5-2015

Applications of Algebraic Geometric Codes to Polar Coding

Sarah E. Anderson
Clemson University

Applications of algebraic geometric codes to polar coding

A Dissertation
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
Mathematical Sciences

---

by
Sarah E. Anderson
May 2015

Accepted by:
Dr. Gretchen L. Matthews, Committee Chair
Dr. Jim Brown
Dr. Shuhong Gao
Dr. Felice Manganiello

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

In 2009, Arikan developed polar codes as the first explicit construction of symmetric capacity achieving codes for binary discrete memoryless channels (DMCs) with low encoding and decoding complexity. In this construction, a kernel matrix

$$G = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \in \mathbb{F}_2^{2\times 2}$$

is considered, and $G^{\otimes n}$ is used to encode a block of 2^n channels. As the number of channels grows, each channel becomes either a noiseless channel or a pure-noise channel, and the rate of this polarization is related to the kernel matrix used.

Korada, Şaşoğlu, and Urbanke considered larger binary matrices as kernels and characterized the speed of polarization by introducing a quantity called the exponent [12]. Şaşoğlu also explored the polarization phenomenon for nonbinary alphabets [23], and polar codes were generalized to arbitrary discrete memoryless channels by Şaşoğlu, Telatar, and Arikan [24]. Mori and Tanaka generalized the arguments of Korada et. al. to kernels over arbitrary finite fields [19] and showed that kernels constructed from Reed-Solomon codes give the largest exponent when the code length

is at most the size of the field [20]. They also stated that a Hermitian code over a field of even characteristic of sufficient size gives a larger exponent than the Reed-Solomon matrix over the same field. Their work, along with the BCH code employed by Korada et. al. [12], suggests algebraic geometric codes are good candidates for constructing kernel matrices.

Algebraic geometric codes were introduced by V.D. Goppa in the 1970s ([10], [11]). Hermitian codes are well-known and studied, and in Chapter 4, we construct kernel matrices from one-point and two-point Hermitian codes as well as from Suzuki codes. In addition, we demonstrate that multipoint code kernels arise naturally from shortening those associated with one-point codes, which leads to a discussion of three-point Hermitian codes.

In Chapter 2, we introduce basics of coding theory and algebraic geometric codes. We end the chapter with an application of algebraic geometric codes to stopping sets. Chapter 3 gives an overview of polar coding, including a discussion of the kernel matrix and exponent. Chapter 4 presents kernels constructed from algebraic geometric codes and ends with a study of triples of rational places on the Hermitian curve.

Chapter 2

Coding Theory

When information is sent across a channel, it may be corrupted by noise in the channel. Coding theory is the study of how one detects, or even corrects, errors that occur due to noise. Error-correcting codes protect information from distortion caused by noise. Figure 2.1 demonstrates how a message m is encoded to be sent across a noisy channel and the received message c' is decoded by the receiver. Applications of coding theory include ISBNs, bar codes, UPCs, flash memories, CDs, DVDs, QR codes, and much more.

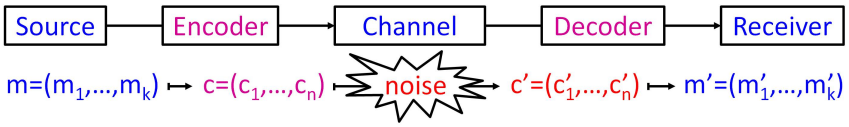

Figure 2.1: Basic encoder and decoder

2.1 Background

Let A be a finite set. Typically, we take A to be a field, and refer to A as an alphabet. More generally, one might consider alphabets which are rings or even those with no algebraic structure.

Definition 1 *Let n be a positive integer. A code, $\mathcal{C}$, is a set of elements of A^n. The length of the code $\mathcal{C}$ is n. Elements of the code $\mathcal{C}$ are called codewords.*

We will restrict our attention to block codes, where a block of k bits is encoded to a codeword of length n. The length n may also be referred to as the block length. To emphasize the choice of alphabet, we sometimes say C is a code over the alphabet A, or just over A for short.

Definition 2 *The minimum distance of a code $\mathcal{C}$ of length n is*

$$d(\mathcal{C}) := \min\{d(x,y) \mid x,y \in \mathcal{C}, x \neq y\}$$

where $d(x,y) = |\{i \in 1,\ldots,n : x_i \neq y_i\}|$ denotes the Hamming distance between codewords $x = (x_1,\ldots,x_n)$ and $y = (y_1,\ldots,y_n)$.

Definition 3 *Let A be a field. A linear code $\mathcal{C}$ of length n is an A-subspace of A^n for some positive integer n. The dimension of $\mathcal{C}$, usually denoted by k, is the dimension of $\mathcal{C}$ as an A-vector space; that is, the dimension of $\mathcal{C}$ is $k := \dim_A(\mathcal{C})$.*

Given a prime power q, we will let $\mathbb{F}_q$ denote the finite field with q elements, and $\mathbb{F}_q^n$ denote the set of $1 \times n$ vectors with entries in $\mathbb{F}_q$ where n is a positive integer. Given $u \in \mathbb{F}_q^n$, u_i denotes the i^{th} coordinate of u. Let $\mathbb{F}_q^{m\times n}$ denote the set of $m \times n$ matrices with entries in $\mathbb{F}_q$ where n and m are positive integers.

We refer to a linear code of length n and dimension k as an $[n, k]$-code. The information rate of an $[n, k]$-code is

$$R = \frac{k}{n}.$$

Definition 4 *Let $\mathcal{C}$ be an $[n, k]$-code. A generator matrix G of $\mathcal{C}$ is a $k \times n$ matrix whose rows form a basis of $\mathcal{C}$; that is,*

$$G = \begin{bmatrix} -- & g_1 & -- \\ -- & g_2 & -- \\ & \vdots & \\ -- & g_{k-1} & -- \\ -- & g_k & -- \end{bmatrix}$$

where $\{g_1, g_2, \ldots, g_k\}$ is a basis for $\mathcal{C}$.

Definition 5 *Let $\mathcal{C}$ be an $[n, k]$-code over $\mathbb{F}_q$. The dual code of $\mathcal{C}$, denoted $\mathcal{C}^\perp$, is defined by*

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

where $\langle x, c \rangle = \sum_{i=1}^{n} x_i c_i$.

Definition 6 *Let $\mathcal{C}$ be an $[n, k]$-code. A parity check matrix of $\mathcal{C}$ is a generator matrix of $\mathcal{C}^\perp$.*

Definition 7 *Two codes $\mathcal{C}_1$ and $\mathcal{C}_2$ of length n over $\mathbb{F}_q$ are isometric if and only if there exists a vector space isomorphism $\sigma : \mathcal{C}_1 \to \mathcal{C}_2$ that preserves Hamming distance; that is, $d(c_1, c_2) = d(\sigma(c_1), \sigma(c_2))$ for all $c_1, c_2 \in \mathcal{C}$.*

Note if $\mathcal{C}$ is a linear code, then the minimum distance of $\mathcal{C}$ satisfies

$$d(\mathcal{C}) = \min\{w(c) \mid 0 \neq c \in \mathcal{C}\}$$

where $w(c) := d(c, 0)$ is the weight of the codeword c.

We refer to a linear code of length n, dimension k, and minimum distance d as an $[n, k, d]$-code over $\mathbb{F}_q$. If $\mathcal{C}$ is an $[n, k, d]$-code, then $\mathcal{C}$ can correct up to $\lfloor \frac{d-1}{2} \rfloor$ errors; that is, for all $w \in \mathbb{F}_q^n$ with $d(w, c) \leq \lfloor \frac{d-1}{2} \rfloor$ for some $c \in \mathcal{C}$, then c is the unique codeword satisfying $d(w, c) \leq \lfloor \frac{d-1}{2} \rfloor$.

One of the main problems of coding theory is to construct codes whose dimension is large (which guarantees a large information rate) and whose minimum distance is large (which guarantees correction of many errors). However, in some sense these are competing goals as demonstrated by with the Singleton Bound below.

Proposition 1 *(Singleton Bound) For an $[n, k, d]$-code $\mathcal{C}$,*

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

Proof

Let $\mathcal{C}$ be an $[n, k, d]$-code. Consider the linear subspace $W \subseteq \mathbb{F}_q^n$ given by

$$W := \{(a_1, \ldots, a_n) \in \mathbb{F}_q^n \mid a_i = 0 \text{ for all } i \geq d\}.$$

Let $a \in W$. Then $w(a) \leq d-1$. Hence, $a \notin \mathcal{C}$ and $\mathcal{C} \cap W = \{0\}$. Note $\dim_{\mathbb{F}_q}(W) = d-1$. Thus,

$$\begin{aligned} k + (d-1) &= \dim_{\mathbb{F}_q}(\mathcal{C}) + \dim_{\mathbb{F}_q}(W) \\ &= \dim_{\mathbb{F}_q}(\mathcal{C} + W) + \dim_{\mathbb{F}_q}(\mathcal{C} \cap W) \\ &= \dim_{\mathbb{F}_q}(\mathcal{C} + W) \\ &\leq n. \end{aligned}$$

Therefore, $k + d \leq n + 1$. $\square$

Codes such that $k + d = n + 1$ are called MDS (maximum distance separable) codes. Next, we will review Reed-Solomon codes, which are examples of MDS codes.

Let $\mathbb{F}_q$ be a finite field of q elements, where q is a power of a prime. Let $n = q - 1$ and $\beta \in \mathbb{F}_q$ be a primitive element of $\mathbb{F}_q^* := \mathbb{F}_q \setminus \{0\}$. For an integer k such that $1 \leq k \leq n$, consider the k-dimensional vector space

$$\mathcal{L}_k := \{f \in \mathbb{F}_q[X] \mid \deg(f) \leq k-1\}$$

and the evaluation map $\mathrm{ev} : \mathcal{L}_k \to \mathbb{F}_q^n$ given by

$$\mathrm{ev}(f) := (f(\beta), f(\beta^2), \ldots, f(\beta^n)) \in \mathbb{F}_q^n.$$

Definition 8 *Given n, k, β, and q as above, the Reed-Solomon code C_k over $\mathbb{F}_q$ is*

$$C_k := \left\{(f(\beta), f(\beta^2), \ldots, f(\beta^n)) : f \in \mathcal{L}_k\right\}.$$

Proposition 2 *The Reed-Solomon code C_k is an $[q-1, k, q-k]$ code over $\mathbb{F}_q$.*

Proof

Note that C_k has length $n = q-1$ and dimension k as ev is injective. Let $c \in C_k \setminus \{0\}$. Then

$$\begin{aligned} w(c) &= n - |\{i \in \{1, \ldots, n\} : f(\beta^i) = 0\}| \\ &\geq n - \deg(f) \\ &\geq n - (k-1). \end{aligned}$$

Thus, $d \geq n+1-k$. By the Singleton Bound, we know $d \leq n+1-k$. Hence, Reed-Solomon codes are MDS codes over $\mathbb{F}_q$. Furthermore, $d = n+1-k = q-1+1-k = q-k$ since $n = q-1$. Therefore, C_k is an $[q-1, k, q-k]$ code over $\mathbb{F}_q$. $\square$

Notice that Reed-Solomon codes are short in comparison with the size of the alphabet $\mathbb{F}_q$ since the length of the codes are $n = q-1$. Reed-Solomon codes are special cases of algebraic geometry codes, which we will review in the next section.

2.2 Preliminaries

We will first introduce preliminaries to algebraic geometry before introducing AG codes and some applications. For a more thorough review, please consult [7], [17, Chapter 1], or [26]. Note that all definitions and theorems come from [26] unless otherwise noted.

Definition 9 *An algebraic function field F/K of one variable over K is an extension field $F \supseteq K$ such that F is a finite algebraic extension of $K(x)$ for some element $x \in F$ which is transcendental over K.*

We will refer to F/K as a function field.

Definition 10 *A place P of the function field F/K is a maximal ideal of some valuation ring $\mathcal{O}$ of F/K. Any element $t \in P$ such the $P = t\mathcal{O}$ is called a prime element for P. Let $\mathbb{P}_F := \{P \mid P \text{ is a place of } F/K\}$ denote the set of places F/K. Note that t is sometimes called a uniformizer.*

Definition 11 *A discrete valuation of F/K is a function $v : F \to \mathbb{Z} \cup \{\infty\}$ such that*

1. *$v(x) = \infty$ if and only if $x = 0$;*
2. *$v(xy) = v(x) + v(y)$ for all $x, y \in F$;*
3. *$v(x + y) \geq \min\{v(x), v(y)\}$ for all $x, y \in F$;*
4. *there exists an element $z \in F$ such that $v(z) = 1$;*
5. *$v(a) = 0$ for any $0 \neq a \in K$.*

Definition 12 *Let $P \in \mathbb{P}_F$ and $\mathcal{O}_P$ be the corresponding valuation ring. Let $\mathcal{O}_P^* := \{x \in \mathcal{O} \mid \text{there is a } w \in \mathcal{O} \text{ with } xw = 1\}$. Choose a prime element t for P. Then every $z \in F \setminus \{0\}$ has a unique representation $z = t^n u$ with $u \in \mathcal{O}_P^*$ and $n \in \mathbb{Z}$. Set $v_P(x) := n$ and $v_P(0) := \infty$ to define a function $v_P : F \to \mathbb{Z} \cup \{\infty\}$.*

Remarks:

1. This definition does not depend on the choice of t. If t' is another prime element for P, then $P = t\mathcal{O} = t'\mathcal{O}$. Hence, $t = t'w$ for some $w \in \mathcal{O}_P^*$. Thus, $t^n u = (t'^n w^n)u = t'^n(w^n u)$ where $w^n u \in \mathcal{O}_P^*$.
2. The function v_P is a discrete valuation of F/K.

Note that if we let $P \in \mathbb{P}_F$ and $\mathcal{O}_P$ be the corresponding valuation ring, then $F_P := \mathcal{O}_P/P$ is a field since P is a maximal ideal. The quotient field $F_P := \mathcal{O}_P/P$ is

called the residue class field of the place P. Let $x \in \mathcal{O}_P$. Define $x(P) \in F_P$ to be the residue class of x modulo P, and set $x(P) := \infty$ if $x \in F \setminus \mathcal{O}_P$. Since $K \subseteq \mathcal{O}_P$ and $K \cap P = \{0\}$, the residue class map $\mathcal{O}_P \to \mathcal{O}_P/P$ induces a canonical embedding of K into $\mathcal{O}_P/P$. Therefore, we will consider K as a subfield of $\mathcal{O}_P/P$ via this embedding.

Definition 13 *Let $P \in \mathbb{P}_F$ and $\mathcal{O}_P$ be the corresponding valuation ring. Let $F_P := \mathcal{O}_P/P$ be the residue class field of the place P. The degree of the place P is*

$$deg(P) := [F_P : K].$$

A place P of F/K of degree one is sometimes called a rational place, or a K-rational place.

Definition 14 *Let $z \in F$ and $P \in \mathbb{P}_F$. The place P is a zero of z of order m if and only if $v_P(z) = m > 0$. The place P is a pole of z of order m if and only if $v_P(z) = -m < 0$.*

Definition 15 *The free abelian group on the set of places of F/K is the divisor group of F/K, denoted by $\mathcal{D}_F$. The elements of $\mathcal{D}_F$ are called divisors of F/K.*

Notice that a divisor is a formal sum

$$D = \sum_{P \in \mathbb{P}_F} n_P P$$

where $n_P \in \mathbb{Z}$ and almost all $n_P = 0$, i.e. all but finitely many. The support of a divisor $D = \sum_{P \in \mathbb{P}_F} n_P P$ is

$$\text{supp}(D) := \{P \in \mathbb{P}_F \mid n_P \neq 0\}.$$

A divisor of the form $D = P$ with $P \in \mathbb{P}_F$ is called a prime divisor. Two divisors $D = \sum_{P \in \mathbb{P}_F} n_P P$ and $D' = \sum_{P \in \mathbb{P}_F} n'_P P$ may be added coefficientwise; that is,

$$D + D' = \sum_{P \in \mathbb{P}_F} (n_P + n'_P) P.$$

The zero element of the divisor group $\mathcal{D}_F$ is the divisor

$$0 := \sum_{P \in \mathbb{P}_F} r_P P$$

where $r_P = 0$ for all $P \in \mathbb{P}_F$.

For $Q \in \mathbb{P}_F$ and $D = \sum_{P \in \mathbb{P}_F} n_P P \in \mathcal{D}_F$, we define $v_Q(D) := n_Q$. A partial ordering on $\mathcal{D}_F$ is defined by

$$D_1 \leq D_2 \text{ if and only if } v_P(D_1) \leq v_P(D_2)$$

for any $P \in \mathbb{P}_F$. A divisor $D \geq 0$ is called positive (or effective). The degree of a divisor is defined by

$$\deg(D) := \sum_{P \in \mathbb{P}_F} v_P(D) \deg(P).$$

Since any nonzero element $x \in F$ has only finitely many zeros and poles in $\mathbb{P}_F$, we may define the following divisors related to the function $x \in F$.

Definition 16 *Let $x \in F \setminus \{0\}$. Let Z denote the set of zeros of x in $\mathbb{P}_F$, and N denote the set of poles of x in $\mathbb{P}_F$. Then define*

$$(x)_0 := \sum_{P \in Z} v_P(x) P,$$

the zero divisor of x;

$$(x)_\infty := \sum_{P \in N} -v_P(x)P,$$

the pole divisor of x; and

$$(x) := (x)_0 - (x)_\infty,$$

the principal divisor of x.

The group $\mathcal{P}_F := \{(x) \mid x \in F \setminus \{0\}\}$ is called the group of principal divisors of F/K.

Definition 17 *For divisors $A \in \mathcal{D}_F$, we write $A \sim B$ if A and B are linearly equivalent; that is, the difference between A and B is a principal divisor.*

Definition 18 *For a divisor $A \in \mathcal{D}_F$, define*

$$\mathcal{L}(A) := \{x \in F \setminus \{0\} \mid (x) \geq -A\} \cup \{0\}.$$

We sometimes call $\mathcal{L}(A)$ the Riemann-Roch space of A.

Note that if

$$A = \sum_{i=1}^{r} n_i P_i - \sum_{j=1}^{s} m_j Q_j$$

where $n_i, m_j > 0$, then $\mathcal{L}(A)$ consists of all elements $x \in F \setminus \{0\}$ such that

1. x has zeros of order $\geq m_j$ at Q_j for $j = 1, \ldots, s$, and
2. x may have poles only at the places $P_1, \ldots, P_r$, with the pole order at P_i being bounded by n_i for $i = 1, \ldots, r$, together with the 0 function.

Lemma 1 *The Riemann-Roch space of A, $\mathcal{L}(A)$, is a vector space over $K \setminus \{0\}$.*

Proof Let $x, y \in \mathcal{L}(A)$ and $a \in K$. For any place $P \in \mathbb{P}_F$,

$$v_P(x+y) \geq \min\{v_P(x), v_P(y)\} \geq -v_P(A)$$

and

$$v_p(ax) = v_P(a) + v_P(x) \geq -v_P(A).$$

Thus, $x+y, ax \in \mathcal{L}(A)$. $\square$

Define $\ell(A) := \dim_K(\mathcal{L}(A))$. Note that if $\deg(A) < 0$, then $\ell(A) = 0$.

