[11] L. R. Goke and G. J. Lipovski, "Banyan networks for partitioning multiprocessor syst.," in *Proc. 1st Ann. Symp. Comput. Architecture*, Dec. 1973, pp. 21–28.
[12] M. C. Pease, III, "The indirect binary *n*-cube microprocessor array," *IEEE Trans. Comput.*, vol. C-26, May 1977, pp. 458–473.
[13] H. J. Siegel, R. J. McMillen, and P. T. Mueller, Jr., "A survey of interconnection methods for reconfigurable parallel processing systems," *AFIPS Conf. Proc.*, vol. 48, 1979 NCC, pp. 387–400.
[14] C.-L. Wu and T.-Y. Feng, "The universality of the shuffle-exchange network," *IEEE Trans. Comput.*, vol. C-30, pp. 324–332, May 1981.
[15] B. J. Smith, "Architecture and applications of HEP multiprocessor computer," *Real-Time Signal Processing IV, SPIE*, Aug. 1981, pp. 241–248.
[16] D. Gajski, D. L. Kuck, D. Lawrie, and A. Sameh, "Cedar—A large scale multiprocessor," *Proc. 1983 Int. Conf. Parallel Processing*, Aug. 1983, pp. 524–529.
[17] A. S. Acampora, M. J. Karol, and M. G. Hluchyj, "Terabit lightwave networks: The Multihop approach," *AT&T Tech. J.*, p. 21, Nov./Dec. 1987.
[18] R. A. Thompson, "Architectures with improved signal-to-noise ratio in photonic systems with fiber-loop delay lines," *IEEE J. on Selected Areas in Commun.*, vol. 6, Aug. 1988, pp. 1096–1106.
[19] V. P. Heuring, H. F. Jordan, and J. P. Pratt, "A bit serial architecture for optical Computing," Center for Optoelectronic Computing Systems, Univ. of Colorado, Boulder, CO, 80309-0525, Technical Report 88-01a, Jan. 1988.
[20] S. V. Ramanan and H. F. Jordan, "Photonic architectures for performing perfect shuffle on a time-division multiplexed signal," Center for Optoelectronic Computing Systems, Univ. of Colorado, Boulder, CO, 80309-0525, Technical Report 89-03, Jan. 1989.
[21] F. T. Stone, J. E. Watson, D. T. Moser, and W. J. Minford, "Performance and yield of pilot-line quantities of Lithium Niobate switches," in *SPIE Conf. Proc. OE/Fibers '89*, Boston, MA, Sept. 1989.
[22] D. B. Sarrazin, "Fiber-optic delay line storage in digital optic computers based on directional couplers," Center for Optoelectronic Computing Systems, Univ. of Colorado, Boulder, CO, 80309-0525, Technical Report 89-04, Jan. 1989.

Capacity and Zero-Error Capacity of Ising Channels

TOBY BERGER, FELLOW, IEEE, AND
FLAVIO BONOMI, MEMBER, IEEE

***Abstract*—Simple binary symmetric one- and two-dimensional channel systems called Ising channels are investigated. The one-dimensional version is a certain binary finite-state channel with input intersymbol interference; we determine its zero-error capacity together with lower and upper bounds to its Shannon capacity. For the two-dimensional channel system we obtain a lower bound for the zero-error capacity and develop an efficient zero-error coding scheme.**

I. Introduction

In Section II we determine the zero-error capacity of a certain finite-state channel with intersymbol interference at the input which we call the one-dimensional zero-temperature Ising channel. We also bound its Shannon capacity and obtain insight into its optimal input probability distribution.

In Section III we discuss the two-dimensional Ising channel, which is related to the famous Ising model [1] of statistical physics and random field theory. It possesses a parameter T, usually called the temperature. We develop some efficient zero-error codes for the two-dimensional Ising channel with $T=0$.

Manuscript received September 11, 1986; revised May 15, 1989. This work was presented in part at the International Symposium on Information Theory, Brighton, England, June 1985.
T. Berger is with Cornell University, Phillips Hall, Ithaca, NY 14853.
F. Bonomi is with AT&T Bell Laboratories, Room 3M306, Holmdel, NJ 07733.
IEEE Log Number 8933106.

II. One-Dimensional Ising Channel

Let Z denote the integers, and define

$$D_1 = \{ i \in Z \colon i \text{ is odd} \}$$
$$D_2 = \{ i \in Z \colon i \text{ is even} \}.$$

Two families of random variables $\{x_i, i \in D_1\}$ and $\{y_i, i \in D_2\}$ are defined. We view $\{x_i, i \in D_1\}$ as the sequence of channel inputs, and $\{y_i, i \in D_2\}$ as the sequence of outputs (see Fig. 1). It is convenient here to modify the indexing of these two families by replacing $i \in D_1$ by $j=(i+1)/2$ and $i \in D_2$ by $j = i/2$; then both families are indexed by the signed integers. The link between the two families is the channel description, which is based

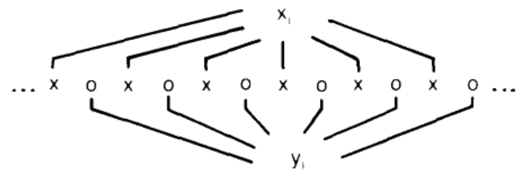

Fig. 1. One-dimensional Ising channel.

on the elementary transition probabilities $P(y_i|x_i, x_{i+1})$ defined by

$$P(1|1,1) = P(0|0,0) = 1-\epsilon$$
$$P(0|1,1) = \epsilon$$
$$P(1|0,0) = \epsilon$$
$$P(1|0,1) = P(0|0,1) = P(1|1,0) = P(0|1,0) = 1/2.$$

The parameter ϵ is related to the temperature parameter T of an associated Ising model. We discuss this relationship in some detail when we treat the two-dimensional Ising channel in Section III. In what follows we will discuss mainly the case $\epsilon = 0$, which corresponds to $T = 0$. Both the Shannon capacity and the zero-error capacity of the channel system based on this elementary channel element will be considered.

