Table I presents the parameters $[n, k, d]$ for the five new codes C_i, and for the A_{ij}, $1 \leq i \leq 5$, $1 \leq j \leq 6$, required by Construction XX. Of the twenty codes A_{ij}, $1 \leq j \leq 4$, all but A_{11} and A_{13} are cyclic. A_{13} is a 5-dimensional subcode of A_{23} that contains the vector of weight 31. Likewise, $A_{11} = A_{12} + A_{13}$ is a 10-dimensional subcode of A_{21} containing the vector of weight 31. Of the remaining eighteen A_{ij}, $1 \leq j \leq 4$, only ten are distinct. Table II shows the roots of the check polynomial for these ten cyclic codes. For example, A_{12} and A_{22} are the same cyclic $[31, 6, 15]$ code with check polynomial of degree 6 and having roots 1 and a, where a is a primitive root of GF(32). A_{23} is also a $[31, 6, 15]$ cyclic code whose check polynomial has roots 1 and a^5. $A_{21} = A_{22} + A_{23}$ is a $[31, 11, 11]$ cyclic code whose check polynomial has roots 1, a and a^5. The check polynomials in Table II were selected to complement the generating polynomials listed in [3, App. D]. The ten codes A_{ij}, $5 \leq j \leq 6$, are, for the most part, trivial. The existence of the $[12, 7, 4]$ and $[18, 9, 6]$ codes is guaranteed by the tables in [1] and [4].

Clearly Construction XX has further applications. The five codes described here were selected to have $d \leq 29$, in order to enable comparison with the information in the table of [1].

References

[1] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*. New York: Elsevier/North Holland, 1977.

[2] N. J. A. Sloane, S. M. Reddy, and C. L. Chen, "New binary codes," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 503–510, July 1972.

[3] W. W. Peterson and E. J. Weldon, Jr., *Error-Correcting Codes*, 2nd Ed. Cambridge: MIT, 1972.

[4] H. J. Helgert and R. D. Stinaff, "Minimum distance bounds for binary linear codes," *IEEE Trans. Inform. Theory*, vol. IT-19, pp. 344–356, May 1973.

Improvements of Winograd's Result on Computation in the Presence of Noise

RUDOLF AHLSWEDE

***Abstract*—Winograd's result concerning Elias' model of computation in the presence of noise can be stated without reference to compntation. If a code φ: $\{0,1\}^k \rightarrow \{0,1\}^n$ is min-preserving ($\varphi(a \wedge b) = \varphi(a) \wedge \varphi(b)$ for $a, b \in \{0,1\}^k$) and ϵn-error correcting, then the rate $k/n \rightarrow 0$ as $k \rightarrow \infty$. This result is improved and extended in two directions.**

1) **For min-preserving codes with *fixed* maximal (and also average) error probability on a binary symmetric channel again $k/n \rightarrow 0$ as $k \rightarrow \infty$ (strong converses).**
2) **Second, codes with lattice properties without reference to computing are studied for their own sake. Already for monotone codes ($\varphi(a) \leq \varphi(b)$ for $a \leq b$) the results in direction 1) hold for maximal errors.**

These results provide examples of coding theorems in which entropy plays no role, and they can be reconsidered from the viewpoint of multiuser information theory.

I. Introduction

J. von Neumann [1] raised the question of how systems with unreliable components can be used efficiently in a reliable way. The most important example known, in which this program was carried out successfully, is Shannon's information theory. In another direction, Elias and Winograd studied the question whether reliable computation is possible at a positive rate in the presence of noise [3], [4].[1,2] Surprisingly, this depends on the Boolean operation used. If error correcting codes are used, for "+" the answer is positive [3] and for "$\wedge$" it is negative [4]. Here we show that the rate is zero also if the noise is modeled by a binary symmetric channel (BSC) and if maximal (resp. average) error probabilities tending to zero are used as performance criterion.[3] Moreover, we prove strong converses ([6]); that is, the rate is even zero for an arbitrary large (but < 1) probability of error. More abstractly, those problems can be studied as coding problems with algebraic constraints that have no reference to computing. Here we concentrate on lattice properties.

The proofs are based on elementary combinatorics. Standard Fano-type arguments for weak converses fail. In contrast to Shannon's theory, parameters of codes are not related to entropy here. Thus the present contribution can be viewed as another example in support of our abstract coding theory [6].

II. Elias' Model

Two strings of data

$$X_1 = (x_{11}, \cdots, x_{1k}), \quad X_2 = (x_{21}, \cdots, x_{2k})$$

are to be encoded by two separate encoders E_1, E_2 into strings of length n

$$E_1(X_1) = Y_1 = (y_{11}, \cdots, y_{1n})$$
$$E_2(X_2) = Y_2 = (y_{21}, \cdots, y_{2n}).$$

The quantity of interest is

$$f(X_1, X_2) = (f_1(x_{11}, x_{21}), f_2(x_{12}, x_{22}), \cdots).$$

However, the computation is done by a computer, operating on a bit-by-bit basis on Y_1, Y_2. In the absence of noise its computation would be

$$F(Y_1, Y_2) = (F_1(y_{11}, y_{21}), \cdots) \triangleq Z = (z_1, \cdots, z_n).$$

In the presence of noise it produces instead

$$Z^* = (z_1^*, \cdots, z_n^*) = Z + \text{noise}.$$

Here the noise is that of a binary symmetric channel. The decoder accepts Z^* as its input and performs the function D to obtain

$$U \triangleq D(Z^*) = (u_1, \cdots, u_k).$$

The whole system performs reliably if $U = f(X_1, X_2)$. Note that F can be different from f.

