp\left\{n\left(|r - I(Y \wedge S)|^+ + \frac{\varepsilon}{4}\right)\right\} '' \tag{26}$$

and the events

$$'' \sum_{i=1}^{M} f_i^{(y^n)}(U_i) > \exp\left\{n\left(|r - I(X \wedge YS)|^+ + \frac{\varepsilon}{4}\right)\right\} '' \tag{27}$$

$(y^n \in \mathcal{T}_{Y|S}^n(s^n))$.

Here we use the following facts.

1) When $r - I(Y \wedge S) \ge 0$ and $r - I(X \wedge YS) \ge 0$, then the product of the RHS's in (26) and (27) is $M^2 \exp\{-n(D(P_{XYS} \| P_X \times P_Y \times P_S) - (\varepsilon/2))\} < 2^{-(n\varepsilon/2)} M^2$ [by (22)], because $M = \exp\{nr\}$ and $D(P_{XYS} \| P_X \times P_Y \times P_S) = (H(YS) + H(X) - H(XYS)) + (H(Y) + H(S) - H(YS)) = I(X \wedge YS) + I(Y \wedge S)$.
2) When $r - I(Y \wedge S) < 0$ [or $r - I(X \wedge YS) < 0$], then the RHS in (26) [or in (27)] equals $2^{(n\varepsilon/4)}$. On the other hand the LHS of (27) [and of (26)] is at most M. Thus, their product $2^{(n\varepsilon/4)} M < 2^{-(n\varepsilon/2)} M^2$.

Thus, to obtain (18) for all (fixed) $s^n \in \mathcal{S}^n$ it suffices to show that the events (26) and (27) have double exponentially small probabilities.

The former is done by setting in the Proposition $\tilde{U}_0 =$ constant, $\tilde{U}_i = V_i$, $g_i = f_i$, $m = M$, $a = \exp\{-n(I(Y \wedge S) - \frac{\varepsilon}{8})\}$, and $b = \exp\{n(|r - I(Y \wedge S)|^+ + \frac{\varepsilon}{4} - r)\}$. The latter is done by setting $\tilde{U}_0 =$ constant, $\tilde{U}_i = U_i$, $g_i = f_i^{(y^n)}$, $m = M$, $a = \exp\{-n(I(X \wedge YS) - \frac{\varepsilon}{8})\}$, and $b = \exp\{n(|r - I(X \wedge YS)|^+ + \frac{\varepsilon}{4} - r)\}$ for $y^n \in \mathcal{T}_{Y|S}^n(s^n)$.

In both cases, we use the fact that for $I = I(Y \wedge S)$ or $I(X \wedge YS)$ $(r - I)^+ + \frac{\varepsilon}{4} - r = |r - I|^+ + \frac{\varepsilon}{4} - |r - I| - I \ge -I + \frac{\varepsilon}{4} > -(I - \frac{\varepsilon}{8})$. ■

IV. Proof of Theorem

For fixed positive α and β choose $P_X \in \mathcal{P}(n, \mathcal{X})$ and $P_Y \in \mathcal{P}(n, \mathcal{Y})$ with $\min_x P_X(x) \ge \alpha$ and $\min_y P_Y(y) \ge \beta$. Also choose positive ξ, ζ_1, ζ_2, and ζ according to Lemma 1 so small that the decoder with rules O)–III) is well-defined.

Next let

$$\zeta^* = \min\{\zeta, \zeta_1, \zeta_2\} \tag{28}$$

and choose ε, δ, r and sufficiently large n such that $M = 2^{nr}$ is an integer and

$$\varepsilon < \frac{\xi}{2}, \quad 0 < \varepsilon < \delta \le r \le \frac{1}{11}\zeta^*. \tag{29}$$

Then, by *Lemma 2*, we get sets of codewords $\mathcal{U}$, $\mathcal{V}$ with rate-vector (r, r) and properties (18) and (19).

The code $\mathcal{U} \times \mathcal{V}$ is decodable by Lemma 1. It remains to be seen that for every $s^n \in \mathcal{S}^n$ the average probability of decoding error is exponentially small.

