THE METRIC STRUCTURE OF CODES FOR THE BINARY SYMMETRIC CHANNEL

A. J. THOMASIAN
UNIVERSITY OF CALIFORNIA, BERKELEY

1. Introduction and summary

The main result of this paper, theorem 4, is that there exists an infinite sequence of binary digits which can be used to generate sliding parity check codes for the binary symmetric channel (BSC), and the error probability for such codes is close, in a certain sense, to minimum *uniformly* in all of the parameters: block length, probability of error for a single digit, and rate not too much less than capacity. This result is obtained by studying the metric structure which can be required of codes, and then relating this to the error probability. The paper is essentially self-contained.

Let B^n be the set of 2^n ordered n-tuples from $B = B^1 = \{0, 1\}$ where an element of B is called a binary digit or bit. For any $n \geqq 1$ we define a metric on B^n by $\overline{ab}$ = the number of coordinates in which a and b differ. An n-code C_n is a nonempty subset of B^n; an element $c \in C_n$ is called a codeword. A code is any set which is an n-code for some n. The code C_n in a sequence of codes $C_1, C_2, \cdots$ is always an n-code. The probability law of a BSC is defined as follows: for each $a \in B^n$ there is a probability distribution on B^n given by

$$P\{b|a\} = p^{\overline{ab}}q^{n-\overline{ab}}, \tag{1}$$

where $0 < p < 1/2$ and $q = 1 - p$. A statement which is made for all p means for all p such that $0 < p < 1/2$. Here $P\{b|a\}$ is the probability that the channel output is b given that the channel input is a. When $n = 1$ we have $P\{0|0\} = P\{1|1\} = q$, $P\{1|0\} = P\{0|1\} = p$ so that the probability p of an error in a single digit does not depend on what that digit is, hence the "symmetric" designation.

A decoder for an n-code C_n is a set D of disjoint subsets of B^n and a 1 to 1 correspondence between C_n and D. Assume for the moment that the elements of C_n and D are ordered by indices so that $C_n = \{c_1, \cdots, c_s\}$, $D = \{D_1, \cdots, D_s\}$ and c_k corresponds to D_k under the 1 to 1 correspondence. In application, the sender and receiver first decide on C_n and D and then use these repeatedly. The

This research was supported by the Information Systems Branch of the Office of Naval Research under Contract Nonr-222(53).

sender feeds a $c_k \in C_n$ into the channel and the receiver observes $b \in B^n$ at the output with probability

$$p^{\overline{c_k b}} q^{n-\overline{c_k b}}. \tag{2}$$

If $b \in D_j$ then the receiver decides that c_j was sent. Thus the probability that the receiver makes an error, if c_k is sent, is $P\{D_k^c | c_k\}$. Let

$$e_p(C_n) = \min_{D} \max_{k} P\{D_k^c | c_k\}. \tag{3}$$

Thus $e_p(C_n)$ is the smallest probability of error that can be guaranteed for all codewords of C_n if a best decoder, for this purpose, is used and the probability of a single digit being received in error is p.

The rate $R_n = R(C_n)$ of an n-code C_n is defined by $R_n = (1/n) \log$ (the number of codewords in C_n), where all logarithms are to the base 2, so that C_n has 2^{nR_n} codewords. The function $H(x) = -x \log x - (1 - x) \log (1 - x)$, $H(0) = 0$, $H(1/2) = 1$, is restricted to the interval $0 \leqq x \leqq 1/2$ so that it is continuous and strictly increasing and has a unique inverse. The well-known channel capacity for the BSC is $1 - H(p)$.

One of the results obtained is an upper bound for $e_p(C_n)$. The decoder used for this purpose associates to c_k, the set

$$D_k = \{b | \overline{c_k b} \leqq np_n' \text{ and } \overline{c_k b} < \overline{c_j b} \text{ for all } j \neq k\}, \tag{4}$$

where p_n' is defined by $R_n = 1 - H(p_n')$. Thus D_k consists of those b, within a sphere of radius np_n' of c_k, which are strictly closer to c_k than to any other codeword. Note that this decoder depends only on C_n and not on p. Clearly these D_k are disjoint, so that

$$e_p(C_n) \leqq \max_{k} P\{D_k^c | c_k\}. \tag{5}$$

In this notation Shannon's coding theorem for the BSC is

THEOREM 1. (a) *Direct half: For any p and R satisfying $0 < R < 1 - H(p)$ there exists a sequence of codes $C_1, C_2, \cdots$ such that $R(C_n) \rightarrow R$ and $e_p(C_n) \rightarrow 0$.*

