On the error-correcting capability of linear codes*

(Presented by the academician A. N. Kolmogorov 1 X 1968)

M. V. Kozlov

originally published in *Doklady Akademii nauk SSSR*, 1969, Vol. 186(2), p. 275-278

Let $GF(2)$, be the Galois field with two elements[1]; Let V^n be the set of all vectors $\mathbf{v} = (v^1, \ldots, v^n)$ with components $v^i \in GF(2)$. V^n forms the n-dimensional vector (coordinate) space over $GF(2)$ [2]. Any homomorphism of V^k into V^n, $k > n$ gives an (k, n) linear code (see [4]). The image of V^k by this representation gives the code space and its elements are the codewords. It is convenient to represent a (k, n) linear code with the help of a matrix $(a_i^j), i = 1, \ldots, k;\ j = 1, \ldots, n$ over $GF(2)$ of rank k, whose rows $\mathbf{a}_1, \ldots, \mathbf{a}_k$ form a basis of the code space.

The error-correcting capability of a (k, n)-linear code can be characterized with the minimum of the weights [3] of non-zero codewords: If the minimum equals $2t + 1$, then any two codewords are distinguished by no less than $2t + 1$ positions, therefore a distortion of any codeword with less than t positions (substitution of 0 by 1 and 1 by 0) does not lead to a loss of information. One of the problems in coding theory consists in the finding of theoretical bounds for the error-correcting capability of linear codes. Here there are number of results (see [4, 1]), from which we consider the following due to Varshamov and Gilbert:

If

$$\sum_{0 \le i \le t-2} C_{n-1}^i < 2^{n-k}, \tag{1}$$

then there exists a (k, n)-code with minimal weight $\geq t$. However until now the exact upper bound of the minimal weights of linear codes remains unknown. In this present note, we proove that the minimal weights of most of (k, n)-linear codes are gathered around the smallest solution (in t) of inequality (1). We pass to the exact formula of the assertion.

Let us consider a uniform probabilistic measure on the set of all binary $k \times n$ matrices, by considering that $a_i^j,\ i = 1, \ldots, k;\ j = 1, \ldots, n$, are mutually independent random variables, taken from the values 0 and 1, with uniform probabilities. We introduce the random variable η_n, equal to the minimum weight of a code space with generator matrix (a_i^j) and we define $\beta_n(t) = P\{\eta_n > t\}$ its distribution function. Hereafter we will suppose that $k = [nR]$ [4], where

*Translated from Russian by P. Loidreau who expresses many thanks to his dictionary

[1] The elements of the field are 0 and 1, the addition and the multiplication are determined by the relations $0 \oplus 0 = 1 \oplus 1 = 0$, $0 \oplus 1 = 1$, $0 \cdot 0 = 0 \cdot 1 = 0$, $1 \cdot 1 = 1$

[2] The operations in V^n are given by the relations $\mathbf{v} \oplus \mathbf{u} = (v^1, \ldots, v^n) \oplus (u^1, \ldots, u^n) = (v^1 \oplus u^1, \ldots, v^n \oplus u^n)$, $w \cdot \mathbf{v} = (w \cdot v^1, \ldots, w \cdot v^n)$, where $w, v^i,\ u^i \in GF(2)$

[3] The number of non-zero coordinates of vector $\mathbf{a}$ is called its weight $w(\mathbf{a})$

[4] $[a]$ is the integer part of the number a

$0 < R < 1$ is fixed, and we will interest ourselves to the asymptotic behaviour of $\beta_n(t)$ under $n \to \infty$. It is possible to describe an asymptotic expression for the smallest solution t_n of the inequality (1).

$$t_n = np + \frac{1}{2}\left(\log_2 \frac{1-p}{p}\right)^{-1} \log_2 n + O(1), \tag{2}$$

where $p < 1/2$ is the root of the equation, $1 - R = H(p)$ [5]. With calculation (2), we rewrite the result of Varshamov-Gilbert under the following form.

Theorem 1 (Varshamov, Gilbert) *If the difference*

$$\left[np + \frac{1}{2}\left(\log_2 \frac{1-p}{p}\right)^{-1} \log_2 n\right] - s_n \tag{3}$$

is limited from below by some constant c_p [6], then $\beta_n(s_n) > 0$.

