# Covering Codes

René Struik

# Covering Codes

# Covering Codes

## Proefschrift

ter verkrijging van de graad van doctor aan de Technische Universiteit Eindhoven, op gezag van de Rector Magnificus, prof.dr. J.H. van Lint, voor een commissie aangewezen door het College van Dekanen in het openbaar te verdedigen op dinsdag 25 oktober 1994 om 16.00 uur

door

**Marinus Struik**

geboren te Rotterdam

Dit proefschrift is goedgekeurd door de promotoren
prof.dr. J.H. van Lint
en
prof.dr.ir. H.C.A. van Tilborg

*To the memory of my father*

# Acknowledgements

I would like to express my gratitude to my supervisors, Prof.dr. J.H. van Lint and Prof.dr. H.C.A. van Tilborg, who gave me the opportunity to carry out my doctoral research in the Discrete Mathematics group of the Eindhoven University of Technology. I enjoyed very much working in their group. I would like to thank them for their advice, encouragement, and patience. I would also like to thank them for the enthusiasm with which they taught their courses in combinatorics, coding theory, and cryptology. It was by these courses, and their enthusiasm, that I became interested in discrete mathematics.

Special thanks go to Prof.dr. G.D. Cohen, Prof.dr. H.F. Mattson Jr., Dr. P. Solé, and, of course, my supervisors, for their careful reading of (parts of) my thesis. Their suggestions and constructive comments proved most valuable.

I thank all the members of the Discrete Mathematics group for providing a pleasant and stimulating working environment.

Finally, I would like to thank the Netherlands Organization for Scientific Research (NWO) for its financial support.

# Contents

# Introduction

Coding theory studies ways to improve the reliability of information transmission between electronic devices, such as computers. If two devices exchange digital information, part of the data can be corrupted due to imperfections of the channel, i.e. transmission errors can occur (the so-called noise). As a result, the information that is received differs from the information that is sent. If one sends messages as such, channel errors would have a great impact on the quality of the received signal. In order to prevent a degradation of the received information, so-called redundancy is built into the signal, i.e. the transmitted sequence consists of more than the necessary information. Thus, even in the presence of noise, encoded messages are easily distinguishable from each other. A well-known example of messages with built-in redundancy is everyday language. The words of our language are only a small portion of all possible strings of letters. Consequently, a misprint in a (long) word results into a string that resembles the original word more than any other word we know. Hence, misprints and spelling flaws can be easily recognized and corrected. This example illustrates that, by using redundancy, encoded messages can be interpreted in the right way, even in the presence of occasional errors. Another example where the introduction of redundancy improves the quality of information transmission is the repetition code. Here, one transmits every symbol of a message three times. At reception of the encoded message, one tries to recover each message symbol by comparing its three transmissions and interpreting the symbol that occurs most of the time (i.e. at least twice) as the symbol sent. If at most one in every three symbols is corrupted by channel errors, then this encoding procedure guarantees a perfect recovery of the original message. In general, encoding schemes such as the ones described above are called error-correcting codes. Designing error-correcting codes that allow a high quality of data transmission with as few redundant symbols as possible is the main subject of coding theory.

Error-correcting codes add redundancy to a signal, so as to allow a perfect recovery of information. Sometimes, however, it is not necessary to achieve perfect recovery of a signal. At first sight, this seems undesirable. But, there are many situations where a limited degradation of the signal is acceptable. An important example is digital quantization, a technique of approximating analog data by a fixed set of signals. This technique is applied in digital signal processing of audio signals or video signals. Here, samples of analog data are represented by suitably chosen reference values. These values are chosen from a discrete set and can, therefore, be represented digitally. Another example is the communication of speech via a telephone network. Here, the signal properties of the received speech and that of the original differ considerably. This degradation is, however, acceptable for voice communication. The examples described above have in common that arbitrary signals are approximated. This approximation process allows a shorter representation of data, at the expense of a limited loss of accuracy. In general, the approximation schemes described above are called (digital) quantizers. Obviously, one wants to design a digital quantizer in such a way that the maximum (or average) distortion of analog signals is minimized. The quality of the approximation depends on the choice of the codewords and on the function that maps the data to these codewords.

Covering codes can be viewed as quantizers for digital data. Each digital sequence is approximated by a codeword from the covering code that agrees with this sequence in as many positions as possible. This approximation process allows a shorter description of data, at the expense of a limited loss of accuracy. The maximum number of errors that one makes in this approximation process (the maximum distortion) is called the covering radius of the code. It should be noted that, in contrast with lossless data compression, where shorter descriptions of data are obtained by removing redundancy, the approximation of data using a covering code leads to lossy data compression, since it is in general not possible to reconstruct the data completely.

As an illustration, we now describe a well-known problem from recreational mathematics, viz. the so-called football pool problem, and show its connection to covering codes. Assume $n$ football matches, which have not been played yet. Each match has three possible outcomes (win, loose, draw), so in total there are $3^n$ possible outcomes of these matches. Obviously, there are also $3^n$ possible forecasts, i.e. lists containing the $n$ predicted outcomes. In many countries, it is a popular game to bet on the outcome of these $n$ matches. People may turn in any number of forecasts. Each forecast costs a fixed amount of money. After the matches have been played, part of the money raised in this way is returned to those participants who submitted a forecast with none or only a few errors. The football pool problem can now be described as follows: what is the minimum number of forecasts one has to submit in order to be sure that, no matter what the outcomes of the matches will be, one is guaranteed to have at least one forecast of these $n$ matches that contains at most one error? In terms of coding theory: design a covering code with covering radius one and as few codewords as possible.

Error-correcting codes and covering codes are designed for different purposes: the first one provides improved reliability, the second one provides a shorter description. Despite their different application area, there are many connections between these two types of codes. To illustrate these connections, we give an example of a code which can be used in either way. Suppose one wants to transmit binary 4-tuples over a noisy channel. A channel error causes a 0 to be interpreted as a 1, and vice versa. If the 4-tuples are sent as such, then channel errors will greatly affect the reliability of the information. To allow correction of a single error, the 4-tuples are encoded to 7-tuples by adding 3 redundant symbols at the end, according to the following table:

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 0: | (0000,000), | 4: | (0100,101), | 8: | (1000,110), | 12: | (1100,011), |
| 1: | (0001,111), | 5: | (0101,010), | 9: | (1001,001), | 13: | (1101,100), |
| 2: | (0010,011), | 6: | (0110,110), | 10: | (1010,101), | 14: | (1110,000), |
| 3: | (0011,100), | 7: | (0111,001), | 11: | (1011,010), | 15: | (1111,111). |

It is not hard to see that the 7-tuples of the table, which are called codewords, differ from each other in at least three positions. Therefore, a single transmission error can be corrected. As an example, the message 1001 represents the number 9 and is therefore encoded as the string 1001001. If the 7-tuple 1101001 is received, then this string resembles the codeword 1001001 more than any other codeword, and is therefore decoded accordingly. The original message 1001 is obtained via a simple table-lookup (or by ignoring the last three bits). A property of the code is that, no matter which 7-tuple is received, there is always exactly one codeword that differs from this 7-tuple in at most one position. Therefore, the code is also a covering code with covering radius one. Hence, by approximating a 7-tuple by a codeword — thereby introducing at most one error — one can describe this 7-tuple by a 4-tuple.

This example demonstrates that there are many relations between error-correcting codes and covering codes. Both types of codes involve a number of basic parameters. A code of length $n$ is a collection of $n$-tuples. The cardinality of a code is its number of elements. A code has minimum distance $d$, if any two distinct codewords differ in at least $d$ positions. A code has covering radius $r$, if any arbitrary $n$-tuple differs from at least one codeword in at most $r$ positions. So, the code from the previous example is a binary code of length $n{=}7$ with cardinality $M{=}16$, minimum distance $d{=}3$, and covering radius $r{=}1$. It is clear that changing a codeword on less than $d/2$ positions results in a word that resembles the original codeword more than any other codeword; therefore, it can be interpreted correctly. Hence, the minimum distance $d$ measures the error-correcting capability of the code. As already indicated, the covering radius $r$ measures the maximum distortion that occurs when an arbitrary $n$-tuple is approximated by a suitable codeword.

When designing codes, two aspects are important, viz. the quality of the code and its implementation cost. Error-correcting codes add redundancy to a signal. This redundancy is minimal if the code has as many codewords as possible, given its length and minimum distance. Covering codes provide a shorter description of data. This description has minimal length if the code has as few codewords as possible, given its length and covering radius. Therefore, the information rate, i.e. the extent to which data is expanded, resp. reduced, determines the quality of both types of codes. For error-correcting codes the information rate should be as large as possible, for covering codes it should be as small as possible. Encoding/decoding are usually to be performed by small electronic devices with limited memory size. Therefore, the code's implementation cost — both in memory terms and in terms of encoding/decoding efficiency — is of practical importance. These two aspects, viz. quality and implementation's efficiency, can be conflicting. Many record-breaking codes have been found using heuristic search algorithms such as simulated annealing. Although these codes have a good quality, their structure is usually poor or absent, whereas structure is essential for an efficient implementation. In general, a systematic and constructive approach to designing codes offers the perspective of achieving both aims at the same time, since it usually results in codes with a lot of (mathematical) structure.

In this thesis we discuss covering codes. For these codes, we present bounds as well as constructions. Usually, we assume that the design parameters of a covering code, i.e. its length $n$ and its covering radius $r$, are fixed. The bounds we give are then lower bounds on the cardinality $K(n,r)$ of any covering code with length $n$ and covering radius $r$. Equivalently, we could have studied the minimum covering radius $r$ that can be attained by any covering code with fixed length $n$ and fixed information rate $k/n$. Where possible, we have tried to give connections between covering codes, error-correcting codes, and other structures from discrete mathematics. We followed two approaches. In the first approach, we took bounds or constructions that are known for error-correcting codes as a starting point for studying similar bounds and constructions for covering codes. In the second approach, we followed the other direction of research, mainly to show that many results from the literature have analogues in the theory of error-correcting codes. Our approach shows that there are intricate relations between covering codes and error-correcting codes.

## Overview

In Chapter 1 we review some notions from coding theory that will be used throughout the rest of this thesis. Moreover, we show how many well-known results from coding theory can be obtained in a uniform way. The upper bounds on the covering radius of a code obtained by Tietäväinen and Delsarte follow as a special case.

In Chapter 2 we discuss a number of lower bounds on the size of binary covering codes with a prescribed length and covering radius. It is shown that most of the lower bounds for covering codes can be described as direct analogues of the well-known Johnson bound for error-correcting codes. The most important result of the chapter is an improvement of the Van Wee bound for binary linear codes. Many other results in this chapter highlight relations between bounds for covering codes and bounds for error-correcting codes. Some of the result of this chapter can also be found in [80].

In Chapter 3 we present lower bounds on linear covering codes by studying their dual structure. We show how the parameters of a linear covering code impose restrictions on the form of the dual code. In particular, we obtain restrictions on both the weight distribution and the intersection of different words in the dual code. Using these restrictions, and coding theory, we show that many sparse linear covering codes simply cannot exist. In particular, we prove a conjecture by Brualdi, Pless, and Wilson. Part of the results of this chapter can also be found in [79].

In Chapter 4 we consider constructions for sparse covering codes. Starting point is a generalization of the direct sum construction that has proved to produce good results for error-correcting codes. We show that this generalization can also be used to obtain good covering codes, i.e. covering codes with relatively few codewords. Although part of the results are not new, their proof and description mostly is. The most important new result of this chapter is a simple construction of an infinite sequence of covering codes with covering radius two, minimum distance four, and density approaching 1, thus improving the previously best known construction, which has density 9/8. Another result is a generalization of the concept of normal code, which allows many simplified proofs of properties of these codes.

# Chapter 1

# Coding Theory — A Quick Review

## 1.1 Introduction

In this chapter we quickly review some notions from coding theory that will be used throughout the rest of this thesis. The chapter is not intended to be an introduction to coding theory: we merely give definitions and leave out the underlying motivation for the different concepts. For these, we refer to one of the many good textbooks on coding theory, see e.g. [58, 64]. The main concepts are treated in a fairly concise way. Yet, we have taken care to make the chapter mostly self-contained. The chapter is organized as follows. In Section 1.2 we introduce the basic concepts that play a role in the study of codes. This section provides the background needed in order to be able to understand the remaining chapters. In the rest of the chapter we quickly discuss some of the more advanced topics in coding theory. In Section 1.3 we give some properties of Krawtchouk polynomials. These polynomials play an essential role in Section 1.4, which introduces weight enumerators and the duality theorem, and in Section 1.5, which considers relations between the dual distance, the weights in a code, and the covering radius. Almost all results obtained in this chapter are well-known. The main contribution of Section 1.5 is a uniform description of these results. Except for the first section, we restrict ourselves to binary codes.

## 1.2 Basic Concepts

In this thesis we mainly consider codes over the binary field $\mathbb{F}_2 = \{0, 1\}$. In this section we give definitions over $\mathbb{F}_q$, the field with $q$ elements. We adopt the notations of [58, 64]. The set of all $n$-tuples over the field $\mathbb{F}_q$ forms a vector space denoted by $\mathbb{F}_q^n$. The Hamming distance $d(\mathbf{x}, \mathbf{y})$ between two words $\mathbf{x}, \mathbf{y} \in \mathbb{F}_q^n$ is defined by $d(\mathbf{x}, \mathbf{y}) := |\{i \mid x_i \neq y_i\}|$. The weight $wt(\mathbf{x})$ of a word $\mathbf{x} \in \mathbb{F}_q^n$ is defined by $wt(\mathbf{x}) := d(\mathbf{x}, \mathbf{0})$. The support of a word $\mathbf{x} \in \mathbb{F}_q^n$ is defined by $\operatorname{supp}(\mathbf{x}) := \{i \mid x_i \neq 0\}$. The support of a set of words is the union of the supports of its elements.

A $q$-ary code $C$ of length $n$ is a nonempty subset of $\mathbb{F}_q^n$. If $C$ has cardinality $M$, $C$ is called an $(n, M)$ code. A linear code of length $n$ is a linear subspace of $\mathbb{F}_q^n$. If $C$ is a subspace of dimension $k$, then $\mathcal{C}$ is called an $[n,k]$ code. Its translates are called cosets of the linear code. The vectors of minimum weight in such a coset are called coset leaders. If $\mathcal{C}$ is an $[n,k]$ code, then its dual code $\mathcal{C}^\perp$ is the $[n,n-k]$ code defined by

$$\mathcal{C}^\perp := \{\mathbf{y} \in \mathbb{F}_q^n \mid \langle \mathbf{x}, \mathbf{y} \rangle = 0 \text{ for all } \mathbf{x} \in \mathcal{C}\},$$

where $\langle \mathbf{x}, \mathbf{y} \rangle$ denotes the standard inner product of $\mathbf{x}$ and $\mathbf{y}$.
A generator matrix $G$ for an $[n,k]$ code $\mathcal{C}$ is a $k \times n$ matrix for which the rows are a basis of $\mathcal{C}$. A generator matrix $H$ for its dual code $\mathcal{C}^\perp$ is called a parity check matrix of $\mathcal{C}$. Thus codewords of $\mathcal{C}$ are characterized by

$$\mathbf{x} \in \mathcal{C} \quad \Leftrightarrow \quad \mathbf{x}H^T = \mathbf{0}.$$

If $\mathcal{C}$ is a $q$-ary $[n,k]$ code with parity check matrix $H$, then for every word $\mathbf{x} \in \mathbb{F}_q^n$ the vector $\mathbf{x}H^T$ is called the syndrome of $\mathbf{x}$.
The minimum distance $d$ of a code $\mathcal{C}$ is the minimum value of $d(\mathbf{x}, \mathbf{y})$ over all pairs of different codewords $\mathbf{x}, \mathbf{y} \in \mathcal{C}$. The covering radius $r$ of a code $\mathcal{C}$ is the maximum value of $d(\mathbf{x}, \mathcal{C})$ over all words $\mathbf{x} \in \mathbb{F}_q^n$. Here $d(\mathbf{x}, \mathcal{C})$ is defined by $d(\mathbf{x}, \mathcal{C}) := \min\{d(\mathbf{x}, \mathbf{c}) \mid \mathbf{c} \in \mathcal{C}\}$. For linear codes, the covering radius is the highest weight of any coset leader of the code. Let $\mathcal{C}$ be a code of length $n$ with covering radius $r$. We say that a word $\mathbf{x} \in \mathbb{F}_q^n$ is covered by the codeword $\mathbf{c} \in \mathcal{C}$, if $d(\mathbf{x}, \mathbf{c}) \leq r$. For linear codes, we say that a syndrome is covered, if any word with this syndrome is covered. The sphere $B_s(\mathbf{x})$ of radius $s$ around $\mathbf{x} \in \mathbb{F}_q^n$ is defined by $B_s(\mathbf{x}) := \{\mathbf{y} \in \mathbb{F}_q^n \mid d(\mathbf{x}, \mathbf{y}) \leq s\}$. The cardinality of this sphere is denoted by $V_q(n,s)$. (If $q$=2, we simply write $V(n,s)$.) From the definition of the covering radius it is immediately clear that a code $\mathcal{C}$ in $\mathbb{F}_q^n$ has covering radius $r$, if for every word $\mathbf{x} \in \mathbb{F}_q^n$ there exists at least one codeword $\mathbf{c} \in \mathcal{C}$ such that $\mathbf{x} \in B_r(\mathbf{c})$.
Sometimes, it is useful to consider the operations extending, puncturing, and shortening. We will only define these operations for binary codes, so $q$=2. For any code $\mathcal{C}$ of length $n$, its extended code $\overline{\mathcal{C}}$ is defined by

$$\overline{\mathcal{C}} := \{(c_1, \ldots, c_n, \sum_{i=1}^{n} c_i) \mid (c_1, \ldots, c_n) \in \mathcal{C}\}.$$

Its punctured code $\mathcal{C}[i]$ is the code one obtains by deleting the $i$th coordinate from every codeword of $\mathcal{C}$. Its shortened code is the code one obtains by considering all codewords that end on the same position and subsequently deleting this last position.

When we refer to the dual distance, dual covering radius, etc., of a linear code $\mathcal{C}$, we consider the respective parameters of its dual code $\mathcal{C}^\perp$. When we have some additional information on the code, we can add this to the parameter description of the code. For example, an $[n,k,d]r$ code denotes an $[n,k]$ code with minimum distance $d$ and covering radius $r$, an $[n,k;W]$ code denotes an $[n,k]$ code for which all codewords have weights in the set $W$, an $(n,M)r$ code denotes an $(n,M)$ code with covering radius $r$, and so on. Sometimes we use the function $d[n,k]$, which denotes the largest achievable minimum distance for any $[n,k]$ code. In the rest of this chapter, as in most of this thesis, we restrict ourselves to binary codes. For these codes, a table of bounds for $d[n,k]$ is provided in [7].

## 1.3 Krawtchouk polynomials

In this section we introduce a sequence of orthogonal polynomials, the so-called Krawtchouk polynomials. These polynomials play an important role in several parts of coding theory. We mention several properties of these polynomials that will be used in the rest of this chapter. For details we refer to [58, Section 1.2].

For $k = 0, 1, 2, \ldots$ the Krawtchouk polynomial $K_k(x; n)$ is defined by

$$K_k(x;n) := \sum_{j=0}^{k} (-1)^j \binom{x}{j}\binom{n-x}{k-j}, \quad \text{where } x \in \mathbb{R}. \tag{1.1}$$

If the parameter $n$ is clear from context, then we simply write $K_k(x)$ instead of $K_k(x;n)$. Notice that $K_k(n-x;n) = (-1)^k K_k(x;n)$.

From Equation (1.1) it follows directly that

$$\sum_{\mathbf{y}:wt(\mathbf{y})=k} (-1)^{\langle \mathbf{x},\mathbf{y}\rangle} = K_k(i;n) \quad \text{if } \mathbf{x} \in \mathbb{F}_2^n \text{ has weight } i. \tag{1.2}$$

Equation (1.2) proves to be useful in the rest of this chapter. We will also need certain relations between the Krawtchouk polynomials and some information on the locations of the zeros of these polynomials.

Krawtchouk polynomials satisfy certain orthogonality relations which prove to be useful in coding theory.
It is clear from the definition that $K_k(x;n)$ is a polynomial of degree $k$ in variable $x$. The Krawtchouk polynomials satisfy the following relation:

$$\sum_{i=0}^{n} \binom{n}{i} K_k(i) K_l(i) = \delta_{kl} \binom{n}{k} 2^n. \tag{1.3}$$

It follows, that the polynomials $\{K_k(x;n)\}_{k=0}^{n}$ form an orthogonal basis of the vector space of all polynomials in $\mathbb{R}[x]$ of degree at most $n$ with inner product

$$\langle f(x), g(x)\rangle_n := \sum_{i=0}^{n} \binom{n}{i} f(i) g(i). \tag{1.4}$$

In the rest of the chapter we will make extensive use of this orthogonality relation. For later use we mention that a simple calculation shows that

$$\langle (n-x) f(x), g(x)\rangle_n = n \langle f(x), g(x)\rangle_{n-1}. \tag{1.5}$$

Apart from orthogonality relation (1.3), the Krawtchouk polynomials also satisfy another kind of orthogonality relation, viz.

$$\sum_{i=0}^{n} K_k(i) K_i(l) = \delta_{kl} 2^n. \tag{1.6}$$

Krawtchouk polynomials have real zeros which satisfy an interlacing property.
The polynomial $K_k(x;n)$ has $k$ distinct zeros in the interval $(0,n)$. These zeros are symmetrical with respect to $\frac{1}{2}n$, since $K_k(n-x;n) = (-1)^k K_k(x;n)$. If $x_1 < \cdots < x_k$ are the zeros of $K_k(x;n)$ and if $y_1 < \cdots < y_{k-1}$ are the zeros of $K_{k-1}(x;n)$, then the zeros have the following interlacing property:

$$0 < x_1 < y_1 < x_2 < \cdots < x_{k-1} < y_{k-1} < x_k < n. \tag{1.7}$$

In particular we find that if $x_{k,n}^{(1)}$ is the smallest zero of $K_k(x;n)$, then

$$x_{k,n-1}^{(1)} < x_{k,n}^{(1)} < x_{k-1,n-1}^{(1)}. \tag{1.8}$$

(All these results can be easily proved by induction on $k+n$, using Equation (1.1) and the identity $K_k(x;n) = K_k(x;n-1) + K_{k-1}(x;n-1)$, which is implied by this equation.)
In general the exact location of the zeros of $K_k(x;n)$ is not known, but asymptotically it is known [64, p. 563] that if $0 < \tau < \frac{1}{2}$, $n \to \infty$, and $k/n \to \tau$, then $x_{k,n}^{(1)}/n \to \frac{1}{2} - \sqrt{\tau(1-\tau)}$.

The Krawtchouk polynomials of degree at most two are

$K_0(x;n) = 1$ without zeros;
$K_1(x;n) = n - 2x$ with zero $\frac{1}{2}n$;
$K_2(x;n) = 2x^2 - 2nx + \binom{n}{2}$ with zeros $\frac{1}{2}(n \pm \sqrt{n})$.

## 1.4 Weight and Distance Enumerators

When studying the properties of a code, it is often necessary to have some more detailed information on the weights of the codewords and the distances in the code. When studying the covering radius of a code, we need to have this information for the translates of the code. For these purposes we introduce the weight and distance enumerator of a code.

Let $C$ be a code of length $n$ and let $A_i$ be the number of codewords of weight $i$. The sequence $\{A_i\}_{i=0}^n$ is called the weight distribution of code $C$. The polynomial $A(z) := \sum A_i\, z^i$ is called the weight enumerator of code $C$. The weight enumerator $A(z)$ of a linear code $C$ and the weight enumerator $B(z)$ of its dual code $C^\perp$ are related via the so-called MacWilliams identities

$$B(z) = \frac{1}{|C|}(1+z)^n\, A\left(\frac{1-z}{1+z}\right). \tag{1.9}$$

For a proof of this statement we refer to Theorem 1.2.
Comparing the coefficients of $z^0, z^1, z^2, \ldots$ on both sides of this equation, we obtain an explicit relation between the weight enumerator $\{A_i\}_{i=0}^n$ of code $C$ and the sequence $\{B_j\}_{j=0}^n$, viz.

$$B_j = \frac{1}{|C|}\sum_{i=0}^{n} A_i K_j(i;n), \quad (j = 0, \ldots, n). \tag{1.10}$$

Here the numbers $\{K_j(i;n)\}_{j=0}^n$ are the Krawtchouk coefficients defined by Equation (1.1).

The sequence $\{B_j\}_{j=0}^n$ defined by Equation (1.10) is called the MacWilliams transform or dual of $\{A_i\}_{i=0}^n$. The MacWilliams transform of the weight distribution $\{A_i\}_{i=0}^n$ only has an interpretation if $C$ is a linear code. If $C$ is not linear, we can still consider this sequence, though. First, however, we define the weight distribution of translates of code $C$ and the distance distribution of this code.

For all $\mathbf{x} \in \mathbb{F}_2^n$ let $A_i(\mathbf{x})$ be the number of codewords at distance $i$ from $\mathbf{x}$. It follows, that the code $\mathbf{x} + C$ has weight distribution $\{A_i(\mathbf{x})\}_{i=0}^n$. We denote the corresponding weight enumerator by $A_{\mathbf{x}}(z)$.
The distance enumerator $A_C(z)$ of code $C$ is defined by

$$A_C(z) := \sum_{i=0}^n A_i(C)z^i = \frac{1}{|C|}\sum_{\mathbf{x}\in C} A_{\mathbf{x}}(z). \tag{1.11}$$

If $A_{\mathbf{x}}(z)$ does not depend on the actual choice of $\mathbf{x} \in C$, then code $C$ is called distance invariant. Linear codes are an example of distance invariant codes. Notice that if $C$ is a distance invariant code and $\mathbf{0} \in C$, then the weight and distance enumerator of $C$ are the same.

The MacWilliams transform of $A_{\mathbf{x}}(z)$ is denoted by $B_{\mathbf{x}}(z) = \sum B_j(\mathbf{x})z^j$; the MacWilliams transform of $A_C(z)$ by $B_C(z)$. From the definition of the distance enumerator it follows directly that

$$B_C(z) := \sum_{i=0}^n B_i(C)z^i = \frac{1}{|C|}\sum_{\mathbf{x}\in C} B_{\mathbf{x}}(z). \tag{1.12}$$

All coefficients of the polynomial $B_C(z)$ are nonnegative.

**Theorem 1.1** Let $C$ be a code of length $n$. Let $A(z)$ be the distance enumerator of $C$ with dual distance enumerator $B(z)$. Let $A_{\mathbf{x}}(z)$ be the weight enumerator of $\mathbf{x} + C$ with dual enumerator $B_{\mathbf{x}}(z)$. Then

1. $B_j(\mathbf{x}) = |C|^{-1} \sum_{\mathbf{y}:wt(\mathbf{y})=j} (-1)^{\langle \mathbf{x},\mathbf{y}\rangle} \sum_{\mathbf{c}\in C}(-1)^{\langle \mathbf{c},\mathbf{y}\rangle}$,

2. $B_j = |C|^{-2} \sum_{\mathbf{y}:wt(\mathbf{y})=j} \left(\sum_{\mathbf{c}\in C}(-1)^{\langle \mathbf{c},\mathbf{y}\rangle}\right)^2 \geq 0.$

**Proof:**

1. From Equation (1.2) and the definition of $B_j(\mathbf{x})$ we infer that

$$\begin{aligned} |\mathcal{C}|B_j(\mathbf{x}) &= \sum_{i=0}^{n} A_i(\mathbf{x})K_j(i;n) \\ &= \sum_{\mathbf{c}\in\mathcal{C}} K_j(d(\mathbf{x},\mathbf{c});n) \\ &= \sum_{\mathbf{c}\in\mathcal{C}} \sum_{\mathbf{y}:wt(\mathbf{y})=j} (-1)^{\langle \mathbf{c}-\mathbf{x},\mathbf{y}\rangle} \\ &= \sum_{\mathbf{y}:wt(\mathbf{y})=j} (-1)^{\langle -\mathbf{x},\mathbf{y}\rangle} \sum_{\mathbf{c}\in\mathcal{C}} (-1)^{\langle \mathbf{c},\mathbf{y}\rangle}. \end{aligned}$$

2. From Property 1 of this theorem and the definition of $B_j$ we infer that

$$\begin{aligned} |\mathcal{C}|^2 B_j &= |\mathcal{C}| \sum_{\mathbf{x}\in\mathcal{C}} B_j(\mathbf{x}) \\ &= \sum_{\mathbf{y}:wt(\mathbf{y})=j} \left( \sum_{\mathbf{c}\in\mathcal{C}} (-1)^{\langle \mathbf{c},\mathbf{y}\rangle} \right)^2 \geq 0. \end{aligned}$$

□

If $\mathcal{C}$ is a linear code, then the sequence $\{B_j\}_{j=0}^{n}$ has a natural interpretation.

**Theorem 1.2** Let $\mathcal{C}$ be a linear code of length $n$ with weight enumerator $A(z)$ and let $B(z)$ be the MacWilliams transform of $A(z)$. Then $B(z)$ is the weight enumerator of the dual code $\mathcal{C}^{\perp}$.

**Proof:** Let $\mathbf{y} \in \mathbb{F}_2^n$ be a vector of weight $j$ and consider the summation

$$\sum_{\mathbf{c}\in\mathcal{C}} (-1)^{\langle \mathbf{c},\mathbf{y}\rangle}.$$

If $\mathbf{y} \in \mathcal{C}^{\perp}$, then the inner product $\langle \mathbf{c},\mathbf{y}\rangle$ always assumes the value 0. If $\mathbf{y} \notin \mathcal{C}^{\perp}$, then the inner product $\langle \mathbf{c},\mathbf{y}\rangle$ assumes the values 0 and 1 equally often in this summation, since $\mathcal{C}$ is a linear code. It follows, that the summation has value $|\mathcal{C}|$ if $\mathbf{y} \in \mathcal{C}^{\perp}$ and value 0 if $\mathbf{y} \notin \mathcal{C}^{\perp}$. Using Property 2 of Theorem 1.1, we infer that $B_j$ is equal to the number of codewords of weight $j$ in $\mathcal{C}^{\perp}$. □

Let $\mathcal{C}$ be a code of length $n$ with dual distance distribution $\{B_j\}_{j=0}^{n}$ and let $N(\mathcal{C}) := \{1 \leq j \leq n \mid B_j \neq 0\}$. The smallest integer in this set is called the dual distance $d'$ of code $\mathcal{C}$. Notice, that if $\mathcal{C}$ is a linear code, then its dual distance is the minimum distance of $\mathcal{C}^{\perp}$.

Sometimes it is useful to consider, instead of the set $N(C)$, the *annihilator polynomial* of code $C$ defined by

$$\sigma(x) = \frac{2^n}{|C|} \prod_{j \in N(C)} \left(1 - \frac{x}{j}\right). \tag{1.13}$$

The next lemma, a direct consequence of Theorem 1.1, proves to be useful in the rest of the chapter.

**Lemma 1.3** Let $C$ be a code of length $n$. Then

1. $B_j = 0 \Leftrightarrow \sum_{\mathbf{c} \in C} (-1)^{\langle \mathbf{c}, \mathbf{y} \rangle} = 0$ for all $\mathbf{y} \in \mathbb{F}_2^n$ with weight $j$,

2. $B_j = 0 \Leftrightarrow B_j(\mathbf{x}) = 0$ for all $\mathbf{x} \in \mathbb{F}_2^n$. □

The dual distance of a code is closely related to that of its punctured and shortened codes. From Property 1 of Lemma 1.3 it follows directly that puncturing a code with dual distance $d'$ does not decrease the dual distance. The following lemma relates the dual distance of a code with that of its shortened codes.

**Lemma 1.4** Let $C$ be an $(n, M, d)$ code with dual distance $d' > 1$. Let $C_0 := \{\mathbf{c} \mid (\mathbf{c}, 0) \in C\}$ and let $C_1 := \{\mathbf{c} \mid (\mathbf{c}, 1) \in C\}$. Then $C_0$ and $C_1$ have parameters $(n-1, M/2, d)$ and dual distance (at least) $d' - 1$.

**Proof:** From Property 1 of Lemma 1.3 we infer that for every vector $(\mathbf{x}, x_n) \in \mathbb{F}_2^n$ of weight $0 < wt((\mathbf{x}, x_n)) < d'$ we have

$$\sum_{(\mathbf{c}, c_n) \in C} (-1)^{\langle (\mathbf{c}, c_n), (\mathbf{x}, x_n) \rangle} = \sum_{\mathbf{c} \in C_0} (-1)^{\langle \mathbf{c}, \mathbf{x} \rangle} + (-1)^{x_n} \sum_{\mathbf{c} \in C_1} (-1)^{\langle \mathbf{c}, \mathbf{x} \rangle} = 0. \tag{1.14}$$

Substitution of $\mathbf{x} = \mathbf{0}$ and $x_n = 1$ in Equation (1.14) yields $|C_0| = |C_1| = M/2$. If $0 < wt(\mathbf{x}) < d' - 1$, then Equation (1.14) holds both for $x_n = 0$ and for $x_n = 1$. Hence

$$\sum_{\mathbf{c} \in C_0} (-1)^{\langle \mathbf{c}, \mathbf{x} \rangle} = \sum_{\mathbf{c} \in C_1} (-1)^{\langle \mathbf{c}, \mathbf{x} \rangle} = 0 \text{ if } 0 < wt(\mathbf{x}) < d' - 1.$$

Using Property 1 of Lemma 1.3 once again, we infer that $C_0$ and $C_1$ have dual distance at least $d' - 1$. □

## 1.5 More on the Code Parameters

In this section we derive certain relations between the parameters of a code. We consider relations between the dual distance, the weights in the code, and the covering radius. Most results have been originally obtained by Delsarte [25] and can also be found in [64]. We show that all results follow from one simple lemma (Lemma 1.5), thus offering a uniform approach.

The next lemma shows that the weight distribution of a code satisfies certain linear equations.

**Lemma 1.5** Let $\mathcal{C}$ be a code of length $n$. Let $\beta(x) := \sum \beta_j K_j(x;n)$ be a polynomial in $\mathbb{R}[x]$ for which $\beta_j = 0$ if $j \in N(\mathcal{C})$. Then

$$\sum_{i=0}^{n} A_i(\mathbf{x})\beta(i) = |\mathcal{C}|\beta_0 \text{ for all } \mathbf{x} \in \mathbb{F}_2^n.$$

**Proof:** Let $\mathbf{x} \in \mathbb{F}_2^n$. From Lemma 1.3 we infer that $B_j(\mathbf{x}) = 0$ for all $0 \neq j \notin N(\mathcal{C})$. Moreover, by definition of $B_j(\mathbf{x})$ we have $B_0(\mathbf{x}) = 1$. Therefore

$$\begin{aligned}\sum_{i=0}^{n} A_i(\mathbf{x})\beta(i) &= \sum_i A_i(\mathbf{x}) \sum_j \beta_j K_j(i;n) \\ &= \sum_j \beta_j \sum_i A_i(\mathbf{x}) K_j(i;n) \\ &= |\mathcal{C}| \sum_j \beta_j B_j(\mathbf{x}) = |\mathcal{C}|\beta_0.\end{aligned}$$

□

It is clear from Lemma 1.5 that the weight distribution of a code $\mathcal{C}$ satisfies $(n+1) - s'$ linearly independent equations, where $s' := |N(\mathcal{C})|$. If the number of unknowns in the weight distribution is at most $d'$, then this weight distribution can be uniquely determined from the weights in the code.

**Theorem 1.6** Let $\mathcal{C}$ be a code of length $n$ with dual distance $d'$. Let $\mathbf{x} \in \mathbb{F}_2^n$. If the set $W := \{i \mid A_i(\mathbf{x}) \neq 0\}$ has size $s \leq d'$, then the weight distribution $\{A_i(\mathbf{x})\}_{i=0}^n$ is uniquely determined by the set $W$.

**Proof:** Let $W = \{w_1, \ldots, w_s\}$ and assume that $s \leq d'$. Let $1 \leq j \leq s$ and let $L_j(x)$ be the polynomial of minimal degree that is 1 on $w_j$ and 0 on $w_i$ $(i \neq j)$, i.e.

$$L_j(x) = \prod_{\substack{i=1 \\ i \neq j}}^{s} \left(\frac{x - w_i}{w_j - w_i}\right) = \sum_{i=0}^{s-1} \lambda_i^{(j)} K_i(x;n).$$

This so-called Lagrange polynomial has degree $s-1 < d'$. Hence we can apply Lemma 1.5 and find that

$$A_{w_j}(\mathbf{x}) = \sum_{i=0}^{n} A_i(\mathbf{x})L_j(i) = \lambda_0^{(j)}|\mathcal{C}|.$$

□

From the proof of Theorem 1.6 it is immediately clear, that if $s_0$ out of the $s$ nonzero coefficients of the weight distribution are known and if $s - s_0 \leq d'$, then we can still determine the weight distribution of the code. Therefore Theorem 1.6 remains valid, if we redefine parameter $s$ to be the number of unknowns in the weight distribution.

**Corollary 1.7** Let $\mathcal{C}$ be a code of length $n$ with dual distance $d'$ and distance distribution $\{A_i\}_{i=0}^{n}$. If the set $W$ of nonzero distances in the code has size $s \leq d'$, then code $\mathcal{C}$ is distance invariant and the distance distribution is uniquely determined by the set $W$.

**Proof:** Let $s \leq d'$. Let $\mathbf{c}$ be a codeword of $\mathcal{C}$ and let $W_{\mathbf{c}} := \{i \mid A_i(\mathbf{c}) \neq 0\}$. By definition of the distance enumerator we have $W_{\mathbf{c}} \subseteq W \cup \{0\}$. Since $A_0(\mathbf{c}) = 1$, the weight distribution $\{A_i(\mathbf{c})\}_{i=0}^{n}$ has at most $d'$ unknowns. Using Theorem 1.6, we find that the weight enumerator $A_{\mathbf{c}}(z)$ is uniquely determined by the set $W$. Since $A_{\mathbf{c}}(z)$ does not depend on the actual choice of $\mathbf{c} \in \mathcal{C}$, code $\mathcal{C}$ is distance invariant. □

Sometimes the code has even more combinatorial structure.

**Definition 1.8** Let $S$ be a set of $v$ elements and let $\mathcal{B}$ be a collection of subsets of $S$, each with cardinality $k$. The pair $(S, \mathcal{B})$ is called a $t$-$(v, k, \lambda)$ design, or $t$-design, if for every $T \subset S$ with cardinality $|T| = t$ there are exactly $\lambda$ elements $B$ of $\mathcal{B}$ such that $T \subset B$.

**Theorem 1.9** Let $\mathcal{C}$ be a code of length $n$ with dual distance $d'$. Let $\mathbf{x} \in \mathbb{F}_2^n$. If the set $W := \{i > 0 \mid A_i(\mathbf{x}) \neq 0\}$ has size $s = d' - t$, where $t > 0$, then the codewords of each weight in $\mathbf{x} + \mathcal{C}$ form a $t$-design, provided that this weight is at least $t$. The parameters of this design are uniquely determined by the set $W$.

**Proof:** Let $t = d' - s > 0$. Let $T$ be any set of $t$ different coordinates of the code. Let $\mathcal{C}_T$ be the code obtained from $\mathbf{x} + \mathcal{C}$ by considering the codewords that have ones on the positions of $T$ and subsequently deleting these positions. Since code $\mathcal{C}_T$ can be obtained from $\mathbf{x} + \mathcal{C}$ by shortening it $t$ times, we infer from Lemma 1.4 that $\mathcal{C}_T$ has cardinality $|\mathcal{C}_T| = 2^{-t}|\mathcal{C}|$ and dual distance at least $d' - t$. Now we consider the weight distribution $\{a_i\}_{i=0}^{n-t}$ of $\mathcal{C}_T$. By definition of $\mathcal{C}_T$ we have $\{i + t \mid a_i \neq 0\} \subset W$, hence $\mathcal{C}_T$ has at most $s = d' - t$ weights. From Theorem 1.6 we obtain that the weight distribution of $\mathcal{C}_T$ is uniquely determined by the set $W$. It follows that $a_i$, the number of codewords of weight $t + i$ in $\mathbf{x} + \mathcal{C}$ that have ones on the positions of $T$, does not depend on the actual choice of $T$, i.e. the codewords of weight $t + i$ in $\mathbf{x} + \mathcal{C}$ form a $t$-$(n, t + i, a_i)$ design. □

Delsarte [25, Theorem 2.2] proved that if a code has dual distance $d'$ and if it has $s$ nonzero distances, then $s \geq \lfloor (d'-1)/2 \rfloor$. This bound was referred to as the dual MacWilliams inequality in [24, Equation (5.37)]. We will need a slightly stronger result.

**Theorem 1.10** Let $C$ be a code of length $n$ with dual distance $d'$. Let $\mathbf{x} \in \mathbb{F}_2^n$ and let $W := \{i \mid A_i(\mathbf{x}) \neq 0\}$. Then $|W| \geq \lfloor (d'-1)/2 \rfloor + 1$.

**Proof:** See the appendix at the end of this chapter. □

Lemma 1.5 shows that the weight distribution of a code satisfies certain linear equations with coefficients induced by some polynomial $\beta(x)$. By a proper choice of this polynomial one obtains an upper bound on the covering radius of the code, as was shown by Tietäväinen [81]. We will prove a small extension of his result.

**Theorem 1.11** Let $C$ be a code of length $n$. Let $\beta(x) := \sum \beta_j K_j(x;n)$ be a nonzero polynomial in $\mathbb{R}[x]$ for which $\beta_j = 0$ if $j \in N(C)$. Suppose $\beta(x) \leq 0$ for all integers in the interval $(\theta, n]$. Then $C$ has covering radius at most $\theta$ in each of the following two cases:

1. $\beta_0 > 0$,
2. $\beta_0 = 0$ and $\beta(x)$ has at most $\lfloor (d'-1)/2 \rfloor$ integral zeros in the interval $(\theta, n]$.

If $C$ is a self-complementary code (i.e. $C$ is invariant under the translation $\mathbf{x} \rightarrow \mathbf{x} + \mathbf{1}$), then the bound on the covering radius remains valid if we replace the constraints on the interval $(\theta, n]$ by constraints on the smaller interval $(\theta, n-\theta)$.

**Proof:** Let $\mathbf{x} \in \mathbb{F}_2^n$. From Lemma 1.5 we infer that

$$\sum_{i=0}^{n} A_i(\mathbf{x})\beta(i) = |C|\beta_0. \tag{1.15}$$

1. If $\beta_0 > 0$, then it follows from Equation (1.15) that not all the numbers $A_i(\mathbf{x})$ with $i \leq \theta$ can be zero, hence $d(\mathbf{x}, C) \leq \theta$. Therefore code $C$ has covering radius at most $\theta$.
2. Let $\beta_0 = 0$ and suppose that $d(\mathbf{x}, C) > \theta$. It follows from Equation (1.15) that $A_i(\mathbf{x})\beta(i) = 0$ for all $i > \theta$, i.e. $\beta(x)$ is zero on the set $W := \{i \mid A_i(\mathbf{x}) \neq 0\}$. By Lemma 1.10 this set has cardinality $|W| \geq \lfloor (d'-1)/2 \rfloor + 1$ and hence $\beta(x)$ has more than $\lfloor (d'-1)/2 \rfloor$ integral zeros on the interval $(\theta, n]$. This proves the statement.

If $C$ is a self-complementary code, then $A_i(\mathbf{x}) = A_{n-i}(\mathbf{x})$. Therefore the result remains valid if we replace the constraints on the interval $(\theta, n]$ by constraints on the smaller interval $(\theta, n-\theta) = (\theta, n] \cap [0, n-\theta)$. □

**Remark 1.12** In [81] the same result was proved, but only for polynomials $\beta(x)$ of degree at most $d'-1$ for which $\beta_0 > 0$.

Any polynomial $\beta(x)$ that satisfies the conditions of Theorem 1.11 yields an upper bound on the covering radius of a code $C$ as a function of the set $N(C)$.

**Example 1.13** Let $C$ be a self-complementary code of length $n$ with dual distance $d' \geq 3$. We want to find an upper bound on the covering radius of this code by considering the polynomial $\beta(x) = \beta_0 K_0(x) + \beta_1 K_1(x) + K_2(x) = \beta_0 + \beta_1(n-2x) + (2x^2 - 2nx + \frac{1}{2}n(n-1))$. Choose $\beta_0 \geq 0$ and $\beta_1$ in such a way that $\beta(\theta) = \beta(n-\theta) = 0$, where $\theta < \frac{1}{2}n$. Since $\beta(x) < 0$ on the interval $(\theta, n-\theta)$, the assumptions of Theorem 1.11 are satisfied and hence $\theta$ is an upper bound on the covering radius of $C$. For each $\beta_0$ in the range $[0, \frac{1}{2}n]$ we find that $\beta_1 = 0$ and $\theta(n-\theta) = \frac{1}{2}\beta_0 + \frac{1}{4}n(n-1)$, i.e. $\theta_{1,2} = \frac{1}{2}(n \pm \sqrt{n - 2\beta_0})$. Clearly, one obtains the best bound if $\beta_0 = 0$, i.e. if $\beta(x) = K_2(x)$. It follows, that $C$ has covering radius $r(C) \leq \frac{1}{2}(n - \sqrt{n})$.

**Remark 1.14** This result was originally proved by Helleseth et al. [36] and was referred to as the Norse bound in [17]. In fact, the proof of [36, Theorem 3] already uses the polynomial $\beta(x) = K_2(x;n)$ implicitly. The bound in the above example is tight, since the first order Reed-Muller code $\mathcal{R}(1,m)$ of length $n = 2^m$, $m$ even, has dual distance $d' = 4$ and covering radius $r = \frac{1}{2}(n - \sqrt{n})$. If $\mathbf{x}$ has maximal distance to this Reed-Muller code, then the coset $\mathbf{x} + \mathcal{R}(1,m)$ only contains words of weights $w_{1,2} = \frac{1}{2}(n \pm \sqrt{n})$.

Theorem 1.11 yields an upper bound on the covering radius of a code $C$ as a function of the set $N(C)$. The best upper bounds known[1] on the covering radius were obtained by Tietäväinen [81] and Delsarte [24] and depend on the dual distance $d' = \min N(C)$, resp. on the number $s' := |N(C)|$. First we give Tietäväinen's bound.

**Theorem 1.15** [81] Let $C$ be a code of length $n$ with dual distance $d' > 1$. Let $x(k,n)$ be the smallest zero of the Krawtchouk polynomial $K_k(x;n)$. Then $C$ has covering radius $r$ with

$$r \leq \begin{cases} x(t,n) & \text{if } d' = 2t, \\ x(t,n-1) & \text{if } d' = 2t+1. \end{cases}$$

**Proof:** Our proof is based upon an application of Theorem 1.11. For any polynomial $\beta(x)$ we denote its Krawtchouk expansion by $\beta(x) = \sum \beta_i K_i(x;n)$. In order to apply Theorem 1.11, we need to know $\beta_0$. Recall from Section 1.3 that the Krawtchouk polynomials $\{K_k(x;n)\}_{k=0}^{n}$ form an orthogonal basis of the vector space of all polynomials in $\mathbb{R}[x]$ of degree at most $n$ with inner product $\langle f(x), g(x)\rangle_n$. Using the detailed orthogonality relation (1.3), we find that $\langle \beta(x), 1\rangle_n = \beta_0 2^n$.

1. Let $\alpha_1$ be the smallest zero of $K_t(x) := K_t(x;n)$. Let $\beta(x) := -K_t^2(x)/(x-\alpha_1)$. Since $K_t(x)$ has degree $t$ and $K_t(x)/(x-\alpha_1)$ is a polynomial of degree $t-1$, we find that
$$\beta_0 2^n = \langle \beta(x), 1\rangle_n = -\langle K_t(x), K_t(x)/(x-\alpha_1)\rangle_n = 0.$$

[1] Very recently an asymptotically superior bound was reported, see [61].

Notice that $\beta(x) \leq 0$ on $(\alpha_1, n]$ and has $t-1$ distinct roots on this interval. It follows, that if $C$ has dual distance $d' \geq 2t$, then $\beta(x)$ satisfies the assumptions of Theorem 1.11 with $\theta = \alpha_1$. Therefore code $C$ has covering radius $r \leq \alpha_1 = x(t,n)$.