Our one-dimensional Ising channel is a particular case of a binary channel with memory, specifically intersymbol interference, at the input. The state s_n of the channel at position (or time instant) n is simply the input at n,

$$s_n = x_n.$$

Fig. 2 describes the particular channel with memory in question. As in [2] define

$$C = \lim_{N \to \infty} C_N$$

where

$$C_N = N^{-1} \max_{P_N} \min_{s_0} I_{P_N}(X^N; Y^N | s_0)$$

with P_N being the joint input distribution on sequences of length N, s_0 being the initial state, and

$$I_{P_N}(X^N; Y^N|s_0) = \sum_{x} \sum_{y} P_N(x) P_N(y|x, s_0) \log P_N(y|x, s_0) \Big/ \sum_{\tilde{x}} P_N(\tilde{x}) P_N(y|\tilde{x}, s_0).$$

0018-9448/90/0100-0173$01.00 ©1990 IEEE

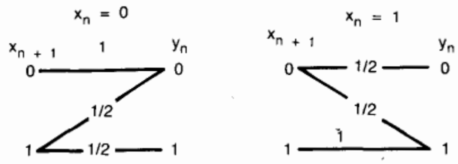

Fig. 2. Channel with memory.

Here, $y = (y_0, \cdots, y_{N-1})$ and $x = (x_1, \cdots, x_N)$ and

$$P_N(y|x, s_0) = \sum_{s_N} P_N(y, s_N|x, s_0)$$

where $P_N(y, s_N|x, s_0)$ can be calculated recursively from

$$P_N(y, s_N|x, s_0) = \sum_{s_{N-1}} P(y_{N-1}, s_N|x_N, s_{N-1}) P_{N-1}(\hat{y}, s_{N-1}|\hat{x}, s_0),$$

with $\hat{y} = (y_0, \cdots, y_{N-2})$ and $\hat{x} = (x_1, \cdots, x_{N-1})$.

Similarly, define

$$\overline{C} = \lim_{N \to \infty} \overline{C}_N = N^{-1} \max_{P_N} \max_{s_0} I_{P_N}(X^N; Y^N|s_0).$$

Again following [2], let

$$q_N(s_N|x, s_0) = \sum_{y} P_N(y, s_N|x, s_0)$$

and say the channel is *indecomposable* if for every $\epsilon > 0$ there exists an N_0 such that for $N \geq N_0$,

$$|q_N(s_N|x, s_0') - q_N(s_N|x, s_0)| < \epsilon$$

for all s_N, x, s_0', and s_0.

Theorem 1 [2, Theorem 4.6.4]: For an indecomposable finite state channel,

$$C = \overline{C}.$$

The simple finite state channel of Fig. 2 is indecomposable since $s_N = x_N$ and hence does not depend on s_0 once x_N is given. The following theorem provides asymptotically tight upper and lower bounds for the capacity of a general finite state channel.

Theorem 2 [2, Theorem 4.6.1]: For a finite state channel with A states,

$$\lim_{N \to \infty} C_N = \sup_N [C_N - \log A/N]$$

$$\lim_{N \to \infty} \overline{C}_N = \inf_N [\overline{C}_N - \log A/N].$$

Thus for every N

$$C_N - \log A/N \leq C \leq \overline{C} \leq \overline{C}_N + \log A/N.$$

We now specialize this general result to the particular channel we are considering and derive improved bounds for that case. In our case $A = 2$, and, based on symmetry considerations,

$$C_N = \overline{C}_N, \qquad \text{for every } N,$$

so

$$C = \overline{C} = C.$$

The sequence $\{C_N\}$ is not in general monotonically increasing, but we can prove the following.

Theorem 3: For any stationary binary channel with intersymbol interference at the input with $s_N = x_N$ and for any N

$$C_{2N} \geq C_N.$$

Proof: Let

$$P_{2N}(x) = P_N(x_1) P_N(x_2)$$

where

$$x = (x_1, \cdots, x_{2N}) \qquad x_1 = (x_1, \cdots, x_N)$$
$$x_2 = (x_{N+1}, \cdots, x_{2N}),$$

and P_N is the input distribution achieving C_N. Similarly, let

$$y_1 = (y_0, \cdots, y_{N-1}) \qquad y_2 = (y_N, \cdots, y_{2N-1}).$$

Let X_1, X_2, Y_1, Y_2 be the ensembles of sequences x_1, x_2, y_1, y_2, respectively. Since P_{2N} does not necessarily achieve C_{2N},

$$\begin{aligned}
2NC_{2N} &\geq \min_{s_0} I(X_1X_2; Y_1Y_2|s_0) \\
&= \min_{s_0} [I(X_1; Y_1Y_2|s_0) + I(X_2; Y_1Y_2|X_1, s_0)] \\
&\geq \min_{s_0} [I(X_1; Y_1|s_0) + I(X_2; Y_2|X_1, s_0)] \\
&\geq \min_{s_0} I(X_1; Y_1|s_0) + \min_{s_0} I(X_2; Y_2|(x_N, \cdots, x_1), s_0) \\
&= \min_{s_0} I(X_1; Y_1|s_0) + I(X_2; Y_2|x_N) \\
&\geq \min_{s_0} I(X_1; Y_1|s_0) + \min_{x_N} I(X_2; Y_2|x_N) \\
&= NC_N + NC_N,
\end{aligned}$$

