An Upper Bound on the Zero-Error List-Coding Capacity

Erdal Arikan

***Abstract*—We present an upper bound on the zero-error list-coding capacity of discrete memoryless channels. Using this bound, we show that the list-3 capacity of the 4/3 channel is at most 0.3512 b, improving the best previous bound. The relation of the bound to earlier similar bounds, in particular, to Körner's graph-entropy bound, is discussed.**

***Index Terms*—Zero-error capacity, list-coding, perfect-hashing, graph-entropy, Shannon capacity of graphs.**

I. Introduction

In ordinary point to point communications, the communication system delivers to the destination a single estimate of the transmitted message. Such a system is said to be a zero-error system if the estimate is always correct. Zero-error systems of this type were first studied by Shannon [1]. Elias [2] considered a more general type of system in which L estimates (L fixed) of the transmitted message are delivered to the destination and an error is said to occur if and only if all L estimates are wrong. The major problem of information-theoretic interest about such systems is to determine the zero-error list-L capacity C_L, i.e., the highest possible rate of communication under the zero error list-L condition. Unfortunately, no formula or algorithm is known for computing C_L. The aim of this correspondence is to give an upper bound on C_L.

We consider a system consisting of a finite discrete memoryless channel K with input alphabet I, output alphabet J, and transition probability matrix $[P(j|i)]$, where $P(j|i)$ is the probability that output letter j is received when input letter i is transmitted. We write $P_N(y|x)$ to denote the probability that $y \in J^N$ is received when $x \in I^N$ is transmitted; since the channel is memoryless, $P_N(y|x) = \prod_{n=1}^{N} P(y_n|x_n)$.

A block code $\mathscr{C}$ is employed in the system, mapping M messages into codewords $x(1),\cdots,x(M)$, with each codeword a sequence of length N from I. When a codeword is transmitted through K, the receiver observes the channel output y, and generates the list $\mathscr{L}(y) = \{m : P_N[y|x(m)] > 0\}$ of all messages that may have been transmitted. $\mathscr{C}$ is called a list-L code if, for each $y, \mathscr{L}(y)$ contains at most L messages. Thus, for a list-L code, the receiver can identify the transmitted message as one of at most L alternatives.

In general, the codewords of a list-L code do not have to be distinct. However, in a list-L code at most $L - 1$ codewords can be identical to any given codeword. So, if we discard repeated codewords from a list-L code, the size of the code is reduced at most by a factor of $1/L$. Since we shall be interested in asymptotic code rates for fixed L, there is no loss of generality in assuming, as we shall do henceforth, that all codewords in the codes under consideration are distinct. (This allows identification of codewords with messages and simplifies the notation considerably.)

Manuscript received April 12, 1993; revised November 16, 1993. This work was supported by the TÜBITAK under project TBAG 1053. This work was presented in part at the DIMACS/IEEE Workshop on Quantization and Coding, Rutgers—The State University, New Brunswick, NJ, and at the 1993 IEEE International Symposium on Information Theory, San Antonio, TX, January 1993.

The author is with the Department of Electrical Engineering, Bilkent University, Ankara 06533, Turkey.

IEEE Log Number 9403848.

The list-L capacity of K is defined by

$$C_L = \limsup_{N\to\infty} \frac{1}{N} \log M(N, L)$$

where $M(N, L)$ is the maximum possible size for a list-L code of length N.[1]

The upper bound on C_L given in this correspondence is an extension of earlier bounds by Shannon [1], Elias [2], Fredman and Komlós [3], Körner [4], Körner and Marton [5], [6]. These bounds have in common the use of the information-theoretic mutual information function.

To obtain the basic mutual-information bound on C_L, consider the above system again. Let $\mathscr{C}$ be a list-L code. Let $R = (1/N)\log M$ denote the rate of $\mathscr{C}$. Suppose a codeword X is chosen equiprobably from $\mathscr{C}$ and transmitted through K. Let Y denote the resulting channel output. Then, $NR = H(X) = H(X|Y) + I(X;Y) \le \log L + I(X;Y)$, where the equalities follow from the definitions of entropy and mutual information functions (see, e.g., [8] for the definitions), and the inequality follows by noting that there are at most L possibilities for X when Y is given. We may upper bound $I(X;Y)$ by NC where C is the ordinary Shannon capacity [8, p. 74] of K. Then, considering a sequence of list-L codes with increasing block lengths and with rates approaching C_L, we obtain $C_L \le C$.

