DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.
DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.
Binomial moments of the distance distribution and the probability of undetected error

A. Barg *  A. Ashikhmin †

Abstract

In [1] K. A. S. Abdel-Ghaffar derives a lower bound on the probability of undetected error for unrestricted codes. The proof relies implicitly on the binomial moments of the distance distribution of the code. We use the fact that these moments count the size of subcodes of the code to give a very simple proof of the bound in [1] by showing that it is essentially equivalent to the Singleton bound. We discuss some combinatorial connections revealed by this proof. We also discuss some improvements of this bound.

Finally, we analyze asymptotics. We show that an upper bound on the undetected error exponent that corresponds to the bound of [1] improves known bounds on this function.

1 Introduction

Let $C$ be a code of size $M$ over a $q$-ary alphabet $F_q$. Suppose $C$ is used for transmission over a $q$-ary symmetric channel with crossover probability $\epsilon < \frac{q-1}{q}$. Let $P(C, \epsilon)$ be the probability that a transmitted codeword $c \in C$ is corrupted in the channel by an error vector $e$ so that $c + e$ is also a codeword in $C$. Such an error cannot be detected by the decoder; therefore, the quantity $P(C, \epsilon)$ is called the probability of undetected error. The set of $n+1$ numbers

$$A_w = \frac{1}{M} \sum_{c \in C} | \{ c' \in C : \text{dist}(c, c') = w \}|$$

is called the (average) distance distribution of $C$. Then clearly,

$$P(C, \epsilon) = \sum_{w=1}^{n} A_w \left( \frac{\epsilon}{q-1} \right)^w (1 - \epsilon)^{n-w}$$

$$= \bar{A} \left( 1 - \epsilon, \frac{\epsilon}{q-1} \right),$$

where $\bar{A}(x, y) = \sum_{w=1}^{n} A_w x^{n-w} y^w$ is the distance function of $C^1$. The problem is to derive lower bounds on the function

$$P(q, n, M, \epsilon) = \min_{q\text{-ary codes of size } M} P(C, \epsilon).$$

---

*Bell Laboratories, Lucent Technologies, 600 Mountain Avenue 2C-375, Murray Hill, NJ 07974.
†Los Alamos National Laboratories.

1"we write $\bar{A}$ because the sum in the standard definition of the distance function $A$ includes the term $w = 0$. 
In [1] K.A.S. Abdel-Ghaffar derives a lower bound on this quantity. We observe that his proof relies on binomial moments of the distance distribution of codes. This enables us to give a very simple proof of Abdel-Ghaffar’s theorem. Though the approach discussed below applies to general codes, we first focus on the linear case because it is especially intuitive and reveals useful combinatorial insights, see Sect. 2. The general case is treated in Section 3. Section 4 deals some improvements of this bound and their asymptotic analysis. In particular, we give an improvement of known upper bounds on the exponent of the probability of undetected error.

2 The linear case

Let $C$ be an $[n,k]$ linear $q$-ary code\(^2\). The following result is proved in [1] (Theorem 2 and Eq. (12)).

**Theorem 1** Let $0 \leq \epsilon \leq (q-1)/q$. Let $C$ be a linear code of size $q^k$ with no all-zero coordinates.

Then

$$P(q, n, q^k, \epsilon) \geq \sum_{w=n-k+1}^{n} \binom{n}{w}(q^k+n-w-1)\left(\frac{\epsilon}{q-1}\right)^w \left(1 - \frac{q^\epsilon}{q-1}\right)^{n-w}. \quad (4)$$

This theorem improves some previously known bounds on the probability of undetected error and, in particular, a bound for linear codes from [2], [3].

A. Proof of Theorem 1. Our observation is that in the linear case this theorem is equivalent to the obvious fact that any submatrix of the parity-check matrix of an $[n,k]$ linear code $C$ has rank at most $n - k$. We begin with reformulating this fact. Let $N = \{1,2,\ldots,n\}$ and let $H$ be a parity-check matrix of $C$ with columns $(h_e, e \in N)$. For any $E \subseteq N$ let $C(E)$ be the linear subcode of $C$ with zeros outside $E$. Let $H(E) = (h_e, e \in E)$.