Definition 19 *The genus g of F/K is defined by*

$$g := \max\{\deg(A) - \ell(A) + 1 \mid A \in \mathcal{D}_F\}.$$

Note that the genus of F/K is a nonnegative integer. To see this fact let $A = 0$. Then $\deg(0) - \ell(0) + 1 = 0$. Thus, $g \geq 0$.

Definition 20 *Let F/K be a function field of genus g. A divisor W of F/K is a canonical divisor if*

$$\deg(W) = 2g - 2 \text{ and } \ell(W) \geq g.$$

Theorem 1 *(Riemann-Roch Theorem) Let F/K be a function field of genus g. Let W be a canonical divisor of F/K. For any divisor $A \in \mathcal{D}_F$,*

$$\ell(A) = \deg(A) + 1 - g + \ell(W - A).$$

Lemma 2 *Let F/K be a function field of genus g. If A is a divisor of F/K of degree*

greater than or equal to $2g-1$*, then*

$$\ell(A) = \deg(A) + 1 - g.$$

Let $\mathbb{N}_0$ denote the set of nonnegative integers.

Definition 21 *For a rational place* P *of* F/K*, the Weierstrass semigroup* $H(P)$ *of* P *is*

$$H(P) := \{\alpha \in \mathbb{N}_0 \mid \textit{there exists } f \in F \textit{ such that } (f)_\infty = \alpha P\}.$$

Notice that $H(P)$ is a monoid under addition. Indeed, if $\alpha, \beta \in H(P)$, then there exists $f, g \in F$ such that $(f)_\infty = \alpha P$ and $(g)_\infty = \beta P$. Then $(fg)_\infty = (f)_\infty + (g)_\infty = \alpha P + \beta P = (\alpha + \beta)P$. Thus, $\alpha + \beta \in H(P)$.

Definition 22 *For a rational place* P *of* F/K*, the Weierstrass gap set* $G(P)$ *of* P *is*

$$G(P) := \mathbb{N}_0 \setminus H(P).$$

The elements of $G(P)$ *are often called gaps at* P *or gap numbers of* P*.*

Theorem 2 *(Weierstrass Gap Theorem) Suppose that* F/K *has genus* $g > 0$ *and* P *is a place of degree one. Then there are exactly* g *gap numbers* $\alpha_1 < \alpha_2 < \ldots < \alpha_g$ *of* P*. Moreover,*

$$\alpha_1 = 1 \textit{ and } \alpha_g = 2g - 1.$$

Definition 23 *For rational places* $P_1, P_2, \ldots, P_m$ *of* F/K*, the Weierstrass gap set* $G(P_1, P_2, \ldots, P_m)$ *of the m-tuple* $(P_1, P_2, \ldots, P_m)$ *is*

$$G(P_1, P_2, \ldots, P_m) := \mathbb{N}_0^m \setminus H(P_1, P_2, \ldots P_m).$$

Definition 24 *[4] Let $P_1, P_2, \ldots, P_m$ be rational places of F/K. An m-tuple $(\alpha_1, \alpha_2, \ldots, \alpha_m)$ of natural numbers is said to be a pure gap at $(P_1, P_2, \ldots, P_m)$ if*

$$\begin{aligned}\ell(\alpha_1 P_1 + \alpha_2 P_2 + \cdots + \alpha_m P_m) &= \ell((\alpha_1 - 1)P_1 + \alpha_2 P_2 + \cdots + \alpha_m P_m) \\ &= \ell(\alpha_1 P_1 + (\alpha_2 - 1)P_2 + \cdots + \alpha_m P_m) \\ &\vdots \\ &= \ell(\alpha_1 P_1 + \alpha_2 P_2 + \cdots + (\alpha_m - 1)P_m).\end{aligned}$$

2.3 Algebraic geometric codes

Algebraic geometric codes were introduced by V.D. Goppa in the 1970s and provided a generalization of Reed-Solomon codes ([10], [11]). Reed-Solomon codes discussed earlier are special cases of algebraic geometry codes.

Definition 25 *Let $D = P_1 + \cdots + P_n$, where $P_1, \ldots, P_n$ are pairwise distinct places of $F/\mathbb{F}_q$ of degree 1, and let A be a divisor of $F/\mathbb{F}_q$ such that $supp(D) \cap supp(A) = \emptyset$. The algebraic geometric code $C(D, A)$ associated with the divisors D and A is defined by*

$$C(D, A) = \{(f(P_1), \ldots, f(P_n)) \mid f \in \mathcal{L}(A)\} \subseteq \mathbb{F}_q^n.$$

If $|supp(A)| = m$, then $C(D, A)$ is called an m-point code; if $m \geq 2$, $C(D, A)$ is known as a multipoint code.

Theorem 3 *The code $C(D, A)$ is an $[n, k, d]$ code with parameters*

$$k = \ell(A) - \ell(A - D) \text{ and } d \geq n - \deg(A).$$

Proof

Consider the code $C(D, A)$ where $D = P_1 + \cdots + P_n$ such that $\text{supp}(D) \cap \text{supp}(A) = \emptyset$, and the evaluation map $\text{ev}_D : \mathcal{L}(A) \to \mathbb{F}_q^n$ given by

$$\text{ev}_D(f) := (f(P_1), \ldots, f(P_n)) \in \mathbb{F}_q^n.$$

The evaluation map is a surjective linear map with kernel

$$\ker(\text{ev}_D) = \{f \in \mathcal{L}(A) \mid v_{P_i}(f) > 0 \text{ for } i = 1, \ldots, n\} = \mathcal{L}(A - D).$$

Thus, $k = \ell(C(D, A)) = \ell(A) - \ell(A - D)$. Now, suppose $C(D, A) \neq 0$, and pick $f \in \mathcal{L}(A)$ such that $w(\text{ev}_D(f)) = d$. Then exactly $n - d$ places $P_{i_1}, \ldots, P_{i_{n-d}}$ in the support of D are zeros of f. Thus,

$$f \in \mathcal{L}\left(A - \left(P_{i_1}, \ldots, P_{i_{n-d}}\right)\right).$$

Hence,

$$0 \geq \deg\left(A - \left(P_{i_1}, \ldots, P_{i_{n-d}}\right)\right) = \deg(A) - (n - d).$$

Therefore, $d \geq n - \deg(A)$. $\square$

Corollary 1 *Suppose that the degree of A is strictly less than n. Then*

1. *The code $C(D, A)$ is an $[n, k, d]$ code with $k = \ell(A) \geq \deg(A) + 1 - g$. Thus,*

$$k + d \geq n + 1 - g.$$

2. *In addition, if $2g - 2 < \deg(A) < n$, then $k = \deg(A) + 1 - g$.*

3. If $\{f_1, \ldots, f_k\}$ is a basis for $\mathcal{L}(A)$, then

$$\begin{bmatrix} f_k(P_1) & f_k(P_2) & \cdots & f_k(P_n) \\ f_{k-1}(P_1) & f_{k-1}(P_2) & \cdots & f_{k-1}(P_n) \\ \vdots & \vdots & & \vdots \\ f_1(P_1) & f_1(P_2) & \cdots & f_1(P_n) \end{bmatrix}$$

is a generator matrix for $C(D, A)$.

Recall that the Singleton Bound gives $k + d \leq n + 1$. As stated in Corollary 1, the code $C(D, A)$ has parameters satisfying $n + 1 - g \leq k + d \leq n + 1$.

Example 1 *Let $F/\mathbb{F}_q$ be of genus 0. Then F is the rational function field over $\mathbb{F}_q$. Let $\mathbb{F}_q = \{0, \alpha, \alpha^2, \ldots, \alpha^{q-1}\}$. Let P_i be the rational place $x - \alpha^i$ of $\mathbb{F}_q$ for $0 \leq i \leq q-1$. Let $A = (k-1)P_\infty$ where P_∞ denotes the infinite place of $\mathbb{F}_q(x)$. Then $C(P_0 + \ldots + P_q, (k-1)P_\infty)$ is the Reed-Solomon code C_k over $\mathbb{F}_q$.*

Example 2 *Let $F = \mathbb{F}_{q^2}(x, y)$ be the function field of the Hermitian curve*

$$y^q + y = x^{q+1}$$

where q is a power of a prime. The Hermitian code over $\mathbb{F}_{q^2}$ of length q^3 is $C(D, aP_\infty)$, where

$$D = \sum_{\alpha,\beta \in \mathbb{F}_{q^2}, \beta^q + \beta = \alpha^{q+1}} P_{\alpha,\beta}$$

and $P_{\alpha,\beta}$ is a common zero of $x - \alpha$ and $y - \beta$. Note that $g = \frac{q(q-1)}{2}$.

We will now discuss several applications of AG codes. Section 2.4 will categorize stopping sets over certain curves; Chapter 4 will discuss polar coding and AG code kernels as well as triples of points of the Hermitian curve.

2.4 Stopping Sets

Combinatorial structures called stopping sets govern the performance of a linear code over the binary erasure channel when coupled with an iterative decoding algorithm.[1]

Let $C := C(H)$ be an $[n, k, d]$-code over $\mathbb{F}_q$ with parity-check matrix H. Let $[n] = \{1, \ldots, n\}$ denote the set of column indices of H.

Definition 26 *A stopping set S of the code $C(H)$ is a subset of $[n]$ such that the restriction of H to S does not contain a row of weight 1. The stopping distance $s(H)$ of $C(H)$ is the minimum size of a nonempty stopping set.*

Note that the stopping set and stopping distance depend on the choice of parity-check matrix H. We only consider the parity-check matrix H^* consisting of all the nonzero codewords of the dual code $C^\perp$.

Let F be a function field over $\mathbb{F}_q$ of genus g. Consider divisors $G = mP_\infty$, where $0 < m < n$, and $D = P_1 + \cdots + P_n$ with disjoint support, where P_i are places of F of degree 1.

The structure of an AG code reveals information about its stopping sets. In 2013, Zhang, Fu, and Wan [31] showed that dimensions of Riemann-Roch spaces can be used to decide if S is a stopping set, as shown in Figure 2.2. Notice that $A = \{i_1, \ldots, i_j\}$ is not a stopping set of $C(D, mP_\infty)^\perp$ if and only if

$$\mathrm{wt}((f(P_{i_1}), \ldots, f(P_{i_j}))) = 1$$

[1] This section is joint work with A. Omairi.

for some $f \in \mathcal{L}(mP_\infty)$ if and only if

$$f(P_{i_k}) \neq 0 \text{ and } f(P_{i_j}) = 0$$

for all $i_j \in A \setminus \{i_k\}$ if and only if

$$(f) \geq \sum_{j \in A \setminus \{i_k\}} P_j - mP_\infty \text{ and } v_{p_{i_k}}(f) = 0$$

if and only if

$$f \in \mathcal{L}\left(mP_\infty - \sum_{j \in A \setminus \{i_k\}} P_j\right) \setminus \mathcal{L}\left(mP_\infty - \sum_{j \in A} P_j\right).$$

Lemma 3 *[31, Theorem 6] Let m be a positive integer less than n. A subset $A \subseteq [n]$ is a stopping set of $C(D, mP_\infty)^\perp$ if and only if*

$$\mathcal{L}\left(mP_\infty - \sum_{j \in A} P_j\right) = \mathcal{L}\left(mP_\infty - \sum_{j \in A \setminus \{i\}} P_j\right).$$

for all $i \in A$.

Corollary 2 *[31, Corollary 7] Let m be a positive integer less than n. Consider $C(D, mP_\infty)^\perp$. Any subset of $[n]$ with cardinality greater than or equal to $m + 2$ is a stopping set of $C(D, mP_\infty)^\perp$. Any non-empty subset of $[n]$ with cardinality less than or equal to $m - 2g + 1$ is not a stopping set of $C(D, mP_\infty)^\perp$.*

Proof

Let A be subset of $[n]$ with cardinality greater than or equal to $m + 2$. Then

$\deg\left(mP_\infty - \sum_{j\in A} P_j\right) < 0$ and $\deg\left(mP_\infty - \sum_{j\in A\setminus\{i\}} P_j\right) \leq 0$. Hence,

$$\ell\left(mP_\infty - \sum_{j\in A} P_j\right) = \ell\left(mP_\infty - \sum_{j\in A\setminus\{i\}} P_j\right) = 0$$

and $\mathcal{L}\left(mP_\infty - \sum_{j\in A} P_j\right) = \mathcal{L}\left(mP_\infty - \sum_{j\in A\setminus\{i\}} P_j\right)$. Thus, A is a stopping set by Lemma 3.

Let B be non-empty subset of $[n]$ with cardinality less than or equal to $m-2g+1$. Then $\deg\left(mP_\infty - \sum_{j\in A} P_j\right) \geq 2g+1$ and $\deg\left(mP_\infty - \sum_{j\in A\setminus\{i\}} P_j\right) \geq 2g+2$. By Lemma 2, $\ell\left(mP_\infty - \sum_{j\in A} P_j\right) \neq \ell\left(mP_\infty - \sum_{j\in A\setminus\{i\}} P_j\right) = 0$ and $\mathcal{L}\left(G - \sum_{j\in A} P_j\right) \neq \mathcal{L}\left(G - \sum_{j\in A\setminus\{i\}} P_j\right)$. Thus, A is not a stopping set by Lemma 3. □

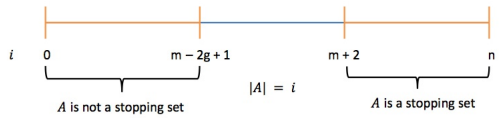

Figure 2.2: The cardinality of A plays a role in whether or not A is a stopping set of $C(D, mP_\infty)^\perp$.

Example 3 *Let m be a positive integer less than n. Consider the Reed-Solomon code C_k over $\mathbb{F}_q$. Then $g = 0$ and $m = n-k-1$. Using Corollary 2, we can categorized all the stopping sets. Any subset of $[n]$ with cardinality greater than or equal to $n-k+1$ is a stopping set. Any non-empty subset of $[n]$ with cardinality less than or equal to $n-k$ is not a stopping set. Hence, the stopping distance is $n-k+1$.*

Let E be an elliptic function field over $\mathbb{F}_q$ with rational place O. Endow $E(\mathbb{F}_q)$ with a group structure and the zero element O. Let $\{P_1, P_2, \ldots, P_n\}$ be a subset of

the set $E(\mathbb{F}_q)$ and $D = P_1 + \cdots + P_n$ such that the supports of D and O are disjoint. Let m be a positive integer less than n. Note $g = 1$. Using Corollary 2, we know any subset of $[n]$ with cardinality greater than or equal to m is a stopping set $C(D, mO)^{\perp}$, and any non-empty subset of $[n]$ with cardinality less than or equal to $m+1$ is not a stopping set $C(D, mO)^{\perp}$.

Using Lemma 4, one can classify stopping sets of size $m+1$ and m in Corollary 3.

Lemma 4 *[31, Corollary 9] Let A be a subset of $[n]$. If*

$$K - G + \sum_{j \in A} P_j \sim E$$

for some effective divisor E with $supp(E) \cap \{P_i \mid i \in A\} = \emptyset$ where K is a canonical divisor on X, then A is stopping set.

Corollary 3 *[31, Theorem 10] Let m be a positive integer less than n. Consider $C(D, mO)^{\perp}$. Any $A \subseteq [n]$ with cardinality $m+1$ is a stopping set if and only if for all $i \in A$,*

$$\sum_{j \in A \ \setminus \{i\}} P_j \neq O.$$

Any $A \subseteq [n]$ with cardinality m is a stopping set if and only if

$$\sum_{j \in A} P_j = O.$$

.

Similar techniques can be used to classify stopping sets of AG codes over hyperelliptic function fields of genus 2. Let E be an hyperelliptic function field of

genus 2 over $\mathbb{F}_q$ with rational place O. Endow $E(\mathbb{F}_q)$ with a group structure and the zero element O. We will consider $C(D, mO)^{\perp}$.

Corollary 4 *Let m be a positive integer less than n. Consider $C(D, mO)^{\perp}$. Any subset of $[n]$ with cardinality greater than or equal to $m+2$ is a stopping set. Any non-empty subset of $[n]$ with cardinality less than or equal to $m-3$ is not a stopping set.*

Proof

From Corollary 2, any subset of $[n]$ with cardinality greater than or equal to $m+2$ is a stopping set. Since $g = 2$, any non-empty subset of $[n]$ with cardinality less than or equal to $m - 2(2) + 1 = m - 3$ is not a stopping set by Corollary 2.

Proof

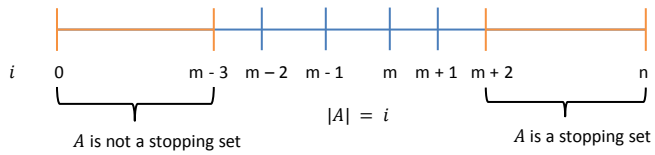

Figure 2.3: Stopping sets of hyperelliptic function fields with genus 2

Corollary 5 *Let m be a positive integer less than n. Consider $C(D, mO)^{\perp}$. Any $A \subseteq [n]$ with cardinality $m+1$ is a stopping set if and only if for all $i \in A$,*

$$\sum_{j \in A \setminus \{i\}} P_j \neq O.$$

Any $A \subseteq [n]$ with cardinality $m-2$ is a stopping set if and only if

$$\sum_{j \in A} P_j = O.$$

.

Proof

Suppose $|A| = m+1$ and that A is not a stopping set. Then by Lemma 3 there exists some $i \in A$ such that

$$\mathcal{L}\left(mO - \sum_{j \in A \setminus \{i\}} P_j\right) \neq \mathcal{L}\left(mO - \sum_{j \in A} P_j\right).$$

Hence, there exists some nonzero $f \in \mathcal{L}(mO)$ such that

$$f \in \mathcal{L}\left(mO - \sum_{j \in A \setminus \{i\}} P_j\right) \setminus \mathcal{L}\left(mO - \sum_{j \in A} P_j\right).$$

Since $|A| = m+1$,

$$\deg\left(mO - \sum_{j \in A \setminus \{i\}} P_j\right) = m - m = 0$$

and

$$(f) \geq -mO + \sum_{j \in A \setminus \{i\}} P_j.$$

Since $\deg((f)) = 0$ and $\deg\left(-mO + \sum_{j \in A \setminus \{i\}} P_j\right) = 0$,

$$(f) = -mO + \sum_{j \in A \setminus \{i\}} P_j = \sum_{j \in A \setminus \{i\}} (P_j - O).$$

Hence, $(f) = O$ since $f \neq 0$. So

$$\sum_{j \in A \setminus \{i\}} P_j = O.$$

Now, suppose $|A| = m - 2$ and that A is a stopping set. Then by Lemma 3, for all $i \in A$,

$$\mathcal{L}\left(mO - \sum_{j \in A \setminus \{i\}} P_j\right) = \mathcal{L}\left(mO - \sum_{j \in A} P_j\right).$$

Notice that since $g = 2$,

$$\deg\left(mO - \sum_{j \in A \setminus \{i\}} P_j\right) = m - (m-3) = 3 = 2g - 1.$$

Then by the Riemann-Roch Theorem, there exists some nonzero f such that

$$f \in \mathcal{L}\left(mO - \sum_{j \in A \setminus \{i\}} P_j\right) = \mathcal{L}\left(mO - \sum_{j \in A} P_j\right),$$

so

$$(f) = mO - \sum_{j \in A} P_j == \sum_{j \in A \setminus \{i\}} (O - P_j).$$

Then $(f) = O$ since $f \neq 0$. Thus,

$$\sum_{j \in A} P_j = O.$$

Conversely, pick A such that $\sum_{j \in A} P_j = O$. Since $2O$ is a canonical divisor, we have

$$2O - mO + \sum_{j \in A} P_j \sim 2O.$$

By Lemma 4, since $2O$ is an effective divisor, A is a stopping set. $\square$

Note it is still undetermined if there are stopping sets of size $m-1$ and m. This remains a topic of further investigation.