Elias makes the following assumptions.

1) X_1 and X_2 are encoded independently.
2) In the absence of noise D is bijective.
3) F operates bit by bit.

These assumptions mean that Y_i only carries information about X_i $(i = 1, 2)$ and does not carry information about logical combinations of the two blocks. Thus, none of the desired computation $f(X_1, X_2)$ is done in the encoder, which is also assumed to be noiseless here, or in the decoder. This is to say that all calculations will be done by the computer.

[1]Thanks are due to A. El Gamal for drawing our attention to the work of these authors.

[2]A referee kindly points out that [11] and [12] give related results and extensions in a different direction, respectively.

[3]This result was conjectured in [4, Abstract].

Manuscript received May 12, 1983; revised July 17, 1984.

The author is with the Fakultät für Mathematik, Universität Bielefeld, 4800 Bielefeld, Federal Republic of Germany.

0018-9448/84/1100-0872$01.00 ©1984 IEEE

III. Winograd's Result and Proof

Theorem: Suppose that

$$f(X_1, X_2) = f_1(x_{11}, x_{21}), \cdots, f_k(x_{1k}, x_{2k}) = (x_{11} \wedge x_{21}), \cdots, (x_{1k} \wedge x_{2k})$$

and

$$F(Y_1, Y_2) = (F_1(y_{11}, y_{21}), \cdots, F_n(y_{1n}, y_{2n})) = (y_{11} \wedge y_{21}, y_{12} \wedge y_{22}, \cdots, y_{1n} \wedge y_{2n}).$$

Then, in order to correct ϵn $(0 < \epsilon < 1)$ errors, necessarily $R = k/n \to 0$ $(n \to \infty)$.

Lemma: Let $D^{-1}(X_1 \wedge X_2) = E_1(X_1) \wedge E_2(X_2)$ and let

$$\bigvee_{\text{all } X} D^{-1}(X) = \mathbf{1}^n$$

then

$$D^{-1}(X) = E_1(X) = E_2(X), \qquad \text{for all } X$$

and

$$X_1 \geq X_2 \Rightarrow D^{-1}(X_1) \geq D^{-1}(X_2).$$

Here $\mathbf{1}^n = (1,1,\cdots,1)$ and $X_1 \geq X_2 \Leftrightarrow x_{1t} \geq x_{2t}$ for $t = 1,\cdots,k$ in $\{0,1\}$.

Proof of the Theorem: Given an error correcting code with additional monotonicity properties on the lattice $\{0,1\}^n$, if d denotes the Hamming distance and if W denotes the weight of a sequence, then

$$d(Z_1, Z_2) = W(Z_1) + W(Z_2) - 2W(Z_1 \wedge Z_2) \geq 2s + 1, \quad s \triangleq \epsilon n. \tag{1}$$

Also

$$Z_1 \geq Z_2 \Rightarrow W(Z_1) = W(Z_2) + d(Z_1, Z_2). \tag{2}$$

Consider the chain

$$\mathbf{0}^k \leq \cdots \leq \mathbf{1}^i\mathbf{0}^{k-i} \leq \cdots \leq \mathbf{1}^k.$$

By the lemma

$$D^{-1}(\mathbf{1}^k) \geq D^{-1}(\mathbf{1}^{k-1}0) \geq \cdots \geq D^{-1}(\mathbf{0}^k). \tag{3}$$

Equations (2) and (3) imply

$$n \geq W[D^{-1}(\mathbf{1}^k)] = \sum_{i=0}^{k} d(D^{-1}(\mathbf{1}^{k-i}\mathbf{0}^i), D^{-1}(\mathbf{1}^{k-i-1}\mathbf{0}^{i+1})) + W(D^{-1}(\mathbf{0}^k))$$

and thus by (1)

$$n \geq k(2s+1) + W(D^{-1}(\mathbf{0}^k)) \geq k(2s+1).$$

Q.E.D.

Observation: Winograd's proof only uses that $\varphi \triangleq D^{-1}$ is monotonic and injective. It does not use the property $\varphi(a \wedge b) = \varphi(a) \wedge \varphi(b)$.

IV. Lattice Codes

To better understand the coding problem treated, we state it in purely combinatorial terms without any references to computing. Henceforth, the number of codewords M will be a power of 2: $M = 2^k$.

Recall that an (M, n, t)-error correcting code is a set of words

$$\{u_1, \cdots, u_M\} \subset \{0,1\}^n$$

with

$$d(u_i, u_j) \geq 2t + 1, \qquad \text{for } i \neq j. \tag{4}$$

An $(M, n, \lambda_{\max})$-code (resp. $(M, n, \lambda_{\text{av}})$-code) for the BSC with the transmission matrix

$$w = \begin{pmatrix} 1-\epsilon & \epsilon \\ \epsilon & 1-\epsilon \end{pmatrix}$$

is the set of pairs $\{(v_i, D_i): 1 \leq i \leq M\}$, where $D_i \cap D_j = \phi$ $(i \neq j)$, $v_i \in \{0,1\}^n$, $D_i \subset \{0,1\}^n$, and

$$\max w(D_i^c | v_i) \leq \lambda_{\max}, \qquad \left(\text{resp. } \frac{1}{M} \sum_{i=1}^{M} w(D_i^c | v_i) \leq \lambda_{\text{av}}\right). \tag{5}$$

Motivated by the computing problem in the presence of noise as described earlier, Elias [3] considered the logical operation "+" for f_i and F_i. In this case the codewords $\{v_1, \cdots, v_M\}$ carry an additional algebraic structure: they form a group under modulo 2 componentwise addition or, equivalently, a subspace of $GF(2)^n$. Moreover, he proved that those group (or linear) codes achieve the capacity of the BSC. Replacing "+" by other logical operations, algebraic conditions are imposed on the codes.

