April 1990 LIDS-P-1966

CODING FOR CHANNELS WITH PARTIALLY LOCALIZED ERRORS*

by

L. A. Bassalygo, S. I. Gelfand, M. S. Pinsker**

ABSTRACT

A model of a communication system over the binary channel is studied. It is assumed that the encoder knows possible positions and the maximum multiplicity of errors. Upper and lower bounds for the asymptotically optimal rate are derived.

*The authors are with the Institute for Problems of Information Transmission; 19, Ermolova str., Moscow 101447, USSR. The last version of the paper was supported at M.I.T. by an A. P. Sloan Foundation grant number 88-10-1.

**This version of the paper was written while M. S. Pinsker was visiting M.I.T, Laboratory for Information and Decision Systems.

CODING FOR CHANNELS WITH PARTIALLY LOCALIZED ERRORS

BASSALYGO L.A., GELFAND S.I., PINSKER M.S.

Abstract. A model of a communication system over the binary channel is studied. It is assumed that the encoder knows possible positions and the maximum multiplicity of errors. Upper and lower bounds for the asymptotically optimal rate are derived.

1. Introduction

In this paper we continue to study various models of channels with localized errors when encoder has some information about possible configuration of errors in the channel. This study was initiated in [1], and in this introduction we remind shortly the basic model of a channel with localized error.

The channel is binary, with input and output alphabets $\{0,1\}$. We consider block codes of length n. The idea of localized errors is that, prior to the encoding of a block, the encoder possesses some information about possible configurations of errors that can occur during the transmission of this block. Namely, the encoder knows that some positions in the block are errorless (perfect), and the errors might occur only on remaining (dubious) positions. On the other hand, the decoder does not having any *a priori* information about which positions are dubious, so that the decoding should be, in a sense, universal.

Making all this rigorous, the notion of a code correcting up to ℓ localized errors (the number of dubious positions does not exceed ℓ) introduced in [1], and some upper and lower bounds for the size of such a code were proved. These bounds yield, in particular, the formula $R = 1 - h(\lambda)$ for the optimal rate R in the case $n \to \infty$, $\lambda = \ell/n = \text{const}$..

The generalization we consider in this paper deals with the case when a further restriction on the number of errors in a given block is imposed. Namely, we assume that some was ℓ dubious positions are known to the encoder, and, moreover, the number of errors accuring on these dubious positions does not exceed some constant $t \leq \ell$; here ℓ and t are two parameters. Again, we assume that the remaining $n - \ell$ positions are perfect (no errors on these positions), and that the decoder has no a priori information about which positions are dubious. The rigorous definition (see next section) leads to the notion of a length n code correcting $\leq t$–on–ℓ localized errors. Our aim is to derive some asymptotic bounds for the size of such a code.

Let us note here that special cases of this problem are coding for the ordinary binary symmetric channel (when $\ell = n$) and coding for channels with localized errors from [1] (when $t = \ell$). As the exact asymptotic rate for the binary symmetric channel is at present unknown, we can not hope to get any exact formulas for the asymptotic optimal rate for the more general case we consider here. However, we will try to convince the reader that

The authors are with the Institute for Problems of Information Transmission; 19, Ermolova str., Moscow, 101447, USSR. The last version of the paper was supported at M.I.T. by an A.P. Sloan Foundation grant number 88-10-1

the situation in our case is, in some sense, not worse than the situation for the binary symmetric channel.

We stress once more that our problem is purely combinatorial with no probabilities involved.

2. The formulation of the problem.

By a *configuration* in a block of length n we mean a subset $E \subset \{1, \dots, n\}$; the *multiplicity* of a configuration is the number of elements $|E|$ in E. We say that a binary vector $e = (e_1, \dots, e_n)$ of length n is *subordinate* to a configuration E if $e_i = 0$ for $i \notin E$. For a given configuration E denote by $\mathcal{V}_t(E)$ the set of all vectors of the Hamming weight $\leq t$ that are subordinate to E. Denote also by $\mathcal{E}_\ell$ the set of all configuration of multiplicity $\leq \ell$.

