# On the number of minimal codewords in codes generated by the adjacency matrix of a graph

Sascha Kurz
sascha.kurz@uni-bayreuth.de
Mathematisches Institut, Universität Bayreuth, Germany

**Abstract**

Minimal codewords have applications in decoding linear codes and in cryptography. We study the number of minimal codewords in binary linear codes that arise by appending a unit matrix to the adjacency matrix of a graph.

## 1 Introduction

Given a linear code its minimal codewords are those whose supports, i.e., the set of nonzero coordinates, do not properly contain the support of another nonzero codeword. They have applications e.g. in secret sharing schemes [9, 13], two-party computation [10], and decoding algorithms [2, 12]. Complete decoding is an NP-hard problem [7], so that it is no surprise that determining the set of minimal codewords is also a hard problem. Their number can grow exponentially in the dimension or the length of the code. The cases where all codewords are minimal are called minimal codes (or intersecting codes in the binary situation). They have e.g. applications in combinatorics [16]. Indeed, the set of minimal codewords is only known for a few classes of linear codes, including $q$-ary Hamming codes, see e.g. [6, 1]. For Reed–Muller codes the problem is only partially solved, see e.g. [8, 15] and the references cited therein.

Here we consider the concatenation of a unit matrix and the adjacency matrix of a graph as a generator matrix of a linear code and study the sets of minimal codewords. For some graph classes we can characterize the sets of minimal codewords and count them. We can fully solve the problem for complete multipartite graphs, paths, and cycles. We also state some lower and upper bounds for the number of minimal codewords in terms of graph parameters. For small numbers of vertices we determine the maximum and minimum number of minimal codewords of connected graphs. It turns out that the minimum number of minimal codewords is always attained by paths. In [3] graphs where associated to linear codes via their cycle space and the corresponding sets of minimal codewords are studied.

## 2 Preliminaries

An $[n,k]_q$ linear code $C$ is a $k$-dimensional subspace of $\mathbb{F}_q^n$. Given a vector $x \in \mathbb{F}_q^n$, the support of $x$ is defined as $\operatorname{supp}(x) = \{i : x_i \neq 0, 1 \leq i \leq n\}$. A $k \times n$ matrix $G$ whose rows form a basis for $C$ is called a generator matrix. If $G = [I_k|A]$, where $I_k$ is the $k \times k$ identity matrix, then it is said to be systematic or in standard form. A nonzero codeword $c \in C$ is minimal if there does not exist a

nonzero codeword $c'$ such that $\mathrm{supp}(c') \subsetneq \mathrm{supp}(c)$. Otherwise (including the case $c = \mathbf{0}$), we call the codeword $c$ non-minimal. Note that a codeword and its nonzero scalar multiples have the same support. We say that two codewords are equivalent if one is a scalar multiple of the other. We use the notation $M(C)$ for the number of non-equivalent minimal codewords of $C$. In the following we will mainly consider binary codes. For some known properties of minimal codewords we refer the interested reader e.g. to [6] and the references cited therein.

Given a graph $\mathcal{G} = (V, E)$ with vertex set $V$ and edge set $E$, we denote by $C(\mathcal{G})$ the binary linear code that is generated by $[I_{\#V}|A]$, where $A$ is an adjacency matrix of $\mathcal{G}$. As a shorthand we use the notation $M(\mathcal{G})$ for $M(C(\mathcal{G}))$. Note that $M(\mathcal{G})$ does not depend on the labeling of the vertices. In the remaining part of the paper we will give brief definitions for most of the used notions from graph theory. For standard definitions like e.g. a connected graph or its connectivity components we refer the reader to some standard text book on graph theory like e.g. [17].

First we observe that we can restrict ourselves to the study of $M(\mathcal{G})$ for connected graphs:

**Lemma 2.1.** *Let $\mathcal{G} = (V, E)$ consist of $r \geq 1$ connectivity components $\mathcal{G}_1 = (V_1, E_1)$, $\mathcal{G}_r = (V_r, E_r)$, i.e., $\cup_{i=1}^r V_i = V$ and $\cup_{i=1}^r E_i = E$. Then, $M(\mathcal{G}) = \sum_{i=1}^r M(\mathcal{G}_i)$.*

*Proof.* The statement is obvious from the direct sum $C(\mathcal{G}) = \oplus_{i=1}^r C(\mathcal{G}_i)$. □

**Lemma 2.2.** *For every graph $\mathcal{G} = (V, E)$ we have $M(\mathcal{G}) \geq \#V$.*

*Proof.* If $G = [I_{\#V}|A]$ is a generator matrix of $C(\mathcal{G})$, then it is easy to check that the $\#V$ rows of $G$ give minimal codewords. □

This trivial lower bound is attained with equality for graphs without edges, i.e., $M((V, \emptyset)) = \#V$. In order to obtain a more interesting problem, we define $m(n)$ as the minimum of $M(\mathcal{G})$, where $\mathcal{G}$ is a connected graph with $n$ vertices, i.e., we ask for the minimum number of minimal codewords a graph with $n$ vertices can give. Similarly, let $M(n)$ denote the maximum of $M(\mathcal{G})$, where $\mathcal{G}$ is a graph, not necessarily connected, with $n$ vertices.

More generally, let $M_q(n, k)$ be the maximum and $m_q(n, k)$ the minimum of $M(C)$ for all $[n, k]_q$ codes $C$. Bounds and some exact values on $M_q(n, k)$ and $m_q(n, k)$ can be found in [2, 4, 5, 6, 11]. Obviously, we have

$$m_2(2n, n) \leq m(n) \leq M(n) \leq M_2(2n, n).$$

Let $C$ be a linear $[k + t, k]_2$ code with systematic generator matrix $G$. By $g^i$ we denote the $i$th row of $G$, where $1 \leq i \leq k$. For each subset $S \subseteq \{1, \ldots, k\}$ let $c^S$ denote the sum of the rows of $G$ with indices in $S$, i.e., $c^S = \sum_{i \in S} g^i \in C$. For each codeword $c \in C$ let $c_S \in \mathbb{F}_2^k$ denote the systematic part of $c$, i.e., the restriction of $c$ to the first $k$ coordinates $c_1, \ldots, c_k$. Similarly, for each codeword $c \in C$ let $c_I \in \mathbb{F}_2^t$ denote the information bits, i.e., the restriction of $c$ to the last $t$ coordinates $c_{k+1}, \ldots, c_{k+t}$. Next, we study some properties of minimal codewords in general binary linear codes.

