

- Wright's hypergeometric function," *Fibonacci Quarterly*, vol. 29, pp. 52-56, Feb. 1991.
- [17] H. Hochstadt, *The Functions of Mathematical Physics*. New York: Wiley-Interscience, 1971.

On the Capacity Region of the Discrete Additive Multiple-Access Arbitrarily Varying Channel

John A. Gubner, *Member, IEEE*

Abstract—The discrete additive multiple-access arbitrarily varying channel (AVC) with two senders and one receiver is considered. Necessary and sufficient conditions are given for its deterministic-code average-probability-of-error capacity region under a state constraint to have a nonempty interior. In the case that no state constraint is present, the capacity region is characterized exactly. In the case of the noiseless mod-2 adder AVC using state constraint function $l(s) = s$ and subject to a state constraint L less than or equal to 0.13616917, the capacity region is shown to be a 45-degree triangle whose legs have length $1 - h(L)$, where h denotes the binary entropy function.

Index Terms—Additive channel, multiple-access, arbitrarily varying channel, state constraint, capacity region.

I. INTRODUCTION

A general multiple-access arbitrarily varying channel (AVC) with two senders and one receiver is a transition probability W from $X \times Y \times S$ into Z , where X , Y , S , and 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 $z \in Z$ given that the channel input symbol from sender 1 is $x \in X$, the channel input symbol from sender 2 is $y \in Y$, and that the channel state is $s \in S$. When block codes of length n are used, we say the AVC is subject to state constraint L if the state-selection mechanism can generate only those state sequences $s = (s_1, \dots, s_n)$ that satisfy a time-average constraint of the form

$$\frac{1}{n} \sum_{k=1}^n l(s_k) \leq L, \quad (1)$$

where l is a given nonnegative constraint function defined on S and satisfying $\min_s l(s) = 0$. Note that if $L \geq \max_s l(s)$, then all state sequences s satisfy (1); in this case we say that the state constraint is not present, or inactive.

Definition (Additive AVC): Let G be a finite nontrivial commutative group. Suppose that $X = Y = Z = G$. We say that W is an additive AVC if

$$W(z | x, y, s) = V_s(z - x - y),$$

for some transition probability V from S into G .

General multiple-access AVC's subject to a state constraint have been studied in [6]. There, both forward and converse results were proved that enable one to give inner and outer bounds on the capacity region. To obtain meaningful inner bounds, one must

Manuscript received November 6, 1990; revised December 20, 1991. This work was supported in part by the Air Force Office of Scientific Research under Grant AFOSR-90-0181. This work was presented in part at the IEEE International Symposium on Information Theory, Budapest, Hungary, June 24-28, 1991.

The author is with the Department of Electrical and Computer Engineering, University of Wisconsin-Madison, 1415 Johnson Drive, Madison, WI 53706-1691.

IEEE Log Number 9108024.

exhibit input probability distributions for which certain inequalities are nonvacuous. We show that for the additive AVC such input distributions always exist.

In the absence of state constraints, we exactly characterize the capacity region of the additive AVC.

In the special case of the noiseless mod-2 adder AVC with $l(s) = s$ and state constraint $L \leq 0.13616917$, the capacity region is shown to be a 45° triangle whose legs have length $1 - h(L)$, where h denotes the binary entropy function defined in Theorem 3.

Additive AVC's with one sender and one receiver were considered in [4, Section V], but under the assumption that the channel symbols come from a finite subset of \mathbb{R}^d rather than a finite commutative group G . This is in contrast to the results of [4, Section IV] concerning a restricted form of additive AVC called a group adder AVC, which is an additive AVC for which $S = G$ and $V_s(t) = \mu(t - s)$ for some probability distribution μ on G . In an earlier paper [3, Section IV] Csiszár and Narayan analyzed the single-user noiseless mod-2 adder AVC.

II. STATEMENT OF RESULTS

In order to state our results, we need the following notation. Let $\mathcal{D}(S)$ denote the set of probability distributions on S . For $r \in \mathcal{D}(S)$, let rV denote the distribution on G defined by $(rV)(t) = \sum_s r(s)V_s(t)$. Let $H(rV)$ denote the entropy of rV . Let

$$\mathcal{D}^L(S) \triangleq \left\{ r \in \mathcal{D}(S) : \sum_{s \in S} l(s)r(s) \leq L \right\}.$$

Note that if $L \geq \max_s l(s)$, then $\mathcal{D}^L(S) = \mathcal{D}(S)$. We now state our main results.