A binary code of length n and size M for a channel with localized errors of multiplicity $\leq \ell$ is a pair (φ, ψ) where φ is the encoding mapping, $\varphi(m, E) \in \{0,1\}^n$ for $m \in \{1, \dots, M\}$, $E \in \mathcal{E}_\ell$, and ψ is the decoding mapping, $\psi(e) \in \{1, \dots, M\}$ for $e \in \{0,1\}^n$. We say that such a code *corrects all* $\leq t$–on–ℓ (partially) localized errors if for any messages $m \in \{1, \dots M\}$, any configuration $E \in \mathcal{E}_\ell$ and any vector $e \in \mathcal{V}_t(E)$ we have $\psi(\varphi(E, m) \oplus e) = m$. The interpretation of this definition is clear: the encoder knows $\leq \ell$ positions on which errors can occur (the configuration E) and the decoder corrects all possible combinations of $\leq t$ errors at these positions (all e's from $\mathcal{V}_t(E)$).

The rate of a code is defined, of course, as $R = n^{-1} \log_2 M$. The problem to find, for given n, ℓ and t, the maximum size $M(n, \ell, t)$ of a code correcting $\leq t$–on–ℓ localized errors. The asymptotic version of the problem is to find, for fixed λ and τ, the maximum value $R = R(\lambda, \tau)$ of the rate of such code with $\ell = \lambda n$, $t = \tau n$ as $n \to \infty$.

3. Results

We are interested in this paper mainly with asymptotic bounds (that is, bounds for $R(\lambda, \tau)$). We begin with lower (existence) bounds.

Introduce, for $0 \leq \tau \leq \lambda \leq 1$, the function $L(\lambda, \tau)$ by the formula

$$L(\lambda, \tau) = \max_{0 \leq \alpha \leq \min(\tau, 1-\lambda)} \max_{\alpha \leq \beta \leq \min(2\tau - \alpha, \lambda)} \left[(1 - \lambda) \left(\frac{\alpha}{1 - \lambda} \right) + \lambda h \left(\frac{\beta}{\lambda} \right) \right].$$

THEOREM 1. *For any $0 \leq \tau \leq \lambda \leq 1$ we have*

$$R(\lambda, \tau) \geq 1 - L(\lambda, \tau).$$

Properties of the function $L(\lambda, \tau)$ are given in the following lemma whose (easy) proof we omit.

LEMMA 1.

- **i** *$L(\lambda, \tau)$ is a nondecreasing function of λ, τ.*
- **ii** *$L(\lambda, \tau) \leq 1$; $L(\lambda, \tau) = 1$ for $\lambda \geq 1/2$, $\tau \geq 1/4$; $L(\lambda, \tau) < 1$ for all other λ, τ.*
- **iii** *$L(\lambda, \tau) = h(\lambda)$ for $\lambda \leq 1/2$, $\lambda(1 - \lambda) \leq \tau \leq \lambda$.*

In particular, Theorem 1 provides a non-trivial (that is, non–zero) lower bound for $R(\lambda,\tau)$ in the region

$$\{\lambda,\tau):\quad \textit{either } \lambda<1/2,\ \textit{or } \tau<1/4\}$$

on the (λ,τ)–plane.

Upper (non–existence) bounds are given by the following theorem.

THEOREM 2. *If (λ,τ) satisfies the condition $\lambda(1-\lambda)\le\tau\le\lambda$ then*

$$R(\lambda,\tau)\le 1-h(\lambda).$$

COROLLARY 1. *$R(\lambda,\tau)=0$ (for $\tau\le\lambda\le 1$) if and only if $\lambda\ge 1/2,\ \tau\ge 1/4$.*