Generally, we define algebraic codes (M, n, t, φ) (resp. (M, n, λ, φ)-codes, $\lambda = \lambda_{\max}$, or λ_{av}) by requiring that $\{u_1, \cdots, u_M\}$ (resp. $\{v_1, \cdots, v_M\}$) be parameterized by

$$\{u_1, \cdots, u_M\} = \{\varphi(z): z \in \{0,1\}^k\}$$

$$\left(\text{resp. } \{v_1, \cdots, v_M\} = \{\varphi(z): z \in \{0,1\}^k\}\right). \tag{6}$$

In this notation linear codes are obtained by an isomorphic vector space embedding of $GF(2)^k$ into $GF(2)^n$.

Motivated by Winograd's theorem, we are concerned here with maps φ, which preserve lattice properties of $\{0,1\}^k$, such as monotonicity

$$a \leq b \Rightarrow \varphi(a) \leq \varphi(b) \tag{7}$$

or preservation of the min-operation "$\wedge$"

$$\varphi(a \wedge b) = \varphi(a) \wedge \varphi(b), \quad a \wedge b = (a_1 \wedge b_1, \cdots, a_k \wedge b_k) \cdots. \tag{8}$$

We adopt the following notation: For a codelength M and a block length n

$$R \triangleq \frac{\log M}{n} = \frac{k}{n} \tag{9}$$

is the rate of the code. For a specified class Φ of maps, we define the optimal rates

$$r_n^{\Phi}(\epsilon) = \max\left\{\frac{\log M}{n}: \exists (M, n, t, \varphi)\text{-code for some } \varphi \in \Phi \text{ and } t = \epsilon n\right\} \tag{10}$$

$$R_n^{\Phi}(\lambda_{\max}) = \max\left\{\frac{\log M}{n}: \exists (M, n, \lambda_{\max}, \varphi)\text{-code for some } \varphi \in \Phi\right\} \tag{11}$$

$$\bar{R}_n^{\Phi}(\lambda_{\text{av}}) = \max\left\{\frac{\log M}{n}: \exists (M, n, \lambda_{\text{av}}, \varphi)\text{-code for some } \varphi \in \Phi\right\}. \tag{12}$$

The class of all ψ: $\{0,1\}^k \to \{0,1\}^n$ for some k, n that are injective and $\wedge$-preserving is denoted by Ψ. In this notation Winograd's theorem can be considered as follows. For $0 < \epsilon < 1$

$$r_n^{\Psi}(\epsilon) \leq (2\epsilon)^{-1} n^{-1} \to 0 \qquad (n \to \infty). \tag{13}$$

That is, an error correcting code cannot have a positive rate (as

$n \to \infty$), if ϵn $(0 < \epsilon < 1)$ many errors must be corrected. This "negative" result has not been established anywhere in the literature for (M, n, λ, Ψ)-codes. Next, we study the growth of $R_n^{\Psi}(\lambda_{\max})$ and show that it differs from that of $r_n^{\Psi}(\epsilon)$.

V. A Lower Bound on $R_n^{\Psi}(\lambda_{\max})$

Theorem 1: $R_n^{\Psi}(\lambda_{\max}) \ge d(\epsilon, \lambda) \cdot (\log n)^{-1}$ for a suitable $d(\epsilon, \lambda) > 0$.

Proof: Choose $n = b \cdot k$ with b to be specified later. For $x \in \{0,1\}$, let $(x)^b \triangleq (x, x, \cdots, x)$ be of length b and define φ: $\{0,1\}^k \to \{0,1\}^n$ as

$$\varphi(x^k) = \varphi(x_1, \cdots, x_k) = \left((x_1)^b, \cdots, (x_k)^b\right). \qquad (14)$$

Obviously, φ is injective and $\wedge, \vee, \ge$ -preserving. Thus, particularly $\varphi \in \Psi$. For the codewords $\{\varphi(x^k): x^k \in \{0,1\}^k\}$, define decoding sets D_{x^k} by maximum likelihood decoding with respect to the BSC (declare an error in case of ties). By symmetry the individual error probabilities are all equal and thus $\lambda_{\max} = \lambda_{\text{av}}$. We calculate now $\lambda_{\max}$. First observe that there are exactly $\binom{k}{l}$ codewords that differ from a fixed codeword $\varphi(x^k)$ in exactly l blocks. The two-codeword error probability of $\varphi(x_k)$ and such a codeword is less than $e^{-c(\epsilon) l b}$ for a suitable constant $c(\epsilon) > 0$. Therefore

$$\begin{aligned}\lambda_{\max} &\le \sum_{l=1}^{k} \binom{k}{l} e^{-c(\epsilon) l b} \le \sum_{l=1}^{k} k^l e^{-c(\epsilon) l b} \\ &\le k e^{-c(\epsilon) b}\left[1 - k e^{-c(\epsilon) b}\right]^{-1}, \qquad \text{if } k e^{-c(\epsilon) b} < 1 \\ &\le \lambda, \qquad \text{if } k e^{-c(\epsilon) b} < \tfrac{1}{2}\lambda.\end{aligned}$$

This condition holds for any b with

$$1 + c(\epsilon)^{-1}\left[\log k - \log\frac{\lambda}{2}\right] \ge b \ge c(\epsilon)^{-1}\left[\log k - \log\frac{\lambda}{2}\right]$$

and therefore

$$\begin{aligned}R_n^{\Psi}(\lambda) &\ge \frac{k}{n} = \frac{1}{b} \\ &\ge \left(1 - c(\epsilon)^{-1}\log\frac{\lambda}{2} + c(\epsilon)^{-1}\log k\right)^{-1} \\ &\ge \frac{c(\epsilon)}{2\log k}, \qquad \text{for } k \ge k_0(\epsilon, \lambda)\end{aligned}$$

which gives the result, because $n \ge k$.