Theorem 1: The deterministic-code average-probability-of-error capacity region under state constraint L of an additive multiple-access AVC V has a nonempty interior, if and only if there is no $r \in \mathcal{D}^L(S)$ such that rV is the uniform distribution on G . Furthermore, the capacity region is always contained in the 45° triangle,

$$\left\{ (R_1, R_2) : R_1 \geq 0, R_2 \geq 0, \right. \\ \left. \text{and } R_1 + R_2 \leq \log |G| - \max_{r \in \mathcal{D}^L(S)} H(rV) \right\}, \quad (2)$$

where $|G|$ denotes the cardinality of the set G .

Remark: Since $\mathcal{D}^L(S)$ is compact and since H is continuous,

$$\log |G| > \max_{r \in \mathcal{D}^L(S)} H(rV), \quad (3)$$

if and only if there is no $r \in \mathcal{D}^L(S)$ such that rV is the uniform distribution on G .

Theorem 2: In the absence of state constraints, the capacity region of the additive multiple-access AVC V is always given by (2), where $\mathcal{D}^L(S)$ is replaced by $\mathcal{D}(S)$.

Proof: Theorem 2 follows from Theorem 1, the preceding Remark, ([7, Theorem 1, p. 214], which says that if the deterministic-code average-probability-of-error capacity region has a nonempty interior, then it is equal to the random-code average-probability-of-error capacity region), and [6, Section IV], which shows that the random-code average-probability-of-error capacity region of the additive AVC is given by (2). We give an independent proof in Section V. \square

We now consider the noiseless mod-2 adder AVC, which is defined as follows. The group G is taken to be the set $\{0, 1\}$ under mod-2 addition. For the set of channel states, we take $S = G$. The noiseless mod-2 adder AVC is obtained by taking $V_s(t) = \delta(t - s)$, where $\delta(0) = 1$ and $\delta(1) = 0$. In discussing this channel, we always take the state-constraint function to be $I(s) = s$ so that (1) is equivalent to the requirement that the fraction of 1's in the sequence s be less than or equal to L .

Theorem 3: For $L \leq 0.13616917$, the deterministic-code average-probability-of-error capacity region under state constraint L of the noiseless mod-2 adder AVC is

$$\{(R_1, R_2) : R_1 \geq 0, R_2 \geq 0, \text{ and } R_1 + R_2 \leq 1 - h(L)\}, \quad (4)$$

where $h(L) \triangleq -L \log L - (1 - L) \log(1 - L)$ is the binary entropy function.

III. PRELIMINARIES FOR THE PROOFS

In this section, some useful quantities associated with a general multiple-access AVC W are introduced. Simplified expressions for these quantities are presented when W is an additive AVC as defined in Section I. For later use in the proof of Theorem 1, the hypotheses of Theorem A in the Appendix are also simplified for an additive AVC.

Given any $p \in \mathcal{D}(X)$, $q \in \mathcal{D}(Y)$, and $r \in \mathcal{D}(S)$, it is understood that the average mutual informations $I(X \wedge Z)$, $I(Y \wedge Z | X)$, $I(S \wedge Z)$, and $I(S \wedge Z | X)$ are computed using the necessary marginal and conditional distributions obtained from

$$P_{XYSZ}(x, y, s, z) = p(x)q(y)r(s)W(z | x, y, s).$$

Note that these mutual informations can be regarded as continuous functions of the 4-tuple (p, q, r, W) . With this in mind, we define

$$I^L(X \wedge Z) \triangleq \inf_{r \in \mathcal{D}^L(S)} I(X \wedge Z)$$

and

$$I^L(Y \wedge Z | X) \triangleq \inf_{r \in \mathcal{D}^L(S)} I(Y \wedge Z | X).$$

In working with additive AVC's, it is convenient to define the convolution of two distributions p and $q \in \mathcal{D}(G)$ by

$$(p * q)(y) \triangleq \sum_{x \in G} p(x)q(y - x), \quad y \in G.$$

Note that $p * q \in \mathcal{D}(G)$ and that convolution is both commutative and associative.