because P_N is the marginal distribution of both X_1 and X_2. It follows that

$$C_{2N} \geq C_N, \qquad \text{for every } N. \qquad \square$$

An immediate corollary is

$$C_{j2^k} \geq C_j, \qquad \text{for every } j \geq 0 \text{ and } k \geq 0.$$

Theorems 1 and 3 together provide us with a specialized lower bound to $C = \overline{C} = C$.

Theorem 4: For a channel with intersymbol interference at the input whose state is the most recent input,

$$C \geq C_j, \qquad \text{for any } j.$$

The simple proof of this theorem is omitted. Progress toward an improved upper bound can be obtained as follows.

Theorem 5: For any channel with intersymbol interference at the input and with a finite number of states A

$$\overline{C} \leq \overline{C}_N + U_N, \qquad \text{for every } N$$

where

$$U_N = \sum_{i=N}^{i=\infty} I(Y_i^{2i-1}; x_i|s_0)/2i$$

and Y_i^{2i-1} is the ensemble of sequences $y_i^{2i-1} = (y_i, \cdots, y_{2i-1})$.

Again, the simple proof of this theorem is omitted. Although the new upper bound obtained in Theorem 5 is an improvement on the upper bound of Theorem 2, it does not lend itself to an easy evaluation. Accordingly, we used the upper bound from Theorem 2 in the following numerical evaluations for the case $T = 0$.

The Arimoto–Blahut algorithm [3], [4] was used to determine $C_N = \overline{C}_N$ for the zero-temperature one-dimensional Ising channel up to $N = 7$. Note that in this case we have $N+1$ input bits and N output bits, but x_0 is given the fixed value s_0. The results of the computation are shown in Table I.

TABLE I
C_N FOR $1 \le N \le 7$

N	C_N
1	0.3219
2	0.507
3	0.5094
4	0.5211
5	0.5225
6	0.5293
7	0.5301

Using the value computed for C_7, we obtain the following numerical bounds for C:

$$C_7 \le C \le \bar{C}_7 + 1/7$$

i.e.,

$$0.5301 \le C \le 0.6723.$$

We now provide further insight into the nature of the optimal input distribution by showing that a Markov chain input distribution achieves a higher mutual information through the one-dimensional Ising channel with $\epsilon = 0$ than does the product input distribution. We make the following assumption about the nature of the optimal input distribution.

Assumption: For every i and N

$$P(x_i, x_{i+1}, \cdots, x_{i+N} = v) = P(x_i, x_{i+1}, \cdots, x_{i+N} = v^c)$$

where $v \in \{0,1\}^{N+1}$, and v^c is the complement of v. In particular,

$$P(x_i = 0) = P(x_i = 1) = 1/2, \qquad \text{for every } i.$$

This assumption implies that

$$P(y_j = 0) = 1/2, \qquad \text{for every } j.$$

In what follows we want to compare the values of average mutual information $I(y_1, \cdots, y_n; x_1, \cdots, x_{n+1}) = I(y_n; x_{n+1})$ obtained in the two cases:

1) $P(x_1, \cdots, x_{n+1}) = P(x_1)P(x_2) \cdots P(x_{n+1})$,
2) $P(x_1, \cdots, x_{n+1}) = P(x_1)P(x_2|x_1) \cdots P(x_{n+1}|x_n)$.

Let

$$y_n = (y_1, \cdots, y_n) \qquad x_n = (x_1, \cdots, x_n)$$

and

$$\Delta I(y_n; x_{n+1}) = I^{(1)}(y_n; x_{n+1}) - I^{(2)}(y_n; x_{n+1})$$

where the superscript stands for the assumption on the input distribution. Using elementary information theory, we can obtain the inequality

$$\Delta I(y_n; x_{n+1}) \le (n-2)\left[H^{(1)}(y_k|y_{k-1}) - H^{(2)}(y_k|y_{k-1}, x_{k-1})\right] - (n-1)(0.5 - p_1)$$

and

$$\lim_{n \to \infty} \Delta I(y_n; x_{n+1})/n \le \left[H^{(1)}(y_k|y_{k-1}) - H^{(2)}(y_k|y_{k-1}, x_{k-1})\right] - (0.5 - p_1),$$

where we have assumed the stationarity of the Markov chain input distribution and

$$P(x_{k+1} = 1|x_k = 0) = P(x_{k+1} = 0|x_k = 1) = p_1.$$

Computations for $p_1 = 0.45$ yield

$$\lim_{n \to \infty} \Delta I(y_n; x_{n+1})/n \le -0.03 < 0.$$

TABLE II
COMPARISON OF $I_N^{(1)}$ AND $I_N^{(2)}$

N	$I_N^{(i)}$	$I_N^{(ii)}$	$\Delta I_N/N$	$\max I_N$
2	0.5	1.0	0.5	1.0
3	0.9544	1.2801	0.3257	1.3215
4	1.40	1.7601	0.09	2.0103
5	1.8565	2.2625	0.0812	2.5289
6	2.3074	2.7714	0.0773	3.0846
7	2.7582	3.2830	0.0749	3.6137

