Configurations in Binary Linear Codes

Martin Dowd
MartDowd@aol.com

It has long been a great mystery of coding theory, why there is a gap of a factor of $\sqrt{2}$ between the maximum length of double error correcting codes given by the sphere packing bound, and the length of the best known linear double error correcting codes. In [Dowd 88] it was observed that a better theory of these packings would seem to be a prerequisite to a theory of the analog of the Erdos-Turan conjecture for linear double error correcting codes. Some theorems were proved about configurations in such codes.

In this paper some further facts about such configurations are proved, and some conjectures made relevant to the main questions. We adopt the following notational conventions. $\mathcal{F}_q$ denotes the finite field of order q, for a prime power q. A binary code of length n is a subset of the vector space $\mathcal{F}_2^n$ over $\mathcal{F}_2$. Such a code is linear if it is a subspace; if k is the dimension the redundancy r is defined to be $n-k$, and the code is said to be an $[n,k]$ code. Codewords are generally denoted v, w, etc. Positions are generally denoted i, j, etc., with $1 \leq i \leq n$. As usual, v_i denotes the element of $\mathcal{F}_2$ in position i of v. A vector in $\mathcal{F}_2^n$ will be called a bit vector, of length n.

Generator matrices will be considered to be $k \times n$, and parity check matrices $n \times r$.

A vector v may be identified with its "support" $\{i : v_i = 1\}$. The operations $v \cap w$, $v \cup w$, and v^c are defined on the vectors via this identification. The Hamming weight $|v|$ of a vector is the cardinality of the support. If $S \subseteq \{1, \ldots, n\}$ the vector $v\restriction S$ is a vector of length $|S|$, defined in the obvious way. If $\sigma \subseteq \mathcal{F}_2^n$ then $\sigma\restriction S$ is defined to be $\{v\restriction S : v \in \sigma\}$.

An isomorphism $\pi : \chi \mapsto \psi$ of two codes in $\mathcal{F}_q^n$ is a permutation of the positions, such that $\pi[\phi] = \psi$, where $\pi(v)_{\pi(i)} = v_i$. An automorphism of χ is an isomorphism of χ to itself.

This paper includes the results of running various computer programs. Coding theory has long been an area of mathematics where this is commonplace ([Wagner], [Chen 70], [Simonis 87], [Chen 91], [Bou&Var]). An applicable comment may be found in the introduction to appendix 1 of [Harary]: "It is very useful to have diagrams of graphs available for the accumulation of data leading to conjectures".

Suppose χ is a code containing 0, of minimum distance at least d where $d = 2\delta + 1$ is odd. Define the A configuration determined by χ to be the set of weight d vectors. Define $A(n)$ to be the largest size of an A configuration in a code of length n. In the coding theory literature, $A(n, w, d)$ denotes the maximum size of a subset of $\mathcal{F}_2^n$, where each vector has Hamming weight w, and the Hamming distance between two vectors is at least d. The value is of interest only for even d; $A(n)$ is $A(n, d, d+1)$.

For $v \in \chi$ let $\alpha_v = \{u - v : d(u,v) = d, u \in \chi\}$. Let $\bar{A}(\chi)$ be the average over $v \in \chi$ of $|\alpha_v|$. The following "codewise" version of the Johnson bound is stated in [B&T]. For convenience a proof is given; it is a variation of the proof of the Johnson bound in [M&S].

Theorem 1. With notation as above,

$$|\chi|\left(1 + n + \binom{n}{2} + \cdots \binom{n}{\delta} + \frac{\binom{n}{\delta+1} - \binom{d}{\delta}\bar{A}(\chi)}{\lfloor n/(\delta+1) \rfloor}\right) \leq 2^n.$$

Proof: For $v \in \chi$ let $\hat{\alpha}_v = \{u \in \chi : d(u,v) = d\}$, so that $\alpha_v = \hat{\alpha}_v - v$; in particular $|\alpha_v| = |\hat{\alpha}_v|$. For $v \in \chi$ let $\theta_v = \{u \in \mathcal{F}_2^n : d(u,v) = \delta + 1\}$. Let $\sigma_i = \{u \in \mathcal{F}_2^n : d(u,\chi) = i\}$. If $u \in \theta_v$ then clearly $d(u,\chi) \leq \delta + 1$; on the other hand $d(u,\chi) < \delta$ cannot hold, else $d(v,w) < 2\delta + 1$ for some $w \in \chi$. If $u \in \hat{\alpha}_v$

let $\theta_{vu} = \{w \in \theta_v : d(u,w) = \delta\}$; by translating to v it is readily verified that $|\theta_{vu}| = \binom{d}{\delta}$. If $w \in \theta_{vu}$ then $w \in \sigma_\delta$. If $w \in \theta_v$ and $w \in \sigma_\delta$ then $w \in \theta_{vu}$ for a unique u. It follows that $|\theta_v \cap \sigma_\delta| = \binom{d}{\delta}|\alpha_v|$, whence $|\theta_v \cap \sigma_{\delta+1}| = \binom{n}{\delta+1} - \binom{d}{\delta}|\alpha_v|$. One readily verifies that if $u \in \theta_{v_1} \cap \theta_{v_2}$ then $(v_1 - u) \cap (v_2 - u) = \emptyset$, whence u can belong to at most $\lfloor n/(\delta+1) \rfloor$ of the θ_v. Consider the incidence matrix with a row for each $u \in \sigma_{\delta+1}$, and a column for each $v \in \chi$, where the matrix entry is 1 iff $u \in \theta_v$. Counting 1's by columns and rows,

$$|\sigma_{\delta+1}| \left\lfloor \frac{n}{\delta+1} \right\rfloor \geq |\chi| \left(\binom{n}{\delta+1} - \binom{d}{\delta} \bar{A}(\chi) \right).$$

Since $|\sigma_i| = |\chi|\binom{n}{i}$ for $i \leq \delta$, the theorem follows.

Suppose χ is a code with A configuration α, $v \in \alpha$, $1 \leq i \leq n$, and $t \in \mathcal{F}_2^n$ with $|t| = \delta$. Let

- $\alpha_i = \{v \in \alpha : i \in v\}$;
- $\alpha_t = \{v \in \alpha : t \subseteq v\}$;
- $\nu_v = \{w \in \alpha : |w - v| = d+1\}$;
- $\nu_{vi} = \nu_v \cap \alpha_i$, where $i \in v$;
- $\nu_{vij} = \nu_v \cap \alpha_{\{i,j\}}$, where $i, j \in v$.