For this let us first fix s^n. It suffices to prove that for $\mathcal{A}_\varepsilon(s^n)$, $\mathcal{B}_\varepsilon(s^n)$, and $\mathcal{C}_\varepsilon(s^n)$ and all $(u, v) \in \mathcal{A}_\varepsilon(s^n) \cap [\mathcal{B}_\varepsilon(s^n) \times \mathcal{C}_\varepsilon(s^n)] \triangleq \mathcal{D}_\varepsilon(s^n)$ (say), $W^n(D_{uv}^c | u, v, s^n)$ is exponentially small, because by (18) and (19)

$$\frac{1}{M^2} \sum_{(u,v) \in \mathcal{U} \times \mathcal{V}} W^n(D_{uv}^c | u, v, s^n) \le \frac{1}{M^2} \sum_{(u,v) \in \mathcal{D}_\varepsilon(s^n)} W^n(D_{uv}^c | u, v, s^n) + 3 \cdot 2^{-(n\varepsilon/4)}.$$

So let us fix $(u, v) \in \mathcal{D}_\varepsilon(s^n)$ and use the decomposition

$$D_{uv}^c = E_0 \cup E_1 \cup E_2 \cup E_3 \tag{30}$$

where an output sequence z^n falls into E_0, E_1, E_2, or E_3, when for the fixed s^n and (u, v) the decoding rules O), I), II), or III) are violated, respectively.

We now upperbound the probabilities of these E_i. For this it is convenient to use the abbreviation

$$\gamma(n) \triangleq (n+1)^{|\mathcal{X}|^2 |\mathcal{Y}|^2 |\mathcal{S}| |\mathcal{Z}|}. \tag{31}$$

Suppose then that

$$(u, v, s^n) \in \mathcal{T}_{XYS}^n, \tag{32}$$

then by (15) and since $(u, v) \in \mathcal{D}_\varepsilon(s^n)$

$$D(P_{XYS} \| P_X \times P_Y \times P_S) \le \varepsilon. \tag{33}$$

Now, define

$$\mathcal{Q}_0 \triangleq \{(X, Y, S, Z): D(P_{XYSZ} \| P_X \times P_Y \times P_S \times W) \ge \xi \text{ and (32) holds }\}. \tag{34}$$

Then, by (4), (27), (33), and (34)

$$\begin{aligned}
&W^n(E_0|u, v, s^n)\\
&= \sum_{(X,Y,S,Z)\in\mathcal{Q}_O} W^n(T^n_{Z|XYS}(u, v, s^n)|u, v, s^n)\\
&\le \gamma(n) \max_{(X,Y,S,Z)\in\mathcal{Q}_O} \exp\{-nD(P_{Z|XYS}\|W|P_{XYS})\}\\
&\quad \text{[by (4.4)]}\\
&\le \gamma(n) \max_{(X,Y,S,Z)\in\mathcal{Q}_0} \exp\{-n(D(P_{Z|XYS}\|W|P_{XYS})\\
&\quad - \varepsilon + D(P_{XYS}\|P_X \times P_Y \times P_S))\} \text{ [by (33)]}\\
&= \gamma(n) \max_{(X,Y,S,Z)\in\mathcal{Q}_0} \exp\{-n(D(P_{XYSZ}\|P_X \times P_Y\\
&\quad \times P_S \times W) - \varepsilon)\}\\
&\le \gamma(n)2^{-n(\xi-\varepsilon)}\\
&\le \gamma(n)2^{-(n\xi/2)}. \qquad (35)
\end{aligned}$$

To upperbound $W^n(E_1|u, v, s^n)$ we define

$$\begin{aligned}
\mathcal{Q}_1 \triangleq \{&(X, X', Y, Y', S, Z)\text{: (32) holds and}\\
&I(XYZ \wedge X'Y'|S) \ge \zeta\} \qquad (36)
\end{aligned}$$

and

$$\begin{aligned}
&\mathcal{J}_1(X, X', Y, Y', S)\\
&\triangleq \{(u', v') \in \mathcal{U} \times \mathcal{V}\text{: } u' \ne u\\
&\quad v' \ne v \text{ and } (u, u', v, v', s^n) \in \mathcal{T}^n_{XX'YY'S}\}. \qquad (37)
\end{aligned}$$

Thus by the definition of E_1 and decoding rule I)

$$\begin{aligned}
E_1 \subset &\bigcup_{(X,X',Y,Y',S,Z)\in\mathcal{Q}_1} \bigcup_{(u',v')\in\mathcal{J}_1(X,X',Y,Y',S)}\\
&\cdot \mathcal{T}^n_{Z|XX'YY'S}(u, u', v, v', s^n). \qquad (38)
\end{aligned}$$

However, by the definitions of $\mathcal{B}_\varepsilon(s^n)$, $\mathcal{C}_\varepsilon(s^n)$, and $\mathcal{D}_\varepsilon(s^n)$ the existence of $(u', v') \in \mathcal{J}_1(X, X', Y, Y', S)$ implies that $I(X \wedge X'YY'S) \le 3r + \varepsilon$ and $I(Y \wedge XX'Y'S) \le 3r + \varepsilon$, which yields

$$I(XY \wedge X'Y'|S) = I(X \wedge X'Y'|S) + I(Y \wedge X'Y'|XS) \le 6r + 2\varepsilon. \qquad (39)$$