We now consider stopping sets of algebraic geometric codes from function fields with larger genus. Let $F = \mathbb{F}_{q^2}(x,y)$ be the function field of the Hermitian curve $y^q + y = x^{q+1}$ where q is a power of a prime. Recall, the Hermitian code over $\mathbb{F}_{q^2}$ of length q^3 is $C(D, mP_\infty)$, where m is a positive integer, $D = \sum_{\alpha,\beta \in \mathbb{F}_{q^2}, \beta^q+\beta=\alpha^{q+1}} P_{\alpha,\beta}$ and $P_{\alpha,\beta}$ is a common zero of $x-\alpha$ and $y-\beta$. Note

$$g = \frac{q(q-1)}{2}.$$

Corollary 6 *Let m be a positive integer less than n. Consider the Hermitian code $C(D, mP_\infty)^\perp$ over $\mathbb{F}_{q^2}$ Any subset of $[n]$ with cardinality greater than or equal to $m+2$ is a stopping set of $C(D, mP_\infty)^\perp$. Any non-empty subset of $[n]$ with cardinality less than or equal to $m - q^2 + q + 2$ is not a stopping set of $C(D, mP_\infty)^\perp$.*

Proof

By Corollary 2, any subset of $[n]$ with cardinality greater than or equal to $m+2$ is a stopping set of $C(D, mP_\infty)^\perp$. Since $g = \frac{q(q-1)}{2}$, any non-empty subset of $[n]$ with cardinality less than or equal to $m - 2\frac{q(q-1)}{2} + 1 = m - q^2 + q + 2$ is not a stopping set of $C(D, mP_\infty)^\perp$ by Corollary 2. $\square$

We say A is a collinear set if the columns of the parity check matrix indexed by A are represented by

$$[\cup_{i \in \{1,\ldots,b\}} \{P_{\alpha_i,\beta} \mid \beta \in K_{\alpha_i}\}] \cup \{P_1, \ldots, P_t\}$$

where b is a nonnegative integer, $P_j \in \{P_{\alpha,\beta} \in \mathbb{P}_F \mid \beta \in K_\alpha\}$ for some $\alpha \in \mathbb{F}_{q^2}$, and $K_\alpha := \{\beta \in \mathbb{F}_{q^2} \mid \beta^q + \beta = \alpha^{q+1}\}$. The follow result allows us to determine if there exists a collinear stopping set for size i for $m - q^2 + q + 2 \leq i \leq m + 1$.

Theorem 4 *[15, Theorem 3.6] Consider the Hermitian function field H over $\mathbb{F}_{q^2}$, and the divisor*

$$S := rP_\infty + \sum_{\beta \in K_\alpha} k_\beta P_{\alpha,\beta}$$

of H where $\alpha \in \mathbb{F}_{q^2}, r \in \mathbb{Z}$, and $k_\beta \in \mathbb{Z}$ for each $\beta \in K_\alpha := \{\beta \mid \beta^q + \beta = \alpha^{q+1}\}$. The dimension of the space $\mathcal{L}(S)$ is given by

$$\ell(S) = \sum_{i=0}^{q} \max\left\{ \left\lfloor \frac{r - iq}{q+1} \right\rfloor + \sum_{\beta \in K_\alpha} \left\lfloor \frac{k_\beta + i}{q+1} \right\rfloor + 1, 0 \right\}.$$

Theorem 5 *Let b and r be nonnegative integers with $0 \leq r \leq q-1$. Let m be a positive integer less than n. Consider the Hermitian code $C(D, mP_\infty)^\perp$ over $\mathbb{F}_{q^2}$. There exists a collinear stopping set of size $bq+t \leq m$ if and only if $m - bq < q^2 - q - 1$ and either $t \geq \lfloor \frac{m-bq}{q+1} \rfloor + 2$ or $t = 0$.*

Proof

Suppose A is the set such that the columns of the parity matrix indexed by A are represented by

$$[\cup_{i \in \{1,\ldots,b\}} \{P_{\alpha_i,\beta} \mid \beta \in K_{\alpha_i}\}] \cup \{P_1, \ldots, P_t\}$$

where $P_j \in \{P_{\alpha,\beta} \in \mathbb{P}_F \mid \beta \in K_\alpha\}$ for some $\alpha \in \mathbb{F}_{q^2}$ and $K_\alpha := \{\beta \in \mathbb{F}_{q^2} \mid \beta^q + \beta = \alpha^{q+1}\}$. Let A' denote the set such that the columns of the parity check matrix indexed by A' are represented by $[\cup_{i \in \{1,\ldots,b\}} \{P_{\alpha_i,\beta} \mid \beta \in K_{\alpha_i}\}]$, and let A'' denote the set such that the columns of the parity check matrix indexed by A'' are represented by $\{P_1, \ldots, P_t\}$.

Let $f \in \mathcal{L}\left(mP_\infty - \sum_{j\in A} P_j\right) = \mathcal{L}\left(mP_\infty - \sum_{j\in A'} P_j - \sum_{j\in A''} P_j\right)$. Then

$$(f) \geq \sum_{j\in A} P_j - mP_\infty = \sum_{j\in A'} P_j + \sum_{j\in A''} P_j - mP_\infty.$$

Hence,

$$\left(f \prod_i \frac{1}{x-\alpha_i}\right) \geq \sum_{j\in A''} P_j - (m-bq)P_\infty.$$

Therefore,

$$\mathcal{L}\left(mP_\infty - \sum_{j\in A'} P_j - \sum_{j\in A''} P_j\right) \cong \mathcal{L}\left((m-bq)P_\infty - \sum_{j\in A''} P_k\right).$$

Recall from Lemma 3 that A is a stopping set if

$$\mathcal{L}\left(mP_\infty - \sum_{j\in A} P_j\right) \cong \mathcal{L}\left(mP_\infty - \sum_{j\in A\setminus\{k\}} P_j\right)$$

for all $k \in A$. Thus, we would like to show

$$\mathcal{L}\left((m-bq)P_\infty - \sum_{j\in A''} P_k\right) \cong \mathcal{L}\left(mP_\infty - \sum_{j\in A\setminus\{k\}} P_j\right)$$

for all $k \in A$.

Case 1: Suppose $k \in A'$. Note if $t = 0$, then this is the only case we need to check. Let $f \in \mathcal{L}\left(mP_\infty - \sum_{j\in A\setminus\{k\}} P_j\right)$. Then

$$(f) \geq \sum_{j\in A\setminus\{k\}} P_j - mP_\infty.$$

Thus,

$$\left(f\prod_i \frac{1}{x-\alpha_i}\right) \geq \sum_{j\in A''} P_j - P_k - (m-bq)P_\infty.$$

Then

$$\mathcal{L}\left(mP_\infty - \sum_{j\in A\setminus\{k\}} P_j\right) \cong \mathcal{L}\left(P_k + (m-bq)P_\infty - \sum_{j\in A''} P_j\right).$$

Note $\ell\left((m-bq)P_\infty - \sum_{j\in A''} P_k\right) = \ell\left(P_k + (m-bq)P_\infty - \sum_{j\in A''} P_j\right)$ if and only if $(1,\beta)\in H(P_k,P_\infty)$, which holds if and only if $\beta \leq m-bq < q^2-q-1$ [9].

Case 2: Suppose $k\in A''$. By Theorem 4,

$$\begin{aligned}\ell\left((m-bq)P_\infty - \sum_{j\in A''} P_k\right) &= \sum_{i=0}^{q} \max\left\{\left\lfloor\frac{m-bq-iq}{q+1}\right\rfloor + \sum_{j\in A''}\left\lfloor\frac{-1+i}{q+1}\right\rfloor + \sum_{j\notin A''}\left\lfloor\frac{0+i}{q+1}\right\rfloor + 1, 0\right\}\\ &= \max\left\{\left\lfloor\frac{m-bq}{q+1}\right\rfloor - t + 1, 0\right\} + \sum_{i=1}^{q}\max\left\{\left\lfloor\frac{m-bq-iq}{q+1}\right\rfloor + 1, 0\right\}.\end{aligned}$$

Also, by Theorem 4,

$$\begin{aligned}&\ell\left((m-bq)P_\infty - \sum_{j\in A''\setminus\{k\}} P_k\right)\\ &= \sum_{i=0}^{q} \max\left\{\left\lfloor\frac{m-bq-iq}{q+1}\right\rfloor + \sum_{j\in A''\setminus\{k\}}\left\lfloor\frac{-1+i}{q+1}\right\rfloor + \sum_{j\notin A''\setminus\{k\}}\left\lfloor\frac{0+i}{q+1}\right\rfloor + 1, 0\right\}\\ &= \max\left\{\left\lfloor\frac{m-bq}{q+1}\right\rfloor - t + 2, 0\right\} + \sum_{i=1}^{q}\max\left\{\left\lfloor\frac{m-bq-iq}{q+1}\right\rfloor + 1, 0\right\}.\end{aligned}$$

Thus, $\ell\left((m-bq)P_\infty - \sum_{j\in A''} P_k\right) = \ell\left((m-bq)P_\infty - \sum_{j\in A''\setminus\{k\}} P_k\right)$ if and only if

$$\max\left\{\left\lfloor\frac{m-bq}{q+1}\right\rfloor - t + 1, 0\right\} + \sum_{i=1}^{q}\max\left\{\left\lfloor\frac{m-bq-iq}{q+1}\right\rfloor + 1, 0\right\}$$
$$= \max\left\{\left\lfloor\frac{m-bq}{q+1}\right\rfloor - t + 2, 0\right\} + \sum_{i=1}^{q}\max\left\{\left\lfloor\frac{m-bq-iq}{q+1}\right\rfloor + 1, 0\right\},$$

which holds if and only

$$\left\lfloor\frac{m-bq}{q+1}\right\rfloor - t + 2 \leq 0.$$

Therefore, A is stopping set if and only if $m - bq < q^2 - q - 1$ and $t \geq \lfloor\frac{m-bq}{q+1}\rfloor + 2$. □

Example 4 *Let $q = 4$ and $m = 40$. We know A is not a stopping set if $|A| \leq 29$ and A is a stopping set if $|A| \geq 42$ as shown in Figure 2.4. Thus, it remains to consider index sets A if $30 \leq |A| \leq 41$. Table 2.1 shows the existence of collinear stopping sets as guaranteed by Theorem 5. Table 2.2 shows the existence of non-collinear stopping sets for $|A| = 37$ and 41 using SAGE.*

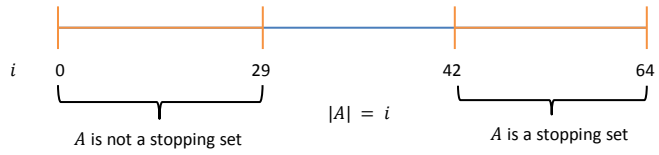

Figure 2.4: Hermitian Example: $q = 4$ and $m = 40$

Example 5 *Let $q = 5$ and $m = 80$. We know A is not a stopping set if $|A| \leq 61$ and A is a stopping set if $|A| \geq 82$ as shown in Figure 2.5. Table 2.3 shows the existence*

i	r	s	t	Existence of a collinear stopping set of size i
30	12	7	2	No collinear stopping set
31	12	7	3	No collinear stopping set
32	**8**	**8**	**0**	**Collinear stopping set**
33	8	8	1	No collinear stopping set
34	8	8	2	No collinear stopping set
35	**8**	**8**	**3**	**Collinear stopping set**
36	**8**	**8**	**4**	**Collinear stopping set**
37	4	9	1	No collinear stopping set
38	**4**	**9**	**2**	**Collinear stopping set**
39	**4**	**9**	**3**	**Collinear stopping set**
40	**4**	**9**	**2**	**Collinear stopping set**
41	0	10	1	No collinear stopping set

Table 2.1: Hermitian Example ($q = 4$ and $m = 40$): Collinear stopping sets

of a collinear stopping set as guaranteed by Theorem 5. Table 2.4 shows the existence of non-collinear stopping sets for $|A| = 68, 71, 72, 76,$ *and* 81 *using SAGE.*

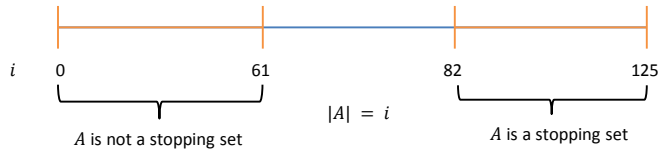

Figure 2.5: Hermitian Example: $q = 5$ and $m = 80$

i	Existence of a stopping set of size i
30	Undetermined
31	Undetermined
32	Yes
33	Undetermined
34	Undetermined
35	Yes
36	Yes
37	**Yes (Non-collinear)**
38	Yes
39	Yes
40	Yes
41	**Yes (Non-collinear)**

Table 2.2: Hermitian Example ($q = 4$ and $m = 40$): Non-collinear stopping sets

i	r	s	t	Existence of a collinear stopping set
62	20	12	2	No collinear stopping set
63	20	12	3	No collinear stopping set
64	20	12	4	No collinear stopping set
65	**15**	**13**	**0**	**Collinear stopping set**
66	15	13	1	No collinear stopping set
67	15	13	2	No collinear stopping set
68	15	13	3	No collinear stopping set
69	**15**	**13**	**4**	**Collinear stopping set**
70	**15**	**13**	**5**	**Collinear stopping set**
71	10	14	1	No collinear stopping set
72	10	14	2	No collinear stopping set
73	**10**	**14**	**3**	**Collinear stopping set**
74	**10**	**14**	**4**	**Collinear stopping set**
75	**10**	**14**	**5**	**Collinear stopping set**
76	5	15	1	No collinear stopping set
77	**5**	**15**	**2**	**Collinear stopping set**
78	**5**	**15**	**3**	**Collinear stopping set**
79	**5**	**15**	**4**	**Collinear stopping set**
80	**5**	**15**	**5**	**Collinear stopping set**
81	0	16	1	No collinear stopping set

Table 2.3: Hermitian Example ($q = 5$ and $m = 80$): Collinear stopping sets

i	Existence of a stopping set
62	Undetermined
63	Undetermined
64	Undetermined
65	Yes
66	Undetermined
67	Undetermined
68	**Yes (Non-collinear)**
69	Yes
70	Yes
71	**Yes (Non-collinear)**
72	**Yes (Non-collinear)**
73	Yes
74	Yes
75	Yes
76	**Yes (Non-collinear)**
77	Yes
78	Yes
79	Yes
80	Yes
81	**Yes (Non-collinear)**

Table 2.4: Hermitian Example ($q = 5$ and $m = 80$): Non-collinear stopping sets

Chapter 3

Polar coding

In this chapter, we will give an introduction to polar codes, which were constructed by Arikan [1] in 2009 as the first explicit construction of symmetric capacity achieving codes for binary discrete memoryless channels with low encoding and decoding complexity. First, we will give some background to the problem and Shannon's theorem.

Let $W : \mathcal{X} \to \mathcal{Y}$ be a discrete channel. The input alphabet of W is $\mathcal{X}$, and the output alphabet of W is $\mathcal{Y}$, both of which are finite. The channel W is defined by transition probabilities

$$W(y|x)$$

for all $x \in \mathcal{X}$ and $y \in \mathcal{Y}$. Let $W_{\mathcal{X}}(x)$ be the probability that $x \in \mathcal{X}$ is sent across W.

Definition 27 *A channel is memoryless if for consecutive uses of the channel with inputs $x_1, x_2, \ldots, x_n$, the probability of obtaining output $y_1, y_2, \ldots, y_n$ is*

$$W(y_1, \ldots, y_n|x_1, \ldots, x_n) = W(y_1|x_1)W(y_2|x_2)\cdots W(y_n|x_n).$$

Definition 28 *The channel transition matrix is the matrix A whose ij^{th} entry is*

$W(y_i|x_j)$. *A channel is said to be symmetric if all rows of its channel transition matrix are permutations of each other, and all columns are permutations of each other.*

Example 6 *The binary symmetric channel (BSC) with crossover probability* $1-p$ *is defined by the transition probabilities*

$$W(0|0) = p,$$
$$W(0|1) = 1-p,$$
$$W(1|1) = p,$$

and

$$W(1|0) = 1-p$$

where $1/2 \leq p \leq 1$. *It is often convenient to express* W *as in Figure 3.1.*

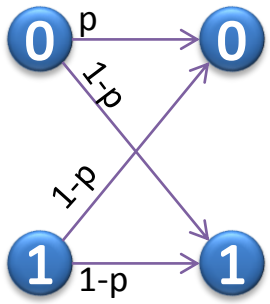

Figure 3.1: Binary symmetric channel

Definition 29 *Let* $W : \mathcal{X} \to \mathcal{Y}$ *be a discrete memoryless channel (DMC). The rate*

of the channel W is defined as

$$I(\mathcal{X};\mathcal{Y}) = \sum_{y\in\mathcal{Y}}\sum_{x\in\mathcal{X}} W(y|x)W_{\mathcal{X}}(x)\log_q\left(\frac{W(y|x)}{W_{\mathcal{X}}(x)}\right),$$

and the channel capacity of the channel W is defined as

$$C = \max_{W_{\mathcal{X}}} I(\mathcal{X};\mathcal{Y}).$$

Example 7 *For the BSC with $p = 0.95$ such that $W_{\mathcal{X}}(0) = W_{\mathcal{X}}(1) = 0.50$, we have*

$$I(\mathcal{X};\mathcal{Y}) = 2(0.50)(0.95)\log_2\left(\frac{0.95}{0.50}\right) + 2(0.50)(0.05)\log_2\left(\frac{0.05}{0.50}\right) = 0.7136.$$

Theorem 6 *(Shannon's Theorem) Suppose W has a channel capacity $C > 0$. Then for every $\varepsilon > 0$ and $R < C$, for every large N, there exists a block code of length N and rate $R_N \geq R$ such that there exists a decoding algorithm with probability of decoding error less than ε.*

For many years, a primary goal of coding theorists has been to find an explicit construction of capacity achieving codes. Coding theorists have searched for powerful codes and efficient implementations via algebraic geometric (AG) codes, low-density parity-check codes, and random codes as well as polar codes. In 2009, polar codes were developed by Arikan [1] as an explicit construction of symmetric capacity achieving codes for binary discrete memoryless channels with low encoding and decoding complexity. Arikan employs the n^{th} Kronecker power of the matrix

$$G_2 := \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \in \mathbb{F}_2^{2\times 2}$$

for encoding a block of 2^n symbols where n is taken to be a positive integer. The matrix G_2 is called the kernel matrix or kernel for short. Synthetic channels are created so that as the number of channels grows, meaning as n goes to infinity, each channel becomes either a noiseless channel or a pure-noise channel. This will be discussed more precisely in the next section, where we formally introduce polar coding.

