stichting
mathematisch
centrum

MC

AFDELING MATHEMATISCHE STATISTIEK SW 19/73 OCTOBER

EDWARD C. van der MEULEN

ON A PROBLEM BY AHLSWEDE REGARDING THE CAPACITY REGION OF CERTAIN MULTI-WAY CHANNELS

Prepublication

2e boerhaavestraat 49 amsterdam

Printed at the Mathematical Centre, 49, 2e Boerhaavestraat, Amsterdam.

The Mathematical Centre, founded the 11-th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O), by the Municipality of Amsterdam, by the University of Amsterdam, by the Free University at Amsterdam, and by industries.

AMS(MOS) subject classification scheme (1970): 94A15

On a problem by Ahlswede regarding the capacity region of certain multi-way channels *)

Edward C. van der Meulen **)

Abstract

A new Fano-type estimate is proved for multi-way channels. It yields upper bounds on the numbers of codewords and the products of these numbers in a code for a channel with s senders and r receivers in the case all senders send messages simultaneously to all receivers. The result is based on a new approach, which makes an approximation argument previously used by Ahlswede (1971b) and Ulrey (1973) unnecessary. It also leads to better Fano-estimates than the ones obtained by these authors. Our approach is canonical for weak converses of coding theorems for multi-way channels and applies also in other communication situations.

---

*) This paper is not for review; it has been submitted for publication in a journal.

**) From January 1, 1974: Department of Statistics, University of Rochester, Rochester, New York 14627, USA.

1. Introduction and Summary.

Ahlswede (1971b) characterized the capacity region of a channel with two senders and one receiver, and of a channel with three senders and one receiver. In that paper, Ahlswede proposed as a Problem to find a simpler proof than the one he gave of the weak converse to the coding theorem for the discrete memoryless channel with three senders and one receiver. His proof turned out to be surprisingly complicated. It involved an approximation argument, based on a Markov-type inequality, and a careful handling of the convex hull of six sets of triples of rates. The approximation argument was not needed in the case of a channel with two senders and one receiver. It arose in the case of a channel with three or more senders, since it is not possible to expurgate from a code for a two-way channel having small average error probability, one which has uniformly small error probability and essentially the same code lengths, as was shown by Ahlswede (1971a).

Subsequently, Ahlswede (1972) found an alternate characterization of the capacity region of the discrete memoryless channel with two senders and one receiver. The proof of this result is based on a new approach to the random coding problem for this channel, which admits non-stationary sources. One advantage of the new characterization is that it makes the proof of the weak converse much simpler than before. Using the same approach, Ahlswede (1972) characterized also the capacity region of a channel with two senders and two receivers for the communication situation in which both senders send messages simultaneously to both receivers.

Recently, Ulrey (1973) has extended the methods and results of Ahlswede (1972) to a channel with s senders and r receivers, in the case all senders send messages simultaneously to all receivers. In particular Ulrey found a characterization of the capacity region of a channel with $s \geq 3$ senders and one receiver, and of a channel with $s \geq 2$ senders and $r \geq 2$ receivers. His characterization of the capacity region of a channel with three senders and one receiver is simpler than the one obtained by Ahlswede (1971b). This new characterization makes it possible to deal more easily with the convex combinations of rate points as compared with the proof of the converse in Ahlswede (1971b). However, the extra difficulty of how to translate a statement about average error into a corresponding statement about maximal error, which comes in only in the case $s \geq 3$, remains. Ulrey, like Ahlswede, resolved this difficulty by resorting to an approximation argument.

In this note we present a simple proof of the inequalities which bound from above the codelengths and their products in the case of a channel with three or more senders. Our proof does not make use of the approximation argument at all, but involves only a careful use of the well-known inequalities of information theory. Our proof is not only simpler than the one given by Ahlswede (1971b) and Ulrey (1973), but also leads to better Fano-estimates of the codelengths. In an earlier paper (van der Meulen, 1971) we have used an argument similar to the one presented here for the derivation of upper bounds on the numbers of codewords in the case of the discrete memoryless channel with two senders and one receiver.