This bound may be tightened by observing that C_L depends on the transition probabilities of K only through the channel adjacency function ϕ_K, defined as follows. For any $n \ge 1$ and $S \subset I^n$,

$$\phi_K(S) = \begin{cases} 1 & \text{if there exists } y \in J^n \text{ s.t. } P_n(y|x) > 0 \text{ for all } x \in S; \\ 0 & \text{otherwise.} \end{cases}$$

Thus, $\phi_K(S) = 1$ if and only if the sequences in S are adjacent in the sense that there is a common channel output sequence reachable from all of them. (Note that, since K is memoryless, ϕ_K is determined by its values on subsets of I.)

It is easy to see that $\mathscr{C}$ is a list-L code for K if and only if $\phi_K(S) = 0$ for each $S \subset \mathscr{C}$ with more than L elements. Thus, if K' is any other channel with the same input alphabet as K and $\phi_{K'} \le \phi_K$, then $C_L(K) \le C_L(K')$. This observation leads to the Shannon–Elias bound [1], [2]:

$$C_L(K) \le \min_{K' : \phi_{K'} \le \phi_K} C(K'). \tag{1}$$

The bound (1) turns out to be rather weak in many examples, apparently because the channel output Y (whichever admissible K' is considered) carries more than enough information necessary to identify the transmitted X as one of L possible alternatives. That list-L codes fail to achieve rates as high as C (unlike codes designed for an average probability of error criterion) may be attributed to the rigid combinatorial constraints that they must satisfy.

A more general framework for obtaining bounds on C_L, which allows exploitation of the combinatorial constraints on the structure of list-L codes, is to choose K' from the class of multiinput channels with side information, as we shall do in the next section and as previously done (in a different notation) in the papers

[1] It is not known if the lim sup can be replaced by lim for any $L \ge 2$. For $L = 1$, this is possible [1].

[3]–[6]. In Section III we show that for the example of the 4/3 channel the bound developed in Section II improves earlier bounds on its list-3 capacity. In general, by a b/l channel we mean a channel K with a b-letter input alphabet I such that $\phi_K(S) = 1$ if and only if $S \subset I$ has not more than l elements. Application of the same bound to arbitrary b/l channels is considered in [10].

Finally, we would like to note that zero-error list-coding is closely related to perfect-hashing, which is a method of information storage and retrieval (cf. [7] for a general discussion of hashing). Körner and Marton [5] give the following formal definition of perfect hashing. Call a set of sequences of length t over a b-letter alphabet *k-separated* if for every k tuple of sequences there exists a coordinate in which they all differ. For fixed t, b, k, let $N(t, b, k)$ denote the largest possible size for such a set of sequences. A main problem of interest in perfect hashing is to determine the numbers

$$C_{b,k} = \limsup_{t\to\infty} \frac{1}{t} \log N(t, b, k).$$

It can be seen that $C_{b,k}$ equals the list-$(k-1)$ capacity C_{k-1} of a $b/(k-1)$ channel. Thus, the bound developed in Section II readily yields upper bounds on $C_{b,k}$, and in some distances improves earlier such bounds, as demonstrated in Section III for $(b, k) = (4, 4)$ and in [10] for several other (b, k).

II. The New Bound

Throughout this section, let K be the channel specified in Section I. To obtain a bound on $C_L(K)$, we consider an alternative communication system with a discrete memoryless channel K' that has input alphabet I^t, output alphabet J', and transition probabilities $[P(j|i, h)]$, $j \in J'$, $i = (i_1, \cdots, i_m) \in I^m$, $h = (h_1, \cdots, h_k) \in I^k$, $m + k = t$. We assume that the h input of the channel is provided to the receiver in the system as side-information, i.e., when (i, h) is transmitted, the receiver observes h (in addition to the channel output j). The parameters m and k are arbitrary integers satisfying $m \geq 1$ and $k \geq 0$, respectively. Let $\mathscr{K}_{m,k}$ denote the class of all such channels for fixed m, k.