VI. Weak Converses via Chains

One readily verifies

$$\varphi \wedge \text{-preserving} \Rightarrow \varphi \text{ monotonic}. \qquad (15)$$

Thus $\Psi \subset \mathscr{M}$, the class of injective, monotonic maps. Actually Winograd's proof, which is based on the properties of a chain of codewords, uses monotonicity (in the form $x_1 \ge x_2 \to D^{-1}(x_1) \ge D^{-1}(x_2)$ of the Lemma) rather than the $\wedge$-preserving property. Here we analyze how far this argument holds. Later we derive sharper results by looking at antichains.

Theorem 2: (Weak converse for maximum error.) For any null sequence $(\lambda_n)_{n=1}^{\infty}$, we have $\lim_{n\to\infty} R_n^{\mathscr{M}}(\lambda_n) = 0$.

Proof: Consider any chain of length k in $\{0,1\}^k$, such as

$$e_1, e_1 \vee e_2, e_1 \vee e_2 \vee e_3, \cdots, e_1 \vee e_2 \vee \cdots \vee e_k$$

where $e_i = (0,0,0,1,0,\cdots,0)$ with a one in the ith position. Then $\varphi(e_1) \le \varphi(e_1 \vee e_2) \le \cdots \le \varphi(e_1 \vee \cdots \vee e_k)$, and therefore there exists $c_1, \cdots, c_k \in \{0,1\}^n$ with

$$c_i \wedge c_j = \mathbf{0} = (0, \cdots, 0), \qquad i \ne j \qquad (16)$$

and

$$\varphi(e_1 \vee \cdots \vee e_l) = c_1 \vee c_2 \vee \cdots \vee c_l, \qquad l = 1, \cdots, k. \qquad (17)$$

For two successive codewords the two-codeword error probabilities satisfy for a suitable constant $f(\epsilon) > 0$

$$\begin{aligned}\lambda_n &\ge \max\left(\lambda(\varphi(e_1 \vee \cdots \vee e_l)), \lambda(\varphi(e_1 \vee \cdots \vee e_{l+1}))\right) \\ &\ge e^{-f(\epsilon) c_{l+1}}. \end{aligned} \qquad (18)$$

Therefore,

$$c_{l+1} \ge -\frac{1}{f(\epsilon)}\log\lambda_n,$$

$$n \ge \sum_{l=1}^{k} c_l \ge \frac{k}{f(\epsilon)}\log\lambda_n,$$

and

$$\frac{k}{n} \le -f(\epsilon)[\log\lambda_n]^{-1} \to 0, \qquad (n \to \infty). \qquad (19)$$

Q.E.D.

Next we establish a somewhat weaker result for average errors.

Theorem 3: ("Very" weak converse for average errors.) For any null sequence $(\lambda_n)_{n=1}^{\infty}$ with $0 \le \lambda_n \le e^{-\delta n}$, $\delta > 0$, $n \in \mathbb{N}$

$$\lim_{n\to\infty} \bar{R}_n^{\mathscr{M}}(\lambda_n) = 0.$$

To apply the previous argument, we must find a subcode of small maximal error probability that contains a chain of sufficient length.

In the multiuser information theory [8], the attempt to extract a suitable maximal error code from an average error code led to the combinatorial problem of Zarankiewicz. The situation here is similar. We are led to another combinatorial problem that was solved by Erdös [9] (Theorem 2.3 in [10]). For arbitrary $\mathscr{B} \subset \{0,1\}^k$, $|\mathscr{B}| = B$, what is the guaranteed length $L(B,k)$ of the longest chain in $\mathscr{B}$?

Define N_k by

N_k = maximal integer with

$$\begin{cases} \displaystyle\sum_{s=1}^{N_k} 2\binom{k}{\frac{k}{2}+s} + \binom{k}{\frac{k}{2}} \le B, & k \text{ is even} \\ \displaystyle\sum_{s=1}^{N_k} 2\binom{k}{\frac{k-1}{2}+s} \le B, & k \text{ is odd}. \end{cases} \qquad (20)$$

Erdös proved

$$L(B,k) \ge \begin{cases} 2N_k + 1, & \text{if } k \text{ is even} \\ 2N_k, & \text{if } k \text{ is odd}. \end{cases} \qquad (21)$$

Proof: Consider the subcode

$$\mathscr{C} = \left\{\varphi(x^k): x^k \in \{0,1\}^k\right.$$

and

$$\left. w\left(D_{x^k}^c \mid \varphi(x^k)\right) \le 2\lambda_n\right\}$$

which by a pigeonhole argument satisfies

$$|\mathscr{C}| \ge \frac{1}{2}2^k. \qquad (22)$$

Since $\binom{k}{k/2} \le 2^k/\sqrt{k}$, (22), (20), and (21) imply the existence of a chain $\mathscr{C}_0 \subset \mathscr{C}$ with

$$|\mathscr{C}_0| \geq \text{const.} \sqrt{k}. \tag{23}$$

The argument given in the proof of Theorem 2 can now be applied to the chain $\mathscr{C}_0$. The corresponding quantities $c_1, \cdots, c_{\text{const.}\sqrt{k}}$ satisfy

$$e^{-\delta k} \geq e^{-f(\epsilon) c_i}.$$

Thus,

$$c_i \geq \frac{\delta k}{f(\epsilon)}$$

$$n \geq \sum_{i=1}^{\text{const.}\sqrt{k}} c_i \geq \text{const.} \sqrt{k} \frac{\delta k}{f(\epsilon)}$$

and finally

$$\frac{k}{n} \leq \text{const.}\ k^{-1/2}.$$

Q.E.D.

