Algebraic, Combinatorial and Geometric Aspects of Some Error-Correcting Codes

Dissertation

zur

Erlangung der naturwissenschaftlichen Doktorwürde
(Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

gianira nicoletta alfarano

aus

Italien

Promotionskommission

Prof. Dr. Joachim Rosenthal (Vorsitz)

Prof. Dr. Andrew Kresch

Prof. Dr. Heide Gluesing-Luerssen

Zürich, 2022

Abstract

This thesis pertains to the theory of error-correcting codes. It consists of two parts: in the first part we focus on some algebraic and geometric aspects of *minimal linear codes* and in the second part on some algebraic and combinatorial constructions of *convolutional codes*.

The minimal codewords of a linear code are those whose supports, i.e., the sets of nonzero coordinates, do not properly contain the support of other nonzero codewords. They have been used in coding theory and cryptography in decoding algorithms, in secret sharing schemes and in secure two-party computation. However, efficiently determining the minimal codewords in a given linear code is a hard problem. To cope with this, minimal codes, whose nonzero codewords are all minimal, have been constructed. In this thesis, we investigate minimal codes endowed with the Hamming and the rank metric. We introduce their geometric correspondence with cutting blocking sets for the Hamming metric case and linear cutting blocking sets for the rank-metric case. We exploit their algebraic and geometric properties in order to derive bounds on their parameters, existence results and constructions.

The second part of the thesis is dedicated to the study of convolutional codes. Convolutional codes have been introduced in 1955 by Elias as a generalization of classical linear block codes to the polynomial setting. Indeed, they are defined as modules over the ring of polynomials over a finite field having polynomial vectors as their codewords. This setup makes convolutional codes particularly suitable for application in streaming systems. In contrast to the well-developed algebraic theory of block codes, there are only a few algebraic constructions of “good” convolutional codes, i.e., codes that can correct as many errors as possible defined over a relatively small field. In this thesis, we investigates algebraic constructions of optimal convolutional codes, which are called *maximum distance profile (MDP)* codes, where the field size is comparable or smaller than the other known constructions. In order to do so, we first analyze in general the property of some polynomial matrices of having a right polynomial inverse. Finally, we propose a combinatorial construction of convolutional codes starting from *difference triangle sets*, which are collections of sets of integers such that any integer can be written in at most one way as difference of two elements in the same set.

Acknowledgments

First of all, I would like to thank my supervisor, Joachim Rosenthal, for giving me the opportunity to pursue the doctoral degree. I thank Joachim for his support and encouragement throughout the entire period, for the discussions, the coffee breaks, the riddles in the train to Neuchâtel, and, more importantly, for being the first person to put trust in me.

I would like to thank the referees and the Ph.D. committee for their careful review of this thesis.

I am extremely grateful to my research group, that at the moment consists of Henri, Julia, Niklas and Simran. I want to thank them all for great discussions and for the time spent together. I want to thank my other academic siblings, Karan and Violetta, for being the best pirate fellows ever. I thank Karan, for all the dinners and lunches he prepared for me, for the time we spent together, for being such a great friend. I thank Vio for everything we have shared, including our living together in Dublin, the lockdown, the Disney movies during the weekends, our discussions.

I would like to thank all my other co-authors, Eimear, Martino, Anina, Altan, Javier, Diego, Alberto, Veronica, Emina and Antonia for the wonderful collaborations and all the things I have learned from them. In particular, I want to thank Javier for the three months he spent in Zurich and for all the things he taught me, during my first year of Ph.D. I would like also to express my special gratitude to Martino, for the frequent visits in Paris, for his support, for always having trusted me in research and for the friendship we created. Even if we did not collaborate yet, I want to thank also Ferdinando and Matteo, for the fruitful discussions, the insights, the mutual mathematical respect, the private conversations and the complains: they are now part of my daily life.

I am also very grateful to Bettina, Carsten, Grit and Franziska. Everyone knows that the Math Institute would not work without them. They have all made my life easier and they really helped me in solving all my troubles.

I would like to thank the whole Math Institute: Alberto, Andres, Benedetta, Cristina, Fernando, Francesco, Genta, Jacopo, Krishna, Lorenzo, Nicola, Noemi, Ödul, Severin, Tabea and all the others. We have become more friends than colleagues and I could not wish for better friends than them. Thanks for the great atmosphere, the gossips and the time spent together.

I could not forget to thank the Salerstrasse's crue: Francesco, Giulia and Ivan. They have listened to my daily complains and we have shared the best moments of my life in Zurich,

together. I feel very lucky to have met them!

I thank Megghi, for being present in my life since the moment we met. I have to thank her for all the love and support, I don't know anyone that understands me like she does.

I would like to thank Giuseppe, for his friendship, support, the competition, for being the person that accept me as I am and with whom I can truly be myself in any aspect. Since the moment we met in Trento we have shared a lot of moments for which I'm really grateful to him.

I want to express my deepest gratitude to my family, who have always supported me and believed in me. They make me happy even if we are a weird family. I thank all my aunts, uncles, cousins and my brother-in-law, Manuel, for their endless support, help and constant presence. Most important, I want to thank the women of my life: my mum and my sister, Serena. They are the best part of me, my rocks, my safe place and words are not enough to thank them for everything they have done for me.

Finally, I would like to express my deepest gratitude to Alessandro, for the great collaborations, his patience and for make me feel suitable for this life. But more importantly, I want to thank him for all his love and support. Even if we have been far, I never felt alone and he has always been with me since the beginning of this adventure. I could not have done this without him.

Contents

Preface

The theory of error-correcting codes has inspired many mathematicians who were interested in applying techniques from algebra and discrete mathematics in order to progress on questions in information processing. Coding theory lies at the intersection of several disciplines in pure and applied mathematics such as algebra, number theory, probability theory, statistics, combinatorics, complexity theory, and statistical physics, which all have helped in the past to increase our knowledge in communication theory.

The *algebraic* theory of error correction originates in Shannon's seminal paper [141] from the 40's. After that, it has been quickly developed, especially thanks to Hamming's work [83], that introduced the mathematical framework of block codes endowed with the *Hamming distance*. The latter is the function $d : \mathbb{F}_q^n \times \mathbb{F}_q^n \to \mathbb{N}_0$, defined for every $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n) \in \mathbb{F}_q^n$ as $d(x, y) = |\{i \in \{1, \dots, n\} \mid x_i \neq y_i\}|$, i.e. the number of coordinates in which x and y differ. We say that a (linear block) code is a subspace of the vector space $\mathbb{F}_q^n$ over the finite field $\mathbb{F}_q$ and endowed with the metric d. The elements of a code are called *codewords*. The metric defined by the Hamming distance is still the most studied one in coding theory, but it is not the only metric used. In this thesis, among the other distance functions, we will also consider the rank metric. For two positive integers $n, m \in \mathbb{N}$ with $n \leq m$, the *rank distance* defined on the space of matrices $\mathbb{F}_q^{n\times m}$ is given by

$$d_R(A, B) = \mathrm{rk}(A - B), \qquad \text{for any } A, B \in \mathbb{F}_q^{n\times m}.$$

In the same way, an equivalent notion of *rank metric* on $\mathbb{F}_{q^m}^n$ is induced by the function

$$d_r(u, v) = \dim_{\mathbb{F}_q} \langle u_1 - v_1, \dots, u_n - v_n \rangle_{\mathbb{F}_q},$$

for $u = (u_1, \dots, u_n), v = (v_1, \dots, v_n) \in \mathbb{F}_{q^m}^n$. An $\mathbb{F}_q$-linear *(matrix) rank-metric code* is a an $\mathbb{F}_q$-subspace $\mathcal{C}$ of $\mathbb{F}_q^{n\times m}$ endowed with the distance d_R and an $\mathbb{F}_{q^m}$-linear *(vector) rank-metric code* is an $\mathbb{F}_{q^m}$-subspace $\mathcal{C}$ of $\mathbb{F}_{q^m}^n$ endowed with the distance d_r. Due to the relevance in applications to network coding, the interest in rank-metric codes has intensified over the past years, and many recent papers have been devoted to their study. In addition to their centrality in applications, rank-metric codes display distinctive combinatorial features. They were first introduced and studied from a theoretical viewpoint in 1978 by Delsarte in [63], although similar notions already appeared in 1951 [87], and later rediscovered by Gabidulin [74] in 1985 and Roth [137] in 1991.

Another central object of this thesis is the family of *convolutional codes*. Convolutional codes were introduced in 1955 by Peter Elias, in his seminal paper [67]. David Forney initially introduced the algebraic tools for the description of convolutional codes in [70]. They can be considered as a generalization of the classical block codes; indeed they are defined as modules over the polynomial ring over a finite field having polynomial vectors as their elements. The polynomial setup leads to a convolutional structure, which makes convolutional codes very suitable for application in streaming systems where the acceptable delay is rather tight and causes the need for sequential encoding and decoding. In this framework, block codes with long block length are not practicable and block codes with short block length only provide a very low correctness. Convolutional codes are particularly suitable for the erasure channel, which is the most used channel in multimedia traffic. An erasure channel is a communication channel model where sequential packets are either received or lost (at a known location). Thanks to their flexibility of grouping the blocks of information, depending on the erasures location, convolutional codes play an important role in this channel. Moreover, in this case, the decoding procedure is easy. Indeed, the main idea is to decode the part of the sequence where the distribution of erasures allows a complete correction, by using elementary linear algebra.

This thesis is divided in two parts, which cover different aspects of some block codes and convolutional codes.

Part I: The first part covers some algebraic and geometric perspectives of *minimal linear codes*, which are linear codes whose all codewords are *minimal*. A codeword c in a code $\mathcal{C} \subseteq \mathbb{F}_q^n$ is minimal if for every other codeword c' it holds the following

$$\{i \in \{1,\dots,n\} \mid c'_i \neq 0\} \subseteq \{i \in \{1,\dots,n\} \mid c_i \neq 0\} \text{ if and only if } c = \lambda c', \text{ for some } \lambda \in \mathbb{F}^*,$$

i.e. the set of the nonzero coordinates of c does not contain the set of the nonzero coordinates of any other linearly independent codeword.

The study of minimal codewords of a linear code finds application in combinatorics [61, 98], in the analysis of the Voronoi region for decoding purposes [19, 1] and in secret sharing schemes [115, 116, 19]. In particular, secret sharing schemes were introduced independently by Shamir and Blakley in 1979 [140, 39]. They are protocols used for distributing a secret among a certain number of participants. In its original framework, a secret sharing scheme works as follows: a dealer gives a share of a secret to n players in such a way that any subset of at least t players can reconstruct the secret, but no subset of less than t players can. This is also called (n,t)-threshold scheme protocol. A more general construction, based on linear codes, was first investigated by McEliece and Sarwate in 1981 [117], where Reed-Solomon codes were used. Later, several authors used other linear error-correcting codes to construct the same protocol [93, 115, 116, 65].

The set of subsets of participants which are able to recover the secret is called *access structure*. It is common to consider only subsets which do not admit proper subsets of participants able to recover the secret: we may refer to their collection as *minimal access structure*. For example, in an (n, t)-threshold scheme protocol, the access structure is given by all subsets of at least t participants, whereas the minimal access structure is given by all subsets of exactly t participants.

In [115], Massey relates the secret sharing protocol to minimal codewords: in particular, the minimal access structure in his secret sharing protocol is given by the support of the minimal codewords of a linear code $\mathcal{C}$, having first entry equal to 1. However, finding in an efficient way the minimal codewords of a general linear code is a difficult task. For this reason, the class of minimal codes has been introduced.

In this thesis, the focus is on the algebraic and geometric structure of minimal codes in the Hamming and rank metrics. This part is organized as follows.

In Chapter 2, we introduce the formal background and the basic notions on coding theory in the Hamming and rank metric, as well as the main tools we will use from finite geometry.

Chapter 3 has been built upon the publications [5] by the author in collaboration with Borello and Neri and [6] by the author in collaboration with Borello, Neri and Ravagnani. In this chapter, we introduce a correspondence between minimal codes endowed with the Hamming metric and *cutting blocking sets*. Motivated by the fact that minimal codes are *asymptotically good* and the lack of constructions of short minimal codes, we propose some geometric constructions. Employing an algebraic approach to their study, we then obtain new lower bounds for both the minimum distance and the length of such a code. This improves on known results and excludes the existence of minimal codes for several new parameter sets. Most of the methods developed in this chapter can be applied to any linear code, but they give the most explicit results when combined with the minimality property of the underlying code.

Chapter 4 is based on [6], by the author, Borello, Neri and Ravagnani. In this chapter, we investigate the geometric correspondence between rank-metric codes and q-systems. We then introduce the concept of *linear cutting blocking set* which is the q-analogue of a cutting blocking set. In this setting, linear cutting blocking sets correspond to *minimal rank-metric codes*. In order to define a notion of minimality for rank-metric codes, we analyze a natural notion of *rank support*, developing then a theory for the supports in the rank metric which is the q-analogue of the theory of the supports in the Hamming metric.

Part II: The second part of the thesis covers some algebraic and combinatorial aspects of convolutional codes. Although convolutional codes are a natural generalization of block codes, only few constructions of them have been presented in the literature; see for instance [70, 79, 80, 124, 134]. This is in contrast to the well-developed theory of block codes.

Motivated by the previous lack of algebraic and combinatorial constructions of convolutional codes, we focus in particular on *maximum distance profile (MDP) convolutional codes*, with the final aim of providing a new concrete construction. These codes are characterized by the property that they can correct the maximum number of errors per time interval.

In Chapter 6 we recall the technical background for the following chapters.

Chapter 7 is based on the publication [8], by the author and Lieb, where we investigate an important property for polynomial matrices: we say that a matrix whose entries are polynomials with coefficients in a finite field is *left prime* if it admits a right polynomial inverse. In [79], a characterization for MDP convolutional codes generated by a left prime polynomial matrix or, equivalently, defined as the kernel of a left prime polynomial matrix has been provided. From that moment, this criterion has been used for costructing MDP convolutional codes; see for instance [153], [12], [101]. Unfortunately, none of these works discuss the left primeness of the constructed matrices. In this chapter, we find some conditions that implies the left prime property and we then show that all the previous constructions actually produce noncatastrophic codes.

Chapter 8 is dedicated to the presentation of the algebraic construction of MDP convolutional codes obtained in [11], by the author in collaboration with Napp, Neri and Requena. For this, we select different matrices coming from Vandermonde ones as the coefficients of the polynomial generator matrix of the convolutional code. In this way, the resulting convolutional code is, under some constraints that will be detailed in the chapter, MDP. Due to the use of Vandermonde matrices, the codes constructed with this method can be considered as a natural extension of *generalized Reed-Solomon* block codes to the context of convolutional codes. For this reason, we call them *weighted Reed-Solomon (WRS) convolutional codes.*

Chapter 9 contains the results published in [10] by the author, Lieb and Rosenthal. It is devoted to present a combinatorial construction of low-density parity-check (LDPC) convolutional codes using difference triangle sets. These codes are defined as the kernel of a sparse matrix, i.e. with few nonzero entries. We provide a theoretical construction for such codes and we study the sufficient field size to avoid the presence of cycles in the associated Tanner graph, making them suitable for an efficient decoding algorithm.

Part I

Minimal Codes in the Hamming and Rank Metrics

Chapter 1

Introduction

In a linear code, a codeword is **minimal** if its support does not contain the support of any codeword other than its scalar multiples. A linear code is minimal if all its codewords are minimal.

In [19], Ashikhmin and Barg gave a sufficient condition for a linear code to be minimal.

Lemma 1.1. Let $\mathcal{C} \subseteq \mathbb{F}_q^n$ be an $[n,k]_q$ code, w_{min}, w_{max} be the minimum and the maximum Hamming weights in $\mathcal{C}$, respectively. Then $\mathcal{C}$ is minimal if

$$\frac{w_{min}}{w_{max}} > \frac{q-1}{q}. \tag{AB}$$

The Ashikhmin-Barg Lemma gave rise to several works with the aim of constructing minimal codes, see for example [47, 162, 64, 66]. However, condition (AB) is only sufficient. Some constructions of families of minimal codes not satisfying the condition (AB) were first presented in [54, 49]. In [85], a necessary and sufficient condition for an $\mathbb{F}_q$-linear code to be minimal was given: an $[n,k]_q$ code $\mathcal{C}$ is minimal if and only if, for every pair of linearly independent codewords $a, b \in \mathcal{C}$, we have

$$\sum_{\lambda \in \mathbb{F}_q^*} \mathrm{wt}(a + \lambda b) \neq (q-1)\mathrm{wt}(a) - \mathrm{wt}(b).$$

In the same paper, the authors constructed an infinite family of minimal linear codes not satisfying the condition (AB). This construction was generalized to finite fields with odd characteristic by Bartoli and Bonini, in [26]. In [42], Bonini and Borello investigated the geometric generalization of the construction in [26], highlighting a first link between minimal codes and cutting blocking sets. Moreover, different types of recent constructions of minimal codes based on weakly regular bent and plateaued functions have been also presented in [118, 119, 120].

In the last decade, minimal codes have been the subject of intense mathematical research. First results on minimal codes were presented in [48] and [147], where in the former paper the main motivation arises from secure two-party computation. In [48], an upper bound on the rate of a minimal codes is established and other bounds on the minimum and maximum weight of minimal codes can be found in [54].

The aim of this part of the thesis is to combine the existing literature on minimal codes with the personal contributions in [5, 7, 6], in order to provide a complete and original overview on the topic.

Organization: In Chapter 2 we provide the background needed for understanding this part of the thesis. In Chapter 3, we survey the contributions from [5] by Alfarano, Borello and Neri and [7], by Alfarano, Borello, Neri and Ravagnani. This is the most dense chapter in the whole thesis. In [5] we propose a geometric interpretation of minimal codes in the classical context. We give a characterization of minimal linear codes in terms of *cutting blocking sets*. A remarkable property of minimal codes that we show is that they form an asymptotically good family. However, our proof is nonconstructive, hence this naturally poses the problem of *explicitly* constructing families of minimal codes of short length for a given dimension, which is equivalent to constructing small cutting blocking sets in a given projective space. Problems of this type are very natural and yet wide open challenges in the realm of extremal combinatorial structures; see e.g. [41, 21, 25]. An important contribution in this direction is [68], where the authors construct small cutting blocking sets in $\mathrm{PG}(k-1,q)$, under the assumption that the characteristic of the field is strictly greater than $k-1$ and the field size is at least $2k-3$. Because of the constraints imposed on the field size, the construction of [68] is of limited applicability in coding theory and does not address the problem of constructing asymptotically good families of minimal codes (where q is fixed and k tends to infinity together with the code length). More recently, a construction of cutting blocking sets in $\mathrm{PG}(3,q)$ and $\mathrm{PG}(5,q)$, which are smaller than the previously known ones, has been given in [28]. This construction produces minimal codes of dimension respectively 4 and 6 over a finite field of arbitrary size.

In [7], we proposed three different approaches of strong combinatorial flavour to the study of minimal codes, each of which has a particular application. Most methods apply more generally to arbitrary linear codes, but give the best and most explicit results when combined with the minimality property of the underlying code.

The idea behind the first approach is to associate to a code a multivariate polynomial, which we call the *support polynomial*. This allows us to capture the combinatorics of the nonzero codewords of a code in an algebraic fashion, characterizing the inclusion relations among supports as the nonvanishing of a polynomial of bounded degree. We then study the support polynomial using tools from algebraic combinatorics, most notably the Alon-Füredi Theorem. As an application of this method, we obtain new lower bounds for both the minimum distance and the length of a minimal code. This improves on known results and excludes the existence of minimal codes for several new parameter sets.

The second approach uses instead ideas from statistics. More precisely, we regard the weight of a nonzero codeword as a discrete random variable and use Pless' equations, along with classical inequalities, to compare its mean and variance. All of this establishes inequalities between the maximum and minimum weight in a linear code, which are sharp for certain code families. In turn, these yield a new upper bound for the minimum distance of a minimal code and exclude the existence of such codes for yet other parameter sets.

Finally, the third approach is based on the correspondence between minimal codes and cutting blocking sets in finite geometry. We first reduce the problem of constructing short minimal codes to that of constructing cutting blocking sets of small cardinality. Then we show how to use the theory of *spreads* in projective spaces to obtain cutting blocking sets whose parameters can be computed explicitly. The applications of this geometric approach are twofold: On the one hand, we obtain new explicit constructions of short minimal codes; on the other hand, we establish a recursive upper bound for the least length of a minimal code over $\mathbb{F}_q$ having prescribed dimension.

In Chapter 4 we focus on codes endowed with the rank metric, following [6], by Alfarano, Borello, Neri and Ravagnani. In the last decade, especially thanks to the advent of network coding [100, 2, 145, 95], the novel class of rank-metric codes has been the subject of intense mathematical research. Interesting progress has been recently made in the attempt of understanding the connection between rank-metric codes and finite geometry [129], yet this link is still not fully understood and rather unexplored.

The starting point of our investigation is a connection between rank-metric codes and the q-analogues of projective systems. This link has been observed already in [129]. Among the various new results, we show that the maximum rank of a nondegenerate rank-metric code $\mathcal{C} \subseteq \mathbb{F}_{q^m}^n$ is $\min\{m, n\}$, a quite simple property that nonetheless has interesting consequences in the theory of anticodes and minimal rank-metric codes.

We then apply the theory of q-systems to show how one can associate a Hamming-metric code to a given rank-metric code. This correspondence translates various properties of a rank-metric code into the homonymous properties in the Hamming metric. In particular, the Hamming-metric code associated to the simplex rank-metric code is (essentially) the classical simplex code.

The interplay between the rank and the Hamming metric also motivates us to investigate one of the best-known parameters of a code, namely, its *total weight*. We identify a suitable rank-metric analogue of the total Hamming weight of a code and show that it has a constant value for all nondegenerate rank-metric codes with the same dimension and length. We then compute its asymptotic behaviour as the field size q tends to infinity, as well as the asymptotic behaviour of its variance under certain assumptions. This illustrates the general behaviour of these parameters over large finite fields.

Several applications of the above-mentioned results and concepts can be seen in theory of minimal rank-metric codes, a research line which is seemingly unexplored. We call a rank-metric code *minimal* if all its codewords have minimal *rank support*. Minimal rank-metric codes are the natural analogues (in the rank-metric) of minimal Hamming-metric codes.

The stepping stone in our approach is a characterization of minimal rank-metric codes via q-systems. The correspondence described above between rank-metric codes and these geometric/combinatorial structures induces a correspondence between minimal rank-metric codes and *linear cutting blocking sets*. The latter concept can be regarded as the q-analogue of the classical notion of a cutting blocking set.

The description of minimal rank-metric codes via the q-analogues of cutting blocking sets allows us to establish a lower bound for their length. More precisely, we find that a minimal rank-metric code $\mathcal{C} \subseteq \mathbb{F}_{q^m}^n$ of dimension k must satisfy

$$n \geq k + m - 1. \tag{1.1}$$

We also show that a nondegenerate rank-metric code is minimal if and only if the associated Hamming-metric code is minimal (under the correspondence described earlier). This result naturally connects the theories of minimal codes in the two metrics and makes it possible to transfer/compare results across them.

A major, rather curious difference between minimal codes in the rank and in the Hamming metric appears to be in the role played by the field size q with respect to bounds and existence results. While in the Hamming metric the field size q is a crucial parameter (e.g., minimal codes do not exist for lengths that are too small compared to a suitable multiple of the field size), most of the bounds and existence results we derive for minimal rank-metric codes do not depend on q, even when this quantity explicitly shows up in the computations.

Our main contributions to the theory of minimal codes in the rank metric lies in existence results and constructions, which we now describe very briefly. We start by giving simple examples of minimal rank-metric codes (the simplex rank-metric code and nondegenerate codes of very large length). Next, we propose a general construction of 3-dimensional minimal rank-metric codes based on the theory of scattered linear sets. The construction also proves that our lower bound for the length of a minimal rank-metric code is sharp for some (infinite) parameter sets. We then establish a general existence result for minimal rank-metric codes based on a combinatorial argument. More precisely, we show that a minimal rank-metric code $\mathcal{C} \subseteq \mathbb{F}_{q^m}^n$ of dimension $k \geq 2$ exists whenever $m \geq 2$ and

$$n \geq 2k + m - 2. \tag{1.2}$$

Comparing (1.1) with (1.2) we see that, in general, the existence of minimal rank-metric codes remains an open question only for $k - 1$ values of n (for any fixed m, k and q).

Chapter 2

Preliminaries on Minimal Codes

In this chapter we provide the background needed for this first part of the thesis. We start with a short introduction to classical linear codes and then we briefly discuss some preliminary notions of finite geometry. The interested reader is referred to [109, 157, 138] for more details on the theory of error-correcting codes and to [24] for a detailed first introduction to finite geometry.

2.1 Codes in the Hamming Metric

Let n be a positive integer and $\mathbb{F}_q$ be the finite field with q elements.

Definition 2.1. The **Hamming weight** of a vector $v = (v_1, \dots, v_n) \in \mathbb{F}_q^n$ is defined as

$$\mathrm{wt}^{\mathrm{H}}(v) = |\{1 \leq i \leq n \mid v_i \neq 0\}|.$$

The Hamming weight induces a metric on $\mathbb{F}_q^n$, known as **Hamming distance** and defines as

$$\begin{aligned} d^{\mathrm{H}} : \mathbb{F}_q^n \times \mathbb{F}_q^n &\to \mathbb{N}, \\ (u, v) &\mapsto \mathrm{wt}^{\mathrm{H}}(u - v). \end{aligned}$$

On this metric it has foundations of classical coding theory have been based in the last 70 years.

Definition 2.2. Let k, n be two positive integers, with $1 \leq k \leq n$. An linear code $\mathcal{C}$ is an $\mathbb{F}_q$-subspace of the vector space $\mathbb{F}_q^n$ of dimension k equipped with the Hamming distance. We define the **minimum distance** of the code $\mathcal{C} \subseteq \mathbb{F}_q^n$ the integer

$$d^{\mathrm{H}}(\mathcal{C}) := \min\{d^{\mathrm{H}}(u, v) \mid u, v \in \mathcal{C},\ u \neq v\}.$$

We denote a linear code $\mathcal{C} \subseteq \mathbb{F}_q^n$ of dimension k by $[n, k]_q$ code and, whenever the minimum distance d of $\mathcal{C}$ is known, we refer to it by $[n, k, d]_q$ linear code. A matrix $G \in \mathbb{F}_q^{k \times n}$ is called a

generator matrix for an $[n,k]_q$ code $\mathcal{C}$ if the rows of G form a basis of the vector space $\mathcal{C}$.

Since all the codes we will consider in this part of the thesis are linear, we will simply omit this word.

Definition 2.3. For every vector $v \in \mathbb{F}_q^n$, we define the support of v as the set of nonzero entries in v, i.e.,

$$\sigma^{\mathrm{H}}(v) = \{1 \leq i \leq n \mid v_i \neq 0\}.$$

Hence, clearly,

$$|\sigma^{\mathrm{H}}(v)| = \mathrm{wt}^{\mathrm{H}}(v).$$

The parameters n, k, d of a code $\mathcal{C} \subseteq \mathbb{F}_q^n$ with dimension k and minimum distance d are linked by the following elegant inequality:

$$d \leq n - k + 1.$$

This bound is known as Singleton bound; see [146]. The proof is quite easy, since one can observe that by deleting $d-1$ of the positions in all codewords, we have a code of length $n-d+1$ and since the minimum distance between two distinct codewords of $\mathcal{C}$ was given by d, in the new smaller code we must still have all distinct codewords. Thus, the dimension of the new code remains k.

Definition 2.4. The codes attaining the Singleton bound are called **maximum distance separable (MDS)** codes.

As we explained, an $[n,k]_q$ code $\mathcal{C}$ can be represented via the generator matrix $G \in \mathbb{F}_q^{k\times n}$, i.e.

$$\mathcal{C} = \{uG \mid \ u \in \mathbb{F}_q^k\}.$$

On the other hand, the code $\mathcal{C}$ can be seen as the kernel of another matrix.

Definition 2.5. A matrix $H \in \mathbb{F}_q^{(n-k)\times n}$ is a **parity-check matrix** for an $[n,k]_q$ code $\mathcal{C}$ if

$$\mathcal{C} = \{v \in \mathbb{F}_q^n \mid Hv^\top = 0\}.$$

For every $u, v \in \mathbb{F}_q^n$, denote by $u \cdot v$ the standard dot product between u and v, i.e.

$$u \cdot v = \sum_{i=1}^{n} u_i v_i.$$

Then we have the following definition.

Definition 2.6. Let $\mathcal{C}$ be an $[n,k]_q$ linear code. The **dual code** $\mathcal{C}^\perp$ is an $[n, n-k]_q$ linear code, defined as

$$\mathcal{C}^\perp = \{v \in \mathbb{F}_q^n \mid u \cdot v = 0 \text{ for all } u \in \mathcal{C}\}.$$

Note that a parity-check matrix for $\mathcal{C}$ is a generator matrix of $\mathcal{C}^\perp$.

Definition 2.7. An $[n,k]_q$ linear code is **nondegenerate** if there exists no coordinate position is identically zero. Furthermore, $\mathcal{C}$ is called **projective** if in one (and thus in all) generator matrix G of $\mathcal{C}$ no two columns are proportional. Note that a projective code is necessarily nondegenerate.

The Definition 2.7 immediately implies that if an $[n,k]_q$ linear code $\mathcal{C}$ is nondegenerate, then the minimum distance $d^\perp$ of its dual code $\mathcal{C}^\perp$ is at least 1.

At this point it is a natural question to ask when two linear code with the same length, dimension and minimum distance defined over the same field are actually *equivalent*. For the purpose of this thesis we are only interested in the following class of equivalence.

Definition 2.8. Let $\mathcal{G}$ be the subgroup of the group of linear automorphisms of $\mathbb{F}_q^n$ generated by the permutations of coordinates and by the multiplication of the i-th coordinate by an element in $\mathbb{F}_q^*$. Two codes $\mathcal{C}$ and $\mathcal{C}'$ are **(monomially) equivalent** if there exists $\sigma \in \mathcal{G}$ such that $\mathcal{C}' = \sigma(\mathcal{C})$.

We finally recall the following operations on codes. For any matrix $G \in \mathbb{F}_q^{k\times n}$, and set $I \subseteq \{1,\dots,n\}$ satisfying $0 < |I| < n$, we denote by $G^I \in \mathbb{F}_q^{k\times|I|}$ the submatrix whose columns are those of G indexed by I, and by $\overline{I}$ the complement of I in $\{1,\dots,n\}$, i.e. $\overline{I} = \{1,\dots,n\} \setminus I$.

Let $\mathcal{C}$ be an $[n,k,d]_q$ and $I \subseteq \{1,\dots,n\}$. We can **puncture** $\mathcal{C}$ by deleting the same coordinates $i \in I$ in each codeword. The resulting code is still linear, its length is $n - |I|$, and we denote it by $\mathcal{C}^I$ and call it **pucturing** on I. If G is a generator matrix for $\mathcal{C}$, then a generator matrix for $\mathcal{C}^I$ is $G^{\overline{I}}$.

Let C be an $[n,k,d]_q$ code and let $S \subseteq \{1,\dots,n\}$. Consider the set

$$\mathcal{C}(S) := \{c \in \mathcal{C} \mid \sigma^{\mathrm{H}}(c) \subseteq \overline{S}\},$$

i.e. the set of codewords which are 0 on S; this set is a subcode of C. Puncturing $C(S)$ on S gives a code over $\mathbb{F}_q$ of length $n - |S|$, called the code **shortened** on S and denoted C_S .

The following result is well-known.

Theorem 2.9. [88, Theorem 1.5.7] Let C be an $[n,k,d]_q$ code. Let $S \subseteq \{1,\dots,n\}$ be a set of s coordinates. Then:

1. $(C^\perp)_S = (C^S)^\perp$ and $(C^\perp)^S = (C_S)^\perp$.

2. If $s < d$ then $\mathcal{C}^S$ and $(C^\perp)_S$ have dimension k and $n-k-s$, respectively.

3. If $s = d$ and S is the set of coordinates where a minimum weight codeword is nonzero, then $\mathcal{C}^S$ and $(C^\perp)_S$ have dimensions $k-1$ and $n-d-k+1$, respectively.

2.2 Projective Systems

In this section we consider linear codes from a geometrical point of view, as detailed in [156]. We first give some background of fundamentals of finite projective geometry. For a detailed introduction we refer to the recent book by Ball [22]. Let $\mathrm{PG}(k,q)$ be the finite projective geometry of dimension k and order q. Due to a result of Veblen and Young [159], all finite projective spaces of dimension greater than two are isomorphic, and they correspond to Galois geometries. The space $\mathrm{PG}(k,q)$ can be easily seen as the vector space of dimension $k+1$ over the finite field $\mathbb{F}_q$. In this representation, the one-dimensional subspaces correspond to the points, the two-dimensional subspaces correspond to the lines, etc. Formally, we have

$$\mathrm{PG}(k,q) := \left(\mathbb{F}_q^{k+1} \setminus \{0\}\right)/\!\sim,$$

where

$$u \sim v \text{ if and only if } u = \lambda v \text{ for some } \lambda \in \mathbb{F}_q.$$

It is not hard to show by elementary counting that the number of points of $\mathrm{PG}(k,q)$ is given by

$$\theta_q(k) := \frac{q^{k+1}-1}{q-1}.$$

A ***d*-flat** Π in $\mathrm{PG}(k,q)$ is a subspace isomorphic to $\mathrm{PG}(d,q)$; if $d=k-1$, the subspace Π is called a **hyperplane**. It is clear that $\theta_q(k)$ is also the number of hyperplanes in $\mathrm{PG}(k,q)$.

Central to the geometric point of view of linear codes is the idea of a projective system.

Definition 2.10. A **projective** $[n,k,d]_q$ **system** $\mathcal{M}$ is a finite set of n points (counted with multiplicity) of $\mathrm{PG}(k-1,q)$ that do not all lie on a hyperplane and such that

$$d = n - \max\{|H \cap \mathcal{M}| \,:\, H \subseteq \mathrm{PG}(k-1,q),\ \dim(H) = k-2\}.$$

Projective $[n,k,d]_q$ systems $\mathcal{M}$ and $\mathcal{M}'$ are **equivalent** if there exists a projective isomorphism $\phi \in \mathrm{PGL}(k-1,q)$ mapping $\mathcal{M}$ to $\mathcal{M}'$ which preserves the multiplicities of the points.

Let $\mathcal{C}$ be an $[n,k]_q$ code with $k \times n$ generator matrix G. Note that multiplying any column of G by a nonzero field element yields a generator matrix for a code which is equivalent to $\mathcal{C}$. Consider the (multi)set of one-dimensional subspaces of $\mathbb{F}_q^n$ spanned by the columns of G. In this way the columns may be considered as a multiset $\mathcal{M}$ of points in $\mathrm{PG}(k-1,q)$ (where the multiplicity depends on how many times a certain column appears in the generator matrix, up to scalar multiple).

For any nonzero vector $v=(v_1,v_2,\dots,v_k)$ in $\mathbb{F}_q^k$, it follows that the projective hyperplane

$$v_1x_1 + v_2x_2 + \cdots + v_kx_k = 0$$

contains $|\mathcal{M}| - w$ points of $\mathcal{M}$ if and only if the codeword vG has weight w. Therefore, linear nondegenerate $[n,k,d]_q$ codes and projective $[n,k,d]_q$ systems are equivalent objects. Indeed, the procedure described above gives a correspondence between $[n,k,d]_q$ codes up to (monomial) equivalence and projective $[n,k,d]_q$ systems up to (projective) equivalence [156, Theorem 1.1.6]. This can be formally stated as follows. We denote by (Φ, Ψ) the correspondence

$$\{ \text{ classes of nondeg. } [n,k,d]_q \text{ codes } \} \longleftrightarrow \{ \text{ classes of projective } [n,k,d]_q \text{ systems } \}.$$

More specifically, for a class of nondegenerate $[n,k,d]_q$ code $[\mathcal{C}]$, $\Phi([C])$ is the (equivalence class of the) multiset obtained by taking the columns with multiplicities of any generator matrix of any representative of $[C]$, while Ψ is the functor that does the inverse operation. Given an equivalence class of multisets $[\mathcal{M}]$ in $\mathrm{PG}(k-1,q)$, it returns the class containing the code whose generator matrix has the points of $\mathcal{M}$, taken with multiplicities, as columns.

2.3 Codes in the Rank Metric

For a vector $v \in \mathbb{F}_{q^m}^n$ and an ordered basis $\Gamma = \{\gamma_1, \dots, \gamma_m\}$ of the field extension $\mathbb{F}_{q^m}/\mathbb{F}_q$, let $\Gamma(v) \in \mathbb{F}_q^{n \times m}$ be the matrix defined by

$$v_i = \sum_{j=1}^{m} \Gamma(v)_{ij} \gamma_j.$$

Note that $\Gamma(v)$ is constructed by simply transposing v and then expanding each entry over the basis Γ. The **Γ-support** of a vector $v \in \mathbb{F}_{q^m}^n$ is the column space of $\Gamma(v)$. It is denoted by $\sigma_\Gamma(v) \subseteq \mathbb{F}_q^n$. The following result can be obtained by a standard linear algebra argument.

Proposition 2.11. Let $v \in \mathbb{F}_{q^m}^n$.

1. We have $\sigma_\Gamma(v) = \sigma_\Gamma(\alpha v)$ for all nonzero $\alpha \in \mathbb{F}_{q^m}$ and all bases Γ.

2. The Γ-support of v does not depend on the choice of the basis Γ.

3. For all matrices $A \in \mathbb{F}_q^{n \times n}$ we have $\Gamma(vA) = A^\top \Gamma(v)$.

Proof.

1. The map

$$m_\alpha : \mathbb{F}_{q^m} \to \mathbb{F}_{q^m}, \; v \mapsto \alpha v$$

is an $\mathbb{F}_q$-linear automorphism. With respect to the basis Γ it is represented by a matrix $A_\Gamma \in \mathrm{GL}_m(q)$, and hence $\Gamma(\alpha v) = \Gamma(v) A_\Gamma$ has the same column space of $\Gamma(v)$.

2. Change of basis acts in the same way. $\Gamma(v) = \Gamma'(v)A$, where $A \in \mathrm{GL}_m(q)$ is the change-of-basis matrix such that $\Gamma = \Gamma' A$.

3. By definition, it suffies to check that for all $i \in \{1, \dots, n\}$ we have

$$(vA)_i = \sum_{j=1}^{m} (A^\top \Gamma(v))_{ij} \gamma_j.$$

On the one hand we have $(vA)_i = \sum_{\ell=1}^{n} v_\ell A_{\ell i}$. On the other hand,

$$\begin{aligned} \sum_{j=1}^{m} (A^\top \Gamma(v))_{ij} \gamma_j &= \sum_{j=1}^{m} \sum_{\ell=1}^{n} A_{\ell i} \Gamma(v)_{\ell j} \gamma_j \\ &= \sum_{\ell=1}^{n} A_{\ell i} \sum_{j=1}^{m} \Gamma(v)_{\ell j} \gamma_j \\ &= \sum_{\ell=1}^{n} v_\ell A_{\ell i}, \end{aligned}$$

as desired.

□

Definition 2.12. In the sequel, for $v \in \mathbb{F}_{q^m}^n$ we let $\sigma^{\mathrm{rk}}(v) := \sigma_\Gamma(v)$ be the **(rank) support** of v, where Γ is *any* basis of $\mathbb{F}_{q^m}/\mathbb{F}_q$. The support is well-defined by Proposition 2.11. The **rank (weight)** of a vector v is the $\mathbb{F}_q$-dimension of its support, denoted by $\mathrm{rk}(v)$.

Rank-metric codes and their fundamental parameters are defined as follows. In this thesis, we follow [74] and only concentrate on rank-metric codes that are linear over $\mathbb{F}_{q^m}$.

Definition 2.13. A **(rank-metric) code** is an $\mathbb{F}_{q^m}$-linear subspace $\mathcal{C} \subseteq \mathbb{F}_{q^m}^n$. Its elements are called **codewords**. The integer n is the **length** of the code. The **dimension** of $\mathcal{C}$ is the dimension as an $\mathbb{F}_{q^m}$-vector space and the **minimum (rank) distance** of a nonzero code $\mathcal{C}$ is

$$d(\mathcal{C}) := \min\{\mathrm{rk}(v) \,:\, v \in \mathcal{C},\, v \neq 0\}.$$

We also define the minimum distance of the zero code to be $n+1$. We say that $\mathcal{C}$ is an $[n,k,d]_{q^m/q}$ code if it has length n, dimension k and minimum distance d. When the minimum distance is not known or is irrelevant, we write $[n,k]_{q^m/q}$. A **generator matrix** of an $[n,k]_{q^m/q}$ code is a matrix $G \in \mathbb{F}_{q^m}^{k \times n}$ whose rows generate $\mathcal{C}$ as an $\mathbb{F}_{q^m}$-linear space. Finally, the **(rank) support** of an $\mathbb{F}_{q^m}$-linear rank-metric code $\mathcal{C}$ is the sum of the supports of its codewords, i.e.,

$$\sigma^{\mathrm{rk}}(\mathcal{C}) = \sum_{v \in \mathcal{C}} \sigma^{\mathrm{rk}}(v),$$

where the sum is intended as sum of vector spaces.

The support of a rank-metric codes is determined by the supports of any set of generators, as the following simple result shows.

Proposition 2.14. For every $v, w \in \mathbb{F}_{q^m}^n$, we have $\sigma^{\mathrm{rk}}(v+w) \subseteq \sigma^{\mathrm{rk}}(v) + \sigma^{\mathrm{rk}}(w)$. Moreover, if $\mathcal{C} = \langle c_1, \dots c_t \rangle_{\mathbb{F}_{q^m}} \subseteq \mathbb{F}_{q^m}^n$ is a rank-metric code, then $\sigma^{\mathrm{rk}}(\mathcal{C}) = \sigma^{\mathrm{rk}}(c_1) + \cdots + \sigma^{\mathrm{rk}}(c_t)$.

Recall that a (**linear**, **rank-metric**) **isometry** of $\mathbb{F}_{q^m}^n$ is an $\mathbb{F}_{q^m}$-linear automorphism φ of $\mathbb{F}_{q^m}^n$ that preserves the rank weight, i.e., such that $\mathrm{rk}(v) = \mathrm{rk}(\varphi(v))$ for all $v \in \mathbb{F}_{q^m}^n$. It is known that the isometry group of $\mathbb{F}_{q^m}^n$, say $\mathcal{G}(q,m,n)$, is generated by the (nonzero) scalar multiplications of $\mathbb{F}_{q^m}$ and the linear group $\mathrm{GL}_n(q)$; see e.g. [34]. More precisely, $\mathcal{G}(q,m,n) \cong \mathbb{F}_{q^m}^* \times \mathrm{GL}_n(q)$, which (right-)acts on $\mathbb{F}_{q^m}^n$ via

$$\begin{array}{rcl} (\mathbb{F}_{q^m}^* \times \mathrm{GL}_n(q)) \times \mathbb{F}_{q^m}^n & \longrightarrow & \mathbb{F}_{q^m}^n \\ ((\alpha, A), v) & \longmapsto & \alpha v A. \end{array}$$

Definition 2.15. Rank-metric codes $\mathcal{C}, \mathcal{C}' \subseteq \mathbb{F}_{q^m}^n$ are (**linearly**) **equivalent** if there exists $\varphi \in \mathcal{G}(q,m,n)$ such that $\mathcal{C}' = \varphi(\mathcal{C})$.

Observe that, by $\mathbb{F}_{q^m}$-linearity, when studying linear equivalence of $[n,k]_{q^m/q}$ codes the action of $\mathbb{F}_{q^m}^*$ is trivial. In particular, $[n,k]_{q^m/q}$ codes $\mathcal{C}$ and $\mathcal{C}'$ are equivalent if and only if there exists $A \in \mathrm{GL}_n(q)$ such that

$$\mathcal{C}' = \mathcal{C} \cdot A := \{vA \,:\, v \in \mathcal{C}\}.$$

We conclude this section with the definition of dual code, which we will use often throughout the thesis.

Definition 2.16. The **dual** of a rank-metric code $\mathcal{C} \subseteq \mathbb{F}_{q^m}^n$ with respect to the standard dot product is the rank-metric code

$$\mathcal{C}^\perp = \{v \in \mathbb{F}_{q^m}^n \,:\, u \cdot v = 0 \text{ for all } u \in \mathcal{C}\} \subseteq \mathbb{F}_{q^m}^n.$$

Recall moreover that $\dim_{\mathbb{F}_{q^m}}(\mathcal{C}) + \dim_{\mathbb{F}_{q^m}}(\mathcal{C}^\perp) = n$ for all rank-metric codes $\mathcal{C} \subseteq \mathbb{F}_{q^m}^n$.

Chapter 3

Three Combinatorial Perspectives on Minimal Codes

The results provided in this chapter are based on the publications [5] by Alfarano, Borello, Neri and [7] by Alfarano, Borello, Neri and Ravagnani. Here, we decided to add the proofs of some results that do not appear in the published papers.

Since this is the most dense chapter of the thesis, for convenience of the reader we list the main contributions, pointing to the corresponding statements.

- In Theorem 3.8, we establish a one-to-one correspondence between equivalence classes of nondegenerate minimal codes and equivalence classes of cutting blocking sets.

- In Theorem 3.18, we show that minimal codes are asymptotically good.

- As an application of methods from algebraic combinatorics, in particular the Alon-Füredi Theorem and the Combinatorial Nullstellensatz:

  1. a lower bound on the minimum distance of a minimal code (Theorem 3.44);
  2. a structural result on the maximal codewords in a linear code (Theorem 3.49);
  3. a lower bound on the block length of a minimal code (Theorem 3.51).

- Combining ideas from coding theory and statistics with the algebraic combinatorial approach outlined above:

  4. an upper bound on the minimum distance of a minimal code and a constraint on its parameters (Corollary 3.63);
  5. a result connecting the relative difference between maximum and minimum weights in a linear code with its block length (Proposition 3.66).

- Using methods from projective geometry, most notably the theory of spreads:

6. a construction of cutting blocking sets from spreads in finite geometry and of the corresponding minimal codes (Theorems 3.73 and 3.75);
7. an inductive construction of cutting blocking sets of small cardinality and of the corresponding minimal codes (Proposition 3.78 and Theorem 3.79);
8. two new general constructions of short minimal codes (Constructions 4 and 5).

3.1 Minimal Codes

We start this section with introducing the family of minimal codes.

Definition 3.1. A codeword c is an $[n,k]_q$ linear code $\mathcal{C}$ is **minimal** if for every $c' \in \mathcal{C}$,

$$\sigma^{\mathrm{H}}(c) \subseteq \sigma^{\mathrm{H}}(c') \iff c = \lambda c' \text{ for some } \lambda \in \mathbb{F}_q.$$

If all the codewords of $\mathcal{C}$ are minimal then $\mathcal{C}$ is called **minimal code**.

In [19], Ashikhmin and Barg gave a sufficient condition for a linear code to be minimal.

Lemma 3.2. Let $\mathcal{C}$ be an $[n,k]_q$ code, w_{min}, w_{max} be the minimum and the maximum Hamming weights in $\mathcal{C}$, respectively. Then $\mathcal{C}$ is minimal if

$$\frac{w_{min}}{w_{max}} > \frac{q-1}{q}. \tag{AB}$$

The Ashikhmin-Barg Lemma gave rise to several works with the aim of constructing minimal codes, see for example [47, 162, 64, 66]. However, condition (AB) is only sufficient. Some constructions of families of minimal codes not satisfying the condition (AB) were first presented in [54, 49]. In [85], a necessary and sufficient condition for an $\mathbb{F}_q$-linear code to be minimal was given: an $[n,k]_q$ code $\mathcal{C}$ is minimal if and only if, for every pair of linearly independent codewords $a, b \in \mathcal{C}$, we have

$$\sum_{\lambda \in \mathbb{F}_q^*} \mathrm{wt}(a + \lambda b) \neq (q-1)\mathrm{wt}(a) - \mathrm{wt}(b). \tag{3.1}$$

Remark 3.3. Following the notation of Definition 3.1, in a minimal code $\mathcal{C}$ any nonzero codeword c is minimal, but also **maximal** (i.e., every other codeword $c' \in \mathcal{C}$ with $\sigma^{\mathrm{H}}(c') \supseteq \sigma^{\mathrm{H}}(c)$ is a multiple of c).

The following simple result states that every minimal codeword c in a $[n,k]_q$ code $\mathcal{C}$ has weight upper bounded by $n-k+1$.

Proposition 3.4. Let $\mathcal{C}$ be an $[n,k]_q$ code. Every minimal codeword $c \in \mathcal{C}$ has $\mathrm{wt}^{\mathrm{H}}(c) \leq n-k+1$.

Proof. 1. Puncturing C on the nonzero positions of c, one gets a new code whose length is $n - \mathrm{wt}^{\mathrm{H}}(c)$ and whose dimension is $k-1$. Therefore $k-1 \leq n - \mathrm{wt}^{\mathrm{H}}(c)$.

2. Since c and c' are maximal, $\sigma^{\mathrm{H}}(c) \cap \sigma^{\mathrm{H}}(c') \neq \emptyset$. Moreover, for any $\alpha \in \mathbb{F}_q^*$, $c + \alpha c'$ has to be linearly independent from both c and c'. In particular, its support cannot contain $\sigma^{\mathrm{H}}(c)$ or $\sigma^{\mathrm{H}}(c')$. Hence, for any $\alpha \in \mathbb{F}_q^*$ there exists an index $i_\alpha \in \sigma^{\mathrm{H}}(c) \cap \sigma^{\mathrm{H}}(c')$ such that $c_{i_\alpha} + \alpha c'_{i_\alpha} = 0$. Since we have $i_\alpha \neq i_\beta$ for all $\alpha, \beta \in \mathbb{F}_q^*$ with $\alpha \neq \beta$, then $|\sigma^{\mathrm{H}}(c) \cap \sigma^{\mathrm{H}}(c')| \geq q - 1$, as desired. □

3.2 Cutting Blocking Sets and Minimal Codes

The concept of a *cutting blocking set* was introduced in [42] with the goal of constructing a family of minimal codes. However, the same objects were known earlier under various names and in different contexts. In [59] these are called *N-fold strong blocking sets* and are used for constructing small saturating sets in projective spaces over finite fields. In [68], cutting blocking sets are referred to as *generator sets* and are constructed as union of disjoint lines.

First we recall some basic background on blocking sets.

Definition 3.5. Let t, r, N be positive integers with $r < N$. A **t-fold r-blocking set** in $\mathrm{PG}(N, q)$ is a set $\mathcal{M} \subseteq \mathrm{PG}(N, q)$ such that for every $(N - r)$-flat Λ of $\mathrm{PG}(N, q)$ we have $|\Lambda \cap \mathcal{M}| \geq t$. When $r = 1$, we will refer to it as a **t-fold blocking set**. When $t = 1$, we will refer to it as an **r-blocking set**. Finally, **blocking sets** are the ones with $r = t = 1$.

Definition 3.6. Let r, N be positive integers with $r < N$. An r-blocking set $\mathcal{M}$ in $\mathrm{PG}(N, q)$ is called **cutting** if for every pair of $(N - r)$-flats Λ, Λ' of $\mathrm{PG}(N, q)$ we have

$$\mathcal{M} \cap \Lambda \subseteq \mathcal{M} \cap \Lambda' \iff \Lambda = \Lambda'.$$

Moreover, a cutting r-blocking set $\mathcal{M}$ is called **minimal** if for every $P \in \mathcal{M}$, the set $\mathcal{M} \setminus \{P\}$ is not a cutting r-blocking set.

The following result gives a different characterization of cutting blocking sets. The result follows also from [42, Theorem 3.5].

Proposition 3.7. A set $\mathcal{M} \subseteq \mathrm{PG}(N, q)$ is a cutting r-blocking set if and only if for every $(N - r)$-flat Λ of $\mathrm{PG}(N, q)$ we have $\langle \mathcal{M} \cap \Lambda \rangle = \Lambda$.
In particular, a cutting r-blocking set in $\mathrm{PG}(N, q)$ is an $(N - r + 1)$-fold blocking set.

Proof. ($\Leftarrow$) Let Λ, Λ' be $(N - r)$-flats of $\mathrm{PG}(N, q)$, such that $\mathcal{M} \cap \Lambda \subseteq \mathcal{M} \cap \Lambda'$. Then $\Lambda = \langle \mathcal{M} \cap \Lambda \rangle \subseteq \langle \mathcal{M} \cap \Lambda' \rangle = \Lambda'$, and since Λ and Λ' have the same dimension, we get $\Lambda = \Lambda'$, i.e. $\mathcal{M}$ is a cutting r-blocking set.

($\Rightarrow$) Suppose by contradiction that there exists an $(N - r)$-flat Λ such that $\langle \Lambda \cap \mathcal{M} \rangle = \Delta \subsetneq \Lambda$. Then, for every $(N - r)$-flat Λ' containing Δ we have $\Lambda' \cap \mathcal{M} \supseteq \Delta \cap \mathcal{M} = \Lambda \cap \mathcal{M}$. And therefore, $\mathcal{M}$ is not a cutting r-blocking set.

□

The following theorem has been proved in [5] and it is the main result in this section. It explain how the correspondence between nondegenerate codes and projective systems restrict to a correspondence between equivalence classes of minimal codes and equivalence classes of cutting blocking sets.

Theorem 3.8. Equivalence classes of $[n, k, d]_q$ minimal codes are in correspondence with equivalence classes of projective $[n, k, d]_q$ systems $\mathcal{M}$ such that $\mathcal{M}$ is a cutting blocking set via (Φ, Ψ). Furthermore, via the same pair of functors (Φ, Ψ), equivalence classes of $[n, k, d]_q$ reduced minimal codes are in correspondence with projective $[n, k, d]_q$ systems $\mathcal{M}$ such that $\mathcal{M}$ is a minimal cutting blocking set and every point in $\mathcal{M}$ has multiplicity 1.

Proof. The first statement follows from the definitions of the two objects. Hyperplanes $\langle v\rangle^\perp$ in $\mathrm{PG}(k-1, q)$ correspond to linearly independent codewords vG of $\mathcal{C}$. For any pair of hyperplanes $H = \langle v\rangle^\perp$ and $H' = \langle v'\rangle^\perp$ we have $\mathcal{M} \cap H \subseteq \mathcal{M} \cap H'$ if and only if $\sigma^{\mathrm{H}}(vG) \supseteq \sigma^{\mathrm{H}}(v'G)$, where G is any generator matrix of $\mathcal{C}$ and $\mathcal{M}$ is the associated projective system. Moreover, since puncturing on a coordinate of a code whose generator matrix is G coincides to removing the corresponding point from the multiset $\mathcal{M}$, we get the second statement. □

Observe that reduced minimal codes correspond to multisets $\mathcal{M}$ with no multiplicity. In particular, in order to construct minimal codes, by Theorem 3.8 we only need to construct classical sets, without multiplicity.

3.3 Bounds on Length and Distance of Minimal Codes

It is natural to ask for which values R we can produce minimal codes of rate $k/n = R$. It is in general easier to construct minimal codes with very small rate, such as simplex codes or related codes as in [26, 42]. However, a priori it is not clear if one can do it for arbitrary rates. In particular, for a given dimension k one would like to determine what is the smallest length n (and hence the largest rate $R = k/n$) such that an $[n, k]_q$ minimal code exists. In this section we provide some partial answers to these questions, proving some bounds on the length and the minimum distance of a minimal code for a fixed dimension. The characterization given in Theorem 3.8 plays a crucial role in dealing with these problems.

The following result shows that minimal codes have relatively large length with respect to their dimension and field size; see also Remark 3.60.

Theorem 3.9. Let $\mathcal{C}$ be an $[n, k]_q$ minimal code. Then

$$n \geq (k-1)q + 1.$$

Proof. If $k = 1$ there is nothing to prove, hence we assume $k \geq 2$. Choose a generator matrix, and the corresponding projective $[n, k]_q$ system $\mathcal{M}$ in $\Pi = \mathrm{PG}(k-1, q)$. Consider the set S of incident point-hyperplane pairs (P, Λ) in Π, where $P \in \mathcal{M}$ and, for every P denote by $m(P)$ the multiplicity of P. Summing over all the points of $\mathcal{M}$ we obtain

$$|S| = \sum_{P \in \mathcal{M}} m(P)\theta_q(k-2) = n\theta_q(k-2), \tag{3.2}$$

since $\theta_q(k-2)$ is the number of hyperplanes through a point.
On the other hand, summing over the set Γ of all the hyperplanes of Π we get

$$|S| = \sum_{H \in \Gamma} \sum_{P \in H} m(P) \geq \sum_{H \in \Gamma} (k-1) = (k-1)\theta_q(k-1), \tag{3.3}$$

where the inequality follows from the fact that $(\mathcal{M}, m)$ is in particular a $(k-1)$-fold blocking set in Π, by Proposition 3.7. Combining (3.2) and (3.3), we obtain

$$n \geq \left\lceil (k-1)\frac{\theta_q(k-1)}{\theta_q(k-2)} \right\rceil,$$

We then conclude observing that $\left\lceil (k-1)\frac{\theta_q(k-1)}{\theta_q(k-2)} \right\rceil = (k-1)q + \left\lceil \frac{k-1}{\theta_q(k-2)} \right\rceil = (k-1)q + 1.$ □

Notice that the result presented in Theorem 3.9 was independently and simultaneously shown also in [106, Theorem 1.4].

As a consequence, we get an asymptotic improvement of a result by Chabanne, Cohen and Patey [48]. In that work, they showed that the rate R of an $[n, Rn]_q$ minimal code for n large enough satisfies $R \leq \log_q(2)$, calling this bound the *Maximal bound*.

Corollary 3.10. If $\mathcal{C}$ is a minimal code of rate R, asymptotically it holds $R \leq \frac{1}{q}$.

Proof. Let $\mathcal{C}$ be a minimal code of rate R. Then, by Theorem 3.9

$$R = \frac{k}{n} \leq \frac{n+q-1}{qn} \longrightarrow \frac{1}{q},$$

as n goes to infinity. □

Remark 3.11. The previous bound is not tight in general. More precisely, in [5] we conjectured a new lower bound for the length n of an $[n, k]_q$ minimal code, which was then proved in [148].

Theorem 3.12 (see [5, 148])**.** Let $\mathcal{C}$ be an $[n, k]_q$ minimal code with $k \geq 2$. We have

$$n \geq (k-1)(q-1) + 1 + \sum_{i=1}^{k-1} \left\lceil \frac{(k-1)(q-1)+1}{q^i} \right\rceil.$$

In Section 3.5 we will further improve the bound in Theorem 3.12 using methods from algebraic combinatorics; see Theorem 3.51.

We now state an important result relating the minimum distance with the dimension of a minimal code and the size of the underlying field, which was originally proved by Cohen, Mesnager and Patey in [54]. In the next sections we will improve this result.

Theorem 3.13. Let $\mathcal{C}$ be an $[n,k,d]_q$ minimal code with $k \geq 2$. Then $d \geq k+q-2$.

Remark 3.14. The bound in Theorem 3.13 is not sharp. However, it can be used to get new bounds on the length of a minimal code, combining Theorem 3.13 with known upper bounds on the minimum distance. It is easy to observe that using the Singleton bound does not improve on Theorem 3.9. However, if q is small, we can get better results using the Griesmer bound [81].

Corollary 3.15. Let $\mathcal{C}$ be an $[n,k]_q$ minimal code. Then

$$n \geq \sum_{i=0}^{k-1} \left\lceil \frac{k+q-2}{q^i} \right\rceil .$$

Proof. It follows combining Theorem 3.13 with the Griesmer bound. □

Remark 3.16. Observe that for some sets of parameters Corollary 3.15 gives a better lower bound on the length of minimal codes than the one of Theorem 3.9, while for other sets of parameters the converse holds. For instance, it is easy to see that for $q=2$, Corollary 3.15 is always better. Viceversa, when $q \geq k \geq 4$, Theorem 3.9 provides better results.

Furthermore, numerical results with MAGMA show that the bound in Corollary 3.15 is not sharp. For example, for $q=2$ and $k=4$, the minimum possible length of a minimal code is 9, while the above bound gives 8.

3.3.1 Asymptotic Performance of Minimal Codes

We recall that there is an existence result that holds asymptotically, i.e. we can actually ensure the existence of minimal codes of arbitrary length n of a fixed rate R that only depends on q. This existence result is not constructive, and it was shown by Chabanne, Cohen and Patey [48].

Theorem 3.17 (Minimal Bound [48])**.** For any rate $R = k/n$ such that

$$0 \leq R \leq \frac{1}{2}\log_q\left(\frac{q^2}{q^2-q+1}\right),$$

there exists an infinite sequence of $[n,k]_q$ minimal codes.

The most important consequence of Theorem 3.13 is that it allows to show that minimal codes are asymptotically good. Let us recall that a family of codes is said **asymptotically good** if it contains a sequence $C = (\mathcal{C}_1, \mathcal{C}_2, \dots)$ of linear codes, where $\mathcal{C}_n$ is an $[n, k_n, d_n]_q$ code such that the rate R and the relative distance δ of $\mathcal{C}_n$, that is

$$R := \liminf_{n\to\infty} \frac{k_n}{n} \qquad \text{and} \qquad \delta := \liminf_{n\to\infty} \frac{d_n}{n},$$

are both positive.

In general, we would like ideally both rate and relative distance of a code to be as large as possible, since the rate measures the number of information coordinates with respect to the length of the code and the relative distance measures the error correction capability of the code. Determining the rate and the relative distance for a class of codes is in general a difficult task. For example, it is still unknown if the family of cyclic codes is asymptotically good. However, some families of asymptotically good codes are known to exist. For example, codes that meet the Asymptotic Gilbert-Varshamov bound, binary quasi-cyclic codes [51, 3], self-dual codes [4], group codes [44].

A direct consequence of Theorem 3.13 and of the Minimal Bound of Theorem 3.17 is the following result.

Theorem 3.18. Minimal codes are asymptotically good.

Remark 3.19. As already mentioned in Chapter 1, we are particularly interested in finding lower bounds on the length of minimal codes or, equivalently, lower bounds on the size of cutting blocking sets in projective spaces. From this point of view, it is not restrictive to only consider projective codes, which correspond to projective systems in which all the points have multiplicity one.

It immediately follows from the definitions that a cutting blocking set $\mathcal{M}$ in $\mathrm{PG}(N,q)$ is necessarily an N-fold blocking set. The following theorem is obtained by combining a well-known result of Beutelspacher (which gives a lower bound on the cardinality of an N-fold blocking set in $\mathrm{PG}(N,q)$ when $N \leq q$) and the correspondence between minimal codes and cutting blocking sets.

Theorem 3.20 (see [36, Theorem 2])**.** Let $\mathcal{C}$ be an $[n,k]_q$ minimal code. If $k-1 \leq q$, then $n \geq (q+1)(k-1)$.

The above results uses the fact that cutting blocking sets in $\mathrm{PG}(k-1,q)$ are in particular $(k-1)$-fold blocking sets. Beutelspacher also characterized $(k-1)$-fold blocking sets in $\mathrm{PG}(k-1,q)$ with cardinality $(q+1)(k-1)$, under the further assumption that $k \leq \sqrt{q}+2$. Recall that, when q is a square, a T-dimensional **Baer subspace** of $\mathrm{PG}(N,q)$ is a subgeometry isomorphic to $\mathrm{PG}(T,\sqrt{q})$.

Theorem 3.21 (see [36, Theorem 3])**.** Let $4 \leq k \leq \sqrt{q}+2$ and let $\mathcal{M}$ be a $(k-1)$-fold blocking set in $\mathrm{PG}(k-1,q)$. Then $|\mathcal{M}| \geq (q+1)(k-1)$. Moreover, equality holds if and only if one of the following scenarios occurs:

1. $\mathcal{M}$ is the set of points on $k-1$ mutually skew lines.

2. $k=\sqrt{q}+2$ and $\mathcal{M}$ is the point set of a 3-dimensional Baer subspace of $\mathrm{PG}(k-1,q)$.

3. $q = 4$, $k = 4$, and $\mathcal{M}$ is the complement of a hyperoval in a plane of $\mathrm{PG}(k-1, q)$, where an hyperoval is a set of $q + 2$ points in a plane, no three of which are collinear.

Lemma 3.22. In $PG(2, q)$ a set $\mathcal{M}$ is a cutting blocking set if and only if it is a 2-fold blocking set.

Proof. Clearly, a cutting blocking set is a 2-fold blocking set, as shown in Proposition 3.7. On the other hand, if $\mathcal{M}$ is a 2-fold blocking set, then for every line ℓ in $PG(2, q)$, $\langle \ell \cap \mathcal{M} \rangle = \ell$, since $|\ell \cap \mathcal{M}| \geq 2$. We conclude again by Proposition 3.7. □

Moreover, in $\mathrm{PG}(2, q)$ one can always construct a 2-fold blocking set of size $3q$, or equivalently a $[3q, 3]_q$ minimal code, by considering the union of three lines that do not intersect in the same point. When q is a square, one can construct a cutting blocking set as union of two disjoint Baer subplanes, producing a minimal code of length $2q + 2\sqrt{q} + 2$. We thus survey the known results on the cardinality of 2-fold blocking sets in $\mathrm{PG}(2, q)$, which turn out to be an accurate estimates also for the length of minimal codes of dimension 3.

Theorem 3.23 (see [23, Theorem 3.1])**.** Suppose that $\mathcal{M}$ be a 2-fold blocking set in $\mathrm{PG}(2, q)$. The following hold.

1. If $q < 9$, then $|\mathcal{M}| \geq 3q$.
2. If $q > 4$ is a square, then $|\mathcal{M}| \geq 2q + 2\sqrt{q} + 2$.
3. If $q > 19$, $q = p^{2d+1}$, then $|\mathcal{M}| \geq 2q + p^d \left\lceil \frac{(p^{d+1}+1)}{(p^d+1)} \right\rceil + 2$.
4. If $q = 11, 13, 17, 19$ is not a square, then $|\mathcal{M}| \geq \frac{(5q+7)}{2}$.

The bounds in Theorem 3.23, parts (3) and (4), are believed not to be sharp; see [23, page 133]. In particular, we are not aware of any construction of 2-fold blocking sets achieving these sizes.

3.4 First Constructions of Minimal Codes

In this section we provide a general construction of reduced minimal codes based on the geometric point of view. For this family of codes, we also determine the weight distribution, using basic combinatorial results in finite geometry.

Remark 3.24. In the literature, there is one *general* construction of small cutting blocking sets we are aware of, which we briefly sketch in this remark.

It was proposed by Fancsali and Sziklai in [68] and it works as follows. One chooses any $2k - 3$ distinct points on the rational normal curve in $\mathrm{PG}(k-1, q)$ and takes the union of the tangent lines at these points. The resulting set is a cutting blocking set, under the assumption that the characteristic of the field is at least k. We call it the **rational normal tangent set**. The corresponding codes are minimal $[(2k-3)(q+1), k]_q$ codes whose minimum distance was proved to be at least kq in [28]. The drawback of this construction is the constraint on both the size (q) and the characteristic (p) of the underlying field, reading $q \geq 2k-3$ and $p \geq k$. For a fixed value of q, the approach of [68] constructs cutting blocking sets in $\mathrm{PG}(k-1,q)$ for only a finite number of values of k.

In the following, we present a second construction that instead works for every choice of the parameters k and q which we provided in in [5] The same construction can be found also in [106, 27].

We start with two auxiliary lemmas, based on avoiding results in finite projective spaces.

Lemma 3.25. Let q be a prime power, k, r be integers such that $1 \leq r \leq k$. Let $P_1, \dots, P_r \in \mathrm{PG}(k-1,q)$ be points not on the same $(r-2)$-flat. Then, the number of hyperplanes H avoiding $P_1, \dots, P_r$ is $q^{k-r}(q-1)^{r-1}$.

