TUCS

Sanna Ranto

Identifying and Locating-Dominating Codes in Binary Hamming Spaces

Turku Centre *for* Computer Science

TUCS Dissertations
No 86, April 2007

Identifying and Locating-Dominating Codes in Binary Hamming Spaces

by

Sanna Ranto

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 of the University on May 16th, 2007, at 12 noon

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

2007

SUPERVISOR

PROFESSOR IIRO HONKALA
Department of Mathematics
University of Turku
FIN-20014 Turku
Finland

REVIEWERS

PROFESSOR OLIVIER HUDRY
Centre National de la Recherche Scientifique
Département Informatique et Réseaux
École Nationale Supérieure des Télécommunications
46 rue Barrault, 75634 PARIS cedex 13
France

PROFESSOR MARK KARPOVSKY
Department of Electrical and Computer Engineering
Boston University
8 Saint Mary’s Street
Boston, Massachusetts 02215
USA

OPPONENT

CHARGÉ DE RECHERCHE ANTOINE LOBSTEIN
Centre National de la Recherche Scientifique
Département Informatique et Réseaux
École Nationale Supérieure des Télécommunications
46 rue Barrault, 75634 PARIS cedex 13
France

ISBN 978-952-12-1880-4
ISSN 1239-1883
Painosalama Oy
Turku, Finland
2007

Acknowledgements

First of all I want to thank my supervisor Professor Iiro Honkala for his continuous support, and patience when long breaks in research took place. Many discussions with Iiro, his collaboration, suggestions for research topics, and careful proofreading made this work possible.

I also want to thank Dr. Tero Laihonen for many inspiring discussions, suggestions for research topics, and for fruitful collaboration.

Professor Aimo Tietäväinen encouraged me to begin my postgraduate studies, and I am grateful for that.

I extend my thanks to co-author Professor Geoffrey Exoo, and the reviewers of this thesis, Professor Olivier Hudry and Professor Mark Karpovsky.

I am grateful for financial support from Turku Centre for Computer Science (TUCS) and Department of Mathematics. I acknowledge the staff at these institutions for creating pleasant working environment.

I acknowledge and thank the Nokia Foundation and the August and Lyydia Heino scholarship fund for partial financial support of my work.

Finally, I want to thank my husband Kalle for his loving support and for reading my thesis. The thesis is dedicated to my lovely sons Samuli, Olli, and Vilho, who, throughout this process supplied numerous kisses and hugs.

Turku
April 2007 Sanna Ranto

Contents

Chapter 1

Introduction

We consider a connected graph $G = (V, E)$ with a vertex set V and an edge set E. A *code* is a nonempty subset of the vertex set V, and a vertex belonging to a code is called a *codeword*. A code is called an *identifying code*, if all the vertices of V have different nonempty codeword neighbourhoods. A code is called a *locating-dominating code* if all the vertices of V not belonging to the code have different nonempty codeword neighbourhoods. Usually, the neighbourhood is the ball of radius r. This means that the codeword neighbourhood of a vertex x in the vertex set V consists of codewords at distance at most r from x. In this case we also speak about *r-identifying* and *r-locating-dominating* codes. The problem is to determine the smallest number of codewords in identifying and locating-dominating codes in different graphs. A code attaining the smallest possible cardinality is called *optimal*.

The research on identifying codes started in 1998 from the paper of Karpovsky, Chakrabarty and Levitin [40]. Since then the theory of identifying codes has been investigated in various graphs such as infinite grids [1, 8, 10, 14, 27, 28], general graphs [12, 19, 51], trees [4, 7] and paths and cycles [3, 22]. Locating-dominating codes were introduced by Slater [60], see also [16, 58]. Other results on locating-dominating codes can be found in [12, 16, 25, 29, 54, 59, 61, 62]. Finding an identifying or locating-dominating code of an optimal cardinality and determining whether or not a given code is identifying or locating-dominating are computationally difficult problems. For these and other complexity results, see [9, 11, 15, 38, 39]. There are many other papers written about identifying and locating-dominating codes than the ones mentioned here. The reader can consult the up-to-date internet bibliography [50] on identifying and locating-dominating codes maintained by Antoine Lobstein.

In this thesis, we consider identifying and locating-dominating codes in binary Hamming spaces or hypercubes. In the binary hypercube, the vertices are labelled by the 2^n binary words of length n, and two vertices are adjacent if and only if their labels differ in exactly one position. Identifying codes in binary Hamming spaces have been considered in [5, 6, 24, 26, 31, 32, 37, 40, 44–46, 52]. In [40] there are lower bounds for the cardinalities of r-identifying codes and some constructions which use covering codes. In [5] r-identifying codes are constructed for $r \geq 2$. In [37] it is shown that for a constant $\rho \in [0,1)$ we have $\lim_{n\to\infty} n^{-1} \log_2 M_{\lfloor \rho n \rfloor}(n) = 1 - H(\rho)$, where $M_r(n)$ is the smallest cardinality of an r-identifying code of length n and $H(x) = x\log_2 x - (1-x)\log_2(1-x)$ is the binary entropy function. In [52] it is shown that the optimal cardinality of 1-identifying codes is a monotonically increasing function with respect to code length. References [44, 45] are about codes which identify sets of vertices of size at most ℓ, when $\ell \geq 3$. In [45] an infinite sequence of optimal $(1, \leq \ell)$-identifying codes is constructed by showing that the problem of finding an optimal $(1, \leq \ell)$-identifying code is equivalent to the problem of finding an optimal $(2\ell - 1)$-fold 1-covering. In [44] an infinite optimal sequence of strongly $(1, \leq \ell)$-identifying codes is constructed. Robust identifying codes are considered in binary Hamming spaces in [26, 31, 32, 46].

Identifying codes can be applied in fault diagnosis of multiprocessor systems. The purpose of fault diagnosis is to locate malfunctioning processors by testing the system. Processors are the vertices of the graph and the links between the processors are represented as the edges of the graph. A set of processors is chosen for checking themselves and their neighbours and report to an outside controller. Each chosen processor sends a message 1 if there is something wrong in the processor itself or some of its neighbours and 0 otherwise. The processors are chosen so that they form an identifying code. Depending on the code the controller can identify one or more malfunctioning processors in each test run. This is the initial application of identifying codes from [40]. A recent application area of identifying codes is sensor networks, like environmental monitoring. For more information about current applications, see [2, 26, 42, 43, 57].

The initial application of locating-dominating codes in [60] is the safeguard analysis of a facility using a fire and intruder alarm system. Now the vertices of a graph represent rooms, corridors, courtyards etc. Each edge connects two areas that are physically adjacent or they are within a sight or sound connection with each other. Sensors are placed in some of the vertices. These sensors each send a message to a controller. In this case there are three possible messages: 0 if no problem is detected, 1 if there is a problem in some neighbouring vertex and 2 if a problem is at the vertex itself.

The structure of the thesis is the following. In Chapter 2, we define identifying codes precisely and introduce some notations that are needed later.

In Chapter 3 we consider r-identifying codes which can identify one fault. The chapter begins with a lower bound for r-identifying codes. This bound improves on the previously known lower bounds for r-identifying codes [40]. We give a construction of 1-identifying codes which gives the smallest known cardinalities for many lengths. We construct r-identifying codes, for $r \geq 2$, by taking a direct sum of r 1-identifying codes. This construction partly solves an open problem in [6]. The results of this chapter are from [18, 49].

In Chapter 4 we consider identifying codes which can identify any at most ℓ faulty vertices when $\ell \geq 2$ is fixed. The chapter begins with lower bounds and general notions and continues with short optimal codes. The main results of this chapter are the constructions in Section 4.3. For example, we have a construction which gives us two infinite families of optimal $(1, \leq 2)$-identifying codes. We also prove that the direct sum of r $(1, \leq \ell)$-identifying codes is an $(r, \leq \ell)$-identifying code. This chapter is based on the papers [34, 47, 49, 56]. We also give results which are not included in these papers.

We consider strongly identifying codes in Chapter 5. In this variant of identifying codes a malfunctioning processor in a fault diagnosis system may or may not send the correct message. Because we cannot trust the received information we have to assume somewhat more than in the basic identifying case. The results of this chapter are from [33, 35, 47, 48].

In Chapter 6 we consider linear identifying codes. An identifying code is linear if it is a subspace of a Hamming space. For linear 1-identifying codes the optimal cardinalities are solved in all cases: for all lengths and for any number ℓ of faulty vertices. The optimal cardinalities are also solved for linear strongly 1-identifying codes. Results of this chapter are from [55].

In Chapter 7 we consider locating-dominating codes in Hamming spaces. We first consider locating-dominating codes which can locate one faulty vertex among non-codewords. We also give two different definitions for locating-dominating codes locating more than one faulty vertex. In both cases, we prove lower bounds and give some constructions. This chapter is based on [36].

In the Appendix, we list some identifying and locating-dominating codes found using a computer.

Chapter 2

Preliminaries

Let $\mathbf{F} = \{0,1\}$ denote the binary field. The binary *Hamming space* $\mathbf{F}^n$ or binary *hypercube* is the n-fold Cartesian product of $\mathbf{F}$. A *code* is a subset of $\mathbf{F}^n$. The elements of $\mathbf{F}^n$ are called *words* or *vectors*, and the elements of a code are called *codewords*. The cardinality of a set X is denoted by $|X|$.

The *Hamming distance* between words $\mathbf{x} = (x_1, \ldots, x_n)$, $\mathbf{y} = (y_1, \ldots, y_n) \in \mathbf{F}^n$ is the number of places in which they differ, that is,

$$d(\mathbf{x}, \mathbf{y}) := |\{i \mid x_i \neq y_i\}|.$$

The set of non-zero coordinates of a word $\mathbf{x} \in \mathbf{F}^n$ is called the *support* of $\mathbf{x}$ and is denoted by $supp(\mathbf{x})$. The cardinality of the support of $\mathbf{x}$ is called the *weight* of $\mathbf{x}$ and is denoted by $w(\mathbf{x})$. For $\mathbf{x} \in \mathbf{F}^n$ we denote

$$\begin{aligned} B_r(\mathbf{x}) &= \{\mathbf{y} \in \mathbf{F}^n \mid d(\mathbf{x}, \mathbf{y}) \leq r\}, \\ S_r(\mathbf{x}) &= \{\mathbf{y} \in \mathbf{F}^n \mid d(\mathbf{x}, \mathbf{y}) = r\}. \end{aligned}$$

The set $B_r(\mathbf{x})$ is called the *Hamming sphere* or *ball of radius r* centred at $\mathbf{x}$. The size of the Hamming sphere of radius r in $\mathbf{F}^n$ does not depend on the choice of the centre and it is denoted by $V(n,r)$. We have

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

The next lemma will often be used.

Lemma 2.1. *Let* $\mathbf{x}, \mathbf{y} \in \mathbf{F}^n$. *Then*

$$|B_1(\mathbf{x}) \cap B_1(\mathbf{y})| = \begin{cases} n+1 & \textit{if } \mathbf{x} = \mathbf{y}, \\ 2 & \textit{if } 1 \leq d(\mathbf{x}, \mathbf{y}) \leq 2, \\ 0 & \textit{otherwise}. \end{cases}$$

The intersection of three different Hamming spheres of radius one contains at most one element.

Proof. Denote by $\mathbf{e}_i$ the word which support is $\{i\}$ and moreover $\mathbf{e}_0 = \mathbf{0}$. For $0 \leq i \leq n$ we have $\mathbf{x} + \mathbf{e}_i \in B_1(\mathbf{x}) \cap B_1(\mathbf{y})$ if and only if $\mathbf{x} + \mathbf{e}_i = \mathbf{y} + \mathbf{e}_k$ for some $0 \leq k \leq n$. This is equivalent to $\mathbf{x} = \mathbf{y} + \mathbf{e}_i + \mathbf{e}_k$. This implies the first claim.

Suppose that $1 \leq d(\mathbf{x}, \mathbf{y}) \leq 2$, then for $\mathbf{e}_i \neq \mathbf{e}_k$ we have $\mathbf{x} + \mathbf{e}_i, \mathbf{x} + \mathbf{e}_k \in B_1(\mathbf{x}) \cap B_1(\mathbf{y})$. If for some $\mathbf{z} \in \mathbf{F}^n$ we have $|B_1(\mathbf{x}) \cap B_1(\mathbf{y}) \cap B_1(\mathbf{z})| = 2$, then $B_1(\mathbf{z})$ has to include those two words which belong to $B_1(\mathbf{x}) \cap B_1(\mathbf{y})$. That means, for some $\mathbf{e}_j$ and $\mathbf{e}_h$ we have $\mathbf{x} + \mathbf{e}_i = \mathbf{y} + \mathbf{e}_k = \mathbf{z} + \mathbf{e}_j$ and $\mathbf{x} + \mathbf{e}_k = \mathbf{y} + \mathbf{e}_i = \mathbf{z} + \mathbf{e}_h$. This is equivalent to $\mathbf{x} = \mathbf{z} + \mathbf{e}_i + \mathbf{e}_j = \mathbf{z} + \mathbf{e}_k + \mathbf{e}_h$. Thus, $\{\mathbf{e}_i, \mathbf{e}_k\} = \{\mathbf{e}_h, \mathbf{e}_j\}$. If $\mathbf{e}_i = \mathbf{e}_h$, then $\mathbf{y} = \mathbf{z}$. If $\mathbf{e}_i = \mathbf{e}_j$, then $\mathbf{x} = \mathbf{z}$. It is clear that in the intersection of three different words there is one word if the words are suitably chosen. □

Let $C \subseteq \mathbf{F}^n$ be a code. For any $X \subseteq \mathbf{F}^n$ we denote

$$I_r(X) = I_r(C; X) = \bigcup_{\mathbf{x} \in X} (B_r(\mathbf{x}) \cap C).$$

The set $I_r(X)$ is called the *I-set* of X. If $r = 1$, we denote $I(X) = I(C; X)$ for short. We also denote $I_r(\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_\ell\}) = I_r(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_\ell) = I_r(C; \mathbf{x}_1, \ldots \mathbf{x}_\ell)$. The set X that we try to identify is called the *fault pattern*.

Definition 2.2. Let r and ℓ be non-negative integers. We say that $C \subseteq \mathbf{F}^n$ is an *$(r, \leq \ell)$-identifying code* if for all $X, Y \subseteq \mathbf{F}^n$, $X \neq Y$, such that $|X| \leq \ell$ and $|Y| \leq \ell$ we have $I_r(X) \neq I_r(Y)$. If $\ell = 1$, we say C is an *r-identifying* code.

The *symmetric difference* of sets A and B is

$$A \triangle B = (A \setminus B) \cup (B \setminus A).$$

In other words, we can say that $C \subseteq \mathbf{F}^n$ is $(r, \leq \ell)$-identifying if and only if for all $X, Y \subseteq \mathbf{F}^n$, $|X| \leq \ell$ and $|Y| \leq \ell$ we have

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

The smallest cardinality of an $(r, \leq \ell)$-identifying code of length n is denoted by $M_r^{(\leq \ell)}(n)$. If $\ell = 1$, we denote the smallest cardinality by $M_r(n)$. A code attaining the smallest cardinality is called *optimal*.

If $\ell = 0$, the requirement for an identifying code is empty and the code is trivial. If $r = 0$, then obviously $M_0^{(\leq \ell)}(n) = 2^n$. From now on we assume that $r \geq 1$ and $\ell \geq 1$.

In the constructions of identifying codes we often encounter covering codes. For further information of covering codes, see [13]. Let us recall some definitions from therein. We say that $\mathbf{x} \in \mathbf{F}^n$ *r-covers* $\mathbf{y} \in \mathbf{F}^n$ if $d(\mathbf{x}, \mathbf{y}) \leq r$.

Definition 2.3. A code $C \subseteq \mathbf{F}^n$ has *covering radius* r if r is the smallest integer such that for all $\mathbf{x} \in \mathbf{F}^n$ there is a codeword $\mathbf{c} \in C$ that r-covers $\mathbf{x}$.

The smallest cardinality of a binary code with length n and covering radius r is denoted by $K(n,r)$. There is a well-known conjecture concerning the covering radius, see [13, p. 352]. According to the conjecture the best r-covering codes are asymptotically perfect.

Conjecture 2.4. *For a fixed r*

$$\lim_{n\to\infty} \frac{K(n,r)V(n,r)}{2^n} = 1.$$

Next we define multiple covering codes.

Definition 2.5. A code $C \subseteq \mathbf{F}^n$ is a *μ-fold r-covering* if every word in $\mathbf{F}^n$ is r-covered by at least μ distinct codewords of C.

The smallest cardinality of a μ-fold r-covering of length n is denoted by $K(n, \mu, r)$.

The *sum* of vectors $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_n)$ is

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, \ldots, x_n + y_n).$$

The *sum* of sets $X, Y \subseteq \mathbf{F}^n$ is

$$X + Y = \{\mathbf{x} + \mathbf{y} \mid \mathbf{x} \in X, \mathbf{y} \in Y\}.$$

The *direct sum* of codes $C_1 \subseteq \mathbf{F}^{n_1}$ and $C_2 \subseteq \mathbf{F}^{n_2}$ is

$$C_1 \oplus C_2 = \{(\mathbf{c}_1, \mathbf{c}_2) \in \mathbf{F}^{n_1+n_2} \mid \mathbf{c}_1 \in C_1,\ \mathbf{c}_2 \in C_2\}.$$

We denote by $\pi(\mathbf{u})$ the parity check bit of $\mathbf{u}$, i.e.,

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

Chapter 3

On identifying codes

In this chapter, we consider r-identifying codes when $r = 1$ and $r \geq 2$. In the case $r = 1$, our main purpose is to construct longer 1-identifying codes from 1-identifying codes that are 2-fold 1-coverings. For $r \geq 2$, we prove $M_r(\sum_{i=1}^{r} n_i) \leq \prod_{i=1}^{r} M(n_i)$. We begin by introducing a lower bound for r-identifying codes which improves on the previously known lower bounds for $r \geq 2$ and n large enough.

This chapter is based on [18] except Theorem 3.3 with corollaries is from [49].

3.1 Lower bounds

The following lower bound for r-identifying codes is from [40, Theorem 1 (iii)]. We mention the result here without the proof.

Theorem 3.1 (Karpovsky et al. [40]). *Let K be the smallest integer such that for a certain integer t $(1 \leq t \leq \min\{K, V(n,r)\})$ the following conditions are true:*

$$\sum_{j=0}^{t-2} \binom{K-1}{j} < V(n,r) \leq \sum_{j=0}^{t-1} \binom{K-1}{j}$$

$$2^n \leq \sum_{j=1}^{t-1} \binom{K}{j} + \left\lfloor \frac{K}{t} \left(V(n,r) - \sum_{j=0}^{t-2} \binom{K-1}{j} \right) \right\rfloor .$$

Then $M_r(n) \geq K$.

The previous lower bound coincides with the next lower bound from [40, Theorem 3] for all r and n large enough. There are at most $|C|$ words r-covered at most by one codeword in an r-identifying code C. All the other words are covered at least by two codewords. Thus, by counting the number of pairs $\{\mathbf{x}, \mathbf{c}\}$ where $\mathbf{x} \in \mathbf{F}^n$, $\mathbf{c} \in C$ and $d(\mathbf{c}, \mathbf{x}) \leq r$ in two ways we get the next lower bound from [40].

Theorem 3.2 (Karpovsky et al. [40]).

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

The next theorem improves on the lower bounds of the previous theorems for $r \geq 2$. As a corollary, we also get the best known lower bound for 1-identifying codes from [40] (see also [6]).

Theorem 3.3. *Let $C \subseteq \mathbf{F}^n$ be r-identifying and $m = \max\{|I_r(\mathbf{x})| : \mathbf{x} \in \mathbf{F}^n\}$. Denote*

$$f_r(x) = \frac{(x-2)(\binom{2r}{r}-1)}{\binom{2r}{r}+\binom{x}{2}-1}.$$

We have

$$|C| \geq \frac{2^n(2+f_r(v))}{V(n,r)+f_r(v)+1}.$$

where $v = m$, if $m \geq 2+2\binom{2r}{r}$, and $v = 3$ otherwise.

Proof. Let $C \subseteq \mathbf{F}^n$ be an r-identifying code. Denote by V_i the words r-covered by exactly i codewords. There are at most $K = |C|$ words which are r-covered by one codeword. All the other words are r-covered at least by 2 codewords. Let $\mathbf{x} \in \mathbf{F}^n$ be r-covered by exactly two codewords, $I_r(\mathbf{x}) = \{\mathbf{c}_1, \mathbf{c}_2\}$. Now $1 \leq d(\mathbf{c}_1, \mathbf{c}_2) \leq 2r$. When $d(\mathbf{c}_1, \mathbf{c}_2) = 2r$ there are $\binom{2r}{r}$ words covering both of these codewords. If $d(\mathbf{c}_1, \mathbf{c}_2) < 2r$, then by [13, Theorem 2.4.8] we know that there are at least $\binom{2r}{r}$ words covering both of these words. Hence, by the definition of identifying codes, for each word which is r-covered by two codewords there are at least $\binom{2r}{r}-1$ words which are r-covered by three or more codewords. On the other hand, if $\mathbf{y} \in \mathbf{F}^n$ is r-covered by $i \geq 3$ codewords, then there can be at most $\binom{i}{2}$ words $\mathbf{z}$ 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 $\mathbf{x} \in V_2$ and $\mathbf{y} \in V_i$ $(i \geq 3)$ and $I_r(\mathbf{x}) \subseteq I_r(\mathbf{y})$, we have

$$(\binom{2r}{r}-1)|V_2| \leq \sum_{i=3}^{m} \binom{i}{2}|V_i|. \tag{3.1}$$

For any positive real number a we get by counting in two ways the number of pairs $\{\mathbf{x}, \mathbf{c}\}$, where $\mathbf{x} \in \mathbf{F}^n$ and $\mathbf{c} \in C$ such that $d(\mathbf{x}, \mathbf{c}) \leq r$

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

We require that a satisfies $i-2-a-\frac{a}{\binom{2r}{r}-1}\binom{i}{2}\geq 0$ for all $3\leq i\leq m$. Hence, we get

$$a\leq\frac{(i-2)(\binom{2r}{r}-1)}{\binom{2r}{r}+\binom{i}{2}-1}=:f_r(i) \qquad (3.2)$$

The function f_r is decreasing when $i\geq 2+\sqrt{2\binom{2r}{r}}$, f_r is increasing for $3\leq i\leq 2+\sqrt{2\binom{2r}{r}}$ and $f_r(3)=f_r(2+2\binom{2r}{r})$. Thus, when $m\geq 2+2\binom{2r}{r}$ we can choose $a=f_r(m)$ and otherwise $a=f_r(3)$. This implies (in both cases)

$$K\cdot V(n,r)\geq K+(2+f_r(v))(2^n-K),$$

from which the claim follows. □

Remark 3.4. Let A be a set of codewords in $B_r(\mathbf{0})$ such that every word in $B_{2r}(\mathbf{0})$ is r-identified by using only the words of A. Suppose $C\subseteq\mathbf{F}^n$ is an r-identifying code. If there is $\mathbf{y}\in\mathbf{F}^n$ such that $|I_r(\mathbf{y})|>|A|$, then the code $C':=(C\setminus I_r(\mathbf{y}))\cup D$, where $D:=\{\mathbf{a}+\mathbf{y}\mid\mathbf{a}\in A\}$, is still r-identifying and $|C'|<|C|$. If the code C' r-covers again some word more than $|A|$ times we continue with the same process. In each step the cardinality of the code is getting smaller and every code is r-identifying. Hence, the process will stop at some point and the result will be an r-identifying code which r-covers each word in $\mathbf{F}^n$ at most $|A|$ times. Consequently, we can use $m\leq|A|$ in the previous theorem. Moreover, if C is r-identifying and attains $M_r(n)$, then we immediately know that $|I_r(\mathbf{x})|\leq|A|$ for all $\mathbf{x}\in\mathbf{F}^n$.

By [40, Theorem 4] we know that $S_1(\mathbf{0})$ 1-identifies $B_2(\mathbf{0})$. By substituting $r=1$ and $m=|S_1(\mathbf{0})|=n$ to the Theorem 3.3 we get the following best known lower bound for 1-identifying codes due to Karpovsky, Chakrabarty and Levitin [40] (see also [6]).

Corollary 3.5 (Karpovsky et al. [40]).

$$M_1(n)\geq\frac{n2^{n+1}}{2+n+n^2}=\frac{n2^n}{V(n,2)}.$$

In [5, Construction 3] it is proven that all the words of weight two except those which supports are $\{i,i+1\}$ for $i=1,\dots,n-1$ and $\{1,n\}$ (when $n\geq 7$) 2-identify all the words in the set $B_4(\mathbf{0})$. Hence, $\binom{n}{2}-n$ words in $S_2(\mathbf{0})$ 2-identify words in $B_4(\mathbf{0})$ when $n\geq 7$. (When we say the set $A\subseteq\mathbf{F}^n$ *r-identifies* the set $B\subseteq\mathbf{F}^n$, we mean that for all $\mathbf{x}\in B$ and $\mathbf{y}\in\mathbf{F}^n$ we have $I_r(A;\mathbf{x})\neq I_r(A;\mathbf{y})$.) Taking $m=\binom{n}{2}-n$ in Theorem 3.3 one obtains the following.

Corollary 3.6. *For $n\geq 7$ we have*

$$M_2(n)\geq\frac{2^{n+2}(n^3-6n^2+17n-24)}{n^5-5n^4+5n^3-11n^2+114n-56}.$$

In [40, Theorem 5] it is proven that the set $S_r(\mathbf{0})$ r-identifies all the words in $B_{2r}(\mathbf{0})$ provided that $r < n/2$. Choosing $m = \binom{n}{r}$ implies the next result.

Corollary 3.7. *When $\binom{n}{r} \geq 2+2\binom{2r}{r}$ and $r < n/2$, we have*

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

We can apply also other constructions of [5] as the set A in the remark above. Using these results we get the following corollaries.

Corollary 3.8. *Let*

$$R = 2\binom{\lceil n/2 \rceil}{r-1}\lceil n/2 \rceil + 2\binom{\lceil n/2 \rceil}{r}.$$

When $V(n,r-1)+R \geq 2+2\binom{2r}{r}$ we have

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

Corollary 3.9. *When $2r-1$ divides n and*

$$R = \binom{n}{r} - \binom{2r-1}{r}\left(\frac{n}{2r-1}\right)^r$$

we have

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

The lower bounds of Theorem 3.1 and Theorem 3.2 coincide (see [40]) for every r when n is large enough. It is easy to see that

$$\frac{2^n(2+f_r(x))}{V(n,r)+f_r(x)+1} > \frac{2^{n+1}}{V(n,r)+1}$$

for all $x > 2$ and $n,r \geq 1$. Hence, the lower bound of Theorem 3.3 is always stronger than the lower bound of Theorem 3.2. Thus, for every fixed r there is n_0 such that for all $n \geq n_0$ we improve on the lower bound of Theorem 3.1. For example, Theorem 3.3 and the corollaries improve Theorem 3.1 for $r = 2$ when $n \geq 10$, $r = 3$ when $n \geq 20$, $r = 4$ when $n \geq 29$, and $r = 5$ when $n \geq 37$.

In the next section we use 2-fold 1-covering codes that are 1-identifying to construct new 1-identifying codes. The proof for the lower bound of 1-identifying codes that are 2-fold 1-covered goes similarly to the proof of the previous theorem. For the proof see [18]. In 2-fold 1-covering code that is 1-identifying there can be $(n+1)$-fold 1-covered codewords.

Theorem 3.10. *If $C \subseteq \mathbf{F}^n$ $(n \geq 2)$ is a 2-fold 1-covering and a 1-identifying code, then*

$$|C| \geq \left\lceil \frac{2^{n+1}(n+1)}{n^2+n+2} \right\rceil. \tag{3.3}$$

3.2 Constructions for 1-identifying codes

In Theorem 3.11 we apply the classical $(\pi(\mathbf{u}), \mathbf{u}, \mathbf{u}+\mathbf{v})$-construction to 2-fold 1-coverings of length n that are 1-identifying and get 2-fold 1-covering and 1-identifying codes of length $2n+1$.

Theorem 3.11. *Suppose $C \subseteq \mathbf{F}^n$ is a 2-fold 1-covering and a 1-identifying code. Then*

$$D = \{(\pi(\mathbf{u}), \mathbf{u}, \mathbf{u}+\mathbf{v}) \mid \mathbf{u} \in \mathbf{F}^n, \mathbf{v} \in C\} \subseteq \mathbf{F}^{2n+1}$$

is a 2-fold 1-covering which is 1-identifying.

Proof. By [13, Theorem 14.4.3] we know that D is a 2-fold 1-covering in $\mathbf{F}^{2n+1}$.

Let $\mathbf{x} = (a, \mathbf{u}, \mathbf{u}+\mathbf{v}) \in \mathbf{F}^{2n+1}$. We divide the words of $\mathbf{F}^{2n+1}$ into two classes depending on the first bit.

I If $a = \pi(\mathbf{u})$, then $I(\mathbf{x}) = \{(a, \mathbf{u}, \mathbf{u}+\mathbf{c}) \mid \mathbf{c} \in C,\ d(\mathbf{c}, \mathbf{v}) \leq 1\}$.

II If $a \neq \pi(\mathbf{u})$, then $I(\mathbf{x}) = A \cup \{(a, \mathbf{u}', \mathbf{u}'+\mathbf{c}) \mid \mathbf{c} \in C,\ d(\mathbf{u}', \mathbf{u}) = 1,\ \mathbf{u}'+\mathbf{c} = \mathbf{u}+\mathbf{v}\}$, where $A = \{(\pi(\mathbf{u}), \mathbf{u}, \mathbf{u}+\mathbf{v})\}$ if $\mathbf{v} \in C$ and otherwise A is empty.

If $|I(\mathbf{x})| \geq 3$, then it is clear by Lemma 2.1 that $\mathbf{x}$ is identified. Thus, suppose $|I(\mathbf{x})| = 2$ this implies that also $|I(\mathbf{v})| = 2$. We shall show that there does not exist a word $\mathbf{y}$ ($\mathbf{y} \neq \mathbf{x}$) such that $I(\mathbf{y}) = I(\mathbf{x})$.

Assume first that $a = \pi(\mathbf{u})$. Now $I(\mathbf{x}) = \{(a, \mathbf{u}, \mathbf{u}+\mathbf{c}_1), (a, \mathbf{u}, \mathbf{u}+\mathbf{c}_2)\}$ where $\mathbf{c}_1, \mathbf{c}_2 \in C$, $d(\mathbf{v}, \mathbf{c}_1) \leq 1$ and $d(\mathbf{v}, \mathbf{c}_2) = 1$. Because the first $(n+1)$-bits of the words in the I-set of $\mathbf{x}$ are the same this implies that if $I(\mathbf{y}) = I(\mathbf{x})$, then also $\mathbf{y}$ belongs to the class I. Thus, $\mathbf{y} = (a, \mathbf{u}, \mathbf{u}+\mathbf{w})$ where $d(\mathbf{w}, \mathbf{c}_1) \leq 1$ and $d(\mathbf{w}, \mathbf{c}_2) \leq 1$. If $I(\mathbf{y}) = I(\mathbf{x})$, then also $I(\mathbf{v}) = I(\mathbf{w})$. This is impossible, because C is 1-identifying and, clearly, $\mathbf{v} \neq \mathbf{w}$.

If $a \neq \pi(\mathbf{u})$ and $\mathbf{v} \notin C$, then $I(\mathbf{x}) = \{(a, \mathbf{u}_1, \mathbf{u}_1+\mathbf{c}_1), (a, \mathbf{u}_2, \mathbf{u}_2+\mathbf{c}_2)\}$ where $d(\mathbf{u}_i, \mathbf{u}) = 1$ and $\mathbf{u}+\mathbf{v} = \mathbf{u}_i+\mathbf{c}_i$ for $i = 1,2$. Thus, $d(\mathbf{v}, \mathbf{c}_1) = d(\mathbf{v}, \mathbf{c}_2) = 1$. If $I(\mathbf{y}) = I(\mathbf{x})$, then also $\mathbf{y}$ belongs to the class II because the n bits following the first bit are changing in the words of $I(\mathbf{y})$. Now $\mathbf{y} = (a, \mathbf{u}', \mathbf{u}'+\mathbf{w})$ where $\mathbf{u}'$ is the unique word in $(B_1(\mathbf{u}_1) \cap B_1(\mathbf{u}_2)) \setminus \{\mathbf{u}\}$ and $\mathbf{u}'+\mathbf{w} = \mathbf{u}_1+\mathbf{c}_1 = \mathbf{u}_2+\mathbf{c}_2$, thus $\mathbf{w}$ is the unique word in $(B_1(\mathbf{c}_1) \cap B_1(\mathbf{c}_2)) \setminus \{\mathbf{v}\}$. If $I(\mathbf{y}) = I(\mathbf{x})$, then also $I(\mathbf{w}) = I(\mathbf{v})$. That is impossible.

Suppose finally $a \neq \pi(\mathbf{u})$ and $\mathbf{v} \in C$. Then $I(\mathbf{x}) = \{(\pi(\mathbf{u}), \mathbf{u}, \mathbf{u}+\mathbf{v}), (a, \mathbf{u}', \mathbf{u}'+\mathbf{c}')\}$ where $d(\mathbf{u}, \mathbf{u}') = 1$, $d(\mathbf{c}', \mathbf{v}) = 1$ and $\mathbf{u}'+\mathbf{c}' = \mathbf{u}+\mathbf{v}$. If $I(\mathbf{x}) = I(\mathbf{y})$, then $\mathbf{y} = (\pi(\mathbf{u}), \mathbf{u}', \mathbf{u}'+\mathbf{c}')$ because the first bit is changing in the words of $I(\mathbf{x})$. But if $I(\mathbf{x}) = I(\mathbf{y})$, then also $I(\mathbf{v}) = I(\mathbf{c}')$ and this cannot happen. □

The next remark is from [56].

Remark 3.12. In general, the $(\pi(\mathbf{u}),\mathbf{u},\mathbf{u}+\mathbf{v})$-construction does not directly apply to 1-identifying codes. Let $C=\{000,110,101,011\}\subseteq\mathbf{F}^3$. The code C is 1-identifying. Applying the $(\pi(\mathbf{u}),\mathbf{u},\mathbf{u}+\mathbf{v})$ -construction to the code C we get a code C' of length 7. In C' we have $I(0000011)=I(1000011)=\{0000011\}$ and therefore C' is not 1-identifying.

Example 3.13. Theorem 3.10 shows that for the lengths $2,3$ and 4 the codes that are presented next are optimal 2-fold 1-coverings which are 1-identifying.

1. Words $01,10,11$ form a 2-fold 1-covering and a 1-identifying code of length 2 and cardinality 3.
2. Words 100, 010, 011, 101, 111 form a 2-fold 1-covering and a 1-identifying code of length 3 and cardinality 5.
3. Words 0010, 0001, 1100, 1010, 0110, 0101, 1101, 1011 form a 2-fold 1-covering and a 1-identifying code of length 4 and cardinality 8.
4. The following words form a 2-fold 1-covering and a 1-identifying code of length 6 and cardinality 22 (lower bound by Theorem 3.10 is 21).

   000000 011111 110010 101101 000001 011110
   101000 110111 110110 101001 110100 101011
   011100 000011 001110 010001 100110 111001
   011010 000101 001001 111110

Combining Theorem 3.11 with the small codes of the previous example we get the next corollary.

Corollary 3.14. *For any $k\geq 0$ we have:*

$$\begin{aligned}
M_1(3\cdot 2^k-1) &\leq 3\cdot 2^{3\cdot(2^k-1)-k}\\
M_1(2^{k+2}-1) &\leq 5\cdot 2^{2^{k+2}-k-4}\\
M_1(5\cdot 2^k-1) &\leq 2^{5\cdot 2^k-2-k}\\
M_1(7\cdot 2^k-1) &\leq 11\cdot 2^{7\cdot 2^k-k-6}
\end{aligned}$$

Let a be the ratio between the sphere bound and the cardinality of a code $C\subseteq\mathbf{F}^n$, i.e.,

$$a=\frac{|C|}{2^n/(n+1)}.$$

Applying the $(\pi(\mathbf{u}),\mathbf{u},\mathbf{u}+\mathbf{v})$-construction to the code (and then to the resulting code and so on) we get an infinite family of codes, whose corresponding ratio is always a. For example, starting with the first code in Example 3.13 we get for all $k\geq 0$ and $n=3\cdot 2^k-1$:

$$\frac{M_1(n)}{2^n/(n+1)}\leq 2.25.$$

By Corollary 3.5 we know that $2 \lesssim a$ when n tends to infinity.

For higher lengths, good (meaning $a \approx 2$) initial codes can be obtained using the codes

$$C = \{\mathbf{c}+\mathbf{b} \mid \mathbf{c} \in A,\ \mathbf{b} \in B_1(\mathbf{0})\}$$

where A is a 1-fold 2-covering given in [64]. Clearly, C is a 2-fold 1-covering in $\mathbf{F}^n$ and it is also 1-identifying by [40]. The asymptotic results from [37, 40] for the size of usual 1-identifying codes are obtained (using C) in the same way also for 2-fold 1-coverings which are 1-identifying.

3.3 Constructions for r-identifying codes

In Theorem 3.18 we answer, when $r_1 = r_2 = 1$, the question posed in [6] whether or not $M_{r_1+r_2}(n_1+n_2) \leq M_{r_1}(n_1)M_{r_2}(n_2)$. The direct sum of two 1-identifying codes does *not* always give a 2-identifying code, as it is considered in [40, Theorem 9] and [41, Theorem 5]. For example, the direct sum of the 1-identifying code $C = \{000,\ 001,\ 101,\ 100\}$ with itself is not 2-identifying, since for example $I_2(000010) = \{000000, 000001, 000100, 001000, 100000\} = I_2(010000)$. We show that the result for the cardinalities of 2-identifying codes can still be retrieved by showing that it is always possible to arrange initial codes such that no problem occurs. The result generalizes for all radii r in such a way that the direct sum of r 1-identifying codes can be used to construct an r-identifying code.

In Theorem 3.20 we generalize [6, Theorem 1 and Theorem 2] for all radii r. Previously, in [6] only radius 1 is considered. This theorem also shows that [40, Corollary 7] and [41, Corollary 6] cannot be proved by using the direct sum of an r-identifying code and $\mathbf{F}^k$ unless $k \geq r+1$.

As we already noticed, the direct sum of two 1-identifying codes does not always lead to a 2-identifying code. In the next lemma a 1-identifying code is arranged in such a way that it can be used in the direct sum of r 1-identifying codes to produce an r-identifying code.

Lemma 3.15. *Suppose $n \geq 3$ and $C \subseteq \mathbf{F}^n$ is a 1-identifying code such that*

$$\begin{aligned} \exists \mathbf{x} \in \mathbf{F}^n \setminus C,\ \exists \mathbf{y} \in C: &\quad I(\mathbf{x}) = \{\mathbf{y}\} \\ \forall \mathbf{u}_1,\ldots,\mathbf{u}_{n-1} \in S_1(\mathbf{x}) \setminus \{\mathbf{y}\}: &\quad \mathbf{u}_i \notin C \\ \forall \mathbf{v}_1,\ldots,\mathbf{v}_{n-1} \in S_1(\mathbf{y}) \setminus \{\mathbf{x}\}: &\quad \mathbf{v}_i \in C,\ I(\mathbf{u}_i) = \{\mathbf{v}_i\}, I(\mathbf{y}) = \{\mathbf{y},\mathbf{v}_1,\ldots,\mathbf{v}_{n-1}\}. \end{aligned}$$

Then also the code $C' = C \triangle \{\mathbf{x},\mathbf{y}\}$ is 1-identifying. Notice that $|C'| = |C|$.