A block code $\mathscr{C}'$ of length N for a channel $K' \in \mathscr{K}_{m,k}$ is any subset of I^{Nt}, the set of t tuples over I^N. We write the codewords of such a code in the form $(x, z) = (x_1, \cdots, x_m, z_1, \cdots, z_k)$, where $x_r, z_s \in I^N$, $r = 1, \cdots, m$, $s = 1, \cdots, k$. The sequence x_r is transmitted via the rth i input, and z_s via the sth h input of K'. When a codeword (x, z) is sent, the receiver observes the channel output y and the side-information z, and produces the list $\mathscr{L}(y, z) = \{(x', z) \in \mathscr{C}': P_N(y|x', z) > 0\}$ of all possible codewords that may have been transmitted. $\mathscr{C}'$ is called a list-L' code for K' if $\mathscr{L}(y, z)$ contains not more than L' elements for every possible y and z.

We introduce some notation before proceeding. Let T be a set of m tuples over I^N. Let z be a k tuple over I^N. We use the notation $\phi_{K'}(T|z)$ as a shorthand for $\phi_{K'}(S)$ where $S = T \times \{z\} = \{(x, z): x \in T\}$. We write $[T]$ to denote the set of all words in I^N that appear as coordinates of m tuples in T. More precisely, if the elements of T are denoted by $x_u = (x_{u1}, \cdots, x_{um})$, $x_{ur} \in I^N$, $u = 1, \cdots, |T|$, $r = 1, \cdots, m$, then $[T]$ is the set of all such x_{ur}. We write $[z]$ to denote $\{z_1, \cdots, z_k\}$, the set of coordinates of z. For any finite set S, $|S|$ denotes the cardinality of S.

For any set $U \subset I^{Nm}$ and any $z \in I^{Nk}$, we define $\mathscr{K}_{m,k}(U, z)$ as the set of all $K' \in \mathscr{K}_{m,k}$ such that, for any $T \subset U$ with $|T| \geq 2$, $\phi_{K'}(T|z) \leq \phi_K([T] \cup [z])$. Note that $\mathscr{K}_{m,k}(U, z)$ is nonempty, always containing the trivial channel K' whose output identically equals its input.

Lemma 1: Let $\mathscr{C}$ be a list-L code for $K, \mathscr{C}'_m$ any subset of $\mathscr{C}^m$, and z any point in $\mathscr{C}^k$. Then, $\mathscr{C}' = \mathscr{C}'_m \times \{z\} = \{(x, z): x \in \mathscr{C}'_m\}$ is a list-L^m code for every $K' \in \mathscr{K}_{m,k}(\mathscr{C}'_m, z)$.

Proof: $\mathscr{C}'_m \times \{z\}$ is a list-L^m code for K' if (and only if) $\phi_{K'}(T|z) = 0$ for every $T \subset \mathscr{C}'_m$ with $|T| \geq L^m + 1$. Suppose, for a proof by contradiction, that there exists $T \subset \mathscr{C}'_m$ such that $|T| \geq L^m + 1$ and $\phi_{K'}(T|z) = 1$. Then, $\phi_K(S) = 1$ for $S = [T] \cup [z]$, since $K' \in \mathscr{K}_{m,k}(\mathscr{C}'_m, z)$. But S is a subset of $\mathscr{C}$, a list-L code for K; so, $\phi_K(S) = 1$ implies $|S| \leq L$. Also, $|T| \leq |S|^m$, since T is a set of m tuples over S. Thus, $|T| \leq L^m$, a contradiction, and the proof is complete. □

Let $\mathscr{C}, \mathscr{C}'_m, z, K'$ be as in the hypothesis of the lemma. Let X denote a random variable from the equiprobable distribution on $\mathscr{C}'_m$, and Y the output of K' when (X, z) is transmitted. That is, suppose that $P_X(x) = 1/|\mathscr{C}'_m|$ for $x \in \mathscr{C}'_m$, and $P_{Y|X}(y|x) = P_N(y|x, z)$, where P_N is the transition probability for K'. Then, we have

$$\begin{aligned} \log|\mathscr{C}'_m| &= H(X) = H(X|z) = H(X|Yz) + I(X; Y|z) \\ &\leq \log L^m + I(X; Y|z) \qquad (2) \end{aligned}$$

where the second equality follows by the independence of X and z (a constant) and the inequality by Lemma 1.

