IV. PROOF OF THEOREM: THE INDUCTION BEGINNING

For the value $t = 1$, only the cases 1 and 3 occur. This simplifies matters, because there is now no need to inform the decoder whether case 2 or case 3 occured. On the other hand, since $\binom{n}{1}$ is linear in n, rather accurate estimates are needed and in particular the inequality in Proposition 2 is too crude for the present purposes. Instead, we use (12) and (13), consequences of Lemma 2.

For general n, we devide the interval of transmission of length n into $b + 1$ successive blocks such that the 0th block has length $r \overset{\Delta}{=} \lfloor \log_1(q-1)n \rfloor$ and all others have length D or $D - 1$ (and so $\lfloor \frac{n-r}{D} \rfloor \le b \le \lceil \frac{n-r}{D-1} \rceil$).

Case 1: The one error can only be in the 0th block, where we send (as previously) 0's. In the other blocks, all sequences can be sent. Therefore, we can transmit q^{n-r} messages, where

$$q^{n-r} = q^{n - \lfloor \log_q(q-1)n \rfloor} \ge \frac{q^n}{(q-1)n+1} = \frac{q^n}{S_1}.$$

Case 3: Let the position of a possible error be in the jth block.

In the 0th block we use only sequences with at most $r - 2$ 0's (to distinguish for the decoder this case from case 1) and we also encode that an error may occur in the jth block. We cannot waste even one position! This we achieve by partitioning the $q^r - r(q-1) - 1$ sequences with at most $r - 2$ 0's into b sets $P_1, \cdots, P_b$ such that

$$\left| |P_i| - \frac{q^r - r(q-1) - 1}{b} \right| \le 1,$$

for $i = 1, 2, \cdots, b$.

Words in P_j inform the decoder that an error may occur in the jth block. There the sender uses a code meeting the bounds in (12) or (13) and in the remaining blocks all sequences can be used.

Therefore, we can transmit at least

$$\left\lfloor \frac{q^r - r(q-1) - 1}{b} \right\rfloor \frac{q^{n-r}}{(q-1)D+1} \ge \frac{q^n}{S_1}(1 + o(1))$$

messages, where $o(a) \to o$ as $D \to \infty$ and $b \to \infty$. The proof is complete. $\square$

REFERENCES

[1] L. A. Bassalygo, S. I. Gelfand, and M. S. Pinsker, "Coding for channels with localized errors," *Proc. 4th Soviet–Swedish Workshop in Inform. Theory*, Gotland, Sweden, 1989, pp. 95–99.

[2] ——, "Coding for channels with partially localized errors," *IEEE Inform. Theory*, vol. 37, no. 3, pp. 880–884, May 1991.

[3] ——, "Simple methods of deduction of lower bounds in coding theory," *Probl. Peredac. Inform.*, vol. 27, no. 4, pp. 3–8, 1991.

[4] R. Ahlswede, "Coding for channels with localized errors: The nonbinary cases," Preprint 89-020 of SFB 343, Universität Bielefeld.

[5] I. I. Dumer and V. A. Zinovyev, "New maximum codes over GF(4)," *Probl. Peredac. Inform.*, vol. 14, no. 3, pp. 24–34, 1978.

[6] G. A. Kabatyansky, "The construction of code correcting single localized error," *Proc. III Int. Workshop on Algebraic and Combinat. Coding Theory*, Tyrnovo, Bulgaria, June, 1992, pp. 22–28, in print.

[7] R. Ahlswede, L. A. Bassalygo, and M. S. Pinsker, "Asymptotically dense nonbinary codes correcting a constant number of localized errors," *Proc. III Int. Workshop on Albegraic and Combinat. Coding Theory*, Tyrnovo, Bulgaria, June 1992, pp. 22–28, in print.

[8] G. A. Kabtyansky and V. I. Panchenko, "Packings and coverings of Hamming spaces with unit spheres," *Probl. Peredac. Inform.*, vol. 24, no. 4, pp. 3–16, 1988.

