New Binary Linear Codes*

T. Aaron Gulliver
Department of Electrical and Electronic Engineering
University of Canterbury
Christchurch, New Zealand
gulliver@elec.canterbury.ac.nz
and
Patric R. J. Östergård
Department of Computer Science and Engineering
Helsinki University of Technology
P.O. Box 5400, 02015 HUT, Finland
patric.ostergard@hut.fi.

Abstract

In this paper, nineteen new binary linear codes are presented which improve the bounds on the maximum possible minimum distance. These codes belong to the class of quasi-cyclic (QC) codes, and have been constructed using a stochastic optimization algorithm, tabu search. Six of the new codes meet the upper bound on minimum distance and so are optimal.

1 Introduction

A binary linear $[n, k, d]$ code is a subspace of F_2^n of dimension k with minimum Hamming distance d. The maximum possible minimum Hamming distance, given n and k, is denoted by $d_2(n, k)$. A linear code which has minimum distance equal to $d_2(n, k)$ is called *optimal*. A related problem, given k and d, is to find the smallest value of n for which there exists an $[n, k, d]$ code; this value is denoted by $n_2(k, d)$. For given values of q and k,

*This research was supported in part by the Academy of Finland.

solving one of these problems is equivalent to solving the other. The values of $n_2(k,d)$ have been determined for $k \leq 7$, and many values of $n_2(k,d)$ for $k \leq 10$ are known [1]. Brouwer [1] maintains a table of upper and lower bounds on $d_2(n,k)$ for $n \leq 256$. In this paper, bounds for nineteen values of $d_2(n,k)$ are improved or established through the construction of quasi-cyclic (QC) codes. Recently, non-degenerate QC codes were investigated by Heijnen, Van Tilborg, Verhoeff, and Weijs [5] and five lower bound improvements were obtained. In this paper we also consider degenerate codes, and by using a stochastic optimization algorithm, tabu search, for the code construction we obtain the new bounds.

The next section describes the class of codes considered. Section III discusses the construction algorithm and the stochastic search method used. Finally, Section IV gives the construction results, and lists the codes which have improved lower bounds on $d_2(n,k)$ or established exact values of this function.

2 Quasi-Cyclic Codes

A code is called quasi-cyclic if a cyclic shift of a codeword by p positions results in another codeword. The blocklength, n, of a QC code is a multiple of p, so that $n = mp$. A class of QC codes can be constructed from $m \times m$ circulant matrices (with a suitable permutation of coordinates). In this case, the generator matrix, G, can be represented as

$$G = [B_0, B_1, B_2, \ldots, B_{p-1}],$$

where

$$B_i = \begin{bmatrix} b_{0,i} & b_{1,i} & b_{2,i} & \cdots & b_{m-2,i} & b_{m-1,i} \\ b_{m-1,i} & b_{0,i} & b_{1,i} & \cdots & b_{m-3,i} & b_{m-2,i} \\ b_{m-2,i} & b_{m-1,i} & b_{0,i} & \cdots & b_{m-4,i} & b_{m-3,i} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ b_{1,i} & b_{2,i} & b_{3,i} & \cdots & b_{m-1,i} & b_{0,i} \end{bmatrix}.$$

The algebra of $m \times m$ circulant matrices over GF(2) is isomorphic to the algebra of polynomials in the ring $\mathrm{GF}(2)[x]/(x^m - 1)$ if B_i is mapped onto the polynomial

$$b_i(x) = b_{0,i} + b_{1,i}x + b_{2,i}x^2 + \cdots + b_{m-1,i}x^{m-1},$$

formed from the entries in the first row of B_i [7]. The set of polynomials $\{b_i(x)\}$ are called the *defining polynomials*. Let $b_0(x)$ denote the all-zero polynomial, and let $M = |\{b_i(x) \mid i > 0\}|$. A subset of p of these polynomials

$$\{b_{j_0}(x), b_{j_1}(x), \cdots, b_{j_{p-1}}(x)\}, 1 \leq j_i \leq M,$$

($j_a \neq j_b$ when $a \neq b$) defines an $[mp, k]$ QC code, where $k = m - \deg(h(x))$ and

$$h(x) = \frac{x^m - 1}{\gcd\{x^m - 1, b_{j_0}(x), b_{j_1}(x), \cdots, b_{j_{p-1}}(x)\}}.$$

The polynomial $h(x)$ is called the order of the QC code [9]. Codes for which $k < m$ are called *degenerate*.