**Lemma 2** Let $E \subseteq N$. Then $|C(E)| \geq \max(1, q^{w-(n-k)})$.

**Proof:** Indeed,

$$\dim C(E) = |E| - \text{rank } H(E). \quad (5)$$

where the last term denotes the number of linearly independent columns with numbers in $E$. 

If the bound in this lemma holds for every $E$ with equality, then $C$ is an MDS code. Since bound (4) is equivalent to this lemma, it holds with equality if and only if $C$ is an MDS code.

**Proof of Theorem 1:** Following [5], let us define numbers $B_w$ from the expansion

$$\tilde{A}(x) = \sum_{w=1}^{n} A_w x^{n-w} y^w = \sum_{w=1}^{n} B_w (x-y)^{n-w} y^w.$$ 

Equating coefficients of these generating functions, we obtain relations between $A_w$ and $B_w$:

$$B_w = \sum_{j=1}^{w} \binom{n-j}{n-w} A_j, \quad A_w = \sum_{j=1}^{w} (-1)^j \binom{n-j}{n-w} B_j, \quad 0 \leq w \leq n. \quad (6)$$

\(^2\)In this case $F_q$ is the field of $q$ elements.
Clearly, $B_w \geq 0$. Using this definition, rewrite (2) as follows:

$$P(C, \epsilon) = \sum_{w=1}^{n} B_w \left( 1 - \frac{\epsilon q}{q-1} \right)^{n-w} \left( \frac{\epsilon}{q-1} \right)^w.$$ \tag{7}

A second ingredient of the proof is the following

**Lemma 3** [4],[5] We have

$$B_w = \sum_{E \subseteq C} (|C| - 1),$$

where the summation is over all $w$-subsets of $N$.

**Proof:** This is a double counting argument. Let $E \subseteq N$, $|E| = w$. Let $c \in C \setminus \{0\}$, $\text{wt}(c) = j$, $\text{supp}(c) \subseteq E$. We can choose such a subset $E$ in $\binom{n-j}{w}$ different ways, or

$$\sum_{E \subseteq C} (|C| - 1) = \sum_{j=1}^{w} \binom{n-j}{n-w} A_j.$$ \tag{8}

The lemma now follows from (6).

Now let us use (7) to complete the proof of the theorem. Namely, by Lemma 2 we have

$$B_w \geq \binom{n}{w} \left( q^{-w} - (n-k-1) \right)$$

for $n-k+1 \leq w \leq n$. For $1 \leq w \leq n-k$ just replace $B_w$ by 0.

Definition (6) implies the following general property of the numbers $B_w$.

**Lemma 4** Let $1 \leq s < w \leq n$. Then

$$B_w \geq B_s \binom{n-s}{n-w} / \binom{w+1}{s-1}.$$  

**Proof:**

$$B_w = \sum_{i=1}^{w} \binom{n-i}{n-w} A_i \geq \sum_{i=1}^{w-1} \binom{n-i}{n-(w-1)} \frac{n-(w-1)}{w-1-i-1} A_i$$

$$\geq \frac{n-w+1}{w-1} B_{w-1} \geq \cdots \geq \frac{(n-w+1)(n-w+2)\ldots(n-s)}{(w-1)(w-2)\ldots s} B_s$$

$$= \frac{(n-s)(n-s-1)\ldots(n-s-(w-s)+1)}{(w-1)\ldots(w-1-(w-s)+1)} B_s = B_s \binom{n-s}{n-w} / \binom{w-1}{s-1}.$$  

**B. AN IMPROVEMENT FOR THE BINARY CASE.** Our proof shows that the linear case treated by Theorem 1 is essentially nothing else than a reformulation of Lemma 2. It is much simpler than the general case in [1]. However, below we argue that our approach yields the result of [1] in full generality.