Proof. It follows from a simple calculation using inclusion-exclusion principle. Since the number of hyperplanes is $\theta_q(k-1)$, and the number of hyperplanes containing at least i points among the P_j's is equal to $\theta_q(k-1-i)$, we get that the number of hyperplanes avoiding all the P_j's is

$$\begin{aligned}
&\theta_q(k-1) - \sum_{i=1}^{r}(-1)^{i-1}\binom{r}{i}\theta_q(k-1-i) \\
&= \frac{1}{q-1}\sum_{i=0}^{r}\binom{r}{i}(-1)^i(q^{k-i}-1) \\
&= \frac{1}{q-1}\left(q^{k-r}\sum_{i=0}^{r}\binom{r}{i}(-1)^i q^{r-i} - \sum_{i=0}^{r}\binom{r}{i}(-1)^i\right) \\
&= q^{k-r}(q-1)^{r-1}.
\end{aligned}$$

□

Lemma 3.26. Let $P_1, \dots, P_k \in \mathrm{PG}(k-1,q)$ be points in general position. Then, the number of hyperplanes containing $P_1, \dots, P_s$ and avoiding $P_{s+1}, \dots, P_k$ is $(q-1)^{k-s-1}$

Proof. Let $\Lambda := \langle P_1, \dots, P_s \rangle$, then the number of hyperplanes containing Λ and avoiding $P_{s+1}, \dots, P_k$ is in correspondence with the number of hyperplanes in $\mathrm{PG}(k-1)/\Lambda \cong PG(k-s,q)$ avoiding $P_{s+1}, \dots, P_k$. Such number is, by Lemma 3.25, equal to $(q-1)^{k-s-1}$. □

Theorem 3.27. Let $P_1, \dots, P_k$ be points in general position in $\mathrm{PG}(k-1,q)$. For $1 \leq i < j \leq k$, consider the line $\ell_{i,j} := \langle P_i, P_j \rangle$. Then, $\mathcal{M} := \bigcup_{i,j} \ell_{i,j}$ is a minimal cutting blocking set.

Proof. Let H be a hyperplane in $PG(k-1,q)$. Since the points $P_1, \dots, P_k$ are in general position, there exists at least one point among them, say P_1 that is not in H. Consider the intersection

$H \cap \mathcal{M}$, which does not contain P_1. Hence H meets the lines $\ell_{1,j}$'s in $k-1$ distinct points $Q_2, \dots Q_k$, i.e. $\{Q_j\} = H \cap \ell_{1,j}$ for $j \in \{2, \dots, k\}$. Take the flat $\Lambda := \langle \mathcal{M} \cap H \rangle$, and observe that

$$\langle \Lambda, P_1 \rangle \supseteq \langle P_1, Q_j \rangle = \ell_{1,j},$$

However, $P_j \in \ell_{1,j}$ for every $j \in \{2, \dots, k\}$, and this implies $\langle \Lambda, P_1 \rangle \supseteq \langle P_1, \dots, P_k \rangle = \mathrm{PG}(k-1, q)$. Hence, necessarily $\dim(\Lambda) = k-2$ and by Proposition 3.7, $\mathcal{M}$ is a cutting blocking set.

It is left to prove that $\mathcal{M}$ is minimal. Suppose we remove one of the points P_i's from $\mathcal{M}$, say P_1, getting $\tilde{\mathcal{M}} := \mathcal{M} \setminus \{P_1\}$. Take a $(k-3)$-flat $\Lambda \subseteq \langle P_2, \dots, P_k \rangle$ avoiding the points $P_2, \dots, P_k$. By Lemma 3.25 such a hyperplane always exists. Hence $H := \langle \Lambda, P_1 \rangle$ is a hyperplane such that $H \cap \tilde{\mathcal{M}} \subseteq \Lambda$, and by Proposition 3.7, $\tilde{M}$ is not minimal. Similarly, choose a point in $\mathcal{M} \setminus \{P_1, \dots, P_k\}$ and remove it from $\mathcal{M}$. Without loss of generality, we can choose $Q_{1,2} \in \ell_{1,2} \setminus \{P_1, P_2\}$ and consider $\tilde{M} := \mathcal{M} \setminus \{Q_{1,2}\}$. Take the space $H := \langle Q_{1,2}, P_3, \dots, P_k \rangle$. It is easy to see that $\langle H, P_1 \rangle = \langle H, P_2 \rangle = \mathrm{PG}(k-1, q)$, and hence H is a hyperplane. Moreover, $H \cap \mathcal{M} = \{Q_{1,2}, P_3, \dots, P_k\}$, therefore $\dim(H \cap \tilde{\mathcal{M}}) = \dim(\langle P_3, \dots, P_k \rangle) = k-3$, and by Proposition 3.7 $\tilde{\mathcal{M}}$ can not be a cutting blocking set. □

From this construction, called **tetrahedron**, one obtains a family of $[(q-1)\binom{k}{2} + k, k, (q-1)(k-1)+1]_q$ minimal codes. As a consequence of Theorem 3.23, when $k=3$ this construction provides a minimal 2-fold blocking set in $\mathrm{PG}(2, q)$ for any $q < 9$.

The next result analyzes the reduced minimal code obtained in Theorem 3.27, giving the full description of its weight distribution.

Theorem 3.28. The code associated to the minimal cutting blocking set of Theorem 3.27 is a $[\binom{k}{2}(q-1) + k, k]_q$ reduced minimal code $\mathcal{C}$, whose weights are exactly

$$f_{q,k}(r) := \frac{1}{2}(k-r)((k+r-1)q - 2k + 4),$$

for every $r \in \{0, \dots, k-1\}$. Furthermore, the weight distribution of $\mathcal{C}$ is given by

$$A_i(\mathcal{C}) = \sum_{\{r \mid f_{q,k}(r) = i\}} \binom{k}{r} (q-1)^{k-r}.$$

Proof. By the equivalence (Φ, Ψ) between codes and projective systems, the dimension of the code obtained by $\mathcal{M}$ is clearly k and its length is $n = \binom{k}{2}(q-1) + k$. Now, for a hyperplane $H = \langle v \rangle^{\perp}$, the weight of its $q-1$ associated codewords (i.e. all the nonzero multiples of vG, where G is the generator matrix obtained from $\mathcal{M}$) is $n - |\mathcal{M} \cap H|$. Therefore, it is determined by $|H \cap \mathcal{M}|$. By the symmetric properties of $\mathcal{M}$, the quantity $|H \cap \mathcal{M}|$ only depends on the integer

$$r := |\{i \in \{1, \dots, k\} \mid P_i \in H\}|.$$

In this case, without loss of generality we can assume that $P_1, \dots, P_r \in H$, and $P_{r+1}, \dots, P_{k-r} \notin H$. Hence, H contains all the lines $\ell_{i,j}$ for $0 \leq i < j \leq r$, and it intersects all the lines $\ell_{i,j}$ in $\{P_i\}$, for $0 \leq i \leq r < j \leq k$, and in $\{Q_{i,j}\}$ for $r+1 \leq i < j \leq k$. Moreover, observe that the points $Q_{i,j}$ are all pairwise distinct. Therefore, the weight of the codeword associated to H is equal to

$$\begin{aligned} f_{q,k}(r) &= \binom{k}{2}(q-1) + k - |\mathcal{M} \cap H| \\ &= \binom{k}{2}(q-1) + k - \Big| \bigcup_{1 \leq i \leq r < j \leq k} \ell_{i,j} \Big| - \Big| \bigcup_{r+1 \leq i < j \leq k} \{Q_{i,j}\} \Big| \\ &= \binom{k}{2}(q-1) + k - \binom{r}{2}(q-1) - r - \binom{k-r}{2} \\ &= \frac{1}{2}(k-r)((k+r-1)q - 2k + 4). \end{aligned}$$

The numbers $A_i(\mathcal{C})$ follow from Lemma 3.26, taking into account that for every hyperplane we need to count $q-1$ distinct codewords, which correspond to all the nonzero multiples. □

Example 3.29. We explain now in details the situation for $k = 3$. The construction of the minimal cutting blocking set of Theorem 3.28 corresponds to the union of three lines ℓ_1, ℓ_2, ℓ_3 in the projective plane $\mathrm{PG}(2, q)$ with trivial intersection, that is $\ell_1 \cap \ell_2 \cap \ell_3 = \emptyset$. We write $\{P_{i,j}\} = \ell_i \cap \ell_j$ for $1 \leq i < j \leq 3$. Here hyperplanes are lines and for any line ℓ there are three possibilities: it can coincide with one of the lines ℓ_i's, it can contain one of the $P_{i,j}$'s, or none of them. The three cases give weights $f_{q,3}(2) = 2q-1$, $f_{q,3}(1) = 3q-2$ and $f_{q,3}(0) = 3q-3$. This code for $q \geq 3$ is a three-weight code with weight distribution $A_0 = 1$, $A_{2q-1} = 3(q-1)$, $A_{3q-3} = (q-1)^3$ and $A_{3q-2} = 3(q-1)^2$, and for $q = 2$ it is a two-weight code with weight distribution $A_0 = 1$, $A_3 = 4$ and $A_5 = 3$.

Remark 3.30. The family of codes described in Theorem 3.27 has been constructed independently also in [106, Proposition 4.4.] and in [27, Section III]. However, in both these papers the authors did not study the reducedness, nor find the weight distributions. Moreover, in [27] the construction has been provided only for $q \geq k+2$. In particular, it is important to highlight the fact that the computation of the weight distribution of this family of codes gives a partial answer to an open problem in [27, Open Problem 1]. This suggests that the geometric point of view allows to analyze better the properties of minimal codes.

3.4.1 Minimal Codes of Dimension 4

Here we exhibit a special construction for minimal codes of dimension 4, using cutting blocking sets in $\mathrm{PG}(3, q)$ which have size smaller than the ones provided in Theorem 3.27.

Construction 1. Let $P_1, P_2, P_3, P_4 \in \mathrm{PG}(3,q)$ be points in general position. Up to change of coordinates, we can assume them to be the (representatives of the) standard basis vectors. Consider the lines $\ell_i = \langle P_i, P_{i+1} \rangle$ for $i \in \{1,2,3,4\}$ and the indices taken modulo 4. For the line $m_1 := \langle P_1, P_3 \rangle$, consider the sheaf of planes $\{H_\alpha \mid \alpha \in \mathbb{F}_q^*\}$ containing it, given by $H_\alpha := \{[x,y,z,\alpha y] \mid [x,y,z] \in \mathrm{PG}(2,q)\}$, where we have removed the planes $\langle \ell_1, \ell_2 \rangle$ and $\langle \ell_3, \ell_4 \rangle$. For the line $m_2 := \langle P_2, P_4 \rangle$, we do the same, and take the sheaf of planes $\{K_\alpha \mid \alpha \in \mathbb{F}_q^*\}$ containing it, given by $K_\alpha := \{[x,y,\alpha x, z] \mid [x,y,z] \in \mathrm{PG}(2,q)\}$, where we have removed the planes $\langle \ell_1, \ell_4 \rangle$ and $\langle \ell_2, \ell_3 \rangle$. Now, for every $\alpha \in \mathbb{F}_q^*$ compute $H_\alpha \cap K_\alpha = \{[x,y,\alpha x, \alpha y] \mid [x,y] \in \mathrm{PG}(1,q)\}$. We fix a $\beta \in \mathbb{F}_q^*$, and take the point

$$Q_{\beta,\alpha} := [1, \beta, \alpha, \beta\alpha].$$

Note that $Q_{\beta,\alpha} \in (H_\alpha \cap K_\alpha) \setminus (m_1 \cup m_2)$ for every $\alpha \in \mathbb{F}_q^*$. Moreover, the points $Q_{\beta,\alpha}$ are all on the line $\ell_\beta := \langle [1,\beta,0,0],[0,0,1,\beta] \rangle = \{[x, \beta x, y, \beta y] \mid [x,y] \in \mathrm{PG}(1,q)\}$.

With this notation we define $\mathcal{M}_\beta$ to be the set

$$\mathcal{M}_\beta := \ell_1 \cup \ell_2 \cup \ell_3 \cup \ell_4 \cup \{Q_{\beta,\alpha} \mid \alpha \in \mathbb{F}_q^*\}.$$

Theorem 3.31. The set $\mathcal{M}_\beta$ is a minimal cutting blocking set in $\mathrm{PG}(3,q)$, for every $\beta \in \mathbb{F}_q^*$.

Proof. Let H be a hyperplane of $\mathrm{PG}(3,q)$, that is a plane. We call $\mathcal{N}$ the union of the four lines. First, it is easy to see that if H contains a line ℓ_i, then $\langle H \cap \mathcal{M}_\beta \rangle$ is a hyperplane, since it contains at least another point not on ℓ_i. Suppose that H meets a line ℓ_i in only one point R_i distinct from P_i and P_{i+1}. Without loss of generality, we can assume $i = 1$. Hence $\langle \mathcal{M}_\beta \cap H \rangle \supseteq \langle \mathcal{N} \cap H \rangle =: \Lambda$. Now, observe that $\langle \Lambda, P_1 \rangle$ contains at least the line ℓ_1, a point on ℓ_2 distinct from P_2 and another point on ℓ_4 different from P_1. Hence

$$\langle \Lambda, P_1 \rangle \supseteq \langle \ell_1, \ell_2, \ell_4 \rangle \supseteq \langle P_1, P_2, P_3, P_4 \rangle = \mathrm{PG}(3,q),$$

which implies $\dim(\Lambda) = 2$. It remains to analyze the only case left, which is $\mathcal{N} \cap H = \{P_1, P_3\}$ (the case $\mathcal{N} \cap H = \{P_2, P_4\}$ is symmetric). In this case, necessarily $H = H_\alpha$, for some $\alpha \in \mathbb{F}_q^*$, and so $\langle H \cap \mathcal{M}_\beta \rangle = \langle P_1, P_3, Q_{\beta,\alpha} \rangle = H_\alpha = H$. This shows that $\mathcal{M}_\beta$ is a cutting blocking set.

It remains to prove that $\mathcal{M}_\beta$ is minimal. Clearly, we can not remove any of the points $Q_{\beta,\alpha}$'s, since $\mathcal{M}_\beta \setminus \{Q_{\beta,\alpha}\}$ meets H_α only in P_1 and P_3. The same happens if we remove one of the points P_i's. Indeed, $\mathcal{M}_\beta \setminus \{P_1\}$ meets H_α only in $\{P_3, Q_{\beta,\alpha}\}$, for every $\alpha \in \mathbb{F}_q^*$ (and symmetrically with $\mathcal{M}_\beta \setminus \{P_3\}$). The same happens with the hyperplanes K_α's if we remove P_2 or P_4. It is left to prove that if we remove a point R on one of the lines, say ℓ_1, the resulting set $\mathcal{M}_\beta \setminus \{R\}$ is not cutting. Take the point P_3 and consider the sheaf of planes containing the line $\langle P_3, R \rangle$. Every plane of this sheaf meets the line ℓ_4 in exactly one point. Hence, the sheaf is parametrized by the points on the line ℓ_4, and we can write it as $\{H_S \mid S \in \ell_4\}$, where clearly $H_S = \langle P_3, R, S \rangle$. Consider now the intersection between H_S and $\tilde{\mathcal{N}} := \mathcal{N} \setminus \{R\}$, i.e. the union of all the four lines without the point R. If $S = P_1$ then $H_{P_1} \cap \tilde{\mathcal{N}} = (\ell_1 \setminus \{R\}) \cup \{P_3\}$, which spans a hyperplane. It is not difficult to see that in all the remaining q cases it spans a line. However, every H_S meets the line ℓ_β in exactly a point. Hence it contains at most one of the $Q_{\beta,\alpha}$'s. However, we have q hyperplanes H_S and only $q-1$ points. Therefore, necessarily there exists $S \in \ell_4$ such that $H_S \cap \mathcal{M}_\beta = \{P_3, R, S\}$ and thus $\mathcal{M}_\beta \setminus \{R\}$ is not cutting.

□

Corollary 3.32. For every $\beta \in \mathbb{F}_q^*$, Construction 1 produces a $[5q-1, 4, 3q-2]_q$ reduced minimal code $\mathcal{C}_\beta$.

Proof. Using the characterization result of Theorem 3.8, clearly the code obtained by the minimal cutting blocking set M_β via (Φ, Ψ) is a $[5q-1, 4]_q$ reduced minimal code. It is left to determine the minimum distance of $\mathcal{C}_\beta$, which corresponds via (Φ, Ψ) to the value $(5q-1) - \max\{|H \cap \mathcal{M}_\beta| : \dim(H) = 2\}$. Any hyperplane H can contain at most two of the lines ℓ_i's and ℓ_β, since every three of them span the whole space $\mathrm{PG}(3, q)$. If it contains none of them, then $|\mathcal{M}_\beta \cap H| \leq 5$. If H contains only one of the ℓ_i's then $|\mathcal{M}_\beta \cap H| \leq q + 3$. In the case H contains only ℓ_β we also have $|\mathcal{M}_\beta \cap H| \leq q + 3$. Finally, the only case in which H contains a pair of lines is when $H = \langle \ell_i, \ell_{i+1} \rangle$, for $i \in \{1, 2, 3, 4\}$ (where the indices are taken modulo 4). In this case, we can see that H does not contain any of the points $Q_{\alpha,\beta}$, and therefore, $|H \cap M_\beta| = 2q + 1$. For every prime power q, the maximum among these values is given by $2q + 1$, and this concludes the proof. □

We conclude this subsection with explanatory examples.

Example 3.33. According to Corollary 3.32, Construction 1 for $q = 2$ and $\beta = 1$ gives rise to a minimal $[9, 4, 4]_2$ code, whose generator matrix is

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

It was proved by computer search with Magma [45] that 9 is the shortest length that a minimal code of dimension 4 can have over $\mathbb{F}_2$. Moreover, always with Magma we observed that this is the unique $[9, 4]_2$ minimal code up to equivalence.

Example 3.34. For $q = 3$ and $\beta = 2$, Construction 1 gives the $[14, 4, 7]_3$ reduced minimal code $\mathcal{C}_2$ whose generator matrix is

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

3.4.2 Minimal Codes of Dimension 5

Here we show another special construction for minimal codes of dimension 5, using cutting blocking sets in $\mathrm{PG}(4,q)$ whose size is smaller than the one provided in Theorem 3.27. When $q=2$, we provide also an alternative construction for minimal codes of dimension 5 as minimal blocking sets in $\mathrm{PG}(4,2)$.

Construction 2. Let P_1, P_2, P_3, P_4, P_5 be five points in general position in $\mathrm{PG}(4,q)$. Without loss of generality, we can assume that they are the (representatives of the) standard basis vectors. Consider the lines $\ell_i = \langle P_i, P_{i+1}\rangle$ for $i \in \{1,2,3,4,5\}$, where the indices are taken modulo 5. Consider now for $i \in \{1,2,3,4\}$ a point $Q_i \in \ell_i \setminus \{P_i, P_{i+1}\}$, and define the lines $m_1 := \langle Q_1, Q_3\rangle$, $m_2 := \langle Q_2, Q_4\rangle$ and $m_3 := \langle Q_1, Q_4\rangle$.

With this notation, we define the set $\mathcal{M} := \ell_1 \cup \ell_2 \cup \ell_3 \cup \ell_4 \cup \ell_5 \cup m_1 \cup m_2 \cup m_3$. We will refer to the above construction also as the *pentagonal construction*.

Theorem 3.35. The set $\mathcal{M}$ defined in Construction 2 is a cutting blocking set in $\mathrm{PG}(4,q)$.

Proof. We first write $\mathcal{N} = \ell_1\cup\ell_2\cup\ell_3\cup\ell_4\cup\ell_5$ and $\mathcal{N}' = m_1\cup m_2\cup m_3$. Let H be a hyperplane, define the spaces $\Lambda := \langle H\cap\mathcal{M}\rangle$ and $\Lambda_1 := \langle H\cap\mathcal{N}\rangle$ and consider the number r of the P_i's that are also in H. Clearly $r \in \{0,1,2,3,4\}$. If $r=4$ it is clear that $\langle H\cap\mathcal{N}\rangle = H$. If $r=0$, then it is easy to see that $\langle \Lambda_1, P_1\rangle$ contains $\mathcal{N}$, and hence it is the whole $\mathrm{PG}(4,q)$. Therefore $\dim(\Lambda_1) = \dim(\Lambda) = 3$. Also if $r=1$, that is $P_1 \in H$, then $\langle \Lambda_1, P_2\rangle$ turns out to be the whole space, hence $\dim(\Lambda)=3$. Now assume that $r=3$. Then we have two possibilities for the indices of these points. They can be consecutive (modulo 5), say P_1, P_2, P_3, in which case H contains their span plus a point on ℓ_4. Clearly this implies $\dim(\Lambda) = \dim(\Lambda_1) = 3$. The second case is when the indices are of the form $i, i+1, i+3$, i.e. $\langle \ell_i, P_{i+3}\rangle \subseteq H$. Then H intersects at least one line $m_j = \langle Q_t, Q_s\rangle$ skew to ℓ_i in another point R distinct from Q_1, Q_2, Q_3, Q_4. Consider then $\langle \Lambda, Q_s\rangle \supseteq \langle \ell_i, m_j, P_{i+3}\rangle = \mathrm{PG}(4,q)$. Hence also in this case $\dim(\Lambda)=3$. It remains to show the case $r=2$. If the indices of these two points are consecutive, then H contains a line ℓ_i and two more points, one on ℓ_{i+2} and one on ℓ_{i+3}. Clearly in this case $\langle \Lambda_1, P_{i+3}\rangle \supseteq \langle \ell_i, \ell_{i+2}, \ell_{i+3}\rangle = \mathrm{PG}(4,q)$, and we conclude also in this case. Suppose now that the two points in H are P_i and P_{i+2}. Then H will also intersect the line ℓ_{i+3} in a point R, and at least a line $m_j = \langle Q_t, Q_s\rangle$ in a point S, which is different from Q_t and Q_s. Then it is easy to see that also in this case $\langle \Lambda, Q_s\rangle = \mathrm{PG}(4,q)$, which finally shows that $\mathcal{M}$ is cutting. □

Corollary 3.36. Construction 2 produces a $[8q-3, 5, 4q-3]_q$ minimal code.

Proof. The fact that from Construction 2 we obtain a $[8q-3,5]_q$ minimal code, simply follows from the characterization result of Theorem 3.8. The minimum distance can be computed observing that a hyperplane H can contain at most 4 lines among the defining lines of $\mathcal{M}$, and this happens only in five cases: $H_1 = \langle \ell_1, \ell_2, m_2, m_3\rangle$, $H_2 = \langle \ell_3, \ell_4, m_1, m_3\rangle$, $H_3 = \langle \ell_1, \ell_2, \ell_3, m_1\rangle$, $H_4 = \langle \ell_2, \ell_3, \ell_4, m_2\rangle$ and $H_5 = \langle \ell_1, \ell_4, \ell_5, m_3\rangle$. In these cases we have $|\mathcal{M}\cap H_i| = 4q$, and the weights of the associated codewords are $8q-3-4q=4q-3$. In all the other cases, it is not difficult to see that any other hyperplane contains a smaller number of points of $\mathcal{M}$. Hence, the minimum distance of the code is $4q-3$. □

In the binary case, the pentagonal construction gives the $[13,5,5]_2$ reduced minimal code whose generator matrix is

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

By Magma computations, we can observe that 13 is the shortest length for a binary minimal code of dimension 5.

For $q=3$ the code obtained is $[21,5,9]_3$, but with the aid of Magma we found a $[20,5,9]_3$ minimal code. Hence, in general Construction 2 does not provide the smallest cutting blocking set in $\mathrm{PG}(4,q)$.

In this sequel, we provide a construction of minimal codes of dimension 5 over $\mathbb{F}_2$, using cutting blocking sets in $\mathrm{PG}(4,2)$, different from the pentagonal construction. We will refer to it as the *hexagonal construction*.

Construction 3. Let $\{P_1,P_2,P_3,P_4,P_5,P_6\}$ be a projective frame in $\mathrm{PG}(4,2)$. Without loss of generality, we can assume P_1,P_2,P_3,P_4,P_5 to be the (representatives of the) standard basis vectors and $P_6=[1,1,1,1,1]$. Consider the lines $\ell_i=\langle P_i,P_{i+1}\rangle$ for $i\in\{1,2,3,4,5,6\}$, where the indices are taken modulo 6. Let $Q:=[1,0,1,0,1]$.

The set $\mathcal{M}:=\ell_1\cup\ell_2\cup\ell_3\cup\ell_4\cup\ell_5\cup\ell_6\cup\{Q\}$ is a minimal cutting blocking set in $\mathrm{PG}(4,2)$. This is not difficult to verify by hand or computer search.

This construction produces the $[13,5,5]_2$ reduced minimal code generated by the following matrix:

$$G_2=\begin{pmatrix} 1&0&0&0&0&1&1&0&0&0&1&0&1\\ 0&1&0&0&0&1&1&1&0&0&1&1&0\\ 0&0&1&0&0&1&0&1&1&0&1&1&1\\ 0&0&0&1&0&1&0&0&1&1&1&1&0\\ 0&0&0&0&1&1&0&0&0&1&0&1&1 \end{pmatrix}$$

With the aid of Magma we observed that the code constructed in this way and the one obtained from the pentagonal construction are the only two $[13,5,5]_2$ minimal codes up to equivalence.

The hexagonal construction can be adapted to $q = 3$. It gives a $[20, 5, 9]_3$ minimal codes, which is the shortest code that we could obtain. Unfortunately, it seems difficult to generalize it for minimal codes of dimension 5 over $\mathbb{F}_q$, for $q > 3$.

3.5 Algebraic Combinatorial Approach

This section develops an algebraic combinatorial approach to study minimal codes. The method uses a generator matrix of a linear code to build a multivariate polynomial "machinery". This allows us to study the maximal codewords of a code by applying classical results on the number of roots of multivariate polynomials over finite grids. As an application of our method, with the aid of Alon's Combinatorial Nullstellensatz [16] and the Alon-Füredi Theorem [17], we improve known lower bounds on the minimum distance and the length of minimal codes.

It is interesting to observe that the results contained in this section are mainly exploiting the fact that in a minimal code all the codewords are *maximal*, as already observed in Remark 3.3. Although in a minimal code this code property is equivalent to all codewords being minimal, the focus on maximal codewords is crucial for deriving both the lower bound on the minimum distance (Theorem 3.44) and the lower bound on the length (Theorem 3.51) of minimal codes.

3.5.1 Combinatorial Nullstellensatz and Alon-Füredi Theorem

We start by surveying tools from algebraic combinatorics that will be applied repeatedly in the sequel. Among these are Alon's Combinatorial Nullstellensatz and the Alon-Füredi Theorem.

Notation 3.37. We state the results of this subsection and of the next one for an arbitrary field $\mathbb{F}$. In Subsection 3.5.3 we will resume focusing on the case $\mathbb{F} = \mathbb{F}_q$ and on linear codes.

For a multivariate polynomial $p \in \mathbb{F}[x_1, \ldots, x_k]$ and a subset $A \subseteq \mathbb{F}^k$, denote by $V_A(p)$ the set of zeros of p in A, and by $U_A(p)$ the nonzeros of p in A, i.e.,

$$\begin{aligned} V_A(p) &= \{v \in A \mid p(v) = 0\}\,, \\ U_A(p) &= \{u \in A \mid p(u) \neq 0\}\,. \end{aligned}$$

The Alon–Füredi Theorem [17, Theorem 5] gives a lower bound on the cardinality of $U_A(p)$ when A is a finite grid and p is not identically zero on A. Equivalently, it provides an upper bound on the number of zeros of p. We recall it for convenience of the reader.

Theorem 3.38 (Alon–Füredi Theorem [17])**.** Let $A = A_1 \times \ldots \times A_k \subseteq \mathbb{F}^k$ be a finite grid with $A_i \subseteq \mathbb{F}$ and $|A_i| = n_i$, where $n_1 \geq n_2 \geq \ldots \geq n_k \geq 2$. Let $p \in \mathbb{F}[x_1, \ldots, x_k]$ be a polynomial that is not identically 0 on A, and let $\bar{p}$ be the polynomial p modulo the ideal $(f_1(x_1), \ldots, f_k(x_k))$, where $f_i(x_i) = \prod_{a \in A_i}(x_i - a)$. Then

$$|U_A(p)| \geq (n_s - \ell) \prod_{i=1}^{s-1} n_i,$$

where ℓ and s are the unique integers satisfying $\deg \bar{p} = \sum_{i=s+1}^{k}(n_i - 1) + \ell$, with $1 \leq s \leq k$ and $1 \leq \ell \leq n_s - 1$.

The above theorem relies on the fact that the polynomial p is not identically zero on the finite grid we are interested in. However, when dealing with polynomials that are not explicitly given, this property is not always easy to verify. In this direction, the celebrated Alon's Combinatorial Nullstellensatz helps determining a sufficient condition for a polynomial to be nonzero on a finite grid. We state it here for completeness.

Theorem 3.39 (Combinatorial Nullstellensatz [16])**.** *Let $p \in \mathbb{F}[x_1, \ldots, x_k]$ and let $\deg p = \sum_{i=1}^{k} r_i$, for some $r_1, \ldots, r_k \in \mathbb{N}$. Suppose that the coefficient of the monomial $x_1^{r_1} x_2^{r_2} \cdots x_k^{r_k}$ in p is nonzero. Let $A := A_1 \times \ldots \times A_k \subseteq \mathbb{F}^k$ be a grid with $|A_i| \geq r_i + 1$ for all $i \in [k]$. Then $U_A(p) \neq \emptyset$.*

3.5.2 The Support Polynomials

We denote by $g^{(i)}$ the i-th column vector of a matrix $G \in \mathbb{F}^{k \times n}$. Moreover, we consider the vector $x = (x_1, \ldots, x_k)$ whose entries are algebraically independent variables over $\mathbb{F}$.

Definition 3.40. The **support polynomial** associated with a matrix $G \in \mathbb{F}^{k \times n}$ and a subset $I \subseteq [n]$ is

$$p_{G,I}(x) := \prod_{i \in I} x \cdot g^{(i)} \in \mathbb{F}[x_1, \ldots, x_k].$$

In our approach, support polynomials are crucial for the study of minimal codes (taking as G a generator matrix of an $[n, k, d]_q$ code and as I a subset of a codeword's support). However, for the moment we focus on general properties of support polynomials that do not necessarily arise from codes. The following result is straightforward and its proof is omitted.

Proposition 3.41. *Let $G \in \mathbb{F}^{k \times n}$ and $I \subseteq [n]$.*

1. *For every $A \in \mathrm{GL}(k, \mathbb{F})$*

$$p_{AG,I}(x) = p_{G,I}(xA) = (p_{G,I} \circ L_A)(x),$$

where L_A denotes the linear map associated to the matrix A, that is $v \longmapsto vA$.

2. *For every $\tau \in \mathcal{S}_n$*

$$p_{G,\tau(I)}(x) = p_{GP_\tau,I}(x),$$

where P_τ is the permutation matrix associated to τ, such that

$$(v_1,\dots,v_n)P_\tau = \left(v_{\tau(1)},\dots,v_{\tau(n)}\right).$$

3. For every $v \in \mathbb{F}^n$
$$p_{GD_v,I}(x) = \Big(\prod_{i\in I} v_i\Big) p_{G,I}(x),$$
where D_v denotes the diagonal matrix whose diagonal is v.

We now study a support polynomial in connection with the rowspace of the matrix $G \in \mathbb{F}^{k\times n}$ that defines it. We first show how the zeros and nonzeros of support polynomials are related when we choose matrices with the same rowspace. Let $G_1, G_2 \in \mathbb{F}^{k\times n}$ be two matrices such that $\mathrm{rowsp}(G_1) = \mathrm{rowsp}(G_2)$. It is easy to see that there exists $A \in \mathrm{GL}(k,\mathbb{F})$ such that

$$U_{\mathbb{F}_q^k}(p_{G_1,I}) = U_{\mathbb{F}_q^k}(p_{G_2,I}) \cdot A := \left\{uA \mid u \in U_{\mathbb{F}_q^k}(p_{G_1,I})\right\},$$
$$V_{\mathbb{F}_q^k}(p_{G_1,I}) = V_{\mathbb{F}_q^k}(p_{G_2,I}) \cdot A := \left\{vA \mid v \in V_{\mathbb{F}_q^k}(p_{G_1,I})\right\}.$$

Indeed, any matrix A with $G_2 = AG_1$ satisfies the desired properties. Moreover, the nonzeros of a support polynomial are closely related to the support of vectors belonging to the rowspace of the defining matrix. This is shown by the following simple result, whose proof is omitted.

Lemma 3.42. Let $G \in \mathbb{F}^{k\times n}$ be a matrix. For all $I \subseteq [n]$ we have

$$U_{\mathbb{F}^k}(p_{G,I}) = \left\{u \in \mathbb{F}^k \mid \sigma^{\mathrm{H}}(uG) \supseteq I\right\}.$$

In particular, $U_{\mathbb{F}^k}(p_{G,I}) \neq \emptyset$ if and only if there exists $c \in \mathrm{rowsp}(G)$ with $\sigma^{\mathrm{H}}(c) \supseteq I$.

3.5.3 Minimum Distance of Minimal Codes

In this subsection we investigate the support polynomials of generator matrices of linear codes and their sets of zeros. As a corollary of our results, we establish a conjecture from [5].

We start with the following lemma, whose proof directly follows from Lemma 3.42 and Remark 3.3.

Lemma 3.43. Let $G \in \mathbb{F}_q^{k\times n}$ be a generator matrix of an $[n,k]_q$ code $\mathcal{C}$. Let $c = uG$ be a maximal codeword of $\mathcal{C}$ and $I := \sigma^{\mathrm{H}}(c)$. Then

$$U_{\mathbb{F}_q^k}(p_{G,I}) = \{\lambda u \mid \lambda \in \mathbb{F}_q^*\}.$$

In particular, if $\mathcal{C}$ is a minimal code then the above statement holds for every nonzero codeword.

Theorem 3.44. Let $\mathcal{C}$ be an $[n,k,d]_q$ code, and let c be a maximal codeword. Then $\mathrm{wt}^{\mathrm{H}}(c) \geq (q-1)(k-1)+1$. In particular, if $\mathcal{C}$ is minimal then $d \geq (q-1)(k-1)+1$.

Proof. Let $c = (c_1, \dots, c_n) \in \mathcal{C}$ be a maximal codeword of weight w and let $I := \sigma^{\mathrm{H}}(c)$, i.e., $c_i \in \mathbb{F}_q^*$ if and only if $i \in I$. Take a generator matrix $G \in \mathbb{F}_q^{k \times n}$ for $\mathcal{C}$ and consider the polynomial $p_{G,I}(x) \in \mathbb{F}_q[x_1, \dots, x_k]$. Observe that $p_{G,I}$ does not vanish identically on $\mathbb{F}_q^k$. Indeed, let $u \in \mathbb{F}_q^k$ be the vector such that $uG = c$. Then $p_{G,I}(u) = \prod_{i \in I} c_i \neq 0$. This also ensures that $\deg p_{G,I} = w$. Since c is a maximal codeword, by Lemma 3.43 we have $U_{\mathbb{F}_q^k}(p_{G,I}) = \{\lambda u \mid \lambda \in \mathbb{F}_q^*\}$, which has cardinality $q-1$.

On the other hand, let $\bar{p}_{G,I}$ denote the reduction of the polynomial $p_{G,I}$ modulo the ideal $(\{x_i^q - x_i \mid i \in [k]\})$. By Theorem 3.38 we have

$$|U_{\mathbb{F}_q^k}(p_{G,I})| \geq (q-\ell)q^{s-1},$$

where ℓ and s are the unique integers satisfying $\deg \bar{p}_{G,I} = (q-1)(k-s)+\ell$, with $1 \leq s \leq k$ and $1 \leq \ell \leq q-1$.

Thus, combining this with the exact value of $|U_{\mathbb{F}_q^k}(p_{G,I})|$, we obtain $q-1 = |U_{\mathbb{F}_q^k}(p_{G,I})| \geq (q-\ell)q^{s-1}$, from which we deduce $s = 1$. Therefore,

$$w = \deg p_{G,I} \geq \deg \bar{p}_{G,I} = (q-1)(k-1)+\ell \geq (q-1)(k-1)+1,$$

as desired. □

Remark 3.45. The Alon-Füredi Theorem (Theorem 3.38) gives a lower bound on the number of nonzeros of a multivariate polynomial in a finite grid in terms of the degree of the polynomial and the size of the grid. This result has been used in coding theory for deriving the minimum distance of generalized Reed-Muller codes in [38], although similar arguments were proposed in [77, 105]. It is interesting to observe that in our Theorem 3.44 the Alon-Füredi Theorem is applied in the "opposite" direction, i.e., we use it to derive a lower bound on the degree of the support polynomial associated to a maximal codeword, knowing the number of its nonzeros.

3.5.4 Maximal Codewords in Linear Codes

In this subsection we use support polynomials to study the structure of maximal codewords in a linear code $\mathcal{C}$. In particular, we show that for any maximal codeword $c \in \mathcal{C}$ there exist several codewords whose support contains a large subset of the support of c. This property will be crucial for deriving a lower bound on the length of minimal codes in Subsection 3.5.5.

For $v = (v_1, \dots, v_k) \in \mathbb{F}_q^k$, define

$$f_v(x) := \prod_{i=1}^{k} \Big(\prod_{s \in \mathbb{F}_q \setminus \{v_i\}} (x_i - s) \Big).$$

Next, consider the ideal $I_q := (x_1^q - x_1, \dots, x_k^q - x_k)$ and denote by $\bar{f}$ the reduction of a polynomial

$f \in \mathbb{F}_q[x_1, \dots, x_k]$ modulo I_q. It is easy to check that for every $v \in \mathbb{F}_q^k$ we have $\bar{f}_v = f_v$. One can also easily prove that the set $\{f_v \mid v \in \mathbb{F}_q^k\}$ is an $\mathbb{F}_q$-basis for the space $\mathbb{F}_q[x_1, \dots, x_k]/I_q$. Moreover, regarding the polynomials f_v's as maps from $\mathbb{F}_q^k$ to $\mathbb{F}_q$, the set $\{f_v \mid v \in \mathbb{F}_q^k\}$ is an $\mathbb{F}_q$-basis of $\{\varphi : \mathbb{F}_q^k \longrightarrow \mathbb{F}_q\}$. This is due to the following well-known result; see for example [123, Section 5.4.1].

Proposition 3.46. The evaluation map on $\mathbb{F}_q[x_1, \dots, x_k]$ induces the isomorphism of $\mathbb{F}_q$-vector spaces

$$\mathbb{F}_q[x_1, \dots, x_k]/I_q \cong \{\varphi : \mathbb{F}_q^k \longrightarrow \mathbb{F}_q\}. \tag{3.4}$$

In particular, for every $p \in \mathbb{F}_q[x_1, \dots, x_k]$ there exist unique $\mu_v \in \mathbb{F}_q$ for $v \in U_{\mathbb{F}_q^k}(p)$, such that

$$\bar{p} = \sum_{v \in U_{\mathbb{F}_q^k}(p)} \mu_v f_v.$$

Proposition 3.47. Let $\mathcal{C}$ be an $[n, k]_q$ code and let $c = (c_1, \dots, c_n) \in \mathcal{C}$ be a maximal codeword of $\mathcal{C}$ with weight w and support $I := \sigma^{\mathrm{H}}(c)$. Let w_1 be the unique integer in $[q-1]$ such that $w_1 \equiv w \mod (q-1)$. Then, for any $A \in \mathrm{GL}(k, q)$ such that the first row of $A^{-1}G$ is equal to c, we have $\bar{p}_{G,I}(x) = p_c(xA)$, where

$$p_c(x) = \Big(\prod_{i=1}^{w} c_i\Big) x_1^{w_1} \prod_{i=2}^{k} (1 - x_i^{q-1}).$$

Proof. We first prove the statement in the case where the first row of G is equal to c. Observe that $\bar{p}_c = p_c$, that is, the polynomial p_c is already reduced modulo I_q. Therefore, by the isomorphism given in (3.4), we only need to show that $p_{G,I}(v) = p_c(v)$ for every $v \in \mathbb{F}_q^k$. By definition of p_c we have

$$p_c(v) = \begin{cases} \lambda^{w_1} \prod_{i=1}^{w} c_i & \text{if } v = \lambda e_1, \\ 0 & \text{otherwise.} \end{cases}$$

On the other hand, by the choice of G and Lemma 3.43 we have

$$p_{G,I}(\lambda e_1) = \prod_{i=1}^{w} (\lambda c_i) = \lambda^{w_1} \prod_{i=1}^{w} c_i,$$

where the last inequality follows using the identity $\lambda^q = \lambda$. Moreover, $p_{G,I}(v) = 0$ for every $v \notin \{\lambda e_1 \mid \lambda \in \mathbb{F}_q^*\}$.

The general case follows from the previous one. We first transform G into $A^{-1}G$, where the first row of $A^{-1}G$ is equal to c. This implies that $p_{A^{-1}G,I}(x) = p_c(x)$. Then, using Proposition 3.41, we find $p_{G,I}(x) = p_{A^{-1}G,I}(xA) = p_c(xA)$. □

Notation 3.48. In the remainder of the section we write $x = (x_1, \dots, x_k)$ and for $\alpha =$

$(\alpha_1, \dots, \alpha_k) \in \mathbb{N}^k$ we denote by x^α the monomial $x_1^{\alpha_1} \cdots x_k^{\alpha_k}$. Moreover, we let $\|\alpha\| := \alpha_1 + \ldots + \alpha_k$. Finally, for a polynomial $p(x) \in \mathbb{F}_q[x_1, \dots, x_k]$ and a monomial x^α, we denote by $[x^\alpha]p(x)$ the coefficient of the monomial $x^\alpha = x_1^{a_1} \cdots x_k^{a_k}$ in $p(x)$.

The following result on maximal codewords will be crucial in the next subsection for deriving a lower bound on the length of a minimal code.

Theorem 3.49. Let $\mathcal{C}$ be an $[n,k]_q$ code and let $c = (c_1, \dots, c_n) \in \mathcal{C}$ be a maximal codeword. For every $j \in \sigma^{\mathrm{H}}(c)$ there exist $I_j \subseteq \sigma^{\mathrm{H}}(c) \setminus \{j\}$ of cardinality $(q-1)(k-1)$ and a codeword $z \in \mathcal{C}$ such that $\sigma^{\mathrm{H}}(z) \cap \sigma^{\mathrm{H}}(c) \supseteq I_j$.

Proof. Let $c \in \mathcal{C}$ be a nonzero codeword with support $I := \sigma^{\mathrm{H}}(c)$ and weight $w = (q-1)(k-1) + w_1$. By Theorem 3.44 we have $w_1 \geq 1$.

Assume first that $w \leq (q-1)k$, which implies $1 \leq w_1 \leq q-1$. We choose a generator matrix G for $\mathcal{C}$ whose first row is equal to c, and assume that $c_i = 1$ for every $i \in I$. This can be done without loss of generality, up to replacing the code with an equivalent one. By Proposition 3.47 we have

$$\bar{p}_{G,I}(x) = p_c(x) = x_1^{w_1} \prod_{i=2}^{k} (1 - x_i^{q-1}).$$

Let $j \in I$ and assume that the j-th column of G is $(1, 0, \dots, 0)^\top$. Define $\mathcal{L}_{I,j} := \{L \subseteq I : j \notin L, |L| = (q-1)(k-1)\}$ and $\beta := (w_1, q-1, q-1, \dots, q-1)$. We have

$$\begin{aligned} (-1)^{k-1} &= [x^\beta]\bar{p}_{G,I}(x) \\ &= [x^\beta]p_{G,I}(x) \\ &= \sum_{L \in \mathcal{L}_{I,j}} [x_2^{q-1} \cdots x_k^{q-1}] p_{G,L}(x). \end{aligned}$$

The first equality follows from direct inspection of $p_c(x)$. The second equality is due to the fact that the degree of $p_{G,I}$ is equal to $(q-1)(k-1) + w_1$, which is also the degree of $\bar{p}_{G,I}$. The third equality follows from the fact that the coefficients of x_1 in the matrix G are all equal to 1. Therefore, there exists $I_j \in \mathcal{L}_{I,j}$ such that $[x_2^{q-1} \cdots x_k^{q-1}] p_{G,I_j}(x) \neq 0$. Let $x' := (x_2, \dots, x_k)$ and consider the polynomial $f(x') := p_{G,I_j}(0, x_2, \dots, x_k)$. This polynomial has degree $|I_j| = (q-1)(k-1)$ and $[x_2^{q-1} \cdots x_k^{q-1}] f(x') = [x_2^{q-1} \cdots x_k^{q-1}] p_{G,I_j}(x) \neq 0$. Hence, by Theorem 3.39, there exists $v \in \mathbb{F}_q^{k-1}$ such that $f(v) = p_{G,I_j}(0, v) \neq 0$. By Lemma 3.42, this implies that the codeword $z := (0, v)G \in \mathcal{C}$ satisfies $\sigma^{\mathrm{H}}(z) \supseteq I_j$.

Now assume that $w > (q-1)k$, from which $w_1 \geq q$. Let us write $w_1 = a(q-1) + b$ with $1 \leq b \leq q-1$. Since $w > (q-1)k$, we have $a \geq 1$. Denote the vector $\beta := (b, q-1, q-1, \dots, q-1)$.

Consider the set $T := \{\alpha \in \mathbb{N}^k : \|\alpha\| = a(q-1), \alpha_i \equiv 0 \mod (q-1) \text{ for every } i \in [k]\}$. Then

$$\begin{aligned}(-1)^{k-1} &= [x^\beta]\bar{p}_{G,I}(x) \\ &= \sum_{\alpha \in T} [x^{\beta+\alpha}]p_{G,I}(x).\end{aligned}$$

This means that there exists $\gamma \in T$ such that $[x^{\beta+\gamma}]p_{G,I}(x) \neq 0$. Define $\mathcal{L}_{I,j}^{(\gamma)} := \{L \subseteq I : |L| = w - b - \gamma_1, j \notin L\}$, $\gamma' := (0, \gamma_2, \dots, \gamma_k)$, and $\beta' := (0, q-1, \dots, q-1)$. We have

$$[x^{\beta+\gamma}]p_{G,I}(x) = \sum_{L \in \mathcal{L}_{I,j}^{(\gamma)}} [x^{\beta'+\gamma'}]p_{G,L}(x)$$

and there exists $K \in \mathcal{L}_{I,j}^{(\gamma)}$ such that $[x^{\beta'+\gamma'}]p_{G,K}(x) \neq 0$. At this point we can consider the set $\mathcal{X}_K := \{M \subseteq K : |M| = (q-1)(k-1)\}$ and write

$$[x^{\beta'+\gamma'}]p_{G,K}(x) = \sum_{M \in \mathcal{X}_K} \lambda_M \big([x^{\beta'}]p_{G,M}(x)\big)$$

for some $\lambda_M \in \mathbb{F}_q$. Since this sum is nonzero, there exists M such that $[x^{\beta'}]p_{G,M}(x) \neq 0$. As in the previous case, we use Theorem 3.39 and Lemma 3.42 to deduce that there exists a codeword $z \in \mathcal{C}$ such that $\sigma^{\mathrm{H}}(z) \supseteq M$. □

The following follows as a special case of Theorem 3.49.

Corollary 3.50. Let $\mathcal{C}$ be an $[n,k,d]_q$ minimal code with $d = (q-1)(k-1)+1$, and let $c \in \mathcal{C}$ be a codeword of minimum weight d. Then, for every $j \in \sigma^{\mathrm{H}}(c)$, there exists a codeword $z \in \mathcal{C}$ such that $\sigma^{\mathrm{H}}(z) \cap \sigma^{\mathrm{H}}(c) = \sigma^{\mathrm{H}}(c) \setminus \{j\}$.

3.5.5 The Length of Minimal Codes

As an application of Theorem 3.49, we derive the following lower bound on the length of a minimal code.

Theorem 3.51. Let $\mathcal{C}$ be an $[n,k,d]_q$ minimal code. We have $n \geq (q+1)(k-1)$.

Proof. Let $c \in \mathcal{C}$ be a codeword of minimum weight d with support $I := \sigma^{\mathrm{H}}(c)$. Up to considering an equivalent code, we can assume $c_i = 1$ for every $i \in I$. Since c is in particular a maximal codeword of $\mathcal{C}$, by Theorem 3.49 there exists a codeword $z \in \mathcal{C}$ such that $|I \cap \sigma^{\mathrm{H}}(z)| \geq (q-1)(k-1)$. Let $J := I \cap \sigma^{\mathrm{H}}(z)$ and for every $\lambda \in \mathbb{F}_q^*$ define $J_\lambda := \{j \in J \mid z_j = \lambda\}$. Clearly, $J = \bigcup_{\lambda \in \mathbb{F}_q^*} J_\lambda$ and the union is disjoint. Thus by the generalized pigeonhole principle there exists $\lambda' \in \mathbb{F}_q^*$ such that

$$|J_{\lambda'}| \geq \left\lceil \frac{|J|}{q-1} \right\rceil \geq k-1.$$

Now consider the codeword $z - \lambda' c$. Its support is $\sigma^{\mathrm{H}}(z - \lambda' c) = (\sigma^{\mathrm{H}}(c) \cup \sigma^{\mathrm{H}}(z)) \setminus J_{\lambda'}$ and

$$\begin{aligned} \mathrm{wt}^{\mathrm{H}}(z - \lambda' c) &= |\sigma^{\mathrm{H}}(c)| + |\sigma^{\mathrm{H}}(z)| - |J| - |J_{\lambda'}| \\ &\leq d + \mathrm{wt}^{\mathrm{H}}(z) - q(k-1). \end{aligned}$$

Combining this with $\mathrm{wt}^{\mathrm{H}}(z - \lambda' c) \geq d$ we get $\mathrm{wt}^{\mathrm{H}}(z) \geq q(k-1)$. Furthermore, by Proposition 3.4 we have $\mathrm{wt}^{\mathrm{H}}(z) \leq n - k + 1$, from which $n \geq (q+1)(k-1)$. □

Remark 3.52. The lower bound of Theorem 3.51 is an improvement on the bound in Theorem 3.12. Indeed, we have

$$(q+1)(k-1) \geq \sum_{i=0}^{k-1} \left\lceil \frac{(q-1)(k-1)+1}{q^i} \right\rceil . \tag{3.5}$$

To see this, we first observe that proving (3.5) is equivalent to proving that $\sum_{i=1}^{k-1} \lceil t/q^i \rceil \leq 2k-3$, where $t := (q-1)(k-1)+1$. Let $r \in \mathbb{N}_{>0}$ be such that $q^r \leq t < q^{r+1}$. We write the integer t in its q-adic expansion, i.e.,

$$t = \sum_{i=0}^{r} a_i q^i$$

for some $a_i \in \{0, 1, \ldots, q-1\}$ and for all $i \in \{0, \ldots, r\}$. We distinguish two cases.
Case I: $t = q^r$. We have

$$\begin{aligned} \sum_{i=1}^{k-1} \left\lceil \frac{t}{q^i} \right\rceil = \sum_{i=1}^{r} \left\lceil \frac{q^r}{q^i} \right\rceil + \sum_{i=r+1}^{k-1} \left\lceil \frac{q^r}{q^i} \right\rceil &= \sum_{i=1}^{r} q^{r-i} + (k-r-1) \\ = \frac{t-1}{q-1} + (k-r-1) = 2k-2-r \leq 2k-3&, \end{aligned}$$

which proves the statement.
Case II: $t > q^r$. Write

$$\begin{aligned} \sum_{i=1}^{k-1} \left\lceil \frac{t}{q^i} \right\rceil &= \sum_{i=1}^{r} \left\lceil \frac{a_r q^r + \cdots + a_0}{q^i} \right\rceil + \sum_{i=r+1}^{k-1} \left\lceil \frac{t}{q^i} \right\rceil \\ &= \sum_{i=1}^{r} \left\lceil \frac{a_r q^r + \cdots + a_0}{q^i} \right\rceil + k - r - 1. \end{aligned} \tag{3.6}$$

Observe that, for each $i \in [r]$,

$$\begin{aligned}\left\lceil \frac{a_r q^r + \cdots + a_0}{q^i} \right\rceil &= \frac{a_r q^r + \cdots + a_i q^i}{q^i} + \left\lceil \frac{a_{i-1} q^{i-1} + \cdots + a_0}{q^i} \right\rceil \\ &\leq \sum_{j=i}^{r} a_j q^{r-j} + 1. \end{aligned} \tag{3.7}$$

Combining (3.6) with (3.7) we obtain

$$\begin{aligned}\sum_{i=1}^{k-1} \left\lceil \frac{t}{q^i} \right\rceil &\leq \sum_{i=1}^{r} \left(1 + \sum_{j=i}^{r} a_j q^{j-i} \right) + (k-r-1) \\ &= \sum_{j=1}^{r} a_j \left(\sum_{i=1}^{j-1} q^i \right) + (k-1) \\ &= \frac{1}{(q-1)} \left(t - \sum_{j=0}^{r} a_j \right) + (k-1) \\ &\overset{(*)}{<} \frac{1}{(q-1)}(t-1) + k - 1 \\ &= 2k-2, \end{aligned}$$

where $(*)$ follows from the fact that $q^r < t < q^{r+1}$.

Remark 3.53. Observe that Theorem 3.51 is an improvement on the bound of Theorem 3.20, since it does not require the extra assumption that $k \leq q+1$.

We conclude this section with a detailed example building on [5, Example 5.11].

Example 3.54. We fix $q = 3$, $k = 4$ and take the minimal $[14, 4, 7]_3$ code $\mathcal{C}$ whose generator matrix is

$$G := \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 & 1 & 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 & 1 & 1 \\ 0 & 1 & 2 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 0 & 1 & 2 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 2 \end{pmatrix}.$$

Let $I := \{8, 9, \ldots, 14\}$ be the support of the codeword c given by the first row of G and let $x = (x_1, x_2, x_3, x_4)$. We compute the associated support polynomial

$$p_{G,I}(x) = x_1(x_3^2 - x_1^2)(x_1^2 - x_4^2)((x_4 - x_2)^2 - (x_3 - x_1)^2).$$

An easy calculation shows that the reduction of $p_{G,I}(x)$ modulo $I_3 = (x_1^3 - x_1, x_2^3 - x_2, x_3^3 - x_3, x_4^3 - x_4)$ is

$$\bar{p}_{G,I}(x) = x_1(1 - x_2^2)(1 - x_3^2)(1 - x_4^2) = p_c(x),$$

as we can also deduce from Proposition 3.47.

Moreover, since $\mathcal{C}$ is a minimal code, we can actually see that for every $j \in I$ there exists a codeword $z^{(j)} \in \mathcal{C}$ such that $\sigma^{\mathrm{H}}(z^{(j)}) \supseteq I \setminus \{j\}$, as stated in Theorem 3.49. These 7 codewords (up to their nonzero scalar multiples) are

$$\begin{aligned} z^{(8)} &= (0,0,0,0,2,1,2,0,1,1,1,1,1,2), \\ z^{(9)} &= (0,1,2,1,2,0,1,2,0,1,1,1,2,2), \\ z^{(10)} &= (1,2,0,1,2,0,1,1,1,0,1,2,1,2), \\ z^{(11)} &= (2,1,0,2,2,2,0,1,2,1,0,2,2,2), \\ z^{(12)} &= (1,2,0,1,1,1,0,1,2,1,2,0,2,1), \\ z^{(13)} &= (0,2,1,2,2,2,0,1,2,1,1,1,0,2), \\ z^{(14)} &= (0,1,2,1,1,1,0,1,2,1,1,1,1,0). \end{aligned}$$

Finally, note that, in order to derive the lower bound on the length of minimal codes given in Theorem 3.51, we use in its proof that each of the codewords $z^{(j)}$ has weight at least $q(k-1) = 9$. However, in this case only $z^{(8)}$ has weight 9, while all the other codewords have weight 11.

Remark 3.55. It is natural to ask whether the bound of Theorem 3.51 is sharp or not. As showed in Theorem 3.8, minimal codes of dimension k over $\mathbb{F}_q$ correspond to cutting blocking sets in $\mathrm{PG}(k-1,q)$ and a cutting blocking set is in particular a $(k-1)$-fold blocking set. When we restrict to the case $4 \leq k \leq \sqrt{q}+2$, Theorem 3.21 characterizes a $(k-1)$-fold blocking set $\mathcal{M}$ in $\mathrm{PG}(k-1,q)$ of cardinality $(q+1)(k-1)$. This only happens in three cases.
Case I: $\mathcal{M}$ is the union of $k-1$ disjoint lines. In this case $\mathcal{M}$ cannot be a cutting blocking set. To see this, write $\mathcal{M} = \ell_1 \cup \ldots \cup \ell_{k-1}$. Pick $P_1 \in \ell_1, \ldots, P_{k-1} \in \ell_{k-1}$ and let $\Lambda := \langle P_1, \ldots, P_{k-1} \rangle$. If $\dim(\Lambda) \leq k-3$, then Λ is contained in a $(k-3)$-flat Λ'. Consider the sheaf of hyperplanes containing Λ'. They are $q+1$ and only $k-1$ of them contain other points of $\mathcal{M}$ in addition to $P_1, \ldots, P_{k-1}$. Since $k-1 < q+1$ there is at least one hyperplane H such that $H \cap \mathcal{M} = \{P_1, \ldots, P_{k-1}\}$ and $\langle H \cap \mathcal{M} \rangle \subseteq \Lambda' \neq H$. This implies that, in this case, $\mathcal{M}$ is not a cutting blocking set. Suppose then that $\dim(\Lambda) = k-2$. Fix $P_1, \ldots, P_{k-3}$ and consider the flat $\Gamma := \langle P_1, \ldots, P_{k-3}, \ell_{k-2} \rangle$. If $\dim(\Gamma) < k-2$, then there exists $Q_{k-2} \in \ell_{k-2} \cap \langle P_1, \ldots, P_{k-3} \rangle$. Thus, if we replace P_{k-2} by Q_{k-2}, we get that $\dim(\Lambda) < k-2$, and we can conclude as done before that $\mathcal{M}$ is not cutting. Hence, assume $\dim(\Gamma) = k-2$. In this case $\Gamma \cap \ell_{k-1} \neq \emptyset$. Take $Q_{k-1} \in \Gamma \cap \ell_{k-1}$. If $Q_{k-1} \in \langle P_1, \ldots, P_{k-3} \rangle$, we substitute P_{k-1} with Q_{k-1} and get again $\dim(\Lambda) < k-2$, which implies $\mathcal{M}$ not being cutting. Therefore, assume that the space $\langle P_1, \ldots, P_{k-3}, Q_{k-1} \rangle$ is a hyperplane in Γ. Since Γ also contains ℓ_{k-2}, there exists $R_{k-2} \in \ell_{k-2} \cap \langle P_1, \ldots, P_{k-3}, Q_{k-1} \rangle$. Thus, replacing P_{k-1} with Q_{k-1} and P_{k-2} with R_{k-2}, we again obtain that $\dim(\Lambda) < k-2$ and $\mathcal{M}$ is not cutting.

Case II: $k = \sqrt{q}+2$ and $\mathcal{M}$ is a 3-dimensional Baer subspace. If $k \geq 5$, then $\langle \mathcal{M} \rangle \neq \mathrm{PG}(k-1, q)$, so $\mathcal{M}$ cannot be a cutting blocking set. For the remaining case, where $k = q = 4$, one can observe that for a (hyper)plane H in $\mathrm{PG}(3, q)$, H intersects $\mathcal{M}$ in a Baer subplane or in a Baer subline. In the latter case, one has $\langle \mathcal{M} \cap H \rangle \neq H$, and so $\mathcal{M}$ is not a cutting blocking set. The fact that for $k = q = 4$ a 3-dimensional Baer subspace $\mathcal{M}$ cannot be cutting could be also deduced from Example 3.64, since the cardinality of $\mathcal{M}$ is 15.

Case III: $q = k = 4$ and $\mathcal{M}$ is the complement of a hyperoval in a plane of $\mathrm{PG}(3, q)$. In this case $\langle \mathcal{M} \rangle \neq \mathrm{PG}(3, q)$ and $\mathcal{M}$ cannot be a cutting blocking set.

Therefore, when $4 \leq k \leq \sqrt{q}+2$, the bound in Theorem 3.51 is never sharp.

Corollary 3.56. Let $\mathcal{C}$ be a minimal $[n, k]_q$ code with $3 \leq k \leq \sqrt{q}+2$. Then $n \geq (q+1)(k-1)+1$, unless $q = 2$ and $k = 3$.

Proof. The case $k \geq 4$ has been discussed in Remark 3.55. When $k = 3$, cutting blocking sets are equivalent to 2-fold blocking set. Using Theorem 3.23, one can easily verify that the only case in which a 2-fold blocking set has cardinality $2q + 2$ is when $q = 2$. □

3.6 Statistical Approach

Most bounds for minimal codes we are aware of involve either (q, n, k), or (q, k, d). Bounds involving all the four parameters (q, n, k, d) can in turn be obtained combining these with classical bounds for Hamming-metric codes, such as the Singleton or the Griesmer bound.

In this section, we develop a method to establish new inequalities that directly involve all the four parameters of a minimal code, namely (q, n, k, d). As an application, we obtain an upper bound for the minimum distance d of a minimal code in terms of (q, n, k). As we will see in the examples, this bound excludes the existence of minimal codes with parameter sets that do not violate any of the known bounds.

Our approach combines Theorem 3.44 with ideas from statistics, interpreting the weight of the codewords of a linear code as a discrete random variable and computing/estimating its mean and variance. As simple corollaries of our bounds, we recover classical results on constant-weight codes.

Throughout this section, $\mathcal{C}$ denotes a nondegenerate code. Our results can be made more precise when $\mathcal{C}$ is projective, i.e., when in one (and thus in all) generator matrix G of $\mathcal{C}$ no two columns are proportional. Note that a projective code is also nondegenerate.

3.6.1 Mean and Variance of the Nonzero Weights in a Linear Code

We start with an upper bound for the sum of the squares of the weights in a nondegenerate linear code. The proof uses one of the Pless' identities.

Lemma 3.57. Let $\mathcal{C}$ be a nondegenerate $[n,k]_q$ code. We have

$$\sum_{c\in\mathcal{C}} \mathrm{wt}^{\mathrm{H}}(c)^2 \geq q^{k-2}\, n\,(q-1)\,[n(q-1)+1].$$

Moreover, equality holds if and only if $\mathcal{C}$ is projective.

Proof. For $i \in \{0,\dots,n\}$ we denote by $W_i(\mathcal{C}^\perp)$ the number of codewords of weight i in the dual code $\mathcal{C}^\perp$. Since $\mathcal{C}$ is nondegenerate, we have $W_1(\mathcal{C}^\perp)=0$. Moreover, $\mathcal{C}$ is projective if and only if $W_2(\mathcal{C}^\perp)=0$. Using Pless' identities [88, Theorem 7.2.3(P1)] we can write

$$\sum_{c\in\mathcal{C}} \mathrm{wt}^{\mathrm{H}}(c)^2 = \sum_{\nu=0}^{2}\left(\nu!\, S(2,\nu)\, q^{k-\nu}(q-1)^{\nu}\binom{n}{n-\nu}\right) + 4W_2(\mathcal{C}^\perp)S(2,2)q^{k-2},$$

where $S(a,b)\geq 0$ is the Stirling number of the second kind indexed by (a,b). Therefore

$$\sum_{c\in\mathcal{C}} \mathrm{wt}^{\mathrm{H}}(c)^2 \geq \sum_{\nu=0}^{2}\left(\nu!\, S(2,\nu)\, q^{k-\nu}(q-1)^{\nu}\binom{n}{n-\nu}\right),$$

with equality if and only if $\mathcal{C}$ is projective. The lemma now follows from the fact that $S(2,0)=0$ and $S(2,1)=S(2,2)=1$. □

The next step consists in defining the mean and variance of the nonzero weights in a linear code and to study the latter via Lemma 3.57.

Notation 3.58. For a code $\mathcal{C}$, let

$$\begin{aligned}\mathbb{E}(\mathcal{C}) &:= (q^k-1)^{-1}\sum_{c\in\mathcal{C}} \mathrm{wt}^{\mathrm{H}}(c),\\ \mathrm{Var}(\mathcal{C}) &:= (q^k-1)^{-1}\sum_{c\in\mathcal{C}} \mathrm{wt}^{\mathrm{H}}(c)^2 - \mathbb{E}(\mathcal{C})^2.\end{aligned}$$

We now compute/estimate these two quantities.

Theorem 3.59. Let $\mathcal{C}$ be a nondegenerate $[n,k]_q$ code. Let $\ell = n(q-1)/(q^k-1)$. We have $\mathbb{E}(\mathcal{C}) = q^{k-1}\ell$ and $\mathrm{Var}(\mathcal{C}) \geq q^{k-2}\,\ell(1-\ell)$. Moreover, equality holds if and only if $\mathcal{C}$ is projective.

Proof. Since $\mathcal{C}$ is nondegenerate, we have

$$\mathbb{E}(\mathcal{C}) = (q^k-1)^{-1}\sum_{i=1}^{n}(q^k-q^{k-1}) = n(q^k-q^{k-1})/(q^k-1) = q^{k-1}\ell.$$

Combining this with Lemma 3.57 we obtain

$$\begin{aligned}\mathrm{Var}(\mathcal{C}) &\geq \frac{q^{k-2}n(q-1)[n(q-1)+1]}{q^k-1} - q^{2k-2}\,\frac{n^2(q-1)^2}{(q^k-1)^2} \\ &= q^{k-2}\ell[n(q-1)+1] - \ell^2 q^{2k-2} \\ &= q^{k-2}\ell(1-\ell),\end{aligned}$$

as desired. □

Remark 3.60. The quantity $\mathbb{E}(\mathcal{C})$ in Notation 3.58 expresses the **average weight** of $\mathcal{C}$ and can be used to extend Theorem 3.9 as follows. Suppose that $\mathcal{C}$ is a nondegenerate $[n,k]_q$ code with maximum weight w. Using $\mathbb{E}(\mathcal{C}) \leq w$ one obtains

$$n \geq \left\lceil (n-w)\,\frac{q^k-1}{q^{k-1}-1} \right\rceil \geq (n-w)q+1.$$

In particular, if $\mathcal{C}$ is minimal then $w \leq n-k+1$ by Proposition 3.4, from which Theorem 3.9 follows.

As an immediate consequence of Theorem 3.59 we obtain the well-known fact that constant-weight codes have large length; see e.g. [43].

Corollary 3.61. Let $\mathcal{C}$ be a constant-weight $[n,k,d]_q$ code. Then we have $n \geq (q^k-1)/(q-1)$. Moreover, if $\mathcal{C}$ is projective then $n = (q^k-1)/(q-1)$ and $d = q^{k-1}$.

Proof. Without loss of generality, $\mathcal{C}$ is nondegenerate. By Theorem 3.59 we have $0 \geq q^{k-2}\ell(1-\ell)$, from which $\ell \geq 1$. If $\mathcal{C}$ is projective, then $\ell = 1$ and therefore $d = \mathbb{E}(\mathcal{C}) = q^{k-1}$, as claimed. □

3.6.2 Bounds

By applying the result of the previous subsection, we can finally derive an upper bound for the minimum distance of a code $\mathcal{C}$ as a function of q, n, k and the maximum weight in $\mathcal{C}$.

Theorem 3.62. Let $\mathcal{C}$ be a nondegenerate $[n,k,d]_q$ code of maximum weight $w > d$. Let $\ell = n(q-1)/(q^k-1)$. We have $w > n(q^k - q^{k-1})/(q^k-1)$ and

$$d \leq \left\lfloor q^{k-1}\ell - \frac{q^{k-2}\ell(1-\ell)}{w-q^{k-1}\ell} \right\rfloor. \tag{3.8}$$

Moreover, equality holds in (3.8) if and only if $\mathcal{C}$ is a projective two-weight code.

