The Minimum Distance of Graph Codes

Høholdt, Tom; Justesen, Jørn

Published in:
Lecture Notes in Computer Science

Link to article, DOI:
10.1007/978-3-642-20901-7_12

Publication date:
2011

The Minimum Distance of Graph Codes

Tom Høholdt[1] and Jørn Justesen[2]

[1] Department of Mathematics, The Technical University of Denmark, Bldg. 303 DK 2800 Lyngby Denmark, T.Hoeholdt@mat.dtu.dk
[2] jorn@justesen.info

Abstract. We study codes constructed from graphs where the code symbols are associated with the edges and the symbols connected to a given vertex are restricted to be codewords in a component code. In particular we treat such codes from bipartite expander graphs coming from Euclidean planes and other geometries. We give results on the minimum distances of the codes.

Keywords : Graph codes, Euclidean and projective geometry.

1 Introduction

In 1981 Tanner [1] introduced a construction of error-correcting codes based on graphs and since then a considerable number of results have been obtained [2], [3], [4], [5] and [6]. The recent textbook by Roth [8] contains a thorough presentation of the subject. In this paper we consider some classes of graph codes and in particular codes from bipartite expander graphs based on finite geometries. In this case the vertices of the graph are labeled by the points and lines of a finite geometry, and there is an edge connecting a line vertex to any vertex labeled by a point on the line. The code symbols are associated with the edges, and the symbols connected to a given vertex are restricted to be codewords in a component code over the field that is used for constructing the geometry.

In Section 2 we recall the construction of the codes and we give basic bounds on their parameters. In Section 3 the lower bound on the minimum distance is improved by considering the properties of eigenvectors of the adjacency matrix of the graph. In Section 4 we specialize to bipartite graphs from finite geometries, and in Section 5 we show that the bound obtained is tight for a special class of graph codes. Section 6 contains the conclusion.

2 Basic Parameters and Bounds

We recall the construction of codes based on graphs.

2.1 General n-regular graphs

Let $G = (V, E)$ be an n-regular connected graph, without loops and multiple edges, with vertex set V and edge set E. Let $|V| = m_1$, $|E| = \frac{m_1 n}{2} = L$ and let C_1 be a (n, k, d) code over the finite field $\mathbb{F}_q$. We now construct a code C of length L over $\mathbb{F}_q$ by associating $\mathbb{F}_q$ symbols with the edges of the graph (in some selected order) and demanding that the symbols connected to a vertex in V shall be a codeword in C_1. It is clear that C is a linear code of length L and if we let K denote the dimension of C we have that $L - K \leq m_1(n - k)$ and therefore

Lemma 1. *The rate $R = \frac{K}{L}$ satisfy*

$$R \geq 2r - 1 \text{ , where } r = \tfrac{k}{n} \text{ is the rate of the component code.}$$

In the following we shall use the *adjacency matrix* matrix of the graph so we recall the definition:

Definition 1. *Let $z_1, z_2, \ldots, z_{m_1}$ be the vertices of the graph G. The adjacency matrix $A = (a_{ij}), i, j = 1, 2, \ldots, m_1$ is defined by*

$$a_{ij} = \begin{cases} 1 \text{ if } z_i \text{ is connected to } z_j \\ 0 \qquad\qquad \text{else} \end{cases}$$

2.2 Bipartite graphs

With bipartite graphs the construction is as follows. Let $G = (V, E)$ be an n-regular connected bipartite graph, without multiple edges, with vertex set $V = V_1 \cup V_2$ such that $V_1 \cap V_2 = \varnothing$ and $|V_1| = |V_2| = m$. A bipartite graph is n-regular if each vertex of V_1 is connected to n vertices of V_2, and each vertex of V_2 is connected to n vertices of V_1.