2. Let $\alpha_1$ be the smallest zero of $K_t(x) := K_t(x; n-1)$. Let $\beta(x) = (x-n)K_t^2(x)/(x-\alpha_1)$. Since $K_t(x)$ has degree $t$ and $K_t(x)/(x-\alpha_1)$ is a polynomial of degree $t-1$, we find that

$$\begin{aligned}\beta_0 2^n = \langle \beta(x), 1\rangle_n &= \langle (x-n)K_t(x), K_t(x)/(x-\alpha_1)\rangle_n \\ &= -n\langle K_t(x; n-1), K_t(x; n-1)/(x-\alpha_1)\rangle_{n-1} = 0.\end{aligned}$$

Here we used Equation (1.5). Notice that $\beta(x) \leq 0$ on $(\alpha_1, n]$ and has $t$ distinct roots on this interval. It follows, that if $C$ has dual distance $d' \geq 2t+1$, then $\beta(x)$ satisfies the assumptions of Theorem 1.11 with $\theta = \alpha_1$. Therefore code $C$ has covering radius $r \leq \alpha_1 = x(t, n-1)$. □

Notice that the upper bounds on the covering radius can only be attained if all the zeros of the Krawtchouk polynomial $K_t(x)$ in the proof are integers. When applying Theorem 1.15, the exact value of $x(k,n)$ is usually not known. In that case one can use the estimate for $x(k,n)$ given in Section 1.3. In [81] Tietäväinen separately considered upper bounds on the covering radius of self-complementary codes. Those bounds follow from Theorem 1.15 as well.

Sometimes it is useful to describe the properties of a polynomial via a related polynomial.

**Lemma 1.16** Let $\beta(x) = \sum_{i=0}^{n} \beta_i K_i(x)$ and let $\gamma(x) = \sum_{i=0}^{n} \gamma_i K_i(x)$. Then $\gamma_i = \beta(i)$ for all $i$ iff $\beta_j = 2^{-n}\gamma(j)$ for all $j$.

**Proof:** Suppose $\gamma(x) = \sum_{i=0}^{n} \beta(i) K_i(x)$. Then $\gamma(x) = \sum_{k=0}^{n} \beta_k \sum_{i=0}^{n} K_k(i) K_i(x)$. From Equation (1.6) we infer that

$$\gamma(j) = \sum_{k=0}^{n} \beta_k \sum_{i=0}^{n} K_k(i) K_i(j) = \sum_{k=0}^{n} \beta_k \delta_{kj} 2^n = \beta_j 2^n \quad \text{for all } j, 0 \leq j \leq n.$$

The converse statement follows from the fact that $\gamma(x)$ is uniquely determined by the function values $\gamma(0), \ldots, \gamma(n)$. □

The polynomial $\gamma(x) = \sum_{i=0}^{n} \beta(i) K_i(x)$ is called the Fourier transform of $\beta(x)$.

Now we are ready to prove Delsarte's bound.

**Theorem 1.17** [24] Let $C$ be a code of length $n$. Then $C$ has covering radius at most $s' := |N(C)|$. Moreover, for every $\mathbf{x} \in \mathbb{F}_2^n$ the weight enumerator $\{A_i(\mathbf{x})\}_{i=0}^{n}$ is uniquely determined by the first $s'$ coefficients of this sequence.

**Proof:** Our proof is based upon an application of Theorem 1.11 and Lemma 1.5. Let $\sigma(x)$ be the annihilator polynomial of code $\mathcal{C}$. Recall from (1.13) that this polynomial is zero on $N(\mathcal{C})$ and has degree $s'$.
Let $\beta(x) = \sum \beta_i K_i(x;n)$ and let $\gamma(x)$ be the Fourier transform of $\beta(x)$.
If we choose $\gamma(x) = \sigma(x)$, then $\gamma(x)$ has degree $s'$, $\gamma(j) = 0$ for all $j \in N(\mathcal{C})$, and $\gamma(0) = 2^n|\mathcal{C}|^{-1}$. By Lemma 1.16 we now have $\beta_0 = |\mathcal{C}|^{-1} > 0$, $\beta_j = 0$ for all $j \in N(\mathcal{C})$, and $\beta(s'+1) = \ldots = \beta(n) = 0$. It follows, that $\beta(x)$ satisfies the conditions of Theorem 1.11 with $\theta = s'$. Therefore code $\mathcal{C}$ has covering radius at most $s'$. By Lemma 1.5 we have

$$\sum_{j=0}^{s'} \gamma_j A_j(\mathbf{x}) = 1.$$

If we choose $\gamma(x) = x^i\ \sigma(x) = \sum \gamma_j^{(i)} K_j(x)$ $(i > 0)$, then we find, similarly, that $\beta_j = 0$ for all $j \in N(\mathcal{C})$ and $\beta_0 = 2^{-n}\gamma(0) = 0$. Now Lemma 1.5 yields

$$\sum_{j=0}^{s'+i} \gamma_j^{(i)} A_j(\mathbf{x}) = 0 \text{ if } i > 0.$$

It follows that the weight enumerator of $\{A_i(\mathbf{x})\}_{i=0}^{n}$ is uniquely determined by the first $s'$ coefficients of this sequence. □

Lemma 1.5 showed that the weight distribution of a code satisfies certain linear equations. Below we will show that, in fact, these equations are characterized by this lemma. First, however, we give a slightly stronger result, which will prove to be useful in Chapter 2, page 34.

**Definition 1.18** Let $f : \mathbb{F}_2^n \to \mathbb{R}$. The Hadamard transform of $f$ is the function $F$ defined by

$$F(\mathbf{y}) := \sum_{\mathbf{x}} (-1)^{\langle \mathbf{x},\mathbf{y}\rangle} f(\mathbf{x}) \quad \text{for all } \mathbf{y} \in \mathbb{F}_2^n.$$

**Theorem 1.19** Let $\mathcal{C}$ be a code of length $n$ with annihilator polynomial $\sigma(x)$ and let $\lambda(x) := \sum \lambda_i K_i(x;n)$ be a polynomial in $\mathbb{R}[x]$. Let $f(\mathbf{x}) := \sum_{i=0}^{n} \lambda_i A_i(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{F}_2^n$ and let $F$ be the Hadamard transform of $f$. Then

1. $F(\mathbf{y}) = \lambda(wt(\mathbf{y})) \sum_{\mathbf{c} \in \mathcal{C}} (-1)^{\langle \mathbf{c},\mathbf{y}\rangle}$,
2. If $f(\mathbf{x}) \geq \beta$ for all $\mathbf{x} \in \mathbb{F}_2^n$, then $|F(\mathbf{y})| \leq |\mathcal{C}|\lambda(0) - \beta \cdot 2^n$ for all $\mathbf{y} \neq \mathbf{0}$,
3. If $f(\mathbf{x}) = \beta$ for all $\mathbf{x} \in \mathbb{F}_2^n$, then $\sigma(x) \mid \lambda(x)$ and $\beta = 2^{-n}|\mathcal{C}|\lambda(0)$.

**Proof:**

1. Let $F_i$ be the Hadamard transform of the function $\mathbf{x} \rightarrow A_i(\mathbf{x})$. From Equation (1.2) we infer that
$$\begin{aligned} F_i(\mathbf{y}) &= \sum_{\mathbf{x}} (-1)^{\langle \mathbf{x}, \mathbf{y} \rangle} A_i(\mathbf{x}) \\ &= \sum_{\mathbf{c} \in \mathcal{C}} \sum_{\mathbf{z}:wt(\mathbf{z})=i} (-1)^{\langle \mathbf{c} - \mathbf{z}, \mathbf{y} \rangle} \\ &= \sum_{\mathbf{z}:wt(\mathbf{z})=i} (-1)^{\langle \mathbf{z}, \mathbf{y} \rangle} \sum_{\mathbf{c} \in \mathcal{C}} (-1)^{\langle \mathbf{c}, \mathbf{y} \rangle} \\ &= K_i(wt(\mathbf{y}); n) \sum_{\mathbf{c} \in \mathcal{C}} (-1)^{\langle \mathbf{c}, \mathbf{y} \rangle}. \end{aligned}$$
The result now follows from the observation that $F(\mathbf{y}) = \sum \lambda_i F_i(\mathbf{y})$.

2. Suppose $f(\mathbf{x}) \geq \beta$ for all $\mathbf{x} \in \mathbb{F}_2^n$. Observe that
$$\sum_{\mathbf{x} \in \mathbb{F}_2^n} (-1)^{\langle \mathbf{x}, \mathbf{y} \rangle} = \begin{cases} 2^n & \text{if } \mathbf{y} = \mathbf{0}, \\ 0 & \text{otherwise.} \end{cases}$$
It follows, that for all nonzero vectors $\mathbf{y}$ in $\mathbb{F}_2^n$ we have
$$|F(\mathbf{y})| = |\sum_{\mathbf{x}} (-1)^{\langle \mathbf{x}, \mathbf{y} \rangle} (f(\mathbf{x}) - \beta)| \leq \sum_{\mathbf{x}} |f(\mathbf{x}) - \beta| = F(\mathbf{0}) - \beta \cdot 2^n. \qquad (1.16)$$
By Property 1 of this theorem we have $F(\mathbf{0}) = |\mathcal{C}|\lambda(0) = |\mathcal{C}| \sum_{i=0}^{n} \lambda_i \binom{n}{i}$, hence the result follows.

3. Suppose $f(\mathbf{x}) = \beta$ for all $\mathbf{x} \in \mathbb{F}_2^n$. From (1.16) we infer that $\beta = 2^{-n}|\mathcal{C}|\lambda(0)$ and that $F(\mathbf{y}) = 0$ for all $\mathbf{y} \neq \mathbf{0}$. By Property 1 of this theorem this implies, that if $\mathbf{y} \in \mathbb{F}_2^n$ has weight $j > 0$ and if
$$\sum_{\mathbf{c} \in \mathcal{C}} (-1)^{\langle \mathbf{c}, \mathbf{y} \rangle} \neq 0,$$
then $\lambda(j) = 0$. From Property 2 of Lemma 1.3 it now follows, that $\sigma(x) \mid \lambda(x)$. □

By stating the properties of the polynomial $\lambda(x)$ in terms of those of its Fourier transform, it can easily be verified that Property 3 of Theorem 1.19 is the converse of Theorem 1.5. In particular, we obtain the following corollary.

**Corollary 1.20** Let $\mathcal{C}$ be a code of length $n$. let $\lambda(x) := \sum \lambda_i K_i(x; n)$ be a polynomial in $\mathbb{R}[x]$ and let $\theta \in \mathbb{R}$. Then
$$\sum_{i=0}^{n} \lambda_i A_i(\mathbf{x}) = \theta \quad \text{for all } \mathbf{x} \in \mathbb{F}_2^n \qquad (1.17)$$
if and only if $\sigma(x) \mid \lambda(x)$ and $\theta = 2^{-n}|\mathcal{C}|\lambda(0)$.

**Remark 1.21** Codes that satisfy Equation (1.17) with $\theta \neq 0$ are called perfect weighted coverings by some authors, see e.g. [18]. This class of codes contains the perfect codes, nearly perfect codes, and uniformly packed codes. It can easily be shown that Lloyd's theorem for perfect codes [58, Theorem 7.1.8] and generalizations of this theorem [58, Theorem 7.3.5] all follow from Corollary 1.20, using Theorem 1.17.

## 1.6 Appendix A

Proof of Theorem 1.10:
Let $t := \lfloor (d'-1)/2 \rfloor$. We will show that the only polynomial in $\mathbb{R}[x]$ of degree at most $t$ that is zero on $W$ is the zero polynomial, thus proving that $|W| \geq t+1$.
Let $\alpha_0, \ldots, \alpha_t \in \mathbb{R}$ and suppose that

$$\sum_{i=0}^{t} \alpha_i K_i(x) = 0 \text{ on } W. \tag{1.18}$$

To prove the theorem we will show that $\alpha_0 = \cdots = \alpha_t = 0$. Let $0 \leq j \leq t$. From (1.18) we infer that

$$\sum_{w=0}^{n} K_j(w) A_w(\mathrm{x}) \sum_{i=0}^{t} \alpha_i K_i(w) = \sum_{i=0}^{t} \alpha_i \sum_{w=0}^{n} A_w(\mathrm{x}) K_i(w) K_j(w) = 0. \tag{1.19}$$

We now consider the last summation in more detail. For all $i, 0 \leq i \leq t$, let

$$S_{ij} = \sum_{w=0}^{n} A_w(\mathrm{x}) K_i(w) K_j(w). \tag{1.20}$$

The polynomial $\beta(x) := K_i(x)K_j(x)$ has degree $i+j \leq 2t < d'$. Denote the Krawtchouk expansion of this polynomial by $\beta(x) = \sum \beta_h K_h(x)$. Using the detailed orthogonality relation (1.3), we find that $\langle \beta(x), 1 \rangle_n = \langle \beta(x), K_0(x) \rangle_n = \beta_0 2^n$. Therefore

$$\beta_0 2^n = \langle \beta(x), 1 \rangle_n = \langle K_i(x) K_j(x), 1 \rangle_n = \langle K_i(x), K_j(x) \rangle_n = \delta_{ij} \binom{n}{j} 2^n. \tag{1.21}$$

Combining Lemma 1.5 with Equations (1.20) and (1.21), we infer that

$$S_{ij} = \delta_{ij} \binom{n}{j} |C|. \tag{1.22}$$

From Equations (1.19), (1.20), and (1.22) we now obtain the following result:

$$\sum_{i=0}^{t} \alpha_i \sum_{w=0}^{n} A_w(\mathrm{x}) K_i(w) K_j(w) = \alpha_j \binom{n}{j} |C| = 0. \tag{1.23}$$

It follows, that $\alpha_j = 0$. □

# Chapter 2

# Lower Bounds on Covering Codes

## 2.1 Introduction

A basic question concerning the covering radius of codes is how to determine $K(n, r)$, the minimum cardinality of any code of length $n$ with covering radius $r$. For linear codes this question amounts to determining the minimum dimension $k$ such that an $[n, k]$ code exists with covering radius (at most) $r$.

In this chapter we are interested in finding lower bounds on the size of a code with a prescribed covering radius. The cardinality $|C|$ of any block code $C$ of length $n$ and covering radius $r$ satisfies the inequality

$$|C| \sum_{i=0}^{r} \binom{n}{i} \geq 2^n.$$

Equality in this so-called *Sphere Covering Bound* holds only if $C$ is a *perfect* code, i.e. has minimum distance $d = 2r + 1$. A necessary condition for this to occur is $(r + 1) \mid (n + 1)$. Whenever $(r + 1) \quad (n + 1)$, we can improve this sphere covering bound. This improved bound, the so-called Van Wee bound [84], can be derived in exactly the same way as the well-known Johnson bound for error-correcting codes [64, p. 532]. Recently, Brouwer and Tolhuizen [6] obtained an improvement of the Johnson bound for binary linear codes. Since the proofs of the Johnson bound and the Van Wee bound are closely related, one might also expect improvements of the Van Wee bound for binary linear codes. In fact, improvements of the Van Wee bound for binary linear codes have already been reported in [46]. We follow a similar approach, but obtain many more results, only using some simple observations.

The existence of a perfect code of length $n$ and minimum distance $d = 2r + 1$ imposes a much stronger condition on the parameters $r$ and $n$ than the divisibility constraint $(r + 1) \mid (n + 1)$, viz. there should exist an $(r + 1) - (n, 2r + 1, 1)$ design. If such a design does not exist, then we can, again, improve the sphere covering bound.

In this chapter we will derive bounds that improve upon the sphere covering bound and compare the results with the Johnson bound for error-correcting codes.

## 2.2 Preliminaries

In this section we introduce some notions to facilitate the exposition in the rest of the chapter.

We use the following notations. The all-zero vector and the all-one vector are denoted by $\mathbf{0}$, resp. $\mathbf{1}$. The vector $\mathbf{e}_i$ denotes the vector with all entries equal to zero, except for the $i$th coordinate, which is one. In addition, for any two subsets $U, V$ of $\mathbb{F}_2^n$ we define $U + V := \{u + v \mid u \in U, v \in V\}$. Recall from the previous chapter that the sphere $B_s(\mathbf{x})$ of radius $s$ around $\mathbf{x} \in \mathbb{F}_2^n$ is defined by $B_s(\mathbf{x}) := \{\mathbf{y} \in \mathbb{F}_2^n \mid d(\mathbf{x}, \mathbf{y}) \leq s\}$. The cardinality of this sphere is denoted by $V(n, s)$.

The notion of packing and covering design will prove to be useful in this chapter.

**Definition 2.1** Let $S$ be a set of $v$ elements and let $\mathcal{B}$ be a collection of subsets of $S$ (the so-called blocks), each with cardinality $k$. The pair $(S, \mathcal{B})$ is called a $t$-$(v, k)$ packing [covering] design, if for every $T \subset S$ with cardinality $|T| = t$ there is at most [at least] one element $B$ of $\mathcal{B}$ such that $T \subset B$. The maximum number of blocks of any $t$-$(v, k)$ packing design is denoted by $f_e(v, k, t)$; the minimum number of blocks of any $t$-$(v, k)$ covering design by $f_c(v, k, t)$.

Notice that $f_e(v, k, t) \leq f_c(v, k, t)$ with equality if and only if there exists a $t$-design with parameters $t$-$(v, k, 1)$ (cf. Definition 1.8).

**Lemma 2.2** For $t > 0$ the numbers $f_e(v, k, t)$ and $f_c(v, k, t)$ satisfy the following recurrence relations:

$$\begin{aligned} f_e(v,k,t) &\leq \left\lfloor \tfrac{v}{k}\, f_e(v-1,k-1,t-1) \right\rfloor, \\ f_c(v,k,t) &\geq \left\lceil \tfrac{v}{k}\, f_c(v-1,k-1,t-1) \right\rceil. \end{aligned}$$

Moreover, for $t = 0$ we have $f_e(v, k, t) = f_c(v, k, t) = 1$.

**Proof:** Let $(S, \mathcal{B})$ be a $t$-$(v, k)$ packing design with $f_e(v, k, t)$ blocks ($t > 0$). The number of blocks that contain a fixed element of $S$ is at most $f_e(v-1, k-1, t-1)$. Therefore the sum of the cardinalities of the blocks of $(S, \mathcal{B})$ satisfies the inequality $k \cdot f_e(v, k, t) \leq v \cdot f_e(v-1, k-1, t-1)$. Since the numbers $f_e(v, k, t)$ are all integers, the result follows. The statements for $f_c(v, k, t)$ can be proved similarly. □

The notion of multiset will prove to be useful in Section 2.4.

**Definition 2.3** Let $X$ be a finite set. A function $f : X \to \mathbb{N} \cup \{0\}$ is called a multiset (on $X$). For any $x \in X$, $f(x)$ is called the multiplicity of point $x$ in multiset $f$. For any multiset $f$ on $X$ and for any set $S \subseteq X$, the projection $(f)_S$ of $f$ on $S$ is defined by $(f)_S(x) := f(x)$ if $x \in S$, and $(f)_S(x) := 0$ otherwise. The cardinality $|f|$ of a multiset $f$ is defined by $|f| := \sum_x f(x)$.

The next lemma will be used in Section 2.5.

**Lemma 2.4** [84, Lemma 6] Let $s, t \geq 0$. For all $\mathbf{x}, \mathbf{y} \in \mathbb{F}_2^n$ the quantity $|B_s(\mathbf{x}) \cap B_t(\mathbf{y})|$ only depends on $d(\mathbf{x}, \mathbf{y})$ and is non-increasing in $d(\mathbf{x}, \mathbf{y})$.

**Proof:** The quantity $|B_s(\mathbf{x}) \cap B_t(\mathbf{y})|$ only depends on $d(\mathbf{x}, \mathbf{y})$, since it is invariant under permutations of the coordinate positions and translations in $\mathbb{F}_2^n$. Let $\mathbf{x} = (\mathbf{0}, 0, \mathbf{0})$, $\mathbf{y}_1 = (\mathbf{1}, 0, \mathbf{0})$, and $\mathbf{y}_2 = (\mathbf{1}, 1, \mathbf{0})$ with $wt(\mathbf{y}_2) = i$. Let $\mathcal{A} := B_s(\mathbf{x}) \cap B_t(\mathbf{y}_1)$ and $\mathcal{B} := B_s(\mathbf{x}) \cap B_t(\mathbf{y}_2)$. To prove the lemma we show that $|\mathcal{A}| \geq |\mathcal{B}|$.
We have $\mathcal{A} \setminus \mathcal{B} = \{\mathbf{z} \in \mathbb{F}_2^n \mid wt(\mathbf{z}) \leq s, z_i = 0, d(\mathbf{z}, \mathbf{y}_1) = t\}$. Similarly, we have $\mathcal{B} \setminus \mathcal{A} = \{\mathbf{z} \in \mathbb{F}_2^n \mid wt(\mathbf{z}) \leq s, z_i = 1, d(\mathbf{z}, \mathbf{y}_2) = t\}$. Puncturing a vector $\mathbf{w} \in \mathbb{F}_2^n$ on position $i$ is a linear mapping; the result is denoted by $\mathbf{w}'$. Puncturing $\mathbf{y}_1, \mathbf{y}_2$, and $\mathbf{z} \in \mathbb{F}_2^n$ on position $i$ yields the vectors $\mathbf{y}'$ and $\mathbf{z}'$ in $\mathbb{F}_2^{n-1}$. Now $(\mathcal{A} \setminus \mathcal{B})' = \{\mathbf{z}' \in \mathbb{F}_2^{n-1} \mid wt(\mathbf{z}') \leq s, d(\mathbf{z}', \mathbf{y}') = t\}$. Similarly, $(\mathcal{B} \setminus \mathcal{A})' = \{\mathbf{z}' \in \mathbb{F}_2^{n-1} \mid wt(\mathbf{z}') \leq s-1, d(\mathbf{z}', \mathbf{y}') = t\}$. Therefore $(\mathcal{A} \setminus \mathcal{B})' \supseteq (\mathcal{B} \setminus \mathcal{A})'$, hence $|\mathcal{A} \setminus \mathcal{B}| = |(\mathcal{A} \setminus \mathcal{B})'| \geq |(\mathcal{B} \setminus \mathcal{A})'| = |\mathcal{B} \setminus \mathcal{A}|$. This proves that $|\mathcal{A}| \geq |\mathcal{B}|$. □

## 2.3 The Johnson Bound

Most lower bounds on the size of a covering code that we discuss in this chapter turn out to be direct analogues of the well-known Johnson bound for error-correcting codes [64, p. 532]. For this reason we discuss the Johnson bound in detail.

The size of any code $(n, M, 2e+1)$ code $C$ satisfies the trivial inequality

$$|C| \sum_{i=0}^{e} \binom{n}{i} \leq 2^n. \tag{2.1}$$

Equality in this so-called *Sphere Packing Bound* holds only if $C$ is a perfect code; often better upper bounds are known. We will discuss the Johnson bound, which improves on the sphere packing bound and gives a strong condition on the parameters of codes meeting the sphere packing bound with equality.

Before we derive the Johnson bound, we first give a definition.

**Definition 2.5** A constant weight code with parameters $(n, d, w)$ is a code of length $n$ with minimum distance at least $d$ for which all codewords have weight $w$. The maximum size of any $(n, d, w)$ constant weight code is denoted by $A(n, d, w)$.

Two different words in an $(n, d, w)$ code intersect in at most $w - \lceil d/2 \rceil$ positions, hence an $(n, d, w)$ code is a $t$-packing design for any $t \geq (w+1) - \lceil d/2 \rceil$. Using Lemma 2.2, we immediately obtain the following result.

**Lemma 2.6** $A(n,d,w) \le \lfloor \frac{n}{w} A(n-1,d,w-1)\rfloor$ if $d \le 2w$, and $A(n,d,w) = 1$ if $d > 2w$.

Now we are ready to derive the Johnson bound.

Let $C$ be an $(n, M, d)$ code with $d = 2e+1$. Let $C_{e+1} := \{\mathbf{x} \in \mathbb{F}_2^n \mid d(\mathbf{x}, C) = e+1\}$. Since $C$ has minimum distance $d = 2e+1$, we have

$$|C| \sum_{i=0}^{e} \binom{n}{i} + |C_{e+1}| \le 2^n. \tag{2.2}$$

The Johnson bound can be derived by estimating the cardinality of $C_{e+1}$, the set of words at distance $e+1$ from the code.

We obtain an estimate on the cardinality of $C_{e+1}$ by estimating in two ways the cardinality of set $S$ defined by

$$S := \{(\mathbf{c}, \mathbf{x}) \mid \mathbf{c} \in C, \mathbf{x} \in C_{e+1}, d(\mathbf{c}, \mathbf{x}) = e+1\}.$$

Let $A_d$ be the average value of $A_d(\mathbf{c})$ over all $\mathbf{c} \in C$. Let $\mu$ be the average value of $A_{e+1}(\mathbf{x})$ over all $\mathbf{x} \in C_{e+1}$. Then we obtain

$$\mu|C_{e+1}| = \sum_{\mathbf{x}\in C_{e+1}} A_{e+1}(\mathbf{x}) = |S| = \sum_{\mathbf{c}\in C} \left(\binom{n}{e+1} - A_d(\mathbf{c})\binom{d}{e}\right) = |C| \left(\binom{n}{e+1} - A_d\binom{d}{e}\right). \tag{2.3}$$

Combining (2.2) with (2.3), we get the following bound on the size of code $C$:

$$|C| \left\{\sum_{i=0}^{e}\binom{n}{i} + \frac{1}{\mu}\left(\binom{n}{e+1} - A_d\binom{d}{e}\right)\right\} \le 2^n. \tag{2.4}$$

We now determine estimates for parameters $A_d$ and $\mu$. By definition of the numbers $A(n, d, w)$ we have

$$A_i(\mathbf{x}) \le A(n,d,i) \text{ for all } \mathbf{x} \in \mathbb{F}_2^n. \tag{2.5}$$

Using Lemma 2.6, we find that

$$A_d \le A(n,d,d) \text{ and } \mu \le A(n,d,e+1) = \lfloor n/(e+1)\rfloor. \tag{2.6}$$

Sustituting estimate (2.6) for parameters $A_d$ and $\mu$ in Equation (2.4), we get the following upper bound on the size of code $C$, known as the *Johnson bound:*

$$|C| \left\{\sum_{i=0}^{e}\binom{n}{i} + \frac{1}{\lfloor \frac{n}{e+1}\rfloor}\left(\binom{n}{e+1} - A(n,d,d)\binom{d}{e}\right)\right\} \le 2^n. \tag{2.7}$$

By Lemma 2.6 we have $A(n,d,d) \le \binom{n}{e+1}/\binom{d}{e}$, hence the Johnson bound always improves upon the sphere packing bound. The bounds coincide iff there exists an $(e+1)$-$(n, 2e+1, 1)$ design.

In the remainder of this chapter we will derive lower bounds on the size of a covering code, which bear a strong similarity with the Johnson bound (2.7). In order to compare these bounds with the Johnson bound, it is useful to consider the following two specialized versions of the Johnson bound:

- From Lemma 2.6 we infer that $\binom{d}{e}A(n,d,d) \le \binom{n}{e}\lfloor\frac{n-e}{e+1}\rfloor$. Combining this estimate with the Johnson bound (2.7), we get the following, original, version of the Johnson bound:
$$|C|\left\{\sum_{i=0}^{e}\binom{n}{i}+\frac{\binom{n}{e}}{\lfloor\frac{n}{e+1}\rfloor}\left(\frac{n+1}{e+1}-\left\lfloor\frac{n+1}{e+1}\right\rfloor\right)\right\}\le 2^n. \tag{2.8}$$
- From Lemma 2.6 we infer that $\binom{d}{e-1}A(n,d,d) \le \binom{n}{e-1}A(n-e+1,d,e+2)$. Combining this estimate with the Johnson bound (2.7), we get the following version of the Johnson bound:
$$|C|\left\{\sum_{i=0}^{e}\binom{n}{i}+\frac{\binom{n+1}{e}}{\frac{n+1}{e+2}\lfloor\frac{n}{e+1}\rfloor}\left(\frac{(n+1-e)(n-e)}{(e+2)(e+1)}-f_e(n-e+1,e+2,2)\right)\right\}\le 2^n. \tag{2.9}$$
(Here we used the equality $A(n-e+1,2e+1,e+1)=f_e(n-e+1,e+2,2)$.)

In Section 2.4 we will derive the direct analogue of bound (2.8) for covering codes, a bound known as the Van Wee bound. In Section 2.7 we will derive the direct analogue of bound (2.9) for covering codes, a bound known as the Zhang bound. In Section 2.5 and Section 2.8 we consider improvements of either bound for binary linear codes.

## 2.4 The Van Wee Bound

In the previous section we discussed the well-known Johnson bound for error-correcting codes. Here we derive a similar bound for covering codes, the so-called Van Wee bound [84].

Let $C$ be an $(n,M,d)r$ code. Let $C_r := \{\mathbf{x}\in\mathbb{F}_2^n \mid d(\mathbf{x},C)=r\}$. Let $f(\mathbf{x}) := |B_r(\mathbf{x})\cap C|$ and let $\mathcal{A}(\mathbf{x}) := f(\mathbf{x})-1$ for all $\mathbf{x}\in\mathbb{F}_2^n$. Since $C$ has covering radius $r$, $\mathcal{A}$ is indeed a multiset and we have
$$|C|\sum_{i=0}^{r}\binom{n}{i}=2^n+|\mathcal{A}|. \tag{2.10}$$
The Van Wee bound can be derived by estimating the cardinality of multiset $\mathcal{A}$, the set of words that are covered more than once, counting multiplicities.

We obtain an estimate on the cardinality of $\mathcal{A}$ by estimating in two ways the double summation

$$S := \sum_{\mathbf{x} \in \mathcal{C}_r} \sum_{\mathbf{z}: d(\mathbf{z},\mathbf{x}) \leq 1} \mathcal{A}(\mathbf{z}).$$

Let $\varepsilon$ be the average cardinality of $(\mathcal{A})_{B_1(\mathbf{x})}$ over all $\mathbf{x} \in \mathcal{C}_r$. Let $\mu$ be the maximum size of $B_1(\mathbf{z}) \cap \mathcal{C}_r$ over all $\mathbf{z} \in \mathbb{F}_2^n$ with $\mathcal{A}(\mathbf{z}) > 0$. Then we obtain

$$\mu|\mathcal{A}| \geq \sum_{\mathbf{z} \in \mathbb{F}_2^n} \mathcal{A}(\mathbf{z}) \cdot |B_1(\mathbf{z}) \cap \mathcal{C}_r| = S = \sum_{\mathbf{x} \in \mathcal{C}_r} |(\mathcal{A})_{B_1(\mathbf{x})}| = \varepsilon|\mathcal{C}_r|. \tag{2.11}$$

From Equation (2.11) we infer that the cardinality of $\mathcal{A}$ satisfies the inequality

$$|\mathcal{A}| \geq \frac{\varepsilon}{\mu}|\mathcal{C}_r|. \tag{2.12}$$

Combining (2.10) with (2.12) and making use of the lower bound on $|\mathcal{C}_r|$ given by

$$|\mathcal{C}_r| \geq 2^n - |\mathcal{C}| \sum_{i=0}^{r-1} \binom{n}{i}, \tag{2.13}$$

we get the following bound on the size of code $\mathcal{C}$:

$$|\mathcal{C}| \left\{ \sum_{i=0}^{r} \binom{n}{i} - \frac{\varepsilon}{\mu + \varepsilon} \binom{n}{r} \right\} \geq 2^n. \tag{2.14}$$

We now determine estimates for parameters $\varepsilon$ and $\mu$.

First we estimate the quantity $|(\mathcal{A})_{B_1(\mathbf{x})}|$ for all $\mathbf{x} \in \mathcal{C}_r$.
Let $d(\mathbf{x}, \mathcal{C}) = r$ and define $T_r(\mathbf{x}) := \{\mathbf{y} \in \mathbb{F}_2^n \mid \mathbf{x} - \mathbf{y} \in \mathcal{C} \text{ and } wt(\mathbf{y}) \leq r+1\}$. Notice that $T_r(\mathbf{x})$ contains only vectors of weights $r$ and $r+1$. Define $a_i := |\{\mathbf{y} \in T_r(\mathbf{x}) \mid y_i = 1\}|$ for all $i, 1 \leq i \leq n$, and $a_\infty := |\{\mathbf{y} \in T_r(\mathbf{x}) \mid wt(\mathbf{y}) = r\}|$. For all codewords $\mathbf{c} \in \mathcal{C}$ we have $d(\mathbf{x} + \mathbf{e}_i, \mathbf{c}) \leq r$ iff $d(\mathbf{x}, \mathbf{c}) \leq r+1$ and $x_i \neq c_i$. Therefore $a_i = |B_r(\mathbf{x} + \mathbf{e}_i) \cap \mathcal{C}| = f(\mathbf{x} + \mathbf{e}_i)$ and $a_\infty = |B_r(\mathbf{x}) \cap \mathcal{C}| = f(\mathbf{x})$.
Since $\sum_{i=1}^{n} a_i + a_\infty = (r+1)|T_r(\mathbf{x})|$, it follows that

$$|(\mathcal{A})_{B_1(\mathbf{x})}| = (r+1)\left(|T_r(\mathbf{x})| - \frac{n+1}{r+1}\right), \quad \text{if } d(\mathbf{x}, \mathcal{C}) = r. \tag{2.15}$$

Since $\mathcal{C}$ has covering radius $r$, all $a_i$ and $a_\infty$ are positive, hence $\operatorname{supp} T_r(\mathbf{x}) = \{1, \ldots, n\}$ and $|T_r(\mathbf{x})| \geq \left\lceil \frac{n+1}{r+1} \right\rceil$. Together with Equation (2.15) this implies that $\varepsilon$ satisfies the inequality

$$\varepsilon \geq (r+1)\left(\left\lceil \frac{n+1}{r+1} \right\rceil - \frac{n+1}{r+1}\right). \tag{2.16}$$

Now we estimate the maximum value of $|B_1(\mathbf{z}) \cap \mathcal{C}_r|$ over all $\mathbf{z} \in \mathbb{F}_2^n$ with $\mathcal{A}(\mathbf{z}) > 0$.

Let $\mathbf{z} \in I\!F_2^n$ with $\mathcal{A}(\mathbf{z}) > 0$, i.e. $f(\mathbf{z}) > 1$. Let $\tilde{T}(\mathbf{z}) := \{\mathbf{y} \in I\!F_2^n \mid \mathbf{z}-\mathbf{y} \in \mathcal{C} \text{ and } wt(\mathbf{y}) \leq r\}$. All vectors $\mathbf{z} + \mathbf{e}_i$ with $i \in \operatorname{supp}\tilde{T}(\mathbf{z})$ are at distance $< r$ from code $\mathcal{C}$. If $d(\mathbf{z}, \mathcal{C}) = r$, then $|\operatorname{supp}\tilde{T}(\mathbf{z})| \geq r + \lceil d/2 \rceil$, since $\mathbf{z}$ is covered at least twice and the distance between different codewords is at least $d$. Similarly, if $d(\mathbf{z}, \mathcal{C}) = r - 1$, then $|\operatorname{supp}\tilde{T}(\mathbf{z})| \geq r - 1 + \lceil d/2 \rceil$. By definition of $\mu$ it follows, that

$$|B_1(\mathbf{z}) \cap \mathcal{C}_r| \leq \mu \leq (n+1) - (r + \lceil d/2 \rceil), \quad \text{if } \mathbf{z} \in I\!F_2^n \text{ and } \mathcal{A}(\mathbf{z}) > 0. \tag{2.17}$$

Substituting the estimates (2.16) for $\varepsilon$ and (2.17) for $\mu$ in Equation (2.14) and taking $d = 1$, we get the following lower bound on the size of code $\mathcal{C}$, known as the *Van Wee bound:*

$$|\mathcal{C}| \left\{ \sum_{i=0}^{r} \binom{n}{i} - \frac{\binom{n}{r}}{\lceil \frac{n-r}{r+1} \rceil} \left( \left\lceil \frac{n+1}{r+1} \right\rceil - \frac{n+1}{r+1} \right) \right\} \geq 2^n. \tag{2.18}$$

**Remark 2.7** Notice the strong similarity between the Van Wee bound and the original version of the Johnson bound for error-correcting codes (2.8). This correspondence is no coincidence: the original version of the Johnson bound and the Van Wee bound can be derived in exactly the same way.

## 2.5 An Improvement of the Van Wee Bound for Binary Linear Codes

The Van Wee bound improves on the sphere covering bound, whenever $(r+1) \quad (n+1)$. If $\mathcal{C}$ is a linear code, we can sometimes improve this bound further. The main idea is that the situation $|T_r(\mathbf{x})| = \lceil \frac{n+1}{r+1} \rceil$ cannot occur too often, because of the linearity of the code. We will see that the actual minimum distance of the code also plays a role. The improvement of the Van Wee bound for binary linear codes is entirely based upon a better estimate for $\varepsilon$ than estimate (2.16). We use the terminology introduced in the previous section.

Let $\mathcal{C}$ be an $[n,k,d]r$ code and let $n + 1 = a(r+1) - b$, with $0 \leq b < r + 1$. From the proof of the Van Wee bound we infer that all sets $T_r(\mathbf{x})$, with $d(\mathbf{x}, \mathcal{C}) = r$, have cardinality $|T_r(\mathbf{x})| \geq a$. We estimate how often the equality $|T_r(\mathbf{x})| = a$ can occur.
First we define $\mathcal{C}_r^{(1)} := \{\mathbf{x} \in \mathcal{C}_r \mid |T_r(\mathbf{x})| = a\}$. Let $\alpha$ and $\beta$ be defined by $\alpha := |\mathcal{C}_r^{(1)}|$ and $\alpha + \beta = |\mathcal{C}_r|$. Parameter $\varepsilon$ in the proof of the Van Wee bound can now be estimated by

$$\varepsilon \geq \frac{\alpha b + \beta(r+1+b)}{\alpha + \beta}. \tag{2.19}$$

We can write this in a form similar to Equation (2.16) and get

$$\varepsilon \geq (r+1)\left( \left\lceil \frac{n+1}{r+1} \right\rceil - \frac{n+1}{r+1} + \frac{\beta}{\alpha+\beta} \right). \tag{2.20}$$

Our improvement of the Van Wee bound for binary linear codes is now obtained by substituting estimate (2.20) for $\varepsilon$ and estimate (2.17) for $\mu$ in Equation (2.14), once we have obtained a suitable upper bound for parameter $\alpha$ (or a lower bound for parameter $\beta$). Notice that in the Van Wee bound $\beta$ is always taken to be zero. For linear codes we can often obtain a better estimate for parameters $\beta$ and $\alpha$.

We now determine estimates for the parameters $\alpha$ and $\beta$. We distinguish two cases, depending on the parity of $a$.

**Case 1:** $a$ is odd.

Let $d(\mathbf{x}, \mathcal{C}) = r$ and suppose $|T_r(\mathbf{x})| = a$. Let $\mathbf{s} := \sum T_r(\mathbf{x})$, the sum of the vectors in $T_r(\mathbf{x})$. If $s_i = 0$, then $a_i = f(\mathbf{x} + \mathbf{e}_i)$ is *even*, otherwise $a_i$ is *odd*. Hence the number of zeros in vector $\mathbf{s}$ equals the number of $i, 1 \leq i \leq n$, for which $a_i$ is even. Since all $a_i$ and $a_\infty$ are positive, $\sum_{i=1}^{n} a_i + a_\infty = (r+1)|T_r(\mathbf{x})|$, and $|T_r(\mathbf{x})| = a$, at most $b$ numbers $a_i$ can be greater than one. Therefore vector $\mathbf{s}$ has weight $wt(\mathbf{s}) \geq n - b$. The code is linear and $a$ is odd, hence vectors $\mathbf{s}$ and $\mathbf{x}$ are in the same coset of code $\mathcal{C}$. So $\mathbf{x} + \mathcal{C} = \mathbf{z} + \mathbf{1} + \mathcal{C}$ with $\mathbf{z} := \mathbf{1} + \mathbf{s}$ of weight $wt(\mathbf{z}) \leq b$. Using the linearity of the code, we infer that $\mathbf{x} \in (B_b(\mathbf{1}) + \mathcal{C}) \cap \mathcal{C}_r = (B_b(\mathbf{1}) \cap \mathcal{C}_r) + \mathcal{C}$. This implies that $\alpha$ and $\beta$ can be estimated by

$$\alpha = |\mathcal{C}_r^{(1)}| \leq |(B_b(\mathbf{1}) \cap \mathcal{C}_r)| \cdot |\mathcal{C}| \quad \text{and} \quad \alpha + \beta = |\mathcal{C}_r|. \tag{2.21}$$

In particular, if $\mathcal{C}$ contains the all-one vector, i.e. is self-complementary, and if $b < r$, then $\alpha = 0$. In general we always have

$$|B_b(\mathbf{1}) \cap \mathcal{C}_r| \leq V(n, b) - V(r, b) \ + \ 1. \tag{2.22}$$

This follows from Lemma 2.4 in the following way: Since code $\mathcal{C}$ has covering radius $r$, we have $d(\mathbf{1}, \mathbf{c}) \leq r$ for some $\mathbf{c} \in \mathcal{C}$. Now $|B_b(\mathbf{1}) \cap B_{r-1}(\mathbf{c})|$ is minimal if $d(\mathbf{1}, \mathbf{c}) = r$ (apply Lemma 2.4). Therefore

$$|B_b(\mathbf{1}) \cap B_{r-1}(\mathbf{c})| \geq \sum_{i<j,\ i+j \leq b} \binom{n-r}{i}\binom{r}{j} \geq V(r, b) \ - \ 1. \tag{2.23}$$

Combining Equations (2.21) and (2.22) we get the following estimate for parameters $\alpha$ and $\beta$ :

$$\begin{aligned} &\alpha + \beta = |\mathcal{C}_r|, \\ &\alpha \leq \min\{(V(n, b) - V(r, b) + 1) \cdot |\mathcal{C}| \ , \ |\mathcal{C}_r|\}. \end{aligned} \tag{2.24}$$

**Case 2:** $a$ is even.

Let $e := \lfloor (d-1)/2 \rfloor$. We consider the case $0 \leq b \leq e$.
Let $d(\mathbf{x}, \mathcal{C}) = r$ and suppose $|T_r(\mathbf{x})| = a$. Let $\mathbf{c} := \sum T_r(\mathbf{x})$, the sum of the vectors in $T_r(\mathbf{x})$. Since $\mathcal{C}$ is a linear code and $a$ is even, the vector $\mathbf{c}$ is a codeword of $\mathcal{C}$. As in Case 1, we find

that $wt(\mathbf{c}) \geq n-b$. Notice that if $c_i = 0$, then the vector $\mathbf{x}+\mathbf{e}_i$ is covered an *even* number of times, otherwise an *odd* number of times. Code $\mathcal{C}$ has minimum distance $d \geq 2e+1$, hence $\mathcal{C}$ contains at most one codeword of weight at least $n-e$. Since $wt(\mathbf{c}) \geq n-b$ and $b \leq e$, the vector $\mathbf{c}$ does not depend on the actual choice of $\mathbf{x} \in \mathcal{C}_r^{(1)}$. We consider two cases.

(I) If $\mathcal{C}$ does not contain the all-one vector, then $wt(\mathbf{c}) \leq n-1$. Now $c_i = 0$ for some $i, 1 \leq i \leq n$, hence $\mathbf{x}+\mathbf{e}_i$ is covered an even number of times, i.e. at least twice. This means that $\mathbf{x}+\mathbf{e}_i$ has positive multiplicity in multiset $\mathcal{A}$. Since $\mathbf{c}$ does not depend on the actual choice of $\mathbf{x} \in \mathcal{C}_r^{(1)}$, in fact the whole set $\mathbf{e}_i + \mathcal{C}_r^{(1)}$ has positive multiplicity in $\mathcal{A}$; therefore $\alpha = |\mathcal{C}_r^{(1)}| \leq |\mathcal{A}|$.

(II) If $\mathcal{C}$ contains the all-one vector, then $\mathbf{c} = \mathbf{1}$. Now all vectors $\mathbf{x}+\mathbf{e}_i$ are covered an odd number of times, i.e. all $\mathcal{A}(\mathbf{x}+\mathbf{e}_i)$ are even. Since $|T_r(\mathbf{x})| = a$, we have $|(\mathcal{A})_{B_1(\mathbf{x})}| = b$. Therefore $\mathcal{A}(\mathbf{x}) \equiv b(\mathbf{mod}\ 2)$, i.e. $|B_r(\mathbf{x}) \cap \mathcal{C}| \equiv (b+1)\ (\mathbf{mod}\ 2)$. If $b$ is odd, then $\mathbf{x}$ is covered an even number of times, i.e. at least twice. This means that $\mathbf{x}$ has positive multiplicity in $\mathcal{A}$. Since $\mathbf{c}$ does not depend on the actual choice of $\mathbf{x} \in \mathcal{C}_r^{(1)}$, we have $\alpha = |\mathcal{C}_r^{(1)}| \leq |\mathcal{A}|$. In fact, the argument works for all odd $b$, $0 \leq b < d$, since $\mathcal{C}$ contains the all-one vector. If $b$ is even, then we cannot expect an improvement of the sphere covering bound, since this occurs for the perfect Hamming and Golay codes (take $b = 0$).

Depending upon whether $\mathcal{C}$ contains the all-one vector, parameters $\alpha$ and $\beta$ can now be estimated by

$$\begin{aligned}
&\alpha + \beta = |\mathcal{C}_r|, \\
&\alpha \leq \min(|\mathcal{A}|, |\mathcal{C}_r|) \ \text{ if } \left\{ \begin{array}{l} \mathbf{1} \notin \mathcal{C},\ \ 0 \leq b \leq \lfloor (d-1)/2 \rfloor \\ \mathbf{1} \in \mathcal{C},\ \ 0 \leq b < d, \ \text{ and } \ b \ \text{ is odd} \end{array} \right. , \\
&\alpha \leq |\mathcal{C}_r| \ \text{ in all other cases.}
\end{aligned} \tag{2.25}$$

In general it is not known whether $\mathcal{C}$ contains the all-one vector. If it is known that $\mathcal{C}$ does not contain any codewords of weight at least $n-b$, then we always have $\mathcal{C}_r^{(1)} = \emptyset$, i.e. $\alpha = 0$. □

**Remark 2.8** Notice that our improvement of the Van Wee bound for binary linear codes works best if $a$ is odd. For $a$ even, the minimum distance of the code is involved, but generally not known. When applying this bound, we may safely assume that the linear code has distance $d \geq 3$, however. This can be seen as follows: when applying our bound, we always consider the smallest length $n$ such that an $[n, n-m]r$ code possibly exists. This code should have minimum distance $d \geq 3$, since otherwise we could have deleted all the double columns and zero-columns from a parity check matrix of this code to obtain a shorter linear code with the same redundancy $m$ and covering radius $r$, but with minimum distance $d \geq 3$. In the next section we follow the procedure described above and hence we may assume that $d \geq 3$. We will not make any further assumptions on the minimum distance. Consequently, when we consider codes with $a$ even, we only obtain improvements on the Van Wee bound if $b = 1$. Our bound can also be used to rule out the existence of many quasi-perfect codes. We do not consider quasi-perfect codes separately, though.

We followed an approach similar to that in [46]. Our bound and the improvement of the Van Wee bound for binary linear codes obtained by Hou correspond in the following sense:

- for $a$ odd, Hou obtained the same bound. It is remarkable that in [46] estimates on the size of set $B_b(\mathbf{1}) \cap C_r$ were only given for $b = 0, 1$, and 2. The next section proves that estimate (2.22) for $|B_b(\mathbf{1}) \cap C_r|$, although simple, already gives rise to many new bounds. Most of the times when we improved a previously known lower bound, $b$ was three, four, or five, but also higher values occurred. For $b > 2$ estimate (2.22) can be sharpened using Equation (2.23). During computations, this sharpening did not produce further improvements, though.

- for $a$ even, Hou only considered codes with $n + 1 = a(r + 1) - 1$ and obtained as estimate
$$\alpha \leq ub := \frac{1}{3}\left(\binom{n}{r} + \binom{n-1}{r-1}\right)|\mathcal{C}|. \tag{2.26}$$
In computations, estimate (2.25) for the value of $\alpha$ is always better than estimate (2.26) in [46].