The Maximal Error Capacity of Arbitrarily Varying Channels for Constant List Sizes

Rudolf Ahlswede

Abstract—The capacity of an arbitrarily varying channel (AVC) for list codes of arbitrarily small list rate under the maximal error criterion has previously been determined. Here, the following sharper result is proved: For an AVC $\mathcal{A}$, any rate R below the list code capacity $C_l(\mathcal{A})$ is achievable with the list size $L(\mathcal{A}, R) = \left\lfloor \frac{\log|\mathcal{Y}|}{C_l(\mathcal{A}) - R} \right\rfloor + 1$, where $\mathcal{Y}$ is the output alphabet. For the average error criterion, the corresponding result was conjectured by Pinsker and proved by Ahlswede and Cai.

Index Terms—Arbitrarily varying channel, list codes, maximal error, balanced hypergraph packing.

I. KNOWN RESULTS

An AVC is defined here by a sequence $\mathcal{A} = (\{P(\cdot | \cdot | s^n) : s^n \in S^n\})_{n=1}^\infty$ of sets of transmission probabilities, where for a finite input alphabet $\mathcal{X}$, a finite output alphabet $\mathcal{Y}$ and a finite set $W = \{w(\cdot | \cdot | s) : s \in S\}$ of stochastic $|\mathcal{X}| \times |\mathcal{Y}|$-matrices

$$P(y^n | x^n | s^n) = \prod_{t=1}^n w(y_t | x_t | s_t), \qquad (1)$$

for all $x^n = (x_1, x_2, \cdots, x_n) \in \mathcal{X}^n = \prod_1^n \mathcal{X}$, for all $y^n \in \mathcal{Y}^n$, and for all $s^n \in S^n$. Let L be a positive integer. An (n, N, L) list code is a system

$$\{(u_i, \mathcal{D}_i) : 1 \le i \le N\},$$

where $u_i \in \mathcal{X}^n$, $\mathcal{D}_i \subset \mathcal{Y}^n$ and

$$\sum_{i=1}^N 1_{\mathcal{D}_i}(y^n) \le L, \qquad \text{for all } y^n \in \mathcal{Y}^n. \qquad (2)$$

$1_\mathcal{B}$ denotes the indicator function of a set $\mathcal{B}$.

We speak of an (n, N, L, λ) code for $\mathcal{A}$, if in addition for all $s^n \in S^n$

$$P(\mathcal{D}_i | u_i | s^n) \ge 1 - \lambda, \qquad \text{for } i = 1, 2, \cdots, N. \qquad (3)$$

We call a number $C_l(\mathcal{A})$ the list code capacity of $\mathcal{A}$, if the two conditions hold

1) For any $\epsilon > 0$, $\delta > 0$ and $\lambda \in (0, 1)$ there is an $(n, \exp\{n(C_l(\mathcal{A}) - \delta)\}, \exp\{n\epsilon\}, \lambda)$ code for all large n.
2) For any $\delta > 0$ and $\lambda \in (0, 1)$ there is no ϵ, $0 < \epsilon < \delta$, such that $(n, \exp\{n(C_l(\mathcal{A}) + \delta)\}, \exp\{n\epsilon\}, \lambda))$ codes exist for all large n.

The so-called row-convex hull of W is defined as

$$\bar{\bar{W}} = \{w(\cdot | \cdot) : w(\cdot | x) \in \text{conv} \{w(\cdot | x | s) : s \in S\} \text{ for all } x \in \mathcal{X}\}. \qquad (4)$$

Let $\mathcal{P}(\mathcal{X})$ stand for the set of probability distributions on $\mathcal{X}$. Denoting by $I(P, w)$ the mutual information for input distribution P and channel matrix w, we can introduce

$$\bar{\bar{C}} = \max_{P \in \mathcal{P}(\mathcal{X})} \min_{w \in \bar{\bar{W}}} I(P, w). \qquad (5)$$

Theorem [1]: $C_l(\mathcal{A}) = \bar{\bar{C}}$.

Remark: The proof of [1] is based on Lemma 2 in conjunction with a binning idea ([1] Lemma 4).