**Lemma 2.3.** *Let $\emptyset \neq S \subseteq \{1, \ldots, k\}$. If there exists a subset $\emptyset \neq T \subsetneq S$ with $c_I^T = \mathbf{0}$, then $c^S$ is non-minimal.*

*Proof.* Since $\mathrm{supp}\left(c_I^{S \setminus T}\right) = \mathrm{supp}\left(c_I^S\right)$ and $\mathrm{supp}\left(c_S^{S \setminus T}\right) \subsetneq \mathrm{supp}\left(c_S^S\right)$, we have $\mathrm{supp}\left(c^{S \setminus T}\right) \subsetneq \mathrm{supp}\left(c^S\right)$. □

**Lemma 2.4.** *Let $\emptyset \neq S \subseteq \{1, \ldots, k\}$. The codeword $c^S$ is non-minimal iff there exists a subset $\emptyset \neq T \subsetneq S$ with $\mathrm{supp}(c_I^T) \subseteq \mathrm{supp}(c_I^S)$.*

*Proof.* Since $S \neq \emptyset$ we have $c^S \neq \mathbf{0}$. Thus, if $c^S$ is non-minimal, there exists a subset $\emptyset \neq T \subsetneq S$ with $\mathrm{supp}(c^T) \subsetneq \mathrm{supp}(c^S)$, so that $\mathrm{supp}(c_I^T) \subseteq \mathrm{supp}(c_I^S)$. For the other direction let $\emptyset \neq T \subsetneq S$ with $\mathrm{supp}(c_I^T) \subseteq \mathrm{supp}(c_I^S)$. If $\mathrm{supp}(c_I^T) \neq \mathrm{supp}(c_I^S)$, then $\mathrm{supp}(c_I^T) \subsetneq \mathrm{supp}(c^S)$ implies $\mathrm{supp}(c^T) \subsetneq \mathrm{supp}(c_I^S)$ so that $c^S$ is non-minimal by definition. If $\mathrm{supp}(c_I^T) = \mathrm{supp}(c_I^S)$, then $c_I^{S \setminus T} = \mathbf{0}$ and we can apply Lemma 2.3. □

**Corollary 2.5.** *Let $c^S$ be a minimal codeword. Then, we have $1 \leq \#S \leq t+1$. Moreover, if $\#S = t+1$, then $c_I^S = \mathbf{0}$.*

*Proof.* The largest cardinality of a set of linearly independent vectors in $\mathbb{F}_2^t$ is $t$. Thus, if $\#S \geq t+1$, then there exists a subset $T \subseteq S$ with $c_I^T = \mathbf{0}$ and $\#T \leq t+1$. We finally apply Lemma 2.3 to conclude $\#S \leq t+1$. □

**Lemma 2.6.** *Let $\emptyset \neq S \subseteq \{1, \ldots, k\}$ be a subset such that $c_I^S = \mathbf{0}$. Then, $c^S$ is minimal iff $c_I^T \neq \mathbf{0}$ for all $\emptyset \neq T \subsetneq S$.*

*Proof.* Since $S \neq \emptyset$ we have $c^S \neq \mathbf{0}$. If $c^S$ is non-minimal, then there exists a subset $\emptyset \neq T \subsetneq S$ with $\mathrm{supp}(c^T) \subsetneq \mathrm{supp}(c^S)$. Since $c_I^S = \mathbf{0}$ this implies $c_I^T = \mathbf{0}$. For the other direction we apply Lemma 2.3. □

We have already observed that $c^S$ is minimal for all subsets $S \subseteq \{1, \ldots, n\}$ of cardinality 1. In a code $C(\mathcal{G})$ obtained from a graph $\mathcal{G}$ also the case of cardinality $\#S = 2$ can be characterized easily:

**Lemma 2.7.** *Let $\mathcal{G} = (V, E)$ be a graph and $C = C(\mathcal{G})$ be its associated code. For $S = \{v_1, v_2\}$ the codeword $c^S$ is minimal iff $v_1$ and $v_2$ have a common neighbor $v_3$ (where we assume that the vertices $v_1$, $v_2$, and $v_3$ are pairwise different).*

*Proof.* If $v_1$ and $v_2$ do not have common neighbors, then $\mathrm{supp}\big(c^{\{v_1\}}\big) \subsetneq c^S$, so that $c^S$ is non-minimal. If $v_1$ and $v_2$ have a common neighbor $v_3$ then $c_I^S$ has a one at position $v_3$ while $c^{\{v_1\}}$ and $c^{\{v_2\}}$ have a one at position $v_3$, so that $c^S$ is minimal. □

A path between two vertices $u$ and $v$ is a sequence of distinct vertices $[v_0, \ldots, v_l]$, such that $v_0 = u$, $v_l = v$, and $\{v_i, v_{i+1}\}$ is an edge for all $0 \leq i < l$. Such a path is called a shortest path if $l$ is minimal. We also call $l$ the length of the path. In a connected graph the length of the shortest path between two vertices gives a metric, i.e., the distance between two vertices is the length of a shortest path connecting them. The diameter of a connected graph is the maximum distance between pairs of vertices. Graphs of diameter 1 are called complete graphs and we will determine the corresponding number $M(\mathcal{G})$ of minimal codewords in Proposition 3.1. For graphs with diameter 2 we have:

**Corollary 2.8.** *For a graph $\mathcal{G} = (V, E)$ with diameter 2 we have $M(\mathcal{G}) \geq \binom{\#V+1}{2} - \#E$.*

