

Weight Enumerator and the Covering Radius of Codes

Fuad M. Shareef

Learning Development & Continuing Education
Queen Mary, University of London, Mile End, E1 4NS
E-mail: f.shareef@qmul.ac.uk

Abstract: The covering radius of a code is a fundamental parameter that is closely related to the quality of the code. Determining the exact value of the covering radius of a binary linear code is a NP-complete problem. The question arises as to whether extra information on the code parameters can help in determining its covering radius. In this paper¹ we are concerned with this question and in particular with the potential relationship between the covering radius of a code and its weight enumerator. This is motivated by the observation that knowledge of the weight enumerator sheds a light on the way codewords are distributed. We shall give an upper bound on the covering radius of arbitrary linear codes in terms of their parameters and show that inequivalent codes with the same weight enumerator do not necessarily have equal covering radius.

Keywords: *Weight Enumerator,*

Introduction

The covering radius of a code is one of its fundamental parameters and good covering codes have a number of applications in data compression, cellular telecommunications [1] and interrelations with other areas of Mathematics. It is bounded above by the external distance of the code [2], but cannot be determined by the weight enumerator [3]. The problem of finding the exact covering radius of a binary linear code has been classified a NP-complete problem [4], i.e. no polynomial time deterministic algorithm is known that solves this problem.

The rest of the paper is organised as follows. In section 2 we review basic definitions and relevant results and in Section 3 we prove the main theorem of this paper about a new upper bound on the value of covering radius in terms of the code parameters. We have finally given some counter examples in section 4.

Standard definitions

A linear $[n, k]$ code C is a k -dimensional vector subspace of $\text{GF}(q)^n$, where $\text{GF}(q)$ is the field of q elements (for binary $q=2$) and n is the length of the code. A linear code can be described by its generator matrix. The

¹ Much of the material in this paper is from the authors' PhD theses [12].

generator matrix of a $[n, k]$ code C is a $k \times n$ matrix, whose rows form a basis of the linear code C and usually denoted by G . An $[n, k, d]$ code is an $[n, k]$ with minimum (non-zero) distance d , where d is the Hamming distance,

$$d(C) = \min\{d(x, y) : x, y \in C, x \neq y\} \text{ and } d(x, y) = |\{i : 1 \leq i \leq n, x_i \neq y_i\}|.$$

The elements of C are called *Codewords*. The *weight* $wt(x)$ of a codeword is defined to be the number of non-zero entries of x , that is $wt(x) = d(x, 0)$. The *weight enumerator* of the code C is defined by the following polynomial

$$W_C(x, y) = \sum_{i=0}^n A_i x^i y^{n-i},$$

where $A_i (i = 0, 1, 2, \dots, n)$ denote the number of codewords of weight i .

There are several equivalent ways of defining the covering radius of codes. Here we adopt the following. The *covering radius* of a block code of length n is, the minimal number $R = R(C)$ such that all vectors in the containing space are within Hamming distance $R(C)$ of some codeword; thus for a code over $GF(q)^n$:

$$R(C) = \max\{\min\{d(x, c) : c \in C\} : x \in GF(q)^n\},$$

where $d(x, c)$ is the distance between the words x and the codeword c .

If we define the inner product of two words $\langle x, y \rangle$ in $GF(q)^n$ in the usual way, that is,

$$\langle x, y \rangle = x_1 y_1 + x_2 y_2 + \dots + x_n y_n,$$

then the *dual code* C^\perp can be defined as follows.

$$C^\perp := \{y \in GF(q)^n : \forall_{x \in C} [\langle x, y \rangle = 0]\}.$$

The dimension of C^\perp equal to $n-k$, and a generator matrix for C^\perp is called a *parity check matrix* for C and denoted by H .

That is, $C := \{x \in GF(q)^n : xH^T = 0\}$ and $GH^T = 0$.

A code C is *self-orthogonal* if C is contained in C^\perp and *self-dual* if $C = C^\perp$.