The method of binning was discovered independently by Slepian and Wolf. Furthermore, the list size exp $\{n\epsilon\}$ can be improved. We quote from [1, p. 835]:

"Actually one could continue to reduce the list size from l to $\log l$ and so on. But thus one achieves the capacity only for larger and larger block lengths. It would be of interest to obtain results for a *constant* list size but those results would have to be obtained by a different approach."

Actually, we use again the first part of the earlier proof.

List Code Lemma ([1, Lemma 2]): One can construct for any $P \in \mathcal{P}(\mathcal{X})$, $\epsilon > 0$, and $n \geq n(\epsilon)$ an (n, N, L, λ) code for $\mathcal{A}$ with

$$N \geq \exp\{nH(P) - f(P, |\mathcal{X}|)\log n\},$$
$$L \leq \exp\{n \max_{w \in \mathcal{W}} H(w^* \mid Q) + ng(\epsilon)\},$$
$$\lambda \leq \exp\{-nE(\epsilon, P)\},$$

where $\lim_{\epsilon \to 0} g(\epsilon) = 0$, $E(\epsilon, P) > 0$ for $\epsilon > 0$, $Q = P \cdot w$ and $P(x)w(y \mid x) = Q(y)w^*(x \mid y)$ for $x \in \mathcal{X}$, $y \in \mathcal{Y}$.

A. The Missing Concept

In [2] we discussed packings, coverings, and partitions in hypergraphs $\mathcal{H} = (\mathcal{V}, \mathcal{E})$. Also, for the last two concepts the notion "c-balanced" was introduced. The following result has found particularly many applications.

Covering Lemma 3 ([2, p. 250]): A hypergraph $\mathcal{H} = (\mathcal{V}, \mathcal{E})$ with

$$D \triangleq \max_{v \in \mathcal{V}} \deg(v), \quad d \triangleq \min_{v \in \mathcal{V}} \deg(v) > 0$$

has a c-balanced covering with exactly k edges, if

1) $k \geq |\mathcal{E}|d^{-1}(\log|\mathcal{V}| + 1)$
2) $c \leq k \leq c|\mathcal{E}|D^{-1}$
3) $\exp\{[h(ck^{-1}) + ck^{-1}\log(D|\mathcal{E}|^{-1})]k + \log|\mathcal{V}|\} < \frac{1}{2}$.

Since $D \geq d$, a) and b) can hold only if $c > \log|\mathcal{V}|$. In typical applications in information theory $|\mathcal{V}|$ depends exponentially on the blocklength n and thus c has to grow with the block length.

However, a more general concept, namely that of a packing with some—but controlled—overlap has not been focused upon. It catches the essential structure of list codes and deserves an analysis for its own sake.

We call $\mathcal{F} \subset \mathcal{E}$ a c-balanced packing of $\mathcal{H}$ if for all $v \in \mathcal{V}$,

$$|\{E: E \in \mathcal{F}, v \in E\}| \leq c. \tag{6}$$

Generally speaking coverings are easier to handle than packings, because overlap is allowed, however for $c > 1$ c-balanced packings are easier to handle than c-balanced coverings, because it is not required that $\mathcal{F}$ covers $\mathcal{V}$! This has the effect that condition a) in Covering Lemma 3 can be dropped and "constant" c's are not automatically excluded.

Indeed the following result holds.

Packing Lemma: A hypergraph $\mathcal{H} = (\mathcal{V}, \mathcal{E})$ has a c-balanced packing with k edges if b) and c) of Covering Lemma 3 hold.

Proof: We just have to repeat the old arguments. Choose edges $E^{(1)}, \cdots, E^{(k)}$ independently at random according to the uniform distribution on $\mathcal{E}$. Next Estimate $\Pr(\{E^{(1)}, \cdots, E^{(k)}\}$ is not c-balanced).