Let us make some further remarks on the linear case. From Lemma 2 and the proof of Theorem 1 we see that the bound can be improved if we can get better estimates on the dimension of subcodes. As remarked above, the equality in (4) is attained on MDS codes. In the binary case the only existing linear MDS codes are codes with parameters $[n,k]$, where $k \in \mathcal{K} = \ldots$.
{\{n - 1, 1, n, 0\}}. Each of these codes is unique, and therefore, the value of \( P(2, n, 2^k, \epsilon) \) for \( k \in K \) is known exactly. Otherwise, the estimate on the numbers \( B_w \) is related to the difficult problem of estimating from above the number of subsets \( E, |E| = n - k \), of linearly independent columns in \( H \), i.e., the number of information sets in the dual vcode \( C^\perp \). However, we can observe that in the binary case the bound on \( P(2, n, 2^k, \epsilon) \) for \( k \notin K \) is strictly greater than the right-hand side of (4). Namely, Let \( C \) be a binary linear \([n, k, d]\) code, \( k \neq 0, 1, n, n - 1 \), and \( A_d \) the number of words of weight \( d \) in it. Then

\[
P(2, n, 2^k, \epsilon) \geq A_d (1 - 2\epsilon)^{n - d} + \sum_{w=n-k+1}^{n} \binom{n}{w}(2^{w-n+k} - 1)\epsilon^w(1 - 2\epsilon)^{n-w},
\]

since \( B_d = A_d \) and \( d \leq n - k \). Elaborating on this argument, we get the following.

**Proposition 5** Let \( C \) be a binary linear \([n, k]\) code, \( k \geq 2 \), and let \( d^* = \lceil \frac{n}{3}(n - k + 2) \rceil \). Then

\[
P(2, n, 2^k, \epsilon) \geq \sum_{w=d^*}^{n-k} \binom{n-d^*}{w} \epsilon^w(1 - 2\epsilon)^{n-w} + \sum_{w=n-k+1}^{n} \binom{n}{w}(2^{w-n+k} - 1)\epsilon^w(1 - 2\epsilon)^{n-w}.
\]

**Proof:** Let \( k \geq 2 \). By the Griesmer bound,

\[
n \geq \sum_{i=0}^{k-1} \left\lfloor \frac{d}{2^i} \right\rfloor \geq \frac{d}{2} + \frac{d}{2} + (k - 2) = \frac{3}{2}d + k - 2,
\]

where \( d \) is the distance of \( C \). This implies \( d \leq d^* := \lceil \frac{n}{3}(n - k + 2) \rceil \). Hence \( B_{d^*} \geq 1 \). Now apply Lemma 4 to estimate \( B_w \) for \( d^* \leq w \leq n - k \) and Lemma 2 to estimate them for \( n-k+1 \leq w \leq n \).

The comparison performed in [1] shows that this bound is in many cases the best known for binary linear codes. Any further information about a class of codes (estimates of the minimum distance, automorphisms, etc.) can further improve the bound for this class.

**C. COMBINATORIAL CONNECTIONS.** The numbers \( B_w \) are related to the binomial moments of the weight spectrum. They appear already in [4]. This paper observes that they also count the cumulative size of subcodes in the dual code of \( C \). Lemma 3 found in [5] gives a nice combinatorial interpretation to these numbers. Later they played a role in [6],[7],[8] and conceivably in many other works.

Since our approach in the linear case links \( P(C, \epsilon) \) and dependence relations among the columns of the parity-check matrix of \( C \), we may also directly express \( P(C, \epsilon) \) via the generating function of linearly (in)dependent subsets of columns in the dual code \( C^\perp \). Moreover, since these generating functions for \( C \) and \( C^\perp \) are related, we can also express \( P(C, \epsilon) \) via the rank distribution in \( C \) itself. Namely, let \( G \) be the generator matrix of \( C \). Let

\[
U_C(x, y) = \sum_{w=0}^{n} \sum_{\alpha=0}^{k} U_{w\alpha}^C x^w y^\alpha,
\]

where

\[
U_{w\alpha}^C = \{|E \subset N : |E| = w, \text{rank } G(E) = \alpha \}.
\]
By definition, let \( U_{00}^C = \delta_{0,0} \).

