Minimum distance bounds for s-regular codes

Dedicated to Jaap Seidel on the occasion of his 80th birthday

William J. Martin

Department of Mathematics and Statistics
University of Winnipeg
Winnipeg, Canada R3B 2E9
William.Martin@UWinnipeg.ca

Abstract. A code $C \subseteq F^n$ is s-regular provided, for every vertex $x \in F^n$, if x is at distance at most s from C then the number of codewords $y \in C$ at distance i from x depends only on i and the distance from x to C. If ρ denotes the covering radius of C and C is ρ-regular, then C is said to be completely regular.

Suppose C is a code with minimum distance d, strength t as an orthogonal array, and dual degree s^*. We prove that $d \leq 2t+1$ when C is completely regular (with the exception of binary repetition codes). The same bound holds when C is $(t+1)$-regular. For unrestricted codes, we show that $d \leq s^* + t$ unless C is a binary repetition code.

1 Codes and their parameters

Let $q \geq 2$, $F = \{0, 1, \ldots, q-1\}$ and $n \geq 1$. The *Hamming graph* $H(n, q)$ has vertex set $X = F^n$ and distance given by the Hamming metric

$$\mathrm{dist}(x, y) = |\{h : x_h \neq y_h\}|.$$

Let A denote the adjacency matrix of this graph and E_j the matrix representing orthogonal projection onto the eigenspace V_j belonging to eigenvalue $\theta_j = n(q-1) - qj$ $(0 \leq j \leq n)$. Cf. [1, Theorem 9.2.1].

To any *code* $C \subseteq X$ $(|C| > 1)$, we associate its *outer distribution matrix* D with rows indexed by X and columns indexed by $\{0, \ldots, n\}$ and with (x, i) entry given by

$$D_{xi} = |\{y \in C : \mathrm{dist}(x, y) = i\}|,$$

i.e., the number of *codewords* at distance i from x. The rank of D is s^*+1 where s^* is the *dual degree* of C. The distance of $x \in X$ to C is

$$\mathrm{dist}(x, C) = \min\{\mathrm{dist}(x, y) : y \in C\}.$$

The *minimum distance* of C is

$$d = \min\{\mathrm{dist}(x, y) : x, y \in C,\ x \neq y\}$$

and the *covering radius* of C is given by

$$\rho = \max\{\mathrm{dist}(x, C) : x \in X\}.$$

The following inequalities are well-known [2, Sec. 5.2.3], [5]:

$$\rho \leq s^*, \qquad d \leq 2\rho + 1.$$

A code $C \subseteq X$ is *completely regular* [2, p67] provided for each $x \in X$ and each $0 \leq i \leq n$, D_{xi} depends only on i and $\mathrm{dist}(x, C)$. That is, D has only $\rho + 1$ distinct rows.

Theorem 1 (Delsarte). *If $d \geq 2s^* - 1$, then C is completely regular.*

Codes with $d = 2s^* + 1$ are *perfect* and codes with $d = 2s^*$ or $d = 2s^* - 1$ are *uniformly-packed.* These important classes have been well-studied (see [1, Sec. 11.1] and, e.g., [5]).

A code C is *s-regular* [5] provided the entries D_{xi} depend only on i and $\mathrm{dist}(x, C)$ whenever $\mathrm{dist}(x, C) \leq s$.

Theorem 2 ([2, Theorem 5.11], [5]). *If $d \geq s^*$, then C is $(d - s^*)$-regular.*

Every code forms a (possibly trivial) orthogonal array. List the codewords as rows of a $|C| \times n$ array of symbols over F. We say that C forms an *orthogonal array* of strength u if, with any choice of u columns, all q^u possible u-tuples over F occur equally often. In this case, each u-tuple clearly occurs $\lambda = |C|/q^u$ times. The largest integer u for which this condition holds is called the *strength* of C and is denoted by t. When C is an additive code, $t + 1$ is the minimum distance of the dual code $C^\perp$.