If $v, w \in \alpha$ then $|v - w| \geq d+1$, so ν_v is the neighborhood of v within α. Also, $|v - w| = d+1$ iff $|v \cap w| = \delta$. A ν_v is partitioned into the 10 classes ν_{vij}, for $i, j \in v$, $i < j$; ν_{vi} is the union of 4 of these.

A number of further definitions used below will now be given.

- An A_1 configuration is any α_i occurring in some code. $A_1(\chi)$ is the maximum of $|\alpha_i|$ for an α_i in χ. $A_1(n)$ is the maximum of $|\alpha_i|$ for an α_i in a code of length n.
- An N configuration is any ν_v occurring in some code. $N(\alpha)$ is the maximum of $|\nu_v|$ for a $v \in \alpha$. $N(n)$ is the maximum of $|\nu_v|$ for a ν_v in a code of length n.
- An N_1 configuration is any ν_{vi} occurring in some code. $N_1(\alpha)$ is the maximum of $|\nu_{vi}|$ for $v \in \alpha$ and i a posdition of v. $N_1(n)$ is the maximum of $|\nu_{vi}|$ for a ν_{vi} in a code of length n.
- If χ is required to be linear, L is appended to the subscript, for A, A_1, N, or N_1, configurations or functions of n.
- Let ν'_v denote ν_v, restricted to v^c; and similarly for ν'_{vi} and ν'_{vij}.

The following theorem gives various facts about quantities just defined; some parts are given in [Dowd 88]; some are well-known.

Theorem 2.

a. ν_v is the disjoint union of the sets $\alpha_t - \{v\}$; hence

$$\sum_{t \subseteq v} |\alpha_t| = |\nu_v| + \binom{d}{\delta}.$$

b.

$$\binom{d}{\delta}|\alpha| = \sum_t |\alpha_t|.$$

c.

$$\binom{d}{\delta}^2 |\alpha| \leq \binom{n}{\delta} \left(N(\alpha) + \binom{d}{\delta} \right).$$

d. Two members of α_t intersect in t. Hence $|\alpha_t| \leq (n - \delta)/(\delta + 1)$.

e.

$$N(n) \leq \binom{d}{\delta} \left\lfloor \frac{n-d}{\delta+1} \right\rfloor.$$

f. $|\alpha| \leq (n/d)A_1(\chi)$.

g.

$$|\alpha_i| \le \frac{n-1}{d-1}\left[\frac{\delta-1}{d-1}N_1(\alpha)+1\right].$$

Proof: For part a, if $w \in \alpha_t - \{v\}$ then $w \in \nu_v$. If $w \in \nu_v$ then $w \cap v = t$ for a unique t, and $w \in \alpha_t - \{v\}$. For part b, counting pairs $\langle v, t\rangle$, the left side counts v first, and the right side counts t first. For part c, let i_t denote $|\alpha_t|$. Using parts a and b,

$$\sum_t i_t^2 = \sum_t i_t \,|\{v \in \alpha : t \subseteq v\}| = \sum_{v\in\alpha}\sum_{t\subseteq v} i_t = \sum_{v\in\alpha}|\nu_v| + \sum_t i_t.$$

For fixed $c = \sum_{v\in\alpha}|\nu_v|$, $\sum_t i_t$ is maximized, subject to $\sum_t (i_t^2 - i_t)$ equaling c, when the i_t are equal, say to the common value i, and it follows that

$$\binom{n}{\delta}(i^2 - i) = c.$$

By part b,

$$\binom{n}{\delta} i = \binom{d}{\delta}|\alpha|.$$

Finally, $c \le |\alpha| N(\alpha)$, and part c follows. For part d, by definition two members intersect in at least t, and they cannot intersect in a larger set. Part e follows by parts a and d. For part f, counting 1's in the incidence matrix of α, $\sum_i |\alpha_i| = d|\alpha|$. Since $|\alpha_i| \le A_1(\chi)$, $d|\alpha| \le nA_1(\chi)$. For part g, given an α_i, let α_i' be the incidence matrix with the common position i deleted. Let p_l be the number of rows which are 1 in column l. Let $q_j = |\nu_{vi}|$, where v is row j. Then

$$\sum_l p_l = (d-1)|\alpha_i|, \qquad (\delta-1)\sum_j q_j = 2\sum_i \binom{p_i}{2}, \qquad \text{and} \quad q_j \le N_1(\alpha).$$

Given $c = (\delta-1)\sum_j q_j$, $\sum_l p_l$ is maximized subject to $2\sum_l \binom{p_l}{2} = c$ when the p_l are all equal, say to p. At equality $2(n-1)\binom{p}{2} = c \le (\delta-1)|\alpha_i|N_1(\alpha)$ and $(n-1)p = (d-1)|\alpha_i|$; part g follows.

Some parts of the theorem can be strengthened using average rather than maximum values; this will be omitted here. From the theorem $A(n) \le (n/d)A_1(n)$, $A_L(n) \le (n/d)A_{1L}(n)$, etc.

From hereon only $\delta = 2$, $d = 5$ will be considered. In this case, $|\nu_v| \le 10\lfloor (n-5)/3 \rfloor$. The bound is achieved in a 3-(n,5,1) design; these exist for $n = 4^m + 1$ where $m \ge 1$ (see BJL, theorem 6.9). In such a design $|\alpha_t| = (n-2)/3$ for any pair t, and the claim follows. In such a design

$$|\alpha| = \frac{1}{10}\binom{n}{3} = \frac{1}{60}n(n-1)(n-2).$$

It is a question of interest how large a code exists, which has this as an A configuration.

Theorem 2.g becomes $|\alpha_i| \le (n-1)/4[N_1(\alpha)/4 + 1]$. This yields the usual bound $(n-1)(n-2)/12$ on $A_1(\chi)$ for an arbitrary code χ, when $N_1 = 4(n-5)/3$.

Configurations almost this large occur in the Preparata codes. These are defined when $n = 4^m - 1$ where $m \ge 2$. The weight 5 vectors form a 2-(n,5,$(n-3)/3$) design, and $|\alpha| = (1/60)n(n-1)(n-3)$ ([M&S], theorem 15.33). It follows that $|\nu_v| = (10/3)(n-6)$.