*Proof.* For each subset $S \subseteq V$ of cardinality 1 the codeword $c^S$ is minimal, which gives $\#V$ minimal codewords. Now consider the $\binom{\#V}{2}$ subset $S = \{u, v\}$ of cardinality 2. If $\{u, v\}$ is not an edge in $\mathcal{G}$, then $u$ and $v$ are at distance 2 in $\mathcal{G}$. In other words, $u$ and $v$ have a common neighbor, so that we can apply Lemma 2.7 to deduce the minimality of $c^S$. Since $\#V + \binom{\#V}{2} = \binom{\#V+1}{2}$, we obtain the stated lower bound. □

## 3 The value of $M(\mathcal{G})$ for some graph classes

Given the number $n$ of vertices, we will always set $V = \{1, \dots, n\}$ in this subsection. By $K_n$ we denote the complete graph on $n$ vertices, i.e., $V = \{1, \dots, n\}$ and $E = \{\{i, j\} : 1 \leq i < j \leq n\}$. Obviously, we have $M(K_1) = 1$ and $M(K_2) = 2$.

**Proposition 3.1.** *For each integer $n \geq 3$ we have*

$$M(K_n) = 2^{n-1} + \binom{n}{2}.$$

*Proof.* For some subset $\emptyset \neq S \subseteq \{1, \dots, n\}$ we can easily describe $c_I^S$. If $\#S \equiv 0 \pmod 2$, then $c_I^S$ equals 1 at position $v$ iff $v \in S$ and 0 otherwise. In the other case, $\#S \equiv 1 \pmod 2$, we have that $c_I^S$ equals 1 at position $v$ iff $v \notin S$ and 0 otherwise. So, if the cardinality of $S$ is even and at least four, then we can choose at subset $T \subsetneq S$ of cardinality $T$ with $\operatorname{supp}(C_I^T) \subsetneq \operatorname{supp}(c_i^S)$, i.e., the codeword $c^S$ is non-minimal. If $\#S = 2$, then we can apply $n \geq 3$ and Lemma 2.7 to deduce that $c^S$ is minimal. This gives $\binom{n}{2}$ cases of minimal codewords.

If $\#S = 1$, then $c^S$ is minimal, which amounts to $n = \binom{n}{1}$ cases. Now let $\#S$ be odd and at least 3. We have $c_I^S = \mathbf{0}$ iff $S = \{1, \dots, n\}$. The only proper subset $T$ with $c_I^T = \mathbf{0}$ is $T = \emptyset$. Now let $\emptyset \neq T \subsetneq S$. If $\#T \equiv 1 \pmod 2$, then $\operatorname{supp}\left(c_I^T\right) \not\subseteq \operatorname{supp}(c_I^S)$, since $T \backslash S \neq \emptyset$. If $\#T \equiv 0 \pmod 2$, then $\operatorname{supp}\left(c_I^T\right) \not\subseteq \operatorname{supp}(c_I^S)$, since $T \neq \emptyset$. Thus, $c^S$ is minimal and there are $\sum_{1 \leq i \leq n \,:\, i \text{ odd}} \binom{n}{i}$ cases in total.

Thus, we have $M(K_n) = \binom{n}{2} + \sum_{1 \leq i \leq n \,:\, i \text{ odd}} \binom{n}{i}$, which can be simplified further. Since $\sum_{i=0}^{n} \binom{n}{i} = 2^n$ and $\sum_{i=0}^{n} (-1)^i \binom{n}{i} = 0$ the sum of odd binomial coefficients $\sum_{1 \leq i \leq n \,:\, i \text{ odd}} \binom{n}{i}$ equals $2^{n-1}$, so that we obtain the proposed formula. □

For two positive integers we denote by $K_{a,b}$ the complete bipartite graph with vertex classes of size $a$ and $b$, respectively, i.e., for $A = \{1, \dots, a\}$ and $B = \{a+1, \dots, a+b\}$ we define the graph via $V = A \cup B$ and $E = \{\{\alpha, \beta\} : \alpha \in A, \beta \in B\}$.

**Proposition 3.2.** *For positive integers $a, b$ we have*

$$M(K_{a,b}) = a + b + \binom{a}{2} + \binom{b}{2}.$$

*Proof.* For some subsets $A' \subseteq \{1, \dots, a\}$ and $B' \subseteq \{a+1, \dots, a+b\}$ with $S := A' \cup B' \neq \emptyset$ we can easily describe $c_I^S$. The value of $c_I^S$ at a position $\alpha \in A$ equals $\#B'$ rem 2, i.e., the remainder of $\#B'$ divided by 2. Similarly, the value of $c_I^S$ at a position $\beta \in B$ equals $\#A'$ rem 2.

Every non-zero codeword can be written as $c^S$ for some subset $\emptyset \neq S \subseteq \{1, \dots, a+b\}$. We decompose $S = A' \cup B'$, where $A' \subseteq A$ and $B' \subseteq B$. If $\#S \geq 3$ and $\#A' \geq 2$, then let $\tilde{A} \subseteq A'$ with $\#\tilde{A} = \#A' - 2$. With this, we have $\operatorname{supp}\left(c_I^{\tilde{A} \cup B'}\right) \subseteq \operatorname{supp}(c_I^S)$, i.e., $c^S$ is not a minimal codeword. If $\#S \geq 3$ and $\#A' \leq 1$, then $\#B' \geq 2$ and we can choose $\tilde{B} \subseteq B'$ with $\#\tilde{B} = \#B' - 2$. Since $\operatorname{supp}\left(c_I^{A' \cup \tilde{B}}\right) \subseteq \operatorname{supp}(c_I^S)$ we again conclude that $c^S$ is not a minimal codeword. If $\#S = 1$, then $c^S$ is a minimal codeword. If $\#S = 2$, then we can apply Lemma 2.7 and conclude that $c^S$ is minimal iff either $S \subseteq A$ or $S \subseteq B$. □

Note that $K_2 = K_{1,1}$.