We end this introduction with an explanation of notation to be used. Recall that given $u \in \mathbb{F}_q^n$, u_i denotes the i^{th} coordinate of u. For $1 \leq i \leq j \leq n$, it is often convenient to write $u_i^j := (u_i, \ldots, u_j) \in \mathbb{F}_q^{j-i+1}$. For two vectors $u, w \in \mathbb{F}_2^n$, $u \oplus w$ denotes the componentwise sum in $\mathbb{F}_2^n$.

Given an $m \times n$ matrix A with entries in a field $\mathbb{F}$, A_{ij} denotes the entry of A in the i^{th} row and j^{th} column, and $\text{Row}_i A$ denotes the i^{th} row of A; here, i is referred to as the row index. The j^{th} column of A is denoted by $\text{Col}_j A$. A $n \times n$ matrix L is lower triangular if $L_{ij} = 0$ for $i < j$. A $n \times n$ matrix U is upper triangular if $U_{ij} = 0$ for $i > j$.

The Kronecker product, denoted by $\otimes$, of two matrices $A \in \mathbb{F}^{n \times m}$ and $B \in \mathbb{F}^{p \times k}$ is defined as

$$A \otimes B = \begin{bmatrix} A_{11}B & \ldots & A_{1n}B \\ \vdots & \ldots & \vdots \\ A_{m1}B & \ldots & A_{mn}B \end{bmatrix} \in \mathbb{F}_q^{np \times mk}$$

Let $A^{\otimes n} := \underbrace{A \otimes \cdots \otimes A}_{n}$ for $n \geq 1$ and $A^{\otimes 0} := [1]$, a 1×1 matrix.

3.1 Channel polarization

In this section, we review polar coding over a q-ary discrete memoryless channel (DMC). Throughout this section, let q be a prime power, $\mathcal{X} := \mathbb{F}_q$, and $W : \mathcal{X} \to \mathcal{Y}$ be a q-ary DMC with transition probabilities $W(y|x)$ for all $x \in \mathcal{X}$ and $y \in \mathcal{Y}$. Two important quantities associated with the channel W are the capacity and the Bhattacharyya parameter; the standard definitions are as follows. The Bhattacharyya distance between $x, x' \in \mathcal{X}$ is

$$Z_{x,x'} = \sum_{y \in \mathcal{Y}} \sqrt{W(y|x)W(y|x')}.$$

Let

$$I(W) := I(\mathcal{X}; \mathcal{Y}) \tag{3.1}$$

when $W_{\mathcal{X}}(x) = \frac{1}{q}$ for all $x \in \mathcal{X}$. The reliability of W is described by the Bhattacharyya parameter

$$Z(W) = \frac{1}{q(q-1)} \sum_{x,x' \in \mathcal{X}, x \neq x'} Z_{x,x'}(W)$$

of W. By [1, Proposition 1], we know for a binary-DMC W,

$$I(W) \geq \log_2 \left(\frac{2}{1 + Z(W)} \right)$$

and

$$I(W) \leq \sqrt{1 + Z(W)^2}.$$

Thus, using the bounds above we could expect $I(W) \approx 0$ if and only if $Z(W) \approx 1$, and $I(W) \approx 1$ if and only if $Z(W) \approx 0$.

Let G be a $l \times l$ matrix over $\mathbb{F}_q$. Set $N = l^n$ for some positive integer n. Let

R_N be the $N \times N$ matrix with columns given by

$$\mathrm{Col}_{il^{n-1}+j}[R_N] = e_{(j-1)l+(i+1)}$$

for $0 \leq i \leq l^{n-1}$ and $1 \leq j \leq l^{n-1}$, where e_k is the standard basis vector of length N whose only nonzero entry is a one in the k^{th} coordinate; that is, R_N is the $N \times N$ permutation matrix such that

$$[u_1, u_2, \ldots, u_N]R_N = [u_1, u_{l+1}, \ldots, u_{N-(l-1)}, u_2, u_{l+2}, \ldots, u_{N-(l-2)}, \ldots, u_l, u_{2l}, \ldots, u_N],$$

where $[u_1, u_2, \ldots, u_N] \in \mathbb{F}_q^{N \times N}$.

A block of $N = l^n$ channels is produced from the channel W by combining and splitting channels as we describe now. Let W^N denote N independent uses of the channel W. To begin, N independent copies of W are combined to form the channel $W_N : \mathcal{X}^N \to \mathcal{Y}^N$ with transition probabilities

$$W_N(y_1^N|u_1^N) = W^N(y_1^N|(u_1^N)^T(B_N G^{\otimes n})) = \prod_{i=1}^{N} W(y_i|((u_1^N)^T B_N G^{\otimes n})_i)$$

where B_N is defined recursively by

$$B_N = R_N(I_l \otimes B_{N/l})$$

and $B_l = I_l$, the $l \times l$ identity matrix.

Next, the channel W_N is split into N channels $W_N^{(i)} : \mathcal{X} \to \mathcal{Y}^N \times \mathcal{X}^{i-1}$, $1 \leq i \leq N$, which are defined by the transition probabilities

$$W_N^{(i)}(y_1^N, u_1^{i-1}|u_i) := \sum_{u_{i+1}^N \in \mathcal{X}^{N-i}} \frac{1}{q^{l^n-1}} W_N(y_1^N|u_1^N).$$

This gives rise to the N channels

$$W_N^{(1)} : \mathcal{X} \to \mathcal{Y}^N$$
$$W_N^{(2)} : \mathcal{X} \to \mathcal{Y}^N \times \mathcal{X}$$
$$W_N^{(3)} : \mathcal{X} \to \mathcal{Y}^N \times \mathcal{X}^2$$
$$\vdots$$
$$W_N^{(N)} : \mathcal{X} \to \mathcal{Y}^N \times \mathcal{X}^{N-1}.$$

The next example illustrates this.

Example 8 *Taking $q = 2$ and $G = G_2$ yields Arikan's original construction. Set $W_1 := W$ for a binary discrete memoryless channel W. At the first step of the recursion, consider $W_2 : \mathcal{X}^2 \to \mathcal{Y}^2$ by the transition probabilities*

$$W_2(y_1, y_2|u_1, u_2) := W(y_1|u_1 \oplus u_2)W(y_2|u_2).$$

We may then split W_2 into two new channels $W_2^{(1)}$ and $W_2^{(2)}$. Then the channel

$$W_2^{(1)} : \mathcal{X} \to \mathcal{Y}^2$$

has transition probabilities

$$W_2^{(1)}(y_1^2|u_1) := \frac{1}{2} \sum_{u_2 \in \mathcal{X}} W_2(y_1^2|u_1^2),$$

and the channel

$$W_2^{(2)} : \mathcal{X} \to \mathcal{Y}^2 \times \mathcal{X}$$

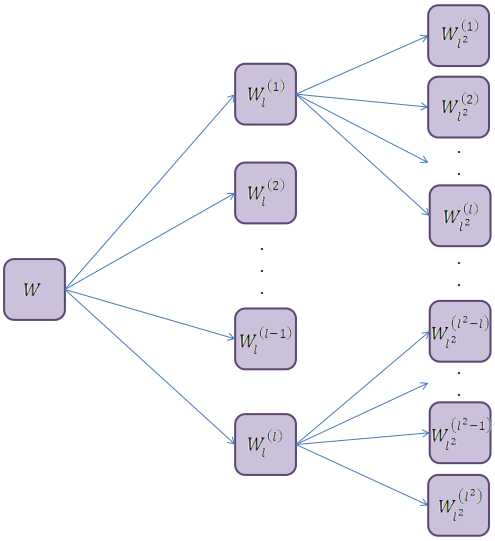

Figure 3.2: Polarization with $l \times l$ kernel matrix

has transition probabilities

$$W_2^{(2)}(y_1^2, u_1|u_2) := \frac{1}{2}W_2(y_1^2|u_1^2).$$

As we discuss below, the properties of the channels $W_N^{(i)}$ depend on the kernel matrix G and on the size of the input alphabet, q.

Theorem 7 *[1, Theorem 1] Let W be a binary DMC. For any fixed $\delta \in (0,1)$, as N goes to infinity the fraction of channels in the set $\left\{W_N^{(i)} : 1 \leq i \leq N\right\}$ such that*

$$I(W_N^{(i)}) \in (1-\delta, 1]$$

approaches $I(W)$ and the fraction of channels in the set $\left\{W_N^{(i)} : 1 \leq i \leq N\right\}$ such that

$$I(W_N^{(i)}) \in [0, \delta)$$

approaches $1 - I(W)$.

This phenomenon is known as polarization. We say that a kernel matrix polarizes a channel W if the behavior described in Theorem 7 holds.

Example 9 *Figure 3.3 shows channel polarization for the BSC W with $p = .95$. Recall from Example 7, $I(W) = 0.7136$. Using G_2, after the first step of channel combining and splitting, we create two new channels $W_2^{(1)}$ and $W_2^{(2)}$. For these new channels,*

$$I(W_2^{(1)}) = 0.5866 \text{ and } I(W_2^{(2)}) = 0.1900.$$

Again using G_2, we create four new channels with

$$I(W_4^{(1)}) = 0.7556,$$
$$I(W_4^{(2)}) = 0.3441,$$
$$I(W_4^{(3)}) = 0.3021,$$

and

$$I(W_4^{(4)}) = 0.0361.$$

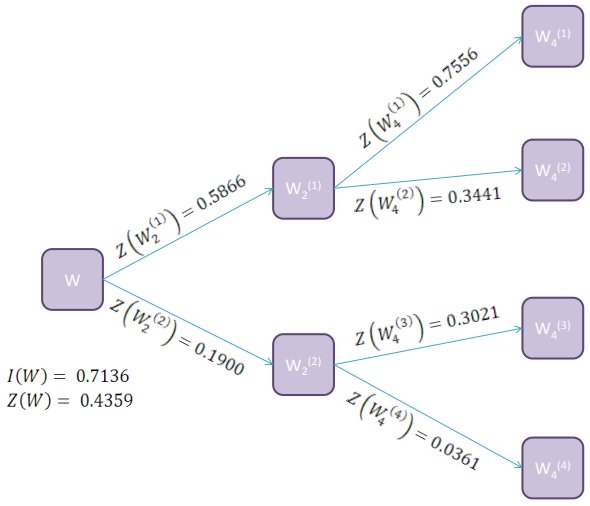

Figure 3.3: Polarization with BSC with $p = .95$

The rate of polarization of a kernel is known as the exponent and its definition is below. In preparation, consider a sequence of independent and identically distributed (i.i.d.) random variables $\{B_n \mid n \geq 1\}$ uniformly distributed over the set $\{1, \ldots, N\}$. Let $Z_n := Z(W'_n)$ where the channels W'_i, $i \geq 0$, are defined recursively by

$$W'_0 = W, \text{ and } W'_{n+1} = (W'_n)_N^{(B_{n+1})}.$$

Definition 30 *[19, Theorem 19] Let W be a q-ary DMC with $0 < I(W) < 1$, and consider an $l \times l$ matrix G with entries in $\mathbb{F}_q$. The exponent of G is the value $E(G)$ such that for any fixed $\beta < E(G)$,*

$$\liminf_{n\to\infty} \Pr[Z_n \leq 2^{-l^{n\beta}}] = I(W),$$

and for any fixed $\beta > E(G)$,

$$\liminf_{n\to\infty} \Pr[Z_n \geq 2^{-l^{n\beta}}] = 1.$$

The exponent $E(G)$ is also called the rate of polarization of G.

It follows that for any fixed rate $0 < R < I(W)$ and $0 < \beta < E(G)$, there exists a sequence $\{\mathcal{A}_N\}$ of sets $\mathcal{A}_N \subseteq \{1, \ldots, N\}$ such that $|\mathcal{A}_N| \geq NR$ and

$$\sum_{i \in \mathcal{A}_N} Z(W_N^{(i)}) = o(2^{-l^{n\beta}}).$$

Definition 31 *[1] Let W be a DMC. Fix N, the code block length, and K, the number of information bits. For (W, N, K), choose an information set $\mathcal{A}_N \subseteq \{1, \ldots, N\}$ such that $|\mathcal{A}_N| = K$ and*

$$Z(W_N^{(i)}) \leq Z(W_N^{(i')}) \quad \textit{for all } i \in \mathcal{A}_N, i' \in \mathcal{A}_N^c.$$

A polar code is defined by $(W, N, K, \mathcal{A}_N, u_{\mathcal{A}_N^c})$, where $u_{\mathcal{A}_N^c}$ are considered the frozen bits.

Example 10 *Return to example 9 using G_2 for the BSC W with $p = .95$. If $K = 2$, then $\mathcal{A}_4 = \{3, 4\}$.*

We encode $u_1^N \in \mathcal{X}^N$ for $N = l^n$ as $x_1^N := u_1^N B_N G^{\otimes n}$ where $B_N = R_N(I_l \otimes B_{N/l})$. Note that $u_1^N = u_{\mathcal{A}} \oplus u_{\mathcal{A}^c}$, where $u_{\mathcal{A}^c}$ are the frozen bits.

Example 11 *Let $N = 8$, and consider using G_2 as the kernel matrix over $\mathbb{F}_2$. Then*

$$G_2^{\otimes 3} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix},$$

$$B_8 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

and

$$B_8 G_2^{\otimes 8} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

Figure 3.4 demonstrates encoding 8*-bits.*

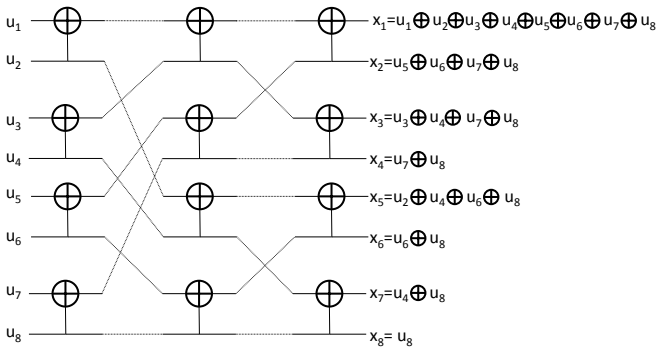

Figure 3.4: 8-bit encoding diagram

We decode channel output y_1^N using a successive cancellation (SC) decoder. The SC decoder observes y_1^N and the frozen bits $u_{\mathcal{A}^c}$ and will output $\hat{u}_1^N$. If $i \in \mathcal{A}^c$, then u_i is a frozen bit and the decoder sets

$$\hat{u}_i := u_i.$$

If $i \notin \mathcal{A}^c$, then the decoder receives the previous decisions $\hat{u}_1^{i-1}$ and computes

$$L_N^{(i)}(y_1^N, \hat{u}_1^{i-1}) := \frac{W_N^{(i)}(y_1^N, \hat{u}_1^{i-1} \mid 0)}{W_N^{(i)}(y_1^N, \hat{u}_1^{i-1} \mid 1)}.$$

Then the decoder sets

$$\hat{u}_i := \begin{cases} 0 & \text{if } L_N^{(i)}(y_1^N, \hat{u}_1^{i-1}) \geq 1 \\ 1 & \text{otherwise.} \end{cases}$$

Note that the decoder will not change any previous decisions.

3.2 Polarization and the exponent

In this section, we further explore polarization with arbitrary kernels and the exponent.

3.2.1 Matrices that polarize

It is important to note that not all $l \times l$ matrices polarize a given channel W. The next result describes some circumstances in which a lower triangular matrix L over a finite field $\mathbb{F}_q$ polarizes a q-ary DMC.

Lemma 5 *[19, Corollaries 13 and 14] Let $W : \mathcal{X} \to \mathcal{Y}$ be a q-ary DMC where $\mathcal{X} = \mathbb{F}_q$. Consider a nonsingular lower triangular matrix L whose entries are elements of $\mathbb{F}_q$. Assume that L is not diagonal.*

1. *If q is a prime, then L polarizes the channel W.*

2. *Suppose that q is a prime power. Let k denote the largest row index of L such that $Row_k L$ has at least two nonzero elements. If there exists $j \in \{1, \ldots, k-1\}$*

such that L_{kj} is a primitive element of $\mathcal{X}$, then L polarizes the channel W.

We set out to translate these properties to an arbitrary matrix G as demonstrated in the next two results.

Theorem 8 *Let q be a prime and $\mathcal{X}$ be a finite field of order q. If G is a nonsingular matrix and no column permutation of G is upper triangular, then G polarizes any DMC W with input alphabet $\mathcal{X}$.*

Proof

Because G is nonsingular, there exists an LU factorization $G = ULP$ where U is an upper triangular matrix, L is a lower triangular matrix, and P is a permutation matrix. Since no column permutation of G is upper triangular, L is not diagonal. Hence, Lemma 5 applies, and L polarizes W. The statistical properties of channels $W_N^{(i)}$ are invariant under the operation $G \mapsto U^{-1}GP^{-1} = L$. Consequently, G polarizes W as L does. $\square$

Theorem 9 *Let q be a prime power and $\mathcal{X} = \mathbb{F}_q$. Assume that G is a nonsingular matrix and no column permutation of G is upper triangular. Let k denote the largest row index of G such that $Row_k G$ has at least two nonzero elements. If there exists $j \in \{1, \ldots, k-1\}$ such that $G_{kk}^{-1} G_{kj}$ is a primitive element of $\mathcal{X}$, then G polarizes any DMC W with input alphabet $\mathcal{X}$.*

Proof

As in the proof of Theorem 8, write $G = ULP$ where U is an upper triangular matrix, L is a lower triangular matrix, and P is a permutation matrix. Observe that L is not a diagonal matrix as no column permutation of G is upper triangular. Let k denote the largest row index such that $Row_k G$ has at least two nonzero elements.

First, consider the case where k corresponds to the last row of L. Notice that the entries on last row of G are nonzero scalar multiples of those in the last row of L; hence, we only need to multiply the last row of G by G_{kk}^{-1} to obtain the condition $L_{kk} = 1$. If there exists a primitive element of $\mathcal{X}$ to the left of the diagonal in the last row of G_{kk}^{-1}G, then L polarizes W according to Lemma 5. Thus, G polarizes W as L does.

Next, consider the case where k does not correspond to last row of L. By the definition of k and the fact that L is nonsingular, the rows of L with index greater than k must only have nonzero entries along the diagonal. Thus, the rows of G with index greater than k must also only have nonzero entries along the diagonal. We can also note that $\text{Row}_k G$ must have a nonzero entry to the left of the diagonal, since it has more than one nonzero entry in L. In applying Gaussian elimination to $\text{Row}_k G$ using the rows below k, the only entries affected are entries to the right of the diagonal since rows of G below $\text{Row}_k G$ only have nonzero entries along the diagonal. Then multiply G by G_{kk}^{-1} to satisfy the condition $L_{kk} = 1$. Hence, if there is a primitive element of $\mathcal{X}$ to the left of the diagonal in $\text{Row}_k(G_{kk}^{-1}G)$, then L satisfies Lemma 5. Therefore, G polarizes W as L does. □

A natural question to consider is if Theorem 9 provides a characterization of those matrices which polarize q-ary DMCs. When q is a prime, this is the case.