Before we present the upper bound of the covering radius, we start with a theorem of Delsarte [2]. We define the *external distance* of a code to be the number of non-zero weights in the dual code.

Theorem 2.1 (Delsarte [2]). If a code has covering radius R and external distance m , then $R \leq m$.

Now by the Mac Williams identities [5] the weight enumerator of a code determines the weight enumerator of the dual code and so determines the external distance, which is an upper bound for the covering radius by Delsarte's bound in Theorem 2.1. However, this bound is not always attained.

Example 2.1: The extended Golay code is self-dual and has weights 0, 8, 12, 16 and 24. So its external distance is 4, which is also the covering radius.

Example 2.2: The self-dual code $D_{14} \{14,7,4\}$ has weights 0, 4,6, 8,10, and 14, so its external distance is 5. But the covering radius is equal to three; this is an example where the Delsarte bound is not attained.

Delsarte [2] defines C as having a *strength* s if each s -subset of coordinates of the code containing all q -ary s -tuples a constant number of times. We say a code has *full support* when it has strength $s=1$; the cardinality of its support is equal to its length, and this means that the generator matrix has no zero columns.

A new upper bound

Theorem 3.1. Let C be a linear code of length n , having full support, over $GF(q)$. Then the covering radius $R(C)$ is at most $\left(\frac{q-1}{q}\right)n$. Equality holds if and only if each maximal set of pairwise linearly dependent columns of any generator matrix of C has cardinality divisible by q .

Proof. We first show that, for any word w in $GF(q)^n$, the average distance from w to a codeword in C is $\left(\frac{q-1}{q}\right)n$. This average is

$$\frac{1}{|C|} \sum_{c \in C} d(w, c) = \frac{1}{q^k} \sum_{c \in C} \sum_{i=1}^n d(w_i, c_i)$$

where $k = \dim(C)$. Now reverse the order of summation. Since C has full support, each position has all q possible symbols a constant number of times. So each position contributes $\left(\frac{q-1}{q}\right)q^k$ to the sum; that is, the new inner sum is

$$\sum_{c \in C} d(w_i, c_i) = \left(\frac{q^k}{q}\right)(q-1)$$

and hence the average distance from w to the codewords of C is

$$\frac{n}{q^k} \left(\frac{q^k}{q}\right)(q-1) = \left(\frac{q-1}{q}\right)n$$

as claimed. In particular, any word lies within distance $\left(\frac{q-1}{q}\right)n$ from a codeword, and the bound is proved. Equality holds if and only if there is a word w which lies at a distance $\left(\frac{q-1}{q}\right)n$ from every codeword. Call two non-

zero column vectors over $GF(q)$ of length k equivalent if one is a scalar multiple of the other. There are $\binom{q^k - 1}{q - 1}$ equivalence classes. and each equivalence class corresponds to a unique point of the projective space $PG(k-1, q)$. Let $S(c)$ denote the support of the one-dimensional subcode $c \in C$. For each of the $\binom{q^k - 1}{q - 1}$ possible equivalence classes p , let $T(p)$ be the set of coordinate positions where a column in p occurs in a fixed generator matrix for C . Each non-zero $S(c)$ is a sum of certain $T(p)$. That is, the $T(p)$ partition $\{1, \dots, n\}$ and therefore also $S(c)$. For, c is a linear combination $\sum a_i r_i$ of the rows r_i of the generator matrix; and $S(c)$ is the disjoint union of those $T(p)$ for which the corresponding linear combination

$\sum a_i p_i$ of the entries of p is not zero. That is.