Theorem 3 (Delsarte [2]). *The strength of a code C with characteristic vector χ is equal to the smallest non-negative integer t such that $E_{t+1}\chi \neq 0$.*

2 The Hamming lattice

Let $\hat{F} = F \cup \{*\}$. The elements of $\hat{F}^n$ are called *subwords.* Now consider the partially ordered set $\mathcal{L} = (\hat{F}^n, \preceq)$ where $x \preceq y$ if and only if, for all coordinates i, $y_i \neq x_i$ implies $x_i = *$. Then $\mathcal{L}$ is a regular meet semilattice [3]. This is called the *Hamming lattice.* The *support* of $x \in \mathcal{L}$ is $\mathrm{supp}(x) = \{i : x_i \neq *\}$; the *rank* of x is $\mathrm{rk}(x) = |\mathrm{supp}(x)|$ and the *weight* of x is $\mathrm{wt}(x) = |\{i : x_i \in F - \{0\}\}|$. The j^{th} *fibre* of $\mathcal{L}$ is the set of all objects having rank j:

$$\mathcal{L}_j = \{x \in \mathcal{L} : \mathrm{rk}(x) = j\}.$$

A rank one subword $s \in \mathcal{L}_1$ will be written $s = bj$ where $b \in F$ and $1 \leq j \leq n$. This indicates that $s_j = b$ and $s_k = *$ for $k \neq j$. As usual, $\vee$ will denote the join in $\mathcal{L}$ with the understanding that $s \vee t$ is sometimes undefined. Let W_j denote the incidence matrix of the j^{th} fibre versus the top fibre X.

Lemma 1. *Every eigenvector $v \in V_j$ $(j > 0)$ can be uniquely represented in the form*

$$v = \xi W_j$$

with the conditions

$$\sum_{\substack{u \preceq s \\ \mathrm{supp}(s) = \mathrm{supp}(u) \cup \{h\}}} \xi_s = 0$$

for each $u \in \mathcal{L}_{j-1}$ and each $h \notin \mathrm{supp}(u)$.

Proof. For a subword s of rank k, let w_s denote the row of W_k indexed by s. Let s be any subword of rank $j - 1$. Then for any fixed coordinate $h \notin \mathrm{supp}(s)$,

$$\sum_{b=0}^{q-1} w_{s \vee bh} = w_s.$$

As $w_s \perp V_j$ [3, Eq. (9)], we therefore have the *balance equations*

$$\sum_{b=0}^{q-1} \xi_{s \vee bh} = 0.$$

Let M be the matrix with rows indexed by all ordered pairs (s, h) where s is a subword of rank $j-1$, $h \notin \mathrm{supp}(s)$ and columns indexed by all subwords z of rank j. In row (s, h), column z, we place a one if $s \preceq z$ and $\mathrm{supp}(z) = \mathrm{supp}(s) \cup \{h\}$ and a zero otherwise. Our balance equations can then be expressed as

$$M\xi = 0.$$

Let $\mathbf{A}$ denote the adjacency matrix of the Hamming graph $H(j, q)$. With the columns of M grouped according to support, it is not hard to see that

$$M^T M = \bigoplus_{i=1}^{\binom{n}{j}} [jI + \mathbf{A}].$$

Of course, $\mathbf{A}$ has smallest eigenvalue $-j$ with multiplicity $(q-1)^j$. Thus the nullity of M is $\binom{n}{j}(q-1)^j$ as claimed. $\square$

3 Determining a code from local information

Let $C \subseteq X$. Then X is partitioned according to distance from C. Let

$$C_i = \{x \in X : \mathrm{dist}(x, C) = i\}$$

and let χ_i denote the characteristic vector of C_i as a subset of X. Neumaier [7] proved that C is a completely regular code if and only if the partition $\pi = \{C_0, \ldots, C_\rho\}$ is equitable: every vertex in C_i is at distance one from a constant number of vertices in C_j. The matrix B of constants is clearly tridiagonal:

$$B = \begin{pmatrix} \alpha_0 & \beta_0 & & 0 \\ \gamma_1 & \alpha_1 & \beta_1 & \\ & \ddots & \ddots & \beta_{\rho-1} \\ 0 & & \gamma_\rho & \alpha_\rho \end{pmatrix}$$