For the distribution function $\beta_n(t)$, it is known (see [5, 3]) that

$$\beta_n(t) \geq 1 - 2^{k-n} \sum_{0 \leq i \leq t} C_n^i,$$

from which it is possible to derive

Theorem 2 (Gallager, Kochelev) *If the difference* (3) *tends to $+\infty$ then $\beta_n(s_n) \to 1$, under $n \to +\infty$.*

The following theorem complements Theorem 2

Theorem 3 *If the difference* (3) *tends to $-\infty$, then $\beta_n(s_n) \to 0$, under $n \to +\infty$.*

Theorem 3 is an immediate consequence of the following:

Theorem 4 *Uniformly for $t \leq \tau_n = \left[np + \frac{1}{2}\left(\log_2 \frac{1-p}{p}\right)^{-1} \log_2 n\right] - c'_p$ we have the relation*

$$\beta_n(t) = [1 + O(2^{-\delta\sqrt{n}})] \exp\left\{-2^{k-n} \sum_{i=0}^{t} C_n^i\right\}, \quad n \to \infty. \tag{4}$$

Corollary 1 *If the difference $\left[np + \frac{1}{2}\left(\log_2 \frac{1-p}{p}\right)^{-1} \log_2 n\right] - s_n \geq c''_p$, then $\beta_n(s_n) > 0$. In other words, there exists a (k,n)-code with minimum weight at least $np + \frac{1}{2}\left(\log_2 \frac{1-p}{p}\right)^{-1} \log_2 n + O(1)$.*

We pass to the proof of Theorem 4. We introduce the notation D_ν^t for the events $\{w(x_\nu^1 \mathbf{a}_1 \oplus \cdots \oplus x_\nu^k \mathbf{a}_k) > t\}$, consisting in this: the weight of the linear combination of the lines $\mathbf{a}_i$ of the matrix (a_i^j), $i = 1, \ldots, k$; $j = 1, \ldots, n$, with coefficients x_ν^i is greater than t; here, $x_\nu^1, \ldots, x_\nu^k$ are the binary writing of the number ν, $1 \leq \nu \leq 2^k - 1$. Then

$$\beta_n(t) = P\{\eta_n > t\} = P\{\bigcap_{1 \leq \nu \leq 2^k - 1} D_\nu^t\}.$$

[5] $H(p) = -p \log_2 p - (1-p) \log_2 (1-p)$

[6] c_p, c'_p, c''_p are constants depending only on p

Lemma 1 ([2]) *Let $G_\nu,\ \nu = 1, \dots, N$ be an arbitrary number of events and*

$$S_r = \sum_{1 \le \nu_1 < \nu_2 < \cdots < \nu_r \le N} P\{G_{\nu_1} G_{\nu_2}, \dots, G_{\nu_r}\}. \tag{5}$$

Then

$$P\{\cup_{\nu=1}^N G_\nu\} = S_1 - S_2 + S_3 - \cdots + (-1)^{N-1} S_N.$$

with this

$$P\{\cup_{\nu=1}^N G_\nu\} \ge S_1 - S_2 + \cdots - S_{2m}, \tag{6}$$

$$P\{\cup_{\nu=1}^N G_\nu\} \le S_1 - S_2 + \cdots - S_{2m} + S_{2m+1}, \tag{7}$$

where m is any integer such that $2m + 1 \le N$.

By using (6) and (7) with the events $G_\nu = \overline{D}_\nu^t,\ N = 2^k - 1$, and by using the proof of lemma 4 below, we arrive at the following expression for $\beta_n(t)$:

$$\beta_n(t) = \sum_{r=0}^{2m} \frac{[-u_n(t)]^r}{r!} + O\left(\frac{[u_n(t)]^{2m+1}}{(2m+1)!} + 2^{-\epsilon n} e^{u_n(t)} \right), \tag{8}$$

where $u_n(t) = 2^{k-n} \sum_{0 \le i \le t} C_n^i,\ m = O(\sqrt{m})$. If $u_n(\tau_n) \le \frac{1}{2} m$, then the residual term in (8) tends uniformly to zero for $t \le \tau_n$ under $n \to \infty$. Further

$$\begin{aligned} \log_2 u_n(\tau_n) &= -\tfrac{1}{2} \log_2 n + n \left[R - 1 + H\left(\tfrac{\tau_n}{n}\right) \right] + O(1) = \\ &= n \left[H\left(\tfrac{\tau_n}{n}\right) - H\left(\tfrac{t_n}{n}\right) \right] + O(1) = \tfrac{1}{2} \log_2 n - \left(\log_2 \tfrac{1-p}{p} \right) c'_p + O(1). \end{aligned}$$

By choosing the constant c'_p sufficiently large, we obtain the fulfilment of the condition $u_n(\tau_n) \le \frac{1}{2} m$ whatever be m of the form $O(\sqrt{n})$. Thus it remains to proove lemma 4, preceded by the following two propositions.