Theorem 10 *Let q be a prime, and $\mathcal{X}$ be a finite field of size q. Suppose that G is a nonsingular matrix with entries in $\mathcal{X}$. If G polarizes any q-ary DMC W with input alphabet $\mathcal{X}$, then no column permutation of G is upper triangular.*

Proof

Suppose that a column permutation of G is upper triangular. Write $G = ULP$ where

U is an upper triangular matrix, L is a lower triangular matrix, and P is a permutation matrix. Then L is a diagonal matrix. Applying an argument similar to that of [12, Lemma 1], we see that L does not polarize W. Hence G does not polarize W. $\square$

However, as the following example shows when q is a prime power that is not prime, whether or not a matrix polarizes a channel is channel dependent.

Example 12 *Consider the DMC $W : \mathbb{F}_4 \to \mathbb{F}_2$ defined by the transition probabilities*

$$W(0|0) = W(0|\alpha) = W(1|1) = W(1|\alpha^2) = 1,$$

where α is a primitive element of $\mathbb{F}_4$. According to Equation (3.1), $I(W) = \log_4(2)$. Using the kernel matrix G_2 as in Section 1, we can combine and split W into two channels $W_2^{(1)}$ and $W_2^{(2)}$. Observe that G_2 does not fit the form of Theorem 9. Calculating capacities as in Equation (3.1) gives $I(W_2^{(1)}) = I(W_2^{(2)}) = \log_4(2)$; thus, G_2 does not polarize the channel W.

Next, consider the DMC $W' : \mathbb{F}_4 \to \mathbb{F}_2$ defined by the transition probabilities

$$W'(0|0) = W'(1|1) = W'(\alpha|\alpha) = W'(\alpha^2|\alpha^2) = \frac{3}{4}$$

and

$$W'(0|1) = W'(1|0) = W'(\alpha|\alpha^2) = W'(\alpha^2|\alpha) = \frac{1}{4},$$

where α is a primitive element of $\mathbb{F}_4$. Note, $I(W') = \frac{3}{4}\log_4(3) + \frac{1}{4}\log_4(1)$. Again, using the kernel matrix G_2, we can construct $W_2^{'(1)}$ and $W_2^{'(2)}$ such that

$$I(W_2^{'(1)}) = \frac{20}{32}\log_4\left(\frac{5}{2}\right) + \frac{12}{32}\log_4\left(\frac{3}{2}\right)$$

and

$$I(W_2^{'(2)}) = \frac{9}{16}\log_4\left(\frac{18}{5}\right) + \frac{6}{16}\log_4(2) + \frac{1}{16}\log_4\left(\frac{2}{5}\right).$$

Hence, G_2 polarizes the channel W'.

Therefore, we have constructed two DMCs W and W' with input alphabet $\mathcal{X} = \mathbb{F}_4$ such that the kernel matrix G_2 polarizes one channel but not the other. It is important to also note that when q is not prime a multi-level code construction may be used as defined in [24].

Further information on polarization over finite fields may be found in [21].

3.2.2 Probability of error using SC decoding

The exponent gives a bound on the best achievable probability of block error under SC decoding. Let P_e be the best achievable probability of block error under SC decoding for polar coding over W using kernel G.

Theorem 11 *[2, Theorem 1] For polar coding on a binary-DMC W with kernel G_2 at any fixed rate $0 < R < I(W)$ with block length $N = 2^n$, and any fixed $\beta < \frac{1}{2}$,*

$$P_e = o(2^{-2^{n\beta}}).$$

We can also consider the probability of block error using polar coding over $\mathbb{F}_q$ with an arbitary kernel matrix. Let W be a q-ary DMC. If G is a matrix that polarizes according to Theorems 8 and 9, then the exponent helps bound the block error probability under successive cancellation (SC) decoding. Using techniques similar to [1] and [22], the following result holds.

Theorem 12 *Consider polar coding over a q-ary DMC W using kernel G at a fixed*

rate $0 < R < I(W)$ *with block length* $N = l^n$. *Assume* G *polarizes* W. *Then*

$$P_e = O(2^{-l^{n\beta}})$$

for $0 < \beta < E(G)$.

Proof

For any q-ary DMC $W : \mathcal{X} \to \mathcal{Y}$ with fixed rate $0 < R < I(W)$ and $0 < \beta < E(G)$, there exist a sequence $\{\mathcal{A}_N\}$ of sets $\mathcal{A}_N \subseteq \{1, \ldots, N\}$ such that $|\mathcal{A}_N| \geq NR$ and

$$Z(W_N^{(i)}) < 2^{-l^{n\beta}}$$

for all $i \in \{1, \ldots, N\}$. Consider the block error event $\mathcal{E} = \cup_{i \in \mathcal{A}_N} \mathcal{B}_i$ where

$$\mathcal{B}_i = \{(u_1^N, y_1^N) \in \mathcal{X}^N \times \mathcal{Y}^N | \hat{u}_1^{i-1} \neq u_1^{i-1}, \hat{u}_i = u_i\},$$

so that block error probability of decoding is

$$P_e = P(\mathcal{E}) = P(\cup_{i \in \mathcal{A}_N} \mathcal{B}_i).$$

Let

$$\mathcal{E}_v = \left\{(u_1^N, y_1^N) \in \mathcal{X}^N \times \mathcal{Y}^N | W_N^{(i)}(y_1^N, u_1^{i-1} | u_i) \leq W_N^{(i)}(y_1^N, u_1^{i-1} | u_i + v)\right\}.$$

Thus,

$$\mathcal{B}_i \subseteq \cup_{v \in \mathcal{X}} \mathcal{E}_v.$$

Then

$$\begin{aligned}
P(\mathcal{B}_i) &= \sum_{v \in \mathcal{X}} P(\mathcal{E}_v) \\
&= \sum_{v \in \mathcal{X}} \sum_{u_1^N, y_1^N} \frac{1}{q^N} \left(W_N(y_1^N | u_1^N) 1_{\mathcal{E}_v}(u_1^N, y_1^N) \right) \\
&\leq \sum_{v \in \mathcal{X}} \sum_{u_1^N, y_1^N} \frac{1}{q^N} W_N(y_1^N | u_1^N) \sqrt{\frac{W_N^{(i)}(y_1^N, u_1^{i-1} | u_i + v)}{W_N^{(i)}(y_1^N, u_1^{i-1} | u_i)}} \\
&= (q-1) Z(W_N^{(i)}).
\end{aligned}$$

Hence,

$$\begin{aligned}
P(\mathcal{E}) &= P(\cup_{i \in \mathcal{A}_n} \mathcal{B}_i) \\
&\leq \sum_{i \in \mathcal{A}_n} (q-1) Z(W_N^{(i)}) \\
&\leq N(q-1) Z(W_N^{(i)}) \\
&\leq N(q-1) 2^{-l^{n\beta}}.
\end{aligned}$$

□

3.2.3 The exponent and partial distances

The exponent of a matrix can be found using partial distances, a method introduced Korada, Şaşoğlu, and Urbanke for the binary case [12] and explored for larger alphabets by Mori and Tanaka [19].

Definition 32 *For $i = 1, \ldots, l$, the i^{th} partial distance of an $l \times l$ matrix $G =$*

$[g_1^T, \ldots, g_l^T]^T$ *over* $\mathbb{F}_q$ *is*

$$D_i := d\left(g_i, \langle g_{i+1}, \ldots, g_l \rangle\right),$$

the minimum Hamming distance between the vector g_i *and the* $\mathbb{F}_q$*-vector space* $\langle g_{i+1}, \ldots, g_l \rangle$ *spanned by* $g_{i+1}, \ldots, g_l \in \mathbb{F}_q^l$.

Lemma 6 *([12, Theorem 4], [19, Theorem 19]) If* G *is an* $l \times l$ *matrix over* $\mathbb{F}_q$*, then the exponent of the polar code with kernel* G *is*

$$E(G) = \frac{1}{l \ln(l)} \sum_{i=1}^{l} \ln(D_i).$$

Consider Arikan's original kernel matrix

$$G_2 := \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \in \mathbb{F}_2^{2\times 2}.$$

Then $D_1 = 1$ and $D_2 = 2$. Applying Lemma 6, we see $E(G_2) = \frac{1}{2}$. Korada, Şaşoğlu, and Urbanke used repeated shortening of a BCH code to create the first binary kernel G_3 with exponent exceeding $E(G_2) = \frac{1}{2}$. [12]

$$G_3 = \begin{bmatrix}
1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 1 \\
0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 \\
0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 \\
0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \\
0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\
1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\
1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 \\
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \\
1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\
1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1
\end{bmatrix}$$

One may verify that the set of partial distances is

$$\{16, 8, 8, 8, 8, 6, 6, 4, 4, 4, 4, 2, 2, 2, 2, 1\}.$$

Applying Lemma 6 yields the exponent of G_3 which is $E(G_3) = 0.51828$.

This fact, together with the above lemma, leads one to consider kernels that are generator matrices of linear codes. The partial distances of the kernel may then be estimated by bounds on the minimum distances of the nested codes. As we see in the next chapter, algebraic geometric codes lend themselves naturally to this con-

struction.

Chapter 4

Algebraic geometric kernels

Let F be a function field over $\mathbb{F}_q$ of genus g. Recall from Chapter 2, an algebraic geometric (AG) code $C(D, A)$ is constructed using divisors A and $D = P_1 + \cdots + P_n$ on F with disjoint supports, where the P_i are distinct places of F of degree 1. The algebraic geometric code $C(D, A)$ is

$$C(D, A) = \{(f(P_1), \ldots, f(P_n)) | f \in \mathcal{L}(A)\} \subseteq \mathbb{F}_q^n,$$

where

$$\mathcal{L}(A) = \{f \in F \mid (f) \geq -A\} \cup \{0\}$$

is the Riemann-Roch space of A. An especially useful property of AG codes is their nested structure. Given divisors A and B with $\text{supp}(A) \cap \text{supp}(D) = \emptyset = \text{supp}(B) \cap \text{supp}(D)$,

$$A \leq B \;\Rightarrow\; \mathcal{L}(A) \subseteq \mathcal{L}(B) \;\Rightarrow\; C(D, A) \subseteq C(D, B).$$

In this chapter, we employ AG codes in the construction of polar code kernels. We see that the nesting plays a key role in the construction as well as in the analysis of

the exponent.

4.1 Construction of kernels using AG codes

4.1.1 Kernel construction and the exponent

Let $F/\mathbb{F}_q$ be a function field of genus g and $P_1, \ldots, P_n$ be places of F of degree one where $n \geq 2g$. Suppose there exist a sequence of divisors

$$A_1 \leq \cdots \leq A_n$$

so that the supports of $D := P_1 + \cdots + P_n$ and A_i are disjoint for all i, $1 \leq i \leq n$, and

$$C(D, A_1) \subsetneqq C(D, A_2) \subsetneqq \cdots \subsetneqq C(D, A_n) = \mathbb{F}_q^n. \tag{4.1}$$

Then there exists an $n \times n$ generator matrix of $C(D, A_n)$ such that for each i, $1 \leq i \leq n$, the submatrix

$$\begin{bmatrix} \mathrm{Row}_{n-i+1} G \\ \vdots \\ \mathrm{Row}_n G \end{bmatrix}$$

of G is a generator matrix for $C(D, A_i)$.

A sequence of divisors satisfying (4.1) can be constructed as follows. Fix a divisor $D = P_1 + \cdots + P_n$, where each P_i is a place of F of degree one, and a place P of $F/\mathbb{F}_q$ of degree one not in the support of D. First, let

$$0 = \alpha_1 < \cdots < \alpha_{n-g}$$

be the least $n - g$ elements of the Weierstrass semigroup at P. Then, by the Weier-

strass Gap Theorem,

$$\alpha_i = i + g - 1$$

for $g + 1 \leq i \leq n - g$. Next, set

$$\alpha_n = n + 2g - 1.$$

According to the Riemann-Roch Theorem,

$$l(\alpha_n P) - l(\alpha_n P - D) = n,$$

because both $\alpha_n P$ and $\alpha_n P - D$ have degrees at least $2g - 1$. Notice that for all positive integers α,

$$l(\alpha P) - l(\alpha P - D) \leq l((\alpha + 1)P) - l((\alpha + 1)P - D) \leq l(\alpha P) - l(\alpha P - D) + 1.$$

Moreover,

$$l((n-1)P) - l((n-1)P - D) = n - g,$$

as the divisor $(n-1)P-D$ has negative degree. As a result, there exists $n \leq \alpha_{n-g+1} < \cdots < \alpha_{n-1} < \alpha_n = n + 2g - 1$ such that

$$l(\alpha_i P) - l(\alpha_i P - D) \neq l(\alpha_{i-1} P) - l(\alpha_{i-1} P - D).$$

For $1 \leq i \leq n$, set

$$A_i := \alpha_i P;$$

note that

$$A_i := (i + g - 1)\, P$$

for $g+1 \leq i \leq n-g$. Then the one-point codes from the sequence of divisors

$$\alpha_1 P \leq \ldots \leq \alpha_g P \leq 2gP \leq (2g+1)P \leq \ldots \leq (n-1)P \leq \alpha_{n-g+1} P \leq \ldots \leq \alpha_n P$$

satisfy (4.1). We will consider the kernel matrix

$$G = \begin{bmatrix} f_n(P_1) & f_n(P_2) & \cdots & f_n(P_n) \\ f_{n-1}(P_1) & f_{n-1}(P_2) & \cdots & f_{n-1}(P_n) \\ \vdots & \vdots & & \vdots \\ f_1(P_1) & f_1(P_2) & \cdots & f_1(P_n) \end{bmatrix},$$

where for each i, $1 \leq i \leq n$, $\{f_1, \ldots, f_i\}$ is a basis for $\mathcal{L}(\alpha_i P)$.

More generally, we may consider such a matrix where $\{f_1, \ldots, f_i\}$ is a basis for $\mathcal{L}(A_i)$ and the A_i are sequence of nested divisors which give rise to a sequence of nested codes satisfying (4.1) for all i.

Theorem 13 *The exponent of the polar code with kernel G constructed using the code $C(D, \alpha_n P)$ with nested codes $C(D, \alpha_i P)$ as above satisfies*

$$E(G) \geq \frac{1}{n \ln(n)} \left[\ln((n-g)!) + \sum_{i=n-g+1}^{n} \ln(d_i) \right],$$

where d_i denotes the minimum distance of $C(D, \alpha_i P)$.

Proof Notice that $C(D, \alpha_i P)$ is a code over $\mathbb{F}_q$ of length n and dimension i; let d_i denote its minimum distance. If $\alpha_i < n$, then

$$n+1-i-g \leq d_i \leq n+1-i.$$

Then we may bound the partial distances D_i of G by

$$D_i \geq d_i \geq n - \alpha_i.$$

This bound combined with Lemma 6 yields the desired result. $\square$

Remark 1 *1. Theorem 13 provides a lower bound on the exponent, as the partial distances associated with these matrices are not necessarily nondecreasing. Certainly, the matrix itself could be manipulated to satisfy this, but doing so would obscure the structure given by the AG code and associated Riemann-Roch spaces. This structure may prove useful in further studies, such as into shortened AG code kernels, which will be discussed in Section 4.2.*

2. *One may use multi-level code construction with the AG code kernels constructed above [24]; however, a manipulation of the kernel will ensure that these kernels polarize according to Theorem 9. In the construction, we may assume*

$$Row_n G = (1, \ldots, 1)$$

by taking $f_1 = 1$. Since $C(D, A_1) \neq C(D, A_2)$, there exists $j < n$ such that $G_{n-1,j} \neq G_{n-1,n}$. Now, replace f_1 with

$$f_1' := (\alpha - 1)\left(G_{n-1,j} - G_{n-1,n}\right)^{-1}\left(f_2 - G_{n-1,n}\right) + 1,$$

where α is a primitive element of $\mathbb{F}_q$, to create a new matrix G'; that is, G' is

an $n \times n$ matrix with

$$Row_i G' := \begin{cases} Row_i G & \text{if } 1 \leq i \leq n-1 \\ (f_1'(P_1), \ldots, f_1'(P_n)) & \text{if } i = n. \end{cases}$$

One may check that $G'_{nj} = \alpha$ and $G'_{nn} = 1$; hence, the new kernel G' polarizes by Theorem 9. The exponent of G' may be bounded as well. Indeed, note that the proof of Theorem 13 applies except for the term d_n. Even so, $D_n \geq 2$ since $G'_{nj} = \alpha$ and $G'_{nn} = 1$.

3. *In general, it is hard to determine minimum distances of AG codes and exact values are only known for a few cases such one-point and two-point Hermitian codes [[8], [29]] (See sections 4.1.2 and 4.2.1). Thus, one may bound the exponent by*

$$E(G) \geq \frac{1}{n \ln(n)} \left[\ln((n-g)!)\right].$$

An immediate corollary of Theorem 13 is the exponent of a kernel based on a Reed-Solomon code; this is computed by Mori and Tanaka [19]. Here, we take F to be the rational function field over $\mathbb{F}_q$. Applying the construction above yields a matrix $G_{RS} \in \mathbb{F}_q^{q \times q}$ whose submatrices correspond to generator matrices of Reed-Solomon codes over $\mathbb{F}_q$. That is, let $\mathbb{F}_q = \{0, \alpha, \alpha^2, \ldots, \alpha^{q-1}\}$, then

$$G_{RS} = \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 & 0 \\ \alpha^{(q-2)(q-2)} & \alpha^{(q-3)(q-2)} & \cdots & \alpha^{q-2} & 1 & 0 \\ \alpha^{(q-2)(q-3)} & \alpha^{(q-3)(q-3)} & \cdots & \alpha^{q-3} & 1 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ \alpha^{(q-2)} & \alpha^{(q-3)} & \cdots & \alpha & 1 & 0 \\ 1 & 1 & \cdots & 1 & 1 & \gamma \end{bmatrix}$$

where γ is nonzero element in $\mathbb{F}_q$ [19]. Note that G_{RS} will polarize if $\gamma \neq 0$ when q is a prime and γ is a primitive element when q is a prime power [19].

Corollary 7 *The exponent of a Reed-Solomon kernel G_{RS} over $\mathbb{F}_q$ is*

$$E(G_{RS}) = \frac{\ln(q!)}{q\ln(q)}.$$

Proof This follows directly from Theorem 13 using the fact that F has genus $g = 0$. $\square$

Another consequence of Theorem 13 is the asymptotic behavior of exponents of kernels constructed from codes over maximal function fields. Recall that a function field over $\mathbb{F}_q$ of genus g is said to be maximal provided its number of places of degree one meets the Hasse-Weil bound; that is, the number of places of $F/\mathbb{F}_q$ of degree one is $q + 1 + 2g\sqrt{q}$.