For an additive AVC, it is readily verified that

$$I^L(X \wedge Z) = \inf_{r \in \mathcal{D}^L(S)} H(p * q * rV) - H(q * rV), \quad (5)$$

and that

$$I^L(Y \wedge Z | X) = \inf_{r \in \mathcal{D}^L(S)} H(q * rV) - H(rV). \quad (6)$$

To simplify the hypotheses of Theorem A, we proceed as follows. For an additive AVC, (A.1) simplifies to

$$\sum_{s \in S} r(s)H(p * q * V_s) > H(q * rV), \quad \text{for all } r \in \mathcal{D}^L(S), \quad (7)$$

and (A.2) simplifies to

$$\sum_{s \in S} r(s)H(q * V_s) > H(rV), \quad \text{for all } r \in \mathcal{D}^L(S). \quad (8)$$

Observe that the following conditions are sufficient to guarantee (7) and (8), respectively:

$$\min_{s \in S} H(p * q * V_s) > \max_{r \in \mathcal{D}^L(S)} H(q * rV) \quad (9)$$

and

$$\min_{s \in S} H(q * V_s) > \max_{r \in \mathcal{D}^L(S)} H(rV). \quad (10)$$

Thus, for an additive AVC, Theorem A can be applied if (9) and (10) hold.

IV. PROOF OF THEOREM 1

The converse result, that the capacity region of an additive AVC is a subset of the triangle in (2), was proved in [6, Sections III and IV]. From this it follows that if the capacity region has a nonempty interior, then (3) must hold; i.e., there can be no $r \in \mathcal{D}^L(S)$ with rV uniform. Our contribution in this paper is the proof of the forward result, that if there is no $r \in \mathcal{D}^L(S)$ with rV uniform, then the triangle in (2) contains nonempty open rectangles that are contained in the capacity region.

Before proceeding, we summarize a few facts about regular probability distributions on a finite commutative group G . Recall that a distribution p is said to be regular if for every pair of distributions q and q' ,

$$p * q = p * q' \quad \text{implies} \quad q = q';$$

otherwise, p is said to be nonregular. First note that the point mass concentrated on the additive identity element is regular. Next, the uniform distribution, u , is always nonregular; this follows from the fact that $u * q = u$ for all q . Thus, $\mathcal{D}(G)$ always contains both regular and nonregular distributions. We also point out that if p is regular, then the convex combination $\lambda p + (1 - \lambda)u$ is regular for all $\lambda \in (0, 1]$. As a consequence, even though the uniform distribution is nonregular, we can always approximate it by a strictly positive regular distribution. We also need the fact that if p is regular, then $p * q = u$ implies $q = u$.

The first step in the proof is to assume that rV is not uniform for any $r \in \mathcal{D}^L(S)$, or equivalently, on account of the remark following the statement of Theorem 1, that $H(rV) < \log |G|$ for all $r \in \mathcal{D}^L(S)$.

The first step in applying Theorem A is to show that there is a distribution q such that (10) holds. If q is close enough to the uniform distribution, we claim that (10) will hold. To see this, note that the right-hand side of (10) is strictly less than $\log |G|$, while $H(q * V_s)$ is a continuous function of q ; if q is close to uniform, then $H(q * V_s)$ will be close to $\log |G|$ for all s in the finite set S .

The next step is to show that there is a distribution p such that (9) holds. To do this, we make the additional assumption that q is regular. (This explains why we did not take q uniform in the preceding paragraph.) We claim that the right-hand side of (9) is strictly less than $\log |G|$. If this were not the case, there would be an $r \in \mathcal{D}^L(S)$ with $q * rV$ uniform. Since q is regular, we would then have rV uniform; but we have assumed this does not happen. Thus, taking p to be any distribution close to uniform will satisfy (9). For later use we assume p is strictly positive. In fact, we can even take p to be uniform since we do not need p to be regular.

To conclude the proof, observe that since p is strictly positive and since q is regular, it follows from the lemma in the Appendix that the quantity $I^L(X \wedge Z)$ defined in (5) is positive. Similarly, since we may assume q is positive, the quantity $I^L(Y \wedge Z | X)$ defined in (6) is also positive.

