CORRELATED SOURCES HELP TRANSMISSION OVER AN ARBITRARILY VARYING CHANNEL

Rudolf Ahlswede and Ning Cai

Abstract

It is well–known that the deterministic code capacity (for the average error probability criterion) of an arbitrarily varying channel (AVC) either equals its random code capacity or zero. Here it is shown that if two components of a correlated source are additionally available to the sender and receiver, respectively, the capacity always equals its random code capacity.

Keywords and phrases: Common randomness, arbitrarily varying channel, side information, positivity of average error capacity

1. Introduction and Result

A well–known dichotomy [1] in the behaviour of the deterministic code capacity (for the average error probability criterion) $\overline{C}(\mathcal{W})$ for an arbitrarily varying channel (AVC) with set of transmission matrices $\mathcal{W}$ is this: either $\overline{C}(\mathcal{W}) = 0$ or else $\overline{C}(\mathcal{W}) = C_R(\mathcal{W})$, the random code capacity $\min\limits_{W \in \overline{\mathcal{W}}} C(W)$, where $C(W)$ is the ordinary capacity of the discrete memoryless channel (DMC) W and $\overline{\mathcal{W}}$ is the convex closure of $\mathcal{W}$. It can happen that $C_R(\mathcal{W}) > 0$ and $\overline{C}(\mathcal{W}) = 0$.

We prove here that in the presence of a correlated source $(U^n, V^n)_{n=1}^\infty$ (which is independent of the message) with $I(U \wedge V) > 0$ and access of the sender to U^n and of the receiver to V^n this cannot happen!

Clearly, if $(U^n)_{n=1}^\infty$ and $(V^n)_{n=1}^\infty$ have positive common randomness, $CR(U,V)$ then the result is readily established by the elimination technique of [1]. Here the common randomness $CR(U,V)$ of a correlated source $(U^n, V^n)_{n=1}^\infty$ is defined as the maximal real R such that for all $\varepsilon > 0$ and sufficiently large n there exist functions $\phi : \mathcal{U}^n \to \{1, 2, \dots, 2^{n(R-\varepsilon)}\}$ and $\Phi : \mathcal{V}^n \to \{1, 2, \dots, 2^{n(R-\varepsilon)}\}$ with $\Pr\bigl(\phi(U^n) \neq \Phi(V^n)\bigr) < \varepsilon$ (c.f. [2]).

So the interesting case arises when

(a) the common randomness $CR(U,V)$ of the correlated source is 0

and

(b) $\overline{C}(\mathcal{W}) = 0$, but $C_R(\mathcal{W}) > 0$.

Now we know not only that a positive $CR(U,V)$ can help to make the channel capacity positive, but from [3] also that a positive channel capacity $\overline{C}(W)$ can be used to make the common randomness positive. However, we are confronted with the situation where *both quantities are* 0 !

Actually we can always find binary–valued functions f and g so that $\bigl(f(U_t), g(V_t)\bigr)_{t=1}^\infty$ satisfies $I\bigl(f(U) \wedge g(V)\bigr) > 0$, if $I(U \wedge V) > 0$.

So we can assume that we are given a binary correlated source with alphabets $\mathcal{U} = \mathcal{V} = \{0,1\}$.

Quite surprisingly, we found a complete solution of the problem.

Theorem. *For an AVC $\mathcal{W}$ let the sender observe $U^n = U_1, \dots, U_n$ and let the receiver observe $V^n = V_1, \dots, V_n$, where $(U^n, V^n)_{n=1}^\infty$ is a memoryless correlated source (which is independent of the message) with generic pair of RV's (U,V) having mutual information $I(U \wedge V) > 0$. Then the capacity $\overline{C}\bigl(\mathcal{W}, (U,V)\bigr)$ for deterministic codes and the average error criterion equals the random capacity $C_R(\mathcal{W})$.*

2. Code concepts and an auxiliary channel

An (n, M, λ) code is a system $\left\{\left(g_m^n(u^n), \mathcal{D}_m(v^n)\right)_{m=1}^M, u^n \in \mathcal{U}^n, v^n \in \mathcal{V}^n\right\}$

$$g_m^n(u^n) \in \mathcal{X}^n \quad \text{for} \quad u^n \in \mathcal{U}^n, \; \mathcal{D}_m(v^n) \subset \mathcal{Y}^n \quad \text{for} \quad v^n \in \mathcal{V}^n,$$
$$\mathcal{D}_m(v^n) \cap \mathcal{D}_{m'}(v^n) = \varnothing \quad \text{for} \quad m \neq m',$$

and

$$\frac{1}{M} \sum_{m=1}^M \sum_{u^n, v^n} P_{UV}^n(u^n, v^n) \cdot W^n\left(\mathcal{D}_m(v^n) | g_m^n(u^n), s^n\right) > 1 - \lambda \tag{2.1}$$

for $s^n \in \mathcal{S}^n$, if $\left\{W(\cdot|\cdot, s) : s \in \mathcal{S}\right\} = \mathcal{W}$. The corresponding capacity is denoted by $\overline{C}\left(\mathcal{W}, (U, V)\right)$.