Thus, from (36)–(39), we have

$$\begin{aligned}
&W^n(E_1|u, v, s^n)\\
&\le \sum_{(X,X',Y,Y',S,Z)\in\mathcal{Q}_1} \sum_{(u',v')\in\mathcal{J}_1(X,X',Y,Y',S)}\\
&\quad \cdot W^n(T^n_{Z|XX'YY'S}(u, u'v, v', s^n)|u, v, s^n)\\
&\le \gamma(n)M^2 \max_{\substack{(X,X',Y,Y'S,Z)\in\mathcal{Q}_1\\ \mathcal{J}_1(X,X',Y,Y'S)\ne\varnothing}}\\
&\quad \cdot \exp\{n(H(Z|XX'YY'S) - H(Z|XYS))\}\\
&\le \gamma(n)M^2 \max_{(X,X',Y,Y',S,Z)\in\mathcal{Q}_1} \exp\{n(-I(Z \wedge X'Y'|XYS)\\
&\quad + 6r + 2\varepsilon - I(XY \wedge X'Y'|S))\} \text{ [by (39)]}\\
&= \gamma(n)M^2 \max_{(X,X',Y,Y',S,Z)\in\mathcal{Q}_1} \exp\{n(-I(XYZ \wedge X'Y'|S)\\
&\quad + 6r + 2\varepsilon)\}\\
&\le \gamma(n) \exp\{n(8r + 2\varepsilon - \zeta)\} \text{ [by (36)]} \qquad (40)
\end{aligned}$$

which is exponentially small by (28) and (29).

Finally, we upperbound $W^n(E_2|u, v, s^n)$. By symmetry $W^n(E_3|u, v, s^n)$ can be bounded analogously.

Define now

$$\begin{aligned}
\mathcal{Q}_2 \triangleq \{&(X, X', Y, S, Z)\text{: (32) holds and}\\
&I(XYZ \wedge X'|S) \ge \zeta_1\} \qquad (41)
\end{aligned}$$

and

$$\begin{aligned}
&\mathcal{J}_2(X, X', Y, S)\\
&\triangleq \{u'\text{: } u' \ne u \text{ and } (u, u', v, s^n) \in \mathcal{T}^n_{XX'YS}\}. \qquad (42)
\end{aligned}$$

Then, by the definition of E_2 and decoding rule II)

$$\begin{aligned}
E_2 \subset &\bigcup_{(X,X',Y,S,Z)\in\mathcal{Q}_2} \bigcup_{u'\in\mathcal{J}_2(X,X',Y,S)}\\
&\cdot \mathcal{T}^n_{Z|XX'YS}(u, u', v, s^n). \qquad (43)
\end{aligned}$$

Moreover, by (42), $\mathcal{J}_2(X, X', Y, S) \ne \varnothing$ implies the existence of RV Y' and codewords $u' \ne u$, $v' \ne v$ such that

$$(u, u', v, v', s^n) \in \mathcal{T}^n_{XX'YY'S} \qquad (44)$$

because one can pick any $u' \in J_2(X, X', Y, S)$ and any $v' \ne v$ and find the corresponding Y'.

In other words $\mathcal{J}_2(X, X', Y, S) \ne \varnothing$ implies the existence of a RV Y' such that $\mathcal{J}_1(X, X', Y, Y', S)$ [defined in (37)] is not empty, which by (39) implies that

$$I(XY \wedge X'|S) \le 6r + 2\varepsilon. \qquad (45)$$

From here and from (41)–(43), we have

$$\begin{aligned}
&W^n(E_2|u, v, s^n)\\
&\le \sum_{(X,X',Y,S,Z)\in\mathcal{Q}_2} \sum_{u'\in\mathcal{J}_2(X,X',Y,S)}\\
&\quad \cdot W^n(T^n_{Z|XX'YS}(u, u'v, s^n)|u \cdot v \cdot s^n))\\
&\le \gamma(n)M \max_{\substack{(X,X',Y,S,Z)\in\mathcal{Q}_2\\ \mathcal{J}_2(X,X',Y,S)\ne\varnothing}}\\
&\quad \cdot \exp\{n(H(Z|XX'YS) - H(Z|XYS))\}\\
&\le \gamma(n)M \max_{(X,X',Y,S,Z)\in\mathcal{Q}_2} \exp\{n(-I(Z \wedge X'|XYS)\\
&\quad + 6r + 2\varepsilon - I(XY \wedge X'|S))\} \text{ [by (45)]}\\
&= \gamma(n)M \max_{(X,X',Y,S,Z)\in\mathcal{Q}_2} \exp\{n(-I(XYZ \wedge X'|S)\\
&\quad + 6r + 2\varepsilon)\}\\
&\le \gamma(n) \exp\{n(7r + 2\varepsilon - \zeta_1)\}, \text{ [by (41)]} \qquad (46)
\end{aligned}$$