2. The main result.

We state and prove our result directly for the general case of s senders. For simplicity we restrict to the case of one receiver. We adhere as much as possible to the notation developed by Ulrey (1973). We assume familiarity with his paper, in particular his sections 1 and 2 up to Lemma 1. We then have the following Lemma, which is an improvement of Lemma 1 of Ulrey (1973).

Lemma. Given a code - $(n, \vec{N}, \lambda)$ for (P, T_{s1}); denote it by $\{M(\bar{i}), A(\bar{i}))|\bar{i} \in \bar{I}\}$ where $A(\bar{i}) \subset Y(1)$ for all $\bar{i} \in \bar{I}$. Let $p_k^t(\cdot)$ and $p^t(\cdot)$ be as defined in respectively (2.5) and (2.2) of Ulrey (1973). Let $D \subset \{1,\ldots,s\}$, $D \neq \phi$. Then

$$\log\left(\prod_{k\in D} N_k\right) \le \frac{\sum_{t=1}^{n} R_D(p^t) + 1}{1-\lambda}. \tag{2.6}$$

Proof. For ease of notation it is assumed that $D = \{1,\ldots,d\}$ for some d, $1 \le d \le s$. First consider the case $d \le s-1$. Let

$$I_D = \{(j_1,\ldots,j_d) \mid j_k \text{ an int.}, 1 \le j_k \le N_k, k = 1,\ldots,d\}. \tag{2.7}$$

For each $\bar{j} \in I_D$ define

$$I_{\bar{j}} = \{\bar{i} \in \bar{I} \mid \bar{i} = (i_1,\ldots,i_s) \text{ where } i_k = j_k \text{ for } k = 1,\ldots,d\}. \tag{2.8}$$

Let

$$B(j) = \bigcup_{\bar{i}\in I_{\bar{j}}} A(\bar{i}). \tag{2.9}$$

Then clearly

(2.10) $\quad B(\bar{j}) \supset A(\bar{i})$

for each $\bar{j} \in I_D$, whenever $\bar{i} \in I_{\bar{j}}$. Let $\tilde{N} = \prod_{k=1}^{d} N_k$, $\bar{N} = \prod_{k=d+1}^{s} N_k$.

Also let $\tilde{X} = \prod_{k=1}^{d} X(k)$, and $\bar{X} = \prod_{k=d+1}^{s} X(k)$. Furthermore define

$$\tilde{\mathcal{M}} = \{\tilde{M} \mid \tilde{M} \text{ is an } n \times d \text{ matrix where } \tilde{M}_k^t \in X_k^t \text{ for all } t = 1,\ldots,n \text{ and } k = 1,\ldots,d\} \tag{2.11}$$

and

$$\bar{\mathcal{M}} = \{\bar{M} \mid \bar{M} \text{ is an } n \times (s-d) \text{ matrix where } \bar{M}_k^t \in X_k^t \text{ for all } t = 1,\ldots,n \text{ and } k = d+1,\ldots,s\}. \tag{2.12}$$

We recall that we are given a fixed code, consisting of a collection $\{u_k(i_k) \mid 1 \le i_k \le N_k,\ k = 1,\ldots,s\}$ where $u_k(i_k) \in X_k$ for all $1 \le i_k \le N_k$ and $k = 1,\ldots,s$. Next define a p.d. $\tilde{\mu}(\cdot)$ on $\tilde{\mathcal{M}}$ by

$$\tilde{\mu}(\tilde{M}) = \begin{cases} \dfrac{1}{\tilde{N}} & \text{if for each } k,\ 1 \le k \le d,\ \tilde{M}_k = u_k(i_k) \text{ for some } i_k,\ 1 \le i_k \le N_k \\ 0 & \text{otherwise.} \end{cases} \tag{2.13}$$

Also define a p.d. $\bar{\mu}(\cdot)$ on $\bar{\mathcal{M}}$ by

$$(2.14)\qquad \bar{\mu}(\bar{M}) = \begin{cases} \frac{1}{\bar{N}} \text{ if for each } k,\ d+1 \le k \le s, \bar{M}_k = u_k(i_k) \text{ for} \\ \quad\text{for some } i_k,\ 1 \le i_k \le N_k \\ \\ 0 \ \text{ otherwise.} \end{cases}$$