The Johnson bound when $d = 5$ is

$$|\chi|\left(1 + n + \binom{n}{2} + \frac{\binom{n}{3} - 10\,\bar{A}(\chi)}{\lfloor n/3 \rfloor}\right) \le 2^n.$$

Since 3-(n,5,1) designs exist, this does not yield an improvement almost everywhere to the sphere packing bound. The nonexistence of perfect codes ([M&S]) does show that the sphere packing bound for double error

correcting codes is met only finitely often. The Preparata codes (which are examples of "nearly perfect" codes) show, however, that for nonlinear codes of minimum weight 5 the sphere packing bound is nearly tight.

The Johnson bound does yield an improvement sometimes. The bound $A(n) \leq \lfloor \frac{n}{5} \lfloor \frac{n-1}{4} \lfloor \frac{n-2}{3} \rfloor \rfloor \rfloor$ is well-known ([M&S], corollary 17.5). The weaker bound $\frac{1}{10}\binom{n}{2} \lfloor \frac{n-2}{3} \rfloor$ follows by theorem 2.c and 2.e. It also follows by theorem 2.g, and the fact that for $d = 5$, $N_1(n) \leq 4 \lfloor \frac{n-5}{3} \rfloor$. The importance of theorem 2 lies in the hope that for linear codes, better bounds on $N(n)$ or $N_1(n)$ can be obtained. Indeed, a slight improvement is readily obtained.

Theorem 3. Suppose $n - 5 = 3l + t$. Then $N_{1L}(n) \leq 4l - 3$ if $n \equiv 2$ or $n \equiv 3 \bmod 6$; and $N_{1L}(n) \leq 4l - 1$ if $n \equiv 0 \bmod 6$.

Proof: For the first claim, suppose two ν_{vij} have size l, for some v, i. the sum of all the vectors in the two classes has weight at most $2t$ in the positions of v^c, and weight 2 in the positions of v. For the second claim, suppose all four ν_{vij} have size l. Let a denote the number of weight 3 columns. Counting flags, $12\,l = 4(3\,l + 1 - a) + 3\,a$, whence $a = 4$. The sum of all the rows has weight 0 in v, and weight 4 in v^c.

[B&T] contains a bound on $A_L(n)$ in the case that n is congruent to 2, 3, or 4, mod 6. In the case $n \equiv 2$, the bound is $\frac{1}{10}\left(\binom{n}{2}\frac{n-5}{3} + 1\right)$. This is better than the bound resulting from theorems 2 and 3. Indeed, [B&T] yields $A_L(254) \leq 267462$, which by the Johnson bound implies that there is no code of redundancy 15. Theorems 2 and 3 yield $A_L(254) \leq 267462$. The argument of [B&T] makes use of the fact that at most one α_t in the entire code can have size l.

If in the case $n \equiv 2$, theorem 3 could be improved to $4l - 4$, the bound on $A_L(n)$ would be improved to $\frac{1}{10}\left(\binom{n}{2}\frac{n-5}{3}\right)$. It suffices to show that there is no N_1 configuration, where there is a class of size l, and 3 classes of size $l - 1$, since as already observed there can be at most one class of size l. However, there does not seem to be an easy way of showing this.

We conjecture that in fact, theorem 3 is a weak bound.

Conjecture 4. $N_{1L}(n)$ is $\leq c_1(n - 5)$ almost everywhere, for a constant c_1 smaller than 4/3.

By theorem 2, this conjecture would yield an upper bound on $A_L(n)$, better by a constant factor than the bound for arbitrary A configurations. This in turn would yield an upper bound on the length of a linear double error correcting code of redundancy r, better by a constant factor than the sphere packing bound. We conjecture that both of these latter bounds are pessimistic, in that the constants derived from c_1 by theorems 1 and 2 are too high.

We define a partial linear space to be an incidence matrix (matrix over $\mathcal{F}_2$), where two columns are incident to at most one row. Many authors require as well that each column to be incident to at least two rows, but we do not; indeed, columns may be all 0. Note that the requirement may equally be stated as, two rows are incident to at most one column; the requirement is that no "rectangle" of 1's occur.

Theorem 5.

a. An N_1' configuration is a partial linear space, of constant row weight 3, together with a partition of the rows into 4 or fewer parts, such that in each part the rows are disjoint.
b. An N_1' configuration is an N_{1L}' configuration iff the following holds. If S is a subset of the rows, let s_j be the number of rows of S in part j, for $1 \leq j \leq 4$. Then the weight of $\sum S$ must be at least 0,3,3,4,4, according to whether the number of even s_j is 0,1,2,3,4 respectively.

Proof: For part a, as already observed, in $\nu_{vi'}$ two rows can intersect in at most one position, and in each ν_{vij} two rows must be disjoint. Conversely, positions 1-5 may be added to the start of each row, with 1 in position 1 of all rows, and 1 in position $i + 1$ of part i for $1 \leq i \leq 4$. The result has minimum distance 5. For part b, for ν_{vi} to be an N_{1L} configuration, the weight of $\sum S$ in the added positions is 0,2,2,4,4, according to whether the number of even s_j is 0,1,2,3,4 respectively. If $\sum S$ is not zero, the total weight must be at least 5, and also for $v + \sum S$.

Corollary 6. The rank of ν_{vi} equals that of ν'_{vi}.

Proof: A linear dependence in ν'_{vi} must have all even s_j, and arises from a linear dependence in ν_{vi}.

Note also that the rank of ν_{vi} or ν_v increases by 1 if v is added. This follows because any sum of even weight vectors has even weight (consider the positions of v).

The partition of theorem 5.a can be considered to be a coloring; in each column the vertices must have all different colors. Note that as a corollary of theorem 5.b, an N'_{1L} configuration must generate a weight 3 code. The latter restriction is not sufficient, though; an example will be given below. Note that a linear combination is the same thing as a set of vertices, and a column is 1 in the sum iff it contains 1 or 3 of the vertices.

The above facts give some hope that conjecture 4 represents a useful approach to improving the sphere packing bound.

Known linear double error correcting codes where $n^2 \geq r$ are uncommon. Suppose $n = 2^s + 1$ and $r = 2s$. Since $n|2^r - 1$ there is a primitive nth root of unity ξ in $\mathcal{F}_{2^r}$. Let g be the minimal polynomial of ξ, and let χ be the cyclic code with g as generator polynomial. χ is thus the code whose parity check matrix may be written $[1, \xi, \xi^2, \ldots, \xi^{n-1}]$. Some parts of the following theorem are stated in [Chen 91], citing the Hartmann-Tzeng bound as proof; we give a direct proof (which has been known to the author for some time).