Proof. The first inequality follows from the fact that $w > \mathbb{E}(\mathcal{C})$, since $\mathcal{C}$ is not constant-weight. Using the inequality of Bhatia–Davis [37] we obtain

$$\mathrm{Var}(\mathcal{C}) \leq (w-\mathbb{E}(\mathcal{C}))(\mathbb{E}(\mathcal{C})-d).$$

Since $\mathcal{C}$ is nondegenerate and not constant-weight we have $\mathbb{E}(\mathcal{C}) = q^{k-1}\ell < w$. Therefore we conclude by Theorem 3.59.

The second part of the statement follows from the fact that the bound of Theorem 3.59 is sharp if and only if $\mathcal{C}$ is projective, and that the Bhatia–Davis inequality is met with equality if and only if the underlying distribution takes only two values; see [37, Proposition 1]. □

As an application of Theorem 3.62 we obtain the following bound for the minimum distance of a minimal code.

Corollary 3.63. Let $\mathcal{C}$ be a minimal nondegenerate $[n,k,d]_q$ code. If $\mathcal{C}$ is not constant-weight, then $n-k+1 > n(q^k-q^{k-1})/(q^k-1)$ and

$$d \leq \left\lfloor \frac{n(q-1)q^{k-2}[n-1-q(k-1)]}{n(q^{k-1}-1)-(k-1)(q^k-1)} \right\rfloor. \tag{3.9}$$

In particular, we have

$$q^{k-2}n^2 - Bn + C \geq 0, \tag{3.10}$$

where

$$\begin{cases} B = q^{k-2} + (k-1)(2q^{k-1}-1) + \dfrac{q^{k-1}-1}{q-1}, \\ C = (k-1)^2(q^k-1) + \dfrac{(k-1)(q^k-1)}{q-1}. \end{cases}$$

Proof. The maximum weight of $\mathcal{C}$ satisfies $w \leq n-k+1$ by Proposition 3.4. Combining this with Theorem 3.62 one gets $n-k+1 > n(q^k-q^{k-1})/(q^k-1)$ and

$$d \leq \left\lfloor q^{k-1}\ell - \frac{q^{k-2}\ell(1-\ell)}{n-k+1-q^{k-1}\ell} \right\rfloor, \tag{3.11}$$

where $\ell = n(q-1)/(q^k-1)$. Lengthy computations show that the RHS of (3.11) is equal to the RHS of (3.9). The second part of the statement follows by combining (3.9) with Theorem 3.44, after lengthy computations. □

Example 3.64. There is no minimal $[16,4]_4$ code. To see this, observe that if such a code existed, then (3.10) would give $-42 \geq 0$, a contradiction. So the minimum length of a minimal code of dimension 4 over $\mathbb{F}_4$ is at least 17.

Consider the parameters $(q,n,k) = (4,17,4)$ and suppose that there exists an $[17,4]_4$ nondegenerate minimal code $\mathcal{C}$. Since $n < (q^k-1)/(q-1)$, $\mathcal{C}$ cannot be constant weight by Corollary 3.61. Therefore by Corollary 3.63 we conclude that $d \leq 10$. The existence of a minimal nondegenerate $[17,4,11]_4$ code is therefore excluded by Corollary 3.63, but it is not excluded by any of the other known bounds for the parameters of minimal codes. Note moreover that, by

Theorem 3.44, we have $d \geq 10$. Therefore the minimum distance of a putative $[17,4]_4$ nondegenerate minimal code is exactly 10 (when the largest minimum distance of an "unrestricted" $[17,4]_4$ linear code is instead known to be 12).

Remark 3.65. The constraints imposed by Corollary 3.63 and Theorem 3.51 are in general incomparable. More precisely, each of the two results excludes the existence of some minimal codes that are not excluded by the other.

One can see that Corollary 3.63 improves on Theorem 3.51 if and only if (3.10) is violated when specialized to $n = (q+1)(k-1)$. After lengthy computations, one sees that this happens if and only if

$$k < \frac{q^{k-1} + 2q^{k-2} - 2q^{k-3} + q - 2}{(q^{k-3}+1)(q-1)}.$$

By manipulating this inequality, it can be checked that Corollary 3.63 improves on Theorem 3.51 for the parameter set $\{(k,q) \mid 3 \leq k \leq q+3,\ q \geq 3\}$. When q is at least 3, this is an improvement also on Corollary 3.56. On the other hand, Theorem 3.51 provides a strictly sharper estimate than Corollary 3.63 if and only if (3.10) is satisfied for $n = (q+1)(k-1) - 1$. For instance, this happens for the parameter set $\{(k,q) : k \geq 2q\}$.

We include in Table 3.1 three collections of parameter sets that are excluded by Theorem 3.51 and Corollary 3.63. The first column contains parameters that are excluded by both results, while the other two contain parameters that are excluded by either Theorem 3.51 or Corollary 3.63 (and not by both).

Some parameters of minimal codes excluded by both Theorem 3.51 and Corollary 3.63	Some parameters of minimal codes excluded by Theorem 3.51 and not by Corollary 3.63	Some parameters of minimal codes excluded by Corollary 3.63 and not by Theorem 3.51
$[8,4]_2$	$[17,7]_2$	$[16,5]_3$
$[15,5]_3$	$[31,9]_3$	$[16,4]_4$
$[24,6]_4$	$[44,10]_4$	$[25,6]_4$
$[35,7]_5$	$[99,21]_4$	$[36,7]_5$
$[63,9]_7$	$[65,12]_5$	$[26,4]_7$

Table 3.1: Code parameters for which the existence of minimal codes is excluded by Theorem 3.51 and/or Corollary 3.63.

3.6.3 Other Applications

In this short subsection we illustrate how Theorem 3.62 can be applied to study codes that are not necessarily minimal. We start with a generalization of Corollary 3.61. More precisely, we show that the relative difference between the maximum and minimum weight of a code, $(w-d)/n$, gives a lower bound on the code's length. In other words, if the maximum and minimum weight of a code are relatively close to each other, then the code length is necessarily large.

Proposition 3.66. Let $\mathcal{C}$ be a nondegenerate $[n,k,d]_q$ code having maximum weight w. We have

$$\frac{1}{n} \leq \frac{q-1}{q^k-1} + \frac{1}{4}\left(\frac{w-d}{n}\right)^2 \frac{q^k-1}{q^{k-2}(q-1)}.$$

Note that in the extreme case where $w=d$ we recover Corollary 3.61.

Proof of Proposition 3.66. Using Popoviciu's inequality for the variance, along with Theorem 3.59, we find

$$q^{k-2}\ell(1-\ell) \leq \mathrm{Var}(\mathcal{C}) \leq \frac{1}{4}(w-d)^2, \tag{3.12}$$

where $\ell = n(q-1)/(q^k-1)$. By substituting this value of ℓ into (3.12) one obtains

$$\frac{n(q-1)q^{k-2}}{q^k-1} - \frac{n^2(q-1)^2q^{k-2}}{(q^k-1)^2} \leq \frac{1}{4}(w-d)^2.$$

Multiplying both sides of the previous inequality by $(q^k-1)/(n^2q^{k-2}(q-1))$ and re-arranging the terms produces the desired result. □

Remark 3.67. Note that combining the inequality in Proposition 3.66 with the Ashikhmin-Barg condition $w < \frac{q}{q-1}d$ (which is, as we have already mentioned, a sufficient condition for a code to be minimal) we get the following second degree equation in n:

$$n^2 - \frac{q^k-1}{q-1}n + \frac{(q^k-1)^2d^2}{4q^{k-2}(q-1)^4} > 0.$$

If the discriminant of the left hand-side is negative, that is if $d > (q-1)q^{k/2-1}$, we get $n \geq d > (q-1)q^{k/2-1}$. Otherwise, $n \geq \frac{1}{2}\frac{q^k-1}{q-1}$. So in both cases the length is exponential in k, which means that minimal codes satisfying the Ashikhmin-Barg condition are long.

A second application of Theorem 3.62 consists in obtaining constraints on the parameters of a code having few weights. A classical result about these codes is the following theorem by Delsarte.

Theorem 3.68 (see [62]). Let $\mathcal{C}$ be an $[n,k]_q$ code and $s = |\{\omega(c) \mid c \in \mathcal{C},\, c \neq 0\}|$. We have

$$q^k \leq \sum_{i=0}^{s} \binom{n}{i}(q-1)^i.$$

Specializing to $s=2$, the previous theorem shows that, for example, any two-weight $[n,k]_q$ code satisfies

$$q^k \leq 1 + n(q-1) + \frac{n(n-1)}{2}(q-1)^2. \tag{3.13}$$

This result however does not take into account *which* values the weight distribution can take. Exploiting this information, Theorem 3.62 provides in general different constraints on n than those in (3.13). We illustrate this with an example.

Example 3.69. Following the notation of Theorem 3.62 and Theorem 3.68, let $(q, k, s, d, w) = (2, 8, 2, 16, 24)$. We look for a nondegenerate binary two-weight code $\mathcal{C}$ of dimension 8 having weights 16 and 24. The constraints imposed on n by Theorem 3.62 imply $34 \leq n \leq 45$, where the upper bound is met with equality if $\mathcal{C}$ is projective. The constraint imposed by (3.13) is instead $n \geq 23$. It is known that there exists a projective binary two-weight code of parameters $(n, k, d, w) = (45, 8, 16, 24)$.

3.7 Geometric Approach

As already illustrated in Theorem 3.8, minimal codes are in one-to-one correspondence with cutting blocking sets. In this section we focus on this point of view on minimal codes, exploiting their geometric characterization to construct new, general and infinite families of minimal codes. In particular, we provide a construction of cutting blocking sets derived from Desarguesian $(r-1)$-spreads of $\mathrm{PG}(rt-1, q)$. In turn, this leads to an inductive construction of small cutting blocking sets or, equivalently, of minimal codes with short length. In contrast to previous approaches, our construction works over any (possibly very small) finite field.

3.7.1 Constructing Minimal Codes from Blocking Sets

In this subsection we generalize the notion of cutting blocking sets, which we will use to construct cutting blocking sets from blocking sets which have weaker properties. We will call such blocking sets ℓ-**cutting**, and we define them as follows. Note that this section cannot be found in the published paper [7].

Definition 3.70. Let ℓ, N be positive integers with $\ell \leq N$. A (blocking) set $\mathcal{B} \subseteq \mathrm{PG}(N, q)$ is said to be ℓ-**cutting** if for any hyperplane H of $\mathrm{PG}(N, q)$, $\dim(\langle H \cap \mathcal{B} \rangle) \geq N - \ell$.

Observe that the 1-cutting blocking set are simply cutting blocking sets, while N-cutting blocking sets are just blocking sets. Hence, the notion of ℓ-cutting blocking sets connects these two notions in a more general framework.

Theorem 3.71. Let $r \geq 1$ be an integer and let $\mathcal{B}_1, \dots, \mathcal{B}_r$ be pairwise distinct ℓ-cutting blocking sets in $\mathrm{PG}(k-1, q)$. Suppose that for every $(k-3)$-flat $\Lambda \subseteq \mathrm{PG}(k-1, q)$ there exists $i \in \{1, \dots, r\}$ such that $\dim(\langle \Lambda \cap \mathcal{B}_i \rangle) < k - \ell - 1$. Then $\mathcal{M} := \mathcal{B}_1 \cup \ldots \cup \mathcal{B}_r$ is a cutting blocking set.

Proof. Let H be a hyperplane in $\mathrm{PG}(k-1, q)$. Assume that $H \cap \mathcal{M}$ is contained in a $(k-3)$-flat Λ. Then $\Lambda \supseteq H \cap \mathcal{M}$ and in particular, Λ contains each $\langle H \cap \mathcal{B}_i \rangle$, which by definition has dimension at least $k - \ell - 1$. However, this contradicts the hypothesis. □

As a particular case of previous theorem, we have the following result.

Theorem 3.72. Let $r \geq 2$ be an integer and let $\mathcal{B}_1, \dots, \mathcal{B}_r$ be pairwise distinct blocking sets in $\mathrm{PG}(k-1,q)$. Suppose that for every $(k-3)$-flat $\Lambda \subseteq \mathrm{PG}(k-1,q)$ there exists $i \in \{1,\dots,r\}$ such that $\Lambda \cap \mathcal{B}_i = \emptyset$. Then $\mathcal{M} := \mathcal{B}_1 \cup \ldots \cup \mathcal{B}_r$ is a cutting blocking set.

Proof. Let H be a hyperplane in $\mathrm{PG}(k-1,q)$. Assume that $H \cap \mathcal{M}$ is contained in a $(k-3)$-flat Λ. Then $\Lambda \supseteq H \cap \mathcal{M}$ and in particular, Λ contains each $H \cap \mathcal{B}_i$, which is non-empty by definition of blocking set. Therefore, Λ meets all the $\mathcal{B}_i$'s, which contradicts the hypothesis. □

3.7.2 Minimal Codes from Spreads

We start by recalling the definition of t-spread in $\mathrm{PG}(k-1,q)$, which we will use to obtain a new construction of minimal codes. A t-**spread** S of $\mathrm{PG}(k-1,q)$ is a partition of $\mathrm{PG}(k-1,q)$ in t-flats. It is well known that such a t-spread exists if and only if $t+1$ divides k; see [139]. In particular, a 1-spread of $\mathrm{PG}(k-1,q)$ is a partition of its points into disjoint lines and it is also called a **linespread**. It exists if and only if k is even.

An algebraic representation of an $(r-1)$-spread of $\mathrm{PG}(2r-1,q)$ can be obtained as follows. Let $\gamma \in \mathbb{F}_{q^r}$ be a primitive element and let $M \in \mathbb{F}_q^{r\times r}$ be the companion matrix of the minimal polynomial of γ over $\mathbb{F}_q$. It is well known that $\mathbb{F}_{q^r} \cong \mathbb{F}_q[\gamma] \cong \mathbb{F}_q[M] = \{0\} \cup \{M^i : 1 \leq i \leq q^r - 1\}$ as $\mathbb{F}_q$-algebras. For $i \in [q^r-1]$ define $V_i := \{[x : xM^i] \mid x \in \mathrm{PG}(r-1,q)\}$, $V_0 := \{[x : 0] \mid x \in \mathrm{PG}(r-1,q)\}$ and $V_{q^r} := \{[0 : y] \mid y \in \mathrm{PG}(r-1,q)\}$. Then the set $\{V_0, \dots, V_{q^r}\}$ is an $(r-1)$-spread of $\mathrm{PG}(2r-1,q)$.

Theorem 3.73. Let S be the $(r-1)$-spread of $\mathrm{PG}(2r-1,q)$ defined above and let $\mathcal{B} = V_0 \cup V_i \cup V_j \cup V_{q^r} \subseteq \mathrm{PG}(2r-1,q)$, with $0 < i < j < q^r$. Suppose that for every $s > 1$ dividing r we have $j - i \not\equiv 0 \mod \left(\frac{q^s-1}{q-1}\right)$. Then $\mathcal{B}$ is a cutting blocking set.

Proof. For ease of exposition we switch to vector notation, in which we represent V_0, V_i, V_j, V_{q^r} as elements of the Grassmannian $\mathrm{Gr}_q(r,2r)$. In this representation we have $V_0 = \mathrm{rowsp}(I_r \mid 0)$, $V_{q^r} = \mathrm{rowsp}(0 \mid I_r)$, $V_i = \mathrm{rowsp}(I_r \mid M^i)$ and $V_j = \mathrm{rowsp}(I_r \mid M^j)$. Let H be a hyperplane in $\mathbb{F}_q^{2r}$. We want to show that $\langle H \cap \mathcal{B}\rangle = H$, or, equivalently, that $\langle H \cap \mathcal{B}\rangle = \langle H \cap V_0\rangle + \langle H \cap V_{q^r}\rangle + \langle H \cap V_i\rangle + \langle H \cap V_j\rangle$ has dimension at least $2r-1$. Observe first that if H contains one among the V_ℓ's, say V_0, then there is nothing to prove, since $\langle H \cap V_0\rangle + \langle H \cap V_i\rangle$ has already dimension (at least) $2r-1$. Hence we can assume that H intersect both V_0 and V_{q^r} in an $(r-1)$-dimensional subspace. Then the space $\langle H \cap V_0\rangle + \langle H \cap V_{q^r}\rangle$ has dimension $2r-2$. We can write the intersection spaces as

$$\begin{aligned} H \cap V_0 &= \mathrm{rowsp}(\,X_1 \mid \;0\;\,), & \qquad H \cap V_i &= \mathrm{rowsp}(X_2 \mid X_2M^i), \\ H \cap V_{q^r} &= \mathrm{rowsp}(\;0\; \mid X_3\,), & \qquad H \cap V_j &= \mathrm{rowsp}(X_4 \mid X_4M^j), \end{aligned}$$

for some $X_1, X_2, X_3, X_4 \in \mathbb{F}_q^{(r-1)\times r}$ of rank $r-1$.

Suppose by contradiction that $\langle H \cap \mathcal{B} \rangle$ has dimension exactly $2r-2$. This implies

$$\mathrm{rowsp}\begin{pmatrix} X_2 & X_2M^i \\ X_4 & X_4M^j \end{pmatrix} \subseteq \mathrm{rowsp}\begin{pmatrix} X_1 & 0 \\ 0 & X_3 \end{pmatrix},$$

which in turn implies

$$\mathrm{rowsp}(X_2) = \mathrm{rowsp}(X_4) = \mathrm{rowsp}(X_1), \quad \mathrm{rowsp}(X_3) = \mathrm{rowsp}(X_2M^i) = \mathrm{rowsp}(X_4M^j).$$

Without loss of generality, we can assume that $X_1 = X_2 = X_4 =: X$, which reduces the above condition to

$$\mathrm{rowsp}(X_3) = \mathrm{rowsp}(XM^i) = \mathrm{rowsp}(XM^j).$$

Thus, there exists a matrix $A \in \mathrm{GL}(r-1,q)$ such that

$$AX - XM^{j-i} = 0. \tag{3.14}$$

The matrix equation in (3.14), where the matrix X is the unknown, is a *Sylvester equation*. This is known to have a unique solution if the minimal polynomials of A and M^{j-i} are coprime; see e.g. [86, Theroem 2.4.4.1]. Observe that the minimal polynomial of M^{j-i} is irreducible of degree r, since M^{j-i} corresponds to the element γ^{j-i} and by the assumption on $j-i$ in the statement we have $\mathbb{F}_q[\gamma^{j-i}] = \mathbb{F}_{q^r}$. Moreover, the minimal polynomial of A has degree at most $r-1$, and hence it is coprime with the one of M^{j-i}'s. Therefore (3.14) has a unique solution, which is clearly $X=0$. This leads to a contradiction and concludes the proof. □

Remark 3.74. Observe that in the particular case of $r = 2$, the construction provided in Theorem 3.73 works because the algebraic spread S that we use is **regular**. This means that S is disjoint union of $q-1$ reguli. A **regulus** is a collection of $q+1$ disjoint lines such that, if a line meets 3 of them, then it meets all of them. And the regulus is uniquely identified by any three lines in it. Hence, if we take any three lines from the spread and the forth one which does not belong to the spread that they generate, then any line that intersects the first three lines cannot intersect also the forth one. Then, using Theorem 3.72, we can deduce that these four lines form a cutting blocking set. In Theorem 3.73, once we fix V_0, V_i, V_{q^2}, the unique regulus defined by those three lines is given by $\{V_0, V_{q^2}, V_i\} \cup \{V_r : r-i \equiv 0 \mod q+1\}$.

We now concentrate on the more general case of $(r-1)$-spreads in $\mathrm{PG}(rt-1,q)$. These can be constructed using the so-called **field reduction**; see [139, 99]. This technique identifies points in $\mathrm{PG}(t-1,q^r)$ with $(r-1)$-flats in $\mathrm{PG}(rt-1,q)$. The idea is exactly the same as for the algebraic $(r-1)$-spread of $\mathrm{PG}(2r-1,q)$ described above. Let γ be a primitive element in $\mathbb{F}_{q^r}$ and let M be the companion matrix of the minimal polynomial of γ over $\mathbb{F}_q$. As already explained, there is an isomorphism $\mathbb{F}_{q^r} \cong \mathbb{F}_q[M] = \{0\} \cup \{M^i : 1 \leq i \leq q^r - 1\}$, which we call ϕ. We can then extend it to vectors in $\mathbb{F}_{q^r}^t$ componentwise, obtaining an injective map

$$\begin{array}{rcl} \varphi : \mathbb{F}_{q^r}^t & \longrightarrow & \mathbb{F}_q^{r\times rt} \\ (v_1,\dots,v_t) & \longmapsto & (\phi(v_1) \mid \dots \mid \phi(v_t)). \end{array}$$

This map can in turn be extended to a map $\bar{\varphi} : \mathrm{PG}(t-1,q^r) \longrightarrow \mathrm{Gr}_q(r,tr)$, the Grassmannian, defined by $P = [v] \longmapsto \mathrm{rowsp}(\varphi(v))$. Note that $\bar{\varphi}$ is well-defined since it does not depend on the choice of the representative v for the point P. Indeed, for a nonzero scalar multiple of v, say $\gamma^i v$, we have $\varphi(\gamma^i v) = M^i\varphi(v)$ and since M^i is invertible, $\mathrm{rowsp}(M^i\varphi(v)) = \mathrm{rowsp}(\varphi(v))$. It is then well-known that $\mathrm{Im}(\bar{\varphi})$ is a (vectorial) r-spread of $\mathbb{F}_q^{rt}$, which naturally gives rise to a projective $(r-1)$-spread of $\mathrm{PG}(rt-1,q)$. Such a spread is known as **Desarguesian spread**; see [139].

In the sequel we will need the following special points in $\mathrm{PG}(t-1,q^r)$: $P_\ell := [e_\ell]$ for $\ell \in [t]$ and $Q_{\ell,m,i} := [u_{\ell,m,i}]$, where $u_{\ell,m,i} := e_\ell + \gamma^i e_m$ for $1 \le \ell < m \le t$ and $i \in [q^r-1]$. These will be used in the next result to extend the construction of Theorem 3.73 from two to t **blocks**.

Theorem 3.75. For each pair of integers (ℓ,m) such that $1 \le \ell < m \le t$, let $j_{\ell,m}, i_{\ell,m} \in [q^r-1]$ be integers with the following property: for all $s > 1$ dividing r, $j_{\ell,m} - i_{\ell,m} \not\equiv 0 \mod \left(\frac{q^s-1}{q-1}\right)$. Define the set

$$\mathcal{T} := \left(\bigcup_{1\le\ell\le t} \bar{\varphi}(P_\ell)\right) \cup \left(\bigcup_{1\le\ell<m\le t} (\bar{\varphi}(Q_{\ell,m,i_{\ell,m}}) \cup \bar{\varphi}(Q_{\ell,m,j_{\ell,m}}))\right).$$

Then the projectivization of $\mathcal{T}$ is a cutting blocking set in $\mathrm{PG}(rt-1,q)$.

Proof. Once again we work in vector notation. Let H be a hyperplane in $\mathbb{F}_q^{rt}$. Let $a := |\{\ell : \bar{\varphi}(P_\ell) \subseteq H\}|$. Then $0 \le a \le t-1$ and $\dim(\langle H \cap \mathcal{T}\rangle) \ge ra + (r-1)(t-a)$. Without loss of generality assume that $\{\ell : \bar{\varphi}(P_\ell) \subseteq H\} = [a]$. Hence $\langle H \cap \mathcal{T}\rangle$ contains the span of the first $a \cdot r$ standard basis vectors. By taking the quotient on this span, we reduce ourselves to proving the same statement for $a = 0$, replacing t by $t-a$. Therefore we can also assume $a = 0$ without loss of generality.

We have that $\Lambda := \langle H \cap \left(\bigcup_\ell \bar{\varphi}(P_\ell)\right)\rangle$ has dimension $(r-1)t$. For all integers $1 \le \ell < m \le t$, define

$$\begin{aligned} \mathcal{S}_{\ell,m} &:= \bar{\varphi}(P_\ell) \cup \bar{\varphi}(P_m) \cup \bar{\varphi}(Q_{\ell,m,i_{\ell,m}}) \cup \bar{\varphi}(Q_{\ell,m,j_{\ell,m}}), \\ \Pi_{\ell,m} &:= \langle \bar{\varphi}(P_\ell) \cup \bar{\varphi}(P_m)\rangle = \langle e_i : (\ell-1)r+1 \le i \le \ell r, \text{ or } (m-1)r \le i \le mr\rangle. \end{aligned}$$

Then $H \cap \Pi_{\ell,m}$ is a hyperplane in $\Pi_{\ell,m} \cong \mathbb{F}_q^{2r}$. Moreover, using the same argument as in the proof of Theorem 3.73, there exists a vector $v_{\ell,m} \in (H \cap \Pi_{\ell,m}) \cap \mathcal{S}_{\ell,m} \subsetneq H \cap \mathcal{S}_{\ell,m}$ such that $v_{\ell,m} \notin \langle H \cap (\bar{\varphi}(P_\ell) \cup \bar{\varphi}(P_m))\rangle$. Observe that the support of $v_{\ell,m}$ is contained only in the ℓ-th and the m-th blocks and that we can write $v_{\ell,m} = w_{\ell,m}^{(\ell)} + w_{\ell,m}^{(m)}$, where $w_{\ell,m}^{(j)} \in \langle \bar{\varphi}(P_j)\rangle$, i.e., it has support contained only in the j-th block, for $j \in \{\ell, m\}$. Now consider the $t-1$ vectors

$v_{1,2},\dots,v_{1,t}$. Since $a=0$, none of the $v_{1,i}$'s belongs to Λ. It is left to show that for each $i\geq 3$ we have $v_{1,i}\notin\Gamma_{i-1}:=\Lambda+\langle v_{1,2},\dots,v_{1,i-1}\rangle$. By contradiction, suppose that $v_{1,i}\in\Gamma_{i-1}$. Let $\rho_i:\mathbb{F}_q^{rt}\to\mathbb{F}_q^r$ denote the projection on the i-th block. We have

$$w_{1,i}^{(i)}=\rho_i(v_{1,i})\in\rho_i(\Gamma_i)=\langle H\cap\bar{\varphi}(P_i)\rangle,$$

since, by construction, the i-th block of any vector in Γ_{i-1} is equal to the i-th block of some element in $\bar{\varphi}(P_i)\cap H$. Therefore, also the vector $w_{1,i}^{(1)}=v_{1,i}-w_{1,i}^{(i)}$ belongs to H. This means that $v_{1,i}\in H\cap(\bar{\varphi}(P_1)\cup\bar{\varphi}(P_i))\subseteq\Lambda$, which leads to a contradiction. □

Remark 3.76. The construction of Theorem 3.75 for $r=t=2$ (or, equivalently, the one of Theorem 3.73 for $r=2$) coincides with the construction of cutting blocking sets of [59, Theorem 3.7], which consists of 4 disjoint lines in $\mathrm{PG}(3,q)$. Therefore, Theorem 3.75 can be viewed as a generalization of that result.

Example 3.77. We explicitly construct a cutting blocking set in $\mathrm{PG}(5,q)$ as explained in Theorem 3.75, with $r=2$ and $t=3$. We take as γ a primitive element of $\mathbb{F}_{q^2}$ whose minimal polynomial over $\mathbb{F}_q$ is $x^2-p_1x-p_0$. We have $P_1=[1:0:0]$, $P_2=[0:1:0]$, $P_3=[0:0:1]$ and choose the following points in $\mathrm{PG}(2,q^2)$: $Q_{1,2,q^2-1}=[1:1:0]$, $Q_{1,2,1}=[1:\gamma:0]$, $Q_{1,3,q^2-1}=[1:0:1]$, $Q_{1,3,1}=[1:0:\gamma]$, $Q_{2,3,q^2-1}=[0:1:1]$, $Q_{2,3,1}=[0:1:\gamma]$. Therefore the set $\mathcal{T}$ is

$$\begin{aligned}\mathcal{T}=&\{(x,y,0,0,0,0):x,y\in\mathbb{F}_q\}\\&\cup\{(0,0,x,y,0,0):x,y\in\mathbb{F}_q\}\cup\{(0,0,0,0,x,y):x,y\in\mathbb{F}_q\}\\&\cup\{(x,y,x,y,0,0):x,y\in\mathbb{F}_q\}\cup\{(x,y,y,p_0x+p_1y,0,0):x,y\in\mathbb{F}_q\}\\&\cup\{(x,y,0,0,x,y):x,y\in\mathbb{F}_q\}\cup\{(x,y,0,0,y,p_0x+p_1y):x,y\in\mathbb{F}_q\}\\&\cup\{(0,0,x,y,x,y):x,y\in\mathbb{F}_q\}\cup\{(0,0,x,y,y,p_0x+p_1y):x,y\in\mathbb{F}_q\}.\end{aligned}$$

The projectivization of $\mathcal{T}$ gives the desired cutting blocking set in $\mathrm{PG}(5,q)$.

3.7.3 Inductive Constructions of Cutting Blocking Sets

As already observed in the Introduction, of particular interest is the study of minimal codes of small length for a given dimension. Formally, for a fixed positive integer k and a prime power q, we are interested in determining the value of

$$m(k,q):=\min\{n\in\mathbb{N}_{\geq 1}\mid\text{ there exists a minimal }[n,k]_q\text{ code}\}.$$

This function has been explicitly studied in [106], where it was observed that $m(2,q) = q+1$ and that

$$q(k-1)+1 \leq m(k,q) \leq (q-1)\binom{k}{2} + k, \tag{3.15}$$

where the upper bound is *constructive* (the tetrahedron from page 24). The same results were independently obtained in [5], where shorter minimal codes are constructed for $k \in \{3,4,5\}$. In this notation, Theorem 3.51 improves on the lower bound in (3.15), reading

$$m(k,q) \geq (q+1)(k-1).$$

We already obtained improvements on this bound in Corollary 3.56 and Corollary 3.63, as shown in Table 3.1.

In [48] it has been shown that the upper bound on $m(k,q)$ in (3.15) is far from being tight. More precisely, one has

$$m(k,q) \leq \frac{2k}{\log_q\left(\frac{q^2}{q^2-q+1}\right)}, \tag{3.16}$$

indicating that, in principle, for a fixed q and k large enough one might construct much shorter minimal codes. In particular, a natural problem is that of finding, for a fixed q, an infinite family of minimal codes over $\mathbb{F}_q$ whose length is linear in k. This problem is naturally motivated by the goal of *explicitly* constructing asymptotically good minimal codes. Indeed, while these codes are known to be asymptotically good, the proofs in [54, 5] are not constructive, as well as the bound in (3.16). We are currently unaware of any *explicit* general construction of minimal codes whose length is unbounded for a *fixed* q, and that are asymptotically shorter than the tetrahedron; see also the discussion in Remark 3.24.

In the sequel, we introduce two new families of minimal codes whose lengths are shorter than the one of the tetrahedron by a factor 2 and by a factor $\frac{9}{4}$, respectively. We start with a result that represents a first step towards inductive constructions of cutting blocking sets.

Proposition 3.78. Let $\mathcal{B} = \mathcal{B}_1 \cup \ldots \cup \mathcal{B}_r$ be a cutting blocking set in $\mathrm{PG}(N,q)$. For each $i \in [r]$, let $\Gamma_i := \langle \mathcal{B}_i \rangle \cong \mathrm{PG}(n_i,q)$ for some $n_i \leq N$ and let $\mathcal{B}'_i \subseteq \Gamma_i$ be the isomorphic image of a cutting blocking set in $\mathrm{PG}(n_i,q)$. Then $\mathcal{B}' := \mathcal{B}'_1 \cup \ldots \cup \mathcal{B}'_r$ is a cutting blocking set.

Proof. Let H be a hyperplane in $\mathrm{PG}(N,q)$. We want to show that $\langle H \cap \mathcal{B}' \rangle = H$. By hypothesis we have that

$$H = \langle H \cap \mathcal{B} \rangle = \langle H \cap \mathcal{B}_1 \rangle + \ldots + \langle H \cap \mathcal{B}_r \rangle.$$

Consider the spaces $\Lambda_i := H \cap \langle \mathcal{B}_i \rangle$, $i \in [r]$. Clearly, $\Lambda_i \supseteq \langle H \cap \mathcal{B}_i \rangle$ for all i. We now examine two cases separately.

Case I: $\Lambda_i = \langle \mathcal{B}_i \rangle$, that is, H contains $\langle \mathcal{B}_i \rangle$. In this case H also contains $\mathcal{B}'_i$ and $\langle H \cap \mathcal{B}'_i \rangle = \langle \mathcal{B}'_i \rangle = \langle \mathcal{B}_i \rangle = \Lambda_i$.

Case II: Λ_i is a hyperplane in $\langle \mathcal{B}_i \rangle$. By hypothesis, $\mathcal{B}'_i$ is a cutting blocking set in $\langle \mathcal{B}_i \rangle$, and hence $\langle H \cap \mathcal{B}'_i \rangle \supseteq \langle \Lambda_i \cap \mathcal{B}'_i \rangle = \Lambda_i$.

Therefore in both cases we have

$$\begin{aligned} \langle H \cap \mathcal{B}' \rangle = \langle H \cap \mathcal{B}'_1 \rangle + \ldots + \langle H \cap \mathcal{B}'_r \rangle &\supseteq \Lambda_1 + \ldots + \Lambda_r \\ &\supseteq \langle H \cap \mathcal{B}_1 \rangle + \ldots + \langle H \cap \mathcal{B}_r \rangle = H, \end{aligned}$$

concluding the proof. □

We are now ready to combine the above result with Theorem 3.75 and derive a recursive upper bound on $m(k,q)$.

Theorem 3.79. For all positive $a, b \in \mathbb{N}$,

$$m(ab,q) \leq a^2 m(b,q).$$

Proof. By Theorem 3.75 we know that we can construct a cutting blocking set in $\mathrm{PG}(ab-1,q)$ with the aid of a $(b-1)$-spread. More precisely, we only need to take a^2 disjoint $(b-1)$-flats $\Gamma_1, \ldots, \Gamma_{a^2} \cong \mathrm{PG}(b-1,q)$ from the spread. By Proposition 3.78, for each of them we can take the isomorphic image of a cutting blocking set in $\mathrm{PG}(b-1,q)$ with minimum cardinality $m(b,q)$. Therefore, we finally obtain a cutting blocking set in $\mathrm{PG}(ab-1,q)$ of cardinality $a^2 m(b,q)$. □

Observe that the proof of Theorem 3.79 gives an explicit way of constructing a minimal $[a^2 m(b,q), ab]_q$ code, provided that there exists already a construction for an $[m(b,q), b]_q$ minimal code. We illustrate how this construction works with the following example.

Example 3.80. We fix $k = 6 = 3 \cdot 2$ and assume q to be a square. Observe that under these assumptions we know the exact values of $m(2,q)$ and $m(3,q)$. Namely, we have $m(2,q) = q+1$ and

$$m(3,q) = \begin{cases} 3q & \text{if } q = 4, \\ 2(q+\sqrt{q}+1) & \text{if } q \geq 9. \end{cases}$$

Now we can use Theorem 3.79 in two ways. On the one hand, we deduce that

$$m(6,q) \leq 9 \cdot m(2,q) = 9(q+1).$$

Such a construction is obtained by taking 9 lines from a linespread in $\mathrm{PG}(5,q)$ as explained also in Example 3.77. On the other hand, by interchanging the roles of 2 and 3 we obtain

$$m(6,q) \leq 4 \cdot m(3,q) = \begin{cases} 12(q+1) & \text{if } q = 4, \\ 8(q+\sqrt{q}+1) & \text{if } q \geq 9. \end{cases}$$

The corresponding cutting blocking set is constructed by first selecting 4 planes in $\mathrm{PG}(5,q)$ via Theorem 3.73, and then by choosing, in each of these planes, a minimal 2-fold blocking set: when $q=4$, we take 3 lines not intersecting all in the same point; when $q \geq 9$, we choose 2 disjoint Baer subplanes. It is easy to check that for $q<64$ the cutting blocking set consisting of 9 lines is smaller, while for $q \geq 64$ the 8 Baer subplanes give rise to a cutting blocking set with smaller cardinality. Notice that both constructions produce a smaller cutting blocking set than the tetrahedron, which contains $15q-9$ points. For instance, let us consider the case $q=4$. The 9 lines give rise to a minimal $[45,6]_4$ code, the 8 Baer subplanes lead to a minimal $[56,6]_4$ code, while the tetrahedron provides a $[66,4]_4$ code. If we take $q=64$, then the three constructions produce minimal codes whose parameters are $[585,6]_{64}$, $[584,6]_{64}$ and $[966,6]_{64}$, respectively.

Remark 3.81. Very recently, a construction of cutting blocking sets in $\mathrm{PG}(5,q)$ as union of seven disjoint lines has been given in [28]. This gives an improvement on the known upper bound for $m(6,q)$. In the same work, a construction of a cutting blocking set in $\mathrm{PG}(3,q^3)$ of size $3(q^3+q^2+q+1)$ has been obtained as union of three suitable disjoint q-order subgeometries. These results together yield the following bounds:

$$\begin{aligned} m(4,q^3) &\leq 3(q^3+q^2+q+1), \\ m(6,q) &\leq 7(q+1). \end{aligned}$$

The proof of Theorem 3.79, which constructs minimal codes of dimension $k=ab$, heavily relies on the existence of a smaller minimal code, whose dimension divides k. Clearly, this recursive construction does not cover all dimensions, as for instance it does not provide any nontrivial minimal code of prime dimension. While for $k=5$ one can rely on the construction provided in [5, Construction 2], which gives a $[8q-3,5]_q$ minimal code, for primes greater than 5 we are not (yet) able to construct any *short* minimal code different from the tetrahedron. Also, we are not (yet) able to construct short minimal codes of odd dimension, unless the latter is divisible by 3 and q is a square. When k is odd one can construct minimal codes taking several $(r-1)$-flats in $\mathrm{PG}(k-1,q)$, where r is the smallest prime dividing k. However, when such a prime is big, the resulting code turns out to be quite long.

The discussion in the previous paragraph motivates us to look for alternative constructions of minimal codes, with the ultimate goal of covering a larger dimension range. Our next move in this direction is an inductive result that allows us to construct a cutting blocking set in $\mathrm{PG}(k,q)$ starting from a smaller one in $\mathrm{PG}(k-1,q)$. The following result has already been shown in [59, Construction A]. We include a proof for completeness.

Proposition 3.82 (see [59, Theorem 3.10])**.** Let $\mathcal{B}'$ be a cutting blocking set in $\mathrm{PG}(k-1,q)$. Fix a hyperplane $\Lambda \subseteq \mathrm{PG}(k,q)$ and take an isomorphic image $\mathcal{T}$ of $\mathcal{B}'$ in Λ. Moreover, select k points $P_1,\ldots,P_k \in \langle\mathcal{T}\rangle$ not lying all in the same $(k-2)$-flat and a point $P \in \mathrm{PG}(k,q)\setminus\Lambda$.

Define the lines $\ell_i := \langle P_i, P\rangle$. Then the set

$$\mathcal{B} := \mathcal{T} \cup \left(\bigcup_{i=1}^{k} \ell_i \setminus \{P_i\} \right)$$

is a cutting blocking set in $\mathrm{PG}(k,q)$. In particular, for every $k \in \mathbb{N}_{\geq 1}$ we have

$$m(k+1,q) \leq m(k,q) + (q-1)k + 1.$$

Proof. Let H be a hyperplane in $\mathrm{PG}(k,q)$. If $H = \Lambda$, then clearly $\langle H \cap \mathcal{B}\rangle = H$. If $H \neq \Lambda$, then we have that $\Lambda_0 := H \cap \Lambda$ is a hyperplane in Λ. Hence $\langle H \cap \mathcal{T}\rangle = \langle \Lambda_0 \cap \mathcal{T}\rangle = \Lambda_0$. Moreover, H meets each of the lines ℓ_i's in a point Q_i. Observe that not all of them can lie in Λ, because otherwise we would have $Q_i = P_i$ for every i and $H = \Lambda$. Therefore, there exists a point $Q_i \in (\ell_i \cap H) \setminus \langle \mathcal{T}\rangle$. This implies that $Q_i \in \langle H \cap \mathcal{B}\rangle \setminus \Lambda_0$ and we can conclude that $\langle H \cap \mathcal{B}\rangle = H$. □

Proposition 3.82 shows how to construct a cutting blocking set in $\mathrm{PG}(k,q)$ which contains a copy of a cutting blocking set $\mathcal{T}$ in $\mathrm{PG}(k-1,q)$. This is achieved by adding $(q-1)k+1$ points to $\mathcal{T}$. Moreover, among cutting blocking sets containing a copy of a smaller cutting blocking set (of codimension 1), the construction of Proposition 3.82 is optimal, as shown by the following result.

Proposition 3.83. Let $\mathcal{B} \subseteq \mathrm{PG}(k,q)$ be a cutting blocking set such that it contains (an isomorphic image of) a cutting blocking set $\mathcal{B}'$ of $\mathrm{PG}(k-1,q)$. Then

$$|\mathcal{B}| \geq |\mathcal{B}'| + (q-1)k + 1.$$

Proof. Let $\mathcal{B}$ be a cutting blocking set in $\mathrm{PG}(k,q)$ and suppose it contains a copy $\mathcal{B}'$ of a cutting blocking set in $\mathrm{PG}(k-1,q)$. Then $\mathcal{B}'$ is contained in a hyperplane H. By the correspondence between linear codes and projective systems, we have

$$d \leq |\mathcal{B}| - |\mathcal{B} \cap H| \leq |\mathcal{B}| - |\mathcal{B}'|.$$

Combining this with Theorem 3.44 we obtain the desired inequality. □

Remark 3.84. Proposition 3.83 shows that the inductive construction from Proposition 3.82 gives rise to a cutting blocking set that is minimal among all the cutting blocking sets containing a given cutting blocking set of codimension 1. It is interesting to observe that starting from $\mathrm{PG}(1,q)$ and iterating this construction k times, one obtains the tetrahedron, which is, therefore, minimal among the cutting blocking sets in $\mathrm{PG}(k-1,q)$ containing an isomorphic copy of a cutting blocking set of $\mathrm{PG}(i,q)$ for each $i \leq k-2$. Note that its cardinality is $\sim \frac{1}{2}qk^2$ for k large.

All of this seems to suggest that in order to obtain cutting blocking sets in $\mathrm{PG}(k-1,q)$ of size $m(k,q)$ (or at least linear in k) one should look at sets that do not contain (isomorphic copies of) smaller cutting blocking sets.

3.7.4 Explicit Constructions of Short Minimal Codes

In this final subsection we combine the results obtained so far to construct minimal codes of short length. To our best knowledge, these constructions produce the shortest known minimal codes, for infinitely many dimensions and field sizes. In particular, the construction applies to all those pairs (k,q) for which the rational normal tangent set of [68] cannot be constructed in $\mathrm{PG}(k-1,q)$.

Construction 4. Assume that $k=2t$, for some $t\in\mathbb{N}_{\geq 1}$. We use the construction from Theorem 3.75, selecting $t^2=\frac{k^2}{4}$ disjoint lines from a linespread. The union of these t^2 lines is a cutting blocking set in $\mathrm{PG}(k-1,q)$, and we denote the corresponding code by $\mathcal{C}_{k,q}$.

Proposition 3.85. The code $\mathcal{C}_{k,q}$ of Construction 4 is minimal with parameters $[(q+1)\frac{k^2}{2},k,q(k-1)]_q$.

Proof. The minimality of $\mathcal{C}_{k,q}$ trivially follows from the fact that the associated projective system is a cutting blocking set; see Theorem 3.75. The length of the code $\mathcal{C}_{k,q}$ coincides with the cardinality of the cutting blocking set, which is $(q+1)\frac{k^2}{4}$. Therefore it remains to show that $d=q(k-1)$. By the correspondence between projective systems and linear codes and Definition 2.10, we have that $d=n-s=(q+1)t^2-s$, where $k=2t$ and

$$s:=\max\{|\bar{H}\cap\bar{\mathcal{T}}| \,:\, \bar{H}\subseteq\mathrm{PG}(k-1,q),\, \dim(\bar{H})=k-2\},$$

where $\bar{\mathcal{T}}$ is the projectivization of the set $\mathcal{T}$ defined in Theorem 3.75. We switch to vector notation and let $\mathcal{A}_{\ell,m}=\{\bar{\varphi}(P_\ell),\bar{\varphi}(P_m),\bar{\varphi}(Q_{\ell,m,i_{\ell,m}}),\bar{\varphi}(Q_{\ell,m,j_{\ell,m}})\}$ for all $1\leq\ell<m\leq t$. Let H be a hyperplane of $\mathbb{F}_q^k=\mathbb{F}_q^{2t}$. Define the set $H_{\mathcal{T}}:=\{i:\bar{\varphi}(P_i)\subseteq H\}$ and the integers $a:=|H_{\mathcal{T}}|$ and $a_{\ell,m}:=|\{A\in\mathcal{A}_{\ell,m}:A\subseteq H\}|$ for $1\leq\ell<m\leq t$. Moreover, let b denote the number of lines forming $\bar{\mathcal{T}}$ that are fully contained in the projectivization $\bar{H}$ of H. Since each of the lines forming $\bar{\mathcal{T}}$ either intersects $\bar{H}$ in a point, or it is contained in $\bar{H}$, we have

$$s=(q+1)b+t^2-b=qb+t^2. \tag{3.17}$$

Therefore, finding the maximum of s is the same as finding the maximum value of b. Now observe that a cannot be equal to t, as otherwise H would contain a basis of $\mathbb{F}_q^{2t}$. Moreover, we have that $a_{\ell,m}\in\{0,1,4\}$. Indeed, by construction, any two subspaces in $\mathcal{A}_{\ell,m}$ span the same 4-dimensional subspace, and if H contains two of them, then it contains all of them. It is readily seen that we have

$$\begin{aligned} b &= a + \sum_{\substack{\ell,m\in H_{\mathcal{T}},\\ \ell<m}} (a_{\ell,m}-2) + \sum_{\substack{\ell\in H_{\mathcal{T}},m\notin H_{\mathcal{T}},\\ \ell<m}} (a_{\ell,m}-1) \\ &\quad + \sum_{\substack{\ell\notin H_{\mathcal{T}},m\in H_{\mathcal{T}},\\ \ell<m}} (a_{\ell,m}-1) + \sum_{\substack{\ell,m\notin H_{\mathcal{T}},\\ \ell<m}} (a_{\ell,m}) \\ &= a + \sum_{\substack{\ell,m\in H_{\mathcal{T}},\\ \ell<m}} 2 + \sum_{\substack{\ell,m\notin H_{\mathcal{T}},\\ \ell<m}} (a_{\ell,m}) \le a + \binom{a}{2} + \binom{t-a}{2} = a^2 + \binom{t-a}{2} =: f_t(a), \end{aligned}$$

where the second equality and the inequality both follow from the fact that $a_{\ell,m}$ can only be equal to $0, 1$ or 4. The function f_t is a quadratic polynomial in a with second derivative equal to $3 > 0$. Hence, the maximum in the interval $[0, t-1]$ is attained in one of the two interval extremes. One can see that this happens when $a = t-1$, from which $b \le (t-1)^2$. Finally, combining this with (3.17) we have $s = qb + t^2 \le q(t-1)^2 + t^2 = (q+1)t^2 - q(2t-1)$ and $d \ge (q+1)t^2 - s = q(2t-1) = q(k-1)$.

On the other hand, we can take any hyperplane H' containing $\bar{\varphi}(P_i)$, for each $i \in [t-1]$. The projectivization of such a hyperplane contains exactly $b = (t-1)^2$ lines forming $\bar{\mathcal{T}}$, and therefore $n - |\bar{H}' \cap \bar{\mathcal{T}}| = q(k-1)$. □

Example 3.86. Let $k = 6$ and take the cutting blocking set obtained in Example 3.77. This is a cutting blocking set arising from Construction 4. When $q = 2$, we take γ to be a root of $x^2 + x + 1$ and obtain a minimal $[27, 6]_2$ code $\mathcal{C}_{6,2}$ whose generator matrix is

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

Our second construction combines Theorem 3.75 with the concept of a Baer subplane.

Construction 5. Assume that $k = 3t$ for some $t \in \mathbb{N}$ and that q is a square. We first use the construction from Theorem 3.75 by selecting $t^2 = \frac{k^2}{9}$ disjoint planes from a 2-spread. Then we choose two disjoint Baer subplanes in each of these planes. The union of the selected $2t^2$ Baer subplanes is a cutting blocking set in $\mathrm{PG}(k-1, q)$, and we denote the corresponding code by $\mathcal{D}_{k,q}$.

Proposition 3.87. The code $\mathcal{D}_{k,q}$ of Construction 5 is a minimal code of parameters $[(q+\sqrt{q}+1)\frac{2k^2}{9}, k, d]_q$, where $d \ge q\big(\frac{4}{3}k - 2\big)$.

Proof. The minimality of $\mathcal{D}_{k,q}$ trivially follows from the fact that the associated projective system $\mathcal{B}$ is a cutting blocking set (Theorem 3.75 and Proposition 3.78). The length of the code $\mathcal{D}_{k,q}$ coincides with the cardinality of the cutting blocking set, which is $(q+\sqrt{q}+1)\frac{2k^2}{9}$. We only need to prove that $d \geq q\left(\frac{4}{3}k-2\right)$. We let $k=3t$ and proceed as before, finding an upper bound on

$$s := \max\{|\bar{H} \cap \mathcal{B}| \,:\, \bar{H} \subseteq \mathrm{PG}(k-1,q),\, \dim(\bar{H}) = k-2\}.$$

Observe that $\mathcal{B}$ is obtained by first forming the cutting blocking set $\bar{\mathcal{T}}$ as in Theorem 3.75, which is the union of t^2 planes $\Lambda_1, \dots, \Lambda_{t^2}$, and then selecting two disjoint Baer subplanes $\mathcal{B}_{i,1}, \mathcal{B}_{i,2}$ in each Λ_i. Let $\bar{H}$ be a hyperplane in $\mathrm{PG}(k-1,q)$ and let b denote the number of planes Λ_i that are fully contained in $\bar{H}$. With this notation, we have

$$\begin{aligned} |\mathcal{B} \cap \bar{H}| &= 2b(q+\sqrt{q}+1) + \sum_{i\,:\,\Lambda_i \not\subseteq \bar{H}} |\mathcal{B}_{i,1} \cap \bar{H}| + |\mathcal{B}_{i,2} \cap \bar{H}| \\ &\leq 2(q+\sqrt{q}+1)b + 2(\sqrt{q}+1)(t^2-b), \end{aligned} \tag{3.18}$$

where the last inequality follows from the fact that a hyperplane $\bar{H}$ meets a Baer subplane in either 1, $\sqrt{q}+1$ or $q+\sqrt{q}+1$ points. Moreover, arguing as in the proof of Proposition 3.85, one proves that $b \leq (t-1)^2$. Combining this with (3.18) we obtain that $s \leq 2(q+\sqrt{q}+1)t^2 - 2q(2t-1)$ and finally $d = n - s \geq q\left(\frac{4}{3}k-2\right)$. □

We conclude with a remark that summarizes the code lengths obtained from the constructions and results of this section.

Remark 3.88. For every positive integer k and every prime power q, we have provided explicit constructions of minimal $[n_{k,q}, k]_q$ codes with $n_{k,q}$ equal to

$$\begin{cases} \frac{1}{4}(q+1)k^2 & \text{if } k \equiv 0 \mod 2, \\ \frac{2}{9}(q+\sqrt{q}+1)k^2 & \text{if } k \equiv 0 \mod 3 \text{ and } q \text{ is a square}, \\ \frac{2}{9}(q+\sqrt{q}+1)(k-1)^2 + (q-1)(k-1)+1 & \text{if } k \equiv 1 \mod 3 \text{ and } q \text{ is a square}, \\ \frac{1}{4}(q+1)(k+1)^2 - (2k+q-2) & \text{otherwise.} \end{cases}$$

The first length is given by Construction 4, the second length is given by Construction 5, and the last two lengths are obtained by combining Proposition 3.82 with these two constructions. It is easy to see that the minimum distance d of any code obtained using Proposition 3.82 meets the bound of Theorem 3.44 with equality, i.e., $d = (q-1)(k-1)+1$.

Chapter 4

Linear Cutting Blocking Sets and Minimal Codes in the Rank Metric

This chapter contains the results provided in [6], by Alfarano, Borello, Neri and Ravagnani. Also in this case, we decided to add some proofs that are not contained in the original paper.

4.1 The Geometry of Rank-Metric Codes

In this section we study the geometric structure of rank-metric codes and their connection with the theory of q-systems, introducing fundamental tools that will be needed later. We also describe one-weight and simplex codes in the rank metric.

4.1.1 Geometric Characterization of Rank-Metric Codes

We start by introducing the natural analogue of the notion of "nondegenerate" code in the rank-metric setting.

Definition 4.1. An $[n,k]_{q^m/q}$ rank-metric code $\mathcal{C}$ is (**rank-**)**nondegenerate** if $\sigma^{\mathrm{rk}}(\mathcal{C}) = \mathbb{F}_q^n$. We say that $\mathcal{C}$ is (**rank-**)**degenerate** if it is not nondegenerate. Moreover, we call $\dim(\sigma^{\mathrm{rk}}(\mathcal{C}))$ the **effective length** of the code $\mathcal{C}$.

Proposition 4.2. Let $\mathcal{C} \subseteq \mathbb{F}_{q^m}^n$ be a rank-metric code. The following are equivalent.

1. $\mathcal{C}$ is rank-nondegenerate.
2. For every $A \in \mathrm{GL}_n(q)$, the code $\mathcal{C} \cdot A$ is Hamming-nondegenerate.
3. The $\mathbb{F}_q$-span of the columns of any generator matrix of G has $\mathbb{F}_q$-dimension n.
4. $d(\mathcal{C}^{\perp}) \geq 2$.

Proof. $\underline{(1) \Rightarrow (2)}$: Assume that $\mathcal{C} \cdot A$ is Hamming-degenerate for some $A \in \mathrm{GL}_n(q)$. Then there exists $1 \leq i \leq n$ with $(vA)_i = 0$ for all $v \in \mathcal{C}$. In particular, $\sigma^{\mathrm{rk}}(vA) \subseteq V := \langle e_j \,:\, j \neq i \rangle$. Using Proposition 2.11, we see that $\sigma^{\mathrm{rk}}(\mathcal{C})$ is contained in an $(n-1)$-dimensional subspace of $\mathbb{F}_q^n$, hence $\mathcal{C}$ is rank-degenerate.

$\underline{(2) \Rightarrow (4)}$: Let $\Gamma := \{\gamma_1, \ldots, \gamma_m\}$ be an $\mathbb{F}_{q^m}/\mathbb{F}_q$ basis. If $d(\mathcal{C}^\perp) = 1$, then there exists $v \in \mathcal{C}^\perp$ with $\mathrm{rk}(\Gamma(v)) = 1$. Therefore there exists $A \in \mathrm{GL}_n(q)$ with $v = (0, \ldots, 0, 1) \in (\mathcal{C} \cdot A)^\perp$. Thus $\mathcal{C} \cdot A$ is a Hamming-degenerate code.

$\underline{(4) \Rightarrow (1)}$: A rank-degenerate code $\mathcal{C}$ is equivalent to a code $\mathcal{C} \cdot A$ in which all codewords have a 0 in the last component. Hence $(0, \ldots, 0, 1) \in (\mathcal{C} \cdot A)^\perp$ and $d(\mathcal{C}^\perp) = 1$.

$\underline{(2) \Rightarrow (3)}$: Let G be a generator matrix of $\mathcal{C}$. Since $\mathcal{C} \cdot A$ is Hamming-nondegenerate for any $A \in \mathrm{GL}_n(q)$, the columns of G are linearly independent over $\mathbb{F}_q$. This implies that $n = \dim(\sigma^{\mathrm{rk}}(\mathcal{C}))$ is equal to the dimension of the $\mathbb{F}_q$-space of the columns of G.

$\underline{(3) \Rightarrow (1)}$: This immediately follows from the definition of rank-nondegenerate code. □

Remark 4.3. By Proposition 4.2, a degenerate code can be isometrically embedded in $\mathbb{F}_{q^m}^{n'}$, where $n' = \dim(\sigma^{\mathrm{rk}}(\mathcal{C}))$.

The following result shows that the parameters of a nondegenerate code must obey certain constraints.

Proposition 4.4. (see [91, Corollary 6.5]) Let $\mathcal{C}$ be an $[n,k]_{q^m/q}$ nondegenerate rank-metric code. Then $n \leq km$.

Proof. Let $\{c_1, \ldots, c_k\}$ be a set of generators for $\mathcal{C}$. Then, by Proposition 2.14, $\sigma^{\mathrm{rk}}(\mathcal{C})$ is generated by $\sigma^{\mathrm{rk}}(c_i)$ for $i = 1, \ldots, k$. Since $\dim(\sigma^{\mathrm{rk}}(c_i)) \leq m$ for all i and $\sigma^{\mathrm{rk}}(\mathcal{C}) = \mathbb{F}_q^n$, we conclude that $n \leq km$. □

Our next move is to identify geometric objects able to capture the structure of rank-metric codes. We re-formulate the definition of q-analogue of a projective system proposed in [129] as follows.

Definition 4.5. An $[n,k,d]_{q^m/q}$ **system** is an n-dimensional $\mathbb{F}_q$-space $\mathcal{U} \subseteq \mathbb{F}_{q^m}^k$ with the properties that $\langle \mathcal{U} \rangle_{\mathbb{F}_{q^m}} = \mathbb{F}_{q^m}^k$ and

$$d = n - \max\left\{\dim_{\mathbb{F}_q}(\mathcal{U} \cap H) \,:\, H \text{ is an } \mathbb{F}_{q^m}\text{-hyperplane of } \mathbb{F}_{q^m}^k\right\}. \tag{4.1}$$

Note that (4.1) can be re-written as

$$\min\left\{\dim_{\mathbb{F}_q}(\mathcal{U} + H) \,:\, H \text{ is an } \mathbb{F}_{q^m}\text{-hyperplane in } \mathbb{F}_{q^m}^k\right\} - m(k-1).$$

When the parameters are not relevant, we simply call such an object a q**-system**.

Two $[n,k]_{q^m/q}$ systems $\mathcal{U},\mathcal{V}$ are said to be **equivalent** if there exists an $\mathbb{F}_{q^m}$-isomorphism $\phi:\mathbb{F}_{q^m}^k\to\mathbb{F}_{q^m}^k$ such that $\phi(\mathcal{U})=\mathcal{V}$.

The following simple result is a geometric formulation of one of the *Standard Equations* (stated in our context), which will be of great help throughout the chapter. Recall that for integers $a\geq b\geq 0$ and a prime power Q, the symbol

$$\binom{a}{b}_Q$$

denotes the number of b-dimensional subspaces of an a-dimensional space over $\mathbb{F}_Q$. This quantity is called a **Gaussian binomial coefficient**.

Lemma 4.6. (The Standard Equations) Let $\mathcal{U}$ be an $[n,k]_{q^m/q}$ system and let Λ_r be the set of all r-dimensional $\mathbb{F}_{q^m}$-subspaces of $\mathbb{F}_{q^m}^k$. We have

$$\sum_{H\in\Lambda_r}|H\cap(\mathcal{U}\setminus\{0\})|=(q^n-1)\binom{k-1}{r-1}_{q^m}. \tag{4.2}$$

Proof. Every vector in $\mathcal{U}\setminus\{0\}$ belongs to exactly $\binom{k-1}{r-1}_{q^m}$ r-dimensional subspaces in Λ_r. Therefore,

$$\sum_{H\in\Lambda_r}|H\cap(\mathcal{U}\setminus\{0\})|=\sum_{u\in\mathcal{U}\setminus\{0\}}|\{H\in\Lambda_r\,:\,u\in H\}|=(q^n-1)\binom{k-1}{r-1}_{q^m},$$

which is the desired result. □

In the remainder of this section we describe the 1-to-1 correspondence between equivalence classes of nondegenerate $[n,k,d]_{q^m/q}$ codes and equivalence classes of $[n,k,d]_{q^m/q}$ systems. We denote the set of equivalence classes of nondegenerate $[n,k,d]_{q^m/q}$ codes by $\mathcal{C}[n,k,d]_{q^m/q}$, and the set of equivalence classes of $[n,k,d]_{q^m/q}$ systems by $\mathcal{U}[n,k,d]_{q^m/q}$. Next, we define a map

$$\Phi:\mathcal{C}[n,k,d]_{q^m/q}\to\mathcal{U}[n,k,d]_{q^m/q}$$

as follows: Given an equivalence class $[\mathcal{C}]\in\mathcal{C}[n,k,d]_{q^m/q}$, let $\Phi([\mathcal{C}])$ be the equivalence class of the $\mathbb{F}_q$-span of the columns of a generator matrix of $\mathcal{C}$. Vice versa, given an equivalence class $[\mathcal{U}]\in\mathcal{U}[n,k,d]_{q^m/q}$, fix an $\mathbb{F}_q$-basis $\{g_1,\ldots,g_n\}$ of $\mathcal{U}$ and let $\Psi([\mathcal{U}])$ be the equivalence class of the code generated by the matrix having the g_i's as columns. In Theorem 4.8 we will show that Φ and Ψ are the inverse of each other.

We recall that the minimum rank distance of a code $\mathcal{C}$ coincides with the minimum $\mathbb{F}_q$-dimension of the linear space generated over $\mathbb{F}_q$ by the entries of $v\in\mathcal{C}$. In particular, $d^{\mathrm{rk}}(\mathcal{C})\leq d^{\mathrm{H}}(\mathcal{C})$. More precisely, the rank of a vector can be rewritten as

$$\mathrm{rk}(v)=\min\{\mathrm{wt}^{\mathrm{H}}(vA)\,:\,A\in\mathrm{GL}_n(q)\}. \tag{4.3}$$

We will also repeatedly use the following simple fact: Let $V, H \subseteq W$ be nonzero finite dimensional vector spaces over $\mathbb{F}_q$ and let $\mathcal{B}$ be the set of $\mathbb{F}_q$-bases of V; then

$$\max\{|B \cap H| \,:\, B \in \mathcal{B}\} = \dim(V \cap H). \tag{4.4}$$

Finally, we will often use the following characterization of the rank of a vector.

Lemma 4.7. Let $\mathcal{C}$ be a nondegenerate $[n,k]_{q^m/q}$ code and let G be a generator matrix of $\mathcal{C}$. For any nonzero $v \in \mathbb{F}_{q^m}^k$ we have

$$\mathrm{rk}(vG) = n - \dim_{\mathbb{F}_q}(\mathcal{U} \cap \langle v \rangle^{\perp}), \tag{4.5}$$

where $\mathcal{U}$ is the $[n,k]_{q^m/q}$ system generated by the $\mathbb{F}_q$-span of the columns of G.

Proof. Using (4.3) we see that for all nonzero $u \in \mathbb{F}_{q^m}^k$ we have

$$\mathrm{rk}(uG) = \min\{\mathrm{wt}^{\mathrm{H}}(uGA) \,:\, A \in \mathrm{GL}_n(q)\} = \min\{n - |\{i \,:\, (GA)_i \in \langle u \rangle^{\perp}\}|\},$$

where $(GA)_i$ is the i-th column of GA and $\langle u \rangle^{\perp}$ is the dual of the 1-dimensional space generated by u. As A ranges over $\mathrm{GL}_n(q)$, the columns of GA range over all bases of $\mathcal{U}$. Therefore we conclude by the identity in (4.4). □

Note that previous lemma can be viewed as the q-analogue of the treatment of Hamming weights in [156]. The following result has already been shown in [129]. We include a complete proof for sake of completeness.

Theorem 4.8. The maps Φ and Ψ are well-defined and are the inverse of each other. In particular, they give a 1-to-1 correspondence between equivalence classes of nondegenerate $[n,k,d]_{q^m/q}$ rank-metric codes and equivalence classes of $[n,k,d]_{q^m/q}$ systems.

Proof. We prove a series of properties separately.

- $\Phi([\mathcal{C}])$ does not depend on the choice of the generator matrix G. Indeed, if G' is another generator matrix for $\mathcal{C}$ then there is an $\mathbb{F}_{q^m}$-linear map φ, such that $\varphi(G) = G'$. The same map sends the $\mathbb{F}_q$-columnspace of G into the $\mathbb{F}_q$-columnspace of G'.

- $\Phi([\mathcal{C}])$ does not depend on $\mathcal{C}$ but only on its equivalence class. To see this, let $\mathcal{C}'$ be a code linearly equivalent to $\mathcal{C}$, then there is a matrix $A \in \mathrm{GL}_n(q)$ such that $\mathcal{C}' = \mathcal{C} \cdot A$. Hence, if G is a generator matrix for $\mathcal{C}$, then GA is a generator matrix for $\mathcal{C}'$ and they have the same $\mathbb{F}_q$-columnspace. Hence, the map Φ does not depend on the choice of the representative.

- $\Phi([\mathcal{C}]) \in \mathcal{U}[n,k,d]_{q^m/q}$. To see this, let $[n',k',d']$ be the parameters of $\Phi([\mathcal{C}])$. We need to show that $(n,k,d) = (n',k',d')$. Since $\mathcal{C}$ has dimension k over $\mathbb{F}_{q^m}$ we have $k = k'$.

In order to prove that $n = n'$, we use the fact that $\mathcal{C}$ is nondegenerate by assumption. More precisely, let G be a generator matrix for $\mathcal{C}$. Since $\mathcal{C}$ is nondegenerate, by Proposition 4.2, $n = \dim(\sigma^{\mathrm{rk}}(\mathcal{C}))$ is equal to the dimension of the $\mathbb{F}_q$-space of the columns of G, that is n'.

Finally, denote by G a generator matrix of $\mathcal{C}$. By Lemma 4.7, for all nonzero $v \in \mathbb{F}_{q^m}^k$ we have

$$\mathrm{rk}(vG) = n - \dim_{\mathbb{F}_q}(\Phi([\mathcal{C}]) \cap \langle v\rangle^{\perp}).$$

As v ranges over the nonzero vectors in $\mathbb{F}_{q^m}^k$, $\langle v\rangle^{\perp}$ ranges over all $\mathbb{F}_{q^m}$-hyperplanes in $\mathbb{F}_{q^m}^k$. Therefore $d = d'$ by definition of d'.

- $\Psi([\mathcal{U}])$ does not depend on $\mathcal{U}$ but only on its equivalence class. To see this, assume $\mathcal{U}'$ is an $[n,k]_{q^m/q}$ system equivalent to $\mathcal{U}$, hence, there is an $\mathbb{F}_{q^m}$-isomorphism ϕ, such that $\phi(\mathcal{U}) = \mathcal{U}'$. In particular, if $\{g_1,\dots,g_n\}$ is a basis of $\mathcal{U}$ and $\{g_1',\dots,g_n'\}$ is a basis of $\mathcal{U}'$, then $\phi(\{g_1',\dots,g_n'\}) = \{g_1',\dots,g_n'\}$. In particular, let G be the matrix whose i-th column is given by g_i and G' be the matrix whose i-th column is given by g_i', then there is a matrix $A \in \mathrm{GL}_n(q)$, such that $G' = GA$. Hence, the rank-metric codes generated by G and G' are linearly equivalent. So, we conclude that Ψ does not depend on the choice of $\mathcal{U}$.