Thus we conclude that p_1 can be chosen such that the Markov chain input distribution achieves a higher mutual information rate through the channel than does the product distribution. Shown in Table II are the information rates using the equiprobable product input distribution and a Markov chain input distribution with optimal value for p_1 for every N. Using elementary information theory one can show that, for the equiprobable input distribution,

$$\lim_{n \to \infty} \left(I^{(1)}(y_n; x_{n+1})/n\right) < 0.5.$$

Further insight into the nature of the optimal distribution is offered by the analysis of the optimal distribution as computed by the Arimoto–Blahut algorithm. The results for $N = 3$, 4, 5, and 6 (see Table III) show, not surprisingly, that the optimal distribution assigns most of the probability to codewords characterized by sequences of pairs 11 or 00. This property of codewords plays a basic role in the following discussion of zero-error codes.

TABLE III
MOST PROBABLE CODEWORDS FOR THE OPTIMAL DISTRIBUTION

	$N=3$		$N=4$		$N=5$		$N=6$	
	w	$P(w)$	w	$P(w)$	w	$P(w)$	w	$P(w)$
	000	0.2998	0000	0.2435	00000	0.1463	000000	0.1102
	100	0.1001	0011	0.2394	11000	0.1220	110000	0.1064
	110	0.1001	1100	0.2394	00100	0.1077	001100	0.0980
	001	0.1001	1111	0.2435	11100	0.1220	111100	0.1064
	011	0.1001			00011	0.1220	000011	0.1064
	111	0.2998			11011	0.1077	110011	0.0980
					00111	0.1220	001111	0.1064
					11111	0.1463	111111	0.1102
Total		1.000		0.9658		0.9960		0.8420

We now prove that the zero-error capacity for the one-dimensional channel with $\epsilon = 0$ ($T = 0$) is exactly $1/2$. By the definition of that channel, it is clear that of the elementary channel inputs $\{00, 01, 10, \text{and } 11\}$, only 00 and 11 are nonadjacent.

Input sequences x and x^* are *nonadjacent* or *output-distinguishable* unless $x_i x_{i+1}$ and $x_i^* x_{i+1}^*$ are adjacent for $1 \le i \le n-1$. A set D of input patterns is nonadjacent if all pairs of elements of D are nonadjacent.

Theorem 6: The maximum zero-error achievable rates for a one-dimensional zero-temperature Ising channel with inputs $x \in \{0,1\}^k$ are

$$R_k = 1/2, \qquad \text{for } k \text{ even}$$

and

$$R_k = 1/2 - 1/2k, \qquad \text{for } k \text{ odd}.$$

Proof: Let D^k be a maximal set of nonadjacent input patterns of length k. Define

$$D_0^k = \{ x = (x_1, \cdots, x_k) \in D^k : x_1 = x_2 = 0 \}$$
$$D_1^k = \{ x = (x_1, \cdots, x_k) \in D^k : x_1 = x_2 = 1 \}$$
$$D_*^k = \{ x = (x_1, \cdots, x_k) \in D^k : x_1 \neq x_2 \}.$$

If either

$$D_0^k = \varnothing \quad \text{or} \quad D_1^k = \varnothing,$$

then the first output location y_1 cannot be used to distinguish among the input patterns. Consequently, we could disregard the first input and output locations x_1 and y_1 and reduce to a system with $k-1$ input locations. In such a case,

$$|D^k| = |D^{k-1}|$$

where D^{k-1} is any maximal nonadjacent set for inputs of length $k-1$.

Let $D_0^k \neq \varnothing$ and $D_1^k \neq \varnothing$. We have that

$$D^k = D_0^k \cup D_1^k \cup D_*^k.$$

If there exists an element

$$x^* = (x_1^*, \cdots, x_k^*) \in D_*^k,$$

its subpattern $(x_2^*, \cdots, x_k^*)$ must be distinguishable from the analogous subpatterns of any element in D_0^k and D_1^k. Define

$$x = (x_1, x_2^*, \cdots, x_k^*)$$

with $x_1 = x_2^*$. Then $(D_*^k - \{x^*\}) \cup (D_0^k \cup D_1^k \cup \{x\})$ is a new maximal set of nonadjacent input patterns. Since this replacement procedure can be repeated until D_*^k is exhausted, there is always a maximal set in each pattern of which the leftmost pair is either 00 or 11. Let us denote it by

$$\tilde{D}^k = \tilde{D}_0^k \cup \tilde{D}_1^k.$$

The sets $\tilde{D}_0^k$ and $\tilde{D}_1^k$ must be maximal sets of nonadjacent patterns with their leftmost pair fixed to 00 or 11, respectively, for otherwise $\tilde{D}^k$ would not be maximal. Note that the output y_2 cannot be used to discriminate inside either $\tilde{D}_0^k$ or $\tilde{D}_1^k$. Accordingly, the elements in the sets $\tilde{D}_0^k$ and $\tilde{D}_1^k$ can be distinguished only via the patterns on the locations indexed $3, 4, \cdots, k$.

This reduces our problem to the consideration of a maximal set of input patterns $\tilde{D}^{k-2}$ on $k-2$ locations. Again, we have either of the following condition:

a) it is possible to find a maximal nonadjacent set of patterns with a repetition in the leftmost pair, or
b) $|\tilde{D}^{k-2}| = |\tilde{D}^{k-3}|$, where $\tilde{D}^{k-3}$ is a maximal set of nonadjacent patterns for a system of length $k-3$.

