TUCS

Ville Junnila

On Identifying and Locating-Dominating Codes

Turku Centre for Computer Science

TUCS Dissertations
No 137, June 2011

On Identifying and Locating-Dominating Codes

Ville Junnila

To be presented, with the permission of the Faculty of Mathematics and Natural Sciences of the University of Turku, for public criticism in Auditorium XXI on June 22, 2011, at 12 noon.

University of Turku
Department of Mathematics
FI-20014 Turku, Finland

2011

Supervisor

DOCENT TERO LAIHONEN
Department of Mathematics
University of Turku
Finland

Reviewers

PROFESSOR RYAN MARTIN
Department of Mathematics
Iowa State University
USA

DOCTOR JULIEN MONCEL
Department QLIO
Institut Universitaire de Technologie de Rodez
France

Opponent

DOCTOR RALF KLASING
Centre National de la Recherche Scientifique (CNRS)
Laboratoire Bordelais de Recherche en Informatique (LaBRI)
Université Bordeaux 1
France

ISBN 978-952-12-2598-7
ISSN 1239-1883

Acknowledgements

First of all, I want to express my deepest gratitude to my supervisor Docent Tero Laihonen for his constant guidance and support. His natural instinct for fruitful research topics as well as his broad knowledge in discrete mathematics have been crucial in preparing this thesis. In addition, his personality and sense of humor made the collaboration a pleasurable experience. I would also like to thank Professor Iiro Honkala as the leader and promoter of our research group for providing an excellent research environment.

I wish to thank my co-author Doctor Sanna Ranto for productive scientific collaboration and her willingness to help with practical matters. My co-author Professor Geoffrey Exoo should also be acknowledged. Without his indisputable talent with computers, some of the results of this thesis would have been completely unreachable. I thank Professor Ryan Martin and Doctor Julien Moncel for agreeing to review the thesis. Their insightful comments and suggestions have improved the thesis significantly.

I am grateful for the steady funding provided by the Turku Centre for Computer Science (TUCS), the Academy of Finland and the Department of Mathematics. The working environment at the Department of Mathematics has been excellent, not least because of the warm and friendly people working here. In particular, I would like to thank my office mates Jarmo Hemminki, Tommi Meskanen and Mikko Pelto for numerous discussions concerning mathematics and related topics. Special thanks are reserved for Eeva Suvitie for relentlessly being herself.

Finally, I would like to thank my parents for their love and support.

Turku, May 2011 Ville Junnila

Contents

Chapter 1

Introduction

In this thesis, we study identifying and locating-dominating codes. We begin the introduction by considering the background and history of these codes in Section 1.1. The mathematical definitions needed in the thesis are presented in Section 1.2. Finally, in Section 1.3, we describe the structure of the thesis.

1.1 Background

The concept of identifying codes was introduced in 1998 by Karpovsky, Chakrabarty and Levitin [60]. The original motivation for studying these codes comes from fault diagnosis in multiprocessor systems. In multiprocessor systems, where processors are linked to the other ones, identifying codes can be defined as follows. Assign to a set of processors the following task: each processor outputs a single value *true* if the processor itself or one of the neighbouring ones are malfunctioning, and otherwise *false* is outputted. We say that this set of chosen processors forms an identifying code if the malfunctioning processor(s) can be located using solely the information provided by the selected processors. In addition to fault diagnosis in multiprocessor systems, identifying codes find their motivation also in various other applications such as environmental monitoring [5], routing in wireless networks [61], sensor networks [69], and fire and intruder alarm systems [71].

Besides the motivating applications above, identifying codes also form a mathematically interesting and rich field for basic research. The previous definition of identifying codes in the case of multiprocessor systems can be generalized in a natural way to graphs by considering processors and links between them as vertices and edges of a graph, respectively. Since the seminal paper by Karpovsky *et al.* [60], the study in the field has been intensive. In particular, identifying codes have been studied in various graphs such as

- binary Hamming spaces [8, 9, 34, 35, 38, 51, 62, 67],
- infinite grids such as the square grid, triangular grid, hexagonal grid, and king grid [4, 14–16, 21, 23–25, 44, 45, 52],
- trees [1, 7, 10, 13], cycles [6, 36, 77], and paths [6, 70].

Identification has also been considered in general graphs. For example, various extremal cardinalities for identifying codes have been studied in [19].

Locating-dominating codes are closely related to the identifying codes. In the literature, they are also called locating-dominating sets. The only difference in the definition is that in the case of locating-dominating codes the reporting vertices (or processors) — instead of just outputting the value *true* or *false* — can also distinguish whether the processor itself or one of the neighbouring ones is faulty. The concept of locating-dominating codes was introduced by Slater [68, 71, 72]. As can be expected, the motivating applications for locating-dominating codes are similar to the ones of identifying codes.

The research on locating-dominating codes has been active. Location-domination has been studied, for example, in the following graphs: binary Hamming spaces [50], infinite grids [40, 46, 73], trees [10, 11], cycles [6, 20] and paths [6, 41]. Moreover, locating-dominating codes have also been considered in general graphs, for example, in [18, 19]. For an extensive listing of papers concerning identifying and locating-dominating codes, we refer to the internet bibliography [64] maintained by Antoine Lobstein.

An identifying or locating-dominating code with the smallest cardinality in a given finite graph is called *optimal*. In the case of infinite graphs a more sophisticated method is required to measure the sizes of codes. However, in this case, the concept of optimal codes is defined analogously. Determining optimal codes in graphs is one of the most natural questions in the field. However, it is usually difficult to determine these optimal codes. In fact, in [17], it is shown that algorithmically finding an optimal identifying or locating-dominating code in a given graph is an NP-hard problem in general. However, there exist graphs in which this problem can be efficiently solved. For example, in [1] and [71], it has been shown that an optimal identifying and locating-dominating code, respectively, can be found in trees in linear time. Moreover, in [74], efficient algorithms for finding approximations of optimal codes have been presented.

In the literature, several variations of identifying and locating-dominating codes have also been studied. For example, strongly identifying codes (see [42, 49]), robust identifying codes (see [39, 43, 48, 63]) and fault-tolerant locating-dominating codes (see [73]) are among these variations. However, in the thesis, we concentrate on the original versions of the identifying and locating-dominating codes.

It should also be noted that identifying and locating-dominating codes are closely related to the subjects of covering codes (see [22]) and dominating sets (see [37]) in coding and graph theory, respectively. The definitions of covering codes and dominating sets are similar to the ones of identifying and locating-dominating codes, although now it is enough to detect that faulty vertices exist instead of actually locating them. Hence, identification and location-domination can be viewed as subclasses of covering codes and dominating sets, respectively.

1.2 Definitions

Let $G = (V, E)$ be a simple, connected and undirected graph with V as the set of vertices and E as the set of edges. Let u and v be vertices in V. If u and v are adjacent, then the edge joining u and v is denoted by uv. The *distance* $d(u, v)$ is the number of edges on any shortest path from u to v. For the rest of the thesis (unless otherwise stated), we assume that r is a non-negative integer. We say that u *r-covers* v if the distance $d(u, v)$ is at most r. The *ball of radius r centered at u* is defined as

$$B_r(u) = \{x \in V \mid d(u, x) \leq r\}.$$

Furthermore, if X is a subset of V, then we define

$$B_r(X) = \bigcup_{x \in X} B_r(x).$$

A nonempty subset of V is called a *code*, and its elements are called *codewords*. Let $C \subseteq V$ be a code. An *I-set* (or an *identifying set*) of the subset X of V with respect to the code C is defined as

$$I_r(C; X) = I_r(X) = B_r(X) \cap C.$$

If $X = \{x_1, x_2, \ldots, x_\ell\}$, then we denote $B_r(X) = B_r(x_1, x_2, \ldots, x_\ell)$ and $I_r(X) = I_r(x_1, x_2, \ldots, x_\ell)$. For a positive integer μ, we say that a code $C \subseteq V$ is a μ-fold r-covering code (or in short a μ-fold r-covering) in G if $|I_r(C; u)| \geq \mu$ for every vertex $u \in V$. If $\mu = 1$, then we say in short that C is an r-covering code (or an r-covering) in G.

The following definition of identifying codes is from [60].

Definition 1.2.1. Let ℓ be a positive integer. A code $C \subseteq V$ is said to be *$(r, \leq \ell)$-identifying* if for all $X, Y \subseteq V$ such that $|X| \leq \ell$, $|Y| \leq \ell$ and $X \neq Y$ we have

$$I_r(C; X) \neq I_r(C; Y).$$

Although we defined the concept of identifying codes also for sets of vertices ($\ell > 1$), in this thesis we focus on the case with single vertices ($\ell = 1$). For the various results concerning the case $\ell > 1$, we refer the interested reader to the papers listed in [64]. If $\ell = 1$, then we simply say that C is an *r-identifying code* in G. Furthermore, if $r = \ell = 1$, then C is said to be *identifying*.

If A and B are subsets of V, then the *symmetric difference* of A and B is defined as $A \triangle B = (A \setminus B) \cup (B \setminus A)$. We say that the vertices $u \in V$ and $v \in V$ are *r-separated* by a codeword of C (or in short by a code C) if the symmetric difference $I_r(C; u) \triangle I_r(C; v)$ is nonempty. Now, in the case $\ell = 1$, the definition of r-identifying codes can be reformulated as follows: a code $C \subseteq V$ is r-identifying if each vertex is r-covered by at least one codeword and each pair of vertices is r-separated by C.

Let $G = (V, E)$ be a finite graph. The smallest cardinality of an $(r, \leq \ell)$-identifying code in G is denoted by $M_{(r,\leq\ell)}(G)$. Furthermore, if ℓ is equal to one, then we write in short $M_{(r,\leq 1)}(G) = M_r(G)$. Notice that the value $M_{(r,\leq\ell)}(G)$ is not always defined since any $(r, \leq \ell)$-identifying codes in G do not necessarily exist. As an example of such a case, one can consider a complete graph. An $(r, \leq \ell)$-identifying code in G attaining the smallest cardinality is called *optimal*.

The following definition of locating-dominating codes was first introduced by Slater in [68, 71, 72] (for $r = 1$) and later generalized by Carson in [11] (for $r > 1$).

Definition 1.2.2. A code $C \subseteq V$ is said to be *r-locating-dominating* in G if for all distinct vertices $u, v \in V \setminus C$ the set $I_r(C; u)$ is nonempty and

$$I_r(C; u) \neq I_r(C; v).$$

If $r = 1$, then we simply say that C is a *locating-dominating code*.

The definition of locating-dominating codes can also be generalized for sets of vertices (as in the case of identifying codes). In fact, there exist two different, natural generalizations of locating-dominating codes for sets of vertices. However, in this thesis, only the case of single vertices is considered. For the more general definitions, the interested reader is referred to [50].

For a finite graph $G = (V, E)$, the smallest cardinality of an r-locating-dominating code is denoted by $M_r^{LD}(G)$. Notice that the value $M_r^{LD}(G)$ is always defined since there exists an r-locating-dominating code in any graph. Indeed, the whole vertex set V is always an r-locating-dominating code in G. An r-locating-dominating code attaining the smallest cardinality is called *optimal*.

Besides finite graphs, we can also study identification and location-domination in infinite graphs. Namely, in Chapters 5 and 6, we consider infinite grids with the vertex set $\mathbb{Z}^2$. Naturally, we also need a way to measure codes in these infinite grids. For this purpose, we first denote

$$Q_n = \{(x, y) \in \mathbb{Z}^2 \mid |x| \leq n, |y| \leq n\},$$

where n is a positive integer. The *density* of a code $C \subseteq \mathbb{Z}^2$ is then defined as

$$D(C) = \limsup_{n \to \infty} \frac{|C \cap Q_n|}{|Q_n|}.$$

The notion of density can now be used to measure sizes of codes in infinite grids. Furthermore, we say that an r-identifying or r-locating-dominating code is *optimal* in infinite grid, if there do not exist, respectively, any r-identifying or r-locating-dominating codes with smaller density.

1.3 Structure of the thesis

In Chapter 2, which is based on the papers [29], [30], [31], [32] and [33], we consider r-identifying codes in binary Hamming spaces. In Section 2.2, we present new lower bounds, which are currently the best known ones, for r-identifying codes with $r \geq 2$. Then, in Sections 2.3, 2.4 and 2.5, we consider three conjectures, which have been stated in the papers [9] and [60]. In these sections, we show various results related to these conjectures. Finally, in Section 2.6, we give some new constructions for r-identifying codes (improving the known upper bounds).

In Chapter 3, which is based on the paper [59], r-identifying codes in cycles and paths are considered. Previously, r-identifying codes have been studied in [6], [36], [70] and [77]. In Sections 3.2 and 3.3, the optimal cardinalities of r-identifying codes, respectively, in cycles and paths are determined in all the remaining open cases.

In Chapter 4, which is based on the papers [27] and [28], we consider r-locating-dominating codes in cycles and paths. In the case of paths, it has been shown by Bertrand *et al.* in [6] that for a path $\mathcal{P}_n$ of length n we have $M_r^{LD}(\mathcal{P}_n) \geq (n+1)/3$. Furthermore, they conjectured that for any $r \geq 2$ there exists an infinite family of n for which the lower bound can be attained. In Section 4.1, we show that this conjecture holds. In fact, we prove a stronger result according to which $M_r^{LD}(\mathcal{P}_n) = \lceil (n+1)/3 \rceil$ for all $n \geq n_r$ when $r \geq 3$ and n_r is large enough ($n_r = \mathcal{O}(r^3)$). In Section 4.2, for cycles $\mathcal{C}_n$ of length n, we prove similar results stating that $n/3 \leq M_r^{LD}(\mathcal{C}_n) \leq n/3 + 1$ if $n \equiv 3 \pmod 6$ and $M_r^{LD}(\mathcal{C}_n) = \lceil n/3 \rceil$ otherwise. Moreover, it is shown for $r = 3$ and $r = 4$ that we have $M_r^{LD}(\mathcal{C}_n) = n/3 + 1$ if $n \equiv 3 \pmod 6$. Furthermore, in Conjecture 4.2.17, we conjecture that the previous result also holds for general r.

Previously, a 2-identifying code in the hexagonal grid with density 4/19 has been presented in [16]. Improving the previously known lower bounds, Martin and Stanton [66] proved that the density of any 2-identifying code in the hexagonal grid is at least 1/5. In Chapter 5, which is based on the papers [57] and [58], we improve this lower bound by showing that the 2-identifying code with density 4/19 is actually optimal.

In Chapter 6, which is based on the papers [55] and [56], we introduce a new way to consider identification in infinite grids with the vertex set $\mathbb{Z}^2$. Namely, we consider identifying codes in infinite grids where the neighbourhood of a vertex is determined by a Euclidean ball with a given radius. In addition, we give lower bounds for identifying codes in these grids as well as general code constructions. We also find optimal identifying codes for a couple of infinite grids with small radii. We end the chapter by determining optimal identifying codes in the king grid with slightly modified balls.

Previously, a sequential version of identification called adaptive identification has been introduced in [2] and [3]. In these papers, adaptive identification is considered in torii of finite square and king grids, and as a further research it is suggested to study adaptive identification in different graphs. In Chapter 7, which is based on the paper [54], we consider this in the case of binary Hamming spaces.

Chapter 2

Identification in binary Hamming spaces

In this chapter, which is based on the papers [29], [30], [31], [32] and [33], we consider r-identifying codes in binary Hamming spaces. These papers also contain results concerning $(r, \leq \ell)$-identifying codes with $\ell > 1$. However, as mentioned in the introduction, in this thesis we concentrate on the case with $\ell = 1$. We begin the chapter by presenting some preliminary definitions and results in Section 2.1. Then, in Section 2.2, a new lower bound for r-identifying codes with $r \geq 2$ is presented. In Sections 2.3, 2.4 and 2.5, we consider three conjectures which have gathered interest in the past. Finally, in Section 2.6, we present some code constructions for r-identifying codes with the smallest known cardinalities.

2.1 Preliminaries

For the rest of the chapter, let n be a positive integer. The binary *Hamming space* $\mathbb{F}^n$ is the n-fold Cartesian product of the binary field $\mathbb{F} = \{0, 1\}$. The elements of $\mathbb{F}^n$ are called *words*. A nonempty subset of $\mathbb{F}^n$ is called a code of length n. Let $\mathbf{x}$ and $\mathbf{y}$ be words belonging to $\mathbb{F}^n$. The *Hamming distance* $d(\mathbf{x}, \mathbf{y})$ between words $\mathbf{x}$ and $\mathbf{y}$ is the number of coordinate places in which they differ. The set of non-zero coordinates of the word $\mathbf{x}$ is called the *support* of $\mathbf{x}$ and is denoted by *supp*$(\mathbf{x})$. The *weight* of $\mathbf{x}$ is the cardinality of the support of $\mathbf{x}$ and is denoted by $w(\mathbf{x})$. The *Hamming ball* of radius r centred at $\mathbf{x}$ is denoted by $B_r(\mathbf{x})$ and consists of the words that are r-covered by $\mathbf{x}$. The set $S_r(\mathbf{x})$ consists of the words that are exactly at distance r from $\mathbf{x}$, i.e. $S_r(\mathbf{x}) = B_r(\mathbf{x}) \setminus B_{r-1}(\mathbf{x})$. The size of a Hamming ball of radius r in $\mathbb{F}^n$ does not depend on the choice of the centre and it is denoted by $V(n, r)$.

Furthermore, we have

$$V(n, r) = \sum_{i=0}^{r} \binom{n}{i}.$$

Assuming $\mathbf{x} = (a_1, a_2, \ldots, a_n) = a_1 a_2 \cdots a_n$ and $\mathbf{y} = (b_1, b_2, \ldots, b_n) = b_1 b_2 \cdots b_n$, we can define the *concatenation* of the words as follows:

$$(\mathbf{x}, \mathbf{y}) = (a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_n) = a_1 a_2 \cdots a_n b_1 b_2 \cdots b_n.$$

The *sum* of the vectors $\mathbf{x}$ and $\mathbf{y}$ is defined as

$$\mathbf{x} + \mathbf{y} = (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n).$$

Let $C_1 \subseteq \mathbb{F}^n$ and $C_2 \subseteq \mathbb{F}^n$ be codes. Then their *direct sum*

$$C_1 \oplus C_2 = \{(\mathbf{x}, \mathbf{y}) \mid \mathbf{x}, \mathbf{y} \in \mathbb{F}^n\}$$

is a code in $\mathbb{F}^{n+m}$. We also denote $\mathbf{x} + C = \{\mathbf{x} + \mathbf{c} \mid \mathbf{c} \in C\}$. The function $\pi(\mathbf{u})$ is used for denoting the parity of $\mathbf{u}$ as follows:

$$\pi(\mathbf{u}) = \begin{cases} 0 & \text{if } w(\mathbf{u}) \text{ is even} \\ 1 & \text{if } w(\mathbf{u}) \text{ is odd.} \end{cases}$$

In this chapter, the size $M_r(\mathbb{F}^n)$ of an optimal r-identifying code in $\mathbb{F}^n$ is written in short as $M_r(n)$.

Let us then present some auxiliary results that will be needed in the chapter. The following estimation for $M_1(n)$ proves valuable in Sections 2.3 and 2.4.

Theorem 2.1.1. *For $n \geq 2$, we have*

$$M_1(n) \leq \frac{9}{2} \cdot \frac{2^n}{n+1}.$$

Proof. Let $n = 3 \cdot 2^s - 1$ with s being a non-negative integer. Then according to [34, Corollary 1] there exists a code C such that it is 1-identifying and a 2-fold 1-covering with cardinality

$$|C| = \frac{9}{4} \cdot \frac{2^n}{n+1}.$$

Therefore, the claim clearly holds for all $n = 3 \cdot 2^s - 1$ with $s \geq 0$.

Consider now the length $3 \cdot 2^s - 1 + k$ with k being an integer such that $0 < k < 3 \cdot 2^s$, i.e. $3 \cdot 2^s - 1 < 3 \cdot 2^s - 1 + k < 3 \cdot 2^{s+1} - 1$. Because C is a 2-fold 1-covering, we know by [9, Theorem 1] that the code $C \oplus \mathbb{F}^k$ is 1-identifying. Now the number of words in the 1-identifying code $C \oplus \mathbb{F}^k$ is

$$2^k |C| \leq \frac{9}{2} \cdot \frac{2^{n+k}}{(n+k)+1}$$

Thus, the claim holds for all $n \geq 2$. $\square$

The following upper bound for $M_r(n)$ has previously been shown in [34, Theorem 3].

Theorem 2.1.2 ([34])**.** *For $i = 1, 2, \ldots, r$, let n_i be a positive integer. Then we have*

$$M_r\left(\sum_{i=1}^{r} n_i\right) \leq \prod_{i=1}^{r} M_1(n_i).$$

2.2 Lower bounds for r-identifying codes

In what follows, we are going to improve the known lower bounds on r-identifying codes for $r \geq 2$. The main underlying idea in the earlier results presented in [9] and [35] was to find values as small as possible for $m = \max\{|I_r(\mathbf{x})| : \mathbf{x} \in \mathbb{F}^n\}$ using partial constructions. (Besides these results, there is also a bound by Karpovsky *et al.* [60]; see Theorem 2.2.5.) In this section, we approach the problem in a different manner. Namely, for $r \geq 2$, we improve the lower bound by concentrating on the function $P_r(n, i)$ defined below instead of the value m.

Let $\mathbf{x} \in \mathbb{F}^n$ and i be an integer such that $i \geq 3$. Define then

$$\begin{aligned} P_r(n, i, \mathbf{x}) = \max_{C \subseteq \mathbb{F}^n} |\{\, \mathbf{y} \in \mathbb{F}^n \mid\ & C \text{ is an } r\text{-identifying code satisfying} \\ & |I_r(C; \mathbf{x})| = i, I_r(C; \mathbf{y}) \subseteq I_r(C; \mathbf{x}),\ |I_r(C; \mathbf{y})| = 2\}|. \end{aligned}$$

In other words, $P_r(n, i, \mathbf{x})$ denotes the maximum number of words $\mathbf{y}$ such that $I_r(C; \mathbf{y}) \subseteq I_r(C; \mathbf{x})$ and $|I_r(C; \mathbf{y})| = 2$, where C is an r-identifying code satisfying $|I_r(C; \mathbf{x})| = i$. Clearly, $P_r(n, i, \mathbf{0}) = P_r(n, i, \mathbf{x})$ for every $\mathbf{x} \in \mathbb{F}^n$ because $\mathbb{F}^n$ is vertex transitive. Therefore, we can denote $P_r(n, i, \mathbf{0}) = P_r(n, i)$. The definition of $P_r(n, i)$ is somewhat complicated. However, it arises naturally from the proof of the following theorem (see the inequality (2.1)). We will examine the function more closely after Theorem 2.2.1.

Theorem 2.2.1. *Define*

$$a = \min_{i=3,\ldots,V(n,r)} \left\{ 2 + \frac{(i-2)(\binom{2r}{r} - 1)}{\binom{2r}{r} + P_r(n, i) - 1} \right\}.$$

Then we have

$$M_r(n) \geq \frac{a \cdot 2^n}{V(n, r) + a - 1}.$$

Proof. Let $C \subseteq \mathbb{F}^n$ be an r-identifying code. Denote by V_i the words which are r-covered by exactly i codewords. Let $\mathbf{x} \in \mathbb{F}^n$ be a word r-covered by exactly two codewords (if any such words $\mathbf{x}$ exist). By [22, Theorem 2.4.8] we know that there are at least $\binom{2r}{r}$ words in $\mathbb{F}^n$ covering both of these codewords (and one of them is $\mathbf{x}$). Therefore, for each word which is r-covered by exactly two codewords there are at least $\binom{2r}{r}-1$ words which are r-covered by at least *three* codewords, since the code C is r-identifying. On the other hand, if $\mathbf{y} \in \mathbb{F}^n$ is r-covered by $i \geq 3$ codewords, then there are at most $P_r(n,i)$ words $\mathbf{z} \in \mathbb{F}^n$ such that $I_r(\mathbf{z}) \subseteq I_r(\mathbf{y})$ and $|I_r(\mathbf{z})| = 2$. Hence, by counting in two ways the number of pairs $\{\mathbf{x}, \mathbf{y}\}$ such that $\mathbf{x} \in V_2$, $\mathbf{y} \in V_i$ $(i \geq 3)$ and $I_r(\mathbf{x}) \subseteq I_r(\mathbf{y})$, we have

$$\left(\binom{2r}{r} - 1\right)|V_2| \leq \sum_{i=3}^{V(n,r)} P_r(n,i)|V_i|. \tag{2.1}$$

Denoting $|C| = K$, it is immediate that at most K words are r-covered by a single codeword. In other words, we have $|V_1| \leq K$.

Now, by counting in two ways the number of pairs $\{\mathbf{x}, \mathbf{c}\}$, where $\mathbf{x} \in \mathbb{F}^n$ and $\mathbf{c} \in C$ is r-covered by $\mathbf{x}$, and by using the inequality (2.1), we have

$$\begin{aligned} K \cdot V(n,r) &= \sum_{i=1}^{V(n,r)} i|V_i| \\ &= a \cdot 2^n - (a-1)|V_1| - (a-2)|V_2| + \sum_{i=3}^{V(n,r)} (i-a)|V_i| \\ &\geq a \cdot 2^n - (a-1)K + \sum_{i=3}^{V(n,r)} \left(i - a - \frac{a-2}{\binom{2r}{r} - 1} P_r(n,i) \right) |V_i|. \end{aligned}$$

By the definition of a, we know that $i - a - \frac{a-2}{\binom{2r}{r}-1} P_r(n,i) \geq 0$ for all $3 \leq i \leq V(n,r)$. Thus, we have

$$K \cdot V(n,r) \geq a \cdot 2^n - (a-1)K.$$

The claim immediately follows from this inequality. □

In applying Theorem 2.2.1, we need to find as good upper bounds for $P_r(n,i)$ as possible. Since we are considering r-identifying codes, we immediately know that $P_r(n,i) \leq \binom{i}{2}$. This estimate provides useful upper bound for small i. On the other hand, it is also clear that $P_r(n,i) \leq V(n,2r)$, since only words in $B_{2r}(\mathbf{0})$ are able to r-cover codewords in $B_r(\mathbf{0})$. (Actually, we can further say that $P_r(n,i) \leq V(n,2r) - 1$, since the word $\mathbf{0}$ is r-covered by $i(\geq 3)$ codewords.) This upper bound works better with bigger i. Together these two estimates imply that

$$P_r(n,i) \leq \min\left\{ \binom{i}{2}, V(n,2r) \right\}. \tag{2.2}$$

In what follows, we present two ways to improve the bound $V(n, 2r)$ for $P_r(n, i)$. The first approach, which is based on Theorem 2.2.2, concentrates on bounding the number of words of weight $2r - 1$ and $2r$ that contribute to the value $P_r(n, i)$. For the second method, assume that w is an integer such that $r \leq w \leq 2r$. Theorem 2.2.4 provides then an upper bound for the number of words in $B_w(\mathbf{0})$ that are r-covered by at most two codewords of $B_r(\mathbf{0})$ when there are exactly i codewords in $B_r(\mathbf{0})$. These two approaches will then be combined (as explained later).

In the following, we define two auxiliary functions, namely $F_r(n, w)$ and $f_r(n, w)$. The relation between these functions and the considered function $P_r(n, i)$ is examined after Theorem 2.2.2. Let now $C \subseteq \mathbb{F}^n$ be an r-identifying code and w be an integer such that $2r - 1 \leq w \leq 2r$. Then define

$$F_r(n, w) = F_r(C; n, w) = \{\mathbf{a} \in S_w(\mathbf{0}) \subseteq \mathbb{F}^n \mid I_r(\mathbf{a}) \subseteq I_r(\mathbf{0}),\ |I_r(\mathbf{a})| = 2\}.$$

In other words, $F_r(n, w)$ consists of the words $\mathbf{a} \in S_w(\mathbf{0})$ such that $I_r(\mathbf{a})$ is a subset of $I_r(\mathbf{0})$ with exactly two codewords. Define also

$$\begin{aligned} f_r(n, w) = \max_{D \subseteq B_r(\mathbf{0})} |\{&I_r(D; \mathbf{x}) \mid \mathbf{x} \in S_w(\mathbf{0}) \subseteq \mathbb{F}^n, \\ &I_r(D; \mathbf{x}) \subseteq I_r(D; \mathbf{0}),\ |I_r(D; \mathbf{x})| = 2\}|. \end{aligned}$$

Clearly, for any r-identifying code $C \subseteq \mathbb{F}^n$ we have $|F_r(n, w)| \leq f_r(n, w)$. (Notice also that the value $f_r(n, w)$ remains unchanged if the word $\mathbf{0}$ is replaced by an arbitrary word $\mathbf{y} \in \mathbb{F}^n$.)

Theorem 2.2.2. *Let $C \subseteq \mathbb{F}^n$ be an r-identifying code. If k and w are integers such that $2r + 1 \leq k \leq n$ and $2r - 1 \leq w \leq 2r$, then*

$$|F_r(n, w)| \leq \frac{f_r(k, w)}{\binom{k}{w}} \binom{n}{w}.$$

Proof. Let $\mathbf{y} \in \mathbb{F}^n$ be a word of weight k. Define

$$H(\mathbf{y}) = \{\mathbf{x} \in \mathbb{F}^n \mid \mathit{supp}(\mathbf{x}) \subseteq \mathit{supp}(\mathbf{y})\}.$$

Let us now consider pairs $\{\mathbf{y}, \mathbf{x}\}$, where $\mathbf{y}$ is a word of weight k and $\mathbf{x} \in H(\mathbf{y}) \cap F_r(n, w)$. Since $2r - 1 \leq w \leq 2r$, each word in $B_r(\mathbf{0})$ that is r-covered by a word in $S_w(\mathbf{0}) \cap H(\mathbf{y})$ belongs to $H(\mathbf{y})$. Therefore, for each word $\mathbf{y}$ of weight k, there exists at most $f_r(k, w)$ different words in $H(\mathbf{y}) \cap F_r(n, w)$. Thus, by counting in two ways the number of pairs $\{\mathbf{y}, \mathbf{x}\}$, we have

$$\binom{n - w}{k - w} |F_r(n, w)| \leq \binom{n}{k} f_r(k, w).$$

Furthermore, we have

$$
\begin{aligned}
|F_r(n,w)| &\le f_r(k,w)\frac{\binom{n}{k}}{\binom{n-w}{k-w}} = f_r(k,w)\frac{\binom{n}{k}\binom{k}{w}}{\binom{n-w}{k-w}\binom{k}{w}} \\
&= f_r(k,w)\frac{\binom{n}{w}\binom{n-w}{k-w}}{\binom{n-w}{k-w}\binom{k}{w}} = \frac{f_r(k,w)}{\binom{k}{w}}\binom{n}{w}.
\end{aligned}
$$

□

Theorem 2.2.2 tells us that the ratio of $|F_r(n,w)|$ to $|S_w(\mathbf{0})| = \binom{n}{w}$ is at most $f_r(k,w)/\binom{k}{w}$ when $n \ge k$ and $2r-1 \le w \le 2r$. Therefore, the value $f_r(k,w)$ for small $k(< n)$ provides an upper bound for the number of words in $F_r(n,w)$. Furthermore, the number of words of weight w that contribute to the value $P_r(n,i)$ is at most

$$\max_{C \subseteq \mathbb{F}^n} |\{F_r(C;n,w) \mid C \text{ is } r\text{-identifying}\}|$$

and, therefore, is bounded from above by $\left(f_r(k,w)/\binom{k}{w}\right)\binom{n}{w}$. Thus, if we know the values $f_r(k_1, 2r-1)$ and $f_r(k_2, 2r)$ with k_1 and k_2 being positive integers, then we have for $n \ge \max\{k_1, k_2\}$ that

$$P_r(n,i) \le \sum_{j=0}^{2r-2}\binom{n}{j} + \frac{f_r(k_1, 2r-1)}{\binom{k_1}{2r-1}}\binom{n}{2r-1} + \frac{f_r(k_2, 2r)}{\binom{k_2}{2r}}\binom{n}{2r}. \qquad (2.3)$$

The following theorem gives an easy upper bound for $f_r(2r+1, 2r)$.

Theorem 2.2.3. *We have*

$$f_r(2r+1, 2r) \le 2r.$$

Proof. Assume to the contrary that $f_r(2r+1, 2r) \ge 2r+1$, i.e. $f_r(2r+1, 2r) = 2r+1$ since $f_r(2r+1, 2r) \le \binom{2r+1}{2r} = 2r+1$. Let $D \subseteq B_r(\mathbf{0})$ be a set such that it attains this value. Now there exist at least three codewords in $S_r(\mathbf{0})$ (or we are done). Therefore, there exist two codewords $\mathbf{c}_1, \mathbf{c}_2 \in S_r(\mathbf{0})$ such that $supp(\mathbf{c}_1) \cap supp(\mathbf{c}_2)$ is nonempty, i.e. $|supp(\mathbf{c}_1) \cup supp(\mathbf{c}_2)| < 2r$. Hence, there exist words $\mathbf{x}_1, \mathbf{x}_2 \in S_{2r}(\mathbf{0})$ such that $\{\mathbf{c}_1, \mathbf{c}_2\}$ is included in $I_r(D;\mathbf{x}_1)$ and $I_r(D;\mathbf{x}_2)$. This is a contradiction, since we assumed that each word in $S_{2r}(\mathbf{0})$ is r-covered by a different set of codewords of size two. □

It should be remarked that the upper bound for $f_r(2r+1, 2r)$ in the previous theorem can be attained. For example, when $r = 2$, it is easy to verify that the set $D = \{00101, 00110, 01001, 01010\}$ attains the value $f_r(5,4) = \binom{5}{4} - 1 = 4$.

Notice that (when n grows) most of the words in $B_{2r}(\mathbf{0})$ belong to $S_{2r}(\mathbf{0})$. Hence, it is natural to concentrate on the values $f_r(n, 2r)$ needed in applying Theorem 2.2.2. The following values provide significant improvements over Theorem 2.2.3:

$$f_2(9,4) = 60,\ \ f_3(9,6) = 42 \text{ and } f_4(10,8) = 24. \tag{2.4}$$

These values have been obtained by extensive computer searches. The method used in the computations is explained in [33].

Using the previous values, we are able to significantly decrease the last term in the equation (2.3). When $r = 2$, it is also straightforward to check that $f_2(5,3) \leq 9$. Thus, when $r = 2$ and $n \geq 9$, we have by the equation (2.3) that ($k_1 = 5, k_2 = 9$)

$$P_2(n,i) \leq \min\left\{\binom{i}{2}, \sum_{j=0}^{2}\binom{n}{j} + \frac{9}{10}\binom{n}{3} + \frac{60}{126}\binom{n}{4}\right\}. \tag{2.5}$$

Actually, this inequality together with Theorem 2.2.1 provides the best known lower bounds for $M_r(n)$ when $r = 2$ and $n \geq 9$ (see Table 2.1).

The considerations above provided an efficient way to estimate the number of words of weight $2r-1$ and $2r$ contributing to the value $P_r(n,i)$. The following theorem, on the other hand, gives an upper bound for the number of words in $B_w(\mathbf{0})$ ($r \leq w \leq 2r$) that are r-covered by at most two codewords in $B_r(\mathbf{0})$ when there are exactly i codewords in $B_r(\mathbf{0})$.

Theorem 2.2.4. *Assume that C is an r-identifying code in $\mathbb{F}^n$. Let w be an integer such that $r \leq w \leq 2r$ and i be the number of codewords in the ball $B_r(\mathbf{0})$. Define*

$$f_{r,b}(n,i) = \min\left\{\sum_{k=b}^{r}\sum_{j=0}^{\lfloor\frac{k-b}{2}\rfloor}\binom{r+b}{k-j}\binom{n-r-b}{j}, i\right\} \text{ and}$$

$$D_{n,r,w}(i_1,\ldots,i_r) = V(n,w) + \sum_{b=1}^{w-r}\frac{2\binom{n}{r+b}}{f_{r,b}(n, i_1+\cdots+i_r)-2}$$
$$-\sum_{b=1}^{w-r}\left(\frac{\sum_{k=b}^{r}\left(i_k\sum_{j=0}^{\lfloor(k-b)/2\rfloor}\binom{k}{n-k}\binom{n-k}{r+b-k+j}\right)}{f_{r,b}(n,i)-2}\right).$$

Then the number of words in $B_w(\mathbf{0})$ that are r-covered by at most two codewords in $B_r(\mathbf{0})$ is at most

$$\max\{D_{n,r,w}(i_1,\ldots,i_r) \mid i_1+\cdots+i_r = i \text{ and } 0 \leq i_j \leq \binom{n}{j} \text{ for } 1 \leq j \leq r\}.$$

Proof. Let $C \subseteq \mathbb{F}^n$ be an r-identifying code. Let k and b be integers such that $1 \leq k \leq r$ and $1 \leq b \leq k$. Let us then count the number of words of weight $r+b$ that a word of weight k r-covers. If a word $\mathbf{x} \in \mathbb{F}^n$ of weight k r-covers a word $\mathbf{y} \in \mathbb{F}^n$ of weight $r+b$, then there are at most $\lfloor \frac{k-b}{2} \rfloor$ positions such that the bits in $\mathbf{x}$ and $\mathbf{y}$ in the corresponding positions are 1 and 0, respectively. Thus, each word of weight k now r-covers

$$\sum_{j=0}^{\lfloor \frac{k-b}{2} \rfloor} \binom{k}{j} \binom{n-k}{r+b-k+j}$$

words of weight $r+b$. In a similar way, it can be showed that each word of weight $r+b$ r-covers

$$\sum_{j=0}^{\lfloor \frac{k-b}{2} \rfloor} \binom{r+b}{k-j} \binom{n-r-b}{j}$$

words of weight k. Therefore, each word of weight $r+b$ r-covers

$$f_{r,b}(n) = \sum_{k=b}^{r} \sum_{j=0}^{\lfloor \frac{k-b}{2} \rfloor} \binom{r+b}{k-j} \binom{n-r-b}{j}$$

words in $B_r(\mathbf{0})$.

Define

$$T_r(j,w) = |\{\, \mathbf{x} \in S_w(\mathbf{0}) \mid I_r(\mathbf{x}) \subseteq I_r(\mathbf{0}),\ |I_r(\mathbf{x})| = j\}|,$$

and denote

$$i_k = |I_r(\mathbf{0}) \cap S_k(\mathbf{0})|, \text{ where } 1 \leq k \leq r.$$

Notice that $i = i_1 + \cdots + i_r$. Now denote $f_{r,b}(n,i) = \min\{f_{r,b}(n), i\}$. Notice that the value $f_{r,b}(n,i)$ now tells us the maximum number of codewords in $B_r(\mathbf{0})$ that each word of weight $r+b$ r-covers. (Actually, here the integer i could be replaced by the sum $i_b + \cdots + i_r$, but it would complicate the analysis of the function $D_{n,r,w}(i_1, \ldots, i_r)$ and did not provide any improvements in the numerical cases we considered.)

By counting in two ways the number of pairs $\{\mathbf{x}, \mathbf{c}\}$ with $\mathbf{x} \in S_{r+b}(\mathbf{0})$

and $\mathbf{c} \in I_r(\mathbf{x}) \cap B_r(\mathbf{0})$, we have

$$\sum_{k=b}^{r} \left(i_k \sum_{j=0}^{\lfloor (k-b)/2 \rfloor} \binom{k}{j} \binom{n-k}{r+b-k+j} \right) = \sum_{j=0}^{f_{r,b}(n,i)} jT_r(j, r+b)$$
$$\leq 2 \sum_{j=0}^{2} T_r(j, r+b) + f_{r,b}(n,i) \sum_{j=3}^{f_{r,b}(n,i)} T_r(j, r+b)$$
$$= 2 \left(\binom{n}{r+b} - \sum_{j=3}^{f_{r,b}(n,i)} T_r(j, r+b) \right) + f_{r,b}(n,i) \sum_{j=3}^{f_{r,b}(n,i)} T_r(j, r+b)$$
$$= 2 \binom{n}{r+b} + (f_{r,b}(n,i) - 2) \sum_{j=3}^{f_{r,b}(n,i)} T_r(j, r+b).$$

Consequently, we obtain that

$$\sum_{j=3}^{f_{r,b}(n,i)} T_r(j, r+b) \geq \frac{\sum_{k=b}^{r} \left(i_k \sum_{j=0}^{\lfloor (k-b)/2 \rfloor} \binom{k}{j} \binom{n-k}{r+b-k+j} \right) - 2\binom{n}{r+b}}{f_{r,b}(n,i) - 2}.$$

Now we have

$$V(n, 2r) - \sum_{b=1}^{w-r} \sum_{j=3}^{f_{r,b}(n,i)} T_r(j, r+b)$$
$$\leq V(n, 2r) - \sum_{b=1}^{w-r} \left(\frac{\sum_{k=b}^{r} \left(i_k \sum_{j=0}^{\lfloor (k-b)/2 \rfloor} \binom{k}{n-k} \binom{n-k}{r+b-k+j} \right) - 2\binom{n}{r+b}}{f_{r,b}(n,i) - 2} \right)$$
$$= D_{n,r,w}(i_1, \ldots, i_r).$$

For given $i_1, \ldots, i_r$ the above inequality provides an upper bound for the number of words in $B_w(\mathbf{0})$ which are r-covered by at most two codewords. Hence, when we maximize the function $D_{n,r}(i_1, \ldots, i_r)$ over all different choices of $i_1, \ldots, i_r$ such that $i_1 + \cdots + i_r = i$ and $i_j \leq \binom{n}{j}$, the claim immediately follows. $\square$

In applying Theorem 2.2.4, we have to be able to solve the following optimization problem for fixed i ($3 \leq i \leq V(n, r)$):

$$\max\{D_{n,r,w}(i_1, \ldots, i_r) \mid i_1 + \cdots + i_r = i \text{ and } i_j \leq \binom{n}{j} \text{ for } j = 1, \ldots, r\}.$$

Indeed, this problem can be solved quite efficiently using the following procedure:

1. Calculate the coefficients of i_k in $D_{n,r,w}(i_1,\ldots,i_r)$. (Notice that the sum $i_1+\cdots+i_r$ is equal to the fixed constant i.)

2. Sort i_k in decreasing order regarding the coefficients of i_k. Let the sorted list be $i_{j_1},\ldots,i_{j_r}$.

3. Let s be the largest integer such that
$$\sum_{k=1}^{s-1}\binom{n}{j_k}\le i.$$
Now the function $D_{n,r,w}(i_1,\ldots,i_r)$ is maximized by choosing $i_{j_s}=i-\sum_{k=1}^{s-1}\binom{n}{j_k}$, $i_{j_k}=\binom{n}{j_k}$ for $k=1,\ldots,s-1$ and $i_{j_{s+1}}=\cdots=i_r=0$.

We have now presented two ways (Theorem 2.2.2 and 2.2.4) to improve the upper bound (2.2) for $P_r(n,i)$. When $r=2$, the best known lower bounds are obtained by using only Theorem 2.2.2 (see the equation (2.5)). However, when $r>2$, to obtain the best lower bounds we combine the two methods explained above. For example, when $r=3$, we obtain the following inequality by combining Theorem 2.2.2 and 2.2.4:
$$P_3(n,i)\le\min\left\{\binom{i}{2},\max_{i_1+i_2+i_3=i}D_{n,3,5}(i_1,i_2,i_3)+\frac{42}{84}\binom{n}{6}\right\},\qquad(2.6)$$
where $0\le i_j\le\binom{n}{j}$ for all $j=1,2,3$. This inequality improves the known lower bounds, when $n\ge 19$ (see Table 2.1).

When $r=4$ and $r=5$, the known lower bounds are improved in a similar way to the inequality (2.6), i.e. we use Theorem 2.2.2 to estimate the number of words of weight $2r$ contributing to the value $P_r(n,i)$ and Theorem 2.2.4 for smaller weights. This method improves the known lower bounds for $r=4$ when $n\ge 28$ and for $r=5$ when $n\ge 37$. (Notice that when $r=4$ we have the value $f_4(10,8)=24$ by extensive computer searches and when $r=5$ we have the estimate $f_5(11,10)\le 10$ by Theorem 2.2.3.) In particular, we have $M_5(37)\ge 542868$ (the best previously known bound is 539088).

As we have seen, Theorem 2.2.1 improves lower bounds when $r\ge 2$ and n is large enough. With small n the best known lower bounds are provided by the third part of Theorem 1 in [60] (by Karpovsky *et al*). For completeness and to cover efficiently also the case $r\ge n/2$ (see [8]) this result is rephrased in the following theorem.

Theorem 2.2.5. *Let $C\subseteq\mathbb{F}^n$ be an r-identifying code. Then we have*
$$|C|\cdot V(n,r)\ge\sum_{i=1}^{s}i\binom{|C|}{i}+(s+1)\left(2^n-\sum_{i=1}^{s}\binom{|C|}{i}\right),\qquad(2.7)$$

where s is the largest integer such that

$$\sum_{i=1}^{s} \binom{|C|}{i} \leq 2^n. \tag{2.8}$$

If $n/2 \leq r \leq n-1$, then we have

$$|C| \cdot V(n, n-r-1) \geq \sum_{i=1}^{s} i\binom{|C|}{i} + (s+1)\left(2^n - \sum_{i=0}^{s} \binom{|C|}{i}\right), \tag{2.9}$$

where s is the largest integer such that

$$\sum_{i=0}^{s} \binom{|C|}{i} \leq 2^n. \tag{2.10}$$

Proof. Denote again by V_i the words which are r-covered by exactly i codewords. Let $C \subseteq \mathbb{F}^n$ be an r-identifying code. Counting the number of pairs $\{\mathbf{x}, \mathbf{c}\}$ where $\mathbf{x} \in \mathbb{F}^n$, $\mathbf{c} \in C$ and $d(\mathbf{x}, \mathbf{c}) \leq r$, we get

$$|C| \cdot V(n, r) = \sum_{i=0}^{V(n,r)} i|V_i|.$$

Clearly, $|V_0| = 0$. To bound from below the right hand side of the equation, we make for small $i = 1, 2, \ldots$ the cardinalities $|V_i|$ as large as possible. Trivially, $|V_i| \leq \binom{|C|}{i}$. But up to which i can we do this? Clearly, up to s defined in (2.8). The rest of the words (i.e. the words in V_i with $i \geq s+1$) are covered by at least $s+1$ times. This yields (2.7).