Theorem 7.

a. If s is odd then χ contains a single cycle of weight 3 vectors.
b. If s is even then χ contains no weight 3 vectors.
c. χ contains no weight 4 vectors.

Proof: The map $x \mapsto x^{2^s}$ is an automorphism of $\mathcal{F}_{2^r}$ which maps ξ to ξ^{-1}. Suppose $1 + \xi^i = \xi^j$ where $i \neq 0$, $j \neq 0$, $j \neq i$. Then $1 + \xi^{-i} = \xi^{-j}$, whence

$$\xi^i + \xi^{-i} = (1 + \xi^i)(1 + \xi^{-i}) = \xi^j \xi^{-j} = 1,$$

so $1 + \xi^i + \xi^{2i} = 0$ and ξ^i is a cube root of 1. But n is divisible by 3 iff s is odd; and in this case $j = 2i$ follows, proving claims 1 and 2. Suppose $1 + \xi^i = \xi^j(1 + \xi^k)$ where $i \neq 0$, $j \neq 0$, $j \neq i$. $k \neq 0$. Then

$$\xi^i + \xi^{-i} = (1 + \xi^i)(1 + \xi^{-i}) = \xi^j(1 + \xi^k)\xi^{-j}(1 + \xi^{-k}) = \xi^k + \xi^{-k}.$$

Since $i \neq 0$, if $i = k$ then $j = 0$, contradictory to hypothesis; thus, $\xi^i + \xi^k = (\xi^i + \xi^k)^{-1}$, so $\xi^i + \xi^k = 1$, so $1 + \xi^i = \xi^{j+i}$. If s is even this is impossible, and if s is odd j must equal i, contradictory to hypothesis.

To the author's knowledge, this is the best known infinite family of linear double error correcting codes. An exhaustive computer search can be performed for cyclic double error correcting codes. We performed such a search, for $n^2 \geq 2^r$ and $n \leq 1025$. The search, which is similar to that of [Chen 70], showed that the only such codes are those of theorem 7.

Before describing the search, some facts about "cyclotomy classes" are reviewed. Some of these hold for any unit u in any commutative ring R. Let $U = \{1, u, \ldots u^{s-1}\}$ be the cyclic subgroup of the group of units, generated bu u. If $v \in R$ then $|vU|$ divides $|U|$. Indeed, let i be least such that $vu^i = v$. If $s = qi + j$ where $j < i$ then $v = vu^s = vu^i \cdots u^i u_j = vu^j$. It follows that $j = 0$. If v is a unit than clearly $|vU| = |U|$, in fact for any set U.

In the case of binary codes of length n, $R = \mathcal{Z}_n$, $u = 2$, and for any $e \in \mathcal{Z}_n$ there is a cyclotomy class C_e. Cyclic codes may be specified using the field $\mathcal{F}_{2^{r_1}}$ where $r_1 = |C_1|$. To search up to $n = 1000$ or higher this is unsatisfactory; for example for $n = 875$, $r_1 = 300$. However, for double error correcting codes where n respects the sphere packing bound, and $n^2 \geq 2^r$, there is at most one possible r. A search may be carried out for lists $e_1, \ldots, e_t$ of cyclotomy class representatives, such that $r_1 + \cdots + r_t = r$, where $r_j = |C_{e_j}|$.

To obtain the parity check matrix, the "reduced" exponent $e_i/\gcd(e_i, n)$ may be used in the field $\mathcal{F}_{2^{r_i}}$. If $n_i = n/\gcd(e_i, n)$ this part of the matrix repeats with a period of n_i.

It may be verified that all relevant parity check matrices are obtained, using facts from section 7.5 of [M&S]. Let ξ be a primitive nth root of unity in $\mathcal{F}_{2^{r_1}}$. Let $e_1, \ldots, e_t$ be a system of representatives of the cyclotomy classes. Let g_i denote the generator polynomial $\prod_{j \in C_i}(x - \xi^j)$ where $C_i = C_{e_i}$. It suffices to consider generator polynomials of the form $g_{i_1} \cdots g_{i_s}$ where the i_j are distinct. For such a polynomial the redundancy is $\sum r_{i_j}$ where $r_i = |C_i|$ (the degree of the polynomial). The parity check matrix has a "column" of width r_{i_j} for each j; its rows are the powers of $\xi^{e_{i_j}}$. Remaining details are left to the reader.

This search for $n \leq 1025$ required 32.5 seconds of CPU time on a 2GHz Pentium. Checking for weight 3 and 4 vectors makes use of a hash table of sums of pairs of rows of the parity check matrix. The primitive polynomials required were obtained using "PrimPoly" [O'Conner].

Linear double error correcting codes with $n^2 \geq 2^r$, which are not necessarily cyclic, are known.

- For $r \equiv 2 \pmod 4$, codes of length $n = 2^r$ exist; see [M&S] or [Chen 91].
- [Wagner] gives a [23,14] code.
- [Chen 91] gives $[33, 23]$ and $[47, 36]$ codes.
- [Gulliver and Bhargava] mentions a quasi-cyclic $[257, 241]$ code.

The value $n_L(r)$ is defined in [M&S] is the largest n for which there exists a linear double error correcting code of redundancy r. From Brouwer's online tables [AEB], we have the following.

r:	4	5	6	7	8	9	10	11	12	13	14	15	16
$n_L(r)$:	5	6	8	11	17	23	33	47−57	65−88	81−124	128−165	151−181	257−359

Email from Dr. Markus Grassl [Grassl] was in agreement with the lower bounds. The search for cyclic codes mentioned above failed to find any codes which would improve the lower bound for $r = 13$ or $r = 15$. As mentioned above, the upper bound for $n = 15$ follows using results from [B&T] ([AEB] states that it follows by the Johnson bound). The upper bound for 16 was obtained by the Johnson bound.

If linear double error correcting codes with $n^2 > (1+\epsilon)2^r$ exist for some $\epsilon > 0$ they have yet to be found. It is clearly of interest to determine further facts about the derived configurations defined above. Indeed, it is of interest to consider them in known codes, providing examples for study.

For example, results above indicate that good codes will be "level", meaning that the α_t (and other configurations) will have a small size spectrum. Examining configurations in known codes provides an example of the amount of symmetry, and as will be seen this might be described as "moderate".