**Corollary 3.3.** *For the star graph $K_{1,n-1}$ we have $M(K_{1,n-1}) = n + \binom{n-1}{2} = \frac{n^2-n+2}{2}$ for all $n \geq 2$.*

Next we want to consider graph arising if we join the centers of two stars by an edge:

**Proposition 3.4.** *Let $\mathcal{G}$ be a graph with two vertices $u$ and $v$ that are joined by an edge. The other $a + b$ vertices, where $a, b \geq 1$, are vertices of degree 1, where $a$ of them have $u$ as their unique neighbor and the other $b$ of them have $v$ as their unique neighbor. With this, we have $M(\mathcal{G}) = 2 + a + b + \binom{a+1}{2} + \binom{b+1}{2}$.*

*Proof.* By $n = 2 + a + b$ we denote the number of vertices of the graph and by $C$ the code $C(\mathcal{G})$ associated to $\mathcal{G}$. For each subset $S \subseteq \{1, \ldots, n\}$ of cardinality 1 the codeword $c^S$ is minimal. For subsets $S$ of cardinality 2 we apply Lemma 2.7. Counting the number of pairs of vertices with a common neighbor gives $\binom{a+1}{2} + \binom{b+1}{2}$ choices. It remains to show that $c^S$ is non-minimal if $\#S \geq 3$. First we note $c_I^T = c_I^S$ if $S$ arises from $T$ by adding two of the $a$ neighbors of $u$ of degree 1. So, $c^S$ is non-minimal in that case. By symmetry, the same is true for the $b$ neighbors of $v$ of degree 1. So, let $x$ be an arbitrary neighbor of $u$ of degree 1 and $y$ be an arbitrary neighbor of $v$ of degree 1. It suffices to consider $S \subseteq \{x, u, v, y\}$. In the following table we consider all choices for $S$ and abbreviate $c_I^S$ by just four binary entries. The second and third entry correspond to vertex $u$ and vertex $v$, respectively. The first entry corresponds to vertex $x$ or any other neighbor of $u$ of degree 1, noting that those entries are all equal. Similarly, the fourth entry corresponds to vertex $y$ or any other neighbor of $v$ of degree 1.

| $S$ | $c_I^S$ | $S$ | $c_I^S$ |
|---|---|---|---|
| $\{u\}$ | $(1,0,1,0)$ | $\{v\}$ | $(0,1,0,1)$ |
| $\{x\}$ | $(0,1,0,0)$ | $\{y\}$ | $(0,0,1,0)$ |
| $\{u,v\}$ | $(1,1,1,1)$ | $\{x,y\}$ | $(0,1,1,0)$ |
| $\{x,u\}$ | $(1,1,1,0)$ | $\{v,y\}$ | $(0,1,1,1)$ |
| $\{x,v\}$ | $(0,0,0,1)$ | $\{u,y\}$ | $(1,0,0,0)$ |
| $\{x,u,v\}$ | $(1,0,1,1)$ | $\{u,v,y\}$ | $(1,1,0,1)$ |
| $\{x,u,y\}$ | $(1,1,0,0)$ | $\{x,v,y\}$ | $(0,0,1,1)$ |
| $\{x,u,v,y\}$ | $(1,0,0,1)$ | | |

The proof is finished by the easy but a bit tedious task to check that for all $S \subseteq \{x, u, v, y\}$ with $\#S \geq 3$ there exists a subset $\emptyset \neq T \subsetneq S$ with $\operatorname{supp}(c_I^T) \subseteq \operatorname{supp}(c_I^S)$, so that we can apply Lemma 2.4 to conclude that $c^S$ is non-minimal. □

For an integer $r \geq 1$ and positive integers $a_1, \ldots, a_r$ we denote by $K_{a_1,\ldots,a_r}$ the complete multipartite graph, i.e., the vertex set of the $n = \sum_{i=1}^r a_i$ vertices is partitioned into $r$ classes such that two vertices are connected by an edge iff the come from different classes.

**Proposition 3.5.** *For each complete multipartite graph $\mathcal{G} = K_{a_1,\ldots,a_r}$ with $r \geq 3$ we have*

$$M(\mathcal{G}) = n + \binom{n}{2} + \sum_{U \subseteq \{1,\ldots,r\} \,:\, \#U \equiv 1 \pmod 2, \#U \geq 3} \prod_{i \in U} a_i,$$

*where $n = \sum_{i=1}^r a_i$.*

*Proof.* Let us denote the $r$ vertex classes by $A_1, \ldots, A_r$. Given a non-empty subset $S$ of the vertex set, we set $A_i' = A_i \cap S$ for all $1 \leq i \leq r$. Let $v \in A_j$ for some $1 \leq j \leq r$. Then, the value of $c_I^S$ at position $v$ is given by $\#(S \backslash A_j)$ rem 2. If $\#S = 1$, then $c^S$ is minimal. If $\#S = 2$, then we can use Lemma 2.7 to deduce that $c^S$ is minimal. If $\#S \geq 3$ and $\#(S \cap A_j) \geq 2$ for some index $1 \leq j \leq r$, then we can choose a two-element subset $U \subseteq S \cap A_j$ and use $c_I^{S \backslash U} \subseteq c_I^S$ to conclude that $c^S$ is non-minimal. It remains to consider the cases where $\#S \geq 3$ and $\#(S \cap A_i) \leq 1$ for all $1 \leq i \leq r$. Similar as in the proof of Proposition 3.1, we easily conclude that $c^S$ is minimal iff $\#S \equiv 1 \pmod 2$. □

We remark that $K_n = K_{1,\ldots,1}$, where the complete multipartite graph has exactly $n$ vertex classes with cardinality 1 each.

For each integer $n \geq 2$ we denote by $P_n$ the graph whose edges are given by $E = \{\{i, i+1\} : 1 \leq i \leq n-1\}$. The graph $P_n$ is also called a path of order $n$.