The problem is to find a subset which gives the largest minimum distance, d. Only codes for $8 \leq m \leq 17$ are considered in this paper because it was found that codes with large values of $m - k$ have poor distance properties. In addition, only codes for $8 \leq k \leq 13$ are presented, as this represents the practical computational limit of the technique used.

3 The Construction Algorithm

It is not necessary to check the weight of every codeword in a QC code in order to determine d. Only a subset, $N < M$, of the codewords need be considered since the Hamming weight of $i_t(x)b_s(x) \bmod (x^m - 1)$ is equal to the weight of $i_t(x)x^l b_s(x) \bmod (x^m - 1)$ for all $l \geq 0$. Note that this argument also applies to the set of defining polynomials.

To simplify the process of searching for good codes, the weights of this subset of codewords can be stored in an array. A matrix, D, can be formed from the arrays for the subset of defining polynomials which need to be considered

$$D = \begin{array}{c|cccccc} & b_1(x) & b_2(x) & \cdots & b_s(x) & \cdots & b_y(x) \\ \hline i_1(x) & w_{11} & w_{12} & \cdots & w_{1s} & \cdots & w_{1y} \\ i_2(x) & w_{21} & w_{22} & \cdots & w_{2s} & \cdots & w_{2y} \\ \vdots & \vdots & \vdots & & \vdots & & \vdots \\ i_t(x) & w_{t1} & w_{t2} & \cdots & w_{ts} & \cdots & w_{ty} \\ \vdots & \vdots & \vdots & & \vdots & & \vdots \\ i_z(x) & w_{z1} & w_{z2} & \cdots & w_{zs} & \cdots & w_{zy} \end{array}$$

where $i_t(x)$ is the tth information polynomial, $b_s(x)$ is the sth generator polynomial, and w_{ts} is the Hamming weight of $i_t(x)b_s(x) \bmod (x^m - 1)$. Since $i_s(x)$ and $b_t(x)$ correspond to the same subset of defining polynomials, D is a square ($y = z = N$), symmetric (by letting $i_t(x) = b_t(x)$ for all $1 \leq t \leq N$) matrix.

The complete weight distribution for a QC code composed of any set of $b_t(x)$ can be constructed from D. The search for a good code consists of finding p columns of D with a large row sum, since the weight of a minimum distance codeword must be contained in these sums.

Having decided on the values of k, m, and p (and thus also $n = mp$), the entries of the integer matrix D can be calculated and the problem formulated as a combinatorial optimization problem. Namely, we want to find

$$\max_{S} \min_{1 \leq j \leq N} \sum_{s \in S} w_{j,s}, \tag{1}$$

where $S \subseteq \{1, 2, \ldots, N\}$ and $|S| = p$. More generally, we could take a multiset S with p elements, but it has turned out that for the new codes found in such studies, no defining polynomial occurs more than once, so (also because this made the optimization procedure perform better) S is here indeed required to be a set.

The optimization method used in this work is *tabu search* [2]. This is a stochastic heuristic which can produce good near-optimal solutions to difficult optimization problems with a reasonable amount of computational effort (but it cannot be used to prove or disprove optimality of solutions found). For an extensive survey of stochastic optimization methods in coding theory, see [6]. Tabu search is based on local search, which means that starting from an initial solution, a series of solutions is obtained so that every new solution only differs slightly from the previous solution (is in the *neighborhood* of the previous solution). To evaluate the quality of solutions, a *cost function* is needed. Tabu search always proceeds to a best possible solution in the neighborhood of the current solution. However, to avoid the search from looping on a subset of moves, (attributes of) recent moves are stored in a so-called tabu list, and inverses of these moves are then not allowed for a certain period of time (here, for a predefined number of moves, L).

Tabu search has here been applied to the problem of finding QC codes,

defined as a minimization problem, in the following way. First, the problem is not formulated as generally as in (1), as the desired minimum distance, d, of the code is fixed. A solution is any set $S \subseteq \{1, 2, \ldots, N\}$ of p columns, the neighborhood of a solution is the set of solutions obtained by replacing one column with a column that is not in the code, and the cost function is of the form

$$C = \sum_{i=1}^{N} \max(0, d - \sum_{j \in S} w_{i,j})^a,$$

where both $a = 1$ and $a = 2$ have been used. A (globally optimal) solution with cost 0 clearly corresponds to a code with minimum distance at least d. The tabu list is simply the indexes of the new columns. Thus, if a column is replaced by another, the new column must not be replaced during the L next moves.