The codes of theorem 7, with r a multiple of 4, are clearly of interest; we will call such a code Ξ_r. It is of interest to try to prove facts about these codes in general. In the absence of general facts, computations must be done for each particular r, using bit vectors for the powers of the primitive root (the use of Jacobi logarithms, q.v. see [L&N], is essentially the same method).

If n is a prime additional facts hold. The only known value for r are 8, 16, or 32, when n is the Fermat prime 17, 257, or 65537. In fact, these are the only known values where n is a prime power. By the proof of theorem II.17 of [H&W], if 2^{s2^t} for s odd is a prime power, then $2^{2^t} + 1$ is a power of the same prime. It is open whether Fermat numbers are square-free [Rivera].

For any r there are $\phi(n)/r$ classes of primitive roots, where ϕ is the Euler totient function. Any one of them can be used to generate the parity check matrix; the resulting codes are isomprphic. The automorphism group of Ξ_r (considered as acting on $\mathcal{Z}_n$) contains the group generated by $x \mapsto x + 1$ and $x \mapsto 2x$ (indeed, any binary cyclic code admits these automorphisms [M&S]). If e is relatively prime to n the stabilizer of $\{0, e\}$ contains the identity, together with $x \mapsto -x + e$.

When n is a Fermat prime, all classes are classes of primitive elements, and there are $(n-1)/r$ of them. There are $(n-1)/r$ orbits of pairs under the automorphism group. When $n = 17$ there are 2 classes. It is readily verified that C_1 is the quadratic residues, C_3 the nonresidues, and Ξ_8 is a quadratic residue code.

On the other hand, it is useful to be able to compute with bit vectors, up to some maximum value of r, for example 32, so that the rows of a parity check matrix fit in a word. Let g be a minimal polynomial of degree r, with ξ a root, and ξ of multiplicative order n.

Given a primitive polynomial, each element of $\mathcal{F}_{2^r}$ may be given as a length r bit vector. Let H_1 be the parity check matrix where row i, $0 \leq i < n$, is this bit vector for ξ^i. Let H_2 be the parity check matrix where row i is the polynomial x^i, reduced modulo g. It is well known that the codes C_1 and C_2 for H_1 and H_2 are identical. Indeed, $f \in C_1$ iff $f(\xi) = 0$ iff $g|f$ iff $f \equiv 0 \bmod g$ iff $f \in C_2$.

Given a primitive polynomial, a minimal polynomial for ξ can be obtained as follows. Using repeated squaring, compute the bit vector for $\xi = \rho^{(2^r-1)/n}$ where ρ is the primitive root, and multiplication is polynomial multiplication modulo the primitive polynomial. Compute an $(r+1)$ by r matrix where row i is the bit vector for ξ^i. Using column Gaussian elimination, reduce rows 0 to $r-1$ to a permutation matrix. The linear combination of the rows which equals 0 is readily determined, and is a minimal polynomial.

A similar method can be used to determine all minimal polynomials for ξ, once one is known. For each conjugacy class C_e where $\gcd(e, n) = 1$, write the powers $1, \xi^e, \xi^{2e}, \ldots, \xi^{re}$ as the rows of a matrix, and proceed as above. The idempotent of the code generated by the minimal polynomial can be computed, using theorems 8.6 and 8.7 of [M&S]. These do not appear to be of much interest, when Ξ_r is not the quadratic residue code Ξ_8.

One of the two minimal polynomials for Ξ_8 has weight 5, which prompts the question of whether every Ξ_r has a weight 5 minimal polynomial. Using the above enumeration, it was determined that Ξ_r for $r \leq 32$ has one when r is 4, 8, 16, 20, or 32.

If χ is a code on $\mathcal{F}_2^n$, the code may be "extended" by adding a position, which will be denoted ∞. Letting χ^e denote the extended code, for $v \in \chi$, the extended codeword v^e in χ^e is obtained by setting v^e_∞ to $\sum_i v_i$. Note that $\sum_i v_i + v^e_\infty = 0$.

It is well known (see e.g. [M&S]) that for a quadratic residue code χ (in particular Ξ_8), $PSL_2(n)$ acts doubly transitively on χ^e. The proof uses idempotents, and it does not seem to be straightforward to apply it to Ξ_{16} or Ξ_{32}, so a computer search was done, to see whether the usual action of $PSL_2(n)$ preserved the code. It does not.

We describe the computation. Given a minimal polynomial, the parity check matrix obtained by powering x (the bit vector with 1 in position 1 and 0 elsewhere) is in the form of an $r \times r$ identity matrix, with a $k \times r$ matrix below it, which we call H_1. The matrix H_1, with a $k \times k$ identity matrix to the right of it, is a generator matrix, which may be "stored implicitly", with only H_1 stored explicitly. A row of the generator matrix may be created from the row index, and transformed by the action of the "inverse" permutation. The resulting row may be checked for membership in the code by multiplying by the parity check matrix on the right.

Generators for $PSL_2(n)$ may be taken as the cycle on $\mathcal{F}_n$, the map $x \mapsto x^2$, and the map $x \mapsto x^{-1}$, which transposes 0 and ∞. In the computation, the rows may be left as length n, rather than $n+1$. In the action of the inverse permutation, the entry in position 0 is replaced by the parity check (sum) bit. The map on the other positions may be determined using a primitive root mod n ([O'Conner] contains a useful subroutine for determining this).

One can conclude some facts about doubly transitive actions using theorem 5.3 of [Cameron]. Note that no minimum distance 5 code which contains a weight 5 vector can admit an action of A_n. Making a list of the n up to 65538 on which a doubly transitive group can act, it may be seen that there is no doubly transitive action on Ξ_r^e when r is 12, 20, 24, or 28. Further facts can doubtless be proved; but we omit this.

Simonis [Simonis 92] shows that X_8^e and X_8 are unique. It is worth noting that uniqueness of χ implies a doubly transitive action on χ^e. Let $\chi_{\hat{i}}$ denote χ "punctured" at position i, that is, with position i deleted from all codewords.

If π is an isomorphism from χ_i^e to χ, let $\pi^e : \chi^e \mapsto \chi^e$ be the extension of π where $\pi(i) = \infty$. It is clear that π^e is an automorphism of χ^e. Indeed if $v \in \chi^e$, v_p is v punctured at i, $v_m = \pi(v_p)$, and v_e is the extension of v_m, then $\pi^e(v) = v_e$, because v_i and $v_{e\infty}$ both equal the parity of the other bits. Similar arguments may be found in [Simonis 00].