We can proceed as before until we are left with one of the following cases:

a′) no input location is left,
b′) one input location and one output location are left.

We claim now that, for any k and any maximal set D^k, conditions b) and b′) can occur at most once. More specifically, for k even neither one can occur, while for k odd exactly one out of b) and b′) will occur. Without loss of generality as regards condition b) we need only consider cases wherein it occurs only at the initial step (i.e., at locations 1 and 2). Similarly, we need to be concerned about condition b′) occurring only at the last step. We can now analyze separately the cases of k even and k odd.

For k even either condition b) at the first step and condition b′) at the last step both occur or neither of them occurs. However, in the former case we could shift all patterns to the left by one input location without modifying the nonadjacency property of the set. Being then free to choose any pattern on the rightmost pair, we can use the output location between them as an additional discriminating output, which contradicts the maximality assumption. For k odd if condition b) occurs at the first step then a′) will occur at the last, whereas if b) does not occur at the first step then b′) will occur at the last. We conclude that for k even a maximal set of nonadjacent input patterns D^k can be determined such that each $x \in D^k$ has the "repetition" code form. For k odd the first or last output location cannot be used for discrimination, so that the size of a maximal set D^k is

$$|D^k| = |D^{k-1}|$$

with D^{k-1} a maximal nonadjacent set of length $k-1$ input patterns.

The cardinality of the set containing all binary "repetition" words of length k is $2^{k/2}$, so for k even

$$|D^k| = 2^{k/2},$$

and

$$R_k = k^{-1} \log |D^k| = 1/2.$$

For k odd we have

$$R_k = k^{-1} \log |D^k| = k^{-1} \log |D^{k-1}| = 1/2 - 1/2k. \qquad \square$$

Sending $k \to \infty$ yields $C_0 = 1/2$ as the zero-error capacity.

III. Two-Dimensional Ising Channel

Consider the integer lattice, $\Lambda = \{(i,j): i, j \text{ integers}\}$, and the partition

$$\Lambda = D_1 \cup D_2$$

with $D_1 = \{(i,j) \in \Lambda: i+j \text{ odd}\}$ the set of the odd diagonals and $D_2 = \{(i,j) \in \Lambda: i+j \text{ even}\}$ the set of the even diagonals of Λ. Let

$$x = \{x_t, t \in D_1\} \qquad y = \{y_t, t \in D_2\}$$

be two families of random variables with bi-dimensional index, i.e., two random fields. In what follows we shall restrict attention to the binary case, $x_t \in \{0,1\}$, $y_t \in \{0,1\}$.

We consider x and y as the input and output field, respectively, of a two-dimensional C_Λ described as follows. Let the set of neighbors of $t \in \Lambda$ be

$$N(t) = \{t + (0,1), t + (0,-1), t + (1,0), t + (-1,0)\},$$

and let y_S and $x_{N(t)}$ be the restrictions of y to $S \subset D_2$ and of x to $N(t)$, respectively. Then define

$$C_\Lambda = [\{0,1\}^{D_1}, \nu(y|x), \{0,1\}^{D_2}]$$

where $\{0,1\}^{D_1}$ and $\{0,1\}^{D_2}$ are the spaces in which x and y, respectively, take values, and ν is the channel transition measure defined by

$$\nu(y_S|x) = \prod_{t \in S} q(y_t | x_{N(t)})$$

for every finite subset $S \subset D_2$. Here, q is the conditional probability for a single output y_t given the neighboring input, $x_{N(t)}$. It is specified as follows. First, we group the possible values of $x_{N(t)}$

into the subsets $\{A_i\}_{i=0,\cdots,4}$ defined by

$$A_i = \left\{ x_{N(t)}:\ \sum_{j \in N(t)} x_j = i \right\}.$$

Following the parlance of the Ising model [1] of statistical physics, let

$$\epsilon = e^{-\beta J},$$

where $\beta^{-1} = kT$, with k a physical constant, T the absolute temperature, and J a constant determining the strength of the interaction. Then Fig. 3 shows the conditional probability q for the elementary channel under consideration.

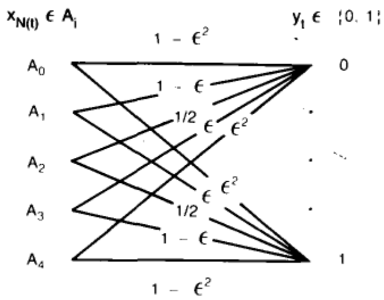

Fig. 3. Line diagram representation of $q(y_t|x_{N(t)})$.

The case $T=0$ is shown in Fig. 4, where $M_0 = A_0 \cup A_1$, $M_1 = A_3 \cup A_4$, $M_2 = A_2$. The channel in Fig. 4 operates like a *voter system*; a majority of ones or zeros in the input neighborhood imposes a one or a zero with probability one at the output. Total uncertainty prevails about y_t when $x_{N(t)}$ is comprised of two zeros and two ones.

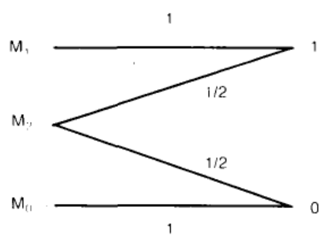

Fig. 4. Elementary channel for $T=0$.

It is possible to show [5] that C_Λ has a capacity (in bits per output site) in the usual Shannon theory sense of the term capacity. This capacity is given by the limit as S expands to cover D_2 of the supremum of $|S|^{-1} I(x; y_S)$ over all input processes $x = \{x_t, t \in D_2\}$.