Let \( E \subseteq N, |E| = w \). The following relation is a version of the fundamental duality for the rank function in matroid theory \([14]\):

\[
k - \text{rank} G(N - E) = w - \text{rank} H(E).
\]

In the linear case it follows from an easy linear-algebraic argument \([8]\): Suppose \( E \) occupies the first part of \( N \) and represent \( G \) in the form \( \begin{bmatrix} A & 0 \\ B & D \end{bmatrix} \), where \( A \) has the maximal possible number of rows. As in Lemma 2, we have \( \dim C(E) = |E| - \text{rank} H(E) \). Also since \( A \) is maximal, \( \text{rk}(B|D) = \text{rk}(D) \), and therefore \( \dim C(E) = k - \text{rank} G(N - E) \).

This implies the following version of the MacWilliams identities:

\[
U_{w0}^C = U_{n-w,k-w+\alpha}^C,
\]

whence we get

\[
U_{C+}(x, y) = \sum_{w=0}^{n} \sum_{\alpha=0}^{n-k} U_{w0}^C x^w y^\alpha = \sum_{w=0}^{n} \sum_{\alpha=0}^{n-k} U_{n-w,k-w+\alpha}^C x^w y^\alpha
\]

\[
= x^n \sum_{w=0}^{n} \sum_{\alpha=0}^{n-k} U_{w,k-n+w+\alpha}^C x^w y^\alpha = x^n y^{n-k} \sum_{w=0}^{n} \sum_{\alpha=0}^{n-k} U_{w,w-\alpha}(xy)^w y^{w-\alpha}
\]

\[
= x^n y^{n-k} U_C(\frac{1}{xy}, y).
\]

The probability of undetected error for \( C \) is expressed via these functions as follows.

**Proposition 6** We have

\[
P(C, \epsilon) = \left( 1 - \frac{\epsilon q}{q - 1} \right)^n U_{C+}(\frac{\epsilon q}{q - 1 - \epsilon q}, \frac{1}{q}) - (1 - \epsilon)^n
\]

\[
= \frac{\epsilon^n}{(q - 1)^n} U_C(\frac{q - 1 - \epsilon q}{q}, \frac{1}{q}) - (1 - \epsilon)^n
\]

\[
= q^k (1 - \epsilon)^n P\left( C_1, \frac{q - 1 - \epsilon q}{q - q \epsilon} \right) + q^{k-n} - (1 - \epsilon)^n.
\]

**Proof:** By (5) and Lemma 2,

\[
B_w = \sum_{\alpha=0}^{n-k} (q^{w-\alpha} - 1) U_{w0}^C = \sum_{\alpha=0}^{n-k} q^{w-\alpha} U_{w0}^C - \binom{n}{w}.
\]

Therefore by (7),

\[
P(C, \epsilon) = \sum_{w=0}^{n-k} \sum_{\alpha=0}^{n-k} (q^{w-\alpha} - 1) U_{w0}^C \left( 1 - \frac{\epsilon q}{q - 1} \right)^n \frac{\epsilon}{q - 1}^w w
\]

\[
= \left( 1 - \frac{\epsilon q}{q - 1} \right)^n \left[ \sum_{w=0}^{n} \sum_{\alpha=0}^{n-k} U_{w0}^C \left( \frac{\epsilon q}{q - 1 - \epsilon q} \right)^w \frac{1}{q} \right] - \sum_{w=0}^{n} \binom{n}{w} \left( \frac{\epsilon}{q - 1 - \epsilon q} \right)^w
\]

\[
= \left( 1 - \frac{\epsilon q}{q - 1} \right)^n U_{C+}(\frac{\epsilon q}{q - 1 - \epsilon q}, \frac{1}{q}) - (1 - \epsilon)^n.
\]
This proves (11). The second equality is immediate from (10). To prove (13) we use (11) for \( C' \).

