LINEAR INTERSECTING CODES

by

G. Cohen* and A. Lempel**

Technical Report #269
March 1983

** Computer Science Department, Technion-IIT, Haifa, Israel
ABSTRACT

We study pairs of binary linear codes $C_1(n, nR_1), C_2(n, nR_2)$ with the property that for any nonzero $c_1 \in C_1$ and $c_2 \in C_2$, there is a coordinate in which both $c_1$ and $c_2$ are nonzero. The main results, including bounds for the achievable rates, pertain to the special case where $C_1 = C_2$. 
1. INTRODUCTION

Let $C$ be a binary linear $(n,k,d)$ code, i.e., a $k$-dimensional subspace of $\mathbb{F}_2^n$ with minimum Hamming distance $d$. That is, for any $c_i, c_j \in C$, $d(c_i, c_j)$, defined as the number of coordinates in which $c_i$ and $c_j$ differ, is at least $d$. Equivalently, $|c| = d(c,0) \geq d$ for all nonzero code words $c$. For $a = (a_1,a_2,\ldots,a_n)$ and $b = (b_1,b_2,\ldots,b_n)$ in $\mathbb{F}_2^n$ define their termwise product as $a \cdot b = (a_1b_1,a_2b_2,\ldots,a_nb_n)$.

Definition 1: Two codes $C_1$ and $C_2$ are $r$-intersecting if for any nonzero $c_1 \in C_1$ and $c_2 \in C_2$ we have $|c_1 \cdot c_2| \geq r$. We write this as $C_1 \cap C_2$; if $C_1 = C_2 = C$ we say that $C$ is an $(n,k,d,r)$, or $r$-intersecting code; if $r=1$ we say "intersecting" and write $C_1 \cap C_2$.

Example 1: $C = \{(000),(110),(101),(011)\}$ is a $(3,2,2,1)$ code.

2. SOME PROPERTIES OF INTERSECTING CODES

Let $d_{\text{Max}} = \max|c|$, $c \in C$, and let $I_0$ be the universal relation. Then, from $d(a,b) = |a| + |b| - 2|a \cdot b|$ the following property is obtained.

(P1) For every $C(n,k,d)$ we have $C \cap C$, with

$$\max(d - \frac{d_{\text{Max}}}{2},0) \leq r \leq \frac{d}{2}.$$  

Deleting the last component from every codeword in $C(n,k,d)$, $d > 1$, one obtains a shortened code $C'(n' = n-1, k' = k, d' \geq d-1)$.

(P2) If $C_1 \cap C_2$ then $C_1 \cap C_2'$ with $r' \geq r-1$. 

Technion - Computer Science Department - Technical Report CS0269 - 1983
Proposition 1 (Lempel-Winogràd): If $C_1(n,k_1,d_1)$ and $C_2(n,k_2,d_2)$ are intersecting then $d_1 \geq k_2$ and $d_2 \geq k_1$.

This was proved in [2]. We give here a proof extending a proposition in [1]. Let $c_2 \in C_2$ with $|c_2| = d_2$. Consider the set $S = \{a \cdot c_2, a \in C_1\}$. The mapping $f$ from $C_1$, $f: a \rightarrow a \cdot c_2$, is linear and injective because the existence of a nonzero $c_1$ in $\text{Ker } f$ would give $c_1 \cdot c_2 = 0$ contradicting $C_1 \cap C_2$. Hence $|S| = |C_1| = 2^{k_1}$. But every $s \in S$ can be identified with a subset of $\text{supp}(c_2) = \{i, c_{2i} = 1\}$. Noting that $|\text{supp}(c_2)| = d_2$, this gives $2^{k_1} \leq 2^{d_2}$.

The proof of $d_1 \geq k_2$ is similar.

Corollary 1: If $C_1$ and $C_2$ are $r$-intersecting, $r \geq 1$, then

\[ d_1 \geq k_2 + r - 1 \]
\[ d_2 \geq k_1 + r - 1. \]

Proof: Apply Proposition 1 to $C_1^{(r-1)}$ and $C_2^{(r-1)}$, obtained from $C_1$ and $C_2$, respectively, by shortening $r-1$ times, and which, by (P2), are intersecting.