When $T=0$, C_Λ has a nonzero zero-error capacity C_0, also measured in bits per site. Let $K_0(n)$ denote the largest number of patterns a code on n input locations can have such that no two of them are adjacent. Then $C_0 = \lim_{n\to\infty} \sup C_0(n)$, where $C_0(n) = n^{-1}\log_2 K_0(n)$.

Exact determination of C_0 appears to be a challenging task. In what follows we bound C_0 by developing codes that outperform simple repetition codes. Hence repetition coding is not optimum in two dimensions in the zero-error sense the way it was shown in Section II to be optimum for the one-dimensional Ising channel with $T=0$.

To gain insight it is useful to start from the analysis of some fundamental finite subsystems. We will consider sequentially:

a) the basic channel element (Fig. 5(a)),
b) the S_0 system (Fig. 5(b)),
c) the S_1 system (Fig. 5(c)).

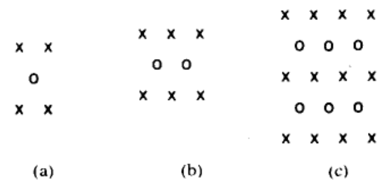

Fig. 5. Channel systems. (a) Basic channel element. (b) S_0 system. (c) S_1 system.

The two nonadjacent input patterns for the basic channel element can be obtained simply by assigning three out of the four input bits either the value 0 or the value 1. The fourth location can have an arbitrary value. Thus

$$C_0(4) = (\log_2 2)/4 = 1/4.$$

For the S_0 system all the nonadjacent input patterns can be obtained by grouping the input bits as in Fig. 6 and assigning a value 0 or 1 to all the members of each group. The number of

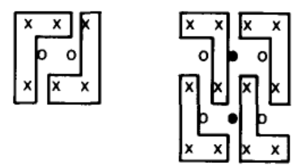

Fig. 6. Grouping of inputs for S_0 and S_1 systems.

nonadjacent patterns so determined is optimal, and consequently

$$C_0(6) = (\log 4)/6 = 1/3.$$

If for the S_1 system we grouped the input locations as in Fig. 6 and assigned zeros and ones to all the elements of each group, we would obtain a set of $2^4 = 16$ nonadjacent patterns and a rate $R = 4/12 = 1/3$. Note that the output bits marked in black in Fig. 6 would be irrelevant for discriminating among these input patterns. It is possible, however, to improve upon this "repetition" code by exploiting the marked output bits to produce 32 nonadjacent input patterns as shown in Table IV. The rate for this code is

$$R_1 = (\log 32)/12 = 0.4166.$$

It is felt that R_1 is the maximum achievable zero-error coding rate on S_1.

Observe that 28 out of the 32 input patterns in Table IV induce a single output pattern, while the remaining four have two possible images. It is easy to verify that each input pattern not listed in Table IV is adjacent to at least one of the listed input patterns. The zero-error code in Table IV is therefore at least a local maximum on S_1. Since 12 inputs surround only six outputs, we have

$$0.4166 \le C_0(12) \le 6/12 = 0.5.$$

We now consider infinite subsystems of the two-dimensional channel. The computation of achievable rates for such subsystems provides lower bounds to the zero-error capacity, and illuminates some of the basic mechanisms involved in the limiting

TABLE IV
ZERO-ERROR CODEBOOK FOR CHANNEL S_1

Number	Input	Output	Labels		Number	Input	Output	Labels	
1	0000 0000 0000	000 000	2	2	17	0011 0111 0011	011 011	0	2
2	0000 1000 1100	000 100	0	2	18	0011 0001 0011	001 001	2	0
3	1100 1000 0000	100 000	0	2	19	0001 1011 1111	001 111	0	1L
4	1110 0100 0000	110 000	0	1L	20	0011 0111 1111	011 111	0	2
5	1100 1000 1100	100 100	0	2	21	1111 0111 0011	111 011	0	2
6	1100 1110 1100	110 110	2	0	22	1111 1111 1111	111 111	2	2
7	0000 0001 0011	000 001	2	0	23	1100 1001 0011	100 001	0	0
8	0000 0010 0111	000 011	1U	0	24	0011 0111 1100	011 110	0	0
9	0000 1010 1111	000 111	0	0	25	1110 1100 1000	110 100	1U	1L
10	1100 0110 0011	110 011	0	0	26	1000 1100 1110	100 110	1L	1U
11	1100 1110 1111	110 111	2	0	27	0111 0011 0001	011 001	1L	1U
12	0011 0001 0000	001 000	2	0	28	0001 0011 0111	001 011	1U	1L
13	0011 1001 1100	001 100	0	0	29	0000 0100 0110	000 X10	1L	1U
14	1111 0101 0000	111 000	0	0	30	1111 1011 1001	111 X01	1L	1U
15	1111 1101 1000	111 100	1U	0	31	0110 0010 0000	01X 000	1L	1U
16	1111 1110 1100	111 110	2	0	32	1001 1101 1111	10X 111	1L	1U

process. Two expansion schemes will be analyzed, stripe expansion and slanted expansion, in both of which a new S_1 block is added at each step, together with the necessary extra output locations.