Although this basic approach worked well, it turned out that this search method became even more effective using the following alternative neighborhood, based on an idea from [8]. During the search, the row sums of the current solution are kept in memory and updated after each move. Now, we go cyclically through this array of sums, until a sum, in row i, is encountered that is smaller than d. Then the neighborhood consist of all possible replacements of column j in the current solution by column j' whenever $w_{i,j'} > w_{i,j}$. This reduces the size of the neighborhood and makes the search more effective.

The values of L used were in the range $p/10 \leq L \leq p/5$. If a code was not found within 5000–10000 iterations, the search was restarted from a new random initial solution. Depending on the computer time available, as many as 100 restarts were performed per instance. The total number of iterations to find a code varied between about one thousand and a few hundred thousand.

4 Construction Results

The defining polynomials of the nineteen new QC codes are compiled in Table 1 (in octal form), and the weight distributions in Table 2. These codes have improved the lower bounds on the maximum minimum distance given in [1]. Six exact values are established, namely $d_2(189, 9) = 92, d_2(207, 9) = 100, d_2(225, 9) = 110, d_2(45, 10) = 18, d_2(91, 12) = 40$ and $d_2(140, 12) = 64$,

Table 1: Defining Polynomials for the New Binary QC Codes

n	k	m	d	$b_i(x)$
144	9	9	68	127, 57, 133, 15, 23, 13, 137, 277, 53, 63, 73, 253, 135, 257, 43, 25
162	9	9	77	165, 1, 17, 25, 133, 67, 75, 43, 137, 153, 253, 7, 357, 117, 147, 31, 47, 23
171	9	9	81	173, 377, 177, 71, 75, 37, 25, 67, 273, 337, 165, 31, 127, 3, 125, 133, 35, 157, 357
180	9	9	87	57, 51, 31, 53, 137, 127, 153, 133, 147, 253, 177, 277, 21, 25, 55, 43, 23, 13, 27, 75
189	9	9	92°	253, 21, 31, 165, 117, 173, 47, 25, 13, 57, 23, 1, 377, 267, 175, 7, 65, 133, 273, 147, 17
207	9	9	100°	31, 75, 25, 57, 23, 47, 1, 55, 125, 3, 155, 137, 51, 73, 377, 167, 153, 7, 175, 3, 127, 157, 273
225	9	9	110°	63, 137, 35, 1, 125, 153, 277, 377, 51, 7, 177, 147, 75, 337, 53, 273, 117, 113, 357, 57, 37, 267, 67, 155, 165
45	10	15	18°	175, 1067, 11173
160	10	10	74	273, 145, 125, 111, 117, 325, 53, 127, 133, 255, 3, 353, 253, 357, 153, 43
170	10	10	80	11, 157, 135, 225, 77, 115, 43, 573, 55, 253, 527, 53, 567, 165, 233, 111, 107
132	11	11	60	75, 57, 1257, 457, 243, 557, 23, 337, 533, 775, 133, 255
154	11	11	70	107, 371, 667, 1737, 553, 271, 455, 45, 1677, 767, 331, 733, 1257, 345
165	11	11	76	345, 35, 33, 5, 563, 333, 57, 117, 277, 23, 365, 331, 1267, 265, 1777
168	11	14	77	7775, 331, 1335, 6667, 2375, 1127, 1347, 1031, 1621, 625, 657, 105
91	12	13	40°	1255, 1135, 1537, 3273, 3553, 1477, 63
135	12	15	60	1661, 27757, 123, 4631, 1753, 16557, 3737, 473, 3245
140	12	14	64°	1045, 3775, 5167, 677, 2155, 2543, 1177, 5, 2133, 1531
176	12	16	80	14347, 24517, 15623, 3445, 12467, 31673, 24555, 7227, 6267, 32675, 2331
112	13	14	48	3311, 5157, 2263, 3257, 2531, 4447, 3063, 12577