Each eigenvalue of B is an eigenvalue of $H(n, q)$ [4, Section 5.2]. The largest is $\tau_0 = n(q-1)$ and the second largest is $\tau_1 = \theta_{t+1}$. Let $\omega = [\omega_0, \ldots, \omega_\rho]$ be a right eigenvector of B with eigenvalue τ_1 having $\omega_0 > 0$. Then we have [6]

$$\omega_0 > \omega_1 > \cdots > \omega_\rho. \tag{1}$$

Now if P is the 01-matrix with i^{th} column equal to χ_i, then $AP = PB$, so that ω lifts to $P\omega$ with

$$A(P\omega) = P(B\omega) = \theta_{t+1}(P\omega).$$

Since C is completely regular, the column space of D has $\{\chi_0, \ldots, \chi_\rho\}$ as a basis, and thus $P\omega \in \operatorname{colsp}(D)$ forcing $P\omega = E_{t+1}\chi$ (up to multiplication by a positive scalar).

The remainder of the paper is devoted to analysing this eigenvector as a function on subwords of rank $t+1$. Assume C is a completely regular code in $H(n,q)$ with characteristic vector χ and having strength t. Let $v = E_{t+1}\chi = P\omega$ and write $v = \xi W_{t+1}$ as in Lemma 1.

For $x \in X$ and $T \subseteq \{1, \ldots, n\}$ with $|T| = t+1$, let $\xi_{x_{|T}}$ denote the s-entry of ξ where $s \preceq x$ is the subword obtained from x by replacing each x_i, $i \notin T$, by $*$. Since

$$v_x = \sum_{w \preceq x} \xi_w = \sum_{|T|=t+1} \xi_{x_{|T}} = \omega_i \tag{2}$$

for $x \in C_i$, knowledge of the ξ-values gives us C.

Theorem 4. *If C is a completely regular code, then either C is a binary repetition code or $d \leq 2t+1$.*

Proof. For $x, y \in X$, the *interval* determined by x and y is the set

$$[x,y] = \{z \in X : \operatorname{dist}(x,z) + \operatorname{dist}(z,y) = \operatorname{dist}(x,y)\}.$$

The induced subgraph on $[x,y]$ is a binary Hamming graph.

Assume $0 \in C$. Let U be a fixed set of $t+1$ coordinates. For a word y with support U and weight $t+1$, we find

$$\sum_{z \in [0,y]} (-1)^{\mathrm{wt}(z)} \xi_{z_{|U}} = \sum_{z \in [0,y]} (-1)^{\mathrm{wt}(z)} v_z = \sum_{h=0}^{t+1} (-1)^h \binom{t+1}{h} \omega_h,$$

since, for each $R \neq U$, all terms of the form $\xi_{z_{|R}}$ cancel. For a given set U, there are $(q-1)^{t+1}$ such equations, which we will call the *cube equations*. Let τ denote the common right-hand side value of these equations.

Since each cube equation involves a unique subword of weight $t+1$, they are independent. Moreover, the left-hand side of each such equation can be viewed as an eigenvector for $H(t+1,q)$ with eigenvalue $-(t+1)$ (where the coordinate set $\{1,\ldots,n\}$ is replaced by U). Thus each such vector is orthogonal to the left-hand side of each balance equation. In summary, the balance equations together with the cube equations give us $\binom{n}{t+1}q^{t+1}$ linearly independent equations which uniquely determine the values ξ_s for $\mathrm{supp}(s)=U$. Since U was arbitrary, we are able to determine the vector ξ, which, in turn, determines v, the distance partition and C.

Straightforward computation verifies that the unique solution to the above linear system is $\xi_s=\eta_h$ for all subwords s having support U and weight h, where

$$\eta_h=\tau(1-q)^{-h}\left(\frac{q-1}{q}\right)^{t+1}.$$

Now C is an orthogonal array of strength t. So there exists a codeword x which differs from 0 in at least t of the coordinates in U. This forces $\eta_t=\eta_0$ or $\eta_{t+1}=\eta_0$. Thus $q=2$. Moreover, the analysis is independent of the choice of U, so we may write down ξ in its entirety. Finally, using the inequality (1) and the identities (2), we see that the codewords maximize the sum $\sum\xi_{x_{|T}}$ over $x\in X$. Now it is easy to see that C must be a binary repetition code. □

Looking back at the above analysis, we notice that the hypothesis that C is completely regular is unnecessarily strong. In order to obtain the constant τ, we merely required that v_x depend only on $\mathrm{dist}(x,C)$ for $\mathrm{dist}(x,C)\le t+1$. Thus it is enough that C is $(t+1)$-regular.