- $\Psi([\mathcal{U}]) \in \mathcal{C}[n,k,d]_{q^m/q}$. To see this, let $\{g_1,\dots,g_n\}$ be an $\mathbb{F}_q$-basis of $\mathcal{U}$ and let $\mathcal{C}$ be the $[n',k',d']$ code whose generator matrix G has g_i as i-th column. Then, obviously, the length of $\mathcal{C}$ is $n' = n$. The rows of G are linearly independent over $\mathbb{F}_{q^m}$ otherwise there is $x \in \mathbb{F}_{q^m}^k$ such that $xg_i^{\top} = 0$ for all i. Hence, x defines an hyperplane containing $\mathcal{U}$, which contradicts the fact that $\langle\mathcal{U}\rangle_{\mathbb{F}_{q^m}} = \mathbb{F}_{q^m}^k$. This ensures that the dimension k' of $\mathcal{C}$ is equal to k. For a matrix $G \in \mathbb{F}_{q^m}^{k\times n}$ we denote by G_i the i-th column of G. Now, for the distance, observe that for all $v \in \mathbb{F}_{q^m}^k$,

$$\begin{aligned} d' &= \min\{\mathrm{wt}^{\mathrm{H}}(vGA) \,:\, A \in \mathrm{GL}_n(q)\} \\ &= \min\{n - |\{i \,:\, (GA)_i \in \langle v\rangle^{\perp}\}|\} \\ &= n - \max\{\dim_{\mathbb{F}_q}(\mathcal{U} \cap H) \,:\, H \text{ is an } \mathbb{F}_{q^m}\text{-hyperplane in } \mathbb{F}_{q^m}^k\}, \end{aligned}$$

  where the last equality follows from Equation (4.4). Finally, since the g_i's are linearly independent over $\mathbb{F}_q$, $\mathcal{C}$ is nondegenerate by Proposition 4.2.

All of this establishes the desired result. □

We also observe that combining Lemma 4.7 with Remark 4.3 one obtains the following lower bound for the minimum distance of a rank-metric code.

Corollary 4.9. Let $\mathcal{C}$ be an $[n,k,d]_{q^m/q}$ code. Then

$$d \geq \dim_{\mathbb{F}_q}(\sigma^{\mathrm{rk}}(\mathcal{C})) - (k-1)m.$$

As an application of Theorem 4.8, we show that a nondegenerate rank-metric code always have a codeword of rank $\min\{n,m\}$. Note that this the largest rank a codeword can possibly have.

Notation 4.10. We denote by $w^{\mathrm{rk}}(\mathcal{C})$ the maximum rank of the codewords of a rank-metric code $\mathcal{C} \subseteq \mathbb{F}_{q^m}^n$.

Proposition 4.11. Let $\mathcal{C}$ be a nondegenerate $[n,k]_{q^m/q}$ code, then $w^{\mathrm{rk}}(\mathcal{C}) = \min\{n,m\}$. In particular, if $n=m$ then an $[n,k]_{q^n/q}$ code is nondegenerate if and only if $w^{\mathrm{rk}}(\mathcal{C}) = n$.

Proof. Since $w^{\mathrm{rk}}(\mathcal{C}) \leq m$, if $w^{\mathrm{rk}}(\mathcal{C}) = n$ then the statement is trivially true, so we may assume that $w^{\mathrm{rk}}(\mathcal{C}) < n$. Let $\mathcal{U}$ be any $[n,k]_{q^m/q}$ system associated with $\mathcal{C}$ via Theorem 4.8. By Lemma 4.7 we have that $\dim(H \cap \mathcal{U}) \geq n - w^{\mathrm{rk}}(\mathcal{C})$ for each $\mathbb{F}_{q^m}$-hyperplane H of $\mathbb{F}_{q^m}^k$. Denote by Λ the set of all $\mathbb{F}_{q^m}$-hyperplanes of $\mathbb{F}_{q^m}^k$. Then we have

$$(q^n-1)\binom{k-1}{1}_{q^m} = \sum_{H\in\Lambda} |H \cap (\mathcal{U} \setminus \{0\})| \geq (q^{n-w^{\mathrm{rk}}(\mathcal{C})}-1)\binom{k}{1}_{q^m},$$

where the first equality follows from Lemma 4.6. The above inequality is equivalent to

$$(q^n-1)(q^{(k-1)m}-1) \geq (q^{n-w^{\mathrm{rk}}(\mathcal{C})}-1)(q^{km}-1).$$

Dividing both sides by $(q^{(k-1)m}-1)$, we obtain

$$\begin{aligned} q^n - 1 &\geq (q^{n-w^{\mathrm{rk}}(\mathcal{C})}-1)\left(q^m + \frac{q^m-1}{q^{(k-1)m}-1}\right) \\ &= q^{n+m-w^{\mathrm{rk}}(\mathcal{C})} - q^m + \frac{(q^{n-w^{\mathrm{rk}}(\mathcal{C})}-1)(q^m-1)}{q^{(k-1)m}-1} \\ &\geq q^{n+m-w^{\mathrm{rk}}(\mathcal{C})} - q^m. \end{aligned}$$

Since $n - w^{\mathrm{rk}}(\mathcal{C}) \geq 1$, this implies $m \leq w^{\mathrm{rk}}(\mathcal{C})$. Since, clearly, $w^{\mathrm{rk}}(\mathcal{C}) \leq m$, then they must be equal. □

As an application of Proposition 4.11, we recover the characterization of optimal $\mathbb{F}_{q^m}$-linear anticodes given in [130, Theorem 18] with a new and concise proof.

Corollary 4.12. Let $\mathcal{C}$ be an $[n,k]_{q^m/q}$ code with $k = w^{\mathrm{rk}}(\mathcal{C})$. If $m \geq n$, then $\mathcal{C}$ has a basis made of vectors with entries in $\mathbb{F}_q$.

Proof. We prove the result by induction on $n-k$. The case $n=k$ is immediate. Now assume that $n \geq k+1$ and that $\mathcal{C}$ has $k = w^{\mathrm{rk}}(\mathcal{C})$. Fix a generator matrix G for $\mathcal{C}$. Since $k < n$, by Proposition 4.11 there exists $A \in \mathrm{GL}_n(q)$ such that the last column of $G \cdot A$ is zero. Denote by G' the matrix obtained from $G \cdot A$ by deleting its last column. The code generated by G' has

$k = w^{\mathrm{rk}}(\mathcal{C})$ and therefore, by the induction hypothesis, has a basis made of vectors with entries in $\mathbb{F}_q$. This means that there exists $B \in \mathrm{GL}_k(q)$ such that BG' (and thus BGA) has entries in $\mathbb{F}_q$. Therefore $BG = BGAA^{-1}$ has entries in $\mathbb{F}_q$ as well. □

We conclude this subsection by surveying the connection between the generalized rank weights of an $[n,k]_{q^m/q}$ rank-metric code and any corresponding $[n,k]_{q^m/q}$ system. The definitions given here are equivalent to those of [129]. We denote the set of Frobenius-closed subspaces of $\mathbb{F}_{q^m}^n$ by $\Lambda_q(n,m)$, that is,

$$\Lambda_q(n,m) := \left\{\mathcal{V} \leq \mathbb{F}_{q^m}^n \,:\, \theta(\mathcal{V}) = \mathcal{V}\right\},$$

where $\theta : x \longmapsto x^q$ is the q-Frobenius automorphism in $\mathbb{F}_{q^m}$ (extended component-wise to vectors). It is known that $\Lambda_q(n,m)$ corresponds to the set of subspaces of $\mathbb{F}_{q^m}^n$ that have a basis of vectors in $\mathbb{F}_q^n$; see [78, Theorem 1].

Definition 4.13. Let $\mathcal{C}$ be an $[n,k]_{q^m/q}$ code. For every $r = 1,\ldots,k$, the **r-th generalized rank weight** of $\mathcal{C}$ is the integer

$$d_r(\mathcal{C}) := \min\left\{\dim(\mathcal{V}) \,:\, \mathcal{V} \in \Lambda_q(n,m),\ \dim(\mathcal{V} \cap \mathcal{C}) \geq r\right\}.$$

The following result was shown in [129]. We state and prove it here for completeness.

Theorem 4.14. Let $\mathcal{C}$ be an $[n,k,d]_{q^m/q}$ nondegenerate code and let $\mathcal{U}$ be any $[n,k,d]_{q^m/q}$ system associated to $\mathcal{C}$. For any $r = 1,\ldots,k$ the r-th generalized rank weight is given by

$$\begin{aligned} d_r(C) &= n - \max\left\{\dim_{\mathbb{F}_q}(\mathcal{U} \cap H) \,:\, H \text{ is an } \mathbb{F}_{q^m}\text{-subspace of codim. } r \text{ of } \mathbb{F}_{q^m}^k\right\} \\ &= \min\left\{\dim_{\mathbb{F}_q}(\mathcal{U} + H) \,:\, H \text{ is an } \mathbb{F}_{q^m}\text{-subspace of codim. } r \text{ of } \mathbb{F}_{q^m}^k\right\} - m(k-r). \end{aligned}$$

In particular, the minimum rank distance of $\mathcal{C}$ is given by

$$d = n - \max\{\dim_{\mathbb{F}_q}(\mathcal{U} \cap H) \,:\, H \text{ is an } \mathbb{F}_{q^m}\text{-hyperplane of } \mathbb{F}_{q^m}^k\}.$$

In order to prove Theorem 4.14, we first recall the notion of generalized Hamming weight. Given an $[n,k]_{q^m/q}$ nondegenerate code $\mathcal{C}$, for every $r = 1,\ldots,k$, the **r-th generalized Hamming weight** of $\mathcal{C}$ is defined as

$$d_r^{\mathrm{H}}(\mathcal{C}) = \min\{|\sigma^{\mathrm{H}}(V)| \,:\, V \subseteq \mathcal{C}, \dim(V) = r\}.$$

It is easy to see that for an $[n,k]_{q^m/q}$ rank-metric code $\mathcal{C}$ one has

$$d_r(\mathcal{C}) = \min\{d_r^{\mathrm{H}}(\mathcal{C} \cdot A) \,:\, A \in \mathrm{GL}_n(q)\}; \tag{4.6}$$

see e.g. [111, Theorem 2]. Recall also the following well-known result.

Lemma 4.15 (see [156, Theorem 1.1.14])**.** *Let $\mathcal{C}$ be an $[n,k]_{q^m/q}$ code and G be a generator matrix for $\mathcal{C}$. Then*

$$d_r^{\mathrm{H}}(\mathcal{C}) = \min\{n - |\{i : G_i \in H\}| : H \leq \mathbb{F}_{q^m}^k,\ \dim H \leq k-r\},$$

where G_i denotes the i-th column of G.

Proof of Theorem 4.14. Let G be a generator matrix of $\mathcal{C}$. Then, by the previous Lemma and Equation (4.6) we obtain that

$$d_r(\mathcal{C}) = n - \max\{|\{i : (GA)_i \in H\}| : A \in \mathrm{GL}_n(q),\ H \leq \mathbb{F}_{q^m}^k,\ \dim H \leq k-r\}.$$

Let $\mathcal{U}$ be the $\mathbb{F}_q$-span of the columns of G, i.e. $\mathcal{U}$ is an $[n,k]_{q^m/q}$ system corresponding to the equivalence class of $\mathcal{C}$. Note that, by Equation (4.4), for a fixed $H \subseteq \mathbb{F}_{q^m}^k$ with $\dim H \leq k-r$ we have that

$$\max\{|\{i : (GA)_i \in H\}| : A \in \mathrm{GL}_n(q)\} = \dim_{\mathbb{F}_q}(\mathcal{U} \cap H).$$

This concludes the proof. □

4.1.2 Simplex and One-Weight Codes in the Rank Metric

In this subsection we use the geometric approach on rank-metric codes to define simplex codes as the natural counterpart of simplex Hamming-metric codes. In particular, this allows to characterize one-weight codes in the rank metric, recovering the results of [129] in this context.

Lemma 4.16. *Let a,b,c,d be positive integers such that $a \leq b$ and $c \leq d$, and let $t \geq 2$ be an integer. Suppose that $(t^a-1)(t^b-1) = (t^c-1)(t^d-1)$. Then $a=c$ and $b=d$.*

Proof. By contradiciton, assume that $(a,b) \neq (c,d)$. Moreover, without loss of generality we can assume $a \leq c$. Since $(a,b) \neq (c,d)$, then we need to have $a < c \leq d$ (if $a=c$ clearly also $b=d$). Moreover, we also have that $b > a$, otherwise the equality is not possible. By expanding the equality $(t^a-1)(t^b-1)-(t^c-1)(t^d-1)=0$, and dividing by t^a, we get

$$t^b - t^{b-a} - t^{c+d-a} + t^{c-a} + t^{d-a} - 1 = 0.$$

All the exponents of t appearing above are positive integers, hence we get a contradiction, since the left hand side is equal to $-1 \mod t$. □

Proposition 4.17. *Let $k \geq 2$, let $\mathcal{C}$ be a $[km,k]_{q^m/q}$ code, and let G be a generator matrix of $\mathcal{C}$. The following are equivalent.*

1. *$\mathcal{C}$ is nondegenerate.*

2. *The $\mathbb{F}_q$-span of the columns of G is $\mathbb{F}_{q^m}^k$.*

3. $\mathcal{C}$ is a one-weight code (with minimum distance m).

4. $d(\mathcal{C}^\perp) > 1$.

5. $d(\mathcal{C}^\perp) = 2$.

6. $\mathcal{C}$ is linearly equivalent to a code whose generator matrix is

$$\left(\, I_k \; \middle| \; \alpha I_k \; \middle| \; \cdots \; \middle| \; \alpha^{m-1} I_k \, \right), \tag{4.7}$$

where $\alpha \in \mathbb{F}_{q^m}$ satisfies $\mathbb{F}_{q^m} = \mathbb{F}_q(\alpha)$.

Proof. (1) ⇒ (2): If $\mathcal{C}$ is nondegenerate, then its support has dimension km, which is also the dimension of the associated $[km, k]_{q^m/q}$ system.

(2) ⇒ (6): The code $\mathcal{C}$ has effective length km and $\mathcal{U} = \mathbb{F}_{q^m}^k$ as corresponding $[n, k]_{q^m/q}$ system. Hence, $\mathcal{U}$ has a basis given by $\mathcal{B} = \{\alpha^i e_j : 0 \leq i \leq m-1, 0 \leq j \leq k-1\}$. Thus, $\mathcal{C}$ belongs to the same equivalence class of the code whose generator matrix is (4.7).

(6) ⇒ (5): Without loss of generality, we can assume that $\mathcal{C}$ is the code whose generator matrix G is (4.7). Since $\mathcal{C}$ is nondegenerate, by Proposition 4.2 we have $d(\mathcal{C}^\perp) > 1$. Moreover, G is a parity check matrix for $\mathcal{C}^\perp$ and from that it is easy to see that the vector $v = \alpha e_1 - e_{k+1}$ belongs to $\mathcal{C}^\perp$ and has rank weight 2. Thus $d(\mathcal{C}^\perp) = 2$.

(5) ⇒ (4): Clear.

(4) ⇒ (1): The equivalence between (4) and (1) holds for every rank-metric code, by Proposition 4.2.

(2) ⇒ (3): Let $\mathcal{C}$ be the $[n, k]_{q^m/q}$ code generated by G. By hypothesis, the $[n, k]_{q^m/q}$ system corresponding to $\mathcal{C}$ is $\mathcal{U} = \mathbb{F}_{q^m}^k$. Moreover, for every nonzero $v \in \mathbb{F}_{q^m}^k$, by (4.5) it holds that

$$\mathrm{rk}(vG) = km - \dim_{\mathbb{F}_q}(\mathcal{U} \cap \langle v \rangle^\perp) = km - (k-1)m = m.$$

(3) ⇒ (1): Let $\mathcal{C}$ be a $[km, k]_{q^m/q}$ code. Let $n \leq km$ be its effective length, that is, $n = \dim(\sigma^{\mathrm{rk}}(\mathcal{C}))$. This means that $\mathcal{C}$ can be isometrically embedded in $\mathbb{F}_{q^m}^n$, obtaining a code $\mathcal{C}'$. Then $\mathcal{C}'$ is a nondegenerate $[n, k]_{q^m/q}$ code with the same weight distribution as $\mathcal{C}$. In particular, $\mathcal{C}'$ is a one-weight code as well. Fix a generator matrix for $\mathcal{C}'$ and consider the associated $[n, k]_{q^m/q}$ system, which we call $\mathcal{U}$. Since $\mathcal{C}'$ is a one-weight code, we have $|H \cap (\mathcal{U} \setminus \{0\})| = (q^a - 1)$ for some a, for every $\mathbb{F}_{q^m}$-hyperplane of $\mathbb{F}_{q^m}^k$. Therefore, if we denote by Λ the set of all the $\mathbb{F}_{q^m}$-hyperplanes in $\mathbb{F}_{q^m}^k$, we have

$$\sum_{H \in \Lambda} |H \cap (\mathcal{U} \setminus \{0\})| = \binom{k}{1}_{q^m} (q^a - 1).$$

Moreover, by applying Equation (4.2) to the right-hand side, we obtain

$$(q^{km}-1)(q^{a}-1)=(q^{(k-1)m}-1)(q^{n}-1).$$

By Lemma 4.16, we have $a=(k-1)m$ and $n=km$. Hence $\mathcal{C}$ itself is nondegenerate. □

We call **simplex rank-metric code** a code that satisfies any of the equivalent conditions in Proposition 4.17. Note that Proposition 4.17 also implies the following characterization of one-weight codes in the rank metric, which is the analogue of the main result of [43].

Corollary 4.18 (Classification of one-weight rank-metric codes)**.** Let $k \geq 2$ and let $\mathcal{C}$ be an $[n,k,d]_{q^m/q}$ one-weight code. Then, the effective length of $\mathcal{C}$ is km and $d=m$. That is, $\mathcal{C}$ is isometric to a simplex rank-metric $[km,k,m]_{q^m/q}$ code.

Proof. If $n \leq km$, as shown in the proof of Proposition 4.17, it has to be $n=km$ and $\mathcal{C}$ is a simplex rank-metric code. Assume now $n>km$. Since the effective length of an $[n,k]_{q^m/q}$ is always at most km, then we can isometrically embed $\mathcal{C}$ in a $[km,k]_{q^m/q}$ code $\mathcal{C}'$, with the same weight distribution. By Proposition 4.17, $\mathcal{C}'$ has to be a simplex rank-metric code. □

We remark that there is a strong analogy between simplex rank-metric codes and their homonyms in the Hamming metric, which is confirmed by both their weight distributions and by geometric characterization.

Indeed, by Corollary 4.18, simplex rank-metric codes are the only nondegenerate one-weight codes in the rank-metric, just like simplex codes in the Hamming metric, up to repetition. In fact, simplex codes in the Hamming metric are the only projective one-weight codes (where **projective** means that no two columns of one, and thus any, generator matrix are linearly dependent).

From a geometric point of view, simplex codes in the Hamming metric have a generator matrix whose columns are formed by all the points of $\mathrm{PG}(k-1,q)$. In the rank-metric, simplex codes are associated to the $[km,k]_{q^m/q}$ system $\mathbb{F}_{q^m}^k$, which is the natural analogue in the rank metric.

We conclude by observing that a definition of simplex code in the rank metric has been recently proposed in [112] (the definition has been given for sum-rank-metric codes, which specialize to rank-metric codes by taking a single matrix block). The simplex codes defined in [112] are different from the simplex codes considered in this thesis. For example, one can check they are not one-weight in general. From a geometric viewpoint, the definition of simplex code proposed in this thesis appears therefore more natural.

4.2 From Rank-Metric to Hamming-Metric Codes

In this section we explore various connections between codes in the rank and in the Hamming metric. In particular, we show how to construct a Hamming-metric code from a rank-metric one and describe how the parameters of the two codes relate to each other.

4.2.1 Linear Sets

Linear sets in finite geometry can be viewed as a generalizations of subgeometries. Their name was first proposed by Lunardon in [107], where linear sets are used for special constructions of blocking sets. The very first example of linear set is probably due to Brouwer and Wilbrink; see [46]. The interested reader is referred to [125] for an in-depth treatment of linear sets.

A special family of linear sets, which is of particular interest for this thesis, is the one of scattered linear sets introduced by Blokhuis and Lavrauw in [40]. Recently, Sheekey and Van de Voorde observed a connection between scattered linear sets and rank-metric codes with optimal parameters in [142, 144]; see [126] for a survey on this topic.

Definition 4.19. Let $\mathcal{U}$ be an $[n,k]_{q^m/q}$ system. The $\mathbb{F}_q$-**linear set** in $\mathrm{PG}(k-1,q^m)$ of rank n associated to $\mathcal{U}$ is the set

$$L_{\mathcal{U}} := \{\langle u\rangle_{\mathbb{F}_{q^m}} \,:\, u \in \mathcal{U}\setminus\{0\}\},$$

where $\langle u\rangle_{\mathbb{F}_{q^m}}$ denotes the projective point corresponding to u.

Let $\Lambda = \mathrm{PG}(W,\mathbb{F}_{q^m})$ be the projective subspace corresponding to the $\mathbb{F}_{q^m}$-subspace W of $\mathbb{F}_{q^m}^k$. We define the **weight** of Λ in $L_{\mathcal{U}}$ as the integer

$$\mathrm{wt}_{\mathcal{U}}(\Lambda) := \dim_{\mathbb{F}_q}(\mathcal{U}\cap W).$$

If Λ is an hyperplane, that is, if $\Lambda = \mathrm{PG}(W,\mathbb{F}_{q^m})$ with $W = \langle v\rangle^{\perp}$ for some nonzero $v \in \mathbb{F}_{q^m}^k$, then $\mathrm{wt}_{\mathcal{U}}(\Lambda) = n - \mathrm{rk}(vG)$, where G is a $k\times n$ matrix associated to $\mathcal{U}$; see Lemma 4.7. Observe moreover that for a point $P \in \mathrm{PG}(k-1,q^m)$ we have that $P \in L_{\mathcal{U}}$ if and only if $\mathrm{wt}_{\mathcal{U}}(P) \geq 1$.

Remark 4.20. The original definition of linear sets does not assume the space $\mathcal{U}$ to be a $[n,k]_{q^m/q}$ system, i.e., that $\langle\mathcal{U}\rangle_{\mathbb{F}_{q^m}}$ is the whole space $\mathbb{F}_{q^m}^k$. However, if $\dim_{\mathbb{F}_{q^m}}(\langle\mathcal{U}\rangle_{\mathbb{F}_{q^m}}) = k-i$, one can assume up to equivalence that $\mathcal{U} \subseteq \langle e_1,\ldots,e_{k-i}\rangle_{\mathbb{F}_{q^m}} =: V$, and then study $\mathcal{U}$ in the projective subspace $\mathrm{PG}(k-i-1,q^m)$ induced by V.

For any $[n,k]_{q^m/q}$ system $\mathcal{U}$, the cardinality of the associated linear set $L_{\mathcal{U}}$ satisfies

$$|L_{\mathcal{U}}| \leq \frac{q^n-1}{q-1}. \tag{4.8}$$

A linear set $L_{\mathcal{U}}$ whose cardinality meets (4.8) with equality is said to be **scattered**. Equivalently, a linear set $L_{\mathcal{U}}$ is scattered if and only if $\mathrm{wt}_{\mathcal{U}}(P) = 1$ for each $P \in L_{\mathcal{U}}$. We also observe that

(4.8) can be refined as follows.

Lemma 4.21. Let $\mathcal{U}$ be an $[n,k]_{q^m/q}$ system. Then

$$\sum_{P\in \mathrm{PG}(k-1,q^m)} \frac{q^{\mathrm{wt}_{\mathcal{U}}(P)}-1}{q-1} = \frac{q^n-1}{q-1}.$$

Proof. Let Λ_1 be the set of 1-dimensional $\mathbb{F}_{q^m}$-subspaces of $\mathbb{F}_{q^m}^k$. Then, we have

$$\sum_{P\in \mathrm{PG}(k-1,q^m)} \frac{q^{\mathrm{wt}_{\mathcal{U}}(P)}-1}{q-1} = \frac{1}{q-1}\sum_{V\in\Lambda_1}(q^{\dim_{\mathbb{F}_q}(\mathcal{U}\cap V)}-1) = \frac{1}{q-1}\sum_{V\in\Lambda_1}|V\cap(\mathcal{U}\setminus\{0\})| = \frac{q^n-1}{q-1},$$

where the latter equality follows from Lemma 4.6. □

4.2.2 The Associated Hamming-Metric Code

The notion of a linear set allows us to describe a connection between rank-metric codes and some particular codes in the Hamming metric. This connection was also observed in [143]. For a $[n,k]_{q^m/q}$ system $\mathcal{U}$ and a point $P\in\mathrm{PG}(k-1,q^m)$, define

$$e_{\mathcal{U}}(P) := \frac{q^{\mathrm{wt}_{\mathcal{U}}(P)}-1}{q-1}.$$

The identity of Lemma 4.21 can be written as

$$\sum_{P\in \mathrm{PG}(k-1,q^m)} e_{\mathcal{U}}(P) = \frac{q^n-1}{q-1}. \tag{4.9}$$

Denote by $\mathcal{U}(n,k)_{q^m/q}$ the set of $[n,k]_{q^m/q}$ systems and by $\mathcal{P}(n,k)_{q^m}$ the set of projective $[n,k]_{q^m}$ systems. Define the map

$$\begin{array}{ccc} \mathcal{U}(n,k)_{q^m/q} & \longrightarrow & \mathcal{P}(\frac{q^n-1}{q-1},k)_{q^m}, \\ \mathcal{U} & \longmapsto & (L_{\mathcal{U}},e_{\mathcal{U}}), \end{array}$$

where $(L_{\mathcal{U}},e_{\mathcal{U}})$ denotes the multiset $L_{\mathcal{U}}$ with multiplicity function $e_{\mathcal{U}}$. The parameters $\frac{q^n-1}{q-1}$ and k of the projective system $(L_{\mathcal{U}},e_{\mathcal{U}})$ directly follow from (4.9). It is easy to see that this map is compatible with the equivalence relations on $\mathcal{U}(n,k)_{q^m/q}$ and on $\mathcal{P}(\frac{q^n-1}{q-1},k)_{q^m}$. Indeed, the actions defining the equivalence classes are given in both cases by the group $\mathrm{PGL}(k,q^m)$. We thus constructed a map

$$\mathrm{Ext}^{\mathrm{H}}: \ \mathcal{U}[n,k]_{q^m/q} \ \longrightarrow \ \mathcal{P}[\tfrac{q^n-1}{q-1},k]_{q^m},$$

where $\mathcal{U}[n,k]_{q^m/q}$ and $\mathcal{P}[\frac{q^n-1}{q-1},k]_{q^m}$ denote the set of equivalence classes of $[n,k]_{q^m/q}$ systems and the set of equivalence classes of projective $[\frac{q^n-1}{q-1},k]_{q^m}$ systems, respectively. This maps leaves also the parameter d of the projective $[\frac{q^n-1}{q-1},k]_{q^m}$ system fixed, as the following result shows.

Lemma 4.22. Let $[\mathcal{U}]$ be the equivalence class of $[n,k,d]_{q^m/q}$ systems. Then $[(L_{\mathcal{U}},e_{\mathcal{U}})]$ is the equivalence class of a projective $[\frac{q^n-1}{q-1},k,\frac{q^n-q^{n-d}}{q-1}]_{q^m}$ system. In other words, the map

$$\mathrm{Ext}^{\mathrm{H}}:\mathcal{U}[n,k,d]_{q^m/q}\longrightarrow\mathcal{P}\left[\frac{q^n-1}{q-1},k,\frac{q^n-q^{n-d}}{q-1}\right]_{q^m}$$

is well-defined.

Proof. The fact that the map $\mathrm{Ext}^{\mathrm{H}}$ sends equivalence classes of $[n,k]_{q^m/q}$ systems in equivalence classes of projective $[\frac{q^n-1}{q-1},k]_q$ systems has already been observed above. We only need to show the compatibility between the third parameters. More precisely, we need to show that for a given $[n,k,d]_{q^m/q}$ system $\mathcal{U}$, every element in $\mathrm{Ext}^{\mathrm{H}}([\mathcal{U}])$ is a projective $[\frac{q^n-1}{q-1},k,\frac{q^n-q^{n-d}}{q-1}]_q$ system. Fix the projective $[\frac{q^n-1}{q-1},k,d']_q$ system $(L_{\mathcal{U}},e_{\mathcal{U}})$, and denote by Λ_{k-1} the set of $\mathbb{F}_{q^m}$-hyperplanes of $\mathbb{F}_{q^m}^k$. Then for any $H\in\Lambda_{k-1}$ we have

$$\begin{aligned}\sum_{P\in\mathrm{PG}(H,\mathbb{F}_{q^m})}e_{\mathcal{U}}(P)&=\sum_{P\in\mathrm{PG}(H,\mathbb{F}_{q^m})}\frac{q^{\mathrm{wt}_{\mathcal{U}}(P)}-1}{q-1}\\&=\frac{1}{q-1}\sum_{\substack{V\subseteq H\\ \dim_{\mathbb{F}_{q^m}}(V)=1}}|V\cap(\mathcal{U}\setminus\{0\}|\\&=\frac{1}{q-1}|H\cap(\mathcal{U}\setminus\{0\})|\\&=\frac{q^{\dim_{\mathbb{F}_q}(H\cap\mathcal{U})}-1}{q-1},\end{aligned}$$

where the second to last identity follows from the fact that $\{V\setminus\{0\}:V\subseteq H,\dim_{\mathbb{F}_{q^m}}(V)=1\}$ is a partition of $H\setminus\{0\}$. Therefore we obtain

$$d'=\frac{q^n-1}{q-1}-\max\left\{\frac{q^{\dim_{\mathbb{F}_q}(H\cap\mathcal{U})}-1}{q-1}\ :\ H\in\Lambda_{k-1}\right\}=\frac{q^n-1}{q-1}-\frac{q^{n-d}-1}{q-1}=\frac{q^n-q^{n-d}}{q-1}.\quad\square$$

Definition 4.23. Let $\mathcal{C}$ be a nondegenerate $[n,k,d]_{q^m/q}$ rank-metric code. We will call any Hamming-metric code in $(\Psi^{\mathrm{H}}\circ\mathrm{Ext}^{\mathrm{H}}\circ\Phi)([\mathcal{C}])$ **associated** with $\mathcal{C}$. Note that any such an object is a $[\frac{q^n-1}{q-1},k,\frac{q^n-q^{n-d}}{q-1}]_{q^m}$ code.

The Hamming-metric code associated to $\mathcal{C}$ in the previous definition is clearly not unique. However, the choice of the code is irrelevant when focusing on properties that are invariant under monomial equivalence. Therefore, for ease of notation, in the sequel we denote by $\mathcal{C}^{\mathrm{H}}$ any code that belongs to $(\Psi^{\mathrm{H}}\circ\mathrm{Ext}^{\mathrm{H}}\circ\Phi)([\mathcal{C}])$.

Example 4.24. Let $q = 2$, $n = 4$ and $m = 3$. Consider $\mathbb{F}_8 = \mathbb{F}_2[\alpha]$, where $\alpha^3 + \alpha + 1 = 0$. Moreover, let $\mathcal{C}$ be the $[4, 2, 1]_{8/2}$ code whose generator matrix

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

Take the $[4, 2, 1]_{8/2}$ system $\mathcal{U}$ spanned by the columns of G, i.e., $\mathcal{U} = \{(a, \beta) : a \in \mathbb{F}_2, \beta \in \mathbb{F}_8\}$. The weights of the points in $\mathrm{PG}(1, 8)$ with respect to $\mathcal{U}$ are given by

$$\begin{aligned} \mathrm{wt}_{\mathcal{U}}([1 : a]) &= 1, \qquad \text{for every } a \in \mathbb{F}_8 \\ \mathrm{wt}_{\mathcal{U}}([0 : 1]) &= 3. \end{aligned}$$

Hence, we obtain that $\mathrm{Ext}^{\mathrm{H}}(\mathcal{U}) = (\mathrm{PG}(1, 8), e_{\mathcal{U}})$, where

$$\begin{aligned} e_{\mathcal{U}}([1 : a]) &= 1, \qquad \text{for every } a \in \mathbb{F}_8 \\ e_{\mathcal{U}}([0 : 1]) &= 7. \end{aligned}$$

At this point, any code $\mathcal{C}^{\mathrm{H}} = C \in (\Psi^{\mathrm{H}} \circ \mathrm{Ext}^{\mathrm{H}} \circ \Phi)([\mathcal{C}])$ is monomially equivalent to the $[15, 2, 8]_8$ (Hamming-metric) code whose generator matrix is

$$G_{\mathrm{Ext}} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & \alpha & \alpha^2 & \alpha^3 & \alpha^4 & \alpha^5 & \alpha^6 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

Example 4.25 (Simplex Rank-Metric Code)**.** Take $\mathcal{C}$ to be the $[km, k, m]_{q^m/q}$ simplex rank-metric code, whose corresponding $[km, k, m]_{q^m/q}$ system is $\Phi([\mathcal{C}]) = [\mathbb{F}_{q^m}^k]$. Denote $\mathcal{U} := \mathbb{F}_{q^m}^k$ and consider the weight of each point $P \in \mathrm{PG}(k-1, q^m)$ in $L_{\mathcal{U}}$. For $P = [v]$, we have

$$\mathrm{wt}_{\mathcal{U}}(P) = \dim_{\mathbb{F}_q}(\mathcal{U} \cap \langle v \rangle_{\mathbb{F}_{q^m}}) = \dim_{\mathbb{F}_q}(\langle v \rangle_{\mathbb{F}_{q^m}}) = m.$$

Therefore by applying the map $\mathrm{Ext}^{\mathrm{H}}$ we obtain

$$\mathrm{Ext}^{\mathrm{H}}([\mathcal{U}]) = [(L_{\mathcal{U}}, e_{\mathcal{U}})],$$

where $L_{\mathcal{U}} = \mathrm{PG}(k-1, q^m)$ and

$$e_{\mathcal{U}}(P) = \frac{q^m - 1}{q - 1} \quad \text{for all } P \in \mathrm{PG}(k-1, q^m).$$

In particular, any code in $(\Psi^{\mathrm{H}} \circ \mathrm{Ext}^{\mathrm{H}} \circ \Phi)([\mathcal{C}])$ is monomially equivalent to the concatenation of $\frac{q^m-1}{q-1}$ copies of the $[\frac{q^{km}-1}{q^m-1}, k, q^{(k-1)m}]_{q^m}$ simplex code in the Hamming metric.

Lemma 4.22 shows how the fundamental parameters of a nondegenerate $[n, k, d]_{q^m/q}$ rank-metric code $\mathcal{C}$ relate to those of an associated Hamming-metric code $\mathcal{C}^{\mathrm{H}}$. The connection can be made even more precise. For example, we can say how the weight distributions of the two codes relate to each other.

Theorem 4.26. Let $\mathcal{C}$ be a nondegenerate $[n,k,d]_{q^m/q}$ rank-metric code with rank-weight distribution $\{A_i^{\mathrm{rk}}(\mathcal{C})\}_i$. Then the Hamming-weight distribution of $\mathcal{C}^{\mathrm{H}}$ is $\{A_j^{\mathrm{H}}(\mathcal{C}^{\mathrm{H}})\}_j$ with

$$A_j^{\mathrm{H}}(\mathcal{C}^{\mathrm{H}}) = \begin{cases} A_i^{\mathrm{rk}}(\mathcal{C}) & \text{if } j = \frac{q^n-q^{n-i}}{q-1}, \\ 0 & \text{otherwise.} \end{cases}$$

Proof. Let G be a generator matrix for $\mathcal{C}$ and denote by $\mathcal{U}$ the $\mathbb{F}_q$-span of its columns. Let G_{Ext} be a generator matrix for $\mathcal{C}^{\mathrm{H}}$ whose columns are the elements of the multiset $(L_{\mathcal{U}}, e_{\mathcal{U}})$. Doing the same computations as in Lemma 4.22 we obtain that, for every $u \in \mathbb{F}_{q^m}^k \setminus \{0\}$,

$$\mathrm{wt}^{\mathrm{H}}(uG_{\mathrm{Ext}}) = \frac{q^n-1}{q-1} - \sum_{P \in \mathrm{PG}(H_u,\mathbb{F}_{q^m})} \frac{q^{\mathrm{wt}_{\mathcal{U}}(P)}-1}{q-1} = \frac{q^n - q^{n-\mathrm{rk}(uG)}}{q-1}, \tag{4.10}$$

where $H_u := \langle u \rangle^{\perp}$. □

Remark 4.27. While the connection between the dual of a code $\mathcal{C}$ and the dual of $\mathcal{C}^{\mathrm{H}}$ seems to be difficult to describe explicitly, we remark that their weight distributions (in the rank and Hamming metric, respectively) are linked via the theory of MacWilliams identities; see [109] for a general reference. More precisely, the Hamming weight distribution of $(\mathcal{C}^{\mathrm{H}})^{\perp}$ can be written in terms of the Hamming weight distribution of $\mathcal{C}^{\mathrm{H}}$. By Theorem 4.26, the latter can be written in terms of the rank weight distribution of $\mathcal{C}$ which, in turn, can be expressed in terms of the rank weight distribution of $\mathcal{C}^{\perp}$.

Remark 4.28. Theorem 4.26 generalizes various known results on Hamming-metric codes obtained from linear sets. This is the case of the two-weight Hamming-metric codes arising from maximum scattered linear sets found by Blokhuis and Lavrauw in [40, Section 5], and of the Hamming-metric codes with $h+1$ weights recently presented by Zini and Zullo in [164, Theorem 7.1].

Finally, one can also prove the following result connecting the generalized weights of $\mathcal{C}$ and $\mathcal{C}^{\mathrm{H}}$ (in the respective metrics).

Theorem 4.29. Let $\mathcal{C}$ be a nondegenerate $[n,k,d]_{q^m/q}$ rank-metric code with generalized rank-weights $\{d_i(\mathcal{C})\}_i$. Then the generalized Hamming-weights of $\mathcal{C}^{\mathrm{H}}$ are given by $\{d_i^{\mathrm{H}}(\mathcal{C}^{\mathrm{H}})\}_i$, where

$$d_i^{\mathrm{H}}(\mathcal{C}^{\mathrm{H}}) = \frac{q^n - q^{n-d_i(\mathcal{C})}}{q-1}.$$

4.2.3 The Total Weight of a Rank-Metric Code

In this subsection we continue comparing codes in the rank and in the Hamming metric. Our focus is on the rank-metric analogue of the *total weight*. It is well-known that the latter only depends on the field size and on the code's dimension and effective length. More precisely, if $\mathcal{C} \subseteq \mathbb{F}_q^n$ is a Hamming-nondegenerate code, then

$$\sum_{v \in \mathcal{C}} \mathrm{wt}^{\mathrm{H}}(v) = n(q^k - q^{k-1}). \tag{4.11}$$

This simple result, which has numerous applications in classical coding theory (for example, a simple proof of the Plotkin bound for linear codes), does not have an immediate analogue in the rank metric. Indeed, it is easy to find examples of rank-nondegenerate codes having the same parameters (q, m, n, k) but for which the quantity $\sum_{v \in \mathcal{C}} \mathrm{rk}(v)$ is not a constant.

In this section, we argue that, in the "total weight" context, a convenient analogue of $\mathrm{wt}^{\mathrm{H}}(v)$ is $q^{n-\mathrm{rk}(v)}$. We start by recalling the following q-analogue of the Pless identities; see [88].

Notation 4.30. For a prime power q and integers n, m, k, j, r, let

$$f_q(n, m, k, j, r) := \sum_{\nu=j}^{r} q^{m(k-\nu)} \binom{n-j}{\nu-j}_q \binom{r}{\nu}_q \prod_{\ell=0}^{\nu-1} (q^\nu - q^\ell).$$

Theorem 4.31 (Theorem 30 of [60]). Let $\mathcal{C}$ be an $[n, k, d]_{q^m/q}$ code. Then for all $0 \leq r \leq n$ we have

$$\sum_{v \in \mathcal{C}} q^{r(n-\mathrm{rk}(v))} = \sum_{j=0}^{r} A_j(\mathcal{C}^\perp)\, f_q(n, m, k, j, r).$$

In analogy with Remark 4.27, we observe that a different statement of Pless-type identities can in principle be obtained by combining the correspondence $\mathcal{C} \to \mathcal{C}^{\mathrm{H}}$ with the classical Pless identities for Hamming-metric codes. For the purposes of this section, Theorem 4.31 is what we will need.

In this thesis, we are not only interested in the q-analogue of the total weight of a code, but also in other related quantities. In order to unify their treatment, it is convenient to regard the Hamming/rank weight of the nonzero elements of a code as a discrete random variable, which we simply denote by $\mathcal{C}^*$, $\mathbb{E}^{\mathrm{rk}}$ and $\mathrm{Var}^{\mathrm{rk}}$ for the mean and variance of (a function of) $\mathcal{C}^*$, viewed as a random variable in the sense explained above.

In the sequel, we call a code $\mathcal{C} \subseteq \mathbb{F}_{q^m}^n$ **rank-2-nondegenerate** if $d(\mathcal{C}^\perp) \geq 3$. Codes with this property are the rank-metric analogues of projective codes in the Hamming metric; see page 68. The previous theorem has the following simple consequences.

Corollary 4.32. Let $\mathcal{C}$ be an $[n,k,d]_{q^m/q}$ code. If $\mathcal{C}$ is rank-nondegenerate, then

$$\begin{aligned}\mathbb{E}^{\mathrm{rk}}[q^{n-\mathcal{C}^*}] &= \frac{-q^n+q^{mk}+q^{m(k-1)}(q^n-1)}{q^{mk}-1},\\ \mathrm{Var}^{\mathrm{rk}}[q^{n-\mathcal{C}^*}] &\geq \frac{-q^{2n}+f_q(n,m,k,0,2)}{q^{mk}-1}-\mathbb{E}^{\mathrm{rk}}[q^{n-\mathcal{C}^*}]^2,\end{aligned}$$

where the latter lower bound is attained with equality if and only if $\mathcal{C}$ is rank-2-nondegenerate.

Corollary 4.32 establishes the rank-metric analogue of the formula for the total weight of a Hamming-metric code in (4.11). It also shows that, for a rank-2-nondegenerate code $\mathcal{C}$, the variance of the random variable $q^{n-\mathcal{C}^*}$ only depends on a few code's parameters. While the formulas in Corollary 4.32 are quite involved and not immediate to interpret, their asymptotics as $q \to +\infty$ can be explicitly computed. The estimates describe how the variance behaves over large fields.

Proposition 4.33. Let $\mathcal{C}$ be an $[n,k,d]_{q^m/q}$ code. If $\mathcal{C}$ is rank-nondegenerate then $n \leq km$ and, as $q \to +\infty$,

$$\mathbb{E}^{\mathrm{rk}}[q^{n-\mathcal{C}^*}] \sim \begin{cases} 1 & \text{if } n \leq m-1,\\ q^{n-m} & \text{if } m+1 \leq n \leq km,\\ 2 & \text{if } n = m.\end{cases}$$

If in addition $\mathcal{C}$ is rank-2-nondegenerate, then $n \leq mk/2$ and for $k \geq 3$ and $q \to +\infty$ we have

$$\mathrm{Var}^{\mathrm{rk}}[q^{n-\mathcal{C}^*}] \sim \begin{cases} q^{-m+n+1} & \text{if } k \leq n \leq m-2 \text{ or } m+2 \leq n \leq mk/2,\\ 1 & \text{if } n = m-1,\\ q & \text{if } n = m,\\ q^2 & \text{if } n = m+1.\end{cases}$$

Proof. The first part of the statement easily follows from Proposition 4.4 and Corollary 4.32. To prove the second part, we start by applying the rank-metric Singleton bound [63, 74] to $\mathcal{C}^\perp$, obtaining $m(n-k) \leq n(m-d(\mathcal{C}^\perp)+1) \leq n(m-2)$. This implies $n \leq mk/2$, as desired.

We now turn to the asymptotic estimates. To simplify the notation, write f_q instead of $f_q(n,m,k,0,2)$. Lengthy computations show that

$$f_q = q^{mk} + q^{m(k-1)}(q^n-1)(q+1) + q^{m(k-2)+1}(q^n-1)(q^{n-1}-1).$$

Therefore

$$f_q(n,m,k,0,2) \sim \begin{cases} q^{mk} & \text{if } n \leq m-2, \\ q^{mk+1} & \text{if } n = m, \\ q^{mk+2n-2m} & \text{if } n \geq m+2, \\ 2q^{mk} & \text{if } n = m-1, \\ 2q^{mk+2} & \text{if } n = m+1. \end{cases}$$

From the first part of the statement we also have

$$\mathbb{E}^{\mathrm{rk}}[q^{n-\mathcal{C}^*}]^2 \sim \begin{cases} 1 & \text{if } n \leq m-1, \\ q^{2n-2m} & \text{if } n \geq m+1, \\ 4 & \text{if } n = m. \end{cases}$$

Using $k \geq 3$ (needed in the case $n = m+1$), this easily gives the asymptotics of

$$\mathrm{Var}^{\mathrm{rk}}[q^{n-\mathcal{C}^*}] = \frac{-q^{2n} + f_q(n,m,k,0,2)}{q^{mk}-1} - \mathbb{E}^{\mathrm{rk}}[q^{n-\mathcal{C}^*}]^2$$

for $n = m-1$, $n = m$, and $n = m+1$. To compute the asymptotics in the other cases, write

$$\frac{-q^{2n} + f_q(n,m,k,0,2)}{q^{mk}-1} - \mathbb{E}^{\mathrm{rk}}[q^{n-\mathcal{C}^*}]^2 = \frac{A_q - B_q}{(q^{mk}-1)^2},$$

where $A_q = (q^{mk}-1)(-q^{2n}+f_q)$ and $B_q = (-q^n + q^{mk} + q^{m(k-1)}(q^n-1))^2$.

If $n \leq m-2$ then $f_q \sim q^{mk} + q^{m(k-1)+n+1}$. Therefore $A_q \sim q^{2mk} + q^{m(2k-1)+n+1}$ and $B_q \sim -q^{2mk} + 2q^{m(2k-1)+n}$, from which the desired asymptotic estimate follows.

If $m+2 \leq n \leq mk/2$, then $m+2 \leq n \leq m(k-1)$, because $k \geq 3$. We then have $f_q \sim q^{m(k-2)+2n} + q^{m(k-1)+n+1}$ and thus $A_q \sim q^{m(2k-2)+2n} + q^{m(2k-1)+n+1}$, $B_q \sim q^{2m(k-1)+2n} + 2q^{m(2k-1)+n}$. This again implies the desired asymptotic estimate.

□

4.3 Minimal Rank-Metric Codes: Geometry and Properties

The next two sections are devoted to the theory of minimal codes in the rank metric. In this first section we propose a definition of minimal and establish a 1-1 correspondence between $[n,k]_{q^m/q}$ minimal rank-metric codes and certain $[n,k]_{q^m/q}$ systems. This allows us to investigate the main properties of this new family of codes.

Definition 4.34. Let $\mathcal{C}$ be an $[n,k]_{q^m/q}$ code. A codeword $v \in \mathcal{C}$ is a **minimal codeword** if, for every $v' \in \mathcal{C}$, $\sigma^{\mathrm{rk}}(v') \subseteq \sigma^{\mathrm{rk}}(v)$ implies $v' = \alpha v$ for some $\alpha \in \mathbb{F}_{q^m}$. We say that $\mathcal{C}$ is **minimal**

if all its codewords are minimal.

Lemma 4.35. Let $v \in \mathbb{F}_{q^m}^n$. The following hold.

1. There exists $A \in \mathrm{GL}_n(q)$ such that $\sigma^{\mathrm{rk}}(vA) = \langle e_i \,:\, i \in \sigma^{\mathrm{H}}(vA)\rangle$.
2. Let $I \subseteq \{1,\dots,n\}$. Then, $\sigma^{\mathrm{rk}}(v) \subseteq \langle e_i \,:\, i \in I\rangle$ if and only if $I \supseteq \sigma^{\mathrm{H}}(v)$. In particular,

   $$\sigma^{\mathrm{H}}(v) = \arg\min\{|I| \,:\, \sigma^{\mathrm{rk}}(v) \subseteq \mathcal{E}_I\},$$

   where $\mathcal{E}_I := \langle e_i \,:\, i \in I\rangle$ and $\sigma^{\mathrm{rk}}(v) \subseteq \langle e_i \,:\, i \in \sigma^{\mathrm{H}}(v)\rangle$.

Proof.

1. Let $r = \dim(\sigma^{\mathrm{rk}}(v))$. By Proposition 2.11, there exist a matrix A and a basis Γ of $\mathbb{F}_{q^m}/\mathbb{F}_q$, such that $\Gamma(vA)$ is in Smith normal form. Hence, $\sigma^{\mathrm{H}}(vA) = \{1,\dots,r\}$ and $\sigma^{\mathrm{rk}}(vA) = \langle e_i \,:\, i \in \{1,\dots,r\}\rangle$.
2. Let $I = \sigma^{\mathrm{H}}(v)$ and fix any basis Γ of $\mathbb{F}_{q^m}/\mathbb{F}_q$. The rows indexed by $\{1,\dots,n\} \setminus I$ in $\Gamma(v)$ are identically zero. Hence, $\sigma^{\mathrm{rk}}(v) \subseteq \langle e_i \,:\, i \in I\rangle$. Vice versa, assume that there exists $t \in \sigma^{\mathrm{H}}(v) \setminus I$. Fix any basis Γ of $\mathbb{F}_{q^m}/\mathbb{F}_q$. Since $t \in \sigma^{\mathrm{H}}(v)$, there exists $j \in \{1,\dots,m\}$ such that $\Gamma(v_t)_j \neq 0$. Hence, the vector $a = (\Gamma(v_1)_j,\dots,\Gamma(v_n)_j)$ belongs to $\sigma^{\mathrm{rk}}(v)$ and has a nonzero entry in the t-th coordinate. Thus, $\sigma^{\mathrm{rk}}(v) \not\subseteq \langle e_i \,:\, i \in I\rangle$. The second statement immediately follows. □

4.3.1 Linear Cutting Blocking Sets and the Parameters of Minimal Codes

In this subsection we give a geometric characterization of minimal codes in the rank metric. This will allow us to derive bounds on their parameters.

We start with the q-analogue of the notion of a cutting blocking set.

Definition 4.36. A $[n,k]_{q^m/q}$ system $\mathcal{U}$ is called a **linear cutting blocking set** if for any $\mathbb{F}_{q^m}$-hyperplanes $H, H' \subseteq \mathbb{F}_{q^m}^k$ we have $(\mathcal{U} \cap H) \subseteq (\mathcal{U} \cap H')$ implies $H = H'$. We will say that that $\mathcal{U}$ is a linear cutting $[n,k]_{q^m/q}$ blocking set to emphasize the parameters.

While the term “linear cutting blocking set” might seem not fully consistent with the terminology used so far (since such an object is not a linear set), one can verify that an $[n,k]_{q^m/q}$ system $\mathcal{U}$ is a linear cutting blocking set if and only if its associated linear set $L_{\mathcal{U}}$ is a cutting blocking set in $\mathrm{PG}(k-1,q^m)$. The proof of this fact can be found in Section 4.3.2; see Theorem 4.46. This explains the choice of the terminology.

We will need the following characterization of linear cutting blocking sets.

Proposition 4.37. A $[n,k]_{q^m/q}$ system $\mathcal{U}$ is a linear cutting blocking set if and only if for every $\mathbb{F}_{q^m}$-hyperplane H we have $\langle H \cap \mathcal{U}\rangle_{\mathbb{F}_{q^m}} = H$.

Proof. ($\Leftarrow$) Let H, H' be two $\mathbb{F}_{q^m}$-hyperplanes of $\mathbb{F}_{q^m}^k$ such that $(\mathcal{U}\cap H) \subseteq (\mathcal{U}\cap H')$. Hence, $H = \langle H\cap\mathcal{U}\rangle_{\mathbb{F}_{q^m}} \subseteq \langle H'\cap\mathcal{U}\rangle_{\mathbb{F}_{q^m}} = H'$. Since H and H' are both hyperplanes, they have to be equal.

($\Rightarrow$) Suppose by contradiction that there exists an $\mathbb{F}_{q^m}$-hyperplane H of $\mathbb{F}_{q^m}^k$ such that $\langle H\cap\mathcal{U}\rangle_{\mathbb{F}_{q^m}} = X \subsetneq H$. Then, for every hyperplane $H' \supset X$ we have $(\mathcal{U}\cap H) \subseteq (\mathcal{U}\cap H')$. Since there are at least q^m such hyperplanes different from H, we obtain that $\mathcal{U}$ is not a linear cutting blocking set. □

Corollary 4.38. If $\mathcal{U}$ is a linear cutting $[n,k]_{q^m/q}$ blocking set, then for every $\mathbb{F}_{q^m}$-hyperplane H of $\mathbb{F}_{q^m}^k$ we have $|H\cap\mathcal{U}| \geq q^{k-1}$.

Proof. Let $t := \dim_{\mathbb{F}_q}(H\cap\mathcal{U})$. Then an $\mathbb{F}_q$-basis for $H\cap\mathcal{U}$ is also a set of $\mathbb{F}_{q^m}$-generators for $\langle H\cap\mathcal{U}\rangle_{\mathbb{F}_{q^m}}$. Hence, since $\mathcal{U}$ is a linear cutting blocking set, by Proposition 4.37 we have

$$m(k-1) = \dim_{\mathbb{F}_q}(\langle H\cap\mathcal{U}\rangle_{\mathbb{F}_{q^m}}) \leq mt,$$

which shows that $t \geq k-1$. □

The geometric description of minimal rank-metric codes via linear cutting blocking sets relies on the following characterization of the inclusion of rank supports.

Theorem 4.39. Let G be a generator matrix for a nondegenerate $[n,k]_{q^m/q}$ code, $\mathcal{U}$ be the corresponding $[n,k]_{q^m/q}$ system and $u,v \in \mathbb{F}_{q^m}^k \setminus \{0\}$. Then,

$$\sigma^{\mathrm{rk}}(uG) \subseteq \sigma^{\mathrm{rk}}(vG) \qquad \text{if and only if} \qquad (\langle u\rangle^{\perp}\cap\mathcal{U}) \supseteq (\langle v\rangle^{\perp}\cap\mathcal{U}).$$

Proof. ($\Leftarrow$) Let $x_1,\ldots,x_t$ be an $\mathbb{F}_q$-basis of the space $X := (\langle v\rangle^{\perp}\cap\mathcal{U})$. Let $A \in \mathrm{GL}_n(q)$ be such that

$$GA = (\, x_1 \,|\, \cdots \,|\, x_t \,|\, G' \,),$$

where $G' \in \mathbb{F}_{q^m}^{k\times(n-t)}$. We have $uGA = (0,\ldots,0|uG')$ and $vGA = (0,\ldots,0|vG')$. Moreover, by (4.5) we have $\mathrm{rk}(vGA) = n-t$ and, by Lemma 4.35, $\sigma^{\mathrm{rk}}(vGA) = \langle e_i \,:\, i = t+1,\ldots,n\rangle$ and $\sigma^{\mathrm{rk}}(uGA) \subseteq \langle e_i \,:\, i = t+1,\ldots,n\rangle$. This means that $\sigma^{\mathrm{rk}}(uGA) \subseteq \sigma^{\mathrm{rk}}(vGA)$. Finally, Proposition 2.11 implies $\sigma^{\mathrm{rk}}(uG) \subseteq \sigma^{\mathrm{rk}}(vG)$.

($\Rightarrow$) Assume now that $\sigma^{\mathrm{rk}}(uG) \subseteq \sigma^{\mathrm{rk}}(vG)$. Let $r := \mathrm{rk}(vG)$. By the first part of Lemma 4.35 there exists $A \in \mathrm{GL}_n(q)$ such that $\sigma^{\mathrm{rk}}(vGA) = \langle e_1,\ldots,e_r\rangle$. Hence, $\sigma^{\mathrm{rk}}(uGA) \subseteq \sigma^{\mathrm{rk}}(vGA) = \langle e_1,\ldots,e_r\rangle$. Denote by $x_1,\ldots,x_n$ the columns of GA, which also form a basis of $\mathcal{U}$. In this notation we have $\langle v\rangle^{\perp}\cap\mathcal{U} = \langle x_{r+1},\ldots,x_n\rangle_{\mathbb{F}_q}$. Moreover, by the second part of Lemma 4.35 we have $\sigma^{\mathrm{H}}(uGA) \subseteq \{1,\ldots,r\}$. This implies that $x_i \in \langle u\rangle^{\perp}$ for $i = r+1,\ldots,n$. Hence, $(\langle u\rangle^{\perp}\cap\mathcal{U}) \supseteq (\langle v\rangle^{\perp}\cap\mathcal{U})$. □

By combining Theorem 4.39 and the correspondence stated in Theorem 4.8 we obtain the following.

Corollary 4.40. The correspondence (Φ,Ψ) defined in Section 4.1.1 induces a 1-1 correspondence between minimal rank-metric codes and linear cutting blocking sets.

Corollary 4.40 has several consequences in the theory of minimal codes. The first result we derive concerns the construction of new minimal codes from existing ones.

Corollary 4.41. Let $\mathcal{C}$ be an $[n,k]_{q^m/q}$ minimal rank-metric code with generator matrix G, and let $v \in \mathbb{F}_{q^m}^k$. Then the $[n+1,k]_{q^m/q}$ code $\bar{\mathcal{C}} = \mathrm{rowsp}(G \mid v^\top)$ is minimal.

Proof. Without loss of generality, we may assume that $\mathcal{C}$ is nondegenerate. Let $\mathcal{U}$ be any $[n,k]_{q^m/q}$ system associated to $[\mathcal{C}]$ and let $\bar{\mathcal{U}} = \langle \mathcal{U}, v\rangle_{\mathbb{F}_q}$. If $v \in \mathcal{U}$, then by Proposition 4.2 the code $\bar{\mathcal{C}}$ is degenerate and it is equivalent to the code $\{(\, c \mid 0\,) : c \in \mathcal{C}\}$, which is clearly minimal. Hence, assume that $v \notin \mathcal{U}$. By Proposition 4.2 we have that $\bar{\mathcal{C}}$ is nondegenerate and $\bar{\mathcal{U}}$ is an $[n+1,k]_{q^m/q}$ system associated to $\bar{\mathcal{C}}$. Let H be any $\mathbb{F}_{q^m}$-hyperplane of $\mathbb{F}_{q^m}^k$. Then

$$H \supseteq \langle H \cap \bar{\mathcal{U}}\rangle_{\mathbb{F}_{q^m}} = \langle H \cap (\mathcal{U} + \langle v\rangle_{\mathbb{F}_q})\rangle_{\mathbb{F}_{q^m}} \supseteq \langle H \cap \mathcal{U}\rangle = H,$$

where the latter equality follows from the fact that, since $\mathcal{C}$ is minimal, $\mathcal{U}$ is a linear cutting blocking set by Theorem 4.40. Therefore $\bar{\mathcal{U}}$ is also a linear cutting blocking set and we conclude using Theorem 4.40 again. □

The following two results are also consequences of Corollary 4.40 and provide information about the parameters of a minimal $[n,k]_{q^m/q}$ code.

Corollary 4.42. Let $\mathcal{C}$ be a minimal $[n,k]_{q^m/q}$ code. Then for every $c \in \mathcal{C}$ we have $\mathrm{rk}(c) \leq \dim_{\mathbb{F}_q}(\sigma^{\mathrm{rk}}(\mathcal{C})) - k + 1$. In particular, $w^{\mathrm{rk}}(\mathcal{C}) \leq \dim_{\mathbb{F}_q}(\sigma^{\mathrm{rk}}(\mathcal{C})) - k + 1 \leq n - k + 1$.

Proof. Let $n' = \dim_{\mathbb{F}_q}(\sigma^{\mathrm{rk}}(\mathcal{C}))$ for ease of notation. As observed in Remark 4.3, we can isometrically embed $\mathcal{C}$ in $\mathbb{F}_{q^m}^{n'}$. Moreover, the resulting code is minimal if and only if $\mathcal{C}$ is minimal. Therefore we can assume without loss of generality that $\mathcal{C}$ is nondegenerate of length $n = \dim_{\mathbb{F}_q}(\sigma^{\mathrm{rk}}(\mathcal{C}))$. Let $\mathcal{U}$ be any $[n,k]_{q^m/q}$ system associated to $\mathcal{C}$. By Corollary 4.40, $\mathcal{U}$ is a linear cutting blocking set. From the proof of Corollary 4.38 we get that $\dim_{\mathbb{F}_q}(H \cap \mathcal{U}) \geq k-1$ for every $\mathbb{F}_{q^m}$-hyperplane of $\mathbb{F}_{q^m}^k$, and we conclude using Lemma 4.7. □

Corollary 4.43. If $\mathcal{C}$ is a minimal $[n,k]_{q^m/q}$ code with $k \geq 2$, then $n \geq k + m - 1$.

Proof. Without loss of generality we shall assume that $\mathcal{C}$ is nondegenerate. Therefore by Proposition 4.11 we have $w^{\mathrm{rk}}(\mathcal{C}) = \min\{m,n\}$. Since $k \geq 2$, by Corollary 4.42 we also have $w^{\mathrm{rk}}(\mathcal{C}) \leq n - k + 1 < n$. Therefore $w^{\mathrm{rk}}(\mathcal{C}) = m$ and using again the fact that $w^{\mathrm{rk}}(\mathcal{C}) \leq n - k + 1$ we find $n \geq m + k - 1$, as desired. □

4.3.2 Connections with Hamming-Metric Minimal Codes

It is natural to ask how the notions of minimality in the rank and in the Hamming metric relate to each other. This is the question we address in this subsection. In particular, we prove that a nondegenerate rank-metric code $\mathcal{C}$ is minimal if and only if its associated code(s) $\mathcal{C}^{\mathrm{H}}$ is minimal; see Section 4.2.2 for the notation.

The following result shows that minimality in the Hamming metric implies minimality in the rank metric. We propose two proofs, one in coding theory parlance and the other in the language of projective systems.

Proposition 4.44. Let $\mathcal{C}$ be an $[n,k]_{q^m/q}$ code with the property of being Hamming-minimal. Then $\mathcal{C}$ is rank-minimal.

Proof. Suppose that $\mathcal{C}$ is not a minimal rank-metric code. Then there exist two codewords v, v' that are $\mathbb{F}_{q^m}$-linearly independent such that $\sigma^{\mathrm{rk}}(v) \subseteq \sigma^{\mathrm{rk}}(v')$. By Lemma 4.35, we also have that $\sigma^{\mathrm{H}}(v') = \arg\min\{|I| \,:\, \sigma^{\mathrm{rk}}(v') \subseteq \mathcal{E}_I\}$, where $\mathcal{E}_I := \langle e_i \,:\, i \in I\rangle$. Since $\sigma^{\mathrm{rk}}(v) \subseteq \sigma^{\mathrm{rk}}(v')$, we have $\sigma^{\mathrm{H}}(v) \subseteq \sigma^{\mathrm{H}}(v')$, and therefore $\mathcal{C}$ is not Hamming-minimal. □

Second proof. Without loss of generality, we may assume that $\mathcal{C}$ is nondegenerate. Let G be a generator matrix for $\mathcal{C}$, and let $\mathcal{B}$ be the basis of the associated $[n,k]_{q^m/q}$ system $\mathcal{U}$ formed by the columns of G. Then $\mathcal{M} := \{\langle u\rangle_{q^m} \,:\, u \in \mathcal{B}\}$ is a projective $[n,k]_{q^m}$ system in $\mathrm{PG}(k-1,q^m)$. By Hamming-minimality and Theorem 3.8, it is a cutting blocking set. Hence $\langle \mathrm{PG}(H,\mathbb{F}_{q^m}) \cap \mathcal{M}\rangle = \mathrm{PG}(H,\mathbb{F}_{q^m})$ for every $\mathbb{F}_{q^m}$-hyperplane H of $\mathbb{F}_{q^m}^k$. Let H be an $\mathbb{F}_{q^m}$-hyperplane of $\mathbb{F}_{q^m}^k$ and let

$$V := \langle H \cap \mathcal{U}\rangle = \langle H \cap \langle \mathcal{B}\rangle_{\mathbb{F}_q}\rangle.$$

Then

$$\mathrm{PG}(V,\mathbb{F}_{q^m}) = \langle \mathrm{PG}(H,\mathbb{F}_{q^m}) \cap L_{\mathcal{U}}\rangle \supseteq \langle \mathrm{PG}(H,\mathbb{F}_{q^m}) \cap \mathcal{M}\rangle = \mathrm{PG}(H,\mathbb{F}_{q^m}),$$

showing that $V = H$, We conclude by applying Proposition 4.37. □

Remark 4.45. The converse of Proposition 4.44 is false in general. For example, let $(q,m,n) = (2,3,4)$. Write $\mathbb{F}_8 = \mathbb{F}_2[\alpha]$, where $\alpha^3+\alpha+1 = 0$. The code generated by

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

is rank-minimal but not Hamming-minimal. Moreover, the code $\mathcal{C}\cdot A$ is not Hamming-minimal for any $A \in \mathrm{GL}_3(2)$. Indeed, if this was the case, then there would exist a Hamming-minimal $[4,2]_8$ code, which contradicts [7, Theorem 2.14].

The previous results and examples show that minimality in the rank and in the Hamming metric gives rise to very different concepts. We now show that the correspondence $\mathcal{C} \to \mathcal{C}^{\mathrm{H}}$ is more natural in this context, as it translates rank-minimality precisely into Hamming-minimality.

Theorem 4.46. Let $\mathcal{C}$ be a nondegenerate $[n,k,d]_{q^m/q}$ rank-metric code. Then $\mathcal{C}$ is minimal if and only if $\mathcal{C}^{\mathrm{H}}$ is Hamming-minimal.

Proof. By Corollary 4.40, $\mathcal{C}$ is minimal if and only if any $[n,k]_{q^m/q}$ system $\mathcal{U}$ associated with $\mathcal{C}$ is a linear cutting blocking set. Now, consider the linear set $L_{\mathcal{U}}$. We show that $\mathcal{U}$ is a linear cutting blocking set if and only if $L_{\mathcal{U}}$ is a cutting blocking set in $\mathrm{PG}(k-1,q^m)$. Let H be an $\mathbb{F}_{q^m}$-hyperplane of $\mathbb{F}_{q^m}^k$, then $L_{\mathcal{U}} \cap \mathrm{PG}(H,\mathbb{F}_{q^m}) = L_{\mathcal{U}} \cap L_H = L_{\mathcal{U}\cap H}$, and hence

$$\langle L_{\mathcal{U}} \cap \mathrm{PG}(H,\mathbb{F}_{q^m})\rangle = \langle L_{\mathcal{U}} \cap L_H\rangle = \langle L_{\mathcal{U}\cap H}\rangle.$$

Moreover, for every subset $S \subseteq \mathbb{F}_{q^m}^k$, one has $L_{\langle S\rangle_{\mathbb{F}_{q^m}}} = \langle L_S\rangle$. This implies that $L_{\mathcal{U}}$ is a cutting blocking set in $\mathrm{PG}(k-1,q^m)$ if and only if for every $\mathbb{F}_{q^m}$-hyperplane H of $\mathbb{F}_{q^m}^k$ we have $L_{\langle \mathcal{U}\cap H\rangle_{\mathbb{F}_{q^m}}} = L_H$. Since H and $\langle \mathcal{U} \cap H\rangle_{\mathbb{F}_{q^m}}$ are both $\mathbb{F}_{q^m}$-linear, the linear set that they define coincide with the respective projective subspaces. Therefore $L_{\mathcal{U}}$ is a cutting blocking set in $\mathrm{PG}(k-1,q^m)$ if and only if $\langle H \cap \mathcal{U}\rangle_{\mathbb{F}_{q^m}} = H$ for every $\mathbb{F}_{q^m}$-hyperplane H in $\mathbb{F}_{q^m}^k$, as claimed. We conclude using Theorem 3.8 – which states that a linear code is Hamming-minimal if and only if the associated projective system is a cutting blocking set – and observing that, by definition, $L_{\mathcal{U}}$ is the projective system associated to $\mathcal{C}^{\mathrm{H}}$. □

Theorem 4.46 allows us to transfer results known for minimal codes in the Hamming metric to the rank metric setting. For example, the following is the rank-metric analogue of the characterization in (3.1).