(*b*) *Converse half: For any p and $R > 1 - H(p)$, if $C_1, C_2, \cdots$ is a sequence of codes with $R(C_n) \rightarrow R$, then $e_p(C_n) \rightarrow 1$.*

Our concern is only with the direct half of the theorem. Now $e_p(C_n)$ can be made to go to zero exponentially and the best exponential error rate is known for a certain interval below capacity. Theorem 2 gives a lower bound for the limiting behavior while theorem 3 states that this lower bound is attainable, in the limit, for a certain interval of rates below capacity. This interval is given by

$$1 - H\left(\frac{\sqrt{p}}{\sqrt{p} + \sqrt{q}}\right) < R = 1 - H(p') < 1 - H(p), \tag{6}$$

where

$$p < \frac{\sqrt{p}}{\sqrt{p} + \sqrt{q}} < \frac{1}{2}, \tag{7}$$

so that the interval is always nonempty. The parameter p', defined in terms of R, is a useful one for the statement of the result. For later convenience the sequence of codes in theorem 3 starts with C_2.

THEOREM 2. *For any p and R satisfying $0 < R = 1 - H(p') < 1 - H(p)$, if $C_1, C_2, \cdots$ is a sequence of codes satisfying $R(C_n) \to R$, then*

$$\liminf_n \left(\frac{1}{n}\right) \log e_p(C_n) \geqq H(p') + p' \log p + (1 - p') \log q < 0. \tag{8}$$

THEOREM 3. *For any p and R satisfying* (6) *there exists a sequence of codes $C_2, C_3, \cdots$ such that $R(C_n) \to R$ and*

$$\lim \left(\frac{1}{n}\right) \log e_p(C_n) = H(p') + p' \log p + (1 - p') \log q < 0. \tag{9}$$

If $f(p) = H(p') + p' \log p + (1 - p') \log q$, then $f(p') = 0$ and [the natural logarithm of 2][$(df)/(dp)$] $= (p' - p)/(pq) > 0$ if $p < p'$. Now $H(p) < H(p')$ in theorems 2 and 3 and this implies that $p < p'$ so that $f(p) < 0$ as stated.

Our interest is in theorem 3 and in certain ways that the sequence of codes $C_2, C_3, \cdots$ can be restricted and still make theorem 3 true. A restricted class of codes of interest is the parity check (PC) codes. Given an integer n and a number R_n, $0 < R_n < 1$, such that $nR_n = k$ is also an integer, an n-PC code C_n, with rate R_n, is an n-code which can be generated from some matrix a_{ij}, $i = 1, \cdots, n - k; j = 1, \cdots, k$; where each $a_{ij} \in B$. The 2^k codewords of C_n are generated as follows: given any $(b_1, \cdots, b_k) \in B^k$, extend it to B^n by determining $b_{k+1}, \cdots, b_n$ from

$$b_{k+i} = \sum_{j=1}^{k} a_{ij} b_j \qquad i = 1, \cdots, n - k \tag{10}$$

where mod 2 arithmetic is used, that is, $0 \cdot 0 = 0 \cdot 1 = 1 \cdot 0 = 0, 1 \cdot 1 = 1, 0 + 0 = 1 + 1 = 0, 0 + 1 = 1 + 0 = 1$. The codewords of C_n are the 2^k n-tuples $(b_1, \cdots, b_n)$ obtained in this way as $(b_1, \cdots, b_k)$ ranges over B^k. A special class of PC codes is sliding parity check (SPC) codes. If the generating matrix of an n-PC code satisfies $a_{ij} = a_{i+j-1}$, where $a_1, \cdots, a_{n-1}$ is a given sequence, the resulting code is an n-SPC code. Thus, any binary sequence $a_1, \cdots, a_{n-1}$ determines an n-SPC code for every rate $R_n = 1/n, \cdots, (n - 1)/n$.

Theorems 2 and 3, and the fact that theorem 3 remains true if the codes are required to be SPC codes, were proved by Elias [1], [2], who also obtained bounds on the probability of error. Proofs of theorems 1, 2, 3 and a modified form of Elias' proof showing that theorem 3 remains true if the codes are required to be PC codes can be found in Feinstein [3].

The advantages of using PC codes in the physical implementation of the coding operation are fairly obvious. An n-SPC code requires only the storage of $(n - 1)$ bits, an n-PC code requires only the storage of $(1 - R_n)nR_nn \leqq n^2/4$ bits, and a general n-code requires the storage of $n2^{nR_n}$ bits. If R_n is, say, 1/2, and n is moderately large, say 100 or 1000, then a general n-code would require an astronomical amount of storage, while an n-SPC code requires only a reasonable amount of storage, and some simple algebraic computations when in use.