$$|S(c_i)| = \sum m_{ij} |T(p_j)|$$

The coefficient matrix $M = m_{ij}$ is in fact the incidence matrix of the complement of the point-hyperplane design in $PG(k-1, q)$. For, if we regard the equivalence classes p as points of $PG(k-1, q)$, the set of p for which $T(p)$ is contained in $S(c)$ is precisely the complement of the hyperplane defined by the equation $\sum a_i p_i = 0$. It follows that M is non-singular. So the number $|T(p)|$ can be recovered from the numbers $|S(c)|$ by solving a system of linear equations with coefficient matrix M , that is, inverting the matrix M . Now suppose that the covering radius is $\left(\frac{q-1}{q}\right)n$, that is, there is a word w with $d(w, c) = \left(\frac{q-1}{q}\right)n$ for each $c \in C$. Let $S(w)$ be the support of w ; then, in

the positions of $S(c)$, for each $\lambda \in F(q)$, the entry of w is λ times the entry of c for exactly $\frac{1}{q}|S(c)|$ positions. Hence

$$|S(c_i) \cap S(w)| = \sum m_{ij} \left(\frac{q-1}{q}\right) |T(p_j)|$$

The numbers $|S(m) \cap S(c)|$ are expressed in terms of the numbers $|S(m) \cap T(p)|$ using the same matrix M as before:

$$|S(c_i) \cap S(w)| = \sum m_{ij} |T(p_j) \cap S(w)|,$$

by considering the code punctured on positions outside $S(w)$. Since these equations have a unique solution,

$$|T(p_j) \cap S(w)| = \left(\frac{q-1}{q}\right) |T(p_j)|$$

for all j . Thus q divides $|T(p_j)|$ for all j , as was to be shown. Conversely, if each $|T(p)|$ has a cardinality of a multiple of q , choose a word w such that, in each set $T(p)$, w contains each entry of $\text{GF}(q)$ exactly $\frac{1}{q}|T(p)|$ times. Then the word w is at distance $\left(\frac{q-1}{q}\right)n$ from every codeword, and the covering radius is $\left(\frac{q-1}{q}\right)n //$.

Note that, Delsarte theorem 2.1 is more general than theorem 3.1. And the new upper bound is not a improvement on Delsarte bound. However, the proof and the result of theorem 3.1 have been used to construct the counterexample to prove the conjecture that the weight enumerator of a linear code determines its covering radius.

Counterexamples

In this section we present four pairs of codes. Each pair has the same weight enumerator but distinct covering radii. This makes them counterexamples to the conjecture that the weight enumerator of a linear code determines its covering radius. The first pair of codes has dimension three. The strategy of the construction is as follows. The weight distribution of a code does not specify the weights of individual words. We take a weight distribution where, for convenience, all the weights are distinct, and assign the weights to words in two different ways. Then we solve the equations for the numbers $|T(p)|$, hoping for an example where in one case all the $|T(p)|$ are even and the other case they are not. The reason this works is that not all permutations of seven possible weights give equivalent codes.

Example 4.1. Our codes will have length 36 and non-zero weights 14, 16, 18, 20, 22, 24 and 30. Consider the seven possible columns of length 3, that is, the columns of the matrix

$$\begin{pmatrix} 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Let them occur a, b, c, d, e, f and g times respectively. Then the non-zero weights of the code are given by $a + c + d + e$ and six similar expressions. Now the solutions to the equations

$$a + c + d + e = 14, 14$$

$$a + b + d + f = 18, 18$$

$$a + b + c + g = 22, 22$$

$$c + d + f + g = 20, 20$$

$$b + d + e + g = 24, 30$$

$$b + c + e + f = 16, 16$$

$$a + e + f + g = 30, 24$$

are

$$(a, b, c, d, e, f) = (6, 4, 0, 2, 6, 6, 12)$$

and

$$(3, 7, 0, 5, 6, 3, 2)$$

respectively. By theorem 3.1 the first code has covering radius 18 and the second has a covering radius strictly smaller than 18. In fact, the covering radius of the second code is 17, as can be seen by considering a word in which the numbers of non-zero entries in the sets $T(p)$ are 2, 3, 0, 2, 3, 1 and 6 respectively. Such a word has distance 17 or 19 from every codeword, as is easily checked.