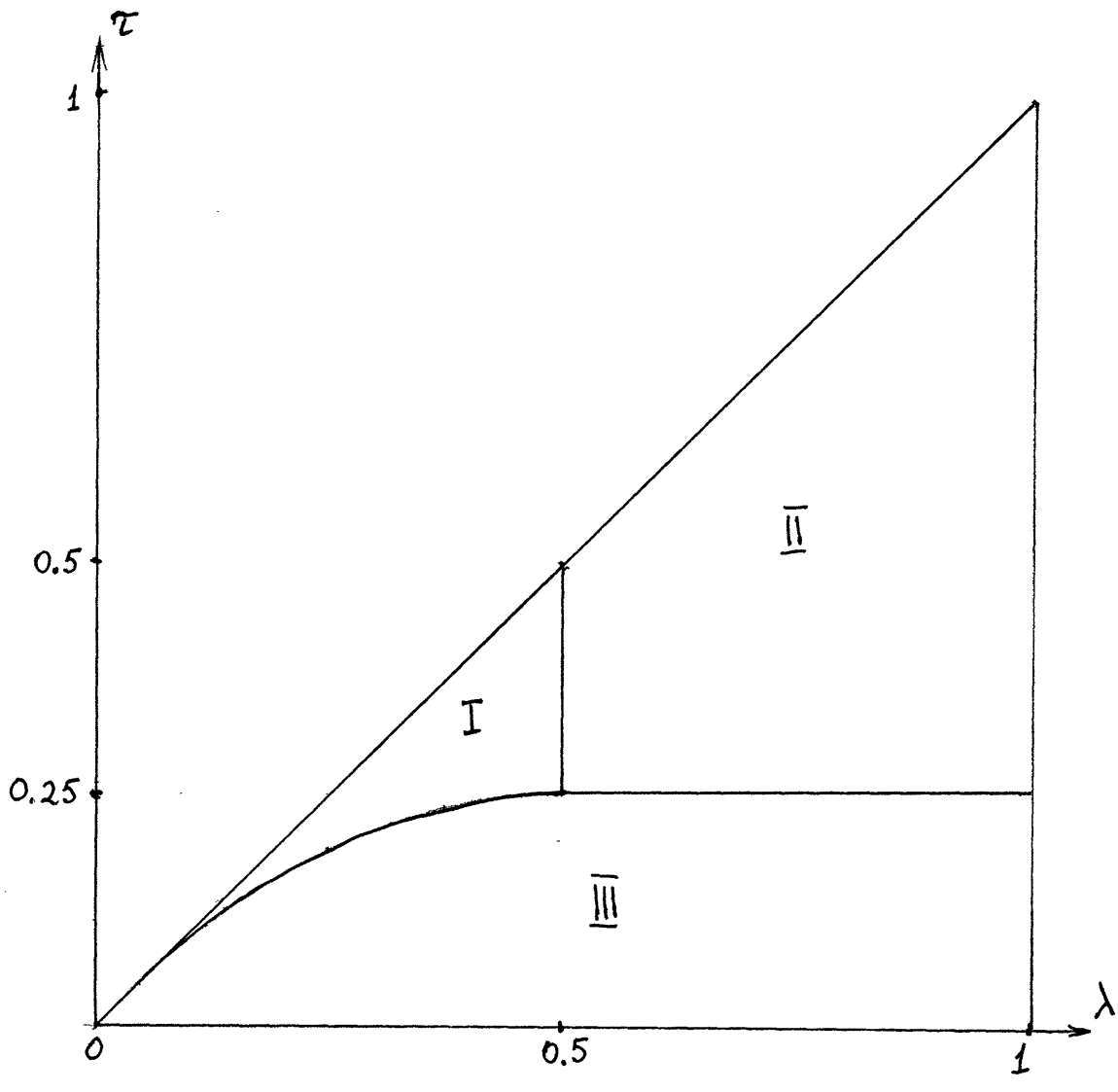
COROLLARY 2. *For $\lambda(1-\lambda)\le\tau\le\lambda$ we have*

$$R(\lambda,\tau)=1-h(\lambda).$$

Let us remark also, that, denoting by $f(\tau)$ any asymptotical upper bound for the optimal rate of ordinary error correcting code, we have clearly