Suppose then $n/2 \leq r \leq n-1$. By [8], we know that an r-identifying code has the property that also the sets $I_{n-r-1}(\mathbf{x})$ are different, but (exactly) one can be empty. Hence, for the radius $n-r-1$, we can count exactly as above, but now $|V_0| \leq 1$ and we use s as defined in (2.10). □

The previous theorem can be used to compute lower bounds for r-identifying codes in the following way: we start our computation from a known lower bound and then increase the size of the code until the equation (2.7) is satisfied. The first value satisfying the equation is then our new lower bound.

In particular, Theorem 2.2.5 gives us that $M_3(5) \geq 9$. On the other hand, we know by [38] that $M_3(5) \leq 10$. The following theorem shows that, indeed, $M_3(5) = 10$.

Theorem 2.2.6. *We have $M_3(5) = 10$.*

Proof. By the considerations above, we know that $9 \leq M_3(5) \leq 10$. Assume then to the contrary that there exists a 3-identifying code $C \subseteq \mathbb{F}^5$ of size 9. By [8], the code C has the property that also the sets $I_1(C; \mathbf{x})$ are different for all $\mathbf{x} \in \mathbb{F}^5$ (although one of these sets can be empty). As before, let V_i denote the set of words which are 1-covered by exactly i codewords of C.

If $|V_j| \geq 1$ for some $j = 4, \ldots, V(5,1)$, then as in (2.9) we get

$$54 = |C| \cdot V(5,1) \geq 1 \cdot 0 + 9 \cdot 1 + 21 \cdot 2 + 1 \cdot 4 = 55,$$

which is a contradiction. Hence, $|V_j| = 0$ for every $j = 4, \ldots, V(5,1)$.

Assume now that $|V_3| \leq 1$. Then, as in the proof of Theorem 2.2.1, we have

$$|V_2| \leq \sum_{i=3}^{V(5,1)} P_r(5,i)|V_i| \leq P_r(5,3) \leq \binom{3}{2} = 3.$$

Since $|V_3| \leq 1$, the number of words in V_2 is at least 21. This observation together with the previous inequality leads to a contradiction. Therefore, $|V_3| \geq 2$. However, this implies that

$$54 = |C| \cdot V(5,1) \geq 1 \cdot 0 + 9 \cdot 1 + 20 \cdot 2 + 2 \cdot 3 \geq 55,$$

which is a contradiction. Thus, there does not exist a 3-identifying code of length 5 with 9 codewords. Hence, we have $M_3(5) = 10$. □

In Table 2.1 we have listed the best known lower bounds for $r = 2, 3$ and $2 \leq n \leq 30$. For the best known upper bounds, we refer to [12] and Section 2.6.

2.3 Results related to the conjecture $M_{r+t}(n+m) \leq M_r(n)M_t(m)$

In [9], the question whether the inequality

$$M_{r+t}(n+m) \leq M_r(n)M_t(m) \tag{2.11}$$

holds is stated as an open problem. For $r = t = 1$, it has been shown in [34] that this inequality indeed holds. In what follows, we approach this conjecture with two different methods. The first method is based on code constructions using direct sums and the second one estimates the sizes of the optimal codes in the inequality. In particular, with the first method we are able to show that $M_{r+1}(n+m) \leq 4M_r(n)M_1(m)$ and using the second approach we obtain that the inequality holds when m is relatively small compared to n.

Let us first start by presenting a theorem that considers the structure of r-identifying codes. For the theorem, recall that a pair of vertices $\mathbf{x}$ and

Table 2.1: Lower bounds (the best previously known bounds in the parentheses) on the cardinalities of r-identifying codes for $r = 2$ and $r = 3$

n	$M_2(n)$	$M_3(n)$
2	-	-
3	f 7	-
4	f 6	f 15
5	a 6	d 10
6	a 8	a 7
7	e 14	a 8
8	a 17	a 10
9	b 27 (a 26)	a 13
10	b 43 (c 41)	a 18
11	b 71 (c 67)	a 25
12	b 118 (c 112)	a 39
13	b 199 (c 190)	a 61
14	b 341 (c 326)	a 95
15	b 590 (c 567)	a 151
16	b 1033 (c 995)	a 241
17	b 1824 (c 1761)	a 383
18	b 3244 (c 3141)	a 608
19	b 5809 (c 5638)	b 974 (a 959)
20	b 10465 (c 10179)	b 1656 (c 1593)
21	b 18949 (c 18471)	b 2839 (c 2722)
22	b 34487 (c 33674)	b 4909 (c 4731)
23	b 63029 (c 61647)	b 8549 (c 8276)
24	b 115664 (c 113288)	b 14985 (c 14562)
25	b 213004 (c 208921)	b 26420 (c 25899)
26	b 393602 (c 386520)	b 46833 (c 45784)
27	b 729508 (c 717218)	b 83425 (c 81749)
28	b 1356002 (c 1334510)	b 149271 (c 146575)
29	b 2526996 (c 2489423)	b 268200 (c 263829)
30	b 4721086 (c 4654848)	b 483728 (c 478179)

Key to the table

a	Theorem 2.2.5 by Karpovsky *et al.* [60]
b	Theorem 2.2.1
c	Theorem 2 in [35]
d	Theorem 2.2.6
e	By computer search in [34]
f	Blass *et al.* in [8]

$\mathbf{y}$ in $\mathbb{F}^n$ are r-separated by a code C if the intersection of the symmetric difference $B_r(\mathbf{x}) \bigtriangleup B_r(\mathbf{y})$ and the code C is nonempty.

Theorem 2.3.1. *Let $C \subseteq \mathbb{F}^n$ be an r-identifying code and k be an integer such that $1 \leq k \leq r$. Assume that $\mathbf{x} \in \mathbb{F}^n$ and the set $(\cup_{i=r-k+1}^{r} S_i(\mathbf{x})) \cap C$ is empty. Then all the pairs of words in $\cup_{i=1}^{k} S_i(\mathbf{x})$ are r-separated by the codewords of $(\cup_{i=1}^{k} S_{r+i}(\mathbf{x})) \cap C$.*

Proof. Assume that $\mathbf{y} \in \cup_{i=1}^{k} S_i(\mathbf{x})$. We first show that $I_r(\mathbf{y}) \cap (\cup_{i=1}^{k} S_{r+i}(\mathbf{x}))$ is nonempty. Assume to the contrary that $I_r(\mathbf{y}) \cap (\cup_{i=1}^{k} S_{r+i}(\mathbf{x})) = \emptyset$. Then, by the assumption $(\cup_{i=r-k+1}^{r} S_i(\mathbf{x})) \cap C = \emptyset$, it follows that

$$I_r(\mathbf{y}) = I_r(\mathbf{y}) \cap B_{r-k}(\mathbf{x}) = I_r(\mathbf{x}),$$

which is a contradiction since it was assumed that the code C is r-identifying.

Now it remains to be shown that all the pairs of words in $\cup_{i=1}^{k} S_i(\mathbf{x})$ are r-separated by the codewords of $(\cup_{i=1}^{k} S_{r+i}(\mathbf{x})) \cap C$. Assume to the contrary that there exist words $\mathbf{y}_1$ and $\mathbf{y}_2$ in $\cup_{i=1}^{k} S_i(\mathbf{x})$ such that $I_r(\mathbf{y}_1) \cap (\cup_{i=1}^{k} S_{r+i}(\mathbf{x})) = I_r(\mathbf{y}_2) \cap (\cup_{i=1}^{k} S_{r+i}(\mathbf{x}))$. Then, by the assumption, it follows that

$$\begin{aligned} I_r(\mathbf{y}_1) &= (I_r(\mathbf{y}_1) \cap (\cup_{i=1}^{k} S_{r+i}(\mathbf{x}))) \cup (I_r(\mathbf{y}_1) \cap B_{r-k}(\mathbf{x})) \\ &= (I_r(\mathbf{y}_2) \cap (\cup_{i=1}^{k} S_{r+i}(\mathbf{x}))) \cup (I_r(\mathbf{y}_2) \cap B_{r-k}(\mathbf{x})) = I_r(\mathbf{y}_2), \end{aligned}$$

which is again a contradiction with the fact that the code C is r-identifying. □

Choosing $k = 1$ in the previous theorem, we obtain the following essential consequence.

Corollary 2.3.2. *Let $C \subseteq \mathbb{F}^n$ be an r-identifying code. Then for all $\mathbf{x} \in \mathbb{F}^n$ there exists $\mathbf{c} \in C$ such that $d(\mathbf{c}, \mathbf{x}) = r$ or $r + 1$.*

For the rest of the section, let t be a positive integer. In the subsequent considerations, we often refer to the following condition for a given code C:

$$\forall\, \mathbf{x}, \mathbf{y} \in \mathbb{F}^n : I_t(C; \mathbf{x}) \setminus I_{t-1}(C; \mathbf{y}) \neq \emptyset. \tag{2.12}$$

A code $C \subseteq \mathbb{F}^n$ is said to be *t-separating* if all the pairs of words in $\mathbb{F}^n$ are t-separated by C. Notice that the only difference between the definitions of t-identifying and t-separating codes is the fact that one I-set is allowed to be empty in the case of t-separating codes. We will also use the following notations:

- The smallest cardinality of a t-identifying code satisfying the condition (2.12) is denoted by $\widehat{M}_t(n)$.

- The smallest cardinality of a t-identifying code which is also $(t-1)$-separating and satisfies the condition (2.12) is denoted by $\widehat{M}_{t,t-1}(n)$.

- The smallest cardinality of a t-identifying code such that for every $\mathbf{x} \in \mathbb{F}^n$ there exists a codeword *exactly* at distance t from $\mathbf{x}$, i.e. $S_r(\mathbf{x}) \cap C \neq \emptyset$, is denoted by $M'_t(n)$.

- We denote by $M''_1(n)$ the smallest cardinality of a 1-identifying and 2-fold 1-covering code. We have $M''_1(n) \leq 2M_1(n)$. Indeed, if $C \subseteq \mathbb{F}^n$ is a 1-identifying code, then $|\{\mathbf{x} \in \mathbb{F}^n \mid |I_1(C;\mathbf{x})| = 1\}| \leq |C|$. Hence, we need to add at most $|C|$ codewords to the code C to get a 2-fold 1-covering.

Theorem 2.3.3. *For $r \geq 0$ and $t \geq 1$ we have*

$$M_{r+t}(n+m) \leq \begin{cases} M_r(n)\widehat{M}_{t,t-1}(m), \\ M'_r(n)\widehat{M}_t(m) \end{cases} \tag{2.13}$$

and

$$M_{r+1}(n+m) \leq M'_r(n)M''_1(m). \tag{2.14}$$

Moreover,

$$M'_r(n) \leq \begin{cases} 2M_r(n), \\ 2^{r+1}M_r(n-r-1), \\ M_r(n) + 2^r K(n-r,r). \end{cases} \tag{2.15}$$

Especially,

$$M_{r+t}(n+m) \leq 2M_r(n)\widehat{M}_t(n) \tag{2.16}$$

$$M_{r+1}(n+m) \leq 4M_r(n)M_1(m). \tag{2.17}$$

Proof. Let us first prove the inequalities (2.13). Let $C_1 \subseteq \mathbb{F}^n$ be an r-identifying code, and $C_2 \subseteq \mathbb{F}^m$ be a t-identifying and $(t-1)$-separating code satisfying the condition (2.12). We will first show that $C = C_1 \oplus C_2 \subseteq \mathbb{F}^{n+m}$ is an $(r+t)$-identifying code. It is easy to see that C is an $(r+t)$-covering code. Therefore, in order to prove that C is $(r+t)$-identifying, it is enough to show that $I_{r+t}(\mathbf{x}) \triangle I_{r+t}(\mathbf{y}) \neq \emptyset$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{F}^{n+m}$ $(\mathbf{x} \neq \mathbf{y})$. Let $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2)$, $\mathbf{y} = (\mathbf{y}_1, \mathbf{y}_2) \in \mathbb{F}^{n+m}$, where $\mathbf{x}_1, \mathbf{y}_1 \in \mathbb{F}^n$, $\mathbf{x}_2, \mathbf{y}_2 \in \mathbb{F}^m$ and $\mathbf{x} \neq \mathbf{y}$.

- Suppose first $\mathbf{x}_1 \neq \mathbf{y}_1$. Then there exists $\mathbf{c}_1 \in I_r(C_1;\mathbf{x}_1) \triangle I_r(C_1;\mathbf{y}_1)$. Without loss of generality, we may assume that $\mathbf{c}_1 \in I_r(C_1;\mathbf{x}_1) \setminus I_r(C_1;\mathbf{y}_1)$. Since the code C_2 satisfies the condition (2.12), there exists a codeword $\mathbf{c}_2 \in C_2$ such that $\mathbf{c}_2 \in I_t(C_2;\mathbf{x}_2) \setminus I_{t-1}(C_2;\mathbf{y}_2)$. Hence, we have $d((\mathbf{c}_1,\mathbf{c}_2),(\mathbf{x}_1,\mathbf{x}_2)) \leq r+t$ and $d((\mathbf{c}_1,\mathbf{c}_2),(\mathbf{y}_1,\mathbf{y}_2)) \geq r+1+t$.

- Suppose then $\mathbf{x}_1 = \mathbf{y}_1$. By Corollary 2.3.2, there exists $\mathbf{c}_1 \in C_1$ such that $d(\mathbf{c}_1, \mathbf{x}_1) = r$ or $r + 1$. Assume first that $d(\mathbf{c}_1, \mathbf{x}_1) = r$. Since C_2 is a t-identifying code and $\mathbf{x}_2 \neq \mathbf{y}_2$, there exists a codeword $\mathbf{c}_2 \in C_2$ such that $\mathbf{c}_2 \in I_t(\mathbf{x}_2) \triangle I_t(\mathbf{y}_2)$. Therefore, $(\mathbf{c}_1, \mathbf{c}_2) \in I_{r+t}(\mathbf{x}) \triangle I_{r+t}(\mathbf{y})$. Assume then that $d(\mathbf{c}_1, \mathbf{x}_1) = r+1$. Since C_2 is also a $(t-1)$-separating code and $\mathbf{x}_2 \neq \mathbf{y}_2$, there exists a codeword $\mathbf{c}_2 \in C_2$ such that $\mathbf{c}_2 \in I_{t-1}(\mathbf{x}_2) \triangle I_{t-1}(\mathbf{y}_2)$. Hence, $(\mathbf{c}_1, \mathbf{c}_2) \in I_{r+t}(\mathbf{x}) \triangle I_{r+t}(\mathbf{y})$. Thus, we have proved that $C_1 \oplus C_2$ is an $(r + t)$-identifying code. This proves the first part of (2.13).

Let $C_3 \subseteq \mathbb{F}^n$ be an r-identifying code such that for every $\mathbf{x} \in \mathbb{F}^n$ there exists a codeword at distance exactly r from it and $C_4 \subseteq \mathbb{F}^m$ a t-identifying code satisfying the condition (2.12). Showing that $C_3 \oplus C_4 \subseteq \mathbb{F}^{n+m}$ is an $(r + t)$-identifying code is similar to the proof described above. However, in the second part of the proof we can assume that there exists a codeword $\mathbf{c}_1 \in \mathbb{F}^n$ such that $d(\mathbf{x}_1, \mathbf{c}_1) = r$.

Let us now move on to the inequality (2.14). It is easy to see that a 1-identifying and 2-fold 1-covering code satisfies the condition (2.12) for $t = 1$. Hence, we have $\widehat{M}_1(n) \leq M_1''(n)$. Therefore, the result immediately follows from (2.13).

In considering the inequalities (2.13) and (2.14), the estimates (2.15) for $M_r'(n)$ prove to be useful. Let us prove the first upper bound $M_r'(n) \leq 2M_r(n)$. Let $C \subseteq \mathbb{F}^n$ be an r-identifying code attaining $M_r(n)$ and $\mathbf{e} \in \mathbb{F}^n$ be a word of weight 1. We will show that the code $C' = C \cup (C + \mathbf{e})$ is an r-identifying code such that for every $\mathbf{x} \in \mathbb{F}^n$ the set $S_r(\mathbf{x}) \cap C'$ is nonempty. Since C is an r-identifying code, the code C' is also r-identifying. Thus, it remains to prove that for every $\mathbf{x} \in \mathbb{F}^n$ the set $S_r(\mathbf{x}) \cap C'$ is nonempty.

Assume then $\mathbf{x} \in \mathbb{F}^n$. If the set $S_r(\mathbf{x}) \cap C \neq \emptyset$, then trivially $S_r(\mathbf{x}) \cap C' \neq \emptyset$. Assume now that $S_r(\mathbf{x}) \cap C = \emptyset$. Consider then the word $\mathbf{y} = \mathbf{x} + \mathbf{e}$. By Theorem 2.3.1, we know that $I_r(C; \mathbf{y}) \cap S_{r+1}(\mathbf{x}) \neq \emptyset$. Therefore, there exists a codeword $\mathbf{c} \in C$ such that

$$d(\mathbf{x}, \mathbf{c}) = r + 1 \text{ and } d(\mathbf{y}, \mathbf{c}) = r.$$

Now for the codeword $\mathbf{c} + \mathbf{e} \in C'$ we know that

$$d(\mathbf{x}, \mathbf{c} + \mathbf{e}) = d(\mathbf{x} + \mathbf{e}, \mathbf{c}) = d(\mathbf{y}, \mathbf{c}) = r.$$

Thus, $S_r(\mathbf{x}) \cap C' \neq \emptyset$ and the claim $M_r'(n) \leq 2M_r(n)$ follows.

The second estimate for $M_r'(n)$ comes from the following observation. If $C \subseteq \mathbb{F}^{n-r-1}$ is an r-identifying code, then $\mathbb{F}^{r+1} \oplus C$ is an r-identifying code by [34, Theorem 4] and it clearly has the property that every word is r-covered (by a codeword) at distance exactly r.

For the last inequality concerning $M'_r(n)$, we take the union of an r-identifying code $C_1 \subseteq \mathbb{F}^m$ and $C_2 \oplus \mathbb{F}^r$, where $C_2 \subseteq \mathbb{F}^{n-r}$ is an r-covering code. As above, we get the desired code.

The inequalities (2.16) and (2.17) are immediate corollaries of the inequalities (2.13), (2.14) and (2.15). □

The inequalities (2.13) resemble the conjecture (2.11), although here we need to require more than just r- and t-identification from the underlying codes. However, it should be noted that there exist *optimal* t-identifying codes which automatically are $(t-1)$-separating and satisfy the condition (2.12). For example, the words $0000, 0011, 0100, 0110, 1000$, and 1001 in $\mathbb{F}^4$ form an optimal 2-identifying code which is also 1-separating and satisfy the condition (2.12). The same also holds for the optimal 3-identifying code in $\mathbb{F}^6$ formed by the words $000010, 001111, 010100, 011001, 101000, 110011$, and 111110.

In [12, Theorem 3] it is proved that when $1 \leq t < m \leq r$ we have

$$M_{r+t}(n+m) \leq 2^m M_r(n). \tag{2.18}$$

Assume first that $t = 1$. Since $C = \mathbb{F}^m \setminus \{1^m\}$ is clearly a 1-identifying and 0-separating code satisfying the condition (2.12), we obtain by (2.13) that $M_{r+1}(n+m) \leq (2^m - 1)M_r(n)$. Then, using (2.17), we have further improvements to (2.18). Namely, by Theorem 2.1.1, we have

$$M_1(m) \leq \frac{9}{2} \cdot \frac{2^m}{m+1} < 2^{m-2} - 1$$

when $m \geq 18$ and, by the tables in [12], this also holds for $m \geq 8$.

In the next theorem we improve (2.18) using (2.16) when $t \geq 2$ and $m \geq 2t$. This is done by giving an upper bound for $\widehat{M}_t(m)$ using a method inspired by Delsarte and Piret [22, p. 320].

Theorem 2.3.4. *Let $m \geq 2t$.*

$$M_{r+t}(n+m) \leq 2 \left\lceil \frac{2^m}{\min\{\binom{m}{t}, 2\binom{m-1}{t}\}} 2m \ln 2 \right\rceil M_r(n).$$

Proof. We first prove that there exists a t-identifying code in $\mathbb{F}^m$ satisfying the condition (2.12) of the cardinality

$$K \leq \left\lceil \frac{2^m}{\min\{\binom{m}{t}, 2\binom{m-1}{t}\}} 2m \ln 2 \right\rceil.$$

Then, by combining this result with (2.16) we get the desired inequality.

We first need two preliminary observations. If $\mathbf{x}, \mathbf{y} \in \mathbb{F}^m$ and $\mathbf{x} \neq \mathbf{y}$, then by [22, Theorem 2.4.8] $|B_t(\mathbf{x}) \triangle B_t(\mathbf{y})| \geq 2\binom{m-1}{t}$. If $\mathbf{x}, \mathbf{y} \in \mathbb{F}^m$ (here $\mathbf{x}$ can be equal to $\mathbf{y}$), then $|B_t(\mathbf{x}) \setminus B_{t-1}(\mathbf{y})| \geq \binom{m}{t}$ because $|B_t(\mathbf{x})| - |B_{t-1}(\mathbf{y})| = \binom{m}{t}$.

Let C be a subset of $\mathbb{F}^m$. Denote by $P_t(C)$ the number of (unwanted) pairs $\{\mathbf{x}, \mathbf{y}\}$ ($\mathbf{x}, \mathbf{y} \in \mathbb{F}^m$) such that

$$I_t(C; \mathbf{x}) \setminus I_{t-1}(C; \mathbf{y}) = \emptyset, \text{ if } \mathbf{x} = \mathbf{y}, \text{ or} \tag{2.19}$$

$$I_t(C; \mathbf{x}) \setminus I_{t-1}(C; \mathbf{y}) = \emptyset \text{ or } I_t(C; \mathbf{x}) \triangle I_t(C; \mathbf{y}) = \emptyset, \text{ if } \mathbf{x} \neq \mathbf{y}. \tag{2.20}$$

We further denote by $\mathcal{C}_K$ the set of all codes of size K in $\mathbb{F}^m$. Now we have

$$\begin{aligned}
\sum_{C \in \mathcal{C}_K} P_t(C) &= \sum_{C \in \mathcal{C}_K} \sum_{\mathbf{x} \in \mathbb{F}^m} \sum_{\substack{\mathbf{y} \in \mathbb{F}^m \\ (2.19) \text{ or } (2.20)}} 1 \\
&= \sum_{\mathbf{x} \in \mathbb{F}^m} \sum_{\mathbf{y} \in \mathbb{F}^m} \sum_{\substack{C \in \mathcal{C}_K \\ (2.19) \text{ or } (2.20)}} 1 \\
&\leq \sum_{\mathbf{x} \in \mathbb{F}^m} \sum_{\substack{\mathbf{y} \in \mathbb{F}^m \\ \mathbf{x} \neq \mathbf{y}}} \binom{2^m - \min\{|B_t(\mathbf{x}) \triangle B_t(\mathbf{y})|, |B_t(\mathbf{x}) \setminus B_{t-1}(\mathbf{y})|\}}{K} \\
&\quad + \sum_{\mathbf{x} \in \mathbb{F}^m} \binom{2^m - |B_t(\mathbf{x}) \setminus B_{t-1}(\mathbf{x})|}{K} \\
&\leq \sum_{\mathbf{x} \in \mathbb{F}^m} \sum_{\mathbf{y} \in \mathbb{F}^m} \binom{2^m - \min\{2\binom{m-1}{t}, \binom{m}{t}\}}{K} \\
&\leq 2^{2m} \binom{2^m - \min\{2\binom{m-1}{t}, \binom{m}{t}\}}{K}.
\end{aligned}$$

Choose now

$$K = \left\lceil \frac{2^m}{\min\{2\binom{m-1}{t}, \binom{m}{t}\}} \ln 2^{2m} \right\rceil.$$

Then, using the fact that $(1 - 1/x)^x < 1/e$ for $x \geq 1$, we have

$$\begin{aligned}
\frac{\sum_{C \in \mathcal{C}_K} P_t(C)}{|\mathcal{C}_K|} &\leq \frac{2^{2m} \binom{2^m - \min\{2\binom{m-1}{t}, \binom{m}{t}\}}{K}}{\binom{2^m}{K}} \\
&= 2^{2m} \prod_{i=0}^{K-1} \frac{2^m - \min\{2\binom{m-1}{t}, \binom{m}{t}\} - i}{2^m - i} \\
&\leq 2^{2m} \left(1 - \frac{\min\{2\binom{m-1}{t}, \binom{m}{t}\}}{2^m}\right)^K \\
&< 1.
\end{aligned}$$

The previous inequality now implies that there exists $C \in \mathcal{C}_K$ such that $P_t(C) = 0$. This means that C is a t-identifying code satisfying the condition (2.12). Thus, the claim follows. $\square$

Assume then that in the previous theorem we have $m = 2t$. Using Stirling's inequality, we get $\widehat{M_t}(m) < 2.24m^{3/2}$. By [8, Section 3], we also know that now a t-identifying code in $\mathbb{F}^m$ is also $(t-1)$-separating. Therefore, by the equation (2.13), $M_{r+t}(n+m) < 2.24m^{3/2} \cdot M_r(n)$ when $m = 2t$. On the other hand, by [60, Theorem 1(1)] we know that $M_t(m) \geq m$, so the coefficient which we have is at most $2.24m^{1/2}$ times bigger than what it could be at the best if the conjecture (2.11) holds.

Let us then consider the second approach, which is based on estimating the sizes of optimal identifying codes. The proof of the following theorem is based on this idea.

Theorem 2.3.5. *Let r and t be integers such that $r \geq 2$ and $t \geq 2$, and let $n \geq 2(t+r)$. If*

$$2t \leq a \leq \frac{2n}{9(t+r)} \left(\frac{2^{r+3}(t-1)!(r-1)!}{9(9(t+r))^r} \right)^{\frac{1}{t}},$$

then we have

$$M_{t+r}(n) \leq M_t(a) M_r(n-a).$$

Proof. Let q and q_0 be integers such that $n = q(t+r)+q_0$ and $0 \leq q_0 < t+r$. Suppose first that $q_0 = 0$. By Theorem 2.1.2, we have

$$M_{t+r}(n) \leq M_1(q)^{t+r}.$$

By Theorem 2.1.1, we have

$$M_1(q)^{t+r} \leq \left(\frac{9}{2}\right)^{t+r} \left(\frac{2^q}{q+1}\right)^{t+r} \leq \left(\frac{9}{2}(t+r)\right)^{t+r} \frac{2^n}{n^{t+r}}.$$

If $2 \leq q_0 < t+r$, then there exists a 1-identifying code of length q_0. Hence, by Theorems 2.1.2 and 2.1.1, we obtain that

$$M_{t+r}(n) \leq M_1(q)^{t+r} M_1(q_0) \leq \left(\frac{9}{2}\right)^{t+r+1} (t+r)^{t+r} \frac{2^n}{n^{t+r}}.$$

Finally, assume that $q_0 = 1$. Now there does not exist a 1-identifying code of length q_0. However, we can write $n = q(t+r-1)+q+1$ and, as above, obtain that

$$\begin{aligned} M_{t+r}(n) \leq M_1(q)^{t+r-1} M_1(q+1) &\leq \left(\frac{9}{2}\right)^{t+r} \frac{2^n}{(q+1)^{t+r-1}} \cdot \frac{1}{q+2} \\ &\leq \left(\frac{9}{2}\right)^{t+r} (t+r)^{t+r} \frac{2^n}{n^{t+r}}. \end{aligned}$$

On the other hand, by [60], we have

$$M_r(n) \geq \frac{2^{n+1}}{V(n,r)+1}.$$

Furthermore, if $r \geq 2$ and $n \geq 2r$, we have

$$M_r(n) \geq \frac{2^{n+1}}{V(n,r)+1} \geq \frac{2^{n+1}}{r\binom{n}{r}} \geq \frac{2^{n+1}(r-1)!}{n^r}. \tag{2.21}$$

Therefore, by the assumption of a, we have

$$\begin{aligned} M_{t+r}(n) &\leq \left(\frac{9}{2}\right)^{t+r+1} (t+r)^{t+r} \frac{2^n}{n^{t+r}} \\ &\leq \frac{2^{n+2}(t-1)!(r-1)!}{a^t n^r} \leq M_t(a) M_r(n-a). \end{aligned} \tag{2.22}$$

□

For example, the previous theorem gives $M_5(507) \leq M_3(7)M_2(500)$ as predicted by the conjecture (2.11). In the previous theorem, it is assumed that $r \geq 2$ and $t \geq 2$. However, the arguments go through also when $r = 1$ or $t = 1$ — the estimate (2.21) just has to be replaced with the lower bound

$$M_1(n) \geq \frac{n2^{n+1}}{n^2+n+2}, \tag{2.23}$$

which is presented in [60].

2.4 Results related to the conjecture $M_{r+1}(n) \leq M_r(n)$

The monotonicity of the size of an optimal r-identifying code in $\mathbb{F}^n$ with respect to the radius r is discussed in the seminal paper [60]. In particular, it is conjectured that the function $M_r(n)$ is increasing with respect to r when n is large enough. The following theorem proves that this conjecture indeed holds. The proof of the theorem is based on ideas similar to the ones of Theorem 2.3.5.

Theorem 2.4.1. *Let $r \geq 1$ be a fixed integer. Then there exists a positive integer n_r such that*

$$M_{r+1}(n) \leq M_r(n) \quad \forall\, n \geq n_r.$$

Proof. Assume first that $r \geq 2$. When $n \geq 2r$, we obtain by the inequality (2.21) that

$$M_r(n) \geq 2(r-1)! \cdot \frac{2^n}{n^r}. \tag{2.24}$$

On the other hand, by (2.22), we have

$$M_{r+1}(n) \leq \left(\frac{9}{2}\right)^{r+2} (r+1)^{r+1} \frac{2^n}{n^{r+1}}.$$

Clearly, there exists a positive integer n_r such that

$$\frac{\frac{9}{2}\left(\frac{9}{2}(r+1)\right)^{r+1}}{n} \leq 2(r-1)! \quad \forall\, n \geq n_r.$$

Thus, by combining the two inequalities above, the claim follows. For $r = 1$, the claim can similarly be proved by replacing (2.24) with (2.23). $\square$

2.5 Results related to the conjecture $M_1(n+1) \leq 2M_1(n)$

In [9], it has been stated as an open problem whether $M_1(n+1) \leq 2M_1(n)$ holds. In the same paper, a weaker result, which says that $M_1(n+1) \leq 3M_1(n)$, is shown. The following theorem improves this result by showing that $M_1(n+1) \leq (2+\varepsilon_n)M_1(n)$, where $\varepsilon_n \to 0$ as $n \to \infty$.

Theorem 2.5.1. *Assume that $n \geq 2$. Then we have*

$$M_1(n+1) \leq \left(2 + \frac{1}{n+1}\right) M_1(n).$$

Proof. Let $C \subseteq \mathbb{F}^n$ be an optimal 1-identifying code attaining $M_1(n)$. Define

$$\begin{aligned} C_1 &= \{\, \mathbf{x} \mid \mathbf{x} \in C\ ,\ |I_1(\mathbf{x})| = 1\} \text{ and} \\ N_1 &= \{\, \mathbf{x} \mid \mathbf{x} \in \mathbb{F}^n,\ \mathbf{x} \notin C,\ |I_1(\mathbf{x})| = 1\}. \end{aligned}$$

Clearly, $|C_1 \cup N_1| \leq M_1(n)$. Assume first $|C_1| \leq M_1(n)/(n+1)$. Let $D_1 = C \oplus \mathbb{F} \subseteq \mathbb{F}^{n+1}$. Denote $O_p = \mathbb{F}^n \oplus \{p\}$ where $p \in \mathbb{F}$. Assume $\mathbf{x} = (\mathbf{x}', a) \in \mathbb{F}^{n+1}$ with $\mathbf{x}' \in \mathbb{F}^n$ and $a \in \mathbb{F}$. Since C is 1-identifying, the set $I_1(D_1; \mathbf{x})$ can coincide only with the I-sets of words in O_{a+1}. If $|I_1(C; \mathbf{x}')| \geq 2$, then the word $\mathbf{x}$ is uniquely identified by its I-set $I_1(D_1; \mathbf{x})$ since each word in O_{a+1} 1-covers a unique word in O_a. It can now be assumed that $|I_1(C; \mathbf{x}')| = 1$.

Assume that $\mathbf{x}' \in N_1$, i.e. $I_1(C; \mathbf{x}') = \{\mathbf{x}' + \mathbf{e}\}$, where $\mathbf{e} \in \mathbb{F}^n$ is a word of weight 1. The only word in O_{a+1} which 1-covers the codeword $(\mathbf{x}' + \mathbf{e}, a)$ is the word $(\mathbf{x}' + \mathbf{e}, a+1)$. However, $|I_1(C; \mathbf{x}' + \mathbf{e})| \geq 2$ and therefore, as above, it can be said that $\mathbf{x}$ is uniquely identified. If $\mathbf{x}' \in C_1$, then clearly

$I_1(D_1;(\mathbf{x}',a)) = I_1(D_1;(\mathbf{x}',a+1))$. But such a problematic case can be solved by adding one codeword to the code D_1 for each such $\mathbf{x}'$. Thus, we have

$$M_1(n+1) \leq \left(2+\frac{1}{n+1}\right) M_1(n).$$

Assume then $|C_1| > M_1(n)/(n+1)$. Let $\mathbf{z} \in \mathbb{F}^n$ be a word of weight 1. Consider then a code $D_2 \subseteq \mathbb{F}^{n+1}$ defined as

$$D_2 = (C \oplus \{0\}) \cup ((C+\mathbf{z}) \oplus \{1\}).$$

Assume that $\mathbf{x} = (\mathbf{x}', a) \in \mathbb{F}^{n+1}$ with $\mathbf{x}' \in \mathbb{F}^n$ and $a \in \mathbb{F}$. If there is at least two codewords in the intersection of $I_1(D_2;\mathbf{x})$ and O_a, then (as above) the word $\mathbf{x}$ is uniquely identified by its I-set $I_1(D_2;\mathbf{x})$.

Assume now that $I_1(D_2;\mathbf{x}) \cap O_a = \{\mathbf{x}\}$. Since C is a 1-identifying code in $\mathbb{F}^n$, it is clear that the word $\mathbf{x}$ and all the words of $\mathbb{F}^{n+1}$ except $(\mathbf{x}', a+1)$ are 1-separated by D_2. Furthermore, since $(\mathbf{x}'+\mathbf{z}, a+1)$ belongs to D_2, also the words $(\mathbf{x}',a)$ and $(\mathbf{x}',a+1)$ are 1-separated by D_2. Thus, $\mathbf{x}$ is uniquely identified by its I-set.

Assume then that $I_1(D_2;(\mathbf{x}',a)) \cap O_a = \{(\mathbf{x}'+\mathbf{e},a)\}$, where $\mathbf{e} \in \mathbb{F}^n$ and $w(\mathbf{e}) = 1$. In order to show that $\mathbf{x}$ is uniquely identified by its I-set, it is enough to show that the words $(\mathbf{x}',a)$ and $(\mathbf{x}'+\mathbf{e},a+1)$ are 1-separated by D_2. If $\mathbf{e} \neq \mathbf{z}$, then $(\mathbf{x}'+\mathbf{e}+\mathbf{z}, a+1) \in D_2$ belongs to the symmetric difference $B_1(\mathbf{x}',a) \triangle B_1(\mathbf{x}'+\mathbf{e},a+1)$ and we are done. Hence, we may assume that $\mathbf{e} = \mathbf{z}$. Although we now have $I_1(D_2;(\mathbf{x}',a)) = I_1(D_2;(\mathbf{x}'+\mathbf{e},a+1))$, each of these problematic cases can be handled by adding one suitable codeword to the set D_2. Moreover, there exists a word $\mathbf{e}' \in \mathbb{F}^n$ of weight 1 such that

$$|\{\mathbf{x} \in \mathbb{F}^n \mid \mathbf{x} \notin C,\ I_1(C;\mathbf{x}) = \{\mathbf{x}+\mathbf{e}'\}\}| \leq \frac{|N_1|}{n}.$$

Therefore, if we choose $\mathbf{z} = \mathbf{e}'$, then we obtain that

$$\begin{aligned} M_1(n+1) &\leq 2M_1(n) + \frac{|N_1|}{n} \leq 2M_1(n) + \frac{M_1(n)-|C_1|}{n} \\ &\leq 2M_1(n) + \frac{M_1(n) - \frac{M_1(n)}{n+1}}{n} = \left(2+\frac{1}{n+1}\right) M_1(n). \end{aligned}$$

□

2.6 Constructions with locating-dominating codes

In this section, we introduce a new direct sum method to find constructions for r-identifying codes with the smallest known cardinalities. This approach is based on the following theorem.

Theorem 2.6.1. *Let $C_1 \subseteq \mathbb{F}^n$ be a 1-identifying code which is also a 2-fold 1-covering and has the property that it is k-locating-dominating for all $1 \leq k \leq r+1 \leq n-2$. Let $C_2 \subseteq \mathbb{F}^m$ be an r-identifying code. Then $C_1 \oplus C_2 \subseteq \mathbb{F}^{n+m}$ is an $(r+1)$-identifying code.*

Proof. It is clear that $C_1 \oplus C_2$ is an $(r+1)$-covering. Let $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2), \mathbf{y} = (\mathbf{y}_1, \mathbf{y}_2) \in \mathbb{F}^{n+m}$, where $\mathbf{x}_1, \mathbf{y}_1 \in \mathbb{F}^n$, $\mathbf{x}_2, \mathbf{y}_2 \in \mathbb{F}^m$, and $\mathbf{x} \neq \mathbf{y}$.

1) If $\mathbf{x}_2 = \mathbf{y}_2$, then we have $\mathbf{x}_1 \neq \mathbf{y}_1$. Thus, there exists $\mathbf{c}_1 \in I_1(C_1; \mathbf{x}_1) \triangle I_1(C_1; \mathbf{y}_1)$. Without loss of generality, we may assume that $\mathbf{c}_1 \in I_1(C_1; \mathbf{x}_1) \setminus I_1(C_1; \mathbf{y}_1)$. Suppose $\mathbf{c}_2 \in I_r(C_2; \mathbf{x}_2)$ and $d(\mathbf{c}_2, \mathbf{x}_2) \geq d(\mathbf{c}, \mathbf{x}_2)$ for all $\mathbf{c} \in I_r(C_2, \mathbf{x}_2)$.

- If $d(\mathbf{x}_2, \mathbf{c}_2) = r$, then $(\mathbf{c}_1, \mathbf{c}_2) \in I_{r+1}(C_1 \oplus C_2; \mathbf{x}) \setminus I_{r+1}(C_1 \oplus C_2; \mathbf{y})$.
- If $0 \leq d(\mathbf{c}_2, \mathbf{x}_2) \leq r-1$, then, by Corollary 2.3.2, there exists $\mathbf{c}_2' \in C_2$ such that $d(\mathbf{c}_2', \mathbf{x}_2) = r+1$. If $\mathbf{x}_1 \in C_1$ (or similarly if $\mathbf{y}_1 \in C_1$), then $(\mathbf{x}_1, \mathbf{c}_2') \in I_{r+1}(C_1 \oplus C_2; \mathbf{x}) \setminus I_{r+1}(C_1 \oplus C_2; \mathbf{y})$. Suppose then $\mathbf{x}_1, \mathbf{y}_1 \notin C_1$ and denote $h = d(\mathbf{c}_2, \mathbf{x}_2)$. Because C_1 is $(r-h+1)$-locating-dominating, there exists $\mathbf{c}_1' \in I_{r-h+1}(C_1; \mathbf{x}_1) \triangle I_{r-h+1}(C_2; \mathbf{y}_1)$, and we have $(\mathbf{c}_1', \mathbf{c}_2) \in I_{r+1}(C_1 \oplus C_2; \mathbf{x}) \triangle I_{r+1}(C_1 \oplus C_2; \mathbf{y})$.

2) Assume then that $\mathbf{x}_2 \neq \mathbf{y}_2$. Because C_2 is an r-identifying code there exists $\mathbf{c}_2 \in I_r(C_2; \mathbf{x}_2) \triangle I_r(C_2; \mathbf{y}_2)$. Without loss of generality, we may assume that $\mathbf{c}_2 \in I_r(C_2; \mathbf{x}_2) \setminus I_r(C_2; \mathbf{y}_2)$. Because C_1 is a 2-fold 1-covering, there exists $\mathbf{c}_1 \in I_1(C_1; \mathbf{x}_1)$ such that $d(\mathbf{c}_1, \mathbf{y}_1) \geq 1$. Now we have $(\mathbf{c}_1, \mathbf{c}_2) \in I_{r+1}(C_1 \oplus C_2; \mathbf{x}) \setminus I_{r+1}(C_1 \oplus C_2; \mathbf{y})$. □

The condition that the identifying code C_1 is a 2-fold 1-covering increases the cardinality only slightly (as can be seen in [34]). The extra requirement that C_1 is also k-locating-dominating for $1 \leq k \leq n-2$ is not demanding cardinality-wise either. Indeed, the best known 1-identifying codes which are also 2-fold 1-coverings given in Table 2.2, are immediately k-locating-dominating for all $1 \leq k \leq n-2$ as well. However, in the following example, it is shown that *not* every 1-identifying and 2-fold 1-covering code is k-locating-dominating for all $1 \leq k \leq n-2$.

Example 2.6.2. The code $\mathbb{F}^6 \setminus (\{000000, 100000\} \cup \{(0, \mathbf{v}) \in \mathbb{F}^6 \mid w(\mathbf{v}) = 3\} \cup \{(1, \mathbf{u}) \in \mathbb{F}^6 \mid \mathbf{w}(\mathbf{u}) = 3\})$ is 1-identifying and 2-fold 1-covering but it is not 3-locating-dominating. However, the smallest known 1-identifying and 2-fold 1-covering code of length 6, which has been presented in [34], has cardinality 22 and it is also k-locating-dominating for all $1 \leq k \leq 4$.

Regardless of the previous example, in the following theorem, it is shown that 1-identifying code is always also 1-, 2- and $(n-2)$-locating-dominating.

Theorem 2.6.3. *Let $n \geq 4$. A 1-identifying code $C \subseteq \mathbb{F}^n$ is always k-locating-dominating for $k = 1$, $k = 2$ and $k = n-2$.*

Proof. By the definition, 1-identifying code is also 1-locating-dominating. By [38], we know that 1-identifying code is also $(n-2)$-identifying and hence also $(n-2)$-locating-dominating.

Let $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$ be two different non-codewords. We will show that $I_2(\mathbf{x}) \triangle I_2(\mathbf{y})$ is nonempty. There exists $\mathbf{c} \in I_1(\mathbf{x}) \triangle I_1(\mathbf{y})$. Without loss of generality, we may assume that $\mathbf{c} \in I_1(\mathbf{x}) \setminus I_1(\mathbf{y})$. If $d(\mathbf{c}, \mathbf{y}) > 2$, we are done. Suppose then that $d(\mathbf{c}, \mathbf{y}) = 2$. This implies that $d(\mathbf{x}, \mathbf{y}) = 1$ or 3.

a) Suppose $d(\mathbf{x}, \mathbf{y}) = 3$. Because $n \geq 4$, there is $\mathbf{z} \in S_1(\mathbf{x}) \setminus S_2(\mathbf{y})$. There exists a codeword $\mathbf{c}' \in C$ such that $d(\mathbf{c}', \mathbf{z}) \leq 1$ and hence $d(\mathbf{c}', \mathbf{y}) \geq 3$. Thus, we have $\mathbf{c}' \in I_2(\mathbf{x}) \setminus I_2(\mathbf{y})$.

b) Suppose $d(\mathbf{x}, \mathbf{y}) = 1$. If there is $\mathbf{c}' \in (C \cap S_2(\mathbf{x})) \setminus S_1(\mathbf{y})$, then we are done. Thus, suppose that if $\mathbf{c}' \in S_2(\mathbf{x}) \cap C$, then $d(\mathbf{c}', \mathbf{y}) = 1$. Denote by $\mathbf{z}$ the unique word in $(S_1(\mathbf{c}) \cap S_1(\mathbf{y})) \setminus \{\mathbf{x}\}$. Now $I_1(\mathbf{z}) = I_1(\mathbf{c})$ unless there is a codeword $\mathbf{c}^*$ ($\mathbf{c}^* \neq \mathbf{c}$) such that $d(\mathbf{c}^*, \mathbf{z}) = 1$ and so $d(\mathbf{c}^*, \mathbf{y}) = 2$ and $d(\mathbf{c}^*, \mathbf{x}) = 3$. This completes the proof. $\square$

In Theorem 2.6.1, we could let k run until $n-1$, or even more. If $k \geq n$, then $C_1 = \mathbb{F}^n \setminus \{\mathbf{0}\}$, and if $k = n-1$, then the code C_1 can be (at the best) a half space as the following theorem shows.

Theorem 2.6.4. *The code $C = \{0\} \oplus \mathbb{F}^{n-1} \subseteq \mathbb{F}^n$ is k-locating-dominating for all $1 \leq k \leq n-1$. When $k = n-1$, the code C is optimal.*

Proof. For all $1 \leq k \leq n-1$ and $\mathbf{x} \in \mathbb{F}^{n-1}$, we have

$$I_k(C; (1, \mathbf{x})) = I_{k-1}(C; (0, \mathbf{x})) = \{0\} \oplus (\mathbb{F}^{n-1} \cap B_{k-1}(\mathbf{x})).$$

Hence, it is clear that for all $1 \leq k \leq n-1$ and $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n \setminus C$, $\mathbf{x} \neq \mathbf{y}$ we have $I_k(\mathbf{x}) \neq I_k(\mathbf{y})$.