\[
P(C', \frac{q - 1 - q \epsilon}{q - q \epsilon}) = \left( 1 - \frac{q - 1 - q \epsilon}{q - q \epsilon} \right)^n U_C \left( \frac{q - 1 - q \epsilon}{q(1 - \epsilon)} \frac{q}{q - 1 - q \epsilon} \frac{1}{q(1 - \epsilon)} \right)
- \left( 1 - \frac{q - 1 - q \epsilon}{q(1 - \epsilon)} \right)^n = \frac{e^n}{(q - 1)^n (1 - \epsilon)^n} U_C \left( \frac{q - 1 - q \epsilon}{\epsilon} \frac{1}{q} \right) - \frac{1}{q^n (1 - \epsilon)^n}.
\]

Further,

\[
q^k (1 - \epsilon)^n P(C', \frac{q - 1 - q \epsilon}{q - q \epsilon}) + q^{k-n} - (1 - \epsilon)^n = \frac{q^k e^n}{(q - 1)^n} U_C \left( \frac{q - 1 - q \epsilon}{\epsilon} \frac{1}{q} \right) - (1 - \epsilon)^n,
\]

as claimed.

Note that the relation between \( P(C, \epsilon) \) and \( P(C', \cdot) \) has appeared in the literature (see Theorem 3.4.1 in [11]). The other parts of this proposition are new.

While the function \( U_C(x, y) \) is a natural object from the coding-theoretic point of view, in combinatorics one studies different, though closely related polynomials, namely the Whitney corank-nullity function and the Tutte polynomial of the code (or of the matroid \( M(C) \) associated with it). These functions play an important role in combinatorics of incidence geometries, see [13], [14]. More on this link in the linear case is found in [8], [10]. In particular, [10] puts forward the problem of estimating the numbers \( U^{\alpha}_{C, \alpha} \). By the above proposition, this would yield new estimates of \( P(C, \epsilon) \).

The Whitney function, by definition, is a polynomial \( R_C(x, y) = \sum_{u=0}^n \sum_{v=0}^k R_{uv} x^u y^v \), where \( R_{uv} := \left\{ E \subseteq N : k - \text{rank} G(E) = u, |E| - \text{rank} G(E) = v \right\} \). The Tutte polynomial \( T_C(x, y) \) can be formally defined as \( T_C(x, y) := R_C(x - 1, y - 1) \). The following proposition is a variation of Greene’s result [9].

**Proposition 7**

\[
P(C, \epsilon) = \frac{e^n}{(q - 1)^n} U_C \left( \frac{q - 1 - q \epsilon}{q - 1 - q \epsilon} \right) - (1 - \epsilon)^n.
\]

**Proof:** Let \( A(z) := A(1, z) = \sum_{i=0}^n A_i z^i \) be another form of the weight enumerator of a code \( C \). The result of [9] states that

\[
A(z) = (1 - z)^k z^{n-k} T_C \left( \frac{1 + (q - 1)z}{1 - z}, \frac{1}{z} \right).
\]

Substitution in (2) completes the proof.

The final topic of this section is a relation of \( P(C, \epsilon) \) and **support weight distributions** of \( C \). The set of support weight distributions is an extension of the notion of the weight distribution of a code, which is related to the functions studied in this section.