For this define

$$g_v^i = \begin{cases} 1, & \text{if } v \notin E^{(i)}, \\ 0, & \text{if } v \in E^{(i)}, \end{cases}$$

and observe that the probability for v to be covered by more than c edges is given by

$$\Pr\left(\sum_{i=1}^k g_v^i < k - c\right) = \Pr\left(\sum_{i=1}^k g_v^i < k(1 - \beta)\right),$$

where $\beta = ck^{-1}$. If we define $p = D|\mathcal{E}|^{-1}$ then condition b) implies that $p \leq \beta$ and the exponential form of Chebychev's inequality gives—as in [2, p. 86, (2.14), pt. I], that

$$\Pr\left(\sum_{i=1}^k g_v^i < k(1 - \beta)\right) < \exp\{[h(\beta) + \beta \log D|\mathcal{E}|^{-1}]k\}. \tag{7}$$

Since

$$\Pr\{(E^{(1)}, \cdots, E^{(k)}) \text{ is not } c\text{-balanced}\}$$
$$\leq \sum_{v \in \mathcal{V}} \Pr\left(\sum_{i=1}^k g_v^i < k(1 - \beta)\right),$$

the result follows. $\qquad\square$

II. APPLICATION OF THE PACKING LEMMA TO THE LIST CODE IN THE LIST CODE LEMMA

Choose the hypergraph $(\mathcal{V}, \mathcal{E})$ with $\mathcal{V} = \mathcal{Y}^n$ and $\mathcal{E} = \{\mathcal{D}_i: 1 \leq i \leq N\}$, where the $\mathcal{D}_i$'s are the decoding sets in the list code with parameters (n, N, L, λ). Notice that for a P assuming the value $\bar{\bar{C}}$

$$|\mathcal{E}|D^{-1} \geq N \cdot L^{-1} \geq \exp\{n\bar{\bar{C}} - f(P, |\mathcal{X}|)\log n - ng(\epsilon)\}. \tag{8}$$

Application of the Packing Lemma gives us a subcode of size $k = \exp\{nR\}$, if R is any rate below $\bar{\bar{C}}$, and of list size $c = L(\mathcal{A}, R)$, provided that conditions b) and c) hold. By (8) and the choices of k and c, b) obviously holds, if ϵ is so small that $g(\epsilon) < \bar{\bar{C}} - R$ and n is large enough. To verify c), we derive an upper bound on the exponent there.

Now,

$$h(ck^{-1})k + c\log(D|\mathcal{E}|^{-1}) + n\log|\mathcal{Y}|$$
$$\leq -c\log c + n \cdot c \cdot R - e^{Rn}(1 - c \cdot e^{-Rn})\log(1 - c \cdot e^{-Rn})$$
$$- n \cdot c\left[\bar{\bar{C}} - \frac{1}{n}f(P, |\mathcal{X}|)\log n - g(\epsilon)\right] + n\log|\mathcal{Y}|$$

and since $-(1 - x)\log(1 - x) \leq 2x$ for small x, we upperbound this by

$$-c\log c + 2c - n \cdot c\left[\bar{\bar{C}} - R - \frac{1}{c}\log|\mathcal{Y}| - 2g(\epsilon)\right],$$

for $n \geq n_1(\epsilon)$.

It suffices to guarantee that the term in square brackets is positive or that

$$c > (\bar{\bar{C}} - R - 2g(\epsilon))^{-1}\log|\mathcal{Y}| > 0.$$

To this end just choose $\epsilon = \epsilon(R)$ small enough.

Remark: The condition in [3] is $c > [E(R)\lambda]^{-1}\log|\mathcal{S}|$, where $E(R)$ is the reliability function of the AVC in the case of random codes and average error.

REFERENCES

[1] R. Ahlswede, "Channel capacities for list codes," *J. Appl. Probab.*, vol. 10, no. 4, pp. 824–836, 1973.

[2] ——, "Coloring hypergraphs: A new approach to multi-user source coding, Pt. I," *J. of Combinat., Inform. Syst. Sci.*, vol. 4, no. 1, pp. 76–115, 1979; "Pt. II," *J. Combinat., Inform. Syst. Sci.*, vol. 5, no. 3, pp. 220–268, 1980.

[3] R. Ahlswede and N. Cai, "Two proofs of Pinsker's conjecture concerning AV channels," *IEEE Trans. Inform. Theory*, vol. 37, no. 6, 1647–1649, Nov. 1991.