Theorem 14 *Let $F/\mathbb{F}_q$ be a maximal function field of genus g, and let G be a generator matrix of an one-point AG code on F of length $n = q + 2g\sqrt{q}$ constructed as in (4.1). Then*

$$\lim_{q\to\infty} E(G) = 1.$$

Proof Suppose $F/\mathbb{F}_q$ is a maximal function field, then

$$g \leq \frac{q - q^{1/2}}{2},$$

and

$$n = q + 2gq^{1/2} \leq q + 2\left(\frac{q - q^{1/2}}{2}\right)q^{1/2} \leq q^{3/2}.$$

In addition,

$$\begin{aligned} n - g &= q + 2gq^{1/2} - g \\ &= q - g(1 - 2q^{1/2}) \\ &\geq q - \left(\frac{q - q^{1/2}}{2}\right)(1 - 2q^{1/2}) \\ &= q^{3/2} + \frac{q}{2} + \frac{q^{1/2}}{2}. \end{aligned}$$

Then

$$E(G) \geq \frac{1}{n\ln(n)}\left(\ln((n-g)!) + \sum_{i=n-g+1}^{n} \ln(d_i)\right)$$

$$\geq \frac{\ln((n-g)!)}{n\ln(n)}$$

$$\geq \frac{1}{q^{3/2}\ln(q^{3/2})} \ln\left(\left(q^{3/2} + \frac{q}{2} + \frac{q^{1/2}}{2}\right)!\right)$$

$$= \frac{1}{q^{3/2}\ln(q^{3/2})} \sum_{i=2}^{q^{3/2}+\frac{q}{2}+\frac{q^{1/2}}{2}} \ln(i)$$

$$\geq \frac{1}{q^{3/2}\ln(q^{3/2})} \int_{1}^{q^{3/2}+\frac{q}{2}+\frac{q^{1/2}}{2}} \ln(x)dx$$

$$= \frac{q^{3/2}+\frac{q}{2}+\frac{q^{1/2}}{2}}{q^{3/2}\ln(q^{3/2})} \ln\left(q^{3/2} + \frac{q}{2} + \frac{q^{1/2}}{2}\right) - \frac{q^{3/2}+\frac{q}{2}+\frac{q^{1/2}}{2}-1}{q^{3/2}\ln(q^{3/2})}$$

$$= \left(1 + \frac{1}{2q^{1/2}} + \frac{1}{2q}\right)\frac{\ln\left(q^{3/2}+\frac{q}{2}+\frac{q^{1/2}}{2}\right)}{\ln(q^{3/2})} - \frac{1}{\ln(q^{3/2})} - \frac{1}{2q^{1/2}\ln(q^{3/2})}$$
$$- \frac{1}{2q\ln(q^{3/2})} + \frac{1}{q^{3/2}\ln(q^{3/2})}.$$

Therefore, $\lim_{q\to\infty} E(G) = 1$. $\square$

In the next subsection, we more closely examine kernels from codes over a particular maximal function field, the Hermitian function field.

4.1.2 Kernels from Hermitian codes

Let $F = \mathbb{F}_{q^2}(x, y)$ be the function field of the curve

$$y^q + y = x^{q+1}$$

where q is a power of a prime; that is, let $F/\mathbb{F}_{q^2}$ be the Hermitian function field. Recall that the Hermitian function field over $\mathbb{F}_{q^2}$ has genus $\frac{q(q-1)}{2}$ and $q^3 + 1$ places of degree one; hence, it is a maximal function field. A Hermitian one-point code is of the form $C(D, aP_\infty)$, where $D = \sum_{\alpha,\beta \in \mathbb{F}_{q^2}, \beta^q + \beta = \alpha^{q+1}} P_{\alpha,\beta}$ and $P_{\alpha,\beta}$ is a common zero of $x - \alpha$ and $y - \beta$. Mori and Tanaka considered generator matrices for Hermitian codes over fields of even characteristic, that is, over $\mathbb{F}_{2^m}$, as kernels of polar codes in [20]. Applying Theorem 13 and the exact distances of one-point Hermitian codes [29] provides a lower bound on the exponent of the resulting kernel for any characteristic. Let $G_H \in \mathbb{F}_{q^2}^{q^3 \times q^3}$ denote a matrix constructed from the Hermitian code $C(D, \alpha_n P_\infty)$ as in (4.1).

Corollary 8 *The exponent of a Hermitian kernel G_H over $\mathbb{F}_{q^2}$ satisfies*

$$E(G_H) \geq \frac{1}{q^3 \ln(q^3)} \ln \left((q^3 - q^2 + q)! \prod_{j=1}^{q-1} \frac{(q^3 - (j-1)q)^{\underline{j}} (q-1)^{\underline{j}} (q^2 - jq)^j}{\prod_{i=1}^{j} (q^2 - jq - i)} \right),$$

where $a^{\underline{i}} := \frac{a!}{(a-i)!}$.

Proof For the Hermitian function field over $\mathbb{F}_{q^2}$, the set of minimum distances of the one-point codes $C(D, \alpha_i P)$ constructed as in (4.1) is

$$\{q^3 - aq - b : 0 \leq b \leq a \leq q-2\} \cup \{q^3 - g - i + 1 : g+1 \leq i \leq q^3 - q^2 - g\}$$
$$\cup\{q^2 - jq - (j+1), \ldots, q^2 - (j+1)q + 1 : 0 \leq j \leq q-1\}$$
$$\cup\{[q^2 - jq]^{j+1} : 0 \leq j \leq q-1\} \cup \{[a]^a : 1 \leq a \leq q-1\},$$

where $g = \frac{q(q-1)}{2}$ and $[a]^t$ denotes the multiset $\{a, \ldots, a\}$ of cardinality t [**?**]. Hence, $E(G_H)$ is bounded below by

$$\frac{1}{q^3} \log_{q^3} \left((q^3 - q^2 + q)!(q-1)!^{q-1} \prod_{j=0}^{q-2} f(j,q) \right)$$

where

$$f(j,q) := \frac{(q^3 - jq)(q^3 - jq - 1)\ldots(q^3 - jq - j)(q^2 - (j+1)q)^{j+1}}{j!(q^2 - (j+1)q - 1)\ldots(q^2 - (j+1)q - (j+1))}.$$

Thus,

$$E(G_H) \geq \frac{1}{q^3} \log_{q^3} \left((q^3 - q^2 + q)! \prod_{j=1}^{q-1} \frac{(q^3 - (j-1)q)^{\underline{j}} (q-1)^{\underline{j}} (q^2 - jq)^j}{\prod_{i=1}^{j} (q^2 - jq - i)} \right).$$

□

Table 1 displays comparisons between the exponents of Reed-Solomon kernels and lower bounds on the exponents of Hermitian kernels. Note that the size of the kernel based on Reed-Solomon codes over $\mathbb{F}_{q^2}$ is $q^2 \times q^2$, while the size of the kernel produced from Hermitian one-point codes over $\mathbb{F}_{q^2}$ is $q^3 \times q^3$.

The table suggests that the exponent of the kernel based on the Hermitian code is greater than that based on the Reed-Solomon code over $\mathbb{F}_{q^2}$, provided $q \neq 2$. Indeed, the proof follows immediately from Theorem 13, Corollary 7, and Corollary 8.

Proposition 3 *Let G_H be a Hermitian kernel over $\mathbb{F}_{q^2}$, and let G_{RS} be a Reed-Solomon kernel also over $\mathbb{F}_{q^2}$. Then*

$$E(G_{RS}) \leq E(G_H)$$

for $q \geq 3$.

Proof Let G_H be a Hermitian kernel over $\mathbb{F}_{q^2}$, and let G_{RS} be a Reed-Solomon kernel also over $\mathbb{F}_{q^2}$. Recall, from Corollary 7, $E(G_{RS}) = \frac{\ln(q^2!)}{q^2 \ln(q^2)}$. Also, since the Hermitian function field over $\mathbb{F}_{q^2}$ has genus $\frac{q(q-1)}{2}$, we may bound $E(G_H)$ using Theorem 13 by

$$E(G_H) \geq \frac{\ln\left(\left(q^3 - \frac{q(q-1)}{2}\right)!\right)}{q^3 \ln(q^3)} = \frac{\ln\left(\left(q^3 - \frac{q^2}{2} + \frac{q}{2}\right)!\right)}{q^3 \ln(q^3)}.$$

Thus, we want to find for what values of q

$$\frac{\ln(q^2!)}{q^2 \ln(q^2)} \leq \frac{\ln\left(\left(q^3 - \frac{q^2}{2} + \frac{q}{2}\right)!\right)}{q^3 \ln(q^3)}.$$

Note

$$\frac{\ln(q^2!)}{q^2 \ln(q^2)} \leq \frac{\ln\left(\left(q^3 - \frac{q^2}{2} + \frac{q}{2}\right)!\right)}{q^3 \ln(q^3)}$$

if and only if

$$\frac{\ln(q^2!)}{2q^2 \ln(q)} \leq \frac{\ln\left(\left(q^3 - \frac{q^2}{2} + \frac{q}{2}\right)!\right)}{3q^3 \ln(q)}$$

if and only if

$$\frac{3q}{2}\ln(q^2!) \leq \ln\left(\left(q^3 - \frac{q^2}{2} + \frac{q}{2}\right)!\right).$$

Since $\ln(x)$ is a continuous and increasing function for $x > 0$,

$$\begin{aligned}
\ln\left(\left(q^3 - \frac{q^2}{2} + \frac{q}{2}\right)!\right) &= \sum_{i=1}^{q^3 - \frac{q^2}{2} + \frac{q}{2}} \ln(i) \\
&= \sum_{i=2}^{q^3 - \frac{q^2}{2} + \frac{q}{2}} \ln(i) \\
&\geq \int_1^{q^3 - \frac{q^2}{2} + \frac{q}{2}} \ln(x)dx \\
&= \left(q^3 - \frac{q^2}{2} + \frac{q}{2}\right)\left(\ln\left(q^3 - \frac{q^2}{2} + \frac{q}{2}\right) - 1\right) + 1.
\end{aligned}$$

Similarly,

$$\begin{aligned}
\frac{3q}{2}\ln\left(q^2!\right) &= \frac{3q}{2}\sum_{i=1}^{q^2} \ln(i) \\
&\leq \frac{3q}{2}\int_1^{q^2+1} \ln(x)dx \\
&= \frac{3q}{2}\left(\left(q^2+1\right)\left(\ln\left(q^2+1\right) - 1\right) + 1\right).
\end{aligned}$$

Computational one may check that

$$\frac{3q}{2}\left(\left(q^2+1\right)\left(\ln\left(q^2+1\right) - 1\right) + 1\right) = \left(q^3 - \frac{q^2}{2} + \frac{q}{2}\right)\left(\ln\left(q^3 - \frac{q^2}{2} + \frac{q}{2}\right) - 1\right) + 1$$

for $q \approx 0.7453$ and $q \approx 7.0545$, and

$$\frac{3q}{2}\left(\left(q^2+1\right)\left(\ln\left(q^2+1\right) - 1\right) + 1\right) < \left(q^3 - \frac{q^2}{2} + \frac{q}{2}\right)\left(\ln\left(q^3 - \frac{q^2}{2} + \frac{q}{2}\right) - 1\right) + 1$$

	m	2	4	6	8
q = 2	Reed-Solomon	0.57312	0.69141	0.77082	0.82226
	Hermitian	0.56216	0.70734	0.80276	0.85930
q = 3	Reed-Solomon	0.64737	0.78120	0.84917	0.88631
	Hermitian	0.65248	0.81459	0.88634	0.91988
q = 5	Reed-Solomon	0.72079	0.84569	0.89648	0.92233
	Hermitian	0.74345	0.88296	0.92819	0.94767

Table 4.1: Lower bounds on exponents of Reed-Solomon and Hermitian kernels over $\mathbb{F}_{q^m}$

for $q \geq 8$. Using Corollary 8, one may check $E(G_{RS}) < E(G_H)$ for $3 \leq q \leq 7$. $\square$

Remark 2 *It should be observed that the kernel matrices in Proposition 3 are over the same field, $\mathbb{F}_{q^2}$, but are not of the same size. Indeed, G_H is a $q^3 \times q^3$ matrix while G_{RS} is of size $q^2 \times q^2$.*

For the purposes of polar coding, it might be just as relevant, if not more so, to compare exponents of matrices of the same size, though over different fields. In this situation, we conclude that the exponent of the Reed-Solomon kernel over $\mathbb{F}_{q^3}$ exceeds the lower bound on the exponent of the Hermitian kernel over $\mathbb{F}_{q^2}$ given in Corollary 8.

4.1.3 Kernels from Suzuki codes

In this subsection, we see that the asymptotic behavior of the exponent in Theorem 14 may occur for kernels constructed from function fields that are not maximal. To do so, we investigate codes from the Suzuki function field, a function field which is not maximal yet is optimal (according to the explicit formulas of Weil).

Let $F = \mathbb{F}_q(x, y)$ be the function field of the Suzuki curve which has defining

equation

$$y^{2^{2r+1}} - y = x^{2^r}(x^{2^{2r+1}} - x)$$

over $\mathbb{F}_q$ where $q = 2^{2r+1}$ and r is a positive integer. Then the genus of F is $g = \sqrt{\frac{q}{2}}(q-1)$, and F has exactly q^2+1 places of degree one. The Suzuki one-point code is of the form $C(D, aP_\infty)$, where $D = \sum_{\alpha,\beta \in \mathbb{F}_{2^{2r+1}}} P_{\alpha,\beta}$ and $P_{\alpha,\beta}$ is a common zero of $x - \alpha$ and $y - \beta$. Theorem 13 then yields the following result. Let G_{Suz} denote a Suzuki kernel over $\mathbb{F}_q$ where $q = 2^{2r+1}$ and r is a positive integer as in (4.1).

Corollary 9 *Let G_{Suz} be a Suzuki kernel over $\mathbb{F}_q$ where $q = 2^{2r+1}$. Then*

$$E(G_{Suz}) \geq \frac{1}{q^2 \ln(q^2)} \ln\left(\left(q^2 - \sqrt{\frac{q}{2}}(q-1)\right)!\right).$$

Proof This follows directly from Theorem 13 using the fact that F has genus $g = \sqrt{\frac{q}{2}}(q-1)$. $\square$

The exact minimum distances of Suzuki one-point codes over $\mathbb{F}_8$ are known according to the work of Chen and Duursma [5]. We further explore this function field in the example below.

Example 13 *Let $F = \mathbb{F}_8(x, y)$ be the function field of the Suzuki curve with defining equation $y^8 - y = x^4(x^8 - x)$, and let α be a primitive element of $\mathbb{F}_8$. A one-point code over this function field, called a Suzuki one-point code, is of length 64. The Suzuki*

one-point codes

$$
\begin{aligned}
C(D, P_\infty) \subsetneqq C(D, 8P_\infty) \subsetneqq C(D, 10P_\infty) \subsetneqq C(D, 12P_\infty) \subsetneqq C(D, 13P_\infty) \\
\subsetneqq C(D, 16P_\infty) \subsetneqq C(D, 18P_\infty) \subsetneqq C(D, 20P_\infty) \subsetneqq C(D, 21P_\infty) \\
\subsetneqq C(D, 22P_\infty) \subsetneqq C(D, 23P_\infty) \subsetneqq C(D, 24P_\infty) \subsetneqq C(D, 25P_\infty) \\
\subsetneqq C(D, 26P_\infty) \subsetneqq C(D, 28P_\infty) \subsetneqq C(D, 29P_\infty) \subsetneqq \cdots \subsetneqq C(D, 63P_\infty) \\
\subsetneqq C(D, 65P_\infty) \subsetneqq C(D, 66P_\infty) \subsetneqq C(D, 67P_\infty) \subsetneqq C(D, 68P_\infty) \\
\subsetneqq C(D, 69P_\infty) \subsetneqq C(D, 70P_\infty) \subsetneqq C(D, 71P_\infty) \subsetneqq C(D, 73P_\infty) \\
\subsetneqq C(D, 75P_\infty) \subsetneqq C(D, 78P_\infty) \subsetneqq C(D, 79P_\infty) \subsetneqq C(D, 81P_\infty) \\
\subsetneqq C(D, 83P_\infty) \subsetneqq C(D, 90P_\infty) \subsetneqq C(D, 91P_\infty) = \mathbb{F}_8^{64}
\end{aligned}
$$

satisfy (4.1).. The resulting kernel matrix is

$$
\begin{array}{c}
\begin{matrix} (0,0) & (0,1) & \dots & (0,\alpha) & \dots & (\alpha^6,\alpha^5) & (\alpha^6,\alpha^6) \end{matrix} \\
\begin{pmatrix}
0 & 1 & \dots & 1 & \dots & (1+\alpha^6)(1+\alpha^5+\alpha^3)^5 & 0 \\
\vdots & \vdots & & \vdots & & \vdots & \vdots \\
0 & 0 & \dots & 0 & \dots & \alpha^6 & \alpha^6 \\
1 & 1 & \dots & 1 & \dots & 1 & 1
\end{pmatrix}.
\end{array}
$$

The partial distances of this matrix are bounded by the exact minimum distances of the Suzuki one-point codes, which are

$$
\begin{aligned}
&64, 56, 56, 52, 51, 48, 46, 44, 43, 42, 42, 40, 39, 38, 36, 35, 34, 33, 32, 31, 30, 29, 28, \\
&28, 26, 25, 24, 24, 22, 21, 20, 20, 18, 18, 16, 16, 16, 13, 12, 12, 12, 12, 8, 8, 8, 8, 8, 8, \\
&8, 8, 7, 7, 6, 6, 4, 4, 4, 4, 4, 3, 3, 2, 2, 1
\end{aligned}
$$

according to [5]. Hence, Theorem 13 implies $E(G_{Suz}) \geq 0.65555$.

Table 2 compares the exponents of Reed-Solomon kernels and lower bounds on the exponents of Suzuki kernels. As with Hermitian kernels, Suzuki kernels yield larger exponents than Reed-Solomon kernels over the same field; however, the larger

	q = 8		q = 32	
Kernel	Exponent	Size of Kernel	Exponent	Size of Kernel
Reed-Solomon	0.63747	8×8	0.73540	32×32
Suzuki	0.65555	64×64	0.73635	1024×1024

Table 4.2: Lower bounds on exponents of Reed-Solomon and Suzuki kernels over $\mathbb{F}_q$ where $q = 2^{2r+1}$

exponent comes at the price of a larger kernel size.

Proposition 4 *Let $q = 2^{2r+1}$ where r is positive integer. Let G_{Suz} be a Suzuki kernel over $\mathbb{F}_q$, and let G_{RS} be a Reed-Solomon kernel also over $\mathbb{F}_q$. Then*

$$E(G_{RS}) \leq E(G_{Suz})$$

for all $q = 2^{2r+1}$ where $r \geq 1$.