which is exponentially small by (28) and (29). ■

Appendix

Proof of Lemma 1: 1) Choose ξ and ξ_2 sufficiently small so that for all u, v, v' with $v \ne v'$,

$$D_{uv} \cap D_{uv'} = \varnothing. \qquad (A.1)$$

It is sufficient to show that in case there is a

$$z^n \in D_{uv} \cap D_{uv'} \ne \varnothing, \qquad (A.2)$$

$\xi + \xi_2$ is bounded from below by a positive number.

By condition non-iii), there exists a positive θ_2 such that for all σ_2: $\mathcal{Y} \to \mathcal{S}$

$$\max_{x,y,y',z} \left|\sum_s \sigma_2(s|y')W(z|x, y, s) - \sum_s \sigma_2(s|y)W(z|x, y', s)\right| > \theta_2. \qquad (A.3)$$

Since we can write for all $\tau_2, \tau_2' : \mathcal{Y} \to \mathcal{S}$,

$$\max_{x,y,y',z} \left| \sum_s \tau_2'(s|y') W(z|x,y,s) - \sum_s \tau_2(s|y) W(z|x,y',s) \right| \text{ as}$$

$$\max_{x,y,y',z} \left| \sum_s \tau_2(s|y') W(z|x,y,s) - \sum_s \tau_2'(s|y) W(z|x,y',s) \right|$$

just by exchanging the two sums and then y and y' we get

$$\begin{aligned}
&\max_{x,y,y',z} \left| \sum_s \tau_2'(s|y') W(z|x,y,s) - \sum_s \tau_2(s|y) W(z|x,y',s) \right| \\
&= \max_{x,y,y',z} \left| \sum_s \frac{\tau_2'(s|y')}{2} W(z|x,y,s) - \sum_s \frac{\tau_2(s|y)}{2} W(z|x,y',s) \right| \\
&\quad + \max_{x,y,y',z} \left| \sum_s \frac{\tau_2(s|y')}{2} W(z|x,y,s) - \sum_s \frac{\tau_2'(s|y)}{2} W(z|x,y',s) \right| \\
&\geq \max_{x,y,y',z} \left\{ \left| \sum_s \frac{\tau_2'(s|y')}{2} W(z|x,y,s) - \sum_s \frac{\tau_2(s|y)}{2} W(z|x,y',s) \right| \right. \\
&\qquad \left. + \left| \sum_s \frac{\tau_2(s|y')}{2} W(z|x,y,s) - \sum_s \frac{\tau_2'(s|y)}{2} W(z|x,y',s) \right| \right\} \\
&\geq \max_{x,y,y',z} \left| \sum_s \frac{\tau_2'(s|y') + \tau_2(s|y')}{2} W(z|x,y,s) - \sum_s \frac{\tau_2'(s|y) + \tau_2(s|y)}{2} W(z|x,y',s) \right|. \qquad \text{(A.4)}
\end{aligned}$$

Applying (A.3) to (A.4) for $\sigma_2 = ((\tau_2' + \tau_2)/2)$, we obtain for all $\tau_2, \tau_2' : \mathcal{Y} \to \mathcal{S}$

$$\max_{x,y,y',s} \left| \sum_s \tau_2'(s|y') W(z|x,y,s) - \sum_s \tau_2(s|y) W(z|x,y',s) \right| > \theta_2. \qquad \text{(A.5)}$$

Next, with the RV's in III) of the decoding rule, (A.2) implies, for $(u, v, v', z^n, s^n, s^n) \in T^n_{XYY'SS'Z}$ (4), (9), (10) and