**Proposition 3.6.** *For each integer $n \geq 1$ we have*

$$M(P_n) = \left\lfloor \frac{(n+1)^2}{4} \right\rfloor = \left\lfloor \frac{n+1}{2} \right\rfloor \cdot \left\lceil \frac{n+1}{2} \right\rceil.$$

*Proof.* Each non-zero codeword of $C(P_n)$ is given as $c^S$ for some subset $\emptyset \neq S \subseteq \{1, \ldots, n\}$. For $\#S = 1$ the codeword $c^S$ is minimal. Given $S$, a maximal distance-2 chain $U$ is a subset of $S$ of the form $U = \{a, a+2, \ldots, b-2, b\}$, where $a \equiv b \pmod 2$ and $a-2, b+2 \notin S$. If $U = \{a, a+2, \ldots, b-2, b\} \neq \emptyset$ is a (maximal) distance-2 chain, then $\operatorname{supp}(c_I^U) = \{n+a-1, n+b+1\} \cap \{n+1, \ldots, 2n\}$. We have $\operatorname{supp}(c_I^U) = \emptyset$ iff $a = 1$ and $b = n$. For a suitable integer $r \geq 1$ let $U_1, \ldots, U_r$ be the unique decomposition of $S$ into maximal distance-2 chains. We directly conclude that the supports of $c^{U_i}$ and $c^{U_j}$ are disjoint for all $1 \leq i < j \leq r$. Thus, $c^S$ cannot be minimal if $r \geq 2$ since $\operatorname{supp}(c^{U_i}) \subseteq \operatorname{supp}(c^S)$ and $\operatorname{supp}(c^{U_i}) \neq \emptyset$ for $1 \leq i \leq r$. Now suppose that $S$ itself is a maximal distance-2 chain, i.e., there exist integers $a$ and $b$ with $S = \{a, a+2, \ldots, b-2, b\}$. Each proper subset $\emptyset \neq T \subsetneq S$ has a decomposition into $r \geq 2$ maximal distance-2 chains $U_1, \ldots, U_r$. Note that $\#\operatorname{supp}\left(c_I^{U_i}\right) \geq 1$ for all $1 \leq i \leq r$. So, if $\operatorname{supp}(c_I^T) \subseteq \operatorname{supp}(c_I^S)$, then we have $r = 2$, $\operatorname{supp}(c_I^T) = \operatorname{supp}(c_I^S) = \{a-1, b+1\}$, and $\{1, n\} \subseteq U_1 \cup U_2$. From the formula for $\operatorname{supp}(c_I^{U_i})$ we conclude $T = U_1 \cup U_2 = \{1, 3, \ldots, a-2, b+2, b+4, \ldots, n\} \not\subseteq S$ – contradiction. Thus, $c^S$ is minimal.

Counting the maximal distance-2 chains gives

$$\begin{aligned} M(P_n) &= \#\{(a,b) : 1 \leq a \leq b \leq n, a \equiv b \pmod 2\} \\ &= \sum_{i=1}^{n} \left\lceil \frac{n+1-i}{2} \right\rceil = \left\lfloor \frac{(n+1)^2}{4} \right\rfloor = \left\lfloor \frac{n+1}{2} \right\rfloor \cdot \left\lceil \frac{n+1}{2} \right\rceil. \end{aligned}$$

□

Note that the formula for $M(P_n)$ is also valid for the case $P_2 = K_2 = K_{1,1}$.

For each integer $n \geq 3$ we denote by $C_n$ the graph whose edges are given by $E = \big\{\{i, i+1\} : 1 \leq i \leq n-1\big\} \cup \big\{\{1, n\}\big\}$. The graph $C_n$ is also called a cycle of order $n$. With $\tau\colon \mathbb{Z} \to \{1, \ldots, n\}$ being the unique mapping with $\tau(z) \equiv z \pmod n$ for all $z \in \mathbb{Z}$, we can also write the edge set of $C_n$ as $\big\{\{\tau(i), \tau(i+1)\} : 1 \leq i \leq n\big\}$. The proof of Proposition 3.6 can be adjusted slightly to determine $M(C_n)$.

**Proposition 3.7.** *For each integer $n \geq 3$ we have*

$$M(C_n) = \begin{cases} \frac{n^2-2n+4}{2} & : \ n \equiv 0 \pmod 2, \\ n^2 - n + 2 & : \ \textit{otherwise.} \end{cases}$$

*Proof.* Each non-zero codeword of $C(C_n)$ is given as $c^S$ for some subset $\emptyset \neq S \subseteq \{1, \ldots, n\}$. For $\#S = 1$ the codeword $c^S$ is minimal. Given $S$, a distance-2 chain $U$ is a subset of $S$ of the form $U = \{\tau(a), \tau(a+2), \ldots, \tau(b-2), \tau(b)\}$, where $a \equiv b \pmod 2$ and $1 \leq a \leq b \leq 2n$. We call $U$ maximal (in $S$) if either $\tau(a-2), \tau(b+2) \notin S$ or $\tau(a-2) = \tau(b)$. The meaning is that neither $\{\tau(a), \tau(a+2), \ldots, \tau(b), \tau(b+2)\}$ nor $\{\tau(a-2), \tau(a), \ldots, \tau(b-2), \tau(b)\}$ is a proper superset of $U$ that is a subset of $S$, i.e., we cannot enlarge $U$ to a strictly larger distance-2 chain. Given a distance-2 chain $U = \{\tau(a), \tau(a+2), \ldots, \tau(b-2), \tau(b)\}$ we have $\operatorname{supp}(c_I^U) = \emptyset$ if $\tau(a-2) = b$ and $\operatorname{supp}(c_I^U) = \{n+\tau(a-1), n+\tau(b+1)\}$ otherwise. For a suitable integer $r \geq 1$ let $U_1, \ldots, U_r$ be the unique decomposition of $S$ into maximal distance-2 chains. We directly conclude that the supports of $c^{U_i}$ and $c^{U_j}$ are disjoint for all $1 \leq i < j \leq r$. Thus, $c^S$ cannot be minimal if $r \geq 2$. Next we show that $c^S$ is minimal iff $S$ is a (maximal) distance-2 chain itself. Each subset $\emptyset \neq T \subsetneq S$ has a decomposition into $r \geq 2$ maximal distance-2 chains $U_1, \ldots, U_r$, where $\#\operatorname{supp}(c_I^{U_i}) = 2$ for all $1 \leq i \leq r$, which contradicts $\cup_{i=1}^r \operatorname{supp}(c_I^{U_i}) \subseteq \operatorname{supp}(c_I^S)$. Thus, it remains to count the number of different maximal distance-2 chains $U = \{\tau(a), \tau(a+2), \ldots, \tau(b-2), \tau(b)\}$.