Thus the implementation of the coding operation for an n-SPC code, once $a_1, \cdots, a_{n-1}$ have been obtained, seems to be feasible for moderately large n; however, the exhaustive method of examining all 2^{n-1} such sequences of $a_1, \cdots, a_{n-1}$ and then selecting the best one, is clearly not a workable method for finding a satisfactory $a_1, \cdots, a_{n-1}$. Very little is known in this direction.

Now to say that theorem 3 remains true when the codes are required to be SPC codes means that for every p and R satisfying (6) there exists a sequence $d_2, d_3, \cdots$, where $d_n \in B^{n-1}$, such that the corresponding sequence of SPC codes satisfies (9) and $R_n \to R$. A natural question to ask is whether there exists one infinite sequence $\alpha = (a_1, a_2, \cdots)$ such that we can let $d_n = (a_1, \cdots, a_{n-1})$. It would be better yet if this same α could be used for all p and R satisfying (6). This will be shown to be true (theorem 5). From a practical point of view this fact still leaves the concern that any such α might produce only terrible codes for moderate n. Thus what we do (theorem 4) is prove that there exists an α such that any SPC code C_n constructed from it has its $e_p(C_n)$ bounded above by a certain simple function of n, R_n, and p; and this bound has the obvious desired asymptotic behavior for any R, p satisfying (6). To the extent that the error bound is satisfactory, and assuming that the restriction on R_n is not objectionable, the result says that there is one α which can be used to design satisfactory SPC codes for all parameter (n, R_n, p) values for the BSC. Unfortunately, we cannot exhibit such an α.

The method of proof has several advantages in addition to being fairly simple and making the results more transparent. For one thing, if a particular n-PC code C_n is being considered, one can compute a certain function, N_d, and then immediately obtain a bound on $e_p(C_n)$, and also immediately make statements like the following: "Regardless of which codeword c is the input, the receiver will not make a mistake if $\overline{cb} \leqq 34$, where b is the received n-tuple; also, the receiver will make a mistake in at most 2 per cent of the cases where $\overline{cb} = 41$." The more usual, and much longer, computational approach is first to get all of the data similar to that in the previous sentence, and then compute $e_p(C_n)$. Statements like the quoted one are particularly powerful in that they don't depend on the probability law of the binary channel and so have some use even when the errors are dependent and nonstationary.

Let $\overline{B}$ be the class of infinite sequences $\alpha = (a_1, a_2, \cdots)$ where each $a_i \in B$. Let Q be the probability measure on the usual σ-field of subsets of $\overline{B}$ which makes the coordinate random variables independent and uniformly distributed, that is, $Q\{\alpha | a_1 = a_1', a_2 = a_2', \cdots, a_n = a_n'\} = (1/2)^n$. The principal result is

THEOREM 4. *For any $\beta > 1$ there is a set $B_\beta \subset \overline{B}$ with $Q(B_\beta) \geqq 1 - (1/\beta)$ such that for every $\alpha = (a_1, a_2, \cdots) \in B_\beta$ the following is true: For any p, R_n, $n \geqq 2$ such that nR_n is an integer and*

$$(11) \qquad 1 - H\left(\frac{\sqrt{p}}{\sqrt{p} + \sqrt{q}}\right) < R_n = 1 - H(p_n') < 1 - H(p),$$

the n-SPC code C_n with rate R_n which is generated from $a_1, \cdots, a_{n-1}$ satisfies

$$(12) \qquad \frac{1}{n}\log e_p(C_n) \leqq \frac{1}{n}\left[\log\left(\frac{q}{p}\frac{\beta}{2}\right) + 5\log n\right] + H(p_n') + p_n'\log p + (1 - p_n')\log q.$$

$Q\{B_\beta\} > 0$ so that B_β is not empty; and since the first term on the right side of (12) goes to 0 as n goes to infinity, such a sequence of codes satisfies theorem 3.

By taking $\beta = 2, 3, \cdots$ and letting $F = \bigcup_{i=2}^{\infty} B_i$, it is clear that theorem 4 immediately implies

THEOREM 5. *There is a set* $F \subset \overline{B}$ *with* $Q(F) = 1$ *such that for any* $\alpha = (a_1, a_2, \cdots) \in F$, *and any* p, R *satisfying* (6), *the following is true*: *If* $\{R_n\}$ *is a sequence of rates with* $R_n \to R$, *and* $\{nR_n\}$ *is always integer-valued, and* C_n *is the* n-SPC *code of rate* R_n *generated from* $(a_1, \cdots, a_{n-1})$, *then* (9) *holds.*

Thus if an $\alpha \in \overline{B}$ is selected at random, according to the Q distribution, then with Q-probability 1 it will generate an asymptotically optimum sequence of SPC codes for every p, R satisfying (6). However, $Q(F) = 1$ implies that there is at least one $\alpha \in F$ with only, say, one 1 in its first 1000 coordinates, and it is easy to show that the n-SPC codes such an α generates for $n \leqq 1000$ are usually terrible. Such possibilities make a result like theorem 4 more useful than theorem 5.