$$R(\lambda,\tau)\le f(\tau). \tag{1}$$

Taking for $f(\tau)$ the Elias bound $f(\tau)=1-h(\rho(\tau))$, where $\rho(\tau)\le 1/2$ is the root of the equation $\rho(1-\rho)=\tau$ (the so called Elias radius), we see that $\tau=\lambda(1-\lambda)$ the bound (1) coincides with the bound given by Theorem 2, and for $\tau<\lambda(1-\lambda)$ the bound (1) is better than the bound in Theorem 2. Hence, what we know about the function $R(\lambda,\tau)$ is summarized at the following picture:



where in the region I ($\lambda(1-\lambda) \le \tau \le \lambda$) we have $R(\lambda,\tau) = 1 - h(\lambda)$, in the region II ($\lambda \ge 1/2,\ \tau \ge 1/4$) we have $R(\lambda,\tau) = 0$, and in the region III we have the bounds $1 - L(\lambda,\tau) \le R(\lambda,\tau) \le 1 - h(\lambda)$.

4. Proof of Theorem 1.

The proof of Theorem 1 is similar to the proof of Theorem 3 from [1]. Let M be the smallest integer satisfying the condition

$$M \ge S(n,\ell,t)/2n. \tag{2}$$

We prove that there exists a code a length n and size M correcting all $\leq t$–on–ℓ localized errors. The encoding φ which satisfies the required condition will be constructed as follows. Let for each $m \in \{1, \dots, M\}$ we are given a family $\{x^m(i),\ i = 1, \dots, n\}$ of n binary vectors, each of length n. We need the following notion of a good vector in the family $\{x^m(i)\}$. A vector $x^m(i)$ is said to be *good* for E if for any $m' \neq m$, for any j and for any $e \in \mathcal{V}_t(E)$ we have

$$t = d(x^m(i) \oplus e,\ x^m(i)) < d(x^m(i) \oplus e,\ x^{m'}(j)).$$

Here $d(\cdot, \cdot)$ is the Hamming distance.

Let us assume now that there exists a family of Mn vectors $\{x^m(i), m \in \{1, \dots, M\}$, $i = 1, \dots, n\}$ satisfying the following condition:

(P) *For any $m \in \{1, \dots, M\}$ and for any configuration $E \in \mathcal{E}_t$ there exists i such that the vector $x^m(i)$ is good for E.*

Using such a family $\{x^m(i)\}$ we construct the code as follows:

$$\varphi(m, E) = x^m(i) \text{ where } x^m(i) \text{ is good for } E;$$
$$\psi(y) = m \text{ where } x^m(i) \text{ is the vector which is closest to } \mathcal{Y}$$
$$\text{among all } Mn \text{ vectors } \{x^m(i)\}.$$

As the family $\{x^m(i)\}$ satisfies the condition (P), the encoding φ is well–defined, and as one can easily see, the constructed code corrects all $\leq t$–on–ℓ errors. We have to verify only that under the condition (2) there exists a family $\{x^m(i)\}$ satisfying (P). Let us compute the number of families $\{x^m(i)\}$ such that the condition (P) does not hold for a given $m_0 \in \{1, \dots, M\}$ and for a given configuration E. Choosing in such a configuration vectors $x^m(j)$, $m \neq m_0$, arbitrarily, we see that the condition (P) will not be satisfied for $m_0 \in \{1, \dots, M\}$ and for E if each of vectors $x^{m_0}(i)$, $1 \leq n$, is close to one of $(M-1)n$ vectors $x^m(j)$, $m \neq m_0$. Namely, denote by $\mathcal{T}(E, \ell, t)$ the set of vectors x of the form $x = x_1 + x_2$, where $x_1 \in \mathcal{V}_t(E)$ and the Hamming weight of x_2 does not exceed t, i.e. $wt(x_2) \leq t$. Then the condition (P) will not be satsified for $m_0 \in \{1, \dots, M\}$ and for E if each of vectors $x^{m_0}(i)$, $1 \leq n$, is of the form $x^m(j) + x$ where $m \neq m_0$ and $x \in \mathcal{T}(E, \ell, t)$. Therefore the number of families $\{x^m(i)\}$ such that the condition (P) does not hold for a given $m_0 \in \{1, \dots, M\}$ and for a given configuration E does not exceed