Table 2: Weight Distributions of the New QC Codes

Code	Weight Distribution
$[144, 9, 68]$	$0^1 68^{225} 72^{159} 76^{45} 80^{63} 84^9 88^9 108^1$
$[162, 9, 77]$	$0^1 77^{126} 78^{120} 79^{54} 80^{45} 85^{36} 86^{72} 87^9 88^{18} 93^{30} 135^1$
$[171, 9, 81]$	$0^1 81^{63} 82^{117} 83^{54} 84^{27} 85^{99} 86^{45} 88^{27} 92^9 93^{12} 97^9 98^{27} 99^{10} 101^9 102^3$
$[180, 9, 87]$	$0^1 87^{192} 88^{189} 95^{45} 96^{57} 103^9 104^9 111^9 135^1$
$[189, 9, 92]$	$0^1 92^{378} 96^{63} 108^{70}$
$[207, 9, 100]$	$0^1 100^{324} 104^{54} 108^{61} 112^9 116^{54} 124^9$
$[225, 9, 110]$	$0^1 110^{324} 112^{117} 126^{60} 128^9 144^1$
$[45, 10, 18]$	$0^1 18^{185} 20^{183} 22^{225} 24^{155} 26^{195} 28^{45} 30^{35}$
$[160, 10, 74]$	$0^1 74^{220} 76^{150} 78^{190} 80^{127} 82^{80} 84^{55} 86^{50} 88^{40} 90^{50} 92^{10} 94^{40} 98^{10} 100^1$
$[170, 10, 80]$	$0^1 80^{516} 88^{380} 96^{115} 104^{10} 120^2$
$[132, 11, 60]$	$0^1 60^{627} 64^{528} 68^{374} 72^{319} 76^{143} 80^{55} 88^1$
$[154, 11, 77]$	$0^1 70^{352} 72^{506} 78^{473} 80^{396} 86^{176} 88^{110} 94^{22} 96^{11} 110^1$
$[165, 11, 76]$	$0^1 76^{616} 80^{473} 84^{429} 88^{254} 92^{154} 96^{88} 100^{33}$
$[168, 11, 77]$	$0^1 77^{380} 78^{238} 80^{294} 85^{420} 86^{168} 88^{168} 93^{196} 94^{105} 96^{49} 101^{28} 126^1$
$[91, 12, 40]$	$0^1 40^{1547} 48^{2275} 56^{273}$
$[135, 12, 60]$	$0^1 60^{835} 64^{1005} 68^{825} 72^{860} 76^{405} 80^{150} 84^{15}$
$[140, 12, 64]$	$0^1 64^{1715} 72^{1680} 80^{700}$
$[176, 12, 80]$	$0^1 80^{1216} 88^{1768} 96^{1002} 104^{104} 112^4 128^1$
$[112, 13, 48]$	$0^1 48^{1743} 56^{4704} 64^{1743} 112^1$

and these are denoted by an ^{o} superscript in Table 1. Numerous other improvements on the bounds in [1] are obtained using, for example, the inequalities

$$d_2(n, k) \geq d_2(n + 1, k) - 1$$

and

$$d_2(n + 1, k) \geq d_2(n, k).$$

The fact that three of the optimal codes listed are degenerate indicates that good QC codes need not have full rank, i.e., $k = m$.

References

[1] A.E. Brouwer, Linear code bound (server), Eindhoven University of Technology, The Netherlands, available electronically in the WWW at <URL:http://www.win.tue.nl/win/math/dw/voorlincod.html>.

[2] F. Glover, "Tabu search—Part I," *ORSA J. Comput.*, vol. 1, pp. 190–206, 1989.

[3] T.A. Gulliver and V.K. Bhargava, "Some best rate $1/p$ and $(p-1)/p$ systematic quasi-cyclic codes," *IEEE Trans. Inform. Theory*, vol. 37, pp. 552–555, 1991.

[4] T.A. Gulliver and P.R.J. Östergård, "Improved bounds for ternary linear codes of dimension 7," *IEEE Trans. Inform. Theory*, vol. 43, pp. 1377–1381, 1997.

[5] P. Heijnen, H. van Tilborg, T. Verhoeff, and S. Weijs, "Some new binary, quasi-cyclic codes," *IEEE Trans. Inform. Theory*, vol. 44, pp. 1994–1996, 1998.

[6] I.S. Honkala and P.R.J. Östergård, "Applications in code design," in *Local Search in Combinatorial Optimization*, E. Aarts and J.K. Lenstra, Eds., New York: Wiley, 1997.

[7] F.J. MacWilliams and N.J.A. Sloane, *The Theory of Error-Correcting Codes*, Amsterdam: North-Holland, 1977.

[8] P.R.J. Östergård, "Constructing covering codes by tabu search," *J. Combin. Des.*, vol. 5, pp. 71–80, 1997.

[9] G.E. Séguin and G. Drolet, "The theory of 1-generator quasi-cyclic codes," Technical Report, Royal Military College of Canada, Kingston, ON, 1991.

[10] G. Solomon and J.J. Stiffler, "Algebraically punctured cyclic codes," *Inform. and Control*, vol. 8, pp. 170–179, 1965.