\[3\] For reader’s convenience we note that in [10] the numbers \( U^{\alpha}_{C, \alpha} \) and \( \alpha^{\alpha}_{C, \alpha} \) are denoted by \( e_{C, \alpha}(C) \) and \( m_{C, \alpha}(C) \), respectively. The notation for \( U^{\alpha}_{C, \alpha} \) in [7] is \( N^\alpha_{\alpha} \).
Let \( C \subseteq C \) be a subcode of \( C \). The subset \( E = \bigcup_{c \in \text{supp}(c)} \) is called the support of \( C \), denoted \( \text{supp} C \). The \( m \)th support weight distribution is the set of numbers \( A_{0}, A_{1}, \ldots, A_{n} \), where
\[
A_{u}^{m} = \{|C \subseteq C : |\text{supp} C| = u, \dim C = m\}.
\]
Support weight distributions have received much attention in the literature, see [15], [7], [8], [10] and references therein. J. Simonis [7, Lemma 1] proved the following identity which is a natural extension of Lemma 3:
\[
\sum_{u=0}^{n-w} \binom{n-u}{w} A_{u}^{k-m} = \sum_{\alpha=0}^{m} \binom{k-\alpha}{k-m} U_{\alpha}^{C}, \quad 0 \leq w \leq n, 0 \leq m \leq k,
\]
where \( \binom{\cdot}{\cdot} \) is a \( q \)-binomial coefficient (Gaussian number). In order to express the numbers \( B_{w} \) (hence, \( P(C, \epsilon) \)) via the support weight distributions of \( C \), we need to solve this set of equations with respect to the numbers \( U_{\alpha}^{C} \). For this we use the orthogonality relation of \( q \)-binomial coefficients [12]
\[
\sum_{u=0}^{n} (-1)^{u-i} q^{u-i} \binom{\epsilon}{\epsilon} \binom{u}{i} \binom{j}{u} = \delta_{ij}.
\]
One obtains
\[
U_{\alpha}^{C} = \sum_{j=0}^{\alpha} (-1)^{j} q^{j(j-1)/2} \binom{k-\alpha+j}{j} \sum_{i=0}^{n} A_{i}^{k-\alpha+j} \binom{n-i}{w}, \quad 0 \leq w \leq n, 0 \leq m \leq k.
\]
Using this in (12), we can express \( P(C, \epsilon) \) in terms of support weight distribution of \( C \) (and, consequently, of \( C^{\perp} \)).

3 The general case

Let us consider unrestricted codes. Our proofs of properties of binomial moments (Lemma 2 and Lemma 3) depend on linearity. However, binomial moments of the distance distribution of a nonlinear code still retain the properties claimed in these lemmas, and the linearity turns out to be irrelevant!

We need two definitions.

\textbf{Definition.} Let \( A \subseteq F_{q}^{m} \) be a set of vectors. Define the support of \( A \) as follows:
\[
\text{supp} A = \{e \in N : \exists_{e, e' \in A} (c \neq c')\}.
\]

\textbf{Definition.} Let \( C \subseteq F_{q}^{m} \) be a code and \( E \subseteq N \). By \( C(E) \) we denote a subcode of \( C \) such that (a), \( \text{supp} C(E) \subseteq E \), and (b), \( \forall c \in C, c \notin C(E) : \text{supp} (C(E) \cup c) \nsubseteq E \).

Thus, \( C(E) \) are maximal subcodes of \( C \) with support in \( E \).
Example. Let $C$ be a code of 6 words:

\[
\begin{align*}
  c_1 &= 000 \\
  c_2 &= 101 \\
  c_3 &= 111 \\
  c_4 &= 010 \\
  c_5 &= 100 \\
  c_6 &= 110,
\end{align*}
\]

and let $E = \{1\}$. Then there are 4 subcodes $C(E) = C_i(E)$, namely $(c_1, c_6), (c_2), (c_3), (c_4, c_6)$.

Generally for a given subset $E$, $|E| = w$, there are $q^{n-w}$ subcodes $C_i(E)$, and $\sum_i |C_i(E)| = |C|$.

If $C$ is linear, we get back to our original definition of $C(E)$.

For the rest of this section we assume that $\supp C = N$.

Let us prove an analog of Lemma 3, which is essentially Lemma 1 in [1], in different notation.

**Lemma 8** Let $A_w$ be as defined in (1). Then

\[
B_w := \sum_{j=1}^{w} \binom{n-j}{n-w} A_j = \frac{1}{|C|} \sum_{E \in (N \choose w)} \sum_{i=1}^{q^{n-w}} |C_i(E)|(|C_i(E)| - 1).
\]

**Proof:** Count in two ways the size of the set

\[
\{(E, (c, c')) | E \in \binom{N}{w}; c, c' \in C, \supp (c, c') \subseteq E, 1 \leq \dist (c, c') \leq w\}.
\]

It is given by the following sum:

\[
\sum_{j=1}^{w} \sum_{c \in C} \sum_{E \in \binom{N}{w}} \sum_{\substack{c' \in C \\supp (c, c') \subseteq E \\dist (c, c') = j}} 1.
\]

This sum is equal to $|C|$ times the lefthandside of the claimed equality. Interchanging the first and the third sums, we observe that it is also equal to the righthandside.