Compared to the improvement of the Johnson bound for binary linear codes [6], our improvement of the Van Wee bound for binary linear codes is less good. From our point of view this indicates that the covering problem is harder to tackle than the corresponding packing problem.

**Example 2.9** To demonstrate the improvement of the Van Wee bound for linear codes we prove that $[30, 18]3$ codes do not exist.
Suppose a $[30, 18]3$ code exists. From Equations (2.10) and (2.13) we infer that $|\mathcal{A}| = (V(30, 3) - 2^{12})|\mathcal{C}| = 430|\mathcal{C}|$ and $|\mathcal{C}_r| \geq (2^{12} - V(30, 2))|\mathcal{C}| = 3630|\mathcal{C}|$. We have $31 = 8 \cdot 4 - 1$, so $a = 8$ and $b = 1$. We may assume that $d \geq 3$, hence application of estimate (2.25) yields $\alpha \leq |\mathcal{A}| = 430|\mathcal{C}|$ and $\beta \geq 3200|\mathcal{C}|$. Substituting the estimates for $\alpha$ and $\beta$ in Equation (2.20) we get $\varepsilon \geq 1 + 4(1 - \frac{430}{3630})$. From Equation (2.17) and $d \geq 3$ we obtain the bound $\mu \leq 26$. Now Equation (2.12) gives rise to the lower bound $|\mathcal{A}| \geq 631|\mathcal{C}|$, in conflict with the exact value $|\mathcal{A}| = 430|\mathcal{C}|$. This finishes the proof.

## 2.6 Implications of the Improved Bound

In this section we compare our improvement of the Van Wee bound for binary linear codes with bounds obtained in the literature.

We will frequently use the functions $l(m, r)$, which denotes the smallest integer $n$ such that an $[n, n - m]r$ code exists, and $t[n, k]$, which denotes the minimum covering radius achievable by any $[n, k]$ code.

Our improvement of the Van Wee bound for binary linear covering codes gives rise to many improvements on the tables for $t[n,k]$, with $n \leq 64$, and to a number of improvements on the tables for $l(m,r)$, with $m \leq 24$ and $r \leq 12$. Below we give an impression of the strength of our bound, compared to improvements on the tables of [8] and [32] mentioned in the literature.

- In [40] Honkala listed seven improvements on $t[n,k]$ within the range of $n \leq 33$. Only two of them cannot be obtained by our simple bound: $t[33,9] \geq 9$, $t[33,15] = 6$.

- In [46, 47] Hou improved several bounds on $t[n,k]$. Since our bound is a strengthening of his bound, all these bounds can also be obtained using our bound. The same remark holds for the bounds mentioned in the paper by Van Wee [84].

- In [90] Zhang and Lo mentioned sixty-five improvements on $t[n,k]$ within the range of $n \leq 64$. Fifty-two of them can also be obtained by our simple bound, however; only thirteen of them cannot be proved in this way:

  $t[34,8] \geq 10$, $t[39,11] \geq 10$, $t[40,10] \geq 11$, $t[47,11] \geq 13$, $t[53,13] \geq 14$,
  $t[54,12] \geq 15$, $t[59,12] \geq 17$, $t[59,15] \geq 15$, $t[60,14] \geq 16$; $t[36,15] = 7$,
  $t[44,8] \geq 14$, $t[50,13] \geq 13$, $t[52,11] \geq 15$.

  (The entry $t[56,18] \geq 13$ in the tables of [90] is in error and should read $t[58,18] \geq 13$. This result also follows from our bound.)

- In [91] Zhang and Lo mentioned twenty improvements on $t[n,k]$ within the range of $n \leq 64$. Ten of them cannot be obtained via our bound:

  $t[36,11] \geq 9$, $t[39,8] \geq 12$, $t[42,7] \geq 14$, $t[47,21] \geq 8$, $t[54,24] \geq 9$,
  $t[56,23] \geq 10$, $t[59,23] \geq 11$, $t[62,23] \geq 12$, $t[63,16] \geq 16$, $t[63,26] \geq 11$.

  (The entry $t[62,63] \geq 8$ in the tables of [91] is clearly in error. It should read $t[62,33] \geq 8$ and also follows from our bound.)

- In [32] Graham and Sloane determined the exact value of $t[n,k]$ for all $k \leq 5$. In general, these exact values cannot be obtained via our bound.

- Several other results mentioned in the literature cannot be obtained by our simple bound. This remark applies to the following bounds: $t[12,6] = 3$ [8, 32], $t[15,6] = 4$ [72], $t[23,15] = 3$ [11], $t[18,11] = 3$ and $t[64,53] = 3$ [88]. However, in Chapter 3 we will present a systematic way to prove those bounds and some new bounds as well.

A comparison of our bound to the bounds mentioned in the literature [8, 11, 32, 40, 46, 47, 72, 84, 88, 90, 91] gives rise to the following improvements on $l(m,r)$ for $m \leq 24$ and $r \leq 12$:

$$\begin{array}{lllll}
l(12,2) \geq 92, & l(12,3) \geq 31, & l(16,3) \geq 75, & l(19,3) \geq 148, & l(20,3) \geq 187, \\
l(21,3) \geq 235, & l(22,3) \geq 295, & l(23,3) \geq 371, & l(24,3) \geq 467, & l(15,4) \geq 32, \\
l(17,4) \geq 44, & l(18,4) \geq 53, & l(19,4) \geq 62, & l(20,4) \geq 73, & l(21,4) \geq 86, \\
l(22,4) \geq 103, & l(23,4) \geq 122, & l(24,4) \geq 144, & l(16,5) \geq 27, & l(19,5) \geq 39, \\
l(21,5) \geq 51, & l(24,5) \geq 76, & l(19,6) \geq 30, & l(23,6) \geq 46, & l(23,7) \geq 37, \\
l(24,7) \geq 40, & l(23,8) \geq 31. & & &
\end{array}$$

We get the following improvements on $t[n,k]$ for $n \leq 64$:

$$\begin{array}{lllll}
t[26,10] = 6, & t[30,7] \geq 9, & t[30,18] = 4, & t[31,16] = 5, & t[33,6] \geq 11, \\
t[36,13] \geq 8, & t[38,19] = 6, & t[39,15] \geq 8, & t[43,26] = 5, & t[45,22] \geq 7, \\
t[46,14] \geq 11, & t[49,16] \geq 11, & t[50,29] = 6, & t[52,25] \geq 8, & t[52,34] = 5, \\
t[58,18] \geq 13, & t[59,17] \geq 14, & t[61,6] \geq 23, & t[61,27] \geq 10, & t[61,42] = 5, \\
t[62,19] \geq 14, & t[64,6] \geq 24, & t[64,7] \geq 23. & &
\end{array}$$

When applying our bound, we find that certain lower bounds on $t[n,k]$ can only be attained by linear codes with a low minimum distance. In particular this holds for the bounds $t[47,6] \geq 16$ ($d \leq 6$), $t[54,6] \geq 19$ ($d \leq 8$), and $t[54,7] \geq 18$ ($d \leq 8$). Here the maximal feasible distance of codes attaining the lower bound is mentioned in brackets. However, if the code is not a quasi-perfect code, i.e. if $d < 2r-1$, we can slightly refine estimate (2.13) on the size of $C_r$ and estimate (2.12) on the size of $\mathcal{A}$, using e.g. the observation that if $d < 2r-1$ then for all codeword pairs $(\mathbf{c},\mathbf{c}')$ with $d(\mathbf{c},\mathbf{c}') = d$ the set $B_{r-1}(\mathbf{c}) \cap B_{r-1}(\mathbf{c}')$ is nonempty. In this way we can show that none of the three mentioned lower bounds can be attained. Since this refinement produces only three new lower bounds, we abstain from details.

$$t[47,6] \geq 17, \quad t[54,6] \geq 20, \quad t[54,7] \geq 19.$$

The comparison given above shows that most bounds for linear covering codes obtained via others methods can also be obtained via our bound. Moreover, we obtained a large number of improvements of previously reported bounds. This is remarkable, since the computation of our lower bound only requires as input the parameters $[n,k,d]r$ of the code. Further improvements of our bound could be obtained if one could drop the restriction on $b$ in the proof of our bound, whenever $a$ is even. In fact, in all but two cases [1] where other methods led to better bounds we had to apply our bound with $a$ even.

Some results mentioned in the literature also hold when the restriction to linear codes is dropped. In this respect the paper by Zhang [89] is especially worth mentioning. This bound is the subject of the next section.

[1] These cases were $t[36,11] \geq 9$ and $t[54,24] \geq 9$ (bounds obtained by Zhang and Lo, see [91]).

## 2.7 The Zhang Bound

Suppose all $\mathbf{x} \in \mathbb{F}_2^n$ satisfy the inequality

$$f(\mathbf{x}) := \sum_{i=0}^{n} \lambda_i A_i(\mathbf{x}) \geq \beta. \tag{2.27}$$

Summing this inequality over all $\mathbf{x} \in \mathbb{F}_2^n$ and making use of the equation

$$\sum\{A_i(\mathbf{x}) \mid \mathbf{x} \in \mathbb{F}_2^n\} = |C|\binom{n}{i}, \tag{2.28}$$

we get the following bound on the size of code $C$:

$$|C|\sum_{i=0}^{n} \lambda_i \binom{n}{i} \geq \beta \cdot 2^n. \tag{2.29}$$

The Zhang bound can be derived by determining inequalities of the form (2.27).

Let $C$ be a binary code of length $n$ with covering radius $r$. Throughout this section, we denote the weight distribution of the translate $\mathbf{x} + C$ by $\{A_i\}_{i=0}^n$, if the vector $\mathbf{x} \in \mathbb{F}_2^n$ is clear from context. If convenient, we identify codewords with their supports.

We obtain inequalities of the form (2.27) by considering functions $\varphi : \mathbb{N} \to [0, \infty)$ with the property that

$$A_{r+1} + A_{r+2} \geq \varphi(A_{r-1} + A_r) \quad \text{for all } \mathbf{x} \in \mathbb{F}_2^n \text{ with } d(\mathbf{x}, C) \geq r - 1. \tag{2.30}$$

Notice that function $\varphi$ yields an estimate on $A_{r+1} + A_{r+2}$ as a function of $A_{r-1} + A_r$. In general it is difficult to apply estimate (2.30) directly, since the exact value of $A_{r-1} + A_r$ is usually not known. Instead of considering estimate (2.30), we therefore consider a weighted version of this inequality. For any positive number $m_1 \in \mathbb{R}$ define

$$m_0 := m_0(m_1) = \min\{m_1\, k \ + \ \varphi(k) \mid k > 0\}. \tag{2.31}$$

Using the definition of $m_0$ and estimate (2.30), we find that all $\mathbf{x} \in \mathbb{F}_2^n$ satisfy the inequality

$$f(\mathbf{x}) := m_0 \sum_{i=0}^{r-2} A_i + m_1(A_{r-1} + A_r) + (A_{r+1} + A_{r+2}) \geq m_0. \tag{2.32}$$

Combining (2.32) with (2.27) and (2.29), we get the following bound on the size of code $C$:

$$|C|\left\{\sum_{i=0}^{r}\binom{n}{i} - \frac{\binom{n+1}{r}}{m_0}\left((m_0 - m_1) - \frac{(n+1-r)(n-r)}{(r+2)(r+1)}\right)\right\} \geq 2^n. \tag{2.33}$$

Equation (2.33) depends on the actual choices for function $\varphi$ and parameter $m_1$.

At this point, the reader might think that the value of parameter $r$ does not play an essential role. We should point out, however, that it does, since it imposes restrictions on the possible choices for function $\varphi$. Later on, when we choose function $\varphi$, we will need that $r$ is the covering radius of code $C$.

First we choose parameter $m_1$. It is clear from (2.31) that $m_0 - m_1 \leq \varphi(1)$. We choose $m_1$ in such a way that $m_0$ is minimal among all pairs $(m_0, m_1)$ for which $m_0 - m_1 = \varphi(1)$, i.e. $m_1$ is the smallest number such that $m_1 + \varphi(1) \leq m_1\ k + \varphi(k)$ for all $k > 0$. Hence

$$m_0 = m_1 + \varphi(1) \text{ and } m_1 = s := \max\{\frac{\varphi(1) - \varphi(k)}{k-1} \mid k > 1\}. \tag{2.34}$$

Notice, that if $\ell$ is the line below the graph $\{(k, \varphi(k)) \mid k > 0\}$ that meets this graph in $k = 1$ and has maximal slope, then $\ell(k) = \varphi(1) - m_1 \times (k-1)$. Substituting (2.34) in Equation (2.33), we get the following lower bound on the size of code $C$:

$$|C| \left\{ \sum_{i=0}^{r} \binom{n}{i} - \frac{\binom{n+1}{r}}{s + \varphi(1)} \left( \varphi(1) - \frac{(n+1-r)(n-r)}{(r+2)(r+1)} \right) \right\} \geq 2^n. \tag{2.35}$$

Now we show that function $\varphi : I\!N \to [0, \infty)$ defined by

$$\varphi(k) := f_c(n+1-k\ r, r+2, 2) \tag{2.36}$$

satisfies the properties of (2.30). Recall that $f_c(v,k,t)$ denotes the minimum cardinality of a code of length $v$ with codewords of weight (at most) $k$ such that every set of $t$ different coordinates is contained in the support of at least one codeword.

Let $d(\mathbf{x}, C) \geq r-1$. For all $s \geq 0$ let $\mathcal{A}_s$ be the collection of codewords of $\mathbf{x} + C$ of weight $s$. Let $X := \{1, \ldots, n\} \setminus \text{supp}(\mathcal{A}_{r-1} \cup \mathcal{A}_r)$. Notice that $A_s = |\mathcal{A}_s|$ and that $|X| \geq n - (r-1)A_{r-1} - rA_r$.
Every pair $(i,j)$ in $X$ with $i \neq j$ is contained in the support of a word of $\mathcal{A}_{r+1} \cup \mathcal{A}_{r+2}$, for otherwise $d(\mathbf{x} + \mathbf{e}_i + \mathbf{e}_j, C) > r$. Therefore the pair $(X, \mathcal{A}_{r+1} \cup \mathcal{A}_{r+2})$ is a 2-covering design. In particular, we find that

$$A_{r+1} + A_{r+2} \geq f_c(n+1-r(A_{r-1} + A_r), r+2, 2) \text{ if } d(\mathbf{x}, C) = r-1. \tag{2.37}$$

If $d(\mathbf{x}, C) = r$, then every coordinate $i \in X$ is in the support of a word of $\mathcal{A}_{r+1}$, for otherwise $d(\mathbf{x} + \mathbf{e}_i, C) > r$. Therefore the pair $(X, \mathcal{A}_{r+1})$ is a 1-covering design and the pair $(X, \mathcal{A}_{r+1} \cup \mathcal{A}_{r+2})$ is a 2-covering design. It follows, that if we define $X' := X \cup \{n+1\}$ and $\mathcal{A}'_{r+1} := \{(\mathbf{a}, 1) \mid \mathbf{a} \in \mathcal{A}_{r+1}\}$, then $(X', \mathcal{A}'_{r+1} \cup \mathcal{A}_{r+2})$ is a 2-covering design. In particular, we find that

$$A_{r+1} + A_{r+2} \geq f_c(n+1-rA_r, r+2, 2) \quad \text{if } d(\mathbf{x}, C) = r. \tag{2.38}$$

Combining (2.37) and (2.38), we find that the function $\varphi$ defined by (2.36) indeed satisfies the properties of (2.30).

If function $\varphi$ is defined by (2.36), then Equation (2.35) is called the *pair covering inequality.*

**Remark 2.10** The pair covering inequality always improves on the sphere covering bound. The bounds coincide iff there exists a 2-design with parameters 2-$(n+1-r, r+2, 1)$. This situation occurs e.g. for perfect codes. This also motivated our choice for $m_1$: if $m_0 - m_1 < \varphi(1)$, then Equation (2.33) would sometimes have been inferior to the sphere covering bound, e.g. for perfect codes. Our choice for function $\varphi$ was motivated by a comparison of inequality (2.35) and the specialized Johnson bound (2.9). The pair covering inequality was originally proved by Zhang [89]. A similar result was obtained by Honkala in [40], although the proofs are different. Both papers generalize methods developed earlier in [84]. The reader who is interested in extensions of the methods presented in this section is invited to consult [91]. In that paper Zhang and Lo derive lower bounds on covering codes by considering, instead of functions $\varphi$ with property (2.30), functions $\varphi$ with the property that $A_{r+2} + A_{r+3} \geq \varphi(A_{r-2} + A_{r-1}, A_r + A_{r+1})$ for all $\mathbf{x} \in \mathbb{F}_2^n$ with $d(\mathbf{x}, C) \geq r-2$. In that case the evaluation of function $\varphi$ becomes quite complicated, however.

To demonstrate the pair covering inequality we give two examples. We use the tables of lower bounds for $f_c(v,k,t)$, i.e. lower bounds for $t$-covering designs, provided in [40].

Recall that $K(n,r)$ denotes the minimum size of any binary code of length $n$ with covering radius $r$.

**Example 2.11** We prove the lower bound $K(33,8) \geq 531$. We use the following table of lower bounds for pair coverings:

| $k$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $\varphi(k) := f_c(34-8k, 10, 2)$ | 10 | 5 | 1 | 0 |

From the table it follows, that the line $\ell$ defined by $\ell(k) := 10 - 5 \times (k-1)$ is below the graph $\{(k, \varphi(k)) \mid k > 0\}$ and meets this graph for $k=1$ and $k=2$. Therefore $5 \times k + f(34-8k, 10, 2) \geq 5k + \ell(k) = 15$ for all $k > 0$. Hence we can use the pair covering inequality with $s = 5$ and $\varphi(1) = 10$ (or $m_0 = 15$ and $m_1 = 5$) and find that $K(33,8) \geq 531$.

This example explains the entry $t[33,9] \geq 9$, which is the first bound we mentioned in Section 2.6 that cannot be obtained by the improvement of the Van Wee bound for linear codes.

**Example 2.12** We prove the lower bound $K(36,6) \geq 32,734$. We use the following table of lower bounds for pair coverings:

| $k$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| $\varphi(k) := f_c(37-6k,8,2)$ | 20 | 13 | 9 | 4 | 1 | 1 | 0 |

From the table it follows, that the line $\ell$ defined by $\ell(k) := 20 - 7 \times (k-1)$ is below the graph $\{(k,\varphi(k)) \mid k > 0\}$ and meets this graph for $k = 1$ and $k = 2$. Therefore $7 \times k + f(37-6k,8,2) \geq 7k + \ell(k) = 27$ for all $k > 0$. Hence we can use the pair covering inequality with $s = 7$ and $\varphi(1) = 20$ (or $m_0 = 27$ and $m_1 = 7$) and find that $K(36,6) \geq 32,734$.

We will return to this example in the next section.

## 2.8 Intersections of Spheres and Hyperplanes

The Zhang bound gives a lower bound on the size of a covering code $C$ and is obtained by determining inequalities of the form (2.27). If $C$ is a linear code, then these inequalities do not only yield a lower bound on the size of code $C$, but also impose restrictions on the weight distribution of its dual code. Using these restrictions, it is sometimes possible to rule out the existence of a linear covering code, with the help of coding theory. First, however, we derive restrictions on arbitrary (nonlinear) codes.

Let $C$ be a code of length $n$ with covering radius $r$. By definition of the covering radius we have

$$\cup\{B_r(\mathbf{c}) \cap S \mid \mathbf{c} \in C\} = S \quad \text{for every subset } S \text{ of } \mathbb{F}_2^n.$$

In this section we derive restrictions on the intersections of $C$ with suitably chosen subsets $S$. We will always choose $S$ to be a hyperplane of $\mathbb{F}_2^n$.

Let $\mathbf{u} \in \mathbb{F}_2^n$ be a vector of weight $w > 0$ and let $H$ be a coset of the hyperplane $\langle \mathbf{u} \rangle^{\perp}$. For all $\mathbf{x} \in \mathbb{F}_2^n$ define $H_i(\mathbf{x}) := |\{\mathbf{y} \in H \mid d(\mathbf{x},\mathbf{y}) = i\}|$. It is clear that

$$H_i(\mathbf{x}) = \begin{cases} K_i^+(w;n) := \sum\limits_{j \text{ even}} \binom{w}{j}\binom{n-w}{i-j} & \text{if } \mathbf{x} \in H, \\ K_i^-(w;n) := \sum\limits_{j \text{ odd}} \binom{w}{j}\binom{n-w}{i-j} & \text{if } \mathbf{x} \notin H. \end{cases} \tag{2.39}$$

This equation forms the basis for the next theorem, which imposes restrictions on the intersections of certain codes with hyperplanes.

**Theorem 2.13** Let $C$ be a binary code of length $n$ and let $\lambda(x) := \sum \lambda_i K_i(x;n)$ be a polynomial in $\mathbb{R}[x]$. Suppose all $\mathbf{x} \in \mathbb{F}_2^n$ satisfy the inequality

$$f(\mathbf{x}) := \sum_{i=0}^{n} \lambda_i A_i(\mathbf{x}) \geq \beta.$$

Let $\mathcal{C}_0 := \mathcal{C} \cap \langle \mathbf{u} \rangle^\perp$ and let $\mathcal{C}_1 := \mathcal{C} \setminus \mathcal{C}_0$, where $\mathbf{u} \in \mathbb{F}_2^n$ has weight $w > 0$. Then

$$|\mathcal{C}_0|\lambda^+(w) + |\mathcal{C}_1|\lambda^-(w) \geq \beta \cdot 2^{n-1} \quad \text{and} \quad |\mathcal{C}_0|\lambda^-(w) + |\mathcal{C}_1|\lambda^+(w) \geq \beta \cdot 2^{n-1},$$

where the polynomials $\lambda^+(x)$ and $\lambda^-(x)$ are defined by

$$\lambda^+(x) := \sum \lambda_i K_i^+(x;n), \text{ resp. } \lambda^-(x) := \sum \lambda_i K_i^-(x;n).$$

(Hence $\lambda(x) = \lambda^+(x) - \lambda^-(x)$.)

**Proof:** Let $H := \langle \mathbf{u} \rangle^\perp$ and let $H^c$ be the complement of this hyperplane. From Equation (2.39) and the definition of $H_i(\cdot)$ we infer that

$$\begin{aligned} \sum_{\mathbf{x} \in H} f(\mathbf{x}) &= \sum_{i=0}^{n} \lambda_i \sum_{\mathbf{x} \in H} A_i(\mathbf{x}) = \sum_{i=0}^{n} \lambda_i \sum_{\mathbf{c} \in \mathcal{C}} H_i(\mathbf{c}) \\ &= \sum_{i=0}^{n} \lambda_i (|\mathcal{C}_0| K_i^+(w;n) + |\mathcal{C}_1| K_i^-(w;n)) \\ &= |\mathcal{C}_0|\lambda^+(w) + |\mathcal{C}_1|\lambda^-(w). \end{aligned}$$

Therefore $|\mathcal{C}_0|\lambda^+(w) + |\mathcal{C}_1|\lambda^-(w) \geq \beta \cdot 2^{n-1}$. The same exercise, but now with $H^c$, yields the other inequality, viz. $|\mathcal{C}_0|\lambda^-(w) + |\mathcal{C}_1|\lambda^+(w) \geq \beta \cdot 2^{n-1}$. □

**Remark 2.14** Let $\mathcal{C}$ be an $(n, M)r$ code. We have $f(\mathbf{x}) := A_0(\mathbf{x}) + \cdots + A_r(\mathbf{x}) \geq 1$ for all $\mathbf{x} \in \mathbb{F}_2^n$, so we can take $\sum \lambda_i x^i = 1 + \cdots + x^r$ and $\beta = 1$ in the above theorem. If we take $w = 1$, then Theorem 2.13 reduces to [17, Theorem 3]. If equality holds in all inequalities involved in proving the above theorem, then we obtain an alternative proof of Corollary 1.20, i.e. $\sigma(x) \mid \lambda(x)$, where $\sigma(x)$ is the annihilator polynomial of the code involved. This follows directly from our theorem, using the observation that $\sigma(j) = 0$ if and only if there is a vector $\mathbf{u} \in \mathbb{F}_2^n$ with weight $j$ such that the hyperplane $\langle \mathbf{u} \rangle^\perp$ partitions $\mathcal{C}$ into parts of unequal size.

**Corollary 2.15** Let $\mathcal{C}$ be an $[n, k]$ code and let $\lambda(x) := \sum \lambda_i K_i(x;n)$ be a polynomial in $\mathbb{R}[x]$. Suppose all $\mathbf{x} \in \mathbb{F}_2^n$ satisfy the inequality

$$f(\mathbf{x}) := \sum_{i=0}^{n} \lambda_i A_i(\mathbf{x}) \geq \beta.$$

Then the nonzero weights $w$ in the dual code $\mathcal{C}^\perp$ satisfy the inequalities

$$|\mathcal{C}| \cdot \lambda^+(w) \geq \beta \cdot 2^{n-1} \quad \text{and} \quad |\mathcal{C}| \cdot \lambda^-(w) \geq \beta \cdot 2^{n-1}, \tag{2.40}$$

where the polynomials $\lambda^+(x)$ and $\lambda^-(x)$ are defined as in Theorem 2.13.

**Proof:** This follows immediately from Theorem 2.13, using the definition of $\mathcal{C}^{\perp}$. □

**Remark 2.16** If one subtracts the inequalities of (2.40) from each other, one obtains the following inequality, which is due to Zhang and Lo [90]:

$$|\mathcal{C}| \cdot |\lambda(w)| \leq |\mathcal{C}| \sum_{i=0}^{n} \lambda_i \binom{n}{i} - \beta \cdot 2^n. \tag{2.41}$$

(Notice that this result also follows from Theorem 1.19.) In a slightly different form, this result can already be found in an earlier paper by Calderbank and Sloane [11] (they take $\sum \lambda_i x^i = 1 + x + \cdots + x^r$ and $\beta = 1$).

It is clear from Corollary 2.15, that if $d^*$ is the smallest positive integer that satisfies (2.40), then $d^*$ is a lower bound on the minimum distance of the code $\mathcal{C}^{\perp}$. In particular, $d^* \leq d[n, n-k]$, where $d[n,k]$ denotes the largest achievable minimum distance of any $[n,k]$ code. (In fact, one can show the slightly stronger inequality $d^* \leq d[n+1, n+1-k]$ to hold, using Corollary 3.9 of the next chapter.)
This observation can be used to rule out the existence of certain linear covering codes that do satisfy the Zhang bound.

**Example 2.17** In the previous example we showed that $K(36,6) \geq 32,734 = 2^{15} - 34$. Suppose $\mathcal{C}$ is a $[36,15]6$ code. By Example 2.12 all $\mathbf{x} \in \mathbb{F}_2^{36}$ satisfy the inequality

$$f(\mathbf{x}) := 27 \times (A_0 + A_1 + A_2 + A_3 + A_4) + 7 \times (A_5 + A_6) + (A_7 + A_8) \geq 27.$$

From Equation 2.41 we infer that the nonzero weights $w$ in the code $\mathcal{C}^{\perp}$ satisfy $9 \leq w \leq 28$. But $d[36,21] < 9$, so evidently $\mathcal{C}^{\perp}$ does not exist, nor does $\mathcal{C}$. Therefore $t[36,15] \geq 7$.

## 2.9 Another Lower Bound

Let $\mathcal{C}$ be a binary code of length $n$ with covering radius $r$. As in Section 2.4, let $\mathcal{A}$ be the multiset defined by $\mathcal{A}(\mathbf{x}) := |B_r(\mathbf{x}) \cap \mathcal{C}| - 1$ for all $\mathbf{x} \in \mathbb{F}_2^n$. Then

$$|\mathcal{C}| \sum_{i=0}^{r} \binom{n}{i} = 2^n + |\mathcal{A}|. \tag{2.42}$$

We derived the Van Wee bound by estimating the cardinality of multiset $\mathcal{A}$, the set of words that are covered more than once, counting multiplicities. The next bound, due to Cohen et al. [17], can be obtained via another estimate for this multiset. Although this bound is in general rather weak compared with the Van Wee bound, we can still obtain new results applying this bound. The proof uses the function $A(n,d)$, which denotes the maximum cardinality of any binary code of length $n$ with minimum distance $d$.

Let $\mathcal{C}_0$ be a maximal $r$-error-correcting subcode of $\mathcal{C}$ and let $\mathcal{C}_1 := \mathcal{C} \setminus \mathcal{C}_0$. Let $\mathbf{c}_0 \in \mathcal{C}_0$ and let $\mathbf{c}_1 \in \mathcal{C}_1$. All words in the set $B_r(\mathbf{c}_0) \cap B_r(\mathbf{c}_1)$ are covered by $\mathbf{c}_1$, but also by the (unique) codeword $\mathbf{c}_0$ of $\mathcal{C}_0$. Hence

$$|\mathcal{A}| \geq \sum_{\mathbf{c}_0 \in \mathcal{C}_0,} \sum_{\mathbf{c}_1 \in \mathcal{C}_1} |B_r(\mathbf{c}_0) \cap B_r(\mathbf{c}_1)|. \tag{2.43}$$

Now we estimate the righthand side of Equation (2.43). The quantity $|B_r(\mathbf{c}_0) \cap B_r(\mathbf{c}_1)|$ only depends on $d(\mathbf{c}_0, \mathbf{c}_1)$ and is non-increasing in $d(\mathbf{c}_0, \mathbf{c}_1)$, cf. Lemma 2.4. Therefore

$$|B_r(\mathbf{c}_0) \cap B_r(\mathbf{c}_1)| \geq \binom{2r}{r}, \text{ if } d(\mathbf{c}_0, \mathbf{c}_1) \leq 2r. \tag{2.44}$$

From the maximality of code $\mathcal{C}_0$ we infer that for each codeword $\mathbf{c}_1 \in \mathcal{C}_1$ there exists a codeword $\mathbf{c}_0 \in \mathcal{C}_0$ with $d(\mathbf{c}_0, \mathbf{c}_1) \leq 2r$. Using this, Equation (2.43), and Equation (2.44), we obtain the bound

$$|\mathcal{A}| \geq (|\mathcal{C}| - |\mathcal{C}_0|)\binom{2r}{r}. \tag{2.45}$$

Combining (2.42) and (2.45) proves the following theorem.

**Theorem 2.18** Let $\mathcal{C}$ be a binary code of length $n$ with covering radius $r$. Then

$$|\mathcal{C}| \left\{ \sum_{i=0}^{r} \binom{n}{i} - \binom{2r}{r} \right\} \geq 2^n - |\mathcal{C}_0| \binom{2r}{r}, \tag{2.46}$$

where $|\mathcal{C}_0| \leq A(n, 2r+1)$.

Under certain conditions, we can take a slightly better estimate for $|\mathcal{C}_0|$ than the one in Theorem 2.18: if $A(n, 2r+1) = 2 \cdot A(n-1, 2r+1)$, then we can use Equation 2.46 with the stronger estimate $|\mathcal{C}_0| \leq A(n, 2r+1) - 1$ instead (unless $\mathcal{C}$ is a perfect code). To see this, we only need to consider the case $|\mathcal{C}_0| = A(n, 2r+1)$. We show that each codeword $\mathbf{c}_1 \in \mathcal{C}_1$ is at distance $\leq 2r$ from at least two codewords of $\mathcal{C}_0$. If so, then the claim follows from the fact that Equation (2.45) can be sharpened to the bound

$$|\mathcal{A}| \geq 2(|\mathcal{C}| - |\mathcal{C}_0|)\binom{2r}{r} \geq (|\mathcal{C}| - |\mathcal{C}_0| + 1)\binom{2r}{r},$$

unless $\mathcal{C}$ is a perfect code. Suppose otherwise, i.e. let $\mathbf{c}_1 \in \mathcal{C}_1$ and suppose that $d(\mathbf{c}_0, \mathbf{c}_1) \leq 2r$ for exactly one codeword $\mathbf{c}_0$ of $\mathcal{C}_0$. Then $\mathcal{C}_0' := (\mathcal{C}_0 \setminus \{\mathbf{c}_0\}) \cup \{\mathbf{c}_1\}$ is another $r$-error-correcting subcode of $\mathcal{C}$ with cardinality $|\mathcal{C}_0'| = |\mathcal{C}_0| = A(n, 2r+1)$. Since $A(n, 2r+1) = 2 \cdot A(n-1, 2r+1)$, exactly half of the codewords of $\mathcal{C}_0$ has a one on a fixed position; the other codewords of $\mathcal{C}_0$ have a zero there. The same remark holds for code $\mathcal{C}_0'$. Since the code $\mathcal{C}_0'$ can be obtained from $\mathcal{C}_0$ by interchanging the codewords $\mathbf{c}_0$ and $\mathbf{c}_1$, it follows that the codewords $\mathbf{c}_0$ and $\mathbf{c}_1$ coincide on every coordinate, i.e. $\mathbf{c}_0 = \mathbf{c}_1$, a contradiction.

To demonstrate this improvement we prove two bounds, viz. $K(9,1) \geq 55$ and $K(5,1) \geq 7$. Neither of these bounds can be obtained by combining Equation (2.46) with the weaker estimate $|\mathcal{C}_0| \leq A(n, 2r+1)$. The bound $K(9,1) \geq 55$ was independently obtained by Habsieger [33] by studying intersections of codes with hyperplanes. The other bound had been established by Stanton et al. [78] using linear programming techniques. Unlike their proofs, our proofs are completely elementary. A table of bounds for $A(n,d)$ is provided in [64].

**Example 2.19** From the tables for $A(n,d)$ we see that $A(9,3) = 40$. Combining Equation (2.46) with the estimate $|\mathcal{C}_0| = 40$, we find that any code $\mathcal{C} \subseteq \mathbb{F}_2^9$ with covering radius one has cardinality $|\mathcal{C}| \geq 54$. In fact this bound is sharp, i.e. if $|\mathcal{C}| = 54$, then equality holds in all inequalities involved in proving Equation (2.46). Since $A(9,3) = 2 \cdot A(8,3)$, we cannot have equalities everywhere, so $K(9,1) \geq 55$.

**Example 2.20** From the tables of $A(n,d)$ we see that $A(5,3) = 4$. Combining Equation (2.46) with the estimate $|\mathcal{C}_0| = 4$, we find that any code $\mathcal{C} \subseteq \mathbb{F}_2^5$ with covering radius one has cardinality $|\mathcal{C}| \geq 6$. Once again, this bound is sharp, i.e. if $|\mathcal{C}| = 6$, then equality holds in all inequalities involved in proving Equation (2.46). Since $A(5,3) = 2 \cdot A(4,3)$, we cannot have equalities everywhere, so $K(5,1) \geq 7$.

**Remark 2.21** Usually, we can improve Equation (2.46) further. For $n \equiv 5 (\textbf{mod } 6)$ improvements were obtained by van Wee [86] and later on by Honkala [44]. By combining arguments used in proving Equation (2.46) with those used in proving the van Wee bound, they obtained improvements of Equation (2.46) yielding the bounds $K(11,1) \geq 177$ [86], resp. $K(17,1) \geq 7399$ [44].

# Chapter 3

# On the Structure of Linear Codes with Covering Radius Two and Three

## 3.1 Introduction

In the previous chapter we discussed general lower bounds on the size of covering codes with a prescribed covering radius. Most lower bounds reported for covering codes [40, 46, 47, 89, 90, 91] are extensions of the so-called Van Wee bound [84], which in turn can be viewed as a direct analogue of the well-known Johnson bound for error-correcting codes [64, p. 532]. For linear codes, a few isolated results relying on ad hoc techniques or on results from computer searches have been reported as well [8, 11, 32, 72, 88].
In this chapter we show how techniques from coding theory can be successfully applied to improve previously reported bounds for linear codes, or to prove in a simple way bounds that were established by computer searches. In particular, we prove a conjecture by Brualdi, Pless, and Wilson [8] (Bound 3.14). Almost all lower bounds reported on linear codes with covering radius two and three can also be derived, in a simpler way, using our methods. The chapter is organized as follows. In Section 3.2 we review some basic coding theory and mention some simple results regarding covering codes. In Section 3.3 we show that a linear code with covering radius two imposes restrictions on the form of its dual code. We consider restrictions on the weight enumerator of the dual code and on the intersections of different codewords. Moreover, we show that any linear code with covering radius two gives rise to a number of (not necessarily linear) codes with covering radius one that can be obtained via the dual code. Using these restrictions, it is sometimes possible to rule out the existence of a covering code, with the help of coding theory. In Section 3.4 we apply the same methods to linear codes with covering radius three. We restrict ourselves to binary codes.

## 3.2 Preliminaries

In this section we review some results from coding theory and some simple results regarding covering codes, which will be used later on.

We denote the $n \times n$ identity matrix by $I_n$, the all-zero matrix by $O$, and the all-one matrix by $J$. We recall from the previous chapter that, for any two subsets $U, V$ of $\mathbb{F}_2^n$, the set $U + V$ is defined as the set $\{u + v \mid u \in U, v \in V\}$. The set of different columns of some matrix $A$ is denoted by $\{A\}$; the vector space spanned by its rows is denoted by $\mathcal{A}$; its rank is denoted by $\mathrm{rk}(A)$. Sometimes it is useful to consider vectors that are only partially specified: we are only interested in some of their entries then. We call such a vector a template; the irrelevant positions are marked with a $\star$-entry. For conciseness, we sometimes use row vectors, where formally column vectors are appropriate, and vice versa. It is always clear from the context, however, whether a vector should be viewed as a row vector or a column vector. Also, we sometimes specify the zero-positions within a matrix by blanks.

We will need the following two results for linear codes.

**Lemma 3.1** [82] Let $C$ be a binary $[n, k, d]$ code. Then deletion of the coordinates corresponding to the nonzero positions of a codeword of weight $w < 2d$ yields a code with parameters $[n - w, k - 1, d - \lfloor w/2 \rfloor]$. □

**Lemma 3.2** [63, 64, pp. 224–225] Let $C$ be a one-weight $[n, k, d]$ code without zero-positions. Then $C$ is a concatenation of simplex codes and $d(2^k - 1) = n\, 2^{k-1}$. In particular $2^{k-1} \mid d$. □

Below we will mention some simple properties regarding covering codes that will be used in the rest of the chapter.

Let $C$ be a binary code of length $n$ with covering radius $r$.
A trivial lower bound on the size of $C$ is given by the *Sphere Covering Bound*

$$|C| \sum_{i=0}^{r} \binom{n}{i} \geq 2^n. \tag{3.1}$$

Van Wee [84] showed that this sphere covering bound can be improved to the bound

$$|C| \left\{ \sum_{i=0}^{r} \binom{n}{i} - \frac{\binom{n}{r}}{\lceil \frac{n-r}{r+1} \rceil} \left( \left\lceil \frac{n+1}{r+1} \right\rceil - \frac{n+1}{r+1} \right) \right\} \geq 2^n. \tag{3.2}$$

As a direct consequence we obtain the following result:

$$\text{If } n \text{ is } \textit{even}, \text{ then } K(n, 1) \geq 2^n/n. \tag{3.3}$$

It can be shown that the Van Wee bound (3.2) for covering codes is the direct analogue of the well-known Johnson bound [64, p. 532] for error-correcting codes: it can be derived in exactly the same way.

The exact value of $K(n,1)$ is known only in a few cases. We will use Equation (3.3), as well as the data of Figure 3.1.

| $n$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| $K(n,1)$ | 1 | 2 | 2 | 4 | 7 | 12 | 16 | 32 | 54 - 64 |

Figure 3.1: Table of $K(n,1)$ for small values of $n$, extracted from [84].

Arbitrary coverings and sphere coverings can be linked, as was shown by Blokhuis and Lam [4]. The result turns out to be very powerful in our context. We will prove a small extension of their result. First we give a definition.

**Definition 3.3** Let $S \subset \mathbb{F}_2^k$ and let $A$ be a binary $k \times n$ matrix. The set $S$ is said to $r$-cover $\mathbb{F}_2^k$ using matrix $A$ iff $\{\mathbf{s} + \mathbf{w}A^T \mid \mathbf{s} \in S \text{ and } wt(\mathbf{w}) \leq r\} = \mathbb{F}_2^k$.

It is clear from the definition that a code with covering radius (at most) $r$ corresponds to an $r$-covering using the identity matrix and that a linear code with parity check matrix $A$ and covering radius (at most) $r$ corresponds to an $r$-covering of the set $\{\mathbf{0}\}$ using matrix $A$.

**Lemma 3.4** If $S$ $r$-covers $\mathbb{F}_2^k$ using $k \times n$ matrix $A$, then the set $C := \{\mathbf{w} \in \mathbb{F}_2^n \mid \mathbf{w}A^T \in S\}$ has covering radius (at most) $r$. In particular, $K(n,r) \leq |S|\ 2^{n-k}$.

**Proof:** Let $\mathbf{x} \in \mathbb{F}_2^n$. Since $\mathbf{x}A^T \in \mathbb{F}_2^k$, we have $\mathbf{x}A^T = \mathbf{s} + \mathbf{w}A^T$ for some vector $\mathbf{s} \in S$ and some vector $\mathbf{w} \in \mathbb{F}_2^n$ with $wt(\mathbf{w}) \leq r$. It follows that $(\mathbf{x} - \mathbf{w})A^T = \mathbf{s}$, hence $d(\mathbf{x}, C) \leq r$. Therefore code $C$ has covering radius (at most) $r$.
To prove the cardinality result, we need to consider two cases. If matrix $A$ has full row-rank, then $C$ has cardinality $|C| = |S|\ 2^{n-k}$. Now suppose $\text{rk}(A) =: a < k$. The set $S$ $r$-covers $\mathbb{F}_2^k$ using matrix $A$. This property still holds if we apply a linear transformation on the elements of $S$ and on the columns of matrix $A$ simultaneously. Therefore we may assume w.l.o.g. that $A^T = (A'^T \mid O^T)$ with matrix $A'$ of full row-rank, i.e. $\text{rk}(A') = a$. For all $\mathbf{t} \in \mathbb{F}_2^{k-a}$ define the set $S(\mathbf{t}) \subseteq \mathbb{F}_2^a$ by $S(\mathbf{t}) := \{\mathbf{s} \in \mathbb{F}_2^a \mid (\mathbf{s}, \mathbf{t}) \in S\}$. Now every set $S(\mathbf{t})$ $r$-covers $\mathbb{F}_2^a$ using matrix $A'$. Matrix $A'$ has full row-rank, hence $K(n,r) \leq |S(\mathbf{t})|\ 2^{n-a}$. In fact $K(n,r) \leq |S|\ 2^{n-k}$, since $|S(\mathbf{t})| \leq |S|\ 2^{-(k-a)}$ for some $\mathbf{t} \in \mathbb{F}_2^{k-a}$. □

**Remark 3.5** In [4] the same result was proved, but only for matrices of full row-rank (so $k \leq n$). Our proof shows that this restriction is not necessary, i.e. we do *not* require that $k \leq n$.

The following lemma, though trivial, has important consequences.

**Lemma 3.6** Let $C \subseteq \mathbb{F}_2^n$ be a code with covering radius $r$. Then its extended code $\overline{C}$ has covering radius $r+1$.

**Proof:** Let $d(\mathbf{x}, C) = r$. Then $d((\mathbf{x}, 0), \overline{C}) + d((\mathbf{x}, 1), \overline{C}) = 2r + 1$, hence $\overline{C}$ has covering radius at least $r+1$. On the other hand it is trivial that the covering radius of $\overline{C}$ is at most one more than the covering radius of $C$. This proves the lemma. □

Binary codes can be slightly modified without changing the covering radius, as is demonstrated by the following lemma.

**Lemma 3.7** Let $C \subseteq \mathbb{F}_2^n$ and let $C' := \{(\sum_{i=1}^{n} c_i, c_2, \ldots, c_n) \mid (c_1, \ldots, c_n) \in C\}$. Then these codes have the same covering radius.

**Proof:** The result follows from Lemma 3.6, since $C$ and $C'$ have equivalent extended codes. □

For linear codes it is more useful to consider the following formulation of Lemma 3.7, which was already mentioned in [9] for codes with an *even* covering radius. We will frequently use this 'inversion property' later on.

**Lemma 3.8** Let $C$ and $C'$ be linear codes with parity check matrices $H = \left(\begin{array}{c|c} 1 & \mathbf{u} \\ \hline 0 & \mathbf{X} \end{array}\right)$, resp. $H' = \left(\begin{array}{c|c} 1 & \mathbf{u}+\mathbf{1} \\ \hline 0 & \mathbf{X} \end{array}\right)$. Then these codes have the same covering radius.

**Proof:** Codes $C$ and $C'$ are related via $C' = \{(\sum_{i=1}^{n} c_i, c_2, \ldots, c_n) \mid (c_1, \ldots, c_n) \in C\}$. Now the result follows from Lemma 3.7. □

**Corollary 3.9** If there exists an $[n, k]r$ code that contains a nonzero codeword of weight $w$ in its dual code, then there also exists an $[n, k]r$ code that contains a codeword of weight $(n+1) - w$ in its dual code.

**Proof:** Assume a word of weight $w$ occurs in the top row of a parity check matrix $H$ of $C$. The claim now follows from Lemma 3.8, after application of suitable row operations on this parity check matrix. □

In the proofs of our results we will always assume that our linear codes have minimum distance at least three. Here we (implicitly) use the following trivial result.

**Lemma 3.10** [8] If there exists an $[n, n-m]r$ code with length $n \leq 2^m - 1$, then there also exists an $[n, n-m]r$ code with minimum distance $d \geq 3$.

**Proof:** The covering radius of a linear code is the smallest integer $r$ such that every syndrome is the sum of at most $r$ columns of a parity check matrix of this code. Therefore deletion of zero-columns and double columns in the parity check matrix does not increase the covering radius, nor does addition of columns. □

# 3.3 Linear Codes with Covering Radius Two

In this section we derive several new lower bounds for linear codes with covering radius two. The key observation is that a linear covering code imposes restrictions on the structure of its dual code. In this way it is possible to transform the problem of designing a 'good' linear covering code into the problem of designing a (dual) linear code with a lot of structure imposed onto it. Techniques from coding theory will sometimes show that such a dual code cannot exist. We will consider restrictions on the weight distribution (one-level constraints) and restrictions on the intersections of different codewords (two-level constraints).

## 3.3.1 One-Level Constraints

Let $C$ be an $[n, m]$ code with generator matrix $H$[1]. If $C$ contains a codeword of weight $w \neq 0$, then we can put the generator matrix into the following 'standard' form:

$$H = \begin{array}{c} \left(\begin{array}{c|c} \quad A_0 \quad & \quad A_1 \quad \\ \hline 0 \;\cdots\; 0 & 1 \;\cdots\; 1 \end{array}\right) \\ \begin{array}{cc} \underset{n-w}{\longleftrightarrow} & \underset{w}{\longleftrightarrow} \end{array} \end{array} \Big\updownarrow m \; .$$

We will often refer to matrix $H$ above and its constituting submatrices $A_0$ and $A_1$.

The next lemma imposes restrictions on the weight distribution of a code with dual covering radius two.

**Lemma 3.11** Let $C$ be an $[n, m]$ code with dual covering radius two. Then the weights $w \neq 0$ in $C$ have the following properties:

1. $w\,(n + 1 - w) \geq 2^{m-1}$,
2. $w\;2^{(n-w)-(m-1)} \geq K(n - w, 1)$,
3. there exists an $[n, m]$ code $C'$ with dual covering radius two that contains a codeword of weight $(n + 1) - w$.

[1] We use the notation $H$ (instead of $G$) for the generator matrix to remind the reader that we will be working in the dual space of a code with covering radius two.

**Proof:** Suppose code $\mathcal{C}$ has a codeword of weight $w \neq 0$. Assume that generator matrix $H$ of code $\mathcal{C}$ is in the standard form, with the bottom row of weight $w$.

1. All the syndromes of $\mathcal{C}^{\perp}$ with template $(\star, 1)$ should be the sum of at most two columns of matrix $H$. Each of the $2^{m-1}$ syndromes of this form is the sum of one of the last $w$ columns of matrix $H$ and at most one of the first $n-w$ columns of this matrix; hence $w + w\,(n-w) = w\,(n+1-w) \geq 2^{m-1}$.

2. By Definition 3.3 the (transposed) columns of matrix $A_1$ 1-cover $\mathbb{F}_2^{m-1}$ using matrix $A_0$. The statement now follows from Lemma 3.4.