Let C_1 be a linear (n, k_1, d_1) code and C_2 a linear (n, k_2, d_2) code both over the finite field $\mathbb{F}_q$. We now construct a code $\mathscr{C}$ of length $L = mn$ over $\mathbb{F}_q$ by associating $\mathbb{F}_q$ symbols with the edges of the graph and demanding that the symbols connected to a vertex of V_1 shall be a codeword of C_1 and that the symbols on the edges connected to a vertex of V_2 shall be a codeword of C_2. More formally we assume an ordering of the edges E of G and for a vertex $u \in V$ let $E(u)$ denote the set of edges incident with u . For a word $\boldsymbol{x} = (x_e)_{e \in E}$ in ${\mathbb{F}_q}^L$ denote by $(\boldsymbol{x})_{E(u)}$ the subword of $\boldsymbol{x}$ that is indexed by $E(u)$, that is $(\boldsymbol{x})_{E(u)} = (x_e)_{e \in E(u)}$. Then the code $\mathscr{C}$ is defined by

$$\mathscr{C} = \{\boldsymbol{c} \in {\mathbb{F}_q}^L : (\boldsymbol{c})_{E(u)} \in C_1 \text{ for every } u \in V_1 \text{ and } (\boldsymbol{c})_{E(u)} \in C_2 \text{ for every } u \in V_2\}$$

It is clear that $\mathscr{C}$ is a linear code. Let K be its dimension. We recall from [3]

Lemma 2. *The rate* $R = \frac{K}{L}$ *of* $\mathscr{C}$ *satisfies*

$$R \geq r_1 + r_2 - 1 \quad \textit{where} \quad r_1 = \frac{k_1}{n} \quad \textit{and} \quad r_2 = \frac{k_2}{n}$$

Proof: The number of linearly independent parity checks is at most $m(n - k_1) + m(n - k_2)$, so $L - K \leq m(n - k_1) + m(n - k_2)$ and since $L = mn$ we get the result. □

Let $x_1, x_2, \ldots, x_m$ be the vertices in V_1 and $y_1, y_2, \ldots, y_m$ the vertices in V_2 and define the $m \times m$ matrix $M = m_{ij}$ by

$$m_{ij} = \begin{cases} 1 \text{ if } x_i \text{ is connected to } y_j \\ 0 \qquad\qquad\quad \text{else} \end{cases}$$

The *adjacency matrix* of the bipartite graph is then

$$A = \begin{pmatrix} 0 & M \\ M^T & 0 \end{pmatrix}$$

2.3 Bounds on the minimum distance

In both cases above we have that each row of A has n 1s, the largest eigenvalue of A is n, and the corresponding eigenvector is the all-ones vector. In the bipartite case also $-n$ is an eigenvalue of A, and the corresponding eigenvector has 1s in the first half of the positions and -1 in the rest. It is known [8] that for a connected graph $-n \leq \lambda_i \leq n$ where λ_i is any eigenvalue and that the second largest eigenvalue λ is closely related to the expansion properties of the graph. Large random graphs and known families with good expansion properties have $\lambda = 2\sqrt{n-1}$ [7]. We quote the following bounds on the minimum distances.

Theorem 1. *The minimum distance* D *of* C *satisfies*

$$D \geq dm_1 \frac{d - \lambda}{2(n - \lambda)} \tag{1}$$

Theorem 2. *The minimum distance* D *of* $\mathscr{C}$ *if* $d_1 = d_2 = d$ *satisfies*

$$D \geq dm \frac{d - \lambda}{n - \lambda} \tag{2}$$

This bound was obtained by Sipser and Spielman in [6], and a slight modification gives the bound in Theorem 1. For proofs see e.g. [8], Chapter 13.

For a complete n-regular graph, where $m_1 = n + 1$, then $\lambda = -1$, the bound of Theorem 1 is

$$D \geq \frac{d(d+1)}{2}$$

which indeed is the right bound for these codes.

We also note that in the case where the bipartite graph is complete, and hence $n = m$ and $\lambda = 0$, we get the usual bound for product codes.