Example 2: Let $C$ be the $(15,6,6)$ code consisting of the even weight codewords of the $(15,7,5)$ double error correcting BCH code [3]. The nonzero weights of $C$ are $|c| = 6, 8, 10$, and, by (P1), $C$ is intersecting. By Corollary 1, it is not 2-intersecting, because $d = k$.

Example 3: For a simplex code $C(2^m-1,m,2^{m-1})$ both bounds in (P1) coincide and $C$ is $2^{m-2}$-intersecting.
3. SOME CONSTRUCTIONS

One can readily verify the following proposition.

**Proposition 2:** From an \((n,k,d,r)\) code with generator matrix \(G\), one can construct an \((n+k+1,k+1,d+1,r)\) code with the following generator matrix. \((U_k\) denotes the unit matrix of order \(k\).)

\[
\begin{pmatrix}
G \\
U_k
\end{pmatrix}
\]

\[
\begin{pmatrix}
0 & 0 & 0 & \cdots & 0 & 1 & 1 & \cdots & 1
\end{pmatrix}
\]

**Example 4:** Starting from the \((15,6,6)\) code of Example 2, one obtains, successively, the \((22,7,7)\), \((30,8,8)\), \((39,9,9)\), etc., intersecting codes.

**Proposition 3:** Consider the three codes \(C_1(n_1,k_1,d_1)\), \(C_2(n_1,k_2 = k_1 + b,d_2,r_2)\) and \(C_3(n_3,b,d_3,r_3)\), where \(C_2\) is the union of \(2^b\) disjoint cosets of \(C_1\). Then there exists an \((n_1 + n_3,k_2,\min\{d_1,d_2 + d_3\}, r_2 + r_3)\) code.

**Proof:** cf. Theorem 9 in Ch. 10 of [3] for the construction of the code. The intersection property is immediate.

**Example 5:** Take \(C_1(15,4,8)\) as in Example 3, \(C_2(15,6,6,1)\) of Example 2, and \(C_3(3,2,2,1)\) of Example 1. This gives an \((18,6,8,2)\) code.
4. NON CONSTRUCTIVE BOUNDS

Choose "randomly" an \((n,k,d)\) code \(C\). For any two codewords \(c_1\) and \(c_2\) in \(C\) the probability that \(c_1 \neq c_2\) is \(1 - (3/4)^n\).

Hence the probability that \(C\) is intersecting is

\[ P_1 = \left( 1 - \left( \frac{3}{4} \right)^n \right) \binom{2^k}{2} . \]

The probability that a randomly chosen subset of \(2^k\) elements of \(\mathbb{F}_2^n\) be linear is

\[ P_2 = \prod_{i=0}^{k-1} \left( \frac{2^n}{2^i} \right) \left( \frac{2^n - 2}{2^i} \frac{2^n}{2^{k-i}} \right)^{-1} . \]

Hence there exists an \((n,k)\) intersecting code if the following condition holds:

\[ \left( \frac{2^n}{2^k} \right) P_1 P_2 > 1. \]

Taking base-2 logarithms, and observing that

\[ \log P_1 > - \binom{2^k}{2} \left( \frac{3}{4} \right)^{n+1} \log e > - \log e \left( 2^{2k-(n+1)\log(4/3)} \right), n \geq 3 \]

and

\[ \log \left( \frac{2^n}{2^k} \right) P_2 > (n-k)k , \]

we see that \((C1)\) holds if \(2k \leq n \log(4/3)\). Denoting the rate \(k/n\) by \(R\) and \(d/n\) by \(\delta\) we have the following result.

**Proposition 4:** For any \(R, 0 \leq R \leq (1/2)\log 4/3\), there exists an infinite sequence of \((n,nR,d \geq nR,1)\) codes.