Proof. Denote by C' the code we get from C by changing $\mathbf{x} \in C$ and $\mathbf{y} \notin C$. Now

$$\begin{aligned} I(C';\mathbf{x}) &= \{\mathbf{x}\} \\ I(C';\mathbf{y}) &= \{\mathbf{x},\mathbf{v}_1,\ldots,\mathbf{v}_{n-1}\} \\ I(C';\mathbf{u}_i) &= \{\mathbf{x},\mathbf{v}_i\}\ \forall\, i = 1,\ldots,n-1 \\ \{\mathbf{v}_i\} &\subseteq I(C';\mathbf{v}_i). \end{aligned}$$

Because the I-sets of the words in $B_1(\mathbf{y})$ are the only ones losing a codeword (adding is not a problem), it is enough to make sure that these I-sets are distinguishable, when proving that C' is 1-identifying.

We see that $I(\mathbf{x})$ and $I(\mathbf{y})$ are unambiguous. The I-sets of $\mathbf{v}_i's$ are unambiguous since every word $\mathbf{z} \in S_3(\mathbf{x})$ that is a neighbour of $\mathbf{v}_i$ is also a neighbour of $\mathbf{v}_k$ for some k $(i \neq k)$. □

The next definition and lemma are needed in the proof of the following theorem. Denote $N = \sum_{i=1}^{r} n_i$ and $N^* = N - n_r$.

Definition 3.16. Assume $\mathbf{c}_i$, $\mathbf{x}_i \in \mathbf{F}^{n_i}$ for $1 \leq i \leq r$. We say that a word $(\mathbf{c}_1, \ldots, \mathbf{c}_r) \in \mathbf{F}^N$ is *in balance with* a word $(\mathbf{x}_1, \ldots, \mathbf{x}_r) \in \mathbf{F}^N$ if for all $1 \leq i \leq r$ we have $d(\mathbf{c}_i, \mathbf{x}_i) \leq 2$.

Lemma 3.17. *Let $C_i \subseteq \mathbf{F}^{n_i}$ be 1-identifying codes for $i = 1, \ldots, r-1$ and $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_{r-1}) \in \mathbf{F}^{N^*}$. There is a codeword $\mathbf{c} \in C_1 \oplus \ldots \oplus C_{r-1}$ such that $d(\mathbf{x}, \mathbf{c}) = r-1$ or there are codewords $\mathbf{c}_1$ and $\mathbf{c}_2$ in $C_1 \oplus \ldots \oplus C_{r-1}$ such that $d(\mathbf{x}, \mathbf{c}_1) = r-2$ and $d(\mathbf{x}, \mathbf{c}_2) = r$. Moreover, the codewords $\mathbf{c}$, $\mathbf{c}_1$ and $\mathbf{c}_2$ are in balance with $\mathbf{x}$.*

Proof. If there are codewords $\mathbf{c}_i \in C_i$ for every $\mathbf{x}_i \in \mathbf{F}^{n_i}$ such that $d(\mathbf{x}_i, \mathbf{c}_i) = 1$, then $d(\mathbf{x}, (\mathbf{c}_1, \ldots, \mathbf{c}_{r-1})) = r-1$ and $(\mathbf{c}_1, \ldots, \mathbf{c}_{r-1})$ is in balance with $\mathbf{x}$. Suppose then there are j words $\mathbf{x}_i$ such that $I_1(\mathbf{x}_i) = \{\mathbf{x}_i\}$. Without loss of generality let $\mathbf{x}_1, \ldots, \mathbf{x}_j$ be such words. When $1 \leq i \leq j$ there is $\mathbf{c}_i^{(2)} \in C_i$ such that $d(\mathbf{x}_i, \mathbf{c}_i^{(2)}) = 2$ because C_i is 1-identifying and the words in $S_1(\mathbf{x}_i)$ must be identified. If j is even, then

$$\mathbf{c}^{(r-1)} = (\mathbf{x}_1, \ldots \mathbf{x}_{j/2}, \mathbf{c}_{j/2+1}^{(2)}, \ldots, \mathbf{c}_j^{(2)}, \mathbf{c}_{j+1}, \ldots, \mathbf{c}_{r-1})$$

is a codeword at distance $r-1$ from $\mathbf{x}$. The codeword $\mathbf{c}^{(r-1)}$ is in balance with $\mathbf{x}$. If j is odd, then

$$\mathbf{c}^{(r)} = (\mathbf{x}_1, \ldots \mathbf{x}_{(j-1)/2}, \mathbf{c}_{(j-1)/2+1}^{(2)}, \ldots, \mathbf{c}_{j-1}^{(2)}, \mathbf{c}_j^{(2)}, \mathbf{c}_{j+1}, \ldots, \mathbf{c}_{r-1})$$

is a codeword at distance r from $\mathbf{x}$ and

$$\mathbf{c}^{(r-2)} = (\mathbf{x}_1, \ldots \mathbf{x}_{(j-1)/2}, \mathbf{c}_{(j-1)/2+1}^{(2)}, \ldots, \mathbf{c}_{j-1}^{(2)}, \mathbf{x}_j, \mathbf{c}_{j+1}, \ldots, \mathbf{c}_{r-1})$$

is a codeword at distance $r-2$ from $\mathbf{x}$. The codewords $\mathbf{c}^{(r)}$ and $\mathbf{c}^{(r-2)}$ are in balance with $\mathbf{x}$. □

Theorem 3.18.

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

Proof. Suppose $C_i \subseteq \mathbf{F}^{n_i}$ for $i = 1,\ldots,r$ are 1-identifying codes and suppose also that the situation of Lemma 3.15 does not exist in any C_i (if there are multiple bad cases, we fix them one by one — it is easy to check that fixing a bad case does not create another). If $n_i = 3$ for some i, then an optimal 1-identifying code $C_i = \{000,110,101,011\}$ is used. We prove that $C = C_1 \oplus \ldots \oplus C_r$ is an r-identifying code. In fact we prove a bit stronger result that for all $\mathbf{x} = (\mathbf{x}_1,\ldots,\mathbf{x}_r)$, $\mathbf{y} = (\mathbf{y}_1,\ldots,\mathbf{y}_r) \in \mathbf{F}^N$, $\mathbf{x} \neq \mathbf{y}$, there is a codeword $\mathbf{c} = (\mathbf{c}_1,\ldots,\mathbf{c}_r)$ in $I_r(C;\mathbf{x}) \bigtriangleup I_r(C;\mathbf{y})$ such that $\mathbf{c}$ is in balance with the word in which I-set it is. We prove the claims by induction on r. The first step, $r = 1$ of induction is trivial.

An induction hypothesis is that $C^* = C_1 \oplus \ldots \oplus C_{r-1}$ is an $(r-1)$-identifying code and for every couple of words $\mathbf{x}^*$, $\mathbf{y}^* \in \mathbf{F}^{N^*}$, $\mathbf{x}^* \neq \mathbf{y}^*$, there is a codeword $\mathbf{c}^* \in I_{r-1}(C^*;\mathbf{x}^*) \bigtriangleup I_{r-1}(C^*;\mathbf{y}^*)$ such that $\mathbf{c}^*$ is in balance with a word in which I-set it is. Denote $\mathbf{x} = (\mathbf{x}^*,\mathbf{x}_r)$ and $\mathbf{y} = (\mathbf{y}^*,\mathbf{y}_r)$ where $\mathbf{x}^* = (\mathbf{x}_1,\ldots,\mathbf{x}_{r-1})$, $\mathbf{y}^* = (\mathbf{y}_1,\ldots,\mathbf{y}_{r-1}) \in \mathbf{F}^{N^*}$ and $\mathbf{x}_r,\mathbf{y}_r \in \mathbf{F}^{n_r}$.

Suppose first that $\mathbf{x}_r \neq \mathbf{y}_r$ **and** $\mathbf{x}^* \neq \mathbf{y}^*$. By the induction hypothesis, we know that there is a codeword $\mathbf{c}^* = (\mathbf{c}_1,\ldots,\mathbf{c}_{r-1}) \in I_{r-1}(C^*;\mathbf{x}^*) \bigtriangleup I_{r-1}(C^*;\mathbf{y}^*)$. Without loss of generality we may assume that $\mathbf{c}^* \in I_{r-1}(C^*;\mathbf{x}^*)$ and by assumption it is in balance with $\mathbf{x}^*$. By adding to the end of $\mathbf{c}^*$ a codeword $\mathbf{c}_r \in C_r$ such that $d(\mathbf{x}_r,\mathbf{c}_r) \leq 2$ we get a codeword that is in balance with $\mathbf{x}$. If at the same time any component of $\mathbf{c}^*$ is changed in such a way that it still is at distance at most two from the corresponding component of $\mathbf{x}$, then again the result codeword is in balance with $\mathbf{x}$.

Since C_r is 1-identifying, we know that there is a codeword $\mathbf{c}_r \in C_r$ such that $\mathbf{c}_r \in I_1(C_r;\mathbf{x}_r) \bigtriangleup I_1(C_r;\mathbf{y}_r)$. If $\mathbf{c}_r \in I_1(C_r;\mathbf{x}_r)$, then $(\mathbf{c}^*,\mathbf{c}_r) \in I_r(C;\mathbf{x}) \setminus I_r(C;\mathbf{y})$. Assume then $\mathbf{c}_r \in I_1(C_r;\mathbf{y}_r)$. There is $\mathbf{v}_r \in I_1(C_r;\mathbf{x}_r)$. If $d(\mathbf{y}_r,\mathbf{v}_r) \geq 1$, then $(\mathbf{c}^*,\mathbf{v}_r) \in I_r(C;\mathbf{x}) \setminus I_r(C;\mathbf{y})$. Suppose $d(\mathbf{v}_r,\mathbf{y}_r) = 0$. Suppose also there is $\mathbf{c}_\mathbf{x} \in C_r$ such that $d(\mathbf{c}_\mathbf{x},\mathbf{x}_r) = 2$ and $d(\mathbf{c}_\mathbf{x},\mathbf{y}_r) = 3$. If such a codeword does not exist, then we are in the situation of Lemma 3.15, and we assumed that it cannot happen.

If $\mathbf{c}^* \in I_{r-2}(C^*;\mathbf{x}^*)$, then $(\mathbf{c}^*,\mathbf{c}_r) \in I_r(C;\mathbf{x}) \setminus I_r(C;\mathbf{y})$. On the other hand, a word $(\mathbf{c}^*,\mathbf{y}_r) \in I_r(C;\mathbf{x})$, and we have $(\mathbf{c}^*,\mathbf{y}_r) \in I_r(C;\mathbf{y})$ if and only if $d(\mathbf{c}^*,\mathbf{y}^*) = r$. Hence, assume now $d(\mathbf{x}^*,\mathbf{c}^*) = r-1$ and $d(\mathbf{y}^*,\mathbf{c}^*) = r$.

If for some j there is a codeword $\mathbf{c}'_j$ such that $d(\mathbf{c}'_j,\mathbf{x}_j) < d(\mathbf{c}_j,\mathbf{x}_j)$, then a codeword $(\mathbf{c}_1,\ldots,\mathbf{c}_{j-1},\mathbf{c}'_j,\mathbf{c}_{j+1},\ldots,\mathbf{c}_{r-1},\mathbf{c}_\mathbf{x}) \in I_r(C;\mathbf{x}) \setminus I_r(C;\mathbf{y})$. Namely, the distance between $\mathbf{y}_j$ and $\mathbf{c}'_j$ decreases by at most two compared to the distance between $\mathbf{y}_j$ and $\mathbf{c}_j$. Suppose then that for any $\mathbf{x}_j$ such a codeword $\mathbf{c}'_j$ does not exist, then $d(\mathbf{x}_j,\mathbf{c}_j) = 1$ and $\mathbf{x}_j \notin C_j$ for all $1 \leq j \leq r-1$. If $\mathbf{c}^*$ is in balance with $\mathbf{y}^*$, then there is k such that $d(\mathbf{y}_k,\mathbf{c}_k) = 2$ (because $d(\mathbf{y}^*,\mathbf{c}^*) = r$). Now there is $\mathbf{c}'_k \in I_1(C_k;\mathbf{y}_k)$ and

$$\mathbf{c} = (\mathbf{c}_1,\ldots,\mathbf{c}_{k-1},\mathbf{c}'_k,\mathbf{c}_{k+1},\ldots,\mathbf{c}_{r-1},\mathbf{c}_r) \in I_r(C;\mathbf{y}) \setminus I_r(C;\mathbf{x}),$$

because $\mathbf{x}_j \notin C_j\ \forall j$. The codeword $\mathbf{c}$ is in balance with $\mathbf{y}$. If $\mathbf{c}^*$ is not in balance with $\mathbf{y}^*$, denote by $j_1,\ldots,j_k$ $(k \geq 1)$ indices for which $d(\mathbf{y}_{j_h},\mathbf{c}_{j_h}) \geq 3$ $(1 \leq h \leq k)$. For all these components there are $\mathbf{c}'_{j_h} \in I_1(C_{j_h};\mathbf{y}_{j_h})$. Because $\mathbf{x}_{j_h} \notin C_{j_h}$ we have

$d(\mathbf{x}_{j_h}, \mathbf{c}'_{j_h}) \geq 1$. Thus,

$$\mathbf{c} = (\mathbf{c}_1, \ldots, \mathbf{c}'_{j_1}, \ldots, \mathbf{c}'_{j_k}, \ldots, \mathbf{c}_{r-1}, \mathbf{c}_r) \in I_r(C;\mathbf{y}) \setminus I_r(C;\mathbf{x}).$$

The codeword $\mathbf{c}$ is in balance with $\mathbf{y}$.

Suppose next that $\mathbf{x}_r = \mathbf{y}_r$ **and** $\mathbf{x}^* \neq \mathbf{y}^*$. If $\mathbf{x}^*$ and $\mathbf{y}^*$ differ at least in two components $\mathbf{x}_i \neq \mathbf{y}_i$ and $\mathbf{x}_j \neq \mathbf{y}_j$ $(i \neq j,\ 1 \leq i, j \leq r-1)$, then we return to the previous case. Namely, we can operate now in an equivalent code

$$C_1 \oplus \ldots \oplus C_{i-1} \oplus C_{i+1} \oplus \ldots \oplus C_r \oplus C_i.$$

By the previous case we find a codeword $\mathbf{c} = (\mathbf{c}_1, \ldots, \mathbf{c}_{i-1}, \mathbf{c}_{i+1}, \ldots, \mathbf{c}_r, \mathbf{c}_i)$ which is in the symmetric difference of the I-sets of $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_{i-1}, \mathbf{x}_{i+1}, \ldots, \mathbf{x}_r, \mathbf{x}_i)$ and $\mathbf{y} = (\mathbf{y}_1, \ldots, \mathbf{y}_{i-1}, \mathbf{y}_{i+1}, \ldots, \mathbf{y}_r, \mathbf{y}_i)$, and it is in balance. By changing the last component back to the ith position we get a solution to the current case.

Thus, we still need to prove that the claims hold when $\mathbf{x}_r \neq \mathbf{y}_r$ **and** $\mathbf{x}^* = \mathbf{y}^*$. There is a codeword $\mathbf{c}_r \in I_1(C_r;\mathbf{x}_r) \bigtriangleup I_1(C_r;\mathbf{y}_r)$. We may assume without loss of generality that $\mathbf{c}_r \in I_1(C_r;\mathbf{x}_r)$. If there is a codeword $\mathbf{c}^* \in C^*$ such that $d(\mathbf{c}^*, \mathbf{x}^*) = r-1$ and $\mathbf{c}^*$ is in balance with $\mathbf{x}^*$, then $(\mathbf{c}^*, \mathbf{c}_r) \in I_r(C;\mathbf{x}) \setminus I_r(C;\mathbf{y})$.

If there does not exist a codeword at distance $r-1$ from $\mathbf{x}^*$, which is in balance with $\mathbf{x}^*$, then by Lemma 3.17 we know that there are codewords $\mathbf{c}^{(r-2)}$ and $\mathbf{c}^{(r)}$ at distances $r-2$ and r from $\mathbf{x}^*$, respectively. These codewords are in balance with $\mathbf{x}^*$.

If $\mathbf{x}_r \in C_r$, then $(\mathbf{c}^{(r)}, \mathbf{x}_r) \in I_r(C_r;\mathbf{x}) \setminus I_r(C_r;\mathbf{y})$. Similar argumentation holds if $\mathbf{y}_r \in C_r$. Suppose now $\mathbf{x}_r, \mathbf{y}_r \notin C_r$. Now $(\mathbf{c}^{(r-2)}, \mathbf{c}_r) \in I_r(C_r;\mathbf{x}) \cap I_r(C_r;\mathbf{y})$ if and only if $d(\mathbf{y}_r, \mathbf{c}_r) = 2$. Thus, suppose $d(\mathbf{y}_r, \mathbf{c}_r) = 2$, this implies $d(\mathbf{x}_r, \mathbf{y}_r) = 1$ or 3.

a) Suppose $d(\mathbf{x}_r, \mathbf{y}_r) = 3$. Assume first that the code length $n_r \geq 4$, then there is $\mathbf{z}_r \in S_1(\mathbf{x}_r) \setminus S_2(\mathbf{y}_r)$. Because $I_1(\mathbf{z}_r)$ is nonempty there is a codeword $\mathbf{c}'_r \in C_r$ such that $d(\mathbf{c}'_r, \mathbf{z}_r) \leq 1$ and $d(\mathbf{c}'_r, \mathbf{y}_r) \geq 3$. Thus, $(\mathbf{c}^{(r-2)}, \mathbf{c}'_r) \in I_r(C;\mathbf{x}) \setminus I_r(C;\mathbf{y})$. If $n_r = 3$, we use the code as assumed to be used in the beginning, and in this code there are no two non-codewords at distance three from each other. If $n_r = 2$, then there are no two non-codewords and this is not a problem.

b) Suppose $d(\mathbf{x}_r, \mathbf{y}_r) = 1$. If there is $\mathbf{c}'_r \in (C_r \cap S_2(\mathbf{x})) \setminus S_1(\mathbf{y})$, then $(\mathbf{c}^{(r-2)}, \mathbf{c}'_r) \in I_r(C;\mathbf{x}) \setminus I_r(C;\mathbf{y})$. Thus, suppose that if $\mathbf{c}'_r \in S_2(\mathbf{x}) \cap C_r$, then $d(\mathbf{c}'_r, \mathbf{y}_r) = 1$. Denote by $\mathbf{z}_r$ a unique word in $(S_1(\mathbf{c}_r) \cap S_1(\mathbf{y}_r)) \setminus \{\mathbf{x}_r\}$. Now $I_1(C_r;\mathbf{z}_r) = I_1(C_r;\mathbf{c}_r)$ unless there is a codeword $\mathbf{c}_\mathbf{z}$ $(\mathbf{c}_\mathbf{z} \neq \mathbf{c}_r)$ such that $d(\mathbf{c}_\mathbf{z}, \mathbf{z}_r) = 1$ and so $d(\mathbf{c}_\mathbf{z}, \mathbf{y}_r) = 2$ and $d(\mathbf{c}_\mathbf{z}, \mathbf{x}_r) = 3$. We get $(\mathbf{c}^{(r-2)}, \mathbf{c}_\mathbf{z}) \in I_r(C;\mathbf{y}) \setminus I_r(C;\mathbf{x})$.

So we can always distinguish between the I-sets of $\mathbf{x}$ and $\mathbf{y}$ for all $\mathbf{x}, \mathbf{y} \in \mathbf{F}^N$, $\mathbf{x} \neq \mathbf{y}$. Moreover, there is a codeword in $I_r(C;\mathbf{x}) \bigtriangleup I_r(C;\mathbf{y})$ which is in balance with the word in which I-set it is. □

Corollary 3.19.

$$M_2(n+m) \leq M_1(n)M_1(m).$$

The next theorem generalizes [6, Theorem 1 and Theorem 2] for all radii r. Previously, in [6] only radius 1 is considered.

Theorem 3.20. *Let $C \subseteq \mathbf{F}^n$ be an r-identifying code and $k \geq 1$. A direct sum $D = C \oplus \mathbf{F}^k$ is r-identifying if and only if*

$$\forall\, \mathbf{x} \in \mathbf{F}^n,\ \exists\, \mathbf{c} \in C: \quad r-k+1 \leq d(\mathbf{x},\mathbf{c}) \leq r. \tag{3.4}$$

Moreover, the direct sum preserves the property (3.4) *in $\mathbf{F}^{n+k}$.*

Proof. Suppose first that there is $\mathbf{x} \in \mathbf{F}^n$ such that $I_r(C;\mathbf{x}) = I_{r-k}(C;\mathbf{x})$. For $\mathbf{a},\mathbf{b} \in \mathbf{F}^k$, $\mathbf{a} \neq \mathbf{b}$, we have $I_r(D;(\mathbf{x},\mathbf{a})) = \{(\mathbf{c},\mathbf{y}) \mid \mathbf{c} \in I_{r-k}(C;\mathbf{x}),\ \mathbf{y} \in \mathbf{F}^k\} = I_r(D;(\mathbf{x},\mathbf{b}))$. This is not possible.

Suppose the condition (3.4) holds. Let $\mathbf{x},\mathbf{x}' \in \mathbf{F}^n$ and $\mathbf{a},\mathbf{b} \in \mathbf{F}^k$. Because C is an r-identifying code it is clear that $I_r(D;(\mathbf{x},\mathbf{a})) \neq I_r(D;(\mathbf{x}',\mathbf{b}))$ when $\mathbf{x} \neq \mathbf{x}'$. Suppose $\mathbf{x} = \mathbf{x}'$ and $\mathbf{a} \neq \mathbf{b}$. By assumption there is $\mathbf{c} \in C$ such that $r-k+1 \leq d := d(\mathbf{x},\mathbf{c}) \leq r$. Now $(\mathbf{c},\mathbf{y}) \in I_r(D;(\mathbf{x},\mathbf{a})) \setminus I_r(D;(\mathbf{x},\mathbf{b}))$ where $\mathbf{y}$ is chosen in such a way that $d(\mathbf{y},\mathbf{a}) \leq r-d$ and $d(\mathbf{y},\mathbf{b}) \geq r-d+1$. This choice is always possible since $d(\mathbf{a},\mathbf{b}) \geq 1$. □

In the previous theorem if $k \geq r+1$ the condition (3.4) is always true so we have the following corollary:

Corollary 3.21. *If C is an r-identifying code and $k \geq 0$, then $C \oplus \mathbf{F}^{r+1+k}$ is r-identifying. Thus, $M_r(n+r+1+k) \leq 2^{r+1+k}M_r(n)$.*

In Table 3.1 we have collected the upper and lower bounds on the smallest cardinalities of 1- and 2-identifying codes.

Table 3.1: Bounds on the cardinalities of 1- and 2-identifying codes.

n	$M_1(n)$	$M_2(n)$
2	3 d	–
3	4 a	a 7 e
4	7 b	a 6 e
5	10 a	b 6 e
6	18 – 19 b	b 8 e
7	32 b	c 14 f
8	56 – 62 c	b 17 – 21 f
9	101 – 115 c	b,d 26 – 36 f
10	183– 236 c	d 41 – 63 f
11	337 – 352 c	d 67 – 148 g
12	623 – 696 c	d 112 – 280 g
13	1158 – 1344 c	d 190 – 504 h
14	2164 – 2784 f	d 326 – 1008 h
15	4063 – 5120 g	d 567 – 1984 i (7,8)
16	7654 – 10240 h	d 995 – 3520 i (5,11)
17	14469 – 20480 g,h	d 1761 – 6688 i (6,11)
18	27434 – 40960 h	d 3141 – 11264 i (7,11)
19	52155 – 65536 g	d 5638 – 21824 i (8,11)
20	99392 – 131072 h	d 10179 – 40480 i (9,11)

Key to the table of $M_1(n)$:

The lower bounds for lengths 4 and 7 are from [6]. The other lower bounds come from [40, Theorem 3] see also [6] and Corollary 3.5.

a Karpovsky et al. [40]

b Blass et al. [6]

c Codes in Appendix

d $M_r(r+1) = 2^{r+1} - 1$ Blass et al. [5, Theorem 5]

f $M_1(n) \leq 4M_1(n-2)$ Blass et al. [6, Theorem 2] (see also Corollary 3.21)

g Corollary 3.14

h $M_1(n) \leq 2M_1(n-1) \Leftrightarrow \forall\, \mathbf{c} \in C\; d(\mathbf{c}, C \setminus \{\mathbf{c}\}) = 1$ Blass et al. [6, Theorem 1] (see also Corollary 3.21)

Key to the table of $M_2(n)$:

a Blass et al. [5]

b Karpovsky et al. [40, Theorem 1(iii)]

c Computer search [17]

d Corollary 3.6

e Blass et al. [5, Theorem 5, Theorem 6], see also the note therein on [17].

f See Appendix.

g Theorem 3.20, see also Appendix.

h Corollary 3.21

i Corollary 3.19, in parentheses there are n and m that are used.

Chapter 4

On codes identifying sets of vertices

In this chapter we consider the situation where the size of the fault pattern can be more than one. In the first section we present lower bounds for $(r, \leq \ell)$-identifying codes. The lower bound for $(1, \leq 2)$-identifying codes turns out to be strict for infinitely many lengths. We also show which size of the fault pattern the code can handle with respect to a code length n and radius r.

In the second section we construct optimal $(1, \leq 2)$-identifying codes for short lengths $4, 5$ and 7.

In the third section we present constructions for $(r, \leq \ell)$-identifying codes. First we present two constructions for $(1, \leq 2)$-identifying codes. The first one of these, $(\pi(\mathbf{u}), \mathbf{u}, \mathbf{u}+\mathbf{v})$-construction, leads to two infinite families of optimal $(1, \leq 2)$-identifying codes. The second construction gives an asymptotically optimal result. At the end of the third section we prove that the direct sum of $\mathbf{F}^r$ and an $(r, \leq \ell)$-identifying code for $1 \leq r \leq 2$ and $\ell \geq 2$ is an $(r, \leq \ell)$-identifying code. In Theorems 4.26 and 4.28 we prove that the direct sum of r $(1, \leq \ell)$-identifying codes is an $(r, \leq \ell)$-identifying code for all $\ell \geq 2$.

Section 4.1 is based on [34] except Theorem 4.7 and Theorem 4.10 are from [49]. The results in Section 4.2 are from [34] except the result concerning $n = 7$ is from [56]. Section 4.3.1 is based on [34] unless otherwise stated. The results of Section 4.3.2 except Theorem 4.29 are from [49]. Theorem 4.19 and Theorem 4.22 for $r = 2$ are only published in this thesis.

4.1 Lower bounds

Suppose C is an $(r, \leq \ell)$-identifying code. There are $2^{|C|}$ subsets and $\sum_{i=0}^{\ell} \binom{2^n}{i}$ fault patterns to be identified. Thus, we have proved the next lower bound from [40].

Theorem 4.1. *If C is an $(r, \leq \ell)$-identifying code, then*

$$|C| \geq \left\lceil \log_2 \sum_{i=0}^{\ell} \binom{2^n}{i} \right\rceil .$$

First we prove some general properties of certain identifying codes.

Lemma 4.2. *Let $C \subseteq \mathbf{F}^n$ be an $(r, \leq \ell)$-identifying code where $\ell \geq 2$. For any two different fault patterns $X = \{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$ and $Y = \{\mathbf{y}_1, \ldots, \mathbf{y}_v\}$ such that $X \not\subseteq Y$, $Y \not\subseteq X$ and $|X \cup Y| \leq \ell$, we have $I_r(X) \not\subseteq I_r(Y)$ and $I_r(Y) \not\subseteq I_r(X)$. In particular, if $I_r(\mathbf{x}) \subseteq I_r(\mathbf{y})$, for some $\mathbf{x}, \mathbf{y} \in \mathbf{F}^n$, then $\mathbf{x} = \mathbf{y}$.*

Proof. If for some X and Y, $I_r(X) \subseteq I_r(Y)$, then $I_r(X \cup Y) = I_r(Y)$, which leads to a contradiction, since $|X \cup Y| \leq \ell$. □

The next theorem generalizes [45, Theorem 1] for all radii $r \geq 1$.

Theorem 4.3. *Assume $r \geq 1$ and $\ell \geq 1$. If $C \subseteq \mathbf{F}^n$ is an $(r, \leq \ell)$-identifying code, then*

$$\forall \mathbf{x} \in \mathbf{F}^n : \quad |I_r(\mathbf{x})| = |B_r(\mathbf{x}) \cap C| \geq 2\ell - 1.$$

Proof. If $\ell = 1$, the claim holds trivially. Suppose $\ell \geq 2$ and assume to the contrary that there is $\mathbf{x} \in \mathbf{F}^n$ such that $|I_r(\mathbf{x})| \leq 2\ell - 2$. We denote $I_r(\mathbf{x}) = \{\mathbf{c}_1, \ldots, \mathbf{c}_k\}$, where $k \leq 2\ell - 2$. We know that $1 \leq d(\mathbf{c}_i, \mathbf{c}_j) \leq 2r$ for every $\mathbf{c}_i, \mathbf{c}_j$, where $i \neq j$ and $1 \leq i, j \leq 2\ell - 2$. Thus, for any couple of codewords $\{\mathbf{c}_i, \mathbf{c}_j\}$ there is a word $\mathbf{y} \in \mathbf{F}^n$, $\mathbf{y} \neq \mathbf{x}$, such that $\{\mathbf{c}_i, \mathbf{c}_j\} \subseteq I_r(\mathbf{y})$. Hence, we find a collection of words Y such that $|Y| \leq \ell - 1$ and $I_r(\mathbf{x}) \subseteq I_r(Y)$, which is a contradiction by Lemma 4.2. □

The next lemma is from [6].

Lemma 4.4. *Let $C \subseteq \mathbf{F}^n$ be a 3-fold 1-covering. Then for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbf{F}^n$, $\mathbf{x} \neq \mathbf{y}$ we have $I(\mathbf{x}) \neq I(\mathbf{y})$ and $I(\mathbf{x}) \neq I(\mathbf{y}, \mathbf{z})$.*

Proof. The claims follow from Lemma 2.1. □

Lemma 4.5. *Let $C \subseteq \mathbf{F}^n$ be a $(1, \leq 2)$-identifying code. There does not exist a* square *of codewords such that $\mathbf{x}, \mathbf{y} \in C$, $d(\mathbf{x}, \mathbf{y}) = 2$, $|I(\mathbf{x})| = |I(\mathbf{y})| = 3$ and $|I(\mathbf{x}) \cap I(\mathbf{y})| = 2$.*

Proof. Suppose to the contrary that for $\mathbf{x}, \mathbf{y} \in C$ we have $d(\mathbf{x}, \mathbf{y}) = 2$, $|I(\mathbf{x})| = |I(\mathbf{y})| = 3$ and $|I(\mathbf{x}) \cap I(\mathbf{y})| = 2$. For $\mathbf{c} \in I(\mathbf{x}) \cap I(\mathbf{y})$ we have $I(\mathbf{x}, \mathbf{c}) = I(\mathbf{y}, \mathbf{c})$, which is a contradiction. □

The next theorem improves on Theorem 4.1. When $r = 1$ and $\ell = 2$ the derived lower bound is shown to be optimal for infinitely many lengths.

Theorem 4.6. *Assume $r \geq 1$ and $\ell \geq 2$. We have*

$$M_r^{(\leq \ell)}(n) \geq K(n, 2\ell - 1, r) \geq (2\ell - 1)\frac{2^n}{V(n,r)}.$$

In particular,

$$M_1^{(\leq 2)}(n) \geq \left\lceil 3\frac{2^n}{n+1} \right\rceil.$$

Moreover, if $n \geq 4$ is even, then

$$M_1^{(\leq 2)}(n) \geq \left\lceil \frac{(3n-8)2^n}{n^2 - 2n - 2} \right\rceil.$$

Proof. Let C be an $(r, \leq \ell)$-identifying code. By Theorem 4.3 we know that all the words of the ambient space are r-covered at least by $(2\ell - 1)$ codewords of C. Hence $M_r^{(\leq \ell)}(n) \geq K(n, 2\ell - 1, r)$. Obviously, $(2\ell - 1) \cdot 2^n / V(n,r)$ is a lower bound on $K(n, 2\ell - 1, r)$. The last assertion follows from [23, Corollary 2]:

$$K(n,3,1) \geq \frac{(3(n-3)+1)2^n}{(n-3)(n+1)+1}$$

for even n where $n \geq 4$. □

The interested reader is referred to [45] for further results on $(1, \leq \ell)$-identifying codes when $\ell \geq 3$.

For $r \geq 2$ and $\ell \geq 2$ we can improve on the previous bound.

Theorem 4.7. *For $r \geq 2$ and $\ell \geq 2$ we have*

$$M_r^{(\leq \ell)}(n) \geq \left\lceil \frac{(2\ell - 1)2^n}{\binom{n}{r} + \binom{n}{r-1}} \right\rceil.$$

Proof. Suppose $C \subseteq \mathbf{F}^n$ is an $(r, \leq \ell)$-identifying code. We shall show that for every $\mathbf{x} \in \mathbf{F}^n$ we have $|I_r(\mathbf{x}) \cap (S_r(\mathbf{x}) \cup S_{r-1}(\mathbf{x}))| \geq 2\ell - 1$. Without loss of generality we can prove that the claim holds for $\mathbf{x} = \mathbf{0}$ and clearly, it then holds for every $\mathbf{x} \in \mathbf{F}^n$. For any two words $\mathbf{c}_1, \mathbf{c}_2$ of weight $1 \leq w(\mathbf{c}_1), w(\mathbf{c}_2) \leq r$ there is a word $\mathbf{y}_1$ of weight two which r-covers both of these words. Clearly, $I_{r-2}(\mathbf{0}) \subseteq I_r(\mathbf{y}_1)$. If $|I_r(\mathbf{0}) \cap (S_r(\mathbf{0}) \cup S_{r-1}(\mathbf{0}))| \leq 2\ell - 2$, then there is a collection of words of weight two, $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{\ell-1}$, which r-cover the whole set $I_r(\mathbf{0})$. This is not possible by Lemma 4.2. The claim follows from a direct calculation

$$|C| \left(\binom{n}{r} + \binom{n}{r-1} \right) \geq (2\ell - 1)2^n.$$

□

Next we provide some nonexistence results.

Theorem 4.8. *Let $r(n,K)$ denote the smallest covering radius among all binary codes of length n and cardinality K. Let $\ell < 2^n$. If $r \geq r(n,\ell)$, then there does not exist an $(r, \leq \ell)$-identifying code of length n.*

Proof. Let $C(\ell)$ be a code realizing the value $r(n,\ell)$. Thus, the Hamming spheres of radius at least $r(n,\ell)$ centered at the codewords cover the entire space. Take $\mathbf{x} \in \mathbf{F}^n$ such that $C(\ell) \neq \mathbf{x} + C(\ell)$ (this is always possible when $C(\ell) \neq \mathbf{F}^n$). Consequently, the sets $C(\ell)$ and $\mathbf{x} + C(\ell)$ of cardinality ℓ cannot be identified by any code $C \subseteq \mathbf{F}^n$. □

Values of $r(n,K)$ can be found for instance from [13, p. 192–200].

Corollary 4.9. *There is no $(1, \leq 2)$-identifying code of length less than four and no $(2, \leq 2)$-identifying code of length less than six.*

If we fix a radius r and a length n, then the following theorem tells what values the cardinality of a fault pattern ℓ *cannot* get.

Theorem 4.10. *Suppose $C \subseteq \mathbf{F}^n$ is an $(r, \leq \ell)$-identifying code. Then*

$$\ell \leq \left\lfloor \frac{n}{2} \right\rfloor - r + 2.$$

Proof. The first step of the proof is to cover the whole set $B_r(\mathbf{0})$ by a small number of words. We take one word of weight $2r$, namely,

$$\mathbf{x} = \overbrace{1\ldots1}^{2r}0\ldots0.$$

Suppose first $2 \nmid n$. We take

$$\begin{aligned} \mathbf{y}_1 &= \overbrace{0\ldots0}^{2r}11000\ldots0 \\ \mathbf{y}_2 &= 0\ldots000110\ldots0 \\ &\vdots \\ \mathbf{y}_{(n-2r-1)/2} &= 0\ldots000\ldots0110. \end{aligned}$$

We also take a word $\mathbf{z} = 0\ldots01$. Now

$$I_r(\mathbf{0}, \mathbf{x}, \mathbf{y}_1, \ldots, \mathbf{y}_{(n-2r-1)/2}, \mathbf{z}) = I_r(\mathbf{x}, \mathbf{y}_1, \ldots, \mathbf{y}_{(n-2r-1)/2}, \mathbf{z})$$

which is impossible if $\ell \geq (n - 2r - 1)/2 + 3$.

Assume then $2 \mid n$. We take

$$\begin{aligned}\mathbf{y}_1 =& \overbrace{0\ldots0}^{2r}11000\ldots0\\ \mathbf{y}_2 =& 0\ldots000110\ldots0\\ &\vdots\\ \mathbf{y}_{(n-2r)/2} =& 0\ldots0000\ldots011.\end{aligned}$$

These words cover all the words of weight r. To cover $B_{r-1}(\mathbf{0})$ we take any word $\mathbf{z}$ of weight one. We have

$$I_r(\mathbf{0},\mathbf{x},\mathbf{y}_1,\ldots,\mathbf{y}_{(n-2r)/2},\mathbf{z}) = I_r(\mathbf{x},\mathbf{y}_1,\ldots,\mathbf{y}_{(n-2r)/2},\mathbf{z}),$$

which is impossible if $\ell \geq (n-2r)/2+3$. □

In [26] it is proved that $(1,\leq 2)$-identifying codes are also 1-edge-robust 1-identifying codes.

If $\mu \geq 3$, a perfect μ-fold 1-covering of length n can find one faulty processor [6] and *detect* a larger number of malfunctioning processors, because then $|I(X)| \geq 2\mu - 2 > \mu = |I(\mathbf{z})|$ for any $\mathbf{z} \in \mathbf{F}^n$ and $X \subseteq \mathbf{F}^n$ where $|X| \geq 2$. This property is clearly not shared by all codes, for example, the 1-identifying code $C = \{\mathbf{x} \in \mathbf{F}^4 | w(\mathbf{x}) \in \{0,2,3\}\}$ gives $I(1000) = I(0000,1000)$. From this point of view the research is done in [30].

4.2 Short codes

In the next theorems we shall obtain the optimal $(1,\leq 2)$-identifying codes of lengths four, five and seven.

Theorem 4.11. $M_1^{(\leq 2)}(4) = 11$.

Proof. The code

$$C = \{0000,0010,0001,1100,1010,1001,0101,0011,1110,1101,0111\}$$

gives the upper bound $M_1^{(\leq 2)}(4) \leq 11$. Every word of $\mathbf{F}^4$ is 3-fold 1-covered by C, thus by Lemma 4.4 single words are distinguishable from each other and from couples of words.

One way of verifying that $I(\mathbf{x},\mathbf{y}) \neq I(\mathbf{x},\mathbf{w})$ for all $\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{w} \in \mathbf{F}^4$, $\{\mathbf{x},\mathbf{y}\} \neq \{\mathbf{z},\mathbf{w}\}$ is by looking at the following matrix A where the first row corresponds to the codeword 0000 and the second to 0010 etc., and the columns correspond to all words $\mathbf{x} = (x_1,x_2,x_3,x_4)$ of the space $\mathbf{F}^4$ in the order of integers J $(0 \leq J \leq 15)$ when they are expressed as $J = \sum_{i=1}^{4} x_i 2^{(i-1)}$. The entry of the matrix is one if the codeword corresponding to that row 1-covers the word of the corresponding column, otherwise the entry is zero. Thus,

$$A=\begin{pmatrix}
1&1&1&0&1&0&0&0&1&0&0&0&0&0&0&0\\
1&0&0&0&1&1&1&0&0&0&0&0&1&0&0&0\\
1&0&0&0&0&0&0&0&1&1&1&0&1&0&0&0\\
0&1&1&1&0&0&0&1&0&0&0&1&0&0&0&0\\
0&1&0&0&1&1&0&1&0&0&0&0&0&1&0&0\\
0&1&0&0&0&0&0&0&1&1&0&1&0&1&0&0\\
0&0&1&0&0&0&0&0&1&0&1&1&0&0&1&0\\
0&0&0&0&1&0&0&0&1&0&0&0&1&1&1&0\\
0&0&0&1&0&1&1&1&0&0&0&0&0&0&0&1\\
0&0&0&1&0&0&0&0&0&1&1&1&0&0&0&1\\
0&0&0&0&0&0&1&0&0&0&1&0&1&0&1&1
\end{pmatrix}.$$

By comparing the logical OR of all pairs of columns we get the claim.