Inequality (2) can be used to obtain upper bounds on the size M of $\mathscr{C}$ by choosing particular forms for $\mathscr{C}'_m$. For example, setting $\mathscr{C}'_m = \mathscr{C}^m$ yields $H(X) = m \log M$. Another possibility, which has yielded better results in applications, is to set $\mathscr{C}'_m = \overline{\mathscr{C}^m} \triangleq \{(x_1, \cdots, x_m) \in \mathscr{C}^m : x_1, \cdots, x_m \text{ are distinct}\}$. Then, $H(X) = \log M^{\underline{m}}$ where $M^{\underline{m}} = \prod_{i=0}^{m-1}(M - i)$. The rest of the paper will be based on this latter choice with the further restriction that $z \in \overline{\mathscr{C}^k}$. The result thus far can be summarized as follows.

Proposition 1: The size M of any list-L code $\mathscr{C}$ for a discrete memoryless channel K satisfies, for any $k \geq 0$, $m \geq 1$

$$\log M^{\underline{m}} \leq m \log L + \min_{z \in \overline{\mathscr{C}^k}} \min_{K' \in \mathscr{K}_{m,k}(\overline{\mathscr{C}^m}, z)} I(X; Y|z) \qquad (3)$$

where X is a random variable from the uniform distribution on $\overline{\mathscr{C}^m}$ and $P_{Y|X}(y|x) = P_N(y|x, z)$ with P_N the transition probability for K'.

Inequality (3) represents the general form of the bound proposed in this correspondence. An equivalent bound is implicit in Körner's work [4]. The bound (3) is not amenable to computation due to its involuted structure. In actual calculations, one finds it necessary to make the range of minimization over K' independent of $\mathscr{C}$. Such a simplified form of the bound is

$$\log M^{\underline{m}} \leq m \log L + \min_{K' \in \mathscr{K}^*_{m,k}} \min_{z \in \overline{\mathscr{C}^k}} I(X; Y|z) \qquad (4)$$

where $\mathscr{K}^*_{m,k}$ is the intersection of $\mathscr{K}_{m,k}(\overline{\mathscr{C}^m}, z)$ over all list-L codes $\mathscr{C}$ for K and all $z \in \overline{\mathscr{C}^k}$.

Another form of the bound is obtained by observing that for fixed K' the minimum over z in (4) can be replaced by an average. This gives

$$\log M^{\underline{m}} \leq m \log L + \min_{K' \in \mathscr{K}^*_{m,k}} I(X; Y|Z) \qquad (5)$$

where Z is a random variable from an arbitrary probability distribution on $\overline{\mathscr{C}^k}$. By choosing the distribution of Z suitably, the bound (5) may be computed relatively easily in specific instances. For example, in [6], the bound (5) was applied to L-uniform channels with Z from the uniform distribution on $\overline{\mathscr{C}^k}$

(and with $m = 1$, $0 \leq k \leq L - 2$).[2] (A channel K is called L-uniform if $\phi_K(S) = 0$ implies $|S| \geq L$.)

Clearly, the bound (5) with a uniform Z may be significantly weaker than (4). Indeed, the main contribution of the present work is the demonstration of this fact. For the 4/3 channel considered in the next section, starting from (4), we derive a bound on its list-3 capacity that improves all previous bounds, in particular, the bound (5) with Z uniform.