The stripe expanding system is shown in Fig. 7. The two output locations between abutting S_1 blocks often can be used to provide information for discriminating among input patterns. Consider, for example, the juxtaposition of patterns 1 and 22 in Table IV. By modifying the values at some of the input locations on the interface, we can obtain four nonadjacent input patterns by forcing either 11, 10, 01, or 00 at the two output locations between blocks (see Table V). The output patterns within the S_1 blocks themselves remain unchanged. Clearly, what is relevant about an input pattern in this regard is the degree to which its rightmost column and/or its leftmost column can be modified without changing the output pattern within the S_1 block.

To characterize the patterns in Table IV in this sense, we introduced the labels shown in the same table in correspondence to each pattern. The symbols used relate to the number of input bits that can change value without affecting the output pattern as follows:

- 0 no input bit can be modified,
- 2 two input bits can be modified,
- 1U the bit on the upper row only can be modified,
- 1L the bit on the lower row only can be modified.

In the label column the position of the entry, right or left, refers to the rightmost or the leftmost column of the S_1 block in

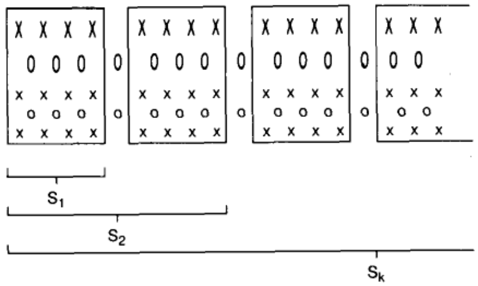

Fig. 7. Stripe expanding system. X = input location; O = output location.

TABLE V
SOME CODEWORDS ON S_2

Input Pattern	Output Pattern
00011111	
	0001111
00001111	
	0001111
00011111	
00011111	
	0001111
00001111	
	0000111
00000111	
00000111	
	0000111
00001111	
	0001111
00011111	
00000111	
	0000111
00001111	
	0000111
00000111	

question. Note that we have used the output bits on the interface for purposes of discrimination only for patterns with the columns on the interface having different values and whenever on both sides it is possible to modify the inputs at the same level, U or L. (It can be shown [6] that this engenders no loss of generality.) Using these considerations and Table IV, we proceed to determine the number of nonadjacent input patterns for a generic system S_k.

Consider first the case $k=2$. In Table VI we give the number of nonadjacent patterns on S_2 whose sixteen leftmost input bits show the indicated S_1 pattern number. The total number of nonadjacent configurations obtained for S_2 in this way is

$$N(S_2)=N(2)=1224,$$

and the corresponding rate is

$$R(S_2)=(\log 1224)/24=0.42739.$$

To obtain a recursive equation for $N(k)$, note that the input pattern on the last two columns of an S_{k-1} block conditions the number of nonadjacent configurations that can be obtained by adding another S_1 block as shown in Table VII.

To determine the coefficients in the system of recursive equations for $N(k)$, we need to know, for each of the four classes in Table VII the distribution of the new patterns for each of these four classes. That information is in Table VIII.

Let Z be the transpose of the matrix in Table VIII. Let $\boldsymbol{N}(n)$ be the column vector

$$\boldsymbol{N}(n)=(N_A(n),N_B(n),N_C(n),N_D(n))^T$$

TABLE VI
NUMBER OF NONADJACENT PATTERNS ON S_2 GIVEN
PATTERN ON LEFTMOST S_1 COMPONENT

Pattern Number	Number of Patterns on S_2	Pattern Number	Number of Patterns on S_2
1	49	17	49
2	49	18	32
3	49	19	39
4	39	20	49
5	49	21	49
6	32	22	49
7	32	23	32
8	32	24	32
9	32	25	39
10	32	26	38
11	32	27	38
12	32	28	39
13	32	29	38
14	32	30	38
15	32	31	38
16	32	32	38

TABLE VII
NUMBER OF NONADJACENT PATTERNS ON RIGHTMOST S_1 COMPONENT
OF S_k GIVEN PATTERN ON RIGHTMOST TWO COLUMNS OF S_{k-1}

Class	Pattern on Last Two Columns	Number of Nonadjacent Patterns Obtainable
A	00 11 00,11 00 11	49
B	10 01 00,11 00 11	39
C	00 11 01 10 00,11,01,10 10 01 11 00	38
D	any other	32

TABLE VIII
CLASS DISTRIBUTION OF PATTERNS ON S_k
GIVEN PATTERN CLASS ON S_{k-1}

Pattern Class on S_{k-1}	Distribution of New Patterns on S_k				
	A	B	C	D	Total
A	11	5	9	24	49
B	9	4	9	17	39
C	9	5	6	18	38
D	8	4	6	14	32

where $N_I(n)$ is the number of nonadjacent patterns in the stripe expanding system S_n belonging to class I, $I=A,B,C,D$. Then

$$\boldsymbol{N}(n+1)=Z\boldsymbol{N}(n), \qquad \text{for every } n\geq 1.$$

and

$$N(n+1)=(1,1,1,1)\boldsymbol{N}(n+1).$$

By the Perron–Frobenius theorem [4], the maximum eigenvalue $\lambda_{\max}$ of Z is real and positive. We computed $\lambda_{\max}=37.9678$. In this case it is possible to use the following asymptotic approximation

$$R(S_\infty)=\lim_{n\to\infty}R(S_n)=\lim_{n\to\infty}(\log N(n))/N_I(n)$$
$$=\lim_{n\to\infty}(\log\lambda_{\max}^n)/(12n)=\log(37.9678)/12=0.43723$$

where $N_I(n)=12n$ is the number of input points for S_n.