By Theorem 4.6, we get $M_1^{(\leq 2)}(4) \geq 11$. □

In the following proof $\overline{\mathbf{x}} := \mathbf{1} + \mathbf{x}$.

Theorem 4.12. *i) If C is a perfect 3-fold 1-covering of length five, then for every $\mathbf{x} \in \mathbf{F}^5$, either $\mathbf{x} \in C$ or $\overline{\mathbf{x}} \in C$ but not both.*

ii) There are exactly two nonequivalent perfect 3-fold 1-coverings of length five.

iii) One of these perfect 3-fold 1-coverings is $(1, \leq 2)$-identifying and the other one is not.

Proof. Without loss of generality, assume that $\mathbf{0} \notin C$, 00100, 00010, 00001 $\in C$, 01000, 10000 $\notin C$. Because all the vectors of weight one are covered by exactly three codewords, the number of codewords of weight two must be six.

To prove i), it suffices to notice that if $\mathbf{1} \notin C$, then the same argument shows that there are three codewords of weight four and six of weight three, and altogether the number of codewords would be more than sixteen, which is the cardinality of the perfect 3-fold 1-covering of length five.

Having proved i), it suffices to consider which vectors of weight two are in C. Because 10000 is covered by three codewords, three of the six codewords of weight two have 1 in the first coordinate, and the same is true for the second coordinate.

If 11000 $\notin C$, there is a unique choice for C:

$$\begin{array}{l} 00100 \\ 00010 \\ 00001 \\ 10100 \\ 10010 \\ 10001 \\ 01100 \\ 01010 \\ 01001 \end{array}.$$

The seven other codewords are the complements of the non-codewords of weight at most two. A routine verification shows that it is indeed a perfect 3-fold 1-covering. We notice that $C_1 = \{00100, 10100, 01100, 11100\} \subseteq C$. Clearly, C is not $(1, \leq 2)$-identifying, because $I(00100, 10100) = C_1 = I(01100, 11100)$.

The other possibility is that $11000 \in C$. Apart from the order of the coordinates, there is again a unique way of choosing the codewords of weight two in such a way that every vector of weight one is covered by exactly three codewords:

$$\begin{array}{l} 00100 \\ 00010 \\ 00001 \\ 11000 \\ 10100 \\ 10010 \\ 01100 \\ 01001 \\ 00011 \end{array}.$$

A routine verification shows that the resulting code $C = \{00100, 00010,$ $00001,$ $11000,$ $10100,$ $10010, 01100, 01001, 00011, 11010, 11001, 10101, 01110, 10111,$ $01111,$ $11111\}$ is a perfect 3-fold 1-covering and $(1, \leq 2)$-identifying, and hence, in particular, inequivalent to the code constructed earlier. □

Corollary 4.13. *There is a unique* $(1, \leq 2)$*-identifying code (up to equivalence) attaining the bound* $M_1^{(\leq 2)}(5) = 16$.

Denote by $\mathscr{H}_3$ the Hamming code of length seven with the parity check matrix

$$H = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}.$$

Let $C_1 = \mathscr{H}_3 + 1011001$ and $C_2 = \mathscr{H}_3 + 0000100$ be two cosets of $\mathscr{H}_3$. Let further P_1 and P_3 be the codes obtained by permuting C_1 using $(7,3)(4,2)$ (the notation (i,j) means that ith and jth columns are interchanged) and $(6,3)(4,1)$, respectively. By P_2 we denote the code obtained from C_2 using the permutation $(1,2)(3,5)$. It is easy to check (with computer) that $U = P_1 \cup P_2 \cup P_3$ is $(1,\leq 2)$-identifying. The following result now follows from Theorem 4.6.

Theorem 4.14. *The code U is a* $(1,\leq 2)$*-identifying code of length seven and hence*

$$M_1^{(\leq 2)}(7) = 48.$$

4.3 Constructions

In this section, we present different ways to construct $(1,\leq 2)$-identifying codes and $(r,\leq \ell)$-identifying codes for $r \geq 2$ and $\ell \geq 2$. In the first construction, we use shorter $(1,\leq 2)$-identifying codes to create longer ones. In the second construction, we get asymptotically good (under Conjecture 2.4) $(1,\leq 2)$-identifying codes using 3-covering codes. In subsection 4.3.2 we prove direct sum constructions for $(r,\leq \ell)$-identifying codes for $r \geq 1$ and $\ell \geq 2$.

4.3.1 Constructions for $(1,\leq 2)$-identifying codes

The construction of the next theorem yields two infinite families of optimal $(1,\leq 2)$-identifying codes.

Theorem 4.15. *If* $C \subseteq \mathbf{F}^n$ *is a* $(1,\leq 2)$*-identifying code, then*

$$C' = \{(\pi(\mathbf{u}),\mathbf{u},\mathbf{u}+\mathbf{v}) \mid \mathbf{u} \in \mathbf{F}^n, \mathbf{v} \in C\},$$

where $\pi(\mathbf{u})$ *denotes a parity check bit on* $\mathbf{u}$*, is a* $(1,\leq 2)$*-identifying code of length* $2n+1$.

Proof. By Theorem 4.3 the code C covers each word at least three times, and by [13, Theorems 3.4.3 and 14.4.3] so does the code C'. By Lemma 4.4 the code C' is 1-identifying and all single words and pairs of words are distinguishable. Thus, we only need to check that all pairs are identified from one another.

Let us divide the words of $\mathbf{F}^{2n+1}$ into two classes by their first bit and consider the codewords which cover a word in each class. Let $\mathbf{x} = (a,\mathbf{u},\mathbf{u}+\mathbf{v}) \in \mathbf{F}^{2n+1}$, where $\mathbf{u},\mathbf{v} \in \mathbf{F}^n$ and $a \in \mathbf{F}$.

I If $a = \pi(\mathbf{u})$, then $I(\mathbf{x}) = \{(\pi(\mathbf{u}),\mathbf{u},\mathbf{u}+\mathbf{c}) \mid \mathbf{c} \in C, d(\mathbf{c},\mathbf{v}) \leq 1\}$.

II If $a \neq \pi(\mathbf{u})$, then $I(\mathbf{x}) = A \cup \{(a,\mathbf{u}',\mathbf{u}+\mathbf{v}) \mid d(\mathbf{u}',\mathbf{u}) = 1, \exists \mathbf{c} \in C : \mathbf{u}+\mathbf{v} = \mathbf{u}'+\mathbf{c}\}$. Here $A = \{(\pi(\mathbf{u}),\mathbf{u},\mathbf{u}+\mathbf{v})\}$ if $\mathbf{v} \in C$, and $A = \emptyset$ if $\mathbf{v} \notin C$.

So in both classes we are interested in codewords $\mathbf{c} \in C$ such that $d(\mathbf{c},\mathbf{v}) \leq 1$. Namely, in the class II the properties $d(\mathbf{u}',\mathbf{u}) = 1$ and $\mathbf{u}+\mathbf{v} = \mathbf{u}'+\mathbf{c}$ imply that also $d(\mathbf{v},\mathbf{c}) = 1$. If $I(\mathbf{x}) = \{(b_i,\mathbf{s}_i,\mathbf{t}_i) \mid i = 1,2,\ldots,k\}$, then in both cases $I(C;\mathbf{v}) = \{\mathbf{s}_i+\mathbf{t}_i \mid i = 1,2,\ldots,k\}$.

Suppose there were words $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$, and $\mathbf{w}$ in $\mathbf{F}^{2n+1}$ such that

$$I(\mathbf{x},\mathbf{y}) = I(\mathbf{z},\mathbf{w}) \text{ and } \{\mathbf{x},\mathbf{y}\} \neq \{\mathbf{z},\mathbf{w}\}, \mathbf{x} \neq \mathbf{y}, \mathbf{z} \neq \mathbf{w}. \tag{4.1}$$

If $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$, and $\mathbf{v}_4$ are $\mathbf{v}'$s of $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$, and $\mathbf{w}$, respectively, then by the previous discussion $I(C;\mathbf{v}_1,\mathbf{v}_2) = I(C;\mathbf{v}_3,\mathbf{v}_4)$. Since C is a $(1,\leq 2)$-identifying code we must have $\{\mathbf{v}_1,\mathbf{v}_2\} = \{\mathbf{v}_3,\mathbf{v}_4\}$. We will show that (4.1) cannot hold. Assume contrary that (4.1) holds.

Because $|I(\mathbf{x})| \geq 3$, we know by (4.1) that at most one of the sets $I(\mathbf{x}) \cap I(\mathbf{z})$ and $I(\mathbf{x}) \cap I(\mathbf{w})$ has cardinality one or less. A similar remark applies to $I(\mathbf{y})$, $I(\mathbf{z})$ and $I(\mathbf{w})$. Hence we can without the loss of generality assume that $|I(\mathbf{x}) \cap I(\mathbf{z})| \geq 2$ and $|I(\mathbf{y}) \cap I(\mathbf{w})| \geq 2$. Depending on which class $\mathbf{x}$ belongs to, also $\mathbf{z}$ belongs to the same class. Indeed, if $\mathbf{x}$ is in the class I, then only last n bits are changing in the codewords of $I(\mathbf{x})$, and, thus, those are the only bits which can change in the codewords of $I(\mathbf{z})$. This means that $\mathbf{z}$ belongs to the class I. Also, if $\mathbf{x}$ belongs to the class II, then the last n bits do not change at all in $I(\mathbf{x})$ and so it is true also for $I(\mathbf{z})$, and hence $\mathbf{z}$ belongs to the class II. Similarly, $\mathbf{y}$ and $\mathbf{w}$ belong to the same class.

1) If $\mathbf{x} = (\pi(\mathbf{u}_1),\mathbf{u}_1,\mathbf{u}_1+\mathbf{v}_1)$ and $\mathbf{y} = (\pi(\mathbf{u}_2),\mathbf{u}_2,\mathbf{u}_2+\mathbf{v}_2)$ are words in the class I, then also $\mathbf{z} = (\pi(\mathbf{u}_3),\mathbf{u}_3,\mathbf{u}_3+\mathbf{v}_3)$ and $\mathbf{w} = (\pi(\mathbf{u}_4),\mathbf{u}_4,\mathbf{u}_4+\mathbf{v}_4)$ are. Since in $I(\mathbf{x})$ and $I(\mathbf{z})$ the codewords begin with the same $n+1$ bits, we get $\mathbf{u}_1 = \mathbf{u}_3$. Similarly $\mathbf{u}_2 = \mathbf{u}_4$. We can assume that $|I(\mathbf{z}) \cap I(\mathbf{y})| \geq 1$ (or that $|I(\mathbf{x}) \cap I(\mathbf{w})| \geq 1$, which is a symmetric case): otherwise $I(\mathbf{z}) = I(\mathbf{x})$ and $I(\mathbf{w}) = I(\mathbf{y})$ and hence $\mathbf{z} = \mathbf{x}$ and $\mathbf{w} = \mathbf{y}$. Since $\mathbf{z}$ and $\mathbf{y}$ are both words in the class I, we get $\mathbf{u}_2 = \mathbf{u}_3$. The fact that $\{\mathbf{v}_1,\mathbf{v}_2\} = \{\mathbf{v}_3,\mathbf{v}_4\}$ now implies that $\{\mathbf{x},\mathbf{y}\} = \{\mathbf{z},\mathbf{w}\}$.

2) Assume $\mathbf{x} = (\pi(\mathbf{u}_1)+1,\mathbf{u}_1,\mathbf{u}_1+\mathbf{v}_1)$ and $\mathbf{y} = (\pi(\mathbf{u}_2)+1,\mathbf{u}_2,\mathbf{u}_2+\mathbf{v}_2)$ and so also $\mathbf{z} = (\pi(\mathbf{u}_3)+1,\mathbf{u}_3,\mathbf{u}_3+\mathbf{v}_3)$ and $\mathbf{w} = (\pi(\mathbf{u}_4)+1,\mathbf{u}_4,\mathbf{u}_4+\mathbf{v}_4)$ are words in the class II. Since in $I(\mathbf{x})$ and $I(\mathbf{z})$ the codewords end with the same n bits as $\mathbf{x}$ and $\mathbf{z}$ we get $\mathbf{u}_1+\mathbf{v}_1 = \mathbf{u}_3+\mathbf{v}_3$, and similarly $\mathbf{u}_2+\mathbf{v}_2 = \mathbf{u}_4+\mathbf{v}_4$. If now $\mathbf{v}_1 = \mathbf{v}_3$ and $\mathbf{v}_2 = \mathbf{v}_4$ we are done, since then $\mathbf{x} = \mathbf{z}$ and $\mathbf{y} = \mathbf{w}$. Suppose therefore that $\mathbf{v}_2 = \mathbf{v}_3$ and $\mathbf{v}_1 = \mathbf{v}_4$. As in the previous case, we can assume that $|I(\mathbf{z}) \cap I(\mathbf{y})| \geq 1$. Now the last n bits must be the same in $I(\mathbf{z})$ and $I(\mathbf{y})$, and thus $\mathbf{u}_3+\mathbf{v}_3 = \mathbf{u}_2+\mathbf{v}_2$ and we get $\mathbf{u}_3 = \mathbf{u}_2$, i.e., $\mathbf{y} = \mathbf{z}$. By Lemma 2.1 the word $\mathbf{z}$ cannot cover the whole $I(\mathbf{x})$, otherwise $\mathbf{z} = \mathbf{x}$. This would imply that $\mathbf{x} = \mathbf{z} = \mathbf{y}$, a contradiction with $\mathbf{x} \neq \mathbf{y}$. So $\mathbf{w}$ must cover at least one word from $I(\mathbf{x})$ which implies $\mathbf{u}_1+\mathbf{v}_1 = \mathbf{u}_4+\mathbf{v}_4$ and now $\mathbf{u}_1 = \mathbf{u}_4$, i.e., $\mathbf{x} = \mathbf{w}$. Therefore $\{\mathbf{x},\mathbf{y}\} = \{\mathbf{z},\mathbf{w}\}$.

3) Suppose finally that $\mathbf{x} = (\pi(\mathbf{u}_1),\mathbf{u}_1,\mathbf{u}_1+\mathbf{v}_1)$ is a word in the class I and $\mathbf{y} = (\pi(\mathbf{u}_2)+1,\mathbf{u}_2,\mathbf{u}_2+\mathbf{v}_2)$ is a word in the class II. Now $\mathbf{z} = (\pi(\mathbf{u}_3),\mathbf{u}_3,\mathbf{u}_3+\mathbf{v}_3)$ is a word in the class I and $\mathbf{w} = (\pi(\mathbf{u}_4)+1,\mathbf{u}_4,\mathbf{u}_4+\mathbf{v}_4)$ is a word in the class

II. Since the first $n+1$ bits are the same in the codewords of $I(\mathbf{x})$ and $I(\mathbf{z})$ we must have $\mathbf{u}_1 = \mathbf{u}_3$. In the codewords of $I(\mathbf{y})$ and $I(\mathbf{w})$ the last n bits are the same, so $\mathbf{u}_2+\mathbf{v}_2 = \mathbf{u}_4+\mathbf{v}_4$. If now $\mathbf{v}_1 = \mathbf{v}_3$ and $\mathbf{v}_2 = \mathbf{v}_4$, we are done since then $\mathbf{x} = \mathbf{z}$ and $\mathbf{y} = \mathbf{w}$. Suppose therefore that $\mathbf{v}_1 = \mathbf{v}_4$ and $\mathbf{v}_2 = \mathbf{v}_3$. As in the previous case, $|I(\mathbf{z}) \cap I(\mathbf{x})| = 2$: otherwise $\mathbf{z} = \mathbf{x}$, $\mathbf{v}_1 = \mathbf{v}_3$ and thus $\mathbf{v}_1 = \mathbf{v}_2 = \mathbf{v}_3 = \mathbf{v}_4$ which together with the equality $\mathbf{u}_2+\mathbf{v}_2 = \mathbf{u}_4+\mathbf{v}_4$ implies that $\mathbf{u}_2 = \mathbf{u}_4$ and hence $\{\mathbf{x},\mathbf{y}\} = \{\mathbf{z},\mathbf{w}\}$. Also $\mathbf{y} \neq \mathbf{w}$, since otherwise $\mathbf{v}_2 = \mathbf{v}_4 = \mathbf{v}_1 = \mathbf{v}_3$ and again $\{\mathbf{x},\mathbf{y}\} = \{\mathbf{z},\mathbf{w}\}$; consequently $d(\mathbf{y},\mathbf{w}) = 2$ and $|I(\mathbf{y})\cap I(\mathbf{w})| = 2$. Now $|I(\mathbf{x})\cap I(\mathbf{w})| = 1$ and $|I(\mathbf{z}) \cap I(\mathbf{y})| = 1$ (these cannot be greater than one, since $\mathbf{x}$ and $\mathbf{w}$, $\mathbf{y}$ and $\mathbf{z}$ belong to different classes).

Let $\mathbf{c}_1$ be the unique codeword in $I(\mathbf{x}) \cap I(\mathbf{w})$, and $\mathbf{c}_2$ be the unique codeword in $I(\mathbf{y})\cap I(\mathbf{z})$. Since all the codewords in $I(\mathbf{x})$ begin with $(\pi(\mathbf{u}_1),\mathbf{u}_1,\ldots)$ and all the codewords in $I(\mathbf{w})$ end in $(\ldots,\mathbf{u}_4+\mathbf{v}_4)$, we know that $\mathbf{c}_1 = (\pi(\mathbf{u}_1),\mathbf{u}_1,\mathbf{u}_4+\mathbf{v}_4)$. Similarly, $\mathbf{c}_2 = (\pi(\mathbf{u}_3),\mathbf{u}_3,\mathbf{u}_2+\mathbf{v}_2)$. Hence $\mathbf{c}_1 = \mathbf{c}_2$, and $I(\mathbf{y}) = I(\mathbf{w})$ and $I(\mathbf{x}) = I(\mathbf{z})$. Therefore $\mathbf{x} = \mathbf{z}$ and $\mathbf{y} = \mathbf{w}$, completing the proof. □

Corollary 4.16. $M_1^{(\leq 2)}(2n+1) \leq 2^n M_1^{(\leq 2)}(n)$.

Corollary 4.17.

$$\begin{aligned} \text{For } k \geq 1: &\quad M_1^{(\leq 2)}(3\cdot 2^k - 1) = 2^{3\cdot 2^k - k - 1}. \\ \text{For } k \geq 3: &\quad M_1^{(\leq 2)}(2^k - 1) = 3\cdot 2^{2^k - k - 1}. \end{aligned}$$

Proof. By Corollary 4.13, we know that $M_1^{(\leq 2)}(5) = 16$. Using Corollary 4.16 recursively and the lower bound from Theorem 4.6 we get the first equality. Similarly, by Theorem 4.14 we get the second claim. □

By Corollary 4.17 we have an optimal $(1,\leq 2)$-identifying code for all lengths greater than or equal to five for which there also exists a perfect 3-fold 1-covering [13, Theorem 14.2.4].

Example 4.18. The code $C = \{00, 01, 10, 11\}$ of length two is a perfect 3-fold 1-covering, but not $(1,\leq 2)$-identifying, since $I(00,11) = I(00,10) = C$. By the construction of Theorem 4.15 we get a perfect 3-fold 1-covering C' of length 5. This code is not $(1,\leq 2)$-identifying, because $\mathbf{c}_1 = 00000$, $\mathbf{c}_2 = 00010$, $\mathbf{c}_3 = 00011$, $\mathbf{c}_4 = 00001$ are codewords in C' and $I(\mathbf{c}_1,\mathbf{c}_2) = I(\mathbf{c}_3,\mathbf{c}_4) = \{\mathbf{c}_1,\mathbf{c}_2,\mathbf{c}_3,\mathbf{c}_4\}$.

The next theorem shows that asymptotically $(1,\leq 2)$-identifying codes reach the lower bound of Theorem 4.6 under the assumption that the Conjecture 2.4 ([13, p. 352]) holds.[1]

[1] Research is done with Tero Laihonen.

Theorem 4.19. *Let $n \geq 5$ and $C' \subseteq \mathbf{F}^n$ be a 3-covering which attains $K(n,3)$. Then the code*

$$C = \{\mathbf{c} \in \mathbf{F}^n \mid 1 \leq d(\mathbf{c}, C') \leq 2\}$$

is $(1, \leq 2)$-identifying.

Proof. For all $\mathbf{x} \in \mathbf{F}^n$ there is a word $\mathbf{c} \in C'$ such that $d(\mathbf{c}, \mathbf{x}) \leq 3$. Let $\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{w} \in \mathbf{F}^n$. If for some $\mathbf{c}' \in C'$ we have $\{\mathbf{x}, \mathbf{y}\} \cap B_3(\mathbf{c}') = \emptyset$ and $\{\mathbf{z}, \mathbf{w}\} \cap B_3(\mathbf{c}') \neq \emptyset$, then it is clear that $I(\mathbf{x}, \mathbf{y}) \neq I(\mathbf{z}, \mathbf{w})$. Hence, it is enough to show that I-sets of different sets $\{\mathbf{x}, \mathbf{y}\}, \{\mathbf{z}, \mathbf{w}\}$ of size at most two are different when $\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{w} \in B_3(\mathbf{c})$ for some $\mathbf{c} \in C'$. We have

$$|I(C; \mathbf{x})| \geq \begin{cases} n & \text{if } \mathbf{x} \in S_0(\mathbf{c}) \cup S_1(\mathbf{c}) \\ 3 & \text{if } \mathbf{x} \in S_2(\mathbf{c}) \cup S_3(\mathbf{c}) \end{cases}.$$

Lemma 4.4 implies that all single words are distinguishable from each other and pairs of words.

Without loss of generality we may assume $\mathbf{c} = \mathbf{0}$, and we denote $S_i(\mathbf{0}) = S_i$. Let us denote $J(\mathbf{x}) = I(\mathbf{x}) \cap (S_1 \cup S_2)$ for $\mathbf{x} \in B_3(\mathbf{0})$. Now $J(\mathbf{x})$ is called the *J-set* of $\mathbf{x}$. We will show that $J(\mathbf{x}, \mathbf{y}) \neq J(\mathbf{z}, \mathbf{w})$ for all $\{\mathbf{x}, \mathbf{y}\} \neq \{\mathbf{z}, \mathbf{w}\}$, where $\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{w} \in B_3(\mathbf{0})$, $\mathbf{x} \neq \mathbf{y}$ and $\mathbf{z} \neq \mathbf{w}$, this implies $I(\mathbf{x}, \mathbf{y}) \neq I(\mathbf{z}, \mathbf{w})$. Assume to the contrary $J(\mathbf{x}, \mathbf{y}) = J(\mathbf{z}, \mathbf{w})$.

Using Lemma 2.1 we see that, if $|J(\mathbf{x})| \geq n$, then there is no set of size two in which $\mathbf{x}$ does not belong to such that this set would cover the whole J-set of $\mathbf{x}$.

i) Assume first $\mathbf{x} \in S_0 \cup S_1$, then without loss of generality $\mathbf{x} = \mathbf{z}$. By Lemma 2.1 we have $1 \leq d(\mathbf{x}, \mathbf{y}) \leq 2$: otherwise $\mathbf{w}$ has to cover the whole J-set of $\mathbf{y}$ which implies $\mathbf{y} = \mathbf{w}$. Similarly, $1 \leq d(\mathbf{x}, \mathbf{w}) \leq 2$. If $\mathbf{y} \in S_0 \cup S_1$, then $\mathbf{y} = \mathbf{w}$ and we are done.

Assume that $\mathbf{x} = \mathbf{z} \in S_0$ and $\mathbf{y}, \mathbf{w} \in S_2 \subseteq C$, now $\mathbf{y} \notin J(\mathbf{z}, \mathbf{w})$ unless $\mathbf{y} = \mathbf{w}$.

Suppose that $\mathbf{x} = \mathbf{z} \in S_1 \subseteq C$ and $\mathbf{y} \in S_2 \subseteq C$. Now $d(\mathbf{x}, \mathbf{y}) = 1$ and in $(J(\mathbf{y}) \setminus \{\mathbf{x}\}) \cap S_1$ there is a codeword that should be covered by $\mathbf{w} \in S_2$. Any possible choice for $\mathbf{w}$ has $J(\mathbf{w})$ which is impossible to cover by $\{\mathbf{x}, \mathbf{y}\}$.

Suppose then $\mathbf{x} = \mathbf{z} \in S_1 \subseteq C$ and $\mathbf{y} \in S_3$, then by the symmetry of the previous case also $\mathbf{w} \in S_3$. Since $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{x}, \mathbf{w}) = 2$ we have either $d(\mathbf{y}, \mathbf{w}) = 2$ or 4. Now $|J(\mathbf{y}) \cap S_2| = |J(\mathbf{w}) \cap S_2| = 3$ and $\mathbf{x} = \mathbf{z}$ covers exactly two codewords from both sets $J(\mathbf{y})$ and $J(\mathbf{w})$. If $d(\mathbf{y}, \mathbf{w}) = 4$, then $(J(\mathbf{w}) \setminus J(\mathbf{x})) \cap S_2 \not\subseteq J(\mathbf{x}, \mathbf{y})$. If $d(\mathbf{y}, \mathbf{w}) = 2$, then $J(\mathbf{w}) \cap J(\mathbf{y}) \cap S_2 = \{\mathbf{c}'\}$. This $\mathbf{c}' \in J(\mathbf{x})$ since the support of $\mathbf{x}$ is the one of the coordinate positions which is in both $\mathbf{y}'$s and $\mathbf{w}'$s supports. Thus, again $J(\mathbf{w}) \not\subseteq J(\mathbf{x}, \mathbf{y})$.

ii) Suppose now that $\mathbf{x} \in S_2$. By symmetry also $\mathbf{y}, \mathbf{z}, \mathbf{w} \in S_2 \cup S_3$. Since $|J(\mathbf{x}) \cap S_1| = 2$ we need two words of S_2 different from $\mathbf{x}$ to cover these words and still $\mathbf{x} \in C$ is not covered. Thus, without loss of generality $\mathbf{z} = \mathbf{x}$. If $\mathbf{y} \in S_2 \subseteq C$, then similarly $\mathbf{w} = \mathbf{y}$. If $\mathbf{y} \in S_3$, then $|(J(\mathbf{y}) \cap S_2) \setminus J(\mathbf{x})| = 2$, but these two words cannot be covered by any $\mathbf{w} \neq \mathbf{y}$.

iii) Finally, let $\mathbf{x} \in S_3$. By symmetry also $\mathbf{y}, \mathbf{z}, \mathbf{w} \in S_3$. Since $|J(\mathbf{x}, \mathbf{y})| \geq 5$ any set $\{\mathbf{z}, \mathbf{w}\} \subseteq S_3, \{\mathbf{z}, \mathbf{w}\} \neq \{\mathbf{x}, \mathbf{y}\}$, cannot cover the set $J(\mathbf{x}, \mathbf{y})$. □

Corollary 4.20. $M_1^{(\leq 2)}(n) \leq K(n,3)\frac{1}{2}(n+n^2)$

Combining Theorem 4.6, Corollary 4.20 and Conjecture 2.4 we get an asymptotic result for $(1, \leq 2)$-identifying codes: When $n \to \infty$

$$M_1^{(\leq 2)}(n) \sim \frac{3 \cdot 2^n}{n}.$$

4.3.2 Direct sum constructions for radii $r \geq 1$

It is proved in [34] that if $C \subseteq \mathbf{F}^n$ is a $(1, \leq 2)$-identifying code, then $C \oplus \mathbf{F}$ is also $(1, \leq 2)$-identifying. In [45] it is proved for $\ell \geq 3$ that a code $C \subseteq \mathbf{F}^n$ is $(1, \leq \ell)$-identifying is and only if it is a $(2\ell - 1)$-fold 1-covering. The direct sum $C \oplus \mathbf{F}$ preserves this property and thus, $C \oplus \mathbf{F} \subseteq \mathbf{F}^{n+1}$ is also a $(1, \leq \ell)$-identifying code. In Theorem 4.22 we prove that for all $\ell \geq 2$ we get a $(2, \leq \ell)$-identifying code $C \subseteq \mathbf{F}^{n+2}$ by taking a direct sum of a $(2, \leq \ell)$-identifying code of length n and $\mathbf{F}^2$. At the same time we also prove the corresponding result for the radius $r = 1$.

Theorems 4.26 and 4.28 generalize the result of Theorem 3.20. We prove that a direct sum of r $(1, \leq \ell)$-identifying codes is $(r, \leq \ell)$-identifying for all $\ell \geq 2$.

Lemma 4.21. *Suppose $r \geq 1$ and $\ell \geq 2$. Let $C \subseteq \mathbf{F}^n$ be an $(r, \leq \ell)$-identifying code. Then for all $\mathbf{x} \in \mathbf{F}^n$ and for every $\mathbf{e} \in S_1(\mathbf{x})$ there is no set of size at most $\ell - 2$, not containing $\mathbf{x}$, that r-covers the set $(C \cap S_r(\mathbf{x})) \setminus B_r(\mathbf{e})$.*

Proof. If for some $\mathbf{x} \in \mathbf{F}^n$ and $\mathbf{e} \in S_1(\mathbf{x})$ there is a set $Y \subseteq \mathbf{F}^n$ such that $|Y| \leq \ell - 2$ and Y r-covers a set $(C \cap S_r(\mathbf{x})) \setminus B_r(\mathbf{e})$. Then $I_r(Y \cup \{\mathbf{x}, \mathbf{e}\}) = I_r(Y \cup \{\mathbf{e}\})$, which is impossible. □

Theorem 4.22. *Let $1 \leq r \leq 2$ and $\ell \geq 2$. If $C \subseteq \mathbf{F}^n$ is an $(r, \leq \ell)$-identifying code, then $D := C \oplus \mathbf{F}^r \subseteq \mathbf{F}^{n+r}$ is $(r, \leq \ell)$-identifying.*

Proof. Let $X, Y \subseteq \mathbf{F}^{n+r}$, $X \neq Y$, $X = \{\mathbf{x}_1, \ldots, \mathbf{x}_{\ell_1}\}$, $Y = \{\mathbf{y}_1, \ldots, \mathbf{y}_{\ell_2}\}$ and $1 \leq \ell_1, \ell_2 \leq \ell$. Let us denote $\mathbf{x}_i = (\mathbf{x}_i^*, \mathbf{x}_i')$ and $\mathbf{y}_j = (\mathbf{y}_j^*, \mathbf{y}_j')$, where $\mathbf{x}_i^*, \mathbf{y}_j^* \in \mathbf{F}^n$ and $\mathbf{x}_i', \mathbf{y}_j' \in \mathbf{F}^r$ for $1 \leq i \leq \ell_1$ and $1 \leq j \leq \ell_2$. Denote $X^* = \{\mathbf{x}_1^*, \ldots, \mathbf{x}_{\ell_1}^*\}$ and $Y^* = \{\mathbf{y}_1^*, \ldots, \mathbf{y}_{\ell_2}^*\}$.

If $X^* \neq Y^*$, then there is $\mathbf{c}^* \in I_r(C; X^*) \triangle I_r(C; Y^*)$. Without loss of generality we can assume $\mathbf{c}^* \in I_r(C; \mathbf{x}_1^*) \setminus I_r(C; Y^*)$. Now $(\mathbf{c}^*, \mathbf{x}_1') \in I_r(D; X) \setminus I_r(D; Y)$.

Suppose $X^* = Y^*$. Assume to the contrary that $I_r(D; X) = I_r(D; Y)$. Because $X \neq Y$, for some i there is $\mathbf{x}_i' \neq \mathbf{y}_h'$ for all h for which $\mathbf{y}_h^* = \mathbf{x}_i^*$. Without loss of generality we can assume $i = 1$. Now $\mathbf{x}_1 \notin Y$ and the words $(\mathbf{x}_1^*, \mathbf{y}_{h_1}'), \ldots, (\mathbf{x}_1^*, \mathbf{y}_{h_j}') \in Y$ do not r-cover codewords at distance r from $\mathbf{x}_1$ ending with $\mathbf{x}_1'$. By Lemma 4.21 we know that such a codeword exists and in Y there must be $\ell - 1$ words which

r-cover these codewords. This implies that every word in Y^* appear only once. This also holds for X^* since $X^* = Y^*$ and $\ell_1 = \ell_2 = \ell$.

By Lemma 4.2 we know that there is a codeword $\mathbf{c}^* \in I_r(C;\mathbf{x}_1^*) \setminus I_r(C;X^* \setminus \{\mathbf{x}_1^*\})$. If $d(\mathbf{c}^*,\mathbf{x}_1^*) \geq 1$, then because $d(\mathbf{x}_1',\mathbf{y}_h') \geq 1$, there is a word $\mathbf{f} \in \mathbf{F}^r$, such that $d(\mathbf{f},\mathbf{x}_1') \leq r - d(\mathbf{c}^*,\mathbf{x}_1^*)$ and $d(\mathbf{f},\mathbf{y}_h') > r - d(\mathbf{c}^*,\mathbf{x}_1^*)$. Thus, $(\mathbf{c}^*,\mathbf{f}) \in I_r(D;X) \setminus I_r(D;Y)$, which is a contradiction.

Suppose therefore that $d(\mathbf{c}^*,\mathbf{x}_1^*) = 0$ and $\mathbf{c}^*$ is the only word in $I_r(C;\mathbf{x}_1^*) \setminus I_r(C;X^* \setminus \{\mathbf{x}_1^*\})$. In particular,

$$d(\mathbf{x}_1^*,X^* \setminus \{\mathbf{x}_1^*\}) \geq r+1. \tag{4.2}$$

Assume there is a word $\mathbf{x}_k \in X \setminus \{\mathbf{x}_1\}$ such that $\mathbf{x}_k \notin Y \setminus \{\mathbf{y}_h\}$, that is $\mathbf{x}_k' \neq \mathbf{y}_s'$, when $\mathbf{x}_k^* = \mathbf{y}_s^*$. As above we get a contradiction unless $\mathbf{x}_k^*$ is the only codeword in $I_r(C;\mathbf{x}_k^*) \setminus I_r(C;X^* \setminus \{\mathbf{x}_k^*\})$, in which case there is $\mathbf{c}_k^* \in I_r(C;\mathbf{x}_1^*) \cap I_r(C;\mathbf{x}_k^*)$ (cf. the third paragraph). Now $I_r(C;(X^* \setminus \{\mathbf{x}_1^*\}) \cup \{\mathbf{c}_k^*\}) = I_r(C;(X^* \setminus \{\mathbf{x}_k^*\}) \cup \{\mathbf{c}_k^*\})$, which is impossible. This means that $X \setminus \{\mathbf{x}_1\} = Y \setminus \{\mathbf{y}_h\}$.

Without loss of generality $h = 1$, and we have $X = \{\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_\ell\}$ and $Y = \{\mathbf{y}_1,\mathbf{x}_2,\ldots,\mathbf{x}_\ell\}$ where $\mathbf{x}_1 \neq \mathbf{y}_1$, $\mathbf{x}_1^* = \mathbf{y}_1^*$ and $\mathbf{x}_1' \neq \mathbf{y}_1'$. Moreover, the set $\{\mathbf{x}_2^*,\ldots,\mathbf{x}_\ell^*\}$ r-covers $I_r(C;\mathbf{x}_1^*) \setminus \{\mathbf{x}_1^*\}$. The assumption $I_r(D;X) = I_r(D;Y)$ implies that the set $\{\mathbf{x}_2,\ldots,\mathbf{x}_\ell\}$ r-covers $I_r(D;\mathbf{x}_1) \bigtriangleup I_r(D;\mathbf{y}_1)$. Suppose $\mathbf{x}_t^* \in X^* \setminus \{\mathbf{x}_1^*\}$ r-covers a codeword $\mathbf{c}_1^* \in (I_r(C;\mathbf{x}_1^*) \setminus I_{r-d(\mathbf{x}_1',\mathbf{y}_1')}(C;\mathbf{x}_1^*)) \setminus I_r(C;X^* \setminus \{\mathbf{x}_1^*,\mathbf{x}_t^*\})$ (such words $\mathbf{x}_t^*$ and $\mathbf{c}_1^*$ always exist). Hence, the word $\mathbf{x}_t = (\mathbf{x}_t^*,\mathbf{x}_t') \in X \cap Y$ r-covers the codewords $(\mathbf{c}_1^*,\mathbf{x}_1') \in I_r(D;\mathbf{x}_1)$ and $(\mathbf{c}_1^*,\mathbf{y}_1') \in I_r(D;\mathbf{y}_1)$ (since $d((\mathbf{x}_1^*,\mathbf{x}_1'),(\mathbf{c}_1^*,\mathbf{y}_1')) \geq r - d(\mathbf{x}_1',\mathbf{y}_1') + 1 + d(\mathbf{x}_1',\mathbf{y}_1') \geq r+1$). We have

$$\begin{array}{rcl} 2r & \geq & d(\mathbf{x}_t,(\mathbf{c}_1^*,\mathbf{x}_1')) + d(\mathbf{x}_t,(\mathbf{c}_1^*,\mathbf{y}_1')) \\ & = & d(\mathbf{x}_t^*,\mathbf{c}_1^*) + d(\mathbf{x}_t',\mathbf{x}_1') + d(\mathbf{x}_t^*,\mathbf{c}_1^*) + d(\mathbf{x}_t',\mathbf{y}_1') \\ & \geq & 2d(\mathbf{x}_t^*,\mathbf{c}_1^*) + d(\mathbf{x}_1',\mathbf{y}_1') \end{array}$$

Hence, $d(\mathbf{x}_t^*,\mathbf{c}_1^*) \leq r - \frac{1}{2}d(\mathbf{x}_1',\mathbf{y}_1')$, which implies

$$\forall\, \mathbf{c}_1^* \in (I_r(C;\mathbf{x}_t^*) \cap (B_r(\mathbf{x}_1^*) \setminus B_{r-d(\mathbf{x}_1',\mathbf{y}_1')}(\mathbf{x}_1^*))) \setminus I_r(X^* \setminus \{\mathbf{x}_1^*,\mathbf{x}_t^*\}) : d(\mathbf{x}_t^*,\mathbf{c}_1^*) \leq r-1. \tag{4.3}$$

This and (4.2) imply $r+1 \leq d(\mathbf{x}_1^*,\mathbf{x}_t^*) \leq 2r-1$. If $r = 1$, this is impossible. From now on $r = 2$ and $d(\mathbf{x}_1^*,\mathbf{x}_t^*) = 3$.

- If $d(\mathbf{x}_1',\mathbf{y}_1') = 2$, then by (4.3) we have $(I_2(C;\mathbf{x}_t^*) \cap I_2(C;\mathbf{x}_1^*)) \setminus I_2(C;X^* \setminus \{\mathbf{x}_1^*,\mathbf{x}_t^*\}) \subseteq I_1(C;\mathbf{x}_t^*) \cap I_2(C;\mathbf{x}_1^*)$. Hence, for every $\mathbf{y}^* \in S_2(\mathbf{x}_1^*)$ such that $d(\mathbf{y}^*,\mathbf{x}_t^*) = 1$ we have

  $$I_2(C;(X^* \setminus \{\mathbf{x}_1^*,\mathbf{x}_t^*\}) \cup \{\mathbf{y}^*\}) = I_2(C;(X^* \setminus \{\mathbf{x}_t^*\}) \cup \{\mathbf{y}^*\}),$$

  which is a contradiction.

- Suppose then $d(\mathbf{x}_1',\mathbf{y}_1')=1$. If $(S_1(\mathbf{x}_1^*)\cap I_2(C;\mathbf{x}_t^*))\setminus I_2(X^*\setminus\{\mathbf{x}_1^*,\mathbf{x}_t^*\})=\emptyset$, then we are done as in the previous case. If there is $\mathbf{c}_2^*\in(S_1(\mathbf{x}_1^*)\cap I_2(C;\mathbf{x}_t^*))\setminus I_2(X^*\setminus\{\mathbf{x}_1^*,\mathbf{x}_t^*\})$, then $(\mathbf{c}_2^*,\mathbf{y}_1'+11)\in I_2(D;\mathbf{x}_1)\setminus I_2(D;Y)$. Namely, $\mathbf{y}_1'+11\neq\mathbf{x}_t'$, otherwise $\mathbf{x}_t$ could not cover any codeword at $S_2(\mathbf{y}_1)$ ending with $\mathbf{y}_1'$.