To end this section, let us note that the Shannon–Elias bound (1) is a special case of (4) with $m = 1$, $k = 0$. Let us also note that, due to the memoryless property of the channels involved, the term $I(X;Y|z)$ in the above bounds can be upper bounded by $\sum_{n=1}^{N} I(X^{(n)}; Y_n | z^{(n)})$, where $X^{(n)} = (X_{1n}, \cdots, X_{mn})$ and $z^{(n)} = (z_{1n}, \cdots, z_{kn})$ are the nth coordinates of the vectors X and z.This yields a single-letter form that may be easier to compute.

III. The 4 / 3 Channel

In this section, we consider a 4/3 channel K, and apply the bound (4) to show that its list-3 capacity satisfies $C_3 \leq 0.3512$ b. This improves the best previous bound $C_3 \leq 3/8$ b, which was obtained by applying (5) with $m = 1$, $k = 0$, and Z uniform [3], [6]. This demonstrates that choosing the random variable Z in (5) from a nonuniform distribution [in particular, concentrating it on a single point as in (4)] may yield better bounds, as might be expected. In the following, all rates will be in bits and all logarithms to base two.

The combinatorial property characterizing list-3 codes for a 4/3 channel is that for any four distinct codewords x_1, x_2, x_3, x_4, there exists a coordinate n such that $x_{1n}, x_{2n}, x_{3n}, x_{4n}$ are distinct. To obtain a bound on C_3 we employ the method of Section II with a channel K' from $\mathscr{K}_{1,2}$. Thus, the inputs of K' are of the form $(i_1, h_1, h_2) \in I^3$, where I denotes the input alphabet of K, and the inputs h_1, h_2 are provided as side-information at the channel output. We specify the output alphabet of K' as $J' = I \cup \{e\}$ where e is a symbol not contained in I, and its transition probabilities as follows:

$$P(j|i_1, h_1, h_2) = \begin{cases} \delta_{je} & \text{if } h_1 = h_2 \\ 1/2 & \text{if } h_1 \neq h_2,\ i_1 \in \{h_1, h_2\},\ j \in I \setminus \{h_1, h_2\} \\ 1 & \text{if } \{i_1, h_1, h_2, j\} = I. \end{cases}$$

Lemma 2: K' specified above belongs to $\mathscr{K}^*_{1,2}$.

Proof: Let $\mathscr{C}$ be an arbitrary list-3 code for K, and $z = (z_1, z_2)$ an arbitrary point in $\mathscr{C}^2$. We must show that, for every $T \subset \mathscr{C}$ with $|T| \geq 2$, $\phi_{K'}(T|z) \leq \phi_K(S)$, where $S = [T] \cup [z]$. We only need consider T for which $\phi_K(S) = 0$. Any such T contains at least two codewords x_1, x_2 such that x_1, x_2, z_1, z_2 are distinct. So, by the defining property of list-3 codes, there exists a coordinate n such that $x_{1n}, x_{2n}, z_{1n}, z_{2n}$ are distinct. Hence, by the way K' has been specified, $\phi_{K'}[(x_{1n}, x_{2n})|(z_{1n}, z_{2n})] = 0$. This implies $\phi_{K'}[(x_1, x_2)|z] = 0$, which in turn implies $\phi_{K'}(T|z) = 0$ (since (x_1, x_2) is a subset of T), completing the proof. □

Henceforth fix $\mathscr{C}$ as a list-3 code for K and $z = (z_1, z_2)$ as a point in $\mathscr{C}^2$. Let N be the length, M the size, R the rate of $\mathscr{C}$. Let X be a random variable equiprobable on $\mathscr{C}$, and Y the random variable observed at the output of K' when (X, z) is transmitted. Thus, $P_{XY}(x, y) = (1/M)P_N(y|x, z)$ for $x \in \mathscr{C}$. By (4), the rate of $\mathscr{C}$ satisfies

$$NR \leq \log 3 + I(X;Y|z). \quad (6)$$

In the rest of this section we develop an upper bound on $I(X;Y|z)$.

For any two sequences u_1, u_2 of equal length, let $d(u_1, u_2)$ denote the number of coordinates n such that $u_{1n} \neq u_{2n}$ (the Hamming distance). Likewise, for any three sequences u_1, u_2, u_3 of equal length, let $d(u_1, u_2, u_3)$ denote the number of coordinates n such that u_{1n}, u_{2n}, u_{3n} are distinct.