Theorem 4.47. Let $\mathcal{C}$ be an $[n,k]_{q^m/q}$ code. Then $\mathcal{C}$ is minimal if and only if

$$\sum_{\lambda\in\mathbb{F}_{q^m}\setminus\{0\}} q^{-\mathrm{rk}(c+\lambda c')} \neq (q^m-1)\cdot q^{-\mathrm{rk}(c)} - q^{-\mathrm{rk}(c')} + 1$$

for all linearly independent $c, c' \in \mathcal{C}$.

Proof. By Theorem 4.46, $\mathcal{C}$ is rank-minimal if and only if any associated-Hamming metric code $\mathcal{C}^{\mathrm{H}}$ is Hamming-minimal. We can now conclude by using (3.1) and (4.10). □

Remark 4.48. It is natural to ask if the best known criterion for Hamming-minimality, namely the *Ashikhmin-Barg condition* of [19, Lemma 2.1], can be transferred to the rank-metric context. The mentioned result states that every $[n,k,d]_{q^m}$ code satisfying $w_{\max}(q^m-1) < q^m d$ is Hamming-minimal, where $w_{\max}$ denotes the maximum Hamming weight of a codeword.

One may naturally try to use the Ashikhmin-Barg condition together with Theorem 4.46 and Theorem 4.26 to obtain a sufficient condition rank-minimality. This can be done as follows.

Let $\mathcal{C}$ be a nondegenerate $[n,k,d]_{q^m/q}$ code. By Corollary 4.43, we may assume without loss of generality that $n \geq m$. By Proposition 4.11, the maximum rank of a codeword in $\mathcal{C}$ is m. Now consider the associated Hamming-metric code $\mathcal{C}^{\mathrm{H}}$. Using Theorem 4.26 we see that the minimum distance of $\mathcal{C}^{\mathrm{H}}$ is $(q^n - q^{n-d})(q-1)^{-1}$ and that the maximum Hamming weight of a codeword in $\mathcal{C}^{\mathrm{H}}$ is $(q^n - q^{n-m})(q-1)^{-1}$. Therefore imposing the Ashikhmin-Barg condition yields the following: A nondegenerate $[n,k,d]_{q^m/q}$ code is rank-minimal if

$$(q^n - q^{n-m})(q^m - 1) < q^m(q^n - q^{n-d}). \tag{4.12}$$

However, it is not difficult to see that (4.12) is only satisfied when $d = m$, that is, when $\mathcal{C}$ is the $[km,k,m]_{q^m/q}$ simplex code; see Proposition 4.17. In other words, the rank metric analogue of the Ashikhmin-Barg condition is trivial.

4.4 Minimal Rank-Metric Codes: Existence and Constructions

In this second section on minimal rank-metric codes we turn to their existence and constructions. In particular, in the light of the geometric characterization of Corollary 4.40 and of the lower bound of Corollary 4.43, we investigate the existence of short minimal codes. We start by showing some simple examples of minimal codes. Then we construct a family of 3-dimensional minimal codes using scattered linear sets, and establish the existence of minimal rank-metric codes for all $n \geq 2k + m - 2$ using a counting argument. The last part of this section is devoted to a new parameter of rank-metric codes, which we call the *linearity index* and use to investigate further the structure of minimal codes.

4.4.1 First Examples of Minimal Rank-Metric Codes

A natural question is whether a simplex rank-metric code is minimal or not. Indeed, in the Hamming-metric simplex codes are among the simplest and best known minimal codes.

Theorem 4.49. Let $\mathcal{C}$ be a $[km,k,m]_{q^m/q}$ simplex rank-metric code. Then $\mathcal{C}$ is minimal.

Proof. By the definition, any $[km,k]_{q^m/q}$ system associated to $\mathcal{C}$ is $\mathbb{F}_{q^m}^k$; see Proposition 4.17. The latter is clearly a linear cutting blocking set, since $H \cap \mathbb{F}_{q^m}^k = H$ for each $\mathbb{F}_{q^m}$-hyperplane H of $\mathbb{F}_{q^m}^k$. □

The following criterion is a sufficiency result to have a minimal rank-metric code.

Proposition 4.50. Let $\mathcal{C}$ be a nondegenerate $[n,k]_{q^m/q}$ code with $n \geq (k-1)m + 1$. Then $\mathcal{C}$ is minimal.

Proof. Let $\mathcal{U}$ be any $[n,k]_{q^m/q}$ system corresponding to $\mathcal{C}$ (up to equivalence) and let H be an $\mathbb{F}_{q^m}$-hyperplane of $\mathbb{F}_{q^m}^k$. By Proposition 4.37, we need to show that $\langle H \cap \mathcal{U} \rangle_{\mathbb{F}_{q^m}} = H$. Since H

is also an $\mathbb{F}_q$-space, we can compute the $\mathbb{F}_q$-dimension of $H \cap \mathcal{U}$ as follows:

$$\begin{aligned}
\dim_{\mathbb{F}_q}(H \cap \mathcal{U}) &= \dim_{\mathbb{F}_q}(H) + \dim_{\mathbb{F}_q}(\mathcal{U}) - \dim_{\mathbb{F}_q}(H + \mathcal{U}) \\
&= (k-1)m + n - \dim_{\mathbb{F}_q}(H + \mathcal{U}) \\
&\geq (k-1)m + (k-1)m + 1 - km \\
&= (k-2)m + 1.
\end{aligned}$$

This implies that $\langle H \cap \mathcal{U} \rangle_{\mathbb{F}_{q^m}}$ has $\mathbb{F}_{q^m}$-dimension strictly greater than $k-2$ and since it is contained in H, it has to be equal to H. □

Proposition 4.50 shows that every nondegenerate $[n,2]_{q^m/q}$ code with $n = m+1$ is minimal. This means that the bound of Corollary 4.43 is sharp for $k = 2$. It is natural to ask if the bound is sharp for other values of k. We will show in Section 4.4.2 that this happens also for $k = 3$.

4.4.2 Three-Dimensional Minimal Rank-Metric Codes

In this section we study minimal $[n,3]_{q^m/q}$ codes. In particular we prove that they exist for every $n \geq m+2$ under the assumption that $m \geq 4$. This also implies that for $k = 3$ and $m \geq 4$ the bound of Corollary 4.43 is sharp.

The first result that we provide links the existence of scattered linear sets with 3-dimensional minimal rank-metric codes.

Theorem 4.51. Let $\mathcal{C}$ be a nondegenerate $[n,3]_{q^m/q}$ code with $n \geq m+2$ and let $\mathcal{U}$ be any $[n,3]_{q^m/q}$ system corresponding to $\mathcal{C}$. If $L_{\mathcal{U}}$ is a scattered linear set, then $\mathcal{C}$ is a minimal rank-metric code.

Proof. Let $\mathcal{C}^{\mathrm{H}} \in (\Psi^{\mathrm{H}} \circ \mathrm{Ext}^{\mathrm{H}} \circ \Phi)([\mathcal{C}])$ be any Hamming-metric code associated with $\mathcal{C}$. By Theorem 4.46, $\mathcal{C}$ is rank-minimal if and only if $\mathcal{C}^{\mathrm{H}}$ is Hamming-minimal, which is in turn equivalent to the fact that $L_{\mathcal{U}}$ is a cutting blocking set in $\mathrm{PG}(2,q^m)$. Consider now the multiplicity function associated to $L_{\mathcal{U}}$ in the projective $[\frac{q^n-1}{q-1}, k]_{q^m}$ system $\mathrm{Ext}^{\mathrm{H}}(\mathcal{U})$. Since $L_{\mathcal{U}}$ is scattered, this means that every point of $L_{\mathcal{U}}$ has multiplicity 1. Let G be any generator matrix of $\mathcal{C}^{\mathrm{H}}$, and let $v \in \mathbb{F}_{q^m}^3 \setminus \{0\}$. Since by Proposition 4.11 the maximum rank of a codeword in $\mathcal{C}$ is m, using Theorem 4.26 we get

$$\mathrm{wt}^{\mathrm{H}}(vG) \leq \frac{q^n - q^{n-m}}{q-1}.$$

Thus,

$$|L_{\mathcal{U}} \cap \langle v \rangle^{\perp}| = \frac{q^n-1}{q-1} - \mathrm{wt}^{\mathrm{H}}(vG) \geq \frac{q^{n-m}-1}{q-1} \geq q+1.$$

In particular, $L_{\mathcal{U}}$ is a $(q+1)$-fold blocking set, and in $\mathrm{PG}(2,q^m)$ this also implies that $L_{\mathcal{U}}$ is cutting. □

Thanks to Theorem 4.51, the existence of certain minimal rank-metric codes reduces to the existence of certain scattered linear sets. There is a well known upper bound on the parameters of these objects, due to Blokhuis and Lavrauw; see [40]. If $\mathcal{U}$ is a $[n,k]_{q^m/q}$ system such that $L_\mathcal{U}$ is scattered, then

$$n \leq \frac{km}{2}. \tag{4.13}$$

In this context, much progress has been made in the study of *maximum scattered linear sets*, which are linear sets whose parameters meet the bound in (4.13) with equality. A construction of such linear sets was first provided by Blokhuis and Lavrauw for k even; see [40]. When instead k is odd and m is even, a construction of linear sets meeting (4.13) for infinitely many parameters was given by Bartoli, Giulietti, Marino and Polverino in [30, Theorem 1.2]. The picture was then completed by Csajbók, Marino, Polverino and Zullo; see [56].

Theorem 4.52 (see [56, Theorem 2.4])**.** Assume that km is even. Then there exists a $[\frac{km}{2},k]_{q^m/q}$ system such that $L_\mathcal{U}$ is scattered.

When km is odd, not much is known yet. One of the few existence results on the maximum rank of a scattered linear set is the following, due to Blokhuis and Lavrauw.

Theorem 4.53 (see [40, Theorem 4.4])**.** Let k,m be positive integers and q be a prime power. There exists an $[ab,k]_{q^m/q}$ system such that $L_\mathcal{U}$ is scattered, whenever a divides k, $\gcd(a,m)=1$ and

$$ab < \begin{cases} \frac{km-k+3}{2} & \text{if } q=2 \text{ and } a=1, \\ \frac{km-k+a+3}{2} & \text{otherwise.} \end{cases}$$

In contrast with the most common line of research in the theory of scattered linear set, in this thesis we are primarily interested in short nondegenerate minimal codes, and thus in linear sets with small rank. For this reason, we state the following simple lemma, whose proof is omitted.

Lemma 4.54. Let $\mathcal{U}$ be an $[n,k]_{q^m/q}$ system such that $L_\mathcal{U}$ is a scattered linear set. If $n>k$, then there exists an $[n-1,k]_{q^m/q}$ system $\mathcal{V}\subseteq\mathcal{U}$ such that $L_\mathcal{V}$ is scattered.

We conclude this subsection by combining the previous three results with each other. This yields the following existence theorem for 3-dimensional minimal rank-metric codes.

Theorem 4.55. Suppose that $m \not\equiv 3,5 \mod 6$ and $m \geq 4$. Then there exists a (nondegenerate) minimal $[m+2,3]_{q^m/q}$ code.

Proof. Observe that by Theorem 4.51 it is enough to prove that there exists an $[m+2,3]_{q^m/q}$ system $\mathcal{U}$ such that $L_\mathcal{U}$ is scattered.

First, assume that m is even. Then, by Theorem 4.52, we have that there exists a $[\frac{3m}{2},3]_{q^m/q}$ system such that $L_\mathcal{U}$ is scattered. Then, since $m+2 \leq \frac{3m}{2}$ whenever $m \geq 4$, using Lemma 4.54 we obtain the desired $[m+2,3]_{q^m/q}$ system.

Now assume that m is odd and $m \not\equiv 0 \mod 3$. Write $m = 3s+i$. We use Theorem 4.53 with $a = 3$ and $b = s+1$, which shows the existence of an $[\frac{m+3-i}{2}, 3]_{q^m/q}$ system $\mathcal{U}$ such that $L_{\mathcal{U}}$ is scattered. If $m \equiv 1 \mod 3$, we get the desired result. □

Remark 4.56. In the remaining cases, finding scattered linear sets of rank $m+2$ in $\mathrm{PG}(2, q^m)$ seems in general a difficult task. For instance, when $m = 5$, the existence of a $[7, 3]_{q^5/q}$ system $\mathcal{U}$ defining a scattered linear set was recently shown in [29, Theorem 5.1], but only in characteristic 2, 3 and 5 and under some restriction on the field size.

4.4.3 Existence Results for Minimal Rank-Metric Codes

In this subsection we establish a general existence result for minimal rank-metric codes. We prove that minimal rank-metric codes exist for all parameter sets (n, m, k) with $m \geq 2$ and $n \geq 2k + m - 2$ (and any q). Combining this with previous results, we then give parameter intervals for which nondegenerate minimal codes exist and do not exist.

Lemma 4.57. Let m, n, k be positive integers and suppose $n \geq k \geq 2$. If

$$\frac{(q^{mn}-1)(q^{m(n-1)}-1)}{(q^{mk}-1)(q^{m(k-1)}-1)} - \frac{1}{2}\sum_{i=2}^{m} \frac{1}{q^m-1}\binom{m}{i}_q \prod_{j=0}^{i-1}(q^n - q^j)\left(\frac{q^{mi}-1}{q^m-1} - 1\right) \tag{4.14}$$

is positive, then there exists a minimal $[n,k]_{q^m/q}$ code.

Proof. We use an argument inspired by the methods of [82] but which is simpler and avoids the graph theory language. Form a set of representatives for the equivalence classes of nonzero vectors in $\mathbb{F}_{q^m}^n$. Call this set $\mathcal{Q}$ and let

$$\mathcal{P} = \{P = \{x, y\} \subseteq \mathcal{Q} \,:\, x \neq y,\, \sigma^{\mathrm{rk}}(x) \subseteq \sigma^{\mathrm{rk}}(y) \text{ or } \sigma^{\mathrm{rk}}(y) \subseteq \sigma^{\mathrm{rk}}(x)\}.$$

The $[n,k]_{q^m/q}$ non-minimal codes are the k-dimensional subspaces $\mathcal{C} \subseteq \mathbb{F}_{q^m}^n$ such that $P \subseteq \mathcal{C}$ for some $P \in \mathcal{P}$. Their number is at most

$$\sum_{P \in \mathcal{P}} |\{\mathcal{C} \subseteq \mathbb{F}_{q^m}^n \,:\, \mathcal{C} \supseteq P\}| = |\mathcal{P}|\binom{n-2}{k-2}_q.$$

Therefore, the minimal $[n,k]_{q^m/q}$ codes are at least

$$\binom{n}{k}_{q^m} - |\mathcal{P}|\binom{n-2}{k-2}_{q^m} = \binom{n-2}{k-2}_{q^m}\left(\frac{(q^{mn}-1)(q^{m(n-1)}-1)}{(q^{mk}-1)(q^{m(k-1)}-1)} - |\mathcal{P}|\right).$$

In particular, a minimal $[n,k]_{q^m/q}$ code exists if

$$\frac{(q^{mn}-1)(q^{m(n-1)}-1)}{(q^{mk}-1)(q^{m(k-1)}-1)} - |\mathcal{P}| > 0. \tag{4.15}$$

Finally, we count the elements of $\mathcal{P}$ as

$$\begin{aligned} 2|\mathcal{P}| &= \sum_{i=1}^{m} |\{(x,y) \in \mathcal{Q}^2 \, : \, x \neq y, \, \mathrm{rk}(y) = i, \, \sigma^{\mathrm{rk}}(x) \subseteq \sigma^{\mathrm{rk}}(y)\}| \\ &= \sum_{i=1}^{m} \sum_{\substack{y \in \mathcal{Q} \\ \mathrm{rk}(y)=i}} |\{x \in \mathcal{Q} \, : \, x \neq y, \, \sigma^{\mathrm{rk}}(x) \subseteq \sigma^{\mathrm{rk}}(y)\}| \\ &= \sum_{i=1}^{m} \frac{1}{q^m - 1} \binom{m}{i}_q \prod_{j=0}^{i-1} (q^n - q^j) \left(\frac{q^{mi} - 1}{q^m - 1} - 1 \right) \\ &= \sum_{i=2}^{m} \frac{1}{q^m - 1} \binom{m}{i}_q \prod_{j=0}^{i-1} (q^n - q^j) \left(\frac{q^{mi} - 1}{q^m - 1} - 1 \right). \end{aligned}$$

Combining this with (4.15) concludes the proof. □

We now give a sufficient condition under which the assumption in Lemma 4.57 is satisfied. This gives us parameter ranges for which minimal codes exist. The next result does not depend on the field size q. This behaviour of minimal rank-metric codes is in sharp contrast with analogous results for minimal codes in the Hamming metric; see e.g. [7, Theorem 2.14].

Corollary 4.58. For every $m, k \geq 2$, there exists a minimal $[2k+m-2, k]_{q^m/q}$ code.

Proof. Fix an integer $n \geq k$ and observe that

$$\frac{(q^{mn} - 1)(q^{m(n-1)} - 1)}{(q^{mk} - 1)(q^{m(k-1)} - 1)} \geq q^{mn + m(n-1) - mk - m(k-1)} = q^{2m(n-k)}.$$

Therefore the quantity in (4.14) can be bounded from below as follows:

$$\begin{aligned} (4.14) &\geq q^{2m(n-k)} - \frac{1}{2(q^m - 1)^2} \sum_{i=2}^{m} \binom{m}{i}_q \cdot q^{\binom{i}{2}} \cdot (q^{mi} - q^m) \prod_{j=0}^{i-1} (q^{n-j} - 1) \\ &> q^{2m(n-k)} - \frac{1}{2(q^m - 1)^2} \sum_{i=2}^{m} \binom{m}{i}_q \cdot q^{\binom{i}{2}} \cdot q^{mi} \prod_{j=0}^{i-1} q^{n-j} \\ &= q^{2m(n-k)} - \frac{1}{2(q^m - 1)^2} \sum_{i=2}^{m} \binom{m}{i}_q \cdot q^{i(m+n)} =: t_q(m,n,k). \end{aligned}$$

Define the function

$$f(q) := \prod_{i=1}^{\infty} \frac{q^i}{q^i - 1}.$$

In the sequel, we will use the following estimates:

$$\binom{a}{b}_q < f(q)\, q^{b(a-b)}, \qquad \text{for } a,b \in \mathbb{N}, \tag{4.16}$$

$$q^{e_1} + \ldots + q^{e_r} \leq \frac{q}{q-1} q^{e_r}, \qquad \text{for } e_i \in \mathbb{Z},\ 0 \leq e_1 < \ldots < e_r. \tag{4.17}$$

We have

$$\begin{aligned} 2(q^m-1)^2 t_q(m,n,k) &= 2(q^m-1)^2 q^{2m(n-k)} - q^{m(m+n)} - \sum_{i=2}^{m-1} \binom{m}{i}_q q^{i(m+n)} \\ &\overset{(4.16)}{>} 2(q^m-1)^2 q^{2m(n-k)} - q^{m(m+n)} - f(q) \sum_{i=2}^{m-1} q^{i(2m+n-i)} \\ &\overset{(4.17)}{>} 2(q^m-1)^2 q^{2m(n-k)} - q^{m(m+n)} - \frac{qf(q)}{q-1} q^{(m-1)(m+n-1)}. \\ &> 2(q^m-1)^2 q^{2m(n-k)} - q^{m(m+n)} - q^{(m-1)(m+n-1)+3}, \end{aligned} \tag{4.18}$$

where the last inequality follows from the fact that $f(q) < 4$ for every prime power q.

We now specialize the argument to $n = 2k+m-2$, proving that $t_q(m, 2k+m-2, k) > 0$ for every $m, k \geq 2$ and prime power q. Using (4.18) we find

$$\begin{aligned} 2(q^m-1)^2 t_q(m, 2k+m-2, k) &> 2(q^m-1)^2 q^{2m(m+k-2)} - q^{2m(m+k-1)} - q^{(m-1)(2m+2k-3)+3} \\ &= 2(q^m-1)^2 q^{2m(m+k-2)} - (1 + q^{-3m-2k+6}) q^{2m(m+k-1)} \\ &\geq 2(q^m-1)^2 q^{2m(m+k-2)} - (1+q^{-4}) q^{2m(m+k-1)} \\ &= q^{2m(m+k-2)-4} \left(2(q^m-1)^2 q^4 - (q^4+1) q^{2m}\right) \end{aligned}$$

Hence $t_q(m, 2k+m-2, k) > 0$ whenever $q^{2m+4} - 4q^{m+4} - q^{2m} + 2q^4 \geq 0$, which holds for every $m \geq 2$ and every prime power q. Therefore there exists a minimal $[2k+m-2, k]_{q^m/q}$ code by Lemma 4.57. □

Remark 4.59. Fix integers $k, m \geq 2$. Then Corollary 4.43 tells us that for any length value $n < k+m-1$ an $[n,k]_{q^m/q}$ minimal code cannot exist, for any field size q. On the other hand, by Corollary 4.58 for $n \geq 2k+m-2$ there exist $[n,k]_{q^m/q}$ minimal codes for every field size q. Therefore the existence of $[n,k]_{q^m/q}$ minimal codes remains in general an open question only for $k+m-1 \leq n \leq 2k+m-3$.

4.4.4 The Linearity Index of a q-System

Given an $[n,k]_{q^m/q}$ system $\mathcal{U}$, one could be interested in understanding how $\mathcal{U}$ is related to $\mathbb{F}_{q^m}$-subspaces of $\mathbb{F}_{q^m}^k$ and not only to $\mathbb{F}_{q^m}$-hyperplanes. This indeed could reveal some additional information on its parameters and whether it can be a linear cutting blocking set or not. In this subsection we define and analyze a new parameter of a projective system and with its aid we generalize the lower bound in Corollary 4.43 for the length of minimal codes.

Let $\mathcal{U}$ be a $[n,k]_{q^m/q}$ system. We introduce a measure for the "linearity" of $\mathcal{U}$ over $\mathbb{F}_{q^m}$. More precisely, we define the **linearity index** of $\mathcal{U}$ as

$$\ell(\mathcal{U}) = \max\{\dim_{\mathbb{F}_{q^m}}(H) \,:\, H \subseteq \mathbb{F}_{q^m}^k \text{ is an } \mathbb{F}_{q^m}\text{-subspace, } H \subseteq \mathcal{U}\}.$$

Observe that the value of $\ell(\mathcal{U})$ is invariant under equivalence of $[n,k]_{q^m/q}$ systems. In particular, it is a well-defined structural parameter of the corresponding equivalence class $[\mathcal{U}]$. The following result relates $\ell(\mathcal{U})$ to the generalized rank weight of a code that gives rise to the q-system $\mathcal{U}$.

Lemma 4.60. Let $\mathcal{C}$ be a nondegenerate $[n,k]_{q^m/q}$ code, and let $\mathcal{U}$ be any corresponding $[n,k]_{q^m/q}$ system. Then

$$\ell(\mathcal{U}) = k - \min\{r \,:\, d_r(\mathcal{C}) = n - (k-r)m\}.$$

Proof. First of all, note that the set $\{r \,:\, d_r(\mathcal{C}) = n-(k-r)m\}$ is nonempty, since $d_k(\mathcal{C}) = n$. By Theorem 4.14,

$$d_r(\mathcal{C}) = n - \max\left\{\dim_{\mathbb{F}_q}(\mathcal{U}\cap H) \,:\, H \text{ is an } \mathbb{F}_{q^m}\text{-subspace of codimension } r \text{ in } \mathbb{F}_{q^m}^k\right\},$$

which is equal to $n-(k-r)m$ if and only if there exists $H \subseteq \mathbb{F}_{q^m}^k$ of codimension r contained in $\mathcal{U}$. □

Lemma 4.60 shows that the parameter ℓ is well-defined in the correspondence of Theorem 4.8. Hence, it is also a well-defined parameter of a nondegenerate code $\mathcal{C}$. Therefore, we will also refer to it as the **linearity index of a code** $\mathcal{C}$, and denote it by $\ell(\mathcal{C})$.

Lemma 4.61. Let $\mathcal{C}$ be a nondegenerate $[n,k]_{q^m/q}$ code with linearity index ℓ. Then, $d_{i+1}(\mathcal{C}) - d_i(\mathcal{C}) = m$ if and only if $i \geq k-\ell(\mathcal{C})$.

Proof. $\underline{(\Leftarrow)}$ This implication follows from the definition of $\ell(\mathcal{C})$.

$\underline{(\Rightarrow)}$ Let $\mathcal{U}$ be any $[n,k]_{q^m/q}$ system associated to $\mathcal{C}$. Let $H \subseteq \mathbb{F}_{q^m}^k$ be a space of codimension i such that $d_i(\mathcal{C}) = n - \dim_{\mathbb{F}_q}(H\cap\mathcal{U})$ and let $t := \dim_{\mathbb{F}_q}(H\cap\mathcal{U})$. Let $\mathcal{V} := H\cap\mathcal{U}$ and observe that $|H'\cap(\mathcal{V}\setminus\{0\})| = (q^{t-m}-1)$ for any hyperplane H' in H. Let Λ be the set of all hyperplanes in H, then by Lemma 4.6, we have

$$\sum_{H'\in\Lambda} |H'\cap(\mathcal{V}\setminus\{0\})| = \binom{k-i}{1}_{q^m}(q^{t-m}-1).$$

Moreover, observe that every nonzero element of $\mathcal{V}$ belongs to exactly $\binom{k-i-1}{1}_{q^m}$ hyperplanes in H. Hence,

$$\sum_{H'\in\Lambda} |H'\cap(\mathcal{V}\setminus\{0\})| = \sum_{v\in\mathcal{V}\setminus\{0\}} |\{H' : v\in H'\}| = \binom{k-i-1}{1}_{q^m}(q^t-1).$$

By a double counting argument, we then have that

$$(q^{(k-i)m}-1)(q^{t-m}-1) = (q^t-1)(q^{(k-i-1)m}-1).$$

By Lemma 4.16, it follows that $t=(k-i)m$. In particular, $\mathcal{V}$ is an $[km,k]_{q^m/q}$ system associated to a simplex code. We conclude then that $H\subseteq\mathcal{V}$ and then $H\subseteq\mathcal{U}$, which implies that $\ell(\mathcal{C})\geq k-i$. □

Proposition 4.62. Let $\mathcal{C}$ be a nondegenerate $[n,k]_{q^m/q}$ code. Then

$$\ell(\mathcal{C})\geq n-k(m-1).$$

Proof. Observe that $\sum_{i=0}^{k-1} d_{i+1}(\mathcal{C})-d_i(\mathcal{C}) = d_k(\mathcal{C})-d_0(\mathcal{C}) = n$. Moreover, by applying Lemma 4.61, we have that

$$\begin{aligned}\sum_{i=0}^{k-1} d_{i+1}(\mathcal{C})-d_i(\mathcal{C}) &= \sum_{i=0}^{k-\ell(\mathcal{C})-1} d_{i+1}(\mathcal{C})-d_i(\mathcal{C}) + \sum_{i=k-\ell(\mathcal{C})}^{k-1} d_{i+1}(\mathcal{C})-d_i(\mathcal{C})\\ &\leq (m-1)(\ell(\mathcal{C})-1)+m\ell(\mathcal{C}) = m(k-1)+\ell(\mathcal{C})-m.\end{aligned}$$ □

The linearity index of a code can help in characterizing and finding improved bounds on the other parameters of a minimal $[n,k]_{q^m/q}$ code.

Lemma 4.63. Let $\mathcal{U}$ be a linear cutting $[n,k]_{q^m/q}$ blocking set. Suppose that there exists an ℓ-dimensional $\mathbb{F}_{q^m}$-subspace T of $\mathbb{F}_{q^m}^k$ such that $T\subseteq\mathcal{U}$. Then $\mathcal{U}/T$ is isomorphic to a linear cutting $[n-\ell m,k-\ell]_{q^m/q}$ blocking set.

Proof. By Proposition 4.37, we need to show that for every $\mathbb{F}_{q^m}$-hyperplane $\bar{H}$ of $\mathbb{F}_{q^m}^k/T$ we have $\langle\bar{H}\cap\mathcal{U}/T\rangle_{\mathbb{F}_{q^m}}=\bar{H}$. The $\mathbb{F}_{q^m}$-hyperplanes of $\mathbb{F}_{q^m}^k/T$ correspond to the $\mathbb{F}_{q^m}$-hyperplanes of $\mathbb{F}_{q^m}^k$ that contain T. Let $\bar{H}$ be an $\mathbb{F}_{q^m}$-hyperplane of $\mathbb{F}_{q^m}^k/T$. Then there exists an $\mathbb{F}_{q^m}$-hyperplane H of $\mathbb{F}_{q^m}^k$ such that $\bar{H}=H/T$. Hence,

$$\langle\bar{H}\cap\mathcal{U}/T\rangle_{\mathbb{F}_{q^m}} = \langle(H\cap\mathcal{U})/T\rangle_{\mathbb{F}_{q^m}} = \langle H\cap\mathcal{U}\rangle_{\mathbb{F}_{q^m}}/T = H/T = \bar{H},$$

where the second last equality follows from the fact that $\mathcal{U}$ is a linear cutting $[n,k]_{q^m/q}$ blocking set and by Proposition 4.37. □

The following result is a generalization of Corollary 4.43.

Proposition 4.64. Let $\mathcal{U}$ be a linear cutting $[n,k]_{q^m/q}$ blocking set and let ℓ be its linearity index. If $k-\ell \geq 2$, then

$$n-k \geq (\ell+1)(m-1).$$

In particular, for every $1 \leq r \leq k - \lfloor \frac{n-k+1}{m-1} \rfloor - 1$, we have $d_r(\mathcal{C}) > n - rm$, where $\mathcal{C}$ is the nondegenerate $[n,k]_{q^m/q}$ code associated to $\mathcal{U}$.

Proof. Let $T \subseteq \mathbb{F}_{q^m}^k$ be an ℓ-dimensional $\mathbb{F}_{q^m}$-subspace contained in $\mathcal{U}$. By Lemma 4.63 we have that $\mathcal{U}/T$ is isomorphic to a linear cutting $[n-\ell m, k-\ell]_{q^m/q}$ blocking set. Therefore, by Corollary 4.43, we obtain

$$n-\ell m \geq k-\ell+m-1,$$

from which we derive the desired inequality.
For the second part, if ℓ does not satisfy the above inequality, i.e. if $\ell \geq \lfloor \frac{n-k+1}{m-1} \rfloor + 1$, then $\mathcal{U}$ cannot contain any ℓ-dimensional $\mathbb{F}_{q^m}$-subspace. This is equivalent to say that $d_{k-\ell}(\mathcal{C}) > n-(k-\ell)m$. □

Remark 4.65. As a consequence of 4.64, it can be immediately seen that in order to construct short minimal rank-metric code, one has to try to construct linear cutting blocking sets not containing $\mathbb{F}_{q^m}$-subspaces. This is also consistent with the construction of minimal $[m+2,3]_{q^m/q}$ codes provided in Section 4.4.2. Indeed, if a $[n,k]_{q^m/q}$ system $\mathcal{U}$ contains a $\mathbb{F}_{q^m}$-subspace H, then in the associated linear set $L_\mathcal{U}$ one has $\mathrm{wt}_\mathcal{U}(P) = m$ for every $P \in \mathrm{PG}(H, \mathbb{F}_{q^m})$. In particular, the associated linear set is far from being scattered.

Proposition 4.64 allows to characterize nondegenerate $[(k-1)m,k]_{q^m/q}$ minimal codes.

Corollary 4.66. Let $k \geq 2$ and $\mathcal{C}$ be a nondegenerate $[(k-1)m,k]_{q^m/q}$ code with linearity index $\ell = \ell(\mathcal{C})$. The following are equivalent.

1. $\mathcal{C}$ is minimal.
2. $\ell < k-2$.
3. $d_2(\mathcal{C}) > m$.

Proof. $\underline{(1) \Rightarrow (2)}$: First observe that ℓ can not be equal to $k-1$. Indeed, if $\ell = k-1$, then the code $\mathcal{C}$ is not k-dimensional. Hence, $k-\ell \geq 2$ and since $\mathcal{C}$ is minimal, then by Proposition 4.64 it holds

$$(k-1)m-k+1 > (\ell+1)(m-1),$$

from which we deduce $\ell < k-2$.

$\underline{(2) \Rightarrow (1)}$: Suppose $\mathcal{C}$ is not minimal and let $\mathcal{U}$ be any associated $[(k-1)m,k]_{q^m/q}$ system. By Proposition 4.37, there exists an $\mathbb{F}_{q^m}$-hyperplane of $\mathbb{F}_{q^m}^k$ such that $\langle H \cap \mathcal{U} \rangle_{\mathbb{F}_{q^m}} =: H'$, with

$\dim_{\mathbb{F}_{q^m}}(H') \leq k-2$. Hence, we obtain

$$\begin{aligned}(k-2)m \geq \dim_{\mathbb{F}_q}(H') &\geq \dim_{\mathbb{F}_q}(H \cap \mathcal{U}) \\ &= \dim_{\mathbb{F}_q}(H) + \dim_{\mathbb{F}_q}(\mathcal{U}) - \dim_{\mathbb{F}_q}(\mathcal{U}+H) \\ &\geq (k-1)m + (k-1)m - km = (k-2)m.\end{aligned}$$

Hence, all the inequalities above are equalities and $H' = H \cap \mathcal{U}$. This implies that $\mathcal{U}$ contains the $(k-2)$-dimensional $\mathbb{F}_{q^m}$-subspace H' and $\ell \geq k-2$.

$\underline{(2) \Leftrightarrow (3)}$: Observe that $\ell \leq k-2$ and $d_2(\mathcal{C}) \geq m$. Then, the statement directly follows from Lemma 4.60. □

Corollary 4.67. Let $k \geq 3$ be an integer. A nondegenerate $[(k-1)m, k]_{q^m/q}$ minimal code exists if and only if $m \geq 3$.

Proof. Suppose that $m \geq 3$ and construct the $[(k-3)m+3(m-1), k]_{q^m/q}$ system $\mathcal{U}'$ as follows. Take $\mathcal{V}' = \langle \alpha^i e_j \,:\, 0 \leq i \leq m-2, k-2 \leq j \leq k\rangle$, where $\alpha \in \mathbb{F}_{q^m}$ is such that $\mathbb{F}_q(\alpha) = \mathbb{F}_{q^m}$. Then, consider $\mathcal{U}' = \{(\, v \mid 0,0,0) \,:\, v \in \mathbb{F}_{q^m}^{k-3}\} \oplus \mathcal{V}'$. By construction $\ell(\mathcal{U}') = k-3$. Moreover, since $m \geq 3$ then $(k-3)m + 3(m-1) \geq (k-1)m$, and we can take any $(k-1)m$-dimensional $\mathbb{F}_q$-subspace $\mathcal{U}$ of $\mathcal{U}'$, which has $\ell(\mathcal{U}) \leq k-3$ and by Corollary 4.66 is minimal.

Assume now $m \leq 2$, and let $\mathcal{C}$ be a nondegenerate a $[(k-1)m, k]_{q^m/q}$ code. Then by Proposition 4.62, we have $\ell(\mathcal{C}) \geq k-m \geq k-2$. Hence, by Corollary 4.66, $\mathcal{C}$ is not minimal. □

Part II

Convolutional Codes

Chapter 5

Introduction

The main aim of coding theory is that of communication through a noisy channel, i.e., a sender wants to send a message (or a sequence of messages) through a channel that possibly adds random noise to the message. Error-correcting codes acts by adding redundancy to the message, in such a way that with a suitable decoding algorithm, it can be corrected.

Convolutional codes were introduced in 1955 by Peter Elias, in his seminal paper [67]. They can be considered as a generalization of the classical block codes to the polynomial setting.

To stress the difference among these two class of error-correcting codes and motivate the generalization to convolutional codes, consider a linear $[n,k]_q$ code $\mathcal{C}$ generated by the matrix $G \in \mathbb{F}_q^{k\times n}$. When a sequence of messages $m_i \in \mathbb{F}_q^k$ has to be encoded via $\mathcal{C}$, the transmitted codewords will be $c_i = m_i G \in \mathbb{F}_q^n$.

Instead of using the constant matrix G as an encoding map, Elias suggested to use more general polynomial matrices of the form $G(z)$ whose entries turns to be elements of the polynomial ring $\mathbb{F}_q[z]$. Formally, a convolutional code can be defined as an $\mathbb{F}_q[z]$-module of $\mathbb{F}_q[z]^n$, generated by the rows of a polynomial matrix $G(z) \in \mathbb{F}_q[z]^{k\times n}$. This allows to consider the information as a whole sequence, then split into blocks of equal length, but the encoded block at any given time depends not only on the information block at that time, but also on a fixed number of previous information blocks. This way, different to block codes, a fixed block will not always be encoded into the same codeword, depending on the position of the block in the whole sequence.

In 1967, Massey and Sain pointed out natural connections to automata theory and systems theory; see [113]. For more details on this connections, the interested reader is referred to the survey [133].

In the next years, Forney developed a mathematical theory of convolutional codes, in which they were defined as k-dimensional linear subspaces of the n-dimensional vector space $\mathbb{F}_q((z))^n$, where $\mathbb{F}_q((z))$ is the field of formal Laurent series; see [71, 70, 73].

Convolutional codes have been extensively implemented in practice with applications in mobile and satellite communication and data streaming. In particular, convolutional codes have been widely investigated over the erasure channel. When considering the erasure channel, which is the most used channel in multimedia traffic, convolutional codes can correct more errors than block codes. An erasure channel is a communication channel where parts of the information sequence are either received or erased and the decoder always knows where the erased parts occured. The advantage of convolutional codes for this type of channel is their flexibility of grouping the blocks of information in an appropriate way, depending on the erasures location, and then decode the part of the sequence with less erasures or where the distribution of erasures allows a complete correction first.

In this part of the thesis, we focus on what we define *algebraic* theory of error-correcting codes, that, in other words, is the theory of algebraic constructions of codes which can correct as many errors as possible and that have an efficient decoding algorithm. The parameter of a block code that determines its error correction capability is the *minimum distance*.

For a convolutional code, there exist different notions of distances. The first one is the *free distance*, which is the analogue of the Hamming distance for two polynomial vectors. The other notion is the one of *column distances*, which measures the error-correction capabilities of the code within a given time interval. We will define them regorously in Chapter 6. Convolutional codes whose column distances increase as rapidly as possible for as long as possible are called *maximum distance profile* (MDP) codes and they have the ability to correct a maximal number of errors per time interval. In [79] an algebraic characterization of MDP convolutional codes over finite fields was given, based on a generator matrix having as structural property the so called *left primeness*. As MDP convolutional codes have the maximal possible growth in the column distances, they can correct the maximal number of errors in a time interval, and therefore are similar to *maximum distance separable* (MDS) block codes within windows of fixed size. However, in contrast to the case of MDS block codes, there are very few algebraic constructions of MDP convolutional codes, all based on a characterization provided in [79]. From a practical point of view, it has been recently shown that MDP codes are very appealing for sequential transmission over the erasure channel and low-delay streaming applications; see [20, 69, 110, 153].

Surprisingly, there exist only two general algebraic constructions that yield two wide classes of MDP convolutional codes (see [12] and [79]), but both require unpractical large field sizes. These classes are built using lower triangular Toeplitz *superregular matrices*, i.e. matrices having the property that the minors that are not *trivially zero* are nonzero; see [13, 79] for a formal definition and details on the relation between superregular matrices and MDP convolutional codes. Due to the difficulty of deriving general constructions, researchers have been focusing on computer search algorithms for finding MDP convolutional codes with small parameters; see [84, 14, 79, 103, 153].

The aim of this part of the thesis is to to combine the existing literature on convolutional codes with the personal contributions in [8, 11, 10], in order to provide a complete and original overview on the topic

Organization In Chapter 6, we provide the background needed for the understanding of this part of the thesis. Chapter 7 is based on the publication [8], by Alfarano and Lieb. We show that if $H(z)$ is a parity-check matrix (resp. $G(z)$ is a generator matrix) of an (n, k, δ) convolutional code $\mathcal{C}$, where $n-k$ divides δ or k divides δ and such that the criterion on the minors of the truncated sliding parity-check matrix H_L^c (resp. generator matrix G_L^c) of $\mathcal{C}$ is satisfied, then $H(z)$ (resp. $G(z)$) is left prime. Observe that if $n-k$ divides δ, we consider the parity-check matrix and if k divides δ, we consider the generator matrix. If k divides δ, our result implies that all (n, k, δ) MDP convolutional codes are necessarily noncatastrophic. If $n-k$ divides δ, it implies that a polynomial matrix $H(z)$ that fulfills the criterion is a parity-check matrix of a convolutional code whose degree equals the sum of the row degrees of $H(z)$ (and of course is noncatastrophic as it has a parity-check matrix). While preparing this thesis, we realized that in the published paper there is a computation mistake which we correct here.

Chapter 8, we survey the results provided in [11], by Alfarano, Napp, Neri and Requena. We use a different approach to derive large classes of MDP convolutional codes and present a new general algebraic construction. Rather than using superregular matrices or generator polynomials of cyclic or quasi-cyclic block codes, we carefully select different modified Vandermonde matrices as the coefficients of the polynomial generator matrix of the convolutional code in such a way that the resulting code is, under some constraints, MDP. Since each modified Vandermonde matrix is the generator (or parity-check) matrix of a generalized Reed-Solomon (GRS) block code, the presented class of codes can be considered as a very natural extension of GRS block codes to the context of convolutional codes. For this reason, we call them *weighted Reed-Solomon (WRS) convolutional codes*. We show that the field size required to build them is significantly smaller than the existing ones in the literature for other classes of MDP codes.

Finally, Chapter 9 is based on [10], by Alfarano, Lieb and Rosenthal. This last part of the thesis is devoted to present a combinatorial construction of low-density parity-check convolutional codes using difference triangle sets. In the last three decades, the area of channel coding gained a lot of attention, due to the fact that many researchers were attracted by the practical realization of coding schemes whose performances approach the Shannon limit. This revolution started in 1993 with the invention of turbo codes and their decoding algorithms [35]. Only few years later, researchers investigated also low-density parity-check (LDPC) block codes and their message passing decoding algorithm. These codes were discovered to be also capable of capacity-approaching performances. The class of LDPC block codes was introduced by Gallager [75], in 1962. Their name is due to the fact that they have a parity-check matrix that is sparse. The analysis of LDPC codes attracted many researchers and a lot of work arose in this direction, starting from the papers of Wiberg [160] and Mackay and Neal [108]. Moreover, in [131, 53] analytical tools were introduced to investigate the limits of the performance of the message passing iterative decoding algorithm, suggested by Tanner already in 1981, [149]. Similarly to LDPC block codes, one can consider LDPC convolutional codes. These codes are defined as the (right) kernel of a sparse sliding parity-check matrix, which allows to still use iterative message passing decoding algorithms. Moreover, it was proven that LDPC convolutional codes are practical in different communication applications, see for instance [122, 32, 31].

In the last few years, some attempts to construct binary LDPC convolutional codes were done. Two types of constructions were mainly investigated. The first one exploits the similarity of quasi-cyclic block codes and time-invariant LDPC convolutional codes, [150, 151, 152]. The second one regards mostly time varying convolutional codes, see for instance [163, 128, 33].

The aim of Chapter 9 is to give a combinatorial construction of LDPC convolutional codes suitable for iterative deoding. In fact, contrary to LDPC block codes for which a lot of combinatorial constructions have been derived (see for example [136, 96, 97, 90, 158]), it is rare to use combinatorial tools for constructing LDPC convolutional codes.

In 1967, Robinson and Bernstein [132] used difference triangle sets for the first time to construct binary recurrent codes, which are defined as the (right) kernel of a binary sliding matrix. At that time, the theory of convolutional codes was not developed yet and the polynomial notation was not used, but now, we may regard recurrent codes as a first prototype of convolutional codes. This was the first time that a combinatorial object has been used to construct convolutional codes. Three years later, Tong in [154], used diffuse difference triangle sets to construct self-orthogonal diffuse convolutional codes, defined by Massey [114]. The aim of these authors was to construct codes suitable for iterative decoding and their result was an adapted version of binary LDPC convolutional codes. In [9], the authors constructed $(n, n-1)_q$ LDPC convolutional codes, whose sliding parity-check matrix is free from 4 and 6-cycles not satisfying the so called full rank condition, starting from difference triangle sets. This was a generalization of the work of Robinson and Bernstein, in which difference triangle sets were used to construct convolutional codes over the binary field, that can only avoid 4-cycles. In 1971, Tong [155] was the first to generalize their construction over $\mathbb{F}_q$, using what we call *weak* difference triangle sets. However, his construction is suitable only for limited rate and in a way that the Tanner graph associated to the parity-check matrix of these codes is free only from 4-cycles.

We give a construction of LDPC convolutional codes for arbitrary rates over arbitrary fields, using difference triangle sets and weak difference triangle sets. In particular, the use of the weak version of these combinatorial objects allows to relax the assumptions required by Robinson, Bernstein and Tong. Indeed, instead of considering sets of nonnegative integers where all the pairwise differences are distinct among all the sets, we may require only that the pairwise differences are distinct in each set. Moreover, we show that using difference triangle sets for this construction produces codes with good distance properties and we provide a bound on the field size that is sufficient to have codes with good distance and to avoid the presence of cycles not satisfying the full rank condition.

Chapter 6

Preliminaries on Convolutional Codes

In this chapter we provide the preliminaries to the first part of the thesis. Most of these notions can be found in [104]. Let $\mathbb{F}_q$ be the finite field with q elements, where q is a prime power, and $\mathbb{F}_q[z]$ the polynomial ring over $\mathbb{F}_q$ in the indeterminate z.

It is well-known that $\mathbb{F}_q[z]$ is a Principal Ideal Domain (PID). Modules over a PID are **free**, namely, they admit a basis. Moreover, two different bases have the same number of elements, called the **rank** of the module. With this premise, we can introduce convolutional codes from a mathematical point of view.

6.1 Definition of Convolutional Codes via Generator and Parity-check Matrices

Definition 6.1 (Convolutional codes)**.** Let $k \leq n$ be two positive integers. An $(n,k)_q$ **convolutional code** $\mathcal{C}$ is a $\mathbb{F}_q[z]$-submodule of $\mathbb{F}_q[z]^n$ of rank k. A $k \times n$ matrix $G(z)$ with entries in $\mathbb{F}_q[z]$ whose rows constitute a basis of $\mathcal{C}$ is called a **generator matrix** for $\mathcal{C}$.

Notice that, given an $(n,k)_q$ convolutional code $\mathcal{C}$, the generator matrix is not unique. Indeed, define a matrix $U(z) \in \mathbb{F}_q[z]^{k\times k}$ with coefficients in $\mathbb{F}_q[z]$ to be **unimodular** if it has a polynomial matrix inverse, i.e. if there exists another matrix $V(z) \in \mathbb{F}_q[z]^{k\times k}$ such that $U(z)V(z) = V(z)U(z) = \mathrm{Id}_k$. By multiplying a generator matrix $G \in \mathbb{F}_q[z]^{k\times n}$ of $\mathcal{C}$ with a unimodular matrix $U(x) \in \mathbb{F}_q[z]^{k\times k}$, we get another generator matrix for $\mathcal{C}$. By elementary arguments it also follows that $U(z)$ is unimodular if and only if its determinant is a nonzero element of $\mathbb{F}_q$.

Definition 6.2. Two generator matrices $G(z)$, $\bar{G}(z)$ of an $(n,k)_q$ convolutional code $\mathcal{C}$ are called **equivalent** if there exists a unimodular matrix $U(z) \in \mathbb{F}_q[z]^{k\times k}$, such that $\bar{G}(z) = U(z)G(z)$.

There exist canonical forms of such equivalence relations, as explained in the following result.

Theorem 6.3. [76, 104] Let $G(z) \in \mathbb{F}_q[z]^{k\times n}$, with $k \leq n$. Then, there exists a unimodular matrix $U(z) \in \mathbb{F}_q[z]^{k\times k}$ such that

$$U(z)G(z) = \begin{bmatrix} h_{1,1}(z) & h_{1,2}(z) & \cdots & h_{1,k}(z) & h_{1,k+1}(z) & \cdots & h_{1,n}(z) \\ & h_{2,2}(z) & \cdots & h_{2,k}(z) & h_{2,k+1}(z) & \cdots & h_{2,n}(z) \\ & & \ddots & \vdots & \vdots & & \vdots \\ & & & h_{k,k}(z) & h_{k,k+1}(z) & \cdots & h_{k,n}(z) \end{bmatrix},$$

where for every $1 \leq i \leq k$, $h_{i,i}(z)$ is a monic polynomial such that $\deg(h_{i,i}) > \deg h_{j,i}$, for all $j < i$. Such a matrix is called the **column Hermite form** of $G(z)$.

If the equivalence relation is induced by right multiplication with a unimodular matrix or by right and left multiplication with unimodular matrices, the canonical forms are called the row Hermite form and the Smith form, respectively. We recall them as well in the following result.

Theorem 6.4. [76, 104] Let $G(z) \in \mathbb{F}_q[z]^{k\times n}$, with $k \leq n$. Then, there exists a unimodular matrix $U(z) \in \mathbb{F}_q[z]^{k\times k}$ such that

$$G(z)U(z) = \begin{bmatrix} h_{1,1}(z) & & & & 0 & \dots & 0 \\ h_{2,1}(z) & h_{2,2}(z) & & & \vdots & & \vdots \\ \vdots & \vdots & \ddots & & \vdots & & \vdots \\ h_{k,1}(z) & h_{k,2}(z) & \cdots & h_{k,k}(z) & 0 & \cdots & 0 \end{bmatrix},$$

where for every $1 \leq i \leq k$, $h_{i,i}(z)$ is a monic polynomial such that $\deg(h_{i,i}) > \deg h_{j,i}$, for all $j < i$. Such a matrix is called the **row Hermite form** of $G(z)$.

Theorem 6.5. [76, 104] Let $G(z) \in \mathbb{F}_q[z]^{k\times n}$, with $k \leq n$. Then, there exists a unimodular matrix $U(z) \in \mathbb{F}_q[z]^{k\times k}$ and a unimodular matrix $V(z) \in \mathbb{F}_q[z]^{n\times n}$ such that

$$S(z) = U(z)G(z)V(z) = \begin{bmatrix} \gamma_1(z) & & & & 0 & \dots & 0 \\ & \gamma_2(z) & & & \vdots & & \vdots \\ & & \ddots & & \vdots & & \vdots \\ & & & \gamma_k(z) & 0 & \dots & 0 \end{bmatrix},$$

where for every $1 \leq i \leq k$, γ_i is a monic polynomial with the property that γ_{i+1} divides γ_i. These polynomials are uniquely determined by $G(z)$ and are called **invariant polynomials** of $G(z)$. $S(z)$ is the **Smith form** of $G(z)$.

Remark 6.6. Two equivalent generator matrices have equal $k\times k$ minors, up to multiplication by a constant, because they differ by multiplication by a unimodular matrix. Hence, the following definition is justified.

Definition 6.7. Let $\mathcal{C}$ be an $(n,k)_q$ convolutional code. The maximal degree of the $k \times k$ minors of one (and hence all) generator matrix of $\mathcal{C}$ is called the **degree** of $\mathcal{C}$ and it is usually denoted by δ. Often, in the literature it is used the notation $(n,k,\delta)_q$ to denote a convolutional code of rank k and degree δ in $\mathbb{F}_q[z]^n$.

For every $1 \leq i \leq k$, the largest degree of any entry in the i-th row of a generator matrix $G(z)$ of $\mathcal{C}$ is called the **i-th row degree** ν_i.

For a generator matrix $G(z)$ of an $(n,k,\delta)_q$ convolutional code, with row degrees $\nu_1,\dots,\nu_k$, we have that $\delta \leq \nu_1 + \dots + \nu_k$. If there is equality, then $G(z)$ is said to be **row-reduced** and $G(z)$ is said a **minimal generator matrix** for $\mathcal{C}$.

Another important property of polynomial matrices is the so-called left primeness, which will be deeply discussed in Chapter 7 in relation to convolutional codes.

Definition 6.8. A polynomial matrix $G(z) \in \mathbb{F}_q[z]^{k\times n}$, with $k \leq n$ is **left prime** or **basic** if in all factorizations $G(z) = U(z)\bar{G}(z)$, with $U(z) \in \mathbb{F}_q[z]^{k\times k}$ and $\bar{G}(z) \in \mathbb{F}_q[z]^{k\times n}$, the left factor $U(z)$ is unimodular.

In the literature, there are many different equivalent notions of a left prime matrix, which we summarize in the following result and will be useful for the next chapter.

Theorem 6.9. [92] Let $G(z) \in \mathbb{F}_q[z]^{k\times n}$, with $k \leq n$. The following are equivalent:

1. $G(z)$ is left prime;
2. The Smith form of $G(z)$ is $[I_k\ 0]$;
3. The row Hermite form of $G(z)$ is $[I_k\ 0]$;
4. $G(z)$ admits a right $n \times k$ polynomial inverse;
5. $G(z)$ can be completed to a unimodular matrix, i.e., there exists $L(z) \in \mathbb{F}_q[z]^{(n-k)\times n}$ such that
$$\begin{bmatrix} G(z) \\ L(z) \end{bmatrix}$$
is unimodular.
6. The ideal generated by all the k-th order minors of $G(z)$ is $\mathbb{F}_q[z]$.
7. For all $u(z) \in \mathbb{F}_q(z)^k$, $u(z)G(z) \in \mathbb{F}_q[z]^n$ implies that $u(z) \in \mathbb{F}_q[z]^k$.
8. $\mathrm{rk}(G(\lambda)) = k$ for all $\lambda \in \overline{\mathbb{F}}_q$, where $\overline{\mathbb{F}}_q$ denotes the algebraic closure of $\mathbb{F}_q$.

Since generator matrices of a convolutional code $\mathcal{C}$ differ by left multiplication with a unimodular matrix, it follows that if a convolutional code admits a left prime generator matrix then all its generator matrices are also left prime.

If $G(z)$ is a left prime generator matrix of an $(n,k)_q$ convolutional code $\mathcal{C}$, then we say that $\mathcal{C}$ is **noncatastrophic**.

Let $\mathcal{C}$ be a noncatastrophic $(n,k,\delta)_q$ convolutional code and $G(z) \in \mathbb{F}_q[z]^{k\times n}$ be a generator matrix of $\mathcal{C}$. Then there exists a matrix $H(z) \in \mathbb{F}_q[z]^{(n-k)\times n}$, such that

$$c(z) \in \mathcal{C} \text{ if and only if } H(z)c(z)^\top = 0. \tag{6.1}$$

Such a matrix $H(z)$ is called a **parity-check matrix** of $\mathcal{C}$ as for the classical block codes. In [161], it has been shown that a convolutional code $\mathcal{C}$ is noncatastrophic if and only if it admits a parity-check matrix.

If $\mathcal{C}$ is a noncatastrophic convolutional code, then it also admits several parity-check matrices. In fact, every convolutional code has several left prime parity-check matrices and several parity-check matrices that are not left prime (in contrast to generator matrices where left primeness is a property of a noncatastrophic code). Indeed, if we consider a left prime parity-check matrix for a convolutional code and multiply it from the left with any polynomial matrix, we obtain another parity-check matrix for the same code. Moreover, in [133], it is shown that if $H(z) \in \mathbb{F}_q[z]^{(n-k)\times n}$ is a left prime and row-reduced parity-check matrix of an $(n,k,\delta)_q$ convolutional code $\mathcal{C}$, then the sum of the row degrees of $H(z)$ is equal to δ. This is not true in general. Indeed, the following example shows that if a not left prime parity-check matrix $H(z)$ of a convolutional code $\mathcal{C}$ is given, one can not obtain the degree of $\mathcal{C}$ as sum of the row degrees of $H(z)$.

Example 6.10. Let $\mathcal{C}$ be a $(3,1)$ convolutional code with with parity-check matrix

$$H(z) = \begin{bmatrix} z(1+z) & 0 & 1+z \\ 0 & 1+z & 1+z \end{bmatrix}.$$

Observe that $\mathcal{C}$ has degree 1 since the matrix

$$\tilde{H}(z) = \begin{bmatrix} z & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

is a left prime and row-reduced parity-check matrix of the same convolutional code, but the sum of the row degrees of $H(z)$ is 3. Moreover, the maximal degree of the full-size minors of $H(z)$ is also 3. This shows that the only way to obtain the degree of the code is by computing an equivalent left prime parity-check matrix. Note also that it does not help that $H(z)$ is row-reduced and that $H(0)$ has full rank.

Let $\mathcal{C} \subseteq \mathbb{F}_q[z]^n$ be an $(n,k,\delta)_q$ convolutional code. Thanks to the canonical isomorphism between $\mathbb{F}_q[z]^n$ and $\mathbb{F}_q^n[z]$, we can define a weight function on $\mathcal{C}$ as follows. Given a codeword

$v(z) = \sum_{i=0}^{r} v_i z^i \in \mathcal{C}$, we define the **weight** of $v(z)$ as

$$\mathrm{wt}(v(z)) := \sum_{i=0}^{r} \mathrm{wt}(v_i) \in \mathbb{N}_0,$$

where $\mathrm{wt}(v_i)$ denotes the Hamming weight of $v_i \in \mathbb{F}_q^n$, i.e. the number of its nonzero components. Finally, the **free distance** of a convolutional code $\mathcal{C}$ is defined as

$$\mathrm{d}_{\mathrm{free}}(\mathcal{C}) := \min\{\mathrm{wt}(v(z)) \mid v(z) \in \mathcal{C},\ v(z) \neq 0\}.$$

The generalized Singleton bound for an $(n, k, \delta)_q$ convolutional code $\mathcal{C}$, derived by Rosenthal and Smarandache in [135], relates the parameters of a convolutional code via the following inequality:

$$\mathrm{d}_{\mathrm{free}}(\mathcal{C}) \leq (n-k)\left(\left\lfloor \frac{\delta}{k} \right\rfloor + 1\right) + \delta + 1. \tag{6.2}$$

A convolutional code whose free distance reaches the bound (6.2) with equality is called **maximum distance separable (MDS) convolutional code**.

Lemma 6.11. [70, 92] Let $H(z) = [h_{i,j}(z)] \in \mathbb{F}_q[z]^{(n-k)\times n}$ with row degrees $\nu_1, \nu_2, \ldots, \nu_{n-k}$ and $[H]_{hr}$ be the highest row degree coefficient matrix defined as the matrix with the i-th row consisting of the coefficients of z^{ν_i} in the i-th row of $H(z)$. Then $H(z)$ is reduced if and only if $[H]_{hr}$ is full row-rank.

The a natural isomorphism between $\mathbb{F}_q[z]^n$ and $\mathbb{F}_q^n[z]$ extends to the space of matrices and allows to consider a generator and a parity-check matrix of a convolutional code as polynomials whose coefficients are matrices. In particular, we will consider $H(z) \in \mathbb{F}_q^{(n-k)\times n}[z]$, such that $H(z) = H_0 + H_1 z + \ldots H_\nu z^\nu$, with $\nu > 0$. With this notation, we can expand the kernel representation $H(z)v(z)^\top$ in the following way:

$$Hv^\top = \begin{bmatrix} H_0 & & & & \\ \vdots & \ddots & & & \\ H_\nu & \cdots & H_0 & & \\ & \ddots & & \ddots & \\ & & H_\nu & \cdots & H_0 \\ & & & \ddots & \vdots \\ & & & & H_\nu \end{bmatrix} \begin{bmatrix} v_0 \\ v_1 \\ \vdots \\ v_r \end{bmatrix} = 0, \tag{6.3}$$

where $r = \deg(v)$. We will refer to the representation of the parity-check matrix of $\mathcal{C}$ in equation (6.3) as **sliding parity-check matrix**.

6.2 MDP Convolutional Codes

MDP Convolutional codes are the central object of Chapters 7 and 8. In this section we briefly define what they are and why their study is important.

In the context of convolutional codes, one aims to build codes which can correct as many errors as possible within windows of different sizes. This property is described by the notion of column distances. More formally, we introduce the following notation. Let $v(z) = \sum_{i=0}^{r} v_i z^i \in \mathbb{F}_q^n[z]$. For any positive integer $j \leq r$, let $v_{[0,j]}(z) := \sum_{i=0}^{j} v_i z^i$.

Definition 6.12. The ***j*-th column distance** d_j^c of an $(n,k,\delta)_q$ convolutional code $\mathcal{C}$ is defined as

$$d_j^c := \min\{\mathrm{wt}(v_{[0,j]}(z)) \mid v(z) \in \mathcal{C}, \quad v_0 \neq 0\}.$$

Moreover, the column distances of $\mathcal{C}$ satisfy the following set of bounds.

Theorem 6.13. [79, Proposition 2.2] For every integer $j \in \mathbb{N}_0$,

$$d_j^c \leq (n-k)(j+1)+1. \tag{6.4}$$

Corollary 6.14. [79, Corollary 2.3] If $d_j^c \leq (n-k)(j+1)+1$ for some $j \in \mathbb{N}_0$, then $d_i^c \leq (n-k)(i+1)+1$ for every $i < j$.

Obviously, $d_j^c \leq \mathrm{d}_{\mathrm{free}}(\mathcal{C})$ for every j. It is easy to see that the maximum index for which the bound (6.4) is achievable is for $j = L$, where

$$L := \left\lfloor \frac{\delta}{k} \right\rfloor + \left\lfloor \frac{\delta}{n-k} \right\rfloor.$$

The $(L+1)$-tuple of numbers $(d_0^c, \dots, d_L^c)$ is called the **column distance profile** of the code $\mathcal{C}$.

Definition 6.15. An $(n,k,\delta)_q$ convolutional code $\mathcal{C}$ whose column distances d_j^c meet the bound of Theorem 6.13 with equality, for all $j = 0, \dots, L$, is called **maximum distance profile** (MDP).

Recall that the encoding map of an $(n,k,\delta)_q$ convolutional code $\mathcal{C}$ is given by the action of a polynomial matrix $G(z)$ and it can be expressed via the multiplication by the following polynomial:

$$G(z) := G_0 + G_1 z + \cdots + G_m z^m,$$

where $G_i \in \mathbb{F}_q^{k \times n}$ and $G_m \neq 0$. In the same way, the parity-check matrix is given by

$$H(z) := H_0 + H_1 z + \cdots + H_\nu z^\nu,$$

with $H_i \in \mathbb{F}_q^{(n-k)\times n}$ and $H_\nu \neq 0$.

Let $\mathcal{C}$ be an $(n,k,\delta)_q$ convolutional code, $G(z)$ be a generator matrix of $\mathcal{C}$ and $H(z)$ be a parity-check matrix for $\mathcal{C}$. For any $j \in \mathbb{N}_0$, we define the **j-th truncated sliding generator matrix** and the **j-th truncated sliding parity-check matrix** as

$$G_j^c := \begin{pmatrix} G_0 & G_1 & \cdots & G_j \\ & G_0 & \cdots & G_{j-1} \\ & & \ddots & \vdots \\ & & & G_0 \end{pmatrix} \in \mathbb{F}_q^{(j+1)k\times(j+1)n},$$

$$H_j^c := \begin{pmatrix} H_0 & & & \\ H_1 & H_0 & & \\ \vdots & \vdots & \ddots & \\ H_j & H_{j-1} & \cdots & H_0 \end{pmatrix} \in \mathbb{F}_q^{(j+1)(n-k)\times(j+1)n},$$

where $G_j = 0$, whenever $j > m$ and $H_j = 0$ whenever $j > \nu$.

These sliding matrices are relevant for the following well-known characterization of MDP convolutional codes.

Theorem 6.16. [79, Corollary 2.3 and Theorem 2.4] Let $G(z) = \sum_{i=0}^m G_i z^i$ and $H(z) = \sum_{i=0}^{\nu} H_i z^i$ be a left prime generator matrix and a left prime parity-check matrix, respectively, of an (n,k,δ) convolutional code $\mathcal{C}$. The following statements are equivalent:

1. $d_j^c(\mathcal{C}) = (n-k)(j+1)+1$,
2. every $(j+1)k \times (j+1)k$ full-size minor of G_j^c formed by columns with indices $1 \le t_1 < \cdots < t_{(j+1)k}$, where $t_{sk+1} > sn$ for $s = 1,\dots,j$, is nonzero ,
3. every $(j+1)(n-k)\times(j+1)(n-k)$ full-size minor of H_j^c formed by columns with indices $1 \le t_1 < \cdots < t_{(j+1)(n-k)}$, where $t_{s(n-k)+1} \le sn$ for $s = 1,\dots,j$, is nonzero.

In particular, $\mathcal{C}$ is MDP if and only if one of the above equivalent conditions holds for $j = L$.

We also recall the following well-known result.

Theorem 6.17. [79, Proposition 2.1] Let $\mathcal{C} \subseteq \mathbb{F}_q[z]^n$ be an $(n,k)_q$ convolutional code. Let $d \in \mathbb{N}$. Then the following properties are equivalent.

1. $d_j^c = d$.
2. None of the first n columns of H_j^c is contained in the span of any other $d-2$ columns and one of the first n columns of H_j^c is in the span of some other $d-1$ columns of that matrix.

Observe that the minors considered in Theorem 6.16 are the only full-size minors of G_j^c and H_j^c that can possibly be non-zero. For this reason, we call these minors **non trivially zero**.

Chapter 7

On the Left Primeness of some Polynomial Matrices with Applications to Convolutional Codes

This short chapter contains the results published in [8] by Alfarano and Lieb. While writing this thesis, the author realized that in the published paper [8] there is a computation mistake. This chapter aims to correct the mentioned work.

The motivation behind this work relays on the fact that several papers that provide a (concrete) construction for MDP convolutional codes, for example [153], [12], [101], are based on the characterization for the parity-check matrix from Theorem 6.16. Unfortunately, in all of them there is no discussion on the left primeness of the constructed matrices. Indeed, in all the mentioned works, only the criterion on the minors of the sliding parity-check is shown to be satisfied.

We then first explain in a Remark 7.2 why the left primeness is not needed in order that this criterion is valid and thus, all of these constructions are correct. However, as we have shown in Chapter 6 in general it is not easy to compute the degree of a convolutional code from a parity-check matrix that is not left prime and hence, it is not a priori clear that the constructed codes have really the degree that is stated in these papers.

Definition 7.1. Let $G(z) = \sum_{i=0}^{m} G_i z^i \in \mathbb{F}_q^{k\times n}[z]$ be a polynomial with $G_m \neq 0$, and let δ be degree of the convolutional code generated by $G(z)$. We say that $G(z)$ has the **MDP property** if the L-th truncated sliding generator matrix G_L^{c} satisfies condition 2 in Theorem 6.16. Let $H(z) = \sum_{i=1}^{\nu} H_i z^i \in \mathbb{F}_q^{(n-k)\times n}$ be a polynomial with $H_\nu \neq 0$ and δ be degree of the convolutional code which has $H(z)$ as parity-check matrix. We say that $H(z)$ has the **MDP property** if the L-th truncated sliding parity-check matrix H_L^{c} satisfies condition 3 in Theorem 6.16.

Remark 7.2. In Theorem 6.16 we assume that $G(z)$ and $H(z)$ are left prime. We will explain the exact role of this property:

1. Considering the corresponding proof in [79], one observes that for the equivalence between conditions 1 and 2 it is in fact enough to assume that G_0 is full rank (which is a consequence of $G(z)$ being left prime). However, both 1 and 2 imply that G_0 is full rank. For 1 this is true, because for $j = 0$ this means that G_0 is the generator matrix of an MDS block code, i.e. in particular full rank. For 2 this follows immediately from the structure of G_j^c. Hence, it is possible to get rid of the assumption that $G(z)$ is left prime. However, note that if $G(z)$ is not left prime, the corresponding code is catastrophic.