It turns out that it suffices to restrict the class of encoding functions g_m^n as follows:

$$g_m^n(u^n) = \left(g_{m,1}(u_1), \dots, g_{m,n}(u_n)\right). \tag{2.2}$$

It is therefore natural to consider an associated AVC $\hat{\mathcal{W}} = \left\{\hat{W}(\cdot, \cdot|\cdot, s) : s \in \mathcal{S}\right\}$ with input letters $g : \mathcal{U} \to \mathcal{X}$ and output letters (y, v). Indeed, we set

$$\hat{W}(y, v|g, s) = P_V(v) \sum_{u=0}^1 P_{U|V}(u|v) W\left(y|g(u), s\right) \quad \text{for} \quad y \in \mathcal{Y}, v \in \{0, 1\}, s \in \mathcal{S}. \tag{2.3}$$

Using (2.2) and (2.3) we can rewrite the left hand side of (2.1) as

$$\begin{aligned}
&\frac{1}{M} \sum_{m=1}^M \sum_{v^n} P_V^n(v^n) \sum_{u^n} P_{U|V}^n(u^n|v^n) W^n\left(\mathcal{D}_m(v^n) | g_m^n(u^n), s^n\right) \\
&= \frac{1}{M} \sum_{m=1}^M \hat{W}^n \left(\bigcup_{v^n \in \mathcal{V}^n} \left(\mathcal{D}_m(v^n) \times v^n\right) | g_m^n, s^n \right) \\
&= \frac{1}{M} \sum_{m=1}^M \hat{W}^n\left(\hat{\mathcal{D}}_m | g_m^n, s^n\right), \quad \text{where}
\end{aligned}$$

$$\hat{\mathcal{D}}_m = \bigcup_{v^n \in \mathcal{V}^n} \left(\mathcal{D}_m(v^n) \times v^n\right). \tag{2.4}$$

We thus have shown (by going from (2.4) backwards) that

$$\overline{C}\left(\mathcal{W}, (U, V)\right) \geq \overline{C}(\hat{\mathcal{W}}). \tag{2.5}$$

Similarly one can show that actually equality holds here, but this is not used in the sequel.

3. Proof of Theorem

Clearly, if $\overline{C}\big(\mathcal{W},(U,V)\big)$ is positive, then we can use in block length $O(\log n)$ a code in the sense of (2.1) to get the common randomness needed to operate the correlated code of rate $\sim C_R(\mathcal{W})$ gained by the elimination technique.

On the other hand, if $\overline{C}\big(\mathcal{W},(U,V)\big) = 0$, then by (2.5) also $\overline{C}(\hat{\mathcal{W}}) = 0$. This implies by [5] that $\hat{\mathcal{W}}$ is symmetrisable in the sense of [4]. It remains to be seen that this in turn implies the existence of a $\overline{W} \in \overline{\mathcal{W}}$ with $C(\overline{W}) = 0$ and therefore $C_R(\mathcal{W}) = 0$.

Let $\mathcal{X} = \{0,1,\dots,a-1\}$, let $\mathcal{G}$ be the set of functions from $\mathcal{U}$ to $\mathcal{X}$, and let $\mathcal{G}^* = \{g^*\} \cup \{g_i : 0 \le i \le a-1\} \subset \mathcal{G}$, where

$$g^*(0) = g^*(1) = 0 \quad \text{and} \quad g_i(u) = i + u \mod a \quad \text{for} \quad u \in \{0,1\}. \tag{3.1}$$

Now by symmetrisability of $\hat{\mathcal{W}}$ there is a stochastic matrix $\tau : \mathcal{G} \to \mathcal{S}$ with

$$P_V(v) \sum_{u=0}^{1} P_{U|V}(u|v) \sum_s \tau(s|g^*) W\big(y|g_i(u),s\big) = P_V(v) \sum_{u=0}^{1} P_{U|V}(u|v) \sum_s \tau(s|g_i) W(y|0,s) \tag{3.2}$$

for all $v = 0,1$, $i \in \mathcal{X}$, and $y \in \mathcal{Y}$.

Cancell $P_V(v)$ and consider (3.2) with $v = 0$ and $v = 1$. Then

$$\sum_{u=0}^{1} P_{U|V}(u|0) \sum_s \tau(s|g^*) W\big(y|g_i(u),s\big) = \sum_{u=0}^{1} P_{U|V}(u|0) \sum_s \tau(s|g_i) W(y|0,s), \tag{3.3}$$

$$\sum_{u=0}^{1} P_{U|V}(u|1) \sum_s \tau(s|g^*) W\big(y|g_i(u),s\big) = \sum_{u=0}^{1} P_{U|V}(u|1) \sum_s \tau(s|g_i) W(y|0,s). \tag{3.4}$$

Clearly the right hand sides of both, (3.3) and (3.4), equal $\sum\limits_s \tau(s|g_i) W(y|0,s)$.