For short component codes, where $d \leq \lambda$, the bound is not useful, but we can get a simple lower bound by the following consideration: Any vertex corresponding to a nonzero codeword on the right side is incident with at least d nonzero edges connecting to vertices in the left set, and these reach at least $d(d-1)$ vertices in the right set with nonzero edges. If the girth of the graph is at least 6, these vertices are distinct, and the minimum distance is always lower bounded by

$$D \geq d(d(d-1)+1) = d(d^2-d+1) \tag{3}$$

If the girth is $g \geq 6$, the argument can be repeated to give

$$D \geq d(1 + d(d-1)\textstyle\sum_{i=0}^{\frac{g/2-3}{2}} (d-1)^{2i}) \text{ if } g/2 \text{ is odd}$$

$$D \geq d(d \sum_{i=0}^{\frac{g/2-2}{2}} (d-1)^{2i}) \tag{4}$$

if $g/2$ is even. This bound also appears in Skachek's thesis [11].

The potential of graph codes is related to the possibility of keeping the component code fixed while the size of the graph increases. In this way the performance can be improved with only a linear increase in decoding complexity. However, for the codes C and $\mathscr{C}$ to have a reasonable rate, the component codes must have high rate, and to get a positive bound from Theorem 1 and Theorem 2, the minimum distance of the component codes has to be larger than λ. The combination of these requirements tends to make the resulting code too long for any realistic application. Thus our emphasis in this paper is to improve the analysis of codes of moderate block length derived from specific good graphs.

3 Improved Lower Bounds on the Minimum Distance

In Section 5 we demonstrate that the bound of Theorem 2 is actually tight in certain cases. However, in some cases of interest, it is possible to get sharper lower bounds. As a first case we consider component codes of different rates. Even though the resulting bound on the minimum distance for fixed overall rate is maximized by choosing component codes with equal distance (see comment after the proof of Theorem 3), the performance with practical decoding algorithms is improved by using unequal distances (as in the case of product codes). Several generalizations of (2) to unequal distances were presented in [9], and [10].

Theorem 3. *The minimum distance D of $\mathscr{C}$ satisfies*

$$D \geq m d_1 \frac{d_2 - \lambda\beta}{n - \lambda\beta} \tag{5}$$

where β is the positive root of

$$\beta^2(\alpha n - d_1) + \beta\lambda(1 - \frac{d_1}{d_2}) + d_1 - n = 0$$

Proof: Let $\grave{E}$ be a set of edges in G that supports a nonzero codeword of C. Let S be the subset of vertices in V_1 incident with $\grave{E}$ and let T be the subset of vertices in V_2 incident with $\grave{E}$. We will get the bound on D from a bound of $|\grave{E}|$. We follow the standard line of proof by defining a vector v as a modified indicator vector for the sets S and T, and then apply a well-known result (see e.g. [8], Lemma 13.6)

$$v^T A v \leq \lambda v^T v \tag{6}$$

Equality holds if and only if v is an eigenvector associated with the eigenvalue λ. We obtain improved bounds by adjusting the coordinates of v to values that are consistent with the the properties of an eigenvector.

Suppose that $|S| = a$ and $|T| = \alpha a$, $\alpha \geq 1$, and let e be the average valency of the vertices in S, thus $\frac{e}{\alpha}$ the average valency in T. Let $v = (v_i)$ be a vector of length $2m$ where

$$v_i = \begin{cases} 1 & \text{if } i \in S \\ -\frac{a}{m-a} & \text{if } i \in V_1 \setminus S \\ \beta & \text{if } i \in T \\ -\frac{\alpha\beta a}{m-\alpha a} & \text{if } i \in V_2 \setminus T \end{cases}$$

where $0 < \beta \leq 1$. By balancing v we assure that the inner product of v with the eigenvectors associated with the largest numerical eigenvalue n is 0. Since the multiplicity of n is one it follows that v is in the space spanned by the eigenvectors of A that are associated with the remaining eigenvalues of A. We can directly calculate the left side of (6) since we know that the number of edges connecting S and T is ae, and thus also the number of edges connecting S and $V_2 \setminus T$, namely $(n-e)a$. Therefore the number of edges connecting $V_1 \setminus S$ and T is $a\alpha(n - e/\alpha)$ and the remaining edges connect $V_1 \setminus S$ and $V_2 \setminus T$. We therefore get