Further for all $t = 1,\ldots,n$, we define a p.d. $\tilde{\mu}^t(\cdot)$ on $\tilde{X}$ by

$$(2.15)\qquad \tilde{\mu}^t(\tilde{x}) = \sum_{\{\tilde{M}|\tilde{M}^t=\tilde{x}\}} \tilde{\mu}(\tilde{M})$$

for all $\tilde{x} \in \tilde{X}$. Similarly, for all $t = 1,\ldots,n$, we define a p.d. $\bar{\mu}^t(\cdot)$ on $\bar{X}$ by

$$(2.16)\qquad \bar{\mu}^t(\bar{x}) = \sum_{\{\bar{M}|\bar{M}^t=\bar{x}\}} \bar{\mu}(\bar{M})$$

for all $\bar{x} \in \bar{X}$. Next we define

$$(2.17)\qquad \tilde{P}(y_n|\tilde{M}) = \sum_{\bar{M} \in \bar{\mathfrak{M}}} P(y_n|\tilde{M},\bar{M})\bar{\mu}(\bar{M})$$

for all $y_n = (y^1,\ldots,y^n) \in \prod_{s=1}^{n} Y = Y_n$ and $\tilde{M} \in \tilde{\mathfrak{M}}$. Here $(\tilde{M},\bar{M}) = M \in \mathfrak{M}$ and

$$(2.18)\qquad P(y_n|\tilde{M},\bar{M}) = \prod_{t=1}^{n} \omega(y^t|M^t).$$

Similarly we define

$$(2.19)\qquad \bar{P}(y_n|\bar{M}) = \sum_{\tilde{M} \in \tilde{\mathfrak{M}}} P(y_n|\tilde{M},\bar{M})\tilde{\mu}(\tilde{M})$$

for all $y_n = (y^1,\ldots,y^n) \in Y_n$ and $\bar{M} \in \bar{\mathfrak{M}}$. For each $\bar{j} \in I_D$ we let

$\tilde{M}(\bar{j}) = (u_1(j_1),\ldots,u_d(j_d))$. We then have the following crucial relations:

$$(2.20)\qquad \frac{1}{\tilde{N}} \sum_{\bar{j} \in I_D} \tilde{P}(B(\bar{j})|\tilde{M}(\bar{j})) = \frac{1}{N} \sum_{\bar{j} \in I_D} \sum_{\bar{i} \in I_{\bar{j}}} P(B(\bar{j})|M(\bar{i}))$$

$$\geq \frac{1}{N} \sum_{\bar{i} \in \bar{I}} P(A(\bar{i})|M(\bar{i}))$$

$$\geq 1 - \lambda.$$

Therefore the system

$$(2.21)\qquad \{(\tilde{u},B(\bar{j}))|\tilde{u} = (u_1(j_1),\ldots,u_d(j_d)) \text{ for all } \bar{j} = (j_1,\ldots,j_d) \in I_D\}$$

is a code $(1, \tilde{N}, \lambda)$ for the one-way channel whose channel probability function is given by (2.17). Consequently Fano's Lemma yields

$$(2.22)\qquad \log \tilde{N} \leq \frac{R(\tilde{\mu}(\cdot),\tilde{P}(\cdot|\cdot)) + 1}{1 - \lambda}$$

where $R(\tilde{\mu}(\cdot),\tilde{P}(\cdot|\cdot))$ is as defined in (2.1) of Ulrey (1973). Now

$$(2.23)\qquad R(\tilde{\mu}(\cdot),\tilde{P}(\cdot|\cdot))$$

$$= \sum_{\tilde{M} \in \tilde{\mathcal{M}}} \sum_{\bar{M} \in \bar{\mathcal{M}}} \sum_{y_n \in Y_n} \tilde{\mu}(\tilde{M})\bar{\mu}(\bar{M})P(y_n|\tilde{M},\bar{M}) \log \frac{\tilde{P}(y_n|\tilde{M})}{\sum_{\tilde{M} \in \tilde{\mathcal{M}}} \tilde{P}(y_n|\tilde{M})\tilde{\mu}(\tilde{M})}$$

$$\leq \sum_{\tilde{M} \in \tilde{\mathcal{M}}} \sum_{\bar{M} \in \bar{\mathcal{M}}} \sum_{y_n \in Y_n} \tilde{\mu}(\tilde{M})\bar{\mu}(\bar{M})P(y_n|\tilde{M},\bar{M}) \log \frac{P(y_n|\tilde{M},\bar{M})}{\sum_{\tilde{M} \in \tilde{\mathcal{M}}} P(y_n|\tilde{M},\bar{M})\tilde{\mu}(\tilde{M})}.$$