$$I(XY'Z \wedge Y|S') < \zeta_2. \qquad \text{(A.6)}$$

By (4), (10), and the log–sum inequality, we have

$$\begin{aligned}
\xi + \zeta_2 &> D(P_{XYSZ} \| P_X \times P_Y \times P_S \times W) + I(XYZ \wedge Y'|S) \\
&= \sum_{x,y,s,z} P_{XYSZ}(x,y,s,z) \cdot \log \frac{P_{XYSZ}(x,y,s,z)}{P_X(x)P_Y(y)P_S(s)W(z|x,y,s)} \\
&\quad + \sum_{x,y,y',s,z} P_{XYY'SZ}(x,y,y',s,z) \cdot \log \frac{P_{Y'|XYSZ}(y'|x,y,s,z)}{P_{Y'|S}(y'|s)} \\
&= \sum_{x,y,y',s,z} P_{XYY'SZ}(x,y,y',s,z) \cdot \log \frac{P_{XYY'SZ}(x,y,y',s,z)}{P_X(x)P_Y(y)P_{Y'S}(y',s)W(z|x,y,s)} \\
&= \sum_{x,y,y',z} \sum_s P_{XYY'SZ}(x,y,y',s,z) \cdot \log \frac{P_{XYY'SZ}(x,y,y',s,z)}{P_X(x)P_Y(y)P_{Y'}(y)P_{S|Y'}(s,y')W(z|x,y,s)} \\
&\geq \sum_{x,y,y',z} P_{XYY'Z}(x,y,y',z) \cdot \log \frac{P_{XYY'Z}(x,y,y',z)}{P_X(x)P_Y(y)P_{Y'}(y') \sum_s P_{S|Y'}(s|y')W(z|x,y,s)}.
\end{aligned} \qquad \text{(A.7)}$$

By Pinsker's inequality [10] and (A.7),

$$\sum_{x,y,y',z} |P_{XYY'Z}(x,y,y',z) - P_X(x)P_Y(y)P_{Y'}(y) \cdot \sum_s P_{S|Y'}(s|y')W(z|x,y,s)| \leq c\sqrt{\xi + \zeta_2}. \qquad \text{(A.8)}$$

Similarly by (9) and (A.6) we have

$$\sum_{x,y,y'z} |P_{XYY'Z}(x,y,y',z) - P_X(x)P_{Y'}(y')P_Y(y) \cdot \sum_s P_{S'|Y}(s|y)W(z|x,y',s)| \leq c\sqrt{\xi + \zeta_2}. \qquad \text{(A.9)}$$

Equations (A.8) and (A.9) together imply

$$\sum_{x,y,y',z} P_X(x)P_Y(y)P_{Y'}(y') \left| \sum_s P_{S|Y'}(s|y')W(z|x,y,s) - \sum_s P_{S'|Y}(s|y)W(z|x,y',s) \right| \leq 2c\sqrt{\xi + \zeta_2} \qquad \text{(A.10)}$$

and, recalling that for all x, y, y' $P_X(x) \geq \alpha$, $P_Y(y) \geq \beta$ and $P_{Y'}(y') = P_Y(y') \geq \beta$

$$\max_{x,y,y',z} \left| \sum_s P_{S|Y'}(s|y')W(z|x,y,s) - \sum_s P_{S'|Y}(s|y)W(z|x,y',s) \right| \leq \frac{2c\sqrt{\xi + \zeta_2}}{\alpha\beta^2}. \qquad \text{(A.11)}$$

Comparing (A.11) with (A.5) for $\tau_2' = P_{S|Y'}$ and $\tau_2 = P_{S'|Y}$, we have $\xi + \zeta_2 \geq 1/4c^{-2}\alpha^2\beta^4\theta_2^2$.

2) Similarly, one can choose ξ and ζ_1 sufficiently small so that for all u, u', v, $u \neq u'$

$$D_{uv} \cap D_{u'v} = \emptyset.$$

3) Choose ξ and ζ sufficiently small so that for all u, v, u', v', $u \neq u'$, $v \neq v'$

$$D_{uv} \cap D_{u'v'} = \emptyset. \qquad \text{(A.12)}$$

As in 1), we shall show that the existence of a

$$z^n \in D_{uv} \cap D_{u'v'} \neq \emptyset \qquad \text{(A.13)}$$

implies that $\xi + \zeta$ is bounded from below by a positive number. Now by the condition non-i), there is a $\theta > 0$ such that for all $\sigma\colon \mathcal{X} \times \mathcal{Y} \to \mathcal{S}$

$$\max_{x, x', y, y', z} \left| \sum_s \sigma(s|x', y') W(z|x, y, s) - \sum_s \sigma(s|x, y) W(z|x', y', s) \right| > \theta. \quad \text{(A.14)}$$