Remark 1: Of course, the proof still works for

$$\lambda_n \leq e^{-n^{1/2+\eta}}, \qquad \eta > 0.$$

Remark 2: Notice that we have not proved the strong converse [6] that says here

$$\lim_{n \to \infty} R_n^{\mathscr{M}}(\lambda) = 0 \quad \left(\text{resp. } \lim_{n \to \infty} \overline{R}_n^{\mathscr{M}}(\lambda) = 0\right) \quad \text{for every constant } \lambda \in (0,1). \tag{24}$$

In contrast to many situations in coding theory, where the question whether both the weak converse and the strong converse hold is often just of academic interest; the strong converse is of great significance for computing problems. If it does not hold in a certain situation, then computing in the presence of noise is possible at a positive rate for certain error probabilities. Unfortunately the strong converse does hold, and thus the desired phenomenon does not occur for the class Ψ. For this class the sharper result can be derived by looking at antichains.

Remark 3: As an instructive exercise in abstract coding theory [6], we propose a problem: does (24) hold for maximal (or even average) errors?

VII. Strong Converse via Antichains

Theorem 4: (Strong converse for maximal error.)

$$\lim_{n \to \infty} R_n^{\Psi}(\lambda) = 0, \qquad \lambda \in (0,1).$$

More precisely,

$$d_1(\epsilon,\lambda)(\log n)^{-1} \leq R_n^{\Psi}(\lambda) \leq d_2(\epsilon,\lambda)(\log n)^{-1}$$

for suitable constants d_1, d_2, and $n \geq 2$.

Proof:

a) Suppose that $\varphi(0,\cdots,0)$ has a one in some components, then because of monotonicity all codewords $\varphi(x^k)$ have a one in those components. They therefore add nothing to the error performance and just decrease the rate. Thus without loss of generality (w.l.o.g.) suppose that

$$\varphi(0,\cdots,0) = (0,\cdots,0) \in \{0,1\}^n. \tag{25}$$

b) Since $e_i \wedge e_j = \mathbf{0}$ $(i \neq j)$, we have $\varphi(e_i \wedge e_j) = \varphi(e_i) \wedge \varphi(e_j) = \mathbf{0}$; that is, $\varphi(e_i)$ $(1 \leq i \leq k)$ have disjoint supports $S_1, \cdots, S_k \subset \{1,2,\cdots,n\}$, and *a fortiori* $\{\varphi(e_i): 1 \leq i \leq k\}$ is an antichain.

By the previous argument we can assume

$$\bigcup_{i=1}^{k} S_i = \{1,2,\cdots,n\}. \tag{26}$$

Thus, for $c_i \triangleq |S_i|$

$$\sum_{i=1}^{k} c_i = n. \tag{27}$$

We will derive the desired bound on the rate by analyzing the error performance of the antichain only.

c) Next, we modify this antichain so that the new supports have all equal cardinalities. Define $\bar{c} \triangleq \sum_{i=1}^{k} c_i/k$. Then $|\{i: c_i > 2\bar{c}\}| 2\bar{c} \leq k\bar{c} = n$ and hence $|\{i: c_i > 2\bar{c}\}| \leq (n/2\bar{c}) = k/2$. For the sake of convenience, set $m = k/2$ and $d = \lfloor 2\bar{c} \rfloor$. Thus we have a subantichain $\{\varphi(a_i): 1 \leq i \leq m\}$ with $c_i \leq d$. To simplify the calculation of the error probability, extend all c_i to d, disregarding the increase in the total length or equivalently the loss in rate. An antichain, which each codeword has by symmetry, is obtained for the same error performance in strict (disregarding ties) maximum likelihood decoding. We upper bound the probability of correct decoding for one codeword by a rough estimate.

d) For the strict maximum likelihood decoding code $\{(\varphi(a_i), D_{\varphi(a_i)}): 1 \leq i \leq m\}$

$$D_{\varphi(a_i)} \subset \left\{ y^n \in \{0,1\}^n : y^n \ngeq \varphi(a_j), \quad \text{for all } j \neq i \right\} \tag{28}$$

because $y^n \geq \varphi(a_j)$ implies $w(y^n|\varphi(a_j)) \geq w(y^n|\varphi(a_i))$. Obviously, from (28)

$$w\left(D_{\varphi(a_i)}|\varphi(a_i)\right) \leq \left(1-(1-\epsilon)^d\right)^{m-1} \tag{29}$$

and necessarily with $\eta \triangleq 1 - \epsilon$

$$\left(1-\eta^d\right)^{m-1} \geq 1-\lambda. \tag{30}$$

This gives a lower bound on d, and thus the desired upper bound on the rate is

$$R = \frac{k}{n} = \frac{1}{\bar{c}} \leq \frac{2}{d},$$

since $d \leq 2\bar{c}$. From (20)

$$\log(1-\eta^d) \geq \frac{\log(1-\lambda)}{m-1}$$

and therefore necessarily from $\log(1-x) \leq -x$,

$$-\eta^d \geq \frac{\log(1-\lambda)}{m-1},$$

$$d \geq \left(\log\left[-\frac{\log(1-\lambda)}{m-1}\right]\right) \log^{-1} \eta.$$