$$v^T A v = \tfrac{2ma\beta}{(m-a)(m-\alpha a)}(me - na\alpha)$$

The inequality (6) and this result give

$$\tfrac{2ma\beta}{(m-a)(m-\alpha a)}(me - na\alpha) \leq \lambda(a + (m-a)\tfrac{a^2}{(m-a)^2} + a\alpha\beta^2 + (m-\alpha a)\tfrac{\alpha^2\beta^2a^2}{(m-\alpha a)^2})$$

and this by a straightforward calculation leads to the following bound on a

$$a \geq \frac{m}{\alpha}\frac{2e\beta - \lambda(1+\alpha\beta^2)}{2\beta n - \lambda(1+\beta^2)} \tag{7}$$

which holds for any positive β. The lower bound on a is met if and only if v is an eigenvector associated with the eigenvalue λ, i.e. $Av = \lambda v$, and a necessary condition for this, where we only look at the upper part of A is

$$\lambda = e\beta - (n-e)\frac{a\alpha\beta}{m - a\alpha} \quad \text{and} \quad \beta\lambda = \frac{e}{\alpha} - (n - \frac{e}{\alpha})\frac{a}{m-a}$$

These two conditions lead to the following expressions for a

$$a = m\frac{e\beta - \lambda}{\alpha(\beta n - \lambda)} \tag{8}$$

$$a = m\frac{\frac{e}{\alpha} - \beta\lambda}{n - \beta\lambda} \tag{9}$$

and by eliminating a we get the equation for β. It can be seen that there is a positive solution less than 1.

Maximizing the right side of (7) with respect to β actually leads to the same equation. Thus this is the sharpest lower bound that can be obtained by this method. The bound can be met if there is a subgraph on S and T with exactly valencies e and $\frac{e}{\alpha}$ (which we expect will rarely be the case). Since $D \geq ea$ the lower bound increases with a and e, we thus get a new lower bound by choosing $\alpha = \frac{e}{d_2}$ since the bound on a decreases with α and then choosing $e = d_1$. □

The bound in Theorem 3 improves the bounds obtained in [9] and [10]. They are respectively

$$D \geq \tfrac{m}{n}(d_1 d_2 - \tfrac{\lambda}{2}(d_1 + d_2)) \text{ where } d_1 \geq d_2 > \tfrac{\lambda}{2}.$$

and

$$D \geq m\tfrac{d_1 d_2 - \lambda\sqrt{d_1 d_2}}{n-\lambda}$$

The comparisons are facilitated by using the approximation $\beta \approx 1/\sqrt{\alpha}$, which can also be used to prove that for fixed rate, i.e. $d_1 + d_2$ fixed, the lower bound is maximum for $d_1 = d_2$.

Example 1. As a case where Theorem 3 gives simple numbers we may take $n = 16, \lambda = 4, d_1 = 8, \alpha = 2$, and consequently $\beta = 2/3$. From (5) we get $D \geq 4m/5$ compared to $m/2$ and $0.78m$ for the earlier bounds.

For Theorem 2 or 3 to hold with equality, the edges connecting vertices in S to $V_2 \setminus T$ must be equally distributed over these vertices (and similarly for edges connecting T to $V_1 \setminus S$). Clearly this is usually not possible because of the integer constraints. In the proof of the following theorem we modify v by distinguishing between the subsets of vertices that are connected to S or T and the remaining vertices. For simplicity we only treat the symmetric case $d_1 = d_2 = d$.