To use this lemma in the expression (7), we need to estimate the sum on the righthandside.

We get the following analog of Lemma 2.

**Lemma 9** Suppose that $\supp (C) = n$. Let $E \in \binom{N}{w}$. Then

\[
\sum_{i=1}^{q^{n-w}} |C_i(E)|(|C_i(E)| - 1) \geq \max\{|C|, |C||q^{w-n}|(|C| - 1)\}.
\]

**Proof:** As remarked above, $\sum_i |C_i(E)| = |C|$. For $|C| \leq q^{n-w}$ estimate each $|C_i(E)|$ by 1. For $|C| > q^{n-w}$, we have

\[
\sum_i |C_i(E)|(|C_i(E)| - 1) = \sum_i |C_i(E)|^2 - |C|.
\]
Using Lagrange multipliers, we see that the sum \( \sum |C_i(E)|^2 \) attains the minimum when all numbers \( |C_i(E)| \) are the same (and equal \( |C|q^{w-n} \)). Hence, we get
\[
\sum_i |C_i(E)||(|C_i(E)| - 1) \geq |C|^2q^{w-n} - |C|.
\]

These two lemmas again lead to the bound of the form (4). We quote this result for future use.

**Proposition 10** Let \( P(q, n, M, \epsilon) \) be defined by (3). Then
\[
P(q, n, M, \epsilon) \geq \sum_{w=n-k+1}^{n} \binom{n}{w} (Mq^{w-n} - 1) \left( \frac{\epsilon}{q - 1} \right)^w \left( 1 - \frac{q\epsilon}{q - 1} \right)^{n-w}.
\]  

The bound in [1] is a little sharper than this because the estimate in Lemma 9 can be improved by using integer parts.

Though it may not be clear at the first glance, linearity is irrelevant to Proposition 6 and the ensuing discussion as well. Its general analog is proved upon defining the average rank function of any code \( C \). We believe that the details can be left to an interested reader.

### 4 Asymptotics

In this section we confine ourselves to binary codes. Let us analyze the asymptotic case. Our source for previous results is [16], which gives a comprehensive overview of the asymptotic bounds for the probability of undetected error. Let \( n \to \infty, M = 2^{Rn}, 0 \leq R \leq 1 \). All logarithms below are base 2. The exponent of the probability of undetected error is defined as
\[
E(n, R, \epsilon) = \left( - \frac{1}{n} \log P(2, n, [2^{Rn}], \epsilon) \right),
\]
\[
E(R, \epsilon) = \lim_{n \to \infty} (n, R, \epsilon)
\]
provided that this limit exists. We are interested in lower bounds on \( \lim \inf_{n \to \infty} E(n, R, \epsilon) \) and upper bounds on \( \lim \sup_{n \to \infty} E(n, R, \epsilon) \). By abuse of notation, below in both cases we write \( E(R, \epsilon) \).

The following theorem is an easy corollary the bound (14).

**Theorem 11**

\[
E(R, \epsilon) \leq \begin{cases} -H(R) - (1 - R) \log \epsilon - R \log(1 - 2\epsilon), & 0 \leq R \leq 1 - 2\epsilon, \\ 1 - R, & 1 - 2\epsilon \leq R \leq 1. \end{cases} \tag{N}
\]

**Proof:** Let us write out the generic term in the sum in (14) in the exponential form, putting \( w = nw \). We get in the exponent
\[
-n[-H(\omega) - (\omega - 1 + R) - \omega \log \epsilon - (1 - \omega) \log(1 - 2\epsilon)].
\]

The minimum of this is attained for \( \omega = 2\epsilon \). Hence this expression gives an upper bound on \( E(R, \epsilon) \) if the corresponding term is within the summation range, or if \( 2\epsilon > 1 - R \). Substituting
\( \omega = 2\epsilon \), we get \( 1 - R \) in the exponent. If \( 2\epsilon \leq 1 - R \), the maximum in the sum is attained for \( \omega = 1 - R \).