Thus, we have shown that there exist distributions p and q such

that (A.1) and (A.2) hold and such that the rectangle defined by (A.3) is nonempty.

V. PROOF OF THEOREM 2

From the proof of Theorem 1, we see that there exist distributions p (nearly uniform and positive) and q (nearly uniform, regular, and positive) such that the open rectangle

$$\{(R_1, R_2) : 0 < R_1 < I^L(X \wedge Z) \text{ and } 0 < R_2 < I^L(Y \wedge Z | X)\} \quad (11)$$

is nonempty and belongs to the capacity region. Clearly, by interchanging the roles of X and Y and p and q , there exist distributions p' (nearly uniform, regular, and positive) and q' (nearly uniform and positive) such that if

$$P_{X'Y'SZ}(x, y, s, z) = p'(x)q'(y)r(s)W(z | x, y, s),$$

then

$$\{(R_1, R_2) : 0 < R_1 < I^L(X' \wedge Z | Y') \text{ and } 0 < R_2 < I^L(Y' \wedge Z)\} \quad (12)$$

is nonempty and belongs to the capacity region. Now, the height of the rectangle in (11) is $I^L(Y \wedge Z | X)$. Since q is nearly uniform, (6) implies that $I^L(Y \wedge Z | X)$ is nearly equal to

$$\log |G| - \max_{r \in \mathcal{D}^L(\mathcal{S})} H(rV), \quad (13)$$

which is the height of the triangle in (2). Similarly, since p' is nearly uniform, the width of the rectangle in (12), $I^L(X' \wedge Z | Y')$, is also nearly equal to (13), which is also the length of the base of the triangle in (2). It follows that any point in the interior of (2) belongs to the convex hull of the open rectangles (11) and (12) if p , q and p' , q' are appropriately selected. Thus, every rate pair in the interior of (2) belongs to the capacity region by the usual time-sharing argument [2, p. 272]. In Section VII, we explain why state constraints cannot be present when using the time-sharing argument.

VI. PROOF OF THEOREM 3

We begin with a few simplifications. First note that $\log |G| = 1$ since $|G| = 2$. Next, since $l(s) = s$, $r \in \mathcal{D}^L(\mathcal{S})$, if and only if $r(1) \leq L$. Also note that for $V_s(t) = \delta(t - s)$, $rV = r$, and so $H(rV) = H(r) = h(r(1))$. Hence, the maximum in (2) is simply $\max_{0 \leq t \leq L} h(t) = h(L)$ when $L \leq 1/2$, since h is increasing on $[0, 1/2]$. It now follows that (4) is simply (2) specialized to the noiseless mod-2 adder AVC. It remains to show that every point in (4) belongs to the capacity region. Since the capacity region is closed, it suffices to consider only interior points of (4). Consider the shaded regions in Fig. 1. We prove that the shaded region at the left belongs to the capacity region. A similar argument interchanging the roles of X and Y will establish that the shaded region at the right also belongs to the capacity region. Since the union of the two shaded regions is the entire triangle, a convex-hull/time-sharing argument is not needed.

Suppose (R_1, R_2) belongs to the shaded region at the left in Fig. 1. We show that there exist distributions p and q such that (R_1, R_2) belongs to the rectangle defined by (A.3) and to which Theorem A applies. To see that this is so, we need the following observations. For the noiseless mod-2 adder AVC with p being the uniform distribution and $0 < q(1) < 1/2$, the rectangle described by (A.3)

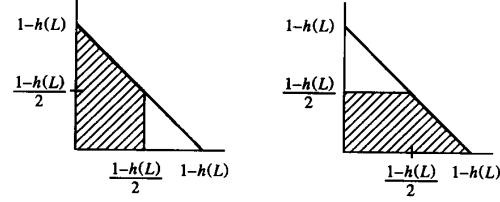


Fig. 1. Decomposition of triangle in (4).

becomes, after a little calculus to evaluate the necessary infima,

$$\{(R_1, R_2) : 0 < R_1 < 1 - h(q(1)(1 - L) + q(0)L) \text{ and } 0 < R_2 < h(q(1)(1 - L) + q(0)L) - h(L)\}. \quad (14)$$

Now, to apply Theorem A to this rectangle, we must also satisfy (9) and (10). Since p is uniform, the left-hand side of (9) is $\log |G|$; since $q(1) < 1/2$, q is regular, and hence, (9) holds. The treatment of (10) is more delicate. In the present situation, (10) simplifies to

$$h(q(1)) > h(L).$$

Thus, we need $L < q(1) < 1/2$.