If $n$ is even, then the case $\tau(a-2) = \tau(b)$ can occur exactly two times, i.e., for the cases $\{1, 3, \ldots, n-1\}$ and $\{2, 4, \ldots, n\}$. Otherwise, we can start at any vertex $1 \leq a \leq n$ and choose $b = a + 2j$, where $0 \leq j \leq (n-4)/2$ since $j = (n-2)/2$ would yield $\tau(a-2) = \tau(b)$. Thus, if $n$ is even, we have $M(C_n) = 2 + n \cdot \frac{n-2}{2} + 2 = \frac{n^2-2n+4}{2}$. If $n$ is odd, then the case $\tau(a-2) = \tau(b)$ occurs iff $U = \{1, \ldots, n\}$. Otherwise, we can start at any vertex $1 \leq a \leq n$ and choose $b = a + 2j$, where $0 \leq j \leq n-2$. Thus, if $n$ is odd, we have $M(C_n) = 1 + n \cdot (n-1) = n^2 - n + 2$ □

Note that the formula for $M(C_n)$ is also valid for the case $C_3 = K_3$.

In a bipartite graph $\mathcal{G}$ we may generalizes the idea of a distance-2 chain as follows. We can build up a new graph $\mathcal{G}'$ with the same vertex set as $\mathcal{G}$. Two vertices in $\mathcal{G}'$ are connected by an edge, by definition, if they are at distance exactly 2 in $\mathcal{G}$. Similar as in the proof of Proposition 3.6 one can show that for each minimal codeword $c^S$ the set $S$ induces a connected subgraph in $\mathcal{G}'$. However, $c^S$ can be non-minimal for the vertex set $S$ of a connected subgraph of $\mathcal{G}'$, i.e., we may only conclude an upper bound on $M(\mathcal{G})$. This e.g. happens in $K_{a,b}$ provided that $a$ and $b$ are large enough.

Another variant of a distance-2 chain can lead to lower bounds.

**Definition 3.1.** *In a graph $\mathcal{G}$ an even path between two vertices $u$ and $v$ is a sequence of distinct vertices $[v_0, \ldots, v_l]$ such that $v_0 = u$, $v_l = v$, $l$ is an even positive integer, and $\{v_i, v_{i+1}\}$ is an edge for all $0 \leq i \leq l-1$. We call $[v_0, \ldots, v_l]$ a shortest even path between $u$ and $b$ if $l$ is minimal.*

Note that even in a connected graph there does not need to exist a shortest even path for two given vertices. Moreover, in the case of existence it does not need to be unique.

**Lemma 3.8.** *Let $\mathcal{G}$ be a graph and $C$ be the associated binary linear code. For each shortest even path $[v_0, \ldots, v_l]$ the codeword $c^S$, where $S = \{0 \leq i \leq l \ : \ i \equiv 0 \pmod 2\}$, is minimal.*

*Proof.* As an abbreviation we set $E = \{0 \leq i \leq l \,:\, i \equiv 0 \pmod 2\}$ and $O = \{0 \leq i \leq l \,:\, i \equiv 1 \pmod 2\}$. Let $o \in O$ and $e \in E$ such that $\{v_o, v_e\}$ is an edge in $\mathcal{G}$. Since $l$ is minimal by assumption we have $e \in \{o-1, o+1\}$. Thus, for each $o \in O$ the vertex $v_o$ has exactly two neighbors $v_e$ where $e \in E$. So, in $c_I^S$ the entries at the positions $v_o$ for $o \in O$ are equal to zero. Now let $\emptyset \neq T \subsetneq S = E$ be an arbitrary subset, $t_{min}$ be the minimal element in $T$, and $t_{max}$ be the maximal element in $T$. If $t_{min} > 0$, then the entry at position $v_{t_{min}-1}$ in $c_I^T$ is equal to one so that $c^T \not\subseteq c^S$. Similarly, if $t_{max} < l$, then the entry at position $v_{t_{max}+1}$ in $c_I^T$ is equal to one so that again $c^T \not\subseteq c^S$. Since $t_{min} = 0$ and $t_{max} = l$ imply $T = S$ the codeword $c^S$ is indeed minimal. □

Note that Lemma 2.7 says that for a two element subset $S \subseteq \{1, \ldots, n\}$ the codeword $c^S$ is exactly minimal if there exists a (shortest) even path of length 2 between the elements of $S$.

**Lemma 3.9.** *Let $[v_0, \ldots, v_l]$ and $\left[v'_0, \ldots, v'_{l'}\right]$ be two shortest even paths. If*

$$\{v_i \,:\, 0 \leq i \leq l, i \equiv 0 \pmod 2\} = \left\{v'_i \,:\, 0 \leq i \leq l', i \equiv 0 \pmod 2\right\},$$

*then $\{v_0, v_l\} = \left\{v'_0, v'_{l'}\right\}$.*

*Proof.* First we note that $l = l'$. Now choose indices $0 \leq i, j \leq l$ such that $v'_0 = v_i$ and $v'_l = v_j$ with $i \equiv j \equiv 0 \pmod 2$. Depending on whether $i < j$ or $i > j$, either $[v_i, v_{i+1}, \ldots, v_j]$ or $[v_i, v_{i-1}, \ldots, v_j]$ is an even path from $v'_0$ to $v'_l$. Since we assume that $l$ is the shortest possible length of such a path, we have $\{i, j\} = \{0, l\}$, i.e., $\{v_0, v_l\} = \left\{v'_0, v'_{l'}\right\}$. □

We remark that in an odd cycle, i.e., $C_n$ where $n$ is odd, the shortest even path between any two neighbored vertices uses all $n$ vertices.