Let us then prove the optimality for $k = n-1$. Let $C \subseteq \mathbb{F}^n$ be any $(n-1)$-locating-dominating code. Let us consider a word $\mathbf{x}$ and its complement word $\mathbf{x} + 11\ldots1$. One of them has to be a codeword otherwise these two words cannot be separated. Because this is true for all words, at least half of the words belong to the code. $\square$

In order to find codes as small as possible satisfying the conditions of C_1 in Theorem 2.6.1, we have used extensive computer searches. For detailed explanation on the computations, we refer the interested reader to [32]. The found codes, which are 1-identifying, 2-fold 1-covering and k-locating-dominating for $1 \leq k \leq n-2$, are presented in Table 2.2. In addition to these codes, it is also easy to verify that the smallest known 1-identifying and 2-fold 1-covering code of length 9 and of cardinality 128 from [34] is k-locating-dominating for $k = 1, \ldots, 7$. Therefore, by applying these codes to Theorem 2.6.1, we obtain the following result.

Table 2.2: The codes, which are 1-identifying, 2-fold 1-covering and k-locating-dominating for $1 \le k \le n-2$, presented. The codewords are binary representations of the listed integers.

n	$\|C\|$	codewords
7	38	84 7 53 15 89 68 8 119 76 34 59 6 112 32 74 45 38 127 97 82 67 22 60 95 102 69 111 17 104 51 120 61 90 24 25 62 65 43
8	70	208 118 128 232 92 215 114 108 66 63 205 204 74 209 24 181 142 231 9 90 101 42 189 15 220 167 6 131 34 75 51 244 203 149 127 110 154 129 85 44 64 164 69 241 226 173 25 155 214 186 121 178 37 56 97 234 20 190 22 228 43 48 199 169 249 255 19 31 115 140
10	249	334 793 723 447 197 569 385 139 450 107 466 430 829 250 487 70 649 1003 171 403 322 401 959 865 561 674 344 939 406 871 935 15 621 790 692 432 705 315 832 35 249 483 452 40 309 359 777 797 976 429 479 773 1014 864 222 601 605 433 272 369 926 213 617 512 517 888 152 109 374 671 582 332 588 932 500 727 8 550 1021 886 763 28 524 786 164 714 712 768 337 883 114 505 276 143 702 488 762 127 89 818 629 1010 179 559 652 1002 224 751 37 667 529 853 328 258 323 632 22 527 826 966 23 564 819 1011 973 142 231 707 163 349 65 952 631 474 680 86 681 133 489 647 801 211 453 238 475 812 112 168 934 26 75 57 189 858 689 710 908 183 852 405 855 936 917 538 814 165 414 558 162 62 384 351 988 985 903 130 5 922 876 220 602 720 694 1020 638 34 69 247 396 262 508 642 106 899 265 252 539 740 669 523 356 442 68 778 446 660 60 610 1013 546 749 410 383 975 592 656 290 19 308 362 304 567 847 157 100 287 921 297 395 969 737 283 606 295 891 144 494 205 933 376 779 230 200 455 792 579 113 736 301

Theorem 2.6.5. *We have* $M_4(n) \leq 38M_3(n-7)$, $M_5(n) \leq 70M_4(n-8)$, $M_6(n) \leq 128M_5(n-9)$ *and* $M_7(n) \leq 249M_6(n-10)$.

By the previous theorem, for example, we get that $M_5(27) \leq 70M_4(19) \leq 58450$ and $M_5(28) \leq 70M_4(20) \leq 120400$ since $M_4(19) \leq 835$ and $M_4(20) \leq 1720$ by [12]. The best previously known bounds are 83840 and 167680, respectively.

The codes of Table 2.2 can also be used for bounding $M_1(n)$ from above. Namely, it has been proved in [34] that if a code $C \subseteq \mathbb{F}^n$ is 1-identifying and 2-fold 1-covering, then the code $D = \{(\pi(\mathbf{u}), \mathbf{u}, \mathbf{u}+\mathbf{v}) \mid \mathbf{u} \in \mathbb{F}^n, \mathbf{v} \in C\} \subseteq \mathbb{F}^{2n+1}$ is 1-identifying and 2-fold 1-covering. (Recall that $\pi(\mathbf{u})$ is used to denote the parity check bit of $\mathbf{u}$.) Combining this result with the codes from our table, we obtain the following improvements on the previous records.

Theorem 2.6.6. *We have* $M_1(17) \leq 17920$ *and* $M_1(21) \leq 254976$.

The best previously known upper bounds for the cardinalities of 1-identifying codes of lengths 17 and 21 are 18558 and 262144, respectively; these results are from [12].

Chapter 3

Identification in cycles and paths

In this chapter, which is based on the paper [59], we consider r-identifying codes in cycles and paths. The chapter begins with some preliminary definitions. In Section 3.1, we also summarise known results, which have previously been presented in [6], [36], [70] and [77]. Then, in Sections 3.2 and 3.3, we determine in the remaining cases the sizes of optimal r-identifying codes in cycles and paths, respectively. The remaining cases have also independently been determined in [20] (using methods different from ours).

3.1 Preliminaries

Let n be an integer such that $n \geq 3$. A cycle $\mathcal{C}_n = (V_n, E_n)$ is a graph such that the set of vertices is defined as $V_n = \{v_i \mid i \in \mathbb{Z}_n\}$ and the set of edges is defined as

$$E_n = \{v_i v_{i+1} \mid i = 0, 1, \ldots, n-2\} \cup \{v_{n-1} v_0\}.$$

Recall that the size of an optimal r-identifying code in a given finite graph G is denoted by $M_r(G)$. The exact values of $M_1(\mathcal{C}_n)$ and $M_2(\mathcal{C}_n)$ have been presented in [36] and [70], respectively. For general r, the following results are known:

- By Bertrand *et al.* [6], we know that if n is even and $n \geq 2r + 4$, then $M_r(\mathcal{C}_n) = n/2$. Moreover, they showed that $M_r(\mathcal{C}_{2r+2}) = 2r + 1$.
- In [36], it is shown that if $n = 2r + 3$, then $M_r(\mathcal{C}_n) = \lfloor 2n/3 \rfloor$.
- If n is odd, $3r + 2 \leq n \leq 8r + 1$, $n \neq 4r + 3$ and $\gcd(2r + 1, n) = 1$, then $M_r(\mathcal{C}_n) = (n + 1)/2$. Moreover, we have $M_r(\mathcal{C}_{4r+3}) = 2r + 3$. These results are also from [36].

- If n is odd, $n \geq 3r + 2$ and $\gcd(2r + 1, n) > 1$, then by [36]

$$M_r(\mathcal{C}_n) = \gcd(2r + 1, n) \left\lceil \frac{n}{2 \gcd(2r + 1, n)} \right\rceil .$$

- Assume that n is odd, $n \geq 3r+2$ and $\gcd(2r+1, n) = 1$. Then, by [77], we know that if $n = 2m(2r+1)+1$ or $n = (2m+1)(2r+1)+2r$ for an integer $m \geq 1$, then $M_r(\mathcal{C}_n) = (n+1)/2+1$, else $M_r(\mathcal{C}_n) = (n+1)/2$.

In conclusion, what remains to be shown is the exact values of $M_r(\mathcal{C}_n)$ when n is odd and $2r + 5 \leq n \leq 3r + 1$. (Notice that there are no r-identifying codes in $\mathcal{C}_n$ when $n \leq 2r + 1$.) These remaining cases are solved in Section 3.2.

Let n be a positive integer. For $n \geq 3$, a path $\mathcal{P}_n = (V_n, E'_n)$ is a graph such that the set of vertices V_n is the same as with the cycles and the set of edges $E'_n = E_n \setminus \{v_{n-1}v_0\}$. Furthermore, we define the path $\mathcal{P}_1 = (V_1, E'_1)$, where $E'_1 = \emptyset$, and the path $\mathcal{P}_2 = (V_2, E'_2)$, where $E'_2 = \{v_1v_2\}$. The exact values of $M_1(\mathcal{P}_n)$ and $M_2(\mathcal{P}_n)$ have been presented, respectively, in [6] and [70]. An infinite family of optimal r-identifying codes have been introduced in [6, Theorem 5] giving the following values: $M_r(\mathcal{P}_{2k(2r+1)+1}) = k(2r + 1) + 1$, where k is a non-negative integer and $r \geq 2$. In Section 3.3, we solve the exact values of $M_r(\mathcal{P}_n)$ for general r and n.

In the following sections, we need the concept of a transversal of a graph, which was found useful in considering identification in cycles in [36]. As can be seen from Section 3.3, it also proves valuable in the case of paths. We say that a code $T \subseteq V$ is a *transversal* of G if for each edge $e = uv \in E$ the vertex u or the vertex v belongs to T. In the literature, a transversal is also known as a *vertex cover* [76, p. 102] or an *edge-covering set* [77] of G.

3.2 Optimal identifying codes in cycles

Let r be a positive integer. In this section, we study r-identifying codes in cycles $\mathcal{C}_n = (V_n, E_n)$, where n is an odd integer and $2r + 5 \leq n \leq 3r + 1$. Throughout the thesis, the indices of the vertices v_i of $\mathcal{C}_n$ are calculated modulo n. Let t be a positive integer. For the following considerations, we define a graph $\mathcal{C}'_{(n,t)} = (V_n, F_n)$, where $F_n = \{v_iv_{i+t} \mid i \in \mathbb{Z}_n\}$. Notice that if C is an r-identifying code in $\mathcal{C}_n$, then C is also a transversal of $\mathcal{C}'_{(n,2r+1)}$ since the adjacent vertices v_i and v_{i+1} ($i \in \mathbb{Z}_n$) are r-separated by C. We also define $Q_t(i) = \{v_i, v_{i+1}, \ldots, v_{i+t-1}\}$. In other words, the set $Q_t(i)$ consists of t consecutive vertices of the cycle $\mathcal{C}_n$ starting from the vertex v_i.

The following lower bound on identifying codes in cycles $\mathcal{C}_n$ has been presented in [36, Theorem 1].

Theorem 3.2.1 ([36]). *Let r be a positive integer and $n \geq 2r+2$. Then*

$$M_r(\mathcal{C}_n) \geq \gcd(2r+1, n) \left\lceil \frac{n}{2\gcd(2r+1,n)} \right\rceil .$$

Let n be an odd integer such that $2r+5 \leq n \leq 3r+1$. Then n can be written as follows: $n = 2r+1+p$, where p is an even integer such that $4 \leq p \leq r$. The following lemma provides a new way to characterize r-identifying codes in cycles with small order (for a given r). Notice that in the following lemma for all $i \in \mathbb{Z}_n$ we have $V_n \setminus B_r(v_i) = Q_p(i+r+1)$, i.e. that the set $Q_p(i+r+1)$ denotes the complement of the ball $B_r(v_i)$.

Lemma 3.2.2. *Let r be a positive integer and $n = 2r+1+p$, where p is an even integer such that $4 \leq p \leq r$. Let T be a transversal of $\mathcal{C}'_{(n,2r+1)}$. If u and v are vertices of $\mathcal{C}_n$ such that $d(u,v) \leq p$, then u and v are r-separated by T. Moreover, the transversal T is an r-identifying code in $\mathcal{C}_n$ if and only if there do not exist $i, j \in \mathbb{Z}_n$ such that*

$$Q_p(i) \cap Q_p(j) = \emptyset \text{ and } T \cap (Q_p(i) \cup Q_p(j)) = \emptyset. \tag{3.1}$$

Proof. Let u and v be vertices of $\mathcal{C}_n$ such that $d(u,v) = d \leq p$. Without loss of generality, we may assume that $u = v_k$ and $v = v_{k+d}$ for some $k \in \mathbb{Z}_n$. Clearly, $v_{k-r} \in B_r(u) \setminus B_r(v)$ and $v_{k+r+1} \in B_r(v) \setminus B_r(u)$. Since T is a transversal of $\mathcal{C}'_{(n,2r+1)}$, then $v_{k-r} \in T$ or $v_{k+r+1} \in T$. Hence, the vertices u and v are r-separated by T.

Assume first that the transversal T is an r-identifying code in $\mathcal{C}_n$. Assume to the contrary that there exist $i, j \in \mathbb{Z}_n$ such that $Q_p(i) \cap Q_p(j) = \emptyset$ and $T \cap (Q_p(i) \cup Q_p(j)) = \emptyset$. Since $B_r(v_{i-r-1}) \triangle B_r(v_{j-r-1}) = Q_p(i) \cup Q_p(j)$, then $I_r(T; v_{i-r-1}) \triangle I_r(T; v_{j-r-1}) = \emptyset$ (a contradiction). Recall from the observation above that $Q_p(i)$ and $Q_p(j)$ denote the complements of the balls $B_r(v_{i-r-1})$ and $B_r(v_{j-r-1})$, respectively. Therefore, the condition (3.1) holds.

Assume then that the condition (3.1) holds. Let $u = v_i$ ($i \in \mathbb{Z}_n$). Let us then show that v_i is r-covered by a vertex of T. Assume to the contrary that $I_r(T; v_i) = \emptyset$. Now $T \cap (Q_p(i-p) \cup Q_p(i)) \subseteq I_r(T; v_i)$ and $Q_p(i-p) \cap Q_p(i) = \emptyset$ (a contradiction). Hence, we have $I_r(T; u) \neq \emptyset$. In addition, the first part of the proof shows that vertices $u, v \in V_n$ are r-separated by T if $d(u,v) \leq p$. Let then $u \in V_n$ and $v \in V_n$ be vertices such that $d(u,v) > p$. Now we have $B_r(u) \triangle B_r(v) = Q_p(i) \cup Q_p(j)$ for some $i, j \in \mathbb{Z}_n$. Since $Q_p(i) \cap Q_p(j) = \emptyset$, we obtain by the condition (3.1) that $I_r(T; u) \triangle I_r(T; v) \neq \emptyset$. Thus, T is an r-identifying code in $\mathcal{C}_n$. □

The following theorem provides exact values for $M_r(\mathcal{C}_n)$ when $2r+5 \leq n \leq 3r+1$ and $\gcd(2r+1, n) = 1$.

Theorem 3.2.3. *Let r be a positive integer and $n = 2r + 1 + p$, where p is an even integer such that $4 \leq p \leq r$. Assume that* $\gcd(2r + 1, n) = 1$. *If $n = 2mp+1$ or $n = (2m+1)p+p-1$ with $m \geq 2$, then $M_r(\mathcal{C}_n) = (n+1)/2+1$, else $M_r(\mathcal{C}_n) = (n + 1)/2$.*

Proof. Recall first that $M_r(\mathcal{C}_n) \geq (n + 1)/2$, by Theorem 3.2.1. As stated earlier, each r-identifying code in $\mathcal{C}_n$ is also a transversal of $\mathcal{C}'_{(n,2r+1)}$. Since the graph $\mathcal{C}'_{(n,2r+1)}$ is a cycle and $\mathcal{C}'_{(n,2r+1)} = \mathcal{C}'_{(n,p)}$, the code

$$T = \{v_{ip} \mid 0 \leq i \leq n - 1, i \text{ is even}\}$$

is a transversal of $\mathcal{C}'_{(n,p)}$. Moreover, the set T is, up to rotations, the unique transversal of $\mathcal{C}'_{(n,p)}$ with $(n + 1)/2$ vertices.

The proof now divides into the following five cases depending on n. (Notice that the cases $n = 2mp$ or $n = (2m + 1)p$ are impossible since n is odd.) In the first three cases, it is shown that T is actually an r-identifying code in C_n implying that $M_r(\mathcal{C}_n) = (n+1)/2$. In the last two cases, we then show that T is not r-identifying in $\mathcal{C}_n$ and, therefore, due to the uniqueness of T we have $M_r(\mathcal{C}_n) \geq (n + 1)/2 + 1$. Furthermore, in these cases, we present a code attaining this improved lower bound.

1) Assume first that $n \leq 4p - 1$. Let us then show that there do not exist $i, j \in \mathbb{Z}_n$ such that $Q_p(i) \cap Q_p(j) = \emptyset$ and $T \cap (Q_p(i) \cup Q_p(j)) = \emptyset$. Assume to the contrary that such i and j exist. Since $T \cap Q_p(i) = \emptyset$ and T is a transversal of $\mathcal{C}'_{(n,p)}$, the sets $Q_p(i - p) \subseteq T$ and $Q_p(i + p) \subseteq T$. The fact that $n \geq |Q_p(i - p) \cup Q_p(i) \cup Q_p(i + p) \cup Q_p(j)| = 4p$ implies a contradiction. Therefore, by Lemma 3.2.2, T is an r-identifying code in $\mathcal{C}_n$. Hence, $M_r(\mathcal{C}_n) = (n + 1)/2$ when $n \leq 4p - 1$.

2) Assume then that $n = 2mp + x$ with $m \geq 2$ and $2 \leq x \leq p - 1$. Let us then show that $T \cap Q_p(i) \neq \emptyset$ for any $i \in \mathbb{Z}_n$. Assume to the contrary that $k \in \mathbb{Z}_n$ is such that $T \cap Q_p(k) = \emptyset$. Since $v_k \notin T$ and $v_{k+1} \notin T$, then $v_{k+p} \in T$ and $v_{k+p+1} \in T$. If the vertex v_{k+p} is such that $v_{k+p+ip} \neq v_0$ for any $i = 0, 1, \ldots, 2m$, then $v_{k+p+2mp} = v_{k+p-x} \in T$ (a contradiction). Otherwise, the vertex v_{k+p+1} is such that $v_{k+p+1+ip} \neq v_0$ for any $i = 0, 1, \ldots, 2m$. Then $v_{k+p+1+2mp} = v_{k+p+1-x} \in T$ (a contradiction). Thus, there does not exist $k \in \mathbb{Z}_n$ such that $T \cap Q_p(k) = \emptyset$. Hence, by Lemma 3.2.2, T is an r-identifying code in $\mathcal{C}_n$ and $M_r(\mathcal{C}_n) = (n + 1)/2$.

3) Assume now that $n = (2m+1)p+x$, where $m \geq 2$ and $1 \leq x \leq p-2$. Since $n = (2m + 2)p - (p - x)$, we can write $n = (2m + 2)p - x'$, where $2 \leq x' \leq p - 1$. In what follows, we show that $T \cap Q_p(i) \neq \emptyset$ for any $i \in \mathbb{Z}_n$. Assume to the contrary that $k \in \mathbb{Z}_n$ is such that $T \cap Q_p(k) = \emptyset$. Then, clearly, $v_{k-1} \in T$ and $v_{k-2} \in T$. If the vertex v_{k-1} is such that $v_{k-1+ip} \neq v_0$ for any $i = 0, 1, \ldots, 2m+2$, then $v_{k-1+(2m+2)p} = v_{k-1+x'} \in T$ (a contradiction). Otherwise, the vertex v_{k-2} is such that $v_{k-2+ip} \neq v_0$ for any

$i = 0, 1, \ldots, 2m+2$. Then $v_{k-2+(2m+2)p} = v_{k-2+x'} \in T$ (a contradiction). Hence, by Lemma 3.2.2, T is an r-identifying code in $\mathcal{C}_n$ and $M_r(\mathcal{C}_n) = (n+1)/2$.

4) Consider then the case $n = 2mp+1$ with $m \geq 2$. It is easy to conclude that

$$T = \{v_0\} \cup \bigcup_{i=1}^{m} Q_p((2i-1)p+1).$$

Therefore, $V_n \backslash T = \bigcup_{i=0}^{m-1} Q_p(2ip+1)$. Thus, by Lemma 3.2.2, the transversal T is not an r-identifying code in $\mathcal{C}_n$. Since T is the unique transversal (up to rotations) of $\mathcal{C}'_{(n,p)}$ with size $(n+1)/2$ and every r-identifying code of $\mathcal{C}_n$ is also a transversal of $\mathcal{C}'_{(n,p)}$, we have $M_r(\mathcal{C}_n) \geq (n+1)/2 + 1$.

Define first sets $A_k = \{v_{k+1}, v_{k+2}, \ldots, v_{k+p-2}\} \cup \{v_{k+p}, v_{k+2p-1}\}$, where k is an integer such that $0 \leq k \leq 2(m-1)p$. Define then a code

$$C_1 = \{v_0, v_{2mp}\} \cup \bigcup_{i=0}^{m-1} A_{2ip}.$$

It is straightforward to verify that C_1 is a transversal of $\mathcal{C}'_{(n,p)}$ and that $C_1 \cap Q_p(i) \neq \emptyset$ for any $i \in \mathbb{Z}_n$. Hence, C_1 is an r-identifying code in $\mathcal{C}_n$. Since $|C_1| = (n+1)/2 + 1$, we have $M_r(\mathcal{C}_n) = (n+1)/2 + 1$.

5) Finally, assume that $n = (2m+1)p + p - 1$ with $m \geq 2$. Now we have

$$T = \bigcup_{i=0}^{m} Q_p(2ip) \text{ and } V_n \setminus T = \bigcup_{i=0}^{m-1} Q_p((2i+1)p) \cup Q_{p-1}((2m+1)p).$$

Then, using similar arguments as in the previous case, we have $M_r(\mathcal{C}_n) \geq (n+1)/2+1$. Define first sets $B_k = \{v_{k+p-3}\} \cup \{v_{k+p}, v_{k+p+1}, \ldots, v_{k+2p-4}\} \cup \{v_{k+2p-2}, v_{k+2p-1}\}$, where k is an integer such that $0 \leq k \leq 2(m-1)p$. Define also a set $B' = \{v_{(2m+1)p-3}\} \cup \{v_{(2m+1)p}, v_{(2m+1)p+1}, \ldots, v_{(2m+1)p+p-2}\}$. Then define a code

$$C_2 = \{v_0\} \cup B' \cup \bigcup_{i=0}^{m-1} B_{2ip}.$$

It is straightforward to verify that C_2 is a transversal of $\mathcal{C}'_{(n,p)}$ and that the set $C_2 \cap Q_p(i)$ is nonempty for any $i \in \mathbb{Z}_n$. Hence, C_2 is an r-identifying code in $\mathcal{C}_n$. Since $|C_2| = (n+1)/2+1$, we have $M_r(\mathcal{C}_n) = (n+1)/2+1$. $\square$

The following theorem provides exact values for $M_r(\mathcal{C}_n)$ when $2r+5 \leq n \leq 3r+1$ and $\gcd(2r+1, n) > 1$. The proof of the theorem is similar to the one of [36, Theorem 9].

Theorem 3.2.4. *Let r be a positive integer and $n = 2r + 1 + p$, where p is an even integer such that $4 \leq p \leq r$. If $\gcd(2r + 1, n) > 1$, then*

$$M_r(\mathcal{C}_n) = \gcd(2r + 1, n) \left\lceil \frac{n}{2\gcd(2r + 1, n)} \right\rceil .$$

Proof. Let $d = \gcd(2r + 1, n) = \gcd(p, n)$ and $n' = n/d$. Notice that n' is odd and $d \geq 3$ since $2 \nmid n$. Recall that $\mathcal{C}'_{(n,2r+1)} = \mathcal{C}'_{(n,p)}$. The graph $\mathcal{C}'_{(n,p)}$ consists of the disjoint union of d cycles on n' vertices. For all $j \in \mathbb{Z}_d$ define the sets

$$T_j = \{v_{j+kp} \mid 0 \leq k \leq n' - 1, k \text{ is even}\}$$

and

$$T'_j = \{v_j\} \cup \{v_{j+kp} \mid 0 \leq k \leq n' - 1, k \text{ is odd}\}.$$

Since n' is odd, we have $|T_j| = |T'_j| = \lceil n'/2 \rceil$. Now define

$$T = T_0 \cup T'_1 \cup \bigcup_{j=2}^{d-1} T_j.$$

Since each T_j and T'_j is a transversal of one of the disjoint subcycles of $\mathcal{C}'_{(n,p)}$, which together form the whole $\mathcal{C}'_{(n,p)}$, the union T of these sets is a transversal of $\mathcal{C}'_{(n,p)}$. Furthermore, the number of vertices in T is equal to $\gcd(2r + 1, n) \lceil n/(2\gcd(2r + 1, n)) \rceil$.

Let us then show that there does not exist $i \in \mathbb{Z}_n$ such that $T \cap Q_p(i) = \emptyset$. Notice that $d \leq p$. Hence, there exists $k \in \mathbb{Z}_{n'}$ such that $\{v_{kp}, v_{kp+1}\} \subseteq Q_p(i)$ or $\{v_{kp+1}, v_{kp+2}\} \subseteq Q_p(i)$. Thus, by the construction of T, we have $T \cap Q_p(i) \neq \emptyset$ for any $i \in \mathbb{Z}_n$. Therefore, by Lemma 3.2.2, T is an r-identifying code in $\mathcal{C}_n$. Thus, the claim follows. $\square$

In conclusion, this completes the work of determining the sizes of optimal r-identifying codes in cycles $\mathcal{C}_n$.

3.3 Optimal identifying codes in paths

In this section, we study r-identifying codes in paths $\mathcal{P}_n = (V_n, E'_n)$. For the following considerations, we define a graph $\mathcal{P}'_{(n,t)} = (V_n, F'_n)$, where t is a positive integer and $F'_n = \{v_i v_{i+t} \mid 0 \leq i \leq n - t - 1\}$. Define also sets $A_1(n) = \{v_{r+1}, v_{r+2}, \ldots, v_{2r}\}$ and $A_2(n) = \{v_{n-2r-1}, v_{n-2r}, \ldots, v_{n-r-2}\}$.

The following lemma characterizes identifying codes in paths.

Lemma 3.3.1. *Let r be a positive integer and $n \geq 2r + 1$. A code $C \subseteq V_n$ is r-identifying in $\mathcal{P}_n$ if and only if the following conditions hold:*

(i) All vertices of V_n are r-covered by a codeword of C.

(ii) *The code C is a transversal of $\mathcal{P}'_{(n,2r+1)}$.*

(iii) *The sets $A_1(n)$ and $A_2(n)$ are subsets of C.*

Proof. Assume first that C is an r-identifying code in $\mathcal{P}_n$. Clearly, each vertex of V_n is r-covered by a codeword of C. For $i = r, r+1, \dots, n-r-2$, the vertices $v_i \in V_n$ and $v_{i+1} \in V_n$ are r-separated by C. Therefore, C is a transversal of $\mathcal{P}'_{(n,2r+1)}$. For $i = 0, 1, \dots, r-1$, we have $B_r(v_i) \triangle B_r(v_{i+1}) = \{v_{i+r+1}\}$. Hence, $A_1(n)$ is a subset of C. Analogously, we have $A_2(n) \subseteq C$.

Assume then that C is a code satisfying the conditions (i), (ii) and (iii). Let u and v be vertices of V_n. In order to prove that C is an r-identifying code in $\mathcal{P}_n$, it is enough to show that the vertices u and v are r-separated by C. Without loss of generality, we may assume that $B_r(u) \cap B_r(v)$ is nonempty and that $u = v_i$ and $v = v_j$ with $i < j$. If $0 \leq i \leq r-1$, then the codeword v_{i+r+1} belongs to $B_r(v_i) \triangle B_r(v_j)$. If $n - r \leq j \leq n-1$, then the codeword v_{j-r-1} belongs to $B_r(v_i) \triangle B_r(v_j)$. Therefore, we may assume that $r \leq i < j \leq n-r-1$. Now the vertices v_{i-r} and v_{i+r+1} belong to $B_r(v_i) \triangle B_r(v_j)$. Since C is a transversal of $\mathcal{P}'_{(n,2r+1)}$, then $v_{i-r} \in C$ or $v_{i+r+1} \in C$. Thus, u and v are r-separated by C. $\square$

For any path $\mathcal{P}_n = (V_n, E'_n)$, define the following subsets of V_n:

$$K_1(\mathcal{P}_n) = \{v_i \mid 0 \leq i \leq n-1, i \text{ is even}\}$$

and

$$K_2(\mathcal{P}_n) = \{v_i \mid 0 \leq i \leq n-1, i \text{ is odd}\}.$$

The following lemma provides a lower bound on the size of a transversal of $\mathcal{P}_n$. The proof of the lemma is trivial.

Lemma 3.3.2. *Let n be a positive integer. If T is a transversal of $\mathcal{P}_n$, then*

$$|T| \geq \left\lfloor \frac{n}{2} \right\rfloor .$$

Moreover, if n is odd, then the unique transversal of $\mathcal{P}_n$ attaining the lower bound is $K_2(\mathcal{P}_n)$.

The following theorem provides exact values for $M_r(\mathcal{P}_n)$ when $n \geq 4r+3$.

Theorem 3.3.3. *Let r be a positive integer and $n = q(2r+1) + p$, where $q \geq 2$ and $1 \leq p \leq 2r+1$. Then we have the following results:*

(i) *Assume that q is even. If $1 \leq p \leq r+1$, then $M_r(\mathcal{P}_n) = q(2r+1)/2 + p$, else $M_r(\mathcal{P}_n) = q(2r+1)/2 + p - 1$.*

(ii) *Assume that q is odd. If $1 \leq p \leq 2r$, then $M_r(\mathcal{P}_n) = (q+1)(2r+1)/2$, else $M_r(\mathcal{P}_n) = (q+1)(2r+1)/2 + 1$.*

Proof. Let C be an r-identifying code in $\mathcal{P}_n$. For a lower bound on $|C|$, we first consider more closely the graph $\mathcal{P}'_{(n,2r+1)}$. Rename the vertices of V_n as follows: $w_k^{(j)} = v_{j+k(2r+1)}$, where j and k are non-negative integers such that $0 \le j \le 2r$ and $0 \le j + k(2r+1) \le n-1$. For $j = 0, 1, \dots, p-1$, define

$$W_j(n) = \{w_k^{(j)} \mid 0 \le k \le q\} \setminus (A_1(n) \cup A_2(n))$$

and, for $j = p, p+1, \dots, 2r$, define

$$W_j(n) = \{w_k^{(j)} \mid 0 \le k \le q-1\} \setminus (A_1(n) \cup A_2(n)).$$

Let j be an integer such that $0 \le j \le 2r$. Define then a graph $\mathcal{S}_j(n) = (W_j(n), H_j(n))$, where the set of edges

$$H_j(n) = \{uv \in F'_n \mid u \in W_j(n), v \in W_j(n)\}.$$

In other words, $\mathcal{S}_j(n)$ is an induced subgraph of $\mathcal{P}'_{(n,2r+1)}$ determined by the vertex set $W_j(n)$. Since only the first or the last vertex of $\{w_k^{(j)} \mid 0 \le k \le q\}$ or $\{w_k^{(j)} \mid 0 \le k \le q-1\}$ can belong to $A_1(n) \cup A_2(n)$, the induced subgraph $\mathcal{S}_j(n)$ is actually a path.

By Lemma 3.3.1, the r-identifying code C is a transversal of $\mathcal{P}'_{(n,2r+1)}$. Therefore, $C \cap W_j(n)$ is a transversal of $\mathcal{S}_j(n)$. Since $\mathcal{S}_j(n)$ is a path, we have that $|C \cap W_j(n)| \ge \lfloor |W_j(n)|/2 \rfloor$ by Lemma 3.3.2. Since the pairwise intersections of the vertex sets $W_j(n)$ are empty, we have

$$|C| \ge |A_1(n)| + |A_2(n)| + \sum_{i=0}^{2r} \left\lfloor \frac{|W_i(n)|}{2} \right\rfloor = 2r + \sum_{i=0}^{2r} \left\lfloor \frac{|W_i(n)|}{2} \right\rfloor . \qquad (3.2)$$

Thus, in order to provide a lower bound for $M_r(\mathcal{P}_n)$, we need to calculate the number of vertices in the sets $W_j(n)$.

Let $n = q(2r+1) + p$, where $q \ge 2$ and $1 \le p \le 2r+1$. Now we have the following two cases to consider.

1) Assume first that $1 \le p \le r+1$. By straightforward calculations, we now have the following results:

(a) For $i = 0, \dots, p-1$, we have $W_i(n) = \{w_0^{(i)}, \dots, w_q^{(i)}\}$ and $|W_i(n)| = q+1$.

(b) For $i = p, \dots, r$, we have $W_i(n) = \{w_0^{(i)}, \dots, w_{q-2}^{(i)}\}$ and $|W_i(n)| = q-1$.

(c) For $i = r+1, \dots, p+r-1$, we have $W_i(n) = \{w_1^{(i)}, \dots, w_{q-2}^{(i)}\}$ and $|W_i(n)| = q-2$.

(d) For $i = p+r, \dots, 2r$, we have $W_i(n) = \{w_1^{(i)}, \dots, w_{q-1}^{(i)}\}$ and $|W_i(n)| = q-1$.

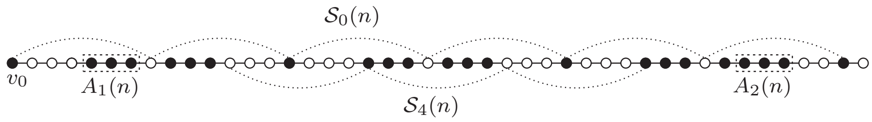


Figure 3.1: The code D_1 illustrated when $r = 3$, $q = 6$, $p = 2$ and $n = 44$. The black dots represent the codewords of D_1.

Notice that the cases (b) and (d) are empty when $p = r + 1$ and the case (c) is empty when $p = 1$. These facts do not affect the calculations of the equation (3.2). Notice also that Lemma 3.3.2 still applies when q is equal to 2 or 3, even though the lengths of the paths $\mathcal{S}_j(n)$ might be equal to 0 or 1.

Assume then that q is even. By the equation (3.2) and the previous calculations, we have $|C| \geq q(2r+1)/2 + p - 1$. Assume that C attains this lower bound. Then the sets $C \cap W_i(n)$ are uniquely determined in the cases (a), (b) and (d), by Lemma 3.3.2. Therefore, it is immediate that the vertex $v_0 \in V_n$ cannot be r-covered by a codeword of C (a contradiction). Hence, $|C| \geq q(2r+1)/2 + p$.

Let us then construct an r-identifying code in $\mathcal{P}_n$ attaining the lower bound. Define

$$D_1 = A_1(n) \cup A_2(n) \cup K_1(\mathcal{S}_0(n)) \cup \bigcup_{i=1}^{2r} K_2(\mathcal{S}_i(n)).$$

The code D_1 is illustrated in Figure 3.1 when $n = 44$ and $r = 3$. Clearly, the code D_1 satisfies the conditions (ii) and (iii) of Lemma 3.3.1. Therefore, it is enough to show that each vertex of V_n is r-covered by a codeword of D_1. By the definitions of $K_1(\mathcal{S}_0(n))$, $K_2(\mathcal{S}_{r+1}(n))$ and $K_2(\mathcal{S}_1(n))$, we know that $k(4r+2) \in D_1$, $k(4r+2)+r+1 \in D_1$ and $k(4r+2)+2r+2 \in D_1$, respectively, when k is an integer such that $1 \leq k \leq q/2 - 1$. Thus, each vertex $v_i \in V_n$ with $3r + 2 \leq i \leq (q-2)(2r+1) + 3r + 2$ is r-covered by a codeword. Since $A_1(n)$ and $A_2(n)$ are subsets of D_1, we also obtain that $v_i \in V_n$ is r-covered by a codeword when $0 \leq i \leq 3r$ or $n - 3r - 1 \leq i \leq n - 1$, respectively. Hence, we have shown that all the vertices of V_n except v_{3r+1} are r-covered by a codeword of D_1. Thus, since v_{3r+1} is r-covered by $v_{2r+2} \in K_2(\mathcal{S}_1(n)) \subseteq D_1$, the condition (i) of Lemma 3.3.1 is satisfied. Hence, D_1 is an r-identifying code in $\mathcal{P}_n$. Moreover, D_1 attains the lower bound. Hence, we have $M_r(\mathcal{P}_n) = q(2r+1)/2 + p$.

Assume now that q is odd. By the equation (3.2) and the results listed above, we have $|C| \geq (q+1)(2r+1)/2$. The code D_1 again satisfies the conditions (ii) and (iii) of Lemma 3.3.1. By considering the set of codewords $K_1(\mathcal{S}_0(n))$, $K_2(\mathcal{S}_1(n))$ and $K_2(\mathcal{S}_{r+1}(n))$ as in the previous case, it can be shown that each vertex of V_n is r-covered by a codeword of D_1. Thus, D_1 is an r-identifying code in $\mathcal{P}_n$ and it attains the obtained lower bound. Hence, we have $M_r(\mathcal{P}_n) = (q+1)(2r+1)/2$.

2) Assume then that $r+2 \leq p \leq 2r+1$. By straightforward calculations, we have the following results:

(a) For $i = 0, \dots, p-r-2$, we have $W_i(n) = \{w_0^{(i)}, \dots, w_{q-1}^{(i)}\}$ and $|W_i(n)| = q$.

(b) For $i = p-r-1, \dots, r$, we have $W_i(n) = \{w_0^{(i)}, \dots, w_q^{(i)}\}$ and $|W_i(n)| = q+1$.

(c) For $i = r+1, \dots, p-1$, we have $W_i(n) = \{w_1^{(i)}, \dots, w_q^{(i)}\}$ and $|W_i(n)| = q$.

(d) For $i = p, \dots, 2r$, we have $W_i(n) = \{w_1^{(i)}, \dots, w_{q-2}^{(i)}\}$ and $|W_i(n)| = q-2$.

The fact that the case (d) is empty when $p = 2r+1$ does not affect the calculation of the equation (3.2).

Assume first that q is even. By the equation (3.2) and the results above, we have $|C| \geq q(2r+1)/2 + p - 1$. Define

$$D_2 = A_1(n) \cup A_2(n) \cup \bigcup_{i=0}^{p-r-2} K_1(\mathcal{S}_i(n)) \cup \bigcup_{i=p-r-1}^{p-1} K_2(\mathcal{S}_i(n)) \cup \bigcup_{i=p}^{2r} K_1(\mathcal{S}_i(n)).$$

Clearly, the conditions (ii) and (iii) of Lemma 3.3.1 are satisfied by D_2. Since $K_1(\mathcal{S}_0(n))$, $K_2(\mathcal{S}_{p-r-1}(n))$ and $K_2(\mathcal{S}_{r+1}(n))$ are subsets of D_2, it can be shown using similar arguments as before that $I_r(D_2; v_i) \neq \emptyset$ for each $i = 0, 1, \dots, n-1$. Thus, D_2 is an r-identifying code and it attains the obtained lower bound. Hence, we have $M_r(\mathcal{P}_n) = q(2r+1)/2 + p - 1$.

Assume then that q is odd. Now we have $|C| \geq (q+1)(2r+1)/2$. Furthermore, assume that $r+2 \leq p \leq 2r$. Define

$$D_3 = A_1(n) \cup A_2(n) \cup \bigcup_{i=0}^{p-r-2} K_2(\mathcal{S}_i(n)) \cup K_1(\mathcal{S}_{p-r-1}(n)) \cup \bigcup_{i=p-r}^{2r} K_2(\mathcal{S}_i(n)).$$

Clearly, the conditions (ii) and (iii) of Lemma 3.3.1 are satisfied by D_3. Since $K_2(\mathcal{S}_0(n))$, $K_1(\mathcal{S}_{p-r-1}(n))$ and $K_2(\mathcal{S}_{p-r}(n))$ are subsets of D_3, it can be shown that $I_r(D_3; v_i) \neq \emptyset$ for each $i = 0, 1, \dots, n-1$. Thus, D_3 is an r-identifying code attaining the lower bound. Therefore, we have $M_r(\mathcal{P}_n) = (q+1)(2r+1)/2$.

Finally, let q be odd and $p = 2r+1$. Assume that the r-identifying code C attains the previously obtained lower bound, i.e. $|C| = (q+1)(2r+1)/2$. Then the sets $C \cap W_i(n)$ are uniquely determined in the cases (a) and (c), by Lemma 3.3.2. Since $p = 2r+1$, the only graph contained in the case (b) is $\mathcal{S}_r(n)$ and the case (d) is empty. Hence, the only case that may contribute a codeword of C to the balls $B_r(v_0)$ and $B_r(v_{n-1})$ is the case (b). Since C attains the lower bound, we have $|C \cap W_r(n)| = |W_r(n)|/2$. Therefore, at least one of the sets $I_r(C; v_0)$ and $I_r(C; v_{n-1})$ is empty. Thus, $|C| \geq (q+1)(2r+1)/2+1$. Define then

$$D_4 = A_1(n) \cup A_2(n) \cup K_1(\mathcal{S}_0(n)) \cup \bigcup_{i=1}^{2r} K_2(\mathcal{S}_i(n)).$$

Clearly, the conditions (ii) and (iii) of Lemma 3.3.1 are satisfied by D_4. Since $K_1(\mathcal{S}_0(n))$, $K_2(\mathcal{S}_1(n))$ and $K_2(\mathcal{S}_{r+1}(n))$ are subsets of D_4, it can be shown that $I_r(D_4; v_i) \neq \emptyset$ for each $i = 0, 1, \ldots, n-1$. Thus, D_4 is an r-identifying code in $\mathcal{P}_n$ and it attains the obtained lower bound. Hence, we have $M_r(\mathcal{P}_n) = (q+1)(2r+1)/2+1$. $\square$

Consider the r-identifying codes in $\mathcal{P}_n$ with $n \leq 4r+2$. Trivially, $M_r(\mathcal{P}_1) = 1$ for any positive integer r. If $2 \leq n \leq 2r$, then there are no r-identifying codes in $\mathcal{P}_n$. The following theorem provides exact values for $M_r(\mathcal{P}_n)$ when $2r+1 \leq n \leq 4r+2$.

Theorem 3.3.4. *Let r be a positive integer. Then we have $M_r(\mathcal{P}_{2r+1}) = 2r$ and $M_r(\mathcal{P}_{4r+2}) = 2r+2$. If $2r+2 \leq n \leq 4r+1$, then $M_r(\mathcal{P}_n) = 2r+1$.*

Proof. Let C be an r-identifying code in $\mathcal{P}_n$. Assume first that $n = 2r+1$. By Lemma 3.3.1, we have $A_1(n) \cup A_2(n) \subseteq C$. Since $A_1(n) \cup A_2(n) = V_n \setminus \{v_r\}$, then $|C| \geq 2r$. Furthermore, it is easy to conclude that the set $A_1(n) \cup A_2(n)$ is actually an r-identifying code in $\mathcal{P}_n$. Therefore, we have $M_r(\mathcal{P}_{2r+1}) = 2r$.

Let then $n = 2r+1+p$, where $1 \leq p \leq r$. Now we have

$$A_1(n) \cup A_2(n) = \{v_p, v_{p+1}, \ldots, v_{2r}\}.$$

Hence, we obtain that $|A_1(n) \cup A_2(n)| = 2r-p+1$. The set of edges of $\mathcal{P}'_{(n,2r+1)}$ is equal to $F'_n = \{v_0v_{2r+1}, v_1v_{2r+2}, \ldots, v_{p-1}v_{2r+p}\}$. Therefore, by Lemmas 3.3.1 and 3.3.2, we have

$$|C| \geq |A_1(n) \cup A_2(n)| + |F'_n| = 2r+1.$$

By Lemma 3.3.1, the code $A_1(n) \cup A_2(n) \cup \{v_0, v_1, \ldots, v_{p-1}\}$ is r-identifying in $\mathcal{P}_n$ attaining the lower bound. Thus, we have $M_r(\mathcal{P}_n) = 2r+1$.

Let now $n = 3r+1+p$, where $1 \leq p \leq r$. We have

$$A_1(n) \cup A_2(n) = \{v_{r+1}, v_{r+2}, \ldots, v_{2r+p-1}\}.$$

Therefore, $|A_1(n) \cup A_2(n)| = r+p-1$. For $i = p-1, p, \ldots, r$, we know that the edges v_iv_{i+2r+1} are such that $v_i \notin A_1(n)$ and $v_{i+2r+1} \notin A_2(n)$. Hence, by similar arguments as before, we have $|C| \geq |A_1(n) \cup A_2(n)| + (r-p+2) = 2r+1$. By Lemma 3.3.1, the set $A_1(n) \cup A_2(n) \cup \{v_p, v_{p+1}, \ldots, v_r\} \cup \{v_{2r+p}\}$ is an r-identifying code in $\mathcal{P}_n$ attaining the obtained lower bound. Thus, we have $M_r(\mathcal{P}_n) = 2r+1$.

Finally, assume that $n = 4r+2$. We have

$$A_1(n) \cup A_2(n) = \{v_{r+1}, v_{r+2}, \ldots, v_{3r}\}.$$

Notice that the sets $B_r(v_0) \cap (A_1(n) \cup A_2(n))$ and $B_r(v_{4r+1}) \cap (A_1(n) \cup A_2(n))$ are empty. Hence, we have $|C| \geq |A_1(n) \cup A_2(n)| + 2 = 2r+2$. On the other hand, the set $\{v_r, v_{3r+1}\} \cup A_1(n) \cup A_2(n)$ is an r-identifying code in $\mathcal{P}_n$ attaining the lower bound. Thus, we have $M_r(\mathcal{P}_{4r+2}) = 2r+2$. □

It is obvious that a cycle $\mathcal{C}_n$ and a path $\mathcal{P}_n$ are closely related to each other. Indeed, the path $\mathcal{P}_n$ only misses the edge $v_{n-1}v_0$. Therefore, a natural question arising is whether there is a link between an optimal r-identifying code in $\mathcal{C}_n$ and $\mathcal{P}_n$. The following theorem concentrates on this question.