Similarly to 1, since we can rewrite

$$\max_{x, x', y, y', z} \left| \sum_s \tau'(s|x', y') W(z|x, y, s) - \sum_s \tau(s|x, y) W(z|x', y', s) \right|$$

as

$$\max_{x, x', y, y', z} \left| \sum_s \tau(s|x', y') W(z|x, y, s) - \sum_s \tau'(s|x, y) W(z|x', y', s) \right|$$

we have for all τ, τ' $\mathcal{X} \times \mathcal{Y} \to \mathcal{S}$

$$\begin{aligned}
&\max_{x, x', y, y', z} \left| \sum_s \tau'(s|x', y') W(z|x, y, s) - \sum_s \tau(s|x, y) W(z|x', y', s) \right| \\
&\quad \geq \max_{x, x', y, y', z} \left\{ \left| \sum_s \frac{\tau'(s|x', y')}{2} W(z|x, y, s) - \sum_s \frac{\tau(s|x, y)}{2} W(z|x', y', s) \right| \right. \\
&\qquad \left. + \left| \sum_s \frac{\tau(s|x', y')}{2} W(z|x, y, s) - \sum_s \frac{\tau'(s|x, y)}{2} W(z|x', y', s) \right| \right\} \\
&\quad \geq \max_{x, x', y, y', z} \left| \sum_s \frac{\tau'(s|x', y') + \tau(s|x', y')}{2} W(z|x, y, s) - \sum_s \frac{\tau'(s|x, y) + \tau(s|x, y)}{2} W(z|x', y', s) \right| \geq \theta \quad \text{(A.15)}
\end{aligned}$$

where the last step uses (A.14) for $\sigma = (\tau' + \tau)/2$.

Next, (A.13) implies that for $(u, u', v, v', s^n, s'^n, z^n) \in \mathcal{T}^n_{XX'YY'SS'Z}$, (4), (5), (6), and

$$I(X'Y'Z \wedge XY|S') < \xi \quad \text{(A.16)}$$

hold. By (4), (6), and the log–sum inequality we get (A.17), shown at the bottom of the page.

Pinsker's inequality is shown in (A.18), at the bottom of the page. Similarly, from (11), (A.16), and Pinsker's inequality we obtain

$$\sum_{x, x', y, y', z} \left| P_{XX'YY'Z}(x, x', y, y', z) - P_{X'}(x') P_{Y'}(y') P_{XY}(x, y) \cdot \sum_s P_{S'|X,Y}(s|x, y) W(z|x', y', s) \right| < c\sqrt{\xi + \zeta}. \quad \text{(A.19)}$$

Thus, by (A.18) and (A.19)

$$\begin{aligned}
&\sum_{x, x', y, y', z} \left| P_X(x) P_Y(y) P_{X'Y'}(x', y') \sum_s P_{S|X'Y'}(s|x', y') \cdot W(z|x, y, s) - P_{X'}(x) P_{Y'}(y) P_{XY}(x, y) \right. \\
&\qquad \left. \cdot \sum_s P_{S'|XY}(s|x, y) W(z|x', y', s) \right| < 2c\sqrt{\xi + \zeta}. \quad \text{(A.20)}
\end{aligned}$$

Next, applying the log–sum inequality to (4), we obtain

$$D(P_{XY} \| P_X \times P_Y) < \xi, \quad \text{(A.21)}$$

and then with Pinsker's inequality

$$\sum_{x, y} |P_{XY}(x, y) - P_X(x) P_Y(y)| < c\sqrt{\xi}. \quad \text{(A.22)}$$

Similarly, it follows from (5) that

$$\sum_{x', y'} |P_{X'Y'}(x', y') - P_{X'}(x') P_{Y'}(y')| < c\sqrt{\xi}. \quad \text{(A.23)}$$

$$\begin{aligned}
\xi + \zeta &> D(P_{XYZS} \| P_X \times P_Y \times P_S \times W) + I(XYZ \wedge X'Y'|S) \\
&= \sum_{x, y, s, z} P_{XYSZ}(x, y, s, z) \log \frac{P_{XYSZ}(x, y, x, s, z)}{P_X(x) P_Y(y) P_S(s) W(z|x, y, s)} \\
&\quad + \sum_{x, x', y, y', s, z} P_{XX'YY'SZ}(x, x'y, y', s, z) \log \frac{P_{X'Y'|XYSZ}(x', y'|x, y, s, z)}{P_{X'Y'|S}(x', y'|s)} \\
&= \sum_{x, x'y, y', s, z} P_{XX'YY'SZ}(x, x', y, y', s, z) \log \frac{P_{XX'YY'SZ}(x, x', y, y', s, z)}{P_X(x) P_Y(y) P_{X'Y'S}(x', y', s) W(z|x, y, s)} \\
&\geq \sum_{x, x', y, y', z} P_{XX'YY'Z}(x, x', y, y', z) \log \frac{P_{XX'YY'Z}(x, x', y, y', z)}{P_X(x) P_Y(y) P_{X'Y'}(x', y') \sum_s P_{S|X'Y'}(s|x', y') W(z|x, y)} \quad \text{(A.17)}
\end{aligned}$$

$$\cdot \sum_{x, x', y, y', z} \left| P_{XX'YY'Z}(x, x', y, y', z) - P_X(x) P_Y(y) P_{X'Y'}(x', y') \sum_s P_{S|X'Y'}(s|x'y') W(z|x, y, s) \right| \leq c\sqrt{\xi + \zeta} \quad \text{(A.18)}$$