**Lemma 3.10.** *Let $\mathcal{G}$ be a connected graph and $\mathcal{T}$ be a spanning tree of $\mathcal{G}$. Considering $\mathcal{T}$ as a bipartite graph, we denote the number of vertices in the two color classes by $a$ and $b$. With this, we have $M(\mathcal{G}) \geq a + b + \binom{a}{2} + \binom{b}{2}$.*

*Proof.* Let $C$ be the code $C(\mathcal{G})$ induced by the graph $\mathcal{G}$. For each of the $n$ subsets $S \subseteq \{1, \ldots, n\}$ of cardinality 1 the codeword $c^S$ is minimal. These are $a + b$ minimal codewords. Now consider two vertices $u$ and $v$ of the same color class in $\mathcal{T}$. Due to the construction of the coloring, there exists an even path between $u$ and $v$ in $\mathcal{T}$, which is also an even path between $u$ and $v$ in $\mathcal{G}$. If the path is not already a shortest even path, then pick one. So, for every two vertices $u$ and $b$ of the same color class (in $\mathcal{T}$) there exists a shortest even path $[v_0, \ldots, v_l]$ between $u$ and $v$ in $\mathcal{G}$, so that Lemma 3.8 implies that $c^S$ is minimal, where $S = \{v_i \,:\, 0 \leq i \leq l, i \equiv 0 \pmod 2\}$. There are $\binom{a}{2} + \binom{b}{2}$ choices and by Lemma 3.9 all of them lead to different minimal codewords $c^S$, where $\#S \geq 2$. □

We remark that the lower bound of Lemma 3.10 is attained with equality in Proposition 3.2, i.e., for complete bipartite graphs, and in Proposition 3.4. In Theorem 4.1 we will use Lemma 3.10 to determine a formula for the minimum number $m(n)$ of minimal codewords of a connected graph with $n$ vertices.

An induced subgraph of a graph $\mathcal{G} = (V, E)$ is a graph whose vertex set is a subset $S \subseteq V$ and whose edges are given by the elements of $E$ where both vertices are contained in $S$. If $\mathcal{G}' = (V', E')$ is an induced subgraph of $\mathcal{G} = (V, E)$ and $c^S$ a minimal codeword in $\mathcal{G}'$, where $S \subseteq V'$, then $c^S$ is also a minimal codeword in $\mathcal{G}$. An odd cycle is an induced subgraph that is isomorphic to $C_l$, where $l$ is odd.

**Proposition 3.11.** *Let $\mathcal{G} = (V, E)$ be a connected graph. If $S \subseteq V$ induces an odd cycle in $\mathcal{C}$, then $c^S$ is a minimal codeword in $C(\mathcal{G})$.*

*Proof.* Let $\mathcal{G}' = (S, E')$ be the subgraph induced by $S$ and $C = C(\mathcal{G}')$ the binary code associated with $\mathcal{G}'$. In $C$ we have $c_I^S = \mathbf{0}$. We can easily check that $c_I^T \neq \mathbf{0}$ for all $\emptyset \neq T \subsetneq S$, so that Lemma 2.6 gives that $c^S$ is minimal in $C$. As noted above, $c^S$ is also minimal in $C(\mathcal{G})$. □

Another lower bound, using more common graph invariants, for $M(\mathcal{G})$ is:

**Lemma 3.12.** *Let $\mathcal{G}$ be a graph with $n$ vertices, maximum degree $\Delta$, and $t$ triangles. Then, we have $M(\mathcal{G}) \geq n + \binom{\Delta}{2} + t$.*

*Proof.* For each of the $n$ subsets $S \subseteq \{1, \ldots, n\}$ of cardinality 1 the codeword $c^S$ is minimal. If $v$ is a vertex of degree $\Delta$, then for any pair $S$ of two neighbors of $v$ we can apply Lemma 2.7 to deduce that $c^S$ is minimal. Note that there are $\binom{\Delta}{2}$ choices. If $S$ consists of the three vertices of a triangle, then $c^S$ is minimal, see Proposition 3.11. □

Next we want to study the special situation where all non-zero codewords are minimal.

**Proposition 3.13.** *If $\mathcal{G}$ is a graph with $n \geq 1$ nodes and $M(\mathcal{G}) = 2^n - 1$, then $\mathcal{G} = K_1$ or $\mathcal{G} = K_3$.*

*Proof.* Due to [16, Theorem 2(iii)] an $[N, k]_2$ code whose non-zero codewords are all minimal satisfies $N \geq 3(k-1)$. In our situation we have $N = 2n$ and $k = n$, so that $n \leq 3$. If $\mathcal{G}$ contains an isolated vertex, then $M(\mathcal{G}) \leq 1 + M(n-1) \leq 1 + 2^{n-1} - 1$, which is strictly less than $2^n - 1$ for $n \geq 2$. Thus, it suffices to consider the connected graphs with up to 3 vertices: $M(P_1) = 1$, $M(P_2) = 2$, $M(P_3) = 4$, and $M(K_3) = 7$, see Proposition 3.6 and Proposition 3.1. □

# 4 Exact values for small parameters

The aim of this subsection is to determine the exact value of $M(n)$ and $m(n)$ for $1 \leq n \leq 10$. Given Lemma 2.1 it suffices to consider connected graphs. We note that there are already 11 716 571 non-isomorphic connected graphs, which we have enumerated using the software package `geng` [14]. For each connected graph $\mathcal{G}$ we determine $M(\mathcal{G})$ by exhaustive enumeration.

| $n$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| $m(n)$ | 1 | 2 | 4 | 6 | 9 | 12 | 16 | 20 | 25 | 30 |
| $M(n)$ | 1 | 2 | 7 | 14 | 26 | 47 | 99 | 190 | 355 | 682 |

The maximum $M(n)$ is attained for $3 \leq n \leq 6$ by a complete graph $K_n$, while the cases $n \in \{7, 8, 9, 10\}$ need other constructions. For $n = 10$ there are 22 isomorphism types of graphs that attain the maximum of 682 minimal codewords. Those graphs are quite diverse, i.e., their number of edges lies between 21 and 32, the minimum degree is either 4 or 5, and the maximum degree ranges from 5 to 9. For $1 \leq n \leq 10$, the minimum value $m(n)$ is attained by a path $P_n$. This observation is also true in general.