Next, it is readily verified that the upper-right corner of the rectangle in (14) lies on the hypotenuse of the triangle in (4). Furthermore, the width of this rectangle varies from 0 to $1 - h(L)$ as $q(1)$ varies from $1/2$ down to 0. However, we require $L < q(1) < 1/2$ for Theorem A to hold. This means that the maximum rectangle width that Theorem A can handle is

$$\sup_{L < q(1) < 1/2} 1 - h(q(1)(1 - L) + q(0)L) = 1 - h(2L(1 - L)).$$

Thus, in order for every interior point of the shaded region at the left in Fig. 1 to be included in a rectangle to which Theorem A applies, we need

$$1 - h(2L(1 - L)) \geq \frac{1}{2} [1 - H(L)],$$

or

$$h(2L(1 - L)) - \frac{1}{2} h(L) \leq \frac{1}{2}. \quad (15)$$

Using a computer to print a few values of $h(2L(1 - L)) - h(L)/2$ shows that (15) holds for $L \leq 0.13616917$.

VII. CONCLUSION AND DISCUSSION

We have given necessary and sufficient conditions in order that the deterministic-code average-probability-of-error capacity region of the discrete additive multiple-access AVC subject to a state constraint have a nonempty interior, and we have exactly determined the capacity region when state constraints are not present. We have also exactly established the capacity region of the mod-2 adder AVC when $l(s) = s$ and $L \leq 0.13616917$.

We now explain why we had to assume that state constraints were not present when we applied the time-sharing argument in Section V. The difficulty arises as follows. Suppose $n = n_1 + n_2$. Then (1) does not imply that both

$$\frac{1}{n_1} \sum_{k=1}^{n_1} l(s_k) \leq L \text{ and } \frac{1}{n_2} \sum_{k=1}^{n_2} l(s_{n_1+k}) \leq L, \quad (16)$$

which is a necessary condition to apply the time-sharing argument to an AVC subject to a state constraint.

The fact that (1) does not imply (16) also appears to prevent one from applying Ahlswede's elimination technique [1], [7], to AVC's subject to a state constraint; thus, our result in Theorem 1 does not appear to be sufficient to allow us to prove the conjecture that the region in (2) is exactly the capacity region when state constraints are present, either by the elimination technique or by time-sharing.

Corrections to [6]

The preceding observations require us to make the following corrections to our prior paper [6].

In [6, Theorem 5.8] the words "closed convex hull," should be replaced by "closure."

In [6, Section V-C] we can no longer conclude that equation (5.8) is correct. Instead, all that we can conclude, using notation defined in [6], is that the closure of

$$\mathcal{R}_X^{1/2}(p^*, q^*, W_a) \cup \mathcal{R}_Y^{1/2}(p^*, q^*, W_a)$$

is a subset of $C(W_a, 1/2)$ and that $C(W_a, 1/2)$ is a subset of the right-hand side of (5.8).

ACKNOWLEDGMENT

The author is grateful to an anonymous reviewer of an earlier version of this correspondence for reminding him that the time-sharing argument can not be applied because (1) does not imply (16).

APPENDIX

Lemma: Let p be any distribution with $p(x) > 0$ for all $x \in G$. For any $q \in \mathcal{D}(G)$, $H(p*q) - H(q) = 0$, if and only if q is the uniform distribution.