With the above facts providing relevant perspective, some computations on the codes Ξ_r were carried out. $A_1(\Xi_r)$ was computed for $r \leq 28$. An outline of the method is as follows; it worked readily for $r \leq 28$ (for which sufficient memory was available) and could be modified to work for $r = 32$.

Let R be a system of representatives of the nonzero cyclotomy classes. Given $i_1 \in R$, let i_2 be any element of $\mathcal{Z}_n$, except those values where ξ^{i_2} equals 1, ξ^{i_1}, or $1+\xi^{i_1}$. Also require that i_1 be no greater than the representative for the class of i_2. For each i_1, i_2, see if $1 + \xi^{i_1} + \xi^{i_2}$ has its flag set in an array of flags for the sums $\xi^{i_3} + \xi^{i_4}$ of pairs. If so, write $\langle 1 + \xi^{i_1} + \xi^{i_2}, x\rangle$ to a file, where x is the class index of i_2; and set a flag that the pair has occurred. After writing this file, write $\langle \xi^{i_3} + \xi^{i_4}, \xi^{i_3}\rangle$ to a file for each flagged pair. Sort the two files; and merge to obtain a file of triples $\langle 1 + \xi^{i_1} + \xi^{i_2}, x, \xi^{i_3}\rangle$.

Let Φ denote the group of automorphisms $\{\phi_{\alpha\beta} : \alpha \in \mathcal{Z}^r, \beta \in \mathcal{Z}_n\}$, where $\phi_{\alpha\beta}(x) = 2^\alpha x + \beta$. The final file above contains at least one representative of each orbit of weight 5 vectors under Φ. It can be used to create a file of weight 5 vectors, one from each orbit and containing 0. A simple method can be used for this, at least for $r \leq 28$, since the input file is of modest size. A useful observation is that given $S = \{i_0 = 0, i_1, i_2, i_3, i_4\}$, the sets in its orbit under Φ which contain 0 are the union of the 5 orbits under the maps $\{\phi_{\alpha 0}\}$, of the sets $S - i_j$.

$A_1(\Xi_r)$ is readily computed from the orbit representatives. The values for r from 4 to 28 are given below. Since Ξ_r is cyclic its automorphism group is transitive, and $A(\Xi_r) = (n/5)A_1(\Xi_r)$. The ratio $A(\Xi_r)/\binom{n}{3}$ is also given.

$A_1(\Xi_r)$	1	10	171	2740	43621	699390	11183551
$A(\Xi_r)/\binom{n}{3}$	.100000	.050000	.050893	.050368	.049969	.050036	.049997

In the case $r = 8$, Φ is transitive on the weight 5 vectors; for $r = 12$ the number of classes is 5, and thereafter increases with increasing r, up to the highest value computed, i.e., 28.

Suppose that $A_L(n) \leq \kappa(n^3/60) + O(n^2)$ for some constant $\kappa \leq 1$; then by theorem 1, for binary linear codes of minimum weight 5, $n + O(1) \leq \lambda \cdot 2^{r/2}$, where $\lambda = \sqrt{1/(1-\kappa/2)}$. For $\kappa = 1$, $\lambda = \sqrt{2} \approx 1.414$, the sphere packing bound. For $\kappa = 1/2$, $\lambda = \sqrt{4/3} \approx 1.155$. Figure 1 is a graph of λ as a function of κ.

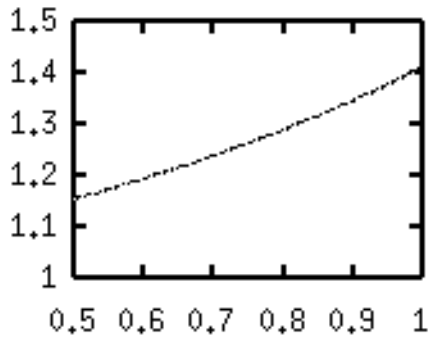

Figure 1

It is evident from the above that the Johnson bound is an overestimate. It can be seen from the proof of theorem 1 that this is due to the fact that, for $u \in \sigma_{\delta+1}$, the average number of v such that $u \in \theta_v$ is less than $n/(\delta + 1)$.

The value of $|\nu_v|$ can also be computed using the list of orbit representatives. For each representative w other than v, each $i \in v$, each $j \in w$, and each $\phi_{\alpha\beta}$ mapping j to i, determine if $\phi_{\alpha\beta}(w)$ intersects v in two positions and if so add it to a list, then eliminate duplicates.

To obtain statistics on $|\nu_v|$, the size of the orbit of v under Φ is also useful. This can be determined by noting that the stabilizer of v is the union over $i \in v$ of the sets $\{\phi(v) : \phi[v] = v \text{ and } \phi(0) = i\}$, each of which is readily computed by considering all α.

The following table gives the range, and average, of $|\nu_v|$ in Ξ_r, for r from 8 to 24; $r = 28$ was omitted because of the time it would have required. The maximum value of $|\nu_{vi}|$ is also given, followed by its ratio to $n - 5$.

$$
\begin{aligned}
r = \ 8: &\quad 16; \quad N_1 = 8\,(.6667)\\
r = 12: &\quad 80{-}107,\ \text{avg} = 98.2; \quad N_1 = 44\,(.7333)\\
r = 16: &\quad 386{-}436,\ \text{avg} = 419.9; \quad N_1 = 184\,(.7302)\\
r = 20: &\quad 1634{-}1800,\ \text{avg} = 1695.8; \quad N_1 = 720\,(.7059)\\
r = 24: &\quad 6652{-}6944,\ \text{avg} = 6821.6; \quad N_1 = 2797\,(.6835)
\end{aligned}
$$

When $r = 20$ the maximum of ν_v is significantly larger than the second largest (1748). When $r/4$ is odd then $5|n$, and the vector $\{0, n/5, \ldots, 4n/5\}$ is in the code. This vector is stabilized by the subgroup of order $5\,r$ of Φ, consisting of the $\phi_{\alpha\beta}$ where β is a multiple of $n/5$. It follows that the ν_{vij} have equal size. In the case of Ξ_{12}, ν_v for this v is the smallest; but in the case of Ξ_{20} it is the largest. In the case Ξ_{28}, the size of ν_{vij} is 2688.