**Theorem 4.1.** *For each integer $n \geq 1$ we have*

$$m(n) = \left\lfloor \frac{(n+1)^2}{4} \right\rfloor = \left\lfloor \frac{n+1}{2} \right\rfloor \cdot \left\lceil \frac{n+1}{2} \right\rceil .$$

*Proof.* Due to Proposition 3.6 it remains to show the corresponding lower bound. To this end we apply Lemma 3.10. Setting $b = n - a$ we obtain the lower bound $n + \binom{a}{2} + \binom{n-a}{2} = a^2 - na + \frac{n(n+1)}{2}$, which is a quadratic polynomial in $a$. Over the real numbers the minimum is attained for $a = n/2$, which gives $m(n) \geq \frac{n(n+2)}{4}$. If $n$ is even this matches the statement. If $n$ is odd we can uproud $\frac{n(n+2)}{4}$ to $\frac{(n+1)^2}{4}$ since $m(n)$ is an integer. □

We remark that all connected graphs $\mathcal{G}$ with $n$ vertices and $M(\mathcal{G}) = m(n)$ are bipartite, since Proposition 3.11 would give an additional minimal codeword that is not counted in Lemma 3.10. If $\mathcal{T}$ is a tree such that all vertices with degree strictly larger than 1 are contained on a path, then it can be easily shown that the lower bound of Lemma 3.10 is attained with equality. If the cardinalities of the two color classes of the bipartite tree $\mathcal{T}$ differ by at most 1, then we have $M(\mathcal{T}) = m(n)$, where $\mathcal{T}$ consists of $n$ vertices. We remark that one can also construct connected graphs $\mathcal{G}$ that contain 4-cycles and satisfy $M(\mathcal{G}) = m(n)$ for their number $n$ of vertices. The determination of $M(n)$ and the description of an infinite family of graphs attaining $M(\mathcal{G}) = M(n)$ is an interesting open problem.

## Acknowledgments

The author wishes to thank Romar dela Cruz for introducing him to the problem of minimal codewords and suggesting the construction of a linear code from the adjacency matrix of a graph.

## References

[1] E. Agrell. Voronoi regions for binary linear block codes. *IEEE Transactions on Information Theory*, 42(1):310–316, 1996.

[2] E. Agrell. On the Voronoi neighbor ratio for binary linear block codes. *IEEE Transactions on Information Theory*, 44(7):3064–3072, 1998.

[3] A. Alahmadi, R. Aldred, R. dela Cruz, S. Ok, P. Solé, and C. Thomassen. The minimum number of minimal codewords in an $[n, k]$-code and in graphic codes. *Discrete Applied Mathematics*, 184:32–39, 2015.

[4] A. Alahmadi, R. E. Aldred, R. de la Cruz, P. Solé, and C. Thomassen. The maximum number of minimal codewords in an $[n, k]$-code. *Discrete Mathematics*, 313(15):1569–1574, 2013.

[5] A. Alahmadi, R. E. Aldred, R. dela Cruz, P. Solé, and C. Thomassen. The maximum number of minimal codewords in long codes. *Discrete Applied Mathematics*, 161(3):424–429, 2013.

[6] A. Ashikhmin and A. Barg. Minimal vectors in linear codes. *IEEE Transactions on Information Theory*, 44(5):2010–2017, 1998.

[7] E. Berlekamp, R. McEliece, and H. Van Tilborg. On the inherent intractability of certain coding problems. *IEEE Transactions on Information Theory*, 24(3):384–386, 1978.

[8] Y. Borissov and N. Manev. Minimal codewords in linear codes. *Serdica Mathematical Journal*, 30(2-3):303–324, 2004.

[9] C. Carlet, C. Ding, and J. Yuan. Linear codes from perfect nonlinear mappings and their secret sharing schemes. *IEEE Transactions on Information Theory*, 51(6):2089–2102, 2005.

[10] H. Chabanne, G. Cohen, and A. Patey. Towards secure two-party computation from the wire-tap channel. In *International Conference on Information Security and Cryptology*, pages 34–46. Springer, 2013.

[11] R. de la Cruz, M. Kiermaier, S. Kurz, and A. Wassermann. On the minimum number of minimal codewords. *arXiv preprint 1912.09804*, 2019.

[12] T.-Y. Hwang. Decoding linear block codes for minimizing word error rate. *IEEE Transactions on Information Theory*, 25(6):733–737, 1979.

[13] J. L. Massey. Minimal codewords and secret sharing. In *Proceedings of the 6th joint Swedish-Russian international workshop on information theory*, pages 276–279. Citeseer, 1993.

[14] B. McKay and A. Piperno. *Practical graph isomorphism II*, volume 60. 2013.

[15] J. Schillewaert, L. Storme, and J. A. Thas. Minimal codewords in reed–muller codes. *Designs, Codes and Cryptography*, 54(3):273–286, 2010.

[16] N. Sloane. Covering arrays and intersecting codes. *Journal of Combinatorial Designs*, 1(1):51–63, 1993.

[17] D. B. West. *Introduction to graph theory*. Prentice Hall Upper Saddle River, 2001.