3. This is a reformulation of Corollary 3.9. □

**Remark 3.12** Notice that Property 1 of Lemma 3.11 is weaker than Property 2 of the same lemma, since it is implied by Property 2, using the sphere covering bound (3.1). Often significantly better bounds for $K(n,1)$ are known than the sphere covering bound, e.g. Equation (3.3).

As an application of Lemma 3.11 we derive two bounds on $l(m,r)$, viz. $l(7,2) = 19$ and $l(2m-1,2) \geq 2^m + 1$ for $m \geq 3$. The bound $l(7,2) = 19$ was proved by Ytrehus [88] using rather involved arguments and the computer. The other bound was conjectured by Brualdi, Pless, and Wilson [8], but up to now only the case $m = 6$ had been settled [9, 88]. The proofs are surprisingly simple.

**Bound 3.13** $l(7,2) = 19$.

**Proof:** Suppose $\mathcal{C}$ is an [18,7] code with dual covering radius two. By Property 1 of Lemma 3.11 we have $w(19-w) \geq 64$, i.e. $5 \leq w \leq 14$. If weight fourteen occurs, then we have $K(4,1) \leq \lfloor 14 \times 2^{-2} \rfloor = 3$ according to Property 2 of Lemma 3.11. But $K(4,1) = 4$, so weight fourteen does not occur. Similarly, if weight thirteen occurs, then we have $K(5,1) \leq 6$, in conflict with the value $K(5,1) = 7$. Hence weight thirteen does not occur either, so $5 \leq w \leq 12$. Property 3 of Lemma 3.11 imposes a further restriction on the set of nonzero weights of the code, viz. $7 \leq w \leq 12$. Since there is no $[18,7,8]$ code, cf. [7], we may assume, by Property 3 of Lemma 3.11, that $\mathcal{C}$ contains a codeword of weight twelve. If we put generator matrix $H$ of code $\mathcal{C}$ into the standard form, with the bottom row of weight twelve, then matrix $A_0$ generates a $[6,6,1]$ code (cf. Lemma 3.1), so $\mathrm{rk}(A_0) = 6$. So w.l.o.g. the following matrix is the generator matrix $H$ of code $\mathcal{C}$ with one additional column adjoined to it (at the right):

$$\overline{H} = \begin{array}{cc} \left(\begin{array}{ccc|ccc|c} & & & & & & 1 \\ & I_6 & & & A_1 & & \vdots \\ & & & & & & 1 \\ \hline 0 & \cdots & 0 & 1 & \cdots & 1 & 0 \end{array}\right) & \updownarrow 6 \\ \underleftrightarrow{\quad\quad 6 \quad\quad}\;\underleftrightarrow{\quad\quad 12 \quad\quad} & \end{array} .$$

Every row of matrix $A_1$ has weight six, since $d(C) = 7$. Since the rows of matrix $A_1$ have even weight, every sum of two different rows of matrix $A_1$ has weight six as well. Hence every row of $\overline{H}$ has a weight divisible by four and binary inner product zero with the other rows of $\overline{H}$. Therefore all weights in the code $\overline{C}$ generated by $\overline{H}$ are divisible by four, i.e. $\overline{C}$ only has two nonzero weights, viz. weights eight and twelve. It follows that $\overline{C}^{\perp}$ must be a uniformly packed code [58, p. 99], but uniformly packed codes with parameters $[19, 12, \geq 3]2$ do not exist [58, p. 105]. Since the number of nonzero weights in $\overline{C}$ is less than the dual distance, we could also have calculated the exact weight distribution of $\overline{C}$, using the MacWilliams identities or Corollary 1.7, and obtained a contradiction in that way. Evidently code $C$ does not exist, hence $l(7,2) \geq 19$.
Gabidulin et al. [30] have given a construction of a $[19,12]2$ code, so in fact $l(7,2) = 19$.

□

**Bound 3.14** $l(2m-1,2) \geq 2^m + 1$ for all $m \geq 3$.

**Proof:** Suppose $C$ is an $[n = 2^m, 2m-1]$ code with dual covering radius two. From Property 1 of Lemma 3.11 we infer that $C$ does not contain the all-one vector if $m \geq 3$. If a codeword of weight $w \neq 0$ occurs in $C$, then we have $K(v,1) \leq w\ 2^{v-(2m-2)}$, with $v + w = n$, according to Property 2 of Lemma 3.11.
If $v$ is *even*, then we have the lower bound $K(v,1) \geq 2^v/v$, cf. Equation (3.3). Thus we obtain the inequality $v\,w \geq \frac{1}{4}n^2$ for even $v$. Since $v+w = n$, in fact equality holds and hence $w = n/2 = 2^{m-1}$. We infer that the even weight subcode of $C$ of dimension $k \geq 2m-2$ is in fact a one-weight code with $d = 2^{m-1}$; hence it satisfies the divisibility constraint $2^{k-1} | \, d$ (cf. Lemma 3.2), i.e. $k \leq m$. However, for $m \geq 3$ we have $k \geq 2m-2 \geq m+1$. Therefore code $C$ does not exist. Hence $l(9,2) \geq 33$, $l(11,2) \geq 65$, $l(13,2) \geq 129$, etc.
This settles a conjecture of Brualdi, Pless, and Wilson [8]. □

Bound 3.14 improves by one the the lower bound on $l(2m-1,2)$ implied by the Van Wee bound. It is not immediately clear whether, for codimension $2m$, we can also improve on the lower bound for $l(2m,2)$ implied by the Van Wee bound. However, often an argument similar to the one given above for odd codimension gives rise to improvements. For example, the sphere covering bound gives $l(16,2) \geq 362$. If equality holds, then all the nonzero weights should be odd, which is clearly impossible. In this way we obtain the bounds $l(16,2) \geq 363$, $l(18,2) \geq 725$, $l(20,2) \geq 1449$, $l(22,2) \geq 2897$, etc., thus improving the previously known bounds by one. Sometimes simply applying the Van Wee bound already yields improvements: in this way one obtains the bounds $l(14,2) \geq 182$ and $l(24,2) \geq 5794$. If both methods do not work in their own right, then we may sometimes fruitfully combine them: if we are able to show that some linear code with dual covering radius two can only have even weights, then we can use a slightly improved version of the Van Wee bound for linear codes that contain the all-one vector. For details we refer to the previous chapter. Below we summarize the results of this paragraph.

**Bound 3.15** $l(14,2) \geq 182$, $l(16,2) \geq 363$, $l(18,2) \geq 725$, $l(20,2) \geq 1449$, $l(22,2) \geq 2897$, $l(24,2) \geq 5794$.

### 3.3.2 Two-Level Constraints

In the previous subsection we derived restrictions on the weight distribution of a code with dual covering radius two. Here we shall also consider restrictions on the intersections of codewords. This enables us to improve on some more lower bounds. Although the techniques are still easy, we have to do somewhat more work to obtain results.

The next lemma imposes restrictions on the division of the supports of codewords over the coordinate positions, when the dual covering radius equals two.

**Lemma 3.16** Let $C$ be an $[n, m]$ code with dual covering radius two. Let $(\mathbf{0}, \mathbf{1})$ be a fixed codeword of $C$ of weight $w \neq 0$. Let $\mathbf{c} = (\mathbf{c}_L, \mathbf{c}_R)$ be another fixed nonzero codeword of code $C$, partitioned in the same way as $(\mathbf{0}, \mathbf{1})$. Then the quantities $a := \min\{wt(\mathbf{c}_L), n + 1 - w - wt(\mathbf{c}_L)\}$ and $x := \min\{wt(\mathbf{c}_R), w - wt(\mathbf{c}_R)\}$ are related via

$$\left\lceil \frac{2^{m-2} - a\,w}{(n+1-w) - 2a} \right\rceil \leq x \leq \left\lfloor \frac{w}{2} \right\rfloor, \tag{3.4}$$

whenever the denominator on the lefthand side is nonzero.

**Proof:** We can put generator matrix $H$ of code $C$ into the following form:

$$H = \begin{array}{c} \left(\begin{array}{ccc|ccc|ccc|ccc} \multicolumn{6}{c|}{A_0'} & \multicolumn{6}{c}{A_1'} \\ \hline 1 & \cdots & 1 & 0 & \cdots & 0 & 1 & \cdots & 1 & 0 & \cdots & 0 \\ \hline 0 & \cdots & 0 & 0 & \cdots & 0 & 1 & \cdots & 1 & 1 & \cdots & 1 \end{array}\right) \\ \begin{array}{cccc} \xleftrightarrow{\qquad} & \xleftrightarrow{\qquad} & \xleftrightarrow{\qquad} & \xleftrightarrow{\qquad} \\ a_1 & a_2 & x_1 & x_2 \end{array} \end{array} \Bigg\updownarrow m \;,$$

where $a_1 = wt(\mathbf{c}_L)$, $x_1 = wt(\mathbf{c}_R)$, $a_1 + a_2 = n - w$, and $x_1 + x_2 = w$.
We may assume that $x_1 \leq x_2$ and $a_1 \leq a_2 + 1$, if necessary by applying the inversion property (Lemma 3.8) and by adding the bottom row of matrix $H$ to the forelast one. All the syndromes with template $(\star, 11)$ should be the sum of at most two columns of matrix $H$, so $x_1 + a_2\ x_1 + a_1\ x_2 \geq 2^{m-2}$. Since $x_1 + x_2 = w$, this is equivalent to $x_1\,(1 + a_2 - a_1) \geq 2^{m-2} - a_1\,w$. Elimination of $a_2$ via the equation $a_1 + a_2 = n - w$ now gives the inequality $x_1\,(n + 1 - w - 2a_1) \geq 2^{m-2} - a_1 w$. Since $a = a_1$ and $x = x_1$, this finishes the proof. □

**Remark 3.17** Notice that one can easily generalize Lemma 3.16 to the case where every (nonzero) syndrome of code $C^{\perp}$ is the sum of at most two columns of matrix $H$ in at least $\mu$ ways (in short: is covered $\mu$ times): in that case one should replace the quantity $2^{m-2}$ in Equation (3.4) by $\mu\,2^{m-2}$. We will not use this strenghtening of Lemma 3.16 in this section. Nevertheless, this generalization will prove to be useful in Section 3.4, when deriving the bound $l(9,3) \geq 17$ (Bound 3.31).

As an application of Lemma 3.16 we derive three bounds on $l(m,r)$, in increasing level of difficulty. First we prove the bound $l(9,2) \geq 34$.

**Bound 3.18** $l(9,2) \geq 34$.

**Proof:** Suppose $C$ is a $[33,9]$ code with dual covering radius two. By Property 1 of Lemma 3.11 we have $w(34-w) \geq 256$, i.e. $12 \leq w \leq 22$. If weight 21 occurs, then we have $K(12,1) \leq 21 \times 2^4 = 336$, according to Property 2 of Lemma 3.11. But $K(12,1) \geq \lceil 2^{12}/12 \rceil = 342$, cf. Equation (3.3), so weight twenty-one does not occur. By Property 3 of Lemma 3.11 weight thirteen does not occur either. Since $d[33,9] \leq 13$, code $C$ must have minimum distance $d(C) = 12$ and we may assume, again by Property 3 of Lemma 3.11, that a codeword of weight twenty-two occurs. Now put generator matrix $H$ of code $C$ into the standard form, with the bottom row of weight twenty-two. Since the minimum distance of $C$ is twelve, matrix $A_0$ has full row-rank; therefore we may assume w.l.o.g. that $A_0 = (I_8 \mid X)$, where $X$ is an $8 \times 3$ matrix.
Lemma 3.16 gives a relation between the weights occurring in code $\mathcal{A}_0$ and the corresponding weights occurring in $\mathcal{A}_1$. In particular we find, that if code $\mathcal{A}_0$ has a codeword of weight $1 \leq a \leq 3$, then the corresponding codeword in $\mathcal{A}_1$ has weight $x = 11$. Therefore three dependent nonzero codewords of $\mathcal{A}_0$ can never all have weights at most three, since their sum is zero. We will use this to prove that $C$ does not exist. From Lemma 3.16 we infer that if a row of $X$ occurs twice, then this row is (111). Furthermore, two rows of $X$ should differ in at least two positions, unless one of them is (111). Therefore matrix $X$ contains the row (111) at least four times. Now matrix $A_0$ contains w.l.o.g. the rows $\mathbf{a}_1 = (10000000, 111)$, $\mathbf{a}_2 = (01000000, 111)$, and $\mathbf{a}_3 = (00100000, 111)$. But this is impossible, since the codewords $\mathbf{a}_1 + \mathbf{a}_2$, $\mathbf{a}_1 + \mathbf{a}_3$, and $\mathbf{a}_2 + \mathbf{a}_3$ of code $\mathcal{A}_0$ all have weight two and are dependent. Evidently code $C$ does not exist. It follows that $l(9,2) \geq 34$. □

The bound $l(6,2) = 13$ was established by Graham and Sloane [32] using a Cray-1 computer. In [8] a non-computer proof for this bound was mentioned, but, for reasons of space, omitted. We give a proof, using the lemmas developed before and some elementary coding theory.

**Bound 3.19** $l(6,2) = 13$.

**Proof:** Suppose $C$ is a [12,6] code with dual covering radius two. From Property 1 of Lemma 3.11 we infer that $w(13-w) \geq 32$, i.e. $4 \leq w \leq 9$. Since $d[12,6] = 4$, the minimum distance of code $C$ is four and we may assume, by Property 3 of Lemma 3.11, that $C$ contains a codeword of weight nine. Now put generator matrix $H$ of code $C$ into the standard form, with the bottom row of weight nine. Lemma 3.16 relates the weights $a$ occurring in code $\mathcal{A}_0$ and the corresponding weights $x$ (with $x \leq 4$) occurring in code $\mathcal{A}_1 + \{\mathbf{0}, \mathbf{1}\}$ via

$$x = 4 \text{ for all } a \neq 2, \text{ and } 2 \leq x \leq 4 \text{ for } a = 2. \tag{3.5}$$

The structure of generator matrix $H$ depends on these relations and on the rank of matrix $A_0$. The rank of matrix $A_0$ satisfies $2 \leq \mathrm{rk}(A_0) \leq 3$. The upper bound follows from the format of the matrix; the lower bound fróm the relations of (3.5). To see this, suppose $\mathrm{rk}(A_0) \leq 1$. Then code $\mathcal{A}_1 + \{\mathbf{0}, \mathbf{1}\}$ contains a four-dimensional linear one-weight code with distance four, which contradicts Lemma 3.2. In fact, the relations of (3.5) imply that $A_0$ generates either a $[3,2,2]$ code or the vector space $\mathbb{F}_2^3$.

Now we are ready to prove the nonexistence of code $\mathcal{C}$. We will use the fact that the $[7,4,3]$ Hamming code is a perfect code and has a 2-transitive automorphism group.

Using the relations of (3.5), we can put generator matrix $H$ into the following form:

$$H = \left(\begin{array}{ccc|ccccccc|cc}
1 & & 1 & \multicolumn{7}{c|}{\mathbf{x}_1} & y_{11} & y_{12} \\
 & 1 & 1 & \multicolumn{7}{c|}{\mathbf{x}_2} & y_{21} & y_{22} \\
\hline
 & & * & 1 & 1 & 1 & 1 & & & & & \\
 & & * & 1 & 1 & & & 1 & 1 & & & \\
 & & * & 1 & & 1 & & 1 & & 1 & & \\
\hline
 & & & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1
\end{array}\right).$$

Notice that $\mathrm{rk}(A_0) = 2$ if all *-entries are zero, and $\mathrm{rk}(A_0) = 3$ otherwise. The submatrix below $\mathbf{x}_1$ and $\mathbf{x}_2$ generates a $[7,4,3]$ Hamming code with covering radius one, so we may assume w.l.o.g. that $wt(\mathbf{x}_1), wt(\mathbf{x}_2) \leq 1$, if necessary after applying suitable row operations on the generator matrix. Notice that the relations of (3.5) guarantee that the upper left $2 \times 3$ submatrix of $H$ is not affected by these linear row operations! We distinguish two cases.

If $\mathbf{x}_1 = \mathbf{0}$, then $(y_{11}, y_{12}) = (11)$ and w.l.o.g. $\mathbf{x}_2 = (1000000)$ and $(y_{21}, y_{22}) = (01)$. Here we used the 2-transitivity of the automorphism group of the Hamming code and $d(\mathcal{C}) = 4$. Now matrix $H$ is completely determined and has the following form:

$$H = \left(\begin{array}{ccc|ccccccc|cc}
1 & & 1 & & & & & & & & 1 & 1 \\
 & 1 & 1 & 1 & & & & & & & & 1 \\
\hline
 & & * & 1 & 1 & 1 & 1 & & & & & \\
 & & * & 1 & 1 & & & 1 & 1 & & & \\
 & & * & 1 & & 1 & & 1 & & 1 & & \\
\hline
 & & & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1
\end{array}\right).$$

An inspection of matrix $H$ reveals that if $\mathrm{rk}(A_0) = 3$ (hence not all *-entries are zero), then e.g. the syndrome (101111) is not covered; if $\mathrm{rk}(A_0) = 2$, then all syndromes but (011110) are covered. (Note that this implies that $l(6,2) \leq 13$.)

If $\mathbf{x}_1 \neq \mathbf{0}$, then we may assume w.l.o.g. that $\mathbf{x}_1 = (1000000)$ and $\mathbf{x}_2 = (0100000)$. Here we used the 2-transitivity of the automorphism group of the Hamming code again. (In fact the *-entries can be in any one of the first three columns, but this does not affect our argument.) Matrix $H$ has w.l.o.g. the following form:

$$H = \left(\begin{array}{ccc|ccccccc|cc}
1 & & 1 & 1 & & & & & & & y_{11} & y_{12} \\
& 1 & 1 & & 1 & & & & & & y_{21} & y_{22} \\
\hline
& & * & 1 & 1 & 1 & 1 & & & & & \\
& & * & 1 & 1 & & & 1 & 1 & & & \\
& & * & 1 & & 1 & & 1 & & 1 & & \\
\hline
& & & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1
\end{array}\right).$$

Let $\mathbf{y}_1 := (y_{11}, y_{21})$ and let $\mathbf{y}_2 := (y_{12}, y_{22})$. We show that not all the syndromes with templates $(\star, 1110)$ or $(\star, 1100)$ are covered. Suppose column four of matrix $H$ differs from column three in the third, fourth, or fifth position. (This happens e.g. when $\text{rk}(A_0) = 2$.) Since all syndromes with template $(\star, 1110)$ should be covered, we have $((10) + \{\mathbf{y}_1, \mathbf{y}_2\}) \cup \{(00), (01)\} = \mathbb{F}_2^2$, i.e. $\{\mathbf{y}_1, \mathbf{y}_2\} = \{(00), (01)\}$. Now matrix $H$ is completely determined. Since the top row of matrix $H$ has weight three and since code $C$ has minimum distance four, we obtain a contradiction. If column four of matrix $H$ does not differ from column three in the third, fourth, or fifth position, then column five of matrix $H$ does, and we can use the same argument to show a contradiction. Therefore code $C$ does not exist. It follows that $l(6, 2) = 13$. □

The smallest linear code with codimension eight and covering radius two has length $l(8, 2)$, where $24 \leq l(8, 2) \leq 26$. The lower bound was proved by Calderbank and Sloane [11]; the upper bound by Gabidulin et al. [30]. We further narrow down this gap to $25 \leq l(8, 2) \leq 26$.

The proof of the lower bound $l(8, 2) \geq 25$ demonstrates that our simple methods can be extremely powerful. The proof is rather tedious and long; therefore it is included at the end of this chapter as an appendix.

**Bound 3.20** $l(8, 2) \geq 25$.

**Remark 3.21** Our bound shows that *linear* $(24, 2^{16})2$ codes do not exist. However, there does exist a *non*linear code with these parameters! This code was constructed recently by Etzion and Greenberg by means of a Preparata code [27] (see also Construction 4.22).

## 3.4 Linear Codes with Covering Radius Three

In this section we extend the methods developed in the previous section to linear codes with covering radius three.

Let $C$ be an $[n, m]$ code with generator matrix $H$. If $C$ contains a codeword of weight $w \neq 0$, then we can put the generator matrix into the following 'standard' form:

$$H = \begin{array}{cc}
\left(\begin{array}{ccc|c|ccc}
\multicolumn{4}{c|}{\mathrm{O}} & \multicolumn{3}{c}{\mathrm{D}} \\ \hline
 & & & 0 & & & \\
 & \mathrm{A} & & \vdots & & \mathrm{E} & \\
 & & & 0 & & & \\ \hline
1 & \cdots & 1 & 1 & 0 & \cdots & 0
\end{array}\right)
&
\begin{array}{l} r_D \\ \\ r_A+1 \\ \\ \end{array}
\\
\begin{array}{cc} \qquad w \qquad & \qquad n-w \qquad \end{array} &
\end{array},$$

where $r_A + r_D + 1 = m$, $r_A = \mathrm{rk}(A) \leq w - 1$, and $r_D = \mathrm{rk}(D) \geq m - w$.
In the sequel we will often refer to matrix $H$ above and its constituting submatrices $A$, $D$, and $E$.

The next lemma imposes restrictions on the weight distribution of certain subcodes induced by a linear code with dual covering radius three.

**Lemma 3.22** Let $\mathcal{C}$ be an $[n, m]$ code with dual covering radius three. Let $\mathcal{D}_w$ be a subcode of $\mathcal{C}$ having zeros on the support of some fixed codeword of weight $w \neq 0$. Then the weights $a \neq 0$ in code $\mathcal{D}_w$ have the following properties:

1. $a\ w\ (n + 1 - w - a) \geq 2^{m-2}$,
2. $a\ w\ 2^{(n-w-a)-(m-2)} \geq K(n - w - a, 1)$,
3. $a\ (n + 1 - w - a) \geq K(w - 1, 1)2^{m-w-1}$.

**Proof:** Suppose code $\mathcal{C}$ has nonzero codewords of weights $a$ and $w$, without overlap. We can put generator matrix $H$ of code $\mathcal{C}$ into the following form:

$$H = \begin{array}{cc}
\left(\begin{array}{ccc|ccc|ccc}
 & X_1 & & & X_2 & & & X_3 & \\ \hline
0 & \cdots & 0 & 1 & \cdots & 1 & 0 & \cdots & 0 \\ \hline
1 & \cdots & 1 & 0 & \cdots & 0 & 0 & \cdots & 0
\end{array}\right)
& m
\\
\begin{array}{ccc} \qquad w \qquad & \qquad a \qquad & \qquad n-w-a \qquad \end{array} &
\end{array}.$$

1. All the syndromes of $\mathcal{C}^{\perp}$ with template $(\star, 11)$ should be the sum of at most three columns of matrix $H$, so $a\ w + a\ w\ (n - w - a)\ = a\ w\ (n + 1 - w - a) \geq 2^{m-2}$.
2. Let $X := \{X_1\} + \{X_2\}$. By Definition 3.3 the set $X$ 1-covers $\mathbb{F}_2^{m-2}$ using matrix $X_3$. The statement now follows from Lemma 3.4, using the inequality $|X| \leq a\ w$.

3. Let $X := \{X_2\} \cup (\{X_2\} + \{X_3\})$, i.e. $X$ is the set of all words $\mathbf{x} \in \mathbb{F}_2^{m-2}$ such that the syndrome $(\mathbf{x}, 10)$ is the sum of at most two columns of matrix $H$. This set has cardinality $|X| \leq a\ (n+1-w-a)$. All the syndromes with template $(\star, 11)$ should be the sum of at most three columns of matrix $H$, hence the set $X$ 1-covers $\mathbb{F}_2^{m-2}$ using the nonzero columns of matrix $X_1$. We may assume that matrix $X_1$ contains a zero-column, if necessary after applying suitable row operations on generator matrix $H$. The statement now follows from Lemma 3.4. □

Using Property 1 of Lemma 3.22 and maximizing the lefthand side with respect to variable $a$, we obtain the following corollary.

**Corollary 3.23** Let $C$ be an $[n, m]$ code with dual covering radius three. Then the weights $w \neq 0$ in code $C$ satisfy

$$w\ (n+1-w)^2 \geq 2^m \quad \text{or} \quad w \geq m.$$

**Remark 3.24** Notice that Property 1 of Lemma 3.22 is weaker than Properties 2 and 3 of the same lemma, since it is implied by both Property 2 and Property 3, using the sphere covering bound (3.1).

As an application of Lemma 3.22 we derive five bounds on $l(m,r)$, in increasing level of difficulty. The first two bounds were already established by others, but are included here to give a simpler proof. The last three bounds are new.

**Bound 3.25** $l(13,3) \geq 38$.

**Proof:** Suppose $C$ is a $[37, 13]$ code with dual covering radius three. From Corollary 3.23 we infer that $C$ has minimum distance $d(C) \geq 13$. But $d[37,13] = 12$, so evidently code $C$ does not exist. Hence $l(13,3) \geq 38$. □

**Remark 3.26** In a recent paper Zhang and Lo [90] proved the same bound with entirely different methods. However, unlike their proof, our proof is completely elementary.

**Bound 3.27** $l(9,3) \geq 16$.

**Proof:** Suppose $C$ is a $[15, 9]$ code with dual covering radius three. Corollary 3.23 implies that $d(C) \geq 4$. Since $d[15,9] = 4$, a codeword of weight four actually occurs. The maximal subcode $\mathcal{D}$ of $C$ having zeros on the support of a fixed codeword of weight four has dimension $\dim(\mathcal{D}) \geq 5$. Property 1 of Lemma 3.22 imposes a restriction on the nonzero weights $a$ of code $\mathcal{D}$, viz. $4a(12-a) \geq 128$, i.e. $4 \leq a \leq 8$. If weight five occurs, then Property 2 of the same lemma implies that $K(6,1) \leq 4 \times 5 \times 2^{-1} = 10$. But $K(6,1) = 12$, so weight five does not occur. Similarly, if weight six [seven] occurs, then we have $K(5,1) \leq 6$ $[K(4,1) \leq 3]$. Since these bounds conflict with the values $K(4,1) = 4$, resp. $K(5,1) = 7$, code $\mathcal{D}$ is an $[11, \geq 5; \{0,4,8\}]$ code. Now $\overline{\mathcal{D}} + \{\mathbf{0}, \mathbf{1}\}$ is a self-dual $[12, 6; \{0,4,8,12\}]$ code. But even self-dual codes of length twelve do not exist [64, p. 626]. Evidently code $C$ does not exist. Hence $l(9,3) \geq 16$. □

**Remark 3.28** This result was originally proved by Simonis [72]. His proof involved a manipulation of the MacWilliams identities, which resulted in a contradiction. Again, our proof is completely elementary. Later on we will improve this bound further and show that $l(9,3) \geq 17$ (Bound 3.31).

**Bound 3.29** $l(12,3) \geq 31$.

**Proof:** Suppose $C$ is a $[30,12]$ code with dual covering radius three. Corollary 3.23 implies that $d(C) \geq 8$. Let $\mathcal{D}_w$ be the maximal subcode of $C$ having zeros on the support of a fixed codeword of weight $w$.
If a codeword of weight eight occurs, then Property 3 of Lemma 3.22 implies that the nonzero weights $a$ of code $\mathcal{D}_8$ satisfy the inequality $a(23-a) \geq 2^3 \times K(7,1) = 128$, i.e. $10 \leq a \leq 13$. If $\mathcal{D}_8$ contains a codeword of weight ten, then Property 2 of Lemma 3.22 gives $K(12,1) \leq 8 \times 10 \times 2^2 = 320$, in conflict with the lower bound $K(12,1) \geq \lceil 2^{12}/12 \rceil = 342$. Hence code $\mathcal{D}_8$ does not contain codewords of weight ten. Similarly, if weight twelve occurs, then we have $K(10,1) \leq 8 \times 12 = 96$, in conflict with the lower bound $K(10,1) \geq \lceil 2^{10}/10 \rceil = 103$. It follows that every nonzero codeword in $\mathcal{D}_8$ has odd weight. This is clearly impossible, since $\mathcal{D}_8$ has dimension at least four. Hence $d(C) \geq 9$.
If a codeword of weight nine occurs, then Property 3 of Lemma 3.22 yields the inequality $a(22-a) \geq 2^2 \times K(8,1) = 128$ for the nonzero weights $a$ in code $\mathcal{D}_9$, a contradiction.
Therefore code $C$ must have minimum distance $d(C) = 10$, since $9 \leq d[30,12] \leq 10$. From Property 3 of Lemma 3.22 we infer that the nonzero weights $a$ in code $\mathcal{D}_{10}$ satisfy $a(21-a) \geq 2 \times K(9,1) \geq 2 \times 54 = 108$, i.e. $9 \leq a \leq 12$. Again, if weight ten [twelve] occurs, then Property 2 of Lemma 3.22 yields $K(10,1) \leq 100$, resp. $K(8,1) \leq 10 \times 12 \times 2^{-2} = 30$. Since these bounds conflict with the lower bounds $K(10,1) \geq 103$, resp. $K(8,1) = 32$, code $\mathcal{D}_{10}$ only has odd nonzero weights. This is impossible, since code $\mathcal{D}_{10}$ has dimension at least two. Evidently code $C$ does not exist. Hence $l(12,3) \geq 31$. □

**Bound 3.30** $l(10,3) \geq 21$.

**Proof:** Suppose $C$ is a $[20,10]$ code with dual covering radius three. From Corollary 3.23 we infer that $d(C) \geq 4$. Let $\mathcal{D}_w$ be the maximal subcode of $C$ having zeros on the support of a fixed codeword of weight $w$.
If a codeword of weight four occurs, then Property 3 of Lemma 3.22 implies that the nonzero weights $a$ of code $\mathcal{D}_4$ satisfy $a(17-a) \geq 2^5 \times K(3,1) = 64$, i.e. $6 \leq a \leq 11$. Therefore code $\mathcal{D}_4$ is a $[16, \geq 6, \geq 6]$ code. Since $d[16,6] = 6$, code $\mathcal{D}_4$ has a codeword of weight six. But then Property 2 of Lemma 3.22 gives $K(10,1) \leq 4 \times 6 \times 2^2 = 96$, in conflict with the lower bound $K(10,1) \geq \lceil 2^{10}/10 \rceil = 103$. Hence $d(C) \geq 5$.
If a codeword of weight five occurs, then Property 3 of Lemma 3.22 yields the inequality $a(16-a) \geq 2^4 \times K(4,1) = 64$ for the nonzero weights $a$ in code $\mathcal{D}_5$, so $a = 8$. Evidently $\mathcal{D}_5$ is a one-weight code with distance $d = 8$. But code $\mathcal{D}_5$ has dimension $k \geq 5$, so it does not satisfy the divisibility constraint $2^{k-1} | \, d$ that holds for one-weight codes (cf. Lemma 3.2). Therefore $d(C) \geq 6$ and, since $d[20,10] = 6$, a codeword of weight six actually occurs.

From Property 3 of Lemma 3.22 we infer that the nonzero weights $a$ of code $\mathcal{D}_6$ satisfy $a(15-a) \geq 2^3 \times K(5,1) = 56$, so $a = 7$ or $a = 8$. Hence $\overline{\mathcal{D}}_6$ is a $[15,4,8]$ code, which is the (unique) simplex code with dual distance three. Therefore all columns of matrix $D$ are distinct and nonzero, i.e. code $\mathcal{D}_6$ also has dual distance three. Now put generator matrix $H$ for code $C$ into the standard form (cf. page 47) with the bottom row of weight six. Since code $\mathcal{D}_6$ has dimension four and dual distance three, at most $6 + \binom{6}{3} = 26 < 32$ syndromes with template $(0000, \star, 1)$ are covered. Evidently code $C$ does not exist. Hence $l(10,3) \geq 21$. □

All the nonexistence proofs treated so far are based upon an application of Lemma 3.22. For the last bound, $l(9,3) \geq 17$, this is not enough. Besides the structure of code $\mathcal{D}$, we also have to take into account the structure of matrix $E$ that appears as a submatrix in generator matrix $H$ in the standard form (cf. page 47). The proof demonstrates how far one can go in analyzing linear covering codes, making use of the tools developed so far.

**Bound 3.31** $l(9,3) \geq 17$.

**Proof:** Suppose $C$ is a $[16,9]$ code with dual covering radius three.
By Corollary 3.23 we have $w \geq 9$ or $w(17-w)^2 \geq 512$, i.e. $d(C) \geq 3$. If a codeword of weight three occurs, then the (maximal) subcode having zeros on the support of this codeword has nonzero weights $a$ satisfying $a(14-a) \geq K(2,1) \times 2^5 = 64$, according to Property 3 of Lemma 3.22. This is clearly impossible, hence $d(C) \geq 4$. Since $d[16,9] = 4$, we have in fact equality. Now put generator matrix $H$ of code $C$ into the standard form, with the bottom row of weight four. We will prove that not all the syndromes with template $(\star, 1)$ are covered. In themselves, the restrictions imposed by Lemma 3.22 on the structure of code $\mathcal{D}$ that is generated by submatrix $D$ of generator matrix $H$ are not strong enough to prove nonexistence; therefore we will first derive some more detailed restrictions that hold for code $\mathcal{D}$. These more detailed restrictions will enable us to prove that code $C$ does not exist. We show that matrix $D$ has rank five and that the structure of code $\mathcal{D}$ can be (partially) determined. Furthermore, we show that it is possible to obtain some information on matrix $E$. We follow the same approach as in the proof of Bound 3.19 ($l(6,2) = 13$).

($i$) Each set of (four) columns induced by a codeword of weight four is independent.
Matrix $D$ has rank at least five. In fact equality holds. To see this, suppose matrix $D$ has rank $k \geq 6$. For all $\mathbf{x} \in \mathbb{F}_2^k \setminus \{\mathbf{0}\}$ the syndromes with template $(\mathbf{x}, \star, 1)$ should be covered, so all nonzero vectors in $\mathbb{F}_2^k$ should be the sum of one or two columns of matrix $D$. This implies that $\mathcal{D}$ is a $[12, \geq 6]$ code with dual covering radius two, in conflict with the bound $l(6,2) = 13$ (Bound 3.19). Therefore matrix $D$ has rank five. This result holds, no matter which codeword of weight four is chosen as the bottom row of generator matrix $H$. Since the total rank of matrix $H$ is nine, the claim follows.

We just proved that the submatrix $A$ of generator matrix $H$ is an invertible matrix, so we may assume w.l.o.g. that $A$ is the identity matrix, i.e. $A = I_3$.

(*ii*) The structure of matrix $E$ can be (partially) determined.
For all $\mathbf{x} \in \mathbb{F}_2^5$ define $S(\mathbf{x}) := \{\mathbf{w}E^T \mid \mathbf{w}D^T = \mathbf{x} \text{ and } wt(\mathbf{w}) \leq 2\}$. All eight syndromes with template $(\mathbf{x}, \star, 1)$ should be the sum of at most three columns of matrix $H$. By Definition 3.3 this implies that for all $\mathbf{x} \in \mathbb{F}_2^5 \setminus \{\mathbf{0}\}$ the set $S(\mathbf{x})$ 1-covers $\mathbb{F}_2^3$ using matrix $A = I_3$, i.e. $S(\mathbf{x}) \subseteq \mathbb{F}_2^3$ has covering radius one. In particular $|S(\mathbf{x})| \geq K(3,1) = 2$ and if $S(\mathbf{x})$ has cardinality two, then $S(\mathbf{x})$ is a perfect $(3,2,3)$ code.
This simple observation turns out to be very useful: if $S(\mathbf{x}) = \{\mathbf{w}_1E^T, \mathbf{w}_2E^T\}$, then matrix $E$ satisfies the linear equation $(\mathbf{w}_1 + \mathbf{w}_2)E^T = \mathbf{1}$. In this way the structure of matrix $E$ can be partially determined.

(*iii*) The structure of code $\mathcal{D}$ can be (partially) determined.
We show that the restrictions imposed upon the weights occurring in code $\mathcal{D}$ and those imposed upon the intersections between different codewords of $\mathcal{D}$ are the same as the restrictions imposed upon the $[12,6]$ code considered in the proof of Bound 3.19.
Code $\mathcal{D}$ is a $[12,5]$ code, since matrix $D$ has rank five, cf. (*i*). Since the bottom row of generator matrix $H$ has ones outside the columns of matrix $D$, we may apply the inversion property (Lemma 3.8) to matrix $D$. From Property 1 of Lemma 3.22 we infer that the nonzero weights $a$ occurring in code $\mathcal{D}$ satisfy $4a(13-a) \geq 128$, i.e. $4 \leq a \leq 9$. Since $d[12,5] = 4$, a codeword of weight four actually occurs and we may assume, by the inversion property, that $\mathcal{D}$ contains a codeword of weight nine. Now put generator matrix $D$ of code $\mathcal{D}$ into the standard form, cf. Section 3.3.1, page 41, with the bottom row of weight nine. The structure of matrix $D$ largely depends on intersection relations, similar to Lemma 3.16, between this codeword of weight nine and all other nonzero codewords of code $\mathcal{D}$. Each nonzero syndrome of $\mathcal{D}^\perp$ is the sum of at most two columns of matrix $D$ in at least two ways (in short: is covered twice), since $S(\mathbf{x})$ has cardinality at least two for all $\mathbf{x} \in \mathbb{F}_2^5 \setminus \{\mathbf{0}\}$, cf. (*ii*). Combining Lemma 3.16 and Remark 3.17 following it, we get the following (strengthened) intersection relation between the codeword of weight nine occurring in $\mathcal{D}$ and all other nonzero codewords in this code:

$$x = 4 \text{ for all } a \neq 2, \text{ and } 2 \leq x \leq 4 \text{ for } a = 2, \tag{3.6}$$

where $a$ and $x$ are defined as in Lemma 3.16. We denote the constituting submatrices of matrix $D$ by $D_0$ and $D_1$ (rather than by $A_0$ and $A_1$). The structure of matrix $D$ depends on the intersection relations (3.6) and on the rank of matrix $D_0$. The rank of the $4 \times 3$ matrix $D_0$ satisfies $1 \leq \text{rk}(D_0) \leq 3$. The upper bound follows from the format of the matrix, the lower bound from the relations of (3.6). To see this, suppose $D_0$ is the all-zero matrix. Then code $\mathcal{D}_1 + \{\mathbf{0}, \mathbf{1}\}$ contains a four-dimensional linear one-weight subcode with distance four, which contradicts Lemma 3.2.

Now we are ready to prove the nonexistence of code $C$. We distinguish four cases, depending on the rank of matrix $D_0$ and on the presence of the all-one vector in code $\mathcal{D}_0$. The proof involves, in increasing level of difficulty, the following cases:

| | |
|---|---|
| Case 1: | $\mathrm{rk}(D_0) = 1$; |
| Case 2: | $\mathrm{rk}(D_0) = 3$; |
| Case 3a: | $\mathrm{rk}(D_0) = 2$ and code $\mathcal{D}_0$ contains the all-one vector; |
| Case 3b: | $\mathrm{rk}(D_0) = 2$ and code $\mathcal{D}_0$ does not contain the all-one vector. |

In each case we prove that either $(i)$ is violated or $d(C) < 4$. In the proof syndromes are always syndromes of code $\mathcal{D}^\perp$. Each of these syndromes is covered iff it is the sum of at most two columns of matrix $D$.

**Case 1:** $\mathrm{rk}(D_0) = 1$.

The largest linear one-weight code with distance four has dimension three. Using the relations of (3.6), we infer that matrix $D_0$ generates a $[3,1,2]$ code; hence matrix $D$ has w.l.o.g. the following form:

$$D = \left(\begin{array}{ccc|ccccccc|cc}
1 & 1 & 0 & \multicolumn{7}{c|}{\mathbf{x}} & \multicolumn{2}{c}{\mathbf{y}} \\
\hline
 & & & 1 & & 1 & & 1 & & 1 & & \\
 & & & 1 & 1 & & & 1 & 1 & & & \\
 & & & 1 & 1 & 1 & 1 & & & & & \\
\hline
 & & & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1
\end{array}\right).$$

The submatrix below $\mathbf{x}$ generates the $[7,4,3]$ Hamming code. The Hamming code is a perfect code and has a 2-transitive automorphism group, so we may assume w.l.o.g. that $wt(\mathbf{x}) \le 1$, i.e. $\mathbf{x} = \mathbf{0}$ or $\mathbf{x} = (1000000)$, if necessary after applying suitable row operations on matrix $D$.

If $\mathbf{x} = \mathbf{0}$, then we have $\mathbf{y} = (11)$, since code $\mathcal{D}$ has minimum distance four. Now matrix $D$ is completely determined. The sets $S((1,111,1))$ and $S((1,111,0))$ both have cardinality two, hence both sets are perfect $(3,2,3)$ codes, cf. $(ii)$. By definition of $S(\cdot)$, the vectors $\mathbf{w}_1 = (110,0000000,00)$ and $\mathbf{w}_2 = (000,0000000,11)$ satisfy $\mathbf{w}_1 E^T = \mathbf{w}_2 E^T = \mathbf{1}$. Now $\mathbf{w}_1 + \mathbf{w}_2$ is contained in both code $\mathcal{D}$ and code $\mathcal{D}^\perp \cap \mathcal{E}^\perp$. Thus we have found a codeword of weight four in $C$ that induces a set of dependent columns in matrix $H$, in conflict with $(i)$.

If $\mathbf{x} = (1000000)$, then the syndrome $(1,111,0)$ is covered twice only if $\mathbf{y} = (00)$. Since $\mathcal{D}$ has minimum distance four, this is not possible.

**Case 2:** $\mathrm{rk}(D_0) = 3$.

Using the relations of (3.6), we infer that matrix $D$ can be put into the following form:

$$D = \left(\begin{array}{ccc|cccc|ccccc}
1 & & & & & & & & & & & \\
 & 1 & & & \multicolumn{2}{c}{X} & & & & Y & & \\
 & & 1 & & & & & & & & & \\
\hline
 & & & 1 & 1 & 1 & 1 & & & & & \\
\hline
 & & & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1
\end{array}\right).$$

Let $X := \left(\mathbf{x}_1^T, \ldots, \mathbf{x}_4^T\right)$. All syndromes with template $(\star, 11)$ are covered exactly twice. It can easily be seen that this implies that matrix $X$ contains two (possibly the same) translates of the repetition code $\{\mathbf{0}, \mathbf{1}\}$ as columns. Therefore we may assume w.l.o.g. that $\mathbf{x}_1 + \mathbf{x}_2 = \mathbf{x}_3 + \mathbf{x}_4 = \mathbf{1}$ and that every row of $X$ has even weight. The sets $S((\mathbf{x}_1, 11))$ and $S((\mathbf{x}_2, 11))$ have cardinality two, hence both sets are perfect $(3, 2, 3)$ codes, cf. $(ii)$. We claim that $\mathbf{w} = (000, 1111, 00000)$ satisfies $\mathbf{w}E^T = \mathbf{0}$. If so, $\mathbf{w}$ is contained in both code $\mathcal{D}$ and code $\mathcal{D}^\perp \cap \mathcal{E}^\perp$; thus we have found a codeword of weight four in $\mathcal{C}$ that induces a set of dependent columns in matrix $H$, in conflict with $(i)$. To prove the claim, we need to distinguish between the (essentially) two possible forms of matrix $X$, i.e. $(\mathbf{x}_1, \ldots, \mathbf{x}_4) = (\mathbf{0}, \mathbf{1}, \mathbf{1}, \mathbf{0})$ or $(\mathbf{x}_1, \ldots, \mathbf{x}_4) = (000, 111, 110, 001)$. In the first case we find, by definition of $S(\cdot)$, that the vectors $\mathbf{w}_1 = (000, 1001, 00000)$ and $\mathbf{w}_2 = (000, 0110, 00000)$ satisfy $\mathbf{w}_1 E^T = \mathbf{w}_2 E^T = \mathbf{1}$; in the other case we find that $\mathbf{w}_1 = (001, 1001, 00000)$ and $\mathbf{w}_2 = (001, 0110, 00000)$ satisfy $\mathbf{w}_1 E^T = \mathbf{w}_2 E^T = \mathbf{1}$. In both cases the vector $\mathbf{w} = \mathbf{w}_1 + \mathbf{w}_2$ satisfies $\mathbf{w}E^T = \mathbf{0}$, as claimed.

**Case 3:** $\operatorname{rk}(D_0) = 2$.

Matrix $D_0$ generates a $[3, 2]$ code, hence $\mathcal{D}_0$ has minimum distance one or two. In the first case we may assume, by the inversion property, that $\mathcal{D}_0$ contains the all-one vector. In the other case $\mathcal{D}_0$ is the $[3, 2, 2]$ even weight code. We consider both cases separately.

**Case 3a:** Code $\mathcal{D}_0$ contains the all-one vector.

Using the relations of (3.6) and the structure of matrix $D_0$, we can put matrix $D$ into the following form:

$$D = \left(\begin{array}{ccc|cccc|ccccc}
1 & & & \multicolumn{4}{c|}{X} & \multicolumn{5}{c}{Y} \\
1 & 1 & 1 & & & & & & & & & \\
\hline
 & & & 1 & 1 & & & 1 & 1 & & & \\
 & & & 1 & 1 & 1 & 1 & & & & & \\
\hline
 & & & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1
\end{array}\right).$$

Let $X := \left(\mathbf{x}_1^T, \ldots, \mathbf{x}_4^T\right)$. All syndromes with template $(\star, 11)$ are covered exactly twice. It can easily be seen that this implies that $\mathbf{x}_1 + \mathbf{x}_2 = \mathbf{x}_3 + \mathbf{x}_4 = \mathbf{1}$ and that every row of $X$ has even weight. The sets $S((\mathbf{x}_1, 111))$ and $S((\mathbf{x}_3, 011))$ have cardinality two, hence both sets are perfect $(3, 2, 3)$ codes, cf. $(ii)$. By definition of $S(\cdot)$, the vectors $\mathbf{w}_1 = (100, 1100, 00000)$ and $\mathbf{w}_2 = (100, 0011, 00000)$ satisfy $\mathbf{w}_1 E^T = \mathbf{w}_2 E^T = \mathbf{1}$. Now $\mathbf{w}_1 + \mathbf{w}_2$ is contained in both code $\mathcal{D}$ and code $\mathcal{D}^\perp \cap \mathcal{E}^\perp$. Thus we have found a codeword of weight four in $\mathcal{C}$ that induces a set of dependent columns in matrix $H$, in conflict with $(i)$.

**Case 3b:** Code $\mathcal{D}_0$ is the even weight code.

Using the relations of (3.6) and the structure of matrix $D_0$, we can put matrix $D$ into the following form:

$$D = \left(\begin{array}{ccc|cc|cc|cc|ccc}
1 & & 1 & \multicolumn{2}{c|}{X_1} & \multicolumn{2}{c|}{X_2} & \multicolumn{2}{c|}{X_3} & \multicolumn{3}{c}{X_4} \\
1 & 1 & & & & & & & & & & \\
\hline
 & & & 1 & 1 & & & 1 & 1 & & & \\
 & & & 1 & 1 & 1 & 1 & & & & & \\
\hline
 & & & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1
\end{array}\right).$$

Notice that all syndromes with templates $(\star, 11, 1)$, $(\star, 01, 1)$, and $(\star, 10, 1)$ are covered exactly twice. We will see that matrix $D$ can be completely determined. Using the structure of matrix $D$, we then derive a contradiction. First we determine the structure of matrices $X_1$, $X_2$, and $X_3$. Using this, matrix $D$ can be completely determined.