Remark. If C_1 is projected onto two complementary sets of coordinates giving codes C_1 and C_2 then the covering radius of C is at least the sum of the covering radii of C_1 and C_2 . For, if we choose words w_1 and w_2 whose distances from C_1 and C_2 are the covering radii of these codes, then the word whose projections are w_1 and w_2 is at distance at least the sum of these distances from C . In general this bound is not sharp. However, we can produce a decomposition related to the binary case in Theorem 3.1 as follows. Let X_1 be a set of coordinates containing one coordinate from each $T(p)$ of odd cardinality, and X_2 the complementary set. Then, by Theorem 3.1, the covering radius of the projection C_2 onto X_2 is half its length. In the example constructed above with covering radius 17, the code C_1 consists of all words of length 4 and even weights; its covering radius is 1. The covering radius of the other projection C_2 is $\frac{1}{2}(36-4)=16$. The covering radius of C is the sum of these two numbers.

Example 4.2. Pless and Sloane [5] determined the weight distributions of the indecomposable binary self-dual codes of length 24. In [3] the covering radii were determined.

Table1: Independent binary self-dual codes of length 24

Code	Weight Distribution						Covering Radius
C	A_2	A_4	A_6	A_8	A_{10}	A_{12}	$C(R)$
A_{24}	0	30	0	639	0	2756	6
B_{24}	0	24	0	663	0	2720	5
C_{24}	0	18	0	687	0	2684	6
D_{24}	0	12	0	711	0	2648	5
E_{24}	0	66	0	495	0	2972	6
F_{24}	0	6	0	735	0	2612	6
G_{24}	0	0	0	750	0	2376	4
H_{24}	0	34	64	239	960	1500	6
I_{24}	0	22	64	287	960	1428	6
J_{24}	0	20	64	295	960	1316	5
K_{24}	0	20	64	295	960	1416	5
L_{24}	0	18	64	303	960	1404	6
M_{24}	0	18	64	303	960	1404	6
N_{24}	0	16	64	311	960	1392	5
O_{24}	0	14	64	319	960	1380	6
P_{24}	0	14	64	319	960	1380	6
Q_{24}	0	12	64	327	960	1368	5
R_{24}	0	12	64	327	960	1368	5
S_{24}	0	10	64	335	960	1356	5
T_{24}	0	10	64	335	960	1356	6
U_{24}	0	8	64	343	960	1344	5
V_{24}	0	6	64	351	960	1332	6
W_{24}	0	6	64	351	960	1332	6
X_{24}	0	4	64	359	960	1320	5
Y_{24}	0	2	64	367	960	1308	6
Z_{24}	0	0	64	375	960	1296	5

The codes labeled S_{24} and T_{24} in Table 1 have the same weight enumerator, but different covering radii: $R(S_{24}) = 5$ and $R(T_{24}) = 6$.

The assumption that the code has full support is equivalent to saying that its dual has minimum weight greater than one. The codes in Example 4.1

necessarily have repeated columns in their generator matrices, which is equivalent to saying that their duals have minimum weight exactly two. By the MacWilliams Identities, if the codes have the same weight distribution, then so do their duals, so Example 4.2 gives two codes whose duals have minimum weight at least three.

Example 4.3. To find another example, we investigated a class of codes where the weight enumerators are very restricted. The strategy is that if we found two codes in the class with different covering radii, then the chance that they would have the same weight enumerator was better. By Gleason's Theorem and its variants, self-dual codes form such a class. The weight enumerator of a doubly-even self-dual code of length 16 is

$$(x^8 + 14x^4y^4 + y^8)^2 = x^{16} + 28x^{12}y^4 + 198x^8y^8 + 28x^4y^{12} + y^{16}$$