Theorem 3.3.5. *Let $n \geq 4r+2$. Then we have $M_r(\mathcal{P}_n) \geq M_r(\mathcal{C}_n) - 1$.*

Proof. Let C be an r-identifying code in a path $\mathcal{P}_n$ of the optimal size $M_r(\mathcal{P}_n)$. Join the ends v_0 and v_{n-1} of the path with an edge forming a cycle $\mathcal{C}_n$. Now consider the code C in the cycle; we obtain that $I_r(x) \neq I_r(y)$ for any $x \neq y$ except $x = v_0$ and $y = v_{n-1}$. Indeed, any two vertices $x, y \in \{v_r, \ldots, v_{n-r-1}\}$ have distinct I-sets; their balls are not affected by the new edge. Any vertex $x = v_i \in \{v_0, \ldots, v_{r-1}\}$ is also distinguished from any $y = v_j$ as long as $i < j$ and $j \neq n-1$ since $I_r(y)$ contains a codeword not belonging to $I_r(x)$. Therefore, by symmetry, $I_r(x) = I_r(y)$ implies that $x = v_0$ and $y = v_{n-1}$. These can be distinguished by adding (if necessary) one more codeword to v_r or v_{n-1-r} giving an r-identifying code of size at most $M_r(\mathcal{P}_n) + 1$ in a cycle. Thus, the assertion follows. □

The bound of the previous theorem can be met (infinitely many times) with equality when n is odd and $\gcd(2r+1, n) = 2r+1$. However, we usually have $M_r(\mathcal{P}_n) > M_r(\mathcal{C}_n) - 1$.

Chapter 4

Location-domination in cycles and paths

In this chapter, which is based on the papers [27] and [28], we consider r-locating-dominating codes in cycles and paths. We begin the chapter by considering r-locating-dominating codes in paths in Section 4.1. As a main result of the section we solve a conjecture stated in [6]; even in a more general form. In Section 4.2, locating-dominating codes in cycles are considered. In this chapter, we focus on location-domination of single vertices in cycles and paths. For the case of sets of vertices, the interested reader is referred to [26].

4.1 Locating-dominating codes in paths

Throughout the section (unless otherwise stated), assume that n is a positive integer. Recall that the size of an optimal r-locating-dominating code in a given finite graph G is denoted by $M_r^{LD}(G)$. Previously, locating-dominating codes in paths have been studied in the papers [6], [41] and [71]. By Slater [71], it is known that $M_1^{LD}(\mathcal{P}_n) = \lceil 2n/5 \rceil$ for any n. In [6], Bertrand *et al.* provided the following lower bound for $r \geq 2$.

Theorem 4.1.1 ([6]). *Let n and r be integers such that $n \geq 1$ and $r \geq 2$. Then we have*

$$M_r^{LD}(\mathcal{P}_n) \geq \left\lceil \frac{n+1}{3} \right\rceil . \tag{4.1}$$

Moreover, in [6], it is conjectured that for any fixed $r \geq 2$, there exist infinitely many values of n such that $M_r^{LD}(\mathcal{P}_n)$ attains the previous lower bound. In [41], it is shown that $M_2^{LD}(\mathcal{P}_n) = \lceil (n+1)/3 \rceil$ for any n. Hence, the conjecture holds when $r = 2$. In Sections 4.1.3 and 4.1.4, we prove that the conjecture also holds when $r \geq 3$. Moreover, we show a stronger result stating that for any $r \geq 3$ we have $M_r^{LD}(\mathcal{P}_n) = \lceil (n+1)/3 \rceil$ for all $n \geq n_r$ when n_r is large enough ($n_r = \mathcal{O}(r^3)$).

In Section 4.1.1, we begin by introducing some basic results concerning r-locating-dominating codes in paths. In Section 4.1.2, we continue by considering r-locating-dominating codes in paths $\mathcal{P}_n$ with small n (for a given r). As a main result concerning locating-dominating codes in paths, we solve the proposed conjecture (in a stronger form) in Section 4.1.3. Finally, in Section 4.1.4, we present optimal 3- and 4-locating-dominating codes in $\mathcal{P}_n$ for all n.

4.1.1 Basics on location-domination in paths

Let C be a nonempty subset of V_n. We first present a useful characterization of r-locating-dominating codes in paths. For this, we need the concept of C-consecutive vertices introduced in [6]. Let i and j be positive integers such that $0 \leq i < j \leq n-1$. We say that (v_i, v_j) is a pair of *C-consecutive vertices* in $\mathcal{P}_n$ if $v_i, v_j \in V_n \setminus C$ and $v_k \in C$ for all integers k such that $i < k < j$. Now we are ready to present the following characterization, which was introduced in [6, Remark 3].

Lemma 4.1.2 ([6]). *Let r be a positive integer. A code $C \subseteq V_n$ is r-locating-dominating in $\mathcal{P}_n$ if and only if each vertex $u \in V_n \setminus C$ is r-covered by a codeword of C and for each pair (u, v) of C-consecutive vertices in $\mathcal{P}_n$ the vertices u and v are r-separated by a codeword of C.*

The following theorem provides a convenient property on the size of the optimal r-locating-dominating codes in $\mathcal{P}_n$.

Theorem 4.1.3. *Let n and r be positive integers. Then we have*

$$M_r^{LD}(\mathcal{P}_n) \leq M_r^{LD}(\mathcal{P}_{n+1}) \leq M_r^{LD}(\mathcal{P}_n) + 1.$$

Proof. Consider first the inequality $M_r^{LD}(\mathcal{P}_n) \leq M_r^{LD}(\mathcal{P}_{n+1})$. Let $C \subseteq V_{n+1} = \{v_0, v_1, \ldots, v_n\}$ be an r-locating-dominating code in $\mathcal{P}_{n+1}$. Assume first that the vertex $v_n \notin C$. Now it is obvious that C is also an r-locating-dominating code in $\mathcal{P}_n$.

Assume then that $v_n \in C$. Denote by X the set of pairs of C-consecutive vertices in $\mathcal{P}_n$. There exists at most one pair $(u, v) \in X$ such that the codeword v_n belongs to the symmetric difference of $I_r(u)$ and $I_r(v)$. If there is no such pair of C-consecutive vertices, then it is clear that $(C \setminus \{v_n\}) \cup \{v_{n-1}\}$ is an r-locating-dominating code in $\mathcal{P}_n$. Assume then that (v_i, v_j) with $i < j$ is the unique pair of C-consecutive vertices such that $v_n \in I_r(v_i) \triangle I_r(v_j)$. Now define $C' = (C \setminus \{v_n\}) \cup \{v_j\}$. Since all the pairs of C-consecutive vertices belonging to $X \setminus \{(v_i, v_j)\}$ are r-separated by a codeword of C', then it is easy to conclude that all the pairs of C'-consecutive vertices are r-separated by a codeword of C' in $\mathcal{P}_n$. Notice that if a vertex is r-covered by v_n, then it is also r-covered by v_j. Therefore, each vertex in V_n is r-covered by a codeword of C'. Thus, by Lemma 4.1.2, C' is an r-locating-dominating code in $\mathcal{P}_n$. In conclusion, we have $M_r^{LD}(\mathcal{P}_n) \leq M_r^{LD}(\mathcal{P}_{n+1})$.

Let then $C \subseteq V_n$ be an r-locating-dominating code in $\mathcal{P}_n$. Since $C \cup \{v_n\}$ is an r-locating-dominating code in $\mathcal{P}_{n+1}$, we immediately obtain that $M_r^{LD}(\mathcal{P}_{n+1}) \leq M_r^{LD}(\mathcal{P}_n) + 1$. □

In what follows, we present a couple of lemmas that are useful in determining the smallest cardinalities of r-locating-dominating codes in paths with a small number of vertices in Section 4.1.2. The first lemma says that if we have an r-locating-dominating code in $\mathcal{P}_n$, then at least r of both the first and the last $2r+1$ vertices of the path are codewords.

Lemma 4.1.4. *Let C be an r-locating-dominating code in $\mathcal{P}_n$ and n be an integer such that $n \geq 2r+1$.*

(i) The intersection $C \cap \{v_0, v_1, \ldots, v_{2r}\}$ contains at least r vertices.

(ii) The intersection $C \cap \{v_{n-2r-1}, v_{n-2r}, \ldots, v_{n-1}\}$ contains at least r vertices.

Proof. Let C be an r-locating-dominating code in $\mathcal{P}_n$. Denote the set $\{v_0, v_1, \ldots, v_r\}$ by $Q_{r+1}(0)$. Assume that there are k codewords in $C \cap Q_{r+1}(0)$ with $0 \leq k \leq r-1$. (Notice that if $k \geq r$, then the case (i) immediately follows.) Now there are $r-k$ pairs (u, v) of C-consecutive vertices such that $u \in Q_{r+1}(0)$ and $v \in Q_{r+1}(0)$. Notice that if (u, v) and (u', v') are such (distinct) pairs of C-consecutive vertices, then the symmetric differences $I_r(u) \triangle I_r(v)$ and $I_r(u') \triangle I_r(v')$ are subsets of $\{v_{r+1}, v_{r+2}, \ldots, v_{2r}\}$ and the intersection of the symmetric differences $I_r(u) \triangle I_r(v)$ and $I_r(u') \triangle I_r(v')$ is empty. Hence, there are at least $r-k$ codewords in $\{v_{r+1}, v_{r+2}, \ldots, v_{2r}\}$. Thus, the claim (i) follows. The case (ii) follows by symmetry. □

The second lemma says that if we have an r-locating-dominating code in $\mathcal{P}_n$, then any set of $3r+1$ consecutive vertices in a path contains at least r codewords.

Lemma 4.1.5. *Let C be an r-locating-dominating code in $\mathcal{P}_n$ and n be an integer such that $n \geq 3r+1$. For $i = 0, 1, \ldots, n-3r-1$, the set*

$$\{v_i, v_{i+1}, \ldots, v_{i+3r}\} \subseteq V_n$$

contains at least r codewords of C.

Proof. Let C be an r-locating-dominating code in $\mathcal{P}_n$ and i be an integer such that $0 \leq i \leq n-3r-1$. Denote $\{v_{i+r}, v_{i+r+1}, \ldots, v_{i+2r}\}$ by

$Q_{r+1}(i+r)$. Assume that there are k codewords in $C \cap Q_{r+1}(i+r)$ with $0 \leq k \leq r-1$. Now there are $r-k$ pairs (u,v) of C-consecutive vertices such that $u \in Q_{r+1}(i+r)$ and $v \in Q_{r+1}(i+r)$. Notice that if (u,v) and (u',v') are such (distinct) pairs of C-consecutive vertices, then it is easy to see that the symmetric differences $I_r(u) \triangle I_r(v)$ and $I_r(u') \triangle I_r(v')$ are subsets of $\{v_i, v_{i+1}, \ldots, v_{i+r-1}\} \cup \{v_{i+2r+1}, v_{i+2r+2}, \ldots, v_{i+3r}\}$ and the intersection of the symmetric differences $I_r(u) \triangle I_r(v)$ and $I_r(u') \triangle I_r(v')$ is empty. Hence, there are at least $r-k$ codewords in $\{v_i, v_{i+1}, \ldots, v_{i+r-1}\} \cup \{v_{i+2r+1}, v_{i+2r+2}, \ldots, v_{i+3r}\}$. Thus, the claim follows. □

4.1.2 Paths with a small number of vertices

In this section, we determine the exact values of $M_r^{LD}(\mathcal{P}_n)$ when $1 \leq n \leq 7r+3$. We also present a new lower bound on $M_r^{LD}(\mathcal{P}_n)$ (improving the previous lower bound of Theorem 4.1.1) for some specific lengths n of the paths.

Consider then the exact values of $M_r^{LD}(\mathcal{P}_n)$ when $1 \leq n \leq 7r+3$. Clearly, we have $M_r^{LD}(\mathcal{P}_1) = 1$. The exact values of $M_r^{LD}(\mathcal{P}_n)$, when $2 \leq n \leq 7r+3$, are given in the following theorem. Previously, in [6], it has been shown that $M_r^{LD}(\mathcal{P}_{3r+1}) = M_r^{LD}(\mathcal{P}_{3r+2}) = r+1$ and $M_r^{LD}(\mathcal{P}_{3r+3}) = r+2$.

Theorem 4.1.6. *Let r be an integer such that $r \geq 2$. Then we have the following results for $2 \leq n \leq 7r+3$:*

1) *If $2 \leq n \leq r+1$, then $M_r^{LD}(\mathcal{P}_n) = n-1$.*
2) *If $r+2 \leq n \leq 2r+1$, then $M_r^{LD}(\mathcal{P}_n) = r$.*
3) *If $2r+2 \leq n \leq 3r+2$, then $M_r^{LD}(\mathcal{P}_n) = r+1$.*
4) *If $n = 3r+3$, then $M_r^{LD}(\mathcal{P}_n) = r+2$.*
5) *If $3r+4 \leq n \leq 4r+2$, then $M_r^{LD}(\mathcal{P}_n) = n-2(r+1)$.*
6) *If $4r+3 \leq n \leq 5r+2$, then $M_r^{LD}(\mathcal{P}_n) = 2r$.*
7) *If $5r+3 \leq n \leq 6r+2$, then $M_r^{LD}(\mathcal{P}_n) = 2r+1$.*
8) *If $6r+3 \leq n \leq 6r+5$, then $M_r^{LD}(\mathcal{P}_n) = 2r+2$.*
9) *If $6r+6 \leq n \leq 7r+3$, then $M_r^{LD}(\mathcal{P}_n) = n-4r-3$.*

Proof. Let C be an r-locating-dominating code in $\mathcal{P}_n$.

1) Assume that $2 \leq n \leq r+1$. Now it is obvious that $B_r(u) = V_n$ for all $u \in V_n$. Hence, it is immediate that $M_r(\mathcal{P}_n) = n-1$.

2) Assume that $r+2 \leq n \leq 2r+1$. Now, by Theorem 4.1.3, we have $M_r^{LD}(\mathcal{P}_n) \geq M_r^{LD}(\mathcal{P}_{r+1}) = r$. On the other hand, using Lemma 4.1.2,

it is easy to verify that $D_2 = \{v_0, v_1, \ldots, v_{r-2}\} \cup \{v_{2r}\}$ is an r-locating-dominating code in $\mathcal{P}_{2r+1}$ with r codewords. Therefore, by Theorem 4.1.3, $M_r^{LD}(\mathcal{P}_n) = r$ when $r + 2 \leq n \leq 2r + 1$.

3) Assume that $2r + 2 \leq n \leq 3r + 2$. Consider first the path $\mathcal{P}_{2r+2}$. It is easy to conclude that each codeword can r-separate at most one pair of C-consecutive vertices in $\mathcal{P}_{2r+2}$. The number of pairs of C-consecutive vertices in $\mathcal{P}_{2r+2}$ is equal to $2r+2-|C|-1$. Therefore, we have the following inequality:

$$|C| \geq 2r + 1 - |C| \iff |C| \geq \frac{2r+1}{2}.$$

Thus, by the previous inequality and Theorem 4.1.3, we have $M_r^{LD}(\mathcal{P}_n) \geq M_r^{LD}(\mathcal{P}_{2r+2}) \geq r+1$. The code $D_3 = \{v_r, v_{r+1}, \ldots, v_{2r-1}\} \cup \{v_{3r}\}$ introduced in [6] is r-locating-dominating in $\mathcal{P}_{3r+2}$. Therefore, $M_r^{LD}(\mathcal{P}_n) = r + 1$ when $2r + 2 \leq n \leq 3r + 2$.

4) In [6], it is shown that $D_4 = \{v_0\} \cup \{v_{r+1}, v_{r+2}, \ldots, v_{2r}\} \cup \{v_{3r+2}\}$ is an r-locating-dominating code in $\mathcal{P}_{3r+3}$. Hence, by Theorem 4.1.1, we have $M_r^{LD}(\mathcal{P}_{3r+3}) = r + 2$.

5) Assume that $3r + 4 \leq n \leq 4r + 2$. Now we can denote $n = 3r + 3 + p$, where $1 \leq p \leq r - 1$. By Lemma 4.1.4, subsets $\{v_0, v_1, \ldots, v_{2r}\}$ and $\{v_{r+p+2}, v_{r+p+3}, \ldots, v_{3r+p+2}\}$ both contain at least r codewords of C. The number of vertices in the intersection of these subsets is equal to $r - p - 1$. Therefore, we have

$$|C| \geq r - p - 1 + 2(r - (r - p - 1)) = r + p + 1.$$

On the other hand, using Lemma 4.1.2, it is straightforward to verify that $D_5 = \{v_1\} \cup \{v_{r+2}, v_{r+3}, \ldots, v_{2r+p}\} \cup \{v_{3r+p+1}\}$ is an r-locating-dominating code in $\mathcal{P}_n$. Thus, $M_r^{LD}(\mathcal{P}_{3r+3+p}) = r + p + 1 = n - 2(r + 1)$ when $3r + 4 \leq n \leq 4r + 2$.

6) Assume that $4r + 3 \leq n \leq 5r + 2$. By Theorem 4.1.3, we have $M_r^{LD}(\mathcal{P}_n) \geq M_r^{LD}(\mathcal{P}_{4r+2}) = 2r$. Then define

$$D_6 = \{v_0\} \cup \{v_{r+2}, v_{r+3}, \ldots, v_{2r}\} \cup \{v_{3r+1}, v_{3r+2}, \ldots, v_{4r-1}\} \cup \{v_{5r+1}\}.$$

The number of vertices in D_6 is equal to $2r$ and, by Lemma 4.1.2, it can be easily verified that D_6 is an r-locating-dominating code in $\mathcal{P}_{5r+2}$. Therefore, by Theorem 4.1.3, $M_r^{LD}(\mathcal{P}_n) = 2r$ when $4r + 3 \leq n \leq 5r + 2$.

7) Assume that $5r+3 \leq n \leq 6r+2$. Let us first show that $M_r^{LD}(\mathcal{P}_{5r+3}) \geq 2r + 1$. Assume to the contrary that C is an r-locating-dominating code in $\mathcal{P}_{5r+3}$ with at most $2r$ codewords. By Lemma 4.1.4, we know that both $\{v_0, v_1, \ldots, v_{2r}\}$ and $\{v_{3r+2}, v_{3r+3}, \ldots, v_{5r+2}\}$ contain at least r codewords of C. Hence, there are no codewords of C in $\{v_{2r+1}, v_{2r+2}, \ldots, v_{3r+1}\}$. Therefore, since all the pairs (u, v) of C-consecutive vertices in $\mathcal{P}_{5r+3}$ such that

$u, v \in \{v_0, v_1, \ldots, v_{2r+1}\}$ are r-separated by a codeword of C, then the codewords of C belonging to $\{v_0, v_1, \ldots, v_{2r+1}\}$ form an r-locating-dominating code in $\mathcal{P}_{2r+2}$ with r codewords. This is a contradiction with the case 3). Thus, by Theorem 4.1.3, $M_r^{LD}(\mathcal{P}_n) \geq M_r^{LD}(\mathcal{P}_{5r+3}) \geq 2r+1$. Define then

$$D_7 = \{v_r, v_{r+1}, \ldots, v_{2r-1}\} \cup \{v_{3r}\} \cup \{v_{4r+2}, v_{4r+3}, \ldots, v_{5r}\} \cup \{v_{5r+2}\}.$$

Using Lemma 4.1.2, it is easy to verify that D_7 is an r-locating-dominating code in $\mathcal{P}_{6r+2}$ with $2r+1$ codewords. Thus, $M_r^{LD}(\mathcal{P}_n) = 2r+1$ when $5r+3 \leq n \leq 6r+2$.

8) Assume that $6r+3 \leq n \leq 6r+5$. By Theorem 4.1.1, we have $M_r^{LD}(\mathcal{P}_n) \geq 2r+2$. Define then

$$\begin{aligned} D_8 = \{v_1, v_{r+1}\} \cup \{v_{r+3}, v_{r+4}, \ldots, v_{2r}\} \cup \{v_{3r+1}, v_{3r+3}\} \\ \cup \{v_{4r+4}, v_{4r+5}, \ldots, v_{5r+1}\} \cup \{v_{5r+3}, v_{6r+3}\}. \end{aligned}$$

By Lemma 4.1.2, D_8 is an r-locating-dominating code in $\mathcal{P}_{6r+5}$ with $2r+2$ vertices. Thus, $M_r^{LD}(\mathcal{P}_n) = 2r+2$ when $6r+3 \leq n \leq 6r+5$.

9) Assume that $6r+6 \leq n \leq 7r+3$. Now we can denote $n = 6r+5+p$, where $1 \leq p \leq r-2$. Consider first the path $\mathcal{P}_{7r+3}$. By Lemma 4.1.4, the subsets $\{v_0, v_1, \ldots, v_{2r}\}$ and $\{v_{5r+2}, v_{5r+3}, \ldots, v_{7r+2}\}$ of V_{7r+3} both contain at least r codewords of C. By Lemma 4.1.5, the same also holds for the subset $\{v_{2r+1}, v_{2r+2}, \ldots, v_{5r+1}\}$. Therefore, $M_r^{LD}(\mathcal{P}_{7r+3}) \geq 3r$. Thus, by Theorem 4.1.3 and the fact that $M_r^{LD}(\mathcal{P}_{6r+5}) = 2r+2$, we have $M_r^{LD}(\mathcal{P}_{6r+5+p}) = 2r+2+p$ when $1 \leq p \leq r-2$. In other words, $M_r^{LD}(\mathcal{P}_n) = n-4r-3$ when $6r+6 \leq n \leq 7r+3$. □

By generalizing the lower bound in the last case of the previous proof, the following theorem is immediately obtained.

Theorem 4.1.7. *Let r be a positive integer and $n = 2(2r+1) + p(3r+1)$ where $p \geq 0$ is an integer. Then we have*

$$M_r^{LD}(\mathcal{P}_n) \geq (p+2)r.$$

Using the notations of the previous theorem, the lower bound of Theorem 4.1.1 implies that

$$M_r^{LD}(\mathcal{P}_n) \geq \left\lceil \frac{n+1}{3} \right\rceil = (p+1)r + 1 + \left\lceil \frac{r+p}{3} \right\rceil.$$

By routine calculations, it can be shown that $(p+2)r > (p+1)r + 1 + \lceil (r+p)/3 \rceil$ if and only if $0 \leq p \leq 2r-6$. Thus, the previous theorem gives improvements on the previously known lower bound when $n = 2(2r+1) + p(3r+1)$ and $0 \leq p \leq 2r-6$.

By applying Theorem 4.1.3 to the previous lower bound, we also obtain new lower bounds for some other values of n. For example, by Theorem 4.1.1, we have $M_5^{LD}(\mathcal{P}_{56}) \geq 19$. However, by Theorem 4.1.7, we have $M_5^{LD}(\mathcal{P}_{54}) \geq 20$ and, therefore, $M_5^{LD}(\mathcal{P}_{56}) \geq M_5^{LD}(\mathcal{P}_{54}) \geq 20$.

The values given by the lower bound of Theorem 4.1.7 are sometimes optimal. For example, when $r = 5$ and $p = 4$, we have $M_5^{LD}(\mathcal{P}_{86}) \geq 30$. On the other hand,

$$\begin{aligned} D_{86} = \{ & v_2, v_6, v_8, v_9, v_{10}, v_{12}, v_{17}, v_{21}, v_{24}, v_{25}, v_{27}, v_{29}, v_{33}, v_{37}, v_{41}, v_{43}, \\ & v_{45}, v_{46}, v_{53}, v_{55}, v_{59}, v_{61}, v_{62}, v_{63}, v_{71}, v_{75}, v_{76}, v_{78}, v_{79}, v_{83}\} \end{aligned}$$

is a 5-locating-dominating code in $\mathcal{P}_{86}$. Therefore, $M_5^{LD}(\mathcal{P}_{86}) = 30$.

4.1.3 Solving a conjecture in long paths

Let r be an integer such that $r \geq 5$. In this section, we show that the size of an optimal r-locating-dominating code in $\mathcal{P}_n$ is equal to $\lceil (n+1)/3 \rceil$ for all $n \geq n_r$ when n_r is large enough ($n_r = \mathcal{O}(r^3)$). The proof of this is based on the result of Theorem 4.1.10 saying that if $n = 3r + 2 + p((r-3)(6r+3) + 3r + 3) + q(6r+3)$, where p and q are non-negative integers, then we have $M_r^{LD}(\mathcal{P}_n) \leq \lceil (n+1)/3 \rceil$. The proof of Theorem 4.1.10 is illustrated in the following example when $r = 5$.

Example 4.1.8. Assume that $r = 5$. Let p and q be non-negative integers. In what follows, we show that if $n = 3r+2+p((r-3)(6r+3)+3r+3)+q(6r+3) = 17 + 84p + 33q$, then $M_5^{LD}(\mathcal{P}_n) \leq \lceil (n+1)/3 \rceil$. In Figures 4.1 and 4.2, first consider the pattern D (the upper dashed box in the figures), which is formed by concatenating the patterns K_1, K_2 and K_3, which are of lengths $6r+3$, $6r+3$ and $3r+3$, respectively. The pattern D is of length $(r-3)(6r+3) + 3r + 3 = 84$ and contains $((r-3)(6r+3) + 3r + 3)/3 = 28$ codewords, i.e. $1/3$ of the vertices of D are codewords. Moreover, it is easy to verify that D is a 5-locating-dominating code in a cycle of length 84 (compare this with Lemma 4.1.9). Similarly, the pattern (the lower dashed box in the figures) formed by K_1 and L_2, which is of length $2(6r+3) = 66$ and contains $(2(6r+3))/3 = 22$ codewords, is a 5-locating-dominating code in a cycle of length 66.

The actual 5-locating-dominating code in $\mathcal{P}_n$ depends on the parity of q. Assume first that q is even, i.e. $q = 2q'$ for some integer q'. The code C_1 is now defined as in Figure 4.1, where the pattern D is repeated p times and the pattern formed by K_1 and L_2 is repeated q' times. Since the patterns D and the one formed by K_1 and L_2 are 5-locating-dominating codes, respectively, in cycles of lengths 84 and 66, it is straightforward to verify that C_1 is a 5-locating-dominating code in $\mathcal{P}_n$ (by Lemma 4.1.2). Similarly, it can be shown that the code C_2 defined in Figure 4.2 is 5-locating-dominating

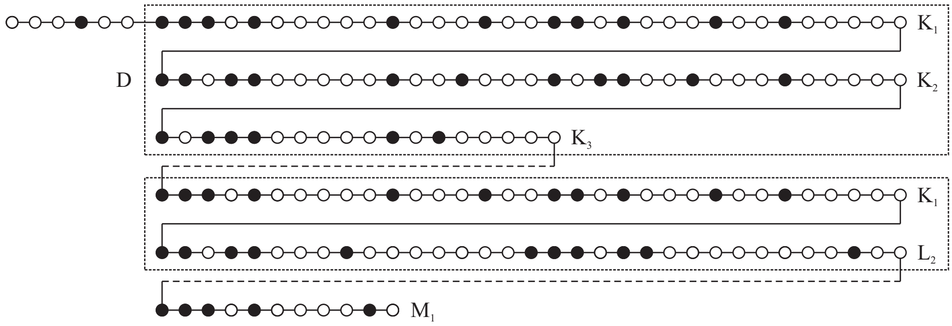


Figure 4.1: The r-locating-dominating code C_1 illustrated when $r = 5$.

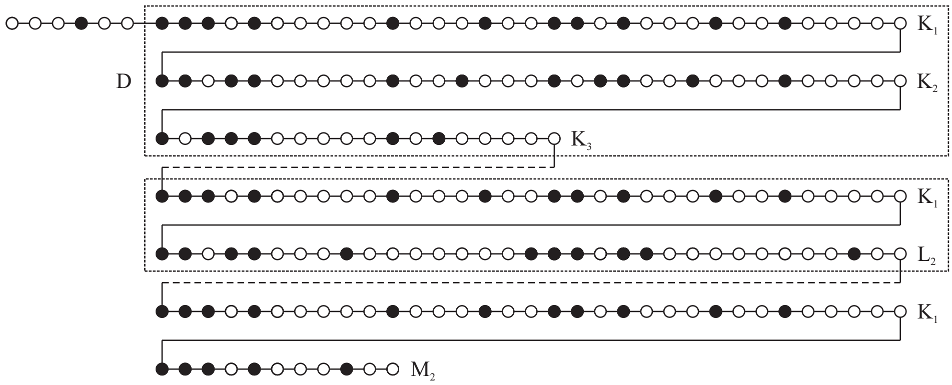


Figure 4.2: The r-locating-dominating code C_2 illustrated when $r = 5$.

in $\mathcal{P}_n$ when q is odd, i.e. $q = 2q' + 1$ for some integer q'. Therefore, if $n = 17 + 84p + 33q$, we have $M_r^{LD}(\mathcal{P}_n) \leq 6 + 28p + 11q = \lceil (n+1)/3 \rceil$.

For the proof of Theorem 4.1.10, we first need to introduce some preliminary definitions and results. Let i and s be non-negative integers. First, for $1 \leq i \leq r-2$, define

$$M_i(s) = \left(\bigcup_{\substack{j=0 \\ j \neq r-i-1}}^{r-1} \{v_{s+j}\} \right) \cup \{v_{s+2r-i}\}$$

and $M_i'(s) = M_i(s) \setminus \{v_{s+2r-i}\}$. Notice that $|M_i(s)| = r$. Furthermore, for

$1 \leq i \leq r-3$, define

$$K_i(s) = M_i'(s) \cup \{v_{s+2r}, v_{s+3r-i}\} \cup \left(\bigcup_{\substack{j=3r+2 \\ j \neq 4r-i}}^{4r} \{v_{s+j}\} \right) \cup \{v_{s+5r-i}, v_{s+5r+2}\},$$

and $K_{r-2}(s) = M_{r-2}'(s) \cup \{v_{s+2r}, v_{s+2r+2}\}$. Notice that for $i = 1, 2, \ldots, r-3$, we have $|K_i(s)| = 2r+1$ and $|K_{r-2}(s)| = r+1$. Finally, define

$$\begin{aligned} L_1(s) = M_1(s) &\cup \left(\bigcup_{j=3r+1}^{4r-1} \{v_{s+j}\} \right) \cup \{v_{s+4r+1}, v_{s+6r+1}\} \\ &\cup \left(\bigcup_{j=6r+3}^{7r+1} \{v_{s+j}\} \right) \cup \{v_{s+8r+3}\} \end{aligned}$$

and, for $2 \leq i \leq r-2$, define

$$L_i(s) = M_i(s) \cup \left(\bigcup_{\substack{j=3r+1 \\ j \neq 4r-i+1}}^{4r+1} \{v_{s+j}\} \right) \cup \{v_{s+6r-i+2}\}.$$

Notice that $|L_1(s)| = 3r+1$ and $|L_i(s)| = 2r+1$ when $2 \leq i \leq r-2$.

As in Example 4.1.8, denote by K_i, L_i and M_i the patterns $\{v_s, v_{s+1}, \ldots, v_{s+\ell-1}\}$ where the codewords are determined by $K_i(s)$, $L_i(s)$ and $M_i(s)$, respectively. The length ℓ of each pattern K_i and L_i is equal to three times the number of codewords in the pattern. For example, the length of the pattern L_1 is equal to $9r+3$ (see the case (iv) below). The length of the pattern M_i is equal to $2r+1$. The following lemma says for general $r \geq 5$ that the patterns K_i, L_i and M_i can be concatenated to form r-locating dominating codes as in Example 4.1.8 (because the beginning of each of them contains $M_i'(s)$).

Lemma 4.1.9. *Let n and s be positive integers, and let r be an integer such that $r \geq 5$. Let C be a code in $\mathcal{P}_n$.*

- *(i) Let i be an integer such that $1 \leq i \leq r-3$. If $K_i(s) \cup M_{i+1}'(s+6r+3) \subseteq C$, then each pair (v_{j_1}, v_{j_2}) of C-consecutive vertices in $\mathcal{P}_n$ such that $s \leq j_1 \leq s+7r+2$ and $s \leq j_2 \leq s+7r+2$ is r-separated by a codeword of C.*
- *(ii) If $K_{r-2}(s) \cup M_1'(s+3r+3) \subseteq C$, then each pair (v_{j_1}, v_{j_2}) of C-consecutive vertices in $\mathcal{P}_n$ such that $s \leq j_1 \leq s+4r+2$ and $s \leq j_2 \leq s+4r+2$ is r-separated by a codeword of C.*

(iii) Let i be an integer such that $2 \leq i \leq r-2$. If $L_i(s) \cup M'_{i-1}(s+6r+3) \subseteq C$, then each pair (v_{j_1}, v_{j_2}) of C-consecutive vertices in $\mathcal{P}_n$ such that $s \leq j_1 \leq s+7r+2$ and $s \leq j_2 \leq s+7r+2$ is r-separated by a codeword of C.

(iv) If $L_1(s) \cup M'_{r-2}(s+9r+3) \subseteq C$, then each pair (v_{j_1}, v_{j_2}) of C-consecutive vertices in $\mathcal{P}_n$ such that $s \leq j_1 \leq s+10r+2$ and $s \leq j_2 \leq s+10r+2$ is r-separated by a codeword of C.

Proof. (i) Let i be an integer with $1 \leq i \leq r-3$ and $C \subseteq V_n$ be a code such that $K_i(s) \cup M'_{i+1}(s+6r+3) \subseteq C$. Consider then the symmetric differences $B_r(v_{j_1}) \triangle B_r(v_{j_2})$, where (v_{j_1}, v_{j_2}) are pairs of C-consecutive vertices such that $s \leq j_1 \leq s+7r+2$ and $s \leq j_2 \leq s+7r+2$. For the following considerations, notice that

$$M'_{i+1}(s+6r+3) = \bigcup_{\substack{j=6r+3 \\ j \neq 7r-i+1}}^{7r+2} \{v_{s+j}\}.$$

Let k be a positive integer. If $s+r \leq k \leq s+2r-i-2$, $s+2r-i \leq k \leq s+2r-2$, $s+4r+2 \leq k \leq s+5r-i-2$ or $s+5r-i+1 \leq k \leq s+5r$, then it is straightforward to verify that the vertex v_{k-r} belongs to the symmetric difference $I_r(v_k) \triangle I_r(v_{k+1})$. If $s+2r+1 \leq k \leq s+3r-i-2$, $s+3r-i+1 \leq k \leq s+3r-1$, $s+5r+3 \leq k \leq s+6r-i-1$ or $s+6r-i+1 \leq k \leq s+6r+1$, then it can be seen that the vertex v_{k+r+1} belongs to the symmetric difference $I_r(v_k) \triangle I_r(v_{k+1})$. Moreover, we have that

$$\begin{aligned}
v_{s+2r} &\in I_r(v_{s+r-i-1}) \triangle I_r(v_{s+r}), \\
v_{s+3r-i} &\in I_r(v_{s+2r-i-1}) \triangle I_r(v_{s+2r-i}), \\
v_{s+r-1} &\in I_r(v_{s+2r-1}) \triangle I_r(v_{s+2r+1}), \\
v_{s+4r-i+1} &\in I_r(v_{s+3r-i-1}) \triangle I_r(v_{s+3r-i+1}), \\
v_{s+2r} &\in I_r(v_{s+3r}) \triangle I_r(v_{s+3r+1}), \\
v_{s+5r-i} &\in I_r(v_{s+3r+1}) \triangle I_r(v_{s+4r-i}), \\
v_{s+3r-i} &\in I_r(v_{s+4r-i}) \triangle I_r(v_{s+4r+1}), \\
v_{s+5r+2} &\in I_r(v_{s+4r+1}) \triangle I_r(v_{s+4r+2}), \\
v_{s+4r-i-1} &\in I_r(v_{s+5r-i-1}) \triangle I_r(v_{s+5r-i+1}), \\
v_{s+6r+3} &\in I_r(v_{s+5r+1}) \triangle I_r(v_{s+5r+3}), \\
v_{s+5r-i} &\in I_r(v_{s+6r-i}) \triangle I_r(v_{s+6r-i+1}) \text{ and} \\
v_{s+5r+2} &\in I_r(v_{s+6r+2}) \triangle I_r(v_{s+7r-i+1}).
\end{aligned}$$

In conclusion, all the pairs (v_{j_1}, v_{j_2}) of C-consecutive vertices in $\mathcal{P}_n$ such that $s \leq j_1 \leq s + 7r + 2$ and $s \leq j_2 \leq s + 7r + 2$ are r-separated by a codeword of C.

The proofs of the cases (ii), (iii) and (iv) are analogous to the first one. $\square$

For a non-negative integer s, define

$$C(s) = \bigcup_{i=0}^{r-3} K_{i+1}(s + i(6r + 3)).$$

Notice that when $r = 5$, $C(s)$ corresponds to the pattern D in Example 4.1.8. The following theorem now proves the conjecture stated in [6, Conjecture 1] when $r \geq 5$.

Theorem 4.1.10. *Let r be an integer such that $r \geq 5$ and $n = 3r + 2 + p((r-3)(6r+3)+3r+3)+q(6r+3)$, where p and q are non-negative integers. Then we have*

$$M_r^{LD}(\mathcal{P}_n) \leq \left\lceil \frac{n+1}{3} \right\rceil .$$

Proof. Let $r \geq 5$ be an integer and $n = 3r + 2 + p((r - 3)(6r + 3) + 3r + 3) + q(6r + 3)$, where p and q are non-negative integers. Assume that q is even, i.e. $q = 2q'$ for some integer q'. Define then

$$\begin{aligned} C_1 = \{v_{r-2}\} \cup \bigcup_{j=0}^{p-1} & C(r + 1 + j((r - 3)(6r + 3) + 3r + 3)) \\ \cup \bigcup_{j=0}^{q'-1} & K_1(r + 1 + p((r - 3)(6r + 3) + 3r + 3) + 2j(6r + 3)) \\ \cup \bigcup_{j=0}^{q'-1} & L_2(r + 1 + p((r - 3)(6r + 3) + 3r + 3) + (2j + 1)(6r + 3)) \\ \cup & M_1(r + 1 + p((r - 3)(6r + 3) + 3r + 3) + q(6r + 3)). \end{aligned}$$

Notice that if $r = 5$, this definition of C_1 coincides with the one of Example 4.1.8. (Recall also the length of the patterns K_i, L_i and M_i as described earlier.) As in the previous example, C_1 is formed by concatenating the patterns K_i, L_i and M_i. Since $M'_i(s) \subseteq K_i(s)$ and $M'_i(s) \subseteq L_i(s)$, Lemma 4.1.9 applies to each occurrence of $K_i(s)$ and $L_i(s)$ in C_1. Therefore, each pair (v_j, v_k) of C_1-consecutive vertices in $\mathcal{P}_n$ such that $r + 1 \leq j \leq n - r - 2$ and $r + 1 \leq k \leq n - r - 2$ is r-separated by a codeword of C_1. Hence, it is easy to see that each pair of C_1-consecutive vertices in $\mathcal{P}_n$ is r-separated by C_1. Since there are no $2r + 1$ consecutive vertices belonging to $V_n \setminus C_1$

in $\mathcal{P}_n$, all the vertices in $\mathcal{P}_n$ are r-covered by a codeword of C_1. Thus, by Lemma 4.1.2, it is easy to conclude that C_1 is an r-locating-dominating code in $\mathcal{P}_n$ with $\lceil (n+1)/3 \rceil$ vertices.

Assume then that q is odd, i.e. $q = 2q' + 1$ for some integer q'. Define then

$$\begin{aligned}
C_2 = \{v_{r-2}\} &\cup \bigcup_{j=0}^{p-1} C(r+1+j((r-3)(6r+3)+3r+3)) \\
&\cup \bigcup_{j=0}^{q'} K_1(r+1+p((r-3)(6r+3)+3r+3)+2j(6r+3)) \\
&\cup \bigcup_{j=0}^{q'-1} L_2(r+1+p((r-3)(6r+3)+3r+3)+(2j+1)(6r+3)) \\
&\cup M_2(r+1+p((r-3)(6r+3)+3r+3)+q(6r+3)).
\end{aligned}$$

Similarly, as in the previous case, it can be shown that C_2 is an r-locating-dominating code in $\mathcal{P}_n$ with $\lceil (n+1)/3 \rceil$ vertices. □

In [53, Theorem 8.3], the following theorem is presented. This theorem turns out useful in future considerations.

Theorem 4.1.11 ([53]). *Let a and b be positive integers such that the greatest common divisor of a and b is equal to 1. Then, for any integer $n > ab - a - b$, there exist such non-negative integers p and q that $n = pa + qb$.*

The length of the path in Theorem 4.1.10 can be written as follows:

$$\begin{aligned}
n &= 3r+2+p((r-3)(6r+3)+3r+3)+q(6r+3) \\
&= 3r+2+3(p((r-3)(2r+1)+r+1)+q(2r+1)).
\end{aligned}$$

The greatest common divisor of $(r-3)(2r+1)+r+1$ and $2r+1$ is equal to 1. Thus, by Theorem 4.1.11, if n' is an integer such that $n' \geq 2r((r-3)(2r+1)+r)$, then there exist non-negative integers p and q such that $n' = p((r-3)(2r+1)+r+1)+q(2r+1)$. Therefore, if n is an integer such that $n \geq 3r+2+3 \cdot 2r((r-3)(2r+1)+r)$ and $n \equiv 2 \pmod 3$, then there exist integers $p \geq 0$ and $q \geq 0$ such that $n = 3r+2+p((r-3)(6r+3)+3r+3)+q(6r+3)$.

Assume that $n \geq 3r+2+6r((r-3)(2r+1)+r)$ and $n = 3k+2$, where k is an integer. Combining the lower bound of Theorem 4.1.1, Theorem 4.1.3 and Theorem 4.1.10 with the previous observation, we obtain that

$$k+1 \leq M_r^{LD}(\mathcal{P}_{3k}) \leq M_r^{LD}(\mathcal{P}_{3k+1}) \leq M_r^{LD}(\mathcal{P}_{3k+2}) \leq k+1.$$

Therefore, $M_r^{LD}(\mathcal{P}_{3k}) = M_r^{LD}(\mathcal{P}_{3k+1}) = M_r^{LD}(\mathcal{P}_{3k+2}) = k+1$. Thus, the following theorem immediately follows.

Theorem 4.1.12. *Let r be a positive integer such that $r \geq 5$. If $n \geq 3r+2+6r((r-3)(2r+1)+r)$, then we have*

$$M_r^{LD}(\mathcal{P}_n) = \left\lceil \frac{n+1}{3} \right\rceil.$$

Theorem 4.1.10 provides one approach to form r-locating-dominating codes in paths using Lemma 4.1.9. However, this lemma can also be applied in other ways. For example, when k is an integer such that $0 \leq k \leq r-3$,

$$\begin{aligned} D(k) = \{v_{r-2}\} \cup L_1(r+1) \cup \left(\bigcup_{j=0}^{k-1} L_{r-2-j}(10r+4+j(6r+3)) \right) \\ \cup M_{r-2-k}(10r+4+k(6r+3)) \end{aligned}$$

is an optimal r-locating-dominating code in $\mathcal{P}_n$ with $n = 12r+5+k(6r+3)$. Notice that the optimal r-locating-dominating codes in paths of these lengths cannot be obtained using Theorem 4.1.10.

4.1.4 The exact values of $M_3^{LD}(\mathcal{P}_n)$ and $M_4^{LD}(\mathcal{P}_n)$

In this section, we solve the exact values of $M_3^{LD}(\mathcal{P}_n)$ and $M_4^{LD}(\mathcal{P}_n)$ for all n. In order to do this, we first need to present some preliminary definitions and results.

Define an infinite path $\mathcal{P}_\infty = (V_\infty, E_\infty)$, where $V_\infty = \{v_i \mid i \in \mathbb{Z}\}$ and $E_\infty = \{v_i v_{i+1} \mid i \in \mathbb{Z}\}$. Define then

$$C = \{v_i \in V_\infty \mid i \equiv 0, 2 \bmod 6\}.$$

In [41], it is stated that if r is an integer such that $r \geq 2$ and $r \equiv 1, 2, 3$ or $4 \pmod 6$, then C is an r-locating-dominating code in $\mathcal{P}_\infty$. This result is rephrased in the following lemma when $r = 3$ and $r = 4$.

Lemma 4.1.13. *Let n and k be integers such that*

$$D = \{v_k, v_{k+2}, v_{k+6}, v_{k+8}, v_{k+12}, v_{k+14}\} \subseteq V_n.$$

If a pair (v_i, v_j) of D-consecutive vertices in $\mathcal{P}_n$ is such that $k+5 \leq i \leq k+13$ and $k+5 \leq j \leq k+13$, then v_i and v_j are 3- *and* 4-*separated by a codeword of D. Moreover, each vertex $v_i \in V_n \setminus D$ such that $k+6 \leq i \leq k+11$ is* 3- *and* 4-*covered by a codeword of D.*

Consider then r-locating-dominating codes in $\mathcal{P}_n$ when $r = 3$. By Theorem 4.1.6, the exact values of $M_3^{LD}(\mathcal{P}_n)$ are known when $1 \leq n \leq 24$. Let p be an integer such that $p \geq 1$. Define

$$D_1(p) = \{v_1\} \cup \left(\bigcup_{i=0}^{p} \{v_{4+6i}, v_{6+6i}\} \right) \cup \{v_{9+6p}, v_{14+6p}, v_{15+6p}, v_{17+6p}\}$$

and

$$D_2(p) = \{v_1\} \cup \left(\bigcup_{i=0}^{p} \{v_{4+6i}, v_{6+6i}\} \right) \cup \{v_{10+6p}, v_{12+6p}, v_{16+6p}, v_{18+6p}, v_{21+6p}\}.$$

It is easy to verify that $D_1(1)$ and $D_2(1)$ are 3-locating-dominating codes in $\mathcal{P}_{26}$ and $\mathcal{P}_{29}$, respectively. Therefore, using Lemma 4.1.13, it is easy to conclude that $D_1(p)$ and $D_2(p)$ are 3-locating-dominating codes in $\mathcal{P}_{20+6p}$ and $\mathcal{P}_{23+6p}$, respectively, when $p \geq 2$. Moreover, by Theorem 4.1.1 and Theorem 4.1.3, we have

$$|D_1(p)| \geq M_3^{LD}(\mathcal{P}_{20+6p}) \geq M_3^{LD}(\mathcal{P}_{19+6p}) \geq M_3^{LD}(\mathcal{P}_{18+6p}) \geq 7 + 2p$$

and

$$|D_2(p)| \geq M_3^{LD}(\mathcal{P}_{23+6p}) \geq M_3^{LD}(\mathcal{P}_{22+6p}) \geq M_3^{LD}(\mathcal{P}_{21+6p}) \geq 8 + 2p.$$

Since $|D_1(p)| = 7+2p$ and $|D_2(p)| = 8+2p$, we have $M_3^{LD}(\mathcal{P}_n) = \lceil (n+1)/3 \rceil$ for any $n \geq 24$. In conclusion, all the values of $M_3^{LD}(\mathcal{P}_n)$ are determined.