Proof Let G_{Suz} be a Suzuki kernel over $\mathbb{F}_q$, and let G_{RS} be a Reed-Solomon kernel also over $\mathbb{F}_q$. Recall, from Corollary 7, $E(G_{RS}) = \frac{\ln(q!)}{q\ln(q)}$ and from Corollary 9

$$E(G_{Suz}) \geq \frac{1}{q^2\ln(q^2)} \ln\left(\left(q^2 - \sqrt{\frac{q}{2}}(q-1)\right)!\right).$$

We want to find for what values of q

$$\frac{\ln(q!)}{q\ln(q)} \leq \frac{1}{q^2\ln(q^2)} \ln\left(\left(q^2 - \sqrt{\frac{q}{2}}(q-1)\right)!\right)$$

Then

$$\frac{\ln(q!)}{q\ln(q)} \leq \frac{1}{q^2\ln(q^2)} \ln\left(\left(q^2 - \sqrt{\frac{q}{2}}(q-1)\right)!\right)$$

if and only if

$$\frac{\ln(q!)}{q\ln(q)} \leq \frac{1}{2q^2\ln(q)} \ln\left(\left(q^2 - \sqrt{\frac{q}{2}}(q-1)\right)!\right)$$

if and only if

$$2q\ln(q!) \leq \ln\left(\left(q^2 - \sqrt{\frac{q}{2}}(q-1)\right)!\right).$$

Since $\ln(x)$ is a continuous and increasing function for $x > 0$,

$$\begin{aligned}
\ln\left(\left(q^2 - \sqrt{\frac{q}{2}}(q-1)\right)!\right) &= \sum_{i=1}^{q^2-\sqrt{\frac{q}{2}}(q-1)} \ln(i) \\
&= \sum_{i=2}^{q^2-\sqrt{\frac{q}{2}}(q-1)} \ln(i) \\
&\geq \int_1^{q^2-\sqrt{\frac{q}{2}}(q-1)} \ln(x)dx \\
&= \left(q^2 - \sqrt{\frac{q}{2}}(q-1)\right)\left(\ln\left(q^2 - \sqrt{\frac{q}{2}}(q-1)\right) - 1\right) + 1.
\end{aligned}$$

Similarly,

$$\begin{aligned}
2q\ln(q) &= 2q\sum_{i=1}^{q} \ln(i) \\
&\leq 2q\int_1^{q+1} \ln(x)dx \\
&= 2q\left((q+1)\left(\ln(q+1) - 1\right) + 1\right).
\end{aligned}$$

Computational one may check that

$$2q\left((q+1)\left(\ln(q+1) - 1\right) + 1\right) = \left(q^2 - \sqrt{\frac{q}{2}}(q-1)\right)\left(\ln\left(q^2 - \sqrt{\frac{q}{2}}(q-1)\right) - 1\right) + 1.$$

for $q \approx 0.5385$ and $q \approx 44.0382$, and

$$2q\left((q+1)\left(\ln(q+1) - 1\right) + 1\right) < \left(q^2 - \sqrt{\frac{q}{2}}(q-1)\right)\left(\ln\left(q^2 - \sqrt{\frac{q}{2}}(q-1)\right) - 1\right) + 1.$$

for $q \geq 45$. One may check $E(G_{RS}) < E(G_H)$ for $q = 8$ and $q = 32$ as shown in Table 4.2. $\square$

Remark 3 *It should be observed kernel matrices in Proposition 4 are over the same field but are not of the same size. Indeed, G_{Suz} is a $q^2 \times q^2$ matrix while G_{RS} is of size $q \times q$.*

As discussed in Remark 2, comparing exponents of matrices of the same size, though over different fields, may also be meaningful for polar coding. In this situation, we conclude that the exponent of the Reed-Solomon kernel over $\mathbb{F}_q$ exceeds the lower bound on the exponent of the Suzuki kernel over $\mathbb{F}_{q^2}$ given in Corollary 9.

The limiting behavior of the exponent in Theorem 14 is not restricted to maximal function fields. In fact, kernels from Suzuki one-point codes display similar asymptotics.

Theorem 15 *Let $q = 2^{2r+1}$ where r is positive integer. Let G_{Suz} be a Suzuki kernel over $\mathbb{F}_q$. Then*

$$\lim_{q \to \infty} E(G_{Suz}) = 1.$$

Proof Let G be a Suzuki kernel over $\mathbb{F}_q$ where $q = 2^{2r+1}$ and r is a positive integer. Then

$$E(G) \geq \frac{1}{q^2} \log_{q^2} \left(\left(q^2 - \sqrt{\frac{q}{2}}(q-1) \right)! \right).$$

Also,

$$\begin{aligned}
\frac{1}{q^2}\log_{q^2}\left(\left(q^2-\sqrt{\frac{q}{2}}(q-1)\right)!\right) &= \frac{1}{q^2}\sum_{i=0}^{q^2-\sqrt{\frac{q}{2}}(q-1)-1}\log_{q^2}(i+1)\\
&\geq \frac{1}{q^2\ln(q^2)}\int_1^{q^2-\sqrt{\frac{q}{2}}(q-1)}\ln(x)dx\\
&= \left(1-\frac{1}{\sqrt{2}q^{1/2}}+\frac{1}{\sqrt{2}q^{3/2}}\right)\frac{\ln\left(q^2-\sqrt{\frac{q}{2}}(q-1)\right)}{\ln(q^2)}\\
&\quad -\left(\frac{1}{\ln(q^2)}-\frac{1}{\sqrt{2}q^{1/2}\ln(q^2)}+\frac{1}{\sqrt{2}q^{3/2}\ln(q^2)}-\frac{1}{q^2\ln(q^2)}\right).
\end{aligned}$$

By L'Hôpital's rule,

$$\begin{aligned}
&\lim_{q\to\infty}\frac{1}{q^2}\log_{q^2}\left(\left(q^2-\sqrt{\frac{q}{2}}(q-1)\right)!\right)\\
&\qquad\geq \lim_{q\to\infty}\left(1-\frac{1}{\sqrt{2}q^{1/2}}+\frac{1}{\sqrt{2}q^{3/2}}\right)\frac{\ln\left(q^2-\sqrt{\frac{q}{2}}(q-1)\right)}{\ln(q^2)}\\
&\qquad\quad -\left(\frac{1}{\ln(q^2)}-\frac{1}{\sqrt{2}q^{1/2}\ln(q^2)}+\frac{1}{\sqrt{2}q^{3/2}\ln(q^2)}-\frac{1}{q^2\ln(q^2)}\right)\\
&\qquad= (1-0+0)(1)-0=1.
\end{aligned}$$

Therefore, the exponent of the Suzuki kernel tends to 1 as $q\to\infty$. □

4.2 Shortening an AG code kernel

The method of shortening can be used to create smaller kernels with large exponent. In [12], Korada, Şaşoğlu, and Urbanke used repeated shortening of a BCH code to create the first binary kernel with exponent exceeding $E(G_2)=\frac{1}{2}$.

To shorten an $l\times l$ kernel G, first find the column j with the longest run of zeros at the top of the column. Then find the row i with the first nonzero element

of column j. Add Row_iG to all the rows with a nonzero element in Col_jG. Finally, remove Col_jG and Row_iG to obtain an $(l-1)\times(l-1)$ matrix. As the next result shows, shortening applied to AG code kernels is a special case of a multipoint code construction.

Theorem 16 *Let $\alpha_1 \leq \cdots \leq \alpha_n$ be integers such that*

$$C(D, \alpha_1 P) \subsetneqq \cdots \subsetneqq C(D, \alpha_n P) = \mathbb{F}_q^n$$

and G be a generator matrix of $C(D, \alpha_n P)$ constructed according to (4.1). Suppose G' is the matrix obtained by shortening applied to the j^{th} column of G. Then

$$C(D - P_j, \alpha_1 P - P_j) \subsetneqq \cdots \subsetneqq C(D - P_j, \alpha_n P - P_j) = \mathbb{F}_q^{n-1},$$

and G' corresponds to the generator matrix of $C(D - P_j, \alpha_n P - P_j)$, which is a two-point code.

Proof Let $\alpha_1 \leq \ldots \leq \alpha_n$ be integers such that

$$C(D, \alpha_1 P) \subsetneqq \cdots \subsetneqq C(D, \alpha_n P) = \mathbb{F}_q^n$$

and $\{f_1, \ldots, f_t\}$ be a basis for $\mathcal{L}(\alpha_t P)$ for all t, $1 \leq t \leq n$. Let G denote a generator matrix of $C(D, \alpha_n P)$ constructed according to (2). Suppose j is the column with the longest run of zeros at the top and $f_i(P_j) \neq 0$ but $f_t(P_j) = 0$ for all $1 \leq t \leq i-1$. Define $\{h_1, \ldots, h_{i-1}, h_{i+1}, \ldots, h_n\}$ as

$$h_s := \begin{cases} f_s & \text{if } f_s(P_j) = 0 \\ f_s + f_i & \text{if } f_s(P_j) = 1. \end{cases}$$

Then $h_s \in \mathcal{L}(\alpha_s P - P_j) \setminus \mathcal{L}(\alpha_{s-1} P - P_j)$. Hence, $\{h_1, \ldots, h_{i-1}, h_{i+1}, \ldots, h_n\}$ is a basis for $\mathcal{L}(a_t P - P_j)$ for all t, $1 \leq t \leq i-1$, $i+1 \leq n$. Thus,

$$\begin{aligned} C(D - P_j, \quad & \alpha_i P - P_j) \subsetneqq \cdots \subsetneqq C(D - P_j, \alpha_{i-1} P - P_j) \\ & \subsetneqq C(D - P_j, \alpha_{i+1} P - P_j) \subsetneqq \cdots \subsetneqq C(D - P_j, \alpha_n P - P_j) = \mathbb{F}_q^{n-1} \end{aligned}$$

is a sequence of codes satisfying (4.1). $\square$

Note that we can apply this method repeatedly, which will result in other multipoint codes.

Example 14 *Consider the following generator matrix G_H for the Hermitian code $C(D, 9P_\infty)$ over $\mathbb{F}_4$, where α is a primitive element of $\mathbb{F}_4$ satisfying $\alpha^2 + \alpha + 1 = 0$:*

$$G_H = \begin{array}{c} \\ X^3Y \\ X^2Y \\ X^3 \\ XY \\ X^2 \\ Y \\ X \\ 1 \end{array} \begin{array}{c} \begin{array}{cccccccc} (0,0) & (0,1) & (1,\alpha) & (1,\alpha^2) & (\alpha,\alpha) & (\alpha,\alpha^2) & (\alpha^2,\alpha) & (\alpha^2,\alpha^2) \end{array} \\ \begin{pmatrix} 0 & 0 & \alpha & \alpha^2 & \alpha & \alpha^2 & \alpha & \alpha^2 \\ 0 & 0 & \alpha & \alpha^2 & 1 & \alpha & \alpha^2 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & \alpha & \alpha^2 & \alpha^2 & 1 & 1 & \alpha \\ 0 & 0 & 1 & 1 & \alpha^2 & \alpha^2 & \alpha & \alpha \\ 0 & 1 & \alpha & \alpha^2 & \alpha & \alpha^2 & \alpha & \alpha^2 \\ 0 & 0 & 1 & 1 & \alpha & \alpha & \alpha^2 & \alpha^2 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} \end{array}.$$

The columns of G_H are indexed by (α, β) such that $P_{\alpha,\beta}$ is a place of degree one of the Hermitian function field over $\mathbb{F}_4$, and the rows are indexed by functions in a basis of the Riemann-Roch space $\mathcal{L}(9P_\infty)$.

For $1 \leq i \leq 7$, the last i rows are indexed by functions which form a basis for $\mathcal{L}(\alpha_i P_\infty)$.

Pick the column with the longest run of zeros on the top, which is the first column of G_H. Since the last row of G_H is the only row with a nonzero entry in the first column, we will remove the last row and the first column of G_H. The resulting kernel is

$$\begin{array}{c} \\ X^3Y \\ X^2Y \\ X^3 \\ XY \\ X^2 \\ Y \\ X \end{array} \begin{array}{c} \begin{array}{ccccccc} (0,1) & (1,\alpha) & (1,\alpha^2) & (\alpha,\alpha) & (\alpha,\alpha^2) & (\alpha^2,\alpha) & (\alpha^2,\alpha^2) \end{array} \\ \begin{pmatrix} 0 & \alpha & \alpha^2 & \alpha & \alpha^2 & \alpha & \alpha^2 \\ 0 & \alpha & \alpha^2 & 1 & \alpha & \alpha^2 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & \alpha & \alpha^2 & \alpha^2 & 1 & 1 & \alpha \\ 0 & 1 & 1 & \alpha^2 & \alpha^2 & \alpha & \alpha \\ 1 & \alpha & \alpha^2 & \alpha & \alpha^2 & \alpha & \alpha^2 \\ 0 & 1 & 1 & \alpha & \alpha & \alpha^2 & \alpha^2 \end{pmatrix} \end{array}$$

which may be obtained from a generator matrix of the two-point Hermitian code $C(D-P_{0,0}, 9P_\infty - P_{0,0})$.

4.2.1 Kernels from Two-Point Hermitian Codes

Let $F = \mathbb{F}_{q^2}(x,y)$ be the Hermitian function field, meaning that it is given by the curve

$$y^q + y = x^{q+1}$$

where q is a power of a prime. A Hermitian two-point code is of the form $C(D, m_1P_\infty + m_2P_{0,0})$, where $D = \sum_{\alpha,\beta\in\mathbb{F}_{q^2},\beta^q+\beta=\alpha^{q+1}} P_{\alpha,\beta}$ and $P_{\alpha,\beta}$ is a common zero of $x-\alpha$ and $y-\beta$. We may assume the two-point Hermitian codes is of this form since the automorphism group of the Hermitian curve is doubly transitive. Applying Theorem 13 and the exact distances of two-point Hermitian codes [8] provides a lower bound on the exponent of the resulting kernel. Let $G_{H2} \in \mathbb{F}_{q^2}^{(q^3-1)\times(q^3-1)}$ denote a matrix

constructed from the two-point Hermitian code $C(D, m_1P_\infty + m_2P_{0,0})$.

Proposition 5 *The exponent of a two-point Hermitian kernel G_{H2} over $\mathbb{F}_{q^2}$ is bounded below by*

$$E(G_{H2}) \geq \frac{1}{(q^3-1)\ln(q^3-1)} \ln \left[\left(\frac{(q^3-1)(q^3-2q+1)!}{\left(\prod_{i=2}^{q-2}\left(\prod_{k=i}^{q-2}(q^3-iq-k)\right)\right)(q^3-(q^2-q+1)q+1)!} \right) \left(\prod_{i=0}^{q-2} \left(\frac{((q-i)(q-1)-i)!}{((q-i)(q-1)-(q-1))!} \right) \right) \left(\prod_{i=1}^{q-2} \left[(q-i)(q-1)\right]^i \right) \left(\prod_{i=2}^{q-1} ((q-1)q) \right) \right].$$

Proof Let $A = \alpha_i P_\infty + (q-1)P_{0,0}$ and $D = P_1 + \cdots + P_n$. We will use the notation $C(\alpha_i, q-1)$ for $C(D, A)$. Note that $\alpha_i = \{sq-1, \ldots, s(q+1) \mid s = 0, 1, \ldots, q-2\}$ and $\alpha_i = i + g - q$ for $g \leq i \leq n-g$. In addition, by Homma and Kim [8], we may determine the minimum distances for $C(\alpha_i, q-1)$. Let $\alpha_i = aq + b$ with $0 \leq b < q$ and $p = q^2 - a$. According to [8], the minimum distances are as follows.

1. If $b = a = 0$, then $d(C(\alpha_i, q-1)) = q^3 - \alpha_i$ - 1.

2. If α_i satisfies either:

   (a) $0 \leq b \leq q-2$ and $q^2 - 1 \leq a \leq b + q^2 + 1$, or

   (b) $b = q-1$ and $q^2 - 1 \leq a \leq q^2 - 2$,

   then $d(C(\alpha_i, q-1)) = q^2 + q - \alpha_i - 2$.

3. If α_i satisfies either:

   (a) $b = 0$ and $1 \leq a \leq q^2 - q$,

   (b) $1 \leq b \leq q-2$ and $b \leq a \leq q^2 - q - 1$, or

   (c) $b = q-1$ and $0 \leq a \leq q^2 - q - 1$,

then $d(C(\alpha_i, q-1)) = q^3 - q - \alpha_i$.

4. If $1 \leq b$, $q \leq p$, and $p + b \leq q$, then $d(C(\alpha_i, q-1)) = pq - q$.

5. If $p \leq q$ and $q < p + b$, then $d(C(\alpha_i, q-1)) = p(q-1) - (b-1)$.

6. If $2 \leq p \leq q - 1$ and $p + b < q$, then $d(C(\alpha_i, q-1)) = p(q-1)$.

7. If $2 \leq p \leq q - 1$ and $p + b = q$, then $d(C(\alpha_i, q-1)) = (p-1)q$.

Using these conditions, we may determine $d(C(\alpha_i, q-1))$ for $1 \leq i \leq n - g$ as shown in the table.

α_i	a	b	p	$d(C(\alpha_i, q-1))$	Condition
0	0	0	q^2	q^3-1	1
$q-1$	0	$q-1$	q^2	q^3-2q+1	3
q	1	0	q^2-1	q^3-2q	3
$q+1$	1	1	q^2-1	q^3-2q-1	3
$2q-1$	1	$q-1$	q^2-1	q^3-3q+1	3
$2q$	2	0	q^2-2	q^3-3q	3
$2q+1$	2	1	q^2-2	q^3-3q-1	3
$2q+2$	2	2	q^2-2	q^3-3q-2	3
$3q-1$	2	$q-1$	q^2-2	q^3-4q+1	3
$3q$	3	0	q^2-3	q^3-4q	3
$3q+1$	3	1	q^2-3	q^3-4q-1	3
$3q+2$	3	2	q^2-3	q^3-4q-2	3
$3q+3$	3	3	q^2-3	q^3-4a-3	3
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$(q-2)q-1$	$q-3$	$q-1$	q^2-q+3	$q^3-(q-1)q+1$	3
$(q-2)q$	$q-2$	0	q^2-q+2	$q^3-(q-1)q$	3
$(q-2)q+1$	$q-2$	1	q^2-q+2	$q^3-(q-1)q-1$	3

α_i	a	b	p	$d(C(\alpha_i, q-1))$	Condition
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$(q-2)q+(q-1)$	$q-2$	$q-1$	q^2-q+2	$q^3-(q-1)q-(q-1)$	3
$(q-1)q$	$q-1$	0	q^2-q+1	q^3-q^2	3
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$(q^2-q-1)q$	q^2-q-1	0	$q+1$	$q^3-(q^2-q)q$	3
$(q^2-q-1)q+1$	q^2-q-1	1	$q+1$	$q^3-(q^2-q)q-1$	3
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$(q^2-q-1)q+(q-1)$	q^2-q-1	$q-1$	$q+1$	$q^3-(q^2-q)q-(q-1)$	3
$(q^2-q)q$	q^2-q	0	q	$q^3-(q^2-q+1)q$	3
$(q^2-q)q+1$	q^2-q	1	q	$q(q-1)$	5
$(q^2-q)q+2$	q^2-q	2	q	$q(q-1)-1$	5
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$(q^2-q)q+(q-1)$	q^2-q	$q-1$	q	$q(q-1)-(q-2)$	5
$(q^2-q+1)q$	q^2-q+1	0	$q-1$	$(q-1)(q-1)$	6
$(q^2-q+1)q+1$	q^2-q+1	1	$q-1$	$(q-2)q$	7
$(q^2-q+1)q+2$	q^2-q+1	2	$q-1$	$(q-1)(q-1)-1$	5
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$(q^2-q+1)q+(q-1)$	q^2-q+1	$q-1$	$q-1$	$(q-1)(q-1)-(q-2)$	5
$(q^2-q+2)q$	q^2-q+2	0	$q-2$	$(q-2)(q-1)$	6
$(q^2-q+2)q+1$	q^2-q+2	1	$q-2$	$(q-2)(q-1)$	6
$(q^2-q+2)q+2$	q^2-q+2	2	$q-2$	$(q-3)q$	7

