THE MAXIMAL ERROR CAPACITY OF AV CHANNELS FOR CONSTANT LIST SIZES

Rudolf Ahlswede

Fakultät für Mathematik
Universität Bielefeld
Postfach 100131
33501 Bielefeld
Germany

Abstract

In [1] we determined the capacity of an arbitrarily varying channel (AVC) for list codes of arbitrarily small list rate under the maximal error criterion. Here we prove a sharper result, which also partially answers a question on page 835 of [1].

Theorem. *For an AVC* $\mathcal{A}$ *any rate* R *below the list code capacity* $C_\ell(\mathcal{A})$ *is achievable with the list size* $L(\mathcal{A},R) = \left\lfloor \frac{\log|\mathcal{Y}|}{C_\ell(\mathcal{A})-R} \right\rfloor + 1$, *where* $\mathcal{Y}$ *is the output alphabet.*

For the average error criterion the corresponding result was conjectured by Pinsker and proved in [3].

Keywords: Arbitrarily varying channel, list codes, maximal error, balanced hypergraph packing.

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

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

for all $x^n = (x_1, x_2, \ldots, x_n) \in \mathcal{X}^n = \prod_1^n \mathcal{X}$, for all $y^n \in \mathcal{Y}^n$, and for all $s^n \in \mathcal{S}^n$.

Let L be a positive integer. An (n, N, L) list code is a system

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

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

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

$1_{\mathcal{B}}$ 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 $s^n \in \mathcal{S}^n$

$$P(\mathcal{D}_i|u_i|s^n) \geq 1 - \lambda \quad \text{for} \quad i = 1, 2, \ldots, N. \tag{3}$$

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

(a) For any $\varepsilon > 0, \delta > 0$ and $\lambda \in (0,1)$ there is an $\big(n, \exp\{n(C_\ell(\mathcal{A}) - \delta)\}, \exp\{n\varepsilon\}, \lambda\big)$ code for all large n .

(b) For any $\delta > 0$ and $\lambda \in (0,1)$ there is no ε , $0 < \varepsilon < \delta$, such that $\big(n, \exp\{n(C_\ell(\mathcal{A}) + \delta)\}, \exp\{n\varepsilon\}, \lambda)\big)$ codes exist for all large n .

The so called row–convex hull of $\mathcal{W}$ is defined as

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

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

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

Theorem ([1]). $C_\ell(\mathcal{A}) = \overline{\overline{C}}$.

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

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

"Actually one could continue to reduce the list size from ℓ to $\log \ell$ 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. *(Lemma 2 of [1])*

One can construct for any $P \in \mathcal{P}(\mathcal{X})$, $\varepsilon > 0$, *and* $n \geq n(\varepsilon)$ *an* (n, N, L, λ) *code for* $\mathcal{A}$ *with*

$$\begin{aligned}
N &\geq \exp\big\{n\ H(P) - f(P, |\mathcal{X}|) \log n\big\}, \\
L &\leq \exp\big\{n \max_{w \in \overline{\overline{\mathcal{W}}}} H(w^*|Q) + n\ g(\varepsilon)\big\}, \\
\lambda &\leq \exp\big\{-n\ E(\varepsilon, P)\big\},
\end{aligned}$$

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

The missing concept.

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

Covering Lemma 3. *([2], page 250)*

A hypergraph $\mathcal{H} = (\mathcal{V}, \mathcal{E})$ *with*

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

has a c–balanced covering with exactly k *edges, if*

(a) $k \geq |\mathcal{E}| d^{-1} \big(\log |\mathcal{V}| + 1\big)$

(b) $c \leq k \leq c|\mathcal{E}| D^{-1}$

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

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 focussed upon. It catches also 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 $E^{(1)}, \dots, E^{(k)}$ independently at random according to the uniform distribution on $\mathcal{E}$. Estimate next $Pr(\{E^{(1)}, \dots, E^{(k)}\}$ is not c–balanced).

For this define again

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

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

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

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 (2.14) on page 86 of [2], Part I — that

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

Since $Pr\{(E^{(1)},\dots,E^{(k)})$ is not c–balanced $\}$ $\leq \sum_{v\in\mathcal{V}} Pr\big(\sum_{i=1}^{k} g_v^i < k(1-\beta)\big)$, the result follows.

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 $\overline{\overline{C}}$

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

Application of the Packing Lemma gives us a subcode of size $k=\exp\{nR\}$, if R is any rate below $\overline{\overline{C}}$, 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 $g(\varepsilon)<\overline{\overline{C}}-R$ and n is large enough. To verify (c) we derive an upper bound on the exponent there.

Now

$$\begin{aligned}
&h(ck^{-1})k+c\log(D|\mathcal{E}|^{-1})+n\log|\mathcal{Y}|\\
&\leq -c\log c+n\cdot c\cdot R-e^{Rn}(1-c\cdot e^{-Rn})\log(1-c\cdot e^{-Rn})\\
&-n\cdot c\big[\overline{\overline{C}}-\frac{1}{n}f(P,|\mathcal{X}|)\log n-g(\varepsilon)\big]+n\log|\mathcal{Y}|
\end{aligned}$$

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

$$-c\log c+2c-n\cdot c\big[\overline{\overline{C}}-R-\frac{1}{c}\log|\mathcal{Y}|-2g(\varepsilon)\big]$$

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

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

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

At this end just choose $\varepsilon=\varepsilon(R)$ small enough.

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

References

[1] R. Ahlswede, "Channel capacities for list codes", Journal of Applied Probability, Vol. 10, No. 4, pp 824–836, 1973.

[2] —, "Coloring hypergraphs: A new approach to multi–user source coding", Pt I, Journ. of Combinatorics, Information and System Sciences, Vol. 4, No. 1, pp 76–115, 1979; Pt II, Journ. of Combinatorics, Information and System Sciences, Vol. 5, No. 3, pp 220–268, 1980.

[3] — and N. Cai, "Two proofs of Pinsker's conjecture concerning AV channels", IEEE Trans. Inf. Theory, Vol. 37, No. 6, 1647–1649, 1991.