Consider then r-locating-dominating codes in $\mathcal{P}_n$ when $r = 4$. By Theorem 4.1.6, the exact values of $M_4^{LD}(\mathcal{P}_n)$ are known when $1 \leq n \leq 31$. Assume now that $p \geq 0$. Define

$$\begin{aligned} D_3(p) = \{v_1, v_5, v_7, v_8\} \cup \left(\bigcup_{i=0}^{p} \{v_{13+6i}, v_{15+6i}\} \right) \\ \cup \{v_{20+6p}, v_{21+6p}, v_{23+6p}, v_{27+6p}\} \end{aligned}$$

and

$$\begin{aligned} D_4(p) = (D_3(p) \setminus \{v_{27+6p}\}) \cup \{v_{28+6p}, v_{31+6p}, v_{34+6p}, v_{36+6p}\} \\ \cup \{v_{39+6p}, v_{42+6p}, v_{47+6p}, v_{49+6p}, v_{50+6p}, v_{53+6p}\}. \end{aligned}$$

It is straightforward to verify that $D_3(0)$, $D_3(1)$, $D_4(0)$ and $D_4(1)$ are 4-locating-dominating codes in $\mathcal{P}_{29}$, $\mathcal{P}_{35}$, $\mathcal{P}_{56}$ and $\mathcal{P}_{62}$, respectively. Therefore, using Lemma 4.1.13, it is easy to conclude that $D_1(p)$ and $D_2(p)$ are 4-locating-dominating codes in $\mathcal{P}_{29+6p}$ and $\mathcal{P}_{56+6p}$, respectively, when $p \geq 2$. Moreover, by Theorem 4.1.1 and Theorem 4.1.3, we have

$$|D_3(p)| \geq M_4^{LD}(\mathcal{P}_{29+6p}) \geq M_4^{LD}(\mathcal{P}_{28+6p}) \geq M_4^{LD}(\mathcal{P}_{27+6p}) \geq 10 + 2p$$

and

$$|D_4(p)| \geq M_4^{LD}(\mathcal{P}_{56+6p}) \geq M_4^{LD}(\mathcal{P}_{55+6p}) \geq M_4^{LD}(\mathcal{P}_{54+6p}) \geq 19 + 2p.$$

Since $|D_3(p)| = 10 + 2p$ and $|D_4(p)| = 19 + 2p$, we obtain that $M_4^{LD}(\mathcal{P}_n) = \lceil (n+1)/3 \rceil$ when $27 + 6p \leq n \leq 29 + 6p$ and $54 + 6p \leq n \leq 56 + 6p$ $(p \geq 0)$.

In conclusion, the values of $M_4^{LD}(\mathcal{P}_n)$ are determined except when $n = 32$, $36 \leq n \leq 38$, $42 \leq n \leq 44$ or $48 \leq n \leq 50$.

By Theorem 4.1.6, we have $M_4^{LD}(\mathcal{P}_{31}) = 12$. Therefore, by Theorem 4.1.3, since $M_4^{LD}(\mathcal{P}_{35}) = 12$, we obtain that $M_4^{LD}(\mathcal{P}_{32}) = 12$. Define then

$$D_{37} = \{v_2, v_3, v_5, v_6, v_{13}, v_{16}, v_{17}, v_{19}, v_{23}, v_{29}, v_{30}, v_{31}, v_{33}\},$$

$$D_{43} = \{v_2, v_3, v_5, v_8, v_{10}, v_{16}, v_{18}, v_{21}, v_{23}, v_{24}, v_{31}, v_{34}, v_{35}, v_{37}, v_{41}\}$$

and

$$D_{49} = \{v_2, v_5, v_6, v_8, v_{13}, v_{16}, v_{19}, v_{20}, v_{26}, v_{27}, v_{30}, v_{33}, v_{38}, v_{40}, v_{41}, v_{42}, v_{48}\}.$$

It is easy to verify that D_{37}, D_{43} and D_{49} are 4-locating-dominating codes, respectively, in $\mathcal{P}_{37}$, $\mathcal{P}_{43}$ and $\mathcal{P}_{49}$ attaining the lower bound of Theorem 4.1.1. Therefore, by Theorem 4.1.3, we also have the optimal 4-locating-dominating codes for the paths $\mathcal{P}_{36}$, $\mathcal{P}_{42}$ and $\mathcal{P}_{48}$. By Theorem 4.1.7, we obtain that $M_4^{LD}(\mathcal{P}_{44}) \geq 16$. On the other hand, we have $M_r^{LD}(\mathcal{P}_{44}) \leq M_r^{LD}(\mathcal{P}_{45}) = 16$. Hence, $M_4^{LD}(\mathcal{P}_{44}) = 16$.

Now the only open values are $M_4^{LD}(\mathcal{P}_{38})$ and $M_4^{LD}(\mathcal{P}_{50})$. By the previous constructions, we know that $M_4^{LD}(\mathcal{P}_{38}) \leq M_4^{LD}(\mathcal{P}_{39}) = 14$ and $M_4^{LD}(\mathcal{P}_{50}) \leq M_4^{LD}(\mathcal{P}_{51}) = 18$. By an exhaustive computer search, we have been able to prove that there are no 4-locating-dominating codes in $\mathcal{P}_{38}$ and $\mathcal{P}_{50}$ with 13 and 17 codewords, respectively. Hence, $M_4^{LD}(\mathcal{P}_{38}) = 14$ and $M_4^{LD}(\mathcal{P}_{50}) = 18$. In conclusion, all the values of $M_4^{LD}(\mathcal{P}_n)$ are determined.

4.2 Locating-dominating codes in cycles

For the rest of the section (unless otherwise stated), let n be a positive integer such that $n \geq 3$. Previously, locating-dominating codes in cycles have been studied in the papers [6], [20] and [72]. By Slater [72], we know that $M_1^{LD}(\mathcal{C}_n) = \lceil 2n/5 \rceil$ for any $n \geq 3$. For radius $r \geq 2$, Bertrand *et al.* [6] (see Theorem 4.2.3) provide the lower bound

$$M_r^{LD}(\mathcal{C}_n) \geq \left\lceil \frac{n}{3} \right\rceil. \tag{4.2}$$

In [6, Theorem 14], it is also shown that for each $r \geq 2$ there exists an infinite family of n such that $M_r^{LD}(\mathcal{C}_n) = \lceil n/3 \rceil$. In particular, it is shown that if r is even, $n > 6$ and $n \equiv 0 \pmod{3r}$, or if r is odd and $n \equiv 0 \pmod{3r+3}$, then the lower bound is attained.

The exact values of $M_r^{LD}(\mathcal{C}_n)$ are determined in [20] for $r = 2$. In particular, it is shown that for $n > 6$ we have

$$M_2^{LD}(\mathcal{C}_n) = \begin{cases} n/3 + 1 & \text{if } n \equiv 3 \pmod 6 \\ \lceil n/3 \rceil & \text{if } n \not\equiv 3 \pmod 6. \end{cases}$$

In Section 4.2.4, we determine the exact values of $M_3^{LD}(\mathcal{C}_n)$ and $M_4^{LD}(\mathcal{C}_n)$. For the summary of the results in these cases, we refer to Theorem 4.2.19. In Section 4.2.3, we prove that for any $r \geq 5$ and $n \geq n_r$ when n_r is large enough ($n_r = \mathcal{O}(r^3)$) we have constructions according to which $M_r^{LD}(\mathcal{C}_n) \leq n/3+1$ if $n \equiv 3 \pmod 6$ and $M_r^{LD}(\mathcal{C}_n) \leq \lceil n/3 \rceil$ otherwise. The latter constructions are optimal by the lower bound (4.2). Using the evidence provided in Sections 4.2.2 and 4.2.4, we conjecture that also the constructions in the case $n \equiv 3 \pmod 6$ are optimal.

In what follows, we begin by introducing some basic results concerning r-locating-dominating codes in cycles in Section 4.2.1. Then, in Section 4.2.2, we proceed by considering r-locating-dominating codes in cycles $\mathcal{C}_n$ with small n (for a given r). In Section 4.2.3, we present constructions for r-locating-dominating codes in cycles for general r and, in Section 4.2.4, we consider r-locating-dominating codes in cycles when $2 \leq r \leq 4$.

4.2.1 Basics on location-domination in cycles

We first present some useful observations concerning r-locating-dominating codes in cycles. Recall the concept of C-consecutive vertices defined in the case of paths in Section 4.1.1. In the case of cycles, we again say that two vertices form a pair of C-consecutive vertices if all the vertices between them are codewords. Formally, assuming i and j are non-negative integers, we say that (v_i, v_j) is a pair of *C-consecutive vertices* in $\mathcal{C}_n$ if $v_i, v_j \in V_n \setminus C$ and $v_k \in C$ for all $k = i+1, i+2, \ldots, j-1$ or for all $k = j+1, j+2, \ldots, i-1$. Recall that the indices of the vertices in $\mathcal{C}_n$ are calculated modulo n. The following lemma is previously presented in [6, Remark 4].

Lemma 4.2.1 ([6])**.** *If $C \subseteq V_n$ is a code in $\mathcal{C}_n$, then each codeword of C can r-separate at most two pairs of C-consecutive vertices.*

In Lemma 4.1.2, a useful characterization of r-locating-dominating codes in paths is presented. The following lemma provides similar characterization in the case of cycles.

Lemma 4.2.2. *A code $C \subseteq V_n$ is r-locating-dominating in $\mathcal{C}_n$ if and only if*

(i) each vertex $u \in V_n \setminus C$ is r-covered by a codeword of C,

(ii) each pair (u, v) of C-consecutive vertices in $\mathcal{C}_n$ is r-separated by C and

(iii) there exists at most one vertex $u \in V_n \setminus C$ such that $I_r(u) = C$.

Proof. If C is an r-locating-dominating code in $\mathcal{C}_n$, then the conditions (i), (ii) and (iii) immediately follow. Assume then that $C \subseteq V_n$ is a code satisfying these three conditions. By the assumption, all the vertices of V_n are r-covered by a codeword of C. Let then u and v be two distinct vertices of V_n. If $I_r(u) = C$, then by the condition (iii), the vertices u and v are r-separated by a codeword.

Hence, we may assume that $I_r(u) \neq C$ and $I_r(v) \neq C$. If the intersection of $I_r(v)$ and $C \setminus I_r(u)$ is nonempty, then the vertices u and v are r-separated by a codeword of C. Otherwise, we have $I_r(v) \subseteq I_r(u)$. Then there exists a non-codeword $w \in V_n$ such that (u, w) is a pair of C-consecutive vertices and the symmetric difference $I_r(u) \triangle I_r(w)$ is a subset of $I_r(u) \triangle I_r(v)$. (Notice that if (u, v) is pair of C-consecutive vertices, then $v = w$.) Therefore, by the condition (ii), we have $I_r(u) \neq I_r(v)$. □

In the previous characterization, the condition (iii) is necessary. Indeed, consider a code $\{v_0, v_2\}$ in $\mathcal{C}_6$ when $r = 2$. Now the conditions (i) and (ii) clearly hold. However, the code is not 2-locating-dominating in $\mathcal{C}_6$ since $I_r(v_1) = I_r(v_4) = \{v_0, v_2\}$. Notice also that if $n \geq 4r + 2$ and the condition (i) holds, then there is no vertex $u \in V_n \setminus C$ such that $I_r(u) = C$.

The following lower bound is presented in [6, Theorem 13]. In order to prove Lemma 4.2.4, we include the proof of the lower bound here.

Theorem 4.2.3 ([6])**.** *For all integers $n \geq 3$ and $r \geq 2$, we have*

$$M_r^{LD}(\mathcal{C}_n) \geq \left\lceil \frac{n}{3} \right\rceil .$$

Proof. Let C be an r-locating-dominating code in $\mathcal{C}_n$. By Lemma 4.2.1, each codeword of C can r-separate at most two pairs of C-consecutive vertices. On the other hand, by Lemma 4.2.2, each pair of C-consecutive vertices has to be r-separated by at least one codeword. Hence, we have $2|C| \geq n - |C|$. This implies the assertion. □

The next lemma immediately follows from the previous proof.

Lemma 4.2.4. *Let n be divisible by three and $r \geq 2$. If C is an r-locating-dominating code in $\mathcal{C}_n$ with $n/3$ codewords, then*

(i) each codeword r-separates exactly two pairs of C-consecutive vertices and

(ii) each pair of C-consecutive vertices is r-separated by exactly one codeword of C.

For future considerations, we introduce the concept of C-block of codewords. Let t be a positive integer. As in Section 3.2, we denote $Q_t(i) = \{v_i, v_{i+1}, \ldots, v_{i+t-1}\}$ ($i \in \mathbb{Z}_n$). Let then $C \subseteq V_n$ be a code. We say that $Q_t(i)$ is a *C-block (of codewords)* if the vertices $v_i, v_{i+1}, \ldots, v_{i+t-1} \in C$ and $v_{i-1}, v_{i+t} \notin C$. Moreover, if $Q_t(i)$ is a C-block of codewords, then the *length* of the C-block is t. Notice that if $Q_t(i)$ is a C-block, then (v_{i-1}, v_{i+t}) is a pair of C-consecutive vertices. Notice also that if $v_{i-1}, v_{i+1} \notin C$ and $v_i \in C$, then we say that $\{v_i\}$ is a C-block of length one.

Now we are ready to present the following two lemmas.

Lemma 4.2.5. *Let n be divisible by three and $r \geq 2$. If C is an r-locating-dominating code in $\mathcal{C}_n$ with $n/3$ codewords, then the length of any C-block of codewords is at most $r-1$.*

Proof. Let C be an r-locating-dominating code in $\mathcal{C}_n$ with $n/3$ codewords. Assume that there exists a C-block $Q_t(i)$ of length $t \geq r+1$. Then it is immediately clear that v_i (and v_{i+t-1}) r-separate at most one pair of C-consecutive vertices. This is a contradiction with Lemma 4.2.4 (i).

Assume then that $Q_r(i)$ is a C-block of length r. Since (v_{i-1}, v_{i+r}) is a pair of C-consecutive vertices, the symmetric difference $I_r(v_{i-1}) \triangle I_r(v_{i+r})$ contains exactly one codeword of C by Lemma 4.2.4 (ii). Therefore, without loss of generality, we may assume that $I_r(v_{i+r}) \setminus I_r(v_{i-1}) = \emptyset$. Since the pairs (v_j, v_{j+1}) of C-consecutive vertices, where $j = i+r, i+r+1, \ldots, i+2r-1$, are r-separated by exactly one codeword of C and $v_{j-r} \in I_r(v_j) \setminus I_r(v_{j+1})$, the vertices $v_{i+2r+1}, v_{i+2r+2}, \ldots, v_{i+3r} \notin C$. Hence, the set $I_r(v_{i+2r})$ is empty (a contradiction). Thus, the claim follows. □

Lemma 4.2.6. *Let n be divisible by three and $r \geq 2$. If C is an r-locating-dominating code in $\mathcal{C}_n$ with $n/3$ codewords, then the number of C-blocks of codewords is even.*

Proof. Let C be an r-locating-dominating code in $\mathcal{C}_n$ with $n/3$ codewords. Assume that $Q_t(i)$ is a C-block (for appropriate integers i and t). Hence, (v_{i-1}, v_{i+t}) is a pair of C-consecutive vertices. This pair is r-separated by a unique codeword. Assume that this codeword belongs to the C-block $Q_{t'}(i')$ (for some appropriate integers i' and t'). Now the pair $(v_{i'-1}, v_{i'+t'})$ of C-consecutive vertices is clearly r-separated by a unique codeword that belongs to the C-block $Q_t(i)$. Therefore, each C-block can be uniquely paired to another C-block. Thus, the number of C-blocks is even. □

4.2.2 Cycles with a small number of vertices

In this section, we consider r-locating-dominating codes in $\mathcal{C}_n$ with small n (for a given r). The following easy theorem gives the exact values of $M_r^{LD}(\mathcal{C}_n)$ when $3 \leq n \leq 2r+1$.

Theorem 4.2.7. *Let n and r be positive integers such that $3 \leq n \leq 2r+1$ and $r \geq 2$. Then we have*

$$M_r^{LD}(\mathcal{C}_n) = n - 1.$$

Proof. Let C be an r-identifying code in $\mathcal{C}_n$. Assume that $|C| \leq n-2$. Then there exist $u, v \in V_n \setminus C$ such that $u \neq v$. Since $B_r(u) = B_r(v) = V_n$, we have $I_r(u) = C$ and $I_r(v) = C$ (a contradiction). Therefore, we have $|C| \geq n-1$. On the other hand, $\{v_0, v_1, \ldots, v_{n-2}\}$ is an r-locating-dominating code in $\mathcal{C}_n$ with $n-1$ codewords. Thus, we have $M_r^{LD}(\mathcal{C}_n) = n-1$. $\square$

The following two theorems consider r-locating-dominating codes in the cycles of length $2r+2$ and $2r+3$.

Theorem 4.2.8. *Let $r \geq 2$. Then we have*

$$M_r^{LD}(\mathcal{C}_{2r+2}) = r+1.$$

Proof. Let C be an r-locating-dominating code in $\mathcal{C}_n$ with $n = 2r+2$. For $v_i \in V_n \setminus C$, consider sets $B'_r(v_i) = V_n \setminus B_r(v_i) = \{v_{i+r+1}\}$. Since C is an r-locating-dominating code in $\mathcal{C}_{2r+2}$, the sets $B'_r(v_i) \cap C$ are unique for all $v_i \in V_n \setminus C$. Assume then that $|C| \leq r$. Since now $|V_n \setminus C| \geq r+2$, there exist (by the pigeonhole principle) vertices $v_i, v_j \in V_n \setminus C$ such that $v_i \neq v_j$ and $B'_r(v_i) \cap C = B'_r(v_j) \cap C$ (a contradiction). Thus, we have $|C| \geq r+1$.

By Lemma 4.2.2, it is straightforward to verify that $\{v_0, v_1, \ldots, v_r\}$ is an r-locating-dominating code in $\mathcal{C}_{2r+2}$. Therefore, we have $M_r^{LD}(\mathcal{C}_{2r+2}) = r+1$. $\square$

Theorem 4.2.9. *Let $r \geq 2$. Then we have*

$$M_r^{LD}(\mathcal{C}_{2r+3}) \geq \left\lceil \frac{2(2r+2)}{5} \right\rceil.$$

Proof. Let C be an r-locating-dominating code in $\mathcal{C}_n$ with $n = 2r+3$. For $v_i \in V_n \setminus C$, consider again the sets $B'_r(v_i) = V_n \setminus B_r(v_i) = \{v_{i+r+1}, v_{i+r+2}\}$. Since C is an r-locating-dominating code in $\mathcal{C}_{2r+3}$, the sets $B'_r(v_i) \cap C$ are unique for all $v_i \in V_n \setminus C$. Hence, at most one of the sets $B'_r(v_i)$ can be empty and at most $|C|$ of them contains only one codeword of C. On the other hand, each codeword can belong to at most two sets $B'_r(v_i)$. Therefore, we have the inequality $|C| + 2(n - 2|C| - 1) \leq 2|C|$. Thus, the claim immediately follows. $\square$

Let $r = 5r' + 1$, where r' is a positive integer. Now, by the previous theorem and the fact that $2r+3 = 5(2r'+1)$, we have $M_r^{LD}(\mathcal{C}_{2r+3}) = M_r^{LD}(\mathcal{C}_{5(2r'+1)}) \geq 2(2r'+1)$. Define then

$$C = \bigcup_{i=0}^{2r'} \{v_{5i}, v_{5i+1}\}.$$

It is straightforward to deduce that C is an r-locating-dominating code in $\mathcal{C}_{2r+3}$ attaining the lower bound of Theorem 4.2.9. Thus, we have an infinite family of radii r for which $M_r^{LD}(\mathcal{C}_{2r+3}) = \lceil 2(2r+2)/5 \rceil$.

Let us then determine the exact values of $M_r^{LD}(\mathcal{C}_{3r})$ and $M_r^{LD}(\mathcal{C}_{3r+3})$. The following theorem, which solves the exact values of $M_r^{LD}(\mathcal{C}_{3r})$ when r is even and $M_r^{LD}(\mathcal{C}_{3r+3})$ when r is odd, have previously been presented in [6].

Theorem 4.2.10 ([6])**.** *Let r be an integer such that $r \geq 3$.*

(i) If r is odd, then $M_r^{LD}(\mathcal{C}_{3r+3}) = r + 1$.

(ii) If r is even, then $M_r^{LD}(\mathcal{C}_{3r}) = r$.

The remaining cases are solved in the following theorem.

Theorem 4.2.11. *Let r be an integer such that $r \geq 3$.*

(i) If r is even, then $M_r^{LD}(\mathcal{C}_{3r+3}) = r + 2$.

(ii) If r is odd, then $M_r^{LD}(\mathcal{C}_{3r}) = r + 1$.

Proof. (i) Let $r \geq 3$ be an even integer. Assume that C is an r-locating-dominating code in $\mathcal{C}_{3r+3}$ with $r+1$ codewords. Let us first show that now each C-block of codewords is of length one. Assume to the contrary that $Q_t(i)$ is a C-block of codewords with $t \geq 2$ (for an appropriate integer i). Now (v_{i-1}, v_{i+t}) is a pair of C-consecutive vertices. The symmetric difference $B_r(v_{i-1}) \triangle B_r(v_{i+t}) = Q_{t+1}(i-r-1) \cup Q_{t+1}(i+r)$ contains at most one codeword, by Lemma 4.2.4 (ii). Without loss of generality, we may assume that $Q_{t+1}(i-r-1) \cap C$ is empty. Since the pairs $(v_{i-r+t-2}, v_{i-r+t-1})$ and $(v_{i-r+t-3}, v_{i-r+t-2})$ of C-consecutive vertices are r-separated, respectively, by the codewords v_{i+t-1} and v_{i+t-2}, the vertices $v_{i-2r+t-2}$ and $v_{i-2r+t-3}$ do not belong to C, by Lemma 4.2.4 (ii). By the considerations above, the symmetric difference $B_r(v_{i-2r+t-3}) \triangle B_r(v_{i-2r+t-2}) = \{v_{i-3r+t-3}, v_{i-r+t-2}\} = \{v_{i+t}, v_{i-r+t-2}\}$ does not contain codewords of C (a contradiction). Hence, each C-block is of length one.

By Lemma 4.2.6, we know that the number of C-blocks is even. Therefore, by the fact that each C-block is of length one, it immediately follows that the number of codewords in C is even. However, this contradicts the assumption that the number of vertices in C is equal to $r+1$. Thus, there does not exist an r-locating-dominating code in $\mathcal{C}_{3r+3}$ with $r+1$ codewords. Hence, we have $M_r^{LD}(\mathcal{C}_{3r+3}) \geq r + 2$. On the other hand, it is straightforward to verify (using Lemma 4.2.2) that $\{v_0, v_1, \ldots, v_r, v_{2r+1}\}$ is an r-locating-dominating code in $\mathcal{C}_{3r+3}$ with $r + 2$ codewords. Thus, we have $M_r^{LD}(\mathcal{C}_{3r+3}) = r + 2$.

(ii) Let $r \geq 3$ be an odd integer. Assume that C is an r-locating-dominating code in $\mathcal{C}_{3r}$ with r codewords. Using similar ideas as in the case (i), it can be shown that each C-block is of length one. Then a contradiction again follows using Lemma 4.2.6. Thus, we have $M_r^{LD}(\mathcal{C}_{3r}) \geq r+1$. On the other hand, it is easy to verify that $\{v_0, v_1, \ldots, v_r\}$ is an r-locating-dominating code in $\mathcal{C}_{3r}$. Therefore, we have $M_r^{LD}(\mathcal{C}_{3r}) = r + 1$. $\square$

4.2.3 Cycles with a large number of vertices

Let r be an integer such that $r \geq 5$. In this section, we prove that for any $n \geq n_r$ when n_r is large enough we have constructions showing $M_r^{LD}(\mathcal{C}_n) \leq n/3 + 1$ if $n \equiv 3 \pmod 6$ and $M_r^{LD}(\mathcal{C}_n) \leq \lceil n/3 \rceil$ if $n \not\equiv 3 \pmod 6$. By Theorem 4.2.3, the latter constructions are optimal.

The following theorem provides a useful relation between the optimal r-locating-dominating codes in cycles and paths.

Theorem 4.2.12. *Let $n \geq 4r + 2$. Then we have $M_r^{LD}(\mathcal{C}_n) \leq M_r^{LD}(\mathcal{P}_{n+1})$.*

Proof. Let C be an r-locating-dominating code in $\mathcal{P}_{n+1}$. Recall that the vertex sets of $\mathcal{P}_{n+1}$ and $\mathcal{C}_n$ are equal to V_{n+1} and $V_n = V_{n+1} \setminus \{v_n\}$, respectively. Assume first that $v_n \notin C$. Now each pair of C-consecutive vertices in $\mathcal{C}_n$ is r-separated by C, since each pair of C-consecutive vertices in $\mathcal{P}_{n+1}$ is r-separated by C. It is also easy to see that all the vertices of $\mathcal{C}_n$ are r-covered by a codeword of C and that there does not exist a vertex $u \in V_n \setminus C$ such that $B_r(u) = C$ (since $n \geq 4r + 2$). Therefore, by Lemma 4.2.2, C is an r-locating-dominating code in $\mathcal{C}_n$.

If $v_0 \notin C$, then the proof is analogous to the previous case. Hence, assume that v_0 and v_n both belong to C. Let then $v_i, v_j, v_k \in V_n \setminus C$ be vertices such that $v_0, v_1, \ldots, v_{i-1} \in C$, $v_{j+1}, v_{j+2}, \ldots, v_n \in C$ and $v_{i+1}, v_{i+2}, \ldots, v_{k-1} \in C$. In other words, (v_j, v_i) and (v_i, v_k) are pairs of C-consecutive vertices. Consider then the code $C' = C \setminus \{v_n\}$ in $\mathcal{C}_n$. It is straightforward to verify that all the pairs except (v_j, v_i) of C'-consecutive vertices in $\mathcal{C}_n$ are r-separated by C'. Moreover, the symmetric difference of $B_r(v_j)$ and $B_r(v_k)$ contains a codeword of C'. Therefore, by Lemma 4.2.2, $C' \cup \{v_i\}$ is an r-locating dominating code in $\mathcal{C}_n$. Thus, in conclusion, we have $M_r^{LD}(\mathcal{C}_n) \leq M_r^{LD}(\mathcal{P}_{n+1})$. □

Assume that $r \geq 5$ and $n \geq 3r + 2 + 6r((r-3)(2r+1) + r)$. Now, by Lemma 4.1.12, we have $M_r^{LD}(\mathcal{P}_n) = \lceil (n+1)/3 \rceil$. Hence, if $n \equiv 1 \pmod 3$, then

$$\left\lceil \frac{n}{3} \right\rceil \leq M_r^{LD}(\mathcal{C}_n) \leq M_r^{LD}(\mathcal{P}_{n+1}) = \left\lceil \frac{n+2}{3} \right\rceil .$$

Therefore, $M_r^{LD}(\mathcal{C}_n) = \lceil n/3 \rceil$. If $n \equiv 3 \pmod 6$, we similarly obtain that $n/3 \leq M_r^{LD}(\mathcal{C}_n) \leq M_r^{LD}(\mathcal{P}_{n+1}) = n/3 + 1$. We also conjecture that $M_r^{LD}(\mathcal{C}_n) = n/3 + 1$ (see Conjecture 4.2.17). In what follows, we give optimal constructions for the remaining cases when $n \equiv 0, 2$ or $5 \pmod 6$.

For this, first recall the definitions of $K_i(s)$, $L_i(s)$, $M_i(s)$ and $M_i'(s)$, and the corresponding patterns K_i, L_i and M_i from Section 4.1.3. Notice that Lemma 4.1.9 can also be used in the case of cycles by replacing each occurrence of $\mathcal{P}_n$ with $\mathcal{C}_n$. Finally, recall the definition of the set $C(s)$ (from Section 4.1.3).

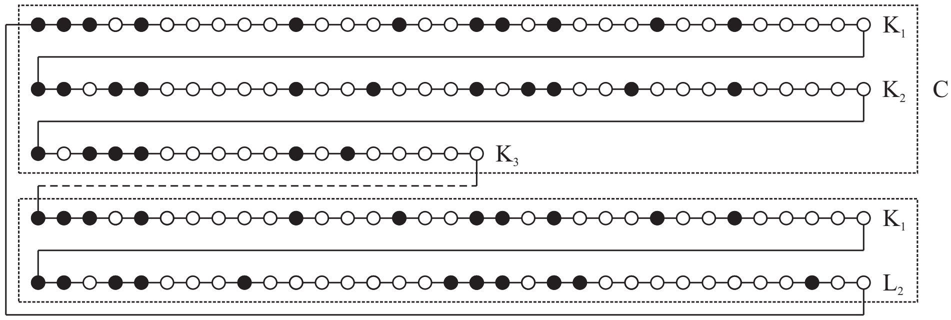


Figure 4.3: The r-locating-dominating code C_0 illustrated when $r = 5$. The pattern C, which is obtained by concatenating the patterns K_1, K_2 and K_3, is repeated p times and the concatenation of K_1 and L_2 is repeated q times.

The following constructions in the case of cycles are quite similar to the ones in the case of paths. However, attention needs to be paid to details. First let $m = p((r-3)(6r+3)+3r+3)+q \cdot 2(6r+3)$, where p and q are non-negative integers. Define then

$$\begin{aligned} C_0 = & \bigcup_{j=0}^{p-1} C(j((r-3)(6r+3)+3r+3)) \\ & \cup \bigcup_{j=0}^{q-1} K_1(p((r-3)(6r+3)+3r+3)+2j(6r+3)) \\ & \cup \bigcup_{j=0}^{q-1} L_2(p((r-3)(6r+3)+3r+3)+(2j+1)(6r+3)). \end{aligned}$$

The code C_0 is illustrated in Figure 4.3 when $r = 5$. Notice that $M_i'(s) \subseteq K_i(s)$ and $M_2'(s) \subseteq L_2(s)$ for any s. Therefore, by Lemma 4.1.9, it is immediate that each pair (v_j, v_k) of C_0-consecutive vertices in $\mathcal{C}_m$ is r-separated by C_0. It is also obvious that all the vertices in $\mathcal{C}_m$ are r-covered by a codeword of C_0 and that there does not exist a vertex $u \in V_m \setminus C_0$ such that $I_r(u) = C_0$. Thus, by Lemma 4.2.2, it is easy to conclude that C_0 is an r-locating-dominating code in $\mathcal{C}_m$ with $m/3$ codewords.

Notice further that the greatest common divisor of $(r-3)(6r+3)+3r+3$ and $2(6r+3)$ is equal to 6. Hence, the greatest common divisor of $1/2 \cdot ((r-3)(2r+1)+r+1)$ and $2r+1$ is equal to 1. Thus, by Theorem 4.1.11, if n' is an integer such that $n' \geq r((r-3)(2r+1)+r-1)$, there exist non-negative integers p and q such that $n' = p/2 \cdot ((r-3)(2r+1)+r+1)+q(2r+1)$. Therefore, if n is an integer such that $n \geq 6r((r-3)(2r+1)+r-1)$ and $n \equiv 0 \pmod 6$, then there exist integers $p \geq 0$ and $q \geq 0$ such that

$n = p((r-3)(6r+3)+3r+3)+q\cdot 2(6r+3)$. Thus, if n is an integer such that $n \geq 6r((r-3)(2r+1)+r-1)$ and $n \equiv 0 \pmod 6$, then by the previous construction we have $M_r^{LD}(\mathcal{C}_n) \leq n/3$.

Let $m = 6r+2+p((r-3)(6r+3)+3r+3)+q\cdot 2(6r+3)$, where p and q are non-negative integers. Define

$$\begin{aligned} C_2 = K_{r-2}(r-1) &\cup \bigcup_{j=0}^{p-1} C(4r+2+j((r-3)(6r+3)+3r+3)) \\ &\cup \bigcup_{j=0}^{q-1} K_1(4r+2+p((r-3)(6r+3)+3r+3)+2j(6r+3)) \\ &\cup \bigcup_{j=0}^{q-1} L_2(4r+2+p((r-3)(6r+3)+3r+3)+(2j+1)(6r+3)) \\ &\cup M_1(4r+2+p((r-3)(6r+3)+3r+3)+2q(6r+3)). \end{aligned}$$

By Lemma 4.1.9, it is immediate that if (v_i, v_j) is a pair of C_2-consecutive vertices in $\mathcal{C}_m$ such that $r-1 \leq i \leq m-r-1$ and $r-1 \leq j \leq m-r-1$, then (v_i, v_j) is r-separated by C_2. Consider then the remaining pairs of C_2-consecutive vertices. For the following considerations, we first recall that $M_1(m-2r) = \{v_{-2r}, v_{-2r+1}, \ldots, v_{-r-3}, v_{-r-1}, v_{-1}\}$ and $K_{r-2}(r-1) = \{v_{r-1}, v_{r+1}, v_{r+2}, \ldots, v_{2r-2}, v_{3r-1}, v_{3r+1}\}$. Now it is easy to see that the pairs (v_{-r-2}, v_{-r}) and (v_{r-2}, v_r) are r-separated by the codeword v_{-1} and the pair (v_{-2}, v_0) is r-separated by the codeword v_{r-1}. Furthermore, for all $i = -r, -r+1, \ldots, -3$ and $j = 0, 1, \ldots, r-3$ the pairs (v_i, v_{i+1}) and (v_j, v_{j+1}) are r-separated by the codewords v_{i-r} and v_{j+1+r}, respectively. Thus, each pair of C_2-consecutive vertices in $\mathcal{C}_m$ is r-separated by C_2. Therefore, by Lemma 4.2.2, it is straightforward to verify that C_2 is an r-locating-dominating code in $\mathcal{C}_m$ with $\lceil m/3 \rceil$ codewords. Thus, as in the previous case, if n is an integer such that $n \geq 6r+2+6r((r-3)(2r+1)+r-1)$ and $n \equiv 2 \pmod 6$, then by the construction above we have $M_r^{LD}(\mathcal{C}_n) \leq \lceil n/3 \rceil$.

Let $m = 12r+5+p((r-3)(6r+3)+3r+3)+q\cdot 2(6r+3)$, where p and q are non-negative integers. Define

$$\begin{aligned} C_5 = K_{r-2}(r) &\cup \bigcup_{j=0}^{p-1} C(4r+3+j((r-3)(6r+3)+3r+3)) \\ &\cup \bigcup_{j=0}^{q} K_1(4r+3+p((r-3)(6r+3)+3r+3)+2j(6r+3)) \\ &\cup \bigcup_{j=0}^{q-1} L_2(4r+3+p((r-3)(6r+3)+3r+3)+(2j+1)(6r+3)) \\ &\cup M_2(4r+3+p((r-3)(6r+3)+3r+3)+(2q+1)(6r+3)). \end{aligned}$$

Again, using Lemmas 4.1.9 and 4.2.2, it can be shown that C_5 is an r-locating-dominating code in $\mathcal{C}_m$ with $\lceil m/3 \rceil$ codewords. Thus, if n is an integer such that $n \geq 12r + 5 + 6r((r-3)(2r+1) + r - 1)$ and $n \equiv 5 \pmod 6$, then by the previous construction $M_r^{LD}(\mathcal{C}_n) \leq \lceil n/3 \rceil$.

Combining the previous results with the lower bound of Theorem 4.2.3, we immediately obtain the following theorem.

Theorem 4.2.13. *Let $r \geq 5$ and $n \geq 12r + 5 + 6r((r-3)(2r+1) + r - 1)$.*

(i) If $n \not\equiv 3 \pmod 6$, then $M_r^{LD}(\mathcal{C}_n) = \lceil n/3 \rceil$.

(ii) If $n \equiv 3 \pmod 6$, then $n/3 \leq M_r^{LD}(\mathcal{C}_n) \leq n/3 + 1$.

In the latter case of the previous theorem, we conjecture that actually $M_r^{LD}(\mathcal{C}_n) = n/3 + 1$ (see Conjecture 4.2.17). Furthermore, if r is an odd integer such that $r \geq 5$, it has been shown in [26] that for the previous theorem to hold it is enough that $n \geq 6r+1+(r-1)(3r+3)$ (i.e. $n = \mathcal{O}(r^2)$).

4.2.4 The exact values of $M_r^{LD}(\mathcal{C}_n)$ when $2 \leq r \leq 4$

In this section, we consider r-locating-dominating codes in $\mathcal{C}_n$ when $2 \leq r \leq 4$. The exact values of $M_2^{LD}(\mathcal{C}_n)$ are determined in [20]. In particular, it is shown that for $n > 6$ if $n \equiv 3 \pmod 6$, then $M_2^{LD}(\mathcal{C}_n) = n/3 + 1$, else $M_2^{LD}(\mathcal{C}_n) = \lceil n/3 \rceil$. In the following theorem, we provide an alternative (and shorter) proof for the lower bound in the case $n \equiv 3 \pmod 6$.

Theorem 4.2.14 ([20])**.** *Let $n \equiv 3 \pmod 6$. Then we have*

$$M_2^{LD}(\mathcal{C}_n) \geq n/3 + 1.$$

Proof. Let C be a 2-locating-dominating code in $\mathcal{C}_n$ with $n/3$ vertices. Now, by Lemma 4.2.5, each C-block is of length one. By Lemma 4.2.6, the number of C-blocks is even. Hence, by combining these two observations, the number of codewords of C is even. This contradicts with the fact that $|C| = n/3$ (an odd integer since $n \equiv 3 \pmod 6$). Thus, we have $M_2^{LD}(\mathcal{C}_n) \geq n/3 + 1$. □

With our new approach, a lower bound similar to the previous theorem can also be proved when $r = 3$ and $r = 4$. The following theorem shows the result for 3-locating-dominating codes.

Theorem 4.2.15. *Let $n \equiv 3 \pmod 6$. Then we have*

$$M_3^{LD}(\mathcal{C}_n) \geq n/3 + 1.$$

Proof. Let C be a 3-locating-dominating code in $\mathcal{C}_n$ with $n/3$ vertices. Notice that each C-block of codewords is now at most of length 2 (by

Lemma 4.2.6). In what follows, we show that the number of C-blocks of length two is even.

Recall that according to Lemma 4.2.4 each pair of C-consecutive vertices is 3-separated by exactly one codeword of C. Assume then that $\{v_i, v_{i+1}\}$ is a C-block of length two. By the previous observation, the set $B_r(v_{i-1}) \triangle B_r(v_{i+2})$ contains exactly one codeword of C. Without loss of generality, we may assume that v_{i-4}, v_{i-3} and v_{i-2} do not belong to C. Then either v_{i+3} or v_{i+5} belongs to C. (Notice that if $v_{i+4} \in C$, then the pair (v_{i+3}, v_{i+5}) of C-consecutive vertices is 3-separated by at least two codewords.)

Assume first that $v_{i+5} \in C$. If now $v_{i+6} \notin C$, then the pair (v_{i+2}, v_{i+3}) of C-consecutive vertices is not r-separated by any codeword of C. Hence, $v_{i+6} \in C$ and further $v_{i+7} \notin C$. Therefore, $\{v_{i+5}, v_{i+6}\}$ is also a C-block of length two. Since the neighbourhoods of the C-blocks $\{v_i, v_{i+1}\}$ and $\{v_{i+5}, v_{i+6}\}$ are symmetrical to each other, these C-blocks of length two can be paired with each other.

Assume then that $v_{i+3} \in C$. Considering now the pairs (v_{i+2}, v_{i+4}), (v_{i+4}, v_{i+5}) and (v_{i+6}, v_{i+7}), we obtain that v_{i+6}, v_{i+7}, v_{i+8} and v_{i+10} do not belong to C. The pairs (v_{i+5}, v_{i+6}) and (v_{i+7}, v_{i+8}) of C-consecutive vertices imply that v_{i+9} and v_{i+11} belong to C. By the fact that now (v_{i+8}, v_{i+10}) is a pair of C-consecutive vertices, we know that either v_{i+12} or v_{i+13} is a codeword of C. If $v_{i+12} \in C$, then $\{v_{i+11}, v_{i+12}\}$ is a C-block and the neighbourhoods of the C-blocks $\{v_i, v_{i+1}\}$ and $\{v_{i+11}, v_{i+12}\}$ are symmetrical to each other. Therefore, these C-blocks of length two can be paired with each other. Assume then that $v_{i+13} \in C$. Consider then the symmetric difference $B_r(v_{i+10}) \triangle B_r(v_{i+12})$, where (v_{i+10}, v_{i+12}) is a pair of C-consecutive vertices. Now either v_{i+14} or v_{i+15} belongs to C. If $v_{i+14} \in C$, then the pair (v_{i+12}, v_{i+15}) of C-consecutive vertices is 3-separated by at least two codewords (a contradiction). Therefore, v_{i+15} belongs to C. Using similar arguments as above, we obtain that $v_{i+16}, v_{i+17}, v_{i+18}, v_{i+19}, v_{i+20}, v_{i+22} \notin C$ and $v_{i+21}, v_{i+23} \in C$. The situation is now analogous to the one in which we considered the pair (v_{i+8}, v_{i+10}) of C-consecutive vertices instead that here we have the pair (v_{i+20}, v_{i+22}).

The previous reasonings can now be repeated. However, since we are operating in a cycle, at some point the repetition has to end. Therefore, for some integer $k \geq 0$ we have that $\{v_i, v_{i+1}\}$ and $\{v_{i+11+12k}, v_{i+12+12k}\}$ are C-blocks with symmetrical neighbourhoods. Clearly, the sets $\{v_i, v_{i+1}\}$ and $\{v_{i+11+12k}, v_{i+12+12k}\}$ do not coincide. Thus, these C-blocks of length two can be paired with each other. In conclusion, each C-block of length two can be uniquely paired to another C-block of length two. Therefore, the number of C-blocks of length two is even.

By Lemma 4.2.6, the number of C-blocks is even. Hence, by the previous considerations, the number of C-blocks of length one is also even. Thus, the number of codewords of C is even. This contradicts with the fact that $|C| = n/3$ is odd. Therefore, we have $M_3^{LD}(\mathcal{C}_n) \geq n/3 + 1$. $\square$

In the following theorem, a lower bound similar to the one in Theorems 4.2.14 and 4.2.15 is presented for 4-locating-dominating codes in cycles.

Theorem 4.2.16. *Let* $n \equiv 3 \pmod 6$. *Then we have*

$$M_4^{LD}(\mathcal{C}_n) \geq n/3 + 1.$$

Proof. Let C be a 4-locating-dominating code in $\mathcal{C}_n$ with $n/3$ vertices. As earlier, we start by showing that the number of C-blocks of length two is even.

Let $\{v_i, v_{i+1}\}$ be a C-block of length two. Without loss of generality, we can again assume that v_{i-5}, v_{i-4} and v_{i-3} do not belong to C. As in the previous proof, we can also conclude that v_{i+5} does not belong to C. Moreover, since v_{i-1} and v_{i+2} are 4-separated by C, either $v_{i+4} \in C$ or $v_{i+6} \in C$ by Lemma 4.2.4.

In what follows, we are going to classify C-blocks of length two into different types depending on their neighbourhood. If $\{v_{i+6}, v_{i+7}, v_{i+8}\}$ is a C-block of length three, then we say that C-block $\{v_i, v_{i+1}\}$ is of type A_1. If $\{v_{i+6}, v_{i+7}\}$ is a C-block of length two, then a contradiction follows since the pair (v_{i+3}, v_{i+4}) of C-consecutive vertices is not 4-separated by a codeword. Assume that $\{v_{i+6}\}$ is a C-block of length one. Then v_{i+4} does not belong to C. If $v_{i+3} \notin C$, then the C-block $\{v_i, v_{i+1}\}$ is said to be of type A_2. Assume further that $v_{i+3} \in C$. If now $v_{i-2} \notin C$, then we say that $\{v_i, v_{i+1}\}$ is of type A_3, else it is of type A_4.

If $\{v_{i+3}, v_{i+4}\}$ is a C-block of length two, then ($v_{i+6} \notin C$ and) $\{v_i, v_{i+1}\}$ is of type A_5. Assume now that $\{v_{i+4}\}$ is a C-block of length one. Then v_{i+6} does not belong to C. If $v_{i-2} \in C$, then the C-block $\{v_i, v_{i+1}\}$ is of type A_6, else it is of type A_7.

For each of the previous types A_i we also have a symmetrical pair A'_i which is considered as a reflection of the neighbourhood of type A_i (between the vertices v_i and v_{i+1}). For example, if $v_{i-4}, v_{i+4}, v_{i+5}, v_{i+6} \notin C$ and $\{v_{i-7}, v_{i-6}, v_{i-5}\}$ is a C-block of length three, then we say that C-block $\{v_i, v_{i+1}\}$ is of type A'_1. By the previous considerations, it is straightforward to verify that each C-block of length two is one of the types A_i or A'_i.