□

Corollary 4.23. *For $\ell\geq 2$ we have:*

$$M_1^{(\leq\ell)}(n+1)\leq 2M_1^{(\leq\ell)}(n).$$

$$M_2^{(\leq\ell)}(n+2)\leq 4M_2^{(\leq\ell)}(n).$$

Theorem 4.24. *Let $r\geq 1$ and $\ell\geq 2$. If $C\subseteq\mathbf{F}^n$ is an $(r,\leq\ell)$-identifying code, then $D:=C\oplus\mathbf{F}^{r+1}\subseteq\mathbf{F}^{n+r+1}$ is $(r,\leq\ell)$-identifying.*

Proof. The first three paragraphs of the proof of Theorem 4.22 goes similarly. Using the same notations we can continue slightly differently. Now we only have the case $d(\mathbf{x}_1^*,\mathbf{c}^*)\geq 0$. Because $d(\mathbf{x}_1',\mathbf{y}_h')\geq 1$ there is a word $\mathbf{f}\in\mathbf{F}^{r+1}$ such that $d(\mathbf{f},\mathbf{x}_1')\leq r-d(\mathbf{c}^*,\mathbf{x}_1^*)$ and $d(\mathbf{f},\mathbf{y}_h')>r-d(\mathbf{c}^*,\mathbf{x}_1^*)$. Thus, $(\mathbf{c}^*,\mathbf{f})\in I_r(D;X)\setminus I_r(D;Y)$. □

The next example shows that Theorem 4.22 cannot be generalized for $(3,\leq 2)$-identifying codes.

Example 4.25. By a computer it can be shown that the code

$$C=\{\mathbf{0}\}\cup(S_3(\mathbf{0})\cap S_2(11111000))\cup(\mathbf{F}^8\setminus B_4(\mathbf{0}))\subseteq\mathbf{F}^8$$

is $(3,\leq 2)$-identifying code of length 8. The code $C\oplus\mathbf{F}^3$ is not $(3,\leq 2)$-identifying since

$$I_3(00000000000,11111000000)=I_3(00000000001,11111000000).$$

Denote $N=\sum_{i=1}^r n_i$ and $N^*=N-n_r$.

Theorem 4.26. *Suppose $r\geq 1$. Let $C_i\subseteq\mathbf{F}^{n_i}$ for $1\leq i\leq r$ be $(1,\leq 2)$-identifying codes. Then $C=C_1\oplus\ldots\oplus C_r$ is an $(r,\leq 2)$-identifying code.*

Proof. Let $X=\{\mathbf{x},\mathbf{y}\}$, $Y=\{\mathbf{z},\mathbf{w}\}\subseteq\mathbf{F}^N$, $X\neq Y$ and $|X|,|Y|\leq 2$. Denote $\mathbf{x}=(\mathbf{x}_1,\ldots,\mathbf{x}_r)$, $\mathbf{y}=(\mathbf{y}_1,\ldots,\mathbf{y}_r)$, $\mathbf{z}=(\mathbf{z}_1,\ldots,\mathbf{z}_r)$ and $\mathbf{w}=(\mathbf{w}_1,\ldots,\mathbf{w}_r)$. If for some k there is $\{\mathbf{x}_k,\mathbf{y}_k\}\neq\{\mathbf{z}_k,\mathbf{w}_k\}$, then because C_k is $(1,\leq 2)$-identifying there is $\mathbf{c}_k\in I_1(C_k;\mathbf{x}_k,\mathbf{y}_k)\ \triangle\ I_1(C_k;\mathbf{z}_k,\mathbf{w}_k)$. Without loss of generality we may assume $\mathbf{c}_k\in I_1(C_k;\mathbf{x}_k)\setminus I_1(C_k;\mathbf{z}_k,\mathbf{w}_k)$. For all $1\leq h\leq r$, $h\neq k$, we have $|I_1(C_h;\mathbf{x}_h)|\geq 3>|\{\mathbf{z}_h,\mathbf{w}_h\}|$. Thus, there is $\mathbf{c}_h\in I_1(C_h;\mathbf{x}_h)$ such that $d(\mathbf{c}_h,\{\mathbf{z}_h,\mathbf{w}_h\})\geq 1$. Hence $(\mathbf{c}_1,\ldots,\mathbf{c}_{k-1},\mathbf{c}_k,\mathbf{c}_{k+1},\ldots,\mathbf{c}_r)\in I_r(C;\mathbf{x})\setminus I_r(C;Y)$.

Suppose that for all $1 \leq k \leq r$ we have $\{\mathbf{x}_k, \mathbf{y}_k\} = \{\mathbf{z}_k, \mathbf{w}_k\}$. This is possible only if $|X| = |Y| = 2$. Because $X \neq Y$ we may assume

$$\begin{aligned} \mathbf{x} &= (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots, \mathbf{x}_r) & \mathbf{z} &= (\mathbf{x}_1, \mathbf{y}_2, \mathbf{z}_3, \ldots, \mathbf{z}_r) \\ \mathbf{y} &= (\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3, \ldots, \mathbf{y}_r) & \mathbf{w} &= (\mathbf{y}_1, \mathbf{x}_2, \mathbf{w}_3, \ldots, \mathbf{w}_r) \end{aligned}$$

where $\mathbf{x}_1 \neq \mathbf{y}_1$, $\mathbf{x}_2 \neq \mathbf{y}_2$ and $\{\mathbf{x}_k, \mathbf{y}_k\} = \{\mathbf{z}_k, \mathbf{w}_k\}$ for $k \geq 3$. Because $\mathbf{x}_h \in \{\mathbf{z}_h, \mathbf{w}_h\}$ for all $3 \leq h \leq r$, then $|I_1(C_h; \mathbf{x}_h) \setminus \{\mathbf{x}_h\}| \geq 2 > |\{\mathbf{z}_h, \mathbf{w}_h\} \setminus \{\mathbf{x}_h\}|$. Thus, for all $3 \leq h \leq r$ there are $\mathbf{c}_h \in I_1(C_h; \mathbf{x}_h)$ such that $d(\mathbf{c}_h, \{\mathbf{z}_h, \mathbf{w}_h\}) \geq 1$. Similarly, we find corresponding codewords for $\mathbf{y}_3, \ldots, \mathbf{y}_r$.

There is $\mathbf{c}_i \in I_1(C_i; \mathbf{x}_i) \setminus I_1(C_i; \mathbf{y}_i)$ for $i = 1, 2$. There are two possible cases

A) $d(\mathbf{x}_i, \mathbf{c}_i) = 1$ and

B) $d(\mathbf{x}_i, \mathbf{c}_i) = 0$, if there is no codeword $\mathbf{c}_i$ such that the case A would hold.

In the latter case we must have $d(\mathbf{x}_i, \mathbf{y}_i) = 2$ and $|I_1(C_i; \mathbf{x}_i) \cap I_1(C_i; \mathbf{y}_i)| = 2$. Now Lemma 4.5 implies that there is $\mathbf{c}_{\mathbf{y}_i} \in I_1(C_i; \mathbf{y}_i) \setminus I_1(C_i; \mathbf{x}_i)$ such that $d(\mathbf{x}_i, \mathbf{c}_{\mathbf{y}_i}) = 3$ and $d(\mathbf{y}_i, \mathbf{c}_{\mathbf{y}_i}) = 1$. If both $\mathbf{x}_1$ and $\mathbf{x}_2$ belong to the case A, then $(\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \ldots, \mathbf{c}_r) \in I_r(C; X) \setminus I_r(C; Y)$. If $\mathbf{x}_1$ belong to the case A and $\mathbf{x}_2$ belong to the case B, then $(\mathbf{c}_1, \mathbf{c}_{\mathbf{y}_2}, \mathbf{c}_3, \ldots, \mathbf{c}_r) \in I_r(C; Y) \setminus I_r(C; X)$. The case $\mathbf{x}_1$ in B and $\mathbf{x}_2$ in A goes similarly. If both $\mathbf{x}_1$ and $\mathbf{x}_2$ belong to the case B, then $(\mathbf{x}_1, \mathbf{c}_{\mathbf{y}_2}, \mathbf{c}_3, \ldots, \mathbf{c}_r) \in I_r(C; Y) \setminus I_r(C; X)$. □

When proving the corresponding result for $\ell \geq 3$ the next lemma is used.

Lemma 4.27. *Let $\ell \geq 3$ and $C \subseteq \mathbf{F}^n$ is an $(1, \leq \ell)$-identifying code. For all $\mathbf{x} \in \mathbf{F}^n$ and $Y \subseteq \mathbf{F}^n$, $|Y| \leq \ell$, we have $\mathbf{c} \in I_1(\mathbf{x}) \setminus \{\mathbf{x}\}$ such that $d(\mathbf{c}, Y) \geq 1$.*

Proof. By [45] or Theorem 4.3 we know that for all $\mathbf{x} \in \mathbf{F}^n$ we have $|I_1(\mathbf{x})| \geq 2\ell - 1$. Thus, $|I_1(\mathbf{x}) \setminus \{\mathbf{x}\}| \geq 2\ell - 2 > \ell \geq |Y|$, which implies the claim. □

Theorem 4.28. *Suppose $\ell \geq 3$. Let $C_i \subseteq \mathbf{F}^{n_i}$ for $1 \leq i \leq r$ be $(1, \leq \ell)$-identifying codes. Then*

$$C = C_1 \oplus \ldots \oplus C_r \subseteq \mathbf{F}^N$$

is an $(r, \leq \ell)$-identifying code.

Proof. By [45] or Theorem 4.3 we know that for all $\mathbf{x} \in \mathbf{F}^{n_i}$ we have $|I_1(C_i; \mathbf{x})| \geq 2\ell - 1$.

We prove by induction on r that $C = C_1 \oplus \ldots \oplus C_r$ is an $(r, \leq \ell)$-identifying code and, moreover, for every $X, Y \subseteq \mathbf{F}^N$, $1 \leq |X|, |Y| \leq \ell$ and $X \neq Y$, there is $\mathbf{c} \in I_r(C; \mathbf{x}) \setminus I_r(C; Y)$ such that $r - 1 \leq d(\mathbf{c}, \mathbf{x}) \leq r$ for some $\mathbf{x} \in X$ or $\mathbf{c} \in I_r(C; \mathbf{y}) \setminus I_r(C; X)$ such that $r - 1 \leq d(\mathbf{c}, \mathbf{y}) \leq r$ for some $\mathbf{y} \in Y$. The first step of induction is trivial, $r = 1$. The induction hypothesis is that the claim holds for $C^* = C_1 \oplus \ldots \oplus C_{r-1}$.

Let $X, Y \subseteq \mathbf{F}^N$, $1 \leq |X|, |Y| \leq \ell$, $X \neq Y$, $X = \{\mathbf{x}_1, \ldots, \mathbf{x}_{\ell_1}\}$ and $Y = \{\mathbf{y}_1, \ldots, \mathbf{y}_{\ell_2}\}$ and denote

$$\mathbf{x}_i = (\mathbf{x}_{i,1}, \ldots, \mathbf{x}_{i,r-1}, \mathbf{x}_{i,r}) = (\mathbf{x}_i^*, \mathbf{x}_{i,r}),\ \mathbf{y}_j = (\mathbf{y}_{j,1}, \ldots, \mathbf{y}_{j,r-1}, \mathbf{y}_{j,r}) = (\mathbf{y}_j^*, \mathbf{y}_{j,r})$$

for $1 \leq i \leq \ell_1$ and $1 \leq j \leq \ell_2$. Denote $X^* = \{\mathbf{x}_1^*, \ldots, \mathbf{x}_{\ell_1}^*\}$ and $Y^* = \{\mathbf{y}_1^*, \ldots, \mathbf{y}_{\ell_2}^*\}$, $X_i = \{\mathbf{x}_{1,i}, \ldots \mathbf{x}_{\ell_1,i}\}$ and $Y_i = \{\mathbf{y}_{1,i}, \ldots, \mathbf{y}_{\ell_2,i}\}$, for $1 \leq i \leq r$.

If $X^* \neq Y^*$, then the induction hypothesis implies that there is a codeword $\mathbf{c}^* \in I_{r-1}(C^*; X^*) \triangle I_{r-1}(C^*; Y^*)$. Without loss of generality we may assume

$$\mathbf{c}^* \in I_{r-1}(C^*; \mathbf{x}_1^*) \setminus I_{r-1}(C^*; Y^*).$$

The induction hypothesis implies, that $r-2 \leq d(\mathbf{c}^*, \mathbf{x}_1^*) \leq r-1$. By Lemma 4.27 we know that there is $\mathbf{c}_r \in I_1(C_r; \mathbf{x}_{1,r}) \setminus \{\mathbf{x}_{1,r}\}$ such that $d(\mathbf{c}_r, Y_r) \geq 1$. Hence, $(\mathbf{c}^*, \mathbf{c}_r) \in I_r(C; X) \setminus I_r(C; Y)$. Moreover, $r-1 \leq d((\mathbf{c}^*, \mathbf{c}_r), \mathbf{x}_1) \leq r$.

Suppose next $X^* = Y^*$ and $X_r \neq Y_r$. Because C_r is $(1, \leq \ell)$-identifying there is $\mathbf{c}_r \in I_1(C_r; X_r) \triangle I_1(C_r; Y_r)$. Without loss of generality we may assume $\mathbf{c}_r \in I_1(C_r; \mathbf{x}_{1,r}) \setminus I_1(C_r; Y_r)$. By Lemma 4.27, for every $1 \leq k \leq r-1$ there is $\mathbf{c}_k \in I_1(C_k; \mathbf{x}_{1,k}) \setminus \{\mathbf{x}_{1,k}\}$ such that $d(\mathbf{c}_k, Y_k) \geq 1$. Hence, $(\mathbf{c}_1, \ldots, \mathbf{c}_{r-1}, \mathbf{c}_r) \in I_r(C; X) \setminus I_r(C; Y)$. Moreover, $(\mathbf{c}_1, \ldots, \mathbf{c}_{r-1}, \mathbf{c}_r) \in S_{r-1}(\mathbf{x}_1) \cup S_r(\mathbf{x}_1)$.

Suppose then $X^* = Y^*$ and $X_r = Y_r$. There is for some k, $\mathbf{x}_{k,r} \neq \mathbf{y}_{h,r}$ for all h for which $\mathbf{x}_k^* = \mathbf{y}_h^*$, otherwise $X = Y$.

- Suppose $|Y^* \setminus \{\mathbf{x}_k^*\}| \leq \ell - 2$, then by Lemma 4.21 there is a codeword $\mathbf{c}^* \in (C^* \cap S_{r-1}(\mathbf{x}_k^*)) \setminus I_{r-1}(Y^* \setminus \{\mathbf{x}_k^*\})$. There is $\mathbf{c}_r \in I_1(C_r; \mathbf{x}_{k,r}) \setminus I_1(C_r; Y_r \setminus \{\mathbf{x}_{k,r}\})$. If $d(\mathbf{c}_r, \mathbf{x}_{k,r}) = 1$, then $(\mathbf{c}^*, \mathbf{c}_r) \in I_r(C; X) \setminus I_r(C; Y)$.

  Suppose $\mathbf{c}_r = \mathbf{x}_{k,r}$ is the only codeword in $I_1(C_r; \mathbf{x}_{k,r}) \setminus I_1(C_r; Y_r \setminus \{\mathbf{x}_{k,r}\})$. This implies $|I_1(C_r; \mathbf{x}_{k,r}) \setminus \{\mathbf{x}_{k,r}\}| = 2\ell - 2$, $|Y_r \setminus \{\mathbf{x}_{k,r}\}| = \ell - 1$, for all $\mathbf{y}_{h,r} \in Y_r \setminus \{\mathbf{x}_{k,r}\}$ we have $d(\mathbf{x}_{k,r}, \mathbf{y}_{h,r}) = 2$ and $|I_1(C_r; \mathbf{x}_{k,r}) \cap I_1(C_r; \mathbf{y}_{h,r})| = 2$. Moreover, we have $d(\mathbf{y}_{h_1,r}, \mathbf{y}_{h_2,r}) = 4$, for $h_1 \neq h_2$. There is $\mathbf{x}_t^* \in X^* = Y^*$ such that $\mathbf{x}_k^* \neq \mathbf{x}_t^*$. Otherwise, $X^* = \{\mathbf{x}_k^*\} = Y^*$ and $X_r = Y_r$ imply that $X = Y$. Suppose there is $\mathbf{y}_{t,r}$ such that $(\mathbf{x}_t^*, \mathbf{y}_{t,r}) \in Y$ and $(\mathbf{x}_k^*, \mathbf{y}_{t,r}) \notin Y$. Let us choose $\mathbf{c}'_r \in I_1(C_r; \mathbf{x}_{k,r}) \cap I_1(C_r; \mathbf{y}_{t,r})$. As mentioned above $d(\mathbf{c}'_r, \mathbf{x}_{k,r}) = d(\mathbf{c}'_r, \mathbf{y}_{t,r}) = 1$ and $d(\mathbf{c}'_r, \mathbf{y}_{h,r}) = 3$ for all $\mathbf{y}_{h,r} \neq \mathbf{x}_{k,r}$ and $\mathbf{y}_{h,r} \neq \mathbf{y}_{t,r}$. We get $(\mathbf{c}^*, \mathbf{c}'_r) \in I_r(C; X) \setminus I_r(C; Y)$, moreover $d((\mathbf{c}^*, \mathbf{c}'_r), (\mathbf{x}_k^*, \mathbf{x}_{k,r})) = r$.

  If such an $\mathbf{y}_{t,r}$ does not exist, then $X^* = Y^* = \{\mathbf{x}_k^*, \mathbf{x}_t^*\}$ and $\mathbf{x}_k^*$ appears $\ell - 1$ times in Y^*. Now $(\mathbf{x}_t^*, \mathbf{x}_{k,r}) \in Y \setminus X$ and $(\mathbf{x}_t^*, \mathbf{x}_{t,r}) \in X \setminus Y$. Moreover, $|Y^* \setminus \{\mathbf{x}_t^*\}| = 1 \leq \ell - 2$, as above we have $\mathbf{c}_t^* \in (C^* \cap S_{r-1}(\mathbf{x}_t^*)) \setminus I_{r-1}(C^*; Y^* \setminus \{\mathbf{x}_t^*\})$. Because $|I_1(C_r; \mathbf{x}_{t,r})| \geq 2\ell - 1 > 3 \geq |(I_1(\mathbf{x}_{t,r}) \cap I_1(\mathbf{x}_{k,r})) \cup \{\mathbf{x}_{t,r}\}|$ there is $\mathbf{c}_{t,r} \in I_1(C_r; \mathbf{x}_{t,r}) \setminus (\{\mathbf{x}_{t,r}\} \cup I_1(C_r; \mathbf{x}_{k,r}))$. As above it is proved, we know that $d(\mathbf{c}_{t,r}, \mathbf{y}_{h,r}) \geq 3$ for all $\mathbf{y}_{h,r} \in Y_r$, $\mathbf{y}_{h,r} \neq \mathbf{x}_{t,r}$. Hence, $(\mathbf{c}_t^*, \mathbf{c}_{t,r}) \in I_r(C; (\mathbf{x}_t^*, \mathbf{x}_{t,r})) \setminus I_r(C; Y)$. Moreover, $d((\mathbf{x}_t^*, \mathbf{x}_{t,r}), (\mathbf{c}_t^*, \mathbf{c}_{t,r})) = r$.

- Suppose then that $|Y^* \setminus \{\mathbf{x}_k^*\}| = \ell - 1$. This implies $|Y^*| = \ell$ and every word in Y^* appears there only once. Because $X^* = Y^*$ the same holds for X^*, as well. Hence, there is $(\mathbf{x}_k^*, \mathbf{x}_{k,r}) \in X \setminus Y$ and $(\mathbf{x}_k^*, \mathbf{y}_{h,r}) \in Y \setminus X$ for some h. By the induction hypothesis there is $\mathbf{c}^* \in I_{r-1}(C^*; X^*) \triangle I_{r-1}(C^*; X^* \setminus \{\mathbf{x}_k^*\})$. This implies $\mathbf{c}^* \in I_{r-1}(C^*; \mathbf{x}_k^*)$ and $r-2 \leq d(\mathbf{c}^*, \mathbf{x}_k^*) \leq r-1$.
  - Suppose $d(\mathbf{c}^*, \mathbf{x}_k^*) = r-1$. Because

    $$|(I_1(C_r; \mathbf{x}_{k,r}) \setminus \{\mathbf{x}_{k,r}\}) \setminus I_1(C_r; \mathbf{y}_{h,r})| \geq 2\ell - 4 > \ell - 2 = |Y_r \setminus \{\mathbf{x}_{k,r}, \mathbf{y}_{h,r}\}|$$

    we know that there is $\mathbf{c}_r \in I_1(C_r; \mathbf{x}_{k,r}) \setminus \{\mathbf{x}_{k,r}\}$ such that $d(\mathbf{c}_r, \mathbf{y}_{h,r}) \geq 2$ and $d(\mathbf{c}_r, \mathbf{y}_{j,r}) \geq 1$ for $j \neq h$. Thus, $(\mathbf{c}^*, \mathbf{c}_r) \in I_r(C; X) \setminus I_r(C; Y)$ and $d((\mathbf{c}^*, \mathbf{c}_r), (\mathbf{x}_k^*, \mathbf{x}_{k,r})) = r$.
  - Suppose $d(\mathbf{c}^*, \mathbf{x}_k^*) = r-2$. We separate cases depending on the distance between $\mathbf{x}_{k,r}$ and $\mathbf{y}_{h,r}$. In every case, we will find a codeword $\mathbf{c}_r$ such that $1 \leq d(\mathbf{x}_{k,r}, \mathbf{c}_r) \leq 2$, $d(\mathbf{y}_{h,r}, \mathbf{c}_r) \geq 3$ and $d(\mathbf{c}_r, \mathbf{y}_{j,r}) \geq 1$, for $j \neq h$. Then $(\mathbf{c}^*, \mathbf{c}_r) \in I_r(C; X) \setminus I_r(C; Y)$ and it satisfies the wanted distance properties. If $d(\mathbf{x}_{k,r}, \mathbf{y}_{h,r}) \geq 4$, then clearly there is $\mathbf{c}_r \in I_1(C_r; \mathbf{x}_{k,r} \setminus \{\mathbf{x}_{k,r}\})$ which satisfies the conditions.

    If $d(\mathbf{x}_{k,r}, \mathbf{y}_{h,r}) = 3$, then $|S_1(\mathbf{x}_{k,r}) \cap B_2(\mathbf{y}_{h,r})| = 3$. Because always $n_i \geq 4$ (otherwise no $(1, \leq \ell)$-identifying code exists), there is $\mathbf{z} \in S_1(\mathbf{x}_{k,r}) \setminus B_2(\mathbf{y}_{h,r})$. Now $|I_1(C_r; \mathbf{z}) \setminus \{\mathbf{x}_{k,r}, \mathbf{z}\}| \geq 2\ell - 3 > \ell - 2 \geq |Y_r \setminus \{\mathbf{x}_{k,r}, \mathbf{y}_{h,r}\}|$ implies that there is $\mathbf{c}_r \in I_1(C_r; \mathbf{z})$ which satisfies the conditions.

    If $d(\mathbf{x}_{k,r}, \mathbf{y}_{h,r}) = 2$, then $|(I_1(C_r; \mathbf{x}_{k,r}) \setminus \{\mathbf{x}_{k,r}\}) \setminus B_2(\mathbf{y}_{h,r})| \geq 2\ell - 4 > \ell - 2 = |Y_r \setminus \{\mathbf{x}_{k,r}, \mathbf{y}_{h,r}\}|$. Thus, we find $\mathbf{c}_r \in I_1(C_r; \mathbf{x}_{k,r})$ which satisfies the conditions.

    If $d(\mathbf{x}_{k,r}, \mathbf{y}_{h,r}) = 1$, then there is $\mathbf{z} \in S_1(\mathbf{x}_{k,r})$, $\mathbf{z} \neq \mathbf{y}_{h,r}$ and

    $$\begin{aligned} |I_1(C_r; \mathbf{z}) \setminus (\{\mathbf{x}_{k,r}, \mathbf{z}\} \cup (I_1(\mathbf{y}_{h,r}) \cap I_1(\mathbf{z})))| &\geq 2\ell - 4 > \ell - 2 \\ &\geq |Y_r \setminus \{\mathbf{x}_{k,r}, \mathbf{y}_{h,r}\}|. \end{aligned}$$

    Thus, there is $\mathbf{c}_r \in I_1(C_r; \mathbf{z}) \cap S_2(\mathbf{x}_{k,r})$, which satisfies the conditions.

□

Combining the results of Theorems 3.18, 4.26 and 4.28 we get the next corollary.

Corollary 4.29. *For $r \geq 1$ and $\ell \geq 1$ we have*

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

In Table 4.1 we have collected lower and upper bounds on the cardinalities of $(1, \leq 2)$- and $(2, \leq 2)$-identifying codes.

Table 4.1: Bounds on the cardinalities of $(1,\leq 2)$- and $(2,\leq 2)$-identifying codes.

n	$(1,\leq 2)$	$(2,\leq 2)$
4	11 a	–
5	16 a	–
6	30 – 32 b	12 – 25 d
7	48 a	14 – 35 d
8	90 – 96 b	22 – 70 d
9	154 – 176 c	35 – 140 b
10	289 – 352 b	56 – 256 (5,5)
11	512 c	94 – 512 (5,6)
12	972 – 1024 b	158 – 768 (5,7)
13	1756 – 2048 b,c	271 – 1536 (6,7)
14	3356 – 4096 b	469 – 2304 (7,7)
15	6144 c	820 – 4608 (7,8)
16	11809 – 12288 b	1446 – 8448 (7,9)
17	21846 – 24576 b,c	2571 – 16384 (6,11)
18	42164 – 49152 b	4600 – 24576 (7,11)
19	78644 – 90112 c	8279 – 49152 (7,12)
20	152304 – 180224 b	14980 – 90112 (9,11)
21	285976 – 360448 b,c	27236 – 180224 (10,11)
22	555411 – 720896 b	49735 – 262144 (11,11)
23	1048576 c	91181 – 524288 (11,12)

The lower bounds for $(1,\leq 2)$-identifying codes come from Theorem 4.6. The lower bounds for $(2,\leq 2)$-identifying codes for $n \geq 7$ come from Theorem 4.7. The lower bound for $(2,\leq 2)$-identifying code for $n = 6$ come from Theorem 4.1. The upper bounds for $(2,\leq 2)$-identifying for $n \geq 9$ come from Corollary 4.29. The used lengths are mentioned in parenthesis.

Key to upper bounds

a	Short code constructions in Section 4.2
b	For $1 \leq r \leq 2$ and $\ell \geq 2 : M_r^{(\leq \ell)}(n+1) \leq 2^r M_1^{(\leq \ell)}(n)$, Theorem 4.22.
c	$(\pi(\mathbf{u}), \mathbf{u}, \mathbf{u}+\mathbf{v})$-construction, Theorem 4.15.
d	See Appendix

Chapter 5

Strongly identifying codes

In this chapter, we consider a modification of identifying codes, strongly identifying codes. Now we want to identify a set of vertices under the assumption that codewords which belong to a fault pattern might be missing from the I-set.

The function of a strongly identifying code is easier to understand by an example. If a processor that belongs to a fault diagnosis system of a multiprocessor mechanism is malfunctioning then it can fail on sending information. If a controller cannot trust on received information some extra conditions are needed before it is possible to identify malfunctioning processors. In this case we need strongly identifying codes. In this model malfunctioning processors send report which may be correct or wrong.

In order to identify fault patterns we require that a code $C \subseteq \mathbf{F}^n$ satisfies the following. Let for any different subsets X and Y of $\mathbf{F}^n$ ($|X|,|Y| \leq \ell$) the sets $I_r(X) \setminus S$ and $I_r(Y) \setminus T$, where $S \subseteq X \cap C$ and $T \subseteq Y \cap C$, be nonempty and distinct. Then obviously, we can always distinguish between X and Y. The sets $I_r(X) \setminus S$ and $I_r(X) \setminus S'$ where $S, S' \subseteq X \cap C$ ($S \neq S'$) are automatically different from each other. This leads to the following definition.

Definition 5.1. Let $C \subseteq \mathbf{F}^n$ be a code and $\ell \geq 1$ an integer. Let further $X \subseteq \mathbf{F}^n$ and $|X| \leq \ell$. Define

$$\mathscr{I}_r(X) = \{U \mid I_r(X) \setminus (X \cap C) \subseteq U \subseteq I_r(X)\}. \tag{5.1}$$

If for all $X_1, X_2 \subseteq \mathbf{F}^n$, where $X_1 \neq X_2$ and $|X_1|, |X_2| \leq \ell$, we have $\mathscr{I}_r(X_1) \cap \mathscr{I}_r(X_2) = \emptyset$, then we say that C is a *strongly* $(r, \leq \ell)$*-identifying code*.

If we replace (5.1) by $\mathscr{I}_r(X) = \{I_r(X)\}$, we get the definition of a (regular) $(r, \leq \ell)$-identifying code. Therefore, a strongly identifying code is always a (regular) identifying code.

We call a strongly $(r, \leq 1)$-identifying code a strongly r-identifying code and we denote $I'_r(\mathbf{y}) = I_r(\mathbf{y}) \setminus \{\mathbf{y}\}$. The smallest cardinality of a strongly $(r, \leq \ell)$-identifying code of length n is denoted by $M_r^{(\leq \ell)SID}(n)$. When $\ell = 1$, we denote

$M_r^{(\leq 1)SID}(n) = M_r^{SID}(n)$. Usually, we omit r from these notations if $r = 1$. A code attaining the smallest cardinality is called *optimal*.

In this chapter we consider strongly 1-identifying codes and strongly $(1, \leq 2)$-identifying codes. In both cases we present constructions and lower bounds. The chapter is based on [35, 48].

5.1 Strongly 1-identifying codes

In this section we construct strongly 1-identifying codes and prove lower bounds for them. Strongly 1-identifying codes lie between 1-identifying codes and $(1, \leq 2)$-identifying codes, as we will see.

5.1.1 Constructions

The next construction is optimal for length 4.

Theorem 5.2. *For $n \geq 4$ we have $M^{SID}(n) \leq 2^{n-1}$.*

Proof. The set $\mathbf{F}^{n-1} \oplus \{0\}$ is a strongly 1-identifying code. Here each codeword $\mathbf{x}$ is covered by exactly n codewords and every non-codeword $\mathbf{y} = \mathbf{y}'1$, where $\mathbf{y}' \in \mathbf{F}^{n-1}$ is covered by the unique codeword $\mathbf{y}'0$. Thus, clearly, $I(\mathbf{y}) \neq I(\mathbf{x}) \neq I'(\mathbf{y})$ and $I'(\mathbf{x}) \neq I'(\mathbf{y})$ for any distinct words $\mathbf{x}$ and $\mathbf{y}$. □

Theorem 5.3. *Let $C \subseteq \mathbf{F}^n$ be a 1-identifying code with the property that $d(\mathbf{c}, C \setminus \{\mathbf{c}\}) = 1$ for all $\mathbf{c} \in C$. Then $D = C \oplus \mathbf{F}$ is strongly 1-identifying.*

Proof. Let $\mathbf{x} = (\mathbf{x}_1, x_2), \mathbf{y} = (\mathbf{y}_1, y_2) \in \mathbf{F}^{n+1}$, where $\mathbf{x}_1, \mathbf{y}_1 \in \mathbf{F}^n$ and $x_2, y_2 \in \mathbf{F}$. If $\mathbf{x}_1 \neq \mathbf{y}_1$, then there is $\mathbf{c} \in I(C; \mathbf{x}_1) \bigtriangleup I(C; \mathbf{y}_1)$. Without loss of generality we can assume that $\mathbf{c} \in I(C; \mathbf{x}_1) \setminus I(C; \mathbf{y}_1)$. If $d(\mathbf{c}, \mathbf{x}_1) = 1$, then $(\mathbf{c}, x_2) \in I'(D; \mathbf{x}) \setminus I(D; \mathbf{y})$. If $d(\mathbf{c}, \mathbf{x}_1) = 0$, then $(\mathbf{c}, x_2 + 1) \in I'(D; \mathbf{x}) \setminus I(D; \mathbf{y})$. If $\mathbf{x}_1 = \mathbf{y}_1$, then $x_2 \neq y_2$. Because $d(C \setminus \{\mathbf{c}\}, \mathbf{c}) = 1$ for all $\mathbf{c} \in C$ we know that there is $\mathbf{c} \in C$ such that $d(\mathbf{c}, \mathbf{x}_1) = 1$. Now $(\mathbf{c}, x_2) \in I'(D; \mathbf{x}) \setminus I(D; \mathbf{y})$. This proves the claim. □

Corollary 5.4. *If C is a strongly 1-identifying code, then its direct sum with $\mathbf{F}$ is as well.*

In Chapter 7, Theorem 7.8, it will be proved that a direct sum of a 1-locating-dominating code and $\mathbf{F}^2$ is a 1-identifying code, which satisfies the property of the previous theorem. Thus, we have the next theorem, where $L(n)$ denotes the smallest cardinality of a 1-locating-dominating code of length n.

Theorem 5.5.

$$M^{SID}(n+3) \leq 8L(n).$$

In the next theorem we show that a $(1,\leq 2)$-identifying code is strongly 1-identifying.

Theorem 5.6. $M^{SID}(n) \leq M_1^{(\leq 2)}(n)$.

Proof. Let $C \subseteq \mathbf{F}^n$ be a $(1,\leq 2)$-identifying code with $M_1^{(\leq 2)}(n)$ codewords. By Theorem 4.3 every word in $\mathbf{F}^n$ is covered by at least three codewords. Thus, for all $\mathbf{x} \in \mathbf{F}^n$ the set $I(\mathbf{x})$ is unique. Moreover, $I'(\mathbf{c}) \neq I'(\mathbf{c}')$ for all $\mathbf{c},\mathbf{c}' \in C$ $(\mathbf{c} \neq \mathbf{c}')$. In particular, if $I'(\mathbf{c}) = \{\mathbf{c}_1,\mathbf{c}_2\} = I'(\mathbf{c}')$, then $I(\mathbf{c}_1,\mathbf{c}) = I(\mathbf{c}_1,\mathbf{c}')$ and this is not possible. □

Theorem 5.7. $M^{SID}(n) \leq 2n \cdot K(n-1,2)$.

Proof. Suppose $n \geq 4$. Let $C' \subseteq \mathbf{F}^{n-1}$ be a code attaining the bound $K(n-1,2)$. Denote $\mathbf{e}_1 = 10\ldots0 \in \mathbf{F}^{n-1}$. A code

$$C := \{\mathbf{c}+\mathbf{a} \mid \mathbf{c} \in C', \mathbf{a} \in B_1(\mathbf{0}) \setminus \mathbf{e}_1\}$$

is 1-identifying with the property that for all $\mathbf{c} \in C$ we have $d(\mathbf{c}, C \setminus \{\mathbf{c}\}) = 1$. Namely, (compare the proof of [40, Theorem 8]) for all $\mathbf{x} \in \mathbf{F}^{n-1}$ there is $\mathbf{c} \in C'$ such that $d(\mathbf{c},\mathbf{x}) \leq 2$. Without loss of generality assume $\mathbf{c} = \mathbf{0}$. If $|I(\mathbf{x})| \geq n$, then $\mathbf{x} = \mathbf{0}$. If $|I(\mathbf{x}) \cap B_1(\mathbf{0})| = 2$, then the support of $\mathbf{x}$ is the union of the supports of the codewords in the I-set. If $|I(\mathbf{x}) \cap B_1(\mathbf{0})| = 1$, then either $\mathbf{x} = \mathbf{e}_1$ and $I(\mathbf{x}) \cap B_1(\mathbf{0}) = \{\mathbf{0}\}$ or $\mathbf{x} \in B_2(\mathbf{0})$ and the support of $\mathbf{x}$ is $\{1\} \cup \{\mathrm{supp}(\mathbf{c}_0) \mid \mathbf{c}_0 \in I(\mathbf{x}) \cap B_1(\mathbf{0})\}$. The claim follows from Theorem 5.3. The result holds also for $n = 3$. □

Definition 5.8. Denote by $W_r(n,k,h)$ the minimum number of codewords in any code C of length n whose codewords all have weight h and which has the property that all the sets $I_r(C;\mathbf{x})$, $\mathbf{x} \in S_k(\mathbf{0})$, are nonempty and different.

Trivially, $W_1(n,2,1) = n-1$. Let us look at the values of $W_1(n,3,2)$.

If we denote by $D(C)$ the smallest number of different codewords of C whose sum is the all-zero word, we get the following theorem.

Theorem 5.9. *Let $\bar{C} = S_2(\mathbf{0}) \setminus C$, $C \subseteq S_2(\mathbf{0})$. The sets $I(\bar{C};\mathbf{x})$ are nonempty and distinct for all words $\mathbf{x}$ of weight three if and only if $D(C) \geq 5$.*

Proof. Clearly, there is a word $\mathbf{x}$ of weight three for which $I(\bar{C};\mathbf{x}) = \emptyset$ if and only if C contains three words whose sum is the all-zero word. Any word $\mathbf{x} \in S_3(\mathbf{0})$ for which $|I(\bar{C};\mathbf{x})| \geq 2$ is obviously uniquely identified. We only need to check whether there are two words of weight three, say $\mathbf{x}_1$ and $\mathbf{x}_2$ with supports $\{i,j,k\}$ and $\{i,j,m\}$, such that $I(\bar{C},\mathbf{x}_1) = \{\mathbf{c}\} = I(\bar{C},\mathbf{x}_2)$ where the support of $\mathbf{c}$ equals $\{i,j\}$. This happens if and only if the words with the supports $\{i,k\}$, $\{j,k\}$, $\{i,m\}$ and $\{j,m\}$ are all in C, i.e., if and only if some four codewords of C add up to the all-zero word. □

According to the theorem above, determining the values of $W_1(n,3,2)$ is equivalent to finding the largest code $C \subseteq \mathbf{F}^n$ whose codewords are all of weight two and $D(C) \geq 5$. Consider an undirected graph whose vertex set is the set of the coordinates $\{1,\ldots n\}$ and an edge is a pair of such coordinates (that is, the support of a word of weight two). Hence, calculating $W_1(n,3,2)$ is equivalent to finding a graph with the maximal number of edges and with the length of the shortest cycle (girth) at least five.

The problem of finding such graphs is well-known and several exact values of $W_1(n,3,2)$ are known (see, e.g., [20, 21, 63] and the references therein). For example, the values of $W_1(n,3,2)$ are 1, 3, 5, 9, 13, 18, 24, 30, 39, 48, 57, 68, 79, 92, 105, 119, 130, 146, 163, 181, 199 for the lengths $n = 3,\ldots 23$, respectively. It also follows that

$$\lim_{n\to\infty} \frac{W_1(n,3,2)}{n^2} = \frac{1}{2}.$$

Example 5.10. The constant weight code $S_2(0)\setminus\{1100000, 0110000, 0011000, 0001100, 0000110, 1000010, 1000001, 0001001\}$ attains the value $W_1(7,3,2) = 13$.

Theorem 5.11. *Suppose $n \geq 7$. If A is a code attaining the value $W_1(n,3,2)$, then every word of weight one is covered by at least three codewords of A.*

Proof. It is easy to see that the code consisting of all the words of length $n \geq 8$ and weight two except the ones with supports $\{1,2\}$, $\{2,3\},\ldots,\{n-1,n\}$, $\{n,1\}$ and $\{1,5\}$ identifies all the words of weight three. From this and Example 5.10 we can deduce that for $n \geq 7$ there are at least $n+1$ words of weight two that do not belong to A.