2. The existence of codes with certain metric properties

It is intuitively reasonable that a code whose codewords are far apart should be a good code. Given an n-code $C_n = \{c_1, \cdots, c_s\}$ and an integer d, $1 \leqq d \leqq n$, let $N_d(c_m)$ equal the number of codewords whose distance from codeword c_m is d. The existence of codes with specified upper bounds for $N_d(c_m)$ will be proved, and then, in lemma 3 and the proof of theorem 4, such bounds will be converted into upper bounds on $e_p(C_n)$.

There is a simple way to guess at the nature of the bounds that can be obtained for $N_d(c_m)$. Say that we wish to construct an n-code with rate R_n, and attempt to do so by selecting 2^{nR_n} elements of B^n at random. If c_m is the mth codeword selected then there is probability $\binom{n}{d}(1/2^n)$ that any particular one of the $2^{nR_n} - 1$ other selections will result in a codeword at distance d from c_m. Thus the expected number of codewords at distance d from c_m is about $2^{nR_n}\binom{n}{d}(1/2^n)$, and this is the form of the bound that we will obtain for $N_d(c_m)$. Define p_n' by $R_n = 1 - H(p_n')$; then an elementary fact, lemma 4b, implies that

$$(13) \qquad 2^{nR_n}\binom{n}{d}\frac{1}{2^n} \leqq 2^{[H(d/n) - H(p'_n)n]},$$

and by use of Stirling's formula this can be converted into an approximate equality for large n. This indicates that the parameter np_n' may be interpreted as the distance d at which $N_d(c_m)$ becomes significant, if the code is selected at random.

One of the nice properties of PC codes is that if $C_n = \{c_1, \cdots, c_s\}$ is an n-PC code then $N_d(c_m)$ is independent of m. In order to prove this, define the weight $w(b)$ of an element $b \in B^n$ to be the number of coordinates of b which equal 1. If $b = (b_1, \cdots, b_n)$ and $b' = (b'_1, \cdots, b'_n)$ are elements of B^n then let $b + b' = (b_1 + b'_1, \cdots, b_n + b'_n)$ where the addition is mod 2. Clearly there will be a contribution of 1 to $w(b + b')$ from the ith coordinate of $(b + b')$ if and only if $b_i \neq b'_i$, so that $\overline{bb'} = w(b + b')$. Thus $N_d(c_m)$ equals the number of times that $w(c_m + c)$ equals d when c ranges over C_n precisely once. Now let C_n be an n-PC code of rate k/n generated by a matrix a_{ij}. Clearly $0 = (0, \cdots, 0) \in C_n$ and $N_d(0) = N_d =$ the number of codewords of weight $d =$ the number of times that

$$w\left(b_1, \cdots, b_k, \sum_{j=1}^{k} a_{1j}b_j, \cdots, \sum_{j=1}^{k} a_{n-k,j}b_j\right) \tag{14}$$

equals d, as $(b_1, \cdots, b_k)$ ranges over B^k precisely once. Let c_m be a particular codeword generated by, say, $(b'_1, \cdots, b'_k)$, then $N_d(c_m)$ is the number of times that

$$\begin{aligned} w\Bigg(b_1 + b'_1, \cdots, b_k + b'_k, \sum_{j=1}^{k} a_{1j}b_j + \sum_{j=1}^{k} a_{1j}b'_j, \cdots, \\ \sum_{j=1}^{k} a_{n-k,j}b_j + \sum_{j=1}^{k} a_{n-k,j}b'_j\Bigg) \\ = w\left(b_1 + b'_1, \cdots, b_k + b'_k, \sum_{j=1}^{k} a_{1j}(b_j + b'_j), \cdots, \sum_{j=1}^{k} a_{n-k,j}(b_j + b'_j)\right) \end{aligned} \tag{15}$$

equals d as $(b_1, \cdots, b_k)$ ranges over B^k precisely once. However $(b_1 + b'_1, \cdots, b_k + b'_k)$ ranges over B^k precisely once if $(b_1, \cdots, b_k)$ ranges over B^k precisely once; so that $N_d(c_m) = N_d$. We henceforth restrict ourselves to SPC codes and we will work with N_d, although some of the results, for example, lemma 3, can be applied to any code.