Denote the sum of the columns of a binary matrix $X$ by $\sum X$.
Let $\mathbf{a} := \sum X_1$, $\mathbf{b} := \sum X_2$, and $\mathbf{c} := \sum X_3$. We show that $\{\mathbf{0}, \mathbf{a}, \mathbf{b}, \mathbf{c}\} = \mathbb{F}_2^2$. First we prove that $\mathbf{a}, \mathbf{b}$, and $\mathbf{c}$ are distinct; then we prove that they are all nonzero.
If $\mathbf{a} = \mathbf{b}$, then we claim that the vector $\mathbf{w} = (000, 1111, 00000)$ satisfies $\mathbf{w}E^T = \mathbf{0}$. If so, $\mathbf{w}$ is contained in both code $\mathcal{D}$ and code $\mathcal{D}^\perp \cap \mathcal{E}^\perp$, since every row of matrix $(X_1|X_2)$ has even weight; thus we have found a codeword of weight four in $C$ that induces a set of dependent columns in matrix $H$, in conflict with (*i*). By symmetry the vectors $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{F}_2^2$ are all distinct then. To prove the claim, first note that w.l.o.g. $X_1 = X_2 = (\mathbf{0}^T, \mathbf{a}^T)$. The sets $S((00, 11, 1))$ and $S(((00, 01, 1))$ have cardinality two, hence both sets are perfect $(3, 2, 3)$ codes, cf. (*ii*). By definition of $S(\cdot)$, we find that for $\mathbf{a} = (10)$, the vectors $\mathbf{w}_1 = (001, 11, 00, 00000)$ and $\mathbf{w}_2 = (001, 00, 11, 00000)$ satisfy $\mathbf{w}_1 E^T = \mathbf{w}_2 E^T = \mathbf{1}$, so $\mathbf{w} = \mathbf{w}_1 + \mathbf{w}_2$ satisfies $\mathbf{w}E^T = \mathbf{0}$. The other possible values for $\mathbf{a}$ give the same result and the claim follows.
If $\mathbf{a} = \mathbf{0}$, then we may assume w.l.o.g. that $X_1 = O$. Since $\mathbf{b} \neq \mathbf{c}$ and since both vectors are nonzero, we have $\{X_2\} + \{X_3\} = \mathbb{F}_2^2$. Consequently, the syndromes $(\mathbf{z}, 11, 0)$ with $\mathbf{z} \notin \{X_4\}$ are covered only once. Since all nonzero syndromes of $\mathcal{D}^\perp$ should be covered twice, this is not possible, cf. (*ii*). Therefore $\mathbf{a} \neq \mathbf{0}$ and by symmetry we obtain $\{\mathbf{0}, \mathbf{a}, \mathbf{b}, \mathbf{c}\} = \mathbb{F}_2^2$.
By adding a suitable linear combination of the last three rows of $D$ to the first two rows, we obtain

$$X_1 = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}, \; X_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \text{ and } X_3 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

We now show that matrix $X_4$ contains three different columns. To see this, first note that $\{X_i\} + \{X_j\} = \mathbb{F}_2^2$ for all $i \neq j, 1 \leq i, j \leq 3$. Now suppose that $X_4 = (\mathbf{x}, \mathbf{x}, \mathbf{y})$. We distinguish two cases. If $\mathbf{x} \neq \mathbf{y}$, then an inspection of matrix $D$ reveals that all syndromes $(\mathbf{z}, \mathbf{x} + \mathbf{y}, 0)$ with $\mathbf{z} \neq \mathbf{x}, \mathbf{y}$ are covered only once. In the other case all syndromes $(\mathbf{z}, \mathbf{w}, 0)$ with $\mathbf{w} \neq \mathbf{0}$ and $\mathbf{z} \neq \mathbf{x}, \mathbf{w} + \mathbf{x}$ are covered only once. Since all nonzero syndromes of $\mathcal{D}^\perp$ should be covered at least twice, this is not possible, cf. (*ii*). Therefore all columns of matrix $X_4$ are distinct.

Irrespective of the particular choices left for matrix $X_4$, matrix $D$ can be put into the following form:

$$D = \left(\begin{array}{ccc|cc|cc|cc|ccc}
1 & & 1 & & 1 & & & & 1 & 1 & & \\
1 & 1 & & & 1 & & 1 & & & & & 1 \\
\hline
 & & & 1 & 1 & & & 1 & 1 & & & \\
 & & & 1 & 1 & 1 & 1 & & & & & \\
\hline
 & & & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1
\end{array}\right).$$

Using the structure of matrix $D$, we now show that our original code $C$ contains a nonzero codeword of weight at most two. Since code $C$ has minimum distance four, we then obtain a contradiction and the proof is finished.
Every syndrome $\mathbf{x} \neq \mathbf{0}$ that is covered exactly twice induces a set $S(\mathbf{x}) = \{\mathbf{w}_1 E^T, \mathbf{w}_2 E^T\}$, which is a perfect $(3, 2, 3)$ code, cf. (*ii*). In this way we obtain an equation that is satisfied by matrix $E$, viz. $\mathbf{w}E^T = \mathbf{1}$ with $\mathbf{w} := \mathbf{w}_1 + \mathbf{w}_2$.

| | syndrome | $\mathbf{x}^{(i)} = \mathbf{w}_1^{(i)} D^T = \mathbf{w}_2^{(i)} D^T$ | | $\mathbf{w}^{(i)} = \mathbf{w}_1^{(i)} + \mathbf{w}_2^{(i)}$ with |
|---|---|---|---|---|
| $i$ | $\mathbf{x}^{(i)}$ | $\mathbf{w}_1^{(i)}$ | $\mathbf{w}_2^{(i)}$ | $\mathbf{w}^{(i)} E^T = \mathbf{1}$ |
| 1 | (11111) | (100,10,00,00,000) | (000,01,00,00,000) | (100,11,00,00,000) |
| 2 | (01011) | (010,00,10,00,000) | (000,00,01,00,000) | (010,00,11,00,000) |
| 3 | (10101) | (001,00,00,10,000) | (000,00,00,01,000) | (001,00,00,11,000) |
| 4 | (11110) | (000,01,00,00,100) | (000,00,01,01,000) | (000,01,01,01,100) |
| 5 | (11010) | (000,00,01,00,010) | (000,01,00,10,000) | (000,01,01,10,010) |
| 6 | (11100) | (000,00,00,01,001) | (000,01,10,00,000) | (000,01,10,01,001) |

Figure 3.2: How certain syndromes of code $\mathcal{D}^\perp$ reveal information on matrix $E$.

Figure 3.2 shows how we can obtain linear restrictions on matrix $E$ from specific syndromes of code $\mathcal{D}^\perp$. Let $W$ be the $6 \times 12$ matrix with as rows the vectors $\mathbf{w}^{(i)}$ from Figure 3.2. By construction matrix $W$ has rank $\mathrm{rk}(W) = 6$ and satisfies the equation

$$W\ (D^T \mid E^T) = (O \mid J).$$

Evidently code $\mathcal{D}^\perp \cap \mathcal{E}^\perp$ has dimension at least five, hence code $\mathcal{D} + \mathcal{E}$ has dimension at most seven. But this is impossible: via elementary row operations on generator matrix $H$ of our original code $C$ we can now construct a nonzero codeword of weight at most two, in conflict with $d(C) = 4$. Therefore code $C$ does not exist. Hence $l(9, 3) \geq 17$. □

## 3.5 Appendix

**Proof** of Bound 3.20 ($l(8, 2) \geq 25$):

Suppose $C$ is a $[24,8]$ code with dual covering radius two. From Property 1 of Lemma 3.11 we infer that $w(25-w) \geq 128$, i.e. $8 \leq w \leq 17$. Since $d[24,8] = 8$, the minimum distance of code $C$ is eight and we may assume, by Property 3 of Lemma 3.11, that a codeword of weight seventeen occurs. Now put generator matrix $H$ of code $C$ into the standard form (cf. page 41), with the bottom row of weight seventeen. Lemma 3.16 relates the weights $a$ occurring in code $\mathcal{A}_0$ and the corresponding weights $x$ (with $x \leq 8$) occurring in code $\mathcal{A}_1 + \{\mathbf{0}, \mathbf{1}\}$ via

$$\begin{array}{rll} x = 8 & \text{for} & a = 0,1,2,6,7; \\ 7 \leq x \leq 8 & \text{for} & a = 3,5; \\ 4 \leq x \leq 8 & \text{for} & a = 4. \end{array} \tag{3.7}$$

The structure of generator matrix $H$ largely depends on these relations and on the structure of matrix $A_0$. We distinguish between the possible ranks of matrix $A_0$. In the actual nonexistence proof for $C$ we will use the fact that

$$\mathcal{A}_0^{\perp} \text{ is a subcode of the } [7,4,3] \text{ Hamming code.} \tag{3.8}$$

Equation (3.8) is obvious when matrix $A_0$ has rank $\mathrm{rk}(A_0) = 7$, since then this matrix spans the vector space $\mathbb{F}_2^7$. If $\mathrm{rk}(A_0) < 7$, then (3.8) can be seen as follows. Consider the case $\mathrm{rk}(A_0) = 6$. Using the relations of (3.7), we can put generator matrix $H$ into the following form:

$$H = \left( \begin{array}{ccc|ccc|ccc} & A_0' & & & X & & & Y & \\ \hline 0 & \cdots & 0 & 1 & \cdots & 1 & 0 & \cdots & 0 \\ \hline 0 & \cdots & 0 & 1 & \cdots & 1 & 1 & \cdots & 1 \end{array} \right), \quad \text{where } X \text{ is a } 6 \times 8 \text{ matrix.}$$

All the syndromes with template $(\star, 11)$ should be covered, hence by Definition 3.3 the columns of matrix $X$ 1-cover $\mathbb{F}_2^6$ using the $6 \times 7$ matrix $A_0'$. From Lemma 3.4 we infer that the code $\mathcal{P} := \{\mathbf{x} \in \mathbb{F}_2^7 \mid A_0'\mathbf{x}^T \in \{X\}\}$ has covering radius one. Since $\mathcal{P}$ has cardinality sixteen, this code is a perfect $(7,16,3)$ code, i.e. a coset of the $[7,4,3]$ Hamming code. Code $\mathcal{P}$ consists of a number of cosets of $\mathcal{A}_0^{\perp}$, hence (3.8) follows. A similar argument proves (3.8) for matrices $A_0$ of rank $\mathrm{rk}(A_0) < 6$. □

For notational convenience we define matrix $AG_2(m)$ as a matrix with as columns all the vectors of $\mathbb{F}_2^m$. The result of puncturing this matrix on the zero-position is denoted by $PG_2(m-1)$. The latter matrix generates the simplex code with distance $2^{m-1}$. In the sequel we will not use the particular order of the columns of these matrices, so the ambiguity in the definitions does not cause any problems.

Now we are ready to prove the nonexistence of code $C$. We distinguish seven cases, depending on the rank of matrix $A_0$ and on the presence of the all-one vector in code $\mathcal{A}_0$. The proof involves, in increasing level of difficulty, the following cases:

Case 1: $\mathrm{rk}(A_0) = 7$;
Case 2a: $\mathrm{rk}(A_0) = 6$ and code $\mathcal{A}_0$ contains the all-one vector;
Case 2b: $\mathrm{rk}(A_0) = 6$ and code $\mathcal{A}_0$ does not contain the all-one vector;
Case 3: $\mathrm{rk}(A_0) = 3$;
Case 4: $\mathrm{rk}(A_0) = 4$;
Case 5a: $\mathrm{rk}(A_0) = 5$ and code $\mathcal{A}_0$ contains the all-one vector;
Case 5b: $\mathrm{rk}(A_0) = 5$ and code $\mathcal{A}_0$ does not contain the all-one vector.

We will always assume that generator matrix $H$ has distinct columns, cf. Lemma 3.10.

**Case 1:** $\mathrm{rk}(A_0) = 7$.

We may assume that $A_0 = I_7$ and that every row of matrix $A_1$ has even weight (apply suitable row operations on the generator matrix). From the relations of (3.7) we infer that every row of matrix $A_1$, and every sum of two different rows of the same matrix, has weight eight. In particular, the binary inner product of any two rows of $A_1$ is zero. It follows that all words of code $\mathcal{A}_1$ have weights divisible by four, so $\mathcal{A}_1$ is a $[17, 7; \{0, 4, 8, 12\}]$ code. Consequently, the relations of (3.7) between the weights $a$ occurring in code $\mathcal{A}_0$ and the corresponding weights $x$ in code $\mathcal{A}_1$ can be sharpened to

$$\begin{array}{ccc} x = 8 & \text{for all} & a \neq 4; \\ x = 4, 8, 12 & \text{for} & a = 4. \end{array} \tag{3.9}$$

(Notice that we do not restrict to $x \leq 8$ here.)
Generator matrix $H$ of code $C$ can be put into the following form:

$$H = \left(\begin{array}{ccc|c|ccc|ccc} & & & 0 & & & & & & \\ & I_6 & & \vdots & & X & & & Y & \\ & & & 0 & & & & & & \\ \hline 0 & \cdots & 0 & 1 & 1 & \cdots & 1 & 0 & \cdots & 0 \\ \hline & & & & 1 & \cdots & 1 & 1 & \cdots & 1 \end{array}\right), \quad \text{where } X \text{ is a } 6 \times 8 \text{ matrix.}$$

From the relations of (3.9) we infer that every row of matrix $Y$, and every sum of two different rows of the same matrix, has weight four. In particular, the binary inner product of any two rows of $Y$ is zero. It follows that all words in the linear span of matrix $Y$ have weights divisible by four. Therefore, all nonzero words in the linear span of matrix $Y$ have weights four or eight. In fact weight eight cannot occur. To see this, suppose $wt(\mathbf{a}Y) = 8$ for some $\mathbf{a} \in \mathbb{F}_2^6$. Code $C$ contains both the codeword $(\mathbf{a}, 0, \mathbf{a}X, \mathbf{a}Y)$ and the codeword $(\mathbf{a}, 1, \mathbf{1} + \mathbf{a}X, \mathbf{a}Y)$. Using the relations of (3.9), we find that $\mathbf{a}X = \mathbf{0}$ or $\mathbf{a}X = \mathbf{1}$, i.e. there exists a codeword of weight sixteen in code $\mathcal{A}_1$. This contradicts the relations of (3.9), however. Therefore matrix $Y$ generates a one-weight code with distance four, as does matrix $X$, using similar arguments. The largest linear one-weight code with distance four has dimension three, cf. Lemma 3.2, hence matrices $X$ and $Y$ both generate $[9, 3; \{0, 4\}]$ codes. From Lemma 3.2 we infer that matrix $Y$ has two zero-columns. Since not all columns of generator matrix $H$ are different, this contradicts our assumptions (cf. Lemma 3.10).

**Case 2** $\mathrm{rk}(A_0) = 6$.

Matrix $A_0$ generates a $[7,6]$ code, hence $\mathcal{A}_0$ has minimum distance one or two. In the first case we may assume, by the inversion property, that $\mathcal{A}_0$ contains the all-one vector. From (3.8) we infer that $\mathcal{A}_0^\perp$ is a subcode of the $[7,4,3]$ Hamming code, i.e. $\mathcal{A}_0$ contains the $[7,3,4]$ simplex code as a subcode. Since code $\mathcal{A}_0$ also contains the all-one vector, it contains the $[7,4,3]$ Hamming code as a subcode. In the other case $\mathcal{A}_0$ is the $[7,6,2]$ even weight code. We consider both cases separately.

**Case 2a:** Code $\mathcal{A}_0$ contains the Hamming code.

The structure of generator matrix $H$ for code $\mathcal{C}$ largely depends on the relations of (3.7) and on the structure of code $\mathcal{A}_0$. First we derive some more detailed constraints, however. From (3.7) we infer that the generator matrix $H$ of code $\mathcal{C}$ can be put into the form

$$H = \left(\begin{array}{c|c|c|c} \qquad I_6 \qquad & \begin{matrix} * \\ \vdots \\ * \end{matrix} & \qquad X \qquad & \qquad Y \qquad \\ \hline 0 \;\cdots\; 0 & 0 & 1 \;\cdots\; 1 & 0 \;\cdots\; 0 \\ \hline & & 1 \;\cdots\; 1 & 1 \;\cdots\; 1 \end{array}\right), \quad \text{where } X \text{ is a } 6 \times 8 \text{ matrix.}$$

We prove that matrix $X$ generates w.l.o.g. a one-weight code with distance four.

From the relations of (3.7) we infer that every row of matrix $X$ has weight four. Every sum of two different rows of this matrix has weight four as well. This also follows from the relations of (3.7), with some effort: let $(\mathbf{a}, \mathbf{x}, \mathbf{y})$ be a codeword of $\mathcal{C}$ with $\mathbf{a} \in \mathbb{F}_2^7$, $\mathbf{x} \in \mathbb{F}_2^8$, and $\mathbf{y} \in \mathbb{F}_2^9$. Notice that $(\mathbf{a}, \mathbf{1}+\mathbf{x}, \mathbf{y})$ is contained in code $\mathcal{C}$ as well. If $wt(\mathbf{a}) = 2$, then $wt(\mathbf{x}) = 4$ follows from the relations of (3.7). If $wt(\mathbf{a}) = 3$, then the relations of (3.7) imply the inequalities $7 \leq wt(\mathbf{x}) + wt(\mathbf{y}) \leq 10$ and $7 \leq (8 - wt(\mathbf{x})) + wt(\mathbf{y}) \leq 10$, i.e. $3 \leq wt(\mathbf{x}) \leq 5$. Hence the sum of two different rows of matrix $X$ has weight four, since every row of matrix $X$ has even weight. Notice that the inner product of two different rows of $X$ is zero. Therefore all weights in the linear span of matrix $X$ are divisible by four and we may assume that $X$ generates a one-weight code with distance four. The largest linear one-weight code with distance four has dimension three, so we may assume that matrix $X$ has rank at most three.

Using the structure of matrices $A_0$ and $X$ and the relations of (3.7), we can put generator matrix $H$ of code $C$ into the following form:

$$H = \left(\begin{array}{ccccccc|cccccccc|ccccccccc}
1 & 1 & 1 & 1 & & & & & & & & & & & & & & & & & & & & \\
1 & 1 & & & 1 & 1 & & \multicolumn{8}{c|}{O} & \multicolumn{9}{c}{Y'} \\
1 & & 1 & & 1 & & 1 & & & & & & & & & & & & & & & & & \\
\hline
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & & 1 & & 1 & & 1 & & 1 & & 1 & & 1 & & 1 & & \\
1 & & & & & & & 1 & 1 & & & 1 & 1 & & & 1 & 1 & & & 1 & 1 & & & \\
 & 1 & & & & & & 1 & 1 & 1 & 1 & & & & & 1 & 1 & 1 & 1 & & & & & \\
 & & & & & & & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & & & & & & & & & \\
\hline
 & & & & & & & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1
\end{array}\right).$$

Notice that we used the fact that the Hamming code is a perfect code and has a two-transitive automorphism group.

All eight syndromes with template $(\star, 0110, 0)$ should be covered. An inspection of matrix $H$ reveals that at most seven syndromes with this template are actually covered, a contradiction.

**Case 2b:** Code $\mathcal{A}_0$ is the even weight code.

We may assume that $A_0 = (I_6 \mid \mathbf{1}^T)$ and that every row of matrix $A_1$ has even weight. Similar to Case 1 ($\text{rk}(A_0) = 7$), we find that $\mathcal{A}_1$ is a $[17, 7; \{0, 4, 8, 12\}]$ code and that the relations of (3.9) hold. From (3.9) we infer that generator matrix $H$ of code $C$ can be put into the form

$$H = \left(\begin{array}{c|c|c|c}
 & 1 & & \\
I_6 & \vdots & X & Y \\
 & 1 & & \\
\hline
0 \ \cdots \ 0 & 0 & 1 \ \cdots \ 1 & 0 \ \cdots \ 0 \\
\hline
 & & 1 \ \cdots \ 1 & 1 \ \cdots \ 1
\end{array}\right), \quad \text{where } X \text{ is a } 6 \times 8 \text{ matrix.}$$

As in Case 1, all weights in the linear span of matrices $X$ and $Y$ are divisible by four. The largest linear one-weight code with distance four has dimension three, so matrices $X$ and $Y$ have rank at most four. We assumed that all columns of generator matrix $H$ are different (cf. Lemma 3.10). From Lemma 3.2 we infer that matrix $Y$, and hence matrix $A_1$, contains the zero-column exactly once. Puncturing matrix $A_1$ on this zero-position yields a code with dual distance $d' \geq 3$. Since the number of nonzero weights in code $\mathcal{A}_1$ is three and $d' \geq 3$, we can compute the weight enumerator of code $\mathcal{A}_1$ using the MacWilliams identities or Corollary 1.7. This results in the weight enumerator $W(z) = 1 + 13z^4 + 99z^8 + 15z^{12}$. In particular, the number of codewords of weight eight in code $\mathcal{A}_1$ is 99. Since $99 > 2 \times (64 - \binom{7}{4}) = 58$, there exists a codeword $(\mathbf{a}, \mathbf{x}) \in C$ with $\mathbf{x} \in \mathcal{A}_1$, $wt(\mathbf{x}) = 8$, and $wt(\mathbf{a}) = 4$. We may assume w.l.o.g. that $\mathbf{a} = (0001111)$.

Using this, the relations of (3.9) and the structure of matrix $X$, we can put generator matrix $H$ of code $C$ into the following form:

$$H = \left(\begin{array}{ccccccc|cccccccc|ccccccccc}
1 & & & & 1 & 1 & 1 & & & & & & & & & & & & & & & & & \\
 & 1 & & 1 & & 1 & 1 & & & & O & & & & & & & & & Y' & & & & \\
 & & 1 & 1 & 1 & & 1 & & & & & & & & & & & & & & & & & \\
\hline
 & & & 1 & & & 1 & 1 & & 1 & & 1 & & 1 & & 1 & & 1 & & 1 & & 1 & & \\
 & & & & 1 & & 1 & 1 & 1 & & & 1 & 1 & & & 1 & 1 & & & 1 & 1 & & & \\
 & & & & & 1 & 1 & 1 & 1 & 1 & 1 & & & & & 1 & 1 & 1 & 1 & & & & & \\
 & & & & & & & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & & & & & & & & & \\
\hline
 & & & & & & & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1
\end{array}\right).$$

It is not immediately evident that the upper left submatrix of $H$ indeed generates a $[7,3,4]$ code. This follows indirectly from the relations of (3.9): matrix $Y'$ cannot contain any codeword of weight eight, since otherwise there would be a codeword of weight sixteen in code $\mathcal{A}_1$, in conflict with the relations of (3.9). Hence $Y'$ generates a $[9,3;\{0,4\}]$ code. Using the relations of (3.9) once again, we find that the upper left submatrix of generator matrix $H$ indeed generates the $[7,3,4]$ simplex code.
Matrix $Y' := (\mathbf{y}_1,\ldots,\mathbf{y}_9)$ generates a $[9,3;\{0,4\}]$ code. From Lemma 3.2 we infer that $Y'$ has two zero-columns. The nonzero binary triples occur exactly once as columns of matrix $Y'$. Since all syndromes with template $(\star, 0000, 0)$ should be covered, we have $\mathbf{y}_8 + \mathbf{y}_9 = \mathbf{1}$. Similarly, considering all syndromes with template $(\star, 0110, 0)$ yields the set equation $\{011, 100, 000\} \cup \{\mathbf{y}_1 + \mathbf{y}_7, \mathbf{y}_2 + \mathbf{y}_8, \mathbf{y}_2 + \mathbf{y}_9, \mathbf{y}_3 + \mathbf{y}_5, \mathbf{y}_4 + \mathbf{y}_6\} = \mathbb{F}_2^3$. Since these sets have the same cardinality and no element occurs more than once, the summation over the elements of either set should yield the same result. Matrix $Y'$ contains all nonzero binary triples exactly once, hence $\mathbf{y}_2 = (111)$. The same exercise for all syndromes with template $(\star, 1010, 0)$ yields $\mathbf{y}_3 = (111)$, so $\mathbf{y}_2 = \mathbf{y}_3$. Since matrix $Y'$ contains all nonzero triples exactly once, we obtain a contradiction.

**Case 3:** $\mathrm{rk}(\mathcal{A}_0) = 3$.

From (3.8) and the rank of matrix $A_0$ we infer that $\mathcal{A}_0^\perp$ is a subcode of the $[7,4,3]$ Hamming code of dimension four. Therefore matrix $A_0$ generates the dual of the Hamming code, i.e. the $[7,3,4]$ simplex code. Using this and the relations of (3.7), we can put generator matrix $H$ of code $C$ into the following form:

$$H = \left(\begin{array}{c|c|c}
PG_2(2) & X & Y \\
\hline
 & PG_2(3) & \\
\hline
 & 1\ 1\ \cdots\ 1\ 1 & 1\ 1
\end{array}\right).$$

Observe that all syndromes with template $(\star, 1)$ are covered. We now show that matrix $X$ generates a code with minimum distance at least five. Suppose the bottom row of matrix $X$ has ones in the positions corresponding to the different vectors $\{\mathbf{x}_i \mid 1 \le i \le s\}$ in the $4 \times 15$ matrix $PG_2(3)$ below $X$. Notice that the relations of (3.7) imply that $s \ge 2$. We distinguish two cases. If the bottom row of matrix $Y$ has weight two, then the four syndromes with template $(\star, 1, \mathbf{x}_1, 0)$ can only be covered if $s \ge 5$. If the bottom row of matrix $Y$ has weight zero or one, then $s \ge 5$ also holds, since otherwise at most three syndromes with template $(\star, 1, \mathbf{x}_1 + \mathbf{x}_2, 0)$ are covered. We see that in both cases the bottom row of matrix $X$ has weight at least five. Since we did not use the structure of the upper left submatrix of $H$, in fact $X$ generates a linear code with distance at least five. Moreover, all codewords in the linear span of $X$ have distance at least five to the (punctured) Reed-Muller code below $X$. Hence the submatrix induced by the fifteen column positions of $X$ (i.e. columns 8 to 22) generates a linear $(15, 2^8, 5)$ code. But there exists only one code with these parameters, viz. the *non*linear Preparata code, a contradiction.

**Case 4:** $\mathrm{rk}(A_0) = 4$.

Matrix $A_0$ generates a $[7,4]$ code, hence $\mathcal{A}_0$ has minimum distance one, two, or three. Using the inversion property (Lemma 3.8), we may assume that code $\mathcal{A}_0$ contains a codeword $\mathbf{a} \in \mathbb{F}_2^7$ of weight two or seven. From (3.8) we infer that $\mathcal{A}_0^\perp$ is a subcode of the $[7,4,3]$ Hamming code, hence $\mathcal{A}_0$ contains the 2-transitive $[7,3,4]$ simplex code as a subcode. Using this and the relations of (3.7), we infer that generator matrix $H$ of code $C$ can be put into the following form:

$$H = \left(\begin{array}{c|c|c|c} PG_2(2) & \multicolumn{2}{c|}{X} & Y \\ \hline \mathbf{a} & 1 \ \cdots \ 1 & 0 \ \cdots \ 0 & \\ \cline{1-3} & AG_2(3) & PG_2(2) & \\ \hline & 1 \ \cdots \ 1 & 1 \ \cdots \ 1 & 1 \quad 1 \end{array}\right).$$

We now show that matrix $X$ generates a code with minimum distance at least five. Suppose the bottom row of matrix $X$ has ones in the positions corresponding to the different vectors $\{\mathbf{x}_i \mid 1 \le i \le s\}$ in the $4 \times 15$ matrix $PG_2(3)$ directly below $X$. We distinguish two cases. If the bottom row of matrix $Y$ has weight two, then $s \ge 2$ and we may assume that $\mathbf{x}_1 \ne (1000)$. Therefore the four templates $(\star, 1, \mathbf{x}_1, 0)$ can only be covered if $s \ge 5$. If the bottom row of matrix $Y$ has weight zero or one, then $s \ge 3$ and we may assume w.l.o.g. that $\mathbf{x}_1 + \mathbf{x}_2 \ne (1000)$. Now the four syndromes with template $(\star, 1, \mathbf{x}_1 + \mathbf{x}_2, 0)$ can only be covered if $s \ge 5$. We see that in both cases the bottom row of matrix $X$ has weight at least five. Since we did not use the structure of the upper left submatrix of $H$, in fact $X$ generates a linear code with distance at least five. Moreover, all codewords in the linear span of $X$ have distance at least five to the (punctured) Reed-Muller code below $X$. Hence the submatrix induced by the fifteen column positions of $X$ (i.e. columns 8 to 22) generates a linear $(15, 2^8, 5)$ code. But there exists only one code with these parameters, viz. the *non*linear Preparata code, a contradiction.

**Case 5:** $\mathrm{rk}(A_0) = 5$.

From (3.8) we infer that $\mathcal{A}_0^\perp$ is a subcode of the $[7,4,3]$ Hamming code, hence $\mathcal{A}_0$ contains the 2-transitive $[7,3,4]$ simplex code as a subcode. Matrix $A_0$ generates a $[7,5]$ code, hence $\mathcal{A}_0$ has minimum distance one or two. By the inversion property (Lemma 3.8) we may assume that code $\mathcal{A}_0$ either contains a codeword of weight seven or has minimum distance two. We consider both cases separately.

**Case 5a:** Code $\mathcal{A}_0$ contains the all-one vector.

Code $\mathcal{A}_0$ contains the 2-transitive $[7,3,4]$ simplex code as a subcode and contains the all-one vector, hence it contains the perfect Hamming code as a subcode. Using this and the relations of (3.7), we can put generator matrix $H$ of code $C$ into the following form:

$$H = \left(\begin{array}{ccccccc|cccc|cccc|cccc|ccccc}
1&1&1&1&&&&&&&&&&&&&&&&&&&&\\
1&1&&&1&1&&\multicolumn{4}{c|}{X_1}&\multicolumn{4}{c|}{X_2}&\multicolumn{4}{c|}{X_3}&\multicolumn{5}{c}{X_4}\\
1&&1&&1&&1&&&&&&&&&&&&&&&&&\\
\hline
1&1&1&1&1&1&1&1&&1&&1&&1&&1&&1&&1&&1&&\\
1&&&&&&&1&1&&&1&1&&&1&1&&&1&1&&&\\
\cline{8-19}
&&&&&&&1&1&1&1&&&&&1&1&1&1&&&&&\\
&&&&&&&1&1&1&1&1&1&1&1&&&&&&&&&\\
\hline
&&&&&&&1&1&1&1&1&1&1&1&1&1&1&1&1&1&1&1&1
\end{array}\right).$$

We will show that this matrix can be transformed to a parity check matrix of an $[18,11]2$ code. Since $l(7,2) = 19$, we then obtain a contradiction.

Let $X_1 := (\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d})$. All syndromes with template $(\star, 1111, 1)$ should be covered, so $\{\mathbf{a}, \mathbf{d}+\mathbf{1}\} \cup (\mathbf{b} + (\mathbb{F}_2^3 \setminus \{\mathbf{0}, \mathbf{1}\})) = \mathbb{F}_2^3$. Since these sets have the same cardinality and no element occurs more than once, the summation over the elements of either set should yield the same result, hence $\mathbf{a} = \mathbf{d}$. The same argument for the syndromes with template $(\star, 0111, 1)$ yields $\mathbf{b} = \mathbf{c}$. By symmetry, this result holds for matrices $X_2$ and $X_3$ as well. Hence the first/last, resp. second/third, column of each of the matrices $X_1, X_2, X_3$ are equal. Now by adding the fourth row of matrix $H$ to the fifth one and by subsequently deleting the fourth row and double columns in the resulting matrix, we get a $7 \times 18$ parity check matrix of a linear code with covering radius two. Since $l(7,2) = 19$ (Bound 3.13), this is not possible.

**Case 5b:** Code $\mathcal{A}_0$ has minimum distance two.

Since code $\mathcal{A}_0$ has minimum distance two, its dual code $\mathcal{A}_0^\perp$ does not contain a zero-position. Using this and (3.8), i.e. $\mathcal{A}_0^\perp$ is a subcode of the $[7,4,3]$ Hamming code, we infer that $\mathcal{A}_0$ is uniquely determined and has parity check matrix

$$A_0^\perp = \left(\begin{array}{cccc|ccc} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ \hline & & & & 1 & 1 & 1 \end{array}\right).$$

In particular, all codewords of $\mathcal{A}_0$ have even weight.
From the relations of (3.7) and the form of matrix $A_0^\perp$ we infer that generator matrix $H$ of code $\mathcal{C}$ can be put into the following form:

$$H = \left(\begin{array}{ccccccc|cccc|cccc|cccc|ccc|cc}
\multicolumn{7}{c|}{Z} & \multicolumn{15}{c|}{X} & \multicolumn{2}{c}{Y} \\
\hline
1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & & 1 & & 1 & & 1 & & 1 & & 1 & & 1 & & 1 & & \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & & & 1 & 1 & & & 1 & 1 & & & 1 & 1 & & & \\
\cline{8-22}
 & & & & & & & 1 & 1 & 1 & 1 & & & & & 1 & 1 & 1 & 1 & & & & & \\
 & & & & & & & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & & & & & & & & & \\
\hline
 & & & & & & & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1
\end{array}\right).$$

Let $(\mathbf{x},\mathbf{y})$ be a word in the linear span of matrix $(X|Y)$ with $\mathbf{y} \in \mathbb{F}_2^2$ and $wt(\mathbf{y}) = 1$. We will prove that this situation never occurs. Since this implies that the last two columns of generator matrix $H$ are equal, we then obtain a contradiction, using Lemma 3.10.

By $\mathcal{R}^*(1,4)$ we denote the punctured Reed-Muller code generated by the $5 \times 15$ matrix below $X$; its extended code will be denoted by $\mathcal{R}(1,4)$.
The proof consists of three steps. In the first step we prove that $d(\mathbf{x}, \mathcal{R}^*(1,4)) \leq 4$. In the second step we show that there can be at most one such a word $(\mathbf{x},\mathbf{y})$ with $wt(\mathbf{y}) = 1$ in the linear span of matrix $(X|Y)$. In the third and last step we show that such words cannot exist.
Before we give the actual proof, we mention some facts on Reed-Muller codes that will be used in the first step of the proof: the Reed-Muller code $\mathcal{R}(1,4)$ with parameters $[16,5,8]$ has covering radius six. If $d(\mathbf{u}, \mathcal{R}(1,4)) = 6$, then $\mathbf{u}$ represents a *bent function* and all distances $d(\mathbf{u},\mathbf{c})$ between $\mathbf{u}$ and some codeword $\mathbf{c} \in \mathcal{R}(1,4)$ are either six of ten. The punctured Reed-Muller code $\mathcal{R}^*(1,4)$ with parameters $[15,5,7]$ has covering radius five. For details we refer to Remark 1.14 or to [64, p. 418].

**Step 1** Let $(\mathbf{x},\mathbf{y})$ be a word in the linear span of matrix $(X|Y)$ with $\mathbf{y} \in \mathbb{F}_2^2$ and $wt(\mathbf{y}) = 1$. Let $(\mathbf{a},\mathbf{x},\mathbf{y})$ be the corresponding codeword of $\mathcal{C}$. Suppose $d(\mathbf{x}, \mathcal{R}^*(1,4)) \geq 5$. The punctured Reed-Muller code $\mathcal{R}^*(1,4)$ with parameters $[15,5,7]$ has covering radius five, hence we may assume that $wt(\mathbf{x}) = 5$. Notice that adding a suitable linear combination of the bottom five rows of generator matrix $H$ is allowed, since this does not affect $wt(\mathbf{y})$. The vector $(\mathbf{x},1)$ has distance six to the Reed-Muller code $\mathcal{R}(1,4)$, hence all distances between $(\mathbf{x},1)$ and the codewords of $\mathcal{R}(1,4)$ are either six or ten. Using the relations of (3.7), we infer that the vectors $\mathbf{a}$, $\mathbf{a}+(1100000)$, $\mathbf{a}+(1010000)$, and $\mathbf{a}+(0110000)$ all have weights 3, 4, or 5. Actually all these vectors have weight four, since they are all codewords of code $\mathcal{A}_0$ and all weights in this code are even. The vector $\mathbf{a}$ has weight four, so the support of $\mathbf{a}$ intersects the supports of the vectors (1100000), (1010000), and (0110000) in exactly one position. It follows that $\mathbf{a}$ has exactly $1\frac{1}{2}$ ones in the first three positions, which is clearly impossible. Evidently $d(\mathbf{x}, \mathcal{R}^*(1,4)) \leq 4$.

**Step 2** Suppose the bottom row of matrix $X$ has ones in the positions corresponding to the different vectors $\{\mathbf{x}_i \mid 1 \leq i \leq s\}$ in the $4 \times 15$ matrix $PG_2(3)$ directly below $X$ and suppose the bottom row of matrix $Y$ has weight one. Notice that the relations of (3.7) imply that $s \geq 3$. If $\mathbf{z} \in \{\mathbf{x}_i + \mathbf{x}_j \mid i \neq j\} \setminus \{(1000), (0100), (1100)\}$, then the four syndromes with template $(\star, 1, \mathbf{z}, 0)$ can only be covered if $s \geq 5$. Similarly, if $\mathbf{z} \in \{\mathbf{x}_i \mid 1 \leq i \leq s\} \setminus \{(1000), (0100), (1100)\}$, then the four syndromes with template $(\star, 1, \mathbf{z}, 0)$ can only be covered if $s \geq 4$. It follows, that if the bottom row of matrix $X$ has weight three, then $\{\mathbf{0}, \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\} = \mathbb{F}_2^2 \times \{(00)\}$. Similarly, if the bottom row of matrix $X$ has weight four, then $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4\} = \mathbb{F}_2^2 \times \{(10)\}$, say. Since we did not use the structure of upper left submatrix $Z$, and since the bottom row of matrix $Y$ has weight one, in fact all vectors $(\mathbf{x}, \mathbf{y})$ in code $\mathcal{A}_1 + \{\mathbf{0}, \mathbf{1}\}$ with $\mathbf{y} \in \mathbb{F}_2^2$ and $wt(\mathbf{y}) = 1$ have this structure, whenever $wt(\mathbf{x}) \leq 4$.

Now suppose $(\mathbf{x}, \mathbf{y})$ and $(\mathbf{x}', \mathbf{y}')$ are distinct words in the linear span of matrix $(X|Y)$ with $wt(\mathbf{y}) = wt(\mathbf{y}') = 1$. In Step 1 we proved that $d(\mathbf{x}, \mathcal{R}^*(1,4)) \leq 4$ and $d(\mathbf{x}', \mathcal{R}^*(1,4)) \leq 4$. Hence we may assume that both $wt(\mathbf{x})$ and $wt(\mathbf{x}')$ are three or four. Now we use the analysis of the structure of vectors of weight three or four given above to obtain a nonzero codeword in $\mathcal{A}_1 + \{\mathbf{0}, \mathbf{1}\}$ of weight less than four, in conflict with the relations of (3.7). Notice that it is enough to show that $\mathbf{x} + \mathbf{x}' \in \mathcal{R}^*(1,4)$, because then $\mathcal{A}_1 + \{\mathbf{0}, \mathbf{1}\}$ contains a nonzero codeword of weight at most two. It is easy to check that indeed $\mathbf{x} + \mathbf{x}' \in \mathcal{R}^*(1,4)$ for all possible choices of $wt(\mathbf{x})$ and $wt(\mathbf{x}')$, since their structure is known. If $wt(\mathbf{x}) = wt(\mathbf{x}') = 3$, then $\mathbf{x} = \mathbf{x}'$. If $wt(\mathbf{x}) = wt(\mathbf{x}') = 4$, then either $\mathbf{x} = \mathbf{x}'$ or $wt(\mathbf{x} + \mathbf{x}') = 8$ and $\mathbf{x} + \mathbf{x}' \in \mathcal{R}^*(1,4)$. If $wt(\mathbf{x}) = 3$ and $wt(\mathbf{x}') = 4$, then $\mathbf{1} + \mathbf{x} + \mathbf{x}'$ has weight eight and is contained in $\mathcal{R}^*(1,4)$.

Therefore, the linear span of matrix $(X|Y)$ contains at most one vector $(\mathbf{x}, \mathbf{y})$ with $\mathbf{y} \in \mathbb{F}_2^2$ and $wt(\mathbf{y}) = 1$.

**Step 3** Suppose $(\mathbf{x}, \mathbf{y})$ is a vector in the linear span of matrix $(X|Y)$ with $\mathbf{y} \in \mathbb{F}_2^2$ and $wt(\mathbf{y}) = 1$. Matrix $Y$ is a $3 \times 2$ matrix and has rank at most two, hence there is a vector $(\mathbf{x}', \mathbf{y})$, with $\mathbf{x} \neq \mathbf{x}'$, in the linear span of matrix $(X|Y)$. This contradicts the result of Step 2, however. Therefore there are no such vectors. □

# Chapter 4

# Constructions for Covering Codes

## 4.1 Introduction

In the previous two chapters we studied lower bounds on the size of covering codes. In Chapter 2 we discussed lower bounds for covering codes, which usually had a direct analogue with an upper bound for error-correcting codes. In Chapter 3 we studied the structure of a linear covering code by means of its dual code. Most results were obtained using techniques from coding theory.
In this chapter we will be interested in finding constructions for covering codes that have few codewords, given their length and covering radius. We will see that, again, one can use constructions for error-correcting codes to design good covering codes. In particular, we will show that one of the constructions for error-correcting codes — a generalization of the direct sum construction — is in particular useful to make some exceptionally good covering codes. We give some other constructions as well. Before we give constructions, we define a measure by which one can judge the quality of a covering code.

## 4.2 Quality Measures for Covering Codes

A trivial lower bound on the size of a covering code is the sphere covering bound. In the previous two chapters we discussed other lower bounds as well (for binary codes). Here we give an upper bound on the minimum size of covering codes. The bound is nonconstructive; it can be obtained using a probabilistic argument.

**Lemma 4.1** Let $K(n,r)$ be the minimum cardinality of any $q$-ary code of length $n$ with covering radius $r$. Then

$$\frac{q^n}{V_q(n,r)} \leq K(n,r) \leq \frac{q^n}{V_q(n,r)}(1+\ln V_q(n,r)).$$

**Proof:** Let $\mathcal{C}$ be a $q$-ary code of length $n$ and let $\mathcal{C}^+ := \{\mathbf{y} \in \mathbb{F}_q^n \mid d(\mathbf{y},\mathcal{C}) > r\}$. By definition, code $\mathcal{D} := \mathcal{C} \cup \mathcal{C}^+$ has covering radius (at most) $r$. Let $E(\mathcal{C}^+)$ be the average cardinality of code $\mathcal{C}^+$ over all codes $\mathcal{C} := \{\mathbf{x}_1, \ldots, \mathbf{x}_M\}$, where $\mathbf{x}_1, \ldots, \mathbf{x}_M$ are arbitrary vectors of $\mathbb{F}_q^n$. By estimating in two ways the cardinality of the set

$$T := \{(\mathbf{y},\mathcal{C}) \mid d(\mathbf{y},\mathcal{C}) > r,\ \mathcal{C} = \{\mathbf{x}_1, \ldots, \mathbf{x}_M\},\ \mathbf{x}_i \in \mathbb{F}_q^n\ (1 \leq i \leq M)\},$$

we find that $E(\mathcal{C}^+) = q^n(1 - q^{-n}V_q(n,r))^M$. Since $\mathcal{D} := \mathcal{C} \cup \mathcal{C}^+$ has covering radius at most $r$, it follows that there exists a $q$-ary code $\mathcal{D}$ of length $n$ with covering radius $r$ and cardinality

$$|\mathcal{D}| \leq M + q^n(1 - q^{-n}V_q(n,r))^M.$$

Now take $M := kp^{-1}$, where $p = q^{-n}V_q(n,r)$ and $k = \ln V_q(n,r)$. A simple calculation shows that

$$|\mathcal{D}| \leq kp^{-1} + q^n(1-p)^{kp^{-1}} \leq kp^{-1} + q^n e^{-k} = (k+1)p^{-1} = \frac{q^n}{V_q(n,r)}(1 + \ln V_q(n,r)).$$

□

**Remark 4.2** A slightly weaker result was already obtained by Cohen and Frankl [15]. If one restricts oneself to *linear* codes, one obtains, perhaps rather surprisingly, almost the same results, see [14] and [3, 15]. One finds, e.g., that the minimum dimension $k$ of any binary $[n,k]r$ code satisfies the inequality

$$n - \log_2 V(n,r) \leq k \leq \lceil n - \log_2 V(n,r) + \log_2(n-k) \rceil.$$

The next theorem is a consequence of Lemma 4.1.

**Theorem 4.3** If $0 \leq \rho \leq \frac{q-1}{q}$, $n \to \infty$, and $r/n \to \rho$, then $\frac{1}{n}\log_q K(n,r) \to 1 - H_q(\rho)$. Here $H_q(x)$ is the $q$-ary entropy function, which is defined by $H_q(0) := 0$ and $H_q(x) := x\log_q(q-1) - x\log_q x - (1-x)\log_q(1-x)$ if $0 < x \leq \frac{q-1}{q}$.

**Proof:** This follows from Lemma 4.1, using the well-known result [58, p. 55] that $\frac{1}{n}\log_q V_q(n,r) \to H_q(\rho)$, if $0 \leq \rho \leq \frac{q-1}{q}$, $n \to \infty$, and $r/n \to \rho$. □

The theorem indicates that, from the point of view of information theory, the capacity region of covering codes is known. It follows that for codes with a fixed information rate the minimum achievable covering radius is asymptotically known. Therefore, an asymptotical measure is not very useful. Instead, we consider another measure for the quality of a covering code, its so-called density.

Let $C$ be a $q$-ary code of length $n$ and let $t \geq 0$. The density $\mu(C,t)$ of $C$ is defined as the average number of codewords that is at distance at most $t$ from a word in the vector space $\mathbb{F}_q^n$, i.e. $\mu(C,t) = q^{-n}|C| \cdot V_q(n,t)$.

Let $\mu_d(n,r) := \min\{\mu(C,r) \mid C \text{ is an } (n,|C|,d)r \text{ code}\}$ and let $\mu_d(r) := \liminf_{n\to\infty} \mu_d(n,r)$. The quantity $\mu_d(n,r)$ reflects the minimum density achievable by any code of length $n$, with minimum distance $d$, and covering radius $r$. When we are not interested in the actual minimum distance of the codes, we delete the distance parameter $d$ as a subscript of the density functions.

By definition, a code $C$ with covering radius $r$ has density $\mu(C,r) \geq 1$. When designing covering codes, one aims for codes with a small cardinality, i.e. codes with a low density. Conversely, if $C$ is $e$-error correcting code of length $n$, then $\mu(C,e) \leq 1$, with equality if and only if $C$ is a perfect code. If $\{C_m\}_{m=1}^{\infty}$ is an infinite sequence of $e$-error correcting codes and $\mu_e(C_m) \to 1$ whenever $m \to \infty$, we call this sequence of codes asymptotically perfect codes. A trivial example of a sequence of (asymptotically) perfect codes is the sequence of Hamming codes. Another example is the class of Preparata codes $\{\mathcal{P}_m\}$, where $\mathcal{P}_m$ has parameters $(2^m-1, 2^{2^m-2m}, 5)3$ and $m \geq 4$ is an even integer. For reasons of symmetry, we call a sequence of codes $\{C_m\}$ with covering radius $r$ asymptotically perfect if $\mu_r(C_m) \to 1$ whenever $m \to \infty$.

Notice that the non-constructive upper bound does *not* give us a fixed upper bound on the density of coverings: we only obtain $\mu_r(n) \leq 1 + \ln V_q(n,r)$. Kabatyanskii [51] proved that $\lim_{n\to\infty} \mu_1(n) = 1$. For the perfect Hamming codes we have equality. Interestingly, for $r > 1$ it is not known whether $\lim_{n\to\infty} \mu_r(n)$ exists.

In the rest of this chapter we will explicitly construct classes of codes with (limiting) densities close to one. We will measure the quality of a covering code by its density alone, thus neglecting complexity questions related to encoding/decoding. It turns out, however, that the codes we construct here often allow implementations with low encoding/decoding cost. These implementations can easily be derived from the various constructions.

## 4.3 The Direct Sum Construction; Generalizations

One of the easiest ways to combine two codes is simply to take their direct sum.

**Definition 4.4** Let $C_1$ and $C_2$ be $q$-ary codes. The direct sum of $C_1$ and $C_2$ is the code $\mathcal{D} := C_1 \times C_2 = \{(\mathbf{u},\mathbf{v}) \mid \mathbf{u} \in C_1, \mathbf{v} \in C_2\}$.

It is trivial that code $\mathcal{D}$ has minimum distance $d(\mathcal{D}) = \min\{d(C_1), d(C_2)\}$ and covering radius $r(\mathcal{D}) = r(C_1)+r(C_2)$. Therefore, the direct sum construction, though simple, generally yields codes with a poor minimum distance and a poor covering radius. However, sometimes one can show that code $\mathcal{D}$ contains a proper subcode with better distance properties than $\mathcal{D}$ itself and (almost) the same covering radius as code $\mathcal{D}$. These proper subcodes can be obtained via a generalization of the direct sum construction. This construction was introduced by Sloane et al. [73] and was mentioned as construction X4 in [64, Chapter 18]. The same construction was reintroduced by Honkala [42] as the *blockwise direct sum* construction. Examples will demonstrate that this generalization yields some codes that are among the best known.