2. Now we consider the equivalence between 1 and 3, which of course is only possible if the code has a parity-check matrix, i.e. is noncatastrophic. If $H(z)$ is not left prime, then there exists an equivalent row-reduced and left prime parity-check matrix for the code $\tilde{H}(z)$, such that $H(z) = U(z)\tilde{H}(z)$ with $U(z) \in \mathbb{F}_q[z]^{(n-k)\times(n-k)}$ and $\deg(\det U(z))) > 0$. Hence, with $U(z) = \sum_i U_i z^i \in \mathbb{F}_q^{(n-k)\times(n-k)}[z]$ and $U_i = 0$ for $i > \deg(U(z))$, one has

$$\begin{bmatrix} H_0 & & 0 \\ \vdots & \ddots & \\ H_j & \cdots & H_0 \end{bmatrix} = \begin{bmatrix} U_0 & & 0 \\ \vdots & \ddots & \\ U_j & \cdots & U_0 \end{bmatrix} \begin{bmatrix} \tilde{H}_0 & & 0 \\ \vdots & \ddots & \\ \tilde{H}_j & \cdots & \tilde{H}_0 \end{bmatrix},$$

   for all $j \in \mathbb{N}_0$. Since $\tilde{H}(z)$ is left prime, $\tilde{H}_0$ is full rank. If $H(z)$ fulfills 3, then all the full-size minors of H_0 are nonzero. Together with $H_0 = U_0\tilde{H}_0$, this implies that U_0 and $\begin{bmatrix} U_0 & & 0 \\ \vdots & \ddots & \\ U_j & \cdots & U_0 \end{bmatrix}$ are full rank. Consequently, H_j^c fulfills 3 if and only if $\tilde{H}_j^c$ fulfills 3, and since $\tilde{H}(z)$ is left prime, $H(z)$ and $\tilde{H}(z)$ are parity-check matrices of an MDP convolutional code whose degree δ is equal to the sum of the row degrees of $\tilde{H}(z)$. Hence, also for the implication from 3 to 1, it is not necessary that the parity-check matrix of the code is left prime.

   However, to construct an MDP convolutional code with a given δ it is necessary to construct it via a left prime parity-check matrix. Otherwise we do not know the degree of the constructed code since it is in general not an easy task to determine the degree of a convolutional code if we only know one of its parity-check matrices which is not left prime, as shown in Example 6.10. In addition, the implication from 1 to 3 is only true if we assume at least that H_0 has full rank (which is a consequence of $H(z)$ being left prime). To see this, consider a parity-check matrix that fulfills 3, i.e. is a parity-check matrix of an MDP convolutional code, and multiply it by $z\mathrm{I}_{n-k}$. The resulting matrix is still a parity-check of the same MDP convolutional code but it has $H_0 = 0$ and hence, can not fulfill 3.

7.1 Left Primeness of Parity-Check and Generator Matrices of MDP Convolutional Codes

In this section, we show for which parameters condition 3 of Theorem 6.16 applied on an $(n,k)_q$ convolutional code $\mathcal{C}$ for $j=L$ implies that the corresponding parity-check matrix of $\mathcal{C}$ is left prime and thus the degree of $\mathcal{C}$ is equal to the sum of the row degrees of this parity-check matrix. Moreover, we show for which parameters condition 2 of Theorem 6.16 for $j=L$ implies that the considered convolutional code is noncatastrophic, i.e. for these parameters every MDP convolutional code is noncatastrophic.

Theorem 7.3. Consider $H(z)\in\mathbb{F}_q[z]^{(n-k)\times n}$ with $\deg(H(z))=\nu$ and set $\delta=(n-k)\nu$ and $r=\left\lfloor\frac{\delta}{k}\right\rfloor$. If the matrix

$$\bar{H} := \begin{bmatrix} H_0 & & \\ \vdots & \ddots & \\ H_\nu & & H_0 \\ & \ddots & \vdots \\ & & H_\nu \end{bmatrix} \in \mathbb{F}_q^{(n-k)(r+\nu+1)\times n(r+1)}$$

has full (row) rank, then $H(z)$ is left prime.

Proof. First note that since $(n-k)(r+\nu+1)=n(r+1)+\delta-k(r+1)<n(r+1)$, $\bar{H}$ has more columns than rows. As $\bar{H}$ has full row rank, the map $\mathbb{F}_q^{n(r+1)}\to\mathbb{F}_q^{(n-k)(r+\nu+1)}$, $v\mapsto\bar{H}v$ is surjective and there exists $\bar{X}=\begin{pmatrix} X_0\\ \vdots\\ X_r\end{pmatrix}\in\mathbb{F}_q^{(r+1)n\times(n-k)}$ with $X_i\in\mathbb{F}_q^{n\times(n-k)}$ for $i=1,\dots,r$ such that $\bar{H}\bar{X}=\begin{pmatrix} I_{n-k}\\ 0_{n-k}\\ \vdots\\ 0_{n-k}\end{pmatrix}$. Defining $X(z)=\sum_{i=0}^{r}X_iz^i$, one gets $H(z)X(z)=\mathrm{I}_{n-k}$ and hence $H(z)$ is left prime.

□

Corollary 7.4. Let $n,k,\delta\in\mathbb{N}$ with $k<n$ and $(n-k)\mid\delta$ and set $\nu=\frac{\delta}{n-k}$. If $H(z)=\sum_{i=0}^{\nu}H_iz^i\in\mathbb{F}_q^{(n-k)\times n}[z]$ has the property that all full-size minors of H_L^c with $L=\left\lfloor\frac{\delta}{k}\right\rfloor+\frac{\delta}{n-k}$ that are not trivially zero are nonzero, then $H(z)$ is a left prime parity-check matrix of an (n,k,δ) MDP convolutional code.

Proof. With the notation of the preceding theorem, one gets $L=r+\nu$ and $\bar{H}$ is a submatrix of H_L^c with the same number of rows. Hence, there is a full-size minor of $\bar{H}$ that is nonzero and

$\bar{H}$ has full (row) rank, which additional implies that H_ν is full rank. Consequently, $H(z)$ is left prime and thus, it is the parity-check matrix of an (n, k, δ) convolutional code, where δ is equal to the sum of the row degrees of $H(z)$, i.e. $\delta = (n-k)\nu$ as H_ν is full rank. Then, Theorem 6.16 implies that this code is MDP. □

Remark 7.5. With the same reasoning one can show that for $n, k, \delta \in \mathbb{N}$ with $k < n$ and $k \mid \delta$ and $m = \frac{\delta}{k}$, if $G(z) = \sum_{i=0}^{m} G_i z^i \in \mathbb{F}_q^{k\times n}[z]$ has the property that all full-size minors of G_L^c with $L = \frac{\delta}{k} + \left\lfloor \frac{\delta}{n-k} \right\rfloor$ that are not trivially zero are nonzero, then $G(z)$ is the generator matrix of a noncatastrophic (n, k, δ) MDP convolutional code.

Remark 7.6. The conditions of the preceding theorem, corollary and remark are not necessary (only sufficient) to ensure that the corresponding polynomial matrix is left prime. As mentioned before, a polynomial matrix is left prime if and only if it has a polynomial right inverse and we provided sufficient conditions in order that this is true.

The following example shows that if $(n-k) \nmid \delta$ (resp. $k \nmid \delta$), then the MDP property on the minors of the sliding parity-check (resp. generator) matrix does in general not imply that the parity-check (resp. generator) matrix of a convolutional code is left prime.

Example 7.7. Let $1 \leq \delta < k$ and $\delta < n-k$, i.e. $L = 0$. We get that $\deg(H(z)) = \lfloor \frac{\delta}{n-k} \rfloor + 1 = 1$, so $H(z) = H_0 + H_1 z$ and $H_L^c = H_0$. If we choose H_0 such that all full-size minors are nonzero and $H_1 = -H_0$, then H_L^c fulfills the MDP property 3 but $H(z) = (z-1)\mathrm{I}_{n-k} H_1$, i.e. $H(z)$ is not left prime and the degree of the code with this parity-check matrix is zero. Hence, this can not be an (n, k, δ) MDP convolutional code. Equivalently, we can show that for such code parameters a generator matrix $G(z) = G_0 + G_1 z$ with $G_0 = -G_1$ having all full-size minors nonzero is not left prime but G_L^c fulfills the MDP criterion 2, i.e. $G(z)$ is the generator matrix of a catastrophic (n, k, δ) MDP convolutional code.

Remark 7.8. In the published paper [8], we also modify Theorem 7.3 by imposing stronger assumptions to get similar results for the case $(n-k) \nmid \delta$. However, this is not correct and so we omit that part in this thesis.

Chapter 8

Weighted Reed-Solomon Convolutional Codes

The results in this chapter can be found in the preprint [11], by Alfarano, Napp, Neri, Requena.

For this chapter we will only consider generator matrices for convolutional codes. We start with a preliminary remark regarding Theorem 6.16; see also Remark 7.2.

Remark 8.1. The original assumption in Theorem 6.16 is that $G(z)$ is left prime. However, as it emerges from the original proof, it is not necessary to show that the convolutional code generated by $G(z)$ is MDP. Having $G(z)$ left prime, indeed, ensure that the code is noncatastrophi. In this way, a similar characterization could be derive from the parity-check matrix. Moreover, notice that in [8], it is shown that in the case $m = \frac{\delta}{k}$, if $G(z) = \sum_{i=1}^{m} G_i z^i$ has the MDP property, then the convolutional code generated by $G(z)$ is noncatastrophic and one can get rid of the assumption of $G(z)$ being left prime.

Remark 8.2. Here we rephrase the MDP property as follows. Each minor of G_L^{c} obtained by selecting the columns with indices as described in Theorem 6.16 is a minor obtained by selecting ℓ_i columns from the i-th columns block, for every $i = 0, \dots, L$, such that

$$\sum_{i=0}^{s} \ell_i \leq (s+1)k \quad \text{for } s = 0, \dots, L-1, \tag{8.1}$$

$$\sum_{i=0}^{L} \ell_i = (L+1)k. \tag{8.2}$$

Note that the remaining full size minors of G_L^{c} not satisfying (8.1) are trivially zero, i.e., are zero independently of the choice of the nonzero entries of G_L^{c}, see [13, 79] for a formal definition.

We conclude this section by recalling the definition of generalized Reed-Solomon codes, which are one of the most studied family of codes in algebraic coding theory, due to their very rich algebraic structure and their suitability for digital implementation in practical applications: they possess optimal distance and admit efficient algebraic decoding algorithms, *e.g.*, Berlekamp-Massey, see [109, Chapters 10 and 11]. We will use their generator matrices in order to construct MDP convolutional codes.

Let $0 < k \leq n$ be two positive integers and consider the set of polynomials with coefficients in $\mathbb{F}_q$ and degree strictly less than k, namely

$$\mathbb{F}_q[x]_{<k} := \{f(x) \in \mathbb{F}_q[x] \mid \deg f < k\}.$$

Definition 8.3. Suppose that $n \leq q$, and consider $\alpha_1, \dots, \alpha_n \in \mathbb{F}_q$ pairwise distinct elements, and $b_1, \dots, b_n \in \mathbb{F}_q^*$. The block code

$$C := \{(b_1 f(\alpha_1), \dots, b_n f(\alpha_n)) \mid f \in \mathbb{F}_q[x]_{<k}\}$$

is called **generalized Reed-Solomon (GRS) code** and it is denoted by $\mathrm{GRS}_k(\alpha, b)$, where $\alpha := (\alpha_1, \dots, \alpha_n)$ and $b = (b_1, \dots, b_n)$.

The canonical generator matrix for a code $C = \mathrm{GRS}_k(\alpha, b)$ has the following form:

$$G := \begin{pmatrix} b_1 & b_2 & \cdots & b_n \\ b_1\alpha_1 & b_2\alpha_2 & \cdots & b_n\alpha_n \\ b_1\alpha_1^2 & b_2\alpha_2^2 & \cdots & b_n\alpha_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ b_1\alpha_1^{k-1} & b_2\alpha_2^{k-1} & \cdots & b_n\alpha_n^{k-1} \end{pmatrix} = V_k(\alpha)\,\mathrm{diag}(b),$$

where $V_k(\alpha)$ is a classical Vandermonde matrix of size $k \times n$ of the form

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ \alpha_1 & \alpha_2 & \cdots & \alpha_n \\ \alpha_1^2 & \alpha_2^2 & \cdots & \alpha_n^2 \\ \vdots & \vdots & \vdots & \vdots \\ \alpha_1^{k-1} & \alpha_2^{k-1} & \cdots & \alpha_n^{k-1} \end{pmatrix}$$

and $\mathrm{diag}(b)$ denotes the diagonal matrix whose diagonal entries are given by $b_1, \dots, b_n$. We call this generator matrix "canonical" since it is obtained by evaluating each monomial of the standard $\mathbb{F}_q$-basis of $\mathbb{F}_q[x]_{<k}$, that is $\{1, x, x^2, \dots, x^{k-1}\}$, in the points $\alpha_1, \dots, \alpha_n$.

8.1 New Construction

In this section we present a new algebraic construction of $(n, k, \delta)_q$ MDP convolutional codes with memory $m = \left\lceil \frac{\delta}{k} \right\rceil$. To this end, we use some generalized Vandermonde matrices as the coefficients of the polynomial matrix $G(z)$ describing the code.

Let k, n be positive integers and let q be a prime power, with $k < n < q$. Let $\alpha := (\alpha_1, \dots, \alpha_n) \in (\mathbb{F}_q^*)^n$, with the α_i's pairwise distinct, and fix γ to be a root of an irreducible polynomial in $\mathbb{F}_q[z]$ of degree s, for some suitable integer s. Clearly, $\mathbb{F}_q(\gamma) \cong \mathbb{F}_{q^s}$.

For any $i \geq 0$, set

$$M_i := \begin{pmatrix} \gamma^{\binom{i+1}{2}k-i}\alpha_1^{(i+1)k-1} & \gamma^{\binom{i+1}{2}k-i}\alpha_2^{(i+1)k-1} & \cdots & \gamma^{\binom{i+1}{2}k-i}\alpha_n^{(i+1)k-1} \\ \vdots & \vdots & & \vdots \\ \gamma^{\binom{i}{2}k+i}\alpha_1^{ik+1} & \gamma^{\binom{i}{2}k+i}\alpha_2^{ik+1} & \cdots & \gamma^{\binom{i}{2}k+i}\alpha_n^{ik+1} \\ \gamma^{\binom{i}{2}k}\alpha_1^{ik} & \gamma^{\binom{i}{2}k}\alpha_2^{ik} & \cdots & \gamma^{\binom{i}{2}k}\alpha_n^{ik} \end{pmatrix}, \tag{8.3}$$

and, for every $i \geq 0$ and $1 \leq j \leq k$,

$$N_{i,j} := \begin{pmatrix} 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \\ \gamma^{\binom{i}{2}k+(j-1)i}\alpha_1^{ik+j-1} & \gamma^{\binom{i}{2}k+(j-1)i}\alpha_2^{ik+j-1} & \cdots & \gamma^{\binom{i}{2}k+(j-1)i}\alpha_n^{ik+j-1} \\ \vdots & \vdots & & \vdots \\ \gamma^{\binom{i}{2}k+i}\alpha_1^{ik+1} & \gamma^{\binom{i}{2}k+i}\alpha_2^{ik+1} & \cdots & \gamma^{\binom{i}{2}k+i}\alpha_n^{ik+1} \\ \gamma^{\binom{i}{2}k}\alpha_1^{ik} & \gamma^{\binom{i}{2}k}\alpha_2^{ik} & \cdots & \gamma^{\binom{i}{2}k}\alpha_n^{ik} \end{pmatrix}. \tag{8.4}$$

For the binomial coefficients, we use the convention that $\binom{a}{b} = 0$ if $a < b$. Observe that for every $i \geq 0$, it holds $N_{i,k} = M_i$. Moreover, $M_i, N_{i,j} \in \mathbb{F}_{q^s}^{k \times n}$ for every $i \geq 1$, while $M_0, N_{0,j} \in \mathbb{F}_q^{k \times n}$. It is easy to see that the matrix M_i is the generator matrix of the $[n,k]_{q^s}$ block code $\mathrm{GRS}_k(\alpha, \alpha^{(ik)})$, where $\alpha^{(ik)} := (\alpha_1^{ik}, \dots, \alpha_n^{ik})$.

Definition 8.4. Let k, n, δ be positive integers with $0 < k \leq n$, $\alpha = (\alpha_1, \dots, \alpha_n) \in (\mathbb{F}_q^*)^n$ and $\gamma \in \mathbb{F}_{q^s}$ be as above. Let $m := \lceil \frac{\delta}{k} \rceil$ and $t := \delta - (m-1)k$, and define

$$G_i := \begin{cases} M_i & \text{if } 0 \leq i \leq m-1, \\ N_{m,t} & \text{if } i = m. \end{cases} \tag{8.5}$$

A convolutional code is called **weighted Reed-Solomon (WRS) convolutional code** if admits $G(z) = \sum_{i=0}^m G_i z^i$ as generator matrix. We will denote such a code by $\mathcal{C}_{k,n}^{\delta}(\gamma, \alpha)$.

We want to study now the codes $\mathcal{C}_{k,n}^{\delta}(\gamma, \alpha)$. We start by providing their parameters.

Proposition 8.5. The code $\mathcal{C}_{k,n}^{\delta}(\gamma, \alpha)$ is an $(n, k, \delta)_{q^s}$ convolutional code. In particular, the generator matrix $G(z)$ given in Definition 8.4 is reduced.

Proof. Clearly the code $\mathcal{C}_{k,n}^{\delta}(\gamma, \alpha)$ is defined over $\mathbb{F}_{q^s}$ and has length n. Moreover, the matrix $G_0 = M_0$ is a Vandermonde matrix and hence it is full rank. This implies that the dimension of the code is k. Let $\tilde{\delta}$ be the degree of $\mathcal{C}^{\delta}_{k,n}$. By definition, the sum of the row degrees of $G(z)$ is δ, and hence $\tilde{\delta} \leq \delta$. In order to show that $\delta = \tilde{\delta}$, it is enough to show that $G(z)$ is reduced, i.e., the leading row coefficient matrix of $G(z)$, denoted by G_∞, is full row rank; see [72] or [92, Theorem 6.3–13]. Recall that $m = \lceil \frac{\delta}{k} \rceil$ and $t = \delta - (m-1)k$. It is easy to see that the matrix G_∞ has the first $k - t$ rows equal to the ones of $G_{m-1} = M_{m-1}$ and the last t rows equal to the ones of $G_m = N_{m,t}$. Such a matrix is a row permutation of

$$\begin{pmatrix}
\gamma^{\binom{m}{2}k+(t-1)m}\alpha_1^{mk+t-1} & \gamma^{\binom{m}{2}k+(t-1)m}\alpha_2^{mk+t-1} & \cdots & \gamma^{\binom{m}{2}k+(t-1)m}\alpha_n^{mk+t-1} \\
\vdots & \vdots & & \vdots \\
\gamma^{\binom{m}{2}k}\alpha_1^{mk} & \gamma^{\binom{m}{2}k}\alpha_2^{mk} & \cdots & \gamma^{\binom{m}{2}k}\alpha_n^{mk} \\
\gamma^{\binom{m}{2}k-m}\alpha_1^{mk-1} & \gamma^{\binom{m}{2}k-m}\alpha_2^{mk-1} & \cdots & \gamma^{\binom{m}{2}k-m}\alpha_n^{mk-1} \\
\vdots & \vdots & & \vdots \\
\gamma^{\binom{m-1}{2}k+t(m-1)}\alpha_1^{(m-1)k+t} & \gamma^{\binom{m-1}{2}k+t(m-1)}\alpha_2^{(m-1)k+t} & \cdots & \gamma^{\binom{m-1}{2}k+t(m-1)}\alpha_n^{(m-1)k+t}
\end{pmatrix},$$

which is full rank, since it is a Vandermonde matrix whose rows are multiplied by powers of γ and whose columns are multiplied by $\alpha_i^{(m-1)k+t}$. □

Definition 8.6. Let k, n and δ be fixed. Consider the matrices $G_i(x)$ as the matrices G_i defined in (8.5) where we have replaced γ by an indeterminate x, and let $G^{\mathrm{c}}_L(x)$ be the corresponding L-th truncated sliding generator matrix. We define the set

$$\begin{aligned}\mathcal{P}(k,n,\delta,\alpha) := \{p(x) \in \mathbb{F}_q[x] \mid\ & p(x) \text{ is a full size minor of } G^{\mathrm{c}}_L(x) \text{ obtained selecting the columns} \\ & \text{with indices } \ 1 \leq j_1 < \cdots < j_{(L+1)k}, \text{ where } j_{rk+1} > rn \text{ for } r = 1, \dots, L\}\,.\end{aligned}$$

Note that the set $\mathcal{P}(k,n,\delta,\alpha)$ represents the full size minors in $G^{\mathrm{c}}_L(x)$ formed as stated in Theorem 6.16.

Theorem 8.7. Let $\mathcal{C}^{\delta}_{k,n}(\gamma,\alpha)$ be the $(n,k,\delta)_{q^s}$ WRS convolutional code with generator matrix $G(z) = \sum_{i=0}^{m} G_i z^i \in \mathbb{F}_{q^s}^{k\times n}[z]$, where the G_i's are defined by (8.5). If $p(\gamma) \neq 0$ for every $p(x) \in \mathcal{P}(k,n,\delta,\alpha)$, then $\mathcal{C}^{\delta}_{k,n}(\gamma,\alpha)$ is an MDP convolutional code.

Proof. As $p(\gamma) \neq 0$ for every $p(x) \in \mathcal{P}(k,n,\delta,\alpha)$, then the condition in Theorem 6.16 is satisfied for $j = L$ and $G(z)$ has the MDP property, i.e., $\mathrm{d}^{\mathrm{c}}_L(\mathcal{C}^{\delta}_{k,n}(\gamma,\alpha)) = (n-k)(L+1)+1$. It follows from [79, Corollary 2.3] that $\mathrm{d}^{\mathrm{c}}_j(\mathcal{C}^{\delta}_{k,n}(\gamma,\alpha)) = (n-k)(j+1)+1$ for $j = 0, 1, \dots, L$ and therefore, by definition, $\mathcal{C}^{\delta}_{k,n}(\gamma,\alpha)$ is an (n,k,δ) MDP convolutional code. □

For a given nonzero polynomial $p(x) \in \mathbb{F}_q[x]$, we denote by $\deg p(x)$ the degree of $p(x)$, and by $\nu(p(x))$ the maximum integer ℓ such that x^ℓ divides $p(x)$. Then we define the integer

$$D(k,n,\delta,\alpha) := \max\{\deg p(x) - \nu(p(x)) \mid 0 \neq p(x) \in \mathcal{P}(k,n,\delta,\alpha)\}.$$

The next result is the main theorem of this section. However, its proof requires several technical lemmas and it can be found in Section 8.2.

Theorem 8.8. Let γ be a root of an irreducible polynomial in $\mathbb{F}_q[z]$ of degree s and let $\mathcal{C}^{\delta}_{k,n}(\gamma,\alpha)$ be the $(n,k,\delta)_{q^s}$ WRS convolutional code whose generator matrix is $G(z) = \sum_{i=0}^{m} G_i z^i \in \mathbb{F}_{q^s}^{k\times n}[z]$, and the G_i's are defined by (8.5). If $s > D(k,n,\delta,\alpha)$, then $\mathcal{C}^{\delta}_{k,n}(\gamma,\alpha)$ is an MDP convolutional code in $\mathbb{F}_{q^s}[z]^n$.

We conclude this section by illustrating with a concrete example how to construct a WRS convolutional code that is also MDP, using the previous theorem.

Example 8.9. We fix the parameters $k = 3$, $n = 5$ and $\delta = 5$. Therefore, we have $m = 2$ and $L = 3$. We then choose a prime power greater than n, that is $q = 7$ and a vector with pairwise distinct nonzero entries $\alpha = (\alpha_1,\alpha_2,\alpha_3,\alpha_4,\alpha_5) = (1,2,3,4,5) \in \mathbb{F}_7^5$. At this point we illustrate how to choose a suitable γ so that the resulting code $\mathcal{C}^5_{3,5}(\gamma,\alpha)$ is MDP. We consider the polynomial version of the 3-th truncated sliding generator matrix $G^{\mathrm{c}}_3(x)$, given by

$$G^{\mathrm{c}}_3(x) = \begin{pmatrix} G_0(x) & G_1(x) & G_2(x) & \\ & G_0(x) & G_1(x) & G_2(x) \\ & & G_0(x) & G_1(x) \\ & & & G_0(x) \end{pmatrix} \in \mathbb{F}_7[x]^{12\times 20},$$

where

$$G_0(x) = \begin{pmatrix} 1 & 4 & 2 & 2 & 4 \\ 1 & 2 & 3 & 4 & 5 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \in \mathbb{F}_7^{3\times 5},$$

$$G_1(x) = \begin{pmatrix} x^2 & 4x^2 & 5x^2 & 2x^2 & 3x^2 \\ x & 2x & 4x & 4x & 2x \\ 1 & 1 & 6 & 1 & 6 \end{pmatrix} \in \mathbb{F}_7[x]^{3\times 5},$$

$$G_2(x) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ x^5 & 2x^5 & 3x^5 & 4x^5 & 5x^5 \\ x^3 & x^3 & x^3 & x^3 & x^3 \end{pmatrix} \in \mathbb{F}_7[x]^{3\times 5}.$$

We now compute the value $D(3,5,5,\alpha)$, which can be checked to be $D(3,5,5,\alpha) = 9$. There are many full size minors of $G^{\mathrm{c}}_3(x)$ from which we can obtain this value. For instance, if we select the columns with indices $\{1,6,7,8,11,12,13,16,17,18,19,20\}$ of $G^{\mathrm{c}}_3(x)$, we have that the full size minor is $p(x) = x^3(3x^9 + 2x^8 + 4x^7 + 5x^5 + x^4 + 4x^3 + x^2 + 4x + 4) \in \mathcal{P}(3,5,5,\alpha)$. Let now choose γ to be a root of an irreducible polynomial of degree $s = 10$ over $\mathbb{F}_7$. Thus, with this choice, the code $\mathcal{C}^5_{3,5}(\gamma,\alpha)$ is an MDP $(5,3,5)$ WRS convolutional code over the field $\mathbb{F}_{7^{10}}$. Its generator matrix is given by $G(z) := G_0(\gamma) + G_1(\gamma)z + G_2(\gamma)z^2$.

8.2 A Multivariate Polynomial Generalization of G^{c}_L

In Section 8.1 we introduced $G^{\mathrm{c}}_L(x)$ in Definition 8.6 as a polynomial generalization of the truncated sliding generator matrix G^{c}_L of WRS convolutional codes, by substituting γ with a variable x. In this section we further generalize its square submatrices by seeing the α_i's defining the generalized Vandermonde matrices as algebraically independent variables y_i's, yelding a multivariate polynomial representation of G^{c}_L. This generalization allows to give a proof of Theorem 8.8, to which this section is dedicated. For the convenience of the reader, in this short introduction we briefly present the idea of the proof.

We denote by B the collection of the involved powers of x, and by Λ the collection of the exponents of y_i's involved in the generalized Vandermonde matrices constituting the matrix $G^{\mathrm{c}}_L(x)$. In other words, B and Λ denote the exponents of the variables. In this way one obtains a polynomial generalization of the square submatrices of G^{c}_L, denoted by $G(x, Y, \mathrm{B}, \Lambda)$, and their minors become multivariate polynomials $p(x, Y)$, where Y denotes the vector formed by the variables y_i's.

Several technical lemmas lead to Theorem 8.19, where we describe the monomial of minimal degree of $p(x, Y)$ in the variable x, which is given by the product of determinants of some particular submatrices of $G(x, Y, \mathrm{B}, \Lambda)$.

By choosing some special values of B and Λ and specializing Y in a suitable vector A of elements in $\mathbb{F}_q$, we obtain that the resulting matrix yields a square submatrix of $G^{\mathrm{c}}_L(x)$ as in Definition 8.6. Moreover, we show that the monomial of minimal degree in $p(x, Y)$ is still nonzero when Y is specialized in A. In particular, the set $\mathcal{P}(k, n, \delta, \alpha)$ defined in Definition 8.6 will consists only of such polynomials $p(x, \mathrm{A})$, which are all nonzero. By carefully choosing the value γ, we then show that the resulting convolutional code $\mathcal{C}^{\delta}_{k,n}(\gamma, \alpha)$ is MDP, by means of Theorem 8.7. In Theorem 8.21 we give the equivalent version of Theorem 8.19 for the monomial of maximum degree in x of the same polynomial $p(x, Y)$. However, when considering $p(x, \mathrm{A})$, such monomial could vanish.

All the results mentioned above are needed to finally prove Theorem 8.8, which states that WRS convolutional codes are MDP.

We start by recalling the definition of generalized Vandermonde matrix. Then, we establish the notation for the remainder of the section.

Definition 8.10. Let $\mathbb{F}_q$ be the finite field with q elements, k, n be positive integers. Let $\lambda = (\lambda_1, \dots, \lambda_k) \in \mathbb{N}^k$ be a vector whose entries are pairwise distinct and $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{F}_q^n$. A $k \times n$ **generalized Vandermonde matrix** is a matrix of the form

$$V(\lambda, \alpha) = \begin{pmatrix} \alpha_1^{\lambda_1} & \alpha_2^{\lambda_1} & \cdots & \alpha_n^{\lambda_1} \\ \alpha_1^{\lambda_2} & \alpha_2^{\lambda_2} & \cdots & \alpha_n^{\lambda_2} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{\lambda_k} & \alpha_2^{\lambda_k} & \cdots & \alpha_n^{\lambda_k} \end{pmatrix} \in \mathbb{F}_q^{k \times n}.$$

The following definition introduces the polynomial matrix which is central to this section.

Definition 8.11. Let $e \in \mathbb{N}$ be a nonegative integer, $(\ell_0, \dots, \ell_e), (k_0, \dots, k_e) \in (\mathbb{N}_{>0})^{e+1}$, such that $\sum_{i=0}^{r} \ell_i \leq \sum_{i=0}^{r} k_i$ for any $r \in \{0, \dots, e-1\}$ and $\sum_{i=0}^{e} \ell_i = \sum_{i=0}^{e} k_i$. For any $j \in \{0, \dots, e\}$, let $y^{(j)} = \left(y_1^{(j)}, \dots, y_{\ell_j}^{(j)}\right)$ be a vector of variables and for any $0 \leq i \leq j \leq e$, let $\lambda^{(i,j)} = (\lambda_1^{(i,j)}, \dots, \lambda_{k_i}^{(i,j)}) \in \mathbb{N}^{k_i}$ be such that the following conditions hold:

(L1) $\lambda_{s-1}^{(i,j)} > \lambda_s^{(i,j)}$, for any $s \in \{2, \dots, k_i\}$.

(L2) $\lambda_1^{(i,j)} > \lambda_{k_{i+1}}^{(i+1,j)}$, for any $0 \leq i \leq j-1$, $1 \leq j \leq e$.

For any $0 \leq i \leq j \leq e$, let $\beta^{(i,j)} = (\beta_1^{(i,j)}, \dots, \beta_{k_i}^{(i,j)}) \in \mathbb{N}^{k_i}$, such that:

(b1) $\beta^{(i,i)} = 0$.

(b2) $\beta_{s-2}^{(i,j)} - \beta_{s-1}^{(i,j)} \geq \beta_{s-1}^{(i,j)} - \beta_s^{(i,j)}$, for any $s \in \{3, \dots, k_i\}$.

(b3) $\beta_{k_i-1}^{(i,j)} - \beta_{k_i}^{(i,j)} \geq \beta_{k_i}^{(i,j)} - \beta_1^{(i+1,j)} + 1 \geq \beta_1^{(i+1,j)} - \beta_2^{(i+1,j)} + 1$, for any $0 \leq i \leq j-1$ and $1 \leq j \leq e$.

(b4) $\left(\beta_{s-2}^{(i,j+1)} - \beta_{s-2}^{(i,j)}\right) - \left(\beta_{s-1}^{(i,j+1)} - \beta_{s-1}^{(i,j)}\right) \geq \left(\beta_{s-1}^{(i,j+1)} - \beta_{s-1}^{(i,j)}\right) - \left(\beta_s^{(i,j+1)} - \beta_s^{(i,j)}\right)$ for any $s \in \{3, \dots, k_i\}$, for any $i \leq j$ and $0 \leq j \leq e-1$.

(b5) $\left(\beta_{k_i}^{(i,j+1)} - \beta_{k_i}^{(i,j)}\right) - \beta_1^{(i+1,j+1)} \geq \beta_1^{(i+1,j+1)} - \beta_2^{(i+1,j+1)}$, for any $0 \leq i \leq j \leq e-1$.

We define

$$A_{i,j}^{\left(\beta^{(i,j)}, \lambda^{(i,j)}\right)} := \operatorname{diag}\left(x^{\beta^{(i,j)}}\right) V\left(\lambda^{(i,j)}, y^{(j)}\right) \in \mathbb{F}_q[x, y^{(j)}]^{k_i \times \ell_j}, \tag{8.6}$$

where

$$\operatorname{diag}\left(x^{\beta^{(i,j)}}\right) = \begin{pmatrix} x^{\beta_1^{(i,j)}} & 0 & \cdots & 0 \\ 0 & x^{\beta_2^{(i,j)}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x^{\beta_{k_i}^{(i,j)}} \end{pmatrix} \in \mathbb{F}[x]^{k_i \times k_i}.$$

To simplify the notation in (8.6), we only write $A_{i,j}$ and specify the vectors $,\lambda^{(i,j)}$, $\beta^{(i,j)}$ only when it is necessary. Let

$$Y := \left(y_1^{(0)}, \dots, y_{\ell_0}^{(0)}, y_1^{(1)}, \dots, y_{\ell_1}^{(1)}, \dots, y_1^{(e)}, \dots, y_{\ell_e}^{(e)}\right)$$

be the vector of all the variables and

$$\mathrm{B} := \left(\beta^{(0,0)}, \dots, \beta^{(0,e)}, \beta^{(1,1)}, \dots, \beta^{(1,e)}, \dots, \beta^{(e,e)}\right),$$

$$\Lambda := \left(\lambda^{(0,0)}, \dots, \lambda^{(0,e)}, \lambda^{(1,1)}, \dots, \lambda^{(1,e)}, \dots, \lambda^{(e,e)}\right)$$

be vectors of exponents. These three vectors uniquely determine the following matrix

$$G(x,Y,\mathrm{B},\Lambda) := \begin{pmatrix} A_{0,0} & A_{0,1} & \cdots & A_{0,e} \\ & A_{1,1} & \cdots & A_{1,e} \\ & & \ddots & \vdots \\ & & & A_{e,e} \end{pmatrix} \in \mathbb{F}_q[x,Y]^{(k_0+\cdots+k_e)\times(\ell_0+\cdots+\ell_e)}. \tag{8.7}$$

In the next example, we provide some tuples satisfying conditions 8.11–8.11 and 8.11–8.11, to get a more intuitive idea of their relations.

Example 8.12. Let $e = 2$ and $(k_0, k_1, k_2) = (4, 4, 4), (\ell_0, \ell_1, \ell_2) = (2, 4, 6) \in \mathbb{N}^3$. Let

$$\begin{array}{rclrclrcl} \lambda^{(0,0)} & = & (3,2,1,0), & \lambda^{(0,1)} & = & (7,6,5,4), & \lambda^{(0,2)} & = & (11,10,9,8) \\ & & & \lambda^{(1,1)} & = & (3,2,1,0), & \lambda^{(1,2)} & = & (7,6,5,4), \\ & & & & & & \lambda^{(2,2)} & = & (3,2,1,0). \end{array}$$

For convenience, we ordered these vectors in row/column blocks. Each row block corresponds to a fixed i, and each column block corresponds to a fixed j. Clearly, all the vectors defined above satisfy condition 8.11, *i.e.*are ordered in a decreasing order. Property 8.11 is referred to vectors in consecutive row blocks but same column block. It states that for any $i \leq j-1 \leq 1$, we require the first entry of $\lambda^{(i,j)}$ to be strictly greater than the last entry of $\lambda^{(i+1,j)}$. In this example, we only need to check the column block defined by $j = 1$ and immediately obtain that $\lambda_1^{(0,1)} > \lambda_4^{(1,1)}$.

Let

$$\begin{array}{rclrclrcl} \beta^{(0,0)} & = & (0,0,0,0), & \beta^{(0,1)} & = & (3,2,1,0), & \beta^{(0,2)} & = & (10,8,6,4), \\ & & & \beta^{(1,1)} & = & (0,0,0,0), & \beta^{(1,2)} & = & (3,2,1,0), \\ & & & & & & \beta^{(2,2)} & = & (0,0,0,0). \end{array}$$

Condition 8.11 is clearly satisfied. Condition 8.11 refers to each vector $\beta^{(i,j)}$. It states that the differences between consecutive entries are non increasing. For instance, consider the vector $\beta^{(0,2)}$. We have

$$\beta_1^{(0,2)} - \beta_2^{(0,2)} \geq \beta_2^{(0,2)} - \beta_3^{(0,2)} \geq \beta_3^{(0,2)} - \beta_4^{(0,2)}.$$

Condition 8.11 refers to two vectors of two different row blocks, but same column block, for instance, $\beta^{(0,2)}$ and $\beta^{(1,2)}$. In this example, this is the only possible pair on which this property can be verified. We have

$$\beta_3^{(0,2)} - \beta_4^{(0,2)} \geq \beta_4^{(0,2)} - \beta_1^{(1,2)} + 1 \geq \beta_1^{(1,2)} - \beta_2^{(1,2)} + 1$$

which is obviously satisfied since $6-4 \geq 4-3+1 \geq 3-2+1$. Condition 8.11 refers to two vectors of the same row block. Consider $\beta^{(0,0)}$ and $\beta^{(0,1)}$. We want that

$$(\beta_1^{(0,1)} - \beta_1^{(0,0)}) - (\beta_2^{(0,1)} - \beta_2^{(0,0)}) \geq (\beta_2^{(0,1)} - \beta_2^{(0,0)}) - (\beta_3^{(0,1)} - \beta_3^{(0,0)}).$$

Indeed, we have $(3-0)-(2-0) \geq (2-0)-(1-0)$, which is trivially true. Finally, condition 8.11 relates two vectors of one row block with one vector of the consecutive row block. Consider $\beta^{(0,1)}$, $\beta^{(0,2)}$ and $\beta^{(1,2)}$. We want

$$\beta_4^{(0,2)} - \beta_4^{(0,1)} - \beta_1^{(1,2)} \geq \beta_1^{(1,2)} - \beta_2^{(1,2)},$$

which is, also in this case, trivially satisfied since $4-0-3 \geq 3-2$.

Note that in Example 8.12, all the inequalities are in fact equalities. This is due on purpose, since this particular case will lead to the construction of a WRS convolutional code.

Next, we illustrate with another example the link between the matrices $G\left(x, Y, \mathrm{B}, \Lambda\right)$ – together with their parameters – and our family of WRS convolutional codes. Indeed, if we take a WRS convolutional code $\mathcal{C}_{k,n}^{\delta}(\gamma, \alpha)$ whose parameters satisfy certain conditions, then some of the full-size submatrices of $G_L^c(x)$ are obtained from $G\left(x, Y, \mathrm{B}, \Lambda\right)$ after carefully choosing the vectors of exponents B and Λ introduced in Definition 8.11, and specializing the vector Y in a suitable vector of elements α_{j_i}'s obtained from α. Observe that the following example only illustrates a special case and it is meant to guide the reader in understanding our approach. The general case is analyzed later, in the proof of Theorem 8.8.

Example 8.13. Let $(\ell_0, \ldots, \ell_e), (k_0, \ldots, k_e) \in \mathbb{N}^{e+1}$ be vectors as in Definition 8.11, with $k_i = k \in \mathbb{N}$, for $i = 0, 1, \ldots, e$. Define the vectors $\mathrm{B} = \left(\beta^{(0,0)}, \ldots, \beta^{(0,e)}, \beta^{(1,1)}, \ldots, \beta^{(1,e)}, \ldots, \beta^{(e,e)}\right)$ and $\Lambda := \left(\lambda^{(0,0)}, \ldots, \lambda^{(0,e)}, \lambda^{(1,1)}, \ldots, \lambda^{(1,e)}, \ldots, \lambda^{(e,e)}\right)$ as

$$\begin{aligned}
\beta^{(i,j)} &= \left(\tbinom{j-i+1}{2}k - (j-i), \ldots, \tbinom{j-i}{2}k + (j-i), \tbinom{j-i}{2}k\right), \\
\lambda^{(i,j)} &= \left((j-i+1)k - 1, \ldots, (j-i)k + 1, (j-i)k\right),
\end{aligned}$$

for each i, j such that $0 \leq j - i \leq e$. Let $Y := (y_1^{(0)}, \ldots, y_{\ell_0}^{(0)}, y_1^{(1)}, \ldots, y_{\ell_1}^{(1)}, \ldots, y_1^{(e)}, \ldots, y_{\ell_e}^{(e)})$ be the vector of variables and consider the matrix $G\left(x, Y, \mathrm{B}, \Lambda\right) \in \mathbb{F}_q[x, Y]^{(e+1)k \times (e+1)k}$.

Now, choose a WRS convolutional code $\mathcal{C}_{k,n}^{\delta}(\gamma, \alpha)$ for suitable γ and α, where we make two assumptions on the parameters. We select n, k and δ such that $\delta = ke$ and $ke < n-k$. The latter assumption implies $L = e$. Consider the generator matrix of $\mathcal{C}_{k,n}^{\delta}(\gamma, \alpha)$ to be $G(z) = \sum_{i=0}^{e} G_i z^i$ as in Definition 8.4, and take the polynomial version of its L-th truncated sliding generator matrix

$$G_L^{\mathrm{c}}(x) = \begin{pmatrix} G_0(x) & G_1(x) & \dots & G_e(x) \\ & G_0(x) & \dots & G_{e-1}(x) \\ & & \ddots & \vdots \\ & & & G_0(x) \end{pmatrix} \in \mathbb{F}_q[x]^{(e+1)k \times (e+1)n}.$$

We now point out that every full size submatrix of $G_L^{\mathrm{c}}(x)$ obtained by taking ℓ_i columns from the i-th column block, can be derived starting from $G\,(x, Y, \mathrm{B}, \Lambda)$ in the following way. Let $J_0, \dots J_e \subseteq \{1, \dots, n\}$ be the corresponding indices of columns that are selected in each block, with $|J_i| = \ell_i$. Then, specialize $y^{(i)} = (y_1^{(i)}, \dots, y_{\ell_i}^{(i)})$ in the elements $\alpha^{(i)} := (\alpha_j : j \in J_i)$. Denote $\mathrm{A} := (\alpha^{(0)}, \dots \alpha^{(e)})$ and observe now that, by construction, the selected submatrix of $G_L^{\mathrm{c}}(x)$ coincides with $G(x, \mathrm{A}, \mathrm{B}, \Lambda)$.

Finally, notice that the conditions on the ℓ_i's given in Definition 8.11 coincides with the conditions for the MDP property in (8.1). However, in this case it is not contemplated that the values ℓ_i's can also be zero. We will study how to obtain the full size submatrices of $G_L^{\mathrm{c}}(x)$ for this general case in the proof of Theorem 8.8.

From Definition 8.11, we have that $\ell_0 \leq k_0$, so we can assume that there exists $r \geq 0$ such that $k_0 = \ell_0 + r$. Moreover, from condition 8.11 it is clear that for any $0 \leq i \leq e$, $A_{i,i} \in \mathbb{F}_q[Y]^{k_i \times \ell_i}$, hence it does not depend on x.

We fix some further notation.

Notation 8.14. For any positive integer i, we denote $[i] := \{1, \dots, i\}$. For any $I \subseteq [\ell_0 + r]$, such that $|I| = \ell_0$, we denote by $A_I(Y)$ the $\ell_0 \times \ell_0$ matrix obtained from $A_{0,0}$ by selecting the rows indexed by I. We denote by I_0 the set of indices $[\ell_0]$.

Moreover, for any $I \subseteq [\ell_0 + r]$ we denote by $\bar{I}$ the complement of I in $[\ell_0 + r]$ and by $G_{\bar{I}}\,(x, Y, \mathrm{B}_{\bar{I}}, \Lambda_{\bar{I}})$ the $((k_0 - \ell_0) + k_1 + \dots + k_e) \times (\ell_1 + \dots + \ell_e)$ submatrix of $G\,(x, Y, \mathrm{B}, \Lambda)$ obtained by erasing the first ℓ_0 columns and the rows indexed by I. This deletion automatically determines two new collections of vectors $\mathrm{B}_{\bar{I}}$ and $\Lambda_{\bar{I}}$. We denote by $\beta_{\bar{I}}^{(0,j)}$ and by $\lambda_{\bar{I}}^{(0,j)}$ the vectors obtained from $\beta^{(0,j)}$ and $\lambda^{(0,j)}$ respectively, after deleting the entries indexed by I. Finally, we set

$$b_{\bar{I}} = \sum_{s \notin I} \beta_s^{(0,1)} \in \mathbb{N}.$$

Remark 8.15. In $\mathrm{B}_{\bar{I}} = (\beta_{\bar{I}}^{(i,j)})_{i,j}$, the deletion of the components indexed by I only regards $\beta^{(0,j)}$.

The results presented in the remainder of this section refer all to a matrix $G\,(x, Y, \mathrm{B}, \Lambda)$ of the form given in Definition 8.11, with component matrices $A^{\left(\beta^{(i,j)}, \lambda^{(i,j)}\right)}$, $0 \leq i, j \leq e$, where $\beta^{(i,j)}, \lambda^{(i,j)}$ satisfy conditions 8.11–8.11, 8.11–8.11 for any i, j.

Lemma 8.16. With the notation above, the following hold:

1. If $\ell_0 = k_0$, then the vectors forming $\mathrm{B}_{\bar{I}_0}$ and $\Lambda_{\bar{I}_0}$ – which define $G_{\bar{I}_0}\left(x, Y, \mathrm{B}_{\bar{I}_0}, \Lambda_{\bar{I}_0}\right)$ – satisfy conditions 8.11–8.11, 8.11–8.11.

2. If $\ell_0 < k_0$, then, for any $I \subseteq [\ell_0 + r]$ of cardinality ℓ_0, the matrix $G_{\bar{I}}\left(x, Y, \mathrm{B}_{\bar{I}}, \Lambda_{\bar{I}}\right)$ can be written as
$$\begin{pmatrix} \operatorname{diag}\left(x^{\beta_{\bar{I}}^{(0,1)}}\right) & 0 \\ 0 & \mathrm{Id} \end{pmatrix} \widetilde{G}_{\bar{I}}(x, Y, \widetilde{\mathrm{B}}_{\bar{I}}, \widetilde{\Lambda}_{\bar{I}}),$$
where Id is the identity matrix and $\widetilde{G}_{\bar{I}}(x, Y, \widetilde{\mathrm{B}}_{\bar{I}}, \widetilde{\Lambda}_{\bar{I}})$ is a $(k_0' + k_2 + \cdots + k_e) \times (\ell_1 + \cdots + \ell_e)$ matrix of the form (8.7), where $k_0' = k_0 + k_1 - \ell_0$ and whose defining vectors in $\widetilde{\mathrm{B}}_{\bar{I}}$ and $\widetilde{\Lambda}_{\bar{I}}$ satisfy conditions 8.11–8.11, 8.11–8.11.

Proof. 1. If $\ell_0 = k_0$, then
$$G_{\bar{I}_0}\left(x, Y, \mathrm{B}_{\bar{I}_0}, \Lambda_{\bar{I}_0}\right) = \begin{pmatrix} A'_{0,0} & A'_{0,1} & \cdots & A'_{0,e-1} \\ & A'_{1,1} & \cdots & A'_{1,e-1} \\ & & \ddots & \vdots \\ & & & A'_{e-1,e-1} \end{pmatrix},$$
where $A'_{i,j} = A_{i+1,j+1}$ for any $0 \leq i, j \leq e-1$, therefore the vectors $\beta^{(i,j)}, \lambda^{(i,j)}$ defining each $A_{i,j}$ clearly satisfy conditions 8.11–8.11, 8.11–8.11.

2. Assume $\ell_0 < k_0$. Note that, for any set of indices $I \subseteq [\ell_0 + r]$, $G_{\bar{I}}\left(x, Y, \mathrm{B}_{\bar{I}}, \Lambda_{\bar{I}}\right)$ is a $(k_0' + k_2 + \cdots + k_e) \times (\ell_1 + \cdots + \ell_e)$, where $k_0' = k_0 + k_1 - \ell_0$. Clearly,
$$G_{\bar{I}}\left(x, Y, \mathrm{B}_{\bar{I}}, \Lambda_{\bar{I}}\right) = \begin{pmatrix} \operatorname{diag}\left(x^{\beta_{\bar{I}}^{(0,1)}}\right) & 0 \\ 0 & \mathrm{Id} \end{pmatrix} \widetilde{G}_{\bar{I}}(x, Y, \widetilde{\mathrm{B}}_{\bar{I}}, \widetilde{\Lambda}_{\bar{I}}),$$
and
$$\widetilde{G}_{\bar{I}}(x, Y, \widetilde{\mathrm{B}}_{\bar{I}}, \widetilde{\Lambda}_{\bar{I}}) = \begin{pmatrix} A'_{0,0} & A'_{0,1} & \cdots & A'_{0,e-1} \\ & A'_{1,1} & \cdots & A'_{1,e-1} \\ & & \ddots & \vdots \\ & & & A'_{e-1,e-1} \end{pmatrix},$$
where $A'_{i,j} = A_{i+1,j+1}$ for any $1 \leq i, j \leq e-1$. Hence, for $i \geq 1$ all the conditions are satisfied. It is left to prove the result for $i = 0$.

Conditions 8.11–8.11 are trivially satisfied, since they are related to the vectors of exponents in the generalized Vandermonde matrices, on which we do not make operations.

If $\widetilde{\mathrm{B}}_{\bar{I}} = (\widetilde{\beta}^{(i,j)})_{i,j}$, then, because of Remark 8.15, it follows that for any $0 \leq j \leq e-1$, we have that

$$\widetilde{\beta}^{(0,j)} = \left(\beta_{\bar{I}}^{(0,j+1)} | \beta^{(1,j+1)}\right) - \left(\beta_{\bar{I}}^{(0,1)} | 0\right), \tag{8.8}$$

where 0 represents the zero vector and the difference is made componentwise. Here, we used the concatenation symbol just to stress that the deletion of components indexed by I only regards the 0-th row.

(b1) $\widetilde{\beta}^{(0,0)} = 0$.

(b2) To show this, we have to consider three cases:

i. If $s \in \{3, \dots, k_0 - \ell_0 - 2\}$ we consider $\widetilde{\beta}^{(0,j)} = \beta_{\bar{I}}^{(0,j+1)} - \beta_{\bar{I}}^{(0,1)}$, with abuse of notation, in the sense of (8.8). For any $1 \le h \le k_0 - \ell_0$, we denote by $\beta_{\bar{I},h}^{(0,j)}$ the h-th entry of $\beta_{\bar{I}}^{(0,j)}$. We need to verify that

$$\beta_{\bar{I},s-2}^{(0,j+1)} - \beta_{\bar{I},s-2}^{(0,1)} - \beta_{\bar{I},s-1}^{(0,j+1)} + \beta_{\bar{I},s-1}^{(0,1)} \ge \beta_{\bar{I},s-1}^{(0,j+1)} - \beta_{\bar{I},s-1}^{(0,1)} - \beta_{\bar{I},s}^{(0,j+1)} + \beta_{\bar{I},s}^{(0,1)}.$$

This is true, since by assumption

$$\beta_{\bar{I},s-2}^{(0,1)} - \beta_{\bar{I},s-1}^{(0,1)} \ge \beta_{\bar{I},s-1}^{(0,1)} - \beta_{\bar{I},s}^{(0,1)}$$

and

$$\beta_{\bar{I},s-2}^{(0,j+1)} - \beta_{\bar{I},s-1}^{(0,j+1)} \ge \beta_{\bar{I},s-1}^{(0,j+1)} - \beta_{\bar{I},s}^{(0,j+1)}.$$

ii. If $s \ge k_0 - \ell_0 + 2$, there is nothing to show, since $\widetilde{\beta}^{(0,j)} = \beta^{(1,j+1)}$ for any $0 \le j \le e - 2$.

iii. In the other case, the result is ensured by condition 8.11 on $\beta^{(0,j)}$.

(b3) Let $1 \le j \le e - 1$. For $i = 0$, we have that

$$\begin{aligned} \widetilde{\beta}_{k_0-1}^{(0,j)} &= \beta_{k_1-1}^{(1,j+1)}, \qquad & \widetilde{\beta}_{k_0}^{(0,j)} &= \beta_{k_1}^{(1,j+1)}, \\ \widetilde{\beta}_{1}^{(1,j)} &= \beta_{1}^{(2,j+1)}, \qquad & \widetilde{\beta}_{2}^{(1,j)} &= \beta_{2}^{(2,j+1)}. \end{aligned}$$

Hence, to verify that

$$\widetilde{\beta}_{k_0-1}^{(0,j)} - \widetilde{\beta}_{k_0}^{(0,j)} \ge \widetilde{\beta}_{k_0}^{(0,j)} - \widetilde{\beta}_{1}^{(1,j)} + 1 \ge \widetilde{\beta}_{2}^{(1,j)} - \widetilde{\beta}_{1}^{(1,j)} + 1,$$

we need to check that

$$\beta_{k_1-1}^{(1,j+1)} - \beta_{k_1}^{(1,j+1)} \ge \beta_{k_1}^{(1,j+1)} - \beta_{1}^{(2,j+1)} + 1 \ge \beta_{2}^{(2,j+1)} - \beta_{1}^{(2,j+1)} + 1,$$

that is true by assumption.

(b4) It follows from the expression (8.8), by using the same reasoning of part 8.11.

(b5) It follows by using the same reasoning of part 8.11.

□

Lemma 8.17. For any $I \subseteq [\ell_0 + r]$, such that $|I| = \ell_0$ and $I_0 = [\ell_0]$, it holds that $\beta^{(0,j)}_{\bar{I}_0} \leq \beta^{(0,j)}_{\bar{I}}$ componentwise, for any j.

Proof. The proof immediately follows from the conditions 8.11–8.11. □

Let $L = (\ell_0, \ldots, \ell_e)$, $K = (k_0, \ldots, k_e)$. We are going to estimate the minimum and the maximum degrees in x of the determinant of the matrix $G(x, Y, \mathrm{B}, \Lambda)$. We define

$$\begin{aligned} \mathrm{d}_{\min}(L, K, \mathrm{B}, \Lambda) &:= \min \deg_x(\det(G(x, Y, \mathrm{B}, \Lambda))) \in \mathbb{N} \cup \{\infty\}, \\ \mathrm{d}_{\max}(L, K, \mathrm{B}, \Lambda) &:= \max \deg_x(\det(G(x, Y, \mathrm{B}, \Lambda))) \in \mathbb{N} \cup \{\infty\}. \end{aligned}$$

In the following lemma, we observe that, whenever Λ is made of vectors satisfying the conditions 8.11–8.11, $\mathrm{d}_{\min}(L, K, \mathrm{B}, \Lambda)$ and $\mathrm{d}_{\max}(L, K, \mathrm{B}, \Lambda)$ only depend on the fixed L, K and B.

Lemma 8.18. Let $\Phi = (\phi^{(i,j)})_{i,j}$ and $\mathrm{B} = (\beta^{(i,j)})_{i,j}$, such that $\phi^{(i,j)}, \beta^{(i,j)}$ satisfy conditions 8.11–8.11 and $\phi^{(i,j)} \geq \beta^{(i,j)}$ componentwise for any i, j. Let $\Lambda = (\lambda^{(i,j)})_{i,j}$ be fixed, such that $\lambda^{(i,j)}$ satisfy conditions 8.11–8.11 for any i, j. Then, for any $\Upsilon = (\upsilon^{(i,j)})_{i,j}$, such that $\upsilon^{(i,j)}$ satisfies conditions 8.11–8.11, we have

$$\begin{aligned} \mathrm{d}_{\min}(L, K, \Phi, \Upsilon) &\geq \mathrm{d}_{\min}(L, K, \mathrm{B}, \Lambda), \\ \mathrm{d}_{\max}(L, K, \Phi, \Upsilon) &\geq \mathrm{d}_{\max}(L, K, \mathrm{B}, \Lambda). \end{aligned}$$

Proof. Observe that the determinant of $G(x, Y, \mathrm{B}, \Lambda)$ is a polynomial in x with coefficients in $\mathbb{F}_q[Y]$ and we can express it via the Leibniz formula. Let $N = \sum_{i=0}^{e} k_i = \sum_{i=0}^{e} \ell_i$ and $\mathcal{S}_N$ be the symmetric group of order N. Then

$$\det G(x, Y, \mathrm{B}, \Lambda) = \sum_{\sigma \in \mathcal{S}_N} \mathrm{sgn}(\sigma) \prod_{i=1}^{N} G(x, Y, \mathrm{B}, \Lambda)_{i,\sigma(i)} = \sum_{\sigma \in \mathcal{S}_N} \mathrm{sgn}(\sigma) G^{\sigma},$$

where $G^{\sigma} = \prod_{i=1}^{N} G(x, Y, \mathrm{B}, \Lambda)_{i,\sigma(i)} = R_{\sigma,\Lambda}(Y) x^{s_{\sigma,\mathrm{B}}}$, for a suitable $s_{\sigma,\mathrm{B}} \in \mathbb{N}$. Here, $R_{\sigma,\Lambda}(Y)$ is a monomial in Y, which can also be 0, depending on σ. Hence,

$$\det G(x, Y, \mathrm{B}, \Lambda) = \sum_{\sigma \in \mathcal{S}_N} R_{\sigma,\Lambda}(Y) x^{s_{\sigma,\mathrm{B}}} = \sum_{\sigma \in Z^{\Lambda}} R_{\sigma,\Lambda}(Y) x^{s_{\sigma,\mathrm{B}}},$$

where $Z^{\Lambda} := \{\sigma \in \mathcal{S}_N \mid G^{\sigma} \neq 0\}$. Obviously, as $\sigma(i) = i \in Z^{\Lambda}$, $Z^{\Lambda} \neq \emptyset$.

Observe that the elements $R_{\sigma,\Lambda}(Y)$, with $\sigma \in Z^{\Lambda}$ are $\mathbb{F}_q$-linearly independent. Indeed, they are monomials in Y and they all involve distinct exponents due to conditions 8.11–8.11 on the components of Λ. This remark is crucial for the rest of the proof.

Let

$$\begin{aligned} P_{\Phi,\Upsilon}(x,Y) &= \det(G(x,Y,\Phi,\Upsilon)) = \sum_{\sigma\in Z^{\Upsilon}} R_{\sigma,\Upsilon}(Y)x^{s_{\sigma,\Phi}}, \\ P_{\mathrm{B},\Lambda}(x,Y) &= \det(G(x,Y,\mathrm{B},\Lambda)) = \sum_{\sigma\in Z^{\Lambda}} R_{\sigma,\Lambda}(Y)x^{s_{\sigma,\mathrm{B}}}. \end{aligned}$$

The only thing to observe is that $R_{\sigma,\Upsilon}(Y)=0$ if and only if $R_{\sigma,\Lambda}(Y)=0$. This is true because of the previous observation. Indeed, these monomials in Y are linearly independent over $\mathbb{F}_q$, hence their nonzeroness depends only on the support of the matrix and not on their exponents. This implies that $Z^{\Upsilon}=Z^{\Lambda}$. In particular, if $R_{\sigma,\Lambda}(Y)x^{\sigma,\mathrm{B}}$ is the monomial in $\mathbb{F}_q[Y][x]$ corresponding to the minimum degree in x of $P_{\mathrm{B},\Lambda}(x,Y)$, then the monomial $R_{\sigma,\Upsilon}(Y)x^{s_{\sigma,\Phi}}$ corresponds to the minimum degree in x of $P_{\Phi,\Upsilon}(x,Y)$. Furthermore, by our assumptions on Φ and B, we have $s_{\sigma,\mathrm{B}}\leq s_{\sigma,\Phi}$. Hence, $\mathrm{d}_{\min}(L,K,\Phi,\Upsilon)\geq \mathrm{d}_{\min}(L,K,\mathrm{B},\Lambda)$. The same argument also holds for the maximum degree. □

We are now ready to present how to determine the monomial of minimum degree in x of $\det(G\left(x,Y,\mathrm{B},\Lambda\right))$.

Theorem 8.19. The determinant of $G\left(x,Y,\mathrm{B},\Lambda\right)$ is nonzero. Moreover, the monomial with minimum degree in x of $\det(G\left(x,Y,\mathrm{B},\Lambda\right))$[1] is given by the product of the $\ell_i\times\ell_i$ minors across the main diagonal, for $i=0,\dots,e$. More precisely, let L_0 be the set of the smallest ℓ_0 row indices of $A_{0,0}$, *i.e.*$L_0=\{1,2,\dots,\ell_0\}$. For any $i=1,\dots,e$ define L_i to be the set of the smallest ℓ_i row indices corresponding to the i-th column block of $G\left(x,Y,\mathrm{B},\Lambda\right)$ after deleting the rows indexed by $\cup_{j=0}^{i-1}L_j$, *i.e.*L_i is given by the first ℓ_i indices in $\{1,\dots,\sum_{j=0}^{i}k_j\}\setminus\cup_{j=0}^{i-1}L_j$. Then, the monomial with minimum degree in x of $\det(G\left(x,Y,\mathrm{B},\Lambda\right))$ is the product, for $i=0,\dots,e$, of the $\ell_i\times\ell_i$ minors whose rows are indexed by L_i and whose columns are the one corresponding to the i-th column block for $i=0,\dots,e$.

Proof. Recall that

$$G\left(x,Y,\mathrm{B},\Lambda\right)=\begin{pmatrix} A_{0,0} & A_{0,1} & \cdots & A_{0,e}\\ & A_{1,1} & \cdots & A_{1,e}\\ & & \ddots & \vdots\\ & & & A_{e,e}\end{pmatrix},$$

with $A_{i,j}=A_{i,j}^{\left(\beta^{(i,j)},\lambda^{(i,j)}\right)}$ and that $\lambda^{(i,j)}$ and $\beta^{(i,j)}$ satisfy the conditions 8.11–8.11 and 8.11–8.11 given in Definition 8.11.

We will prove the result by induction on e.

[1] For monomial with minimum degree in x of $\det(G\left(x,Y,\mathrm{B},\Lambda\right))$, we mean the monomial as an element in $\mathbb{F}_q[Y][x]$, of the form $p(Y)x^b$, where $p(Y)$ is in the ring of coefficients $\mathbb{F}_q[Y]$ and b is the smallest exponent of x involved in the expression of $\det(G\left(x,Y,\mathrm{B},\Lambda\right))$.

Base case $e = 0$: In this case, $k_0 = \ell_0$ and $A_{0,0}$ is a generalized Vandermonde matrix, whose determinant is a nonzero polynomial.

Induction case: Assume that the result is true for all the numbers of blocks up to e and prove it for $e+1$ blocks.

This time we are going to use Laplace formula to compute the determinant of $G(x, Y, \mathrm{B}, \Lambda)$.

$$\begin{aligned}
\det(G(x,Y,\mathrm{B},\Lambda)) &= \sum_{\substack{I \subseteq [\ell_0 + r] \\ |I| = \ell_0}} (\pm 1) \det(A_I(Y)) \det(G_{\bar{I}}(x, Y, \mathrm{B}_{\bar{I}}, \Lambda_{\bar{I}})) \\
&= \sum_{\substack{I \subseteq [\ell_0 + r] \\ |I| = \ell_0}} (\pm 1) \det(A_I(Y)) \det \begin{pmatrix} \operatorname{diag}\left(x^{\beta_{\bar{I}}^{(0,1)}}\right) & 0 \\ 0 & \mathrm{Id} \end{pmatrix} \det(\widetilde{G}_{\bar{I}}(x, Y, \mathrm{B}_{\bar{I}}, \Lambda_{\bar{I}})) \\
&= \sum_{\substack{I \subseteq [\ell_0 + r] \\ |I| = \ell_0}} (\pm 1) \det(A_I(Y)) x^{b_{\bar{I}}} \det(\widetilde{G}_{\bar{I}}(x, Y, \mathrm{B}_{\bar{I}}, \Lambda_{\bar{I}})) \\
&= \det(A_{I_0}(Y)) x^{b_{\bar{I}_0}} \det(\widetilde{G}_{\bar{I}_0}(x, Y, \mathrm{B}_{\bar{I}_0}, \Lambda_{\bar{I}_0})) + \\
&\sum_{\substack{I \subseteq [\ell_0 + r] \\ |I| = \ell_0 \\ I \neq [\ell_0]}} (\pm 1) \det(A_I(Y)) x^{b_{\bar{I}}} \det(\widetilde{G}_{\bar{I}}(x, Y, \mathrm{B}_{\bar{I}}, \Lambda_{\bar{I}})).
\end{aligned}$$

Observe that $\widetilde{G}_{\bar{I}}(x, Y, \mathrm{B}_{\bar{I}}, \Lambda_{\bar{I}})$ is composed by $e \times e$ blocks and by Lemma 8.16, the vectors in $\mathrm{B}_{\bar{I}}$ and $\Lambda_{\bar{I}}$ satisfy conditions 8.11–8.11 and 8.11–8.11, that is $\widetilde{G}_{\bar{I}}(x, Y, \mathrm{B}_{\bar{I}}, \Lambda_{\bar{I}})$ is of the form given in Definition 8.11.

Let $M := \mathrm{d}_{\min}(L_{\bar{I}_0}, K_{\bar{I}_0}, \mathrm{B}_{\bar{I}_0}, \Lambda_{\bar{I}_0})$. Now, from Lemma 8.17 and Lemma 8.18 we have that $M \leq \mathrm{d}_{\min}(L_{\bar{I}}, K_{\bar{I}}, \widetilde{\mathrm{B}}_{\bar{I}}, \widetilde{\Lambda}_{\bar{I}})$, for any $I \subseteq [\ell_0 + r]$, with $|I| = \ell_0$. Hence, we can write $\mathrm{d}_{\min}(L_{\bar{I}}, K_{\bar{I}}, \mathrm{B}_{\bar{I}}, \Lambda_{\bar{I}}) = M + s_I$, where $s_I \in \mathbb{N}$ depends on the chosen I.

Therefore,

$$\begin{aligned}
\det(G(x,Y,\mathrm{B},\Lambda)) = x^{b_{\bar{I}_0}} \det(A_{I_0}(Y)) x^M \left(P_{I_0}(Y) + x Q_{I_0}(x, Y)\right) + \\
+ \sum_{\substack{I \subseteq [\ell_0 + r] \\ |I| = \ell_0 \\ I \neq [\ell_0]}} (\pm 1) x^{b_I} x^{M + s_I} \det(A_I(Y)) \left(P_I(Y) + x Q_I(x, Y)\right),
\end{aligned}$$

where $P_{I_0}(Y), P_I(Y) \neq 0$, for any I.

By definition of $b_{\bar{I}}$ and $\beta^{(i,j)}$, it is evident that $b_{\bar{I}_0} < b_{\bar{I}}$ for any $I \neq I_0$. Hence, we have that M is the minimum degree of the determinant of $G_{\bar{I}_0}\left(x, Y, \mathrm{B}_{\bar{I}_0}, \Lambda_{\bar{I}_0}\right)$ and, by inductive hypothesis, the corresponding monomial is obtained by multiplying the $\ell_i \times \ell_i$ minors across the main diagonal for $i = 1, \ldots, e$. Therefore, the minimum degree in x of $\det(G(x, Y, \mathrm{B}, \Lambda))$ is $b_{\bar{I}_0} + M$ and its corresponding monomial is given by the product of the determinants obtained by selecting the first ℓ_0 rows and columns and the $\ell_i \times \ell_i$ minors across the main diagonal, as explained in the statement. □

We now provide an exhaustive example from which will illustrate Theorem 8.19.