LEMMA 1. *For any β, n, k, d satisfying $\beta > 1$, $n \geqq 2$, $1 \leqq k \leqq n$, $1 \leqq d \leqq n$, there is a set $B_\beta(n, k, d) \subset \overline{B}$ with $Q\{B_\beta(n, k, d)\} \geqq 1 - (1/\beta)$ such that for any $\alpha = (a_1, a_2, \cdots) \in B_\beta(n, k, d)$ the n-SPC code of rate k/n which is generated from $a_1, \cdots, a_{n-1}$ satisfies $N_d \leqq \beta 2^{k-n}\binom{n}{d}$.*

PROOF. Let β, n, k, d be given and let $\bar{b}_1, \cdots, \bar{b}_k, \bar{a}_1, \cdots, \bar{a}_{n-1}$ be random bits, that is,

$$P(\bar{b}_1 = b_1, \cdots, \bar{b}_k = b_k, \bar{a}_1 = a_1, \cdots, \bar{a}_{n-1} = a_{n-1}) = \left(\frac{1}{2}\right)^{n+k-1} \tag{16}$$

for any bits $b_1, \cdots, b_k, a_1, \cdots, a_{n-1}$. Define the random variables $\bar{b}_{k+1}, \cdots, \bar{b}_n$ just as in the definition of an SPC code, that is,

$$\bar{b}_{k+r} = \bar{a}_r\bar{b}_1 + \bar{a}_{r+1}\bar{b}_2 + \cdots + \bar{a}_{r+k-1}\bar{b}_k, \qquad 1 \leqq r \leqq n - k. \tag{17}$$

Now let $\bar{b} = (\bar{b}_1, \cdots, \bar{b}_k)$, $\bar{b}' = (\bar{b}_{k+1}, \cdots, \bar{b}_n)$, and $0 = (0, 0, \cdots, 0)$. Clearly $P\{\bar{b}' = 0 | \bar{b} = 0\} = 1$.

We first prove

(i) If $0 \neq b \in B^k$, $b' \in B^{n-k}$ then $P\{\bar{b}' = b' | \bar{b} = b\} = 1/2^{n-k}$. That is, if we are given $\bar{b} = b \neq 0$ then $\bar{b}_{k+1}, \cdots, \bar{b}_n$ are random bits. Fix $b \neq 0$ and let $P'\{\cdot\} = P\{\cdot | \bar{b} = b\}$ be the distribution conditioned on $\bar{b} = b$. Thus we wish to prove that $P'\{\bar{b}' = b'\} = 1/2^{n-k}$ for all $b' \in B^{n-k}$. Let b_l be the last coordinate of b which is 1, so that $b_l = 1$, $b_{l+1} = \cdots = b_k = 0$, and

$$(18) \qquad \bar{b}_{k+r} = \bar{a}_r b_1 + \bar{a}_{r+1} b_2 + \cdots + \bar{a}_{r+l-2} b_{l-1} + \bar{a}_{r+l-1}.$$

Now

$$(19) \qquad \begin{aligned} &P'\{\bar{b}_{k+1} = b'_{k+1} | \bar{a}_1 = a_1, \cdots, \bar{a}_{l-1} = a_{l-1}\} \\ &\quad = P'\{\bar{a}_1 b_1 + \cdots + \bar{a}_{l-1} b_{l-1} + \bar{a}_l = b'_{k+1} | \bar{a}_1 = a_1, \cdots, \bar{a}_{l-1} = a_{l-1}\} = \frac{1}{2} \end{aligned}$$

for any bits $b'_{k+1}, a_1, \cdots, a_{l-1}$, because the addition is mod 2. Thus $P'(\bar{b}_{k+1} = 0) = P'(\bar{b}_{k+1} = 1) = 1/2$ and in order to prove (i) it will be enough to show that

$$(20) \qquad P'\{\bar{b}_{k+r} = b'_{k+r} | \bar{b}_{k+1} = b'_{k+1}, \cdots, \bar{b}_{k+r-1} = b'_{k+r-1}\} = \frac{1}{2}$$

for all $b' \in B^{n-k}$ and all r with $2 \leqq r \leqq n - k$. But $\bar{b}_{k+1}, \cdots, \bar{b}_{k+r-1}$ are functions of only $\bar{a}_1, \cdots, \bar{a}_{l+r-2}$ so that it is enough to show that

$$(21) \qquad P'\{\bar{b}_{k+r} = b'_{k+r} | \bar{a}_1 = a_1, \cdots, \bar{a}_{l+r-2} = a_{l+r-2}\} = \frac{1}{2}$$

for all bits $b'_{k+r}, a_1, \cdots, a_{l+r-2}$ and all r with $2 \leqq r \leqq n - k$. But this is clearly true from (8) so that (i) is proved.