Suppose to the contrary that there is a word $\mathbf{x}$ of weight one, which is covered by less than three codewords. Let $\{s\}$ be the support of $\mathbf{x}$. Without loss of generality we can assume that $s = 1$. Denote by i the number of codewords of weight two that cover $\mathbf{x}$.

If $i = 0$, then none of the words with supports $\{1, j\}$, for $j = 2,\ldots,n$, is a codeword. There are $n-1$ such words. Hence there is also a word whose support is $\{k,h\}$, $(k,h \geq 2,\ k \neq h)$ and which is not in A. Now the word of weight three with support $\{1,k,h\}$ is not covered at all, a contradiction.

If $i = 1$, then for exactly one j the word with support $\{1, j\}$ is a codeword. Now there must be $n+1-(n-2) = 3$ non-codewords of weight two which begin with zero. If any of them does not contain j in its support, then we are done as in the previous case. Without loss of generality, words with supports $\{j,k_1\}$ and $\{j,k_2\}$ are non-codewords, for some k_1 and k_2, $k_1 \neq k_2$. But now the words of weight three with supports $\{1, j,k_1\}$ and $\{1, j,k_2\}$ cannot be distinguished, because both of them are only covered by the word with support $\{1, j\}$.

If $i = 2$, then for some j_1 and j_2 words with supports $\{1, j_1\}$ and $\{1, j_2\}$ are codewords. Now there are at least four non-codewords of weight two that begin with zero. As in the previous case either j_1 or j_2 must occur in the support of each of them. This means, without loss of generality, that there are non-codewords with supports $\{j_1,k_1\}$ and $\{j_1,k_2\}$, $k_1 \neq j_2$, $k_2 \neq j_2$, $k_1 \neq k_2$. We get a contradiction as in the previous case. □

Theorem 5.12. *For* $n \geq 7$,

$$M^{SID}(n) \leq (W_1(n,3,2)+n-1)K(n,3).$$

Proof. Let A be a set which attains the value $W_1(n,3,2)$, B realizing the value $W_1(n,2,1) = n-1$, and D a code of length n and covering radius three. We show that $(A \cup B)+D$ is strongly identifying code.

Let $\mathbf{x} \in \mathbf{F}^n$ and denote $H(\mathbf{c}) = \{\mathbf{c}+\mathbf{y} \mid \mathbf{y} \in A \cup B\}$ for $\mathbf{c} \in D$. Then $d(\mathbf{x},\mathbf{c}) \leq 3$ if and only if $H(\mathbf{c}) \cap I(\mathbf{x}) \neq \emptyset$ (and if and only if $H(\mathbf{c}) \cap I'(\mathbf{x}) \neq \emptyset$). Therefore, using $I(\mathbf{x})$ (or $I'(\mathbf{x})$) we can find a codeword $\mathbf{c} \in D$ such that $d(\mathbf{x},\mathbf{c}) \leq 3$. Without loss of generality, assume that $\mathbf{c} = \mathbf{0}$.

If $\mathbf{x} = \mathbf{0}$, then we immediately know it, because $I(\mathbf{x})$ and $I'(\mathbf{x})$ both contain at least $n-1 \geq 3$ words of weight one, which uniquely identify $\mathbf{x}$. Assume that $\mathbf{x} \neq \mathbf{0}$. If $I(\mathbf{x})$ (or $I'(\mathbf{x})$) contains $\mathbf{0}$, then we know that $w(\mathbf{x}) = 1$, and by Theorem 5.11, at least three of the words in A cover $\mathbf{x}$, and therefore uniquely identify it. We can now assume that we know $w(\mathbf{x}) \geq 2$. Then $w(\mathbf{x}) = 2$ if and only if $I(\mathbf{x})$ (or $I'(\mathbf{x})$) contains at least one word of weight one. When it is known that $w(\mathbf{x}) = 2$, the words in B uniquely identify $\mathbf{x}$. When we know that $w(\mathbf{x}) = 3$, then words of A uniquely identify $\mathbf{x}$. □

5.1.2 Nonexistence results

Denote by N_i the number of codewords of weight i in a code C.

Lemma 5.13. *Let C be a strongly 1-identifying code of length $n \geq 3$. If $\mathbf{0} \notin C$, then at least $\lceil 2n/3 \rceil$ codewords of weight two are needed to identify all the words of weight one. If $\mathbf{0} \in C$, then we need at least $\lceil 2(n-1)/3 \rceil$ codewords of weight two.*

Proof. Assume first that $\mathbf{0} \notin C$. If s denotes the number of words $\mathbf{x} \in S_1(\mathbf{0})$ such that $|I'(\mathbf{x})| = 1$, then

$$s + 2(n-s) \leq \sum_{\mathbf{x} \in S_1(\mathbf{0})} |I'(\mathbf{x})| = 2N_2.$$

Since $s \leq N_2$, we get $N_2 \geq \lceil 2n/3 \rceil$.

Assume then that $\mathbf{0} \in C$. Then $I'(\mathbf{x}) = \{\mathbf{0}\}$ for at most one $\mathbf{x} \in S_1(\mathbf{0})$. Considering the sets $I'(\mathbf{x}) \setminus \{\mathbf{0}\}$, we similarly get

$$N_2 + 2(n - N_2 - 1) \leq \sum_{\mathbf{x} \in S_1(\mathbf{0})} |I'(\mathbf{x}) \setminus \{\mathbf{0}\}| = 2N_2,$$

and the second claim follows. □

Theorem 5.14.

$$M^{SID}(n) \geq \left\lceil \frac{2^n \cdot \lceil 2n/3 \rceil}{\binom{n}{2} + \lceil 2n/3 \rceil - \lceil 2(n-1)/3 \rceil} \right\rceil .$$

Proof. Assume that C is a code with $M^{SID}(n)$ codewords. Applying Lemma 5.13 to all the words of $\mathbf{F}^n$ we get

$$(2^n - M^{SID}(n))\lceil 2n/3 \rceil + M^{SID}(n)\lceil 2(n-1)/3 \rceil \leq M^{SID}(n)\binom{n}{2},$$

from which the theorem follows. □

Lemma 5.15. *Let $n \geq 4$. If $C \subseteq \mathbf{F}^n$ is an optimal strongly 1-identifying code, then for all $\mathbf{x} \in \mathbf{F}^n$ we have $|I'(\mathbf{x})| \leq n-1$.*

Proof. Suppose to the contrary that for some $\mathbf{x}$ we have $|I'(C;\mathbf{x})| = n$. Without loss of generality we can assume $\mathbf{x} = \mathbf{0}$. We will show that $C \setminus \{\mathbf{y}\}$ is a strongly identifying code, we choose $\mathbf{y}$ in the following way: we take $\mathbf{y} \in S_1(\mathbf{0})$ such that $|I(C;\mathbf{y}) \cap S_2(\mathbf{0})| = 1$, if such a word exists; otherwise we take any $\mathbf{y} \in S_1(\mathbf{0})$. It suffices to show that

$$I'(\mathbf{v}) \neq I(\mathbf{u}) \neq I(\mathbf{v}) \text{ and } I'(\mathbf{v}) \neq I'(\mathbf{u}) \neq I(\mathbf{v}) \tag{5.2}$$

for all $\mathbf{u} \in B_1(\mathbf{y})$ and $\mathbf{v} \in \mathbf{F}^n$. Here and from now on the notations I and I' all refer to the code $C \setminus \{\mathbf{y}\}$. Since $|I'(\mathbf{0})| = n-1 \geq 3$ we may always exclude the cases $\mathbf{u} = \mathbf{0}$ or $\mathbf{v} = \mathbf{0}$.

If also $\mathbf{v} \in B_1(\mathbf{y}) \setminus \{\mathbf{0}\}$ and $\mathbf{v} \neq \mathbf{u}$, then either $\mathbf{u} \neq \mathbf{y}$ or $\mathbf{v} \neq \mathbf{y}$; say $\mathbf{u} \neq \mathbf{y}$. Now $\mathbf{u} + \mathbf{y} \in I'(\mathbf{u}) \cap S_1(\mathbf{0})$ but $\mathbf{u} + \mathbf{y} \notin I(\mathbf{v})$. This implies (5.2) in this case.

Suppose then that $\mathbf{v} \notin B_1(\mathbf{y})$. Let first $w(\mathbf{v}) \geq 3$. Since C is strongly identifying code, we get (5.2) for $\mathbf{u} = \mathbf{y}$. If $\mathbf{u} \neq \mathbf{y}$, $I'(\mathbf{u}) \cap S_1(\mathbf{0}) \neq \emptyset$ and $I(\mathbf{v}) \cap S_1(\mathbf{0}) = \emptyset$. This gives (5.2) for these $\mathbf{v}$. Let then $w(\mathbf{v}) = 2$. Now there clearly is a codeword $\mathbf{a} \in I'(\mathbf{v}) \cap S_1(\mathbf{0})$ such that $\mathbf{a} \notin I(\mathbf{u})$ (because $\mathbf{u} \neq \mathbf{0}$). Let finally $w(\mathbf{v}) = 1$. Since $I'(C;\mathbf{v}) \neq I'(C;\mathbf{y})$ we may assume that $w(\mathbf{u}) = 2$. If $\mathbf{0} \in C$, then $\mathbf{0} \in I'(\mathbf{v})$ but $\mathbf{0} \notin I(\mathbf{u})$. If $\mathbf{0} \notin C$, then the choice of $\mathbf{y}$ guarantees that there is a word of weight two in $I'(\mathbf{v})$ which does not belong to $I(\mathbf{u})$. Therefore, we have (5.2) for all words $\mathbf{u}$ and $\mathbf{v}$. □

In the proof of the next lower bound (cf. [40, Theorem 3] and [6, Theorem 9]) we use the concept of *excess*, see for example [13]: Assume that $C \subseteq \mathbf{F}^n$ has covering radius one. If a vector $\mathbf{x} \in \mathbf{F}^n$ is 1-covered by exactly $i+1$ codewords of C, then we say that the excess $E(\mathbf{x})$ on $\mathbf{x}$ is i. In general, the excess $E(V)$ on a subset $V \subseteq \mathbf{F}^n$ is defined by $E(V) = \sum_{\mathbf{x} \in V} E(\mathbf{x})$.

Theorem 5.16.

$$M^{SID}(6) \geq 18,\ M^{SID}(7) \geq 32,\ M^{SID}(8) \geq 57.$$

For $n \geq 9$,

$$M^{SID}(n) \geq \left\lceil \frac{2^{n+1}(n^2-2n+4)}{n^3-n^2+2n+8} \right\rceil.$$

Proof. Let C be a code realizing $M^{SID}(n)$ with $n \geq 6$. The number of points for which $|I(\mathbf{x})| = 1$ is at most $M^{SID}(n)$. The points $\mathbf{x}$ with $|I(\mathbf{x})| = 2$ are called *sons* and the ones with $|I(\mathbf{x})| > 2$ are called *fathers*. If $|I(\mathbf{x})| = 2$, then there is a unique point $\mathbf{y}$ such that $I(\mathbf{x}) \subset I(\mathbf{y})$ and it is called the father of $\mathbf{x}$. A *family* consists of a father and its sons. The space is partitioned by the families and the points with $|I(\mathbf{x})| = 1$.

Assume that $\mathbf{f}$ is a father and denote by $S = S(\mathbf{f})$ the set of sons of $\mathbf{f}$. Let $|I(\mathbf{f})| = i \geq 3$. We shall examine the average excess of a family, i.e., the function

$$\frac{|S|+i-1}{|S|+1} =: g(i,|S|).$$

Let us bound above the number of sons of a father and thus bound below the average excess of a family. Without loss of generality we can assume that the father $\mathbf{f}$ is the all-zero word. Clearly, a father may have at most $\binom{i}{2}$ sons, but as we shall see, we can often say more.

Assume first that $\mathbf{f} \notin C$. By the previous lemma we know that $i \leq n-1$. Suppose first that $i = n-1$.

Denote by $\mathbf{x}$ the unique word in $S_1(\mathbf{f}) \setminus C$. We show that for all $\mathbf{c} \in I(\mathbf{f})$ all the $n-2$ words in $B_1(\mathbf{c}) \setminus \{\mathbf{f}, \mathbf{c}, \mathbf{x}+\mathbf{c}\}$ cannot be sons of $\mathbf{f}$. Indeed, there must be a codeword $\mathbf{c}' \neq \mathbf{c}$ in $I(\mathbf{c})$, otherwise $I'(\mathbf{c}) = \emptyset$, and if $\mathbf{c}' \neq \mathbf{x}+\mathbf{c}$ we are done, so assume $\mathbf{c}' = \mathbf{x}+\mathbf{c}$ is the unique codeword in $I'(\mathbf{c})$. There has to be a codeword $\mathbf{c}''$ of weight three in $I(\mathbf{c}')$: otherwise $I(\mathbf{c}') = I(\mathbf{c})$. We have $\mathbf{c}'' = \mathbf{c}+\mathbf{x}+\mathbf{z}$ for some $\mathbf{z} \in I(\mathbf{f})$. But then $\mathbf{c}+\mathbf{z}$ cannot be a son, since $|I(\mathbf{c}+\mathbf{x})| \geq 3$. Consequently, $\mathbf{c} \in I(\mathbf{f})$ can have at most $n-3$ sons of $\mathbf{f}$ at distance one from it. Counting in two ways the pairs $(\mathbf{c}, \mathbf{s})$ where $\mathbf{c} \in I(\mathbf{f})$, $\mathbf{s} \in S$ and $d(\mathbf{c}, \mathbf{s}) = 1$ we obtain $2|S| \leq (n-3)|I(\mathbf{f})|$ which implies $|S| \leq \lfloor (n-3)(n-1)/2 \rfloor =: U_1$.

Consider the case $\mathbf{f} \notin C$ and $i = n-2$. Denote by $\mathbf{x}_1$ and $\mathbf{x}_2$ the two words in $S_1(\mathbf{f}) \setminus C$. Let $\mathbf{c} \in I(\mathbf{f})$. Because C is strongly identifying code, $I'(\mathbf{x}_1+\mathbf{c}) \neq I'(\mathbf{x}_2+\mathbf{c})$, and therefore one of these contains a codeword of weight three, which is not contained in the other. Without loss of generality, $\mathbf{x}_1+\mathbf{c}+\mathbf{z} \in C$ for some $\mathbf{z} \in I(\mathbf{f})$, $\mathbf{z} \neq \mathbf{c}$. Then $\mathbf{c}, \mathbf{z}, \mathbf{x}_1+\mathbf{c}+\mathbf{z} \in I(\mathbf{c}+\mathbf{z})$ and hence $\mathbf{c}+\mathbf{z}$ is not a son. Thus, there cannot be $n-3$ sons in $S_1(\mathbf{c})$. Counting as above, we get $|S| \leq \lfloor (n-2)(n-4)/2 \rfloor =: U_2$.

Notice that for values $i = 4, \ldots n$ the function $g(i, \binom{i}{2})$ is decreasing, and $g(3, \binom{3}{2}) = g(6, \binom{6}{2})$. Hence for $i = 3, \ldots, n-3$ we may bound $g(i, \binom{i}{2})$ below by $g_1(n) := g(n-3, \binom{n-3}{2})$, when $n \geq 9$.

Assume now that $\mathbf{f} \in C$ and $|I(\mathbf{f})| = i$. In this case S consists of the sons at distance one and two from $\mathbf{f}$. Since a son is covered by exactly two codewords, there can be only one son at distance one from $\mathbf{f}$. Indeed, if $\mathbf{s}_1, \mathbf{s}_2 \in S$ and $d(\mathbf{f}, \mathbf{s}_1) = d(\mathbf{f}, \mathbf{s}_2) = 1$, then $I'(\mathbf{s}_1) = I'(\mathbf{s}_2) = \{\mathbf{f}\}$. Consequently, $|S| \leq \binom{i-1}{2} + 1$, because there are at most $\binom{i-1}{2}$ sons of weight two.

Notice that $g(i, \binom{i-1}{2} + 1)$ is decreasing on i ($i = 4, \ldots, n$) and we may bound it below by $g_2(n) := g(n, \binom{n-1}{2} + 1)$ for $n \geq 6$.

The minimum of $g(3, \binom{3}{2}), g(4, \binom{4}{2}), g(5, \binom{5}{2})$ and $g_2(3)$ is $g(3, \binom{3}{2}) = 5/4$. So to find the minimum of the above mentioned lower estimates on $g(i, |S|)$ we need to compare the functions $g_1(n), g(n-1, U_1), g(n-2, U_2), g_2(n)$, and $5/4$. When $6 \leq n \leq 8$ the minimum is $5/4$ and for $n \geq 9$ the minimum is $g_2(n)$. Denote by $M(n)$ the minimum.

Consequently, the average excess of a family is at least $M(n)$ and hence for every family $\mathscr{F}$ the excess of it $E(\mathscr{F}) \geq M(n)|\mathscr{F}|$. Since the excess on $\mathbf{F}^n$ by C is $(n+1)M^{SID}(n) - 2^n$ we get

$$(n+1)M^{SID}(n) - 2^n \geq (2^n - M^{SID}(n))M(n).$$

Routine calculations give the claim. □

There is (see, [64, Construction 4.24]) infinite sequences of codes $(C_i)_{i=1}^{\infty}$ of length $n_i \longrightarrow \infty$ and covering radius two such that

$$\lim_{i \to \infty} \frac{K(n_i, 2)V(n, 2)}{2^{n_i}} = 1.$$

Such families are known to exist for the lengths $n_i = 2^i + 5 \cdot 2^{i/2-2} - 2$ where $i \geq 4$ is even and $n_i = 23 \cdot 2^{i/2-4} - 3$, where $i \geq 10$ is even. Combining this result with Theorem 5.7 and Theorem 5.16, we have two infinite sequences of strongly identifying codes such that

$$\frac{2^{n+1}}{n}(1 + f(n)) \leq M^{SID}(n) \leq \frac{2^{n+1}}{n}(1 + g(n)),$$

where $f(n)$ and $g(n)$ tend to 0, when n tends to infinity.

Using the results of [40] we also see that we have an infinite sequence of lengths such that asymptotically the ratio between the smallest cardinalities of identifying codes and strongly identifying codes tends to one.

5.1.3 Short codes

In this section we prove sharp results for the cardinalities of strongly 1-identifying codes for short lengths.

Theorem 5.17. $M^{SID}(3) = 6, M^{SID}(4) = 8.$

Proof. The lower bound on $M^{SID}(3)$ follows from Theorem 5.14 and the upper bound from Theorem 5.3.

The upper bound in the case $n = 4$ comes from Theorem 5.2. Theorem 5.14 gives that $M^{SID}(4) \geq 7$. Assume $M^{SID}(4) = 7$ and let C be a code attaining the value $M^{SID}(4)$. We can assume that neither 0000 nor 1111 is a codeword. Namely, if every word-complement pair contains at least one codeword, there would be at least eight codewords.

By Lemma 5.13 we need at least $\lceil 2 \cdot 4/3 \rceil = 3$ codewords of weight two to identify all words of weight one. Assume that $N_2 = 3$. These codewords can be chosen in two ways. Either there are two words of weight one that are covered by two codewords, or there is a word which is covered by three codewords; in both cases all the other words of weight one are covered by one codeword each. In the first case, if $\mathbf{x}$ is a word of weight one such that $|I'(\mathbf{x})| = 2$, then there is a word $\mathbf{y}$ of weight three such that $I'(\mathbf{y}) = I'(\mathbf{x})$, so we need at least four codewords of weight two. In the second case if $\mathbf{x}$ is the word for which $|I'(\mathbf{x})| = 3$ then for the complement $\overline{\mathbf{x}}$ of $\mathbf{x}$ we have $I'(\overline{\mathbf{x}}) = \emptyset$, so again $N_2 \geq 4$.

Because $I'(0000) \neq \emptyset$ and $I'(1111) \neq \emptyset$, we have $N_1 \geq 1$ and $N_3 \geq 1$. Because $I'(\mathbf{x}) \neq \emptyset$ for all $\mathbf{x}$ of weight two, there is a codeword $\mathbf{c}$ of weight one such that its complement is also a codeword, or $N_1 + N_3 \geq 4$. Without loss of generality we can assume that 1000 and 0111 are codewords. Because $I'(1100)$, $I'(1010)$ and $I'(1001)$ must all be different, we need at least two more codewords of weight one or three. Again $N_1 + N_3 \geq 4$. □

Theorem 5.18. $M^{SID}(5) = 14$.

Proof. A strongly 1-identifying code of length 5 and cardinality 14 is {10000, 01000, 00100, 11000, 01100, 01010, 00011, 11010, 10101, 01101, 01011, 11101, 11011, 10111}.

Let C be a strongly 1-identifying code of length five. To prove the lower bound we can assume that neither 00000 nor 11111 is a codeword. Namely, if every word-complement pair contained at least one codeword, then there would be at least 16 codewords. We will prove that $N_1 + N_3 \geq 7$. By symmetry (considering the code $\{\mathbf{1} + \mathbf{c} \mid \mathbf{c} \in C\}$) we then know that also $N_2 + N_4 \geq 7$, and the claim follows. Because $I'(00000)$ is not empty, $N_1 \geq 1$.

Case 1: $N_1 = 1$. We can assume that $10000 \in C$. Denote $A = \{11000, 10100, 10010, 10001\}$. The sets $I'(\mathbf{x})$, $\mathbf{x} \in A$ are different and, because $I'(00000) = \{10000\}$, each of the sets $I'(\mathbf{x})$ contains at least one codeword of weight three. Hence, C contains at least three codewords of weight three that begin with 1, say $\mathbf{c}_1$, $\mathbf{c}_2$ and $\mathbf{c}_3$. Each of these three codewords covers one word of weight two that is not in A. But still we have three words of weight two that are not covered by them. If $\mathbf{c}_1$, $\mathbf{c}_2$ and $\mathbf{c}_3$ are at distance one from one word in A, then the remaining three words of weight two that are not covered, say $\mathbf{x}_1$, $\mathbf{x}_2$, and $\mathbf{x}_3$, are all at distance one from one word of weight three. Because $I'(\mathbf{x}_1)$, $I'(\mathbf{x}_2)$, and $I'(\mathbf{x}_3)$

must be nonempty and different, there must be at least three more codewords of weight three, and hence $N_1+N_3 \geq 7$. Suppose therefore that two words from the set A are covered by two of the codewords $\mathbf{c}_1$, $\mathbf{c}_2$, and $\mathbf{c}_3$ and two by one. Without loss of generality we can assume that $\mathbf{c}_1$, $\mathbf{c}_2$, and $\mathbf{c}_3$ are $11100, 10110$, and 10011. Now the words $01010, 01001$, and 00101 are not covered. The only possibility to make $I'(01010)$, $I'(01001)$, and $I'(00101)$ all nonempty and different using only two more codewords of weight three is to choose 01101 and 01011 as codewords. But then $I'(01111) = I'(01001)$, so we need all in all at least six codewords of weight three, and hence $N_1+N_3 \geq 7$.

Case 2: $N_1 = 2$. Without loss of generality assume $10000, 01000 \in C$. Because $I'(00000) \neq I'(11000)$ we can assume $11100 \in C$. Because $I'(10010) \neq I'(10001)$, at least one of the words 11010, 11001, 10110, or 10101 must be in C. The first two and the last two cases are symmetric. In the first case $N_1+N_3 \geq 7$, because $I'(00110)$, $I'(00101)$, and $I'(00011)$ are nonempty and different, and we need three more codewords of weight three. In the last case suppose that 10110 is a codeword. Again $I'(00110)$, $I'(00101)$, and $I'(00011)$ must be nonempty and different, which requires three codewords of weight three, but of course 10110 can be used as one of them. Let us assume that $N_3 = 4$. Then of the remaining two codewords of weight three (at least) one has to cover either 01010 or 01001 but not both. There are now three possibilities to choose them: 1) $00111, 01101$, now $I'(01111) = I'(00101)$, 2) $01101, 01011$, now $I'(11011) = I'(00011)$, 3) 01101, 10011, now $I'(01111) = I'(00101)$. So in each case we need at least one more codeword of weight three, and so $N_1+N_3 \geq 7$.

Case 3: $N_1 \geq 3$. By Lemma 5.13 we know that $N_3 \geq 4$, so $N_1+N_3 \geq 7$. □

5.2 Strongly $(1, \leq 2)$-identifying codes

In this section we consider strongly $(1, \leq 2)$-identifying codes. We begin with a lower bound which turns out to be optimal for infinitely many cases. We continue with constructions. In the first construction we prove that the direct sum of a $(1, \leq 2)$-identifying code and $\mathbf{F}$ is a strongly $(1, \leq 2)$-identifying code. This result combined with the infinite optimal families of $(1, \leq 2)$-identifying codes (Corollary 4.17) gives optimal results for strongly $(1, \leq 2)$-identifying codes. In the second construction we prove that $(\pi(\mathbf{u}), \mathbf{u}, \mathbf{u}+\mathbf{v})$-construction can be applied to strongly $(1, \leq 2)$-identifying codes.

Theorem 5.19. *Let $\ell \geq 2$. Then*

$$M^{(\leq \ell)SID}(n) \geq \left\lceil (2\ell - 1)\frac{2^n}{n} \right\rceil.$$

Proof. Let $C \subseteq \mathbf{F}^n$ be a strongly $(1, \leq \ell)$-identifying code. If $\mathbf{x} \notin C$, then $|I(\mathbf{x})| \geq 2\ell - 1$. Indeed, otherwise if $I(\mathbf{x}) = \{\mathbf{c}_1, \ldots, \mathbf{c}_{2\ell-2}\}$ and $\mathbf{x}_i$ $(i = 1, \ldots, \ell - 1)$ is the unique word different from $\mathbf{x}$ at distance one from both $\mathbf{c}_{2i-1}$ and $\mathbf{c}_{2i}$, we have $I(\mathbf{x}_1, \dots, \mathbf{x}_{\ell-1}) = I(\mathbf{x}_1, \dots, \mathbf{x}_{\ell-1}, \mathbf{x})$, which is impossible. Obviously, fewer than $2\ell - 2$ codewords in $I(\mathbf{x})$ is also impossible.

Assume then that $\mathbf{x} \in C$. Suppose that $I(\mathbf{x}) = \{\mathbf{c}_1, \dots, \mathbf{c}_{2\ell-2}, \mathbf{x}\}$ and define $\mathbf{x}_i$ as above for all $i = 1, \dots, \ell - 1$. Now $I(\mathbf{x}_1, \dots, \mathbf{x}_{\ell-1}) = I(\mathbf{x}_1, \dots, \mathbf{x}_{\ell-1}, \mathbf{x}) \setminus \{\mathbf{x}\}$ which is not allowed and hence $|I(\mathbf{x})| \geq 2\ell$.

Thus, we obtain

$$|C|(n+1) \geq 2\ell|C| + (2\ell - 1)(2^n - |C|)$$

which gives the claim. □

As we shall see, this lower bound can often be attained when $\ell = 2$.

The proof of the theorem implies the next corollary.

Corollary 5.20. *Let $C \subseteq \mathbf{F}^n$ be a strongly $(1, \leq \ell)$-identifying code, then for all $\mathbf{x} \in \mathbf{F}^n$ we have $|I'(\mathbf{x})| \geq 2\ell - 1$.*

The interested reader is referred to [44] for further results on strongly $(1, \leq \ell)$-identifying codes when $\ell \geq 3$.

The following theorem turns out to be very useful when we construct (optimal) strongly $(1, \leq 2)$-identifying codes.

Theorem 5.21. *If $C \subseteq \mathbf{F}^n$ is a $(1, \leq 2)$-identifying code, then $D = C \oplus \mathbf{F} \subseteq \mathbf{F}^{n+1}$ is strongly $(1, \leq 2)$-identifying.*

Proof. The code C is 3-fold 1-covering by Theorem 4.3. The direct sum of C and $\mathbf{F}$ preserves this property. Moreover, every codeword is covered by at least four codewords of D, namely, for all $\mathbf{c} \in C$ and $a \in \mathbf{F}$ we have $I(D; (\mathbf{c}, a)) = \{(\mathbf{c}', a) \mid d(\mathbf{c}, \mathbf{c}') \leq 1, \mathbf{c}' \in C\} \cup \{(\mathbf{c}, a+1)\}$. Thus, for all $\mathbf{x} \in \mathbf{F}^{n+1}$ we have $|I'(\mathbf{x})| \geq 3$ and applying Lemma 4.4 we know that in $\mathbf{F}^{n+1}$ single words are distinguishable from each other and pairs of words also in the strong sense.

We still need to prove that for all $\{\mathbf{x}_1, \mathbf{x}_2\} \neq \{\mathbf{y}_1, \mathbf{y}_2\}$, $\mathbf{x}_i, \mathbf{y}_i \in \mathbf{F}^{n+1}$, $\mathbf{x}_1 \neq \mathbf{x}_2$ and $\mathbf{y}_1 \neq \mathbf{y}_2$, we have

$$J(\mathbf{x}_1, \mathbf{x}_2) \neq J(\mathbf{y}_1, \mathbf{y}_2),$$

where $I(D; Z) \setminus \{Z\} \subseteq J(Z) \subseteq I(D; Z)$. Denote $\mathbf{x}_i = (\mathbf{x}_i^*, x_i')$ and $\mathbf{y}_i = (\mathbf{y}_i^*, y_i')$, where $\mathbf{x}_i^*, \mathbf{y}_i^* \in \mathbf{F}^n$ and $x_i', y_i' \in \mathbf{F}$. From now on in this proof the I-set notation relates to the code C.

i) Assume first $\{\mathbf{x}_1^*, \mathbf{x}_2^*\} \neq \{\mathbf{y}_1^*, \mathbf{y}_2^*\}$. Then there is $\mathbf{c}^* \in I(\mathbf{x}_1^*, \mathbf{x}_2^*) \triangle I(\mathbf{y}_1^*, \mathbf{y}_2^*)$. Without loss of generality we may assume that $\mathbf{c}^* \in I(\mathbf{x}_1^*) \setminus I(\mathbf{y}_1^*, \mathbf{y}_2^*)$.

- If $\mathbf{c}^* \notin \{\mathbf{x}_1^*, \mathbf{x}_2^*\}$, then $(\mathbf{c}^*, x_1') \in J(\mathbf{x}_1, \mathbf{x}_2) \setminus J(\mathbf{y}_1, \mathbf{y}_2)$.
- If $\mathbf{c}^* = \mathbf{x}_1^*$ and $\mathbf{c}^* \neq \mathbf{x}_2^*$, then $(\mathbf{c}^*, x_1' + 1) \in J(\mathbf{x}_1, \mathbf{x}_2) \setminus J(\mathbf{y}_1, \mathbf{y}_2)$. (Similarly, the case $\mathbf{c}^* = \mathbf{x}_2^*$ and $\mathbf{c}^* \neq \mathbf{x}_1^*$.)

- Suppose $\mathbf{c}^* = \mathbf{x}_1^* = \mathbf{x}_2^*$, hence $\{\mathbf{x}_1,\mathbf{x}_2\} = \{(\mathbf{c}^*,0),(\mathbf{c}^*,1)\}$. Suppose also there is no other codeword than $\mathbf{c}^*$ in $I(\mathbf{c}^*)\setminus I(\mathbf{y}_1^*,\mathbf{y}_2^*)$ (if there is, we get a solution from it). This implies $d(\mathbf{c}^*,\{\mathbf{y}_1^*,\mathbf{y}_2^*\}) = 2$. Without loss of generality we may assume $|I(\mathbf{c}^*)\cap I(\mathbf{y}_1^*)| \geq |I(\mathbf{c}^*)\cap I(\mathbf{y}_2^*)|$.

  If $\mathbf{y}_1^* \neq \mathbf{y}_2^*$, then there is a codeword $\mathbf{c}_1^*$ such that $d(\mathbf{c}_1^*,\mathbf{c}^*) = d(\mathbf{c}_1^*,\mathbf{y}_1^*) = 1$ and $d(\mathbf{c}_1^*,\mathbf{y}_2^*) > 1$. Hence, $(\mathbf{c}_1^*,y_1'+1)\in J(\mathbf{x}_1,\mathbf{x}_2)\setminus J(\mathbf{y}_1,\mathbf{y}_2)$.

  If $\mathbf{y}_1^* = \mathbf{y}_2^*$, then there is $\mathbf{c}_2^* \in (I(\mathbf{y}_1^*)\setminus\{\mathbf{y}_1^*\})\setminus I(\mathbf{c}^*)$ (otherwise we have a square of codewords: for $\mathbf{c}_1^* \in I(\mathbf{c}^*)\cap I(\mathbf{y}_1^*)$ we have $I(\mathbf{c}^*,\mathbf{c}_1^*) = I(\mathbf{y}_1^*,\mathbf{c}_1^*)$ which is impossible). Thus, $(\mathbf{c}_2^*,y_1')\in J(\mathbf{y}_1,\mathbf{y}_2)\setminus J(\mathbf{x}_1,\mathbf{x}_2)$.

ii) Suppose now that $\{\mathbf{x}_1^*,\mathbf{x}_2^*\} = \{\mathbf{y}_1^*,\mathbf{y}_2^*\}$. We must have $\mathbf{x}_1^* \neq \mathbf{x}_2^*$ and $\mathbf{y}_1^* \neq \mathbf{y}_2^*$ otherwise we cannot have $\{\mathbf{x}_1,\mathbf{x}_2\} \neq \{\mathbf{y}_1,\mathbf{y}_2\}$. We may assume $\mathbf{x}_1^* = \mathbf{y}_1^*$ and $\mathbf{x}_2^* = \mathbf{y}_2^*$. Without loss of generality we can assume $x_1' \neq y_1'$, that is $x_1' = y_1'+1$. There is a codeword $\mathbf{c}^* \in I(\mathbf{x}_1^*)\setminus I(\mathbf{x}_2^*)$. If $d(\mathbf{c}^*,\mathbf{x}_1^*) = 1$, then $(\mathbf{c}^*,x_1')\in J(\mathbf{x}_1,\mathbf{x}_2)\setminus J(\mathbf{y}_1,\mathbf{y}_2)$. If $d(\mathbf{x}_1^*,\mathbf{c}^*) = 0$ (and $I(\mathbf{x}_1^*)\setminus I(\mathbf{x}_2^*) = \{\mathbf{x}_1^*\}$), then there is $\mathbf{c}_2^* \in (I(\mathbf{x}_2^*)\setminus\{\mathbf{x}_2^*\})\setminus I(\mathbf{x}_1^*)$: otherwise there is a square of codewords as above. Now $(\mathbf{c}_2^*,x_2')\in J(\mathbf{x}_1,\mathbf{x}_2)$ and $(\mathbf{c}_2^*,x_2')\in J(\mathbf{y}_1,\mathbf{y}_2)$ if and only if $\mathbf{y}_2 = \mathbf{x}_2$ ($y_2' = x_2'$). Suppose $\mathbf{y}_2 = \mathbf{x}_2$. The assumption $d(\mathbf{x}_1^*,\mathbf{c}^*) = 0$ implies that there is a codeword $\mathbf{c}_3^* \in C$ such that $d(\mathbf{c}_3^*,\mathbf{x}_1^*) = d(\mathbf{c}_3^*,\mathbf{x}_2^*) = 1$. Now $(\mathbf{c}_3^*,x_2'+1)\in J(\mathbf{x}_1,\mathbf{x}_2) \bigtriangleup J(\mathbf{y}_1,\mathbf{y}_2)$, since either $x_2'+1 = x_1'$ or $x_2'+1 = y_1'$, but not both. □

Corollary 5.22. $M^{(\leq 2)SID}(n) \leq 2M^{(\leq 2)}(n-1)$.

We are now in a position to give two infinite families of optimal codes.

Corollary 5.23.

$$\begin{aligned} \text{For } k \geq 1: &\qquad M_1^{(\leq 2)SID}(3\cdot 2^k) = 2^{3\cdot 2^k-k}.\\ \text{For } k \geq 3: &\qquad M_1^{(\leq 2)SID}(2^k) = 3\cdot 2^{2^k-k}. \end{aligned}$$

Proof. By Corollary 4.17 we know that $M_1^{(\leq 2)}(n) = 2^{3\cdot 2^k-k-1}$, if $n = 3\cdot 2^k - 1$ ($k \geq 1$), and $M_1^{(\leq 2)}(n) = 3\cdot 2^{2^k-k-1}$, if $n = 2^k-1$ ($k \geq 3$). Combining this with Corollary 5.22 and the lower bound from Theorem 5.19 we obtain the equations. □

No infinite family of optimal strongly 1-identifying codes is known.

Corollary 5.24. *If C is strongly $(1,\leq 2)$-identifying, then the direct sum with* $\mathbf{F}$ *is as well.*

Before presenting one more optimality result, we show a construction which yields a strongly $(1,\leq 2)$-identifying code of length $2n+1$ from a strongly $(1,\leq 2)$-identifying code of length n.

Theorem 5.25. *Let C be a strongly $(1,\le 2)$-identifying code of length n. The code*

$$C' = \{(\pi(\mathbf{u}),\mathbf{u},\mathbf{u}+\mathbf{c}) \mid \mathbf{u}\in \mathbf{F}^n, \mathbf{c}\in C\},$$

where $\pi(\mathbf{u})$ denotes the parity check bit of $\mathbf{u}$, is a strongly $(1,\le 2)$-identifying code of length $2n+1$.

Proof. By Corollary 5.20, we know that for all $\mathbf{x}\in \mathbf{F}^n$, $|I'(\mathbf{x})|\ge 3$. The construction maintains this property by [13, Theorem 3.4.3 and 14.4.3]. Lemma 4.4 now implies that all single words are distinguishable from each other and pairs of words also in the strong sense. Thus, to prove the claim it suffices to show that for all $\{\mathbf{x},\mathbf{y}\},\{\mathbf{z},\mathbf{w}\}\subseteq \mathbf{F}^{2n+1}$, $\{\mathbf{x},\mathbf{y}\}\neq\{\mathbf{z},\mathbf{w}\}$, $\mathbf{x}\neq\mathbf{y}$ and $\mathbf{z}\neq\mathbf{w}$ we have

$$J(\mathbf{x},\mathbf{y})\neq J(\mathbf{z},\mathbf{w}) \tag{5.3}$$

for all the sets J, where $I(D;Z)\setminus\{Z\}\subseteq J(Z)\subseteq I(D;Z)$, for $Z\subseteq \mathbf{F}^{2n+1}$.

Let us divide the words of $\mathbf{F}^{2n+1}$ into two classes according to their first bit and consider the codewords which cover a word in each class. Let $\mathbf{x}=(a,\mathbf{u},\mathbf{u}+\mathbf{v})\in \mathbf{F}^{2n+1}$.

I If $a=\pi(\mathbf{u})$, then $I(\mathbf{x})=\{(\pi(\mathbf{u}),\mathbf{u},\mathbf{u}+\mathbf{c})\mid \mathbf{c}\in C, d(\mathbf{c},\mathbf{v})\le 1\}$.

II If $a\neq\pi(\mathbf{u})$, then $I(\mathbf{x})=A\cup\{(a,\mathbf{u}',\mathbf{u}+\mathbf{v})\mid d(\mathbf{u}',\mathbf{u})=1,\exists\, \mathbf{c}\in C: \mathbf{u}+\mathbf{v}=\mathbf{u}'+\mathbf{c}\}$. Here $A=\{(\pi(\mathbf{u}),\mathbf{u},\mathbf{u}+\mathbf{v})\}$ if $\mathbf{v}\in C$, and $A=\emptyset$ if $\mathbf{v}\notin C$.