Lemma 3: $I(X; Y|z) \leq \sum_{x \in \mathscr{C}} M^{-1} d(x, z_1, z_2) = \sum_{n=1}^{N} [1 - Q_n(z_{1n}) - Q_n(z_{2n})] d(z_{1n}, z_{2n})$ where $Q_n(\cdot)$ is the empirical distribution of the nth coordinate of the codewords in $\mathscr{C}$, i.e.,

$$Q_n(i) = \frac{\text{number of } x \text{ in } \mathscr{C} \text{ such that } x_n = i}{M}.$$

Proof: For coordinates n where $z_{1n} = z_{2n}$, we have $Y_n = e$. For $z_{1n} \neq z_{2n}$, Y_n can take one of at most two values. So, the number of possible values of Y is at most $2^{d(z_1, z_2)}$. This gives $H(Y|z) \leq d(z_1, z_2)$. On the other hand, for each x, y we have either $P_N(y|x, z) = 0$ or $P_N(y|x, z_1, z_2) = 2^{-[d(z_1, z_2) - d(x, z_1, z_2)]}$. Thus, $H(Y|X, z) = d(z_1, z_2) - \sum_{x \in C} M^{-1} d(x, z_1, z_2)$. Since $I(X; Y|z) = H(Y|z) - H(Y|X, z)$, the inequality follows. The proof is completed by noting that $d(x, z_1, z_2) = \sum_{n=1}^{N} d(x_n, z_{1n}, z_{2n})$ and

$$\sum_{x \in \mathscr{C}} M^{-1} d(x_n, z_{1n}, z_{2n}) = [1 - Q_n(z_{1n}) - Q_n(z_{2n})] d(z_{1n}, z_{2n}).$$

□

Lemma 3 and (6) give the following constraint on the rate and composition of $\mathscr{C}$:

$$NR \leq \log 3 + \sum_{n=1}^{N} [1 - Q_n(z_{1n}) - Q_n(z_{2n})] d(z_{1n}, z_{2n}). \quad (7)$$

To obtain a tight bound on C_3 using (7), we need to show that $\mathscr{C}$ can be chosen with rate close to C_3 and with $Q_n(i)$ not too small for any n, i.

Lemma 4: Given any $\epsilon > 0$, there exist list-3 codes (for the 4/3 channel) of arbitrarily large lengths, with rates $\geq C_3 - \epsilon$, and for which $Q_n(i) \geq 1 - 2^{-(C_3 - 2\epsilon)}$ for all n, i.

Proof: For any $\epsilon > 0$, there exists a finite integer N_ϵ such that every list-3 code with length $N \geq N_\epsilon$ has rate $\leq C_3 + \epsilon$. This follows from the definition of C_3. Fix $\epsilon > 0$, and consider a list-3 code $\mathscr{C}$ with rate $R \geq C_3 - \epsilon$ and length $N > 3N_\epsilon$. The existence of such a code for arbitrarily large N is also guaranteed by the definition of C_3.