There are exactly two inequivalent codes having the above weight enumerator, the code E_{16} and the direct sum of two binary extended Hamming codes of length 8 (that is $A_8 + A_8$). The covering radius of the code $A_8 + A_8$ is 4 and computation showed that the covering radius of the code E_{16} is 4. Hence we have two inequivalent codes having the same weight enumerator and the same covering radius. During the process of finding an example of two codes having the same weight enumerator but different covering radii, we found by trial and error a code R'_{20} and compared it with the code S_{20} . The code R'_{20} is a small modification of R_{20} it is formally self-dual but not self-dual. The two codes S_{20} and R'_{20} have the same weight enumerator but different covering radii. The generator matrix, weight enumerator and covering radius of each code is given below. S_{20} is a linear

{ 20,10,4} code (see [6]over GF(2), with a generator matrix:

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \end{pmatrix}$$

The weight distribution of S_{20} is

$$1,0,0,0,13,0,64,0,242,0,384,0,242,0,64,0,13,0,0,0,1$$

The covering radius of S_{20} is five. R'_{20} is a linear $\{20, 10, 4\}$ code over $GF(2)$, with a generator matrix:

$$\begin{pmatrix} 00111100000000000000 \\ 00001111000000000000 \\ 00000000111100000000 \\ 00000000001111000000 \\ 00000000000000111100 \\ 00000000000000001111 \\ 01001100000000100110 \\ 00000000000000001111 \\ 10101010110000000110 \\ 11101010101010000000 \\ 11000000101010101100 \\ 11000011000011000011 \end{pmatrix}$$

The weight distribution of R'_{20} is

$$1,0,0,0,13,0,64,0,242,0,384,0,242,0,64,0,13,0,0,0,1$$

The covering radius of R'_{20} is 4.

Remark. The first code S_{20} is self dual, whereas the second code R'_{20} is not self dual and not self orthogonal, however its dual has the same weight enumerator (i.e., it is a formally self dual code) and same covering radius. The code R'_{20} found by trial and error and its generator matrix is very similar to the generator matrix of the code R_{20} they only differ by one row.

Example 4.4. The smallest example we know of was given in [8] as an example of code pairs with identical weight enumerators but with dissimilar higher weight enumerators. The codes generated by

$$\begin{pmatrix} 110000 \\ 001100 \\ 000011 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 110000 \\ 101000 \\ 111111 \end{pmatrix}$$

both have weight distribution

$$1,0,3,0,3,0,1,$$

but they have covering radii 3 and 2 respectively.

Conclusion and open problem

We have investigated a suggested alternative relationship between the covering radius of a code and its weight enumerator polynomial. We have given a new upper bound for the value of covering radius of a code in

terms of its parameters. We have also demonstrated, using counterexamples, that the weight enumerator of a code does not determine its covering radius uniquely. However, determining all cases where the upper bound is attained remains an open problem. Future work will be focused on this problem.

Problem: Determine all cases where the bound is attained.

Acknowledgement

I would like to thank Prof. Peter J. Cameron for his continuous help, advice and support during the course of my PhD study. I am also grateful to the reviewer for his useful comments and valuable suggestions.

References

1. G. Cohen, I. Honkala, S. Litsyn, A. Lobstein. (1997) *Covering Codes*, North-Holland, Mathematical Library.
2. P. Delsarte. (1973) Four fundamental parameters of a code and their combinatorial significance. *Inform. and Control*, 23:407-438.
3. F. M. Shareef (2001). *The covering radius of codes and its relation to designs*. PhD thesis, University of London,.
4. E.F. Assmus Jr & F.F. Mattson, Jr (1969). New 5 –designs, *J. Combinatorial Theory(A)*, 27,307-423.
5. V. Pless (1972). A classification of self-orthogonal codes over GF(2). *Discrete Math.*, 3:209-246.
6. V. Pless and N. J. A. Sloane (1975). On the classification and enumeration of self-dual codes. *J. Comb. Theory, Ser. A*, 18:313-335.
7. C. G. Rutherford (2001). *Matroids, codes, and their polynomial links*. PhD thesis, University of London.
8. F.J. MacWilliams and N.J. Sloan (1977). *Theory of error-correcting codes*, North Holland, Amsterdam.