Hence, in both classes we are interested in codewords $\mathbf{c}\in C$ such that $d(\mathbf{c},\mathbf{v})\le 1$. Namely, in the class II the properties $d(\mathbf{u}',\mathbf{u})\le 1$ and $\mathbf{u}+\mathbf{v}=\mathbf{u}'+\mathbf{c}$ imply that also $d(\mathbf{v},\mathbf{c})\le 1$. If $I(\mathbf{x})=\{(b_i,\mathbf{s}_i,\mathbf{t}_i)\mid i=1,2,\dots,k\}$, then in both cases $I(C;\mathbf{v})=\{\mathbf{s}_i+\mathbf{t}_i\mid i=1,2,\dots,k\}$. If $\mathbf{x}\in C'$ and it is removed from $I(\mathbf{x})$, then $\mathbf{v}$ is removed from $I(C;\mathbf{v})$. Notice that if $I(\mathbf{x},\mathbf{y})\setminus\{\mathbf{x},\mathbf{y}\}=\{(b_i,\mathbf{s}_i,\mathbf{t}_i)\}\mid i=1,2,\dots,k\}$, where $\mathbf{x}=(a_1,\mathbf{u}_1,\mathbf{u}_1+\mathbf{v}_1)$ and $\mathbf{y}=(a_2,\mathbf{u}_2,\mathbf{u}_2+\mathbf{v}_2)$, then $I(C;\{\mathbf{v}_1,\mathbf{v}_2\})\setminus\{\mathbf{v}_1,\mathbf{v}_2\}\subseteq\{\mathbf{s}_i+\mathbf{t}_i\mid i=1,2,\dots,k\}\subseteq I(C;\{\mathbf{v}_1,\mathbf{v}_2\})$. Similarly, if $I(\mathbf{x},\mathbf{y})\setminus\{\mathbf{x}\}=\{(b_i,\mathbf{s}_i,\mathbf{t}_i)\mid i=1,2,\dots k\}$, then $I(C;\{\mathbf{v}_1,\mathbf{v}_2\})\setminus\{\mathbf{v}_1\}\subseteq\{\mathbf{s}_i+\mathbf{t}_i\mid i=1,2,\dots,k\}\subseteq I(C;\{\mathbf{v}_1,\mathbf{v}_2\})$.

Suppose one of the inequalities (5.3) does not hold. Denote $\mathbf{x}=(a_1,\mathbf{u}_1,\mathbf{u}_1+\mathbf{v}_1)$, $\mathbf{y}=(a_2,\mathbf{u}_2,\mathbf{u}_2+\mathbf{v}_2)$, $\mathbf{z}=(a_3,\mathbf{u}_3,\mathbf{u}_3+\mathbf{v}_3)$ and $\mathbf{w}=(a_4,\mathbf{u}_4,\mathbf{u}_4+\mathbf{v}_4)$. Since C is a strongly $(1,\le 2)$-identifying code we must have $\{\mathbf{v}_1,\mathbf{v}_2\}=\{\mathbf{v}_3,\mathbf{v}_4\}$ by the previous discussion.

By Theorem 4.15 we know that the code C' is $(1,\le 2)$-identifying. Thus, it is enough to consider pairs where there is at least one codeword. Without loss of generality we may assume $\mathbf{x}\in C'$. Thus, $\mathbf{x}$ belongs to a class I. We must have either $|I'(\mathbf{x})\cap I(\mathbf{z})|\ge 2$ or $|I'(\mathbf{x})\cap I(\mathbf{w})|\ge 2$. Without loss of generality suppose $|I'(\mathbf{x})\cap I(\mathbf{z})|\ge 2$. Because the last n bits are changing in the words of $I'(\mathbf{x})$, this must be true also in $I(\mathbf{z})$, thus, also $\mathbf{z}$ belongs to the class I. This also implies that $\mathbf{u}_1=\mathbf{u}_3$.

If $\mathbf{y}$ belongs to the class II, then we must have $|I(\mathbf{y}) \cap I(\mathbf{w})| \geq 2$, since $\mathbf{z}$ from the class I can cover at most one codeword from $I(\mathbf{y})$. Hence, now also $\mathbf{w}$ belongs to the class II. Thus, $\mathbf{y}$ and $\mathbf{w}$ belong to the same class.

Let first all $\mathbf{x}, \mathbf{y}, \mathbf{z}$ and $\mathbf{w}$ lie in the first class. If $I'(\mathbf{x}) \cap I(\mathbf{w}) \neq \emptyset$, then we have $\mathbf{u}_1 = \mathbf{u}_4$. Similarly, if $I(\mathbf{y}) \cap I(\mathbf{z}) \neq \emptyset$, then $\mathbf{u}_2 = \mathbf{u}_3$. Now $\mathbf{u}_1 = \mathbf{u}_2 = \mathbf{u}_3 = \mathbf{u}_4$ and together with $\{\mathbf{v}_1, \mathbf{v}_2\} = \{\mathbf{v}_3, \mathbf{v}_4\}$, we get $\{\mathbf{x}, \mathbf{y}\} = \{\mathbf{z}, \mathbf{w}\}$, what is a contradiction. If $I'(\mathbf{x}) \cap I(\mathbf{w}) = \emptyset$, then $|I'(\mathbf{x}) \cap I(\mathbf{z})| \geq 3$, which implies by Lemma 2.1 that $\mathbf{x} = \mathbf{z}$. Thus, $|I'(\mathbf{y}) \cap I'(\mathbf{w})| \geq 3$ and $\mathbf{y} = \mathbf{w}$.

Assume next that $\mathbf{x}$ and $\mathbf{z}$ belong to the first class and $\mathbf{y}$ and $\mathbf{w}$ lie in the second. Because $\mathbf{y}$ and $\mathbf{w}$ cannot be codewords, it suffices to verify only the following inequalities of (5.3) $I(\mathbf{x}, \mathbf{y}) \setminus \{\mathbf{x}\} \neq I(\mathbf{z}, \mathbf{w}) \setminus \{\mathbf{z}\}$ and $I(\mathbf{x}, \mathbf{y}) \setminus \{\mathbf{x}\} \neq I(\mathbf{z}, \mathbf{w})$.

Suppose $I(\mathbf{x}, \mathbf{y}) \setminus \{\mathbf{x}\} = I(\mathbf{z}, \mathbf{w}) \setminus \{\mathbf{z}\}$. We have $\mathbf{x} \neq \mathbf{z}$, since otherwise $I(\mathbf{x}, \mathbf{y}) = I(\mathbf{z}, \mathbf{w})$, which is impossible due to the fact that C' is $(1, \leq 2)$-identifying. We must have $I'(\mathbf{x}) \cap I'(\mathbf{z}) \neq \emptyset$ and hence $\mathbf{u}_1 = \mathbf{u}_3$. Because $|I'(\mathbf{e})| \geq 3$ for all $\mathbf{e} \in \mathbf{F}^{2n+1}$, we obtain $|I'(\mathbf{x}) \cap I(\mathbf{w})| = 1$ and $|I'(\mathbf{z}) \cap I(\mathbf{y})| = 1$ (the number of elements in the intersections cannot be greater than one, since $\mathbf{x}$ and $\mathbf{w}$ and also $\mathbf{y}$ and $\mathbf{z}$ belong to different classes). The words in these two intersections must be different, since $\mathbf{x} \neq \mathbf{z}$. Since words $\mathbf{y}$ and $\mathbf{w}$ belong to the class II and we must have $I(\mathbf{y}) \cap I(\mathbf{w}) \neq \emptyset$, the last n bits of the words in $I(\mathbf{y})$ and $I(\mathbf{w})$ are the same. On the other hand, the words in the intersections $I'(\mathbf{x}) \cap I(\mathbf{w})$ and $I'(\mathbf{z}) \cap I(\mathbf{y})$ must also end with those n bits, but now $|I'(\mathbf{x}) \cap I'(\mathbf{z})| \geq 3$, since $\mathbf{u}_1 = \mathbf{u}_3$, and thus $\mathbf{x} = \mathbf{z}$, a contradiction.

Assume finally $I(\mathbf{x}, \mathbf{y}) \setminus \{\mathbf{x}\} = I(\mathbf{z}, \mathbf{w})$. Evidently, $\mathbf{x} \in C'$ and $\mathbf{x} \neq \mathbf{z}$. If $\mathbf{z}$ is a codeword, then in $I(\mathbf{z})$ there are at least two codewords which are not in $I'(\mathbf{x})$, these two words cannot be both in $I(\mathbf{y})$ either ($\mathbf{y}$ and $\mathbf{z}$ are in different classes), so $\mathbf{z} \notin C'$. Now $I(\mathbf{x}, \mathbf{y}) \setminus \{\mathbf{x}\} = I(\mathbf{z}, \mathbf{w}) \setminus \{\mathbf{z}\}$, where $\mathbf{z} \notin C'$ but belongs to the class I. This is impossible by the previous case. □

There does not exist a strongly $(1, \leq 2)$-identifying code of length less than five. Indeed, for length four we notice that there cannot be a strongly $(1, \leq 2)$-identifying code, since always $I(0000, 1111) \setminus \{0000, 1111\} = I(0011, 1100) \setminus \{0011, 1100\}$ regardless of the code.

In the next proof, we denote by $S_i^n(\mathbf{x})$ the words at distance i from a word $\mathbf{x}$ in $\mathbf{F}^n$.

Theorem 5.26. $M^{(\leq 2)SID}(5) = 22$.

Proof. By Theorem 4.11 we have $M^{(\leq 2)}(4) = 11$, and thus we obtain the upper bound using Corollary 5.22.

We prove the lower bound by using Corollary 5.20, $|I'(\mathbf{x})| \geq 3$ for all $\mathbf{x} \in \mathbf{F}^5$. Suppose $M^{(\leq 2)SID}(5) \leq 21$. By [13, p. 383] we need at least 22 codewords to cover each word in $\mathbf{F}^5$ at least four times. Hence we know that there is a non-codeword which is covered by exactly three codewords. Let C be a code attaining the bound $M^{(\leq 2)SID}(5)$. Without loss of generality, assume $00000 \notin C$, and

$I(00000) = \{10000, 01000, 00100\}$. Each word of weight one must be covered by at least three codewords of weight two. Each codeword of weight two covers two words of weight one. Thus, we need at least $\lceil 3 \cdot 5/2 \rceil = 8$ codewords of weight two. If $11111 \notin C$ this is a symmetric situation, so all in all there are now at least 22 codewords.

Suppose $11111 \in C$. Now again there are at least three codewords of weight four. The word 00011 must be covered by three codewords of weight three, those can only be $\{10011, 01011, 00111\} =: T$. The words of the set $A := S_1^3(000) \oplus S_1^2(00)$ must be covered by at least two codewords of weight three. Denote by $G_{\mathbf{a}} = \mathbf{a} \oplus S_1^2(00)$, where $\mathbf{a} \in S_1^3(000)$. The words of weight three always cover exactly two words of A, except that 11100 covers none. Moreover, a word in $S_3^5(00000) \setminus (\{11100\} \cup T)$ covers a word both in $G_{\mathbf{a}}$ and $G_{\mathbf{b}}$ for some **a** and **b** ($\mathbf{a} \neq \mathbf{b}$). Let $\mathbf{c}_1 \in S_3^5(00000)$ cover a word in $G_{\mathbf{a}}$ and $G_{\mathbf{b}}$ and, moreover, $\mathbf{c}_2 \in S_3^5(00000)$ cover the other word in $G_{\mathbf{a}}$ and a word in $G_{\mathbf{d}}$, where $\mathbf{b} \neq \mathbf{a} \neq \mathbf{d}$ and $\mathbf{b} \neq \mathbf{d}$. The distance between the yet uncovered (by others than the words in T) words of $G_{\mathbf{b}}$ and $G_{\mathbf{d}}$ is four. Thus, we need at least seven codewords of weight three in C. Again there are at least 22 codewords. □

In Table 5.1 we have collected lower and upper bounds on the cardinalities of strongly 1-identifying and strongly $(1, \leq 2)$-identifying codes.

Table 5.1: Bounds on the cardinalities of strongly 1-identifying codes and strongly $(1, \leq 2)$-identifying codes.

n	$M^{SID}(n)$	$M^{(\leq 2)SID}(n)$
3	a 6 A	-
4	b 8 B	-
5	c 14 C	22 G
6	a, d 18 – 20 D	32 H
7	d 32 – 38 E	55 – 64 I
8	d 57 – 64 A	96 H
9	d 102 – 128 A	171 – 192 I
10	d 186 – 256 A	308 – 352 I
11	d 341 – 488 F	559 – 704 I, J
12	d 629 – 768 A	1024 H
13	d 1169 – 1536 A	1891 – 2048 I, J
14	d 2182 – 2816 A	3511 – 4096 I
15	d 4091 – 5632 A	6554 – 8192 I, J
16	d 7703 – 10240 A	12288 H
17	d 14552 – 20480 A	23131 – 24576 I, J
18	d 27575 – 40960 A	43691 – 49152 I
19	d 52397 – 81920 A	82783 – 98304 I, J
20	d 99813 – 163840 A, D	157287 – 180224 I
21	d 190565 – 327680 A	299594 – 360448 I, J
22	d 364580 – 630784 D	571951 – 720896 I
23	d 698812 – 905216 E	1094167 – 1441792 I, J
24	d 1341774 – 1572864 D	2097152 I

Lower bounds on $M^{(\leq 2)SID}(n)$ for $n \geq 6$ come from Theorem 5.19 and for $n = 5$ the lower bound comes from Theorem 5.26.

	$M^{SID}(n)$		$M^{(\leq 2)SID}(n)$
a	Theorem 5.14	G	Theorem 5.26
b	Theorem 5.17	H	Corollary 5.23
c	Theorem 5.18	I	Theorem 5.21
d	Theorem 5.16	J	Theorem 5.25
A	Theorem 5.3		
B	Theorem 5.2		
C	Theorem 5.18		
D	Theorem 5.7		
E	Theorem 5.12		
F	Theorem 5.5		

Chapter 6

Linear identifying codes

A binary code $C \subseteq \mathbf{F}^n$ is called *linear* if all the pairwise sums of codewords belong to the code. This means that C is a subspace of $\mathbf{F}^n$. There is a linearly independent set of words of $\mathbf{F}^n$ that generates the code C. The number of these codewords, k, is called the *dimension* of the code. Thus, the cardinality of the code is 2^k.

In this chapter we consider linear identifying codes. We solve the smallest dimensions for all possible lengths and for all $\ell \geq 1$ of linear $(1, \leq \ell)$-identifying and linear strongly $(1, \leq \ell)$-identifying codes. The requirement of linearity increases the number of codewords in most cases compared with the case where linearity is not required. However, we will see that for infinitely many lengths linear $(1, \leq \ell)$-identifying codes and linear strongly $(1, \leq \ell)$-identifying codes for $\ell \geq 3$, are as good as the corresponding codes in the case where linearity is not required. This chapter is based on [55].

The smallest dimension of a binary linear $(r, \leq \ell)$-identifying code of length n is denoted by $k_r^{(\leq \ell)}[n]$. A code attaining the smallest dimension is called *optimal*.

The next lemma will often be used.

Lemma 6.1. *If $C \subseteq \mathbf{F}^n$ is a linear code and $\mathbf{c} \in C$, then for all $\mathbf{x} \in \mathbf{F}^n$*

$$I(\mathbf{x}+\mathbf{c}) = I(\mathbf{x}) + \mathbf{c}.$$

Proof. Clearly, $\mathbf{y} \in I(\mathbf{x}+\mathbf{c})$ if and only if $\mathbf{y} \in C$ and $d(\mathbf{x}+\mathbf{c}, \mathbf{y}) \leq 1$. This is equivalent to the fact that $\mathbf{y}+\mathbf{c} \in C$ (because C is linear) and $d(\mathbf{x}, \mathbf{y}+\mathbf{c}) \leq 1$; that is, $\mathbf{y}+\mathbf{c} \in I(\mathbf{x})$. This is equivalent to $\mathbf{y} \in I(\mathbf{x}) + \mathbf{c}$. □

Lemma 6.1 implies that each codeword is covered by the same number of codewords in a linear code. Namely, suppose C is a linear code. Then for all $\mathbf{c} \in C$ we have $I(\mathbf{0}+\mathbf{c}) = I(\mathbf{0}) + \mathbf{c}$. Thus, each codeword is covered by the same number of codewords as $\mathbf{0}$, which is always a codeword. By the lemma, we also know that if $\mathbf{x}, \mathbf{y} \notin C, \mathbf{e}_i \in S_1(\mathbf{0})$ and both $\mathbf{x}+\mathbf{e}_i$ and $\mathbf{y}+\mathbf{e}_i \in C$, then $|I(\mathbf{x})| = |I(\mathbf{y})|$. Indeed,

$I(\mathbf{x})+\mathbf{x}+\mathbf{e}_i=I(\mathbf{x}+\mathbf{x}+\mathbf{e}_i)=I(\mathbf{e}_i)=I(\mathbf{y}+\mathbf{y}+\mathbf{e}_i)=I(\mathbf{y})+\mathbf{y}+\mathbf{e}_i$. In fact, we get more; namely, $|I(\mathbf{x})|=|I(\mathbf{e}_i)|=|I(\mathbf{y})|$.

Let us recall the definitions from [13]. A code C is a *μ-fold* 1*-covering* if the codewords of C cover each word of $\mathbf{F}^n$ at least μ times. The smallest dimension of a linear μ-fold 1-covering is denoted by $k[n,1,\mu]$. A code attaining the smallest dimension is called *optimal*.

We denote by $k[n,1,\mu,\nu]$ the smallest dimension of a binary linear code of length n such that each non-codeword in $\mathbf{F}^n$ is covered by at least μ codewords and each codeword is covered by at least ν codewords. These covering codes are a generalization of μ-fold 1-coverings, and $k[n,1,\mu]=k[n,1,\mu,\mu]$. The next theorem is the sphere-covering bound for such codes.

Theorem 6.2. *Suppose μ and ν are positive integers. Then*

$$k[n,1,\mu,\nu]\geq\lceil n+\log_2\mu-\log_2(n+1+\mu-\nu)\rceil.$$

Proof. Denote $K=2^{k[n,1,\mu,\nu]}$. The claim follows from the direct calculation

$$\nu\cdot K+\mu(2^n-K)\leq K(n+1).$$

□

A linear k-dimensional code $C\subseteq\mathbf{F}^n$ can always be defined by choosing an $(n-k)\times n$ matrix H — called its *parity check matrix* — such that

$$\mathbf{x}\in C \quad \text{if and only if} \quad H\mathbf{x}^T=0.$$

Let $\mathbf{x}\in\mathbf{F}^n$ be arbitrary, and consider its *syndrome* $H\mathbf{x}^T$. If $H\mathbf{x}^T$ appears in H as the i-th column, then clearly the word obtained by changing the i-th bit in $\mathbf{x}$ belongs to C. Consequently, if $\mathbf{x}$ is a codeword, the number of codewords within distance one from it equals one plus the number of zero columns in H; if $\mathbf{x}$ is a non-codeword, the number of codewords within distance one from it equals the number of times the syndrome of $\mathbf{x}$ appears as a column in H.

The proof of the next theorem gives optimal constructions for all possible lengths of linear codes such that all non-codewords are covered at least μ times and all codewords are covered at least ν times.

Theorem 6.3. *Let μ and ν be positive integers. When $n=(2^r-1)\mu+\nu-1+s$, where $r\geq 0$ and $0\leq s\leq 2^r\mu-1$, then*

$$k[n,1,\mu,\nu]=n-r.$$

Proof. The lower bound comes from Theorem 6.2.

If $r=0$, then by Theorem 6.2, $k[n,1,\mu,\nu]\geq n$. In $\mathbf{F}^n$ codewords are covered at least ν times and, thus, it is the optimal code.

Suppose then that $r \geq 1$. Let H be any $(r \times n)$-matrix which contains the all-zero column at least $\nu - 1$ times and every non-zero column at least μ times. Now $n = (2^r - 1)\mu + \nu - 1 + s$, where $0 \leq s \leq 2^r\mu - 1$. The matrix H has clearly rank r, and the code which has H as its parity check matrix is as required, by the previous discussion. □

6.1 On linear 1-identifying codes

In this section, we first prove a lower bound for linear 1-identifying codes. We also prove a construction which increases the length and dimension of a linear 1-identifying code by one and the result is a linear 1-identifying code. The corresponding result is not known for 1-identifying codes, in general. In Theorem 6.6 we combine these results and get the optimal dimensions for all possible lengths of linear 1-identifying codes.

Theorem 6.4. *For $n \geq 2$ we have*

$$k_1^{(\leq 1)}[n] \geq \lceil n + \log_2 3 - \log_2(n+3) \rceil .$$

Proof. Suppose $C \subseteq \mathbf{F}^n$ is a linear 1-identifying code of length n and cardinality K. We will show that each word in $\mathbf{F}^n$ is covered either by one codeword or by at least three codewords. Because C is 1-identifying we have $|I(\mathbf{x})| \geq 1$ for all $\mathbf{x} \in \mathbf{F}^n$, and $|I(\mathbf{x})| = 1$ for at most K words.

Suppose that for some $\mathbf{y} \in \mathbf{F}^n$ we have $|I(\mathbf{y})| \geq 2$. Then $\{\mathbf{y} + \mathbf{e}_1, \mathbf{y} + \mathbf{e}_2\} \subseteq I(\mathbf{y})$ where $\mathbf{e}_1, \mathbf{e}_2 \in B_1(\mathbf{0})$. Now $\{\mathbf{y} + \mathbf{e}_1, \mathbf{y} + \mathbf{e}_2\} \subseteq I(\mathbf{y} + \mathbf{e}_1 + \mathbf{e}_2)$ and by Lemma 6.1 we know that $|I(\mathbf{y})| = |I(\mathbf{y} + \mathbf{e}_1 + \mathbf{e}_2)|$. This implies a contradiction with identifying property unless $|I(\mathbf{y})| \geq 3$. Since $\mathbf{y}$ was chosen arbitrarily we know that for all $\mathbf{y} \in \mathbf{F}^n$ such that $|I(\mathbf{y})| > 1$ we have $|I(\mathbf{y})| \geq 3$. Now we can calculate

$$K + 3(2^n - K) \leq K(n+1)$$

from which the claim follows. □

For linear 1-identifying codes we get the next nice result which is not known for 1-identifying codes in general.

Theorem 6.5.

$$k_1^{(\leq 1)}[n] \leq k_1^{(\leq 1)}[n-1] + 1.$$

Proof. Let D be a linear 1-identifying code of length $n-1$ and dimension $k_1^{(\leq 1)}[n-1]$. If each codeword is covered by more than one codeword then by [6, Theorem 1] (or Theorem 3.20) the direct sum $D \oplus \mathbf{F}$ is a 1-identifying code of length n. Clearly, the code $D \oplus \mathbf{F}$ is linear.

Suppose that in D each codeword is covered only by itself, this is equivalent with the condition that the minimum distance of D is at least two. Then as noticed in the proof of Theorem 6.4 each non-codeword is covered by at least three codewords. Denote by $C(0) = \{00, 11\} \subseteq \mathbf{F}^2$ and $C(1) = \{01, 10\} \subseteq \mathbf{F}^2$. We will show that the code

$$C = \bigcup_{(x_1,x_2,\ldots,x_{n-1}) \in D} (x_1, x_2, \ldots, x_{n-2}) \oplus C(x_{n-1})$$

is a linear 1-identifying code of length n. If G_D is a generator matrix for D, then the matrix

$$G_C = \begin{pmatrix} & & & & 0 \\ & G_D & & & \vdots \\ & & & & 0 \\ 0 & \ldots & 0 & 1 & 1 \end{pmatrix}$$

is a generator matrix for C and, hence, C is linear. Since the minimum distance of D is at least two, we see from the generator matrix G_C that the minimum distance of C is two. Namely, $0\ldots011 \in C$ and, thus, the minimum distance is at most two. If $\mathbf{x} \in D$ and $\mathbf{x} \neq \mathbf{0}$, then $(\mathbf{x}, 0) \in C$ and the weight of $(\mathbf{x}, 0)$ is the same as the weight of $\mathbf{x}$. If the last bit of $\mathbf{x}$ is one, then the codeword $(\mathbf{x}, 0) + 0\ldots011$ has the same weight as $\mathbf{x}$, otherwise, the weight is one greater. Thus, each codeword of C is covered only by itself.

Suppose $\mathbf{x} = (\mathbf{x}', \mathbf{a}) \in \mathbf{F}^n$, where $\mathbf{x}' \in \mathbf{F}^{n-2}$ and $\mathbf{a} \in \mathbf{F}^2$, is a non-codeword. Assume that $\mathbf{a} \in C(\mathbf{z})$, where $\mathbf{z} \in \mathbf{F}$. Now $(\mathbf{x}', \mathbf{z}) \notin D$ and we know that $|I(D; (\mathbf{x}', \mathbf{z}))| \geq 3$. In particular,

$$I(D; (\mathbf{x}', \mathbf{z})) = \{(\mathbf{c}, \mathbf{z}) \in D \mid d(\mathbf{x}', \mathbf{c}) \leq 1\} \cup A,$$

where $A = \{(\mathbf{x}', \mathbf{z} + 1)\}$ if $(\mathbf{x}', \mathbf{z} + 1) \in D$ and otherwise A is empty. Now

$$I(C; \mathbf{x}) = \{(\mathbf{c}, \mathbf{a}) \in \mathbf{F}^n \mid d(\mathbf{x}', \mathbf{c}) \leq 1, (\mathbf{c}, \mathbf{z}) \in D\} \cup B,$$

where $B = \{(\mathbf{x}', \mathbf{b}) \in \mathbf{F}^n \mid \mathbf{b} \in C(\mathbf{z} + 1)\}$ if $(\mathbf{x}', \mathbf{z} + 1) \in D$ and otherwise B is empty. This means that also in C each non-codeword is covered by at least three codewords. Since in the intersection of three Hamming balls of radius one there is at most one word, we know that each non-codeword is distinguishable. Hence, each word in $\mathbf{F}^n$ has a unique I-set. Thus, C is a linear 1-identifying code of length n and dimension $k_1^{(\leq 1)}[n - \ 1] + 1$. □

Theorem 6.6. *Let* $n = 3(2^r - 1) + s$ *for* $r \geq 1$ *and* $0 \leq s \leq 3 \cdot 2^r - 1$. *Then*

$$k_1^{(\leq 1)}[n] = n - r.$$

Proof. Let $n = 3(2^r - 1)$. By Theorem 6.3 there is a linear code such that each non-codeword is covered by exactly three codewords and each codeword is covered only by itself. Thus, this code is 1-identifying. The dimension of the code is

$n-r$. The upper bounds for $1 \leq s \leq 3\cdot 2^r-1$ follow from Theorem 6.5. The lower bounds follow from Theorem 6.4. □

At least for some lengths there are also other optimal linear 1-identifying codes than what we constructed in the proofs of Theorem 6.6 and Theorem 6.5. Let $n=2^r-1=3(2^{r-2}-1)+2^{r-2}+2$ for $r\geq 3$ and denote by $\mathscr{H}$ the Hamming code of length n. Then the code

$$C=\mathscr{H}\cup(\mathscr{H}+10^{n-1})\cup(\mathscr{H}+010^{n-2})\cup(\mathscr{H}+110^{n-2})$$

is a linear code and each word in $\mathbf{F}^n$ is covered by four codewords, and, thus, it is 1-identifying. The code C is optimal and by repeatedly taking the direct sum with $\mathbf{F}$ we get optimal codes up to the length $3(2^{r-1}-1)-1$. These codes and the codes of Theorem 6.6 are not equivalent, since in the codes of Theorem 6.6 all codewords are covered only by themselves.

6.2 On linear $(1,\leq \ell)$-identifying codes

Theorem 6.7 solves the dimensions of $(1,\leq 2)$-identifying codes for all possible lengths. The optimal results for linear $(1,\leq \ell)$-identifying codes, when $\ell\geq 3$, follow from the results in [45].

Theorem 6.7. *A linear code $C\subseteq \mathbf{F}^n$ is $(1,\leq 2)$-identifying if and only if it is a 5-fold 1-covering. In particular, if $n=5\cdot 2^r-1+s$, where $r\geq 0$ and $0\leq s\leq 5\cdot 2^r-1$, we have*

$$k_1^{(\leq 2)}[n]=k[n,1,5]=n-r.$$

Proof. Suppose first that $|I(C;\mathbf{x})|\geq 5$ for all $\mathbf{x}\in\mathbf{F}^n$. By Lemma 2.1 no set of size at most two where $\mathbf{x}$ does not belong to can cover the whole I-set of $\mathbf{x}$. Thus, C is $(1,\leq 2)$-identifying.

Let C be a linear $(1,\leq 2)$-identifying code. By Theorem 4.3 we know that $|I(\mathbf{x})|\geq 3$ for all $\mathbf{x}\in\mathbf{F}^n$. Let us first consider codewords. Suppose $\mathbf{c}\in C$. Since C is linear, $\mathbf{0}\in C$. Suppose first that $|I(\mathbf{c})|=3$. Then by Lemma 6.1, $I(\mathbf{0})=\{\mathbf{0},\mathbf{c}_1,\mathbf{c}_2\}$ for some codewords $\mathbf{c}_1,\mathbf{c}_2\in S_1(\mathbf{0})$. By linearity, $\mathbf{c}_1+\mathbf{c}_2\in C$. Again by Lemma 6.1, $I(\mathbf{c}_1)=\{\mathbf{0},\mathbf{c}_1,\mathbf{c}_1+\mathbf{c}_2\}$ and $I(\mathbf{c}_1+\mathbf{c}_2)=\{\mathbf{c}_1,\mathbf{c}_2,\mathbf{c}_1+\mathbf{c}_2\}$, and thus $I(\mathbf{0},\mathbf{c}_1)=I(\mathbf{0},\mathbf{c}_1+\mathbf{c}_2)=\{\mathbf{0},\mathbf{c}_1,\mathbf{c}_2,\mathbf{c}_1+\mathbf{c}_2\}$, which is impossible. Hence $|I(\mathbf{c})|\geq 4$ for all $\mathbf{c}\in C$.

If $|I(\mathbf{c})|=4$, then $I(\mathbf{0})=\{\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{0}\}$ where the supports of $\mathbf{x},\mathbf{y}$, and $\mathbf{z}$ are $\{i\}$, $\{j\}$, and $\{k\}$, respectively. By linearity, there are codewords $\mathbf{c}_1,\mathbf{c}_2,\mathbf{c}_3$ and $\mathbf{c}_4$ with supports $\{i,j\}$, $\{i,k\}$, $\{j,k\}$ and $\{i,j,k\}$, respectively. But now $I(\mathbf{0},\mathbf{c}_4)=I(\mathbf{y},\mathbf{c}_2)=\{\mathbf{0},\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{c}_1,\mathbf{c}_2,\mathbf{c}_3,\mathbf{c}_4\}$. Thus, each codeword is covered by at least five codewords.

We claim that also each non-codeword is covered at least by five codewords. First, we show that $|I(\mathbf{x})|\geq 4$ for all $\mathbf{x}\notin C$. Assume on the contrary that $\mathbf{x}\notin C$

and $|I(\mathbf{x})| = 3$. Without loss of generality (by Lemma 6.1) we can assume that $\mathbf{x} \in S_1(\mathbf{0})$. There are two non-codewords $\mathbf{y}$ and $\mathbf{z}$ in $S_1(\mathbf{0})$ such that $\mathbf{x}+\mathbf{y}$ and $\mathbf{x}+\mathbf{z}$ belong to the code, and by linearity also $\mathbf{y}+\mathbf{z} \in C$. By Lemma 6.1, $I(\mathbf{x},\mathbf{y}) = \{\mathbf{0},\mathbf{x}+\mathbf{y},\mathbf{x}+\mathbf{z},\mathbf{y}+\mathbf{z}\} = I(\mathbf{x},\mathbf{z})$, which is a contradiction. Hence, $|I(\mathbf{x})| \geq 4$ for all $\mathbf{x} \notin C$.

Assume now that $\mathbf{x} \notin C$ and $I(\mathbf{x}) = \{\mathbf{x}+\mathbf{e}_1,\mathbf{x}+\mathbf{e}_2,\mathbf{x}+\mathbf{e}_3,\mathbf{x}+\mathbf{e}_4\}$ where $\mathbf{e}_i \in S_1(\mathbf{0})$, for $i = 1,\dots,4$ are different non-codewords. Denote

$$\begin{aligned}\mathbf{y} &= \mathbf{x}+\mathbf{e}_1+\mathbf{e}_2\\ \mathbf{z} &= \mathbf{x}+\mathbf{e}_3+\mathbf{e}_4\\ \mathbf{v} &= \mathbf{y}+\mathbf{e}_3+\mathbf{e}_4.\end{aligned}$$

Using Lemma 6.1 we get

$$\begin{aligned}I(\mathbf{y}) &= \{\mathbf{y}+\mathbf{e}_1 = \mathbf{x}+\mathbf{e}_2, \mathbf{y}+\mathbf{e}_2 = \mathbf{x}+\mathbf{e}_1, \mathbf{y}+\mathbf{e}_3, \mathbf{y}+\mathbf{e}_4\}\\ I(\mathbf{z}) &= \{\mathbf{z}+\mathbf{e}_1, \mathbf{z}+\mathbf{e}_2, \mathbf{z}+\mathbf{e}_3 = \mathbf{x}+\mathbf{e}_4, \mathbf{z}+\mathbf{e}_4 = \mathbf{x}+\mathbf{e}_3\}\\ I(\mathbf{v}) &= \{\mathbf{v}+\mathbf{e}_1, \mathbf{v}+\mathbf{e}_2, \mathbf{v}+\mathbf{e}_3 = \mathbf{y}+\mathbf{e}_4, \mathbf{v}+\mathbf{e}_4 = \mathbf{y}+\mathbf{e}_3\}.\end{aligned}$$

Now $\mathbf{z}+\mathbf{e}_1 = \mathbf{x}+\mathbf{e}_3+\mathbf{e}_4+\mathbf{e}_1 = \mathbf{y}+\mathbf{e}_3+\mathbf{e}_4+\mathbf{e}_2 = \mathbf{v}+\mathbf{e}_2$ and similarly $\mathbf{z}+\mathbf{e}_2 = \mathbf{v}+\mathbf{e}_1$. Thus, we have $I(\mathbf{x},\mathbf{v}) = I(\mathbf{y},\mathbf{z})$ which is impossible, and, hence, $|I(\mathbf{x})| \geq 5$ for all $\mathbf{x} \in \mathbf{F}^n$.

The optimal dimensions follow from Theorem 6.3. □

Let $\ell \geq 3$. It is proved in [45, Theorem 2] that $C \subseteq \mathbf{F}^n$ is $(1,\leq \ell)$-identifying if and only if it is a $(2\ell-1)$-fold 1-covering. Thus, we have the next theorem, the optimal dimensions follow again from Theorem 6.3.

Theorem 6.8. *If $\ell \geq 3$ and $n = (2\ell-1)2^r - 1 + s$, where $r \geq 0$ and $0 \leq s \leq (2\ell-1)2^r - 1$, then*

$$k_1^{(\leq \ell)}[n] = k[n,1,2\ell-1] = n-r.$$

For $\ell \geq 3$, we know by [45, Corollary 1] that

$$M_1^{(\leq \ell)}(n) = \frac{(2\ell-1)2^n}{n+1}$$

if and only if there are integers $i \geq 0$, $\mu_0 > 0$ such that $\mu_0 \mid (2\ell-1)$ and $2\ell-1 \leq 2^i\mu_0$ and $n = \mu_0 2^i - 1$. When $n = 2^r(2\ell-1)-1$ for $r \geq 0$ the above requirements are true and $(2\ell-1)2^n/(n+1) = 2^{n-r}$, and, thus, the linear codes constructed above for lengths $n = 2^r(2\ell-1)-1$ where $r \geq 0$ are optimal also in the case where linearity is not required.

When $\ell = 2$ and $\ell = 3$ we want, in both cases, each word to be covered by at least five codewords, and, thus, in these two cases the optimal codes are the same.

6.3 Linear strongly identifying codes

We defined strongly $(r, \leq \ell)$-identifying codes in Chapter 5. In this section, we solve all the optimal dimensions for linear strongly $(1, \leq \ell)$-identifying codes for any $\ell \geq 1$ and for any possible code length.

The smallest dimension of a linear strongly $(1, \leq \ell)$-identifying code is denoted by $k_1^{(\leq \ell)SID}[n]$.

Theorem 6.9. *For $n \geq 3$ we have*

$$k_1^{(\leq 1)SID}[n] \geq \lceil n + \log_2 3 - \log_2(n+2) \rceil .$$

Proof. Let C be a linear strongly 1-identifying code, and denote $K = |C|$. There are at most K words which are covered exactly by one codeword. Because a strongly 1-identifying code is 1-identifying, we know from the proof of Theorem 6.4 that all the other words are at least 3-fold 1-covered. For all $\mathbf{c} \in C$, $|I(\mathbf{c})| > 1$ since otherwise $I'(\mathbf{c}) = \emptyset$, which is impossible. In fact, codewords are covered by at least four codewords. Assume to the contrary that $|I(\mathbf{c})| = 3$ for $\mathbf{c} \in C$. Then by Lemma 6.1, $|I(\mathbf{0})| = 3$. In particular, $I(\mathbf{0}) = \{\mathbf{0}, \mathbf{c}_1, \mathbf{c}_2\}$, where $\mathbf{c}_1, \mathbf{c}_2 \in S_1(\mathbf{0})$ are codewords. By linearity $\mathbf{c}_1 + \mathbf{c}_2 \in C$ and using again Lemma 6.1 we get $I'(\mathbf{0}) = I'(\mathbf{c}_1 + \mathbf{c}_2) = \{\mathbf{c}_1, \mathbf{c}_2\}$, which is a contradiction.

Thus, we can calculate

$$4K + K + 3(2^n - 2K) \leq K(n+1)$$

from which the claim follows. □

Theorem 6.10. *Let $n = 3(2^r - 1) + 1 + s$, where $r \geq 1$ and $0 \leq s \leq 3 \cdot 2^r - 1$. Then*

$$k_1^{(\leq 1)SID}[n] = n - r.$$

Proof. Let H be an $(r \times n)$-matrix such that there is at most one nonzero binary r-tuple which occurs only once as the column of H and all the other r-tuples (including the all-zero tuple) occur at least three times. As in the proof of Theorem 6.3, we know that the code which has H as its parity check matrix has the following properties. Each codeword is covered at least four times. There are 2^{n-r} or 0 non-codewords which are covered by exactly one codeword. All the other non-codewords are covered by at least three codewords.

The I-sets $I(\mathbf{x})$ for all $\mathbf{x} \in \mathbf{F}^n$ are unique since at distance one from each codeword there is at most one non-codeword which is covered only by this one codeword, and all the other words of $\mathbf{F}^n$ are covered by at least three codewords. By Lemma 2.1 we know that three codewords is enough to make an I-set unique. Also the sets $I'(\mathbf{x})$ for all codewords are distinguishable, since again there are at least three words in $I'(\mathbf{x})$.

The lower bound follows from Theorem 6.9. □

There is also a linear strongly 1-identifying code of length $n = 3$ (Theorem 6.10 does not give this). We have $k_1^{(\leq 1)SID}[3] = 3$ since $\mathbf{F}^3$ covers each codeword four times and there are no non-codewords. Theorem 6.9 gives the lower bound.

Theorem 6.11. *Let $\ell \geq 3$. If $n = (2\ell - 1)2^r + s$, where $r \geq 0$ and $0 \leq s \leq (2\ell - 1)2^r - 1$, then*

$$k_1^{(\leq \ell)SID}[n] = k[n, 1, 2\ell - 1, 2\ell] = n - r.$$

Proof. By [44] we know that C is strongly $(1 \leq \ell)$-identifying if and only if C covers each word at least $2\ell - 1$ times and each codeword at least 2ℓ times. Thus, we have the claim. The optimal dimensions follow from Theorem 6.3. □

Theorem 6.12. *If $n = 5 \cdot 2^r + s$, where $r \geq 0$ and $0 \leq s \leq 5 \cdot 2^r - 1$, then*

$$k_1^{(\leq 2)SID}[n] = k[n, 1, 5, 6] = n - r.$$

Proof. First we prove the lower bound. Suppose C is a linear strongly $(1, \leq 2)$-identifying code. By Theorem 6.7 we know that each word in $\mathbf{F}^n$ must be covered by at least five codewords, because otherwise C cannot be a linear $(1, \leq 2)$-identifying code. If codewords were covered by five codewords, then $I(\mathbf{0}) = \{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4, \mathbf{0}\}$ for some $\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4 \in C$ of weight one. By linearity and Lemma 6.1 it would be $I'(\mathbf{0}) \cup I'(\mathbf{c}_1 + \mathbf{c}_2 + \mathbf{c}_3 + \mathbf{c}_4) = \{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4, \mathbf{c}_1 + \mathbf{c}_2 + \mathbf{c}_3, \mathbf{c}_1 + \mathbf{c}_2 + \mathbf{c}_4, \mathbf{c}_1 + \mathbf{c}_3 + \mathbf{c}_4, \mathbf{c}_2 + \mathbf{c}_3 + \mathbf{c}_4\} = I'(\mathbf{c}_1 + \mathbf{c}_2) \cup I'(\mathbf{c}_3 + \mathbf{c}_4)$. Thus, all codewords must be covered by at least six codewords.