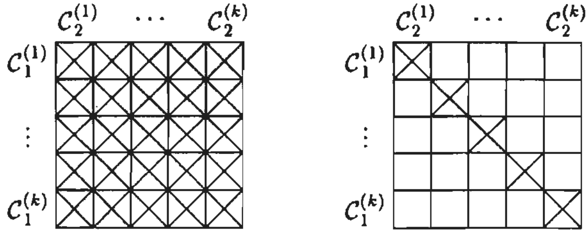

Figure 4.1: The direct sum versus the blockwise direct sum.

**Definition 4.5** Let $C_1$ and $C_2$ be $q$-ary codes, each the union of $k$ subcodes $C_1^{(1)}, \ldots, C_1^{(k)}$, resp. $C_2^{(1)}, \ldots, C_2^{(k)}$. The blockwise direct sum (BDS) of codes $C_1$ and $C_2$ with respect to these subcodes is the code $\mathcal{D} := \cup\{C_1^{(i)} \times C_2^{(i)} \mid 1 \leq i \leq k\}$ — see Figure 4.1.

**Remark 4.6** The subcodes of $C_1$ or $C_2$ can be labelled in $k!$ ways, hence one can form the blockwise direct sum of $C_1$ and $C_2$ in $k!$ ways. The case $k = 1$ corresponds to taking the direct sum of two codes. In contrast to [64, Chapter 18], we do not require the subcodes of $C_1$ or $C_2$ to be disjoint. Usually we will consider partitions, however.

If $C$ is the union of $k$ translates of subcode $C'$, we denote this by $C/C'$. Parameter $k$ can be computed from the cardinalities of codes $C$ and $C'$ or follows from the context (if the subcodes are not disjoint). If $C/C'$ is a partition that is isomorphic to a vector space over $\mathbb{F}_q$, then we call $C/C'$ a linear partition.

The next theorem gives some useful properties of blockwise direct sums.

**Theorem 4.7** Let $\mathcal{D}$ be the blockwise direct sum of the partitions $C_1/C_1'$ and $C_2/C_2'$. Then $\mathcal{D}$ has the following properties:

1. If $C_1$ and $C_2'$ (or $C_2$ and $C_1'$) are systematic codes, then $\mathcal{D}$ is a systematic code.
2. If $C_1'$ is invariant under the translation $\mathbf{x} \rightarrow \mathbf{x} + \mathbf{a}$ and $C_2'$ is invariant under the translation $\mathbf{y} \rightarrow \mathbf{y} + \mathbf{b}$, then $\mathcal{D}$ is invariant under the translation $(\mathbf{x}, \mathbf{y}) \rightarrow (\mathbf{x}, \mathbf{y}) + (\mathbf{a}, \mathbf{b})$.
3. If $C_1$, $C_1'$, $C_2$, and $C_2'$ are linear codes, then $\mathcal{D}$ can be made linear.

4. If $\mathcal{C}_1/\mathcal{C}_1'$ and $\mathcal{C}_2/\mathcal{C}_2'$ are linear partitions, then $(\mathcal{C}_1 \times \mathcal{C}_2)/\mathcal{D}$ can be made a linear partition.

**Proof:** Trivial. □

The next theorem gives a bound on the minimum distance of the blockwise direct sum of two codes.

**Theorem 4.8** [64, p. 585] Let $\mathcal{C}_1$ and $\mathcal{C}_2$ be $q$-ary codes, each the union of $k$ subcodes $\mathcal{C}_1^{(1)}, \ldots, \mathcal{C}_1^{(k)}$, resp. $\mathcal{C}_2^{(1)}, \ldots, \mathcal{C}_2^{(k)}$. Suppose that $\mathcal{C}_1$ has minimum distance $d_1$ and that all its disjoint subcodes have minimum distances at least $d_{11}$. Furthermore, suppose that the respective distances for code $\mathcal{C}_2$ and its disjoint subcodes are $d_2$ and $d_{22}$. Then the blockwise direct sum $\mathcal{D}$ of these two codes has minimum distance $d(\mathcal{D}) \geq \min\{d_{11}, d_{22}, d_1 + d_2\}$.

**Proof:** Trivial. □

**Remark 4.9** Notice that the subcodes of $\mathcal{C}_1$ and $\mathcal{C}_2$ are assumed to be disjoint. If one drops this restriction, Theorem 4.8 remains valid, provided every code is viewed as a multiset, every union is viewed as a union of multisets, and the minimum distance of a multiset is defined to be zero if any word is contained more than once in this multiset. Since we seldom use codes with intersecting subcodes as input to the BDS-construction, we refrain from giving formal definitions here.

The blockwise direct sum construction can be used to obtain new 2-error-correcting codes from old ones. As an example, we consider a construction due to Sloane et al. [73]. The codes one obtains via this construction are among the best known.

**Construction 4.10** [64, p. 587] Let $m \geq 4$ be divisible by four. The Hamming code $\mathcal{H}_m$ of length $2^m - 1$ can be partitioned into $2^{m-1}$ translates of the Preparata code $\mathcal{P}_m$, see [64, p. 474]. The vector space $I\!F_2^n$, with $n = \sqrt{2^m} - 1$, can be partitioned into $2^{m-1}$ translates of the Preparata code $\mathcal{P}_{(m/2)}$ of length $n$. If we apply the BDS-construction with as inputs the partitions $\mathcal{H}_m/\mathcal{P}_m$ and $\overline{I\!F_2^n/\mathcal{P}}_{(m/2)}$, then the resulting code has parameters $(2^m + \sqrt{2^m} - 1, 2^{(2^m+\sqrt{2^m}-1)-2m}, 5)$. The minimum distance follows from Theorem 4.8 and the minimum distances of the Hamming code and the Preparata code, which are three, resp. five.

Theorem 4.8 gives a bound on the minimum distance of the blockwise direct sum of two codes. It is possible to give a bound on the covering radius as well. This bound depends on a notion, called the $k$-norm.

**Definition 4.11** Let $\mathcal{C}$ be a $q$-ary code of length $n$. Let $\mathcal{C}$ be the union of $k$ subcodes $\mathcal{C}^{(1)}, \ldots, \mathcal{C}^{(k)}$, which do not necessarily have to be disjoint. The $k$-norm $N$ of code $\mathcal{C}$ with respect to subcodes $\mathcal{C}^{(1)}, \ldots, \mathcal{C}^{(k)}$ is the maximum value of $\min_i d(\mathbf{x}, \mathcal{C}^{(i)}) + \max_j d(\mathbf{x}, \mathcal{C}^{(j)})$ over all $\mathbf{x} \in \mathbb{F}_q^n$.

**Theorem 4.12** [42] Suppose that $\mathcal{C}_1 \subset \mathbb{F}_q^{n_1}$ has $k$-norm $N_1$ with respect to subcodes $\mathcal{C}_1^{(1)}, \ldots, \mathcal{C}_1^{(k)}$ and that $\mathcal{C}_2 \subset \mathbb{F}_q^{n_2}$ has $k$-norm $N_2$ with respect to subcodes $\mathcal{C}_2^{(1)}, \ldots, \mathcal{C}_2^{(k)}$. Then the blockwise direct sum $\mathcal{D}$ of these two codes has covering radius $r \leq \lfloor (N_1 + N_2)/2 \rfloor$. Moreover, if $k = 2$, then $\mathcal{D}$ has 2-norm $N_1 + N_2$ with respect to subcodes $\mathcal{C}_1^{(1)} \times \mathcal{C}_2^{(1)}$ and $\mathcal{C}_1^{(2)} \times \mathcal{C}_2^{(2)}$.

**Proof:** Let $\mathbf{z} = (\mathbf{x}, \mathbf{y}) \in \mathbb{F}_q^{n_1} \times \mathbb{F}_q^{n_2}$. Let $\alpha$ and $\beta$ be defined by $d(\mathbf{x}, \mathcal{C}_1) = d(\mathbf{x}, \mathcal{C}_1^{(\alpha)})$ and $d(\mathbf{y}, \mathcal{C}_2) = d(\mathbf{y}, \mathcal{C}_2^{(\beta)})$. Then $d(\mathbf{z}, \mathcal{C}_1^{(\alpha)} \times \mathcal{C}_2^{(\alpha)}) + d(\mathbf{z}, \mathcal{C}_1^{(\beta)} \times \mathcal{C}_2^{(\beta)}) = \big(d(\mathbf{x}, \mathcal{C}_1^{(\alpha)}) + d(\mathbf{x}, \mathcal{C}_1^{(\beta)})\big) + \big(d(\mathbf{y}, \mathcal{C}_2^{(\alpha)}) + d(\mathbf{y}, \mathcal{C}_2^{(\beta)})\big) \leq N_1 + N_2$. It follows that $d(\mathbf{z}, \mathcal{D}) \leq \lfloor N_1 + N_2)/2 \rfloor$. The other statement follows from the observation that for $k = 2$ the upper bound $N_1 + N_2$ can be attained (with $\alpha \neq \beta$). □

**Remark 4.13** In fact, Theorem 4.12 also holds under slightly weaker conditions. To this end, Honkala [42] introduced a slightly weaker notion than $k$-normality, the so-called $(k, t)$-subnormality. Since this refinement does not yield stronger results when applying the BDS-construction, we do not describe this refinement in detail.

**Lemma 4.14** Let $\mathcal{C}$ be a $q$-ary code with covering radius $r$ which is the union of $k$ subcodes $\mathcal{C}^{(1)}, \ldots, \mathcal{C}^{(k)}$, each with covering radius at most $r'$. Then $\mathcal{C}$ has $k$-norm $N \leq r + r'$ with respect to $\mathcal{C}^{(1)}, \ldots, \mathcal{C}^{(k)}$.

**Proof:** This follows from Definition 4.11 and the fact that $d(\mathbf{x}, \mathcal{C}) = \min_i d(\mathbf{x}, \mathcal{C}^{(i)})$. □

We will apply the BDS-construction primarily to binary codes. The following observation will prove to be useful later on.

**Lemma 4.15** Let $\mathcal{C}$ be a binary code with $k$-norm $N$ with respect to subcodes $\mathcal{C}^{(1)}, \ldots, \mathcal{C}^{(k)}$. Then $\overline{\mathcal{C}}$ has an *even* $k$-norm $\overline{N}$ with respect to subcodes $\overline{\mathcal{C}}^{(1)}, \ldots, \overline{\mathcal{C}}^{(k)}$, and $\overline{N} = N + 1$ or $\overline{N} = N + 2$.

**Proof:** An extended binary code has an even $k$-norm, since it only contains codewords of even weight. Since $d((\mathbf{x}, 0), \overline{\mathcal{C}}^{(i)}) + d((\mathbf{x}, 1), \overline{\mathcal{C}}^{(i)}) = 2d(\mathbf{x}, \mathcal{C}^{(i)}) + 1$ for all $i, 1 \leq i \leq k$, it follows from Definition 4.11 that the $k$-norm of $\overline{\mathcal{C}}$ is at least $N + 1$. On the other hand it is trivial that the $k$-norm of $\overline{\mathcal{C}}$ is at most $N + 2$. This proves the lemma. □

**Remark 4.16** For $k = 1$ this lemma reduces to Lemma 3.6.

### 4.3.1 Some $k$-Norms

Before we can apply Theorem 4.12, we need to find the $k$-norms of some specific codes.

Let $C$ be a $q$-ary code of length $n$ with covering radius $r$. The vector space $\mathbb{F}_q^n$ is the union of $k$ translates of code $C$, for some $k \leq V_q(n,r) = |B_r(\mathbf{0})|$. The upper bound on $k$ follows from the definition of the covering radius, which implies that $C + B_r(\mathbf{0}) = \mathbb{F}_q^n$. Consequently, $\mathbb{F}_q^n/C$ has $k$-norm $r$, for some $k \leq V_q(n,r)$. Sometimes one can improve this bound on $k$ considerably, especially when $C$ partitions $\mathbb{F}_q^n$, e.g. when $C$ is a linear or a systematic code. In these cases we find $k = q^n/|C|$.

In particular, we obtain the following two norms:

**Norm 4.17** The vector space $\mathbb{F}_2^n$, with $n = 2^m - 1$, can be partitioned into $n+1$ translates of the Hamming code $\mathcal{H}_m$. Since the Hamming code has covering radius one, we find that $\mathbb{F}_2^n/\mathcal{H}_m$ has $2^m$-norm 1. More generally, if $C$ is a block code of length $n$ with covering radius one, then $\mathbb{F}_2^n/C$ has $(n+1)$-norm 1.

**Norm 4.18** Let $C$ be a binary code with parameters $(n, 2^{n-m})2$ that partitions the vector space $\mathbb{F}_2^n$ into $2^m$ translates. Since $C$ has covering radius two, we find that $\mathbb{F}_2^n/C$ has $2^m$-norm 2. More generally, if $\mathbb{F}_2^n$ is the union of $2^m$ translates of a block code $C$ with covering radius two, then $\mathbb{F}_2^n/C$ has $2^m$-norm 2. In particular, this holds if $n+1 \leq \sqrt{2^{m+1}}$, since then $V(n,2) \leq 2^m$.

The next norm will prove to be the key to many constructions for covering codes with a low density.

**Norm 4.19** For all even $m \geq 4$, the Hamming code $\mathcal{H}_m$ of length $n = 2^m - 1$ can be partitioned into $2^{m-1}$ translates of the Preparata code $\mathcal{P}_m$. The $2^{m-1}$-norm of $\mathcal{H}_m/\mathcal{P}_m$ can be determined using the following property of Preparata codes: if $d(\mathbf{x}, \mathcal{P}_m) = 3$, then $\mathbf{x} \in \mathcal{H}_m$, and if $\mathbf{x} \notin \mathcal{H}_m$, then $d(\mathbf{x}, \mathcal{P}_m) < 3$. From this property we infer that $\mathcal{H}_m/\mathcal{P}_m$ has $2^{m-1}$-norm 3.

**Remark 4.20** The partition $\mathcal{H}_m/\mathcal{P}_m$ has $2^{m-1}$-norm 3. Based upon the covering radii of the Hamming code and the Preparata code alone, one would have estimated this norm to be four. This explains the usefulness of Norm 4.19.

The next norm can be determined from the covering radii of the codes involved.

**Norm 4.21** The Hamming code $\mathcal{H}_m$ of length $n = 2^m - 1$ can be partitioned into $2^m$ cosets of the double-error-correcting *BCH*-code $BCH_m(5)$. This *BCH*-code has covering radius three, since it is a quasi-perfect code, cf. [64, p. 279]. The Hamming code has covering radius one, hence $\mathcal{H}_m/BCH_m(5)$ has $2^m$-norm 4.

### 4.3.2 Some Blockwise Direct Sums

In this section we give some examples of the BDS-construction. We will use the $k$-norms computed in the previous section. In each instance of the BDS-construction, we merely indicate which codes are the inputs to this construction. The minimum distances and covering radii of the constructed codes follow by a straightforward application of Theorems 4.8 and 4.12, using the $k$-norms mentioned before. In all cases it can readily be verified that the bound on the covering radius is tight. Unless stated otherwise, the constructed codes have better parameters than previously known.

We will frequently use the functions $n_d(m,r)$, which denotes the smallest integer $n$ such that an $(n, 2^{n-m})r$ code exists. If we restrict ourselves to systematic codes, we use the notation $n_d^*(m,r)$ instead.

**Construction 4.22** Let $m \geq 4$ be even. The partition $\mathcal{H}_m/\mathcal{P}_m$ has $2^{m-1}$-norm 3. The partition $\mathbb{F}_2^n/\mathcal{H}_{m-1}$, where $n = 2^{m-1} - 1$, has $2^{m-1}$-norm 1. Using Lemma 4.15, we infer that the partition $\overline{\mathbb{F}_2^n}/\overline{\mathcal{H}}_{m-1}$ has $2^{m-1}$-norm 2. If we apply the BDS-construction with as inputs the partitions $\mathcal{H}_m/\mathcal{P}_m$ and $\overline{\mathbb{F}_2^n}/\overline{\mathcal{H}}_{m-1}$, then the resulting code $\mathcal{D}_{2m}$ has parameters $(2^m + 2^{m-1} - 1, 2^{(2^m+2^{m-1}-1)-2m}, 4)2$. Therefore $n_4^*(2m,2) \leq \frac{3}{2}2^m - 1$. The density of this class of codes satisfies $\mu(\mathcal{D}_{2m},2) \rightarrow 1\frac{1}{8}$, if $m \rightarrow \infty$. Hence $\mu_4(2) \leq 1\frac{1}{8}$.

**Remark 4.23** This bound is essentially due to Etzion et al. [27]. Although they used a perfect mixed code with covering radius two to construct these codes, their method can be described similar to Construction 4.22: their codes result from the BDS-construction with as inputs the partitions $\overline{\mathcal{H}}_m/\overline{\mathcal{P}}_m$ and $\mathbb{F}_2^n/\mathcal{H}_{m-1}$. The parameters of the resulting codes are the same as those for the codes obtained via Construction 4.22, except for the minimum distance, which is only three in their construction.

The codes one obtains via Construction 4.22 are nonlinear, since they use the Preparata code as a building block. In general, it is not known whether codes with these parameters exist that are linear. For $m = 4$ the answer is negative: Construction 4.22 yields the bound $n_4^*(8,2) \leq 23$. Calderbank and Sloane [11] proved that linear $[23,15]2$ codes do not exist. In fact, $[24,16]2$ codes do not exist either, as was proved by Struik [79]. The best-known linear codes with redundancy eight have been constructed by Gabidulin et al. [30]. They have parameters $[26,18,3]2$ or $[28,20,4]2$, depending upon the minimum distance. For $m = 6$, Construction 4.22 yields the bound $n_4^*(12,2) \leq 95$. The best-known linear codes with redundancy twelve have parameters $[107,95,3]2$ or $[117,105,4]2$, depending upon the minimum distance. These codes have been constructed by Davydov et al. [23], resp. Gabidulin et al. [30]. On the other hand, $n^*(12,2) \geq 91$, since any code with parameters $(90, 2^{78})2$ is a perfect code and perfect codes with these parameters do not exist. Moreover, equality in this bound cannot be attained by a linear code, as was proved by Struik [80]. These examples demonstrate that Construction 4.22 can be extremely powerful.

The codes one obtains via Construction 4.22 have density approximately $1\frac{1}{8}$. Up to now this has been the best result known for codes with covering radius two. The next construction shows that one can do even better: we will construct codes with covering radius two that asymptotically have density 1, i.e. which are asymptotically perfect codes!

**Construction 4.24** Let $m \geq 4$ be even. The partition $\mathcal{H}_m/\mathcal{P}_m$ has $2^{m-1}$-norm 3. Let $\mathcal{C}$ be an $(n_d, 2^{n_d-(m-1)}, d)2$ code which partitions $\mathbb{F}_2^{n_d}$ into $2^{m-1}$ translates. The partition $\mathbb{F}_2^{n_d}/\mathcal{C}$ has $2^{m-1}$-norm 2. If we apply the BDS-construction with as inputs these two partitions, then the resulting code $\mathcal{D}^{(d)}_{2m-1}$ has parameters $(2^m+n_d-1, 2^{(2^m+n_d-1)-(2m-1)}, d)2$. Therefore $n_d^*(2m-1,2) \leq (2^m-1)+n_d$. Now we consider some specific choices for code $\mathcal{C}$. Consider the linear codes with covering radius two and odd codimension constructed by Gabidulin et al. [30]. These codes have parameters $[n, n-(m-1), d]2$, where

$$n = \begin{cases} \frac{5}{4}\sqrt{2^m} - 1 & \text{if } d = 3 \text{ or } d = m = 4, \\ \frac{3}{2}\sqrt{2^m} - 3 & \text{if } d = 4 \text{ and } m \geq 6, \\ \frac{23}{16}\sqrt{2^m} - 3 & \text{if } d = 4 \text{ and } m \geq 10. \end{cases} \tag{4.1}$$

Since these codes are all linear, code $\mathcal{C}$ can be chosen to be any of these codes. Thus we obtain the bounds $n_3^*(2m-1,2) \leq 2^m + \frac{5}{4}\sqrt{2^m} - 2$, for all even $m \geq 4$, and $n_4^*(2m-1,2) \leq 2^m + \frac{23}{16}\sqrt{2^m} - 4$, for all even $m \geq 10$. In both cases, we find that the density of this class of codes satisfies $\mu(\mathcal{D}^{(d)}_{2m-1}, 2) \to 1$, if $m \to \infty$. Hence $\mu_3(2) = \mu_4(2) = 1$. This means that we constructed a sequence of asymptotically perfect codes with covering radius two!

Again, it is not known whether linear codes with these parameters exist. For $m = 4$, Construction 4.24 yields the bounds $n_4^*(7,2) \leq 19$. The best-known linear codes with redundancy seven have parameters $[19,12,3]2$ or $[21,14,4]2$, depending upon the minimum distance, cf. Equation (4.1). It was conjectured in [30], that $[19,12,4]2$ codes do not exist. Construction 4.24 proves that this conjecture is false, if the restriction to linear codes is dropped. For $m = 6$, Construction 4.24 yields the bound $n_4^*(11,2) \leq 72$. The best-known linear codes with redundancy eleven have parameters $[79,68,3]2$ and $[89,78,4]2$, depending upon the minimum distance. This example shows that Construction 4.24 can be extremely powerful, even when the redundancy is small. Notice the strong similarity between Construction 4.24 and Construction 4.10.

**Construction 4.25** [27] Let $m \geq 4$ be even. The partition $\mathcal{H}_m/\mathcal{P}_m$ has $2^{m-1}$-norm 3. Using Lemma 4.15, we infer that the partition $\overline{\mathcal{H}}_m/\overline{\mathcal{P}}_m$ has $2^{m-1}$-norm 4. If we apply the BDS-construction with as inputs these two partitions, then the resulting code $\mathcal{D}_{3m}$ has parameters $(2^{m+1}-1, 2^{(2^{m+1}-1)-3m}, 5)3$. Therefore $n_5^*(3m,3) \leq 2^{m+1}-1$. The density of this class of codes satisfies $\mu(\mathcal{D}_{3m}, 3) \to 1\frac{1}{3}$, if $m \to \infty$. Hence $\mu_5(3) \leq 1\frac{1}{3}$.

Before we can give the next construction, we need to compute another norm.

**Norm 4.26** Let $m \geq 4$ be even. The Hamming code $\mathcal{H}_m$ of length $n = 2^m - 1$ can be partitioned into $2^{m-1}$ translates of the Preparata code $\mathcal{P}_m$. It follows, that the code $\mathcal{H}_m \times \mathbb{F}_2$ can be partitioned into $2^m$ translates of the extended Preparata code $\overline{\mathcal{P}}_m$. Therefore, code $\mathcal{F}_{2m} := \mathcal{H}_m \times (\mathcal{H}_m \times \mathbb{F}_2)$ can be partitioned into $2^m$ translates of the code $\mathcal{D}_{3m}$ obtained via Construction 4.25. Since $\mathcal{F}_{2m}$ has covering radius two, we find that the partition $\mathcal{F}_{2m}/\mathcal{D}_{3m}$ has $2^m$-norm 5.

**Construction 4.27** [27] Let $m \geq 4$ be even. The partition $\mathcal{F}_{2m}/\mathcal{D}_{3m}$ has $2^m$-norm 5, cf. Norm 4.26. The partition $\mathbb{F}_2^n/\mathcal{H}_m$, where $n = 2^m - 1$, has $2^m$-norm 1. Using Lemma 4.15, we infer that the partition $\overline{\mathbb{F}_2^n}/\overline{\mathcal{H}}_m$ has $2^m$-norm 2. If we apply the BDS-construction with as inputs the partitions $\mathcal{F}_{2m}/\mathcal{D}_{3m}$ and $\overline{\mathbb{F}_2^n}/\overline{\mathcal{H}}_m$, then the resulting code $\mathcal{D}_{3m+1}$ has parameters $(3 \cdot 2^m - 1, 2^{(3 \cdot 2^m - 1)-(3m+1)}, 3)3$. Therefore $n_3^*(3m+1, 3) \leq 3 \cdot 2^m - 1$. The density of this class of codes satisfies $\mu(\mathcal{D}_{3m+1}, 3) \to 2\frac{1}{4}$, if $m \to \infty$. Hence $\mu_3(3) \leq 2\frac{1}{4}$.

The codes one obtains via Construction 4.27 have covering radius three and redundancy $3m + 1$, where $m$ is even. Codes with the same covering radius and redundancy can also be constructed, if $m$ is odd. The parameters turn out to be even better.

**Construction 4.28** Let $m \geq 4$ be even. The partition $\mathcal{H}_m/\mathcal{P}_m$ has $2^{m-1}$-norm 3. The partition $\mathcal{H}_{m-1}/BCH_{m-1}(5)$ has $2^{m-1}$-norm 4. If we apply the BDS-construction with as inputs these two partitions, then the resulting code $\mathcal{D}_{3m-2}$ has parameters $(2^m + 2^{m-1} - 2, 2^{(2^m+2^{m-1}-2)-(3m-2)}, 5)3$. Therefore $n_5^*(3m-2, 3) \leq \frac{3}{2}2^m - 2$. The density of this class of codes satisfies $\mu(\mathcal{D}_{3m-2}, 3) \to 2\frac{1}{4}$, if $m \to \infty$. Hence $\mu_5(3) \leq 2\frac{1}{4}$. Although the limiting density is not very impressive, we still find some good codes of small length in this way: $n_5^*(10, 3) \leq 22$, $n_5^*(16, 3) \leq 94$. The bound $n^*(10, 3) \leq 22$ can also be realized by a punctured Golay code with parameters $[22, 12, 6]3$.

**Remark 4.29** In [23] Davydov and Drozhzhina-Labinskaya constructed linear codes with parameters $[3 \cdot 2^m - 1, (3 \cdot 2^m - 1) - (3m + 1), 3]3$ for all $m \geq 7$ and for $m = 5$. These linear codes and the nonlinear codes obtained via Construction 4.27 have the same parameters (if $m$ is even). The codes one obtains via Construction 4.28 have, for the same redundancy, a slightly shorter length than these linear codes and a higher minimum distance: five instead of three. In general it is not known whether linear codes with the parameters obtained via Construction 4.28 exist.

**Construction 4.30** Let $m \geq 4$ be even. The partition $\mathcal{F}_{2m}/\mathcal{D}_{3m}$ has $2^m$-norm 5, cf. Norm 4.26. Using Lemma 4.15, we infer that the partition $\overline{\mathcal{F}}_{2m}/\overline{\mathcal{D}}_{3m}$ has $2^m$-norm 6. If we apply the BDS-construction with as inputs these two partitions, then the resulting code $\mathcal{D}_{5m+1}$ has parameters $(2^{m+2} - 1, 2^{(2^{m+2}-1)-(5m+1)}, 3)5$. Therefore, $n_3^*(5m+1, 5) \leq 2^{m+2} - 1$. The density of this class of codes satisfies $\mu(\mathcal{D}_{5m+1}, 5) \to \frac{2^9}{5!} = 4\frac{4}{15}$. Hence $\mu_3(5) \leq 4\frac{4}{15}$. Though this limiting density is not very impressive, we still find some good codes of small length in this way: $n_3^*(21, 5) \leq 63$, $n_3^*(31, 5) \leq 255$.

The blockwise direct sum construction yields the best results, when it is applied to code partitions. The next two constructions show that we do not have to restrict to partitions, though.

**Construction 4.31** Let $m \geq 4$ be even. The partition $\mathcal{H}_m/\mathcal{P}_m$ has $2^{m-1}$-norm 3. Let $C$ be an $(n, |C|)1$ code, with $n \leq 2^{m-1} - 1$. The covering $\mathbb{F}_2^n/C$ has $(n+1)$-norm 1. Since $n + 1 \leq 2^{m-1}$, $\mathbb{F}_2^n/C$ also has $2^{m-1}$-norm 1. Using Lemma 4.15, we infer that $\overline{\mathbb{F}_2^n}/\overline{C}$ has $2^{m-1}$-norm 2. If we apply the BDS-construction with as inputs the partitions $\mathcal{H}_m/\mathcal{P}_m$ and $\overline{\mathbb{F}_2^n}/\overline{C}$, then the resulting code $\mathcal{D}$ has parameters $(2^m + n, |C| \cdot 2^{2^m-m-1}, 2)2$. Now we consider some specific choices for code $C$. If we take $C = \mathcal{H}_{m-1}$, then this construction coincides with Construction 4.22. If we take $n < 2^{m-1} - 1$, then we get new results. In general we can choose $C$ to be any optimal covering code, i.e. $|C| = K(n, 1)$, and obtain a code with parameters $(2^m + n, K(n, 1) \cdot 2^{2^m-m-1}, 2)2$, where $n + 1 \leq 2^{m-1}$. We give two examples for $m = 4$, due to Etzion et al. [27]. If we choose $C$ to be the (unique) $(5, 7, 1)1$ code, then we obtain a code with parameters $(21, 7 \cdot 2^{11}, 2)2$. If we choose $C$ to be a $(6, 12, 2)1$ code, then we obtain a code with parameters $(22, 12 \cdot 2^{11}, 2)2$.

**Construction 4.32** Let $m \geq 4$ be even. The partition $\mathcal{H}_m/\mathcal{P}_m$ has $2^{m-1}$-norm 3. Let $C$ be an $(n, |C|, d)2$ code, with $n \leq \sqrt{2^m} - 1$. The covering $\mathbb{F}_2^n/C$ has $k$-norm 2, for some $k \leq V(n, 2)$. Since $n + 1 \leq \sqrt{2^m}$, we have $V(n, 2) \leq 2^{m-1}$, so $\mathbb{F}_2^n/C$ also has $2^{m-1}$-norm 2. If we apply the BDS-construction with as inputs the partitions $\mathcal{H}_m/\mathcal{P}_m$ and $\mathbb{F}_2^n/C$, then the resulting code $\mathcal{D}$ has parameters $(2^m+n-1, |C| \cdot 2^{2^m-m-1}, \min\{3, d\})2$. We now consider some specific choices for code $C$. In general we can choose $C$ to be any optimal covering code, i.e. $|C| = K(n, 2)$, and obtain a code with parameters $(2^m + n - 1, K(n, 2) \cdot 2^{2^m-m-1}, 2)2$, where $n+1 \leq \sqrt{2^m}$. We give two examples for $m = 6$. If we choose $C$ to be a $(7, 7, 1)2$ code, then we obtain a code with parameters $(70, 7 \cdot 2^{57}, 1)2$. If we choose $C$ to be a $(8, 12, 2)2$ code, then we obtain a code with parameters $(71, 12 \cdot 2^{57}, 2)2$. Notice, that $n + 1 > \sqrt{2^m}$ in the last example. Nevertheless, we can still apply the construction, since there exists an $(8, 12, 2)2$ code for which 32 translates cover the vector space $\mathbb{F}_2^8$. This can be seen as follows: if one applies the ADS-construction to a $(6, 12, 2)1$ code and the $[3, 1, 3]$ repetition code, then the resulting code has parameters $(8, 12, 2)2 = (6, 12, 2)1 \oplus [3, 1, 3]1$. Since $\mathbb{F}_2^6$ is the union of seven translates of the $(6, 12, 2)1$ code, in fact $\mathbb{F}_2^8$ is the union of (at most) $28 = 4 \cdot (6 + 1)$ translates of this $(8, 12, 2)2$ code.

**Norm 4.33** Let $m \geq 4$ be even. The vector space $\mathbb{F}_2^n$, with $n{=}2^{m-1}{-}1$, can be partitioned into $2^m$ translates of the Hamming code $\mathcal{H}_{m-1}$. It follows, that the code $\mathbb{F}_2^{n+1}$ can be partitioned into $2^m$ translates of the extended Hamming code $\overline{\mathcal{H}}_{m-1}$. Therefore, code $\mathcal{F}_m := \mathcal{H}_m \times \mathbb{F}_2^{n+1}$ can be partitioned into $2^m$ translates of the code $\mathcal{D}_{2m}$ obtained via Construction 4.22. Since $\mathcal{F}_m$ has covering radius one, we find that the partition $\mathcal{F}_m/\mathcal{D}_{2m}$ has $2^m$-norm 3.

**Construction 4.34** Let $m \geq 4$ be even. The partition $\mathcal{F}_m/\mathcal{D}_{2m}$ has $2^m$-norm 3, cf. Norm 4.33. Let $C$ be an $(n, |C|)1$ code, with $n \leq 2^m - 1$. the covering $\mathbb{F}_2^n/C$ has $(n+1)$-norm 1. Since $n + 1 \leq 2^m$, $\mathbb{F}_2^n/C$ also has $2^m$-norm 1. Using Lemma 4.15, we infer that $\overline{\mathbb{F}_2^n}/\overline{C}$ has $2^m$-norm 2. If we apply the BDS-construction with as inputs the partitions $\mathcal{F}_m/\mathcal{D}_{2m}$ and $\overline{\mathbb{F}_2^n}/\overline{C}$, then the resulting code $\mathcal{D}$ has parameters $(\frac{3}{2}2^m+n, |C|\cdot 2^{\frac{3}{2}2^m-m-1}, 1)2$. Now we consider some specific choices for code $C$. If we take $C = \mathcal{H}_m$, then this construction coincides with Construction 4.22. If we take $n < 2^m-1$, then we get new results. In general we can choose $C$ to be any optimal covering code, i.e. $|C| = K(n,1)$, and obtain a code with parameters $(\frac{3}{2}2^m + n, K(n,1)\cdot 2^{\frac{3}{2}2^m-m-1}, 1)2$, where $n+1 \le 2^m$.

**Example 4.35** Using the tables of $K(n,1)$ of [27] we find codes with the following parameters:

$$\begin{array}{ll}
K(31,2) \le K(7,1) \times 2^{19} \le 16 \times 2^{19}, & \\
K(32,2) \le K(8,1) \times 2^{19} \le 32 \times 2^{19}, & K(36,2) \le K(12,1) \times 2^{19} \le 382 \times 2^{19}, \\
K(33,2) \le K(9,1) \times 2^{19} \le 62 \times 2^{19}, & K(37,2) \le K(13,1) \times 2^{19} \le 750 \times 2^{19}, \\
K(34,2) \le K(10,1) \times 2^{19} \le 120 \times 2^{19}, & K(38,2) \le K(14,1) \times 2^{19} \le 1460 \times 2^{19}, \\
K(35,2) \le K(11,1) \times 2^{19} \le 192 \times 2^{19}, & K(39,2) \le K(15,1) \times 2^{19} \le 2048 \times 2^{19}.
\end{array}$$

We conclude this section by giving a BDS-construction, where not all subcodes are translates of each other.

**Construction 4.36** Let $C$ be the $(6,12,2)1$ code of Example 4.81. By inspection we see that $C$ can be partitioned into the following three subcodes, each with covering radius two:

$$\begin{aligned}
C_1 &:= \{(000,100),(111,011),(100,111),(011,000)\}, \\
C_2 &:= \{(000,010),(111,101),(010,111),(101,000)\}, \\
C_3 &:= \{(000,001),(111,110),(001,111),(110,000)\}.
\end{aligned}$$

Consequently, the partition of $C$ into the three subcodes $C_1$, $C_2$, and $C_3$ has 3-norm 3. Using Lemma 4.15, we infer that the partition of $\overline{C}$ into the subcodes $\overline{C}_1$, $\overline{C}_2$, and $\overline{C}_3$ has 3-norm 4. If we apply the BDS-construction with as inputs these two partitions, then the resulting code $\mathcal{D}$ has parameters $(13,48,3)3$.

## 4.4 A Construction Using the Golay Code

In the previous section we used the BDS-construction to obtain a large number of covering codes. These codes are all nonlinear, since their construction involves Preparata codes. Here we shall give some examples of *linear* codes that can be obtained via the BDS-construction. The constructions involve the binary Golay code and some codes derived from this code. Using the relations between these codes, we obtain two new norms. This information will enable us to find some linear covering codes with better parameters than previously known.

Trivially, the Golay code is a subcode of some linear code with covering radius two. We will show that the Golay code is contained in a $[23,16,2]2$ code. Since $[23,15]2$ codes do not exist, this is optimal in some sense. In addition, we will show how the Golay code can be used to construct a $[30, 19, 4]3$ code that is contained in a $[30, 22, 3]2$ code.

First we introduce some notations.
Let $\alpha$ be a primitive element of $GF(8)$. The points of the projective line $PG(1, 8)$ are 1-dimensional subspaces of $(\mathbb{F}_8)^2$ and can be identified, via the vector space isomorphism $GF(8) \cong (\mathbb{F}_2)^3$, with 3-dimensional subspaces of $(\mathbb{F}_2)^6$. Thus the projective points $(0, 1)$, $(1, 0)$, $(1, 1)$, resp. $(1, \alpha)$ can be identified with the 3-dimensional vector spaces $\mathcal{P}_1, \mathcal{P}_2, \mathcal{P}_3$, resp. $\mathcal{P}_4$ in $(\mathbb{F}_2)^6$. These vector spaces are pairwise independent. With each point $\mathbf{p}$ of the projective line $PG(1, 8)$ we associate a matrix $V(\mathbf{p}) := (\alpha^0\mathbf{p}, \alpha^1\mathbf{p}, \cdots, \alpha^6\mathbf{p}, \mathbf{0})$, where each column is considered in its binary representation. The result of puncturing this matrix on the zero-position is denoted by $V^*(\mathbf{p})$. Define $P_1 := V(0, 1)$, $P_2 := V(1, 0)$, $P_3 := V^*(1, 1)$, and $P_4 := V^*(1, \alpha)$.

Now we define the Golay code.
Let $\mathcal{H}$ be the cyclic $[7, 4, 3]$ Hamming code with zero $\alpha$. Let $\mathcal{H}^r$ be obtained by reversing the positions of all codewords of $\mathcal{H}$. The extended Golay code $\mathcal{G}_{24}$ is defined by

$$\mathcal{G}_{24} := \{(\mathbf{a} + \mathbf{x}, \mathbf{b} + \mathbf{x}, \mathbf{a} + \mathbf{b} + \mathbf{x}) \mid \mathbf{a}, \mathbf{b} \in \overline{\mathcal{H}} \text{ and } \mathbf{x} \in \overline{\mathcal{H}^r}\}.$$

This self-dual code has parameters $[24, 12, 8]4$. The binary Golay code $\mathcal{G}_{23}$ with parameters $[23, 12, 7]3$ can be obtained by puncturing $\mathcal{G}_{24}$ on its last coordinate.

The extended Golay code is a self-dual code. Consequently, the vectors $(\mathbf{1}, \mathbf{0}, \mathbf{0})$, $(\mathbf{0}, \mathbf{1}, \mathbf{0})$, $(\mathbf{0}, \mathbf{h}, \mathbf{h})$, and $(\mathbf{h}, \mathbf{0}, \mathbf{h})$, where $\mathbf{h}$ is a codeword of the extended Hamming code $\overline{\mathcal{H}}$, are parity checks of the extended Golay code. Hence all codewords $(\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3)$ of the Golay code satisfy the parity check equations:

$$\begin{gathered}\sum_{i=1}^{3} \mathbf{c}_i P_i^T = \mathbf{0}, \\ wt(\mathbf{c}_1) \text{ and } wt(\mathbf{c}_2) \text{ are even.}\end{gathered} \tag{4.2}$$

Consider the $[23, 15]$ code defined by the equations of (4.2). This code has covering radius three and contains the Golay code as a subcode. We will show, that if one deletes a suitable parity check equation from the equations of (4.2), then one obtains a code with parameters $[23, 16, 2]2$. Moreover, we will show how the Golay code can be used to define a $[30, 19, 4]3$ that is contained in a $[30, 22, 3]2$ code.

Let $\mathcal{D}_7 := \{(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) \mid wt(\mathbf{x}_1 + \mathbf{x}_2) \text{ is even}, \sum_{i=1}^{3} \mathbf{x}_i P_i^T = 0\}$. Using the fact that the vector spaces $\mathcal{P}_1$, $\mathcal{P}_2$, and $\mathcal{P}_3$ are pairwise independent, one easily verifies that $\mathcal{D}_7$ has parameters $[23, 16, 2]2$. Each codeword of the Golay code satisfies the equations of (4.2), hence $\mathcal{D}_7$ contains the Golay code as a subcode.

Let $\mathcal{D}_8 := \{(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4) \mid wt(\mathbf{x}_1)$ and $wt(\mathbf{x}_2)$ are even, $\sum_{i=1}^{4} \mathbf{x}_i P_i^T = 0\}$. Using the fact that the vector spaces $\mathcal{P}_1, \ldots, \mathcal{P}_4$ are pairwise independent, one easily verifies that $\mathcal{D}_8$ has parameters $[30, 22, 3]2$. Observe that, for every codeword $\mathbf{c}$ of the Golay code, the word $(\mathbf{c}, 0)$ is a codeword of $\mathcal{D}_8$. We will use this observation to design a $[30, 19, 4]3$ that is contained in code $\mathcal{D}_8$.

Let $X_1$, $X_2$, and $X_3$ be matrices such that $(X_1 \mid X_2)$ is a parity check matrix of code $\mathcal{D}_8$ and such that $\left(X_1^T \mid X_3^T\right)^T$ is a parity check matrix of the Golay code. Let $\mathcal{D}_{11}$ be the code with parity check matrix

$$D_{11}^{\perp} = \left(\begin{array}{c|c} X_1 & X_2 \\ \hline X_3 & J \end{array}\right),$$

where $J$ denotes the all-one matrix. By construction, code $\mathcal{D}_{11}$ is contained in code $\mathcal{D}_8$ and has parameters $[30, 19, 4]3$. The minimum distance of $\mathcal{D}_{11}$ follows from the fact that the vector spaces $\mathcal{P}_1, \ldots, \mathcal{P}_4$ are pairwise independent; the covering radius follows from the observation that adjoining extra columns to a parity check matrix of the Golay code does not increase the covering radius.

From the properties of codes $\mathcal{D}_7$ and $\mathcal{D}_8$ we obtain the following two norms.

**Norm 4.37** There is a partition $[23, 16, 2]2/[23, 12, 7]3$ with 16-norm 5.

**Norm 4.38** There is a partition $[30, 22, 3]2/[30, 19, 4]3$ with 8-norm 5.

We give two examples of how Norm 4.37 and Norm 4.38 can be used in the BDS-construction to obtain linear covering codes. The codes will have better parameters than previously known. As before, we merely indicate which codes are the inputs to the BDS-construction. In each instance, the minimum distance and the covering radius of the constructed codes follow by a straightforward application of Theorems 4.8 and 4.12. The linearity of the constructed codes follows from Property 3 of Theorem 4.7.

In [9] Brualdi and Pless obtained a $[48, 29, 4]3$ code by applying an ADS-construction on the binary Golay code and a $[26, 18, 3]2$ code, thus proving the bound $l_4(19, 5) \leq 48$. The next construction improves this upper bound by one.

**Construction 4.39** The partition $[23, 16, 2]2/[23, 12, 7]3$ has 16-norm 5, cf. Norm 4.37. Using Lemma 4.15, we infer that the partition $\overline{[23, 16, 2]}/\overline{[23, 12, 7]}$ has 16-norm 6. If we apply the BDS-construction with as inputs these two partitions, then we obtain a code with parameters $[47, 28, 4]5$. Therefore $l_4(19, 5) \leq 47$.

**Remark 4.40** By analyzing the BDS-construction in detail, we see that the nonzero codewords of the $[47, 28, 4]5$ code obtained via Construction 4.39 have either weight four or weight at least seven and that the first situation only occurs once.

The bound $l(12,3) \leq 38$ was mentioned in [9] as a private communication of Dan Ashlock, but his computer proof could not be reproduced [1]. We will construct a $[38,26,4]3$ code, thus showing that this bound remains valid. In fact, we give four constructions that successively improve the bound $l(12,3) \leq 41$ to the bound $l(12,3) \leq 38$. These constructions illustrate to what extent the various generalizations of the direct sum construction can be used to improve bounds.

The direct sum of a $[26,18,3]2$ code and a $[15,11,3]1$ Hamming code is a code with parameters $[41,29,3]3$. If one applies an ADS-construction on these codes, then one obtains a code with parameters $[40,28,3]3 = [26,18,3]2 \dot{\oplus} [15,11,3]1$. Constructions 4.41 and 4.42 yield further improvements. The first construction yields a code with parameters $[39,27,4]3$; the latter one a code with parameters $[38,26,4]3$. We will use Norm 4.17.

**Construction 4.41** The partition $\mathbb{F}_2^{15}/[15,11,3]$ has 16-norm 1. Using Lemma 4.15, we find that partition $[16,15,2]/[16,11,4]$ has 16-norm 2. The partition $[23,16,2]2/[23,12,7]3$ has 16-norm 5, cf. Norm 4.37. If we apply the BDS-construction with as inputs these two partitions, then we obtain a code with parameters $[39,27,4]3$. Therefore $l_4(12,3) \leq 39$.

**Construction 4.42** The partition $\mathbb{F}_2^{7}/[7,4,3]$ has 8-norm 1. Using Lemma 4.15, we infer that the partition $[8,7,2]/[8,4,4]$ has 8-norm 2. The partition $[30,22,3]2/[30,19,4]3$ has 8-norm 5, cf. Norm 4.38. If we apply the BDS-construction with as inputs these two partitions, then we obtain a code with parameters $[38,26,4]3$. Therefore $l_4(12,3) \leq 38$. A parity check matrix for this code is shown in Figure 4.2.

$$\left(\begin{array}{c|c|c|c|c}
00000000 & 10010110 & 1001011 & 1001011 & 00000000 \\
00000000 & 01011100 & 0101110 & 0101110 & 00000000 \\
00000000 & 00101110 & 0010111 & 0010111 & 00000000 \\
\hline
10010110 & 00000000 & 1001011 & 0010111 & 00000000 \\
01011100 & 00000000 & 0101110 & 1011100 & 00000000 \\
00101110 & 00000000 & 0010111 & 0101110 & 00000000 \\
\hline
11111111 & 00000000 & 0000000 & 0000000 & 00000000 \\
00000000 & 11111111 & 0000000 & 0000000 & 00000000 \\
\hline
11101000 & 11101000 & 1110100 & 1111111 & 10010110 \\
00111010 & 00111010 & 0011101 & 1111111 & 01011100 \\
01110100 & 01110100 & 0111010 & 1111111 & 00101110 \\
\hline
00000000 & 00000000 & 0000000 & 0000000 & 11111111
\end{array}\right)$$

Figure 4.2: Parity check matrix of a [38,26,4]3 code.

## 4.5 Amalgamated Direct Sums

In the previous section we used the BDS-construction to design infinite families of covering codes with a low density. In all these constructions we used Preparata codes and the property that the partition $\mathcal{H}_m/\mathcal{P}_m$ has $2^{m-1}$-norm 3 (cf. Norm 4.19). Based upon the covering radii of the Hamming code and the Preparata code alone, one would have estimated this norm to be four. This explains the usefulness of this specific $k$-norm. In general, the problem of finding the $k$-norm of a code $C$ with respect to the subcodes $C^{(1)}, \ldots, C^{(k)}$ is very hard. The case $k$=2 has received considerable attention in the literature. In fact, most papers on covering radius problems deal with this case. For small parameters, this specific instance of the BDS-construction yields some good results; for larger lengths the results are poor. The constructions are all referred to as *amalgamated direct sum* (ADS) constructions. In all constructions, 2-norms play a key-role. To analyze this so-called ADS-construction, a confusing number of notions has been introduced: normality, subnormality, seminormality, strong seminormality, and $(k,t)$-subnormality. Below we will give a uniform description of the main results on the ADS-construction. We restrict ourselves to binary codes.

We denote the even weight subcode of a code $C$ by $C_e$, the odd weight subcode by $C_o$. Sometimes we will use the linear mapping $\sigma$ on $\mathbb{F}_2^n$ defined by $\sigma(x_1, \ldots, x_n) := (\sum_{i=1}^{n} x_i, x_2, \ldots, x_n)$. Notice that $\sigma^2 = \sigma \circ \sigma$ is the identity mapping.

### 4.5.1 Some 2-Norms

In this section we will find a large class of binary codes with a small 2-norm. First we give an example.