Assume that the C-block $\{v_i, v_{i+1}\}$ is of type A'_1. Then v_{i-2}, v_{i-1}, v_{i+2}, v_{i+3}, v_{i+4}, v_{i+5} and v_{i+6} do not belong to C. Considering the pairs (v_{i+2}, v_{i+3}), (v_{i+3}, v_{i+4}), (v_{i+4}, v_{i+5}) and (v_{i+5}, v_{i+6}) of C-consecutive vertices, we have that $v_{i+7}, v_{i+8} \in C$ and $v_{i+9}, v_{i+10} \notin C$. The pair (v_{i+9}, v_{i+10}) of C-consecutive vertices imply that $v_{i+14} \in C$. Therefore, considering the pair (v_{i+6}, v_{i+9}) of C-consecutive vertices, we obtain that the C-block

$\{v_{i+7}, v_{i+8}\}$ of length two is either of type A_1 or A_7. The proof of the following symmetrical result is analogous: if the C-block $\{v_i, v_{i+1}\}$ is of type A_1, then the C-block $\{v_{i-7}, v_{i-6}\}$ of length two is either of type A'_1 or A'_7.

Assume that the C-block $\{v_i, v_{i+1}\}$ is of type A_2. Then the vertices v_{i+2}, v_{i+3}, v_{i+4} and v_{i+5} do not belong to C. Considering the pairs (v_{i+2}, v_{i+3}), (v_{i+3}, v_{i+4}) and (v_{i+4}, v_{i+5}) of C-consecutive vertices, it immediately follows that $v_{i+7}, v_{i+9} \notin C$ and $v_{i+8} \in C$. Since $v_{i+1} \in B_4(v_{i+5}) \triangle B_4(v_{i+7})$, then $v_{i+10}, v_{i+11} \notin C$. Considering the pair (v_{i+7}, v_{i+9}) of C-consecutive vertices, we know that either v_{i+12} or v_{i+13} belong to C. If $v_{i+13} \in C$, then it is straightforward to conclude (using similar arguments as before) that $\{v_{i+13}, v_{i+14}\}$ is a C-block of type A'_2. Otherwise, it can be seen that $v_{i+12}, v_{i+14} \in C$ and $v_{i+13}, v_{i+15}, v_{i+16}, v_{i+17} \notin C$. The situation is now analogous to the one in which we considered the pair (v_{i+7}, v_{i+9}) of C-consecutive vertices instead that here we have the pair (v_{i+13}, v_{i+15}). The previous reasonings can be repeated. However, since we are operating in a cycle, at some point the repetition has to end. Therefore, for some non-negative integer k we have that $\{v_i, v_{i+1}\}$ and $\{v_{i+13+6k}, v_{i+14+6k}\}$ are C-blocks of type A_2 and A'_2, respectively. The following symmetrical result also holds: if $\{v_i, v_{i+1}\}$ is a C-block of type A'_2, then for some non-negative integer k we have that $\{v_{i-13-6k}, v_{i-12-6k}\}$ is a C-block of type A_2.

In the following, we list the results of the previous two paragraphs and other analogous ones, which can be obtained using similar arguments:

- If $\{v_i, v_{i+1}\}$ is a C-block of type A'_1, then $\{v_{i+7}, v_{i+8}\}$ is a C-block either of type A_1 or A_7.
- If $\{v_i, v_{i+1}\}$ is a C-block of type A_2, then for some non-negative integer k we know that $\{v_{i+13+6k}, v_{i+14+6k}\}$ is a C-block of type A'_2.
- If $\{v_i, v_{i+1}\}$ is a C-block of type A'_3, then either $\{v_{i+11}, v_{i+12}\}$ is a C-block of type A'_6 or $\{v_{i+13}, v_{i+14}\}$ is a C-block of type A'_4.
- If $\{v_i, v_{i+1}\}$ is a C-block of type A_4, then $\{v_{i+13}, v_{i+14}\}$ is a C-block either of type A_3 or A_5.
- If $\{v_i, v_{i+1}\}$ is a C-block of type A'_5, then either $\{v_{i+11}, v_{i+12}\}$ is a C-block of type A'_6 or $\{v_{i+13}, v_{i+14}\}$ is a C-block of type A'_4.
- If $\{v_i, v_{i+1}\}$ is a C-block of type A_6, then $\{v_{i+11}, v_{i+12}\}$ is a C-block either of type A_3 or A_5.
- If $\{v_i, v_{i+1}\}$ is a C-block of type A'_7, then $\{v_{i+7}, v_{i+8}\}$ is a C-block either of type A_1 or A_7.

The obvious symmetrical results also hold. For example, if $\{v_i, v_{i+1}\}$ is a C-block of type A'_4, then $\{v_{i-13}, v_{i-12}\}$ is a C-block either of type A'_3 or A'_5.

The results listed above provide an approach to pair C-blocks of length two. The C-block $\{v_i, v_{i+1}\}$ depending on its type is paired with the C-block of length two suggested by the previous results. For example, the C-block $\{v_i, v_{i+1}\}$ of type A_3' is paired with $\{v_{i+11}, v_{i+12}\}$ or $\{v_{i+13}, v_{i+14}\}$ depending on which one of these sets is a C-block. Using the results listed above, it is straightforward to verify that this way each C-block of length two is uniquely paired with another such one. Therefore, the number of C-blocks of length two is even.

By Lemma 4.2.6, the number of C-blocks is even. Hence, since the number of C-blocks of length two is even, the number of C-blocks that are of length one or three is also even. Thus, the number of codewords of C is even. This contradicts with the fact that $|C| = n/3$. Therefore, we have $M_4^{LD}(\mathcal{C}_n) \geq n/3 + 1$. $\square$

Theorems 4.2.11, 4.2.14, 4.2.15 and 4.2.16 suggest the following conjecture.

Conjecture 4.2.17. *Let n be a positive integer such that $n \equiv 3 \pmod 6$. Then for any r we have*

$$M_r^{LD}(\mathcal{C}_n) \geq n/3 + 1.$$

Notice that when n is large enough (with respect to r), we believe that the conjecture holds with equality, i.e. that $M_r^{LD}(\mathcal{C}_n) = n/3 + 1$ (recall the constructions from the previous section). However, we do not have the equality with small n. For example, by Theorem 4.2.7, we have $M_4^{LD}(\mathcal{C}_9) = 8 > 4$.

In what follows, we construct optimal r-locating-dominating codes in $\mathcal{C}_n$ when $3 \leq r \leq 4$. The following lemma, which is analogous to Lemma 4.1.13, is needed in the constructions.

Lemma 4.2.18. *Let n and k be integers such that*

$$D = \{v_k, v_{k+2}, v_{k+6}, v_{k+8}, v_{k+12}, v_{k+14}\} \subseteq V_n.$$

If a pair (v_i, v_j) of D-consecutive vertices in $\mathcal{C}_n$ is such that $k+5 \leq i \leq k+13$ and $k+5 \leq j \leq k+13$, then v_i and v_j are 3- and 4-separated by D. Moreover, for each vertex $v_i \in V_n \setminus D$ such that $k+6 \leq i \leq k+11$ we have $\emptyset \subsetneq I_3(D; v_i) \subsetneq D$ and $\emptyset \subsetneq I_4(D; v_i) \subsetneq D$.

Consider then 3-locating-dominating codes in $\mathcal{C}_n$. The exact values of $M_3^{LD}(\mathcal{C}_n)$ when $3 \leq n \leq 8$ are determined in Theorems 4.2.7 and 4.2.8. Let p be a non-negative integer. Define then

$$D(p) = \bigcup_{i=0}^{p} \{v_{6i}, v_{6i+2}\}.$$

It is straightforward to verify that $D(1)$ and $D(2)$ are 3-locating-dominating codes in $\mathcal{C}_9$, $\mathcal{C}_{10}$, $\mathcal{C}_{11}$, $\mathcal{C}_{12}$ and $\mathcal{C}_{15}$, $\mathcal{C}_{16}$, $\mathcal{C}_{17}$, $\mathcal{C}_{18}$, respectively. Therefore, by combining Lemmas 4.2.2 and 4.2.18, it can be concluded that $D(p)$ is a 3-locating-dominating code in $\mathcal{C}_{6p+3}$, $\mathcal{C}_{6p+4}$, $\mathcal{C}_{6p+5}$ and $\mathcal{C}_{6p+6}$ with $2(p+1)$ codewords when $p \geq 1$. Similarly, it can be shown that $D(p) \cup \{v_{6p+5}\}$ is a 3-locating-dominating code in $\mathcal{C}_{6p+8}$ with $2p+3$ codewords when $p \geq 1$. Furthermore, $D(p) \cup \{v_{6p+5}, v_{6p+8}, v_{6p+10}\}$ is a 3-locating-dominating code in $\mathcal{C}_{6p+13}$ with $2p+5$ codewords when $p \geq 0$. In conclusion, the constructions given above attain the lower bounds of Theorems 4.2.3 and 4.2.15. Thus, the exact values of $M_3^{LD}(\mathcal{C}_n)$ are determined for all n.

Consider now 4-locating-dominating codes in $\mathcal{C}_n$. By Theorems 4.2.7 and 4.2.8, the exact values of $M_4^{LD}(\mathcal{C}_n)$ are known when $3 \leq n \leq 10$. By Lemma 4.2.18, $D_1(p)$ is a 4-locating-dominating code in $\mathcal{C}_{6p+6}$ when $p \geq 2$. Using analogous arguments as above in the case $r = 3$, the following results can be shown:

- The code $D(p) \cup \{v_{6p+5}, v_{6p+7}, v_{6p+8}\}$ is 4-locating-dominating in $\mathcal{C}_{6p+13}$ with $2p+5$ codewords when $p \geq 0$.
- The code $D(p) \cup \{v_{6p+7}\}$ is 4-locating-dominating in $\mathcal{C}_{6p+8}$ with $2p+3$ codewords when $p \geq 1$.
- The code $D(p) \cup \{v_{6p+4}, v_{6p+7}, v_{6p+9}, v_{6p+10}\}$ is 4-locating-dominating in $\mathcal{C}_{6p+15}$ with $2p+6$ codewords when $p \geq 0$.
- The code $D(p) \cup \{v_{6p+4}, v_{6p+6}\}$ is 4-locating-dominating in $\mathcal{C}_{6p+10}$ with $2p+4$ codewords when $p \geq 1$.
- For $p \geq 0$, the code $D(p) \cup \{v_{6p+7}, v_{6p+8}, v_{6p+10}, v_{6p+15}, v_{6p+18}, v_{6p+21}\}$ is 4-locating-dominating in $\mathcal{C}_{6p+23}$ with $2p+8$ codewords.

In conclusion, by Theorems 4.2.3 and 4.2.16, the exact values of $M_4^{LD}(\mathcal{C}_n)$ are determined for all n except 11, 12 or 17. The missing values can be easily determined since it is straightforward to verify that $\{v_0, v_1, v_3, v_4\}$, $\{v_0, v_2, v_4, v_6\}$ and $\{v_0, v_1, v_4, v_7, v_{10}, v_{11}\}$ are 4-locating-dominating codes in $\mathcal{C}_{11}$, $\mathcal{C}_{12}$ and $\mathcal{C}_{17}$, respectively, attaining the lower bound of Theorem 4.2.3.

The following theorem summarizes the previous results on 3- and 4-locating-dominating codes.

Theorem 4.2.19. *Let $n \geq 3$ and $3 \leq r \leq 4$. Then we have the following results:*

(i) $M_r^{LD}(\mathcal{C}_n) = n - 1$ *if* $3 \leq n \leq 2r + 1$.

(ii) $M_r^{LD}(\mathcal{C}_{2r+2}) = r + 1$.

(iii) $M_r^{LD}(\mathcal{C}_n) = n/3 + 1$ *if* $n > 2r + 2$ *and* $n \equiv 3 \pmod 6$.

(iv) $M_r^{LD}(\mathcal{C}_n) = \lceil n/3 \rceil$ *if* $n > 2r + 2$ *and* $n \not\equiv 3 \pmod 6$.

In finding the optimal families of r-locating-dominating codes in the cases $r = 3$ and $r = 4$, some computer searches were applied to obtain the initial codes. A brief explanation of the used algorithms can be found in [27].

Chapter 5

Optimal 2-identifying code in the hexagonal grid

Previously, a 2-identifying code in the hexagonal grid with density $4/19$ has been constructed in [16]. In this chapter, which is based on the papers [57] and [58], we show that this 2-identifying code is optimal, i.e. that there do not exist any 2-identifying codes in the hexagonal grid with density smaller than $4/19$. In Section 5.1, we first start by presenting some preliminary definitions and summarising known results. Then, in Section 5.2, we define the notion of share and explain how it is used in obtaining lower bounds for r-identifying codes. Finally, in Section 5.3, the proof of the lower bound is presented.

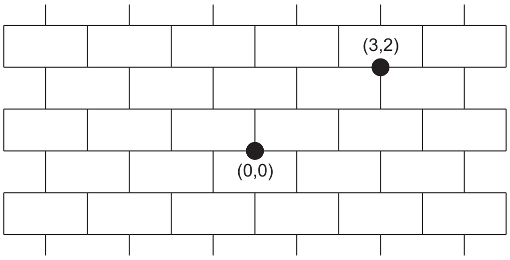
5.1 Preliminaries

We define the hexagonal grid $G_H = (V_H, E_H)$ using the brick wall representation as follows: the set of vertices $V_H = \mathbb{Z}^2$ and the set of edges

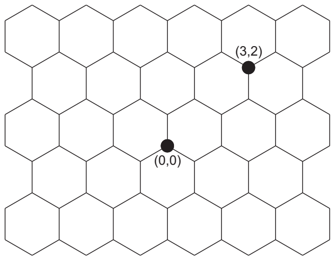
$$E_H = \{\{\mathbf{u} = (i,j), \mathbf{v}\} \mid \mathbf{u}, \mathbf{v} \in \mathbb{Z}^2, \mathbf{u} - \mathbf{v} \in \{(0, (-1)^{i+j+1}), (\pm 1, 0)\}\}.$$

This definition is illustrated in Figure 5.1(a). The hexagonal grid can also be illustrated using the honeycomb representation as in Figure 5.1(b). In both illustrations, lines represent the edges and intersections of the lines represent the vertices of G_H. The labeling of the vertices in the brick wall representation is self-explanatory. This labeling can also be applied to the honeycomb representation, if we visualize the honeycomb to be obtained from the brick wall by squeezing it from left and right. For an example of the labeling of the vertices, we refer to Figure 5.1.

Previously, r-identifying codes in G_H have been studied in various papers. The first results concerning r-identifying codes in G_H have been presented in the seminal paper [60] in the case $r = 1$. Later these results have



(a) Brick wall



(b) Honeycomb

Figure 5.1: The brick wall and the honeycomb representations illustrated.

been improved by showing that there exists a 1-identifying code with density 3/7 (see Cohen *et al.* [23]) and that there do not exist 1-identifying codes in G_H with density smaller than 12/29 (see Cranston and Yu [25]). For general $r \geq 2$, Charon *et al.* [14] showed that each r-identifying code C in G_H has $D(C) \geq 2/(5r+3)$ if r is even and $D(C) \geq 2/(5r+2)$ if r is odd. They also presented a construction for each $r \geq 2$ giving an r-identifying code $C \subseteq V_H$ with $D(C) \sim 8/(9r)$.

For small values of r, the previous constructions have been improved in [16] by Charon *et al.* In particular, it is shown that there exists a 2-identifying code in G_H with density 4/19. In the case $r = 2$, the lower bound 2/11 from [16, Equation (1)] (and the aforementioned general lower bound) is improved in Martin and Stanton [66] by showing that the density of any 2-identifying code in G_H is at least 1/5. In this chapter, we further improve this lower bound to 4/19. In other words, we show that the previously presented 2-identifying code with density 4/19 is optimal.

5.2 Lower bounds using share

Let $G = (V, E)$ be a simple, connected and undirected graph. Assume also that C is a code in G. The following concept of the share of a codeword has been introduced by Slater in [73]. The *share* of a codeword $c \in C$ is defined as

$$s_r(C; c) = s_r(c) = \sum_{u \in B_r(c)} \frac{1}{|I_r(C; u)|}.$$

The notion of share proves to be useful in determining lower bounds of r-identifying codes (as explained in the following paragraph).

Assume that $G = (V, E)$ is a finite graph and D is a code in G such that $B_r(u) \cap D$ is non-empty for all $u \in V$. Now we have $\sum_{c \in D} s_r(D; c) = |V|$, since each vertex $u \in V$ such that $|I_r(u)| = k$ contributes the summand $1/k$ to $s_r(D; c)$ for each of the k codewords $c \in B_r(u)$. Assume further that there exists a positive real number α such that $s_r(D;c) \leq \alpha$ for all $c \in D$. Then we have $|V| \leq \alpha|D|$, which immediately implies that

$$|D| \geq \frac{1}{\alpha}|V|.$$

Assume then that for any r-identifying code C in G we have $s_r(C;c) \leq \alpha$ for all $c \in C$. By the aforementioned observation, we then obtain the lower bound $|V|/\alpha$ for the size of an r-identifying code in G. In other words, by determining the maximum share for any r-identifying code, we obtain a lower bound for the minimum size of an r-identifying code.

The previous reasoning can also be generalized to the case when an infinite graph is considered. In particular, if for any r-identifying code C in G_H we have $s_r(C;c) \leq \alpha$ for all $c \in C$, then it can be shown that the density of an r-identifying code in G_H is always at least $1/\alpha$ (compare to Theorem 5.3.5). The main idea behind the proof of the lower bound (in Section 5.3) is based on this observation, although we use a more sophisticated method by showing that for any 2-identifying code the share is on *average* at most 19/4. In Theorem 5.3.5, we present a formal proof to verify that this method is indeed valid.

In the proof of the lower bound, we need to determine upper bounds for shares of codewords. To formally present a way to estimate shares, we first need to introduce some notations.

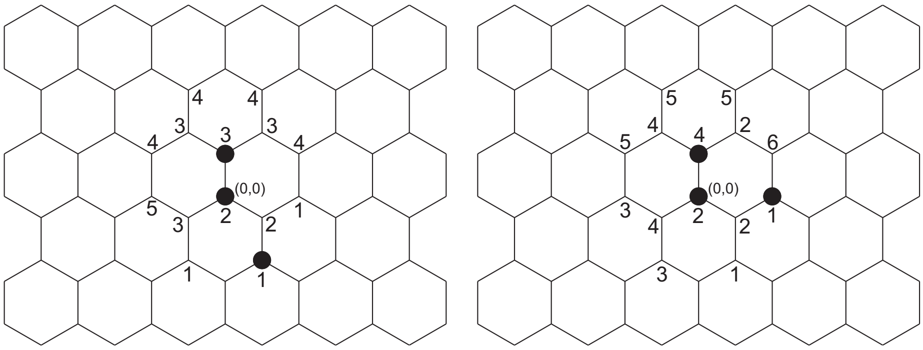
Let $D \subseteq V$ be a code and c be a codeword of D. Consider then the I-sets $I_r(D;u)$ when u goes through all the vertices in $B_r(c)$. (Notice that all of these I-sets do not have to be different.) Denote the different identifying sets by $I_1, I_2, \ldots, I_k$, where k is a positive integer. Furthermore, denote the number of identifying sets equal to I_j by i_j ($j = 1, 2, \ldots, k$). Now we are ready to present the following lemma, which provides a method to estimate the shares of the codewords.

Lemma 5.2.1. *Let C be an r-identifying code in G and let D be a non-empty subset of C. For $c \in D$, using the previous notations, we have*

$$s_r(C;c) \leq \sum_{j=1}^{k} \left(\frac{1}{|I_j|} + (i_j - 1)\frac{1}{|I_j| + 1} \right).$$

Proof. Assume that $c \in D$. Then, for each $j = 1, 2, \ldots, k$, define $\mathcal{I}_j = \{u \in B_r(c) \mid I_j = I_r(D;u)\}$. Now it is obvious that for at most one vertex $u \in \mathcal{I}_j$ we have $I_j = I_r(C;u)$ and the other vertices of $\mathcal{I}_j$ are r-covered by at least $|I_j| + 1$ codewords of C. Hence, the claim immediately follows. $\square$

The previous lemma will be used numerous times in the following presentation. The computations needed in applying this lemma may sometimes be



(a) The first case

(b) The second case

Figure 5.2: The cases of Example 5.2.2 illustrated. The black dots represent codewords of C.

a little bit tedious, but always very straightforward. It is also quite easy to implement an algorithm to compute the upper bound given by the lemma. Furthermore, the use of the previous lemma is illustrated in the following example.

Example 5.2.2. Let C be a 2-identifying code in the hexagonal grid G_H. For the first case (see Figure 5.2(a)), assume that $D = \{(0,0),(0,1),(1,-1)\}$ is a subset of C. Now we have the following facts:

- $I_2(D;\mathbf{u}) = \{(0,0),(1,-1)\}$ for $\mathbf{u} = (-1,-1),(1,-1)$ and $(2,0)$ (the vertices labeled with 1 in the figure),
- $I_2(D;\mathbf{u}) = \{(0,1),(0,0),(1,-1)\}$ for $\mathbf{u} = (0,0)$ and $(1,0)$ (the vertices labeled with 2 in the figure),
- $I_2(D;\mathbf{u}) = \{(0,0),(0,1)\}$ for $\mathbf{u} = (-1,1),(0,1),(1,1)$ and $(-1,0)$ (the vertices labeled with 3 in the figure),
- $I_2(D;\mathbf{u}) = \{(0,1)\}$ for $\mathbf{u} = (-2,1),(-1,2),(1,2)$ and $(2,1)$ (the vertices labeled with 4 in the figure), and
- $I_2(D;(-2,0)) = \{(0,0)\}$ (the vertex labeled with 5 in the figure).

Thus, by Lemma 5.2.1, we obtain that

$$s_2(C;(0,0)) \leq \left(\frac{1}{2} + 2\cdot\frac{1}{3}\right) + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{2} + 3\cdot\frac{1}{3}\right) + 1 = \frac{17}{4}.$$

Similarly, we also have

$$s_2(C;(0,1)) \leq 1 + 4\cdot\frac{1}{2} + 4\cdot\frac{1}{3} + \frac{1}{4} = \frac{55}{12}.$$

For the second case, which is illustrated in Figure 5.2(b), we assume that $D = \{(0,0),(0,1),(2,0)\}$ is a subset of C. As in the previous case (with the aid of the figure), it can be concluded that

$$s_2(C;(0,0)) \leq \frac{13}{3}$$

and

$$s_2(C;(0,1)) \leq \frac{9}{2}.$$

It should be noted that the results of this example will be later used in the proof of Lemma 5.3.3.

5.3 The proof of the lower bound

For the rest of the section, assume that C is a 2-identifying code in G_H. It can be shown that $s_2(\mathbf{c}) \leq 5$ for all $\mathbf{c} \in C$ (see the proof of Lemma 5.3.4). This provides another approach to obtain the lower bound $D(C) \geq 1/5$, which was previously shown in [66]. In order to improve this lower bound, we need to consider the shares of codewords on average. Indeed, we can show that on average the share of a codeword is at most $19/4$. Therefore, as shown in Theorem 5.3.5, we obtain that the density $D(C) \geq 4/19$.

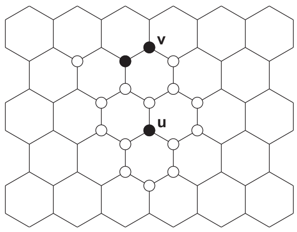
The averaging process is done by introducing a *shifting scheme* designed to even out the shares among the codewords of C. (Notice that the shifting scheme can also be understood as a discharging method.) The rules of the shifting scheme are defined in Section 5.3.1. In Section 5.3.2, we introduce three lemmas, which state the following results:

- If $s_2(\mathbf{c}) > 19/4$ for some $\mathbf{c} \in C$, then at least $s_2(\mathbf{c}) - 19/4$ units of share is shifted from $\mathbf{c}$ to other codewords. (Lemma 5.3.4)
- If share is shifted to a codeword $\mathbf{c} \in C$, then $s_2(\mathbf{c}) \leq 19/4$ and the codeword $\mathbf{c}$ receives at most $19/4 - s_2(\mathbf{c})$ units of share. (Lemmas 5.3.2 and 5.3.3)

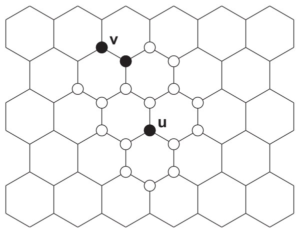
In other words, after the shifting is done, the share of each codeword is at most $19/4$. Using this fact, we are able to prove the main theorem (Theorem 5.3.5) of the paper according to which $D(C) \geq 4/19$. Finally, in Section 5.3.3, we provide the proofs of the three lemmas.

5.3.1 The rules of the shifting scheme

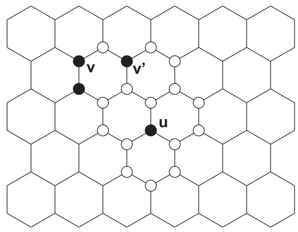
The *rules* of the shifting scheme are illustrated in Figure 5.3. Translations, rotations and reflections (over the line passing vertically through $\mathbf{u}$) can be applied to each rule in such a way that the structure of the underlying graph G_H is preserved. In the rules, share is shifted as follows:



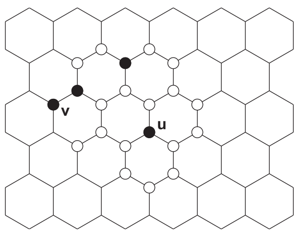
(a) Rule 1



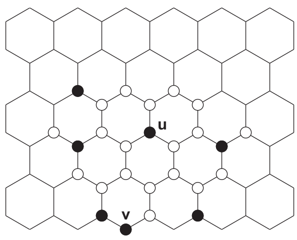
(b) Rule 2



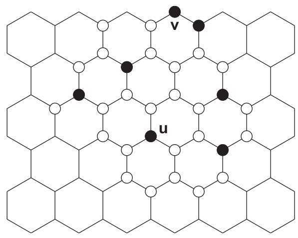
(c) Rule 3



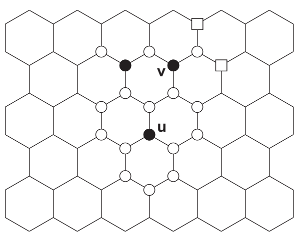
(d) Rule 4



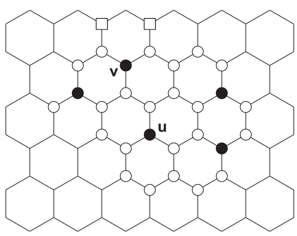
(e) Rule 5



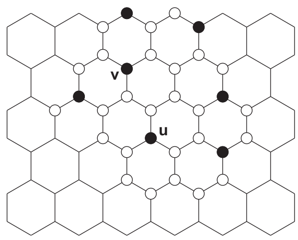
(f) Rule 6



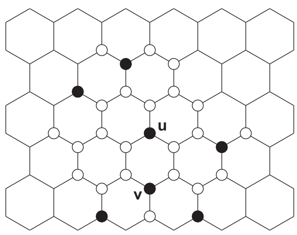
(g) Rule 7



(h) Rule 8



(i) Rule 9



(j) Rule 10

Figure 5.3: The rules of the shifting scheme illustrated. The black dots represent codewords and the white dots represent non-codewords. In the rules 7 and 8, at least one of the vertices marked with a white square is a codeword.

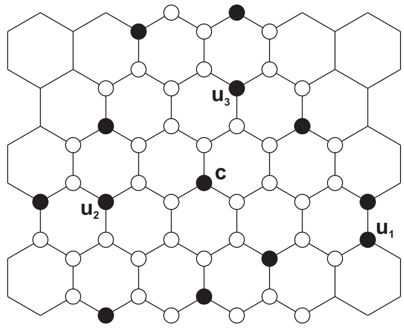


Figure 5.4: An example of the use of the shifting rules.

- In the rules 1, 2, 4 and 7, we shift 1/4 units of share from $\mathbf{u}$ to $\mathbf{v}$.
- In the rule 3, we shift 1/6 and 1/12 units of share from $\mathbf{u}$ to $\mathbf{v}$ and $\mathbf{v}'$, respectively.
- In the rule 5, we shift 1/6 units of share from $\mathbf{u}$ to $\mathbf{v}$.
- In the rules 6, 8, 9 and 10, we shift 1/12 units of share from $\mathbf{u}$ to $\mathbf{v}$.

We also have the following modifications to the previous rules:

- If in the rules 1, 2 and 7 we have $\mathbf{u} + (0, -1) \in C$, then we only shift 1/12 units of share from $\mathbf{u}$ to $\mathbf{v}$ and denote these new rules (respectively) by 1.1, 2.1 and 7.1. Moreover, in the rule 1.1, we shift 1/12 units of share to $\mathbf{v}$ whether $(-3, 2)$ belongs to C or not.
- If in the rule 1 we have $\mathbf{u} + (-3, 2) \in C$, then then we shift 1/4 units of share from $\mathbf{u}$ to $\mathbf{u} + (-1, 2)$ (no share is shifted to $\mathbf{v}$) and denote this new rule by 1.2.
- If in the rule 2 we have $\mathbf{u} + (-3, 1) \in C$, then we shift 1/6 units of share from $\mathbf{u}$ to $\mathbf{v}$ and denote this new rule by 2.2.
- If in the rule 2 we have $\mathbf{u} + (1, 2) \in C$, then we shift 1/12 units of share from $\mathbf{u}$ to $\mathbf{v}$ and denote this new rule by 2.3.

The modified share of a codeword $\mathbf{c} \in C$, which is obtained after the shifting scheme is applied, is denoted by $\bar{s}_2(\mathbf{c})$. The use of the rules is illustrated in the following example.

Example 5.3.1. Consider the codeword $\mathbf{c}$ with the surroundings as illustrated in Figure 5.4. The share of the codeword $\mathbf{c}$ is equal to 5. The rules 6, 8 and 9 apply to the codeword $\mathbf{c}$ and according to the rules 1/12 units of share is shifted from $\mathbf{c}$ to $\mathbf{u}_1$, $\mathbf{u}_2$ and $\mathbf{u}_3$, respectively. (Recall that reflections and rotations can be applied to the constellations in Figure 5.3.)

Hence, after the shifting scheme is applied, we have $\bar{s}_2(\mathbf{c}) = 19/4$ for the modified share. In order to ensure that also $\bar{s}_2(\mathbf{u}_i) \leq 19/4$ for any $i = 1, 2, 3$, we refer to the proofs of Lemmas 5.3.2 and 5.3.3.

5.3.2 The main theorem

The following three lemmas show that $\bar{s}_2(\mathbf{c}) \leq 19/4$ for all $\mathbf{c} \in C$. The proofs of the lemmas are postponed to Section 5.3.3.

Lemma 5.3.2. *Let $\mathbf{c} \in C$ be a codeword such that $\mathbf{c}$ is not adjacent to another codeword and share is shifted to $\mathbf{c}$ according to the previous rules. Then we have $\bar{s}_2(\mathbf{c}) \leq 19/4$.*

Lemma 5.3.3. *Let $\mathbf{c} \in C$ be a codeword such that $\mathbf{c}$ is adjacent to another codeword and share is shifted to $\mathbf{c}$ according to the previous rules. Then we have $\bar{s}_2(\mathbf{c}) \leq 19/4$.*

Lemma 5.3.4. *Let $\mathbf{c} \in C$ be a codeword such that no share is shifted to $\mathbf{c}$ according to the previous rules. Then we have $\bar{s}_2(\mathbf{c}) \leq 19/4$.*

As stated in the previous lemmas, we have $\bar{s}_2(\mathbf{c}) \leq 19/4$ for any $\mathbf{c} \in C$. Now we are ready to prove the main theorem of the paper.

Theorem 5.3.5. *If C is a 2-identifying code in the hexagonal grid G_H, then the density*

$$D(C) \geq \frac{4}{19}.$$

Proof. Assume that C is a 2-identifying code in G_H. Since each vertex $\mathbf{u} \in Q_{n-2}$ with $|I_2(\mathbf{u})| = i$ contributes the summand $1/i$ to $s_2(\mathbf{c})$ for each of the i codewords $\mathbf{c} \in B_2(\mathbf{u})$, we have

$$\sum_{\mathbf{c} \in C \cap Q_n} s_2(\mathbf{c}) \geq |Q_{n-2}|. \tag{5.1}$$

Furthermore, we have

$$\sum_{\mathbf{c} \in C \cap Q_n} s_2(\mathbf{c}) \leq \sum_{\mathbf{c} \in C \cap Q_n} \bar{s}_2(\mathbf{c}) + \frac{19}{4}|Q_{n+6} \setminus Q_n|, \tag{5.2}$$

because shifting shares inside Q_n does not affect the sum and each codeword in $Q_{n+6} \setminus Q_n$ can receive at most 19/4 units of share (by Lemmas 5.3.2 and 5.3.3). Notice also that codewords in Q_n cannot shift share to codewords outside Q_{n+6}. Therefore, combining the equations (5.1) and (5.2) with the fact that $\bar{s}_2(\mathbf{c}) \leq 19/4$ for any $\mathbf{c} \in C$, we obtain

$$\frac{|C \cap Q_n|}{|Q_n|} \geq \frac{4}{19} \cdot \frac{|Q_{n-2}|}{|Q_n|} - \frac{|Q_{n+6} \setminus Q_n|}{|Q_n|}.$$

Since $|Q_k| = (2k+1)^2$ for any positive integer k, it is easy to conclude from the previous inequality that the density $D(C) \geq 4/19$. □

5.3.3 The proofs of the lemmas

In what follows, we provide the proofs of Lemmas 5.3.2, 5.3.3 and 5.3.4.

Proof of Lemma 5.3.2. Notice first that share can be shifted to the codeword $\mathbf{c}$ only according to the rules 3, 7, 7.1, 8, 9 and 10 since $\mathbf{c}$ is not adjacent to another codeword. The main idea of the following proof is to show that $\mathbf{c}$ cannot receive share according to two different rules. In each case, this observation then straightforwardly implies the claim.

Assume first that $\mathbf{c}$ receives share according to the rule 10. Without loss of generality, we may assume that the rule is applied as in Figure 5.3(j) (when $\mathbf{c} = \mathbf{v}$). Now the vertices $\mathbf{c} + (-2, -1)$, $\mathbf{c} + (0, 1)$ and $\mathbf{c} + (2, -1)$ belong to the code C. Since this is not the case with the other rules (see Figure 5.3), they cannot be applied to $\mathbf{c}$. Moreover, choosing $D = \{\mathbf{c}, \mathbf{c} + (-2, -1), \mathbf{c} + (0, 1), \mathbf{c} + (2, -1)\}$ in Lemma 5.2.1, we have $s_2(\mathbf{c}) \leq 9/2$. Thus, since according to the rule 10 share can be shifted to $\mathbf{c}$ only from $\mathbf{c}+(-2, -1)$, $\mathbf{c} + (0, 1)$ and $\mathbf{c} + (2, -1)$ and at most once for each of these codewords, we obtain that $\bar{s}_2(\mathbf{c}) \leq s_2(\mathbf{c}) + 3 \cdot 1/12 \leq 19/4$.

Assume then that share is shifted to $\mathbf{c}$ according to the rule 9. First of all, by the previous paragraph, the rule 10 cannot be applied to $\mathbf{c}$. Since now we have $I_2(\mathbf{c}) = \{\mathbf{c}\}$, it is immediate that $\mathbf{c}$ cannot receive share according to the rules 3, 7, 7.1 or 8. Moreover, it is easy to see that the rule 9 can be applied only once. Therefore, since we have $s_2(\mathbf{c}) \leq 14/3$ by Lemma 5.2.1, we obtain that $\bar{s}_2(\mathbf{c}) \leq s_2(\mathbf{c}) + 1/12 \leq 19/4$.

Assume that $\mathbf{c}$ receives share according to the rule 8 and that the rule is used as in Figure 5.3(h). Now, if the rule 7 or 7.1 was used, then there would exist two codewords in $B_2(\mathbf{c})$ such that the distance between them is equal to 4. By the constellation of the rule 8, this is impossible. Let us then show that neither the rule 3 can be used. Assume to the contrary that share is shifted to $\mathbf{c}$ according to the rule 3. Since there is a pair of adjacent codewords in the constellation of the rule 3, the vertex $\mathbf{c} + (0, 1)$ belongs to C. Furthermore, we have either $\mathbf{c} + (-1, 1) \in C$ or $\mathbf{c} + (1, 1) \in C$ (but not both). If $\mathbf{c} + (-1, 1) \in C$, then we have a contradiction since share is shifted from $\mathbf{c} + (3, 0)$, which is not a codeword in the constellation of the rule 8. On the other hand, if $\mathbf{c} + (1, 1) \in C$, then share is received from $\mathbf{c} + (-3, 0)$. This again leads to a contradiction since $\mathbf{c} + (-2, -1) \in C$. In conclusion, only the rule 8 can be applied to $\mathbf{c}$. Moreover, the rule 8 can be used at most once. If $\mathbf{c} + (-1, 1) \in C$ or $\mathbf{c} + (1, 1) \in C$, then $s_2(\mathbf{c}) \leq 55/12$ or $s_2(\mathbf{c}) \leq 23/6$ by Lemma 5.2.1, respectively. Hence, we have $\bar{s}_2(\mathbf{c}) \leq s_2(\mathbf{c}) + 1/12 \leq 56/12 \leq 19/4$.

Assume that $\mathbf{c}$ receives share according to the rule 3 and that the rule is used as in Figure 5.3(c) (when $\mathbf{c} = \mathbf{v}'$). By the previous considerations, we know that share cannot be shifted to $\mathbf{c}$ according to the rules 8, 9 and 10.

Let us then show that neither the the rules 7 or 7.1 can be applied to $\mathbf{c}$. Assume to the contrary that $\mathbf{c}$ receives share according to the rule 7. Now $\mathbf{c}$ can receive share only from the vertices $\mathbf{c}+(-1,-2)$, $\mathbf{c}+(1,-2)$, $\mathbf{c}+(3,0)$, $\mathbf{c}+(2,1)$, $\mathbf{c}+(-3,0)$ and $\mathbf{c}+(-2,1)$. Now we have the following observations:

- Since $\mathbf{c}+(-1,-2) \notin C$ share cannot be shifted from $\mathbf{c}+(-1,-2)$ and $\mathbf{c}+(1,-2)$.
- Since $\mathbf{c}+(-2,0) \in C$ share cannot be shifted from $\mathbf{c}+(-3,0)$ and $\mathbf{c}+(-2,1)$.
- Since $\mathbf{c}+(-1,-1) \notin C$ and $\mathbf{c}+(1,-1) \notin C$ share cannot be shifted from $\mathbf{c}+(3,0)$ and $\mathbf{c}+(2,1)$.

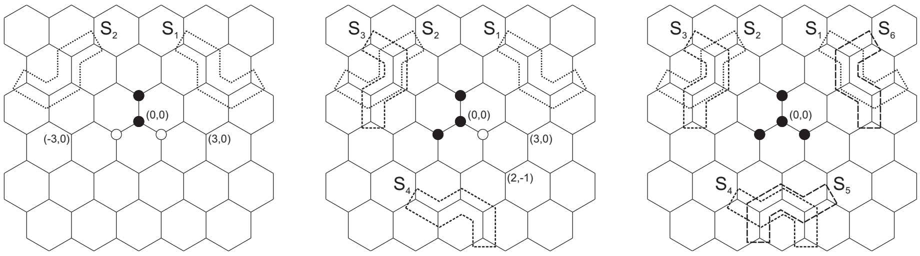
A contradiction now follows from these facts. Hence, only the rule 3 can be applied to $\mathbf{c}$. Moreover, it is easy to see that share can be shifted to $\mathbf{c}$ at most twice according to the rule 3. Therefore, since we have $s_2(\mathbf{c}) \leq 9/2$ by Lemma 5.2.1, we obtain that $\bar{s}_2(\mathbf{c}) \leq s_2(\mathbf{c}) + 2 \cdot 1/12 \leq 19/4$.

Finally, assume that $\mathbf{c}$ receives share according to the rule 7 or 7.1 and that the rule is used as in Figure 5.3(g). As shown above, other rules cannot be applied to $\mathbf{c}$. Moreover, the rules 7 and 7.1 can be applied to $\mathbf{c}$ only once. Furthermore, if $\mathbf{c}+(2,0) \in C$ or $\mathbf{c}+(1,1) \in C$, then we have $s_2(\mathbf{c}) \leq 23/6$ or $s_2(\mathbf{c}) \leq 53/12$ by Lemma 5.2.1, respectively. Thus, we have $\bar{s}_2(\mathbf{c}) \leq s_2(\mathbf{c}) + 1/4 \leq 14/3 \leq 19/4$. □

Proof of Lemma 5.3.3. Since $\mathbf{c}$ is adjacent to another codeword, it is immediate that $\mathbf{c}$ can receive share only according to the rules 1, 1.1, 1.2, 2, 2.1, 2.2, 2.3, 3, 4, 5 and 6. The proof of the lemma is now divided into three cases depending on the number of codewords adjacent to $\mathbf{c}$.

1) Assume first that $\mathbf{c}$ is adjacent to exactly one codeword. Without loss of generality, we may assume that $\mathbf{c}=(0,0)$ and the adjacent codeword is $(0,1)$. Hence, we have $(-1,0) \notin C$ and $(1,0) \notin C$. Now the only possibilities for $\mathbf{c}$ to receive share is from the vertices $(-3,0)$ or $(3,0)$ (the rule 1.2) and from the vertices that belong to $S_1 = \{(5,1),(4,1),(3,1),(3,2),(2,2)\}$ and $S_2 = \{(-5,1),(-4,1),(-3,1),(-3,2),(-2,2)\}$. These observations are illustrated in Figure 5.5(a).

Consider then more closely the set S_1. In what follows, we show that at most $1/4$ units of share is shifted from the vertices of S_1 to $\mathbf{c}$. Notice that if share is shifted only from one vertex of S_1, then we are immediately done. Observe then that if $\mathbf{u} \in S_1$ shifts share to $\mathbf{c}$ according to the rules 1–6 or their modifications, then we have $I_2(\mathbf{u}) = \{\mathbf{u}\}$. Therefore, if two vertices of S_1 shift share, then one of these vertices is $(4,1)$ or $(5,1)$. Assume that share is shifted from $(4,1)$ according to the rule 5 (see Figure 5.3(e)). Now we have $C \cap S_1 = \{(2,2),(4,1)\}$ and $(4,3) \in C$. Therefore, since at most $1/12$



(a) The first case (b) The second case (c) The third case

Figure 5.5: The cases of the proof of Lemma 5.3.3 illustrated.

units of share is shifted from $(2,2)$ according to the rule 2.1, we obtain that no more than $1/6 + 1/12 = 1/4$ units of share can be shifted from S_1 to $\mathbf{c}$. Similarly, if share is shifted from $(5,1)$ according to the rule 6, then it can be shown that $\mathbf{c}$ receives at most $1/4$ units of share from S_1. In conclusion, at most $1/4$ units of share is shifted from the vertices of S_1 to $\mathbf{c}$. Analogously, this statement also holds for the vertices of S_2.

Assume that the rule 1.2 is used. (Clearly, this rule can be used only once.) Without loss of generality, we may assume that $(-3,0) \in C$ and $(1,-1) \in C$. Therefore, since we can choose $D = \{\mathbf{c}, (0,1), (-3,0), (1,-1)\}$ in Lemma 5.2.1, we obtain that $s_2(\mathbf{c}) \leq 15/4$. Thus, since the codewords in each of the sets S_1 and S_2 can shift at most $1/4$ units of share to $\mathbf{c}$, we have $\bar{s}_2(\mathbf{c}) \leq s_2(\mathbf{c}) + 3 \cdot 1/4 \leq 9/2$. Assume then that the rule 1.2 cannot be applied to $\mathbf{c}$. Since $\mathbf{c}$ and $(0,1)$ are 2-separated by C, there exists at least one codeword in the symmetric difference $B_2(\mathbf{c}) \triangle B_2(0,1)$. Thus, without loss of generality, we may assume that $(1,-1) \in C$, $(1,2) \in C$, $(2,0) \in C$ or $(2,1) \in C$. The first part of the proof is then concluded by the following four cases:

- Assume that $(1,-1) \in C$. Now, by the first case of Example 5.2.2, we know that $s_2(\mathbf{c}) \leq 17/4$. Therefore, since share is shifted to $\mathbf{c}$ only from the vertices of the sets S_1 and S_2, we have $\bar{s}_2(\mathbf{c}) \leq s_2(\mathbf{c}) + 2 \cdot 1/4 \leq 19/4$.

- Assume that $(1,2) \in C$. It is straightforward to verify that share can be shifted to $\mathbf{c}$ only from $(-3,1)$, $(-2,2)$, $(-3,2)$, $(-4,1)$ and $(-5,1)$ according to the rules 1.1, 2.3, 3, 5 and 6, respectively. Moreover, it is easy to see that at most one of these rules can be used (and only once). Thus, the codeword $\mathbf{c}$ receives at most $1/6$ units of share. Hence, we have $\bar{s}_2(\mathbf{c}) \leq s_2(\mathbf{c}) + 1/6 \leq 19/4$ since $s_2(\mathbf{c}) \leq 55/12$ by the first case of Example 5.2.2 (and obvious symmetrical argument).

- Assume that $(2,0) \in C$. Consider then more closely the vertices of S_1. Now it is easy to conclude that the vertices $(5,1)$, $(4,1)$ and $(3,1)$

cannot shift share to $\mathbf{c}$. Hence, only either $(3, 2)$ according to the rule 3 or $(2, 2)$ according to the rule 2.2 (but not both) is capable of shifting share to $\mathbf{c}$. In each case, $\mathbf{c}$ receives at most $1/6$ units of share. By the second case of Example 5.2.2, we know that $s_2(\mathbf{c}) \leq 13/3$. Therefore, since at most $1/4$ units of share is received from S_2, we have $\bar{s}_2(\mathbf{c}) \leq s_2(\mathbf{c}) + 1/4 + 1/6 \leq 19/4$.