Thus

$$R \leq \frac{2}{d} \leq 2 \log \eta \frac{1}{\log\log(1-\lambda)^{-1} - \log(m-1)}.$$

Since $\eta = 1-\epsilon$, $m = k/2$, a constant $\rho(\lambda,\epsilon)$ exists for which

$$R = \frac{k}{n} \leq \rho(\lambda,\epsilon)(\log k)^{-1}, \qquad k \geq 4. \tag{31}$$

e) Finally, this bound is expressed in terms of n. Since $R \leq 1$, substitution of Rn for k in (31) yields

$$R \leq \rho \frac{1}{\log R + \log n}. \tag{32}$$

Suppose now that for *every* $\kappa > 0$ $R \geq \kappa/\log n$ for n large, then

$$\frac{\kappa}{\log n} \leq \frac{\rho}{\log \kappa - \log\log n + \log n}$$

or

$$\log \kappa - \log\log n + \log n \leq \frac{\rho}{\kappa}\log n,$$

a contradiction. Thus, $R_n^{\Psi}(\lambda) = O(1/\log n)$ and by Theorem 1 this bound is best within a constant for the first-order term.

Theorem 5: (Strong converse for average error)

$$\lim_{n\to\infty} \bar{R}_n^{\Psi}(\lambda) = 0, \qquad \lambda \in (0,1).$$

Actually,

$$\bar{d}_1(\epsilon,\lambda)(\log n)^{-1} \leq \bar{R}_n^{\Psi}(\lambda) \leq \bar{d}_2(\epsilon,\lambda)(\log n)^{-1}$$

for suitable constants d_1, d_2, and $n \geq 2$.

Proof:

a) *Auxiliary combinatorial lemma:* Paralleling the proof of Theorem 3, we find a sufficiently long maximal error subcode with $\lambda_{\max} = (1+\delta)\lambda$ that now has the structure of an antichain and an additional "disjointness property" like the one used in the proof of Theorem 4. For this purpose, we first derive an auxiliary combinatorial result.

Let $\mathscr{G}$ be a directed graph with vertex set $\mathscr{V}$. For $v \in \mathscr{V}$ denote by $\mathscr{I}(v)$ (resp. $\mathscr{O}(v)$) the set of vertices reaching v (resp. reached from v) by an arrow. Further, let

$$I_{\max} \triangleq \max_{v\in\mathscr{V}} |\mathscr{I}(v)|, \quad \bar{I} \triangleq |\mathscr{V}|^{-1}\sum_{v\in\mathscr{V}} |\mathscr{I}(v)|$$

denote the maximal and the average in degrees.

Lemma 1: Let $\mathscr{V}_1 \cup \mathscr{V}_2$ be a partition of the vertex set $\mathscr{V}$ of a directed graph $\mathscr{G}$ such that

$$|\mathscr{V}_1| > |\mathscr{V}|\left(1-(1-\sigma)\bar{I}I_{\max}^{-1}\right). \tag{33}$$

Then there exists a $v_0 \in \mathscr{V}$ with

$$|\mathscr{O}(v_0) \cap \mathscr{V}_1| \geq \sigma\bar{I}. \tag{34}$$

Proof: Suppose that (34) does not hold. Then

$$\sum_{v\in\mathscr{V}_1} |\mathscr{I}(v)| = \sum_{v\in\mathscr{V}} |\mathscr{O}(v)\cap\mathscr{V}_1| < |\mathscr{V}|\sigma\bar{I},$$

and therefore

$$|\mathscr{V}_2| I_{\max} \geq \sum_{v\in\mathscr{V}_2} |\mathscr{I}(v)| > |\mathscr{V}|(1-\sigma)\bar{I}.$$

Thus

$$|\mathscr{V}_2| > |\mathscr{V}|(1-\sigma)\bar{I}I_{\max}^{-1}$$

in contradiction to $|\mathscr{V}| = |\mathscr{V}_1| + |\mathscr{V}_2|$ and (33).

b) *Definition of the graph:* The following definitions are motivated by the fact that most words x^k are in the "middle" of the lattice and that we are interested in those with "good" codewords. Write $\lambda(x^k) \triangleq w(D_k^c|\varphi(x^k))$ and consider for $0 < \rho < 1/2$ and γ with $(1+\gamma)\bar{\lambda} < 1$ the sets

$$\mathscr{V} = \mathscr{V}(\rho) \triangleq \left\{x^k \in \{0,1\}^k : |d(x^k,\mathbf{0}) - \tfrac{1}{2}k| \leq \rho k\right\} \tag{35}$$

$$\mathscr{V}_1 = \mathscr{V}_1(\gamma,\rho) \triangleq \left\{x^k \in \mathscr{V}(\rho): \lambda(x^k) \leq (1+\gamma)\bar{\lambda}\right\}, \quad \mathscr{V}_2 \triangleq \mathscr{V} - \mathscr{V}_1 \tag{36}$$

$$\mathscr{C}(\gamma,\rho) \triangleq \left\{\varphi(x^k) : x^k \in \mathscr{V}_1(\gamma,\rho)\right\}. \tag{37}$$

Now define a directed graph $\mathscr{G}$ with vertex set $\mathscr{V} = \mathscr{V}(\rho)$ by

$$u \to v \Leftrightarrow u \leq v \tag{38}$$

where u, v differ in exactly one component.

c) *Application of Lemma 1:* To apply the lemma to the graph $\mathscr{G}$ and the partition $\mathscr{V} = \mathscr{V}_1 \cup \mathscr{V}_2$, we must first estimate $|\mathscr{V}|$, $|\mathscr{V}_1|$, $I_{\max}$, and $\bar{I}$.