We evaluate these equations by inserting the values for the g_i and get with the convention $i \oplus j = i + j \mod a$

$$P_{U|V}(0|0) \sum_s \tau(s|g^*) W(y|i,s) + P_{U|V}(1|0) \sum_s \tau(s|g^*) W(y|i \oplus 1,s) = \sum_s \tau(s|g_i) W(y|0,s). \tag{3.5}$$

$$P_{U|V}(0|1) \sum_s \tau(s|g^*) W(y|i,s) + P_{U|V}(1|1) \sum_s \tau(s|g^*) W(y|i \oplus 1,s) = \sum_s \tau(s|g_i) W(y|0,s). \tag{3.6}$$

With the abbreviations

$$\alpha = \sum_s \tau(s|g_i)W(y|0,s),$$
$$z_0 = \sum_s \tau(s|g^*)W(y|i,s), \quad \text{and} \quad z_1 = \sum_s \tau(s|g^*)W(y|i\oplus 1,s)$$

we get therefore the system of two equations

$$\begin{aligned} P_{U|V}(0|0)z_0 + P_{U|V}(1|0)z_1 &= \alpha \\ P_{U|V}(0|1)z_0 + P_{U|V}(1|1)z_1 &= \alpha. \end{aligned} \tag{3.7}$$

Since $I(U\wedge V)>0$ implies $P_{U|V}(\cdot|0)\neq P_{U|V}(\cdot|1)$, we get
$\det\begin{pmatrix} P_{U|V}(0|0) & P_{U|V}(1|0) \\ P_{U|V}(0|1) & P_{U|V}(1|1)\end{pmatrix} \neq 0$
and (3.7) has a unique solution $z_0 = z_1 = \alpha$. Hence, for all $i\in\mathcal{X}$ and all $y\in\mathcal{Y}$

$$\sum_s \tau(s|g^*)W(y|i,s) = \sum_s \tau(s|g^*)W(y|i\oplus 1,s)$$

and $\overline{W}(\cdot|\cdot) \equiv \sum_s \tau(s|g^*)W(\cdot|\cdot,s) \in \overline{\mathcal{W}}$ has identical rows.
Therefore $C(\overline{W}) = 0 = C_R(\mathcal{W})$. Q.E.D.

Remark:

Inspection of the proof shows that actually we proved the following.

If $\overline{U}\,\overline{V}\,\overline{Y}$ forms a Markov chain and $\overline{U}$ and $\overline{V}$ are binary, then $I(\overline{U}\wedge\overline{V})>0$ and $I(\overline{U}\wedge\overline{Y})=0$ imply $I(\overline{V}\wedge\overline{Y})=0$.

Indeed, $P_{\overline{U}\,\overline{V}\,\overline{Y}}(\overline{U}=u,\overline{V}=v,\overline{Y}=y) = P_{\overline{U}}(u)P_{\overline{V}|\overline{U}}(v|u)P_{\overline{Y}|\overline{V}}(y|v)$ and $I(\overline{U}\wedge\overline{Y}) = 0$ imply that for all y

$$\begin{aligned} P_{\overline{V}|\overline{U}}(0|0)P_{\overline{Y}|\overline{V}}(y|0) + P_{\overline{V}|\overline{U}}(1|0)P_{\overline{Y}|\overline{V}}(y|1) &= P_{\overline{Y}}(y) \\ P_{\overline{V}|\overline{U}}(0|1)P_{\overline{Y}|\overline{V}}(y|0) + P_{\overline{V}|\overline{U}}(1|1)P_{\overline{Y}|\overline{V}}(y|1) &= P_{\overline{Y}}(y). \end{aligned} \tag{3.8}$$

Suppose that $I(\overline{U}\wedge\overline{V})>0$, then $P_{\overline{V}|\overline{U}}(\cdot|0)\neq P_{\overline{V}|\overline{U}}(\cdot|1)$ and therefore (3.8) has the unique solution

$$P_{\overline{Y}|\overline{V}}(y|0) = P_{\overline{Y}|\overline{V}}(y|1) = P_{\overline{Y}}(y) \quad \text{for all} \quad y,$$

i.e. $I(\overline{V}\wedge\overline{Y})=0$.

References

[1] R. Ahlswede, "Elimination of correlation in random codes for arbitrarily varying channels", Z. Wahrscheinlichkeitstheorie und verw. Geb. 44, 159–185, 1978.

[2] R. Ahlswede and I. Csiszár, "Common randomness in information theory and cryptography, Part I: Secret sharing", IEEE Trans. Information Theory, Vol. 39, No. 4, 1121–1132, 1993, Part II to appear ibid.

[3] R. Ahlswede and V. Balakirsky, "Identification under random processes", Preprint 95–098, SFB 343 "Diskrete Strukturen in der Mathematik", Universität Bielefeld, *Problemy Peredachi Informatsii*, (special issue devoted to M.S. Pinsker), vol. 32, no. 1, 144–160, Jan.–Mar. 1996.

[4] T. Ericson, "Exponential error bounds for random codes in the arbitrarily varying channel", IEEE Trans. Information Theory, Vol. 31, 42–48, 1985.

[5] I. Csiszár and P. Narayan, "The capacity of the arbitrarily channel revisited: positivity, constraints", IEEE Trans. Information Theory, Vol. 34, No. 2, 181–193, 1988.