Clearly (i) implies that

$$(22) \qquad P\{(\bar{b}, \bar{b}') = 0\} = \frac{1}{2^k}$$

$$(23) \qquad P\{(\bar{b}, \bar{b}') = (b, b')\} = \frac{1}{2^n} \qquad \text{if} \quad b \neq 0.$$

If $T_n = w[(\bar{b}_1, \cdots, \bar{b}_n)]$ then

$$(24) \qquad P\{T_n = d\} \leqq \binom{n}{d} \frac{1}{2^n}$$

because $P\{\bar{b}_1 = b_1, \cdots, \bar{b}_n = b_n\} = 1/2^n$, or 0, for any $(b_1, \cdots, b_n)$ with $w[(b_1, \cdots, b_n)] = d$. Let $A_\beta^c \subset B^{n-1}$ consist of those a such that

$$(25) \qquad P\{T_n = d | \bar{a} = a\} \geqq \beta P\{T_n = d\},$$

where $\bar{a} = (\bar{a}_1, \cdots, \bar{a}_{n-1})$, so that

$$(26) \qquad P\{T_n = d\} \geqq P\{T_n = d | \bar{a} \in A_\beta^c\} P\{\bar{a} \in A_\beta^c\} \geqq \beta P\{T_n = d\} P\{\bar{a} \in A_\beta^c\}$$

hence $P\{\bar{a} \in A_\beta^c\} \leqq 1/\beta$. Thus there exists a set $A_\beta \subset B^{n-1}$ such that $P\{\bar{a} \in A_\beta\} \geqq 1 - 1/\beta$ and such that

$$(27) \qquad P\{T_n = d | \bar{a} = a\} \leqq \beta P\{T_n = d\} \leqq \beta \binom{n}{d} \frac{1}{2^n}$$

for all $a \in A_\beta$. But this is just the conclusion of lemma 1 with $B_\beta(n, k, d)$ = the set of $\alpha' = (a_1, a_2, \cdots) \in \overline{B}$ such that $(a_1, \cdots, a_{n-1}) \in A_\beta$; because, if we take such an α' and generate the n-SPC code of rate k/n from it, then

$$P\{T_n = d | \bar{a} = a\} = \frac{N_d}{2^k} \tag{28}$$

so that lemma 1 is proved.

LEMMA 2. *For any $\beta > 1$ there is a set $B_\beta \subset \overline{B}$ with $Q(B_\beta) \geqq 1 - (1/\beta)$ such that for every $\alpha = (a_1, a_2, \cdots) \in B_\beta$ the following is true: For any n, k satisfying $1 \leqq k \leqq n$, $n \geqq 2$, the n-SPC code of rate k/n which is generated from $a_1, a_2, \cdots, a_{n-1}$ satisfies*

$$N_d \leqq \beta n^4 2^{k-n} \binom{n}{d} \qquad \text{for all } d,\ 1 \leqq d \leqq n. \tag{29}$$

PROOF. Using lemma 1, let

$$B_\beta = \bigcap_{n=2}^{\infty} \bigcap_{k=1}^{n} \bigcap_{d=1}^{n} B_{\beta n^4}(n, k, d) \tag{30}$$

so that

$$\begin{aligned} Q(B_\beta) &\geqq 1 - \sum_n \sum_k \sum_d Q[B^c_{\beta n^4}(n, k, d)] \geqq 1 - \sum_n \sum_k \sum_d \frac{1}{\beta n^4} \\ &= 1 - \sum_{n=2}^{\infty} \frac{1}{\beta n^2} \geqq 1 - \frac{1}{\beta} \end{aligned} \tag{31}$$

and lemma 2 is proved.