$$[(M-1)n|\mathcal{T}(E, \ell, t)|]^n\, 2^{n^2(M-1)}$$

(here and later $|X|$ denotes the cardinality of a finite set X). Therefore the number of families $\{x^m(i)\}$ which do not satisfy the condition (P) does not exceed

$$MC_n^\ell\, [(M-1)n|\mathcal{T}(E, \ell, t)|]^n\, 2^{n^2(M-1)}. \tag{3}$$

Now we have to compute the number of elements in $\mathcal{T}(E, \ell, t)$. It is given by the following simple Lemma, whose proof we leave to the reader.

LEMMA 2.

i *The number of elements* $|\mathcal{T}(E,\ell,t)|$ *in* $\mathcal{T}(E,\ell,t)$ *does not depend on the configuration* E *of given multiplicity* ℓ, *and is given by the formula*

$$|\mathcal{T}(E,\ell,t)| = \sum_{0\le\alpha\le\min\{t,n-\ell\}} C_{n-\ell}^{\alpha} \sum_{\alpha\le b\le\min\{2t-\alpha,\ell\}} C_{\ell}^{b}.$$

ii *Asymptotically, as* $n\to\infty$, $\ell=\lambda n$, $t=\tau n$,

$$|\mathcal{T}(E,\ell,t)| \approx 2^{nL(\lambda,\tau)}.$$

This lemma, together with formula (3) and the condition (2) on M, implies that the number of configurations which do not satisfy the condition (P) is less that the total number 2^{n^2M} of configurations. Hence, there exists at least one configuration that satisfies the condition (P), and Theorem 1 is proved.

5. Proof of Theorem 2

Suppose that λ and τ satisfy the condition $\lambda(1-\lambda)\le\tau\le\lambda$ and let for some $\varepsilon>0$ and for a sufficiently large n there exists a code of size

$$M > 2^{n[1-h(\lambda)+\varepsilon]}, \tag{4}$$

correcting $\le\tau n$–on–λn localized errors. We try to achieve a contradiction, thus proving that in the above region we have $R(\lambda,\tau)\le 1-h(\lambda)$.

For any message $m\in\{1,\dots,M\}$ and for any $E\in\mathcal{E}_\ell$ let $\varphi(m,E)$ be the corresponding codeword. Denote by $\mathcal{B}(m,E)$ the set of all vectors Y of the form $Y=\varphi(m,E)+x$, $x\in\mathcal{V}(E)$. Next, for any m denote $\mathcal{B}_m=\cup_{E\in\mathcal{E}_\ell}\mathcal{B}(m,E)$. The next lemma (which is essentially Lemma 1 from [1]) gives a lower bound for the size of any set $\mathcal{B}_m$.

We need the following combinatorial construction. For any binary vector Y of length n and any configuration E denote by $\mathcal{B}(y,E)$ the set of vectors of the form $Y+x$, $x\in\mathcal{V}(E)$.

LEMMA 3. *Let we are given* Q *binary vectors* $Y_1,\dots,Y_Q$ *of length* n *and* Q *pairwise distinct configurations* $E_1,\dots,E_Q$. *Then the size of the set* $\mathcal{B}=\cup_{q=1}^{Q}\mathcal{B}(Y_q,E_q)$ *satisfies* $|\mathcal{B}|\ge Q$.

Proof of Lemma 3 uses the induction of n. For $n=1$ the statement is obvious.