We shall first derive the coordinates of a hypothetical eigenvector corresponding to sets S and T of minimal (equal) size. To get a useful bound for smaller d we denote the set of vertices in $V_2 \setminus T$ that are connected to S as U_2 and the set of vertices in V_2 not in T or U_2 as W_2. Similarly V_1 is divided into S, U_1, and W_1. The eigenvector v' is assumed to have coordinates 1 in positions corresponding to S and T, u in positions corresponding to U_1 and U_2, and w in the remaining positions. We get the smallest value of a by assuming that the $a(n-d)$ edges from S reach distinct vertices in U_2, and that this is consequently the size of the set. It now follows from the assumption that v' is an eigenvector with eigenvalue λ that $u = (\lambda - d)/(n - d)$. Further $|W_1| = f = m - a(n - d + 1)$,

and since the vector has to be balanced, $w = a(d - \lambda - 1)/f$. Let the number of edges connecting a vertex in U_1 to vertices in W_2 be g. The number of such edges incident with a vertex in W_2, h, then follows. We get the final condition by applying the eigenvalue calculation to a vertex in U_1:

$$1 + (n - g - 1)\frac{\lambda - d}{n - d} + gw = \lambda\frac{\lambda - d}{n - d} \tag{10}$$

The remaining parameter in v' should be selected to minimize a for a given value of m. The minimum is always on the boundary of the range $0 \leq g \leq n - 1$ and $0 \leq h \leq n$. The condition $h = n$ applies for $\lambda + 1 < d$ down to a value close to $d = \lambda$. For smaller d the limit is $g = n - 1$. Both conditions hold when

$$\lambda^2 + \lambda(n - d) - d(n - 1) = 0 \tag{11}$$

which clearly has a solution $\lambda = d - \epsilon$ for a small positive ϵ.

From the properties of such a potential eigenvector we get the following lower bound:

Theorem 4. *The minimum distance of C is lower bounded by $D \geq da$ where*

$$m/a \leq 1 + n - d + \frac{(n - 1)(n - d)(\lambda - d + 1)}{n - d - \lambda^2 + d\lambda} \tag{12}$$

for $(\lambda^2 - n)/(\lambda - 1) < d \leq \lambda + \epsilon$ and

$$m/a \leq 1 + n - d + \frac{(n - d)\lambda(\lambda - d + 1)}{n(d - \lambda)} \tag{13}$$

for $\lambda + 1 \geq d \geq \lambda + \epsilon$, where ϵ is a positive number derived below.

Proof: The expression (13) in the Theorem follows from (10). However to arrive at a solution with positive parameters in the other case we must assume $(\lambda^2 - n)/(\lambda - 1) < d$, which also ensures that the denominator in (12) is positive. To prove that the eigenvectors give actual lower bounds on the minimum distance of the code we assume that a minimum weight codeword defines S and T as before. We then construct the vector v using the value u from the eigenvector and choose the value w' of the remaining coordinates to get a balanced vector. For $d > \lambda$ we minimize the number of additional vertices, f, by letting each have $h = n$ connections to W_2. The result then follows by choosing w to get a balanced vector. For $d = \lambda$ we get $u = 0$, and from $h = n$ we directly get $m/a \leq 2n - d$. For $d < \lambda$ the vector v is inserted in the inequality (5), we get an inequality for a/m which depends on the parameter corresponding to g. The minimum is again always on the boundary. Thus the only remaining variable is a. Calculating the two sides of (5) we find

$$v^T Av = ad + 2a(\lambda - d) + a(n - g - 1)(\lambda - d)^2/(n - d) + 2ag(\lambda - d)w + f(n - h)w^2$$

and

$$\lambda v^T v = \lambda a + \lambda a(\lambda - d)^2/(n - d) + \lambda f w^2.$$

The terms containing a factor w or fw^2 vanish for small a/m. We can then reduce the inequality by the factor $(\lambda - d)$, which is positive. We then find that the inequality (5) cannot be satisfied for very small a as long as $(\lambda^2 - n)/(\lambda - 1) < d$, and consequently A cannot have eigenvalue λ. The smallest value of a that lets (5) be satisfied gives equality and thus v is the eigenvector.

□

Theorem 4 improves on (2) and (3) when the graph is large and d is close to λ as demonstrated in Example 4. As a case of particular interest we mention that for $d = \lambda$, $D \geq dm/(2n - d)$. For very small d, the approach could be extended by considering additional subsets of vertices (reached from S and T in several steps), but the improvements would only apply to very long codes.