**Example 4.43** Let $C_0$ be the $[6,3,3]$ code with parity check matrix $H$ defined by

$$H = \left( \begin{array}{ccc|ccc} 1 & & & 0 & 1 & 1 \\ & 1 & & 1 & 0 & 1 \\ & & 1 & 1 & 1 & 0 \end{array} \right).$$

This code has covering radius two and all vectors $\mathbf{x}$ with $d(\mathbf{x}, C_0) = 2$ have syndrome $\mathbf{s} = (111)$, i.e. are in code $C_1 := \mathbf{1} + C_0$. Similarly, all vectors $\mathbf{x}$ with $d(\mathbf{x}, C_1) = 2$ are in code $C_0$. It follows, that $C := C_0 \cup C_1$ has 2-norm 2 with respect to subcodes $C_0$ and $C_1$. Notice that the 2-norm could not have been smaller than 2, since $C$ has covering radius one. Codes $C$, $C_0$, and $C_1$ can be obtained by puncturing, resp. shortening the $[7,4,3]$ Hamming code on a coordinate.

The $[7,4,3]$ Hamming code is an example of a code from which one can obtain partitions with a relatively small 2-norm. Such codes are called normal codes.

**Definition 4.44** Let $C$ be a binary code of length $n$ with covering radius $r$. Suppose $C$ has 2-norm $N$ with respect to subcodes $C_0$ and $C_1$. Trivially, we have $N \geq 2r$. If $N \leq 2r+1$, then $C$ is called a subnormal code (with respect to subcodes $C_0$ and $C_1$). Coordinate $i$ is called a suitable coordinate, if $C[i]$ has 2-norm $N' < N$ with respect to subcodes $C_0[i]$ and $C_1[i]$. If $C$ is subnormal with respect to subcodes $C_0$ and $C_1$ and has a suitable coordinate, then $C$ is called a normal code. A code that is not normal is called abnormal.

**Remark 4.45** We will only consider partitions of a code $C$ into two subcodes $C_0$ and $C_1$, i.e. we will always assume that $C_0$ and $C_1$ are disjoint sets. Any set $S \subset \mathbb{F}_2^n$ defines a partition of a code $C$ into the two subcodes $C \cap S$ and $C \setminus S$. Conversely, any partition of $C$ into two subsets can be defined via such a separating set. For conciseness, we will sometimes specify a partition of a code into two subcodes via a separating set.

Normal codes were introduced by Graham and Sloane in [32] for linear codes. In [32] and all subsequent papers [12, 17, 28, 40, 42, 67, 85] normal codes are defined in a more restrictive way than we do: in all these papers a code is called normal, if it is subnormal with respect to one of the hyperplanes $\langle e_i \rangle^{\perp}$, where $1 \leq i \leq n$. We will show that our more general definition of normality does not impose any restrictions on later constructions (cf. Construction 4.46). Our definition also shows that the abnormal codes constructed in [28, 55, 85] are not abnormal at all! (if one takes our definition of normality).

The main reason for introducing normal codes is the so-called amalgamated direct sum construction.

**Construction 4.46** Let $C_1 \subset \mathbb{F}_2^{n_1}$ and $C_2 \subset \mathbb{F}_2^{n_2}$ be codes with covering radius $r_1$, resp. $r_2$. Suppose that $C_1$ is a normal code with respect to subcodes $C_1^{(0)}$ and $C_1^{(1)}$ and suppose $i$ is a suitable coordinate, where $1 \leq i \leq n_1$. Moreover, suppose that $C_2$ is a subnormal code with respect to subcodes $C_2^{(0)}$ and $C_2^{(1)}$. By definition, $C_1[i]$ has 2-norm at most $2r_1$ with respect to subcodes $C_1^{(0)}[i]$ and $C_1^{(1)}[i]$. If we apply the BDS-construction with as inputs the codes $C_1[i]$ and $C_2$ with their respective subcodes, then the resulting code $\mathcal{D}$ has parameters $(n, M)r$, where $n = n_1 + n_2 - 1$, $M \leq |C_1^{(0)}| \cdot |C_2^{(0)}| + |C_1^{(1)}| \cdot |C_2^{(1)}|$, and $r \leq r_1 + r_2$. Notice, that if $r(\mathcal{D}) = r_1 + r_2$, then $\mathcal{D}$ is subnormal with respect to subcodes $\mathcal{D}_0 := C_1^{(0)}[i] \times C_2^{(0)}$ and $\mathcal{D}_1 := C_1^{(1)}[i] \times C_2^{(1)}$.

Notice that the BDS-construction with $C_0$ and $C_1$ and their respective subcodes as inputs yields a code $\mathcal{D}'$ with the same cardinality as code $\mathcal{D}$, but now with length $n_1 + n_2$ and covering radius (at most) $r_1 + r_2 + 1$. Code $\mathcal{D}$ can be obtained by puncturing code $\mathcal{D}'$ on coordinate $i$. The normality of code $C_0$ guarantees that puncturing yields a code with covering radius $r_1 + r_2$, which is usually one less than the covering radius of code $\mathcal{D}'$. This is the main motivation for distinguishing between normal and subnormal codes.

**Remark 4.47** Construction 4.46 was introduced by Graham and Sloane [32] as the *amalgamated direct sum* construction. In that paper and subsequent papers code $\mathcal{D}$ is called the amalgamated direct sum (ADS) of codes $\mathcal{C}_1$ and $\mathcal{C}_2$ and is denoted by $\mathcal{C}_1 \dot{\oplus}\, \mathcal{C}_2$. This notation is ambiguous, though, since there might be more than one suitable coordinate. Therefore we denote by $\mathcal{C}_1 \dot{\oplus}\, \mathcal{C}_2$ the collection of all possible amalgamated direct sums of $\mathcal{C}_1$ and $\mathcal{C}_2$.

### 4.5.2 Normal and Subnormal Codes

**Lemma 4.48** Let $\mathcal{C}$ be a binary code of length $n$. Suppose $\mathcal{C}$ has 2-norm $N$ with respect to subcodes $\mathcal{C}_0$ and $\mathcal{C}_1$. Then code $\mathcal{D} := (\mathcal{C}_0 \times \{0\}) \cup (\mathcal{C}_1 \times \{1\})$ has 2-norm $N + 1$ with respect to the hyperplane $\langle \mathbf{e}_{n+1} \rangle^\perp$.

**Proof:** Code $\mathcal{C}$ has 2-norm $N$ with respect to subcodes $\mathcal{C}_0$ and $\mathcal{C}_1$. The binary field $\mathbb{F}_2$ has 2-norm 1 with respect to the singleton-sets $\{0\}$ and $\{1\}$. The lemma now follows from Theorem 4.12. □

By symmetry, the converse of Lemma 4.48 also holds.

**Lemma 4.49** Let $\mathcal{C}$ be a binary code of length $n$ with covering radius $r$. Suppose $\mathcal{C}$ has 2-norm $N$ with respect to the subcodes $\mathcal{C}_0 := \mathcal{C} \cap \langle \mathbf{e}_i \rangle^\perp$ and $\mathcal{C}_1 := \mathcal{C} \setminus \mathcal{C}_0$, where $1 \le i \le n$. Then $\mathcal{C}[i]$ has 2-norm $N - 1$ with respect to the subcodes $\mathcal{C}_0[i]$ and $\mathcal{C}_1[i]$.

As a direct consequence, we obtain the following theorem.

**Theorem 4.50** Let $\mathcal{C}$ be a binary code of length $n$ that is subnormal with respect to one of the hyperplanes $\langle \mathbf{e}_1 \rangle^\perp, \ldots, \langle \mathbf{e}_n \rangle^\perp$. Then $\mathcal{C}$ is a normal code.

In the literature, normal codes are those codes that satisfy the condition of Theorem 4.50. Our class of normal codes is strictly larger. To prove this, we need some additional lemmas.

**Lemma 4.51** Let $\mathcal{C}$ be a binary code with covering radius $r$. Suppose $\mathcal{C}$ has 2-norm $N$ with respect to subcodes $\mathcal{C}_0$ and $\mathcal{C}_1$. Then $\overline{\mathcal{C}}$ has covering radius $r + 1$ and an *even* 2-norm $N + 1$ or $N + 2$ with respect to subcodes $\overline{\mathcal{C}}_0$ and $\overline{\mathcal{C}}_1$.

**Proof:** This follows immediately from Lemma 4.15. □

**Lemma 4.52** Let $\mathcal{C}$ be a binary code and suppose $\mathcal{C}$ has 2-norm $N$ with respect to subcodes $\mathcal{C}_0$ and $\mathcal{C}_1$. Suppose $\mathcal{C}^\sigma$ has 2-norm $N'$ with respect to subcodes $\mathcal{C}_0^\sigma$ and $\mathcal{C}_1^\sigma$. Then $N' \le N + 1$ (and $N' \le N$, if $N$ is odd).

**Proof:** Codes $\mathcal{C}$ and $\mathcal{C}^\sigma$ have equivalent extended codes. The result now follows from Lemma 4.51. □

**Lemma 4.53** Let $\mathcal{C}$ be a binary code of length $n$ with covering radius $r$. Suppose $\mathcal{C}$ has 2-norm $N$ with respect to the even weight subcode $\mathcal{C}_e := \mathcal{C} \cap \langle \mathbf{1} \rangle^\perp$. Then $\mathcal{C}[1]$ has 2-norm $N' < N$ with respect to the subcode $\mathcal{C}_e[1]$.

**Proof:** It is trivial that the 2-norm of $C$ with respect to its even/odd weight subcodes is odd. From Lemma 4.52 and the fact that $N$ is odd we infer, that code $C^\sigma$ has 2-norm $N^\sigma \leq N$ with respect to the subcode $(C_e)^\sigma = C^\sigma \cap \langle \mathrm{e}_1 \rangle^\perp$, i.e. $C^\sigma$ has 2-norm $N^\sigma \leq N$ with respect to the hyperplane $\langle \mathrm{e}_1 \rangle^\perp$. From Lemma 4.49 we infer that $C^\sigma[1]$ has 2-norm $N' = N^\sigma - 1 < N$ with respect to the subcode $(C_e)^\sigma[1]$. Codes $C$ and $C^\sigma$ differ only in the first coordinate, hence $C[1] = C^\sigma[1]$ and $C_e[1] = (C_e)^\sigma[1]$. Consequently, $C[1]$ has 2-norm $N' < N$ with respect to the subcode $C_e[1]$. □

From Lemma 4.53 and Theorem 4.50 we directly obtain the following result.

**Theorem 4.54** Let $C$ be a binary code of length $n$ that is subnormal with respect to one of the hyperplanes $\langle \mathrm{e}_1 \rangle^\perp, \ldots, \langle \mathrm{e}_n \rangle^\perp, \langle \mathbf{1} \rangle^\perp$. Then $C$ is a normal code.

**Remark 4.55** Several papers consider codes which are abnormal with respect to all hyperplanes $\langle \mathrm{e}_i \rangle^\perp$, where $1 \leq i \leq n$. Frankl [55] constructed codes with this property with covering radius one; Van Wee [85] generalized this construction to hold for arbitrary covering radii. All these codes have minimum distance one; constructions with higher minimum distances have been given by Etzion et al. [28]. Although these codes are all abnormal in the 'classical' sense, i.e. abnormal with respect to the hyperplanes $\langle \mathrm{e}_i \rangle^\perp$, with $1 \leq i \leq n$, most of these codes can easily be shown to be subnormal with respect to the hyperplane $\langle \mathbf{1} \rangle^\perp$, i.e. they are normal according to our less restrictive definition! Therefore, most of the 'abnormal' codes explicitly constructed by the above authors are not abnormal in our sens, since they can be mapped via mapping $\sigma$ to a code that is normal with respect to the hyperplane $\langle \mathrm{e}_1 \rangle^\perp$, i.e. is normal in our sense. It should be mentioned, that the constructions in papers [28, 55, 85] can be easily adapted to obtain codes that are abnormal with respect to all the hyperplanes $\langle \mathrm{e}_1 \rangle^\perp, \ldots, \langle \mathrm{e}_n \rangle^\perp, \langle \mathbf{1} \rangle^\perp$. Since we are interested only in normal codes, not in abnormal ones, we do not describe these adaptations here.

The next lemmas prove to be useful later on.

**Lemma 4.56** Code $C$ is subnormal with respect to subcodes $C_0$ and $C_1$, iff $C^\sigma$ is subnormal with respect to subcodes $C_0^\sigma$ and $C_1^\sigma$.

**Proof:** Codes $C$ and $C^\sigma$ have the same covering radius, since they have equivalent extended codes (cf. Lemma 4.51). The implication now follows from Lemma 4.52. The equivalence follows from the fact that the rôles of code $C$ and code $C^\sigma$ can be interchanged, since $(C^\sigma)^\sigma = C$. □

**Lemma 4.57** Code $C$ is normal with respect to the hyperplane $\langle \mathbf{1} \rangle^\perp$, iff $C^\sigma$ is normal with respect to the hyperplane $\langle \mathrm{e}_1 \rangle^\perp$.

**Proof:** This follows from Lemma 4.56 and the observation that $(C_e)^\sigma = C^\sigma \cap \langle \mathrm{e}_1 \rangle^\perp$. □

### 4.5.3 Some Normal Codes

In this section we will find a large class of normal binary codes by considering codes with a small minimum distance and quasi-perfect codes. All results are based upon determining the 2-norm of a code with respect to the hyperplane $\langle \mathbf{1} \rangle^{\perp}$. Abnormality of a code with respect to this hyperplane imposes certain restrictions on the structure of the code. Combinatorial arguments will sometimes show that these restrictions cannot be met. We will use that the 2-norm with respect to the hyperplane $\langle \mathbf{1} \rangle^{\perp}$ is odd.

The next lemma shows that many codes with a small minimum distance also have a small 2-norm.

**Lemma 4.58** Let $C$ be a binary $(n, M)r$ code that is invariant under the translation $\mathbf{x} \rightarrow \mathbf{x} + \mathbf{d}$, for some nonzero vector $\mathbf{d} \in I\!F_2^n$ of *odd* weight $d$. Then $C$ has 2-norm at most $2r + \lceil d/2 \rceil$ with respect to the hyperplane $\langle \mathbf{1} \rangle^{\perp}$. This norm is odd.

**Proof:** Let $\mathbf{x} \in I\!F_2^n$. Let $\mathbf{d}_1, \mathbf{d}_2 \in I\!F_2^n$ be words with disjoint supports such that $\mathbf{d} = \mathbf{d}_1 + \mathbf{d}_2$. Since $d(\mathbf{x} + \mathbf{d}_1, C) \leq r$, we have $\mathbf{x} + \mathbf{d}_1 + \mathbf{e} \in C$, for some vector $\mathbf{e}$ of weight $wt(\mathbf{e}) \leq r$. Since $C = C + \mathbf{d}$, also $\mathbf{x} + \mathbf{d}_2 + \mathbf{e} \in C$. Vector $\mathbf{d}$ has odd weight, so the weights of the vectors $\mathbf{x} + \mathbf{d}_1 + \mathbf{e}$ and $\mathbf{x} + \mathbf{d}_2 + \mathbf{e}$ have different parity. It follows, that $d(\mathbf{x}, C_e), d(\mathbf{x}, C_o) \leq \max\{wt(\mathbf{d}_1 + \mathbf{e}), wt(\mathbf{d}_2 + \mathbf{e})\}$. If we take $wt(\mathbf{d}_1) = \lceil d/2 \rceil$, then we obtain $d(\mathbf{x}, C_e), d(\mathbf{x}, C_o) \leq r + \lceil d/2 \rceil$. Code $C$ has covering radius $r$ and codes $C_e$ and $C_o$ both have covering radius at most $r + \lceil d/2 \rceil$, hence $d(\mathbf{x}, C_e) + d(\mathbf{x}, C_o) \leq 2r + \lceil d/2 \rceil$. □

**Theorem 4.59** Let $C$ be a binary code that is invariant under the translation $\mathbf{x} \rightarrow \mathbf{x} + \mathbf{d}$, for some nonzero vector $\mathbf{d}$ of weight $d \leq 4$. Let $i \in \text{supp}(\mathbf{d})$. Then $C$ is normal with respect to the hyperplane $\langle \mathbf{e}_i \rangle^{\perp}$.

**Proof:** We may permute coordinate positions, so we may assume that the first coordinate is in the support of vector $\mathbf{d}$. Code $C$ is normal with respect to the hyperplane $\langle \mathbf{e}_1 \rangle^{\perp}$, iff $C^{\sigma}$ is normal with respect to the hyperplane $\langle \mathbf{1} \rangle^{\perp}$, cf. Lemma 4.57. Code $C^{\sigma}$ is invariant under the translation $\mathbf{x} \rightarrow \mathbf{x} + \sigma(\mathbf{d})$. Since the first coordinate is in the support of vector $\mathbf{d}$, the weight of $\sigma(\mathbf{d})$ is odd and at most three. The result now follows from Lemma 4.58. □

**Corollary 4.60** All binary linear codes with minimum distance $d \leq 4$ are normal with all coordinates in the support of a codeword of weight at most four acceptable.

**Remark 4.61** This corollary was attributed to C.L.M. van Pul in [45]. In fact, the result already follows from [17, Theorem 24], which states Corollary 4.60, but now for $d \leq 3$. This follows from the following observation: if all linear codes with odd minimum distance $d$ are normal with respect to some suitable coordinate in the support of a codeword of weight $d$, then the same result holds for codes with minimum distance $d + 1$. (Once again, use the technique of mapping code $C$ to code $C^{\sigma}$.) Kilby and Sloane [54, 55] stated that every linear code with distance five is also normal, but this result remains to be proved, as was pointed out in [45].

The next lemma imposes restrictions on the structure of a quasi-perfect code, whenever it is abnormal with respect to the hyperplane $\langle\mathbf{1}\rangle^{\perp}$. This lemma will enable us to prove that almost all quasi-perfect codes are normal.

**Lemma 4.62** Let $C$ be an $(n, M, d)r$ quasi-perfect code , i.e. $d \geq 2r - 1$. Then $C$ is abnormal with respect to the hyperplane $\langle\mathbf{1}\rangle^{\perp}$ if and only if for some vector $\mathbf{z} \in \mathbb{F}_2^n$, the set

$$T_r(\mathbf{z}) := \{\mathbf{y} \in \mathbb{F}_2^n \mid \mathbf{z} - \mathbf{y} \in C \text{ and } wt(\mathbf{y}) \leq r + 1\}$$

consists of $n/r$ words of weight $r$ and with disjoint supports.

**Proof:** Assume that $C$ is abnormal with respect to the hyperplane $\langle\mathbf{1}\rangle^{\perp}$, i.e. $d(\mathbf{x}, C_e) + d(\mathbf{x}, C_o) > 2r+1$ for some $\mathbf{x} \in \mathbb{F}_2^n$. Let $t := d(\mathbf{x}, C)$ and let $\mathbf{c}$ be a codeword with $d(\mathbf{x}, \mathbf{c}) = t$. Moreover, let $\mathbf{z} \in \mathbb{F}_2^n$ be any vector with $d(\mathbf{z}, \mathbf{x}) = r - t$ and $d(\mathbf{z}, \mathbf{c}) = r$. Let $\mathbf{c}'$ be any codeword with $d(\mathbf{z}, \mathbf{c}') \leq r + 1$. By the triangle inequality, we have $d(\mathbf{x}, \mathbf{c}) + d(\mathbf{x}, \mathbf{c}') \leq d(\mathbf{z}, \mathbf{c}) + d(\mathbf{z}, \mathbf{c}') \leq 2r + 1$. If $d(\mathbf{c}, \mathbf{c}')$ is odd, then $d(\mathbf{x}, C_e) + d(\mathbf{x}, C_o) \leq 2r + 1$, in conflict with our assumption. So $d(\mathbf{c}, \mathbf{c}')$ is even. Hence $\mathbf{c} = \mathbf{c}'$ or $d(\mathbf{c}, \mathbf{c}') = 2r$, since $C$ is a quasi-perfect code. Consequently, $d(\mathbf{z}, \mathbf{c}') = r$ and all vectors of $T_r(\mathbf{z})$ have weight $r$ and disjoint supports. Moreover, $\text{supp}T_r(\mathbf{z}) = \{1, \ldots, n\}$, since $d(\mathbf{z}+\mathbf{e}_i, C) = r+1$ for all $i \notin \text{supp}T_r(\mathbf{z})$. It follows, that $T_r(\mathbf{z})$ has cardinality $n/r$.
Conversely, assume that $\mathbf{z} \in \mathbb{F}_2^n$ is such that $T_r(\mathbf{z})$ only contains vectors of weight $r$. Then $d(\mathbf{z}, C_e) + d(\mathbf{z}, C_o) \geq 2r + 3$, so $C$ is abnormal with respect to the hyperplane $\langle\mathbf{1}\rangle^{\perp}$. □

**Remark 4.63** Codes with covering radius one are quasi-perfect codes. For these codes, Lemma 4.62 reduces to [40, Theorem 10].

**Theorem 4.64** All quasi-perfect $(n, M, d)r$ codes with $r \quad n$ are normal with respect to the hyperplane $\langle\mathbf{1}\rangle^{\perp}$.

The next example shows that the constraint $r \quad n$ in Theorem 4.64 cannot be dropped.

**Example 4.65** Let $C$ be the $[10, 5, 4]2$ code with parity check matrix $H$ defined by

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

Let $\mathbf{x}$ be any vector with syndrome $\mathbf{s} = (10000)$. An inspection of matrix $H$ reveals that all distances $d(\mathbf{x}, \mathbf{c})$ between $\mathbf{x}$ and some codeword $\mathbf{c} \in C$ are two, five, six, or ten. Consequently, $d(\mathbf{x}, C_e) + d(\mathbf{x}, C_o) = 7$. It is easily verified that all words $\mathbf{y}$ with another syndrome satisfy $d(\mathbf{y}, \mathcal{C}_e) + d(\mathbf{y}, \mathcal{C}_o) = 5$. Code $\mathcal{C}^\sigma$ is equivalent to the $[10,5,3]2$ code of [32, p. 386]. Notice, that $\mathcal{C}$ is equivalent to its dual code $\mathcal{C}^\perp$, since parity check matrix $H$, with the first row inverted, is a generator matrix for code $\mathcal{C}$. This is the smallest nontrivial example of a linear code which is abnormal with respect to the hyperplane $\langle \mathbf{1} \rangle^\perp$. (Trivial examples are the *even* codes.)

Many quasi-perfect codes have parameters $(n, M, d)r$ , where $r \mid n$. Although we cannot apply Theorem 4.64 to prove normality of these codes, we can often establish the same result using Lemma 4.62 directly. We will show that all $(n, K(n,1))1$ codes are normal, as are all codes with $d \geq 2r$ and all linear quasi-perfect codes. Notice that, in order for a code $\mathcal{C}$ to be normal, it suffices to prove that either $\mathcal{C}$ or $\mathcal{C}^\sigma$ is normal with respect to the hyperplane $\langle \mathbf{1} \rangle^\perp$, cf. Lemma 4.57.

**Theorem 4.66** [40] Let $\mathcal{C}$ be an $(n, K(n,1))1$ code. Then $\mathcal{C}$ is normal with respect to the hyperplane $\langle \mathbf{1} \rangle^\perp$.

**Proof:** Suppose otherwise. Then $B_2(\mathbf{z}) \cap \mathcal{C} = S_1(\mathbf{z})$ for some $\mathbf{z} \in \mathbb{F}_2^n$, cf. Lemma 4.62. Let $d(\mathbf{y}, \mathbf{z}) = 2$ and let $B_1(\mathbf{y}) \cap B_1(\mathbf{z}) = \{\mathbf{a}, \mathbf{b}\}$. Then $\mathcal{C}' = \mathcal{C} \cup \{\mathbf{y}\} \setminus \{\mathbf{a}, \mathbf{b}\}$ has covering radius one. But now $|\mathcal{C}'| < |\mathcal{C}| = K(n,1)$, a contradiction. □

**Theorem 4.67** Let $\mathcal{C}$ be an $(n, M, d)r$ code with $d \geq 2r$. Then $\mathcal{C}^\sigma$ is normal with respect to the hyperplane $\langle \mathbf{1} \rangle^\perp$.

**Proof:** Suppose otherwise. Mapping $\sigma$ maps quasi-perfect codes to quasi-perfect codes, hence we can apply Lemma 4.62 to $\mathcal{C}^\sigma$. From Lemma 4.62 we infer that $r$ divides $n$ and that, for some $\mathbf{z} \in \mathbb{F}_2^n$, the set $\mathcal{S} := \{\mathbf{z}+\mathbf{c} \mid \mathbf{c} \in \mathcal{C}^\sigma \text{ and } d(\mathbf{z}, \mathbf{c}) \leq r+1\}$ consists of $n/r$ words with weight $r$ and disjoint supports. The words of $\mathcal{S}$ partition all coordinate positions, hence the distances between the words of $\mathcal{S}^\sigma$ are $2r-1$ and $2r$ and both distances do occur. Since $\mathcal{S}^\sigma \subset \sigma(\mathbf{z}) + \mathcal{C}$, code $\mathcal{C}$ has minimum distance $2r-1$. We obtain a contradiction, since $d \geq 2r$. □

**Theorem 4.68** Let $\mathcal{C}$ be a linear quasi-perfect code with parameters $[n, k, d]r$. Then either $\mathcal{C}$ or $\mathcal{C}^\sigma$ is normal with respect to the hyperplane $\langle \mathbf{1} \rangle^\perp$.

**Proof:** Suppose $\mathcal{C}$ and $\mathcal{C}^\sigma$ are both abnormal with respect to the hyperplane $\langle \mathbf{1} \rangle^\perp$. Mapping $\sigma$ maps quasi-perfect codes to quasi-perfect codes, hence we can apply Lemma 4.62 to both codes $\mathcal{C}$ and $\mathcal{C}^\sigma$. From Lemma 4.62 we infer that $r$ divides $n$ and that there is a vector $\mathbf{z}$ such that the set $\mathcal{S} := \{\mathbf{z} + \mathbf{c} \mid \mathbf{c} \in \mathcal{C} \text{ and } d(\mathbf{z}, \mathbf{c}) \leq r+1\}$ consists of $n/r$ words of weight $r$ and with disjoint supports. Similarly, there is a vector $\mathbf{w}$ such that the set $\mathcal{T} := \{\mathbf{w} + \mathbf{c} \mid \mathbf{c} \in \mathcal{C}^\sigma \text{ and } d(\mathbf{w}, \mathbf{c}) \leq r+1\}$ consists of $n/r$ words of weight $r$ and with disjoint supports. The words of $\mathcal{T}$ partition the coordinate positions, hence the distances between the words of $\mathcal{T}^\sigma$ are $2r-1$ and $2r$ and both distances do occur. Notice that all words of $\mathcal{T}^\sigma$ have weight $r$ or $r+1$, if $r$ is odd; if $r$ is even, then all words of $\mathcal{T}^\sigma + \mathbf{e}_1$ have weight $r$ or $r+1$. The vectors in set $\mathcal{S}$ partition the coordinate positions. Therefore the sum of the vectors in set $\mathcal{S}$, $\sum \mathcal{S}$, is the all-one vector. Similarly, $\sum \mathcal{T}$ is the all-one vector. Notice that $\sigma(\mathbf{w}) + \mathcal{T}^\sigma \subset \mathcal{C}$, since $\mathbf{w} + \mathcal{T} \subset \mathcal{C}^\sigma$.

Using the linearity of codes $\mathcal{C}$ and $\mathcal{C}^\sigma$, we will derive a contradiction. We distinguish two cases, depending on the parity of the quotient $n/r$.

$n/r$ **is even:** Code $\mathcal{C}$ is a linear code and $n/r$ is even, hence the vector $\mathbf{1} = \sum \mathcal{S}$ is a codeword of $\mathcal{C}$. Similarly, $\mathbf{1} = \sum \mathcal{T}$ is a codeword of $\mathcal{C}^\sigma$. Consequently, $\mathcal{C} \cap \mathcal{C}^\sigma$ contains the codewords $\mathbf{1}$ and $\sigma(\mathbf{1})$. Notice that $n$ is even, so $\sigma(\mathbf{1}) = \mathbf{1} + \mathbf{e}_1$. Since $\mathcal{C}$ is a linear code, we infer that it has minimum distance one. But this implies that $\mathcal{C}$ is normal with respect to the hyperplane $\langle \mathbf{1} \rangle^\perp$ (cf. Lemma 4.58), in conflict with our assumptions.

$n/r$ **is odd:** Code $\mathcal{C}$ is a linear code and $n/r$ is odd, hence the vector $\mathbf{z}$ and the all-one vector are in the same coset of code $\mathcal{C}$, i.e. $\mathbf{z} + \mathcal{C} = \mathbf{1} + \mathcal{C}$. Similarly, $\mathbf{w} + \mathcal{C}^\sigma = \mathbf{1} + \mathcal{C}^\sigma$ and hence $\sigma(\mathbf{w}) + \mathcal{C} = \sigma(\mathbf{1}) + \mathcal{C}$. We distinguish two cases, depending on the parity of $r$ (or $n$). If $r$ is odd, then $\sigma(\mathbf{1}) = \mathbf{1}$, hence vectors $\mathbf{z}$ and $\sigma(\mathbf{w})$ are in the same coset of $\mathcal{C}$. Since the code is linear, this implies that $\mathcal{T}^\sigma = \mathcal{S}$, a contradiction. If $r$ is even, then $n$ is even, so $\sigma(\mathbf{1}) = \mathbf{1} + \mathbf{e}_1$. Consequently, vectors $\mathbf{z}$ and $\sigma(\mathbf{w}) + \mathbf{e}_1$ are in the same coset of $\mathcal{C}$. Since the code is linear, this implies that $\mathcal{T}^\sigma + \mathbf{e}_1 = \mathcal{S}$, a contradiction. □

**Remark 4.69** All normality results mentioned in the papers [28, 48, 85] can easily be proved from our theorems, with the help of Corollary 4.52.

### 4.5.4 Subnormal Codes

In the previous section we found a large class of normal codes. Examples of linear abnormal codes are not known. Though it is commonly believed that all linear codes are normal, this conjecture has not been settled yet. The next result, due to Calderbank [12], is related to this conjecture. We include it for completeness.

**Lemma 4.70** [12] Let $\mathcal{C}$ be an $[n,k]r$ code. If $2^k + 2^{n-k} > 2^{\{(n+1) - \lceil \frac{n+1}{r+1} \rceil\}} + 1$, then $\mathcal{C}$ is subnormal with respect to some hyperplane.

**Proof:** Assume that $\mathcal{C}$ is absubnormal with respect to all hyperplanes of $\mathcal{C}$. We may assume that all hyperplanes pass through the origin, since any hyperplane $H$ of $\mathcal{C}$ and its complement $\mathcal{C} \setminus H$ define the same partitioning of $\mathcal{C}$.

For all $\mathbf{x} \in \mathbb{F}_2^n$ define $P(\mathbf{x}) := \{(\mathbf{c}, \mathbf{c}') \in \mathcal{C} \times \mathcal{C} \mid d(\mathbf{x}, \mathbf{c}) + d(\mathbf{x}, \mathbf{c}') \leq 2r+1\}$. Furthermore, let $V(\mathbf{x}) := \langle \{\mathbf{c}' - \mathbf{c} \mid (\mathbf{c}, \mathbf{c}') \in P(\mathbf{x})\} \rangle$.

We say that hyperplane $H$ fails to separate $P(\mathbf{x})$, if $V(\mathbf{x}) \subset H$. Notice that any hyperplane that fails to separate $P(\mathbf{x})$, also fails to separate $P(\mathbf{y})$, if $\mathbf{x}$ and $\mathbf{y}$ are in the same coset of $C$. This follows, since $C$ is a linear code. By assumption, all hyperplanes of $C$ fail to separate $P(\mathbf{x})$, for some $\mathbf{x} \in \mathbb{F}_2^n$.

We will prove the lemma by estimating in two ways the cardinality of the set

$$T := \{(\mathbf{x}, H) \mid \mathbf{x} \in \mathbb{F}_2^n, \text{ hyperplane } H \text{ fails to separate } P(\mathbf{x})\}.$$

Let $\lambda$ be the minimum dimension of $V(\mathbf{x})$ over all $\mathbf{x} \in \mathbb{F}_2^n$. Using the fact that any $t$-dimensional vector space in $\mathbb{F}_2^n$ contains $2^t - 1$ hyperplanes through the origin, we can easily compute the cardinality of set $T$ and find

$$|C|\ (2^k - 1) \leq |T| = \sum\{2^{k-\dim V(\mathbf{x})} - 1 \mid \mathbf{x} \in \mathbb{F}_2^n\} \leq 2^n \cdot (2^{k-\lambda} - 1). \qquad (4.3)$$

We can rewrite this expression and obtain the following inequality:

$$2^k + 2^{n-k} \leq 2^{n-\lambda} + 1. \qquad (4.4)$$

Now we estimate the dimension of vector space $V(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{F}_2^n$.
Let $t := d(\mathbf{x}, C)$ and let $\mathbf{c}$ be a codeword with $d(\mathbf{x}, \mathbf{c}) = t$. A vector $\mathbf{z} \in \mathbb{F}_2^n$ such that $d(\mathbf{u}, C) \leq d(\mathbf{z}, C)$, for all $\mathbf{u}$ with $d(\mathbf{u}, \mathbf{z}) \leq 1$, is called an orphan. There exists an orphan $\mathbf{z}$ with $d(\mathbf{z}, C) = d(\mathbf{z}, \mathbf{c}) =: u$ and $d(\mathbf{z}, \mathbf{x}) = u - t$, for some $t \leq u \leq r$. Define $T_u(\mathbf{z}) := \{\mathbf{y} \in \mathbb{F}_2^n \mid \mathbf{z} - \mathbf{y} \in C \text{ and } wt(\mathbf{y}) \leq u + 1\}$. For all vectors $\mathbf{y} \in T_u(\mathbf{z})$ we have $d(\mathbf{x}, \mathbf{c}) + d(\mathbf{x}, \mathbf{z} + \mathbf{y}) \leq 2u + 1$, so $(\mathbf{c}, \mathbf{z} + \mathbf{y}) \in P(\mathbf{x})$. It follows, that $\dim V(\mathbf{x}) \geq \dim\langle T_u(\mathbf{z})\rangle - 1$. We will estimate the dimension of $V(\mathbf{x})$ via an estimate for the dimension of $\langle T_u(\mathbf{z})\rangle$.
Notice that set $T_u(\mathbf{z})$ contains only vectors of weights $u$ and $u + 1$. For all $i \notin \operatorname{supp}T_u(\mathbf{z})$ we have $d(\mathbf{z} + \mathbf{e}_i, C) = u + 1$. Since $\mathbf{z}$ is an orphan, we infer that $\operatorname{supp}T_u(\mathbf{z}) = \{1, \ldots, n\}$. It follows, that $|T_u(\mathbf{z})| \geq \lceil\frac{n+1}{u+1}\rceil$. Moreover, there is a minimal subset $T_0$ of $T_u(\mathbf{z})$ with $\operatorname{supp}T_0 = \{1, \ldots, n\}$. Since this set has cardinality $|T_0| \geq \lceil\frac{n+1}{u+1}\rceil$, the vector space $\langle T_u(\mathbf{z})\rangle$ has dimension $\dim\langle T_u(\mathbf{z})\rangle \geq \lceil\frac{n+1}{u+1}\rceil \geq \lceil\frac{n+1}{r+1}\rceil$. Consequently,

$$\dim V(\mathbf{x}) \geq \lambda \geq \left\lceil\frac{n - r}{r + 1}\right\rceil. \qquad (4.5)$$

Substituting estimate (4.5) for $\lambda$ in Equation (4.4), we get the inequality

$$2^k + 2^{n-k} \leq 2^{\{(n+1) - \lceil\frac{n+1}{r+1}\rceil\}} + 1. \qquad (4.6)$$

This completes the proof. □

**Corollary 4.71** Let $C$ be an $[n, k]r$ code. If $k/(n + 1) > r/(r + 1)$ or $(n - k)/(n + 1) > r/(r + 1)$, then $C$ is normal with respect to some hyperplane.

## 4.6 Codes from Geometries and Extension Fields

In this section we construct covering codes from projective geometries and from codes over another alphabet. First, however, we introduce some notions from projective geometry.

### 4.6.1 Fields, Vector Spaces, and Geometries

The field $\mathbb{F}_{q^m}$ is a vector space of dimension $m$ over $\mathbb{F}_q$. With respect to a fixed basis, each element of $\mathbb{F}_{q^m}$ can be represented as a $q$-ary vector of length $m$ by means of the vector space isomorphism $\mathbb{F}_{q^m} \cong (\mathbb{F}_q)^m$. Via this isomorphism, one can view any vector space of dimension $k$ over $\mathbb{F}_{q^m}$ as a vector space of dimension $km$ over $\mathbb{F}_q$. For conciseness, we sometimes exploit these correspondences and consider vectors with coordinates from $\mathbb{F}_{q^m}$, where formally their $q$-ary representations are appropriate, and vice versa. It should always be clear from the context, however, whether a vector with entries from $\mathbb{F}_{q^m}$ should be viewed as a $q$-ary vector or not.

Consider the vector space $(\mathbb{F}_q)^m$. A $k$-dimensional affine subspace or $k$-flat is a coset of a $k$-dimensional linear subspace of $(\mathbb{F}_q)^m$. If $k = m-1$, we call the $k$-flat a hyperplane. The affine geometry of dimension $m$ over the field $\mathbb{F}_q$ is the partially ordered set of all affine subspaces of the vector space $(\mathbb{F}_q)^m$. This geometry is commonly denoted by $AG(m,q)$. The projective geometry of dimension $m$ over the field $\mathbb{F}_q$, denoted by $PG(m,q)$, is the lattice of all linear subspaces of $(\mathbb{F}_q)^{m+1}$. The subspaces of dimension 1 are called (projective) points, those of dimension 2 are called lines, etc. The projective point $\langle(x_1,\dots,x_{m+1})\rangle$ in $(\mathbb{F}_q)^{m+1}$ can be represented by any of the vectors $\lambda(x_1,\dots,x_{m+1})$, where $0 \neq \lambda \in \mathbb{F}_q$. Therefore the set of projective points in $PG(m,q)$ can be identified with a maximal set of pairwise linearly independent vectors (over $\mathbb{F}_q$) in $(\mathbb{F}_q)^{m+1}$, the so-called homogeneous representation of the points of $PG(m,q)$. Projective geometries of dimension one are called projective lines, those of dimension two are called projective planes, etc.

We are mainly interested in certain subsets of affine and projective geometries.
Let $k \geq 1$. A collection $\mathcal{S} := \{V_1,\dots,V_n\}$ of distinct $m$-dimensional vector spaces in $(\mathbb{F}_q)^{km}$ is called a $k$-independent set, if $(\mathbb{F}_q)^{km}$ is spanned by any $k$ distinct vector spaces in $\mathcal{S}$. The collection $\mathcal{S}$ is called an $r$-spanning set in $(\mathbb{F}_q)^{km}$, if each vector in $(\mathbb{F}_q)^{km}$ is contained in the linear span of some collection of $r$ distinct vector spaces in $\mathcal{S}$. Both notions can also be defined projectively. A $k$-independent set in $(\mathbb{F}_q)^k$ is commonly called an arc in $PG(k-1,q)$ (or $n$-arc if it has size $n$). Notice that an $n$-arc in $PG(k-1,q^m)$ can be viewed as a $k$-independent set of size $n$ in $(\mathbb{F}_q)^{km}$, if one considers 1-dimensional vector spaces over $\mathbb{F}_{q^m}$ as $m$-dimensional vector spaces over $\mathbb{F}_q$. The following are examples of arcs in $PG(k-1,q)$ with $k \geq 2$, cf. [64, p. 323]:

$\mathcal{S}_1 := \{\mathrm{e}_1,\dots,\mathrm{e}_k,\mathrm{e}_1+\dots+\mathrm{e}_k\}$ is a $(k+1)$-arc in $PG(k-1,q)$;

$\mathcal{S}_2 := \{(1,x,\dots,x^{k-1}) \mid x \in \mathbb{F}_q\} \cup \{(0,\dots,0,1)\}$ is a $(q+1)$-arc in $PG(k-1,q)$;

$\mathcal{S}_3 := \{(1,x,x^2) \mid x \in \mathbb{F}_q\} \cup \{(0,0,1),(0,1,0)\}$ is a $(q+2)$-arc in $PG(3,q)$, provided $q$ is even.

An $n$-arc in $PG(k-1,q)$ is called a complete arc, if it is not contained in any $(n+1)$-arc in $PG(k-1,q)$. Complete arcs in $PG(0,q)$ and $PG(1,q)$ are trivial: they contain all points. The maximum cardinality of any complete arc in $PG(k-1,q)$ is denoted by $m_k(q)$.

Complete arcs of cardinality $m_3(q)$ in a projective plane are called ovals. For $k \geq 3$, the exact value of $m_k(q)$ is known only in a few cases, though it is conjectured [64, p. 328] that

$$m_k(q) = \begin{cases} q+2 & \text{if } q = 2^m \text{, and } k = 3 \text{ or } k = q-1, \\ \max\{q+1, k+1\} & \text{otherwise.} \end{cases} \tag{4.7}$$

We will only consider projective geometries $PG(k-1, q)$ with $k \leq 4$. For these values of $k$, the conjectured values of $m_k(q)$ are exact [13] and can be realized by one of the arcs $\mathcal{S}_1$, $\mathcal{S}_2$, or $\mathcal{S}_3$.

A line $\ell$ in $PG(k-1, q)$ is called a passant, tangent, resp. secant on some set $\mathcal{S}$ of points in $PG(k-1, q)$, depending on whether $|\ell \cap \mathcal{S}|$ is zero, one, resp. (at least) two. Sometimes certain relations between lines and arcs in $PG(k-1, q)$ hold. We will only consider relations in projective planes. Any $(2^m+1)$-arc $\mathcal{S}$ in $PG(2, 2^m)$ is contained in a unique oval with $2^m + 2$ points. The additional point is called the nucleus of $\mathcal{S}$. An oval in $PG(2, 2^m)$ does not contain any tangents in $PG(2, 2^m)$. An oval in $PG(2, q)$, $q$ odd, contains $q+1$ points; moreover, no three distinct tangents on this oval meet in one point of $PG(2, q)$.

Now we are ready to describe a construction of covering codes using projective geometries.

### 4.6.2 Codes from Projective Geometries

Linear codes with covering radius $r$ and $r$-spanning sets in projective geometries are equivalent objects: the columns of a parity check matrix of a $q$-ary $[n, n-m]r$ code form an $r$-spanning set of points in $PG(m-1, q)$, and vice versa. As an example of a 2-spanning set in $PG(3, 2^m)$ we consider a geometrical construction by Brualdi et al. [8].

**Theorem 4.72** Let $q = 2^m$. Let $V$ be a plane in $PG(3, q)$ and let $\mathcal{O}$ be an oval in $V$. Consider a line $\ell$ in $PG(3, q)$ through some point $N$ of $\mathcal{O}$, $\ell$ not in $V$. Then $\mathcal{S} := (\mathcal{O} \cup \ell) \setminus N$ is a 2-spanning set in $PG(3, q)$ with $2q+1$ points.

**Proof:** Let $x$ be any point of $PG(3, q)$ not in $\mathcal{S}$. We distinguish two cases, depending on whether $x$ is in $V$ or not. If $x$ is in $V$, then $x$ is on $(q/2)+1$ secants of $\mathcal{O}$, since $V$ does not contain any tangents on $\mathcal{O}$. It follows, that $x$ is on a secant of $\mathcal{O} \setminus N$ (if $x \neq N$) or on the line $\ell$ (if $x = N$). If $x$ is not in $V$, then the plane $W$ determined by $x$ and $\ell$ intersects $V$ in a line $\ell'$ which contains $N$ and hence another point $y$ of $\mathcal{O}$, because $V$ does not contain any tangents on $\mathcal{O}$. The line $\ell''$ through $x$ and $y$ is contained in $W$, hence it intersects the line $\ell$ in a point $z \neq N$. It follows that $x$ is on the secant through $y$ and $z$. Hence all points of $PG(3, q)$ are on a line determined by two points of $\mathcal{S}$. □

**Remark 4.73** Theorem 4.72 gives a 2-spanning set with $2q+1$ points in $PG(3, q)$, where $q$ is an even prime power. A slight modification of this theorem yields this result for odd prime powers $q$ as well (unless $q = 3$). We leave out the details. Notice that the trivial construction (taking two nonintersecting lines) gives a 2-spanning set with $2q+2$ points.

### 4.6.3 Codes from Codes over Extension Fields

In this section we construct covering codes from codes over another alphabet.

Let $\mathcal{H}_m(q)$ be the $q$-ary Hamming code of length $a = (q^m - 1)/(q-1)$. The vector space $\mathbb{F}_q^a$ can be partitioned into $q^m$ translates of $\mathcal{H}_m(q)$. We denote this partition by $\mathbb{F}_q^a/\mathcal{H}_m(q)$. Let $f$ be any bijection from $\mathbb{F}_q^a/\mathcal{H}_m(q)$ to the field $GF(q^m)$, e.g. the mapping from $\mathbb{F}_q^a$ to the $q^m$ possible syndromes of $\mathcal{H}_m(q)$. Notice that function $f$ has the property that $f(B_1(\mathbf{x})) = \mathbb{F}_{q^m}$ for all $\mathbf{x} \in \mathbb{F}_q^a$, since the Hamming code and its translates all have covering radius one. Bijection $f$ can be used to map codes over the extension field $\mathbb{F}_{q^m}$ to $q$-ary codes with the same covering radius.

**Lemma 4.74** Let $\mathcal{C}$ be a code of length $n$ over $\mathbb{F}_{q^m}$ with covering radius $r$. Let $a = (q^m - 1)/(q-1)$ and let $f$ be any bijection from $\mathbb{F}_q^a/\mathcal{H}_m(q)$ to the field $\mathbb{F}_{q^m}$. Then code $\mathcal{D}$ defined by

$$\mathcal{D} := \{(\mathbf{x}_1, \ldots, \mathbf{x}_n) \in (\mathbb{F}_q^a)^n \mid (f(\mathbf{x}_1), \ldots, f(\mathbf{x}_n)) \in \mathcal{C}\} \tag{4.8}$$

has covering radius $r$.

**Proof:** Let $(\mathbf{x}_1, \ldots, \mathbf{x}_n) \in (\mathbb{F}_q^a)^n$ and let $\mathbf{y} := (f(\mathbf{x}_1), \ldots, f(\mathbf{x}_n))$. Function $f$ has the property that $f(B_1(\mathbf{x})) = \mathbb{F}_{q^m}$ for all $\mathbf{x} \in \mathbb{F}_q^a$. It follows, that $d((\mathbf{x}_1, \ldots, \mathbf{x}_n), \mathcal{D}) = d(\mathbf{y}, \mathcal{C}) \leq r$. □

**Remark 4.75** It is clear that if $\mathcal{C}$ is linear and if $f$ is a linear function, then code $\mathcal{D}$ defined by (4.8) is linear. In that case this lemma reduces to [8, Lemma 3.1].

In Lemma 4.74 we used Hamming codes and a bijective function $f$ to define a mapping from codes over the extension field $\mathbb{F}_{q^m}$ to $q$-ary codes with the same covering radius. The proof of this lemma only uses a certain property of function $f$, viz. the property that $f(B_1(\mathbf{x})) = \mathbb{F}_{q^m}$ for all $\mathbf{x} \in \mathbb{F}_q^a$. Therefore one can also apply this construction using instead of Hamming codes other codes with covering radius one: suppose $\mathbb{F}_q^a$ is the union of $k$ subcodes $\mathcal{C}_0, \ldots, \mathcal{C}_{k-1}$, each with covering radius one. For all $\mathbf{x} \in \mathbb{F}_q^a$ let $f(\mathbf{x}) := \{0 \leq i < k \mid \mathbf{x} \in \mathcal{C}_i\}$. We call $f$ the characteristic function of $\mathcal{C}_0, \ldots, \mathcal{C}_{k-1}$. Notice that mapping $f$ has the property that $\cup f(B_1(\mathbf{x})) = \mathbb{Z}_k$ for all $\mathbf{x} \in \mathbb{F}_q^a$. (If $k = q^m$, then we identify elements of $\mathbb{Z}_k$ with those of $\mathbb{F}_{q^m}$.) Characteristic function $f$ can be used to map mixed codes to $q$-ary codes with the same covering radius, thus generalizing Lemma 4.74.