- Assume that $(2, 1) \in C$. By the second case of Example 5.2.2 (and symmetry), we obtain that $s_2(\mathbf{c}) \leq 9/2$. Since now share can be shifted to $\mathbf{c}$ only from S_2, we obtain that $\bar{s}_2(\mathbf{c}) \leq s_2(\mathbf{c}) + 1/4 \leq 19/4$.

2) Assume that $\mathbf{c}$ is adjacent to exactly two codewords. Without loss of generality, we may assume that $\mathbf{c} = (0, 0)$ and that the adjacent codewords are $(-1, 0)$ and $(0, 1)$. Let then S_3 and S_4 be sets which are obtained by rotating respectively the sets S_1 and S_2 by $2\pi/3$ (counter-clockwise in the honeycomb representation) around the origin. Again the codewords in each of these sets S_i can shift at most $1/4$ units of share to $\mathbf{c}$. In addition to the previous ones, at most $1/4$ units of share can also be shifted to $\mathbf{c}$ from either $(2, -1)$ or $(3, 0)$ (but not both) according to the rule 1.2. These observations are illustrated in Figure 5.5(b).

Assume first that share is shifted to $\mathbf{c}$ according to the rule 1.2. Without loss of generality, we may assume that $\mathbf{c}$ receives share from the vertex $(3, 0)$. Then we immediately have $(3, 0) \in C$ and $(-1, -1) \in C$. Therefore, we have $s_2(\mathbf{c}) \leq 15/4$ by Lemma 5.2.1. Furthermore, since $(-1, -1) \in C$, the codeword $\mathbf{c}$ does not receive share from the set S_4. Thus, we have $\bar{s}_2(\mathbf{c}) \leq s_2(\mathbf{c}) + 4 \cdot 1/4 \leq 19/4$.

Assume then that the rule 1.2 is not used. Since the vertices $\mathbf{c}$ and $(0, 1)$ are 2-separated by C, there exists a codeword in $B_r(\mathbf{c}) \triangle B_r(0, 1)$. If $(1, -1) \in C$ or $(2, 0) \in C$, then $s_2(\mathbf{c}) \leq 53/15$ (by Lemma 5.2.1) and we are immediately done since at most 1 unit of share can be shifted to $\mathbf{c}$ from the union of the sets S_i. If $(-2, 1) \in C$, then $s_2(\mathbf{c}) \leq 21/5$ and we are done since share cannot be shifted to $\mathbf{c}$ from S_2 and S_3. If $(-1, -1) \in C$ or $(-2, 0) \in C$, then $s_2(\mathbf{c}) \leq 79/20$ (by Lemma 5.2.1) and we are again done (since share is not shifted from S_3). Hence, we may assume that $(-1, 2) \in C$, $(1, 2) \in C$ or $(2, 1) \in C$. Analogously, it can also be assumed that $(-3, 0) \in C$, $(-2, -1) \in C$ or $(0, -1) \in C$ since $\mathbf{c}$ and $(-1, 0)$ are 2-separated by C. Thus, at most two of the sets S_i can shift share to $\mathbf{c}$. Therefore, since $s_2(\mathbf{c}) \leq 17/4$ (choose $D = \{\mathbf{c}, (-1, 0), (0, 1)\}$ in Lemma 5.2.1), we have $\bar{s}_2(\mathbf{c}) \leq s_2(\mathbf{c}) + 2 \cdot 1/4 \leq 19/4$.

3) Finally, assume that all the vertices adjacent to $\mathbf{c}$ are codewords, i.e. $(-1, 0) \in C$, $(0, 1) \in C$ and $(1, 0) \in C$ (see Figure 5.5(c)). Notice that now the rule 1.2 cannot be used. Since there again exists a codeword in $B_r(\mathbf{c}) \triangle B_r(0, 1)$, it is easy to conclude as above that at most $5 \cdot 1/4$ units

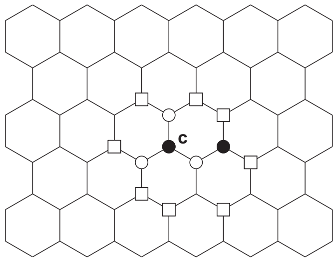


Figure 5.6: The symmetric difference $B_2(\mathbf{c}) \triangle B_2(1,0)$ consists of the squared vertices. At least one of these vertices is a codeword.

of share is shifted to $\mathbf{c}$. Therefore, since $s_2(\mathbf{c}) \leq 67/20$ by Lemma 5.2.1, we have $\bar{s}_2(\mathbf{c}) \leq s_2(\mathbf{c})+5/4 \leq 19/4$. This completes the proof of the lemma. $\square$

Proof of Lemma 5.3.4. Without loss of generality, we may assume that $\mathbf{c} = (0,0)$. Assume first that $|I_2(\mathbf{c})| \geq 2$. If now $\mathbf{c}$ is adjacent to another codeword, then $\bar{s}_2(\mathbf{c}) \leq s_2(\mathbf{c}) \leq 19/4$ (by Lemma 5.2.1). Hence, we may assume that $(-1,0),(1,0),(0,1) \notin C$. Let then $(2,0)$ be a codeword of C. Since the vertices $\mathbf{c}$ and $(1,0)$ are 2-separated by C, there is at least one codeword in the symmetric difference $B_2(\mathbf{c}) \triangle B_2(1,0)$ (see Figure 5.6). Therefore, by Lemma 5.2.1, it is straightforward (albeit tedious) to verify that in all the possible cases $\bar{s}_2(\mathbf{c}) \leq s_2(\mathbf{c}) \leq 19/4$. Indeed, we can choose in Lemma 5.2.1 the set D to consist of the vertices $\mathbf{c}$, $(2,0)$ and a codeword in the symmetric difference $B_2(\mathbf{c}) \triangle B_2(1,0)$.

From now on, we may assume that $I_2(\mathbf{c}) = \{\mathbf{c}\}$. In what follows, we use the notations: $A_1 = \{(-1,1),(0,1),(1,1)\}$, $A_2 = \{(-2,0),(-1,0),(-1,-1)\}$, $A_3 = \{(1,-1),(1,0),(2,0)\}$, $A_1' = \{(-1,2),(1,2)\}$, $A_2' = \{(-3,0),(-2,-1)\}$ and $A_3' = \{(2,-1),(3,0)\}$. These sets are illustrated in Figure 5.7. The proof of the lemma now divides into three cases depending on the number of codewords in the set $\{(-2,1),(2,1),(0,-1)\}$

1) Assume first that $(-2,1),(2,1),(0,-1) \notin C$. Since the vertices $\mathbf{c}$, $(-1,0)$, $(1,0)$ and $(0,1)$ are 2-separated by C, each one of the sets A_1', A_2' and A_3' contains at least one codeword. Hence, each of the sets A_1, A_2 and A_3 contains a vertex whose I-set contains at least three codewords. Indeed, if for example $(-1,2) \in C$, then $(-1,1)$ or $(0,1)$ is such a vertex in A_1. Therefore, we obtain that $s_2(\mathbf{c}) \leq 1+6\cdot 1/2+3\cdot 1/3 = 5$.

Consider then the set A_1' that contains at least one codeword as stated above. Assume first that both $(-1,2) \in C$ and $(1,2) \in C$. If the vertex $(0,2)$ also belongs to C, then the I-sets of all the vertices in A_1 have size at least 3. Then there are at least 5 vertices in $B_2(\mathbf{c})$ which are 2-covered by at least 3 codewords. Therefore, we have $\bar{s}_2(\mathbf{c}) \leq s_2(\mathbf{c}) \leq 1+4\cdot 1/2+5\cdot 1/3 = 14/3$ (and we are done). Hence, suppose that $(0,2) \notin C$. Furthermore, assume first

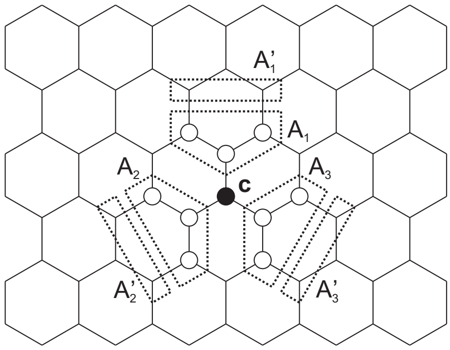
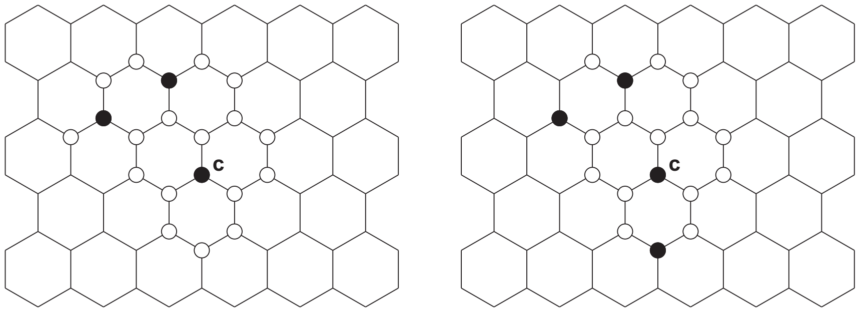


Figure 5.7: The sets A_1, A_2, A_3, A_1', A_2' and A_3' illustrated.

that $(-2,2) \in C$. If also $(-3,1) \in C$, then $|I_2(0,1)| \geq 3$ and $|I_2(-1,1)| \geq 4$. Thus, we have $\bar{s}_2(\mathbf{c}) \leq s_2(\mathbf{c}) \leq 1+5\cdot 1/2+3\cdot 1/3+1/4 = 19/4$. On the other hand, if $(-3,1) \notin C$, then the rule 2.3 can be applied to $\mathbf{c}$ and we obtain that $\bar{s}_2(\mathbf{c}) \leq s_2(\mathbf{c})-1/12 \leq 29/6-1/12 \leq 19/4$. Hence, we may assume that $(-2,2) \notin C$ and $(2,2) \notin C$ (by symmetry). Since the vertices $(-1,2)$ and $(1,2)$ are 2-separated by C, at least one of the vertices $(-3,2)$, $(-2,3)$, $(2,3)$ and $(3,2)$ belongs to C. Hence, we can shift at least 1/4 units of share from $\mathbf{c}$ according to the rule 7. Therefore, we have $\bar{s}_2(\mathbf{c}) \leq s_2(\mathbf{c}) - 1/4 \leq 19/4$.

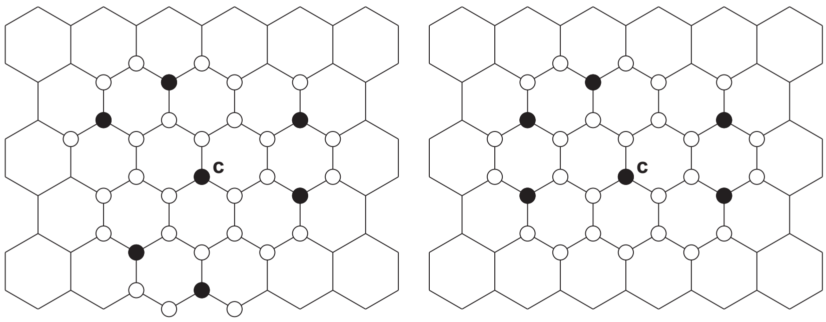
By the considerations above, we may without loss of generality assume that $(-1,2) \in C$ and $(1,2) \notin C$. If $(0,2) \in C$, then 1/4 units of share can be shifted from $\mathbf{c}$ according to the rule 1 or 1.2, and we are done. Thus, suppose that $(0,2) \notin C$. Assume then that $(-2,2) \in C$. If the rule 2 applies to $\mathbf{c}$, then we are immediately done (1/4 units of share is shifted from $\mathbf{c}$). On the other hand, if $(-3,1) \in C$, then by the fact that $|I_2(-1,1)| \geq 4$ we have $s_2(\mathbf{c}) \leq 59/12$ and we are again done since at least 1/6 units of share is shifted from $\mathbf{c}$ according to the rule 2.2. Hence, assume that $(-2,2) \notin C$. Since $(0,1)$ and $(-1,1)$ are 2-separated by C, the vertex $(-3,1)$ belongs to C. Now, if $(-3,2) \in C$, then 1/4 units of share can be shifted from $\mathbf{c}$ according to the rule 3 and we are done. Thus, we may assume that $(-3,2) \notin C$. If $(-4,1) \in C$ and $(-3,0) \notin C$, then the rule 4 can be applied to $\mathbf{c}$ and we are done. Furthermore, if $(-4,1) \in C$, and $(-3,0) \in C$, then instead of the set A_1' consider the set A_3'. Repeating the previous arguments for the set A_3', we obtain that $\bar{s}_2(\mathbf{c}) \leq 19/4$ also in this case. Thus, we may assume that $(-4,1) \notin C$. The surroundings of the codeword $\mathbf{c}$ (obtained above) is illustrated in Figure 5.8(a).

The previous reasoning also applies when we consider the sets A_2' and A_3' instead of A_1'. This leads straightforwardly to the observation that we have only two different neighbourhoods of $\mathbf{c}$ (up to rotations and reflections). These neighbourhoods are illustrated in Figure 5.9.



(a) The first case (b) The second case

Figure 5.8: The surroundings of the codeword **c** illustrated in (a) the first and (b) the second part of the proof of Lemma 5.3.4.



(a) (b)

Figure 5.9: Two cases of the proof of Lemma 5.3.4 illustrated.

Consider first the case in Figure 5.9(a). In what follows, we show that $\mathbf{c}$ shifts $1/12$ units of share to $(-1,2)$ or $(1,3)$, or that we originally have at least 2 vertices in A_1 such that their I-sets have at least 3 codewords. In both cases, the (maximum) share of $\mathbf{c}$ is reduced by at least $1/12$. (Actually, in the latter case, the share is reduced by $1/6$.) This observation can then be generalized to the (other two) symmetrical cases implying that $\bar{s}_2(\mathbf{c}) \leq 5 - 3 \cdot 1/12 = 19/4$. If $(2,2) \in C$, then $(-1,1) \in A_1$ and $(1,1) \in A_1$ are 2-covered by three codewords. Hence, we may assume that $(2,2) \notin C$. If we can shift $1/12$ units of share from $\mathbf{c}$ to $(-1,2)$ according to the rule 8, then we are immediately done. Therefore, we may assume that $(-2,3) \notin C$ and $(0,3) \notin C$. Since $(-1,2)$ and $(1,2)$ are 2-separated by C, the vertex $(2,3)$ belongs to C. Furthermore, since $(-1,2)$ and $(0,2)$ are 2-separated by C, at least one of the vertices $(-1,3)$ and $(1,3)$ is a codeword. Therefore, at least $1/12$ units of share can shifted from $\mathbf{c}$ according to the rule 6 or 9.

Consider then the case in Figure 5.9(b). Let us now show that $\mathbf{c}$ shifts at least $1/6$ units of share to $(-1,-2)$ or $(1,-2)$, or that we originally have more than 2 vertices in $A_2 \cup A_3$ such that their I-sets are at least of size 3. In both cases, the (maximum) share of $\mathbf{c}$ is reduced by at least $1/6$. This result together with the observation in the previous paragraph, then implies that $\bar{s}_2(\mathbf{c}) \leq 5 - 1/6 - 1/12 = 19/4$. If now $(0,-2) \in C$, then we know that $(-2,0) \in A_2$, $(2,0) \in A_3$, and at least one of the vertices $(-1,-1) \in A_2$ and $(1,-1) \in A_3$ are such that their I-sets are at least of size 3. Hence, we may assume that $(0,-2) \notin C$. Since $\mathbf{c}$, $(-1,-1)$ and $(1,-1)$ are 2-separated by C, the vertices $(-2,-2)$ and $(2,-2)$ belong to C. Furthermore, since $(0,-1)$ is 2-covered by a codeword of C, we have $(-1,-2) \in C$ or $(1,-2) \in C$. Therefore, we can shift at least $1/6$ units of share to $(-1,-2)$ or $(1,-2)$ according to the rule 5. In conclusion, if $\mathbf{c}$ is a codeword such that $I_2(\mathbf{c}) = \{\mathbf{c}\}$ and $(-2,1), (2,1), (0,-1) \notin C$, then we have $\bar{s}_2(\mathbf{c}) \leq 19/4$.

2) Assume then that $\mathbf{c}$ is a codeword such that $I_2(\mathbf{c}) = \{\mathbf{c}\}$, and exactly one of the vertices $(-2,1)$, $(2,1)$ and $(0,-1)$ is a codeword. Without loss of generality, we may assume that $(-2,1) \notin C$, $(2,1) \notin C$ and $(0,-1) \in C$. Now, by considering the vertices $(-1,-1) \in A_2$, $(-1,0) \in A_2$, $(1,-1) \in A_3$ and $(1,0) \in A_3$, we obtain that there are at least 3 vertices in $A_2 \cup A_3$ which are 2-covered by at least 3 codewords. Thus, since there is also one such vertex in A_1, we have $s_2(\mathbf{c}) \leq 1 + 5 \cdot 1/2 + 4 \cdot 1/3 = 29/6$. Assume first that both $(-1,2)$ and $(1,2)$ belong to C. If $(-2,2) \in C$, $(0,2) \in C$ or $(2,2) \in C$, there are at least 2 vertices in A_1 with the I-set of size at least 3 and, therefore, we have $s_2(\mathbf{c}) \leq 19/4$. Hence, according to the rule 7.1, $1/12$ units of share can be shifted from $\mathbf{c}$ (since $(-1,2)$ and $(1,2)$ are 2-separated by C) and we are done. Thus, without loss of generality, we may assume that $(-1,2) \in C$ and $(1,2) \notin C$. If now $(-3,1) \in C$ and $(-2,2) \in C$, then

$|I_2(-1,1)| \geq 4$ and there is no problem since $s_2(\mathbf{c}) \leq 1+5\cdot 1/2+3\cdot 1/3+1/4 = 19/4$. Furthermore, if the rules 1.1 or 2.1 can be used, then we are again done. Thus, we may assume that $(-2,2) \notin C$ and $(0,2) \notin C$. Therefore, since $(-1,1)$ and $(0,1)$ are 2-separated by C, we obtain that $(-3,1) \in C$. We have now arrived at the constellation illustrated in Figure 5.8(b). If now one of the vertices $(-4,0)$, $(-3,0)$, $(-3,-1)$, $(-2,-1)$, $(0,-2)$ and $(2,-1)$ belongs to C, then by Lemma 5.2.1 we immediately have $s_2(\mathbf{c}) \leq 19/4$. Hence, we may assume that none of these vertices belongs to C. Thus, since $(-1,0)$, $(-1,-1)$ and $(1,0)$ are 2-separated by C, we have $(-2,-2) \in C$ and $(3,0) \in C$. If now $(3,-1) \in C$ or $(4,0) \in C$, then again $s_2(\mathbf{c}) \leq 19/4$ and we are done. Therefore, since $(-1,0)$ and $(1,-1)$ are 2-separated by C, we have $(2,-2) \in C$ and $1/12$ units of share can be shifted from $\mathbf{c}$ to $(0,-1)$ according to the rule 10. Hence, we have $\bar{s}_2(\mathbf{c}) \leq 19/4$.

3) Finally, assume that $\mathbf{c}$ is a codeword such that $I_2(\mathbf{c}) = \{\mathbf{c}\}$, and that at least two of the vertices $(-2,1)$, $(2,1)$ and $(0,-1)$ belong to C. Then, by Lemma 5.2.1, we have $s_2(\mathbf{c}) \leq 14/3$. This observation completes the proof of the lemma. $\square$

Chapter 6

Identification in $\mathbb{Z}^2$ using Euclidean balls

In this chapter, which is based on the papers [55] and [56], we consider identification in the infinite grid $\mathbb{Z}^2$ with Euclidean balls. In Section 6.1, we first present some preliminary definitions and results. Then, in Section 6.2, we proceed by introducing some code constructions and lower bounds. Finally, we end the chapter with a discussion on identifying codes in the king grid with slightly modified balls.

6.1 Preliminaries

Let $G = (V, E)$ be a simple, connected and undirected graph. Let r be a non-negative integer. Previously, we have denoted a ball of radius r centred at a vertex $u \in V$ simply by $B_r(u)$ since there has not been a reason to specify the underlying graph in the notation. However, in this chapter, we sometimes need to specify the underlying graph. Therefore, we more precisely write $B_r(u) = B_r(G; u)$ if there is a possibility of confusion. Furthermore, if r is equal to 1, then we write in short as follows: $B_1(G; u) = B(G; u) = B(u)$. We use analogous notations also for the identifying sets $I_r(G, C; u)$.

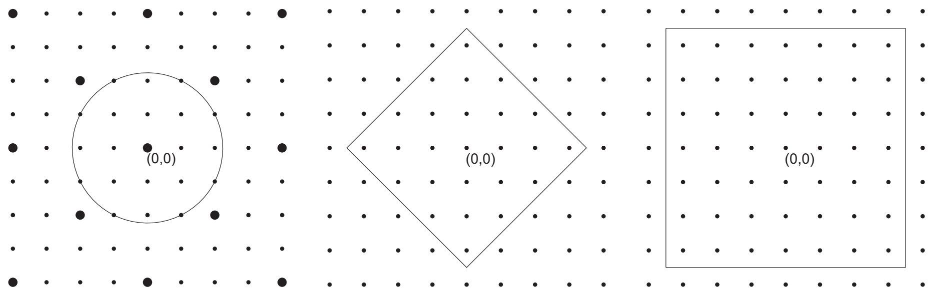
Assume now that the vertex set V is equal to $\mathbb{Z}^2$. Let then t be a positive integer and $\mathbf{u} = (x, y)$ be a vertex in $\mathbb{Z}^2$. The graph $\mathcal{S}_t$ with the ball (i.e. the closed neighbourhood)

$$B(\mathcal{S}_t; \mathbf{u}) = \{(x', y') \in \mathbb{Z}^2 \mid |x - x'| + |y - y'| \leq t\}$$

is called the *square grid*. The graph $\mathcal{K}_t$ with the ball

$$B(\mathcal{K}_t; \mathbf{u}) = \{(x', y') \in \mathbb{Z}^2 \mid |x - x'| \leq t, |y - y'| \leq t\}$$

is called the *king grid*. The graphs $\mathcal{S}_t$ and $\mathcal{K}_t$ are illustrated in Figure 6.1. Notice that now we have $B_t(\mathcal{S}_1; u) = B_1(\mathcal{S}_t; u)$ and $B_t(\mathcal{K}_1; u) = B_1(\mathcal{K}_t; u)$.



(a) $\mathcal{E}_{\sqrt{5}}$ (b) $\mathcal{S}_3$ (c) $\mathcal{K}_3$

Figure 6.1: (a) The ball $B(\mathcal{E}_{\sqrt{5}}; (0,0))$ and the code C_2 (defined in Section 6.3) illustrated. (b) The ball $B(\mathcal{S}_3; (0,0))$ illustrated. (c) The ball $B(\mathcal{K}_3; (0,0))$ illustrated.

Therefore, also the t-identifying codes in $\mathcal{S}_1$ and $\mathcal{K}_1$ coincide with the 1-identifying codes in $\mathcal{S}_t$ and $\mathcal{K}_t$, respectively.

Recall that identifying codes in $\mathcal{S}_t$ and $\mathcal{K}_t$ have been studied, for example, in [14,45,52] and [15,44], respectively. Notice that in these graphs for larger values of t the shape of the ball $B(u)$ resembles a square as can be seen in Figure 6.1. However, this kind of behaviour is not always desired. For example, when applying identifying codes to sensor networks (see [69,73]), it would seem to be more natural for a sensor to check an area with the shape of a Euclidean ball. Hence, in this chapter, we focus on identifying codes in grid graphs, where Euclidean balls are used to define the edges of the graphs.

For the rest of the chapter, assume that r is a positive real number. Let again $V = \mathbb{Z}^2$. The graph $\mathcal{E}_r = (V, E)$ is defined by the edge set E for which the vertices $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{Z}^2$ are adjacent if the Euclidean distance of $\mathbf{u}$ and $\mathbf{v}$ is at most r. Moreover, if the vertices $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{F}^n$ are adjacent, then we say that $\mathbf{u}$ *covers* $\mathbf{v}$. If $\mathbf{u} = (x, y) \in \mathbb{Z}^2$, then we have

$$B(\mathcal{E}_r; \mathbf{u}) = \{(x', y') \in \mathbb{Z}^2 \mid (x - x')^2 + (y - y')^2 \leq r^2\}.$$

Obviously, we have $\mathcal{S}_1 = \mathcal{E}_1$, $\mathcal{K}_1 = \mathcal{E}_{\sqrt{2}}$, $\mathcal{S}_2 = \mathcal{E}_2$ and $\mathcal{K}_2 = \mathcal{E}_{2\sqrt{2}}$. Furthermore, the graph $\mathcal{E}_{\sqrt{5}}$ is illustrated in Figure 6.1.

In the sequel, we will need the following result from [14, Proposition 1].

Theorem 6.1.1 ([14]). *Let $G = (V, E)$ be a simple, connected and undirected graph. Let $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3 \in V$ be three vertices of G and C be an identifying code in G. Then the set*

$$H(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3) = (B(\mathbf{u}_1) \triangle B(\mathbf{u}_2)) \cup (B(\mathbf{u}_1) \triangle B(\mathbf{u}_3)) \cup (B(\mathbf{u}_2) \triangle B(\mathbf{u}_3))$$

contains at least two codewords.

6.2 Identifying codes in the graphs $\mathcal{E}_r$

In what follows, we construct a 1-identifying code for the graph $\mathcal{E}_r$, when $r \geq 1$ is an arbitrary real number, and also provide a lower bound on the density of such codes. For the considerations, we define the *horizontal line* as $L_i^{(h)} = \{(x', i) \mid x' \in \mathbb{Z}\}$ and the *vertical line* as $L_i^{(v)} = \{(i, y') \mid y' \in \mathbb{Z}\}$, where i is an integer. If $\mathbf{u}$ is a vertex in $\mathbb{Z}^2$ and X is a subset of $\mathbb{Z}^2$, then the *sum* of $\mathbf{u}$ and X is defined as $\mathbf{u} + X = \{\mathbf{u} + \mathbf{v} \mid \mathbf{v} \in X\}$. We first present the following lemma.

Lemma 6.2.1. *Let $\mathbf{u} = (x, y)$ be a vertex in $\mathbb{Z}^2$ and $r \geq 1$ be a real number. In $B(x, y) \setminus B(x, y-1)$ there exist $2\lfloor r \rfloor + 1$ vertices, which lie on consecutive vertical lines $L_i^{(v)}$ with $i = x - \lfloor r \rfloor, \ldots, x + \lfloor r \rfloor$.*

Proof. Moving the center $\mathbf{u} = (x, y)$ of a ball to $(x, y-1)$ means that for $i = x - \lfloor r \rfloor, \ldots, x + \lfloor r \rfloor$ the vertex $\mathbf{u}$ covers on $L_i^{(v)}$ exactly one vertex of $\mathbb{Z}^2$ which is not covered by $(x, y-1)$, since the second coordinate decreases by one. The claim immediately follows from this observation. $\square$

Notice that analogous results to the previous lemma hold when the considered pattern is rotated by $\pi/2$, π and $3\pi/2$. For example, when the pattern of the lemma is rotated counter-clockwise by $\pi/2$, we obtain that the set $B(x, y) \setminus B(x+1, y)$ contains vertices on $2\lfloor r \rfloor + 1$ consecutive horizontal lines.

For the construction of the identifying codes in $\mathcal{E}_r$, we first introduce the following sets of vertices

$$C^{(h)} = \{(j, 0) \in \mathbb{Z}^2 \mid j \equiv 0 \bmod 2\} \text{ and } C^{(v)} = \{(0, j) \in \mathbb{Z}^2 \mid j \equiv 0 \bmod 2\}.$$

Then define a code C_k as follows:

$$\begin{aligned} C_k = &\bigcup_{i \in \mathbb{Z}} \Big((C^{(h)} + (0, i \cdot 2k)) \cup (C^{(h)} + (1, k + i \cdot 2k)) \Big) \\ &\bigcup_{i \in \mathbb{Z}} \Big((C^{(v)} + (i \cdot 2k, 0)) \cup (C^{(v)} + (k + i \cdot 2k, 1)) \Big), \end{aligned}$$

where $k \in \mathbb{Z}$ and $k \geq 1$. The following theorem shows that the previous code C_k provides a 1-identifying code for the graph $\mathcal{E}_r$.

Theorem 6.2.2. *Let $r \geq 1$ be a real number.*

(i) If $r^2 - \lfloor r \rfloor^2 \geq 1$, then the code $C_{2\lfloor r \rfloor + 1}$ is identifying in $\mathcal{E}_r$.

(ii) If $r^2 - \lfloor r \rfloor^2 < 1$, then the code $C_{2\lfloor r \rfloor}$ is identifying in $\mathcal{E}_r$.

Proof. (i) Let $\mathbf{u} = (x, y)$ be a vertex in $\mathbb{Z}^2$. Assume first that $r^2 - \lfloor r \rfloor^2 \geq 1$. This assumption implies that the vertices $(x - \lfloor r \rfloor, y - 1)$, $(x - \lfloor r \rfloor, y + 1)$, $(x + \lfloor r \rfloor, y - 1)$ and $(x + \lfloor r \rfloor, y + 1)$ belong to $B(\mathbf{u})$. Therefore, the set $\{(i, j) \in \mathbb{Z}^2 \mid x - \lfloor r \rfloor \leq i \leq x + \lfloor r \rfloor, y - 1 \leq j \leq y + 1\}$ is a subset of $B(\mathbf{u})$. By the construction of $C_{2\lfloor r \rfloor+1}$, one of the $2\lfloor r \rfloor + 1$ consecutive vertical lines is such that every other vertex in the line is a codeword. Hence, the ball $B(\mathbf{u})$ contains a codeword. In other words, each vertex in $\mathbb{Z}^2$ is covered by a codeword.

Let $\mathbf{v} = (x + x', y + y')$ be a vertex in $\mathbb{Z}^2$ such that $\mathbf{v} \neq \mathbf{u}$. Consider then the symmetric difference $B(\mathbf{u}) \triangle B(\mathbf{v})$. In order to prove that $C_{2\lfloor r \rfloor+1}$ is an identifying code in $\mathcal{E}_r$, we have to show that this symmetric difference always contains a codeword. Without loss of generality, we can assume that $x' \geq 0$ and $y' \geq 0$. (Other cases are analogous.) If $B(\mathbf{u}) \cap B(\mathbf{v}) = \emptyset$, then we are done. Hence, assume that $B(\mathbf{u}) \cap B(\mathbf{v}) \neq \emptyset$.

Assume first that $x' \geq 2$ or $y' \geq 2$. Let $y' \geq 2$ (the other case is analogous). Denote then $\mathbf{u}' = (x, y + y')$ and $\mathbf{v}' = (x + x', y')$. Using similar arguments as in the proof of Lemma 6.2.1, we conclude that each vertical line $L_i^{(v)}$ with $x - \lfloor r \rfloor \leq i \leq x + \lfloor x'/2 \rfloor$ contains two consecutive vertices in $B(\mathbf{u}) \setminus B(\mathbf{u}')$. (Recall that $r^2 - \lfloor r \rfloor^2 \geq 1$.) Clearly, these same points are also included in $B(\mathbf{u}) \setminus B(\mathbf{v})$. By symmetry, we can show that each vertical line $L_i^{(v)}$ with $x + \lceil x'/2 \rceil \leq i \leq x + x' + \lfloor r \rfloor$ contains two consecutive vertices in $B(\mathbf{v}) \setminus B(\mathbf{u})$. Thus, we have shown that each vertical line $L_i^{(v)}$ with $x - \lfloor r \rfloor \leq i \leq x + x' + \lfloor r \rfloor$ contains two consecutive vertices in $B(\mathbf{u}) \triangle B(\mathbf{v})$. Therefore, we conclude that there exists a codeword in $B(\mathbf{u}) \triangle B(\mathbf{v})$.

Assume now that $x' \leq 1$ and $y' \leq 1$. Then we have the following cases to consider:

1) Assume that $x' = 0$ and $y' = 1$. Let $L_k^{(v)}$ be a vertical line with $x - \lfloor r \rfloor \leq k \leq x + \lfloor r \rfloor$. By Lemma 6.2.1, the set $L_k^{(v)} \cap (B(\mathbf{v}) \setminus B(\mathbf{u}))$ is nonempty. Let $\mathbf{w} = (k, y+1+a) \in \mathbb{Z}^2$ be a vertex in $B(\mathbf{v}) \setminus B(\mathbf{u})$. Then, by symmetry, a vertex $\mathbf{w}' = (k, y - a) \in \mathbb{Z}^2$ belongs to $B(\mathbf{u}) \setminus B(\mathbf{v})$. It is immediate that the parity of the second coordinates of the vertices $\mathbf{w}$ and $\mathbf{w}'$ are different. Therefore, since one of the vertical lines $L_i^{(v)}$ with $x - \lfloor r \rfloor \leq i \leq x + \lfloor r \rfloor$ is such that every other vertex in the line is a codeword, the symmetric difference $B(\mathbf{u}) \triangle B(\mathbf{v})$ contains a codeword.

2) If $x' = 1$ and $y' = 0$, the proof goes exactly like in the previous case; just replace the vertical lines by horizontal ones.

3) Assume now that $x' = 1$ and $y' = 1$. Let $\mathbf{w} = (k, y + 1 + a) \in L_k^{(v)}$, where $x - \lfloor r \rfloor \leq k \leq x$, be a vertex such that $\mathbf{w} \in B(x, y+1) \setminus B(x, y)$. By symmetry, the vertex $\mathbf{w}' = (k, y - a)$ belongs to $B(x, y) \setminus B(x, y+1)$.

Since $k \leq x$, the vertex $\mathbf{w}' \in B(x,y) \setminus B(x+1,y+1)$. If $\mathbf{w} \in B(x+1,y+1) \setminus B(x,y)$, then the vertical line $L_k^{(v)}$ contains two vertices ($\mathbf{w}$ and $\mathbf{w}'$) in $B(\mathbf{u}) \triangle B(\mathbf{v})$ such that the parity of their second coordinates are different. Assume then that $\mathbf{w} \notin B(x+1,y+1) \setminus B(x,y)$. Hence, by symmetry, the vertex $\mathbf{w}'' = (k, y+1-a) \in B(x,y) \setminus B(x+1,y+1)$. Clearly, the parity of the second coordinates of $\mathbf{w}'$ and $\mathbf{w}''$ are different. Analogous arguments also apply, when we are considering the vertical lines $L_k^{(v)}$ with $x+1 \leq k \leq x+1+\lfloor r \rfloor$. Hence, each line $L_i^{(v)}$ with $x - \lfloor r \rfloor \leq i \leq x+1+\lfloor r \rfloor$ contains two vertices in $B(\mathbf{u}) \triangle B(\mathbf{v})$ such that the parity of the second coordinates of the vertices are different. Thus, there exists a codeword in $B(\mathbf{u}) \triangle B(\mathbf{v})$.

In conclusion, we have shown that $C_{2\lfloor r \rfloor+1}$ is an identifying code in $\mathcal{E}_r$ when $r^2 - \lfloor r \rfloor^2 \geq 1$.

(ii) Let again $\mathbf{u} = (x,y)$ be a vertex in $\mathbb{Z}^2$. Assume then that $r^2 - \lfloor r \rfloor^2 < 1$. Define the set $A = \{(i,j) \in \mathbb{Z}^2 \mid x - \lfloor r \rfloor \leq i \leq x + \lfloor r \rfloor, y-1 \leq j \leq y\} \setminus \{(x - \lfloor r \rfloor, y-1), (x+\lfloor r \rfloor, y-1)\}$. Let us then show that the set A contains a codeword of $C_{2\lfloor r \rfloor}$. If a vertical line $L_i^{(v)}$ with $x - \lfloor r \rfloor + 1 \leq i \leq x + \lfloor r \rfloor - 1$ is such that every other vertex in the line is a codeword, then we are clearly done. Otherwise, we know that the vertical lines $L_{x-\lfloor r \rfloor}^{(v)}$ and $L_{x+\lfloor r \rfloor}^{(v)}$ are such that every other vertex in the lines is a codeword. Hence, by the construction of $C_{2\lfloor r \rfloor}$, either the vertex $(x - \lfloor r \rfloor, y)$ or $(x + \lfloor r \rfloor, y)$ is a codeword. Since $A \subseteq B(\mathbf{u})$, the word $\mathbf{u}$ is covered by a codeword.

Let $\mathbf{v} = (x+x', y+y')$ be a vertex in $\mathbb{Z}^2$ and $\mathbf{v} \neq \mathbf{u}$. We need to show that the symmetric difference $B(\mathbf{u}) \triangle B(\mathbf{v})$ contains a codeword (when $B(\mathbf{u}) \cap B(\mathbf{v}) \neq \emptyset$). Without loss of generality, we can assume that $x' \geq 0$ and $y' \geq 0$. If $x' = 0$ and $y' = 1$, or $x' = 1$ and $y' = 0$, then the proof goes exactly as in the cases 1) and 2) of the part (i), respectively. Assume that $x' = 0$ and $y' \geq 2$. If now a vertical line $L_i^{(v)}$ with $x - \lfloor r \rfloor + 1 \leq i \leq x + \lfloor r \rfloor - 1$ is such that every other vertex in the line is a codeword, then we are done. Otherwise, either the vertex $(x - \lfloor r \rfloor, y)$ or $(x + \lfloor r \rfloor, y)$ in $B(\mathbf{u}) \triangle B(\mathbf{v})$ is a codeword. Therefore, $I(\mathbf{u}) \triangle I(\mathbf{v}) \neq \emptyset$. Similar arguments also apply when $x' \geq 2$ and $y' = 0$. If $x' = 1$ and $y' = 1$, then the proof goes exactly as previously in the case 3), but we just consider the $2\lfloor r \rfloor$ consecutive vertical lines $L_i^{(v)}$ with $x - \lfloor r \rfloor + 1 \leq i \leq x + \lfloor r \rfloor$. If $x' \geq 1$ and $y' \geq 2$, then the proof is similar to the third paragraph of the proof of the part (i), but we just consider the vertical lines $L_i^{(v)}$ with $x - \lfloor r \rfloor + 1 \leq i \leq x + x' + \lfloor r \rfloor - 1$. The case with $x' \geq 2$ and $y' \geq 1$ goes the same way as the previous one. In conclusion, we have shown that $C_{2\lfloor r \rfloor}$ is an identifying code in $\mathcal{E}_r$ when $r^2 - \lfloor r \rfloor^2 < 1$. □

It is easy to conclude that the density of the code C_k satisfies $D(C_k) \leq 1/k$. Therefore, by the previous theorem, we have shown that for any real number $r \geq 1$ there exists an identifying code C in $\mathcal{E}_r$ such that the density

$$D(C) \leq \frac{1}{2\lfloor r \rfloor}.$$

For small values of r, there exist identifying codes with smaller densities. Indeed, since $\mathcal{E}_{\sqrt{2}} = \mathcal{K}_1$ and $\mathcal{E}_{2\sqrt{2}} = \mathcal{K}_2$, we have optimal identifying codes in $\mathcal{E}_{\sqrt{2}}$ and $\mathcal{E}_{2\sqrt{2}}$ with densities 2/9 and 1/8, respectively (see [15]). Recall that $\mathcal{E}_1 = \mathcal{S}_1$ and $\mathcal{E}_2 = \mathcal{S}_2$. It has been shown in [21] that there exists an identifying code with density 7/20 in $\mathcal{S}_1$. Moreover, it was proved in [4] that there are no identifying codes in $\mathcal{S}_1$ with smaller density. There exists an identifying code in $\mathcal{S}_2$ with density 5/29 (see [52]). In [14], it has been shown that there does not exist an identifying code in $\mathcal{S}_2$ with density smaller than 3/20. Later, this lower bound has been improved to 6/37 in [66].

Consider then a lower bound on the density of an identifying code in $\mathcal{E}_r$. In order to provide a lower bound, we first need to present an auxiliary theorem. This theorem is a rephrased version of [47, Theorem 5]. For completeness, we have also included the proof.

Theorem 6.2.3. *Assume that $C \subseteq \mathbb{Z}^2$ is a code. Let $S = \{\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_k\}$ be a subset containing k different points of $\mathbb{Z}^2$. For each $i = 1, 2, \ldots, k$ we choose a real number $w_i \geq 0$, which we call the weight of $\mathbf{s}_i$ and denote by $w(\mathbf{s}_i)$. For all subsets A of S we define*

$$w(A) = \sum_{\mathbf{a} \in A} w(\mathbf{a}).$$

If for all $\mathbf{v} \in \mathbb{Z}^2$ we have $w((\mathbf{v} + C) \cap S) \geq 1$, then the density of C satisfies

$$D(C) \geq \frac{1}{w_1 + w_2 + \cdots + w_k}.$$

Proof. Since S is finite, we can choose a constant h such that $S \subseteq Q_h$. (Recall the definition of Q_n from Section 1.2.) Consider then the sum $\sum_{\mathbf{v} \in Q_{n-h}} w((\mathbf{v} + C) \cap S)$, where $n > h$. Now we have

$$|Q_{n-h}| \leq \sum_{\mathbf{v} \in Q_{n-h}} w((\mathbf{v} + C) \cap S) \leq \sum_{i=1}^{k} w_i f_i(n), \tag{6.1}$$

where $f_i(n)$ denotes the number of pairs $(\mathbf{c}, \mathbf{v})$ such that $\mathbf{c} \in C$, $\mathbf{v} \in Q_{n-h}$ and $\mathbf{s}_i = \mathbf{v} + \mathbf{c}$. Since $\mathbf{v} \in Q_{n-h}$ and $\mathbf{s}_i \in Q_h$, we know that $\mathbf{c} = \mathbf{s}_i - \mathbf{v} \in Q_n$. Hence, there are at most $|C \cap Q_n|$ choices for $\mathbf{c}$. Furthermore, for every $\mathbf{c}$ there is at most one possible choice for $\mathbf{v} \in Q_{n-h}$ such that $\mathbf{s}_i = \mathbf{c} + \mathbf{v}$. Therefore, $f_i(n) \leq |C \cap Q_n|$.

Combining this result with the inequality (6.1), we have

$$|Q_{n-h}| \leq (w_1 + w_2 + \cdots + w_k)|C \cap Q_n|.$$

Thus,

$$\frac{|C \cap Q_n|}{|Q_n|} \geq \frac{|Q_{n-h}|}{|Q_n|} \cdot \frac{1}{w_1 + w_2 + \cdots + w_k}.$$

Now the claim immediately follows, since $|Q_{n-h}|/|Q_n| \to 1$ when $n \to \infty$. $\square$

In what follows, we prove a lower bound on the density of an identifying code in $\mathcal{E}_r$. The lower bound can actually be attained for some graphs $\mathcal{E}_r$ as can be seen from Theorem 6.3.3.

Theorem 6.2.4. *If $C \subseteq \mathbb{Z}^2$ is an identifying code in $\mathcal{E}_r$, then the density satisfies*

$$D(C) \geq \frac{3}{4\lfloor r \rfloor + 4\lfloor b \rfloor + 4\left\lfloor \sqrt{r^2 - (\lfloor b \rfloor + 1)^2} \right\rfloor + 8},$$

where $b = -1/2 + 1/2 \cdot \sqrt{2r^2 - 1}$.

Proof. Let $C \subseteq \mathbb{Z}^2$ be an identifying code in $\mathcal{E}_r$. Denote $\mathbf{u}_1 = (0,0)$, $\mathbf{u}_2 = (-1,0)$, $\mathbf{u}_3 = (0,-1)$ and $\mathbf{u}_4 = (-1,-1)$. Then define the set

$$\begin{aligned} H = {} & (B(\mathbf{u}_1) \triangle B(\mathbf{u}_2)) \cup (B(\mathbf{u}_1) \triangle B(\mathbf{u}_3)) \cup (B(\mathbf{u}_1) \triangle B(\mathbf{u}_4)) \\ & \cup (B(\mathbf{u}_2) \triangle B(\mathbf{u}_3)) \cup (B(\mathbf{u}_2) \triangle B(\mathbf{u}_4)) \cup (B(\mathbf{u}_3) \triangle B(\mathbf{u}_4)). \end{aligned}$$

Let H' be the set of vertices that belong to H and are covered by exactly two of the vertices $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$ and $\mathbf{u}_4$.

Notice that if $\mathbf{v} \in H \setminus H'$, then $\mathbf{v}$ is covered by exactly one or three of the vertices $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$ and $\mathbf{u}_4$. If a codeword $\mathbf{c} \in C$ belongs to $H \setminus H'$, then, by Theorem 6.1.1, there exist at least three codewords in H. On the other hand, if there does not exist any codeword in $H \setminus H'$, then there clearly exist at least two codewords in H'.

Using the notations of Theorem 6.2.3, we choose $S = H$. The weight of a vertex $\mathbf{s} \in H$ is now defined as follows: if $\mathbf{s} \in H'$, then $w(\mathbf{s}) = 1/2$, else $w(\mathbf{s}) = 1/3$. By the considerations in the previous paragraph, we conclude that for every $\mathbf{v} \in \mathbb{Z}^2$ we have $w((\mathbf{v} + C) \cap H) \geq 1$. By Theorem 6.2.3, we have

$$D(C) \geq \frac{1}{1/2 \cdot |H'| + 1/3 \cdot (|H| - |H'|)} = \frac{3}{|H| + 1/2 \cdot |H'|}.$$

For the lower bound, it is now enough to calculate the number of vertices in H and H'.