Again, it is important to determine which patterns can be modified on their first and last rows without affecting the output configuration. Using reasoning paralleling that just described for the striped expansion, we can determine that the asymptotic

The slanted expanding system is shown in Fig. 8. To explicate the mechanisms involved, we consider again the juxtaposition of patterns 1 and 22 of Table IV. Table IX shows the four nonadjacent patterns on S_2' obtained by modifying input bits on the interface.

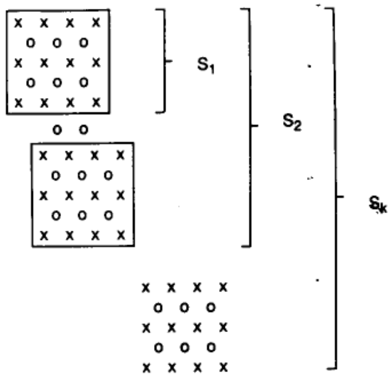

Fig. 8. Slanted expanding system.

TABLE IX
SOME CODEWORDS ON S_2'

Input	Output	Input	Output
0000		0000	
	000		000
0000		0000	
	000		000
0100		0001	
	10		01
1101		0111	
	111		111
1111		1111	
	111		111
1111		1111	
0000		0000	
	000		000
0000		0000	
	000		000
0000		0101	
	00		11
0101		1111	
	111		111
1111		1111	
	111		111
1111		1111	

zero-error rate of the slanted expansion scheme is 0.4321, which is not as high as that of the stripe expansion.

Consider a finite or semiinfinite subsystem S of the two-dimensional channel introduced before. Assume that we can tile the infinite lattice with a collection of S systems as, for instance, in Fig. 9. Here the output bit locations between blocks are represented by small circles.

To obtain a lower bound to the zero-error capacity of the global system, we can disregard the information possibly carried by the output bits between blocks. It is thus clear that, if R_S is a zero-error achievable rate for S, then the infinite system capacity C_0 must satisfy

$$R_S \leq C_0.$$

The best lower bound we have obtained in this way is

$$C_0 \geq 0.43723$$

by using the results of the stripe expansion problem.

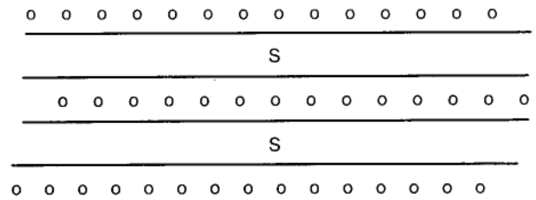

Fig. 9. Tiling of infinite lattice.

We now develop a crude upper bound on C_0. Note that each output location carries at most one bit of information. Therefore,

$$C_0(S) \leq N_0(S)/N_1(S)$$

where $C_0(S)$ is the zero-error capacity for a generic channel system S and $N_1(S)$ and $N_0(S)$ are the numbers of input and output locations of S, respectively. As an example consider the infinite slanted channel system S_∞' analyzed before. If S_k' is the finite slanted system with k blocks, we have $N_1(S_k')=12k$ and $N_0(S_k')=6k+2(k-1)$, so

$$C_0(S_\infty') \leq 0.67.$$

This crude upper bound for the slanted expanding system provides a yardstick for assessing the lower bounding technique discussed earlier.

REFERENCES

[1] E. Ising, "Beitrag zur theorie des ferromagnetismus," *Z. Phys.*, vol. 31, p. 253, 1925.

[2] R. G. Gallager, *Information Theory and Reliable Communications.* New York: Wiley, 1968.

[3] S. Arimoto, "An algorithm for computing the capacity of arbitrary discrete memoryless channels," *IEEE Trans. Inform. Theory*, vol IT-18, pp. 14–20, 1972.

[4] R. E. Blahut, "Computation of channel capacity and rate distortion functions," *IEEE Trans. Inform. Theory*, vol IT-18, pp. 460–473, 1972.

[5] T. Berger and S. Y. Shen, "Communication theory via random fields," presented at 1983 IEEE Int. Symp. on Information Theory, St. Jovite, PQ, Canada, Sept. 1983.

[6] F. Bonomi, "Problems in the information theory of random fields," Ph.D. dissertation. School of Elect. Eng., Cornell Univ., Ithaca, NY, Aug. 1985.

On the Achievable Rate Region of Sequential Decoding for a Class of Multiaccess Channels

ERDAL ARIKAN, MEMBER, IEEE

***Abstract* — The achievable-rate region of sequential decoding for the class of pairwise reversible multiaccess channels is determined. This result is obtained by finding tight lower bounds to the average list size for the same class of channels. The average list size is defined as the expected number of incorrect messages that appear, to a maximum-likelihood decoder, to be at least as likely as the correct message. The average list size bounds developed here may be of independent interest, with possible applications to list-decoding schemes.**

I. INTRODUCTION

The application of sequential decoding to multiaccess channels was considered in [1], where it is shown that all rates in a certain region $\boldsymbol{R}_0$ are achievable within finite average computation per decoded digit. However, the question of whether $\boldsymbol{R}_0$ equals the achievable-rate region $\boldsymbol{R}_{\text{comp}}$ of sequential decoding is left open.

Manuscript received April 26, 1988; revised March 14, 1989. This work was supported in part by the Defense Advanced Research Projects Agency under Contract N000 14-84-0357.

The author is with the Department of Electrical Engineering, Bilkent University, P.K. 8, 06572, Maltepe, Ankara, Turkey.

IEEE Log Number 8933107.

0018-9448/90/0100-0180$01.00 ©1990 IEEE