If there exist n, i such that $Q_n(i) < 1 - 2^{-(C_3 - 2\epsilon)}$, consider the subcode $\mathscr{C}' = \{x \in \mathscr{C} : x_n \neq i\}$. $\mathscr{C}'$ is a list-3 code (any subcode of $\mathscr{C}$ is a list-3 code) with $M_1 = M[1 - Q_n(i)] > 2^{N(C_3 - \epsilon) - (C_3 - 2\epsilon)}$ codewords, where $M = 2^{NR}$ is the number of codewords in $\mathscr{C}$. Let $\mathscr{C}_1$ be the code obtained by deleting the nth coordinate of each codeword in $\mathscr{C}'$. $\mathscr{C}_1$ has length $N - 1$, and it is easy to see that it is also a list-3 code for the 4/3 channel. Thus, $\mathscr{C}_1$ has M_1 codewords and rate $R_1 = [1/(N - 1)] \log M_1 > C_3 - \epsilon + \epsilon/(N - 1)$. Since $R_1 > C_3 - \epsilon$, we may iterate the above procedure with $\mathscr{C}_1$ in place of $\mathscr{C}$. At the end of the kth round, we shall have a code $\mathscr{C}_k$ with length $N_k = N - k$, number of codewords $M_k > M2^{-k(C_3 - 2\epsilon)}$, and rate $R_k > C_3 - \epsilon + k\epsilon/(N - k)$. If this process could continue for more than $2N/3$ rounds, at round $k = \lceil 2N/3 \rceil$ we would have a code with length $\lfloor N/3 \rfloor$ and rate $> C_3 + \epsilon$. But that would contradict the assumption that $N > 3N_\epsilon$. So, the process terminates at some step $k < 2N/3$, yielding a list-3 code with length $N_k = N - k > N/3$, rate $R_k > C_3 - \epsilon$, and for which $Q_n(i) \geq 1 - 2^{-(C_3 - 2\epsilon)}$ for all $n = 1, \cdots, N_k$ and all $i \in I$. Since $N/3$ can be arbitrarily large, this completes the proof. □

[2] The choice $m = 1$ here is not optimum. For example, for the 5/4 channel, $m = 2$ yields a better result.

Proposition 2: The list-3 capacity of the 4/3 channel satisfies $C_3 \le 0.3512$ b.

Proof: Let $\epsilon > 0$ be arbitrary and consider a list-3 code $\mathscr{C}$ with rate $R \ge C_3 - \epsilon$ and $Q_n(i) \ge 1 - 2^{-(C_3 - 2\epsilon)}$ for all n, i. By Lemma 4, such a code exists and its length N can be assumed arbitrarily large. Substituting the parameters for this code into (7), we obtain

$$N(C_3 - \epsilon) \le \log 3 + (2 \cdot 2^{-(C_3 - 2\epsilon)} - 1) d(z_1, z_2).$$

Let $d(\mathscr{C}) = \min\{d(z_1, z_2) : z_1, z_2 \in \mathscr{C},\ z_1 \ne z_2\}$ and let

$$\delta(R) = \limsup_{N \to \infty} \{d(\mathscr{C})/N : \mathscr{C} \text{ is a quaternary code with length } N \text{ and rate } \ge R\}.$$

Taking z_1, z_2 at distance $d(\mathscr{C})$, letting $\epsilon \to 0$, and $N \to \infty$, we get

$$C_3 \le (2^{1-C_3} - 1)\delta(C_3). \tag{8}$$

By the Plotkin bound [8, p. 545] (as modified for a quaternary alphabet), $\delta(R) \le (1 - R/2)(3/4)$. Substituting this into (8) yields $C_3 \le (2^{1-C_3} - 1)(1 - C_3/2)(3/4)$, from which we obtain

$$C_3 \le \sup\{\alpha : \alpha \le (2^{1-\alpha} - 1)(1 - \alpha/2)(3/4)\} < 0.351\,152\,268$$

□

Clearly, the above bound can be improved by using better estimates of $\delta(C_3)$, e.g., the Elias bound in its general form as discussed in [9, p. 410]. We note that a direct combinatorial proof of the inequality (7) is possible.[3] Finally, let us also note that the method used in this section has been generalized to arbitrary b/l channels in [10].

Acknowledgments

The author wishes to thank J. Körner for his many helpful comments and for pointing out an error in an earlier version of this work.