Let now $Y_1,\dots,Y_Q$ and $E_1,\dots,E_Q$ satisfy the condition of the lemma. We construct, for the length $n-1$, two families of vectors $u_1,\dots,u_S$ and $v_1,\dots,v_T$, and two families of configurations $F_1,\dots,F_S$ and $G_1,\dots,G_T$ as follows. We will sort out all $q, 1\le q\le Q$, one by one, constructing for each q either a new pair (u_s,F_s) from the first family, or a new pair (v_t,G_t) from the second family, or two new pairs (u_s,F_s), (v_t,G_t), one from the first family, another from the second one. If $q, 1\le q\le Q$, is such that $n\notin E_q$ and the n–th component of Y_q is 0, then we delete this n–th component, place the shortened

vector into the first family of vectors ($\mathcal{U}$–family), and place the configuration E_q (which belongs to $\{1,\dots,n-1\}$) into the first family of configurations (F–family). If $q, 1 \le q \le Q$, is such that $n \notin E_q$ and the n–th component of Y_q is 1, then we delete this n–th component, place the shortened vector into the second family of vectors (v–family), and place the configuration E_q (which again belongs to $\{1,\dots,n-1\}$) into the second family of configurations (G–family). Finally, if q is such that $n \in E_q$, then we delete the n–th component of Y_q, place the shortened vector both into both the first and into the second family of vectors; also, we place the configuration $E_q \cap \{1,\dots,n-1\}$ of the length $n-1$ into both families of configurations. Denote by Q_0, Q_1, Q_* the number of indices $q, 1 \le q \le Q$, corresponding to the first, second, and third type above (that is, such that $n \notin E_q$ and $(y_q)_n = 0$; $n \notin E_q$ and $(Y_q)_n = 1$; $n \in E_q$ respectively). Clearly,

$$S = Q_0 + Q_*,\ T = Q_1 + Q_*,\ Q = Q_0 + Q_1 + Q_*.$$

By the construction of families $\{u_s, F_s\}$ and $\{v_t, G_t\}$ it is clear that if we complete the vectors from $\cup_{s=1}^{S} \mathcal{B}(U_s, F_s)$ by 0 at the n–th position and the vectors from $\cup_{t=1}^{T} \mathcal{B}(v_t, G_t)$ by 1 at the n–th position we obtain each vector from $\mathcal{B} = \cup_{q=1}^{Q} \mathcal{B}(Y_q, E_q)$ exactly once. Therefore

$$|\cup_{s=1}^{S} \mathcal{B}(u_s, F_s)| + |\cup_{t=1}^{T} \mathcal{B}(v_t, G_t)| = |\mathcal{B}|. \tag{5}$$

To estimate the left–hand–side of (5) we denote by S' the number of distinct configurations among F_s and by T' the number of distinct subsets among G_t. Using the fact that all E_q, $1 \le q \le Q$, are distinct, one easily gets that

$$S' + T' \ge S + T - Q_* = Q_0 + Q_1 + Q_* = Q.$$

Hence, (5) and the induction assumption implies that $|\mathcal{B}| \ge Q$ and Lemma 3 is proved.

Recalling the definition of sets $\mathcal{B}_m$, $1 \le m \le M$, we see that Lemma 3 yields the following estimate

$$|\mathcal{B}_m| \ge \sum_{i=1}^{\ell} C_n^i.$$

Now, if M satisfies (4) and n is sufficiently large, there exists an element $x \in \{0,1\}^n$ which belongs to at least $K = 2^{n\varepsilon/2}$ distinct sets $\mathcal{B}_m$. This means that for this point x there exist K messages $m_1,\dots,m_K$, K configurations $E_1,\dots,E_K \in \mathcal{E}_\ell$ and K vectors $x_1,\dots,x_K$ with $x_k \in \mathcal{V}(E_k)$ such that $x = \varphi(m_k, E_k) + x_k$ for any k, $1 \le k \le K$. As the weight of any x_k does not exceed ℓ, there exist at least $I \ge K/\ell \ge 2^{n\varepsilon/4}$ (for sufficiently large n) distinct messages among m_k such that the corresponding vectors x_k all have the same weight $r \le \ell$. Without loss of generality we can assume that these messages are $m_1,\dots,m_I$.