Next consider for fixed $\bar{M} \in \bar{\mathcal{M}}$ the expression

$$(2.24) \qquad \sum_{\tilde{M} \in \tilde{\mathcal{M}}} \sum_{y_n \in Y_n} \tilde{\mu}(\tilde{M}) P(y_n|\tilde{M},\bar{M}) \log \frac{P(y_n|\tilde{M},\bar{M})}{\sum_{\tilde{M} \in \tilde{\mathcal{M}}} P(y_n|\tilde{M},\bar{M})\tilde{\mu}(\tilde{M})} .$$

By a standard theorem of information theory (see p. 75 of Gallager, 1968) expression (2.24) is less than or equal to

$$(2.25) \qquad \sum_{t=1}^{n} \sum_{\tilde{M}^t \in X^t} \sum_{y^t \in Y} \tilde{\mu}^t(\tilde{M}^t)\omega(y^t|\tilde{M}^t,\bar{M}^t) \log \frac{\omega(y^t|\tilde{M}^t,\bar{M}^t)}{\sum_{\tilde{M}^t \in X^t} \omega(y^t|\tilde{M}^t,\bar{M}^t)\tilde{\mu}^t(\tilde{M}^t)} .$$

By definition expression (2.25) is equal to

$$(2.26) \qquad \sum_{t=1}^{n} R(\tilde{\mu}^t,\omega(\cdot|\cdot,\bar{M}^t)).$$

In the above $\bar{M} = (\bar{M}^1,\ldots,\bar{M}^n)$ was fixed. We now take the expectation of (2.24) with respect to $\bar{\mu}(\cdot)$. We obtain, using (2.25) and (2.26), that

$$(2.27) \qquad R(\tilde{\mu}(\cdot),\tilde{P}(\cdot|\cdot)) \le \sum_{t=1}^{n} R_D(p^t).$$

Here $p^t(M^t) = \tilde{\mu}^t(\tilde{M}^t)\bar{\mu}^t(\bar{M}^t)$, and $R_D(p^t)$ is as defined in (2.3) of Ulrey (1973). Now comparing (2.27) with (2.22) we obtain (2.6) for the case $1 \le d \le s-1$. The case $d = s$ is immediate. Hence the Lemma is proved.

References

Ahlswede, R. (1971a), "On two-way communication channels and a problem by Zarankiewicz," presented at the Sixth Prague Conference on Information Theory, Statistical Decision Functions, and Random Processes.

Ahlswede, R. (1971b), "Multi-way communication channels," presented at the Second International Symposium on Information Theory at Tsahkadsor, Armenian S.S.R.. To appear in Problems of Control and Information Theory.

Ahlswede, R. (1972), "The capacity region of a channel with two senders and two receivers," submitted to Annals of Probability.

Gallager, R.G. (1968), "Information Theory and Reliable Communication," John Wiley and Sons, New York.

Shannon, C.E. (1961), "Two-way communication channels," Proc. Fourth Berkeley Symposium on Math. Statist. Prob. 1, 611-644, University of California Press, Berkeley.

Ulrey, M.L. (1973), "A coding theorem for a channel with s senders and r receivers," submitted for publication.

van der Meulen, E.C. (1971), "The discrete memoryless channel with two senders and one receiver," presented at the Second International Symposium on Information Theory at Tsahkadsor, Armenian S.S.R.. To appear in Problems of Control and Information Theory.

Wolfowitz, J. (1964), "Coding Theorems of Information Theory," second edition, Springer-Verlag, Berlin-Heidelberg-New York.