Proof: Define a distribution on $G \times G$ by setting $P_{XY}(x, y) = p(x)q(y - x)$. Then

$$\begin{aligned} I(X \wedge Y) &= H(Y) - H(Y|X) \\ &= H(p*q) - H(q). \end{aligned}$$

Clearly, $H(p*q) - H(q) = 0$, if and only if X and Y are independent. Since $p(x) > 0$ for all x , X and Y are independent, if and only if $q(y - x) = q(y)$ for all $x, y \in G$. Thus, X and Y are independent, if and only if q is uniform. \square

Theorem A: Let W be a general multiple-access AVC. If $p \in \mathcal{D}(X)$ and $q \in \mathcal{D}(Y)$ are such that (recall paragraph 2 of Section III)

$$I(X \wedge Z) > I(S \wedge Z), \quad \text{for all } r \in \mathcal{D}^L(S), \quad (\text{A.1})$$

and

$$I(Y \wedge Z|X) > I(S \wedge Z|X), \quad \text{for all } r \in \mathcal{D}^L(S), \quad (\text{A.2})$$

then every pair (R_1, R_2) satisfying

$$0 < R_1 < I^L(X \wedge Z) \quad \text{and} \quad 0 < R_2 < I^L(Y \wedge Z|X) \quad (\text{A.3})$$

belongs to the deterministic-code average-probability-of-error capacity region under state constraint L .

Proof: This theorem can be proved by making trivial modifications to the proof of [5, Theorem 5.5]. A similar observation was made in [6, Section V-A], though it was not pointed out there that in this case the modifications do not require that the channel be nonsymmetrizable [6, Definitions 3.3, 3.5, and 3.7]. \square

REFERENCES

- [1] R. Ahlswede, "Elimination of correlation in random codes for arbitrarily varying channels," *Z. Wahrscheinlichkeitstheorie verw. Geb.*, vol. 44, pp. 159-175, 1978.
- [2] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. New York: Academic, 1981.
- [3] I. Csiszár and P. Narayan, "The capacity of the arbitrarily varying channel revisited: Positivity, constraints," *IEEE Trans. Inform. Theory*, vol. 34, pp. 181-193, Mar. 1988.
- [4] —, "Capacity and decoding rules for classes of arbitrarily varying channels," *IEEE Trans. Inform. Theory*, vol. 35, pp. 752-769, July 1989.
- [5] J. A. Gubner, "On the deterministic-code capacity of the multiple-access arbitrarily varying channel," *IEEE Trans. Inform. Theory*, vol. 36, pp. 262-275, Mar. 1990.
- [6] —, "State constraints for the multiple-access arbitrarily varying channel," *IEEE Trans. Inform. Theory*, vol. 37, pp. 27-35, Jan. 1991.
- [7] J.-H. Jahn, "Coding of arbitrarily varying multiuser channels," *IEEE Trans. Inform. Theory*, vol. IT-27, pp. 212-226, Mar. 1981.

Shaping Using Variable-Size Regions

Jay N. Livingston, *Member, IEEE*

Abstract—Constellation shaping is extended to provide shaping gains without resorting to high-dimensional constellations. This is accomplished by dividing the constellation into unequal sized constellations, and selecting these constellations on an equiprobable basis. A design example is provided, demonstrating the simplicity and power of the approach.

Index Terms—Coding, modulation, shaping, signal constellations, and nonequiprobable signaling.

I. INTRODUCTION

Coding schemes for transmission of data over the Gaussian channel have been used over the last decade that lead to improved performance [1]-[4]. The most popular approach is to use *coset codes*, and to attain high code rates coupled with good performance, the shift has been to use higher dimensional constellations. One result of moving to higher dimensions is the ability to achieve what has been called *shaping gains*. This is due to the reduction in average symbol energy that can accompany the use of a constellation whose boundary is not an N -cube. In particular, as the constellation becomes more spherical, it enforces a nonequiprobable distribution on signal points drawn from a constituent two-dimensional constellation. It has been shown [5] that in the limit as $N \rightarrow \infty$, an N -sphere can achieve 1.53 dB of shaping gain, and will enforce a truncated Gaussian distribution on the constituent 2-D constellation. Attention has been focused on these shaping gains, as they can be achieved independently of any gain due to the use of a coset code.

Multidimensional constellations with significant shaping gain were first described by Conway and Sloane in [6]. Other constellations with shaping gain and simple decoding methods were described by

Manuscript received July 3, 1990, revised June 20, 1991. This work was presented in part at the IEEE Globecom '90 Conference on Communications, San Diego, CA, December 1990.

The author is with the Department of Electrical Engineering, Texas A&M University, College Station, TX 77843-3128.
IEEE Log Number 9107508.