Example 8.20. Next, we show an example of how to construct the matrix $G(x, Y, \mathrm{B}, \Lambda)$ given in Definition 8.11 and use Theorem 8.19 to determine the monomial of minimum degree in x of $\det(G\,(x, Y, \mathrm{B}, \Lambda))$. Let $e = 2$ and $(k_0, k_1, k_2) = (3, 4, 5), (\ell_0, \ell_1, \ell_2) = (2, 4, 6) \in \mathbb{N}^3$. The blocks $A_{i,j}$ with $0 \leq i \leq j \leq 2$ composing $G\,(x, Y, \mathrm{B}, \Lambda)$ are of the form

$$\begin{aligned} A_{i,j} &= \operatorname{diag}\left(x^{\beta^{(i,j)}}\right) V\left(\lambda^{(i,j)}, y^{(j)}\right) \\ &= \begin{pmatrix} x^{\beta_1^{(i,j)}} \left(y_1^{(j)}\right)^{\lambda_1^{(i,j)}} & x^{\beta_1^{(i,j)}} \left(y_2^{(j)}\right)^{\lambda_1^{(i,j)}} & \cdots & x^{\beta_1^{(i,j)}} \left(y_{\ell_j}^{(j)}\right)^{\lambda_1^{(i,j)}} \\ x^{\beta_2^{(i,j)}} \left(y_1^{(j)}\right)^{\lambda_2^{(i,j)}} & x^{\beta_2^{(i,j)}} \left(y_2^{(j)}\right)^{\lambda_2^{(i,j)}} & \cdots & x^{\beta_2^{(i,j)}} \left(y_{\ell_j}^{(j)}\right)^{\lambda_2^{(i,j)}} \\ \vdots & \vdots & \ddots & \vdots \\ x^{\beta_{k_i}^{(i,j)}} \left(y_1^{(j)}\right)^{\lambda_{k_i}^{(i,j)}} & x^{\beta_{k_i}^{(i,j)}} \left(y_2^{(j)}\right)^{\lambda_{k_i}^{(i,j)}} & \cdots & x^{\beta_{k_i}^{(i,j)}} \left(y_{\ell_j}^{(j)}\right)^{\lambda_{k_i}^{(i,j)}} \end{pmatrix}. \end{aligned}$$

Consider $Y := (y^{(0)}, y^{(1)}, y^{(2)})$ the vector of variables, where

$$\begin{aligned} y^{(0)} &= (y_1^{(0)}, y_2^{(0)}) = (y_1, y_2), \\ y^{(1)} &= (y_1^{(1)}, y_2^{(1)}, y_3^{(1)}, y_4^{(1)}) = (z_1, z_2, z_3, z_4), \\ y^{(2)} &= (y_1^{(2)}, y_2^{(2)}, y_3^{(2)}, y_4^{(2)}, y_5^{(2)}, y_6^{(2)}) = (w_1, w_2, w_3, w_4, w_5, w_6). \end{aligned}$$

Let $\beta^{(i,j)} = \left(\beta_1^{(i,j)}, \ldots, \beta_{k_i}^{(i,j)}\right)$ be the vector consisting of the powers of x of the rows of $A_{i,j}$, with $0 \leq i \leq j \leq 2$. We only take the values for a fixed column of $A_{i,j}$, since that every column has the same powers. Let

$$\mathrm{B} := (\beta^{(0,0)}, \beta^{(0,1)}, \beta^{(0,2)}, \beta^{(1,1)}, \beta^{(1,2)}, \beta^{(2,2)}),$$

where, in order to lighten the notation, we define

$$\begin{array}{lllllllll} \beta^{(0,0)} & = & (0,0,0) & \beta^{(0,1)} & = & (2,1,0) & \beta^{(0,2)} & = & (9,7,5) \\ & & & \beta^{(1,1)} & = & (0,0,0,0) & \beta^{(1,2)} & = & (3,2,1,0) \\ & & & & & & \beta^{(2,2)} & = & (0,0,0,0,0). \end{array}$$

Now, let $\lambda^{(i,j)} = \left(\lambda_1^{(i,j)}, \ldots, \lambda_{k_i}^{(i,j)}\right)$ be the vector composed by the powers of variables Y of a fixed column of $A_{i,j}$, with $0 \leq i \leq j \leq 2$. In our case, we obtain that

$$\Lambda := \left(\lambda^{(0,0)}, \lambda^{(0,1)}, \lambda^{(0,2)}, \lambda^{(1,1)}, \lambda^{(1,2)}, \lambda^{(2,2)}\right),$$

where

$$\begin{array}{rclrclrcl}
\lambda^{(0,0)} & = & (2,1,0) & \lambda^{(0,1)} & = & (6,5,4) & \lambda^{(0,2)} & = & (11,10,9) \\
 & & & \lambda^{(1,1)} & = & (3,2,1,0) & \lambda^{(1,2)} & = & (8,7,6,5) \\
 & & & & & & \lambda^{(2,2)} & = & (4,3,2,1,0).
\end{array}$$

It is easy to check that the values $\beta^{(i,j)}$ and $\lambda^{(i,j)}$ of the vectors B and Λ, respectively, satisfy the conditions of Definition 8.11.

Hence, the matrix $G\left(x, Y, \mathrm{B}, \Lambda\right)$ is given by

$$G(x,Y,\mathrm{B},\Lambda) = \begin{pmatrix} A_{0,0} & A_{0,1} & A_{0,2} \\ & A_{1,1} & A_{1,2} \\ & & A_{2,2} \end{pmatrix} =$$

$$\left(\begin{array}{cc|cccc|cccccc}
y_1^2 & y_2^2 & x^2z_1^6 & x^2z_2^6 & x^2z_3^6 & x^2z_4^6 & x^9w_1^{11} & x^9w_2^{11} & x^9w_3^{11} & x^9w_4^{11} & x^9w_5^{11} & x^9w_6^{11} \\
y_1 & y_2 & xz_1^5 & xz_2^5 & xz_3^5 & xz_4^5 & x^7w_1^{10} & x^7w_2^{10} & x^7w_3^{10} & x^7w_4^{10} & x^7w_5^{10} & x^7w_6^{10} \\
1 & 1 & z_1^4 & z_2^4 & z_3^4 & z_4^4 & x^5w_1^9 & x^5w_2^9 & x^5w_3^9 & x^5w_4^9 & x^5w_5^9 & x^5w_6^9 \\
\hline
 & & z_1^3 & z_2^3 & z_3^3 & z_4^3 & x^3w_1^8 & x^3w_2^8 & x^3w_3^8 & x^3w_4^8 & x^3w_5^8 & x^3w_6^8 \\
 & & z_1^2 & z_2^2 & z_3^2 & z_4^2 & x^2w_1^7 & x^2w_2^7 & x^2w_3^7 & x^2w_4^7 & x^2w_5^7 & x^2w_6^7 \\
 & & z_1 & z_2 & z_3 & z_4 & xw_1^6 & xw_2^6 & xw_3^6 & xw_4^6 & xw_5^6 & xw_6^6 \\
 & & 1 & 1 & 1 & 1 & w_1^5 & w_2^5 & w_3^5 & w_4^5 & w_5^5 & w_6^5 \\
\hline
 & & & & & & w_1^4 & w_2^4 & w_3^4 & w_4^4 & w_5^4 & w_6^4 \\
 & & & & & & w_1^3 & w_2^3 & w_3^3 & w_4^3 & w_5^3 & w_6^3 \\
 & & & & & & w_1^2 & w_2^2 & w_3^2 & w_4^2 & w_5^2 & w_6^2 \\
 & & & & & & w_1 & w_2 & w_3 & w_4 & w_5 & w_6 \\
 & & & & & & 1 & 1 & 1 & 1 & 1 & 1
\end{array}\right).$$

In order to compute the minimum degree in x of the determinant of $G\left(x, Y, \mathrm{B}, \Lambda\right)$ we use Theorem 8.19. To this end, we need to define the sets L_i for $i = 0, 1, 2$, given in Theorem 8.19. L_0 is the set of the smallest ℓ_0 row indices of $A_{0,0}$, that is $L_0 = \{1, 2\}$, and the rest of sets L_i are given by the first l_i indices in $\{1, \ldots, \sum_{j=0}^{i} k_j\} \setminus \cup_{j=0}^{i-1} L_j$. First, we obtain these previous subsets for any $i = 1, 2$. L_1 is composed by the first $\ell_1 = 4$ indices in

$$\{1, \ldots, k_0 + k_1\} \setminus L_0 = \{1, 2, 3, 4, 5, 6, 7\} \setminus \{1, 2\} = \{3, 4, 5, 6, 7\},$$

that is, $L_1 = \{3, 4, 5, 6\}$; and L_2 is composed by the first l_2 indices in

$$\{1, \ldots, k_0 + k_1 + k_2\} \setminus (L_0 \cup L_1) = \{1, 2, \ldots, 12\} \setminus \{1, 2, 3, 4, 5, 6\} = \{7, 8, 9, 10, 11, 12\},$$

that is, $L_2 = \{7, 8, 9, 10, 11, 12\}$. By applying Theorem 8.19, we have that the monomial of minimal degree in x of $\det(G\,(x, Y, \mathrm{B}, \Lambda))$ is the product of three square minors, whose rows are indexed by L_i for $i = 0, 1, 2$, given in the diagonal of the matrix $G\,(x, Y, \mathrm{B}, \Lambda)$

$$\begin{pmatrix}
y_1^2 & y_2^2 & x^2z_1^6 & x^2z_2^6 & x^2z_3^6 & x^2z_4^6 & x^9w_1^{11} & x^9w_2^{11} & x^9w_3^{11} & x^9w_4^{11} & x^9w_5^{11} & x^9w_6^{11} \\
y_1 & y_2 & xz_1^5 & xz_2^5 & xz_3^5 & xz_4^5 & x^7w_1^{10} & x^7w_2^{10} & x^7w_3^{10} & x^7w_4^{10} & x^7w_5^{10} & x^7w_6^{10} \\
1 & 1 & z_1^4 & z_2^4 & z_3^4 & z_4^4 & x^5w_1^9 & x^5w_2^9 & x^5w_3^9 & x^5w_4^9 & x^5w_5^9 & x^5w_6^9 \\
 & & z_1^3 & z_2^3 & z_3^3 & z_4^3 & x^3w_1^8 & x^3w_2^8 & x^3w_3^8 & x^3w_4^8 & x^3w_5^8 & x^3w_6^8 \\
 & & z_1^2 & z_2^2 & z_3^2 & z_4^2 & x^2w_1^7 & x^2w_2^7 & x^2w_3^7 & x^2w_4^7 & x^2w_5^7 & x^2w_6^7 \\
 & & z_1 & z_2 & z_3 & z_4 & xw_1^6 & xw_2^6 & xw_3^6 & xw_4^6 & xw_5^6 & xw_6^6 \\
 & & 1 & 1 & 1 & 1 & w_1^5 & w_2^5 & w_3^5 & w_4^5 & w_5^5 & w_6^5 \\
 & & & & & & w_1^4 & w_2^4 & w_3^4 & w_4^4 & w_5^4 & w_6^4 \\
 & & & & & & w_1^3 & w_2^3 & w_3^3 & w_4^3 & w_5^3 & w_6^3 \\
 & & & & & & w_1^2 & w_2^2 & w_3^2 & w_4^2 & w_5^2 & w_6^2 \\
 & & & & & & w_1 & w_2 & w_3 & w_4 & w_5 & w_6 \\
 & & & & & & 1 & 1 & 1 & 1 & 1 & 1
\end{pmatrix}$$

This product of determinants produces the monomial in $\mathbb{F}_q[Y][x]$ of minimal degree in x (that is 0) given by

$$y_1y_2z_1z_2z_3z_4(y_1 - y_2) \prod_{1 \le i < j \le 4} (z_i - z_j) \prod_{1 \le i < j \le 6} (w_i - w_j).$$

We also obtain a similar result for the maximum degree. This time the monomial of maximum degree is obtained by taking the product of another set of minors. However, we will omit the proof, since it is technical and it is based on the same idea of the proof of Theorem 8.19. Using induction and Laplace formula for computing the determinant, one reduces to the case of one block less, estimating the degrees of monomials by means of Lemma 8.18.

Theorem 8.21. *Let L_0 be the set of the highest ℓ_0 row indices of $A_{0,0}$, i.e. $L_0 = \{k_0 - \ell_0 + 1, k_0 - \ell_0 + 2, \ldots, k_0\}$. For any $i = 1, \ldots, e$ define L_i to be the set of the highest ℓ_i row indices corresponding to the i-th column block of $G\,(x, Y, \mathrm{B}, \Lambda)$ after deleting the rows indexed by $\cup_{j=0}^{i-1} L_j$, i.e.L_i is given by the highest ℓ_i indices in $\{1, \ldots, \sum_{j=0}^{i} k_j\} \setminus \cup_{j=0}^{i-1} L_j$. Then, the monomial with maximum degree in x of $\det(G\,(x, Y, \mathrm{B}, \Lambda))$ is the product, for $i = 0, \ldots, e$, of the $\ell_i \times \ell_i$ minors whose rows are indexed by L_i and whose columns are the one corresponding to the i-th column block for $i = 0, \ldots, e$.*

Example 8.22. Following the same setting as the one in Example 8.20, we illustrate how Theorem 8.21 works. In this case, we have that the monomial with maximal degree in x of $G\,(x, Y, \mathrm{B}, \Lambda)$ is obtained taking $L_0 = \{2, 3\}$, $L_1 = \{4, 5, 6, 7\}$ and $L_2 = \{1, 8, 9, 10, 11, 12\}$.

Graphically, the product of the three minors are given below

$$\begin{pmatrix}
y_1^2 & y_2^2 & x^2z_1^6 & x^2z_2^6 & x^2z_3^6 & x^2z_4^6 & x^9w_1^{11} & x^9w_2^{11} & x^9w_3^{11} & x^9w_4^{11} & x^9w_5^{11} & x^9w_6^{11} \\
y_1 & y_2 & xz_1^5 & xz_2^5 & xz_3^5 & xz_4^5 & x^7w_1^{10} & x^7w_2^{10} & x^7w_3^{10} & x^7w_4^{10} & x^7w_5^{10} & x^7w_6^{10} \\
1 & 1 & z_1^4 & z_2^4 & z_3^4 & z_4^4 & x^5w_1^9 & x^5w_2^9 & x^5w_3^9 & x^5w_4^9 & x^5w_5^9 & x^5w_6^9 \\
 & & z_1^3 & z_2^3 & z_3^3 & z_4^3 & x^3w_1^8 & x^3w_2^8 & x^3w_3^8 & x^3w_4^8 & x^3w_5^8 & x^3w_6^8 \\
 & & z_1^2 & z_2^2 & z_3^2 & z_4^2 & x^2w_1^7 & x^2w_2^7 & x^2w_3^7 & x^2w_4^7 & x^2w_5^7 & x^2w_6^7 \\
 & & z_1 & z_2 & z_3 & z_4 & xw_1^6 & xw_2^6 & xw_3^6 & xw_4^6 & xw_5^6 & xw_6^6 \\
 & & 1 & 1 & 1 & 1 & w_1^5 & w_2^5 & w_3^5 & w_4^5 & w_5^5 & w_6^5 \\
 & & & & & & w_1^4 & w_2^4 & w_3^4 & w_4^4 & w_5^4 & w_6^4 \\
 & & & & & & w_1^3 & w_2^3 & w_3^3 & w_4^3 & w_5^3 & w_6^3 \\
 & & & & & & w_1^2 & w_2^2 & w_3^2 & w_4^2 & w_5^2 & w_6^2 \\
 & & & & & & w_1 & w_2 & w_3 & w_4 & w_5 & w_6 \\
 & & & & & & 1 & 1 & 1 & 1 & 1 & 1
\end{pmatrix}.$$

We are now ready to give a proof of Theorem 8.8, which is based on the result of Theorem 8.19.

Proof of Theorem 8.8. Let $\mathcal{C}$ be the $(n,k,\delta)_{q^s}$ convolutional code $\mathcal{C}^{\delta}_{k,n}(\gamma,\alpha)$. Set $L := \left\lfloor \frac{\delta}{k} \right\rfloor + \left\lfloor \frac{\delta}{n-k} \right\rfloor$ and let

$$G^{\mathrm{c}}_L := \begin{pmatrix} G_0 & G_1 & \cdots & G_L \\ & G_0 & \cdots & G_{L-1} \\ & & \ddots & \vdots \\ & & & G_0 \end{pmatrix}$$

be the L-th truncated sliding generator matrix. Moreover, for $i = 0, \dots, L$, let $G_i(x)$ be the matrix G_i where we have substituted the element $\gamma \in \mathbb{F}_{q^s}$ with an algebraically independent variable x, and $G^{\mathrm{c}}_L(x)$ be corresponding L-th truncated sliding generator matrix. With this notation, we have that $G_i(\gamma) = G_i$ for $i = 0, \dots, L$ and hence $G^{\mathrm{c}}_L(\gamma) = G^{\mathrm{c}}_L$. Moreover, by Lemma 8.7, we only need to prove that $p(\gamma) \neq 0$ for every $p(x) \in \mathcal{P}(k,n,\delta,\alpha)$.
Now we divide the proof in two distinct cases.

__**Case I:** $\delta = km$__. In this case, every matrix G_i appearing as a block of G^{c}_L is either M_i or the zero matrix. With this setting, we analyze two subcases.

__**Case I-A:** $km < n-k$__. We consider the column blocks of $G^{\mathrm{c}}_L(x)$ indexed by $0, 1, \dots, L$. Let $F(x)$ be a square $(L+1)k \times (L+1)k$ submatrix of $G^{\mathrm{c}}_L(x)$ obtained by selecting $\bar{\ell}_i$ columns from the i-th column block, where the $\bar{\ell}_i$'s satisfy (8.1). Let $I_F := \{t_0, t_1, \dots, t_e\}$ be the set of indices of the column blocks involved in the selection of $F(x)$ (*i.e.* $a \in I_F$ if and only if there exists a column of $F(x)$ which comes from the a-th column block), where we have ordered them as $0 \leq t_0 < \ldots < t_e \leq e$. At this point, one can see that $F(x)$ is a block upper triangular matrix of the form

$$F(x) = \begin{pmatrix} A_{0,0} & A_{0,1} & \cdots & A_{0,e} \\ & A_{1,1} & \cdots & A_{1,e} \\ & & \ddots & \vdots \\ & & & A_{e,e} \end{pmatrix},$$

where $A_{0,j} \in \mathbb{F}_q[x]^{((t_0+1)k)\times \bar{\ell}_{t_j}}$, and $A_{i,j} \in \mathbb{F}_q[x]^{((t_i-t_{i-1})k)\times \bar{\ell}_{t_j}}$ for $1 \leq i \leq e$ and $0 \leq j \leq e$. Now, we are going to show that $F(x)$ is obtained from a matrix of the form $G\,(x, Y, \mathrm{B}, \Lambda)$, after a suitable specialization of Y. Let us define $k_0 := (t_0+1)k$, $k_i := (t_i - t_{i-1})k$ for $1 \leq i \leq e$ and $\ell_j := \bar{\ell}_{t_j}$ for $0 \leq j \leq e$. One can see that for each s such that $0 \leq s \leq e-1$, we have

$$\sum_{j=0}^{s} \ell_j = \sum_{j=0}^{t_s} \bar{\ell}_j \leq (t_s+1)k = (t_0+1)k + \sum_{j=1}^{s}(t_j - t_{j-1})k = \sum_{j=0}^{s} k_j,$$

and that

$$\sum_{j=0}^{e} \ell_j = \sum_{j=0}^{t_e} \bar{\ell}_j = (L+1)k = \sum_{j=0}^{e} k_j.$$

By definition, in each block $A_{i,j}$, every row is a monomial in x with constant degree, and hence it can be written as $\mathrm{diag}(x^{\beta^{(i,j)}})V(\lambda^{(i,j)}, \alpha^{(j)})$, for suitable $\beta^{(i,j)}, \lambda^{(i,j)} \in \mathbb{N}^{k_i}$, and some $\alpha^{(j)} \in \mathbb{F}_q^{\ell_j}$ obtained selecting ℓ_j entries from α. Lengthy computations show that the vectors $\beta^{(i,j)}$'s satisfy the conditions 8.11–8.11 and that the vectors $\lambda^{(i,j)}$'s satisfy the conditions 8.11–8.11. Define now $\widehat{\mathrm{B}} := (\beta^{(i,j)})_{i,j}, \widehat{\Lambda} := (\lambda^{(i,j)})_{i,j}$, $\mathrm{A} = (\alpha^{(j)})_j$. Hence, our submatrix $F(x)$ is obtained from the matrix $G(x, Y, \widehat{\mathrm{B}}, \widehat{\Lambda})$ by evaluating the Y in a suitable vector A of α_i's, *i.e.*$F(x) = G(x, \mathrm{A}, \widehat{\mathrm{B}}, \widehat{\Lambda})$.

By Theorem 8.19, the determinant of $G(x, Y, \widehat{\mathrm{B}}, \widehat{\Lambda})$ is a nonzero polynomial in x and Y, which we denote by $f(x, Y) = \sum_i f_i(Y)x^i$. Therefore, $\det(F(x)) = f(x, \mathrm{A})$. It remains to prove that $f(x, \mathrm{A})$ is still a nonzero polynomial, *i.e.*that there exists at least one index i such that $f_i(\mathrm{A}) \neq 0$. By Theorem 8.19 we know that the monomial of minimum degree in x is $f_b(Y)x^b$, for some $b \in \mathbb{N}$, and it is obtained by multiplying the $\ell_i \times \ell_i$ minors along the main diagonal. By the structure of these minors, we have that for every $i \in \{0, \dots, e\}$, there exist integers $b_i, t_i \in \mathbb{N}$ and a set $J_i \subseteq [n]$ with $|J_i| = \ell_i$, such that the corresponding $\ell_i \times \ell_i$ minor is given by

$$\left(\prod_{j \in J_i} \alpha_j\right)^{t_i} x^{b_i} \det(V(\alpha^{(i)}, (0, 1, \dots, k_i - 1))),$$

where J_i is selected from α in order to get $\alpha^{(i)}$, and where $b = b_0 + \cdots + b_e$. Since the matrices $V(\alpha^{(i)}, (0, 1, \dots, k_i - 1))$ are classical Vandermonde matrices with pairwise distinct and nonzero defining entries, we get $f_b(\mathrm{A}) \neq 0$ and so $f(x, \mathrm{A}) \neq 0$.

Notice that the product of the $\ell_i \times \ell_i$ minors along the main diagonal for $i = 0, \dots, e$ corresponds to the product of the $\bar{\ell}_i \times \bar{\ell}_i$ minors along the main diagonal for $i = 0, \dots, L$, if we use the convention that the determinant of a 0×0 matrix is 1.

This shows that the set $\mathcal{P}(k, n, km, \alpha)$ does not contain the zero-polynomial. Let $p(x) \in \mathcal{P}(k, n, km, \alpha)$. Since $p(x) \neq 0$, we can write $p(x) = x^{\nu(p(x))} p_1(x)$, where $p_1(x) \in \mathbb{F}_q[x]$ and $\deg(p_1(x)) = \deg(p(x)) - \nu(p(x))$. Since $s > \deg(p_1(x))$, we get that $p_1(\gamma) \neq 0$ and therefore $p(\gamma) \neq 0$.

<u>**Case I-B:** $km \geq n-k$</u>. Let $r := \left\lfloor \frac{\delta}{n-k} \right\rfloor = \left\lfloor \frac{km}{n-k} \right\rfloor$. Then we have $L = m+r$ and $G_{m+i}(x) = 0$ for $i = 1, \dots, r$. Therefore, the polynomial version $G_L^c(x) = G_{m+r}^c(x)$ of the L-th truncated sliding generator matrix is given by

$$G_L^c(x) := \begin{pmatrix} M_0(x) & M_1(x) & \cdots & M_m(x) & 0 & \cdots & 0 \\ & M_0(x) & \cdots & M_{m-1}(x) & M_m(x) & & \\ & & \ddots & & \ddots & \ddots & 0 \\ & & & M_0(x) & & \ddots & M_m(x) \\ & & & & M_0(x) & & M_{m-1}(x) \\ & & & & & \ddots & \vdots \\ & & & & & & M_0(x) \end{pmatrix},$$

where $M_i(x)$'s are the matrices defined as in (8.3) in which γ has been replaced by the variable x. We extend this matrix to the matrix

$$\tilde{G}_L^c(x) := \begin{pmatrix} M_0(x) & M_1(x) & \cdots & M_m(x) & M_{m+1}(x) & \cdots & M_{m+r}(x) \\ & M_0(x) & \cdots & M_{m-1}(x) & M_m(x) & & \vdots \\ & & \ddots & & \ddots & \ddots & M_{m+1}(x) \\ & & & M_0(x) & & \ddots & M_m(x) \\ & & & & M_0(x) & & M_{m-1}(x) \\ & & & & & \ddots & \vdots \\ & & & & & & M_0(x) \end{pmatrix}.$$

where we have replaced the 0 blocks in the topright part of $G_L^c(x)$ with the matrices $M_{m+i}(x)$.

Also in this case we choose $L+1$ integers $\bar{\ell}_0, \dots, \bar{\ell}_L$ with the constraint that $\sum_{j=0}^{s} \bar{\ell}_j \leq (s+1)k$, for $s = 0, \dots, L-1$ and $\sum_{j=0}^{L} \bar{\ell}_j = (L+1)k$, and consider a maximal submatrix of $G_L^c(x)$ obtained by selecting $\bar{\ell}_i$ columns from the i-th columns block, for $i = 0, \dots, L$, which we call $F(x)$. Moreover, let $\tilde{F}(x)$ be the corresponding $(L+1)k \times (L+1)k$ submatrix of $\tilde{G}_L^c(x)$. According to Lemma 8.7, we only need to prove that $\det(F(x)) \neq 0$, and then we can conclude as we did for Case **I-A**.

First we observe the following relations between $f(x) := \det(F(x))$ and $\tilde{f}(x) := \det(\tilde{F}(x))$. We can write $f(x) = \sum_{\sigma \in \mathcal{S}_{(L+1)k}} F^\sigma$ and $\tilde{f}(x) = \sum_{\sigma \in \mathcal{S}_{(L+1)k}} \tilde{F}^\sigma$, where $F^\sigma = \prod_i F_{i,\sigma(i)}$ and $\tilde{F}^\sigma = \prod_i \tilde{F}_{i,\sigma(i)}$. We define $\Theta_j := \{\sigma \in \mathcal{S}_{(L+1)k} \mid \deg_x(F^\sigma) = j\}$ and $\tilde{\Theta}_j := \{\sigma \in \mathcal{S}_{(L+1)k} \mid$

$\deg_x(\tilde{F}^\sigma) = j\}$. We have

$$f(x) = \sum_j f_j x^j = \sum_j \left(\sum_{\sigma \in \Theta_j} F^\sigma \right) x^j,$$
$$\tilde{f}(x) = \sum_j \tilde{f}_j x^j = \sum_j \left(\sum_{\sigma \in \tilde{\Theta}_j} \tilde{F}^\sigma \right) x^j.$$

By definition of $G_L^{\mathrm{c}}(x)$ and $\tilde{G}_L^{\mathrm{c}}(x)$, every entry of $G_L^{\mathrm{c}}(x)$ is equal to the corresponding entry of $\tilde{G}_L^{\mathrm{c}}(x)$ or it is equal to 0. Therefore, we get that $F^\sigma \in \{0, \tilde{F}^\sigma\}$, for every $\sigma \in \mathcal{S}_{(L+1)k}$. Hence we can also write

$$\tilde{f}(x) = \sum_j \tilde{f}_j x^j = \sum_j \left(\sum_{\sigma \in \Theta_j} \tilde{F}^\sigma \right) x^j.$$

In particular, if a monomial $\tilde{f}_t x^t$ of a certain degree t in $\tilde{f}(x)$ is obtained from the matrix $\tilde{G}_L^{\mathrm{c}}(x)$ without involving the blocks $G_{m+i}(x)$ for $i = 1, \ldots, r$, then the monomial $f_t x^t$ of the same degree t in $f(x)$ is the same, *i.e.*$f_t = \tilde{f}_t$.

By using the same proof as in Case I-A, we can see that we can use Theorem 8.19 to show that $\tilde{f}(x) = \det(\tilde{F}(x)) \neq 0$ and the monomial $\tilde{f}_M x^M$ of minimum degree M corresponds to the product of the $\bar{\ell}_i \times \bar{\ell}_i$ minors along the main diagonal (observe that here we allow $\bar{\ell}_i$ to be zero, by using the convention that the determinant of a 0×0 matrix to be 1). If we show that these minors do not involve any of the blocks $G_{m+j}(x)$ for $j = 1, \ldots, r$, then we can deduce that $f_M = \tilde{f}_M \neq 0$, and this concludes the proof. Suppose by contradiction that one of the $\bar{\ell}_i \times \bar{\ell}_i$ minors involves one of the blocks $G_{m+j}(x)$ for $j = 1, \ldots, r$. Then it must be $i = m + a$ for some $a = j, \ldots, r$. Moreover, this happens if and only if $(m + r - a + j + 1)k < \sum_{t=a}^r \bar{\ell}_{m+t}$. However, we have $\sum_{t=a}^r \bar{\ell}_{m+t} \leq n(r - a + 1)$ and since $r = \lfloor \frac{km}{n-k} \rfloor$, also $(n-k)r \leq km$. Therefore, we get the following chain of inequalities

$$\begin{aligned} n(r - a + j + 1) &= nr + n(1 - a + j) \leq km + kr + n(1 - a + j) \\ &\leq km + kr + k(1 - a + j) = k(m + r - a + j + 1) \\ &< \sum_{t=a}^r \bar{\ell}_{m+t} \leq n(r - a + 1), \end{aligned}$$

which yields a contradiction.

Case II: $k \nmid \delta$. In this case, $G_m = N_{m,t}$, where $1 \leq t = \delta - (m-1)k \leq k - 1$. At this point, we use a similar argument as done in Case **I-B**. Let $L = m - 1 + \lfloor \frac{k(m-1)+t}{n-k} \rfloor = m - 1 + r$, for some $r \geq 0$. Consider the polynomial matrix $G_L^{\mathrm{c}}(x)$. If $r = 0$, then the matrix $N_{m,t}(x)$ does not appear in $G_L^{\mathrm{c}}(x)$ and we conclude as in Case **I-A**. Therefore, assume $r \geq 1$. Observe that the matrix $N_{m,t}(x)$ is 0 in the first $k - t$ rows, and it coincides with the matrix $M_m(x)$ in the last t rows, We construct the matrix $\tilde{G}_L^{\mathrm{c}}(x)$ from $G_L^{\mathrm{c}}(x)$ by replacing the blocks $N_{m,t}(x)$ by $M_m(x)$

and the topright 0 blocks by matrices $M_{m+i}(x)$, for $i = 1, \dots, r$, obtaining

$$\tilde{G}_L^{\mathrm{c}}(x) := \begin{pmatrix} M_0(x) & M_1(x) & \cdots & M_m(x) & M_{m+1}(x) & \cdots & M_{m+r-1}(x) \\ & M_0(x) & \cdots & M_{m-1}(x) & M_m(x) & & \vdots \\ & & \ddots & & \ddots & \ddots & M_{m+1}(x) \\ & & & M_0(x) & & \ddots & M_m(x) \\ & & & & M_0(x) & & M_{m-1}(x) \\ & & & & & \ddots & \vdots \\ & & & & & & M_0(x) \end{pmatrix}.$$

Now, for the same argument used in Case **I-B**, we just need to prove that none of $\bar{\ell}_i \times \bar{\ell}_i$ minors across the main diagonal involves any of the blocks $M_{m+j}(x)$, for $j \geq 1$, and neither the first $k - t$ rows of $M_m(x)$. By contradiction, assume that this happens in the block $m + a$, for some $a = 0, \dots, r - 1$. This is true if and only if $\bar{\ell}_{m+a} + \ldots + \bar{\ell}_{m+r-1} > \delta + k(r - a)$. However, we have $\sum_{t=a}^{r-1} \bar{\ell}_{m+t} \leq n(r - a)$ and since $r = \lfloor \frac{\delta}{n-k} \rfloor$, also $(n - k)r \leq \delta$. Therefore, we get the following chain of inequalities

$$\begin{aligned} n(r - a) \leq \delta + kr - na &\leq \delta + k(r - a) \\ &< \sum_{t=a}^{r-1} \bar{\ell}_{m+t} \leq n(r - a), \end{aligned}$$

which yields a contradiction. □

Remark 8.23. It is important to point out that the same proof does not work if we try to use Theorem 8.21 instead of Theorem 8.19. Indeed, if k divides δ, whenever we select $\ell_0, \dots, \ell_e$ columns from the $e + 1$ column blocks and we have that one of the ℓ_i's for $i = 1, \dots, e$ is strictly greater than k, then the $\ell_i \times \ell_i$ submatrix that we select is not a classical Vandermonde, hence it is not guaranteed that when we evaluate the variables Y in A we get a nonzero determinant. This happens essentially for every choice of ℓ_i's, except when we take $\ell_i = k$ for every $i = 0, \dots, e$, in which case the minor is clearly nonzero, since it is the product of determinants of classical Vandermonde matrices. In the case that k does not divide δ it is even more clear that we cannot use Theorem 8.21, since the $\ell_e \times \ell_e$ submatrix that we should select will we have some zero rows.

8.3 Field Size for MDP WRS Convolutional Codes

In Section 8.1 we gave a construction of WRS convolutional codes, which we proved in Theorem 8.8 to be MDP, under the assumption that the extension degree s is larger than the value $D(k, n, \delta, \alpha)$. Our goal in this section is to give an estimate on the required field size. In particular, we will prove that for constructing WRS convolutional codes that are MDP we need a field of size q^s, where q is any prime power greater than n and $s = \mathcal{O}(\delta^3)$. It is straightforward to observe that for our base field $\mathbb{F}_q$ from which we take the vector α we need that $q > n$. This is because we require that α has pairwise distinct nonzero elements. The only thing that we need to estimate is the magnitude of the degree extension s, or equivalently, of the integer $D(k, n, \delta, \alpha)$.

Proposition 8.24. Let k, n, δ be integers such that $0 < k < n$, let $m := \lceil \frac{\delta}{k} \rceil$ and let $\alpha \in (\mathbb{F}_q^*)^n$ be a vector of nonzero pairwise distinct elements. Then,

$$D(k, n, \delta, \alpha) \leq (L - m + 1)\binom{\delta}{2} + k^2\binom{m}{3} + \binom{k}{2}\binom{m}{2}.$$

Proof. We provide an upper bound on the value $D(k, n, \delta, \alpha)$ by estimating the maximum degree of the polynomials in $\mathcal{P}(k, n, \delta, \alpha)$. To this end, we take the obvious upper bound in which we consider in each row of $G_L^{\mathrm{c}}(x)$ the maximum degree, and then we sum up all these values. We divide the matrix in row blocks, indexed by $0, 1, \ldots, L$. Observe that in the row block L only the matrix G_0 appears, in the row block $(L-1)$ the highest degrees are given by the row degrees in x of $G_1(x)$, and so on. This means that for each $i \in \{0, \ldots, m-1\}$, the row block $(L-i)$ has highest row degrees in x given by those of $G_i(x)$. These matrices correspond to the matrices $M_i(x)$ and it is straightforward to see that the sum of the row degrees in x of the matrix $M_i(x)$ is

$$w_i := \sum_{j=0}^{k-1} \binom{i}{2} k + ij = \binom{i}{2} k^2 + i\binom{k}{2}.$$

For the remaining blocks, the degrees are the same due to the structure of $G_L^{\mathrm{c}}(x)$ in which all the matrices G_i's appear. Hence, we only can consider $(L - m + 1)$ times the sum of row degrees of the first block. Observe that $G_m(x) = N_{m,t}(x)$, where $t := \delta - k(m-1)$. Hence, the row degrees of this first block will consist of the row degrees of the last t rows of $N_{m,t}(x)$, and the first $k - t$ row degrees of $M_{m-1}(x)$. Summing this quantities we get that the sum of the row degrees in x of the first block is

$$\begin{aligned}
w_{m,t} &:= \sum_{j=0}^{t-1} \left(\binom{m}{2} k + mi\right) + \sum_{j=t}^{k-1} \left(\binom{m-1}{2} + (m-1)j\right) \\
&= \frac{1}{2} t \left(km(m-1) + (t-1)m\right) + \frac{1}{2}(k-t)\left(k(m-1)(m-2) + (k+t-1)(m-1)\right) \\
&= \frac{1}{2}\left(t(\delta-1)m + (k-t)(\delta-1)(m-1)\right) \\
&= \frac{1}{2}(\delta-1)\left(k(m-1) + t\right) \\
&= \binom{\delta}{2}.
\end{aligned}$$

Putting together all these quantities, we get

$$\begin{aligned}
D(k,n,\delta,\alpha) &\leq (L-m+1)w_{m,t} + \sum_{i=0}^{m-1} w_i \\
&= (L-m+1)\binom{\delta}{2} + k^2 \sum_{i=0}^{m-1}\binom{i}{2} + \binom{k}{2}\sum_{i=0}^{m-1} i \\
&= (L-m+1)\binom{\delta}{2} + k^2\binom{m}{3} + \binom{k}{2}\binom{m}{2}.
\end{aligned}$$

□

Next, we show an example considering our WRS convolutional codes where we compute the value of $D(k,n,\delta,\alpha)$ and compare it with the upper bound of Proposition 8.24.

Example 8.25. Suppose that we want to construct a WRS convolutional code with parameters $k=2$, $n=7$ and $\delta=4$. We have that its memory m has to be 2 and also L is equal to 2. Moreover, we need to start with a base field $\mathbb{F}_q$ whose cardinality is $q>7$. In this example we consider the case $q=11$, that is $\mathbb{F}_{11}=\{0,1,\dots,10\}$. We now fix the vector $\alpha\in\mathbb{F}_{11}^7$ to have nonzero pairwise distinct entries as $\alpha := (1,2,3,4,5,6,7)$. In order to determine a suitable γ for constructing the code $\mathcal{C}_{2,7}^4(\gamma,\alpha)$, we first want to compute the value $D(2,7,4,\alpha)$. For this purpose, we are now going to construct the matrix $G_2^{\mathrm{c}}(x)$.

$$\begin{aligned}
G_0(x) &= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}, \\
G_1(x) &= \begin{pmatrix} x & 8x & 5x & 9x & 4x & 7x & 2x \\ 1 & 4 & 9 & 5 & 3 & 3 & 5 \end{pmatrix}, \\
G_2(x) &= \begin{pmatrix} x^4 & 10x^4 & x^4 & x^4 & x^4 & 10x^4 & 10x^4 \\ x^2 & 5x^2 & 4x^2 & 3x^2 & 9x^2 & 9x^2 & 3x^2 \end{pmatrix}.
\end{aligned}$$

The L-th truncated sliding matrix $G_L^{\mathrm{c}}(x)$ is

$$G_2^{\mathrm{c}}(x) = \begin{pmatrix} G_0(x) & G_1(x) & G_2(x) \\ & G_0(x) & G_1(x) \\ & & G_0(x) \end{pmatrix} \in \mathbb{F}_{11}[x]^{6\times 21}.$$

By Theorem 8.7, the matrix $G(z)$ generates a $\mathcal{C}_{2,7}^4(\gamma,\alpha)$ MDP convolutional code if it has the MDP property, *i.e.*, the nontrivial full size minors of G_2^{c} are all nonzero. A necessary field size for this to hold is $\mathbb{F}_{11^s}$ with $s > D(2,7,4,\alpha)$, according to Theorem 8.8. It can be checked that

$D(2,7,4,\alpha) = 4$, that comes from the full size minor $p(x) = 7x^4 + 2x^2 + 7x + 4 \in \mathcal{P}(2,7,4,\alpha)$ of $G_2^{\mathfrak{c}}(x)$ obtained selecting the columns with indices $\{1,8,9,15,16,17\}$. From Proposition 8.24, we have that $D(2,7,4,\alpha) \leq 7$, and we can observe that the bound is bigger than the real value.

Theorem 8.26. The WRS convolutional codes $\mathcal{C}_{k,n}^{\delta}(\gamma,\alpha)$ provide a family of MDP convolutional code over a field of size q^s, where $q > n$ and

$$s > \frac{\delta^3}{2}\left(\frac{1}{(n-k)} + \frac{1}{3k}\right) + \frac{\delta^2}{2}\left(\frac{3}{2} - \frac{1}{(n-k)} - \frac{1}{2k}\right) + \delta\left(\frac{k}{12} - \frac{3}{4}\right).$$

Proof. By Theorem 8.8, we get that the WRS convolutional code $\mathcal{C}_{k,n}^{\delta}(\gamma,\alpha)$ is MDP whenever α is a vector of pairwise distinct elements of $\mathbb{F}_q^*$, and γ is a root of an irreducible polynomial in $\mathbb{F}_q[x]$ of degree $s > D(k,n,\delta,\alpha)$. Then, we need $q > n$ and the result follows from an involved estimates of the bound on $D(k,n,\delta,\alpha)$ given in Proposition 8.24, which produce

$$D(k,n,\delta,\alpha) \leq \frac{\delta^3}{2}\left(\frac{1}{(n-k)} + \frac{1}{3k}\right) + \frac{\delta^2}{2}\left(\frac{3}{2} - \frac{1}{(n-k)} - \frac{1}{2k}\right) + \delta\left(\frac{k}{12} - \frac{3}{4}\right).$$

□

Corollary 8.27. For every k, n, δ positive integers with $0 < k < n-k$, there exists an $(n,k,\delta)_{q^s}$ MDP convolutional code where q is any prime power greater than n and

$$s > \frac{\delta^3}{3t} + \frac{3\delta^2}{4} + \frac{\delta k}{12}.$$

Proof. By Theorem 8.26 we can construct an $(n,k,\delta)_{q^s}$ MDP WRS convolutional code $\mathcal{C}_{k,n}^{\delta}(\gamma,\alpha)$, where $q > n$ and $s > \frac{\delta^3}{2}\left(\frac{1}{(n-k)} + \frac{1}{3k}\right) + \frac{3\delta^2}{4} + \frac{\delta k}{4}$. Since $1/(n-k) \leq 1/k$, we obtain that for $s > \frac{\delta^3}{3k} + \frac{3\delta^2}{4} + \frac{\delta k}{4}$ we can construct an MDP WRS convolutional code with the desired parameters.

□

The above results only provide upper bounds on the values $D(k,n,\delta,\alpha)$. In Table 8.1 we depict the actual values for some small parameters of k, n and δ (found by exhaustive computer search) and compare them with the bound given in Proposition 8.24.

8.3.1 Field Size for Memory 1

In Proposition 8.24 we gave a general upper bound on the value $D(k,n,\delta,\alpha)$, which then produces a sufficient field size for constructing WRS convolutional codes. Here we improve this upper bound when the resulting code has memory $m = 1$, that is when $\delta \leq k$. In order to do so, we will distinguish two cases: when $k < n-k$ and when $k = n-k$.

$[n,k,\delta]$	$D(k,n,\delta,\alpha)$ bound	$D(k,n,\delta,\alpha)$
[2,1,1]	0	0
[2,1,2]	3	2
[3,2,2]	3	3
[3,1,2]	2	1
[3,2,1]	0	0
[4,2,2]	2	2
[4,1,3]	7	4
[5,2,2]	1	1
[6,2,2]	1	1
[6,2,3]	1	1
[7,2,2]	1	1
[7,3,3]	3	2

Table 8.1: Comparison between the values $D(k,n,\delta,\alpha)$ obtained by computer search and the corresponding bounds given in Proposition 8.24, for small parameters n,k,δ.

For the first case, we will assume that $k<n-k$, that is when the WRS convolutional code has memory 1. We first notice that when $1\leq\delta<k<n-k$, we have $L=0$, and hence $D(k,n,\delta,\alpha)=0$. Therefore, the only interesting case is when $\delta=k$.

Proposition 8.28. Suppose that $k<n-k$, Then

$$D(k,n,k,\alpha)\leq\frac{k^2}{4}.$$

Proof. Since we are in the case $k<n-k$, we have $L=m=1$, and hence the matrix $G_L^{\mathrm{c}}(x)$ is of the form

$$G_1^{\mathrm{c}}(x)=\begin{pmatrix}G_0 & G_1(x)\\ 0 & G_0\end{pmatrix}$$

For every $2k\times 2k$ submatrix of $G_L^{\mathrm{c}}(x)$ that we need to consider, we can choose $k-i$ columns from the first block, and $k+i$ columns from the second block, for some $i=0,\dots,k$. Let us fix the value i, and two subsets $I_1,I_2\subseteq[n]$ such that $|I_1|=k-i,|I_2|=k+i$. Let $F(x)$ be the submatrix obtained by selecting the columns indexed by I_1 from the first block, and the columns indexed by I_2 from the second block, and let $f(x):=\det(F(x))$. Then

$$F(x)=\begin{pmatrix}A_0 & B_1(x)\\ 0 & B_0\end{pmatrix},$$

where $A_0\in\mathbb{F}_q^{k\times(k-i)},B_0\in\mathbb{F}_q^{k\times(k+i)},B_1(x)\in\mathbb{F}_q[x]^{k\times(k+i)}$. Computing $f(x)$ using Laplace formula on the first $k-i$ columns, we observe that the determinant will always be given by sum of products of a $(k-i)\times(k-i)$ minor of A_0 times the determinant of the remaining $(k+i)\times(k+i)$ minor in the second column block. Hence, the degree of $f(x)$ is at most the sum of the degrees of the first i rows of $B_1(x)$, that is given by

$$\sum_{j=k-i}^{k-1} j.$$

By Theorem 8.19, the minimum degree monomial of $f(x)$ is given by the product of the top $(k-i)\times(k-i)$ minor of A_0 times the determinant of the remaining submatrix of $\begin{pmatrix} B_1(x) \\ B_0 \end{pmatrix}$. The degree of this monomial is exactly $\sum_{j=0}^{i-1} j$, and thus

$$\deg(f(x)) - \nu(f(x)) \leq \sum_{j=k-i}^{k-1} j - \sum_{j=0}^{i-1} j = \sum_{j=0}^{i-1} (k-i+j-j) = i(k-i).$$

Therefore, we get

$$\begin{aligned} D(k,n,k,\alpha) &= \max\{\deg f(x) - \nu(f(x)) \mid 0 \neq f(x) \in \mathcal{P}(k,n,\delta,\alpha)\} \\ &\leq \max\{i(k-i) \mid i = 0,\dots,k\} \\ &\leq \frac{k^2}{4}. \end{aligned}$$

□

Notice that in the above analysis we only left out the case when $n = 2k$. Also in this case, if $\delta < k$, we have $L = 0$ and hence $D(k,2k,\delta,\alpha) = 0$. Therefore, the only case left to study is $D(k,2k,k,\alpha)$, which we do in the following proposition.

Proposition 8.29. For every positive integer k, we have

$$D(k,2k,k,\alpha) \leq \frac{k^2}{2}.$$

Proof. In this case we have $m = 1$ and $L = 2$. Hence, the polynomial version of the sliding generator matrix is of the form

$$G_2^{\mathrm{c}}(x) = \begin{pmatrix} G_0 & G_1(x) & 0 \\ 0 & G_0 & G_1(x) \\ 0 & 0 & G_0 \end{pmatrix}.$$

By the restriction of (8.1), the only admissible full size submatrices we should consider are those obtained from $G_2^{\mathrm{c}}(x)$ by selecting $\ell_0 = k-i$ columns from the first columns block, $\ell_1 = k-j+i$ columns from the second columns block and $\ell_2 = k+j$ columns from the last columns block, for any i, j such that $0 \leq i, j \leq k$. For a submatrix made with this choice, that we indicate as

$$F(x) = \begin{pmatrix} A_0 & B_1(x) & 0 \\ 0 & B_0 & C_1(x) \\ 0 & 0 & C_0 \end{pmatrix},$$

we can compute the minimum degree in x of the resulting minor $f(x) = \det(F(x))$, that is $\nu(f(x))$. Even though there is a 0-block in the top right corner, we can still use Theorem 8.19, by using the same argument used in the proof of Theorem 8.8, and obtain that $\nu(f(x)) = \sum_{t=0}^{i-1} t + \sum_{t=0}^{j-1} t = \binom{i}{2} + \binom{j}{2}$. The degree of $f(x)$ is going to be obtained by selecting the product of the minors of the following three submatrices: the $\ell_0 \times \ell_0$ submatrix obtained selecting the last ℓ_0 of A_0; the $\ell_1 \times \ell_1$ submatrix obtained by selecting the first i rows of $B_1(x)$ and the last $k - j$ rows of B_0; the $\ell_2 \times \ell_2$ submatrix obtained by selecting the first j rows of $C_1(x)$ together with the whole matrix C_0. This gives the maximum possible degree, but since for some choice of the vector α this could be 0, we have

$$\deg(f(x)) \leq \sum_{t=k-i}^{k-1} t + \sum_{t=k-j}^{k-1} t,$$

and therefore,

$$\deg(f(x)) - \nu(f(x)) \leq \sum_{t=k-i}^{k-1} t + \sum_{t=k-j}^{k-1} t - \binom{i}{2} - \binom{j}{2} = i(k-i) + j(k-j).$$

Thus,

$$\begin{aligned} D(k, 2k, k, \alpha) &= \max\{\deg f(x) - \nu(f(x)) \mid 0 \neq f(x) \in \mathcal{P}(k, n, \delta, \alpha)\} \\ &\leq \max\{i(k-i) + j(k-j) \mid i, j = 0, \ldots, k\} \\ &\leq \frac{k^2}{2}. \end{aligned}$$

□

Remark 8.30. Proposition 8.28 and Proposition 8.29 improve on the more general bound given in Proposition 8.24. Indeed, with the latter result we obtain $D(k, n, k, \alpha) \leq \frac{k(k-1)}{2}$ when $k < n - k$, and $D(k, 2k, k, \alpha) \leq k(k-1)$. Therefore, in both cases we refine the estimate by a factor 1/2.

8.3.2 Comparison

In this subsection we present a comparison of the existing fields sizes required to build MDP convolutional codes for several sets of given parameters (n, k, δ). The compared results are of different nature and need to be distinguished. Some of them are general bounds on the field size but with no associated concrete construction achieving such a bound. Others are conjectures or examples found by computer search. These differences are explained and analyzed in this subsection.

As mentioned before, superregular matrices have been one of the fundamental tools for constructing MDP codes and the required field size to construct the codes has been often given in terms of the field size needed to build the associated superregular Toeplitz matrices. An upper triangular Toeplitz matrix

$$A = \begin{pmatrix} a_0 & \cdots & a_{r-1} & a_r \\ & a_0 & \cdots & a_{r-1} \\ & & \ddots & \vdots \\ & & & a_0 \end{pmatrix} \tag{8.9}$$

is superregular if $a_i \neq 0$ for each $i = 0, 1, \dots, r$ and all the square submatrices of A with no zeros in the diagonal are nonsingular. It can be verified [121] that in order to built an (n, k, δ) MDP convolutional code an upper triangular Toeplitz superregular matrix, A, of size greater or equal than

$$r := \max\{n-k, k\}(L+1) + \min\{n-k, k\} - 1, \tag{8.10}$$

needs to be constructed.

In [12] and [79] two general classes of superregular matrices of any size were presented. The lower bound on the field size required to build the superregular lower triangular Toeplitz matrix $A \in \mathbb{F}^{r\times r}$ in [79] is $|\mathbb{F}| > c^r r^{r/2}$ where $c = \binom{r-1}{\lfloor \frac{r-1}{2} \rfloor}$. For the one provided in [12] the lower bound is given by $|\mathbb{F}| \geq 2^{(2^{(r+2)})}$. For upper bounds on the size of a field $\mathbb{F}$ to ensure the existence (without providing a concrete construction) of a superregular lower triangular Toeplitz matrix over $\mathbb{F}$, see [89] and [102]. Based on examples derived by computer search, it was conjectured in [89, Conjecture 3.5] and [79] that for $r \geq 5$ there exists a superregular lower triangular Toeplitz matrix of order r over the field $\mathbb{F}_{2^{r-2}}$.

We compare these results in Table 8.2 together with some examples in [15] and [110] found by optimized computer search.

Remark 8.31. Here we compare asymptotically the field size needed for our WRS convolutional codes to be MDP with the other two existing general constructions of MDP convolutional provided in [12, 79]. We consider the case in which we fix the rate of the code to be constant $R := \frac{k}{n}$, and express all the field sizes in terms of R, δ and n. First, notice that the parameter r defined above can be approximated by $R^{-1}\delta + n - 1$. Assuming now that the value $R^{-1}\delta + n - 1$ grows, we study the asymptotic behaviours of the field sizes.

By using Stirling approximation formula, we have that

$$c := \binom{r-1}{\lfloor \frac{r-1}{2} \rfloor} \sim \frac{2^{R^{-1}\delta+n}}{\sqrt{8\pi(R^{-1}\delta+n-2)}},$$

leading to the following asymptotic approximation for the field size needed in [79]:

$$|\mathbb{F}| \sim c^r r^{r/2} \sim e \cdot \frac{2^{(R^{-1}\delta+n-1)(R^{-1}\delta+n-\frac{3}{2})}}{\sqrt{\pi}^{(R^{-1}\delta+n-1)}} \sim 2^{(R^{-1}\delta+n-1)(R^{-1}\delta+n-\frac{3}{2}-\log_2(\pi))}. \tag{8.11}$$

Moreover, the field size needed for the construction in [12] is

$$|\mathbb{F}| \sim 2^{2^{R^{-1}\delta+n+1}},$$

which is always asymptotically worse than (8.11). Finally, by Corollary 8.27, we have that for our constructions we need a field size of approximately

$$|\mathbb{F}| \sim 2^{\log_2(n)\frac{11}{24}(\frac{\delta^3}{Rn}+\delta Rn)}. \tag{8.12}$$

In order to compare (8.12) with (8.11), we compare the asymptotics of their logarithms. If δ is constant, then our construction is better, while if n is constant the one of [79] is better. Suppose now that none of δ and n is constant. We have that whenever $\delta = \Theta(n^{1-\epsilon}(\log n)^{\beta})$ for any $0 < \epsilon < 1$ and $\beta \in \mathbb{R}$, our construction beats the field size of [79], while when $\delta = \Theta(n^{1+\epsilon}(\log n)^{\beta})$ for any $\epsilon > 0$ and $\beta \in \mathbb{R}$, the field size of [79] is asymptotically better than ours. Furthermore, when $\delta = \Theta(n(\log n)^{-1-\beta})$ for any $\beta > 0$ our field size is better, while in the case $\delta = \Theta(n(\log n)^{-1+\beta})$ for any $\beta > 0$, the one in [79] is smaller than ours. Finally, in the case that $\delta = \Theta(n(\log n)^{-1})$, the logarithms of the field sizes are asymptotically the same, so one should carefully investigate the smallest order terms.

$[n, k, \delta]$ L, m, μ, r	[12]	[79]	[110]*	[15]*	[89]†	[102]†	$\mathcal{C}^{\delta}_{k,n}$	[79]‡
$[2, 1, 1]$ $2, 1, 1, 3$	2^8	43	2^5	3	3	55	3	–
$[2, 1, 2]$ $4, 2, 2, 5$	2^{32}	434692	2^7	7	11	1261	27	2^3
$[3, 2, 2]$ $3, 1, 2, 8$	2^{512}	$5^8 7^8 2^{12} + 1$	2^{11}	31	233	1981	256	2^6
$[3, 1, 2]$ $3, 2, 1, 8$	2^{512}	$5^8 7^8 2^{12} + 1$	2^{11}	31	233	3961	16	2^6
$[3, 2, 1]$ $1, 1, 1, 4$	2^{32}	$2^4 3^4 + 1$	–	5	5	3	4	2^2
$[4, 2, 2]$ $2, 1, 1, 7$	2^{128}	$\sim 10^{12}$	–	17	77	5545	125	2^5
$[4, 1, 3]$ $4, 3, 1, 15$	$2^{2^{17}}$	$\sim 7 \cdot 10^{61}$	–	–	1338936	232561	3125	2^{13}
$[5, 2, 2]$ $1, 1, 1, 7$	2^{2^9}	$\sim 10^{12}$	–	17	77	35	49	32
$[6, 2, 2]$ $1, 1, 1, 9$	$2^{2^{11}}$	$\sim 7 \cdot 10^{20}$	–	59	751	71	49	128
$[6, 2, 2]$ $1, 2, 1, 9$	$2^{2^{11}}$	$\sim 7 \cdot 10^{20}$	–	59	751	71	49	128
$[7, 2, 2]$ $1, 1, 1, 11$	$2^{2^{13}}$	$\sim 10^{32}$	–	–	8525	126	64	512
$[7, 3, 3]$ $1, 1, 1, 10$	$2^{2^{12}}$	$\sim 10^{26}$	–	127	2495	532	512	256

Table 8.2: Parameters and smallest field sizes of MDP convolutional codes, according to known results in the literature. The columns are marked with * if the result is found by computer search; the results marked with † indicates that they are not constructive; the symbol ‡ means that the correspondent result is based on a conjecture. The symbol – indicates that there are no constructions for such parameters. For the nonconstructive results marked with † we included the smallest field size needed, even if it is not a prime power. The cells with the colored background indicate the best field size for the given parameters

Chapter 9

Construction of LDPC Convolutional Codes via Difference Triangle Sets

The result of this chapter have been published by Alfarano, Lieb and Rosenthal in [10].

9.1 Difference Triangle Sets

A difference triangle set is a collection of sets of integers such that any integer can be written in at most one way as difference of two elements in the same set. Difference triangle sets find application in combinatorics, radio systems, optical orthogonal codes and other areas of mathematics [94, 50, 52]. We refer to [55] for a more detailed treatment. More formally, we define them in the following way, by distinguishing between weak difference triangle sets and difference triangle sets.

Definition 9.1. Let N, M be positive integers. An (N, M)-*weak difference triangle set* (wDTS) is a collection of sets $\mathcal{T} := \{T_1, T_2, \dots, T_N\}$, where for any $1 \leq i \leq N$, $T_i := \{a_{i,j} \mid 1 \leq j \leq M\}$ is a set of nonnegative integers such that $a_{i,1} < a_{i,2} < \cdots < a_{i,M}$ and for $1 \leq i \leq N$ the differences $a_{i,j} - a_{i,k}$, with $1 \leq k < j \leq M$ are distinct. If all the differences in all the sets are distinct, we call $\mathcal{T}$ a (N, M)-*difference triangle set* (DTS).

An important parameter characterizing an (N, M)-(w)DTS $\mathcal{T}$ is the *scope* $m(\mathcal{T})$, which is defined as

$$m(\mathcal{T}) := \max\{a_{i,M} \mid 1 \leq i \leq N\}.$$

A very well-studied problem in combinatorics is finding families of (N, M)-DTSs with minimum scope. In this work, we will use the sets in a (w)DTS as supports of some columns of a sliding parity-check matrix of a convolutional code. We will then relate the scope of the (w)DTS with the degree of the code. Since we want to minimize the degree of the code, it is evident that the mentioned combinatorial problem plays a crucial role also here.

The name “difference triangle" is derived from a way of writing the differences inside the sets composing $\mathcal{T}$ in a triangular form .

Example 9.2 (wDTS)**.** Let $\mathcal{T} = \{\{1,2,4,8\},\{1,3,7,15\},\{1,5,10,16\}\}$. Then $\mathcal{T}$ is a $(3,4)$-wDTS.

The “triangles" associated to $\mathcal{T}$ are the following:

$$\begin{array}{ccccc} 1 & & 2 & & 4 \\ & 3 & & 6 & \\ & & 7 & & \end{array} \qquad \begin{array}{ccccc} 2 & & 4 & & 8 \\ & 6 & & 12 & \\ & & 14 & & \end{array} \qquad \begin{array}{ccccc} 4 & & 5 & & 6 \\ & 9 & & 11 & \\ & & 15 & & \end{array}$$

Example 9.3 (DTS)**.** Let $\mathcal{T} = \{\{1,4,16,20\},\{1,7,12,14\},\{1,9,18,19\}\}$. Then $\mathcal{T}$ is a $(3,4)$-DTS.

The “triangles" associated to $\mathcal{T}$ are the following:

$$\begin{array}{ccccc} 3 & & 12 & & 4 \\ & 15 & & 16 & \\ & & 19 & & \end{array} \qquad \begin{array}{ccccc} 6 & & 5 & & 2 \\ & 11 & & 7 & \\ & & 13 & & \end{array} \qquad \begin{array}{ccccc} 8 & & 9 & & 1 \\ & 17 & & 10 & \\ & & 18 & & \end{array}$$

9.2 LDPC Codes over Arbitrary Finite Fields

LDPC codes are known for their performance near the Shannon-limit over the additive white Gaussian noise channel [108]. Shortly after they were rediscovered, binary LDPC codes were generalized over arbitrary finite fields. This new construction was first investigated by Davey and Mackay in 1998 in [57]. In [58], it was observed that LDPC codes defined over a finite field with q elements can have better performances than the binary ones. An LDPC code is defined as the kernel of an $N \times M$ sparse matrix H with entries in $\mathbb{F}_q$. We can associate to H a bipartite graph $\mathcal{G} = (V, E)$, called *Tanner graph*, where $V = V_s \cup V_c$ is the set of vertices. In particular, $V_s = \{v_1, \dots, v_N\}$ is the set of *variable nodes* and $V_c = \{c_1, \dots, c_M\}$ is the set of *check nodes*. $E \subseteq V_s \times V_c$ is the set of edges, with $e_{n,m} = (v_n, c_m) \in E$ if and only if $h_{n,m} \neq 0$. The edge $e_{n,m}$ connecting a check node and a variable node is labelled by $h_{n,m}$, that is the corresponding *permutation node*. For an even integer $m = 2\ell$, we call a simple closed path consisting of ℓ check nodes and ℓ variable nodes in $\mathcal{G}$ an *m-cycle*. The length of the shortest cycle is called the *girth* of $\mathcal{G}$ or girth of H. It is proved that higher the girth is, the lower the decoding failure of the bit flipping algorithm is. Moreover, in [127] the authors showed that short cycles in an LDPC code may be harmful if they do not satisfy the so called full rank condition (FRC). This is because if the FRC is not satisfied, the short cycles produce low-weight codewords or they form absorbing sets, [18].

Moreover, in [127] and in [18] it is shown that an m-cycle, with $m = 2\ell$ in an LDPC code with parity-check matrix H can be represented, up to permutations, by an $\ell \times \ell$ submatrix of H

of the form

$$A = \begin{bmatrix} a_1 & a_2 & 0 & \cdots & \cdots & 0 \\ 0 & a_3 & a_4 & \cdots & \cdots & \vdots \\ \vdots & & \ddots & & & \vdots \\ \vdots & & & \ddots & & \vdots \\ 0 & & & & a_{2\ell-3} & a_{2\ell-2} \\ a_{2\ell} & 0 & \cdots & \cdots & 0 & a_{2\ell-1} \end{bmatrix}, \tag{9.1}$$

where $a_i \in \mathbb{F}_q^*$. The cycle does not satisfy the FRC if the determinant of A is equal to 0. In this case, the cycle gives an absorbing set. Hence, it is a common problem to construct LDPC codes in which the shortest cycles satisfy the FRC.

In this work, we are interested in the convolutional counterpart of LDPC block codes, which is given by convolutional codes defined over a finite field $\mathbb{F}_q$ as kernel of a sparse sliding parity-check matrix (here with sparse we mean that in particular each H_i is sparse).

9.3 Construction of LDPC Convolutional Codes

In this section, we use difference triangle sets to construct LDPC convolutional codes over $\mathbb{F}_q$. The construction was provided for $(n, n-1)_q$ convolutional codes in [9]. Here, we generalize it for arbitrary n and k.

We will construct a sliding parity-check matrix H as in (6.3), whose kernel defines a convolutional code. Due to the block structure of H, it is enough to consider

$$\mathcal{H} := H_\nu^c = \begin{bmatrix} H_0 & & & \\ H_1 & H_0 & & \\ \vdots & \vdots & \ddots & \\ H_\nu & H_{\nu-1} & \cdots & H_0 \end{bmatrix}, \tag{9.2}$$

since H is then constructed by sliding it. It is easy to see that H does contain a cycle of length 2ℓ not satisfying the FRC if and only if $\mathcal{H}$ does. Assuming that H_0 is full rank, we can perform Gaussian elimination on the matrix

$$\begin{bmatrix} H_0 \\ H_1 \\ \vdots \\ H_\nu \end{bmatrix},$$

which results in the block matrix

$$\bar{H} = \begin{bmatrix} A_0 & | & I_{n-k} \\ A_1 & | & 0 \\ \vdots & & \vdots \\ A_\nu & | & 0 \end{bmatrix}, \tag{9.3}$$

with $A_i \in \mathbb{F}_q^{(n-k)\times k}$ for $i = 1, \dots, \nu$. With abuse of notation, we write H_0 for $[A_0|I_{n-k}]$, and H_i for the matrices $[A_i|0]$.

Remark 9.4. If we define the matrix $\tilde{H}(z) = \sum_{i=0}^{\nu} A_i z^i \in \mathbb{F}_q[z]^{(n-k)\times k}$, then we obtain that $H(z) = [\tilde{H}(z)\ I_{n-k}]$ and hence $H(z)$ has a polynomial right inverse, i.e. $H(z)$ is basic.

Given $n \in \mathbb{N}$, with the following definition we describe how we construct the above mentioned matrix $\bar{H}$ from a (k,w)-wDTS, which then will define an $(n,k)_q$ convolutional code.

Definition 9.5. Let k, n be positive integers with $n > k$ and $\mathcal{T} := \{T_1, \dots, T_k\}$ be a (k,w)-wDTS with scope $m(\mathcal{T})$. Set $\nu = \left\lceil \frac{m(\mathcal{T})}{n-k} \right\rceil - 1$ and define the matrix $\bar{H} \in \mathbb{F}_q^{(\nu+1)(n-k)\times n}$, in which the l-th column has weight w and support T_l, i.e. for any $1 \le i \le (\nu+1)(n-k)$ and $1 \le l \le k$, $\bar{H}_{i,l} \neq 0$ if and only if $i \in T_l$. We say that $\bar{H}$ has support $\mathcal{T}$. The last $n-k$ columns of $\bar{H}$ are given by $[I_{n-k}, 0_{n-k}, \dots, 0_{n-k}]^\top$. Derive the matrix $\mathcal{H}$ by "shifting" the columns of $\bar{H}$ by multiples of $n-k$ and then a sliding matrix H of the form of equation (6.3). Finally, define $\mathcal{C} := \ker(\mathcal{H})$ over $\mathbb{F}_q$.

Observe that if $k = n-1$, we simply get the construction provided in [9, Definition 4].

Proposition 9.6. Let n, k, w be positive integers with $n > k$, $\mathcal{T}$ be a (k,w)-wDTS with scope $m(\mathcal{T})$ and set $\nu = \left\lceil \frac{m(\mathcal{T})}{n-k} \right\rceil - 1$. If $\bar{H}$ has support $\mathcal{T}$, then the corresponding code is an (n,k,δ) convolutional code with $\nu \le \delta \le \nu(n-k)$. Moreover H_ν is full rank if and only if $\delta = \nu(n-k)$.

Proof. As the matrix $H(z)$ defined in Remark 9.4 is basic, δ is the maximal degree of the full-size minors of H, which is clearly upper bounded by $\nu(n-k)$. Moreover, any minor formed by a column with degree ν and suitable columns of the systematic part of H has degree ν, which proves the lower bound.

If H_ν is full rank, it is equal to $[H]_{hr}$, and H is reduced. Hence, δ is equal to the sum of the $n-k$ row degrees that are all equal to ν, i.e. $\delta = \nu(n-k)$. If H_ν is not full rank, there are two possible cases. First, if H_ν contains no all-zero row, then $[H]_{hr} = H_\nu$ is not full rank, and hence δ is strictly smaller than the sum of the row degrees which is $\nu(n-k)$. Second, if H_ν contains a row of zeros, then the sum of the row degrees of H is strictly smaller than $\nu(n-k)$ and thus, also δ is strictly smaller than $\nu(n-k)$. □

Remark 9.7. If $k < n-k$, i.e. the rate of the code is smaller than $1/2$, then (9.3) implies that H_ν cannot be full rank. Moreover, in this case, $[H]_{hr}$ can only be full rank if at least $n-2k$ row degrees of H are zero.