Upper bound: By the proof of Theorem 6.11, we know that a code attaining the dimension $k[n, 1, 5, 6]$ is strongly $(1, \leq 3)$-identifying. This implies that such a code is also a linear strongly $(1, \leq 2)$-identifying code. The optimal dimensions follow from Theorem 6.3. □

The optimal linear strongly $(1, \leq \ell)$-identifying codes are the same for the values $\ell = 2$ and $\ell = 3$.

Let $\ell \geq 3$. By [44, Theorem 3] we know that $M_1^{(\leq \ell)SID}(n) = (2\ell - 1)2^n/n$, when there are integers $i \geq 0$, $\mu_0 > 0$ such that $\mu_0 \mid (2\ell - 1)$, $2\ell - 1 \leq 2^i \mu_0$ and $n = \mu_0 2^i$. If $n = (2\ell - 1)2^r$, when $r \geq 0$, the optimal linear strongly $(1, \leq \ell)$-identifying codes of length n are optimal also in the case where the linearity is not required.

Chapter 7

Locating-Dominating Codes

Let G be a graph with a vertex set V. A subset $S \subseteq V$ is said to be a *dominating set* if every vertex in $V \setminus S$ is adjacent to at least one vertex in S. A dominating set S is defined to be *locating* if for all $u, v \in V \setminus S$ the sets of adjacent vertices in S are different. This is the definition of *locating-dominating sets* in graphs. Locating-dominating sets in graphs are considered for example in [54, 60–62].

We consider locating-dominating sets or codes in binary Hamming spaces, and we generalize the definition to locate-dominate more than one vertex. This chapter is based on [36]. Theorem 7.22 is not published before.

Consider the binary hypercube (or an arbitrary graph). Assume that it is used to model a multiprocessor architecture, and that each node corresponds to a processor and each edge corresponds to a dedicated link between two processors. We choose some of the processors as *codewords*, and ask each of them to check its r-neighbourhood. Each codeword sends us the symbol 2 if it itself is malfunctioning, 1 if it itself is fine, but at least one processor in its r-neighbourhood is malfunctioning, and 0 otherwise. When we have received the reports from all the codewords we wish to locate the malfunctioning processors, under the assumption that there are at most ℓ of them. The integers r and ℓ are given at the outset. This is the classical locating-dominating set problem if $\ell = 1$ (and then the requirement for the code C, i.e., the set of codewords, is that the sets $B_r(\mathbf{x}) \cap C$ are different and nonempty for all $\mathbf{x} \in \mathbf{F}^n \setminus C$). This leads to the following definition.

Definition 7.1. Let r and ℓ be non-negative integers. A code $C \subseteq \mathbf{F}^n$ is an $(r, \leq \ell)$-locating-dominating code of type A — an $(r, \leq \ell)$-LDA for short — if for all $X \subseteq \mathbf{F}^n$ and $Y \subseteq \mathbf{F}^n$ of size at most ℓ,

$$B_r(X) \cap C = B_r(Y) \cap C \tag{7.1}$$

$$X \cap C = Y \cap C \tag{7.2}$$

implies that $X = Y$.

If $X = \emptyset$, then $B_r(X) \cap C = \emptyset$: so, in particular, the previous definition implies that $B_r(Y) \cap C \neq \emptyset$ whenever $Y \neq \emptyset$.

In the second variant we only consider subsets X and Y of $\mathbf{F}^n \setminus C$. Then $X \cap C$ and $Y \cap C$ are empty, (7.2) becomes void, and (7.1) takes the simple form $B_r(X) \cap C = B_r(Y) \cap C$. This corresponds to a two-step testing procedure: we first fix all the codewords that have sent 2, and then immediately perform a second round, during which we only receive 1's and 0's.

Definition 7.2. Let r and ℓ be non-negative integers. A code $C \subseteq \mathbf{F}^n$ is an $(r, \leq \ell)$-locating-dominating code of type B — an $(r, \leq \ell)$-LDB for short — if for all sets $X \subseteq \mathbf{F}^n \setminus C$ and $Y \subseteq \mathbf{F}^n \setminus C$ of size at most ℓ,

$$B_r(X) \cap C = B_r(Y) \cap C$$

implies that $X = Y$.

Both definitions reduce to the definition of an r-locating-dominating code when $\ell = 1$. Clearly, an $(r, \leq \ell)$-LDA is always an $(r, \leq \ell)$-LDB. The smallest cardinalities of an $(r, \leq \ell)$-LDA and an $(r, \leq \ell)$-LDB of length n are denoted by $LA_r^{(\leq \ell)}(n)$ and $LB_r^{(\leq \ell)}(n)$, respectively.

Example 7.3. (i) Let C be the even-weight code of length $n \geq 5$, i.e., it consists of all 2^{n-1} words of length n that have even weight. We shall see in Theorem 7.9 that C is a $(1, \leq \lfloor (n+1)/2 \rfloor)$-LDA. Clearly, C is not $(1, \leq 2)$-identifying (take $X = \{100\ldots0\}$, $Y = \{000\ldots0, 100\ldots0\}$).

(ii) The code $C = \mathbf{F}^{n-1} \oplus \{0\}$, $n \geq 2$, is a $(1, \leq 2^{n-1})$-LDB (and hence a $(1, \leq \ell)$-LDB for all ℓ). Clearly, C is not $(1, \leq 2)$-LDA (take $X = \{(\mathbf{x}, 0)\}$ and $Y = \{(\mathbf{x}, 0), (\mathbf{x}, 1)\}$ for any $\mathbf{x} \in \mathbf{F}^{n-1}$).

(iii) For all $k \leq \ell$ an $(r, \leq \ell)$-LDA is always $(r, \leq k)$-LDA, and an $(r, \leq \ell)$-LDB is an $(r, \leq k)$-LDB.

7.1 On the case $\ell = 1$

In this section we examine the case $\ell = 1$. We denote

$$L(n) = LA_1^{(\leq 1)}(n) = LB_1^{(\leq 1)}(n).$$

From [40] we know that $M_1(n) \leq nK(n, 2)$. This implies trivially that $L(n) \leq nK(n, 2)$.

Theorem 7.4. *For all $n \geq 4$,*

$$L(n+1) \leq (2n-1)K(n, 2).$$

Proof. Let $C \subseteq \mathbf{F}^n$ $(n \geq 4)$ be a code attaining the bound $K(n,2)$. Denote by $\mathbf{e}_1$ the word of weight one whose first coordinate equals one. Let $C_1 = \{\mathbf{c}+\mathbf{e} \mid \mathbf{c} \in C, \mathbf{e} \in S_1(\mathbf{0})\}$ and $C_2 = \{\mathbf{c}+\mathbf{e} \mid \mathbf{c} \in C, \mathbf{e} \in S_1(\mathbf{0}) \setminus \mathbf{e}_1\}$. We claim that $D = (C_1 \oplus \{0\}) \cup (C_2 \oplus \{1\})$ is 1-locating-dominating.

Denote $O_i = \mathbf{F}^n \oplus \{i\}$ for $i = 0,1$. By the construction, $I(\mathbf{x}) \neq \emptyset$ for all $\mathbf{x} \in \mathbf{F}^{n+1}$. Assume that $\mathbf{x},\mathbf{y} \in \mathbf{F}^{n+1} \setminus D$ and $\mathbf{x} \neq \mathbf{y}$. It is shown in [40] that C_1 is 1-identifying. Hence, if $\mathbf{x},\mathbf{y} \in O_0$, then $I(\mathbf{x}) \neq I(\mathbf{y})$.

Suppose $\mathbf{x} \in O_0$ and $\mathbf{y} \in O_1$. Because $\mathbf{x} \notin D$, the structure of C_1 implies that $|I(\mathbf{x}) \cap O_0| \geq 2$ whereas $|I(\mathbf{y}) \cap O_0| \leq 1$. Hence, $I(\mathbf{x}) \neq I(\mathbf{y})$.

Assume next that $\mathbf{x},\mathbf{y} \in O_1$. There is $\mathbf{c} \in C \oplus \{1\}$ such that $d(\mathbf{c},\mathbf{x}) \leq 2$. If $\mathbf{x} = \mathbf{c}$, then $|I(\mathbf{x}) \cap S_1(\mathbf{c})| \geq 3$ and $|I(\mathbf{y}) \cap I(\mathbf{x})| \leq 2$ by Lemma 2.1. If $\mathbf{x} = \mathbf{c}+\mathbf{e}_1$, then $I(\mathbf{x}) \cap O_0 \neq \emptyset$ and thus $I(\mathbf{x}) \neq I(\mathbf{y})$. Assume then that $\mathbf{x} \in S_2(\mathbf{c})$ and that $\mathbf{x}$ covers two elements of $I(\mathbf{c})$. In this case only $\mathbf{y} = \mathbf{c}$ could cover both elements of $I(\mathbf{x}) \cap S_1(\mathbf{c})$, and this is impossible as we have seen. Suppose finally that $\mathbf{x} \in S_2(\mathbf{c})$ and $d(\mathbf{x},\mathbf{c}+\mathbf{e}_1) = 1$. Then $\mathbf{y}$ should cover the unique element in $I(\mathbf{x}) \cap S_1(\mathbf{c})$ and consequently $\mathbf{y} \in S_2(\mathbf{c})$, but then $\mathbf{y}$ covers also another codeword of $I(\mathbf{c})$, which gives $I(\mathbf{x}) \neq I(\mathbf{y})$. □

By [61, Theorem 10] we have the next theorem.

Theorem 7.5.

$$L(n) \geq \left\lceil \frac{2^{n+1}}{n+3} \right\rceil.$$

In the proof of the next theorem we use the notation of excess, see page 52 (cf. [6]). Theorem can also be proved using the technique of the proof of Theorem 3.3.

Theorem 7.6.

$$L(n) \geq \left\lceil \frac{n^2 2^{n+1}}{n^3+2n^2+3n-2} \right\rceil.$$

Proof. Suppose C is an optimal 1-locating-dominating code of length n. Denote by K the cardinality of C.

It is not difficult to check that there can be only four kinds of words in $\mathbf{F}^n$:

1. a word which is covered by exactly one codeword,

2. a codeword, say $\mathbf{c}_1$, which forms a *couple* with another codeword $\mathbf{c}_2$ such that $I(\mathbf{c}_1) = I(\mathbf{c}_2) = \{\mathbf{c}_1,\mathbf{c}_2\}$,

3. a word $\mathbf{x}$ which has $E(\mathbf{x}) \geq 2$, and

4. a word $\mathbf{y}$ which has $E(\mathbf{y}) = 1$ and for which there exists a word $\mathbf{x}$ such that $E(\mathbf{x}) \geq 2$ and $I(\mathbf{y}) \subset I(\mathbf{x})$.

A word of type 3 is called a *father*. A word of type 4 is called a *son*; the word $\mathbf{x}$ such that $I(\mathbf{y}) \subset I(\mathbf{x})$ is called its father (it is easy to see that $\mathbf{x}$ is uniquely defined). A *family* consists of a father and its sons. The families, couples and points with excess zero partition the whole space $\mathbf{F}^n$.

Suppose that a father is covered by i codewords. Then there are at most $\binom{i}{2}$ sons in the family. The average excess on the points in a family whose father is covered by exactly $i \geq 3$ codewords is therefore at least

$$f(i) := \frac{\binom{i}{2} + i - 1}{\binom{i}{2} + 1}.$$

This is a decreasing function on $i \geq 4$; and $f(3) = f(6)$. Assume that $n \geq 6$. Then $f(i) \geq f(n)$ for all $3 \leq i \leq n$. Since C is an optimal code there is no codeword $\mathbf{c} \in C$ such that $|I(\mathbf{c})| = n+1$. Namely, if there were such a codeword, then we could remove it and still get a 1-locating-dominating code. This is impossible since C is optimal.

The excess on $\mathbf{F}^n$ is $K(n+1) - 2^n$. We now estimate it in a different way. There are at most $2K$ words outside the families. The uniquely covered points contribute nothing to excess. We can estimate

$$K(n+1) - 2^n \geq (2^n - 2K)f(n),$$

from which we get the claimed lower bound on K for $n \geq 6$.

When $n = 5$ then the minimum of $f(3), f(4)$ and $f(5)$ is $f(3)$. Now we can estimate $6K - 2^5 \geq (2^5 - 2K)f(3)$. From this we get $K \geq 9$. And, thus, the lower bound holds for $n = 5$. Similarly, when $n = 4$ we notice that the lower bound holds. For $1 \leq n \leq 3$ the lower bound is true by Theorem 7.5. □

Theorem 7.7. *If $C \subseteq \mathbf{F}^n$ is 1-locating-dominating with the property that $d(\mathbf{c}, C \setminus \{\mathbf{c}\}) = 1$ for all $\mathbf{c} \in C$, then $D = C \oplus \mathbf{F}$ is 1-identifying.*

Proof. Let $\mathbf{x}, \mathbf{y} \in \mathbf{F}^{n+1}$, $\mathbf{x} = (\mathbf{x}_1, x_2)$ and $\mathbf{y} = (\mathbf{y}_1, y_2)$, where $\mathbf{x}_1, \mathbf{y}_1 \in \mathbf{F}^n$ and $x_2, y_2 \in \mathbf{F}$. If $\mathbf{x}_1 \neq \mathbf{y}_1$ and $\mathbf{x}_1, \mathbf{y}_1 \notin C$, then there is $\mathbf{c} \in I(C; \mathbf{x}_1) \triangle I(C; \mathbf{y}_1)$. Without loss of generality we can assume $\mathbf{c} \in I(C; \mathbf{x}_1)$. Thus, $(\mathbf{c}, x_2) \in I(D; \mathbf{x}) \setminus I(D; \mathbf{y})$.

Assume $\mathbf{x}_1 \neq \mathbf{y}_1$ and at least the other one of them belong to C, assume $\mathbf{x}_1 \in C$. Then $d(\mathbf{x}_1, \mathbf{y}_1) \geq 1$ and, thus, $(\mathbf{x}_1, y_2 + 1) \in I(D; \mathbf{x}) \setminus I(D; \mathbf{y})$. If $\mathbf{x}_1 = \mathbf{y}_1$, then $x_2 \neq y_2$. By the assumption $d(\mathbf{a}, C \setminus \{\mathbf{a}\}) = 1$ for all $\mathbf{a} \in \mathbf{F}^n$, we know that there is $\mathbf{c} \in C$ such that $d(\mathbf{c}, \mathbf{x}_1) = 1$ and, thus, $(\mathbf{c}, \mathbf{x}_2) \in I(D; \mathbf{x}) \setminus I(D; \mathbf{y})$. □

It is proved in [6] (see also Theorem 3.20) that if $C \subseteq \mathbf{F}^n$ is 1-identifying, then $C \oplus \mathbf{F}^2$ is also 1-identifying. The next theorem shows[1] that we do not need a 1-identifying code in this construction, 1-locating-dominating code is enough.

[1] as was pointed out by Tero Laihonen

Theorem 7.8. *Let $C \subseteq \mathbf{F}^n$ be 1-locating-dominating, then $C \oplus \mathbf{F}^2$ is 1-identifying which satisfies that for all $\mathbf{c} \in C \oplus \mathbf{F}^2$ we have $d(\mathbf{c}, (C \oplus \mathbf{F}^2) \setminus \{\mathbf{c}\}) = 1$.*

Proof. It will be proved in Theorem 7.18 that $C \oplus \mathbf{F}$ is 1-locating-dominating. The direct sum construction also implies that all the codewords are 2-fold 1-covered in $\mathbf{F}^{n+1}$ and in $\mathbf{F}^{n+2}$. The claim follows from the previous theorem. □

7.2 Locating-dominating codes of type A

In this section we consider locating-dominating codes of type A when $\ell \geq 3$. We have the following characterization of the locating-dominating codes of type A; cf. [45].

Theorem 7.9. *Assume that $\ell \geq 3$. A code $C \subseteq \mathbf{F}^n$ is a $(1, \leq \ell)$-LDA if and only if every element of $\mathbf{F}^n \setminus C$ is covered by at least $2\ell - 1$ codewords of C.*

Proof. Assume first that every element of $\mathbf{F}^n \setminus C$ is covered by at least $2\ell - 1$ codewords of C. Assume that X and Y are two subsets of $\mathbf{F}^n$, each of size at most ℓ, and that $X \cap C = Y \cap C$ and

$$B_1(X) \cap C = B_1(Y) \cap C. \tag{7.3}$$

Assume contrary that $\mathbf{x} \in X \setminus C$ does not belong to Y. Without loss of generality $\mathbf{x} = 000\ldots0$. As a non-codeword, $\mathbf{x}$ is covered by at least $2\ell - 1$ codewords, say $\mathbf{c}_1, \ldots, \mathbf{c}_{2\ell-1}$; and without loss of generality, they have been indexed in such a way that $\mathbf{c}_1, \ldots, \mathbf{c}_i \in X$, whereas $\mathbf{c}_{i+1}, \ldots, \mathbf{c}_{2\ell-1} \notin X$ $(0 \leq i \leq \ell - 1)$. The fact that $X \cap C = Y \cap C$ immediately implies that $\mathbf{c}_1, \ldots, \mathbf{c}_i \in Y$ and $\mathbf{c}_{i+1}, \ldots, \mathbf{c}_{2\ell-1} \notin Y$.

By (7.3), the codewords $\mathbf{c}_{i+1}, \ldots, \mathbf{c}_{2\ell-1}$ are therefore covered by the at most $\ell - i$ words of weight two of Y, each covering at most two of the codewords $\mathbf{c}_{i+1}$, $\ldots, \mathbf{c}_{2\ell-1}$. Because $2(\ell - i - 1) < 2\ell - 1 - i$, there are exactly $\ell - i$ words of weight two in Y (together with the remaining i words of weight one).

We first notice that the set X does not contain any word of weight four. Indeed, because $X \cap C = Y \cap C$, and Y does not contain any word of weight four, there is no codeword of weight four in X. Any non-codeword of weight four in X would be covered by at least $(2\ell - 1) - 4$ (> 0) codewords of weight five, which cannot be, because no word in Y could cover a codeword of weight five.

Assume that $\mathbf{y} \in Y \setminus C$ has weight two. We show that $\mathbf{y} \in X$. By assumption, $\mathbf{y}$ is covered by at least $2\ell - 1$ codewords, of which at least $2\ell - 3$ have weight three; and they all must be covered by the words in X. Because there are no words of weight four in X, all these at least $2\ell - 3$ words of weight three must be covered by the at most $\ell - 1 - i$ words of weight two and three in X. Because $2\ell - 3 > \ell - 1 - i$ for all $\ell \geq 3$, and no word of weight two or three other than $\mathbf{y}$ itself covers more than one of these at least $2\ell - 3$ words, we conclude that $\mathbf{y}$ must belong to X.

We have shown that $Y \setminus C \subseteq X$. We already knew that $Y \cap C = X \cap C \subseteq X$, and hence $Y \subseteq X$. But Y has size ℓ, and therefore $Y = X$, contradicting the fact that $\mathbf{x} \in X \setminus Y$.

We have therefore shown that C is a $(1, \leq \ell)$-LDA.

Assume conversely that C is a $(1, \leq \ell)$-LDA, and let $\mathbf{x} \in \mathbf{F}^n \setminus C$ be arbitrary. Without loss of generality, $\mathbf{x} = 000\ldots0$. Assume that the codewords that cover $\mathbf{x}$ are $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_j$. Let $i = \lceil j/2 \rceil$. We can choose i words $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_i \in \mathbf{F}^n \setminus \{\mathbf{x}\}$ that together cover all the words $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_j$. Then $X = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_i, \mathbf{x}\}$ and $Y = \{\mathbf{x}_1, \mathbf{x}_2, \ldots \mathbf{x}_i\}$ satisfy (7.1) and (7.2), and both have cardinality at most $i+1 \leq \ell$, unless $j \geq 2\ell - 1$. Hence every non-codeword is covered by at least $2\ell - 1$ codewords of C. □

Let us denote by $K(n, r, \mu, \nu)$ the smallest cardinality of a code $C \subseteq \mathbf{F}^n$ such that every non-codeword is r-covered by at least μ codewords and every codeword is r-covered by at least ν codewords.

From the proof of Theorem 6.2 we get the following lower bound.

Theorem 7.10. *For $\ell \geq 3$,*

$$LA_1^{(\leq \ell)}(n) = K(n, 1, 2\ell - 1, 1) \geq \frac{(2\ell - 1)2^n}{n + 2\ell - 1}.$$

Proof. By the previous theorem it suffices to prove the inequality. Assume that $C \subseteq \mathbf{F}^n$ attains the bound $K(n, 1, 2\ell - 1, 1)$. We count in two ways the number of pairs $(\mathbf{c}, \mathbf{x})$, where $\mathbf{c} \in C$, $\mathbf{x} \in \mathbf{F}^n$ and $d(\mathbf{c}, \mathbf{x}) \leq 1$. Given $\mathbf{c}$, there are $n+1$ choices for $\mathbf{x}$, so the number of such pairs equals $(n+1)|C|$. Given $\mathbf{x}$, there are at least $2\ell - 1$ choices for $\mathbf{c}$ if $\mathbf{x}$ is not a codeword, and at least one, if $\mathbf{x}$ is a codeword. Hence $(n+1)|C| \geq (2\ell - 1)(2^n - |C|) + |C|$, which gives the claim. □

From [45] we know that

$$\frac{M_1^{(\leq \ell)}(n)}{(2\ell - 1)2^n/(n+1)} \to 1$$

for a given $\ell \geq 3$ when $n \to \infty$. Because $LA_1^{(\leq \ell)}(n) \leq M_1^{(\leq \ell)}(n)$, the previous theorem gives the following corollary.

Corollary 7.11. *For a fixed $\ell \geq 3$,*

$$LA_1^{(\leq \ell)}(n) \sim \frac{(2\ell - 1)2^n}{n}$$

when $n \to \infty$.

We also get the following infinite family of optimal codes.

Theorem 7.12. *Assume that $\ell \geq 3$. Then*

$$LA_1^{(\leq \ell)}((2\ell-1)(2^s-1)) = 2^{(2\ell-1)(2^s-1)-s}$$

for all $s = 1,2,\ldots$.

Proof. By Theorem 6.3 we know that for $n = (2\ell-1)(2^s-1)$ the inequality of Theorem 7.10 is attained by a linear code whose parity check matrix has as its columns $2\ell-1$ copies of every non-zero element in $\mathbf{F}^s$. □

Theorem 7.13. *Let $\ell \geq 3$. If C is a $(1,\leq \ell)$-LDA then the direct sum $C \oplus \mathbf{F}$ is also a $(1,\leq \ell)$-LDA. Thus $LA_1^{(\leq \ell)}(n) \leq 2LA_1^{(\leq \ell)}(n-1)$.*

Proof. The claim follows from Theorem 7.9 since the direct sum preserves the property that each non-codeword is covered by at least $2\ell-1$ codewords. □

7.3 Locating-dominating codes of type B

In this section we consider $(r,\leq \ell)$-locating-dominating codes of type B for $\ell \geq 2$ and $r \geq 1$. Firstly, we prove lower bounds and secondly we introduce some constructions.

7.3.1 Lower bounds

We denote

$$\begin{aligned} N &= \mathbf{F}^n \setminus C, \\ C_i &= \{\mathbf{c} \in C \mid |I(\mathbf{c})| = i\}, \\ N_i &= \{\mathbf{x} \in N \mid |I(\mathbf{x})| = i\} \end{aligned}$$

and $C_{i,j} = \cup_{k=i}^{j} C_k$ and further $N_{i,j} = \cup_{k=i}^{j} N_k$.

Theorem 7.14. *Let $2 \leq \ell \leq n-1$. Then*

$$LB_1^{(\leq \ell)}(n) \geq \left\lceil \frac{\ell}{n+\ell-1} 2^n \right\rceil .$$

Proof. Assume that $C \subseteq \mathbf{F}^n$ is a $(1,\leq \ell)$-LDB where $2 \leq \ell \leq n-1$.

Let us count in two ways the number of pairs $(\mathbf{x},\mathbf{c})$ where $\mathbf{x} \in N_{1,\ell-1}$, $\mathbf{c} \in C_n$ and $d(\mathbf{c},\mathbf{x}) = 1$. Every $\mathbf{x} \in N_{1,\ell-1}$ has at least one $\mathbf{c} \in C_n$ in $I(\mathbf{x})$, since otherwise there would exist $\mathbf{y}_i \in N \cap S_2(\mathbf{x}) \cap B_1(\mathbf{c}_i)$ for all $\mathbf{c}_i \in I(\mathbf{x})$ and, thus,

$$I(\mathbf{y}_1,\ldots,\mathbf{y}_{|I(\mathbf{x})|}) = I(\mathbf{y}_1,\ldots,\mathbf{y}_{|I(\mathbf{x})|},\mathbf{x})$$

which is impossible. On the other hand, each $\mathbf{c} \in C_n$ can have at most one $\mathbf{x} \in N_{1,\ell-1}$ at distance one from it. Therefore,

$$|N_{1,\ell-1}| \leq |C_n|. \tag{7.4}$$

Next we compute in two ways the number of pairs $(\mathbf{x}, \mathbf{c})$ where $\mathbf{c} \in C_1$, $\mathbf{x} \in N_{\ell+1,n}$ and $d(\mathbf{x}, \mathbf{c}) = 1$. All the non-codewords at distance one from $\mathbf{c} \in C_1$ belong to $N_{\ell+1,n}$. Indeed, assume that there were a word $\mathbf{x}$ at distance one from $\mathbf{c}$ such that $I(\mathbf{x}) = \{\mathbf{c}_1, \ldots, \mathbf{c}_i, \mathbf{c}\}$ $(0 \leq i \leq \ell - 1)$. If $i = 0$, let $\mathbf{y}$ be any word other than $\mathbf{x}$ which is at distance 1 from $\mathbf{c}$. Then $I(\mathbf{y}) = I(\mathbf{y}, \mathbf{x})$. Assume that $i \geq 1$. Denoting by $\mathbf{y}_i$ the unique word other than $\mathbf{x}$ such that $d(\mathbf{c}, \mathbf{y}_i) = d(\mathbf{y}_i, \mathbf{c}_i) = 1$, we get the contradiction (note that $\mathbf{y}_i \in N$)

$$I(\mathbf{y}_1, \ldots, \mathbf{y}_i) = I(\mathbf{y}_1, \ldots, \mathbf{y}_i, \mathbf{x}).$$

On the other hand, a non-codeword $\mathbf{x} \in N_{\ell+1,n}$ can have at most n elements of C_1 at distance one from it. Consequently,

$$|C_1| \leq |N_{\ell+1,n}|.$$

Finally, we count the number of pairs $(\mathbf{x}, \mathbf{c})$ such that $\mathbf{x} \in \mathbf{F}^n$, $\mathbf{c} \in C$ and $d(\mathbf{x}, \mathbf{c}) \leq 1$. Now using (7.4) and the previous inequality, we obtain

$$\begin{array}{rcl} |C|(n+1) & \geq & \ell|N| - (\ell-1)|N_{1,\ell-1}| + |N_{\ell+1,n}| + 2|C| - |C_1| + (n-2)|C_n| \\ & \geq & \ell|N| + 2|C|. \end{array}$$

This yields the claim. □

Lemma 7.15. *Let $C \subseteq \mathbf{F}^n$ be a $(1, \leq \ell)$-LDB with $2 \leq \ell \leq n-1$. If j is a non-negative integer satisfying $\ell - j \geq 1$ and $\ell + j \leq n$, then*

$$(\ell - j)|N_{1,\ell+j}| \leq (n - \ell + j + 1)|C_{\ell-j,n-1}| + (\ell - j)|C_n|.$$

Proof. Let us count the number of pairs $(\mathbf{c}, \mathbf{x})$, where $\mathbf{c} \in C_{\ell-j,n}$, $\mathbf{x} \in N_{1,\ell+j}$ and $d(\mathbf{x}, \mathbf{c}) = 1$, in two ways. As mentioned in the proof of Theorem 7.14, if $\mathbf{x} \in N_{1,\ell-1}$, then there is at least one such pair $(\mathbf{c}, \mathbf{x})$.

Let us then concentrate on $\mathbf{x} \in N_{\ell,\ell+j}$. If there is $\mathbf{c} \in I(\mathbf{x})$ such that $\mathbf{c} \in C_n$, then there is again one sought pair. Assume then that for all $\mathbf{c} \in I(\mathbf{x})$ we have $|I(\mathbf{c})| \leq n - 1$. Denote by $M_{\ell,\ell+j}$ the subset of $N_{\ell,\ell+j}$ which consists of all such elements $\mathbf{x}$.

Let $\mathbf{x} \in M_{\ell,\ell+j}$ and without loss of generality $\mathbf{x} = 00\ldots0$. Denote by P a maximal subset of $\{\mathbf{y} \in N \cap S_2(\mathbf{x}) \mid |I(\mathbf{y}) \cap I(\mathbf{x})| = 2\}$ such that its minimum distance is four. Since C is a $(1, \leq \ell)$-LDB, we must have

$$\ell \leq |I(\mathbf{x})| - |P|. \tag{7.5}$$

Indeed, suppose $|I(\mathbf{x})|-|P| \leq \ell-1$. Denote by $\mathbf{y}_i$ a word in $N \cap S_2(\mathbf{x})$ such that $d(\mathbf{y}_i,\mathbf{c}_i)=1$ for $\mathbf{c}_i \in R = I(\mathbf{x}) \setminus I(P)$. Notice that such words $\mathbf{y}_i$ exist for every $\mathbf{c}_i$ because $\mathbf{x} \in M_{\ell,\ell+j}$. Now this leads to

$$I(P \cup \{\mathbf{y}_1,\ldots,\mathbf{y}_{|R|}\}) = I(P \cup \{\mathbf{y}_1,\ldots,\mathbf{y}_{|R|},\mathbf{x}\})$$

which is impossible because

$$|P \cup \{\mathbf{y}_1,\ldots,\mathbf{y}_{|R|},\mathbf{x}\}| = |I(\mathbf{x})|-|P|+1 \leq \ell.$$

Clearly, by (7.5), $0 \leq |P| \leq |I(\mathbf{x})|-\ell \leq j$. Hence, by virtue of (7.5), $|R| = |I(\mathbf{x})|-2|P| \geq \ell-|P| \geq \ell-j$. Since P is maximal, $\{\mathbf{y} \in S_2(\mathbf{x}) \mid |I(\mathbf{y}) \cap R| = 2\} \subseteq C$, and consequently, we get at least $\ell-j$ pairs from every $\mathbf{x} \in M_{\ell,\ell+j}$.

All in all the number of sought pairs is at least

$$\begin{array}{rl} & |N_{1,\ell-1}|+|N_{\ell,\ell+j} \setminus M_{\ell,\ell+j}|+(\ell-j)|M_{\ell,\ell+j}| \\ = & |N_{1,\ell-1}|+|N_{\ell,\ell+j}|+(\ell-j-1)|M_{\ell,\ell+j}|. \end{array}$$

On the other hand, at distance one from $\mathbf{c} \in C_{\ell-j,n-1}$ (resp. $\mathbf{c} \in C_n$) there are at most $n-\ell+j+1$ words (resp. one word) of N and hence possibly of $N_{1,\ell+j}$. Thus, the number of sought pairs is at most

$$(n-\ell+j+1)|C_{\ell-j,n-1}|+|C_n|.$$

Furthermore, one immediately verifies the modification of (7.4) that $|N_{1,\ell-1}|+|N_{\ell,\ell+j} \setminus M_{\ell,\ell+j}| \leq |C_n|$. Multiplying this by $\ell-j-1$ and adding it to the inequality which we get from the estimates of the number of pairs proves the claim. □

Theorem 7.16. *Let* $\sqrt{n+1} \leq \ell \leq n-1$. *Denote*

$$k = \min\left\{ \left\lfloor \frac{\ell^2-n-1}{1+\ell+n} \right\rfloor, n-\ell-1 \right\}.$$

Then

$$LB_1^{(\leq \ell)}(n) \geq \left\lceil \frac{\ell+k+1}{n+\ell+k+1} 2^n \right\rceil.$$

Proof. Let $C \subseteq \mathbf{F}^n$ be a $(1,\leq \ell)$-LDB with $\sqrt{n+1} \leq \ell \leq n-1$. Obviously,

$$\begin{array}{rcl} |C|(n+1) & \geq & (\ell+k+1)|N|-(k+1)|N_{\ell,\ell+k}|-(\ell+k)|N_{1,\ell-1}| \\ & & +|C|+(\ell-k-1)|C_{\ell-k,n-1}|+(n-1)|C_n|. \end{array}$$

Since $\ell \leq n-1$ we get by (7.4) that

$$\begin{array}{rcl} |C|(n+1) & \geq & (\ell+k+1)|N|-(k+1)|N_{\ell,\ell+k}|-(k+1)|N_{1,\ell-1}| \\ & & +|C|+(\ell-k-1)|C_{\ell-k,n-1}|+(n-\ell)|C_n|. \end{array}$$

It is easy to check that the choice of k satisfies

$$\ell - k - 1 \geq \frac{(k+1)(n-\ell+k+1)}{\ell-k}$$

and furthermore

$$n - \ell \geq k+1.$$

Consequently, using Lemma 7.15 for $j = k$ one gets

$$|C|(n+1) \geq (\ell+k+1)|N| + |C|.$$

The claim follows immediately from this. □

7.3.2 Constructions

In this section we give different constructions for locating-dominating codes of type B. In the beginning we consider $r = 1$, in Theorem 7.22 we have $r \geq 2$.

Theorem 7.17. *Let $n \geq 3$. If $\ell \geq \frac{1}{2}(\sqrt{2n^2+2n+1}-1)$, then*

$$LB_1^{(\leq \ell)}(n) = 2^{n-1}.$$

Proof. If $\ell \geq \frac{1}{2}(\sqrt{2n^2+2n+1}-1)$, then in Theorem 7.16 we have $k = n-\ell-1$. Combining this lower bound to Example 7.3(ii) and (iii) we get the claim. □

Theorem 7.18. *If C is a $(1,\leq \ell)$-LDB, then $C \oplus \mathbf{F}$ is also a $(1,\leq \ell)$-LDB. In particular,*

$$LB_1^{(\leq \ell)}(n+1) \leq 2LB_1^{(\leq \ell)}(n).$$

Proof. Suppose C is a $(1,\leq \ell)$-LDB of length n. Divide the words of $\mathbf{F}^{n+1}$ into two layers O_0 and O_1 depending on the last bit of the word, i.e., $O_i = \mathbf{F}^n \oplus \{i\}$ for $i \in \mathbf{F}$. Only a codeword can have in its I-set a codeword from the other layer than where it itself lies. Thus, the I-set of a set of non-codewords on the layer O_i has codewords only from the same layer. So, if $|X| \leq \ell$, then $|X \cap O_i| \leq \ell$ for both $i \in \mathbf{F}$, and because C is a $(1,\leq \ell)$-LDB, $I(X \cap O_i) \subseteq O_i$ uniquely identifies $X \cap O_i$ for both $i \in \mathbf{F}$. □

We denote by $A(n,d,w)$ the smallest cardinality of a code of length n and minimum distance d whose codewords all have weight w. For values of this function, see [53].

Denote $S_i = S_i(\mathbf{0})$.

Theorem 7.19. *For $\ell \leq n/2-1$ we have*

$$LB_1^{(\leq \ell)}(n) \leq K(n,\ell+1)(V(n,\ell) - A(n,6,\ell)).$$

Proof. Let D be a code attaining the value $K(n,\ell+1)$, A a code attaining the value $A(n,6,\ell)$, and $B = B_\ell(00\ldots0)$. We shall show that the code

$$C = D + (B \setminus A) = \{\mathbf{c}+\mathbf{b} \mid \mathbf{c} \in D,\ \mathbf{b} \in B \setminus A\}$$

is a $(1,\leq \ell)$-LDB. In the code C there are at most $K(n,\ell+1)(V(n,\ell) - A(n,6,\ell))$ codewords.

Suppose there are sets $X \subseteq N$ and $Y \subseteq N$ such that $X \neq Y$, $|X|,|Y| \leq \ell$ and $I(X) = I(Y)$. There is a word $\mathbf{x} \in X$ such that $\mathbf{x} \notin Y$. Since $\mathbf{x} \notin C$ there is a codeword $\mathbf{c} \in D$ such that $d(\mathbf{x},\mathbf{c}) = \ell$ or $\ell+1$. Without loss of generality we can assume that $\mathbf{c} = 00\ldots0$.

If $d(\mathbf{x},\mathbf{c}) = \ell$, then $\mathbf{x}$ is covered by ℓ codewords of weight $\ell-1$. None of them can belong to $I(Y)$ since $B_{\ell-1}(00\ldots0) \cap N = \emptyset$ and non-codewords in $S_\ell \setminus \{\mathbf{x}\}$ are at least at distance six from $\mathbf{x}$.

Suppose now that $d(\mathbf{x},\mathbf{c}) = \ell+1$. Then $\ell \leq |I(\mathbf{x}) \cap S_\ell| \leq \ell+1$. If $\mathbf{y} \in Y$ covers a codeword of weight ℓ in $I(\mathbf{x})$, then the weight of $\mathbf{y}$ must be $\ell+1$; and since $\mathbf{y} \neq \mathbf{x}$, $\mathbf{y}$ can cover only one codeword of weight ℓ in $I(\mathbf{x})$. Because $|Y| \leq \ell$, this implies that $|I(\mathbf{x}) \cap S_\ell| = \ell$, $|Y| = \ell$, $Y \subseteq S_{\ell+1}$, and each word in Y covers one codeword from $I(\mathbf{x}) \cap S_\ell$. For $\mathbf{y} \in Y$ we have $|I(\mathbf{y}) \cap S_\ell| = \ell+1$, because the minimum distance of non-codewords of weight ℓ is six. As we have already seen (for $\mathbf{x}$ whose role $\mathbf{y}$ now assumes), this is only possible if $\mathbf{y} \in X$. Consequently, X must contain all the ℓ elements of Y together with $\mathbf{x} \notin Y$, which is impossible, since $|X| \leq \ell$. □

The previous theorem is interesting in the light of Conjecture 2.4. Namely, for every fixed $\ell \geq 2$ Theorems 7.14 and 7.19 give the bounds

$$\frac{\ell}{n}2^n(1+f(n)) \leq LB_1^{(\leq \ell)}(n) \leq \frac{\ell+1}{n}2^n(1+g(n)),$$

where both $f(n)$ and $g(n)$ tend to zero when $n \to \infty$.

Next we construct some $(1,\leq 2)$-LDBs.

Theorem 7.20. *Let $n \geq 5$ and C be a code attaining the bound $K(n,3)$. Then the code*

$$D = C \cup \{\mathbf{c}+\mathbf{x} \mid \mathbf{c} \in C, \mathbf{x} \in S_2(\mathbf{0})\}$$

is a $(1,\leq 2)$-LDB and has cardinality at most $K(n,3)(1+\binom{n}{2})$.

Proof. Suppose there are $X,Y \subseteq \mathbf{F}^n \setminus D$ such that $|X|,|Y| \leq 2$, $X \neq Y$ and $I(X) = I(Y)$. By the definition of D, every $\mathbf{y} \notin D$ has distance 1 or 3 to C.

We first show that any word $\mathbf{y} \in Y$ for which there is a codeword $\mathbf{c} \in C$ such that $d(\mathbf{c},\mathbf{y}) = 1$ also belongs to X. Indeed, such a word $\mathbf{y}$ is covered by n (≥ 5) codewords of D and thus by Lemma 2.1 no set of size at most two where $\mathbf{y}$ does not belong to can cover all those words.