For a general n-regular graph we similarly split the set of vertices into S, U, and W, and the same derivation gives lower bounds on the minimum distance that are half the values of (12) and (13).

4 Expander Graphs from Geometries

Certain bipartite graphs derived from generalized polygons have good expansion properties [4], and hence the codes derived from these have large minimum distances. The generalized polygons are incidence structures consisting of points and lines where any point is incident with the same number of lines, and any line is incident with the same number of points. A generalized N-gon, where N is a natural number, defines a bipartite graph $G = (V, E)$ that satisfies the following conditions:

- For all vertices $u, v \in V$, $d(u, v) \leq N$, where $d(u, v)$ is the length of the minimum path connecting u and v.
- If $d(u, v) = s < N$, then there is a unique path of length s connecting u and v.
- Given a vertex $u \in V$ there exists a vertex $v \in V$ such that $d(u, v) = N$.

We note that this implies that the girth of the bipartite graph is at least $2N$. Most of this paper is concerned with graphs from finite planes, and in this context the 3-gons are derived from finite projective planes. (The definition and properties of these can be found in [14], Chapter 2).

Let M be an incidence matrix for a projective plane over $\mathbb{F}_q$ with $m = q^2 + q + 1$ points with homogeneous coordinates $(x : y : z)$ and $q^2 + q + 1$ lines with homogeneous coordinates $(a : b : c)$ where a point is incident with a line if $ax + by + cz = 0$. The bipartite graph then has adjacency matrix

$$A = \begin{pmatrix} 0 & M \\ M^T & 0 \end{pmatrix}$$

The graph is invariant to an interchange of the two sets of variables $(x : y : z)$ and $(a : b : c)$.

Thus each row of A has $q+1$ 1s so the largest eigenvalue is $q+1$ and the corresponding eigenvector is the all-ones vector. The graph may be seen as a simple expander graph: The eigenvalues are $\pm(q+1)$ and $\pm\sqrt{q}$ (all real since A is symmetric). (See [4].)

Starting from a vertex in the right set, $q+1$ vertices in the left set can be reached in one transition, and $q(q+1)$ vertices in the right set can be reached from these vertices. The graph can be used to define a code by associating a symbol with each edge and letting all edges that meet in a vertex satisfy the parity checks of an (n,k,d) code, where $n=q+1$. Thus the length of the total code is

$$L = mn = (q^2+q+1)(q+1)$$

It is sometimes more convenient to let M be an incidence matrix for an Euclidean plane [14] with $m=q^2$ points, (x,y), and q^2 lines of the form $y=ax+b$. The lines of the form $x=c$ are omitted, and in this way the graph is invariant to an interchange of the two sets of variables.

Thus each row of the adjacency matrix has q 1s and the eigenvalues are $\pm q$, $\pm\sqrt{q}$ and 0 [4].

All edges that meet in a vertex satisfy the parity checks of an (n,k,d) code with $n=q$. Thus the length of the code is

$$L = q^3$$

Example 2. For $q=4$, the projective plane and $(5,3,3)$ component codes give codes of length $L=5\cdot 21=105$. The minimum distance is lower bounded by (2) and (3), which in this case give the same value

$$D \geq 21\cdot 5\cdot (d-2)/(n-2) = 21$$

A subgraph with 7 vertices of degree 3 on each side can be found as a binary sub-plane, and for this reason the lower bound is tight. The rate is lower bounded by $R \geq 2\cdot 3/5 - 1 = 1/5$, but later we shall see that the actual dimension is 29. If the vertices are labeled $(x:y:z)$ and $(a:b:c)$ where the last nonzero coordinate is chosen to be 1, the vertex $(\alpha:1:1)$ is connected to the 5 vertices $(\alpha^2:1:0),(\alpha^2:0:1),(0:1:1),(\alpha:\alpha:1),(1:\alpha^2:1)$. We can find a basis for the component codewords by evaluating z^2, yz, and y^2.(See e.g. [16] p. 69). Thus the generator matrix of the component code becomes

$$G = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & \alpha & \alpha^2 \\ 1 & 0 & 1 & \alpha^2 & \alpha \end{pmatrix}$$

In particular the codeword $(1,1,1,0,0)$ is part of the binary sub-plane.