Proposition 9.8. Let n, k, w be positive integers with $n > k$ and $\mathcal{T}$ be a (k, w)-wDTS. Assume $\bar{H}$ has support $\mathcal{T}$ and consider the convolutional code $\mathcal{C}$ constructed as kernel of the sliding parity-check matrix corresponding to $\bar{H}$. If N is the maximal codeword length, i.e. for any codeword $v(z) \in \mathcal{C}$, $\deg(v) + 1 \leq N/n$, then the sliding parity-check matrix corresponding to $\bar{H}$ has density

$$\frac{wk + n - k}{(n-k)(\nu n + N)}.$$

Proof. To compute the density of a matrix, one has to divide the number of nonzero entries by the total number of entries. The result follows immediately. □

Theorem 9.9. Let $\mathcal{C}$ be an (n, k) convolutional code with parity-check matrix H. Assume that all the columns of $\begin{bmatrix} A_0^\top & \cdots & A_\nu^\top \end{bmatrix}^\top$ defined as in (9.3) have weight w and denote by w_j the minimal column weight of $\begin{bmatrix} A_0^\top & \cdots & A_j^\top \end{bmatrix}^\top$. For $I \subset \{1, \ldots, (n-k)(\nu+1)\}$ and $J \subset \{1, \ldots, n(\nu+1)\}$ we define $[\mathcal{H}]_{I;J}$ as the submatrix of $\mathcal{H}$ with row indices I and column indices J. Assume that for some $\tilde{w} \leq w$ all I, J with $|J| \leq |I| \leq \tilde{w}$ and $j_1 := \min(J) \leq k$ and I containing the indices where column j_1 is nonzero, we have that the first column of $[\mathcal{H}]_{I;J}$ is not contained in the span of the other columns of $[\mathcal{H}]_{I;J}$. Then

(i) $\tilde{w} + 1 \leq \mathrm{d}_{\text{free}}(\mathcal{C}) \leq w + 1$,

(ii) $\min(w_j, \tilde{w}) + 1 \leq d_j^c(\mathcal{C}) \leq w_j + 1$.

Proof. (i) Without loss of generality, we can assume that the first entry in the first row of H_0 is nonzero. Denote the first column of $\mathcal{H}$ by $[h_{1,1}, \ldots, h_{1,(n-k)\nu}]^\top$. Then, $v(z) = \sum_{i=0}^{r} v_i z^i$ with

$$\begin{aligned} v_0 &= [1\ 0 \cdots 0\ \ -h_{1,1} \cdots\ -h_{1,(n-k)}] \quad \text{and} \\ v_i &= [0\ 0 \cdots 0\ \ -h_{1,(n-k)i+1} \cdots\ -h_{1,(n-k)(i-1)}], \end{aligned}$$

for $i \geq 1$ is a codeword with $\mathrm{wt}(v(z)) = w + 1$ as the weight of the first column of $\mathcal{H}$ is equal to w. Hence $\mathrm{d}_{\text{free}} \leq w + 1$.

Assume by contradiction that there exists a codeword $v(z) \neq 0$ with weight $d \leq \tilde{w}$. We can assume that $v_0 \neq 0$, i.e. there exists $i \in \{1, \ldots, n\}$ with $v_{0,i} \neq 0$. We know that $\mathcal{H}v^\top = 0$ and from (9.3) we obtain that there exists $j \in \{1, \ldots, n\}$ with $j \neq i$ and $v_{0,j} \neq 0$ and we can assume that $i \leq k$.

Now, we consider the homogeneous system of linear equations given by $\mathcal{H}v^\top = 0$ and we only take the rows, i.e. equations, where column i of $\mathcal{H}$ has nonzero entries. Moreover, we define $\tilde{v} \in \mathbb{F}^d$ as the vector consisting of the nonzero components of $v_0, v_1, \ldots, v_{\deg(v)}$. We end up with a system of equations of the form $[\mathcal{H}]_{I;J}\tilde{v}^\top = 0$ where $[\mathcal{H}]_{I;J}$ fulfills the assumptions stated in the theorem. But this is a contradiction as $\tilde{v}^\top$ has all components nonzero and therefore $[\mathcal{H}]_{I;J}\tilde{v}^\top = 0$ implies that the first column of $[\mathcal{H}]_{I;J}$ is contained in the span of the other columns

of this matrix.
(ii) The result follows from Theorem 6.17 with an analogue reasoning as in part (i). □

Remark 9.10. With the assumptions of Theorem 9.9, if $\tilde{w} = w$, one has $d_j^c = \mathrm{d}_{\text{free}}$ for $j \geq \nu$. Moreover, if $\bar{H}$ has support $\mathcal{T}$, one achieves higher column distances (especially for small j) if the elements of $\mathcal{T}$ are small.

Corollary 9.11. If $\mathcal{T}$ is a (k, w)-DTS and $\mathcal{C}$ is an (n, k) convolutional code constructed from $\mathcal{T}$ as in Definition 9.5, then one has that:

(i) $\mathrm{d}_{\text{free}}(\mathcal{C}) = w + 1$,

(ii) $d_j^c(\mathcal{C}) = w_j + 1$.

Proof. As already mentioned in [132], matrices $\mathcal{H}$ constructed from a DTS have the property that for every pair of columns, their supports intersect at most once. Since $[\mathcal{H}]_{I;J}$ as defined in Theorem 9.9 has the property that all entries in the first column are non-zero, all other columns have at most one non-zero entry. But this implies that the first column cannot be in the span of the other columns and thus, the requirements of Theorem 9.9 are fulfilled for $\tilde{w} = w$, which proves the corollary. □

Remark 9.12. If $n - k > 1$, it is not necessary to have a DTS to obtain that all columns of $\mathcal{H}$ intersect at most once since one only has to consider shifts of columns by multiples of $n - k$. Therefore, we still need to consider a set $\mathcal{T} = \{T_1, \dots, T_k\}$ such that all the differences $a_{i_1,j_1} - a_{i_1,s_1}$ and $a_{i_2,j_2} - a_{i_2,s_2}$ for $i_1 \neq i_2$ are different, i.e. two differences coming from different triangles of $\mathcal{T}$ have always to be different, but $a_{i,j_1} - a_{i,s_1}$ and $a_{i,j_2} - a_{i,s_2}$, i.e. differences coming from the same triangle, only have to be different if $(n - k) \mid (a_{i,j_1} - a_{i,j_2})$.

Example 9.13. Consider $n = 3$, $k = 1$ and $T_1 = \{1, 2, 3\}$. It holds $2 - 1 = 3 - 2$ but since $3 - 2$ is not divisible by $n - k = 2$, this does not matter and we still get that all columns of $\mathcal{H}$ intersect at most once. For example for $\nu = 1$, we get

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

From Corollary 9.11 we know that if we use a DTS to construct the parity-check matrix of the code, then the values of the nonzero entries are not important to achieve good distance properties. In the following, we present a construction that achieves also quite large distances if one takes the sets in a wDTS as support sets for the columns of the non-systematic part of $\bar{H}$. Moreover, in Section 9.4, we show that this construction ensures that the Tanner graph associated to H is free from cycles of arbitrary length not satisfying the FRC if the size of the underlying field is sufficiently large and the wDTS fulfills some additional properties.

Definition 9.14. Let k, n be positive integers with $n > k$ and $\mathcal{T} := \{T_1, \dots, T_k\}$ be a (k, w)-wDTS with scope $m(\mathcal{T})$. Set $\nu = \left\lceil \frac{m(\mathcal{T})}{n-k} \right\rceil - 1$ and let α be a primitive element for $\mathbb{F}_q$, so that every non-zero element of $\mathbb{F}_q$ can be written as power of α. For any $1 \leq i \leq (\nu + 1)(n - k)$, $1 \leq l \leq k$, define

$$\bar{H}^{\mathcal{T}}_{i,l} := \begin{cases} \alpha^{il} & \text{if } i \in T_l \\ 0 & \text{otherwise} \end{cases}.$$

Obtain the matrix $\mathcal{H}^{\mathcal{T}}$ by “shifting" the columns of $\bar{H}^{\mathcal{T}}$ by multiples of $n - k$ and then a sliding matrix $H^{\mathcal{T}}$ of the form of equation (6.3). Finally, define $\mathcal{C}^{\mathcal{T}} := \ker(\mathcal{H}^{\mathcal{T}})$ over $\mathbb{F}_q$.

Example 9.15. Let $\mathbb{F}_q := \{0, 1, \alpha, \dots, \alpha^{q-2}\}$ and $\mathcal{T}$ be a $(2, 3)$-wDTS, such that $T_1 := \{1, 2, 6\}$ and $T_2 := \{1, 2, 4\}$. Then, with the notation above,

$$\bar{H}^{\mathcal{T}} = \begin{bmatrix} \alpha & \alpha^2 & 1 \\ \alpha^2 & \alpha^4 & 0 \\ 0 & 0 & 0 \\ 0 & \alpha^8 & 0 \\ 0 & 0 & 0 \\ \alpha^6 & 0 & 0 \end{bmatrix},$$

which leads to the following sliding matrix.

$$\mathcal{H}^{\mathcal{T}} = \begin{bmatrix} \alpha & \alpha^2 & 1 & & & & & & & & & & & & & & & \\ \alpha^2 & \alpha^4 & 0 & \alpha & \alpha^2 & 1 & & & & & & & & & & & & \\ 0 & 0 & 0 & \alpha^2 & \alpha^4 & 0 & \alpha & \alpha^2 & 1 & & & & & & & & & \\ 0 & \alpha^8 & 0 & 0 & 0 & 0 & \alpha^2 & \alpha^4 & 0 & \alpha & \alpha^2 & 1 & & & & & & \\ 0 & 0 & 0 & 0 & \alpha^8 & 0 & 0 & 0 & 0 & \alpha^2 & \alpha^4 & 0 & \alpha & \alpha^2 & 1 & & & \\ \alpha^6 & 0 & 0 & 0 & 0 & 0 & 0 & \alpha^8 & 0 & 0 & 0 & 0 & \alpha^2 & \alpha^4 & 0 & \alpha & \alpha^2 & 1 \end{bmatrix}.$$

The code constructed here is a $(3, 2)_q$ convolutional code. In this example, one has $d^c_0 = 2$, $d^c_1 = d^c_2 = d^c_3 = d^c_4 = 3$ and $d_5 = \mathrm{d}_{\text{free}} = 4$.

The next theorem is a generalization of [9, Theorem 12] to any rate.

Theorem 9.16. Let w, n, k be positive integers with $n > k$ and $\mathcal{T}$ be a (k, w)-wDTS with scope $m(\mathcal{T})$ and $q > (\nu + 1)(n - k)(k - 1) + 1 = \lceil \frac{m(\mathcal{T})}{n-k} \rceil (n - k)(k - 1) + 1$. Let $\mathcal{C}^{\mathcal{T}}$ be the $(n, k)_q$ convolutional code defined from $\mathcal{T}$, as defined in Definition 9.14 and consider $\mathcal{H}^{\mathcal{T}}$ as in (9.2). Then, all the 2×2 minors in $\mathcal{H}^{\mathcal{T}}$ that are non-trivially zero are non-zero.

Proof. The only 2×2 minors to check are the ones of the form $\begin{vmatrix} a_1 & a_2 \\ a_3 & a_4 \end{vmatrix}$. By definition of wDTS, the support of any column of $\mathcal{H}^{\mathcal{T}}$ intersects the support of its shift at most once. This ensures that the columns of all these minors are the shift of two different columns of $\bar{H}^{\mathcal{T}}$. Moreover, all the elements in the minor are powers of α. In particular, let $1 \leq i, r \leq (\nu+1)(n-k)$, $1 \leq j, \ell \leq k$ (note that $j < \ell$ or $\ell < j$ according to which columns from $\bar{H}^{\mathcal{T}}$ are involved in the shifts). Hence we have that:

$$\begin{vmatrix} a_1 & a_2 \\ a_3 & a_4 \end{vmatrix} = \begin{vmatrix} \alpha^{ij} & \alpha^{m\ell} \\ \alpha^{(i+r)j} & \alpha^{(m+r)\ell} \end{vmatrix} = \\ \alpha^{ij}\alpha^{(m+r)\ell} - \alpha^{m\ell}\alpha^{(i+r)j} = \alpha^{ij+m\ell}(\alpha^{r\ell} - \alpha^{rj})$$

which is 0 if and only if $r\ell = rj \mod (q-1)$. Since it holds that $0 \leq j < \ell \leq k$ or $0 \leq \ell < j \leq k$ and $1 \leq r \leq (\nu+1)(n-k)$, this cannot happen. □

The following theorem is a generalization of [9, Theorem 13] for any rate. However, in the proof in [9] there is a computation mistake, hence we put the correct version below.

Theorem 9.17. Let w, n, k be positive integers with $n > k$ and $\mathcal{T}$ be a (k, w)-wDTS with scope $m(\mathcal{T})$, $w \geq 3$. Let $\mathcal{C}^{\mathcal{T}}$ be the $(n, k)_q$ convolutional code defined from $\mathcal{T}$, as in Definition 9.14 with $\mathcal{H}^{\mathcal{T}}$ as defined in (9.2) and assume that $(\nu+1)(n-k) > 2$. Assume also that $q = p^N$, where $p > 2$ and

$$N > (\nu+1)(n-k)(k-1) = \left\lceil \frac{m(\mathcal{T})}{n-k} \right\rceil (n-k)(k-1).$$

Then, all the 3×3 minors in $\mathcal{H}^{\mathcal{T}}$ that are non-trivially zero are non-zero.

Proof. We need to distinguish different cases.
Case I. The 3×3 minors are of the form

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ a_7 & a_8 & a_9 \end{vmatrix},$$

with $a_i \neq 0$ for any i. As we observed in Theorem 9.16, in this case all the columns are shifts of three different columns from $\bar{H}^{\mathcal{T}}$, since each column can intersect any of its shifts at most once. Observe that we can write a minor of this form as

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ a_7 & a_8 & a_9 \end{vmatrix} = \begin{vmatrix} \alpha^{ij} & \alpha^{lu} & \alpha^{tm} \\ \alpha^{(i+r)j} & \alpha^{(l+r)u} & \alpha^{(t+r)m} \\ \alpha^{(i+r+s)j} & \alpha^{(l+r+s)u} & \alpha^{(t+r+s)m} \end{vmatrix},$$

where $1 \leq i, l, t \leq (\nu+1)(n-k)$, $r, s \in \mathbb{Z}$ are possibly negative, with $r \neq s$, and $1 \leq j, u, m \leq k$

representing the index of the column from which the selected element comes from (or if the selected elements belongs to the shift of some column, j, u, m are still the indexes of the original column). Due to symmetry in this case we can assume $r, s \in \mathbb{N}$ and $1 \leq i, l, t \leq (\nu+1)(n-k)-3$. Moreover, $-(\nu+1)(n-k)+1 \leq i+r, l+r, t+r \leq (\nu+1)(n-k)-1$ and $-(\nu+1)(n-k) \leq i+r+s, l+r+s, t+r+s \leq (\nu+1)(n-k)$. This determinant is 0 if and only if

$$\begin{aligned} &\alpha^{ru+rm+sm} + \alpha^{rm+rj+sj} + \alpha^{rj+ru+sk} = \\ &\alpha^{ru+rj+sj} + \alpha^{rj+rm+sm} + \alpha^{ru+rm+sk}. \end{aligned} \tag{9.4}$$

Without loss of generality we can assume that $j < u < m$ and it turns out that the maximum exponent in equation (9.4) is $ru + rm + sm$ while the minimum is $ru + rj + sj$. Let $M := ru + rm + sm - (ru + rj + sj)$. It is not difficult to see that the maximum value for M is $((\nu+1)(n-k)-1)(k-1)$ hence this determinant can not be zero because α is a primitive element for $\mathbb{F}_q$ and, by assumption, $q = p^N$, where $N > M$.

<u>**Case II**</u>. The 3×3 minors are of the form

$$\begin{vmatrix} a_1 & a_2 & 0 \\ a_3 & a_4 & a_5 \\ a_6 & 0 & a_7 \end{vmatrix}.$$

As in the first case, we can assume that the minor is given by

$$\begin{vmatrix} \alpha^{ij} & \alpha^{lu} & 0 \\ \alpha^{(i+r)j} & \alpha^{(l+r)u} & \alpha^{(t+r)m} \\ \alpha^{(i+r+s)j} & 0 & \alpha^{(t+r+s)m} \end{vmatrix},$$

with the same bounds on the variables as before. But, in this case $j \neq u, m$ but u can be equal to m. Indeed, the first column intersects the other two in two places, which means that they are not all shifts of the same column. However, the second and third ones can belong to the same column. This determinant is 0 when $\alpha^{ru+sm} + \alpha^{rj+sj} - \alpha^{rm+sm} = 0$. In this case, according to the different possibilities for j, u, m and r, s we check the maximum and the minimum exponent. We present here only the worst case for the field size, which is obtained when $j < u < m$, $r < 0$. We see that the minimum exponent is $rj + sj$ and the maximum is $rj + sm$. We consider $M := rj + sm - rj - sj$ and we check what is the maximum value that M can reach. It is not difficult to see that this is $(\nu+1)(n-k)(k-1)$. When $p = p^N$, with $N > M$, the considered determinant is never 0.

Case III. The 3×3 minors are of the form

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ a_7 & a_8 & 0 \end{vmatrix},$$

with $a_i \neq 0$ for any i. We can assume that, the minor is given by

$$\begin{vmatrix} \alpha^{ij} & \alpha^{lu} & \alpha^{tm} \\ \alpha^{(i+r)j} & \alpha^{(l+r)u} & \alpha^{(t+r)m} \\ \alpha^{(i+r+s)j} & \alpha^{(l+r+s)u} & 0 \end{vmatrix},$$

with the same bounds on the variables as in previous cases. However, this time $1 \leq j < u < m \leq k$. After some straightforward computations, we get that this determinant is 0 if and only if

$$\alpha^{rm+rj+sj} + \alpha^{rj+ru+su} = \alpha^{ru+rj+sj} + \alpha^{ru+rm+su}. \tag{9.5}$$

In the worst case, consider $M := ru + rj + su - (rm + rj + sj) = r(u - m) + s(u - j)$ with $r < 0$. We immediately see that the maximum value that M can reach is $(\nu+1)(n-k)(k-2)+1$, hence this determinant can not be zero because α is a primitive element for $\mathbb{F}_q$ and, by assumption, $q = p^N$, where $N > M$.

Case IV. The 3×3 minors are of the form

$$\begin{vmatrix} a_1 & a_2 & 0 \\ 0 & a_3 & a_4 \\ a_6 & 0 & a_5 \end{vmatrix}.$$

In this case, we can have that the three considered columns come from different shifts of the same one, hence we allow that some (or all) among j, u, m are equal. Arguing as before, we notice that these minors are given by

$$\begin{vmatrix} \alpha^{ij} & \alpha^{lu} & 0 \\ 0 & \alpha^{(l+r)u} & \alpha^{(t+r)m} \\ \alpha^{(i+r+s)j} & 0 & \alpha^{(t+r+s)m} \end{vmatrix} = \alpha^{ij+lu+tm+rm}(\alpha^{ru+sm} + \alpha^{rj+sj}).$$

This determinant is 0 whenever $r(u-j)+s(m-j)-(q-1)/2 = 0 \mod (q-1)$. Analyzing all the possibilities we can have according to r, s being negative or positive and j, u, m being equal or different, after some computations, we obtain that, whenever $q > 2(k-1)((\nu+1)(n-k)-1)+1$, the considered determinant is never 0. And this is the case for our field size assumption. □

Observe that Case IV of Theorem 9.17 corresponds to the lower bound for the field size sufficient to avoid the presence of 6-cycles not satisfying the FRC. Hence, we have the following result.

Corollary 9.18. Let $\mathcal{C}^{\mathcal{T}}$ be an (n, k) convolutional code constructed from a (k, w) wDTS $\mathcal{T}$ and satisfying the conditions of Theorem 9.16 and Theorem 9.17. Then, $d_{free}(\mathcal{C}^{\mathcal{T}}) \geq 3$ and the code is free from 4 and 6-cycles not satisfying the FRC.

Remark 9.19. If $\mathcal{C}^{\mathcal{T}}$ is an (n, k) convolutional code constructed from a (k, w) wDTS $\mathcal{T}$ and satisfying the conditions of Theorem 9.16 and Theorem 9.17, such that H_ν has no zero row and $n - k \leq \min\{3, k\}$, then, it follows from Proposition 9.6 that $\delta = \nu(n - k)$.

Example 9.20. Consider the $(3, 2)_q$ code constructed in Example 9.15. Note that $\nu = 5$, hence, for $q > 11$ we can avoid all the 6-cycles not satisfying the FRC (Case IV of Theorem 9.17).

9.4 Excluding Cycles not Satisfying the FRC

In this section, we give some conditions that ensure that the Tanner graph associated to the sliding parity-check matrix of a convolutional code constructed via a difference triangle set is free of 2ℓ-cycles not satisfying the FRC.

First of all we recall from Subsection 9.2 that a 2ℓ-cycle can be represented by an $\ell \times \ell$ submatrix of $\mathcal{H}$ that up to column and row permutations is of the form

$$A = \begin{bmatrix} a_1 & a_2 & 0 & \cdots & \cdots & 0 \\ 0 & a_3 & a_4 & \cdots & \cdots & \vdots \\ \vdots & & \ddots & & & \vdots \\ \vdots & & & \ddots & & \vdots \\ 0 & & & & a_{2\ell-3} & a_{2\ell-2} \\ a_{2\ell} & 0 & \cdots & \cdots & 0 & a_{2\ell-1} \end{bmatrix}, \tag{9.6}$$

where $a_i \in \mathbb{F}_q^*$.

Remark 9.21. Observe that

$$\begin{bmatrix} A_0 & & \\ \vdots & \ddots & \\ A_\nu & \cdots & A_0 \end{bmatrix} \in \mathbb{F}^{(\nu+1)(n-k)\times(\nu+1)k},$$

hence it is clear that the Tanner graph associated to H can only contain 2ℓ-cycles for

$$\ell \leq \min\{(\nu + 1)(n - k), (\nu + 1)k\}.$$

At first, we will investigate conditions on the wDTS used to construct the convolutional code that ensure that the associated Tanner graph contains no cycles at all independently of the nonzero values of the sliding parity-check matrix and hence also independently of the underlying finite field.

Proposition 9.22. If $\mathcal{C}$ is an (n,k) convolutional code whose parity-check matrix has support $\mathcal{T}$ where $\mathcal{T}$ is a (k,w)-wDTS with the property that none of the differences $a_{i,j}-a_{i,m}$ for $1\leq i\leq k$ and $1\leq m<j\leq w$ is divisible by $n-k$, then each pair of columns that is next to each other in A as in (9.6) consists of shifts of different columns of $\bar{H}$. In particular, at most $\lfloor\frac{\ell}{2}\rfloor$ columns of A can be shifts of the same column of $\bar{H}$.

Proof. The fact that none of the differences in the set is divisible by $n-k$ implies that the support of any column of $\bar{H}$ does not intersect the support of any of its shifts (by multiples of $n-k$). Since the supports of neighbouring columns of A intersect, they have to be shifts of different columns of $\bar{H}$. □

Corollary 9.23. If $\mathcal{C}$ is an (n,k) convolutional code whose parity-check matrix has support $\mathcal{T}$ where $\mathcal{T}$ is a (k,w)-wDTS with the property that $T_1=\cdots=T_k$ and none of the differences $a_{1,j}-a_{1,m}$ for $1\leq m<j\leq w$ is divisible by $n-k$, then the Tanner graph associated to the parity-check matrix H of $\mathcal{C}$ is free from cycles of any size (over every base field) not satisfying the FRC.

Theorem 9.24. Assume that $\mathcal{C}$ is an (n,k) convolutional code constructed from an (k,w)-DTS $\mathcal{T}$ with $a_{i,1}=1$ for all $1\leq i\leq k$, where $(n-k)$ does not divide any of the nonzero differences $a_{i_1,j}-a_{i_2,m}$ for $1\leq i_1,i_2\leq k$ and $1\leq m,j\leq w$. Then, the Tanner graph associated to the parity-check matrix H of $\mathcal{C}$ is free from cycles of any size (over every base field) not satisfying the FRC.

Proof. Assume by contradiction that $\mathcal{H}$ contains up to permutations a submatrix A of the form (9.6). As the supports of the first two columns of A intersect, they have to be shifts of different columns of $\bar{H}$. The supports of such shifts can only intersect once and the entries of this intersection come from the first row of $\bar{H}$. Applying the same reasoning to the intersection of the supports of the second and third column of A, implies that a_2 and a_3 in A both come from the first row of $\bar{H}$ which is not possible. This shows the result. □

Example 9.25. Consider the $(2,3)$-DTS $\mathcal{T}=\{T_1,T_2\}$ with $T_1=\{1,2,5\}$ and $T_2=\{1,3,9\}$. The set of all occurring nonzero differences $a_{i_1,j}-a_{i_2,m}$ is $\{1,2,3,4,6,7,8\}$, i.e. none of them is divisible by 5. Hence the matrix $H(z)=H_0+H_1z$ with $H_0=[\bar{H}_0\ I_5]$ and $H_1=[\bar{H}_1\ 0_5]$, where

$$\bar{H}_0=\begin{bmatrix}1&1\\1&0\\0&1\\0&0\\1&0\end{bmatrix},\qquad \bar{H}_1=\begin{bmatrix}0&0\\0&0\\0&0\\0&1\\0&0\end{bmatrix}$$

and

$$[H]_{hr} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

full rank, i.e. $\delta = 1$, is the parity-check matrix of an $(7,2,1)_q$ convolutional code that is free of cycles of any size for any prime power q.

Next, we want to relax the conditions on the wDTS used for construction of the convolutional code but still exclude cycles in the Tanner graph of the sliding parity-check matrix that do not fulfill the FRC by using the construction from Definition 9.14 and considering sufficiently large field sizes.

To ensure that the considered cycle does not satisfy the FRC, we have to guarantee that $\det A \neq 0$ as an element of $\mathbb{F}_q$. It is easy to check that

$$\det A = \prod_{\substack{i=1 \\ i \text{ odd}}}^{2\ell} a_i \pm \prod_{\substack{i=1 \\ i \text{ even}}}^{2\ell} a_i .$$

Let $\mathcal{T}$ be a (k,w)-wDTS and let $\mathcal{C}^{\mathcal{T}}$ be the convolutional code defined from $\mathcal{T}$, with $\mathcal{H}^{\mathcal{T}}$ as defined in (9.2). Each matrix representation A of a 2ℓ-cycle comes from selecting ℓ rows and ℓ columns of $\mathcal{H}^{\mathcal{T}}$. Moreover, in each column of A, exactly two positions are non-zero. Let α be a primitive element for $\mathbb{F}_q$, let $s_1, \dots, s_\ell \in \mathbb{N}$ be the indexes of the columns of $\mathcal{H}^{\mathcal{T}}$, selected to form the cycle, (we consider s_i also if we select the shift of the i-th column) hence we have that $1 \leq s_h \leq k$.

We can write A in the following form:

$$\begin{bmatrix} \alpha^{r_1 s_1} & \alpha^{r_2 s_2} & 0 & 0 & \cdots & 0 \\ 0 & \alpha^{(r_2+i_1)s_2} & \alpha^{(r_3+i_1)s_3} & 0 & \cdots & 0 \\ 0 & 0 & \alpha^{(r_3+i_1+i_2)s_3} & \alpha^{(r_4+i_1+i_2)s_4} & \cdots & 0 \\ \vdots & & & \ddots & \ddots & \\ \alpha^{(r_1+i_1+\cdots+i_{\ell-1})s_1} & 0 & 0 & 0 & \cdots & \alpha^{(r_\ell+i_1+\cdots+i_{\ell-1})s_\ell} \end{bmatrix},$$

where $i_h \in \mathbb{Z}$ and $|i_h|$ is equal to a difference from $T_{s_{h+1}}$ for $h = 1, \dots \ell - 1$ and $|i_1 + \cdots + i_{\ell-1}|$ is equal to a difference from T_{s_1}. Moreover, $1 \leq r_h + i_1 + \ldots + i_g \leq (\nu+1)(n-k)$ for $h = 1, \dots, \ell$ and $g = 0, \dots, \ell - 1$.

We want to estimate the sufficient field size to have that this determinant is nonzero and therefore, we distinguish two cases.

Case I: Assume that ℓ is odd. In this case, the determinant of a matrix of the form (9.6) is given by

$$\det A = \prod_{\substack{i=1\\ i \text{ odd}}}^{2\ell} a_i + \prod_{\substack{i=1\\ i \text{ even}}}^{2\ell} a_i.$$

Hence, if the characteristic of the field is $p > 2$, it is equal to 0 in $\mathbb{F}_q$ if and only if

$$\alpha^{(i_1+i_2+\cdots+i_{\ell-1})s_1} + \alpha^{i_1s_2+i_2s_3+\cdots+i_{\ell-1}s_\ell} = 0,$$

which is equivalent to

$$(i_1 + i_2 + \cdots + i_{\ell-1})s_1 = i_1s_2 + i_2s_3 + \cdots + i_{\ell-1}s_\ell + \frac{(q-1)}{2} \mod (q-1),$$

and hence

$$i_1(s_2 - s_1) + i_2(s_3 - s_1) + \cdots + i_{\ell-1}(s_\ell - s_1) - \frac{(q-1)}{2} = 0 \mod (q-1).$$

It is then enough to consider q bigger than the maximum value that can be reached by the function

$$1 + 2\sum_{h=1}^{\ell-1} i_h(s_{h+1} - s_1).$$

Now, note that i_h can be also negative but in general, we can say that $|i_h| \leq (\nu+1)(n-k)-1$. Moreover, $|s_i - s_1| \leq k-1$. Hence, if we can ensure that

$$\begin{aligned} q &> 2((\nu+1)(n-k)-1)(\ell-1)(k-1)+1 \\ &= 2(\nu+1)(n-k)(\ell-1)(k-1) - 2(\ell-1)(k-1)+1, \end{aligned}$$

with this construction we have a convolutional code whose sliding parity-check matrix is associated to a Tanner graph free from 2ℓ-cycles, with ℓ odd, not satisfying the FRC.

Remark 9.26. Observe that in Theorem 9.17, we computed a more accurate estimation of the field size for getting rid of the 2ℓ cycles, for $\ell = 3$, namely, $q > 2(\nu+1)(n-k)(k-1)-2(k-1)+1$. The computation above shows that with $q > 4(\nu+1)(n-k)(k-1)-4(k-1)+1$ we do not have 6-cycles not satisfying the FRC. This difference is due to the possibility of a better estimation of the terms in the above inequality.

With the discussion above we have proved the following result.

Theorem 9.27. Let n, k, w be positive integers with $n > k$, $\mathcal{T}$ be a (k, w)-wDTS and $\mathcal{C}^{\mathcal{T}}$ be the $(n, k)_q$ convolutional code constructed from $\mathcal{T}$ with $q = p^N$ and $p > 2$. A sufficient condition for obtaining a code whose sliding parity-check matrix is free from 2ℓ-cycles not satisfying the FRC with ℓ odd is to choose a field size $q > 2(\nu+1)(n-k)(\ell-1)(k-1) - 2(k-1)(\ell-1)+1$, where $\nu = \left\lceil \frac{m(\mathcal{T})}{n-k} \right\rceil - 1$ is the degree of the parity-check matrix of $\mathcal{C}^{\mathcal{T}}$.

Example 9.28. Consider again the code constructed in Example 9.15. From Remark 9.21, we know that the highest length that we can have for a cycle is $10 = 2 \cdot 5$, but for q odd with $q > 41$ all the 10-cycles satisfy the FRC.

<u>**Case II:**</u> Assume that ℓ is even. In this case, the determinant of a matrix of the form (9.6) is given by

$$\det A = \prod_{\substack{i=1\\ i \text{ odd}}}^{2\ell} a_i - \prod_{\substack{i=1\\ i \text{ even}}}^{2\ell} a_i.$$

After some straightforward computation, it is easy to see that this determinant is equal to 0 in $\mathbb{F}_q$ if and only if

$$\alpha^{(i_1+i_2+\cdots+i_{\ell-1})s_1} = \alpha^{i_1 s_2 + i_2 s_3 + \cdots + i_{\ell-1} s_\ell},$$

which is equivalent to

$$(i_1 + i_2 + \cdots + i_{\ell-1})s_1 = i_1 s_2 + i_2 s_3 + \cdots + i_{\ell-1} s_\ell \mod (q-1),$$

and hence

$$f(i,s) := i_1(s_2 - s_1) + i_2(s_3 - s_1) + \cdots + i_{\ell-1}(s_\ell - s_1) = 0 \mod (q-1).$$

for $i := (i_1, \dots, i_{\ell-1})$ and $s := (s_1, \dots, s_\ell)$.

Moreover, we have the following constraints:

1. $-(\nu+1)(n-k)+1 \leq i_h \leq (\nu+1)(n-k)-1$ for $h = 1, \dots, \ell-1$
2. $-k+1 \leq s_{h+1} - s_1 \leq k-1$, for $h = 1, \dots, \ell-1$;

We have to find conditions on the corresponding wDTS to ensure that $f(i,s)$ is nonzero when viewed as an element of $\mathbb{Z}$ and then, we can determine a lower bound for q in order that it is also nonzero modulo $q-1$.

Using Proposition 9.22, we know that if none of the differences in the difference triangle set is divisible by $n-k$, then not all the values $s_1, \dots, s_\ell$ can be identical. In particular, there is at least one $h \in \{2, \dots, \ell\}$ such that $s_h - s_1 \neq 0$.

Theorem 9.29. *Let ℓ be an even integer, k, n, w be integers such that $n > k$, $\mathcal{T}$ be a (k,w)-wDTS and $\mathcal{C}^{\mathcal{T}}$ be the $(n,k)_q$ convolutional code constructed from $\mathcal{T}$. Assume that $\mathcal{T}$ fulfills the conditions of Proposition 9.22 and has the property that $f(i,s)$ is nonzero in $\mathbb{Z}$ for all $s_1, \dots, s_\ell \in \{1, \dots, k\}$ not all equal if $|i_h|$ is equal to a difference from $T_{s_{h+1}}$ for $h = 1, \dots \ell-1$ and $|i_1 + \cdots + i_{\ell-1}|$ is equal to a difference from T_{s_1} and $q > ((\nu+1)(n-k)-1)\left((k-1)\frac{\ell}{2} + (k-2)\frac{\ell-2}{2}\right) + 1$. Then, the Tanner graph associated to the sliding parity-check matrix of $\mathcal{C}^{\mathcal{T}}$ is free from 2ℓ-cycles that do not satisfy the FRC.*

Proof. The conditions of the theorem ensure that $f(i,s)$ is nonzero in $\mathbb{Z}$. Moreover, it follows from Proposition 9.22 that

$$((\nu+1)(n-k)-1)\left((k-1)\frac{\ell}{2}+(k-2)\frac{\ell-2}{2}\right)$$

is an upper bound for $|f(i,s)|$. Hence, the result follows. □

Next, we want to give an example for a convolutional code that fulfills the conditions of the preceding theorem.

Example 9.30. Let $n=7$ and $k=2$ and $T_1=\{1,2,5,9\}$ and $T_2=\{1,2,4,10\}$, i.e. $\nu=1$. Note that T_1 is no difference triangle in the strict sense as $9-5=5-1$ but as $n-k=5$ does not divide $9-5$, we can still use it for the construction of our code (see Remark 9.12). We get

$$\mathcal{H}^{\mathcal{T}}=\begin{bmatrix}
\alpha & \alpha^2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
\alpha^2 & \alpha^4 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & \alpha^8 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
\alpha^5 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \alpha & \alpha^2 & 1 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \alpha^2 & \alpha^4 & 0 & 1 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0\\
\alpha^9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \alpha^8 & 0 & 0 & 0 & 1 & 0\\
0 & \alpha^{20} & 0 & 0 & 0 & 0 & 0 & \alpha^5 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}.$$

Form Remark 9.21 one knows that with these parameters it is not possible to have cycles of length 2ℓ for $\ell>4$. Moreover, from Theorem 9.16, we obtain that we can exclude 4-cycles not fulfilling the FRC if $q>11$ and from Theorem 9.17, that we can exclude 6-cycles not fulfilling the FRC if $q>19$ and q is odd. We will show that with the help of the preceding theorem, we can also exclude 8-cycles in $\mathcal{H}^{\mathcal{T}}$ that do not fulfill the FRC or in other words, all 2ℓ-cycles for any ℓ in $\mathcal{H}^{\mathcal{T}}$ fulfill the FRC for $q>19$. First, from Proposition 9.22, we know that in the matrix A representing any 8-cycle we necessarily have $s_1=s_3$ and $s_2=s_4$ and each column of $\bar{H}^{\mathcal{T}}$ is involved once unshifted and once shifted by 5. We get $f(i,s)=\pm(i_1+i_3)$ and have to exclude that $i_1\neq -i_3$. Considering $\mathcal{H}^{\mathcal{T}}$, we realize that 8-cycles are only possible for $s_1=s_3=1$ and $s_2=s_4=2$ and $i_1\in\{\pm 8,\pm 9\}$ and $i_3\in\{\pm 2,\pm 3\}$. Hence, for $q>9\cdot 2+1=19$, the corresponding convolutional code is free from 8-cycles not fulfilling the FRC and hence, free from 2ℓ-cycles not fulfilling the FRC for any ℓ.

To conclude this section, we will modify our construction from Definition 9.14 in order to further relax the conditions on the underlying wDTS and still ensuring that we have no cycles not fulfilling the FRC. However, this will come with the cost of a larger field size.

Definition 9.31. Let k, n be positive integers with $n > k$ and $\mathcal{T} := \{T_1, \dots, T_k\}$ an (k, w)-wDTS with scope $m(\mathcal{T})$. Set $\nu = \left\lceil \frac{m(\mathcal{T})}{n-k} \right\rceil - 1$ and let α be a primitive element for $\mathbb{F}_q$. Moreover, let P be a prime (with properties that will be determined later). For any $1 \leq i \leq (\nu+1)(n-k)$, $1 \leq l \leq k$, define

$$\bar{H}_{i,l}^{(\mathcal{T})} := \begin{cases} \alpha^{P^i l} & \text{if } i \in T_l \\ 0 & \text{otherwise} \end{cases} .$$

Theorem 9.32. Let k, n, w be positive integers with $n > k$ and $\mathcal{T}$ be a DTS with $a_{i,1} = 1$ for all $1 \leq i \leq k$ and $\mathcal{C}$ be an $(n, k)_q$ convolutional code constructed from $\bar{H}^{(\mathcal{T})}$. If $P > \ell k$ and $q > kP^{(\nu+1)(n-k)} \frac{P^{2\ell}-1}{P^{2\ell}-P^{2\ell-1}} + 1$, then the Tanner graph associated to the sliding parity-check matrix contains no cycles of size 2ℓ not fulfilling the FRC.

Proof. As with the construction from Definition 9.14, we obtain that $\det(A) = 0$ if and only if a certain linear combination $\tilde{f}(i, s)$ of exponents of P with 2ℓ coefficients from $\{1, \dots, k\}$ is zero. As the exponents correspond to row indices before a possible shift and the unshifted columns only intersect in the first row, all exponents that are equal to any other exponent are equal to 1. Moreover, as exponents from the same column of A cannot be the same, at most ℓ exponents can be equal to 1. In summary, we obtain that $\tilde{f}(i, s)$ is of the form $\tilde{f}(i, s) = Px + P^{e_1}x_1 + \cdots + P^{e_t}x_t$ with natural numbers $1 < e_1 < \cdots < e_t \leq (\nu+1)(n-k)$, $t \in \{\ell, \dots, 2\ell\}$, $x_j \in \{-k, \dots, k\} \setminus \{0\}$ for $j = 1, \dots, t$ and $x \in \{-\ell k, \dots, +\ell k\}$. Since m was chosen to be a prime larger than ℓk, $\tilde{f}(i, s)$ is nonzero in $\mathbb{Z}$. Furthermore, $|\tilde{f}(i, s)| \leq k \sum_{i=0}^{2\ell-1} P^{(\nu+1)(n-k)-i} = kP^{(\nu+1)(n-k)} \frac{P^{2\ell}-1}{P^{2\ell}-P^{2\ell-1}}$ and hence it cannot be zero modulo $q - 1$. □

Finally, we illustrate our modified construction with an example.

Example 9.33. If we take the DTS $\mathcal{T} = \{\{1, 2, 5\}, \{1, 3, 8\}\}$ to construct an $(6, 2)_q$ convolutional code, we have $m(\mathcal{T}) = 8$ and $\nu = 1$. If we want that the girth of the corresponding parity-check matrix is at least 12, we have to choose $P > 10$, i.e. $P = 11$. To get the desired property it would be sufficient if the field size is larger than 4.716×10^8. If it is sufficient to have a girth of at least 8, it would be enough to choose $P = 7$ and the sufficient field size decreases to 1.35×10^7.

Bibliography

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

[2] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung. Network information flow. *IEEE Trans. Inform. Theory*, 46(4):1204–1216, 2000.

[3] A. Alahmadi, C. Güneri, H. Shoaib, and P. Solé. Long quasi-polycyclic t-CIS codes. *Adv. Math. Commun.*, 12(1):189–198, 2018.

[4] A. Alahmadi, F. Özdemir, and P. Solé. On self-dual double circulant codes. *Designs, Codes and Cryptography*, 86(6):1257–1265, 2018.

[5] G. N. Alfarano, M. Borello, and A. Neri. A geometric characterization of minimal codes and their asymptotic performance. *Advances in Mathematics of Communications*, 16(1):115, 2022.

[6] G. N. Alfarano, M. Borello, A. Neri, and A. Ravagnani. Linear cutting blocking sets and minimal codes in the rank metric. *Journal of Combinatorial Theory, Series A*, 192:105658, 2022.

[7] G. N. Alfarano, M. Borello, A. Neri, and A. Ravagnani. Three combinatorial perspectives on minimal codes. *SIAM Journal on Discrete Mathematics*, 36(1):461–489, 2022.

[8] G. N. Alfarano and J. Lieb. On the left primeness of some polynomial matrices with applications to convolutional codes. *Journal of Algebra and Its Applications*, 20(11):2150207, 2021.

[9] G. N. Alfarano, J. Lieb, and J. Rosenthal. Construction of rate $(n-1)/n$ non-binary LDPC convolutional codes via difference triangle sets. *Proceedings of IEEE International Symposium on Information Theory*, 2020.

[10] G. N. Alfarano, J. Lieb, and J. Rosenthal. Construction of LDPC convolutional codes via difference triangle sets. *Designs, Codes and Cryptography*, 89(10):2235–2254, 2021.

[11] G. N. Alfarano, D. Napp, A. Neri, and V. Requena. Weighted Reed-Solomon convolutional codes. *arXiv preprint arXiv:2012.11417*, 2020.

[12] P. Almeida, D. Napp, and R. Pinto. A new class of superregular matrices and MDP convolutional codes. *Linear Algebra and its Applications*, 439(7):2145–2157, 2013.

[13] P. Almeida, D. Napp, and R. Pinto. Superregular matrices and applications to convolutional codes. *Linear Algebra and its Applications*, 499:1–25, 2016.

[14] P. J. Almeida and J. Lieb. Complete j-MDP convolutional codes. *IEEE Transactions on Information Theory*, 66(12):7348–7359, 2020.

[15] P. J. Almeida and D. Napp. Superregular matrices over small finite fields. *arXiv preprint arXiv:2008.00215*, 2020.

[16] N. Alon. Combinatorial nullstellensatz. *Combin. Probab. Comput.*, 8(1-2):7–29, 2001.

[17] N. Alon and Z. Füredi. Covering the cube by affine hyperplanes. *European J. Combin.*, 14(2):79–83, 1993.

[18] B. Amiri, J. Kliewer, and L. Dolecek. Analysis and enumeration of absorbing sets for non-binary graph-based codes. *IEEE Transactions on Communications*, 62(2):398–409, 2014.

[19] A. Ashikhmin and A. Barg. Minimal vectors in linear codes. *IEEE Trans. Inform. Theory*, 44(5):2010–2017, 1998.

[20] A. Badr, P. Patil, A. Khisti, W.-T. Tan, and J. Apostolopoulos. Layered constructions for low-delay streaming codes. *IEEE Transactions on Information Theory*, 63(1):111–141, 2016.

[21] S. Ball. Multiple blocking sets and arcs in finite planes. *Journal of the London Mathematical Society*, 54(3):581–593, 1996.

[22] S. Ball. *Finite geometry and combinatorial applications*, volume 82. Cambridge University Press, 2015.

[23] S. Ball and A. Blokhuis. On the size of a double blocking set in $\mathrm{PG}(2, q)$. *Finite Fields and their Applications*, 2(2):125–137, 1996.

[24] S. Ball and Z. Weiner. An introduction to finite geometry. *Preprint*, 162, 2011.

[25] J. Barát and L. Storme. Multiple blocking sets in $\mathrm{PG}(n, q), n \geq 3$. *Des. Codes Cryptogr.*, 33(1):5–21, 2004.

[26] D. Bartoli and M. Bonini. Minimal linear codes in odd characteristic. *IEEE Transactions on Information Theory*, 65(7):4152–4155, 2019.

[27] D. Bartoli, M. Bonini, and B. Güneş. An inductive construction of minimal codes. *arXiv preprint arXiv:1911.09093*, 2019.

[28] D. Bartoli, A. Cossidente, G. Marino, and F. Pavese. On cutting blocking sets and their codes. *arXiv preprint arXiv:2011.11101*, 2020.

[29] D. Bartoli, B. Csajbók, G. Marino, and R. Trombetti. Evasive subspaces. *arXiv preprint arXiv:2005.08401*, 2020.

[30] D. Bartoli, M. Giulietti, G. Marino, and O. Polverino. Maximum scattered linear sets and complete caps in Galois spaces. *Combinatorica*, 38(2):255–278, 2018.

[31] S. Bates, Z. Chen, and X. Dong. Low-density parity-check convolutional codes for ethernet networks. In *PACRIM. 2005 IEEE Pacific Rim Conference on Communications, Computers and signal Processing, 2005.*, pages 85–88. IEEE, 2005.

[32] S. Bates, D. G. Elliott, and R. Swamy. Termination sequence generation circuits for low-density parity-check convolutional codes. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 53(9):1909–1917, 2006.

[33] M. Battaglioni, M. Baldi, F. Chiaraluce, and M. Lentmaier. Girth properties of time-varying SC-LDPC convolutional codes. In *2019 IEEE International Symposium on Information Theory (ISIT)*, pages 2599–2603. IEEE, 2019.

[34] T. P. Berger. Isometries for rank distance and permutation group of Gabidulin codes. *IEEE Trans. Inform. Theory*, 49(11):3016–3019, 2003.

[35] C. Berrou, A. Glavieux, and P. Thitimajshima. Near Shannon limit error-correcting coding and decoding: Turbo-codes. 1. In *Proceedings of ICC'93-IEEE International Conference on Communications*, volume 2, pages 1064–1070. IEEE, 1993.

[36] A. Beutelspacher. On Baer subspaces of finite projective spaces. *Math. Z.*, 184(3):301–319, 1983.

[37] R. Bhatia and C. Davis. A better bound on the variance. *Amer. Math. Monthly*, 107(4):353–357, 2000.

[38] A. Bishnoi, P. L. Clark, A. Potukuchi, and J. R. Schmitt. On zeros of a polynomial in a finite grid. *Combinatorics, Probability and Computing*, 27(3):310–333, 2018.

[39] G. R. Blakley. Safeguarding cryptographic keys. *Proceedings of the 1979 AFIPS National Computer Conference*, 48:313–317, 1979.

[40] A. Blokhuis and M. Lavrauw. Scattered spaces with respect to a spread in $\mathrm{PG}(n, q)$. *Geom. Dedicata*, 81(1-3):231–243, 2000.

[41] A. Blokhuis, P. Sziklai, and T. Szonyi. Blocking sets in projective spaces. In L. Storme and J. de Beule, editors, *Current research topics in Galois geometry*, pages 61–84. Nova Sci. Publ., New York, 2011.

[42] M. Bonini and M. Borello. Minimal linear codes arising from blocking sets. *J. Algebraic Combin.*, 53(2):327–341, 2021.

[43] A. Bonisoli. Every equidistant linear code is a sequence of dual Hamming codes. *Ars Combin.*, 18:181–186, 1983.

[44] M. Borello and W. Willems. Group codes over fields are asymptotically good. *arXiv preprint arXiv:1904.10885*, 2019.

[45] W. Bosma, J. Cannon, and C. Playoust. The Magma algebra system I: The user language. *J. Symbol. Comput.*, 24:235–265, 1997.

[46] A. E. Brouwer and H. A. Wilbrink. Blocking sets in translation planes. *J. Geom.*, 19(2):200, 1982.

[47] C. Carlet, C. Ding, and J. Yuan. Linear codes from highly nonlinear functions and their secret sharing schemes. *IEEE Trans. Inf. Theory*, 51(6):2089–2102, 2005.

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

[49] S. Chang and J. Y. Hyun. Linear codes from simplicial complexes. *Designs, Codes and Cryptography*, 86(10):2167–2181, 2018.

[50] Y. M. Chee and C. J. Colbourn. Constructions for difference triangle sets. *IEEE Transactions on Information Theory*, 43(4):1346–1349, 1997.

[51] C. Chen, W. W. Peterson, and E. Weldon Jr. Some results on quasi-cyclic codes. *Information and Control*, 15(5):407–423, 1969.

[52] Z. Chen, P. Fan, and F. Jin. Disjoint difference sets, difference triangle sets, and related codes. *IEEE Transactions on Information Theory*, 38(2):518–522, 1992.

[53] S.-Y. Chung, T. J. Richardson, and R. L. Urbanke. Analysis of sum-product decoding of low-density parity-check codes using a gaussian approximation. *IEEE Transactions on Information theory*, 47(2):657–670, 2001.

[54] G. D. Cohen, S. Mesnager, and A. Patey. On minimal and quasi-minimal linear codes. In *Cryptography and coding*, volume 8308 of *Lecture Notes in Comput. Sci.*, pages 85–98. Springer, Heidelberg, 2013.

[55] C. J. Colbourn. Difference triangle sets. *Chapter in The CRC Handbook of Combinatorial Designs by CJ Colbourn and J. Dintz*, pages 312–317, 1996.

[56] B. Csajbók, G. Marino, O. Polverino, and F. Zullo. Maximum scattered linear sets and MRD-codes. *J. Algebraic Combin.*, 46(3-4):517–531, 2017.

[57] M. C. Davey and D. J. MacKay. Low density parity check codes over GF (q). In *1998 Information Theory Workshop (Cat. No. 98EX131)*, pages 70–71. IEEE, 1998.

[58] M. C. Davey and D. J. MacKay. Monte Carlo simulations of infinite low density parity check codes over GF(q). In *Proc. of Int. Workshop on Optimal Codes and related Topics*, pages 9–15. Citeseer, 1998.

[59] A. A. Davydov, M. Giulietti, S. Marcugini, and F. Pambianco. Linear nonbinary covering codes and saturating sets in projective spaces. *Adv. in Math. Commun.*, 5(1):119–147, 2011.

[60] J. de la Cruz, E. Gorla, H. H. López, and A. Ravagnani. Weight distribution of rank-metric codes. *Des. Codes Cryptogr.*, 86(1):1–16, 2018.

[61] R. dela Cruz and S. Kurz. On the maximum number of minimal codewords. *Discrete Mathematics*, 344(9):112510, 2021.

[62] P. Delsarte. Four fundamental parameters of a code and their combinatorial significance. *Inform. and Control*, 23(5):407–438, 1973.

[63] P. Delsarte. Bilinear forms over a finite field, with applications to coding theory. *J. Combin. Theory Ser. A*, 25(3):226–241, 1978.

[64] C. Ding. Linear codes from some 2-designs. *IEEE Transactions on information theory*, 61(6):3265–3275, 2015.

[65] C. Ding, D. R. Kohel, and S. Ling. Secret-sharing with a class of ternary codes. *Theoretical Computer Science*, 246(1-2):285–298, 2000.

[66] C. Ding, C. Li, N. Li, and Z. Zhou. Three-weight cyclic codes and their weight distributions. *Discrete Mathematics*, 339(2):415–427, 2016.

[67] P. Elias. Coding for noisy channels. *IRE Conv. Rec.*, 3:37–46, 1955.

[68] S. Fancsali and P. Sziklai. Lines in higgledy-piggledy arrangement. *Electron. J. Comb.*, 21(2), 2014.

[69] S. L. Fong, A. Khisti, B. Li, W. Tan, X. Zhu, and J. Apostolopoulos. Optimal streaming codes for channels with burst and arbitrary erasures. *IEEE Transactions on Information Theory*, 65(7):4274–4292, 2019.

[70] G. Forney. Convolutional codes I: Algebraic structure. *IEEE Transactions on Information Theory*, 16(6):720–738, 1970.

[71] G. Forney. Structural analysis of convolutional codes via dual codes. *IEEE Transactions on Information Theory*, 19(4):512–518, 1973.

[72] G. Forney, Jr. Minimal bases of rational vector spaces, with applications to multivariable linear systems. *SIAM Journal on Control*, 13:493–520, 1975.

[73] G. D. Forney Jr. Convolutional codes II. Maximum-likelihood decoding. *Information and control*, 25(3):222–266, 1974.

[74] E. M. Gabidulin. Theory of codes with maximum rank distance. *Probl. Peredachi Informatsii*, 21(1):3–16, 1985.

[75] R. Gallager. Low-density parity-check codes. *IRE Transactions on Information Theory*, 8(1):21–28, 1962.

[76] F. R. Gantmakher. *The theory of matrices*, volume 131. American Mathematical Soc., 1959.

[77] O. Geil and C. Thomsen. Weighted Reed–Muller codes revisited. *Des. Codes Cryptogr.*, 66(1-3):195–220, 2013.

[78] M. Giorgetti and A. Previtali. Galois invariance, trace codes and subfield subcodes. *Finite Fields Appl.*, 16(2):96–99, 2010.

[79] H. Gluesing-Luerssen, J. Rosenthal, and R. Smarandache. Strongly-MDS convolutional codes. *IEEE Transactions on Information Theory*, 52(2):584–598, 2006.

[80] H. Gluesing-Luerssen and F.-L. Tsang. A matrix ring description for cyclic convolutional codes. *Advances in Mathematics of Communications*, 2(1):55, 2008.

[81] J. H. Griesmer. A bound for error-correcting codes. *IBM J. Research Develop.*, 4:532–542, 1960.

[82] A. Gruica and A. Ravagnani. Common complements of linear subspaces and the sparseness of MRD codes. *arXiv:2011.02993*, 2020.

[83] R. W. Hamming. Error detecting and error correcting codes. *The Bell system technical journal*, 29(2):147–160, 1950.

[84] J. Hansen, J. Østergaard, J. Kudahl, and J. H. Madsen. Superregular lower triangular Toeplitz matrices for low delay wireless streaming. *IEEE Transactions on Communications*, 65(9):4027–4038, 2017.

[85] Z. Heng, C. Ding, and Z. Zhou. Minimal linear codes over finite fields. *Finite Fields Appl.*, 54:176–196, 2018.

[86] R. A. Horn and C. R. Johnson. *Matrix analysis*. Cambridge University Press, 2013.

[87] L.-K. Hua. A theorem on matrices over a sfield and its applications. *Acta Math. Sinica*, 1(2):109–163, 1951.

[88] W. C. Huffman and V. Pless. *Fundamentals of error-correcting codes*. Cambridge University Press, 2010.

[89] R. Hutchinson, R. Smarandache, and J. Trumpf. On superregular matrices and MDP convolutional codes. *Linear Algebra and its Applications*, 428:2585–2596, 2008.

[90] S. J. Johnson and S. R. Weller. Regular low-density parity-check codes from combinatorial designs. In *Proceedings 2001 IEEE Information Theory Workshop (Cat. No. 01EX494)*, pages 90–92. IEEE, 2001.

[91] R. Jurrius and R. Pellikaan. On defining generalized rank weights. *Adv. Math. Commun.*, 11(1):225–235, 2017.

[92] T. Kailath. *Linear systems*, volume 156. Prentice-Hall Englewood Cliffs, NJ, 1980.

[93] E. Karnin, J. Greene, and M. Hellman. On secret sharing systems. *IEEE Transactions on Information Theory*, 29(1):35–41, 1983.

[94] T. Klove. Bounds and construction for difference triangle sets. *IEEE Transactions on Information Theory*, 35(4):879–886, 1989.

[95] R. Kötter and F. R. Kschischang. Coding for errors and erasures in random network coding. *IEEE Trans. Inform. Theory*, 54(8):3579–3591, 2008.

[96] Y. Kou, S. Lin, and M. Fossorier. Construction of low density parity check codes: a geometric approach. In *Proceedings of the 2nd International Symposium on Turbo Codes and Related Topics*, pages 137–140, 2000.

[97] Y. Kou, S. Lin, and M. P. Fossorier. Low-density parity-check codes based on finite geometries: a rediscovery and new results. *IEEE Transactions on Information theory*, 47(7):2711–2736, 2001.

[98] S. Kurz. On the number of minimal codewords in codes generated by the adjacency matrix of a graph. *Discrete Applied Mathematics*, 309:221–228, 2022.

[99] M. Lavrauw and G. Van de Voorde. Field reduction and linear sets in finite geometry. *Topics in finite fields*, 632:271–293, 2015.

[100] S.-Y. R. Li, R. W. Yeung, and N. Cai. Linear network coding. *IEEE Trans. Inform. Theory*, 49(2):371–381, 2003.

[101] J. Lieb. Complete MDP convolutional codes. *Journal of Algebra and Its Applications*, 18(6):1950105, 2019.

[102] J. Lieb. Necessary field size and probability for MDP and complete MDP convolutional codes. *Designs, Codes and Cryptography*, 87(12):3019–3043, 2019.

[103] J. Lieb and R. Pinto. Constructions of MDS convolutional codes using superregular matrices. *Journal of Algebra Combinatorics Discrete Structures and Applications*, 7(1), 2020.

[104] J. Lieb, R. Pinto, and J. Rosenthal. Convolutional codes. *arXiv preprint arXiv:2001.08281*, 2020.

[105] H. H. López, C. Rentería-Márquez, and R. H. Villarreal. Affine cartesian codes. *Des. Codes Cryptogr.*, 71(1):5–19, 2014.

[106] W. Lu, X. Wu, and X. Cao. The parameters of minimal linear codes. *arXiv preprint arXiv:1911.07648*, 2019.

[107] G. Lunardon. Normal spreads. *Geom. Dedicata*, 75(3):245–261, 1999.

[108] D. J. MacKay and R. M. Neal. Near shannon limit performance of low density parity check codes. *Electronics letters*, 32(18):1645–1646, 1996.

[109] F. J. MacWilliams and N. J. A. Sloane. *The theory of error correcting codes*, volume 16. Elsevier, 1977.

[110] R. Mahmood, A. Badr, and A. Khisti. Convolutional codes with maximum column sum rank for network streaming. *IEEE Transactions on Information Theory*, 62(6):3039–3052, 2016.

[111] U. Martínez-Peñas. On the similarities between generalized rank and Hamming weights and their applications to network coding. *IEEE Trans. Inform. Theory*, 62(7):4081–4095, 2016.

[112] U. Martínez-Peñas. Hamming and simplex codes for the sum-rank metric. *Des. Codes Cryptogr.*, 88(8):1521–1539, 2020.

[113] J. Massey and M. Sain. Codes, automata, and continuous systems: Explicit interconnections. *IEEE Transactions on Automatic Control*, 12(6):644–650, 1967.

[114] J. L. Massey. *Threshold decoding.* PhD thesis, Massachusetts Institute of Technology, Research Laboratory of Electronics, 1963.

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

[116] J. L. Massey. Some applications of coding theory in cryptography. *Codes and Ciphers: Cryptography and Coding IV*, pages 33–47, 1995.

[117] R. J. McEliece and D. V. Sarwate. On sharing secrets and Reed-Solomon codes. *Communications of the ACM*, 24(9):583–584, 1981.

[118] S. Mesnager. Linear codes with few weights from weakly regular bent functions based on a generic construction. *Cryptogr. Commun.*, 9:71–84, 2017.

[119] S. Mesnager, F. Özbudak, and A. Sınak. Linear codes from weakly regular plateaued functions and their secret sharing schemes. *Des. Codes Cryptogr.*, 87(2-3):463–480, 2019.

[120] S. Mesnager and A. Sınak. Several classes of minimal linear codes with few weights from weakly regular plateaued functions. *IEEE Trans. Inf. Theory*, 66(4):2296–2310, 2019.

[121] D. Napp and R. Smarandache. Constructing strongly MDS convolutional codes with maximum distance profile. *Advances in Mathematics of Communications*, 10(2):275–290, 2016.

[122] P. Oswald and A. Shokrollahi. Capacity-achieving sequences for the erasure channel. *IEEE Transactions on Information Theory*, 48(12):3017–3028, 2002.

[123] R. Pellikaan, X.-W. Wu, and S. Bulygin. *Codes, cryptology and curves with computer algebra.* Cambridge University Press, 2018.

[124] P. Piret. *Convolutional codes: an algebraic approach.* MIT press, 1988.

[125] O. Polverino. Linear sets in finite projective spaces. *Discrete Math.*, 310(22):3096–3107, 2010.

[126] O. Polverino and F. Zullo. Connections between scattered linear sets and MRD-codes. *Bulletin of the ICA*, 89:46–74, 2020.

[127] C. Poulliat, M. Fossorier, and D. Declercq. Design of regular $(2, d_c)$-LDPC codes over GF(q) using their binary images. *IEEE Transactions on Communications*, 56(10):1626–1635, 2008.

[128] A. E. Pusane, R. Smarandache, P. O. Vontobel, and D. J. Costello. Deriving good LDPC convolutional codes from LDPC block codes. *IEEE Transactions on Information Theory*, 57(2):835–857, 2011.

[129] T. H. Randrianarisoa. A geometric approach to rank metric codes and a classification of constant weight codes. *Des. Codes Cryptogr.*, 88(7):1331–1348, 2020.

[130] A. Ravagnani. Generalized weights: an anticode approach. *J. Pure Appl. Algebra*, 220(5):1946–1962, 2016.

[131] T. J. Richardson and R. L. Urbanke. The capacity of low-density parity-check codes under message-passing decoding. *IEEE Transactions on information theory*, 47(2):599–618, 2001.

[132] J. P. Robinson and A. Bernstein. A class of binary recurrent codes with limited error propagation. *IEEE Transactions on Information Theory*, 13(1):106–113, 1967.

[133] J. Rosenthal. Connections between linear systems and convolutional codes. In *Codes, Systems, and Graphical Models*, pages 39–66. Springer, 2001.

[134] J. Rosenthal and R. Smarandache. Maximum distance separable convolutional codes. *Applicable Algebra in Engineering, Communication and Computing*, 10(1):15–32, 1999.

[135] J. Rosenthal and R. Smarandache. Maximum distance separable convolutional codes. *Appl. Algebra Engrg. Comm. Comput.*, 10(1):15–32, 1999.

[136] J. Rosenthal and P. O. Vontobel. Constructions of LDPC codes using Ramanujan graphs and ideas from Margulis. In *in Proc. of the 38-th Allerton Conference on Communication, Control, and Computing.* Citeseer, 2000.

[137] R. M. Roth. Maximum-rank array codes and their application to crisscross error correction. *IEEE Transactions on Information Theory*, 37(2):328 –336, mar 1991.

[138] R. M. Roth. Introduction to coding theory. *IET Communications*, 47, 2006.

[139] B. Segre. Teoria di Galois, fibrazioni proiettive e geometrie non desarguesiane. *Ann. Mat. Pura Appl.*, 64(1):1–76, 1964.

[140] A. Shamir. How to share a secret. *Communications of the ACM*, 22(11):612–613, 1979.

[141] C. E. Shannon. A mathematical theory of communication. *The Bell system technical journal*, 27(3):379–423, 1948.

[142] J. Sheekey. A new family of linear maximum rank distance codes. *Adv. Math. Commun.*, 10(3):475–488, 2016.

[143] J. Sheekey. (Scattered) Linear Sets are to Rank-Metric Codes as Arcs are to Hamming-Metric Codes. In M. Greferath, C. Hollanti, and J. Rosenthal, editors, *Oberwolfach Report No. 13/2019*, 2019.

[144] J. Sheekey and G. Van de Voorde. Rank-metric codes, linear sets, and their duality. *Des. Codes Cryptogr.*, 88(4):655–675, 2020.

[145] D. Silva, F. R. Kschischang, and R. Kötter. A rank-metric approach to error control in random network coding. *IEEE Trans. Inform. Theory*, 54(9):3951–3967, 2008.

[146] R. Singleton. Maximum distance q-nary codes. *IEEE Transactions on Information Theory*, 10(2):116–118, 1964.

[147] Y. Song and Z. Li. Secret sharing with a class of minimal linear codes. *arXiv preprint arXiv:1202.4058*, 2012.

[148] C. Tang, Y. Qiu, Q. Liao, and Z. Zhou. Full characterization of minimal linear codes as cutting blocking sets. *IEEE Trans. Inform. Theory*, 67(6):3690–3700, 2021.

[149] R. Tanner. A recursive approach to low complexity codes. *IEEE Transactions on information theory*, 27(5):533–547, 1981.

[150] R. M. Tanner. Error-correcting coding system, Oct. 13 1981. US Patent 4,295,218.

[151] R. M. Tanner. *Convolutional codes from quasi-cyclic codes: A link between the theories of block and convolutional codes.* University of California, Santa Cruz, Computer Research Laboratory, 1987.

[152] R. M. Tanner, D. Sridhara, A. Sridharan, T. E. Fuja, and D. J. Costello. Ldpc block and convolutional codes based on circulant matrices. *IEEE Transactions on Information Theory*, 50(12):2966–2984, 2004.

[153] V. Tomás, J. Rosenthal, and R. Smarandache. Decoding of convolutional codes over the erasure channel. *IEEE Transactions on Information Theory*, 58(1):90–108, 2012.

[154] S.-Y. Tong. Systematic construction of self-orthogonal diffuse codes. *IEEE Transactions on Information Theory*, 16(5):594–604, 1970.

[155] S.-Y. Tong. Character-correcting convolutional self-orthogonal codes. *Information and Control*, 18:183–202, 1971.

[156] M. Tsfasman and S. G. Vladut. *Algebraic-geometric codes*, volume 58. Springer Science & Business Media, 2013.

[157] J. H. Van Lint. *Introduction to coding theory*, volume 86. Springer Science & Business Media, 2012.

[158] B. Vasic and O. Milenkovic. Combinatorial constructions of low-density parity-check codes for iterative decoding. *IEEE Transactions on information theory*, 50(6):1156–1176, 2004.

[159] O. Veblen and J. W. Young. *Projective geometry. Vol. 1.* Blaisdell Publishing Co. Ginn and Co. New York-Toronto-London, 1965.

[160] N. Wiberg. *Codes and decoding on general graphs*. PhD thesis, Dept. Electrical Eng., Univ. Linkoping, 1996.

[161] E. V. York. *Algebraic description and construction of error correcting codes: a linear system point of view*. PhD thesis, University of Notre Dame, 1997.

[162] J. Yuan and C. Ding. Secret sharing schemes from three classes of linear codes. *IEEE Transactions on Information Theory*, 52(1):206–212, 2005.

[163] H. Zhou and N. Goertz. Cycle analysis of time-invariant LDPC convolutional codes. In *2010 17th International Conference on Telecommunications*, pages 23–28. IEEE, 2010.

[164] G. Zini and F. Zullo. Scattered subspaces and related codes. *Des. Codes Cryptogr.*, pages 1–21, 2021.