Finally, with (A.20), (A.22), and (A.23), we conclude that

$$\begin{aligned}
&\sum_{x,x',y,y',z} P_X(x)P_{X'}(x')P_Y(y)P_{Y'}(y') \\
&\quad\times\left|\sum_s P_{S|X'Y'}(s|x',y')W(z|x,y,s)\right. \\
&\quad\left.-\sum_s P_{S'|XY}(s|x,y)W(z|x',y',s)\right| \\
&\le \sum_{x,x',y,y',z} P_X(x)P_Y(y)\left(\sum_s P_{S|X',Y'}(s|x',y')W(z|x,y,s)\right) \\
&\quad\cdot|P_{X'}(x')P_{Y'}(y')-P_{X'Y'}(x',y')| \\
&\quad+\sum_{x,x',y,y',z}\left|P_X(x)P_Y(y)P_{X'Y'}(x',y')\right. \\
&\quad\cdot\sum_s P_{S|X',Y'}(s|x',y')W(z|x,y,s) \\
&\quad\left.-P_{X'}(x')P_{Y'}(y')P_{XY}(x,y)\sum_s P_{S'|XY}(s|x,y)W(z|x',y',s)\right| \\
&\quad+\sum_{x,x',y,y',z} P_{X'}(x)P_{Y'}(y)\left(\sum_s P_{S'|XY}(s|x,y)W(z|x',y',s)\right) \\
&\quad\cdot|P_X(x)P_Y(y)-P_{XY}(x,y)| \\
&\le \sum_{x,x',y,y',z}|P_{X'}(x')P_{Y'}(y')-P_{X'Y'}(x',y')|+2c\sqrt{\xi+\zeta} \\
&\quad+\sum_{x,x',y,y',z}|P_X(x),P_Y(y)-P_{XY}(x,y)|\le 2c\sqrt{\xi}|\mathcal{X}||\mathcal{Y}||Z| \\
&\quad+2c\sqrt{\xi+\zeta}\le 4c|\mathcal{X}||\mathcal{Y}||Z|\sqrt{\xi+\zeta} \qquad \text{(A.24)}
\end{aligned}$$

which, together with (A.15) (for $\tau' = P_{S|X'Y'}$, $\tau = P_{S|XY}$), $P_X = P_{X'}$, $P_Y = P_{Y'}$, $\min_x P_X(x) \ge \alpha$, and $\min_y P_Y(y) \ge \beta$, implies

$$\xi+\zeta \ge (4c|\mathcal{X}||\mathcal{Y}||Z|)^{-2}\alpha^4\beta^4\theta^2.$$

■

References

[1] R. Ahlswede, "Multiway communication channels," in *Proc. 2nd Int. Symp. Inform. Theory,* Akademiai Kiado, Budapest, Hungary, 1973, pp. 23–52 (Thakadsor, Armenian SSR 1971).

[2] ——, "The capacity region of a channel with two senders and two receivers," *Ann. Probability,* vol. 2, no. 5, pp. 805–814, 1974.

[3] ——, "Elimination of correlation in random codes for arbitrarily varying channels," *Z. Wahrsch. Verw. Geb.,* vol. 44, pp. 159–175, 1978.

[4] ——, "A method of coding and its application to arbitrarily varying channels," *J. Combin. Inform. Syst. Sci.,* vol. 5, no. 1, pp. 10–35, 1980.

[5] T. Ericson, "Exponential error bounds for random codes in the arbitrarily varying channel," *IEEE Trans. Inform. Theory,* vol. IT-31, pp. 42–48, 1985.

[6] I. Csiszár and P. Narayan, "The capacity of the arbitrarily varying channel revisited: Positivity, constraints," *IEEE Trans. Inform. Theory,* vol. 34, pp. 181–193, 1988.

[7] J. A. Gubner, "On the deterministic-code capacity of the multiple-access arbitrarily varying channel," *IEEE Trans. Inform. Theory,* vol. 36, pp. 262–275, 1990.