Finally, the statistics of $|\alpha_t|$ can be obtained using the list of orbit representatives. A file may be created contining an entry for each v in the list, and each i in a system of conjugacy class representatives, of the elements in the orbit of v inder Φ which contain 0 and i. This file may then be sorted, and duplicates eliminated, yielding $|\alpha_t|$ for one t from each orbit of t under Φ. The size of the orbit of t is $nc/2$ where c is the size of the conjugacy class.

The following table gives the size distribution of $|\alpha_t|$. The distribution is given for $r = 8$ and $r = 12$; only the range of values is given for larger r. Note that the average of $|\alpha_t|$ equals $(10\,|\alpha|)/\binom{n}{2}$, or $4\,A_1(\chi)/(n-1)$ for cyclic codes.

$$
\begin{aligned}
r = \ 8: &\quad 2^{68}3^{68}\\
r = 12: &\quad 9^{520}10^{390}11^{390}12^{780}\\
r = 16: &\quad 39{-}47\\
r = 20: &\quad 161{-}181\\
r = 24: &\quad 662{-}703\\
r = 28: &\quad 2689{-}2773
\end{aligned}
$$

The code Ξ_8 has several properties which distinguish it among the Ξ_r. It is a quadratic residue code, it is unique, and Φ is transitive on the weight 5 vectors. For Ξ_8, $|\alpha| = 34$, $|\alpha_i| = 10$, and $|\nu_v| = 16$. For the pair 01, $|\alpha_t| = 2$, and this also holds for the pairs t in the orbit of 01 under Φ; for the remaining pairs, i.e., the orbit of 03, $|\alpha_t| = 3$.

Theorem 2.f for $d = 5$ becomes $|\alpha| \le (n/5)A_1(\chi)$. For Ξ_8, equality holds; indeed this is true for any binary code with a transitive automorphism group. Theorem 2.c for $d = 5$ becomes $100|\alpha| \le \binom{n}{\delta}(N(\alpha)+10)$. Substituting for $|\alpha|$ yields $15 \le N(\alpha)$, and the bound is not met. This is because the $|\alpha_t|$ are not equal; the sum in the proof of theorem 2.c becomes $68 * 9 + 68 * 4 = 34 * 16 + 68 * 3 + 68 * 2$.

A graph may be associated with an N configuration ν_v, where the nodes represent the positions of v, and the edge from i to j is labelled with $|\nu_{vij}|$. Figure 2 is this graph for ν_v in Ξ_8 (recall that these are isomorphic), with a single edge for size 1, and a double edge for size 2. From hereon the sizes of the ν_{vij} in a ν_{vi} will be considered to be a 4-tuple $\langle s_1, s_2, s_3, s_4\rangle$, where $s_1 \geq s_2 \geq s_3 \geq s_4$. In Ξ_8, there is a ν_{vi} with sizes $\langle 2, 2, 2, 2\rangle$; the other 4 have sizes $\langle 2, 2, 1, 1\rangle$.

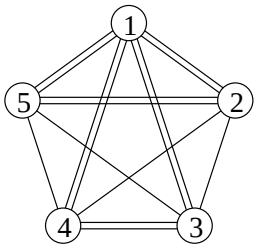

Figure 2

By examining a ν_{vi} with sizes $\langle 2, 2, 2, 2\rangle$ it may be determined that the matrix whose rows are the vectors of ν'_{vi} is the vertex-edge incidence matrix of the cube. The pairs in a ν_{vij} may be described as follows. Remove two opposite faces of the cube; and take a diagonal of each remaining face, so that no two intersect. It may also be determined that the rank of the ν'_{vi} matrix is 7.

For small values of n exhaustive searches can be carried out for N and N_1 configurations. A search may be carried out for ν_{vi}, with a particular list of part sizes, by searching sequences of triples, such that no two intersect in more than one position; and two in the same part are disjoint. Writing triples as bit vectors of length $n-5$, the first n_{01} triples can be 0x7, 0x38, ... The triples in each part can have increasing index in some order of the triples. Also, the first triple in the second part can be taken from a small set of possibilities.

As the triples are searched, the span of the corresponding weight 5 vectors can be maintained. When the span is doubled, it can be checked whether a nonzero vector v with $|v| \leq 4$ occurs. Each triple in the search sequence can be flagged, with an indication of whether its weight 5 vector doubled the span, or was already in it.

If this search is carried out with $n = 17$ and size list $\langle 2, 2, 2, 2\rangle$, all legal ν_{vi} configurations are found to have

- rank 7,
- 8 weight 5 vectors in the span, and
- 2 1's in each column.

Theorem 8. An N'_1 configuration with 2 1's in each column can occur only if $3|(n-5)$. Such a configuration has $2l$ rows where $l = (n-5)/3$. It is thus a a 1-($2l$,3,2) design, or cubic graph with $2l$ vertices. Its span is weight 3 iff each connected component is three edge connected. If it is an N'_{1L} configuration it is triangle-free.

Proof: If r is the number of rows and $c = n-5$ the number of columns then $2c = 3r$, whence $3|(n-5)$; and the configuration satisfies the definition of a cubic graph. The sum of a set S of rows (vertices) is the edges (columns) with one end in S and the other in S^c. If it contains a triangle then in a ν_{vi} from which it is derived, in v^c the weight of the sum of a triangle is 3, and in v the weight is 4. Adding v yields a weight 4 vector.

There are 6 cubic graphs on 8 vertices ([Gordon]); only 2 of them are triangle-free. Both of these occur in ν_{vi} configurations with sizes $\langle 2, 2, 2, 2\rangle$ when $n = 17$. Various 4-colorings occur.

The other possible size lists for $|\nu_{vi}| = 8$ are

4400, 4310, 4220, 4211, 3320, 3311, 3221.

Exhaustive search shows that none exist with these sizes (some cases can readily be proved by hand). It follows that the maximum value for ν_{vi} is 8; no ν_{vi} of size 9 could be an extension of a 2222, because the code would be Ξ_8.

The next value of n for which $n-5$ is divisible by 6 is 23. There is a unique optimal code [Simonis 00] of this length, and dimension 14, called the Wagner code after its discoverer.

The distribution of $|\alpha_t|$ is $1^6 2^{18} 3^{136} 4^{75} 5^{18}$. The distribution of $|\nu_v|$ is $24^{36} 26^{48}$; the counts for each sorted sequence of $|\nu_{vi}|$ are as follows.