Lemma 2 *If the system of vectors $\mathbf{x}_i = (x_i^1, \dots, x_i^k) \in V^k,\ i = 1, \dots, \ell;\ \ell \le k$, are linearly independent, then the random vectors $\mathbf{b}_i = x_i^1 \mathbf{a}_1 \oplus \cdots \oplus x_i^k \mathbf{a}_k,\ i = 1, \dots, \ell$ are mutually independent* [7].

PROOF : Since all variables a_i^j are mutually independent, then it is enough to check that $x_i^1 a_1^1 \oplus \cdots \oplus x_i^k a_k^1,\ i = 1, \dots, \ell$ are mutually independent. Without loss of generality, it is possible to consider that $\ell = k$. We show that the probability of the intersection of events

$$\{x_i^1 a_1^1 \oplus \cdots \oplus x_i^k a_k^1 = y_i\}, \quad i = 1, \dots, k, \tag{9}$$

is equal to the product of probabilities, that is 2^{-k}; here $y_i = 0$ or $1,\ i = 1, \dots, k$. By solving the system of linear equations (9), with the method of elimination of variables we obtain the equivalent event to (9) $\cup_{1 \le i \le k} \{a_i^k = \tilde{y}_i\}$ (where $\tilde{y}_i = 0$ or $1,\ i = 1, \dots, k$) which evidently has probability 2^{-k}, and this prooves lemma 2. ■

[7] Hereafter we say that the vectors $\mathbf{b}_i,\ i = 1, \dots, \ell$ are mutually independent if the totality of their coordinates $b_i^j,\ i = 1, \dots, \ell;\ j = 1, \dots, n.$ are mutually independent

Lemma 3 *Let $\mathbf{b}_1,\dots,\mathbf{b}_\ell$; $\ell \geq 2$, be independent random vectors from V^n. Then uniformly for $t \leq \tau_n$ and $\ell \leq n$*

$$P\{w(\mathbf{b}_1) \leq t,\dots,w(\mathbf{b}_\ell) \leq t, w(\mathbf{b}_1 \oplus \dots \oplus \mathbf{b}_\ell) \leq t\} = O(2^{-\epsilon_1 n}) \left[\sum_{i\leq t} C_n^i 2^{-n}\right]^\ell, \quad (10)$$

where $\epsilon_1 > 0$ does not depend on ℓ.

Proof : We denote by a_ℓ the probability (10), and by γ_ℓ the probability of the same event as in (10) too, but by considering that the components of the vectors $\mathbf{b}_1,\dots,\mathbf{b}_\ell$ are in essence independent random variables taken from the values 0 and 1 with probabilities p_n and $1-p_n$ respectively; here $p_n = (t_n/n)$. Every elementary solution constrained by the vectors $\mathbf{b}_1,\dots,\mathbf{b}_\ell$ is described by a binary $\ell \times n$ matrix. The number of such matrices answering the event in (10) is equal to $a_\ell 2^{\ell n}$, moreover any such matrix has no more than $t\ell$ ones. Thus

$$\gamma_\ell \geq a_\ell 2^{\ell n} p_n^{t\ell}(1-p_n)^{n\ell - t\ell}, \quad (11)$$

if only $p_n \leq 1/2$. Clearly,

$$\gamma_\ell \leq \widehat{P}\{w(\mathbf{b}_1 \oplus \dots \oplus \mathbf{b}_\ell) \leq t\},$$

where the mark $\widehat{\ }$ over P points out that the components of the vectors $\mathbf{b}_1,\dots,\mathbf{b}_\ell$ are taken according to the new distribution. The vector $\mathbf{b}_1 \oplus \dots \oplus \mathbf{b}_\ell$ has independent coordinates each of which has value 1 with probability δ_ℓ, obeying to the recurrence equation

$$\delta_\ell = \delta_{\ell-1}(1-p_n) + (1-\delta_{\ell-1})p_n, \quad \delta_1 = p_n.$$

From here it is easy to deduce, that δ_ℓ increases monotonically with ℓ (towards 1/2). Therefore, under $t \leq \delta_2 n$

$$\gamma_\ell \leq \sum_{i\leq t} C_n^i \delta_2^i (1-\delta_2)^{n-i}, \quad \ell \geq 2. \quad (12)$$