Given an n-code C_n let M_d = the number of $b \in B^n$ such that $w(b) = d$ and such that there exists a $c \in C_n$ with $\overline{bc} \leqq d$. Recalling the decoder definition (4) and assuming $d \leqq np'_n$, for the moment, we see that M_d equals the number of $b \in B^n$, of weight d, which will not be decoded as 0. Let $V_n(d)$ = the volume of a sphere of radius d in B^n, that is,

$$V_n(d) = \sum_{k=0}^{d} \binom{n}{k}. \tag{32}$$

LEMMA 3. *Let C_n be an n-code such that for all d, $N_d \leqq A \binom{n}{d}$. Then for all d, $M_d \leqq A \binom{n}{d} V_n(d)$.*

PROOF. Let b_{ij} be an indexing of all the elements of B^n such that

$$w(b_{ij}) = i, \qquad i = 0, 1, \cdots, n; j = 1, 2, \cdots, \binom{n}{i}. \tag{33}$$

Let $S(x) = 1$ for $0 \leqq x \leqq d$, $S(x) = 0$ otherwise, where d is fixed. Let

$$e_{ij} = \sum_{k=1}^{\binom{n}{d}} S[w(b_{dk} + b_{ij})] \tag{34}$$

and notice that e_{ij} is independent of j. This is intuitively clear from the geometric interpretation and may be proved by observing that $w(b_{dk} + b_{ij})$ is invariant under a simultaneous permutation of the coordinates of both b_{dk} and b_{ij}. In the summations the range of i is $0, 1, \cdots, n$; the range of j is $1, 2, \cdots, \binom{n}{i}$; and the range of k is $1, 2, \cdots, \binom{n}{d}$.

Let $\delta_{ij} = 1$ if $b_{ij} \in C_n$ and 0 otherwise, so that $N_i = \sum_j \delta_{ij} \leqq A \binom{n}{i}$. Now if there is a $c \in C_n$ such that $\overline{cb_{dk}} \leqq d$, then

$$\sum_i \sum_j S[w(b_{dk} + b_{ij})]\delta_{ij} \tag{35}$$

will be $\geqq 1$ so that $M_d \leqq$ the summation over k of (35). Interchanging the order of summation we have $M_d \leqq \sum_i \sum_j e_{ij}\delta_{ij}$, but e_{ij} is independent of j, so that

$$\begin{aligned} M_d &\leqq \sum_i e_{ij} \sum_j \delta_{ij} \leqq A \sum_i \binom{n}{i} e_{ij} = A \sum_i \sum_j e_{ij} \\ &= A \sum_k \sum_i \sum_j S[w(b_{dk} + b_{ij})]. \end{aligned} \tag{36}$$

But $(b_{dk} + b_{ij})$ ranges over B^n precisely once as i, j cover their range so that

$$\sum_i \sum_j S[w(b_{dk} + b_{ij})] = \sum_i \sum_j S[w(b_{ij})] = V_n(d) \tag{37}$$

and lemma 3 is proved.

3. The proof of theorem 4

Now theorem 4 follows quite directly from lemmas 2 and 3 but before proceeding to the proof we collect together some needed analytical details in

LEMMA 4. (a) *If* $P\{S_n = d\} = \binom{n}{d} p^d q^{n-d}$ *and* $np \leqq d \leqq n$ *then*

$$\frac{1}{n} \log P\{S_n \geqq d\} \leqq H\left(\frac{d}{n}\right) + \frac{d}{n} \log p + \left(1 - \frac{d}{n}\right) \log q. \tag{38}$$

(b) *If* $0 \leqq d \leqq n/2$ *then*

$$\frac{1}{n} \log V_n(d) \leqq H\left(\frac{d}{n}\right). \tag{39}$$

(c) *If* $1 \leqq d \leqq n/2$ *then*

$$V_n(d) \geqq \left(\frac{n+1}{d} - 1\right) V_n(d-1). \tag{40}$$

PROOF. (a) If $t \geqq 0$ then

$$P\{S_n - d \geqq 0\} \leqq Ee^{t(S_n - d)} = e^{-td} Ee^{tS_n} = f^n(t) \tag{41}$$

where $f(t) = \exp[-(d/n)t](pe^t + q)$. Setting $df/dt = 0$, the solution t_0 is given by

$$e^{t_0} = \frac{d/n}{1 - d/n} \frac{q}{p} \geqq 1 \tag{42}$$

and part (a) follows by evaluating $f(t_0)$.

(b) If $t \geqq 0$ then

$$\sum_{k=0}^{d} \binom{n}{k} \leqq \sum_{k=0}^{n} \binom{n}{k} e^{t(d-k)} = g^n(t) \tag{43}$$

where $g(t) = \exp [t(d/n)][1 + \exp (-t)]$. Setting $dg/dt = 0$, the solution t_0 is given by

$$e^{t_0} = \frac{n}{d} - 1 \geqq 1 \tag{44}$$

and part (b) follows by evaluating $g(t_0)$.