**Lemma 4.76** Let $\mathcal{C} \subseteq \mathbb{Z}_{k_1} \times \cdots \times \mathbb{Z}_{k_n}$ be a mixed code with covering radius $r$. For all $i, 1 \leq i \leq n$, let $\mathbb{F}_q^{a_i}$ be the union of $k_i$ subcodes $\mathcal{C}_i^{(0)}, \ldots, \mathcal{C}_i^{(k_i-1)}$ with covering radius 1 and let $f_i$ be the characteristic function associated with these subcodes. Then code $\mathcal{D}$ defined by

$$\mathcal{D} := \{(\mathbf{x}_1, \ldots, \mathbf{x}_n) \mid (f_1(\mathbf{x}_1) \times \cdots \times f_n(\mathbf{x}_n)) \cap \mathcal{C} \neq \emptyset\} \tag{4.9}$$

has covering radius $r$.

**Proof:** This follows directly from the observation that for all $i, 1 \leq i \leq n$, characteristic function $f_i$ has the property that $\cup f_i(B_1(\mathbf{x}_i)) = \mathbb{Z}_{k_i}$ for all $\mathbf{x} \in \mathbb{F}_q^{a_i}$. □

**Remark 4.77** Obviously, the construction in Lemma 4.76 can be straightforwardly generalized to yield mappings from mixed codes to other mixed codes. Lemma 4.76 was used by Østergård [66] to obtain a $(10, 120)1$ code from a mixed code $\mathcal{C} \subseteq \mathbb{F}_4 \times (\mathbb{F}_2)^7$ with cardinality $|\mathcal{C}| = 60$ and covering radius 1, which he had found by simulated annealing. The same construction was used by Etzion et al. [27, Construction A] to obtain codes with covering radius two from certain perfect two-error-correcting mixed codes in $\mathbb{F}_{2^{m-1}} \times (\mathbb{F}_2)^{2^m}$ (cf. also Remark 4.23).

**Example 4.78** [8] Let $q = 2^m$. Let $S$ be the $4 \times (2q+1)$ matrix with as columns the points of the 2-spanning set of Theorem 4.72. Matrix $S$ is a parity check matrix of a $q$-ary code with parameters $[2q+1, 2q-3, 3]2$. Let $\alpha$ be a primitive element of $\mathbb{F}_{2^m}$. The matrix $H = (1\ \alpha\ \cdots\ \alpha^{q-2})$ is a parity check matrix of the binary Hamming code of length $q-1$. Let $f$ be the linear function defined by $f(\mathbf{x}) = \mathbf{x}H^T$ for all $\mathbf{x} \in \mathbb{F}_2^{q-1}$. From Lemma 4.74 we infer that code $\mathcal{D}$ defined by (4.8) has parameters $[n, n-4m, 3]2$, with $n = (2^{m+1}+1)(2^m-1)$. It can easily be verified that the Kronecker product $S \otimes H$ is a parity check matrix of code $\mathcal{D}$. Code $\mathcal{D}$ is equivalent to the code with parity check matrix $S' := H \otimes S = (S \mid \alpha S \mid \cdots \mid \alpha^{q-2}S)$.

## 4.7 Piecewise Constant Codes

In this section we construct covering codes from codes with a mixed alphabet.

First we give some definitions.
Let $\mathcal{C} \subseteq V_n := \mathbb{Z}_{k_1} \times \cdots \times \mathbb{Z}_{k_n}$. If not all $k_i$ are the same, then $\mathcal{C}$ is called a mixed code. The Manhattan distance $d_M(\mathbf{x}, \mathbf{y})$ between two words $\mathbf{x}, \mathbf{y} \in V_n$ is defined by $d_M(\mathbf{x}, \mathbf{y}) := \sum_i |x_i - y_i|$. The Manhattan radius $r_M$ of code $\mathcal{C}$ is the maximum value of $d_M(\mathbf{x}, \mathcal{C})$ over all words $\mathbf{x} \in V_n$. Here $d_M(\mathbf{x}, \mathcal{C}) := \min\{d_M(\mathbf{x}, \mathbf{c}) \mid \mathbf{c} \in \mathcal{C}\}$. Notice, that if all $k_i$'s are two, then the Manhattan distance and radius correspond to the usual definition of Hamming distance and covering radius.

Now we are ready to describe the construction of a class of codes introduced by Cohen et al. [17], the so-called piecewise constant codes. The examples are from the same paper.

**Lemma 4.79** Let $\mathcal{C} \subseteq V_n := \mathbb{Z}_{k_1+1} \times \cdots \times \mathbb{Z}_{k_n+1}$. Let $f$ be the weight function, i.e. $f(\mathbf{x}) := wt(\mathbf{x})$. Then code $\mathcal{D}$ defined by

$$\mathcal{D} := \{(\mathbf{x}_1, \ldots, \mathbf{x}_n) \in \mathbb{F}_q^{k_1} \times \cdots \times \mathbb{F}_q^{k_n} \mid (f(\mathbf{x}_1), \ldots, f(\mathbf{x}_n)) \in \mathcal{C}\} \tag{4.10}$$

has covering radius $r(\mathcal{D}) = r_M(\mathcal{C})$.

**Proof:** Let $(\mathbf{x}_1, \ldots, \mathbf{x}_n) \in \mathbb{F}_q^{k_1} \times \cdots \times \mathbb{F}_q^{k_n}$ and let $\mathbf{y} := (f(\mathbf{x}_1), \ldots, f(\mathbf{x}_n))$. For all $i, 1 \leq i \leq n$, function $f$ has the property that $f(B_1(\mathbf{x}_i)) = \{wt(\mathbf{x}_i) - 1, wt(\mathbf{x}_i), wt(\mathbf{x}_i) + 1\} \cap \{0, \ldots, k_i\}$ for all $\mathbf{x}_i \in \mathbb{F}_q^{k_i}$. It follows, that $d((\mathbf{x}_1, \ldots, \mathbf{x}_n), \mathcal{D}) = d_M(\mathbf{y}, \mathcal{C}) \leq r_M$. □

**Example 4.80** Let $\mathcal{C} \subseteq \mathbb{Z}_3 \times \mathbb{Z}_4$ be defined by $\mathcal{C} := \{(0,0), (0,3), (1,0), (2,2)\}$. The codewords of this code can be depicted in the two-dimensional array of Figure 4.3 with as $(i,j)$-th entry $\binom{2}{i} \cdot \binom{3}{j}$. The parameters of the corresponding piecewise constant code $\mathcal{D}$ can easily be determined from this array. Kalbfleisch and Stanton [78] proved — by linear programming techniques — the uniqueness of the $(5,7)1$ code shown in Figure 4.3.

| | | j → | | | |
|---|---|---|---|---|---|
| | | 0 | 1 | 2 | 3 |
| i | 0 | ① | 3 | 3 | ① |
| ↓ | 1 | ② | 6 | 6 | 2 |
| | 2 | 1 | 3 | ③ | 1 |

(00,000)
(00,111)
(10,000)
(01,000)
(11,011)
(11,101)
(11,110)

Figure 4.3: Depiction of code $\mathcal{C}$ and the corresponding $(5,7)1$ code $\mathcal{D}$.

**Example 4.81** Let $\mathcal{C} \subseteq \mathbb{Z}_4 \times \mathbb{Z}_4$ be defined by $\mathcal{C} := \{(0,1), (1,3), (2,0), (3,2)\}$. As in the previous example, this code can be depicted in the two-dimensional array shown in Figure 4.4. From this figure we see that the corresponding piecewise constant code $\mathcal{D}$ has parameters $(6,12,2)1$. In fact, one can show by linear programming that this code is optimal [78]. By inspection we see that $\mathcal{D}$ can be partitioned into the following three subcodes:

| | | j → | | | |
|---|---|---|---|---|---|
| | | 0 | 1 | 2 | 3 |
| | 0 | 1 | ③ | 3 | 1 |
| i | 1 | 3 | 9 | 9 | ③ |
| ↓ | 2 | ③ | 9 | 9 | 3 |
| | 3 | 1 | 3 | ③ | 1 |

Figure 4.4: Depiction of mixed code $\mathcal{C}$.

$$\begin{aligned}
\mathcal{D}_1 &:= \{(000,100), (111,011), (100,111), (011,000)\}, \\
\mathcal{D}_2 &:= \{(000,010), (111,101), (010,111), (101,000)\}, \\
\mathcal{D}_3 &:= \{(000,001), (111,110), (001,111), (110,000)\}.
\end{aligned}$$

Each of these subcodes has covering radius two, since each subcode is equivalent to a translate of the linear code $\langle(111,000),(000,111)\rangle$.

## 4.8 Codes from Block Designs

In this section we show how designs can be used to construct covering codes. We give a construction for an $(11,192)1$ code. This construction is due to Cohen et al. [17]. A similar construction was used in [64, p. 72] to construct an $(11,144,3)$ code. Here we show that, in fact, the latter code occurs as a subcode of the covering code. We identify codewords with subsets by means of their supports.

Consider the Steiner system with parameters 5-$(12,6,1)$. This is a self-complementary design, i.e. if a block occurs in the design, then its complement also occurs. Let $\mathcal{B}$ be the collection of 132 blocks of this design. Different blocks intersect in at most four points, since any collection of five points is contained in exactly one block. Thus, blocks have mutual Hamming distance at least four.

Let $\mathcal{K}$ be any collection of six disjoint pairs of points, together with their complements. This code has minimum distance four.
The code $\mathcal{C} := \mathcal{B} \cup \mathcal{K}$ has minimum distance four as well, since all codewords of $\mathcal{B}$ have weight six and all codewords of $\mathcal{K}$ have weight two or ten. Thus we obtain a $(12,144,4)$ code.

Now we determine the covering radius of this code and a related code.
Let $\mathcal{L}$ be the collection $\{\mathbf{x} = (\mathbf{x}_L, \mathbf{x}_R) \in \mathbb{F}_2^6 \times \mathbb{F}_2^6 \mid wt(\mathbf{x}) = 2, wt(\mathbf{x}_L) = 0 \text{ or } wt(\mathbf{x}_R) = 0\}$, together with their complements. Let $\mathbf{x} = (\mathbf{x}_L, \mathbf{x}_R)$ be any word of $\mathbb{F}_2^{12}$, partitioned into two parts of length six. Since $\mathcal{B}$ is a 5-$(12,6,1)$ design, any word of weight four, five, or six has distance at most two to code $\mathcal{B}$. Any word of weight zero, one, or two has distance at most two to code $\mathcal{K}$. Now consider any word $\mathbf{x} = (\mathbf{x}_L, \mathbf{x}_R)$ of weight three. Since $wt(\mathbf{x}) = 3$, either $0 \leq wt(\mathbf{x}_L) \leq 1$ or $0 \leq wt(\mathbf{x}_R) \leq 1$. In both cases $d(\mathbf{x}, \mathcal{L}) = 1$. Furthermore, $d(\mathbf{x}, \mathcal{K}) \leq 3$. Since the codes $\mathcal{B}$, $\mathcal{K}$, and $\mathcal{L}$ are self-complementary, it follows that code $\mathcal{C}$ has covering radius three and that the code $\mathcal{D} := \mathcal{B} \cup \mathcal{L}$ has covering radius two. Code $\mathcal{D}$ has parameters $(12,192,2)2$ and contains $\mathcal{C}$ as a subcode, for suitable choice of $\mathcal{K}$.

All the words of codes $\mathcal{C}$ and $\mathcal{D}$ have even weight, hence puncturing these codes on any coordinate decreases the covering radius by one. Thus we obtain an $(11,192,1)1$ code containing an $(11,144,3)2$ code as subcode.

# Bibliography

[1] D. Ashlock, private communication.

[2] M. Beveraggi, G.D. Cohen, "On the Density of Best Coverings in Hamming Space," in *Proceedings of the 2nd International Colloquium on Coding Theory and Applications*, G.D. Cohen et al., Ed., Lecture Notes in Computer Science, Vol. 311, New York: Springer-Verlag, 1986, pp. 39-44.

[3] V.M. Blinovskii, "Lower Asymptotic Bound on the Number of Linear Codewords in a Sphere of Given Radius in $\mathbb{F}_q^n$," *Probl. Inform. Transm.*, Vol. 23, pp. 130-132, April-June 1987.

[4] A. Blokhuis, C.W.H. Lam, "More Coverings by Rook Domains," *Journal of Combinatorial Theory*, A **36**, pp. 240-244, 1984.

[5] A.E. Brouwer, A. Schrijver, "The Blocking Number of an Affine Space," *Journal of Combinatorial Theory*, A **24**, pp. 251-253, 1978.

[6] A.E. Brouwer, L.M.G.M. Tolhuizen, "A Sharpening of the Johnson Bound for Binary Linear Codes and the Nonexistence of Linear Codes with Preparata Parameters," *Designs, Codes and Cryptography*, Vol. **3**, pp. 95-98, May 1993.

[7] A.E. Brouwer, T. Verhoeff, "An Updated Table of Minimum-Distance Bounds for Binary Linear Codes," *IEEE Trans. Inform. Theory*, Vol. IT-39, pp. 662-677, March 1993.

[8] R.A. Brualdi, V.S. Pless, R.M. Wilson, "Short Codes with a Given Covering Radius," *IEEE Trans. Inform. Theory*, Vol. IT-35, pp. 99-109, January 1989.

[9] R.A. Brualdi, V.S. Pless, "On the Length of Codes with a Given Covering Radius," in *Coding Theory and Design Theory, Part 1*, D.R. Chaudhuri, Ed., New York: Springer-Verlag, 1990, pp. 9-15.

[10] P.B. Buschbach, Michiel G.L. Gerritzen, H.C.A. van Tilborg, "On the Covering Radius of Binary, Linear Codes Meeting the Griesmer Bound," *IEEE Trans. Inform. Theory*, Vol. IT-31, pp. 465-468, July 1985.

[11] A.R. Calderbank, N.J.A. Sloane, "Inequalities for Covering Codes," *IEEE Trans. Inform. Theory*, Vol. IT-34, pp. 1276-1280, September 1988.

[12] A.R. Calderbank, "Covering Bounds for Codes," *Journal of Combinatorial Theory*, **A**, pp. 117-122, 1992.

[13] L.R.A. Casse, "A Solution to Beniamino Segre's 'Problem $I_{r,q}$' for $q$ even," *Accademia Nazionale dei Lincei, Rendiconti della Classe di Scienze fisiche, matematiche e naturali*, **46**, pp. 13-20, January 1969.

[14] G.D. Cohen, "A Nonconstructive Upper Bound on the Covering Radius," *IEEE Trans. Inform. Theory*, Vol. IT-29, pp. 352-353, May 1983.

[15] G.D. Cohen, P. Frankl, "Good Coverings of Hamming Spaces with Spheres," *Discrete Mathematics*, **56**, pp. 125-131, 1985.

[16] G.D. Cohen, M.G. Karpovsky, H.F. Mattson Jr., J.R. Schatz, "Covering Radius — Survey and Recent Results," *IEEE Trans. Inform. Theory*, Vol. IT-31, pp. 328-343, May 1985.

[17] G.D. Cohen, A.C. Lobstein, N.J.A. Sloane, "Further Results on the Covering Radius of Codes," *IEEE Trans. Inform. Theory*, Vol. IT-32, pp. 680-694, September 1986.

[18] G.D. Cohen, I.S. Honkala, S.N. Litsyn, H.F. Mattson Jr., "Weighted Coverings and Packings," preprint (1994).

[19] A.A. Davydov, L.M. Tombak, "Quasiperfect Linear Binary Codes with Distance 4 and Complete Caps in Projective Geometry," *Probl. Inform. Transm.*, Vol. 25, No. 4, pp. 265-275, October-December 1989.

[20] A.A. Davydov, "Construction of Linear Covering Codes," *Probl. Inform. Transm.*, Vol. 26, No. 4, pp. 317-331, October-December 1990.

[21] A.A. Davydov, "Constructions and Families of $q$-ary Linear Covering Codes and Saturated Sets of Points in Projective Geometry," in *Proceedings of the Fifth Joint Soviet-Swedish Workshop on Information Theory*, Moscow, USSR, January 13-19, 1991, pp. 46-49.

[22] A.A. Davydov, A.Yu. Drozhzhina-Labinskaya, "Constructions of Binary Linear Covering Codes," in *Proceedings of the International Workshop on Algebraic and Combinatorial Theory*, Vaneshta Voda, Bulgaria, June 22-28, 1992, pp. 51-54.

[23] A.A. Davydov, A.Yu. Drozhzhina-Labinskaya, "Constructions, Families and Tables of Binary Linear Covering Codes," to appear in *IEEE Trans. Inform. Theory*.

[24] P. Delsarte, "An Algebraic Approach to the Association Schemes of Coding Theory," Ph.D. Thesis, Université Catholique de Louvain, 1973.

[25] P. Delsarte, "Four Fundamental Parameters of a Code," MBLE Research Lab., Report R 184, January 1972.

[26] R. Dougherty, H. Janwa, "Covering Radius Computations for Binary Cyclic Codes," *Mathematics of Computation*, Vol. 57, pp. 415-434, July 1991.

[27] T. Etzion, G. Greenberg, "Constructions for Perfect Mixed Codes and Other Covering Codes," *IEEE Trans. Inform. Theory*, Vol. IT-39, pp. 209-214, January 1993.

[28] T. Etzion, G. Greenberg, I.S. Honkala, "Normal and Abnormal Codes," *IEEE Trans. Inform. Theory*, Vol. IT-39, pp. 1453-1456, July 1993.

[29] P. Frankl, V. Rödl, "Near Perfect Coverings in Graphs and Hypergraphs," *Europ. J. Combinatorics*, **6**, pp. 317-326, 1985.

[30] E.M. Gabidulin, A.A. Davydov, L.M. Tombak, "Linear Codes with Covering Radius 2 and Other New Covering Codes," *IEEE Trans. Inform. Theory*, Vol. IT-37, pp. 219-224, January 1991.

[31] A. Gersho, R.M. Gray, *Vector Quantization and Signal Compression*, Boston-Dordrecht-London: Kluwer, 1992.

[32] R.L. Graham, N.J.A. Sloane, "On the Covering Radius of Codes," *IEEE Trans. Inform. Theory*, Vol. IT-31, pp. 385-401, May 1985.

[33] L. Habsieger, "Lower Bounds for $q$-ary Coverings by Spheres of Radius One," *Journal of Combinatorial Theory*, A **67**, pp. 199-222, 1994.

[34] H.O. Hämäläinen, I.S. Honkala, M. K. Kaikkonen, S.N. Litsyn, "Bounds for Binary Multiple Covering Codes," *Designs, Codes and Cryptography*, Vol. 3, pp. 251-275, July 1993.

[35] A.R. Hammons Jr., P.V. Kumar, A.R. Calderbank, N.J.A. Sloane, P. Solé, "The $\mathbb{Z}_4$-Linearity of Kerdock, Preparata, Goethals, and Related Codes," *IEEE Trans. Inform. Theory*, Vol. IT-40, pp. 301-319, March 1994.

[36] T. Helleseth, T. Klöve, J. Mykkeltveit, "On the Covering Radius of Binary Codes," *IEEE Trans. Inform. Theory*, Vol. IT-24, pp. 627-628, September 1978.

[37] I.S. Honkala, "Lower Bounds for Binary Covering Codes," *IEEE Trans. Inform. Theory*, Vol. IT-34, pp. 326-329, March 1988.

[38] I.S. Honkala, "A New Construction for Covering Codes," *IEEE Trans. Inform. Theory*, Vol. IT-34, pp. 1343-1344, September 1988.

[39] I.S. Honkala, "Lower Bounds for $q$-ary Covering Codes," *IEEE Trans. Inform. Theory*, Vol. IT-36, pp. 664-671, May 1990.

[40] I.S. Honkala, "Modified Bounds for Covering Codes," *IEEE Trans. Inform. Theory*, Vol. IT-37, pp. 351-365, March 1991.

[41] I.S. Honkala, H.O. Hämäläinen, "Bounds for Abnormal Binary Codes with Covering Radius One," *IEEE Trans. Inform. Theory*, Vol. IT-37, pp. 372-375, March 1991.

[42] I.S. Honkala, "On $(k, t)$-Subnormal Covering Codes," *IEEE Trans. Inform. Theory*, Vol. IT-37, pp. 1203-1206, July 1991.

[43] I.S. Honkala, "On Lengthening of Covering Codes," *Discrete Mathematics* 106/107, pp. 291-295, 1992.

[44] I.S. Honkala, "A Lower Bound on Binary Codes with Covering Radius One," presented at the French-Israeli Workshop of Algebraic Coding Theory, ENST, Paris, July 19-21, 1993.

[45] X-D. Hou, "Some Results on the Norm of Codes," *IEEE Trans. Inform. Theory*, Vol. IT-36, pp. 683-685, May 1990.

[46] X-D. Hou, "New Lower Bounds for Covering Codes," *IEEE Trans. Inform. Theory*, Vol. IT-36, pp. 895-899, July 1990.

[47] X-D. Hou, "An Improved Sphere Covering Bound for the Codes with $n = 3R + 2$," *IEEE Trans. Inform. Theory*, Vol. IT-36, pp. 1476-1478, November 1990.

[48] X-D. Hou, "Binary Linear Quasi-Perfect Codes are Normal," *IEEE Trans. Inform. Theory*, Vol. IT-37, pp. 378-379, March 1991.

[49] R.E. Jamison, "Covering Finite Fields with Cosets of Subspaces," *Journal of Combinatorial Theory*, A **22**, pp. 253-266, 1977.

[50] H. Janwa, "Some New Upper Bounds on the Covering Radius of Binary Linear Codes," *IEEE Trans. Inform. Theory*, Vol. IT-35, pp. 110-122, January 1989.

[51] G.A. Kabatyanskii, V.I. Panchenko, "On Sphere Packings and Coverings of the Hamming Space," *Probl. Inform. Transm.*, Vol. 24, pp. 261-272, October-December 1988.

[52] H.J.L. Kamps, J.H. van Lint, "A Covering Problem," *Colloquia Mathematica Societatis János Bolyai*, 4, pp. 679-685, 1970.

[53] H.J.L. Kamps, J.H. van Lint, "The Football Pool Problem for Five Matches," *Journal of Combinatorial Theory*, **3**, pp. 315-325, 1967.

[54] K.E. Kilby, N.J.A. Sloane, On the Covering Radius Problem for Codes I. Bounds on Normalized Covering Radius," *SIAM J. Alg. Disc. Meth.*, Vol. 8, pp. 604-618, October 1987.

[55] K.E. Kilby, N.J.A. Sloane, On the Covering Radius Problem for Codes II. Bounds on Normalized Covering Radius," *SIAM J. Alg. Disc. Meth.*, Vol. 8, pp. 619-627, October 1987.

[56] P.J.M. van Laarhoven, E.H.L. Aarts, J.H. van Lint, L.T. Wille, "New Upper Bounds for the Football Pool Problem for 6,7 and 8 Matches," *Journal of Combinatorial Theory*, A **52**, pp. 304-312, 1989.

[57] J. Lahtonen, "An Optimal Polynomial for a Covering Radius Problem," private communication.

[58] J.H. van Lint, *Introduction to Coding Theory*, New York-Heidelberg-Berlin: Springer-Verlag, 1982.

[59] J.H. Van Lint Jr., "Covering Radius Problems," Master's Thesis, Eindhoven University of Technology, June 1988.

[60] C.L. Liu, B.G. Ong, G.R. Ruth, "A Construction Scheme for Linear and Nonlinear Codes," Discrete Mathematics, 4, pp. 171-184, 1973.

[61] S. Litsyn, P. Solé, A. Tietäväinen, "New Upper Bounds on the Covering Radius of Codes with Known Dual Distance," in *Proceedings of the 1994 IEEE International Symposium on Information Theory*, Trondheim, Norway, June 27-July 1, 1994, p. 304.

[62] A.C. Lobstein, G.J.M. van Wee, "On Normal and Subnormal $q$-ary Codes," *IEEE Trans. Inform. Theory*, Vol. IT-35, pp. 1291-1295, November 1989.

[63] J.E. MacDonald, "Design Methods for Maximum-Distance Error-Correcting Codes," *IBM J. Res. Devel.*, 4, pp. 43-57, January 1960.

[64] F.J. MacWilliams, N.J.A. Sloane, *The Theory of Error-Correcting Codes*, Amsterdam-New York-Oxford-Tokio: North-Holland, 1977.

[65] H.F. Mattson Jr., "An Improved Upper Bound on Covering Radius," in *Proceedings of AAECC-2*, Lecture Notes in Computer Science, Vol. 228, A. Poli, Ed., New York: Springer-Verlag, 1984, pp. 90-106.

[66] P.R.J. Östergård, "A New Binary Code of Length 10 and Covering Radius 1," *IEEE Trans. Inform. Theory*, Vol. IT-37, pp. 179-180, January 1991.

[67] P.R.J. Östergård, "Upper Bounds for $q$-ary Covering Codes," *IEEE Trans. Inform. Theory*, Vol. IT-37, pp. 660-664, May 1991.

[68] P.R.J. Östergård, "Further Results on $(k,t)$-Subnormal Covering Codes," *IEEE Trans. Inform. Theory*, Vol. IT-38, pp. 206-210, January 1992.

[69] P.R.J. Östergård, "New Upper Bounds for the Football Pool Problem for 11 and 12 Matches," *Journal of Combinatorial Theory*, A **67**, pp. 161-168, 1994.

[70] V. Rödl, "On a Packing and Covering Problem," *Europ. J. Combinatorics*, **6**, pp. 69-78, 1985.

[71] J. Simonis, "Covering Radius: Improving on the Sphere Covering Bound," in *Proceedings of AAECC-6*, Lecture Notes in Computer Science, Vol. 357, T. Mora, Ed., New York: Springer-Verlag, 1989, pp. 377-385.

[72] J. Simonis, "The Minimal Covering Radius t[15,6] of a Six-Dimensional Binary Linear Code of Length 15 is Equal to 4," *IEEE Trans. Inform. Theory*, Vol. IT-34, pp. 1344-1345, September 1988.

[73] N.J.A. Sloane, S.M. Reddy, C-L. Chen, "New Binary Codes," *IEEE Trans. Inform. Theory*, Vol. IT-18, pp. 503-510, July 1972.

[74] N.J.A. Sloane, D.S. Whitehead, "New Family of Single-Error Correcting Codes," *IEEE Trans. Inform. Theory*, Vol. IT-16, pp. 717-719, November 1970.

[75] P. Solé, K.G. Mehrotra, "Generalization of the Norse Bounds to Codes of Higher Strengths," *IEEE Trans. Inform. Theory*, Vol. IT-37, pp. 190-192, January 1991.

[76] P. Solé, P. Stokes, "Covering Radius, Codimension, and Dual-Distance Width," *IEEE Trans. Inform. Theory*, Vol. IT-39, pp. 1195-1203, 1993.

[77] P. Solé, "Short Proofs of the Tietäväinen-Levenshtein Bounds," submitted to *IEEE Trans. Inform. Theory*.

[78] R.G. Stanton, J.G. Kalbfleisch, "Covering Problems for Dichotomized Matchings," *Aequationes Math.*, Vol. 1, pp. 94-103, 1968.

[79] R. Struik, "On the Structure of Linear Codes with Covering Radius Two and Three," to appear in *IEEE Trans. Inform. Theory*, September 1994 issue.

[80] R. Struik, "An Improvement of the Van Wee Bound for Binary Linear Covering Codes," to appear in *IEEE Trans. Inform. Theory*, July 1994 issue.

[81] A. Tietäväinen, "An Upper Bound on the Covering Radius as a Function of the Dual Distance," *IEEE Trans. Inform. Theory*, Vol. IT-36, pp. 1472-1474, November 1990.

[82] H.C.A. van Tilborg, "On the Uniqueness resp. Non-existence of Certain Codes Meeting the Griesmer Bound," *Inform.Contr.*, Vol. 44, pp. 16-35, January 1980.

[83] A. Vardy, T. Etzion, "Some Constructions of Perfect Binary Codes," in *Proceedings of AAECC-10*, Lecture Notes in Computer Science, Vol. 673, G.D. Cohen et al., Ed., New York: Springer-Verlag, 1993, pp. 344-354.

[84] G.J.M. van Wee, "Improved Sphere Bounds on the Covering Radius of Codes," *IEEE Trans. Inform. Theory*, Vol. IT-34, pp. 237-245, March 1988.

[85] G.J.M. van Wee, "More Binary Covering Codes are Normal," *IEEE Trans. Inform. Theory*, Vol. IT-36, pp. 1466-1470, November 1990.

[86] G.J.M. van Wee, "Some New Lower Bounds for Binary and Ternary Covering Codes," *IEEE Trans. Inform. Theory*, Vol. IT-39, pp. 1422-1424, July 1993.

[87] Ø. Ytrehus, T. Helleseth, "There is No Binary $[25, 8, 10]$ Code," *IEEE Trans. Inform. Theory*, Vol. IT-36, pp. 3695-696, May 1990.

[88] Ø. Ytrehus, "Binary [18,11]2 Codes Do Not Exist — Nor Do [64,53]2 Codes," *IEEE Trans. Inform. Theory*, Vol. IT-37, pp. 349-351, March 1991.

[89] Z. Zhang, "Linear Inequalities for Covering Codes: Part I — Pair Covering Inequalities," *IEEE Trans. Inform. Theory*, Vol. IT-37, pp. 573-582, May 1991.

[90] Z. Zhang and C. Lo, "Lower Bounds on $t[n, k]$ from Linear Inequalities," *IEEE Trans. Inform. Theory*, Vol. IT-38, pp. 194-197, January 1992.

[91] Z. Zhang and C. Lo, "Linear Inequalities for Covering Codes: Part II — Triple Covering Inequalities," *IEEE Trans. Inform. Theory*, Vol. IT-38, pp. 1648-1662, November 1992.

# Notation

Our notation follows [58,64]. For elementary notation from coding theory we refer to Chapter 1. Below we mention some notations that are frequently used throughout this thesis or that may not be standard.

| | |
|---|---|
| $\langle \mathbf{x}, \mathbf{y} \rangle$ | standard inner product of vectors $\mathbf{x}$ and $\mathbf{y}$ |
| $\mathbf{0}$, $\mathbf{1}$, $\mathbf{e}_i$ | all-zero vector, all-one vector, $i$th unit vector |
| $I_n$, $O$, $J$ | $n \times n$ identity matrix, all-zero matrix, all-one matrix |
| $A \otimes B$ | kronecker product of matrices $A$ and $B$ |
| $A(n,d)$ | maximum size of any code of length $n$ with minimum distance $d$ |
| $A(n,d,w)$ | maximum size of any constant weight code of length $n$ with all weights $w$, and minimum distance $d$ |
| $d[n,k]$ | maximum distance achievable by any $[n,k]$ code |
| $K(n,r)$ | minimum size of any code of length $n$ with covering radius $r$ |
| $t[n,k]$ | minimum covering radius achievable by any $[n,k]$ code |
| $l(m,r)$ | minimum length of any linear code with redundancy $m$ and covering radius $r$ |
| $n_d^*(m,r)$ | minimum length of any systematic code with redundancy $m$, covering radius $r$, and distance $d$ |
| $\mu_d(n,r)$ | minimum density of any code with length $n$, covering radius $r$, and minimum distance $d$ |
| $\mathcal{H}_m$ | Hamming code of length $n = 2^m - 1$ |
| $\mathcal{R}(r,m)$ | $r$th order Reed-Muller code of length $2^m$ |
| $\mathcal{P}_m$ | Preparata code of length $2^m - 1$, $m$ even |
| $C_e$, $C_o$ | even, resp. odd, weight subcode of code $C$ |
| $C[i]$ | punctured code of $C$, punctured on $i$th coordinate |
| $C/C'$ | a code $C$ that is the union of translates of code $C'$ |
| $C_1 \dot{\oplus} C_2$ | amalgamated direct sum of codes $C_1$ and $C_2$ |
| $f_e(v,k,t)$ | maximum size of any $t$-$(v,k,1)$ packing design |
| $f_c(v,k,t)$ | minimum size of any $t$-$(v,k,1)$ covering design |
| $\cup S$ | union of the elements of set $S$ |
| $\sum S$ | sum of elements of set $S$ |
| $U + V$ | sum of sets $U$ and $V$ |
| $K_k(x;n)$ | Krawtchouk polynomial of degree $k$ |

# Samenvatting

Dit proefschrift beschrijft de resultaten van mijn onderzoek op het gebied van de coderingstheorie. Het betreft onderzoek naar overdekkingscodes. Deze codes kunnen worden gebruikt voor het benaderen van willekeurige digitale informatie door een eindig aantal codewoorden, zonder dat hierdoor al te grote fouten optreden. Door dit benaderingsproces kan informatie efficiënter en dus korter worden beschreven. Dit proces introduceert informatieverlies. Door een geschikt ontwerp van overdekkingscodes is het mogelijk het maximale informatieverlies, gegeven door de overdekkingsstraal, te beperken.

In hoofdstuk 1 behandelen we in het kort enkele onderwerpen uit de coderingstheorie welke in de rest van het proefschrift van pas komen. Tevens bewijzen we een lemma dat ons in staat stelt vele resultaten uit de coderingstheorie op eenvoudige en uniforme wijze te bewijzen. De bovengrenzen voor de overdekkingsstraal welke door Tietäväinen en Delsarte werden behaald blijken als speciaal geval op te treden.

In hoofdstuk 2 behandelen we ondergrenzen voor het aantal codewoorden in overdekkingscodes. We laten zien dat de meeste bekende grenzen voor deze codes een directe analogie hebben met een bekende bovengrens uit de coderingstheorie, welke reeds in de jaren zestig door Johnson werd bewezen. Een belangrijk resultaat van dit hoofdstuk is een verbetering van de zgn. Van Wee ondergrens voor binaire lineaire codes. Veel andere resultaten geven verbanden aan tussen bekende grenzen voor foutencorrigerende codes en overdekkingscodes.

In hoofdstuk 3 bestuderen we ondergrenzen voor het aantal codewoorden in een lineaire code aan de hand van de structuur van zijn zgn. duale code. We laten zien dat de parameters van een overdekkingscode eisen opleggen aan de gewichtsverdeling en de doorsnede van woorden in die duale code. Een combinatie van deze ontwerpeisen met technieken uit de theorie van foutenverbeterende codes levert op dat al te zuinige overdekkingscodes vaak niet mogelijk zijn. Als toepassing van de theorie bewijzen we op eenvoudige wijze een vermoeden van Brualdi, Pless en Wilson.

In hoofdstuk 4 komen constructies voor zuinige overdekkingscodes aan de orde. Uitgangspunt bij deze constructies zijn enkele constructies die bij het ontwerp van foutenverbeterende codes al hun vruchten hebben afgeworpen. We laten zien dat deze constructie ook toegepast kan worden bij het ontwerpen van goede overdekkingscodes, d.w.z. overdekkingscodes met relatief weinig codewoorden. Een deel van de resultaten was reeds bekend, maar wordt hier op een nieuwe manier gepresenteerd met eenvoudiger bewijzen. Het belangrijkste resultaat van het hoofstuk is een eenvoudige constructie van een klasse van overdekkingscodes met overdekkingsstraal twee, minimum afstand vier, en een relatieve dichtheid die dicht bij 1 ligt, hetgeen vrijwel optimaal is. De beste tot nu toe bekende constructie leverde codes op met een relatieve dichtheid van 9/8. Verder geven we o.a. een generalisatie van het begrip normale code, die ons in staat stelt de meeste eigenschappen van deze codes op een eenvoudige wijze te bewijzen.

# Curriculum Vitae

I was born on July 6, 1965, in Rotterdam, the Netherlands. I received my high school education at the Eindhovens Protestants Lyceum and obtained my Gymnasium-$\beta$ diploma in 1983. Afterwards I started my studies in Computer Science at the Eindhoven University of Technology. In May 1988 I graduated with distinction and passed examination in eight additional subjects. My Master's thesis discussed relations between the operational and formal behaviour of functional programming languages. During my studies I was a teaching assistent in numerical mathematics from September 1986 till April 1987, and a trainee at PTT Research Laboratories in Leidschendam from January 1987 till June 1987 (investigating data security systems). From July 1988 till November 1989 I performed my military duties as an army officer at NATO Headquarters in Brussels, Belgium, where I evaluated in projectteams information systems under development throughout NATO. Since December 1989 I have worked as a researcher in the Discrete Mathematics group of the Eindhoven University of Technology, with support from the Netherlands Organization for Scientific Research (NWO). I did research on coding theory and cryptology and taught several courses in combinatorics. The results of my research on coding theory formed the topic of this Ph.D. dissertation.

# Stellingen
# behorende bij het proefschrift
# *Covering Codes*
# van René Struik

1. In 1978, McEliece [1] introduced a probabilistic public-key cryptosystem based on error-correcting codes. This scheme allows implementations which are two to three orders of magnitude faster than RSA [2]. It has two major drawbacks, however: the key size is quite large and, in order for the scheme to be computationally secure, an information rate close to $\frac{1}{2}$ is necessary.
   Covering codes provide an excellent method for data reduction of the ciphertext, thus effectively increasing the information rate. For the optimization of the McEliece scheme discussed in [3] this results in an increase of the information rate of the original scheme (0.619) by over 20%, without affecting the system's security or the implementation's efficiency.

   [1] R.J. McEliece, "A Public-Key Cryptosystem Based on Algebraic Coding Theory," *DSN Progress Report 42-44*, JPL Pasadena, pp. 114-116, 1978.

   [2] *Contemporary Cryptology — The Science of Information Integrity*, G.J. Simmons, Ed., New York: IEEE Press, 1991, pp. 521-522.

   [3] J. van Tilburg, "On the McEliece Public-Key Cryptosystem," in *Advances in Cryptology — CRYPTO'88*, Lecture Notes in Computer Science, Vol. 403, S. Goldwasser, Ed., New York: Springer-Verlag, 1989, pp. 119-131.

2. Algebraic cryptosystems can be described as follows: let $\mathcal{C} \subset \mathbb{F}_2^n$ be a block code, let $\mathcal{F} : \mathbb{F}_2^k \to \mathcal{C}$ be a bijection, and let $\mathcal{Z} \subset \mathbb{F}_2^n$ be a nonempty set of error patterns. A message $\mathbf{m} \in \mathbb{F}_2^k$ is encrypted as the ciphertext $\mathbf{c} \in \mathbb{F}_2^n$ as follows:

$$\mathbf{c} = \mathcal{F}(\mathbf{m}) + \mathbf{z},$$

   where $\mathbf{z}$ is selected at random from the set $\mathcal{Z}$ of error patterns. Decryption of the ciphertext $\mathbf{c}$ is done by removing the error pattern $\mathbf{z}$ using a decoding algorithm for code $\mathcal{C}$ and subsequently applying $\mathcal{F}^{-1}$.
   The McEliece scheme is an algebraic cryptosystem that uses an error-correcting code $\mathcal{C}$ and a maximum-likelihood decoding algorithm for this code. In [1], Rao and Nam proposed a private-key algebraic cryptosystem, using a linear mapping $\mathcal{F}$ and a syndrome decoding algorithm. They claim that their system provides high information rates, small block lengths, efficient encoding/decoding, while being secure against a chosen-plaintext attack. This claim is false: it was shown in [2] that the claimed work-factor $W$ of a chosen-plaintext attack can be reduced to approximately $\sqrt[k]{W}$. The following table summarizes the cost of the chosen-plaintext attack (where $N = |\mathcal{Z}|$):

|  | Rao and Nam [1] | Struik and Van Tilburg [2] |
|---|---|---|
| key size (bits) | $O(nN)$ | $O(nN)$ |
| encryptions (average) | $O(N \log N)$ | $O(kN \log N)$ |
| bit operations | $\Omega(knN^k)$ | $O(knN^2 \log N)$ |
| memory needed (bits) | $O(nN)$ | $O(nN)$ |

Using the Birthday Paradox, one can replace the number $N$ by $\sqrt{N}$, both in the workload and in the average number of encryptions needed.

[1] T.R.N. Rao, K-H. Nam, "Private-Key Algebraic Cryptosystems," in *Advances in Cryptology — CRYPTO'86*, Lecture Notes in Computer Science, Vol. 263, A.M. Odlyzko, Ed., New-York: Springer-Verlag, 1987, pp. 35-48.

[2] R. Struik, J. van Tilburg, "The Rao-Nam Scheme is Insecure Against a Chosen-Plaintext Attack," in *Advances in Cryptology — CRYPTO'87*, Lecture Notes in Computer Science, Vol. 293, C. Pomerance, Ed., New York: Springer-Verlag, 1988, pp. 445-457.

3. In [1,2] modifications are studied of the Rao-Nam scheme, which was shown to be vulnerable to a chosen-plaintext attack (see previous statement). In his Ph.D. dissertation [1], Denny proposes a private-key algebraic cryptosystem using a Preparata code with associated maximum-likelihood decoding algorithm and a nonlinear mapping $\mathcal{F}$. The main result of [1] is the claim that this scheme allows efficient encoding/decoding and is secure against a chosen-plaintext attack, even for small block lengths. This claim is false: it was shown in [3] that the claimed workfactor $W$ of a chosen-plaintext attack can be reduced to approximately $\sqrt[t]{W}$, where $t \approx 2n/\log n$. The following table summarizes the cost of the chosen-plaintext attack (where $N = |\mathcal{Z}|$):

|  | Denny [1] | Struik [3] |
|---|---|---|
| key size (bits) | $O(nN)$ | $O(nN)$ |
| encryptions (average) | $O(N \log N)$ | $O(kN \log N)$ |
| bit operations | $\Omega(n^{2k})$ | $O(n^{\log n})$ |
| memory needed (bits) | $O(nN)$ | $O(nN)$ |

Using the Birthday Paradox and the particular form of mapping $\mathcal{F}$ proposed in [1], the required workload can even be reduced to only $O(n^3)$ operations.

[1] W.F. Denny, "Encryptions Using Linear and Non-Linear Codes: Implementation and Security Considerations," Ph.D. Dissertation, The Center for Advanced Computer Studies, University of Southwestern Louisiana, Lafayette, 1988.

[2] T.R.N. Rao, K-H. Nam, "Private-Key Algebraic-Code Encryptions," *IEEE Trans. Inform. Theory*, Vol. IT-35, pp. 829-833, July 1989.

[3] R. Struik, "On the Rao-Nam Scheme Using Nonlinear Codes," in *Proceedings of the 1991 IEEE International Symposium on Information Theory*, Budapest, 24-28 June, 1991, p. 174.

4. The structure of the $(23, 2^{15}, 3)2$ code discovered by Etzion et al. [1] and that of the $(23, 2^{12}, 7)$ Golay code are quite similar: shortening either code yields the $(16, 2^8, 6)$ Nordstrom-Robinson code and Hamming codes. This strongly suggests that the $(23, 2^{15}, 3)2$ code can be partitioned into eight translates of the binary Golay code. Nevertheless, this is not possible.

   [1] T. Etzion, G. Greenberg, "Constructions for Perfect Mixed Codes and Other Covering Codes," *IEEE Trans. Inform. Theory*, Vol. IT-39, pp. 209-214, January 1993.

5. Davydov et al. [1] describe a method for constructing new binary linear covering codes from old ones. Their method can be generalized to arbitrary codes using concepts from projective geometry.

   It is well-known that each point $\mathbf{p}$ of the projective space $PG(r-1, q^m)$ can be identified with an $m$-dimensional vector space $\mathcal{P}$ over $I\!F_q$. Let $V(\mathbf{p})$ be the matrix with as columns the $q^m$ points of $\mathcal{P}$ and let $V^*(\mathbf{p})$ be the matrix with as columns the $(q^m-1)/(q-1)$ projective points of $\mathcal{P}$. Let $\{\mathbf{p}_1, \ldots, \mathbf{p}_n\}$ be an $n$-arc in $PG(r-1, q^m)$. Define $P_i := V(\mathbf{p}_i)$ for all $i, 1 \leq i \leq n$, and let $P_\infty := (V^*(\mathbf{p}_1) \mid \ldots \mid V^*(\mathbf{p}_r))$.

   The construction given in [1] can now be generalized as follows: let $C$ be a $q$-ary code of length $n$ with covering radius $r$. Let $\mathcal{D}$ be the $q$-ary code defined by

$$\mathcal{D} := \{(\mathbf{x}_1, \ldots, \mathbf{x}_n, \mathbf{x}_\infty) \mid (f(\mathbf{x}_1), \ldots, f(\mathbf{x}_n)) \in C \text{ and } \sum_{i=1}^{n} \mathbf{x}_i P_i^T + \mathbf{x}_\infty P_\infty^T = \mathbf{0}\},$$

   where $f(\mathbf{y})$ is the sum of the coordinates of $\mathbf{y}$. Then code $\mathcal{D}$ has covering radius $r$ and cardinality $|\mathcal{D}| = q^{-mr}|C|q^{N-n}$, where $N = q^m n + r(q^m-1)/(q-1)$.

   [1] A.A. Davydov, A.Yu. Drozhzhina-Labinskaya, "Constructions, Families and Tables of Binary Linear Covering Codes," to appear in *IEEE Trans. Inform. Theory*.

6. Let $C$ be a binary $(n, M)r$ code. The function $\pi : \{1, \ldots, n\} \longrightarrow I\!N \cup \{0\}$ induces a partition of the coordinate positions of $C$. Partition $\pi$ is called sufficient for $C$, if for all $\mathbf{x} \in I\!F_2^n$, there exists a vector $\mathbf{e} \in I\!F_2^n$ with $\mathbf{x} - \mathbf{e} \in C$, $wt(\mathbf{e}) \leq r$, and $wt(\mathbf{e}) = |\pi(\text{supp}(\mathbf{e}))|$. The effective length $n_e$ of code $C$ is the minimum number of blocks in any partition $\pi$ that is sufficient for $C$.

   The effective length enables us to improve constructions by Davydov et al. [1] and Honkala [2]:

   Let $C$ be a binary code of length $n$ with covering radius 2 and effective length $n_e$. Then there exists a code $C'$ of length $N$ with covering radius 2 and cardinality $|C'| = 2^{-2m}|C|2^{N-n}$, where

$$N = \begin{cases} (n+1)2^m - 1 & \text{if } n_e \leq 2^m \leq n, \\ (n+2)2^m - 2 & \text{if } n_e \leq 2^m + 1, \\ (n+2)2^m - 3 & \text{if } n_e \leq 2^m, \\ (n+2)2^m - \sqrt{2^m} - 1 & \text{if } n_e \leq 2^m \text{ and } m \text{ is even.} \end{cases}$$

Moreover, there exists a code $C'$ of length $N = 2n + 2$ with covering radius 2 and cardinality $|C'| = |C|2^{2+(n-n_c)}\lfloor 2^{n_c}/3 \rfloor$.

[1] A.A. Davydov, A.Yu. Drozhzhina-Labinskaya, "Constructions, Families and Tables of Binary Linear Covering Codes," to appear in *IEEE Trans. Inform. Theory.*

[2] I.S. Honkala, "A New Construction for Covering Codes," *IEEE Trans. Inform. Theory,* Vol. IT-34, pp. 1343-1344, September 1988.

7. Let $n^*(m, r)$ be the smallest length of any systematic code with redundancy $m$ and covering radius $r$. Then

$$n^*(2m, 2) \leq \begin{cases} \frac{3}{2}2^m - 1 & \text{if } m \text{ is even,} \\ \frac{3}{2}2^m - 1 + (\sqrt{2^{m+1}} - 2) & \text{if } m \equiv 1(\text{mod } 4), \\ \frac{3}{2}2^m - 1 + (2\sqrt{2^{m+1}} - 2) & \text{if } m \equiv 3(\text{mod } 4). \end{cases}$$

This supplements Construction 4.22 of Chapter 4 of this Ph.D. dissertation.

8. Contrary to popular belief among some coding theorists, the Amalgamated Direct Sum construction produces bad results in general. Therefore, the continuous quest for characterization of normal codes, for which the only application so far has been the Amalgamated Direct Sum construction, is remarkable.

9. The use of group algebras in characterization proofs of perfect code configurations, although elegant, is not necessary.

10. In the long term, the main social impact of computerization will not so much be the replacement of labour by machinery, but — potentially much more dangerous to western society — the replacement of expensive western labour by cheaper labour from outside Europe, especially in the service industry.

[1] Robert B. Reich, "The Work of Nations: Preparing Ourselves for 21st Century Capitalism," New York: Alfred A. Knopf, 1991/2.

11. De studiebeurs is in feite een sociale uitkering en dient daarom ten laste van de begroting van het Ministerie van Sociale Zaken te komen.