Since $X \neq Y$ there is $\mathbf{y} \in Y \setminus X$ such that $d(\mathbf{c},\mathbf{y}) = 3$ for some $\mathbf{c} \in C$. Without loss of generality, $\mathbf{c} = 0000\ldots$ and $\mathbf{y} = 1110\ldots$. Then $\mathbf{y}$ is covered by the codewords $1100\ldots$, $1010\ldots$ and $0110\ldots$ of D. If an element $\mathbf{x}$ of $X \subseteq \mathbf{F}^n \setminus D$ covers at least one of these three words, then $\mathbf{x}$ has weight one or three. If $w(\mathbf{x}) = 3$, then $\mathbf{x}$ $(\neq \mathbf{y})$ covers at most one of them. Consequently, there is a word $\mathbf{x} \in X$ of weight one which covers two of them. But being of weight one, $\mathbf{x}$ must also belong to Y. Without loss of generality, $\mathbf{x} = 1000\ldots$.

Still we have one codeword $\mathbf{c} = 0110\ldots$ which is not in $I(\mathbf{x})$. To cover this word $\mathbf{c}$ we need one more word in the set X. This word cannot have weight one since then it should belong to Y as well, which is impossible. So, without loss of generality, the other word in X is $0111\ldots$. But then $0101\ldots$ and $0011\ldots$ belong to $I(X)$ but not to $I(Y)$, which gives the required contradiction. □

Example 7.21. The repetition code $\{0000000, 1111111\}$ satisfies $K(7,3)$. By the previous theorem we have a $(1,\leq 2)$-LDB of length 7 and cardinality 44. From this we get by Theorem 7.18 a $(1,\leq 2)$-LDB of length 8 and cardinality 88.

Let C be the perfect binary Golay code of length 23, dimension 12 and minimum distance 7 (and covering radius 3). By Theorem 7.20 we can construct a $(1,\leq 2)$-LDB of length 23 with cardinality $2^{12}(1+\binom{23}{2}) = 1040384$.

The next theorem[2] considers locating-dominating codes of type B when $r \geq 2$ and $\ell = 2$.

Theorem 7.22. *Let $r \geq 2$. Let $C' \subseteq \mathbf{F}^n$ be a code attaining $K(n,2r)$. Then*

$$C = C' + B_r(\mathbf{0})$$

is $(r,\leq 2)$-locating-dominating of type B.

Proof. Suppose we are given a set $I_r(X)$ and we have to settle out the set $X \subseteq \mathbf{F}^n \setminus C$. We show that for $\mathbf{c} \in C'$ we can tell what $X \cap B_{2r}(\mathbf{c})$ is by looking at $B_r(\mathbf{c}) \cap I_r(X)$. Without loss of generality assume that $\mathbf{c} = \mathbf{0}$ and $X \subseteq B_{2r}(\mathbf{c}) \setminus B_r(\mathbf{c})$ where $|X| \leq 2$. Let k be the smallest weight of codewords in $B_r(\mathbf{0}) \cap I_r(X)$.

If $|I_r(X) \cap S_k(\mathbf{0})| = \binom{k+r}{k}$, then

$$\mathrm{supp}(\mathbf{x}) = \bigcup_{\mathbf{u} \in I_r(X) \cap S_k(\mathbf{0})} \mathrm{supp}(\mathbf{u})$$

for any word $\mathbf{x} \in X$ of weight $k+r$. If $I_r(X) \setminus I_r(\mathbf{x}) = \emptyset$, then $X = \{\mathbf{x}\}$, otherwise $X = \{\mathbf{x},\mathbf{y}\}$. We already know that $w(\mathbf{x}) < w(\mathbf{y})$. Hence, there is $i \in \mathrm{supp}(\mathbf{y}) \setminus \mathrm{supp}(\mathbf{x})$. Denote $A = \{i \in \mathrm{supp}(\mathbf{y}) \mid i \notin \mathrm{supp}(\mathbf{x})\}$ and $B = \{i \in \mathrm{supp}(\mathbf{y}) \mid i \in \mathrm{supp}(\mathbf{x})\}$. Because $w(\mathbf{y}) - r = |A| + |B| - r \leq |A| + w(\mathbf{x}) - r$ and $|A| \geq 1$, any word $\mathbf{v}$ of weight $w(\mathbf{y}) - r$ in $I_r(X)$ which has

$$|\mathrm{supp}(\mathbf{v}) \cap B| \leq w(\mathbf{x}) - r \text{ and } |\mathrm{supp}(\mathbf{v}) \cap A| \geq 1 \tag{7.6}$$

[2] Research is done with Tero Laihonen.

is at least at distance $r+1$ from $\mathbf{x}$ and at distance r from $\mathbf{y}$. Thus, $\mathbf{v} \in I_r(X) \setminus I_r(\mathbf{x})$. On the other hand, the union of the supports of all the possible choices of words $\mathbf{v}$ satisfying (7.6) includes all the coordinates of $\mathrm{supp}(\mathbf{y})$. Hence,

$$\mathrm{supp}(\mathbf{y}) = \bigcup_{\substack{\mathbf{v} \in I_r(X)\setminus I_r(\mathbf{x}) \\ w(\mathbf{v})=w(\mathbf{y})-r}} \mathrm{supp}(\mathbf{v})$$

Suppose $|I_r(X) \cap S_k(\mathbf{0})| > \binom{k+r}{k}$. Now $\{\mathbf{x},\mathbf{y}\} \subseteq S_{k+r}(\mathbf{0})$. Suppose first that $k > 1$. For $i \in \mathrm{supp}(\mathbf{x}) \bigtriangleup \mathrm{supp}(\mathbf{y})$ we know that

$$|\{\mathbf{c} \in S_k(\mathbf{0}) \cap I_r(X) \mid i \in \mathrm{supp}(\mathbf{c})\}| = \binom{k+r-1}{k-1}. \tag{7.7}$$

If $i \in \mathrm{supp}(\mathbf{x}) \cap \mathrm{supp}(\mathbf{y})$, then

$$|\{\mathbf{c} \in S_k(\mathbf{0}) \cap I_r(X) \mid i \in \mathrm{supp}(\mathbf{c})\}| > \binom{k+r-1}{k-1}.$$

If $i \notin \mathrm{supp}(\mathbf{x}) \cup \mathrm{supp}(\mathbf{y})$, then $|\{\mathbf{c} \in S_k(\mathbf{0}) \cap I_r(X) \mid i \in \mathrm{supp}(\mathbf{c})\}| = 0$. Hence, for some i such that (7.7) is satisfied, we see that

$$\mathrm{supp}(\mathbf{x}) = \bigcup_{\substack{\mathbf{c} \in I_r(X) \cap S_k(\mathbf{0}) \\ i \in \mathrm{supp}(\mathbf{c})}} \mathrm{supp}(\mathbf{c})$$

and

$$\mathrm{supp}(\mathbf{y}) = \bigcup_{\mathbf{c} \in (I_r(X)\setminus I_r(\mathbf{x})) \cap S_k(\mathbf{0})} \mathrm{supp}(\mathbf{c}).$$

If $k=1$, then we can do the same except using $k+1=2$ instead of k (support i have to occur now r times). There are codewords of weight $k+1=2$ when $r \geq 2$. □

Corollary 7.23. *For $r \geq 2$*

$$LB_r^{(\leq 2)}(n) \leq V(n,r)K(n,2r).$$

In Table 7.1 we have collected lower and upper bounds on the cardinalities of 1-locating-dominating codes, $(1,\leq 2)$- and $(1,\leq 3)$-locating-dominating codes of type B for short lengths.

Table 7.1: Values of the cardinalities of locating-dominating codes.

The cases $n = 1$ are trivial. The lower bounds $L(2) \geq 2$ and $L(3) \geq 4$ are easy to check by hand and Example 7.3(iii) then gives the lower bounds for lengths two and three for $\ell \geq 2$.

n	$L(n)$	$LB_1^{(\leq 2)}(n)$	$LB_1^{(\leq 3)}(n)$
1	1	1	1
2	2	2	2
3	4	4	4
4	a 6 A	c 8 D	c 8 D
5	a 10 A	d 11 – 16 D	c 15 – 16 D
6	b 16 – 18 B	d 19 – 29 E	c 26 – 32 D
7	b 28 – 32 C	d 32 – 44 F	c 47 – 64 D
8	b 50 – 61 A	d 57 – 88 G	c 86 – 128 D
9	b 91 – 115 C	d 103 – 176 G	d 140 – 256 D

Key to table

a	Lower bound obtained by a computer search.
b	Theorem 7.6
c	Theorem 7.16
d	Theorem 7.14
A	See Appendix
B	Theorem 7.4
C	$L_1^{(\leq \ell)}(n) \leq M_1^{(\leq \ell)}(n)$
D	Example 7.3(iii)
E	See Appendix
F	Theorem 7.20
G	Theorem 7.18

7.4 Optimal linear locating-dominating codes

A binary code $C \subseteq \mathbf{F}^n$ is said to be *linear* if it is a k-dimensional subspace of $\mathbf{F}^n$. In this section we consider linear $(1, \leq \ell)$-LDBs and $(1, \leq \ell)$-LDAs. We denote by $LB_1^{(\leq \ell)}[n]$ (resp. $LA_1^{(\leq \ell)}[n]$) the smallest dimension of a linear $(1, \leq \ell)$-LDB (resp. $(1, \leq \ell)$-LDA) of length n.

Let us first consider the case $\ell = 1$ and denote $L[n] = LA_1^{(\leq 1)}[n] = LB_1^{(\leq 1)}[n]$.

Theorem 7.24.

$$L[n] \geq \lceil n + \log_2 3 - \log_2(n+5) \rceil .$$

Proof. Suppose C is a linear 1-locating-dominating code of length n and $|C| = K$. There can be K non-codewords which are covered by one codeword each. All the other non-codewords must be covered by at least three times. Namely, if $I(\mathbf{x}) = \{\mathbf{c}_1, \mathbf{c}_2\}$ for some $\mathbf{x} \notin C$, then for $\mathbf{y} = \mathbf{x} + \mathbf{c}_1 + \mathbf{c}_2$, $I(\mathbf{y}) = I(\mathbf{x})$ by Lemma 6.1.

Thus, we have $2K + 3(2^n - 2K) \leq K(n+1)$, from which the claim follows. □

Theorem 7.25. *Let $n = 3 \cdot 2^k - 5 + s$, for $k \geq 1$ and $0 \leq s < 3 \cdot 2^k$. Then $L[n] = n - k$.*

Proof. The lower bound follows from Theorem 7.24.

Because $n \geq 3 \cdot 2^k - 5$, we can choose a $(k \times n)$-matrix H in which every nonzero column appears at least three times except for one nonzero column which appears exactly once. Let C be the code with parity check matrix H. We claim that C is a 1-locating-dominating. Each non-codeword which is covered by at least three codewords is identified by Lemma 2.1. Each non-codeword of N_1 is also identified, because for each $\mathbf{c} \in C$, there is exactly one word of N_1 at distance one from it. □

Theorem 7.26. *For all $n \geq 5$,*

$$LB_1^{(\leq 2)}[n] \geq \lceil n + \log_2 5 - \log_2(n+5) \rceil .$$

Proof. Let $C \subseteq \mathbf{F}^n$ be a linear $(1, \leq 2)$-LDB of cardinality K.

If $I(\mathbf{x}) = \{\mathbf{c}\}$ for some $\mathbf{x} \notin C$, then by the proof of (7.4) we have $|I(\mathbf{c})| = n$ and, thus, by Lemma 6.1, $|I(\mathbf{c})| = n$ for all $\mathbf{c} \in C$. Hence, $K \cdot n + (2^n - K) \leq K(n+1)$, i.e., $K \geq 2^{n-1}$, and the claim follows. So assume that $|I(\mathbf{x})| \geq 2$ for all $\mathbf{x} \notin C$. If $|I(\mathbf{x})| = 2$ for some $\mathbf{x} \notin C$, then $I(\mathbf{x}) = \{\mathbf{x} + \mathbf{e}_i, \mathbf{x} + \mathbf{e}_j\}$, for some $\mathbf{e}_i, \mathbf{e}_j \in S_1(\mathbf{0})$, $\mathbf{e}_i \neq \mathbf{e}_j$. Now $I(\mathbf{x}) \subseteq I(\mathbf{x} + \mathbf{e}_i + \mathbf{e}_j)$, which is impossible unless $\mathbf{x} + \mathbf{e}_i + \mathbf{e}_j \in C$. This would imply that $\mathbf{e}_i, \mathbf{e}_j \in C$ and thus, also $\mathbf{x} \in C$. This is impossible by the assumption.

Similarly, as in Theorem 6.7 we can show that each non-codeword must be covered by at least five codewords. Now we have $K + 5(2^n - K) \leq K(n+1)$, from which the claim follows. □

Theorem 7.27. *Let $n = 5(2^k - 1) + s$, for $k \geq 1$ and $0 \leq s < 5 \cdot 2^k$. Then $LB_1^{(\leq 2)}[n] = n - k$.*

Proof. The lower bound follows from Theorem 7.26. Let H be a $(k \times n)$-matrix where every nonzero k-tuple appears at least five times. Let C be the code which has H as its parity check matrix. Every non-codeword is covered at least by five codewords of C. Thus, any set of size at most 2 cannot cover the I-set of a codeword $\mathbf{x}$ unless this set does not include $\mathbf{x}$. This proves the claim. □

Theorem 7.28. *For $\ell \geq 3$ and $n \geq 2\ell - 1$,*

$$LB_1^{(\leq \ell)}[n] \geq \lceil n + \log_2(2\ell - 1) - \log_2(n + 2\ell - 1) \rceil .$$

Proof. Assume that C is a linear $(1, \leq \ell)$-LDB of length n and cardinality K.

If there is a word $\mathbf{x} \notin C$ such that $|I(\mathbf{x})| = 1$, then the beginning of the proof of Theorem 7.26 (since C is trivially also a $(1, \leq 2)$-LDB) shows that $K \geq 2^{n-1}$, and since we have assumed that $n \geq 2\ell - 1$, we are already done.

Assume now that $|I(\mathbf{x})| \geq 2$ for all $\mathbf{x} \notin C$. We shall show that each noncodeword must be covered by at least $2\ell - 1$ codewords. Assume to the contrary that for some $\mathbf{x} \notin C$, $|I(\mathbf{x})| = j$ where $2 \leq j \leq 2\ell - 2$, and denote $I(\mathbf{x}) = \{\mathbf{x} + \mathbf{e}_1, \ldots, \mathbf{x} + \mathbf{e}_j\}$, where $\mathbf{e}_i \in S_1(\mathbf{0})$. For all $1 \leq i \leq j/2$, let $\mathbf{y}_i = \mathbf{x} + \mathbf{e}_{2i-1} + \mathbf{e}_{2i}$ (which is the unique word other than $\mathbf{x}$ at distance one from both $\mathbf{x} + \mathbf{e}_{2i-1}$ and $\mathbf{x} + \mathbf{e}_{2i}$). If j is odd, define $\mathbf{y}_0 = \mathbf{x} + \mathbf{e}_1 + \mathbf{e}_j$. Let Y consists of all the words $\mathbf{y}_i$, $1 \leq i \leq j/2$, together with $\mathbf{y}_0$, if j is odd. Because C is linear and $\mathbf{x} \notin C$, we know that $Y \subseteq \mathbf{F}^n \setminus C$. Now $I(Y) = I(Y \cup \{\mathbf{x}\})$ gives a contradiction, because $|Y \cup \{\mathbf{x}\}| = \lceil j/2 \rceil + 1 \leq \ell$.

Thus, we have $K + (2\ell - 1)(2^n - K) \leq K(n+1)$, and the claim follows. □

Theorem 7.29. *Let $\ell \geq 3$ and $n = (2\ell - 1)(2^k - 1) + s$, for $k \geq 1$ and $0 \leq s < (2\ell - 1)2^k$. Then $LB_1^{(\leq \ell)}[n] = n - k$.*

Proof. The lower bound follows from Theorem 7.28. Let H be a $(k \times n)$-matrix in which every nonzero k-tuple appears at least $2\ell - 1$ times. The code which has H as a parity check matrix is $(1, \leq \ell)$-LDB by Theorem 7.9, and has the required dimension. □

The previous proof in fact gives the following theorem as an immediate corollary.

Theorem 7.30. *Let $\ell \geq 3$. For all $n \geq 2\ell - 1$ we have*

$$LA_1^{(\leq \ell)}[n] = LB_1^{(\leq \ell)}[n].$$

Theorem 7.31. *For all $n \geq 5$, we have*

$$LA_1^{(\leq 2)}[n] = LB_1^{(\leq 2)}[n].$$

Proof. By Theorems 7.27, 7.29 and 7.30 we know that for all $n \geq 5$, $LB_1^{(\leq 2)}[n] = LB_1^{(\leq 3)}[n] = LA_1^{(\leq 3)}[n]$. Trivially, $LA_1^{(\leq 3)}[n] \geq LA_1^{(\leq 2)}[n] \geq LB_1^{(\leq 2)}[n]$, and the claim follows. □

Appendix

In this chapter we list some identifying and locating-dominating codes found using a computer. The 1- and 2-identifying codes are from [18], see also [17]. The $(2, \leq 2)$-identifying codes presented here have not been published before. The locating-dominating codes are from [36] except the one of length 8.

1-identifying codes

- The binary expressions of the numbers in the next table form a 1-identifying code of length 8 and cardinality 62, [17].

5	10	16	17	21	26	27	29	34	35	36
40	42	54	55	64	68	75	76	79	82	89
94	103	105	109	111	112	116	119	121	124	131
134	137	150	152	153	156	162	164	165	174	175
178	181	185	191	192	193	202	205	207	211	212
214	231	232	243	250	251	252				

- The binary expressions of the numbers in the next table form a 1-identifying code C_9 of length 9 and cardinality 115, [17].

4	7	9	14	15	17	19	28	33	36	38
43	52	56	61	65	66	76	79	80	87	90
91	97	105	108	116	118	119	127	133	138	143
145	149	150	159	162	164	168	171	178	184	187
190	195	200	204	208	214	218	221	224	231	237
238	240	241	258	264	269	273	276	277	279	282
288	296	299	307	317	318	319	321	322	326	346
349	357	358	359	362	371	376	378	383	384	385
390	393	410	412	423	430	432	434	435	439	445
446	451	452	455	459	460	467	470	476	477	489
490	493	501	505	508						

- The codewords covered only by themselves in the previous code C_9 are

$$\begin{aligned} A = \ & \{000001001, 000011100, 010001010, 011100111, \\ & 011101110, 100001101, 110000110, 111110101\}. \end{aligned}$$

The set $B = \{100001001, 011100110, 000011000, 000001010, 111000110, 11111000\}$ 1-covers A. Now $(C_9 \oplus \mathbf{F}) \cup (B \oplus \{\mathbf{0}\})$ is 1-identifying of length 10 and cardinality 236.

- The code $C = \{\mathbf{c} \in \mathbf{F}^5 \mid w(\mathbf{c}) \in \{0,1,4,5\}\} \subseteq \mathbf{F}^5$ is a 2-fold 1-covering and 1-identifying (see [40]) and $|C| = 12$. By Theorem 3.10 it is the smallest possible such a code. Applying the $(\pi(\mathbf{u}), \mathbf{u}, \mathbf{u}+\mathbf{v})$-construction of Theorem 3.11 to C we get a 1-identifying code C_{11} of length 11. From C_{11} we can remove words $\{(\pi(\mathbf{u}), \mathbf{u}, \mathbf{u}+\mathbf{v}) \mid \mathbf{u} \in \mathbf{F}^5,\ w(\mathbf{u}) \text{ even}, \mathbf{v} \in \{00000, 11111\}\}$ and the result is still 1-identifying. The cardinality of the code is 352.

- A direct sum of the code C_{11} of length 11 and $\mathbf{F}$ is a 1-identifying code by Theorem 3.20. By a computer search, we found that $(C_{11} \oplus \mathbf{F}) \setminus A$, where A is of size 72, is also 1-identifying. Thus, $M_1(12) \leq 696$. The set A consists of the binary expressions of the numbers in the following table.

0	2	30	31	197	198	216	217	329
330	340	341	396	398	402	403	588	589
593	594	650	651	660	662	774	775	792
794	960	961	989	990	1116	1121	1122	1178
1188	1190	1302	1320	1322	1488	1517	1518	1550
1584	1586	1736	1781	1782	1860	1913	1914	1922
1980	1982	2113	2141	2203	2327	2509	2513	2575
2761	2773	2885	2905	2947	3301	3433	3697	4093

- The code of length 13 made by $(\pi(\mathbf{u}), \mathbf{u}, \mathbf{u}+\mathbf{v})$-construction from the code of length 6 in Example 3.13 can be improved by removing a set B of size 64. Thus, $M_1(13) \leq 1344$. The set B is the binary expressions of the numbers in the following table.

0	14	205	234	331	364	390	392	583
584	644	651	770	781	961	974	1127	1144
1170	1188	1300	1314	1505	1534	1561	1582	1754
1773	1884	1899	1951	1960	2145	2171	2232	2535
2601	2610	2755	2801	2885	2991	3094	3285	3315
3445	3665	3679	3740	3771	3901	4055	4834	4964
5744	6134	6202	6393	6467	6528	6771	6832	6922
7113								

2-identifying codes

- The binary expressions of the numbers 4, 6, 9, 19, 26, 34, 53, 90, 93, 107, 108, 113, 118, 121 form a 2-identifying code of length 7 and cardinality 14, [17].
- The binary expressions of the numbers 7, 9, 26, 36, 55, 63, 64, 85, 107, 114, 144, 163, 174, 185, 205, 211, 220, 222, 232, 244, 246 form a 2-identifying code of length 8 and cardinality 21, [17].
- The binary expressions of the numbers in the next table form a 2-identifying code of length 9 and cardinality 36, [17].

8	9	22	49	60	76	91	98	116	133	152
166	171	189	195	204	222	232	253	270	272	291
301	327	375	377	383	406	411	416	450	466	469
486	493	506								

- The binary expressions of the numbers in the next table form a 2-identifying code $\mathscr{C}_{10}$ of length 10 and cardinality 63.

10	31	36	49	69	92	114	128	159	175	179
195	206	233	244	262	272	330	343	346	355	364
375	377	397	408	430	435	465	485	521	526	531
571	614	616	629	635	657	679	684	690	724	729
733	762	766	782	786	789	813	824	828	834	860
900	923	930	967	969	993	994	1021			

- By adding to $\mathscr{C}_{10}$ the binary expressions of the numbers 26, 62, 132, 205, 451, 467, 620, 624, 704, 718, 793, 1017 and removing 834 we get a 2-identifying code $\mathscr{D}_{10}$ of cardinality 74. The code $\mathscr{D}_{10}$ has the property that for all $\mathbf{x} \in \mathbf{F}^n$ there is $\mathbf{c} \in \mathscr{D}_{10}$ such that $d(\mathbf{x}, \mathbf{c}) = 2$. Theorem 3.20 shows that $\mathscr{D}_{10} \oplus \mathbf{F}$ is 2-identifying of length 11 and cardinality 148.
- By adding to $\mathscr{C}_{10}$ the binary expressions of the numbers 132, 205, 451, 620, 624, 793, 1017 we get a 2-identifying code $\mathscr{E}_{10}$ of cardinality 70. The code $\mathscr{E}_{10}$ has the property that for all $\mathbf{c} \in \mathscr{E}_{10}$ we have $1 \leq d(\mathbf{c}, \mathscr{E}_{10} \setminus \{\mathbf{c}\}) \leq 2$. Theorem 3.20 shows that $\mathscr{E}_{10} \oplus \mathbf{F}^2$ is a 2-identifying code of length 12 and cardinality 280.

$(2, \leq 2)$-identifying codes

- The binary expressions of the next numbers form a $(2, \leq 2)$-identifying code of length 6 and cardinality 25:

1	12	13	21	23	25	26	30	37	38	39	41	42
43	44	45	48	49	51	54	55	56	57	58	59	

- The binary expressions of the following 35 elements subset of the $(1, \leq 2)$-identifying code of length 7 in Theorem 4.14 is $(2, \leq 2)$-identifying.

11	14	18	29	30	37	40	45	51	52	56	59	62
63	64	65	68	75	76	82	87	90	91	92	93	97
98	106	109	111	113	116	118	119	121				

The direct sum of this code and $\mathbf{F}$ is a $(2, \leq 2)$-identifying code of length 8 and cardinality 70.

Locating-dominating codes

- Codes $\{0000, 1100, 0101, 1110, 1011, 0111\} \subseteq \mathbf{F}^4$ and $\{\mathbf{x} \in \mathbf{F}^5 \mid w(\mathbf{x}) = 1 \text{ or } 4\}$ are optimal 1-locating-dominating. The lower bounds are obtained by a computer.

- The 1-identifying code of length 8 and cardinality 62 on page 87 is 1-locating-dominating even if a word 112 is removed from it.

- The binary expressions of the numbers

0	3	6	9	10	13	17	20	23	26	28	29	30
31	35	36	37	39	40	45	46	48	50	51	55	57
58	59	61										

form a $(1, \leq 2)$-LDB of cardinality 29 and length 6.

Bibliography

[1] Y. Ben-Haim and S. Litsyn. Exact minimum density of codes identifying vertices in the square grid. *SIAM J. Discrete Math.*, 19(1):69–82, 2005.

[2] T. Y. Berger-Wolf, W. E. Hart, and J. Saia. Discrete sensor placement problems in distribution networks. *Math. Comput. Modelling*, 42(13):1385–1396, 2005.

[3] N. Bertrand, I. Charon, O. Hudry, and A. Lobstein. Identifying and locating-dominating codes on chains and cycles. *European J. Combin.*, 25(7):969–987, 2004.

[4] N. Bertrand, I. Charon, O. Hudry, and A. Lobstein. 1-identifying codes on trees. *Australas. J. Combin.*, 31:21–35, 2005.

[5] U. Blass, I. Honkala, and S. Litsyn. On binary codes for identification. *J. Combin. Des.*, 8(2):151–156, 2000.

[6] U. Blass, I. Honkala, and S. Litsyn. Bounds on identifying codes. *Discrete Math.*, 241(1-3):119–128, 2001.

[7] I. Charon, S. Gravier, O. Hudry, A. Lobstein, M. Mollard, and J. Moncel. A linear algorithm for minimum 1-identifying codes in oriented trees. *Discrete Appl. Math.*, 154(8):1246–1253, 2006.

[8] I. Charon, I. Honkala, O. Hudry, and A. Lobstein. General bounds for identifying codes in some infinite regular graphs. *Electron. J. Combin.*, 8(1):Research Paper 39, 21 pp., 2001.

[9] I. Charon, O. Hudry, and A. Lobstein. Identifying and locating-dominating codes: NP-completeness results for directed graphs. *IEEE Trans. Inform. Theory*, 48(8):2192–2200, 2002.

[10] I. Charon, O. Hudry, and A. Lobstein. Identifying codes with small radius in some infinite regular graphs. *Electron. J. Combin.*, 9(1):Research Paper 11, 25 pp., 2002.

[11] I. Charon, O. Hudry, and A. 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.

[12] I. Charon, O. Hudry, and A. Lobstein. Extremal cardinalities for identifying and locating-dominating codes in graphs. *Discrete Math.*, 307(3-5):356–366, 2007.

[13] G. Cohen, I. Honkala, S. Litsyn, and A. Lobstein. *Covering Codes*. Elsevier, Amsterdam, 1997.

[14] G. Cohen, I. Honkala, A. Lobstein, and G. Zémor. New bounds for codes identifying vertices in graphs. *Electron. J. Combin.*, 6:Research Paper 19, 14 pp., 1999.

[15] G. D. Cohen, I. Honkala, A. Lobstein, and G. Zémor. On identifying codes. In A. Barg and S. Litsyn, editors, *Proceedings of the DIMACS Workshop on Codes and Association Schemes*, volume 56, pages 97–109, AMS, Providence, 2001.

[16] C. J. Colbourn, P. J. Slater, and L. K. Stewart. Locating dominating sets in series parallel networks. *Congr. Numer.*, 56:135–162, 1987. Sixteenth Manitoba conference on numerical mathematics and computing (Winnipeg, Man., 1986).

[17] G. Exoo. Computational results on identifying t-codes. Preprint.

[18] G. Exoo, T. Laihonen, and S. Ranto. Improved upper bounds on binary identifying codes. *IEEE Trans. Inform. Theory*, submitted.

[19] A. Frieze, R. Martin, J. Moncel, M. Ruszinkó, and C. Smyth. Codes identifying sets of vertices in random networks. *Discrete Math.*, 307(9-10):1094–1107, 2007.

[20] D. Garnick and N. A. Nieuwejaar. Nonisomorphic extremal graphs without three-cycles or four-cycles. *J. Combin. Math. Combin. Comput.*, 12:33–56, 1992.

[21] D. K. Garnick, Y. H. H. Kwong, and F. Lazebnik. Extremal graphs without three-cycles or four-cycles. *J. Graph Theory*, 17:633–645, 1993.

[22] S. Gravier, J. Moncel, and A. Semri. Identifying codes of cycles. *European Journal of Combinatorics*, 27(5):767–776, 2006.

[23] H. O. Hämäläinen, I. S. Honkala, M. K. Kaikkonen, and S. N. Litsyn. Bounds for binary multiple covering codes. *Des. Codes Cryptogr.*, 3:251–275, 1993.

[24] I. Honkala. On the identifying radius of codes. In I. Honkala and T. Harju, editors, *Proceedings of the Seventh Nordic Combinatorial Conference*, TUCS General Publications, pages 33–43, Turku, 1999.

[25] I. Honkala. An optimal locating-dominating set in the infinite triangular grid. *Discrete Math.*, 306(21):2670–2681, 2006.

[26] I. Honkala, M. G. Karpovsky, and L. B. Levitin. On robust and dynamic identifying codes. *IEEE Trans. Inform. Theory*, 52(2):599–611, 2006.

[27] I. Honkala and T. Laihonen. On identification in the triangular grid. *J. Combin. Theory Ser. B*, 91(1):67–86, 2004.

[28] I. Honkala and T. Laihonen. On identifying codes in the triangular and square grids. *SIAM J. Comput.*, 33(2):304–312, 2004.

[29] I. Honkala and T. Laihonen. On locating-dominating sets in infinite grids. *European J. Combin.*, 27(2):218–227, 2006.

[30] I. Honkala and T. Laihonen. On a new class of identifying codes in graphs. *Inform. Process. Lett.*, 102(2-3):92–98, 2007.

[31] I. Honkala and T. Laihonen. On identifying codes that are robust against edge changes. *Inform. and Comput.*, to appear.

[32] I. Honkala and T. Laihonen. On vertex-robust identifying codes of level three. *Ars Combin.*, to appear.

[33] I. Honkala, T. Laihonen, and S. Ranto. Codes for strong identification. In *Electronic Notes in Discrete Mathematics*, volume 6, pages 279–287, 2001. Proceedings of International Workshop on Coding and Cryptography.

[34] I. Honkala, T. Laihonen, and S. Ranto. On codes identifying sets of vertices in Hamming spaces. *Des. Codes Cryptogr.*, 24(2):193–204, 2001.

[35] I. Honkala, T. Laihonen, and S. Ranto. On strongly identifying codes. *Discrete Math.*, 254(1-3):191–205, 2002.

[36] I. Honkala, T. Laihonen, and S. Ranto. On locating-dominating codes in binary Hamming spaces. *Discrete Math. Theor. Comput. Sci.*, 6(2):265–281, 2004.

[37] I. Honkala and A. Lobstein. On identifying codes in binary Hamming spaces. *J. Combin. Theory Ser. A*, 99(2):232–243, 2002.

[38] I. Honkala and A. Lobstein. On the complexity of the identification problem in Hamming spaces. *Acta Inform.*, 38(11-12):839–845, 2002.

[39] I. Honkala and A. Lobstein. On the density of identifying codes in the square lattice. *J. Combin. Theory Ser. B*, 85(2):297–306, 2002.

[40] M. G. Karpovsky, K. Chakrabarty, and L. B. Levitin. On a new class of codes for identifying vertices in graphs. *IEEE Trans. Inform. Theory*, 44(2):599–611, 1998.

[41] M. G. Karpovsky, K. Chakrabarty, L. B. Levitin, and D. R. Avresky. On the covering of vertices for fault diagnosis in hypercubes. *Inform. Process. Lett.*, 69(2):99–103, 1999.

[42] M. Laifenfeld and A. Trachtenberg. Disjoint identifying-codes for arbitrary graphs. In *Proceedings of International Symposium on Information Theory, 2005. ISIT 2005*, pages 244–248, 2005.

[43] M. Laifenfeld, A. Trachtenberg, and T. Y. Berger-Wolf. Identifying codes and the set cover problem. In *Proceedings of the 44th Annual Allerton Conf. on Communication, Control and Computing*, Monticello, USA, 2006.

[44] T. Laihonen. Optimal codes for strong identification. *European J. Combin.*, 23(3):307–313, 2002.

[45] T. Laihonen. Sequences of optimal identifying codes. *IEEE Trans. Inform. Theory*, 48(3):774–776, 2002.

[46] T. Laihonen. On optimal edge-robust and vertex-robust $(1, \leq l)$-identifying codes. *SIAM J. Discrete Math.*, 18(4):825–834, 2005.

[47] T. Laihonen and S. Ranto. Codes identifying sets of vertices. In *Applied algebra, algebraic algorithms and error-correcting codes (Melbourne, 2001)*, volume 2227 of *Lecture Notes in Computer Science*, pages 82–91. Springer, Berlin, 2001.

[48] T. Laihonen and S. Ranto. Families of optimal codes for strong identification. *Discrete Appl. Math.*, 121(1-3):203–213, 2002.

[49] T. Laihonen and S. Ranto. Codes identifying sets of binary words with large radii. In *Proceedings of International Workshop on Coding and Cryptography WCC 2007*, to appear.

[50] A. Lobstein. Identifying and locating-dominating codes in graphs, a bibliography. Published electronically at `http://perso.enst.fr/∼lobstein/debutBIBidetlocdom.pdf`.

[51] J. Moncel. Constructing codes identifying sets of vertices. *Des. Codes Cryptogr.*, 41(1):23–31, 2006.

[52] J. Moncel. Monotonicity of the minimum cardinality of an identifying code in the hypercube. *Discrete Appl. Math.*, 154(6):898–899, 2006.

[53] E. M. Rains and N. J. A. Sloane. Table of constant weight binary codes. Published electronically at `http://www.research.att.com/∼njas/codes/Andw/`.

[54] D. F. Rall and P. J. Slater. On location–domination numbers for certain classes of graphs. *Congr. Numer.*, 45:97–106, 1984.

[55] S. M. Ranto. Optimal linear identifying codes. *IEEE Trans. Inform. Theory*, 49(6):1544–1547, 2003.

[56] S. M. Ranto, I. S. Honkala, and T. K. Laihonen. Two families of optimal identifying codes in binary Hamming spaces. *IEEE Trans. Inform. Theory*, 48(5):1200–1203, 2002.

[57] S. Ray, D. Starobinski, A. Trachtenberg, and R. Ungrangsi. Robust location detection with sensor networks. *IEEE Journal on Selected Areas in Comminications (Special Issue on Fundamental Performance Limits of Wireless Sensor Networks)*, vi. 22, 2004.

[58] P. J. Slater. Domination and loation in graphs. Research raport 93, National University of Singapore, 1983.

[59] P. J. Slater. Domination and location in acyclic graphs. *Networks*, 17(1):55–64, 1987.

[60] P. J. Slater. Dominating and reference sets in a graph. *J. Math. Phys. Sci.*, 22:445–455, 1988.

[61] P. J. Slater. Locating dominating sets and locating-dominating sets. In *Graph Theory, Combinatories and Applications: Proceedings of the Seventh Quadrennial International Conference on the Theory and Applications of Graphs*, volume 2, pages 1073–1079. Wiley, 1995.

[62] P. J. Slater. Fault-tolerant locating-dominating sets. *Discrete Math.*, 249(1–3):179–189, 2002.

[63] N. J. A. Sloane. The on-line encyclopedia of integer sequences. Published electronically at `http://www.research.att.com/∼ njas/sequences`.

[64] R. Struik. *Covering Codes*. Ph. D. thesis, Eindhoven University of Technology, the Netherlands, 1994.

Turku Centre for Computer Science
TUCS Dissertations

51. **Attila Gyenesei**, Discovering Frequent Fuzzy Patterns in Relations of Quantitative Attributes
52. **Petteri Kaitovaara**, Packaging of IT Services – Conceptual and Empirical Studies
53. **Petri Rosendahl**, Niho Type Cross-Correlation Functions and Related Equations
54. **Péter Majlender**, A Normative Approach to Possibility Theory and Soft Decision Support
55. **Seppo Virtanen**, A Framework for Rapid Design and Evaluation of Protocol Processors
56. **Tomas Eklund**, The Self-Organizing Map in Financial Benchmarking
57. **Mikael Collan**, Giga-Investments: Modelling the Valuation of Very Large Industrial Real Investments
58. **Dag Björklund**, A Kernel Language for Unified Code Synthesis
59. **Shengnan Han**, Understanding User Adoption of Mobile Technology: Focusing on Physicians in Finland
60. **Irina Georgescu**, Rational Choice and Revealed Preference: A Fuzzy Approach
61. **Ping Yan**, Limit Cycles for Generalized Liénard-type and Lotka-Volterra Systems
62. **Joonas Lehtinen**, Coding of Wavelet-Transformed Images
63. **Tommi Meskanen**, On the NTRU Cryptosystem
64. **Saeed Salehi**, Varieties of Tree Languages
65. **Jukka Arvo**, Efficient Algorithms for Hardware-Accelerated Shadow Computation
66. **Mika Hirvikorpi**, On the Tactical Level Production Planning in Flexible Manufacturing Systems
67. **Adrian Costea**, Computational Intelligence Methods for Quantitative Data Mining
68. **Cristina Seceleanu**, A Methodology for Constructing Correct Reactive Systems
69. **Luigia Petre**, Modeling with Action Systems
70. **Lu Yan**, Systematic Design of Ubiquitous Systems
71. **Mehran Gomari**, On the Generalization Ability of Bayesian Neural Networks
72. **Ville Harkke**, Knowledge Freedom for Medical Professionals – An Evaluation Study of a Mobile Information System for Physicians in Finland
73. **Marius Cosmin Codrea**, Pattern Analysis of Chlorophyll Fluorescence Signals
74. **Aiying Rong**, Cogeneration Planning Under the Deregulated Power Market and Emissions Trading Scheme
75. **Chihab BenMoussa**, Supporting the Sales Force through Mobile Information and Communication Technologies: Focusing on the Pharmaceutical Sales Force
76. **Jussi Salmi**, Improving Data Analysis in Proteomics
77. **Orieta Celiku**, Mechanized Reasoning for Dually-Nondeterministic and Probabilistic Programs
78. **Kaj-Mikael Björk**, Supply Chain Efficiency with Some Forest Industry Improvements
79. **Viorel Preoteasa**, Program Variables – The Core of Mechanical Reasoning about Imperative Programs
80. **Jonne Poikonen**, Absolute Value Extraction and Order Statistic Filtering for a Mixed-Mode Array Image Processor
81. **Luka Milovanov**, Agile Software Development in an Academic Environment
82. **Francisco Augusto Alcaraz Garcia**, Real Options, Default Risk and Soft Applications
83. **Kai K. Kimppa**, Problems with the Justification of Intellectual Property Rights in Relation to Software and Other Digitally Distributable Media
84. **Dragoş Truşcan**, Model Driven Development of Programmable Architectures
85. **Eugen Czeizler**, The Inverse Neighborhood Problem and Applications of Welch Sets in Automata Theory
86. **Sanna Ranto**, Identifying and Locating-Dominating Codes in Binary Hamming Spaces

Turku
Centre *for*
Computer
Science

Joukahaisenkatu 3-5 B, 20520 Turku, Finland | www.tucs.fi

University of Turku

- Department of Information Technology
- Department of Mathematics

Åbo Akademi University

- Department of Information Technologies

Turku School of Economics

- Institute of Information Systems Sciences

ISBN 978-952-12-1880-4
ISSN 1239-1883