[8] J. H. Jahn, "Coding of arbitrarily varying multiuser channels," *IEEE Trans. Inform. Theory,* vol. IT-27, pp. 212–226, 1981.

[9] I. Csiszár and J. Körner, "On the capacity of the arbitrarily varying channels for maximum probability of error," *Z. Wahrsch. Verw. Geb.,* vol. 57, pp. 87–101, 1981.

[10] M. S. Pinsker, "Information and information stability of random variable and processes," (in Russian), vol. 7 of the series, *Problemy Peredacii.* San Francisco, CA: Holden-Day, 1964 (SSSR Moscow, Engl. transl.).

Arbitrarily Varying Multiple-Access Channels—Part II: Correlated Senders' Side Information, Correlated Messages, and Ambiguous Transmission

Rudolf Ahlswede and Ning Cai

***Abstract*—We consider an arbitrarily varying multiple-access channel (AVMAC) $\mathcal{W}$ in which the two senders $\mathcal{X}$ and $\mathcal{Y}$ observe, respectively, the components K^m and L^m of a memoryless correlated source (MCS) $\{(K^m, L^m)\}_{m=1}^{\infty}$ with generic rv's (K, L). In Part I of this work [16], it has been shown for the AVMAC *without* the MCS that in order for the achievable rate region for deterministic codes and the average probability of error criterion to be nonempty, it was sufficient if the AVC were $\mathcal{X}$ nonsymmetrizable, $\mathcal{Y}$ nonsymmetrizable, and $\mathcal{XY}$ nonsymmetrizable. (The necessity of these conditions had been shown earlier by Gubner [7].)**

Let $\mathcal{R}_R(\mathcal{W})$ denote the random code achievable rate region of the AVMAC $\mathcal{W}$. In the present paper, the authors, in effect, trade the loss in achievable rates due to symmetrizability off the gains provided by the MCS. Let $\mathcal{R}(\mathcal{W}, (K, L))$ represent the achievable rate region of the AVC $\mathcal{W}$ *with* MCS, for deterministic codes and the average probability of error criterion. There are two main results:

1) **if $I(K \wedge L) > 0$, then $\mathcal{R}(\mathcal{W}, (K, L))$ has a nonempty interior iff $\mathcal{R}_R(\mathcal{W})$ does too and $\mathcal{W}$ is $\mathcal{XY}$ nonsymmetrizable;**
2) **if $I(K \wedge L) > 0, H(K|L) > 0, H(L|K) > 0$, then the MCS can be transmitted over the AVMAC iff $\mathcal{R}_R(\mathcal{W})$ has a nonempty interior and $\mathcal{W}$ is $\mathcal{XY}$ nonsymmetrizable.**

***Index Terms*— Ahlswede's dichotomy, correlated messages, Ericson's symmetrizability, multiple-access AVC, senders' side information.**

I. Introduction

In [10], we have shown how a memoryless correlated source (MCS) helps the transmission over an arbitrarily varying channel (AVC). Precisely, we established the following result.

Theorem AC_1: For an AVC $\mathcal{W}$ let the sender observe $K^n = K_1, \cdots, K_n$ and let the receiver observe $L^n = L_1, \cdots, L_n$ where $(K^n, L^n)_{n=1}^{\infty}$ is a MCS with generic pair of RV's (K, L) having mutual information $I(K \wedge L) > 0$. Then the capacity $C(\mathcal{W}, (K, L))$ for deterministic codes and the average error criterion equals the random capacity $C_R(\mathcal{W})$.

It serves here as a guide to establish coding theorems for arbitrarily varying multiple-access channels (MAC).

For an arbitrarily varying MAC, abbreviated as AVMAC and defined by a set $\mathcal{W} = \{W(\cdot|\cdot,\cdot,s): s \in \mathcal{S}\}$ of stochastic $(\mathcal{X} \times \mathcal{Y} \to \mathcal{Z})$ matrices, it is natural to investigate the effect of the MCS

$$(K^n, L^n)_{n=1}^{\infty} \text{ with } I(K \wedge L) > 0, \qquad (1)$$

when the $\mathcal{X}$-encoder observes K^n and the $\mathcal{Y}$-encoder observes L^n.

Manuscript received January 7, 1997; revised May 15, 1998. Presented at the IEEE International Symposium on Information Theory, Ulm, Germany, June 29–July 4, 1997.

The authors are with the Fakultät für Mathematik, Universität Bielefeld, Bielefeld, 33501 Germany.

Communicated by S. Shamai, Associate Editor for Shannon Theory.

Publisher Item Identifier S 0018-9448(99)01408-X.