(c) From Feller [4], p. 140, (3.6) with $p = 1/2$, it follows that

$$V_n(d-1) \leqq \binom{n}{d-1} \frac{n-d+2}{n-2d+3} \tag{45}$$

so that

$$\begin{aligned} \frac{V_n(d)}{V_n(d-1)} &= 1 + \frac{\binom{n}{d}}{V_n(d-1)} \geqq 1 + \frac{\binom{n}{d}}{\binom{n}{d-1}} \frac{n-2d+3}{n-d+2} \\ &= 1 + \frac{n-d+1}{d} \frac{n-2d+3}{n-d+2} = 1 + \frac{n-d+1}{d}\left(1 - \frac{d-1}{n-d+2}\right) \\ &\geqq 1 + \frac{n-d+1}{d}\left(1 - \frac{d}{n-d+1}\right) = \frac{n+1}{d} - 1 \end{aligned} \tag{46}$$

and lemma 4 is proved.

Proceeding to the proof of theorem 4 we wish to exhibit a B_β, and naturally we use the B_β whose existence is guaranteed by lemma 2. Let $\alpha = (a_1, a_2, \cdots) \in B_\beta$, $n \geqq 2$, $1 \leqq nR_n = n[1 - H(p'_n)] = k \leqq n$, and let C_n be the n-SPC code of rate k/n which is generated from $a_1, \cdots, a_{n-1}$. Let p, p'_n satisfy (11) so that $p < p'_n < \sqrt{p}/(\sqrt{p} + \sqrt{q})$, and let r be the unique integer such that

$$p'_n - \frac{1}{n} < \frac{r}{n} \leqq p'_n < \frac{\sqrt{p}}{\sqrt{p} + \sqrt{q}} < \frac{1}{2}. \tag{47}$$

Label the codeword 0 by c_1 and let D_1 be given by (4) so that

$$P\{D_1^c|0\} \leqq \sum_{d=1}^{r} M_d p^d q^{n-d} + P\{S_n \geqq r + 1\} \tag{48}$$

and from lemmas 2 and 3

$$P\{D_1^c|0\} \leqq \beta n^4 2^{-H(p'_n)n} \sum_{d=1}^{r} \binom{n}{d} p^d q^{n-d} V_n(d) + P\{S_n \geqq r + 1\}. \tag{49}$$

Now, $N_d(c_m)$ is independent of m, and lemma 3 can be applied to c_m, as well as to 0, so that the same bound holds for $P\{D_m^c|c_m\}$ and hence for $e_p(C_n)$. If $1 \leqq d \leqq r$ then from lemma 4(c) and (47).

$$(50)\quad \frac{\binom{n}{d} p^d q^{n-d} V_n(d)}{\binom{n}{d-1} p^{d-1} q^{n-d+1} V_n(d-1)}$$
$$\geqq \left(\frac{n+1}{d} - 1\right)\frac{p}{q}\left(\frac{n+1}{d} - 1\right) \geqq \left(\frac{n}{r} - 1\right)^2 \frac{p}{q} \geqq 1$$

so that the terms in the summation are increasing and since there are at most $n/2$ of them we get

$$(51)\qquad e_p(C_n) \leqq \frac{\beta}{2}\, n^5 2^{-H(p_n')n} \binom{n}{r} p^r q^{n-r} V_n(r) + P\{S_n \geqq r + 1\}.$$

But $2^{-H(p_n')n} V_n(r) \leqq 1$ by lemma 4(b) and (47) so that using $P\{S_n \geqq r\} \leqq P\{S_n \geqq np_n' - 1\}$ we get

$$(52)\qquad e_p(C_n) \leqq \frac{\beta}{2} n^5 P\{S_n \geqq np_n' - 1\}.$$

Now lemma 4(a) and $H[p_n' - (1/n)] \leqq H(p_n')$ imply that

$$(53)\qquad \frac{1}{n} \log e_p(C_n)$$
$$\leqq \frac{1}{n} \log \frac{\beta}{2} n^5 + H\left(p_n' - \frac{1}{n}\right) + \left(p_n' - \frac{1}{n}\right) \log p + \left(1 - p_n' + \frac{1}{n}\right) \log q$$
$$\leqq \frac{1}{n} \log \frac{\beta}{2} n^5 \frac{q}{p} + H(p_n') + p_n' \log p + (1 - p_n') \log q$$

and theorem 4 is proved.

REFERENCES

[1] P. Elias, "Coding for noisy channels," *IRE Convention Record*, Part 4 (1955), pp. 37–44.

[2] ———, "Coding for two noisy channels," *Proc. London Symp. Information Theory*, London, Butterworth Scientific Publications, 1955.

[3] A. Feinstein, *Foundations of Information Theory*, New York, McGraw-Hill, 1958.

[4] W. Feller, *An Introduction to Probability Theory and its Applications*, Vol. 1 (2nd ed.), New York, Wiley, 1957.