It is well-known that

$$|\mathscr{V}| = |\mathscr{V}(\rho)| \geq (1 - e^{-f(\rho)k})2^k \tag{39}$$

for a suitable constant $f(\rho) > 0$. Since

$$\bar{\lambda}2^k \geq \sum_{x^k\in\{0,1\}^k} \lambda(x^k) \geq |\{x^k : \lambda(x^k) > (1+\gamma)\bar{\lambda}\}|(1+\gamma)\bar{\lambda}$$

we have

$$|\{x^k \in \{0,1\}^k : \lambda(x^k) \leq (1+\gamma)\bar{\lambda}\}| \geq \frac{\gamma}{1+\gamma}2^k$$

which together with (39) implies

$$|\mathscr{V}_1| = |\mathscr{V}_1(\gamma,\rho)| \geq \left(\frac{\gamma}{1+\gamma} - e^{-f(\rho)k}\right)2^k. \tag{40}$$

Obviously, from the definitions it follows that

$$I_{\max} = \lfloor(\tfrac{1}{2}+\rho)k\rfloor \tag{41}$$

and

$$(\tfrac{1}{2}-\rho)k \leq \bar{I} \leq (\tfrac{1}{2}+\rho)k. \tag{42}$$

Actually, $\bar{I}$ is almost equal to $(1/2)k$.

Next, we insure condition (33). Choose $\sigma \triangleq \gamma/4(1+\gamma)$, then $0 < \rho \leq 1/4$ and so small that $(1/2-\rho)/(1/2+\rho) \geq 1-\sigma$. Thus,

$$\begin{aligned}\frac{1}{2}\frac{\gamma}{1+\gamma} &= 2\sigma \geq 2\sigma - \sigma^2 \\ &= 1-(1-\sigma)^2 \geq 1-(1-\sigma)\cdot\frac{\frac{1}{2}-\rho}{\frac{1}{2}+\rho} \\ &\geq (1-\sigma)\bar{I}I_{\max}^{-1}\end{aligned}$$

and for $k \geq k_0(\gamma,\rho)$

$$\frac{1}{2}\frac{\gamma}{1+\gamma} \geq e^{-f(\rho)k}.$$

Therefore

$$\frac{\gamma}{1+\gamma} - e^{-f(\rho)k} \geq 1-(1-\sigma)\bar{I}I_{\max}^{-1}. \tag{43}$$

This, $|\mathscr{V}| \leq 2^k$, and (40) imply condition (33).

The lemma guarantees the existence of a $v_0 \in \mathscr{V}$ with

$$|\mathscr{O}(v_0)\cap\mathscr{V}_1| \geq \sigma\bar{I} \geq \frac{1}{4}\frac{\gamma}{1+\gamma}\frac{1}{4}k \tag{44}$$

by the definition of σ, (42), and $\rho \leq 1/4$.

d) *The desired antichain in* $\mathscr{C}(\gamma,\rho)$: For $k' = \gamma k/16(1+\gamma)$ consider a subset

$$\{a_1,\cdots,a_{k'}\} \subset \mathscr{O}(v_0)\cap\mathscr{V}_1 \tag{45}$$

and define the subcode

$$\mathscr{C}^* = \{\varphi(a_i) : 1 \leq i \leq k'\} \subset \mathscr{C}(\gamma,\rho). \tag{46}$$

By the definitions of the graph and the a_i, $\varphi(a_i) \geq \varphi(v_0)$ for all i and $a_i \wedge a_j = v_0$ for $i \neq j$. Since $\varphi \in \psi$, we have

$$\varphi(a_i) \wedge \varphi(a_j) = \varphi(a_i \wedge a_j) = \varphi(v_0). \tag{47}$$

This means that the supports of the $\varphi(a_i)$ contain the support of $\varphi(v_0)$ and are disjoint in its complement. Dropping the components in the support of $\varphi(v_0)$ leads to a code of the same rate and error performance. This is exactly the maximal error code used in the proof of Theorem 4, and the proof of Theorem 5 can be completed in the same way.

Another Proof of Theorem 5: The preceding proof was based on a rather general idea. Using more of the lattice structure, we can provide a simpler proof.

Define $\mathscr{X}_j^k = \{x^k \in \{0,1\}^k\colon\ d(x^k, \mathbf{0}) = j\}$. From the argument leading to (40), for any $0 < \rho < 1/2$ an l exists, $k/2 \le l < (\frac{1}{2} + \rho)k$, such that for

$$\mathscr{V}_1' \triangleq \{x^k\colon x^k \in \mathscr{X}_l^k,\ \lambda(x^k) \le \bar{\lambda}(1+\gamma)\}$$

$$|\mathscr{V}_1'| \ge \delta(\bar{\lambda}, \gamma, \rho)\binom{k}{l} \tag{48}$$

for a suitable δ. The following (44′) replaces (44). From there the proof can be completed as in d.