10,10,10, 9, 9	8
11,10, 9, 9, 9	8
11,10,10,10, 7	4
11,11,10, 8, 8	4
12, 9, 9, 9, 9	2
12,10, 9, 9, 8	8
12,11,11, 7, 7	2
12,10,10,10,10	16
12,12,10,10, 8	32

The average of $|\nu_{vij}|$ is 2.51.

There are 92 ν_{vi} of size 12. For each, consider the size list, the rank, and the count of weight i columns for i from 0 to 4. The frequencies of these data are as follows.

3,3,3,3	10	2,0,12,4,0:	4
3,3,3,3	11	0,4,10,4,0:	6
3,3,3,3	11	1,2,11,4,0:	8
3,3,3,3	11	1,4,8,4,1:	2
4,3,3,2	11	0,5,9,3,1:	8
4,3,3,2	12	0,3,12,3,0:	16
4,3,3,2	12	0,5,8,5,0:	24
4,3,3,2	12	0,6,7,4,1:	8
4,4,2,2	11	1,0,16,0,1:	8
4,4,2,2	12	0,5,8,5,0:	8

A search was done for an N_1 configuration ν_{vi} with $n = 23$ and sizes 4,3,3,3. The search was run for 1/2 hour and none were found. A search for one with sizes 4,4,3,2 ran to completion, and also failed to find any.

To conclude, we make some further observations on arbitrary N_{1L} configurations.

Theorem 9. Suppose in an N'_{1L} configuration, columns i and j have weights w_1 and w_2, where $w_1 w_2$ equals 20, 30, 40, 41, or 42. Then any 1 of column i may be moved in its row to column j, and the result is still an N'_{1L} configuration.

Proof: If a linear combination does not include the flipped row then the weight of the sum doesn't change. If it does, then the weight decreases only if the weight of the sum of all the rows is 11 in columns i and j; and this cannot occur in the cases of the theorem.

Conjecture 10. There is no N'_{1L} configuration where evey column has weight at least 3.

By theorem 9, conjecture 10 implies that conjecture 4 holds, with $c_1 \leq 1$. In the notation of figure 1, if $c_1 \leq 1$ then $\kappa \leq 3/4$, and $\lambda \leq \sqrt{8/5} \approx 1.265$.

If $3|(n-5)$ and $l = (n-5)/3$, it has been observed above that a cubic graph cannot occur as an N'_{1L} configuration with sizes $\langle l, l, 0, 0\rangle$ when l is odd, and can occur with sizes $\langle l/2, l/2, l/2, l/2\rangle$ when l is even. Some other observations include the following.

- $K_{3,3}$ provides an example of an N_1' configuration with a weight 3 span, which is not an N_{1L}' configuration.
- The graph obtained from two copies of $K_{3,3}$ by replacing an edge of each by edges joining their endpoints properly is a 12,12-bipartite cubic graph, which is not 3 edge connnected (example courtesy of Doug West [West]).
- If it could be shown that a cubic graph with sizes $\langle l, l, 0, 0\rangle$ which is an N_{1L}' configuration can have only sufficiently few vectors added and remain an N_{1L}' configuration, then $|\nu_{vi}| \leq 4l - 3$ when $n \equiv 5 \bmod 6$ would follow.
- When $n = 17$ no vertex can be added to either cubic graph, for any vertex partition.
- The claim that l vectors can't be added is a special case of conjecture 9. Such a tripartite configuration is a bipartite cubic graph in each pair of parts.

The facts presented here concerning some configurations in binary linear codes of minimum weight 5 show that they are of interest, and should be studied further. Although the evidence is not conclusive, it is suggestive that the constant c_1 of conjecture 5 might be under .75. In the notation of figure 1, if $c_1 \leq 3/4$ then $\kappa \leq 9/16$, and $\lambda \leq \sqrt{32/23} \approx 1.18$.

References.

[BJL] T. Beth, D. Jungnickel, and H. Lenz, *Design Theory*, Cambridge University Press, 1993.

[Bou&Var] I. G. Bouyukliev and Z. G. Varbanov, " Some results for linear binary codes with minimum distance 5 and 6", IEEE Transactions on Information Theory 51(12): 4387-4391 (2005)

[B&T] A. Brouwer and L. Tolhuizen, "A Sharpening of the Johnson Bound for Binary Linear Codes and the Nonexistence of Linear Codes with Preparata Parameters", Designs, Codes and Cryptography 3 (1993), 95–98.

[Cameron] P. Cameron, "Finite permutation groups and finite simple groups", Bull. London Math. Soc. 13 (1981), 1–22.

[Chen 70] C. L. Chen, "Computer results on the minimum distance of some binary cyclic codes", IEEE Trans. Inform. Theory 16 (1970), 359–360.

[Chen 91] C. L. Chen, "Construction of some binary linear codes of minimum distance five", IEEE Trans. Inform. Theory 37 (1991) 1429–1431.

[Dowd 88] M. Dowd, "Questions related to the Erdos-Turan conjecture", SIAM Journal in Discrete Mathematics 1 (1988), 142–150.

[Grassl] M. Grassl, private communication.

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

[Harary] F. Harary, *Graph Theory*, Addison-Wesley, 1971.

[H&W] G. Hardy and M. Wright, *An Introduction to the Theory of Numbers*, Oxford University Press, 1968.

[M&S] F. MacWilliams and N. Sloane, *The Theory of Error-Correcting Codes*, North Holland, 1977.

[L&N] R. Lidl and H. Niederreiter, *Finite Fields*, Addison-Wesley Publishing Co., 1983.

[O'Conner] http://www.seanerikoconnor.freeservers.com

[Rivera] http://www.primepuzzles.net

[Royle] http://people.csse.uwa.edu.au/gordon/remote/cubics/

[Simonis 87] J. Simonis, "Binary even [25,15,6] codes do not exist" IEEE Trans. Inform. Theory 33 (1987) 151–153.

[Simonis 92] J. Simonis, "The [18, 9, 6] code is unique", Discrete Mathematics 106/107 (1992) 439–448.

[Simonis 00] J. Simonis, "The [23, 14, 5] Wagner code is unique", Discrete Mathematics 213 (2000), 269–282.

[Wagner] T. J. Wagner, "A search technique for quasi-perfect codes", Information and Control 9 (1966) 94–99.

[West] D. West, private communication.