Corollary 1. *If C is s-regular for some $s\ge t+1$, then $d\le 2t+1$ unless C is a binary repetition code.* □

Corollary 2. *For any code C, $d\le s^*+t$, unless C is a binary repetition code.*

Proof. If $d-s^*\ge t+1$, then, C is $(t+1)$-regular by Theorem 2. So either C is a binary repetition code or $d\le 2t+1$ from the previous corollary. But this implies $s^* \leq t$ and $d \geq 2s^*+1$, which forces C to be a perfect code.

The perfect codes over a prime-power alphabet are known: (d, s^*, t) belongs to the set

$$\left\{(7,7,3),(5,5,2),(3,1,q^{m-1})\right\}$$

where $m \geq 2$. The only possible exception here is the case $q = m = 2$, again a binary repetition code.

If q is not a prime power, then any perfect code must have $d \in \{3,5\}$ [1, p355]. If $d = 3$, then $\theta_{t+1} = -1$. Thus, $t = k(q-1)$, and no exceptions arise.

Finally, suppose $d = 5$. Then Theorem 4 gives $t \geq 2$. So $t = 2$ is the only case of concern. Hence we assume that the quotient matrix

$$B = \begin{pmatrix} 0 & n(q-1) & 0 \\ 1 & q-2 & (n-1)(q-1) \\ 0 & 2 & n(q-1)-1 \end{pmatrix}$$

has θ_3 as its second largest eigenvalue. This yields the equation

$$n(q-1) - 3q = \frac{q-4+\sqrt{1+(4n+q-7)(q-1)}}{2}.$$

Eliminating the radical, we find

$$(2n(q-1) - 7q + 4)^2 = 1 + (4n+q-7)(q-1).$$

Reduce both sides modulo $q-1$ to see that $q-1$ divides 8. But then q is a prime power and all perfect codes are known. $\square$

The following codes have $d = s^* + t$:

- Any code with $d = n$ (e.g., all repetition codes);
- the ternary Hamming code of length four;
- $C \cong F' \times F^{n-1}$ where $F' \subseteq F$ and $1 \leq |F'| < q$;
- all words of even weight in $H(3,2)$;
- $C = \{000000, 000111, 111000, 111111\}$ in $H(6,2)$.

Repetition codes with $n \leq 3$ or $q = 2$ are completely regular; the remaining codes with $d = n$ are not. All other examples in the list are uniformly packed. Any code not listed having $d = s^* + t$ must either be uniformly packed with $5 \leq d \leq 8$ or fail to be completely regular.

Acknowledgments

This work was completed while the author was visiting the Department of Combinatorics and Optimization at the University of Waterloo. The author wishes to thank the department for its hospitality and accommodation during this visit. The author's research is supported by the Canadian government through NSERC grant number OGP0155422. Additional support was provided by MITACS and CITO through the Centre for Applied Cryptographic Research.

References

1. A. E. Brouwer, A. M. Cohen and A. Neumaier. Distance-Regular Graphs. Springer-Verlag, Berlin (1989).
2. P. Delsarte. An algebraic approach to the association schemes of coding theory. *Philips Res. Repts. Suppl.* **10** (1973).
3. P. Delsarte. Association schemes and t-designs in regular semilattices. *J. Combin. Th., Ser. A* **20** (1976), 230–243.
4. C. D. Godsil. Algebraic Combinatorics. Chapman and Hall, New York (1993).
5. J. M. Goethals and H. C. A. van Tilborg. Uniformly packed codes. *Philips Res. Repts.* **30** (1975), 9–36.
6. W. J. Martin. Completely regular designs of strength one. *J. Alg. Combin.* **3** (1994), 177–185.
7. A. Neumaier. Completely regular codes. *Discrete Math.* **106/107** (1992), 353–360.