This result has also been obtained by Komlos (unpublished).
with \( R \leq \delta \) (Proposition 1) to obtain \( R \leq 0.283 \) (the solution of \( f(x) = x \)).

Here \( H(x) = -x \log x - (1-x) \log (1-x) \) is the binary entropy function.

Proposition 5: For \( n \) large enough, maximal intersecting codes are such that \( \frac{1}{2} \log \frac{4}{3} \leq R \leq 0.283 \), or \( 0.207 < R \leq 0.283 \).

Proposition 4 can be strengthened as follows. For any \( R \) in the range \( 0 \leq R \leq \frac{1}{2} \log 4/3 \), almost all \((n,nR)\) codes are intersecting. But it is known that almost all \((n,nR)\) codes satisfy the Varshamov-Gilbert bound (see, e.g., [3], Ch. 17); namely,

\[
H(\delta) \geq 1-R,
\]

and almost no codes are better. Combining this with Proposition 1, i.e., \( R \leq \delta \), and letting \( R_0 \) denote the smallest solution of \( x = 1 - H(x) \), we obtain,

\[
R \leq R_0 = H^{-1}(1-R_0) \leq H^{-1}(1-R) \leq \delta, \quad \text{where} \quad R_0 \approx 0.22.
\]

An infinite sequence of \((n,nR_0,d>R_0,1)\) codes is very unlikely to exist. Still, intersecting codes may meet the Varshamov-Gilbert bound, or equivalently, the following may hold.

Conjecture: For \( n \) large enough, maximal (for cardinality) intersecting codes have \( R = \delta = R_0 \).
5. CYCLIC CODES

In this section codes are cyclic, i.e., if \( c = (c_1, c_2, \ldots, c_n) \) is in \( C \), so are the cyclic shifts of \( c: (c_m, c_{m+1}, \ldots, c_n, c_1, \ldots, c_{m-1}) \), \( 2 \leq m \leq n \).

Identifying a codeword \( c \) with \( c(x) = c_1 + c_2 x + \ldots + c_n x^{n-1} \), cyclic codes can be viewed as ideals of \( A = F_2[x]/(x^n-1) \), the ring of binary polynomials with multiplication modulo \( (x^n-1) \). Defining the generator \( g(x) \) of \( C \) as the polynomial of minimal degree in \( C \), one has \( C = g(x) \cdot A \). For more details, see [3].

**Proposition 6:** If \( C_1(n, k_1, d_1) \) and \( C_2(n, k_2, d_2) \) are \( r \)-intersecting cyclic codes then \( d_1 d_2 \geq nr \).

**Proof:** Let \( M_1 \) be the \( m_1 \times n \) matrix whose rows are the minimum weight codewords of \( C_i \), \( i = 1, 2 \). \( M_1 \) contains \( d_1 m_1 \) ONES. Since \( C_1 \) is cyclic, each column of \( M_1 \) has weight \( d_1 m_1/n \). Now consider the \( m_1 m_2 \times n \) matrix \( M \) whose rows are the termwise products \( e_1 \cdot e_2 \), where \( e_1 \) is a row of \( M_1 \) and \( e_2 \) is a row of \( M_2 \). Since \( C_1 \cap C_2 \), the weight of every row of \( M \) is at least \( r \). Hence \( M \) has at least \( m_1 m_2 r \) ONES. But the weight of every column of \( M \) is exactly \( \frac{d_1 m_1}{n} \cdot \frac{d_2 m_2}{n} \) which implies our claim.

**Proposition 7:** If \( C \) is a BCH code of length \( n = 2^m - 1 \) and designed distance \( 2t + 1 < \frac{1}{3} 2^{m/2} + 3 \), then the dual, \( C^* \), of \( C \) is intersecting.

**Proof:** From Theorem 18 of Ch. 9 in [3], the nonzero weights \( w \) of \( C^* \) are in the range

\[
2^{m-1} - (t-1)2^{m/2} < w \leq 2^{m-1} + (t-1)2^{m/2}.
\]

It is easy to verify that if \( 2t+1 \leq \frac{1}{3} 2^{m/2} + 3 \) then \( d^* - \frac{d_{\text{Max}}}{2} \) is positive and the result follows from (P1).
Proposition 8: If \( C_1(n,k_1,d_1) \) and \( C_2(n,k_2,d_2) \) are \( r \)-intersecting then \( k_1 + k_2 \leq n - r + 2 \).