Lemma 2: A $v_0' \in \mathscr{X}_{l-1}^k$ exists such that

$$|\mathscr{O}(v_0) \cap \mathscr{V}_1'| \ge \delta(k - (l-1)). \tag{44'}$$

Proof: Look at the bipartite graph

$$(\mathscr{X}_{l-1}^k, \mathscr{X}_l^k, \mathscr{E}')$$

where

$$(x^k, y^k) \in \mathscr{E}' \Leftrightarrow x^k < y^k.$$

Obviously,

$$|\mathscr{E}'| = \binom{k}{l-1}(k-(l-1)) = \binom{k}{l}l. \tag{49}$$

The number of edges leaving $\mathscr{V}_1'$ is bigger than $\delta\binom{k}{l}l$. Now suppose that for all $v \in \mathscr{X}_{l-1}^k$

$$|\mathscr{O}(v) \cap \mathscr{V}'| < \delta(k-(l-1)).$$

Then the number of edges to $\mathscr{V}_1'$ is smaller than

$$\binom{k}{l-1}\delta(k-(l-1)) = \delta\binom{k}{l}l$$

which is a contradiction. Q.E.D.

References

[1] J. von Neumann, "Probabilistic logics and the synthesis of reliable organisms from unreliable components," in *Automata Studies*, Annals of Mathematical Studies, vol. 34, C. E. Shannon and J. McCarthy, Ed., Princeton, N.J.: Princeton Univ. 1956, pp. 43–98.

[2] E. F. Moore and C. E. Shannon, "Reliable circuits using less reliable relays," *J. Franklin Inst.*, vol. 262, pp. 191–208, 281–297, 1956.

[3] P. Elias, "Computation in the presence of noise," *IBM J. Res. Develop.*, vol. 3, pp. 346–353, 1958.

[4] S. Winograd, "Coding for logical operations," *IBM. J. Res. Develop.*, vol. 6, pp. 430–436, 1962.

[5] S. Winograd and J. D. Cowan, *Reliable Computation in the Presence of Noise.* Cambridge: MIT, 1963.

[6] J. Wolfowitz, "The coding of messages subject to chance errors," *Illinois J. Math.*, vol. 1, no. 4, pp. 591–605, 1957.

[7] R. Ahlswede, "Coloring hypergraphs: A new approach to multi-user source coding," *J. Comb., Inform. Theory, and Syst. Sci.*, part I, vol. 4, no. 1, pp. 76–115, 1979, part II, vol. 5, no. 3, pp. 220–268, 1980

[8] R. Ahlswede, "On two way-communication channels and a problem by Zarankiewicz," in *Sixth Prague Conf. Information Theory, Statistical Decisions, Fct's and Rand. Proc.*, Sep. 1971, Czechoslovakia Academy of Sciences, pp. 23–37, 1973.

[9] P. Erdös, "On a lemma of Littlewood and Offord," *Bull. Amer. Math. Soc.*, vol. 51, pp. 898–902, 1945.

[10] C. Greene and D. J. Kleitman, "Proof techniques in the theory of finite sets," in *Studies in Combinatorics*, G. C. Rota, Ed., Studies in Mathematics, vol. 17, Publ. Math. Assoc. of America, pp. 22–79, 1978.

[11] W. W. Peterson and M. O. Rabin, "On codes for checking logical operations," *IBM J. Res. Develop.*, vol. 3, pp. 163–168, 1959.

[12] S. Winograd, "Redundancy and Complexity of Logical Elements," *Inform. Contr.*, vol. 5, pp. 177–194, 1963.

TABLE I

	Allowable Relative Error					
	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
Lower Bound in [1]	2.06	3.96	7.28	13.09	23.35	*
Q_1	2.91	9.91	31.6	*	*	*
Q_2	1.54	3.4	6.48	11.8	21.1	*
Q_3	1.16	2.3	3.83	5.95	8.98	13.4
Q_4	.95	1.83	2.89	4.21	5.89	8.08
Q_5	.83	1.55	2.39	3.38	4.55	5.98
Q_6	.73	1.37	2.08	2.88	3.8	4.87

Asterisks denote values that were not computed due to underflow in the computer.

A Sequence of Upper and Lower Bounds for the Q Function

THOMAS K. PHILIPS AND AHMED SAHRAOUI

***Abstract*—A sequence of upper and lower bounds for the Q function defined as $Q(x) = 1/\sqrt{2\pi}\int_x^\infty \exp[(-y^2)/2]\,dy$ is developed. These bounds are shown to be tighter than those most commonly used.**

It has been shown [1], for the Q function defined as

$$Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^\infty \exp\frac{(-y^2)}{2}\,dy,$$

that for $x > 0$

$$\frac{1}{\sqrt{2\pi}}\left(1 - \frac{1}{x^2}\right)\frac{\exp(-x^2/2)}{x} < Q(x) < \frac{1}{\sqrt{2\pi}}\frac{\exp(-x^2/2)}{x}.$$

Our aim is to develop tighter bounds. Consider rewriting the Q function as

$$Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^\infty 1 \cdot \exp\frac{(-y^2)}{2}\,dy,$$

and note that if $y/x > 1$, then $P_n(y) = 1 - \{1 - (y/x)\}^n$ is an upper bound for 1 if n is odd and a lower bound for it if n is even. Now define

$$A_n(x) \triangleq \frac{1}{\sqrt{2\pi}}\int_x^\infty P_n(y) \cdot \exp\frac{(-y^2)}{2}\,dy. \tag{1}$$

Manuscript received May 8, 1984; revised June 27, 1984. This work was supported in part from National Science Foundation Grant ECS-83-10771.

The authors are with the Department of Electrical and Computer Engineering, University of Massachusetts, Amherst, MA 01003 USA.

0018-9448/84/1100-0877$01.00 ©1984 IEEE