For the calculations, define the set $T = \{(x,y) \in \mathbb{Z}^2 \mid x \geq 0, y \geq 0\}$. It is clear that a vertex $\mathbf{u} \in T \cap H$ if and only if $\mathbf{u} \in B(0,0) \setminus B(-1,-1)$ and $\mathbf{u} \in T$. Now, by straightforward calculations, we obtain that the number of vertices in $T \cap H$ is equal to

$$\sum_{i=0}^{\lfloor r\rfloor -1} \left(\left\lfloor \sqrt{r^2 - i^2} \right\rfloor - \left\lfloor \sqrt{r^2 - (i+1)^2 - 1} \right\rfloor \right) + \left\lfloor \sqrt{r^2 - \lfloor r \rfloor^2} \right\rfloor + 1 = 2\lfloor r \rfloor + 1.$$

Therefore, by symmetry, the number of vertices in H is equal to $4(2\lfloor r\rfloor + 1) = 8\lfloor r \rfloor + 4$.

Consider then the number of vertices in H'. It is easy to see that the circles of radius r centered at the points $(-1,0)$ and $(0,-1)$ intersect each other in the point (b,b), where $b = -1/2 + 1/2 \cdot \sqrt{2r^2 - 1}$. Then define the set $T_b = \{(x,y) \in \mathbb{Z}^2 \mid 0 \leq x \leq b, y \geq 0\}$. It is clear that a vertex $\mathbf{u} \in T_b \cap H'$ if and only if $\mathbf{u} \in (B(0,0) \cup B(-1,0)) \setminus (B(0,-1) \cup B(-1,-1))$ and $\mathbf{u} \in T_b$. Hence, by straightforward computations, we have

$$\begin{aligned} |T_b \cap H'| &= \sum_{i=0}^{\lfloor b \rfloor} \left(\left\lfloor \sqrt{r^2 - (i+1)^2} \right\rfloor - \left\lfloor \sqrt{r^2 - i^2 - 1} \right\rfloor \right) \\ &= \left\lfloor \sqrt{r^2 - (\lfloor b \rfloor + 1)^2} \right\rfloor + \lfloor b \rfloor - \lfloor r \rfloor + 1. \end{aligned}$$

Therefore, by symmetry, the number of vertices in H' is equal to

$$8 \left(\left\lfloor \sqrt{r^2 - (\lfloor b \rfloor + 1)^2} \right\rfloor + \lfloor b \rfloor - \lfloor r \rfloor + 1 \right).$$

Thus, we have the lower bound on the density

$$D(C) \geq \frac{3}{4\lfloor r \rfloor + 4\lfloor b \rfloor + 4 \left\lfloor \sqrt{r^2 - (\lfloor b \rfloor + 1)^2} \right\rfloor + 8}.$$

□

Let us then consider more closely the lower bound given by the previous theorem. As in the theorem, let $C \subseteq \mathbb{Z}^2$ be an identifying code in $\mathcal{E}_r$ and denote $b = -1/2 + 1/2 \cdot \sqrt{2r^2 - 1}$. Denote further $\lfloor b \rfloor = k \in \mathbb{Z}$. Since now $b < k+1$, we have that $r < \sqrt{1/2 \cdot (2k+3)^2 + 1/2}$. Therefore, we have

$$\sqrt{r^2 - (\lfloor b \rfloor + 1)^2} \leq \sqrt{\left(\sqrt{1/2 \cdot (2k+3)^2 + 1/2} \right)^2 - (\lfloor b \rfloor + 1)^2} = k + 2.$$

Hence, we further obtain that $\left\lfloor \sqrt{r^2 - (\lfloor b \rfloor + 1)^2} \right\rfloor \leq \lfloor b \rfloor + 1$. Thus, the denominator of the lower bound can be estimated as follows:

$$4\lfloor r \rfloor + 4\lfloor b \rfloor + 4 \left\lfloor \sqrt{r^2 - (\lfloor b \rfloor + 1)^2} \right\rfloor + 8 \leq 4\lfloor r \rfloor + 8\lfloor b \rfloor + 12 \leq 4(\sqrt{2} + 1)r + 12.$$

Therefore, we have the following approximation for the lower bound on the density of an identifying code C in $\mathcal{E}_r$:

$$\frac{1}{3.22r+4} \leq \frac{3}{4(\sqrt{2}+1)r+12} \leq D(C).$$

Recall also the previous result stating that for any r there exists an r-identifying code C in $\mathcal{E}_r$ such that

$$D(C) \leq \frac{1}{2\lfloor r \rfloor}.$$

6.3 Identifying codes in the king grids without corners

In this section, we consider identification in a graph closely related to the king grid. These considerations provide two optimal identifying codes in $\mathcal{E}_r$, as is shown in Theorem 6.3.3. The vertex set V is again equal to $\mathbb{Z}^2$. Let then t be a positive integer and $\mathbf{u} = (x, y)$ be a vertex in $\mathbb{Z}^2$. The edge set E of the considered graph $\mathcal{K}'_t$ is such that

$$B(\mathcal{K}'_t; \mathbf{u}) = B(\mathcal{K}_t; \mathbf{u}) \setminus \{(x+t, y+t), (x+t, y-t), (x-t, y+t), (x-t, y-t)\}.$$

The graph $\mathcal{K}'_t$ is called the *king grid without corners*. Notice that $\mathcal{K}'_1 = \mathcal{S}_1$. As mentioned in Section 6.2, there exists an optimal identifying code in $\mathcal{S}_1$ with density 7/20.

Define a code

$$C_t = \bigcup_{i \in \mathbb{Z}} \{(2t \cdot i + \alpha, \alpha) \mid \alpha \in \mathbb{Z} \text{ and } \alpha \text{ is even}\}.$$

The code C_t is illustrated in Figure 6.1(a) when $t = 2$. Clearly, the density $D(C_t)$ is equal to $1/(4t)$. It has been shown in [15] that C_t is an optimal identifying code in $\mathcal{K}_t$. The following theorem shows that C_t is also an identifying code in $\mathcal{K}'_t$. Notice that now the ball in $\mathcal{K}'_t$ is smaller than the one in $\mathcal{K}_t$. In Theorem 6.3.2, we prove that identifying codes in $\mathcal{K}'_t$ with a lower density do not exist.

Theorem 6.3.1. *Let t be an integer such that $t \geq 2$. Then the code C_t is identifying in $\mathcal{K}'_t$.*

Proof. Let $\mathbf{w} = (x, y)$ be a vertex in $\mathbb{Z}^2$. Then define sets

$$A_h(\mathbf{w}) = \{(i, j) \in \mathbb{Z}^2 \mid x \leq i \leq x + 2t - 1, y \leq j \leq y + 1\}$$

and

$$A_v(\mathbf{w}) = \{(i, j) \in \mathbb{Z}^2 \mid x \leq i \leq x + 1, y \leq j \leq y + 2t - 1\}.$$

Let i be an integer. If i is even, then the horizontal line $L_i^{(h)}$ is such that one of the $2t$ consecutive vertices in the line is a codeword of C_t. The same also holds for the vertical lines. Thus, the sets $A_h(\mathbf{w})$ and $A_v(\mathbf{w})$ both contain at least one codeword.

Let $\mathbf{u} = (x_1, y_1)$ and $\mathbf{v} = (x_2, y_2)$ be vertices in $\mathbb{Z}^2$. The I-set $I(\mathbf{u})$ is nonempty, since the ball $B(\mathbf{u})$ contains the set $A_h(\mathbf{w})$ with a suitable choice of $\mathbf{w}$, when $t \geq 2$. In order to prove the claim, we have to show that the symmetric difference $B(\mathbf{u}) \triangle B(\mathbf{v})$ always contains a codeword. Assume first that $|x_1 - x_2| \geq 3$ or $|y_1 - y_2| \geq 3$. Then the symmetric difference $B(\mathbf{u}) \triangle B(\mathbf{v})$ contains the set $A_v(\mathbf{w})$ or $A_h(\mathbf{w})$. Thus, $I(\mathbf{u}) \triangle I(\mathbf{v}) \neq \emptyset$.

Assume now that $|x_1 - x_2| \leq 2$ and $|y_1 - y_2| \leq 2$. Then we have the following cases to consider (other cases are analogous):

1) Assume that $\mathbf{v} = (x_1 + 1, y_1)$ or $\mathbf{v} = (x_1 + 2, y_1)$. Denote $X_1 = \{(x_1 - t, y_1 - t + 1), (x_1 - t, y_1 - t + 2), \ldots, (x_1 - t, y_1 + t - 1)\}$ and $X_2 = \{(x_1 + t + 1, y_1 - t + 1), (x_1 + t + 1, y_1 - t + 2), \ldots, (x_1 + t + 1, y_1 + t - 1)\}$. It is easy to see that $X_1, X_2 \subseteq B(\mathbf{u}) \triangle B(\mathbf{v})$ and $(x_1 - t + 1, y_1 - t), (x_1 + t, y_1 + t) \in B(\mathbf{u}) \triangle B(\mathbf{v})$. Assume first that $x_1 - t$ is even. Then, by the previous considerations, either X_1 contains a codeword or the vertex $(x_1 - t, y_1 - t)$ is a codeword. If X_1 contains a codeword, we are done. Otherwise, the vertex $(x_1 - t, y_1 - t)$ is a codeword. Therefore, by the construction of C_t, the vertex $(x_1 - t + 2t, y_1 - t + 2t) = (x_1 + t, y_1 + t)$, which belongs to $B(\mathbf{u}) \triangle B(\mathbf{v})$, is a codeword. Assume then that $x_1 - t$ is odd. Now $x_1 + t + 1$ is even. The proof in this case is similar to the previous one.

2) Assume that $\mathbf{v} = (x_1 + 1, y_1 + 1)$ or $\mathbf{v} = (x_1 + 2, y_1 + 2)$. Denote $Y_1 = X_1$ and $Y_2 = (0, 1) + X_2$. It is easy to see that $Y_1, Y_2 \subseteq B(\mathbf{u}) \triangle B(\mathbf{v})$ and $(x_1 - t + 1, y_1 - t + 1), (x_1 + t, y_1 + t) \in B(\mathbf{u}) \triangle B(\mathbf{v})$. Assume first that $x_1 - t$ is even. If Y_1 contains a codeword, we are done. Otherwise, the vertex $(x_1 - t, y_1 - t)$ is a codeword. Therefore, the vertex $(x_1 - t + 2t, y_1 - t + 2t) = (x_1 + t, y_1 + t)$ is a codeword. If $x_1 - t$ is odd, then $x_1 + t + 1$ is even and the proof is similar to the first case.

3) Assume that $\mathbf{v} = (x_1 + 2, y_1 + 1)$. The proof is now analogous to the previous case 2).

In conclusion, we have shown that the symmetric difference $I(\mathbf{u}) \triangle I(\mathbf{v})$ is always nonempty. Hence, the claim follows. □

The following theorem provides a lower bound on the density of an identifying code in $\mathcal{K}'_t$.

Theorem 6.3.2. *If C is an identifying code in $\mathcal{K}'_t$, then the density*

$$D(C) \geq \frac{1}{4t}.$$

Proof. Let C be an identifying code in $\mathcal{K}'_t$. Define the vertices $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4 \in \mathbb{Z}^2$ and the sets $H, H' \subseteq \mathbb{Z}^2$ as in the proof of Theorem 6.2.4. Using similar arguments as in the proof of Theorem 6.2.4, we have

$$D(C) \geq \frac{3}{|H| + 1/2 \cdot |H'|}.$$

It is easy to calculate that $|H| = 8t + 4$ and $|H'| = 4(2t - 2)$. Therefore,

$$D(C) \geq \frac{3}{8t + 4 + 1/2 \cdot 4(2t - 2)} = \frac{1}{4t}.$$

$\square$

In conclusion, we have shown that C_t is an optimal identifying code in $\mathcal{K}'_t$. Hence, we have the following theorem concerning identifying codes in $\mathcal{E}_r$ when $r = \sqrt{5}$ or $r = \sqrt{13}$.

Theorem 6.3.3. *The codes C_2 and C_3 are optimal identifying codes in $\mathcal{E}_{\sqrt{5}}$ and $\mathcal{E}_{\sqrt{13}}$, respectively.*

Proof. The claim immediately follows from the fact that $\mathcal{E}_{\sqrt{5}} = \mathcal{K}'_2$ and $\mathcal{E}_{\sqrt{13}} = \mathcal{K}'_3$. $\square$

Chapter 7

Adaptive identification

Recently, a new concept of adaptive identification has been introduced in [3]. In this chapter, which is based on the paper [54], we consider adaptive identification in binary Hamming spaces. We begin the chapter with some preliminary definitions in Section 7.1 and then proceed with the results in Section 7.2.

7.1 Preliminaries

Let $G = (V, E)$ be a simple, connected and undirected graph. First recall (from Section 1.2) that a code in G is said to be an r-covering if each vertex of G is r-covered by a codeword. The minimum cardinality of an r-covering in G is denoted by $\gamma_r(G)$. Furthermore, a code C is called an *r-packing* in G, if the number of vertices in $I_r(C; u)$ is at most one for all $u \in V$. In other words, the r-balls centered at the vertices of C are all pairwise disjoint. The maximum cardinality of an r-packing in G is denoted by $c_r(G)$. If a code C is both an r-covering and r-packing in G, then C is called an *r-perfect code*.

Assume that a given graph G may contain *faulty vertices* and that we can ask whether there is a faulty vertex (or faulty vertices) in $B_r(u)$ for all $u \in V$. The *query* $\mathcal{Q}_r : V \longrightarrow \{0, 1\}$ is equal to 1 for $u \in V$, if there is a faulty vertex in $B_r(u)$, else $\mathcal{Q}_r(u)$ is equal to 0. We also say that a vertex $v \in V$ is *r-covered* by a query $\mathcal{Q}_r(u)$ ($u \in V$), if v belongs to the r-ball $B_r(u)$. Now the problem is to locate the faulty vertices using the queries $\mathcal{Q}_r(u)$. The definition of identifying codes guarantees that if $C \subseteq V$ is an $(r, \leq \ell)$-identifying code in G, then by asking simultaneously all the queries $\mathcal{Q}_r(c)$ for $c \in C$ we can locate in one step all the faulty vertices in G (assuming there are at most ℓ faulty vertices in G).

The definition of identifying codes is based on the fact that all the queries have to be asked simultaneously. However, *adaptive identification*, which was recently introduced in [3], is based on the idea that the queries can be asked one after the other, i.e. that a new query may depend on the answers given by the previous ones. In what follows, we call the identifying codes (in the sense of Definition 1.2.1) *regular* to distinguish them from the adaptive ones.

Let ℓ be the maximum number of faulty vertices in a graph G. Recall that the minimum cardinality of an $(r, \leq \ell)$-identifying code in G is then denoted by $M_{(r,\leq \ell)}(G)$. In adaptive identification, the corresponding value is the minimum of the maximum number of queries required to identify the (at most ℓ) faulty vertices and it is denoted by $a_{(r,\leq \ell)}(G)$. We also say that an algorithm (or a series of queries) $\mathcal{A}$ is *adaptive* $(r, \leq \ell)$*-identifying*, if it can identify the at most ℓ faulty vertices in G using only the queries $\mathcal{Q}_r(u)$ $(u \in V)$.

In Ben-Haim *et al.* [2] and [3], adaptive $(r, \leq 1)$-identification is considered in torii of square and king lattices. They suggest that further study would be needed in these torii when $\ell > 1$ and also in various other graphs such as binary Hamming spaces. In this chapter, we study adaptive identification in binary Hamming spaces $\mathbb{F}^n$ when $\ell = 1$. For the results in various graphs in the case $\ell > 1$, the interested reader is referred to [54].

7.2 Adaptive identification in Hamming spaces

First recall the basic definitions concerning binary Hamming spaces $\mathbb{F}^n$ from Section 2.1. We also need the following lemma, which can be easily proven.

Lemma 7.2.1. *Let* $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$. *Then*

$$|B_1(\mathbf{x}) \cap B_1(\mathbf{y})| = \begin{cases} n+1 & \text{if } \mathbf{x} = \mathbf{y}, \\ 2 & \text{if } 1 \leq d(\mathbf{x}, \mathbf{y}) \leq 2, \\ 0 & \text{otherwise.} \end{cases}$$

The following lemma is needed in the proof of Theorem 7.2.3, which provides a lower bound for $a_{(1,\leq 1)}(\mathbb{F}^n)$.

Lemma 7.2.2. *Let X be a nonempty subset of $\mathbb{F}^n$. Assume that there are 0 or 1 faulty words in X. Then an adaptive $(1, \leq 1)$-identifying algorithm needs at least*

$$\left\lceil \sqrt{\frac{|X|}{2}} \right\rceil$$

queries, which are centered at a word in $\mathbb{F}^n$, to locate the faulty word in X or to conclude that there is none.

Proof. Let X be a nonempty subset of $\mathbb{F}^n$ and let $\mathcal{A}$ be an algorithm that identifies the faulty word in X using queries from $\mathbb{F}^n$. Define then k as the maximum number of words in X that are 1-covered by a 1-ball of $\mathbb{F}^n$, i.e.

$$k = \max_{\mathbf{x} \in \mathbb{F}^n} |B_1(\mathbf{x}) \cap X|.$$

Now we have two approaches for the lower bound on the number of queries used in $\mathcal{A}$:

(1) By the previous definition, a query $\mathcal{Q}_1(\mathbf{x})$ with $\mathbf{x} \in \mathbb{F}^n$ can 1-cover at most k words of X. Assume that the first $\lfloor |X|/k \rfloor - 1$ queries of $\mathcal{A}$ output value 0. Now there still exist at least k words in X that are not 1-covered by any of the previous queries, and we need at least $\lceil \log_2(k+1) \rceil$ queries to locate the faulty word among these uncovered words or to conclude that there is none. Thus, the number of queries used in $\mathcal{A}$ is at least $\lfloor |X|/k \rfloor - 1 + \lceil \log_2(k+1) \rceil$. If $k = 1$, then the claim clearly follows. Otherwise, we need at least $|X|/k$ queries in the algorithm $\mathcal{A}$.

(2) On the other hand, we know by Lemma 7.2.1 that the number of words in the intersection of two different 1-balls of $\mathbb{F}^n$ is at most 2. Let then $\mathbf{x} \in \mathbb{F}^n$ be a word such that the number of words in $B_1(\mathbf{x}) \cap X$ is equal to k. Assume that there exists a faulty word in $B_1(\mathbf{x}) \cap X$. Using similar arguments as in the first lower bound from (1), we obtain that the number of queries used in $\mathcal{A}$ is at least

$$\begin{cases} \lfloor k/2 \rfloor - 1 + \lceil \log_2 2 \rceil & \text{if } 2 \mid k, \\ \lfloor k/2 \rfloor - 1 + \lceil \log_2 3 \rceil & \text{if } 2 \nmid k. \end{cases}$$

Therefore, the number of queries needed is at least $k/2$.

By the considerations above, the number of queries needed in $\mathcal{A}$ is at least $\max\{|X|/k, k/2\}$. Therefore, by straightforward analysis, it can be concluded that (with any choice of k) the number of queries needed is at least

$$\left\lceil \sqrt{\frac{|X|}{2}} \right\rceil .$$

□

The following theorem provides a lower bound for $a_{(1,\leq 1)}(\mathbb{F}^n)$.

Theorem 7.2.3. *We have*

$$a_{(1,\leq 1)}(\mathbb{F}^n) \geq c_1(\mathbb{F}^n) + \left\lfloor \frac{n+1}{8} \right\rfloor .$$

Proof. Let algorithm $\mathcal{A}$ be an adaptive $(1, \leq 1)$-identifying code (in $\mathbb{F}^n$). (Notice that the size of a ball of radius 1 in $\mathbb{F}^n$ is equal to $n+1$.) Assume then that the first $c_1(\mathbb{F}^n) - \lceil (n+1)/8 \rceil$ queries of $\mathcal{A}$ output value 0. (Here the number of queries $c_1(\mathbb{F}^n) - \lceil (n+1)/8 \rceil$ is chosen in such a way that it gives the best possible lower bound using this approach.) Then the number of words that are not 1-covered by the previous queries is at least $\lceil (n+1)/8 \rceil (n+1)$. Therefore, by Lemma 7.2.2, the number of queries used in $\mathcal{A}$ is at least

$$c_1(\mathbb{F}^n) - \left\lceil \frac{n+1}{8} \right\rceil + \left\lceil \sqrt{\frac{\lceil (n+1)/8 \rceil (n+1)}{2}} \right\rceil \geq c_1(\mathbb{F}^n) + \left\lfloor \frac{n+1}{8} \right\rfloor .$$

□

The following theorem provides an upper bound for $a_{(1,\leq 1)}(\mathbb{F}^n)$.

Theorem 7.2.4. *We have*

$$a_{(1,\leq 1)}(\mathbb{F}^n) \leq \gamma_1(\mathbb{F}^n) + \left\lceil \frac{n+1}{2} \right\rceil .$$

Proof. Let $C = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{|C|}\}$ be a 1-covering code in $\mathbb{F}^n$ attaining $\gamma_1(\mathbb{F}^n)$. Denote then by $\mathbf{e}_i$ the word in $\mathbb{F}^n$ that has value 1 in the ith coordinate place and value 0 in all other places, i.e. $supp(\mathbf{e}_i) = \{i\}$. Now the following algorithm is adaptive $(1, \leq 1)$-identifying:

1. For $i = 1, \ldots, |C| - 1$ ask the query $\mathcal{Q}_1(\mathbf{x}_i)$. If $\mathcal{Q}_1(\mathbf{x}_i) = 1$ for any $i = 1, \ldots, |C| - 1$, then the faulty word in $B_1(\mathbf{x}_i)$ can be located as in the following step 2.

2. Assume then that all the previous queries output value 0, meaning that any of these queries do not 1-cover a faulty word. Now we can assume without loss of generality that $\mathbf{x}_{|C|} = \mathbf{0}$. For $i = 1, \ldots, \lceil n/2 \rceil - 1$ ask the query $\mathcal{Q}_1(\mathbf{e}_{2i-1} + \mathbf{e}_{2i})$. Now, if for any i we have $\mathcal{Q}_1(\mathbf{e}_{2i-1} + \mathbf{e}_{2i}) = 1$, then the faulty word can be located using one more suitably chosen query. Hence, assume that any of the previous queries do not 1-cover a faulty word. Now it can be easily seen that we only need two more queries to locate the faulty word in the remaining words or to conclude that there are none.

In conclusion, the previous algorithm uses at most $\gamma_1(\mathbb{F}^n) + \lceil n/2 \rceil + 1 = \gamma_1(\mathbb{F}^n) + \lceil (n+1)/2 \rceil$ queries. □

Let s and n be integers such that $s \geq 3$ and $n = 2^s - 1$. Consider then the binary Hamming space $\mathbb{F}^n$. By [65, Chapter 6] and [75, Chapter 3], we know that now there exists a 1-perfect covering in $\mathbb{F}^n$. Hence, we have $c_1(\mathbb{F}^n) = \gamma_1(\mathbb{F}^n)$. Therefore, for the previous lengths, Theorems 7.2.3 and 7.2.4 can be written as follows:

$$c_1(\mathbb{F}^n) + \frac{n+1}{8} \leq a_{(1,\leq 1)}(\mathbb{F}^n) \leq c_1(\mathbb{F}^n) + \frac{n+1}{2}.$$

Hence, we know the order of growth for $a_{(1,\leq 1)}(\mathbb{F}^n)$.

Chapter 8

Conclusion

In this thesis, we have considered various topics in the fields of identifying and locating-dominating codes. In what follows, we summarise some of the obtained results as well as give suggestions for future research.

In Chapter 2, we studied identifying codes in binary Hamming spaces $\mathbb{F}^n$. First, in Section 2.2, we were able to improve known lower bounds for the sizes of r-identifying codes in $\mathbb{F}^n$ when $r \geq 2$. The main tool in achieving these improvements was the study of the function $P_r(n, i, \mathbf{x})$. More precisely, by providing a good upper bound for this function, we managed to improve known lower bounds on $M_r(\mathbb{F}^n)$. However, there still seems to be room for significant improvement concerning the lower bound and a viable approach in doing so could be to obtain a better understanding of the function $P_r(n, i, \mathbf{x})$. In Sections 2.3, 2.4 and 2.5, we studied three conjectures, which have been stated in the papers [9] and [60], and although we were not able to completely solve these problems, we managed to present several results related to these conjectures. Hence, there is still work to do concerning these conjectures as well as some other ones such as whether $M_r(\mathbb{F}^n) \leq M_r(\mathbb{F}^{n+1})$ holds for general r when n is large enough (the case with $r = 1$ is shown to be true; see [67]).

In Chapters 3 and 4, we considered identifying and locating-dominating codes in cycles $\mathcal{C}_n$ and paths $\mathcal{P}_n$. In the case of identifying codes, we determined the exact values of $M_r(\mathcal{C}_n)$ and $M_r(\mathcal{P}_n)$ in all the remaining open cases. Concerning locating-dominating codes in paths, we solved a conjecture, which states that there exists an infinite family of n such that $M_r^{LD}(\mathcal{P}_n) = \lceil (n+1)/3 \rceil$, by showing that this equality actually holds always when n is large enough. Although this result determines the majority of the exact values of $M_r^{LD}(\mathcal{P}_n)$, there is still room for future research when n is small. In the case of locating-dominating codes in cycles, we were able to prove a result similar to the one of the paths according to which $n/3 \leq M_r^{LD}(\mathcal{C}_n) \leq n/3 + 1$ if $n \equiv 3 \pmod 6$ and $M_r^{LD}(\mathcal{C}_n) = \lceil n/3 \rceil$ otherwise. In the latter case, we also conjectured that the equality $M_r^{LD}(\mathcal{C}_n) = n/3 + 1$ actually holds. It is obvious that a proof for this conjecture would be welcomed.

In Chapter 5, we showed that a 2-identifying code in the hexagonal grid G_H with density 4/19 is optimal by proving that there are no 2-identifying codes in G_H with smaller density. The main idea of the proof in the improved lower bound was to show that on average the share of each codeword in a 2-identifying code in G_H is at most 19/4. This approach, which is based on the concept of share, can also be used for improving lower bounds of r-identifying codes (with small r) in other infinite grids such as square and triangular grids. Indeed, in a forthcoming paper, we have been able to improve known lower bound on 2-identifying code in the square grid.

In Chapter 6, we studied identification in the infinite grid $\mathbb{Z}^2$ using Euclidean balls. We obtained a general lower bound stating that the density of any identifying code C in $\mathcal{E}_r$ satisfies $D(C) \geq 1/(3.22r + 4)$. Furthermore, we also presented a construction according to which we have an identifying code C in $\mathcal{E}_r$ (for any positive real number r) such that $D(C) \leq 1/(2\lfloor r \rfloor)$. As a future research subject, it would be interesting to obtain improvements for both the lower and upper bounds above. In addition to the general results, we also proved that the densities of optimal identifying codes in $\mathcal{E}_{\sqrt{5}}$ and $\mathcal{E}_{\sqrt{13}}$ are 1/8 and 1/12, respectively. Improvements over the general results might also be possible in other cases with small radius.

Finally, in Chapter 7, we briefly discussed adaptive identification, which is a sequential variant of regular identification. In particular, we obtained the following bounds for the maximum number of queries $a_{(r,\leq 1)}(\mathbb{F}^n)$ needed in an adaptive $(r, \leq 1)$-identifying algorithm:

$$c_1(\mathbb{F}^n) + \frac{n+1}{8} \leq a_{(1,\leq 1)}(\mathbb{F}^n) \leq c_1(\mathbb{F}^n) + \frac{n+1}{2}.$$

More extensive coverage on adaptive identification can be found in [54].

Bibliography

[1] David Auger. Minimal identifying codes in trees and planar graphs with large girth. *European J. Combin.*, 31(5):1372–1384, 2010.

[2] Yael Ben-Haim, Sylvain Gravier, Antoine Lobstein, and Julien Moncel. Adaptive identification in graphs. Technical report, Rapport interne Telecom Paris-2007D012, September 2007.

[3] Yael Ben-Haim, Sylvain Gravier, Antoine Lobstein, and Julien Moncel. Adaptive identification in graphs. *J. Comb. Theory Ser. A*, 115(7):1114–1126, 2008.

[4] Yael Ben-Haim and Simon Litsyn. Exact minimum density of codes identifying vertices in the square grid. *SIAM J. Discrete Math.*, 19(1):69–82, 2005.

[5] Tanya Y. Berger-Wolf, William E. Hart, and Jared Saia. Discrete sensor placement problems in distribution networks. *Math. Comput. Modelling*, 42(13):1385–1396, 2005.

[6] Nathalie Bertrand, Irène Charon, Olivier Hudry, and Antoine Lobstein. Identifying and locating-dominating codes on chains and cycles. *European J. Combin.*, 25(7):969–987, 2004.

[7] Nathalie Bertrand, Irène Charon, Olivier Hudry, and Antoine Lobstein. 1-identifying codes on trees. *Australas. J. Combin.*, 31:21–35, 2005.

[8] Uri Blass, Iiro Honkala, and Simon Litsyn. On binary codes for identification. *J. Combin. Des.*, 8(2):151–156, 2000.

[9] Uri Blass, Iiro Honkala, and Simon Litsyn. Bounds on identifying codes. *Discrete Math.*, 241(1-3):119–128, 2001.

[10] Mostafa Blidia, Mustapha Chellali, Frédéric Maffray, Julien Moncel, and Ahmed Semri. Locating-domination and identifying codes in trees. *Australas. J. Combin.*, 39:219–232, 2007.

[11] D. I. Carson. On generalized location-domination. In *Graph theory, combinatorics, and algorithms, Vol. 1, 2 (Kalamazoo, MI, 1992)*, Wiley-Intersci. Publ., pages 161–179. Wiley, New York, 1995.

[12] Irène Charon, Gérard Cohen, Olivier Hudry, and Antoine Lobstein. New identifying codes in the binary Hamming space. *European J. Combin.*, 31(2):491–501, 2010.

[13] Irène Charon, Sylvain Gravier, Olivier Hudry, Antoine Lobstein, Michel Mollard, and Julien Moncel. A linear algorithm for minimum 1-identifying codes in oriented trees. *Discrete Appl. Math.*, 154(8):1246–1253, 2006.

[14] Irène Charon, Iiro Honkala, Olivier Hudry, and Antoine Lobstein. General bounds for identifying codes in some infinite regular graphs. *Electron. J. Combin.*, 8(1):Research Paper 39, 21 pp., 2001.

[15] Irène Charon, Iiro Honkala, Olivier Hudry, and Antoine Lobstein. The minimum density of an identifying code in the king lattice. *Discrete Math.*, 276(1-3):95–109, 2004. 6th International Conference on Graph Theory.

[16] Irène Charon, Olivier Hudry, and Antoine Lobstein. Identifying codes with small radius in some infinite regular graphs. *Electron. J. Combin.*, 9(1):Research Paper 11, 25 pp., 2002.

[17] Irène Charon, Olivier Hudry, and Antoine Lobstein. Minimizing the size of an identifying or locating-dominating code in a graph is NP-hard. *Theoret. Comput. Sci.*, 290(3):2109–2120, 2003.

[18] Irène Charon, Olivier Hudry, and Antoine Lobstein. Possible cardinalities for locating-dominating codes in graphs. *Australas. J. Combin.*, 34:23–31, 2006.

[19] Irène Charon, Olivier Hudry, and Antoine Lobstein. Extremal cardinalities for identifying and locating-dominating codes in graphs. *Discrete Math.*, 307(3-5):356–366, 2007.

[20] Chunxia Chen, Changhong Lu, and Zhengke Miao. Identifying codes and locating-dominating sets on paths and cycles. Submitted, Discrete Appl. Math., 2009.

[21] Gérard Cohen, Sylvain Gravier, Iiro Honkala, Antoine Lobstein, Michel Mollard, Charles Payan, and Gilles Zémor. Improved identifying codes for the grid. *Electron. J. Combin.*, 6: Research Paper 19, Comment, 3 pp., 1999.

[22] Gérard Cohen, Iiro Honkala, Simon Litsyn, and Antoine Lobstein. *Covering codes*, volume 54 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam, 1997.

[23] Gérard Cohen, Iiro Honkala, Antoine Lobstein, and Gilles Zémor. Bounds for codes identifying vertices in the hexagonal grid. *SIAM J. Discrete Math.*, 13(14):492–504, 2000.

[24] Gérard Cohen, Iiro Honkala, Antoine Lobstein, and Gilles Zémor. On codes identifying vertices in the two-dimensional square lattice with diagonals. *IEEE Trans. Comput.*, 50(2):174–176, 2001.

[25] Daniel W. Cranston and Gexin Yu. A new lower bound on the density of vertex identifying codes for the infinite hexagonal grid. *Electron. J. Combin.*, 16(1):Research Paper 113, 16, 2009.

[26] Geoffrey Exoo, Ville Junnila, and Tero Laihonen. On location-domination of set of vertices in cycles and paths. In *Proceedings of the Thirty-Ninth Southeastern International Conference on Combinatorics, Graph Theory and Computing*, volume 202, pages 97–112, 2010.

[27] Geoffrey Exoo, Ville Junnila, and Tero Laihonen. Locating-dominating codes in cycles. *Australas. J. Combin.*, 49:177–194, 2011.

[28] Geoffrey Exoo, Ville Junnila, and Tero Laihonen. Locating-dominating codes in paths. *Discrete Math.*, to appear.

[29] Geoffrey Exoo, Ville Junnila, Tero Laihonen, and Sanna Ranto. Constructions for identifying codes. In *Proceedings of the Eleventh International Workshop on Algebraic and Combinatorial Coding Theory. ACCT 2008.*, pages 92–98, 2008.

[30] Geoffrey Exoo, Ville Junnila, Tero Laihonen, and Sanna Ranto. Improved bounds on binary identifying codes. In *Proceedings of IEEE International Symposium on Information Theory, 2008. ISIT 2008.*, pages 629–633, 2008.

[31] Geoffrey Exoo, Ville Junnila, Tero Laihonen, and Sanna Ranto. Locating vertices using codes. In *Proceedings of the Thirty-Ninth Southeastern International Conference on Combinatorics, Graph Theory and Computing*, volume 191, pages 143–159, 2008.

[32] Geoffrey Exoo, Ville Junnila, Tero Laihonen, and Sanna Ranto. Upper bounds for binary identifying codes. *Adv. in Appl. Math.*, 42(3):277–289, 2009.

[33] Geoffrey Exoo, Ville Junnila, Tero Laihonen, and Sanna Ranto. Improved bounds on identifying codes in binary Hamming spaces. *European J. Combin.*, 31(3):813–827, 2010.

[34] Geoffrey Exoo, Tero Laihonen, and Sanna Ranto. Improved upper bounds on binary identifying codes. *IEEE Trans. Inform. Theory*, 53(11):4255–4260, 2007.

[35] Geoffrey Exoo, Tero Laihonen, and Sanna Ranto. New bounds on binary identifying codes. *Discrete Appl. Math.*, 156(12):2250–2263, 2008.

[36] Sylvain Gravier, Julien Moncel, and Ahmed Semri. Identifying codes of cycles. *European J. Combin.*, 27(5):767–776, 2006.

[37] Teresa W. Haynes, Stephen T. Hedetniemi, and Peter J. Slater. *Fundamentals of domination in graphs*, volume 208 of *Monographs and Textbooks in Pure and Applied Mathematics*.

[38] Iiro Honkala. On the identifying radius of codes. In *Proceedings of the Seventh Nordic Combinatorial Conference*, TUCS General Publications, pages 33–43, Turku, 1999.

[39] Iiro Honkala. An optimal edge-robust identifying code in the triangular lattice. *Ann. Comb.*, 8(3):303–323, 2004.

[40] Iiro Honkala. An optimal locating-dominating set in the infinite triangular grid. *Discrete Math.*, 306(21):2670–2681, 2006.

[41] Iiro Honkala. On r-locating-dominating sets in paths. *European J. Combin.*, 30(4):1022–1025, 2009.

[42] Iiro Honkala. An optimal strongly identifying code in the infinite triangular grid. *Electron. J. Combin.*, 17(1), 2010.

[43] Iiro Honkala, Mark G. Karpovsky, and Lev B. Levitin. On robust and dynamic identifying codes. *IEEE Trans. Inform. Theory*, 52(2):599–611, 2006.

[44] Iiro Honkala and Tero Laihonen. Codes for identification in the king lattice. *Graphs Combin.*, 19(4):505–516, 2003.

[45] Iiro Honkala and Tero Laihonen. On identifying codes in the triangular and square grids. *SIAM J. Comput.*, 33(2):304–312, 2004.

[46] Iiro Honkala and Tero Laihonen. On locating-dominating sets in infinite grids. *European J. Combin.*, 27(2):218–227, 2006.

[47] Iiro Honkala and Tero Laihonen. On a new class of identifying codes in graphs. *Inform. Process. Lett.*, 102(2-3):92–98, 2007.

[48] Iiro Honkala and Tero Laihonen. On identifying codes that are robust against edge changes. *Inform. and Comput.*, 205(7):1078–1095, 2007.

[49] Iiro Honkala, Tero Laihonen, and Sanna Ranto. On strongly identifying codes. *Discrete Math.*, 254(1-3):191–205, 2002.

[50] Iiro Honkala, Tero Laihonen, and Sanna Ranto. On locating-dominating codes in binary Hamming spaces. *Discrete Math. Theor. Comput. Sci.*, 6(2):265–281, 2004.

[51] Iiro Honkala and Antoine Lobstein. On identifying codes in binary Hamming spaces. *J. Combin. Theory Ser. A*, 99(2):232–243, 2002.

[52] Iiro Honkala and Antoine Lobstein. On the density of identifying codes in the square lattice. *J. Combin. Theory Ser. B*, 85(2):297–306, 2002.

[53] Loo Keng Hua. *Introduction to number theory.* Springer-Verlag, Berlin, 1982. Translated from the Chinese by Peter Shiu.

[54] Ville Junnila. Adaptive identification of sets of vertices in graphs. *Discrete Math. Theor. Comput. Sci.*, submitted.

[55] Ville Junnila and Tero Laihonen. Codes for identification in $\mathbb{Z}^2$ with Euclidean balls. In *Preproceedings of the International Workshop on Coding and Cryptography*, pages 298–309, 2009.

[56] Ville Junnila and Tero Laihonen. Identification in $\mathbb{Z}^2$ using Euclidean balls. *Discrete Appl. Math.*, 159(5):355–343, 2011.

[57] Ville Junnila and Tero Laihonen. Optimality of a 2-identifying code in the hexagonal grid. In *Preproceedings of the International Workshop on Coding and Cryptography*, pages 47–56, 2011.

[58] Ville Junnila and Tero Laihonen. Optimal lower bound for 2-identifying code in the hexagonal grid. *Electron. J. Combin.*, submitted.

[59] Ville Junnila and Tero Laihonen. Optimal identifying codes in cycles and paths. *Graphs Combin.*, to appear.

[60] Mark G. Karpovsky, Krishnendu Chakrabarty, and Lev B. Levitin. On a new class of codes for identifying vertices in graphs. *IEEE Trans. Inform. Theory*, 44(2):599–611, 1998.

[61] Moshe Laifenfeld, Ari Trachtenberg, Reuven Cohen, and David Starobinski. Joint monitoring and routing in wireless sensor networks using robust identifying codes. *Mobile Networks and Applications*, 14:415–432, 2009. 10.1007/s11036-008-0105-x.

[62] Tero Laihonen. Sequences of optimal identifying codes. *IEEE Trans. Inform. Theory*, 48(3):774–776, 2002.

[63] Tero Laihonen. On robust identification in the square and king grids. *Discrete Appl. Math.*, 154:2499–2510, 2006.

[64] Antoine Lobstein. Identifying and locating-dominating codes in graphs, a bibliography. Published electronically at `http://perso.enst.fr/~lobstein/debutBIBidetlocdom.pdf`.

[65] F. Jessie MacWilliams and Neil J. A. Sloane. *The theory of error-correcting codes*, volume 16 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam, 1977.

[66] Ryan Martin and Brendon Stanton. Lower bounds for identifying codes in some infinite grids. *Electron. J. Combin.*, 17(1):Research Paper 122, 2010.

[67] Julien Moncel. Monotonicity of the minimum cardinality of an identifying code in the hypercube. *Discrete Appl. Math.*, 154(6):898–899, 2006.

[68] Douglas F. Rall and Peter J. Slater. On location-domination numbers for certain classes of graphs. *Congr. Numer.*, 45:97–106, 1984.

[69] Saikat Ray, David Starobinski, Ari Trachtenberg, and Rachanee Ungrangsi. Robust location detection with sensor networks. *IEEE Journal on Selected Areas in Communications*, 22(6):1016–1025, August 2004.

[70] David L. Roberts and Fred S. Roberts. Locating sensors in paths and cycles: The case of 2-identifying codes. *European J. Combin.*, 29(1):72–82, 2008.

[71] Peter J. Slater. Domination and location in acyclic graphs. *Networks*, 17(1):55–64, 1987.

[72] Peter J. Slater. Dominating and reference sets in a graph. *J. Math. Phys. Sci.*, 22:445–455, 1988.

[73] Peter J. Slater. Fault-tolerant locating-dominating sets. *Discrete Math.*, 249(1–3):179–189, 2002.

[74] Jukka Suomela. Approximability of identifying codes and locating-dominating codes. *Inform. Process. Lett.*, 103(1):28–33, 2007.

[75] Jacobus H. van Lint. *Introduction to coding theory*, volume 86 of *Graduate Texts in Mathematics*. Springer-Verlag, Berlin, third edition, 1999.

[76] Douglas B. West. *Introduction to graph theory*. Prentice Hall Inc., Upper Saddle River, NJ, 1996.

[77] Min Xu, Krishnaiyan Thulasiraman, and Xiao-Dong Hu. Identifying codes of cycles with odd orders. *European J. Combin.*, 29(7):1717–1720, 2008.

Turku Centre for Computer Science
TUCS Dissertations

104. **Denis Shestakov**, Search Interfaces on the Web: Querying and Characterizing
105. **Sampo Pyysalo**, A Dependency Parsing Approach to Biomedical Text Mining
106. **Anna Sell**, Mobile Digital Calendars in Knowledge Work
107. **Dorina Marghescu**, Evaluating Multidimensional Visualization Techniques in Data Mining Tasks
108. **Tero Säntt**i, A Co-Processor Approach for Efficient Java Execution in Embedded Systems
109. **Kari Salonen**, Setup Optimization in High-Mix Surface Mount PCB Assembly
110. **Pontus Boström**, Formal Design and Verification of Systems Using Domain-Specific Languages
111. **Camilla J. Hollanti**, Order-Theoretic Methods for Space-Time Coding: Symmetric and Asymmetric Designs
112. **Heidi Himmanen**, On Transmission System Design for Wireless Broadcasting
113. **Sébastien Lafond**, Simulation of Embedded Systems for Energy Consumption Estimation
114. **Evgeni Tsivtsivadze**, Learning Preferences with Kernel-Based Methods
115. **Petri Salmela**, On Communication and Conjugacy of Rational Languages and the Fixed Point Method
116. **Siamak Taati**, Conservation Laws in Cellular Automata
117. **Vladimir Rogojin**, Gene Assembly in Stichotrichous Ciliates: Elementary Operations, Parallelism and Computation
118. **Alexey Dudkov**, Chip and Signature Interleaving in DS CDMA Systems
119. **Janne Savela**, Role of Selected Spectral Attributes in the Perception of Synthetic Vowels
120. **Kristian Nybom**, Low-Density Parity-Check Codes for Wireless Datacast Networks
121. **Johanna Tuominen**, Formal Power Analysis of Systems-on-Chip
122. **Teijo Lehtonen**, On Fault Tolerance Methods for Networks-on-Chip
123. **Eeva Suvitie**, On Inner Products Involving Holomorphic Cusp Forms and Maass Forms
124. **Linda Mannila**, Teaching Mathematics and Programming – New Approaches with Empirical Evaluation
125. **Hanna Suominen**, Machine Learning and Clinical Text: Supporting Health Information Flow
126. **Tuomo Saarni**, Segmental Durations of Speech
127. **Johannes Eriksson**, Tool-Supported Invariant-Based Programming
128. **Tero Jokela**, Design and Analysis of Forward Error Control Coding and Signaling for Guaranteeing QoS in Wireless Broadcast Systems
129. **Ville Lukkarila**, On Undecidable Dynamical Properties of Reversible One-Dimensional Cellular Automata
130. **Qaisar Ahmad Malik**, Combining Model-Based Testing and Stepwise Formal Development
131. **Mikko-Jussi Laakso**, Promoting Programming Learning: Engagement, Automatic Assessment with Immediate Feedback in Visualizations
132. **Riikka Vuokko**, A Practice Perspective on Organizational Implementation of Information Technology
133. **Jeanette Heidenberg**, Towards Increased Productivity and Quality in Software Development Using Agile, Lean and Collaborative Approaches
134. **Yong Liu**, Solving the Puzzle of Mobile Learning Adoption
135. **Stina Ojala**, Towards an Integrative Information Society: Studies on Individuality in Speech and Sign
136. **Matteo Brunelli**, Some Advances in Mathematical Models for Preference Relations
137. **Ville Junnila**, On Identifying and Locating-Dominating Codes

TURKU
CENTRE *for*
COMPUTER
SCIENCE

Joukahaisenkatu 3-5 B, 20520 Turku, Finland | www. tucs.fi

University of Turku
Faculty of Mathematics and Natural Sciences
- Department of Information Technology
- Department of Mathematics

Turku School of Economics
- Institute of Information Systems Science

Åbo Akademi University
Division for Natural Sciences and Technology
- Department of Information Technologies

ISBN 978-952-12-2598-7
ISSN 1239-1883

Ville Junnila On Identifying and Locating-Dominating Codes

Ville Junnila On Identifying and Locating-Dominating Codes

Ville Junnila On Identifying and Locating-Dominating Codes