References

[1] C. E. Shannon, "The zero error capacity of a noisy channel," *IRE Trans. Inform. Theory*, vol. IT-2, no. 3, pp. 8–19, Sept. 1956.
[2] P. Elias, "Zero error capacity under list decoding," *IEEE Trans. Inform. Theory*, vol. IT-34, pp. 1070–1074, Sept. 1988.
[3] M. Fredman and J. Komlós, "On the size of separating systems and perfect hash functions," *SIAM J. Algebraic Discrete Methods*, vol. 5, no. 1, pp. 61–68, 1984.
[4] J. Körner, "Fredman-Komlós bounds and information theory," *SIAM J. Algebraic Discrete Methods*, vol. 7, no. 4, pp. 560–570, Oct. 1986.
[5] J. Körner and K. Marton, "New bounds for perfect hashing via information theory," *Euro. J. Combinatorics*, vol. 9, pp. 523–530, 1988.
[6] ——, "On the capacity of uniform hypergraphs," *IEEE Trans. Inform. Theory*, vol. IT-36, pp. 153–156, Jan. 1990.
[7] D. E. Knuth, *The Art of Computer Programming, Vol. 3.* Reading, MA: Addison-Wesley, 1973.
[8] R. G. Gallager, *Information Theory and Reliable Communication.* New York: Wiley, 1968.
[9] R. E. Blahut, *Principles and Practice of Information Theory.* Reading, MA: Addison-Wesley, 1987.
[10] E. Arikan, "An improved graph-entropy bound for perfect-hashing," in *Proc. IEEE Int. Symp. Inform. Theory*, p. 314, Trondheim, Norway, June 1994.

[3] Such a proof was communicated to the author by J. Körner.

Bounds on the Zero-Error Capacity of the Input-Constrained Bit-Shift Channel

Victor Yu. Krachkovsky

***Abstract*—New lower and upper bounds on a maximal achievable rate for runlength-limited codes, capable of correcting any combination of bit-shift errors (i.e., a zero-error capacity of the bit-shift channel), are presented. The lower bound is a generalization of the bound obtained by Shamai and Zehavi. It is shown that in certain cases, the upper and the lower bounds asymptotically coincide.**

***Index Terms*—Runlength-limited codes, error correction, zero-error capacity.**

I. Introduction

Let X be a finite alphabet, and let X^n be the set of all n-words $\boldsymbol{x} = (x_1, \cdots, x_n)$, $x_i \in X$. A *constrained system* is a subset of words from X^n that comply with some limitation L. One of the most notable types of limitations is a runlength limitation. Let l, m be a pair of integers, $m > l$. We say that a word $\boldsymbol{x} \in X^n$ over the binary alphabet $X = \{0, 1\}$ is an *(l, m)-runlength limited or $RLL_0(l, m)$-sequence* if the following conditions are satisfied.

1) Every two binary "1"'s in $\boldsymbol{x}$ are separated by at least l "0"'s.
2) Any $m + 1$ consecutive symbols in $\boldsymbol{x}$ contain at least one symbol "1."

If only the first condition is satisfied, we set $m = \infty$ and call $\boldsymbol{x}$ an *$RLL_0(l, \infty)$-sequence*. For the convenience of analysis, we also suppose that

3) $\boldsymbol{x}$ begins by at least l "0"'s.
4) the last symbol in $\boldsymbol{x}$ is "1."

The additional conditions 3) and 4) guarantee a "merging" property for $\boldsymbol{x}$ and do not play any role in asymptotics. The set of all words, satisfying 1)–4), presents a *runlength-limited constrained system*, denoted by $X_L^n \subseteq X^n$. Any subset of M sequences $A_n \triangleq \{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_M\} \subseteq X_L^n$ is called an *runlength-limited block code* of length n and rate $R_n \triangleq 1/n \cdot \log_2 M$. The maximal achievable rate of a runlength-limited block code is called the capacity of the constrained system X_L^n and is denoted by C. Shannon [9] showed that for a broad class of irreducible and deterministic constrained systems (this class also includes runlength-limited systems),

$$C = \log_2 \lambda$$

where λ is the largest positive eigenvalue of a system's characteristic equation.

Runlength-limited codes are used in high-quality digital systems such as optical and magnetic recordings. They could also be used for data transmission over certain narrow-band channels. For noisy channels, runlength-limited codes need to possess some error-correcting ability. In recent times, attention has been given to the problem of designing runlength-limited error-correcting codes for a symmetric memoryless channel (see, for example, [1], [7], [10]). For most applications, however, the

Manuscript received May 28, 1992; revised November 18, 1993.
IEEE Log Number 9403841.

The author is with the Department of Information Systems, St. Petersburg Academy of Aerospace Instrumentation, 190000, B.Morskaya, 67, St. Petersburg, Russia.

0018-9448/94$04.00