It is possible to construct longer codes from generalized N-gons, but it is known [12] that for $N>6$ there are no N-gons with degree $q+1$.

5 Minimum Distances of Codes from Geometries

For the specific codes constructed from graphs derived from finite geometries it is possible to get tighter bounds on the minimum distances, and in some cases we can determine the exact value. Such results provide some insight into the structure of the code and the tightness of the bounds. The use of RS component codes also serves to allow a combination of good rates and distances for moderate code lengths.

When $q = 2^r$, the field $\mathbb{F}_q$ contains a subfield with $q' = 2^s$ symbols whenever s divides r. With the chosen coordinates for the projective plane, the component extended RS code has $q' + 1$ positions with coordinates in the subfield. If the minimum distance of the component code is $q' + 1$, it has a codeword which is 1 in these positions and zero otherwise.

The projective plane contains a subfield projective plane over ${\mathbb{F}_q}'$. With our choice of coordinates, such a plane may be found by taking the vertices that have coordinates in the subfield and the edges incident to these vertices. It now follows from the remark above that if $q' + 1$ is the minimum distance of the component code, the graph code has a codeword which is 1 on the edges corresponding to the subfield plane and zero otherwise.

Since $\mathbb{F}_2$ is a subfield of any field of characteristic 2, there is always a sub-plane with 7 points and lines, and thus for $d = 3$, the minimum distance of the graph code is ≤ 21. In this case the lower bound (3) is satisfied with equality and 21 is the actual minimum distance. Similarly the lower bound

$$(q' + 1)(q'^2 + q' + 1)$$

is reached by a codeword on the sub-plane whenever the component code has $d = q' + 1$.

For $q = 2^{2r}$, $\mathbb{F}_{2^r}$ is a subfield, and in this case the codeword in the subfield plane has weight satisfying both of the bounds (2) and (3). Thus it is seen that this is the case where the two bounds coincide. Actually we have the more general result:

Theorem 5. *For $q = 2^{2r}$ and any d, $2^r + 1 \leq d \leq 2^{2r}$ there is a graph code with generalized RS component codes such that the minimum distance D satisfies Theorem 2 with equality.*

Proof: It is well known that we can order the points of the projective plane in a cyclic way as powers of a non-primitive element of ${\mathbb{F}_q}^3$. Similarly, within this sequence the powers of an element of order $2^{2r}+2^r+1$ are the points in a subfield plane (although these are not the points that have subfield coordinates). The cosets of this cyclic subgroup are other versions of the smaller projective plane. It follows that each line in the original plane is a line in one of the subplanes and has exactly one point in each of the other subplanes. Thus by combining the required number of these cosets we can get graphs of any required degree. By assigning symbols to the edges and choosing the appropriate scaling of the symbols in the component codes, we get a codeword with the weight indicated by Theorem 2. □

Example 3. For $q = 16$, the projective plane and component codes of length 17 give codes of length $L = 4641$. The minimum distance is lower bounded by

$$D \geq 21d(d-4)$$

On each side of the graph, the vertices can be divided into a set of 21 vertices corresponding to the points of a subplane over $q = 4$, and 12 shifts of this set. From unions of such sets we can construct the balanced eigenvectors needed for the lower bound on the minimum distance to be tight. Thus at least for some choice of the mapping of component code symbols on the edges of the graph, the lower bound is tight for $d \geq 5$.

In the Euclidean plane, we get a slightly higher value of the bounds for $d = \sqrt{q}+1$, but (2) does not give an integer value. The configuration of d^2-d+1 points and lines, which support minimum weight codewords in the projective planes, do not exist in Euclidean planes. Thus for $d = \sqrt{q}$ we get $a \geq q - \sqrt{q}+1$, but the bound is not tight. However, there may be codewords supported by the $q-1$ nonzero points of a subplane. Theorem 4 gives $a = \frac{m}{2n-d} = \frac{q^2}{2q-\sqrt{q}}$, which is clearly weaker in this case.