But $\delta_2 = 2p_n(1-p_n) \to 2p(1-p) > p$, while $\tau_n/n \to p$. Therefore, the right term in (12) is $O(2^{-\epsilon_1 n})$, uniformly for $t \leq \tau_n$. By gathering together (11) and (12) and by noticing that uniformly for $t \leq \tau_n$,

$$\sum_{i\leq \ell} C_n^i = 2^{o(n)}[p_n^t(1-p_n)^{n-t}]^{-1},$$

we arrive to the requested result. ■

Lemma 4 *For S_r^t, defined by* (3) *with $G_\nu = \overline{D}_\nu^t$, $N = 2^k - 1$, $r \leq 2m+1$, $m = O(\sqrt{n})$, uniformly for $t \leq \tau_n$ we have the relation*

$$S_r^t = \frac{[u_n(t)]^r}{r!} + O(2^{-\epsilon_2 n}) \sum_{i\leq r} \frac{[u_n(t)]^i}{i!}. \quad (13)$$

PROOF : To simplify the demonstration, we identify the indexes ν_j in the sum (5) for S_r^t with the k-dimensional binary vectors corresponding to the k-size binary writing of the numbers ν_j. We will divide the sum (5) into $r - [\log_2 r]$ terms according to the number of vectors in the linearly independent maximal subsystem from $\nu_1, \ldots, \nu_r$ (such systems contains no less than $[\log_2 r]+1$ vectors, since $\nu_1, \ldots, \nu_r$ are distinct). For the sum $\Sigma^{(r)}$, by solving the case of the linear independence of the vectors $\nu_1, \ldots, \nu_r$ and by using the independence of the events $\overline{D}_{\nu_1}^t, \ldots, \overline{D}_{\nu_r}^t$ (lemma 2), we have

$$\Sigma^{(r)} P\{\overline{D}_{\nu_1}^t, \ldots, \overline{D}_{\nu_r}^t\} = \frac{[u_n(t)]^r}{r!} \prod_{0 \leq j \leq r-1} (1 - 2^{-k+j}).^8, \tag{14}$$

Suppose now that the linearly independent maximal system consists of $\ell < r$ vectors, and for convenience we suppose that $\nu_1, \ldots, \nu_\ell$ are independent. Then

$$\begin{aligned} &P\{\overline{D}_{\nu_1}^t, \ldots, \overline{D}_{\nu_r}^t\} \leq P\{\overline{D}_{\nu_1}^t, \ldots, \overline{D}_{\nu_\ell}^t \overline{D}_{\nu_{\ell+1}}^t\} = \\ &= P\{w(\mathbf{b}_1) \leq t, \ldots, w(\mathbf{b}_\ell) \leq t, w(x^1\mathbf{b}_1 \oplus \cdots \oplus x^\ell \mathbf{b}_\ell) \leq t\}, \end{aligned}$$

where the vectors $\mathbf{b}_1, \ldots, \mathbf{b}_\ell$ are mutually independent, and x^i is equal to 0 or 1, while the number of x^i equal to 1 is greater or equal to 2. We upper bound the number of items in the sum $\Sigma^{(\ell)}$ by the quantity

$$\frac{1}{\ell!} \prod_{0 \leq i \leq \ell-1} (2^k - 2^i) 2^{\ell(r-\ell)} r! = \frac{1}{\ell!} 2^{k\ell} O(2^{r^2}) = \frac{1}{\ell!} 2^{k\ell} O(2^{m^2}).$$

By choosing $m = O(\sqrt{n})$ such that $m^2 \leq \frac{\epsilon_1}{2} n$, and by applying lemma 3, we arrive to the following estimate for $\Sigma^{(\ell)}$:

$$\Sigma^{(\ell)} P\{\overline{D}_{\nu_1}^t, \ldots, \overline{D}_{\nu_r}^t\} = O(2^{-\epsilon_2 n}) \frac{[u_n(t)]^\ell}{\ell!}, \quad \epsilon_2 = \frac{\epsilon_1}{2}. \tag{15}$$

By summing (14) and (15) under $\ell = 1, 2, \ldots, r-1$, we arrive at (13). ■

The author expresses his sincere thanks to A. N. Kolmogorov for the formulation of the problem and the guidance.

References

[1] L. A. Bassalygo. *Some problems in coding theory.* PhD thesis, Moscow State University, 1967. in Russian.

[2] V. Feller. *An introduction to probability theory and applications.* 1964.

[3] V. N. Kochelev. *Problems of Information Transmission,* volume 1. 1965.

[4] W. Peterson. *Error-Correcting codes.* 1964.

[5] R.Gallager. *Low-Density Parity-Check Codes.* 1966.

[6] D. Slepian.

[8] $\prod_{0 \leq j \leq r-1} (2^k - 2^j)$ represents exactly the number of $r \times k$ matrices of rank k (see [6])