Proof: \( \deg(g_1(x)) = n - k_1 \) and \( \deg(g_2(x)) = n - k_2 \). Hence, 
\[
|g_1(x) \cdot (x^{k_2 - 1}g_2(x))| \geq r \implies n - k_1 + n - k_2 + 2 \geq n + r.
\]

Proposition 9: For any \( R \), almost no \((n,nR)\) cyclic code is intersecting.

Proof: As in Proposition 4. \( x^{n-1} \) has \( \approx n(\log n)^{-1} \) factors of degree \( \log n \) and we must pick \( n(1-R)(\log n)^{-1} \) of them to get a polynomial \( f(x) \) of degree \( n - k = n(1-R) \). Hence, there are only approximately 
\[
\left(\frac{n(\log n)^{-1}}{n(1-R)(\log n)^{-1}}\right)^{n^{nR}} \text{ cyclic } (n,nR) \text{ codes. If } P'_1 \text{ denotes the probability that a randomly chosen subset of } 2^{nR} \text{ elements of } \mathbb{F}_2^n \text{ be cyclic, we see that for any fixed } R, \left(\frac{2^{nR}}{2^{nR}}\right)^{P_1 P'_2} \text{ approaches zero when } n \text{ goes to infinity. Thus the following problem remains open.}
\]

Problem: Is there a rate \( R, 0 \leq R \leq R_0 \), for which there exists an infinite sequence of \((n,nR,d \geq nR,1)\) cyclic codes? This is linked with Research Problem (9.2) in [3].

Example 6: In [3], Ch. 8, an infinite family of \((2^m, 2m+1)\) cyclic codes with three different nonzero weights is given for odd \( m \). By (P1), these codes are intersecting. The first few of these codes are: 
\((31,10,12), (127,14,56), (511,18,24)\).

For every \( c_1 \in C, c_1(x) = g(x) \sum_{i \in I} x^i, I \subseteq \{0,1,\ldots, k-1\} \). Hence
\[
(g(x) \sum_{i \in I} x^i) \cdot (g(x) \sum_{j \in J} x^j) \neq 0 \text{ for any nonempty subsets } I, J \text{ of } \{0,1,\ldots, k-1\}.
\]
Setting \( g_i = g \ast (x^i g) = (x^i g) \ast g \) and noticing that if \( j \geq i \) then 
\((x^i g) \ast (x^j g) = x^i g_{j-i} \), we obtain the following result.

**Proposition 10:** \( C \) is intersecting iff for all nonempty \( I, J \subseteq \{0, 1, \ldots, k-1\} \)

\[
\sum_{i \in I, j \in J} \min(i, j) g_{i-j} \neq 0 
\]

6. **THE NONBINARY CASE**

Proposition 1 extends in a straightforward manner to codes over any field \( F_q \). But we can say a little more: if one of the codes is maximum distance separable, i.e., satisfies \( d = n-k+1 \), then Proposition 1 and Corollary 1 are "iff".

**Proposition 11:** If \( C_1(n,k_1,d_1) \) and \( C_2(n,k_2,d_2) \) are codes over \( F_q \) with \( d_1 = n-k_1+1 \) and \( d_2 \geq k_1 + r - 1 \) then \( C_1 \cap C_2 \).

**Proof:** We have

\[ d_1 + d_2 \geq n + r \] and thus \( C_1 \cap C_2 \).

Also, \( k_1 \leq d_2 - r + 1 \leq n - k_2 + 2 - r \) so that \( n - k_1 + 1 = d_1 \geq k_2 + r - 1 \) holds automatically.

For example, any Reed-Solomon code ([3], Ch. 10) is intersecting iff \( d \geq k \).
REFERENCES