Summarizing all this, we come to the following picture. There exists an element $x \in \{0,1\}^n$ such that the sphere of the radius $r \le \ell$ centered at x contains I codewords $\varphi(m_i, E_i)$ for I distinct messages $m_1,\dots,m_I$ and I configurations $E_1,\dots,E_I \in \mathcal{E}_\ell$ with the property that $x + \varphi(m_i, E_i) \in \mathcal{V}(E_i)$ for all $i, 1 \le i \le I$.

Now we use a version of the Johnson lemma which estimates the minimum distance between these $\varphi(m_i, E_i)$. Note here that in the similar situation this lemma is also used in the proof of the Elias bound (see [2]).

LEMMA 4. *Let I points $x_1, \dots, x_I$ lie on some sphere S_r of radius r in $\{0,1\}^n$. Then*

$$\frac{1}{n} d(x_i, x_j) < \frac{I}{I-1} \frac{2r}{n} \left(1 - \frac{r}{n}\right)$$

for at least one pair $i \neq j$.

PROOF: Without loss of generality we can assume that S_r is centered at zero. Write down the binary $(I \times n)$–matrix A whose (i,k)–th entry a_{ik}, $1 \leq i \leq I$, $1 \leq k \leq n$, is the k–th coordinate of the vector x_i. Denote by v_k the number of 1's in the k–th column of this matrix. Counting down the total number of 1's in A we get

$$\sum_{k=1}^{n} v_k = rI. \tag{6}$$

(because the weight of each x_i is r). Next, we compute the average pairwise distance

$$D = \left(\frac{I(I-1)}{2}\right)^{-1} \sum_{1 \leq i < j \leq I} d(x_i, x_j).$$

between distinct x_i's. Any pair of symbols $(0,1)$ in the same column of A (that is, a pair of elements $a_{ik_1} = 0$, $a_{ik_2} = 1$ of the matrix A) puts 1 into the sum in D. Hence,

$$D = \sum_{k=1}^{n} v_k (I - v_k) = I \sum_{k=1}^{n} v_k - \sum_{k=1}^{n} v_k^2 = rI^2 - \sum_{k=1}^{n} v_k^2.$$

One can easily see that under the condition (6) we have $\sum_{k=1}^{n} v_k^2 \geq \frac{r^2 I^2}{n}$. Hence

$$\left(\frac{I(I-1)}{2}\right) D \leq rI^2 \left(1 - \frac{r}{n}\right).$$

As the minimum distance between vectors x_i does not exceed the average distance D, Lemma 4 is proved.

Apply this lemma to our situation. Recalling that $I \geq 2^{n\varepsilon/4}$, $r \leq \ell$ and $\lambda(1-\lambda) \leq \tau$, we see that among the codewords $\varphi(m_i, E_i)$ there exists at least two, $\varphi(m_i, E_i), \varphi(m_j, E_j)$, with the distance between them not exceeding $2t$. As both codewords $\varphi(m_i, E_i), \varphi(m_j, E_j)$ lies on the distance exactly r from x, and $x + \varphi(m_i, E_i) \in \mathcal{V}(E_i)$, $x + \varphi(m_j, E_j) \in \mathcal{V}(E_j)$, we can find two error vectors, $e_i \in \mathcal{V}(E_i)$, $e_j \in \mathcal{V}(E_j)$, both of the weight $\leq t$, such that $y = \varphi(m_i, E_i) + e_i = \varphi(m_j, E_j) + e_j$. Thus, receiving the vector Y, decoder can not distinguish between messages m_i and m_j. This means that our code can not correct all $\leq t$–on–ℓ localized errors, contrary to the assumption at the beginning of the proof. Theorem 2 is proved.

REFERENCES

1. L.A. Bassalygo, S.I. Gelfand, M.S. Pinsker, *Coding for channels with localized errors*, Proc. 4–th Soviet–Swedish Workshop in Information Theory, Gotland, Sweden, (1989).
2. F.J. MacWilliams, N.J.A. Sloane, *The theory of error–correcting codes*, North–Holland, Amsterdam, (1977).