Let us quote the known results with which this result should be compared. References can be looked up in [16]. Let

\[
D(u, v) = -u \log v - (1 - u) \log(1 - v), \quad H(u) = -u \log u - (1 - u) \log(1 - u).
\]

**Theorem 12** Let \( \delta(R) = H^{-1}(1 - R) \). Then

\[
E(R, \epsilon) \geq \begin{cases} 
D(\delta(R), \epsilon), & 0 \leq R \leq 1 - H(\epsilon) \\
1 - R, & 1 - H(\epsilon) \leq R \leq 1.
\end{cases} \quad (L1)
\]

Let \( p(R) \) be any upper bound on the distance of codes of rate \( R \) such that \( p''(R) \geq 0 \) and let \( R_0 = R_0(\epsilon) \) be defined by \( p'(R_0) = (\log(\epsilon/(1 - \epsilon)))^{-1} \). Let \( \tau(R) = -R/ \log(2^{1-R} - 1) \). Then

\[
E(R, \epsilon) \leq \begin{cases} 
D(p(R), \epsilon), & 0 \leq R \leq R_0, \\
D(p(R_0), \epsilon) + R_0 - R, & R_0 \leq R \leq 1, \\
D(\tau(R), \epsilon) - R, & 0 \leq R \leq \log(2 - 2\epsilon), \\
1 - R, & \log(2 - 2\epsilon) \leq R \leq 1.
\end{cases} \quad (L2, L3, L4, LL)
\]

Bound (K) is due to Korzhik. Bounds (L1)-(L4) were proved by Levenshtein. Bound (LL) is due to Leontiev and Levenshtein.

Some remarks may be in order. First, by (K) and (LL) the function \( E(R, \epsilon) \) is known exactly for \( \log(2 - 2\epsilon) \leq R \leq 1 \). Next, the “minimum-distance” bound (L2) is an analog of the minimum-distance bound of Shannon, Gallager and Berlekamp. The straight-line bound (L3) [16] is a segment with slope \(-1\) that can be drawn from any point of the minimum-distance bound. The point \( R_0 \) is chosen for this segment to be tangent to (L2) and therefore is optimal.

These bounds should be compared to the bound of Theorem 11. First, this theorem shows that the interval on which the function \( E(R, \epsilon) \) is known exactly (and equals \( 1 - R \)) is extended to \( 1 - 2\epsilon \leq R \leq 1 \). Next, bound (N) from this theorem is better than the bound (LL) for a certain interval of rates.

Bounds from both theorems are drawn in Fig. 1, where for \( p(R) \) we have chosen the upper bound

\[
p(R) = \frac{1}{2} - \sqrt{H^{-1}(R)(1 - H^{-1}(R))},
\]

which is the best presently known for small code rates. The straight-line bound (L3) still is the best in a certain interval of rates \( R > R_0(\epsilon) \). Finally, the best upper bound currently known is given in the following proposition.

**Proposition 13** Let \( p(R) \) be defined by (15) and let \( R_0 \) be the root of \( p'(R) = (\log(\epsilon/(1 - \epsilon)))^{-1} \).

\[
E(R, \epsilon) = \begin{cases} 
D(p(R), \epsilon), & 0 \leq R \leq R_0, \\
\min\{D(p(R_0), \epsilon) + R_0 - R, -H(R) - (1 - R) \log \epsilon - R \log(1 - 2\epsilon)\}, & R_0 \leq R \leq 1 - 2\epsilon, \\
1 - R, & 1 - 2\epsilon \leq R \leq 1.
\end{cases}
\]
We conclude with the following quote from [16]: “Surprisingly, the question whether the bound $E(R, \epsilon) \leq 1 - R$ is tight for all rates $R$, $1 - H(\epsilon) \leq R \leq 1$, is still open.”

References


Figure 1: Bounds on $E(R,\epsilon)$, $\epsilon = 0.2$. The shaded areas show the segment on which $E(R,\epsilon) = 1 - R$. Straight-line bound $L3$ (dotted) is still slightly better than the other bounds. LL bound is improved everywhere.