Bipartite graphs derived from generalized quadrangles produce longer codes from small component codes. Thus the bound of Theorem 4 may be of interest for such codes.

Example 4. Consider the generalized quadrangle over F_8. In this case there are 585 nodes on each side of the graph. The second eigenvalue is $\sqrt{2q} = 4$. For $d = 3$, the bound (4) gives at least 15 nonzero vertices, and a codeword of this weight can be constructed by taking the F_2 subset of the graph. For $d = 4$ the same bound gives 40 vertices, but from Theorem 4with $n = 9$ we find that at least $\lceil m/14 \rceil = 42$ vertices are nonzero. In this case the integer constraints are not directly satisfied, and a corresponding eigenvector cannot exist, whereas with $a = 45$ it may be possible to get a construction similar to that in the proof of Theorem 4 with $|W| = 7a$.

6 Conclusion

We have derived a new bound on the minimum distance of some graph codes and have analyzed some of these when the underlying graph comes from a finite geometry.

7 Acknowledgement

This work was supported in part by Danish Resarch Council Grant 272-07-0266.

References

1. M. Tanner: "A Recursive Approach to Low Complexity Codes", *IEEE Trans.Inform.Theory*, vol.27, pp. 533-547, September 1981.
2. G. Zémor:"On expander codes", *IEEE Trans.Inform.Theory* (Special Issue on Codes on Graphs and iterative Algorithms), vol.47, pp. 835-837, Februry 2001.
3. A. Barg and G. Zémor: "Error exponents of expander codes", *IEEE Trans.Inform.Theory*, vol.48, pp. 1725-1729, June 2002.
4. M. Tanner: "Explicit Concentrators from Generalized N-Gons", *SIAM J.Alg.Disc.Meth.*, Vol.5, No.3, pp. 287-293, September 1984.
5. M. Tanner: "Minimum-Distance Bounds by Graph Analysis" *IEEE Trans.Inform.Theory*, vol.47, pp.808-821, February 2001.
6. M. Sipser and D.A. Spielman: "Expander Codes", *IEEE Trans.Inform.Theory*, vol.42 No.6, pp. 1710-1722, November 1996.
7. G. Davidoff, P. Sarnak and A. Valette: *Elementary Number Theory, Group Theory, and Ramanujan Graphs*, London Mathematical Society, Student Texts 55, 2003.
8. R.M. Roth: *Introduction to Coding Theory*. Cambridge: Cambridge University Press, 2006.
9. H. Janwa and A.K. Lal: "On Tanner Codes: Minimum Distance and Decoding", *AAECC*, vol.13, pp.335-347, 2003.
10. R.M. Roth and V. Skachek: "Improved Nearly-MDS Expander Codes", *IEEE Trans.Inform.Theory*, vol.52, No.8, pp.3650-3661, August 2006.
11. V. Skachek, "Low-Density-Parity-Check Codes: Constructions and Bounds", Ph.D. Thesis, Technion, Haifa, Israel, January 2007.
12. W.Feit and G.Higman: "The nonexistence of certain generalized polygons", *J.Algebra*, vol.1, pp. 114-131, 1964.
13. H.van Maldeghem: *Generalized Polygons*. Basel, Switzerland: Birkhäuser-Verlag, 1998.
14. I. A. Blake and R. C. Mullin: *The Mathematical Theory of Coding*, New York: Academic Press, 1975.
15. N. Lauritzen: *Concrete Abstract Algebra* Cambridge: Cambridge University Press, 2005.
16. R. E. Blahut: *Algebraic Codes on Lines, Planes, and Curves* Cambridge: Cambridge University Press 2008.
17. T. Høholdt and J. Justesen: "Graph codes with Reed-Solomon component codes", *Proceedings ISIT 2006*, pp. 2022-26, Seattle, Washington, July, 2006.