$(q^2-q+2)q+3$	q^2-q+2	3	$q-2$	$(q-2)(q-1)-2$	5
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$(q^2-q+2)q+(q-1)$	q^2-q+2	$q-1$	$q-2$	$(q-2)(q-1)-(q-2)$	5
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$(q^2-2)q$	q^2-2	0	2	$2(q-1)$	6
$(q^2-2)q+1$	q^2-2	1	2	$2(q-1)$	6
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$(q^2-1)q+(q-3)$	q^2-2	$q-3$	2	$2(q-1)$	6
$(q^2-1)q+(q-2)$	q^2-2	$q-2$	2	q	7
$(q^2-1)q+(q-1)$	q^2-2	$q-1$	2	$2(q-1)-(q-2)$	5

Using these minimum distances, we may bound the exponent by

$$
\begin{aligned}
E(G_{H2}) \geq \frac{1}{(q^3-1)\ln(q^3-1)} \Bigg[& \ln(q^3-1) + \sum_{i=-1}^{1} \ln(q^3-2q-i) + \sum_{i=-1}^{2} \ln(q^3-3q-i) \\
& + \sum_{i=-1}^{3} \ln(q^3-4q-i) + \sum_{i=-1}^{4} \ln(q^3-5q-i) + \ldots + \sum_{i=-1}^{q-3} \ln(q^3-(q-2)q-i) \\
& + \sum_{i=-1}^{q-1} \ln(q^3-(q-1)q-i) + \sum_{i=0}^{q-1} \ln(q^3-(q)q-i) + \sum_{i=0}^{q-1} \ln(q^3-(q+1)q-i) + \ldots \\
& + \sum_{i=0}^{q-1} \ln(q^3-(q^2-q)q-i) + \ln(q^3-(q^2-q+1)q) \\
& + \sum_{i=0}^{q-2} \ln(q(q-1)-i) + \ln((q-1)(q-1)) \\
& + \ln((q-2)q) + \sum_{i=1}^{q-2} \ln((q-1)(q-2)-i) + 2\ln((q-2)(q-1)) \\
& + \ln((q-3)q) + \sum_{i=2}^{q-2} \ln((q-2)(q-1)-i) + \ldots \\
& + (q-2)\ln(2(q-1)) + \ln(q) + \ln(2(q-1)-(q-2)] \\
= & \frac{1}{(q^3-1)\ln(q^3-1)} \Bigg[\ln \Bigg(\Bigg(\frac{(q^3-1)(q^3-2q+1)!}{\left(\prod_{i=2}^{q-2}\left(\prod_{k=i}^{q-2}(q^3-iq-k)\right)\right)(q^3-(q^2-q+1)q+1)!} \Bigg) \\
& \left(\prod_{i=0}^{q-2} \left(\frac{((q-i)(q-1)-i)!}{((q-i)(q-1)-(q-1))!} \right) \right) \left(\prod_{i=1}^{q-2} \left[(q-i)(q-1)\right]^i \right) \left(\prod_{i=2}^{q-1} ((q-1)q) \right) \Bigg) \Bigg] .
\end{aligned}
$$

□

Using bounds on minimum distances, we may compare the bounds on the minimum exponent of two-point Hermitian kernels to Reed-Solomon and one-point Hermitian kernels. Table 4.3 compares these kernels over $\mathbb{F}_9$ and $\mathbb{F}_{16}$. One may note that the kernel and exponent for the two-point Hermitian code $C(D, m_1P_\infty + m_2P_{(0,0)})$ depends on the value of m_2. This dependance gives us freedom when constructing a

	q = 9		q = 16	
Kernel	Exponent	Size	Exponent	Size
Reed-Solomon	0.6474	9×9	0.6914	16×16
One-Point Hermitian	0.6525	27×27	0.7073	64×64
Two-Point Hermitian	0.6622	26×26	0.7235	63×63

Table 4.3: Lower bounds on exponents of Reed-Solomon, One-Point Hermitian, and Two-Point Hermitian kernels over $\mathbb{F}_q$

sequence of nested codes, and we may construct a kernel by picking nested codes that have the largest minimum distance. Tables 4.4 and 4.6 the minimum distance for two-point Hermitain codes for various m_2 over $\mathbb{F}_9$ and $\mathbb{F}_{16}$. Tables 4.5 and 4.8 use these bounds to give bounds for the exponent for two-point Hermitain kernels for various m_2 over $\mathbb{F}_9$ and $\mathbb{F}_{16}$. Included in Tables 4.5 and 4.8 are bounds for the exponent for the kernel constructed by picking the codes that have the largest minimum distance but still are nested. The codes chosen to construct this new kernel, called the mixed kernel, are shown in bold in Tables 4.4 and 4.6. It is also important to note that two-point Hertimian codes may not always give a larger exponent as shown in Table 4.5 for $m_2 = 1$.

Dimension	$m_2 = 3$	$m_2 = 2$	$m_2 = 1$
1	26	26	**26**
2	23	22	**23**
3	21	21	**22**
4	20	20	**20**
5	19	19	**19**
6	18	18	**18**
7	17	17	**17**
8	16	16	**16**
9	15	15	**15**
10	14	14	**14**
11	13	13	**13**
12	12	12	**12**
13	11	11	**11**
14	10	10	**10**
15	9	9	**9**
16	8	8	**8**
17	7	7	**7**
18	6	6	**6**
19	6	**6**	5
20	4	**5**	4
21	3	**4**	4
22	3	**3**	3
23	3	**3**	2
24	2	**2**	2
25	2	**2**	2
26	**2**	1	1

Table 4.4: Minimum distances of two-point Hermitian codes over $\mathbb{F}_9$

	$m_2 = 3$	$m_2 = 2$	$m_2 = 1$
Exponent	0.6556	0.6530	0.6445
Mixed Kernel			
Exponent	0.6622		

Table 4.5: Lower bounds on the exponent of two-point Hermitian codes over $\mathbb{F}_9$

Dimension	$m_2 = 4$	$m_2 = 3$	$m_2 = 2$	$m_2 = 1$
1	63	63	63	**63**
2	59	57	59	**59**
3	56	56	58	**58**
4	55	55	54	**55**
5	54	53	53	**54**
6	52	52	52	**53**
7	51	51	51	**51**
8	50	50	50	**50**
9	49	49	49	**49**
10	48	48	48	**48**
11	47	47	47	**47**
12	46	46	46	**46**
13	45	45	45	**45**
14	44	44	44	**44**
15	43	43	43	**43**
16	42	42	42	**42**
17	41	41	41	**41**
18	40	40	40	**40**
19	39	39	39	**39**
20	38	38	38	**38**
21	37	37	37	**37**
22	36	36	36	**36**
23	35	35	35	**35**
24	34	34	34	**34**
25	33	33	33	**33**
26	32	32	32	**32**
27	31	31	31	**31**
28	30	30	30	**30**
29	29	29	29	**29**
30	28	28	28	**28**
31	27	27	27	**27**
32	26	26	26	**26**
33	25	25	25	**25**
34	24	24	24	**24**
35	23	23	23	**23**
36	22	22	22	**22**
37	21	21	21	**21**
38	20	20	20	**20**
39	19	19	19	**19**

Table 4.6: Minimum distances of two-point Hermitian codes of dimension at most 39 over $\mathbb{F}_{16}$

Dimension	$m_2 = 4$	$m_2 = 3$	$m_2 = 2$	$m_2 = 1$
40	18	18	18	**18**
41	17	17	17	**17**
42	16	16	14	**16**
43	15	15	14	**15**
44	14	14	14	**14**
45	13	13	13	**13**
46	12	12	**12**	12
47	12	**12**	11	11
48	10	**11**	10	10
49	9	**10**	9	10
50	8	**9**	9	8
51	8	**8**	8	7
52	8	**8**	7	6
53	5	**7**	6	6
54	4	**6**	6	6
55	4	**6**	6	4
56	4	**4**	4	3
57	4	**4**	3	3
58	3	**3**	3	3
59	3	**3**	3	3
60	3	**3**	3	2
61	3	**2**	1	2
62	1	**2**	1	2
63	1	**2**	1	2

Table 4.7: Minimum distances of two-point Hermitian codes of dimension at least 40 over $\mathbb{F}_{16}$

	$m_2 = 4$	$m_2 = 3$	$m_2 = 2$	$m_2 = 1$
Exponent	0.7044	0.7178	0.7018	0.7053
Mixed Kernel				
Exponent	0.7235			

Table 4.8: Lower bounds on the exponent of two-point Hermitian codes over $\mathbb{F}_{16}$

4.2.2 Three-point Hermitian kernels and Hermitian triples

A natural next step would be to continue to shorten a kernel until the exponent no longer increases. This shortening process would result in a nesting of multipoint algebraic geometric codes. However, even shortening a two-point Hermitian kernel to a three-point Hermitian kernel proves to be difficult to analyze as not as much is known about three-point Hermitian codes. Since the automorphism group of the Hermitian curve is doubly transitive, the one- and two-point codes $C(D, G)$ do not depend on the choice of support for the divisor G. However, the situation is more intricate for m-point codes with $m > 2$ and leads to the study of triples of rational points on the Hermitian curve.

We will consider triples of rational points on the Hermitian curve X_q which has defining equation $y^q + y = x^{q+1}$ over $\mathbb{F}_{q^2}$. Note, that the associated homogeneous polynomial associated with $f(x, y) = x^{q+1} - y^q - y$ is

$$F(X, Y, Z) = X^{q+1} - Y^q Z - Y Z^q.$$

Let $(a : b : c) := \{(\alpha a, \alpha b, \alpha c) \mid \alpha \in \mathbb{F}_{q^2} \setminus \{0\}\}$. Then F_{q^2}-rational points of X_q are $\{P_{ab} \mid b^q + b = a^{q+1}\}$ and the point at infinity. Thus, the curve X_q is maximal over $\mathbb{F}_{q^2}$, having genus $g = \frac{q(q-1)}{2}$ and $q^3 + 1$ $\mathbb{F}_{q^2}$-rational points. In particular, for each $a \in \mathbb{F}_{q^2}$, there are q distinct $\mathbb{F}_{q^2}$-rational points P_{ab} on X_q and a unique point at infinity P_∞. Let $\mathrm{Aut}(X_q)$ denote the automorphism group of X_q; that it, $\mathrm{Aut}(X_q) = \{\sigma : X_q \to X_q \mid \sigma \text{ is an isomorphism and } \sigma(X_q) = X_q\}$.

We are interested in m-point codes on X_q, which are those of the form

$$C\left(D, \sum_{i=1}^{m} a_i Q_i\right).$$

When $m = 1, 2$, the codes are well-understood and do not depend on the choice of Q_i (see, for instance, [8, 18, 30, 29]). Hence, we consider the case $m = 3$.

Given a prime power q, let

$$T_q := \{(P, Q, R) : P, Q, R \text{ are distinct } \mathbb{F}_{q^2}\text{-rational points on } X_q\},$$

be the set of triples of distinct $\mathbb{F}_{q^2}$-rational points on X_q. Define $\sim$ on T_q by

$$(P, Q, R) \sim (P', Q', R')$$

if and only if $C(D, aP + bQ + cR)$ is isometric to $C(D, aP' + bQ' + cR')$ for all $a, b, c \in \mathbb{N}$ with respect to the Hamming distance. It is immediate that $\sim$ is an equivalence relation on T_q. We wish to determine the equivalence classes of $\sim$ in pursuit of the classification of families of three-point Hermitian codes.

Proposition 6 *[13, Little] Let $(\beta : \delta : \gamma)$ and $(\lambda : \mu : \nu)$ be any two distinct rational points on X_q, and let $\varepsilon \in \mathbb{F}_{q^2}^*$ be any nonzero field element. There exist an automorphism σ of X_q induced by the linear mapping on $\mathbb{P}^2$ defined by the following matrix*

$$M = \begin{bmatrix} \varepsilon(\delta^q\mu^q - \gamma^q\nu^q) & \varepsilon^{q+1}\xi\lambda & \beta \\ \varepsilon(\beta^q\mu^q - \gamma^q\lambda^q) & \varepsilon^{q+1}\xi\mu & \delta \\ \varepsilon(\delta^q\lambda^q - \beta^q\nu^q) & \varepsilon^{q+1}\xi\nu & \gamma \end{bmatrix}$$

where $\xi = -\lambda^q\beta + \nu^q\gamma + \mu^q\delta$. Moreover, every element of $Aut(X_q)$ can be written in this form for some choice of the two rational points and ε.

Lemma 7 *[16, Theorem 5.4] Each equivalence class of $\sim$ contains a triple of the*

form $(P_{0,0}, P_\infty, P_{1,b})$ *or* $(P_{0,0}, P_\infty, P_{0,b})$. *Consequently, there are at most* $q+1$ *classes of triples of rational points on the Hermitian curve over* $\mathbb{F}_{q^2}$.

Note that we can reduce the number of classes of triples even further. One can also show that there exists $\sigma \in \mathrm{Aut}(X_q)$ such that

$$(\sigma(P_{0,0}), \sigma(P_\infty), \sigma(P_{1,b})) = (P_{0,0}, P_\infty, P_{1,c})$$

if and only if b and c are congugates under the map $b \leftrightarrow b^q$.

The next result immediately follows.

Theorem 17 *There are at most* $\left\lfloor \frac{q+1}{2} \right\rfloor + 1$ *classes of triples of rational points on the Hermitian curve over* $\mathbb{F}_{q^2}$.

We now turn to the Weierstrass semigroup to gain further information on three-point Hermitian codes. Recall that the Weierstrass semigroup of an m-tuple $(P_1, \ldots, P_m)$ on X is

$$H(P_1, \ldots, P_m) = \left\{ (\alpha_1, \ldots, \alpha_m) \in \mathbb{N}^m : \exists f \in F \text{ with } (f)_\infty = \sum_{i=1}^{m} \alpha_i P_i \right\}.$$

Hence, $(\alpha_1, \ldots, \alpha_m) \in H(P_1, \ldots, P_m)$ if and only if

$$\mathcal{L}\left(\sum_{i=1}^{m} \alpha_i P_i \right) \neq \mathcal{L}\left((\alpha_j - 1) P_j + \sum_{\substack{i=1 \\ j \neq i}}^{m} \alpha_i P_i \right)$$

for all $j, 1 \leq j \leq m$. Taking $m = 3$ and $(P, Q, R) \in T_q$, we see that $(P, Q, R) \sim (P', Q', R')$

implies $H(P,Q,R) = H(P',Q',R')$. Therefore,

$$\#\{H(P,Q,R) : (P,Q,R) \in T_q\} \leq \#\text{ equivalence classes of triples in } T_q \leq \left\lfloor \frac{q+1}{2} \right\rfloor + 1.$$

In light of this, it is interesting to consider the number of distinct Weierstrass semigroups $H(P,Q,R)$ for $(P,Q,R) \in T_q$. Let $\mathbb{Z}_+$ denote the set of positive integers. Note we may define a partial ordering on $\mathbb{Z}_+^n$ with the relation $\leq$ defined as: $(x_1,\ldots,x_n) \leq (y_1,\ldots,y_n)$ if and only if $x_i \leq y_i$ for all $i \in 1,\ldots,n$. When we refer to a minimal element of $\mathbb{Z}_+^n$ (or a subset of $\mathbb{Z}_+^n$), we mean with respect to this partial ordering. Clearly, $H(P,Q,R)$ is completely determined by P and Q, and the set

$$\Gamma(P,Q,R) := \left\{ v \in \mathbb{Z}_+^3 : \begin{array}{l} v \text{ is minimal in } \{p \in H(P,Q,R) : p_i = v_i\} \\ \text{for some } i, 1 \leq i \leq 3 \end{array} \right\}.$$

If the rational points P, Q, and R are collinear, then $\Gamma(P,Q,R)$ is known [15]; otherwise, for small values of q, one can compute $\Gamma(P,Q,R)$ using kash [6] and Sage [25]. As a result, we obtain the number of Weierstrass semigroups as displayed in Table 4.2.2 [16]. Examples of $\Gamma(P,Q,R)$ for $q = 4$ and 5 are given below.

q	upper bound on # of equivalence classes of triples	# Weierstrass semigroups
4	2	3
5	4	4
7	4	5
8	3	5
9	4	6

Table 4.9: Number of Weierstrass semigroups of triples on the Hermitian curve X_q over $\mathbb{F}_{q^2}$

Example 15 *Consider the Hermitian curve X_4 given by $y^4 + y = x^5$ over $\mathbb{F}_{16}$. Let*

α *be a primitive element of* $\mathbb{F}_{16}$. *One can compute that*

$$\Gamma\left(P_{0,0}, P_\infty, P_{1,\alpha}\right) = \left\{ \begin{array}{l} (1,1,7),(1,2,6),(1,6,2),(1,7,1),(2,1,6),(2,2,3), \\ (2,3,3),(2,6,1),(3,2,2),(6,1,2),(6,2,1),(7,1,1) \end{array} \right\},$$

whereas

$$\Gamma\left(P_{0,0}, P_\infty, P_{0,1}\right) = \{(1,1,6),(1,6,1),(2,2,2),(6,1,1)\}$$

Then there are exactly two Weierstrass semigroups of triples of $\mathbb{F}_{16}$*-rational rational points on* X_4*:*

$$H(P_{0,0}, P_\infty, P_{0,1}) \neq H(P_{0,0}, P_\infty, P_{1,\alpha}).$$

Thus, there are at least two equivalence classes of triples over $\mathbb{F}_{16}$*, and at most* $\left\lfloor \frac{4+1}{2} \right\rfloor + 1 = 3$ *equivalence classes of triples over* $\mathbb{F}_{16}$.

Example 16 *Consider the Hermitian curve* X_5 *given by* $y^5 + y = x^6$ *over* $\mathbb{F}_{25}$. *Let* α *be a primitive element of* $\mathbb{F}_{25}$. *Then there are exactly four Weierstrass semigroups of triples of* $\mathbb{F}_{25}$*-rational points on* X_5*:*

$$H(P_{0,0}, P_\infty, P_{0,1}), H(P_{0,0}, P_\infty, P_{1,\alpha^4}), H(P_{0,0}, P_\infty, P_{1,\alpha^6}), H(P_{0,0}, P_\infty, P_{1,\alpha^7}).$$

Indeed, $(2,2,8) \in \Gamma\left(P_{0,0}, P_\infty, P_{1,\alpha^4}\right)$ *whereas*

$$(2,2,9) \in \Gamma\left(P_{0,0}, P_\infty, P_{1,\alpha^6}\right) \cap \Gamma\left(P_{0,0}, P_\infty, P_{1,\alpha^7}\right).$$

In addition, $\Gamma\left(P_{0,0}, P_\infty, P_{1,\alpha^6}\right) \setminus \Gamma\left(P_{0,0}, P_\infty, P_{1,\alpha^7}\right) = \{(3,3,4),(3,4,3),(4,3,3)\}$, *and* $\Gamma\left(P_{0,0}, P_\infty, P_{1,\alpha^7}\right) \cap \Gamma\left(P_{0,0}, P_\infty, P_{1,\alpha^6}\right) = \{(3,3,3)\}$. *In addition, there are at most* $\left\lfloor \frac{5+1}{2} \right\rfloor + 1 = 4$ *equivalence classes of triples over* $\mathbb{F}_{25}$. *Thus, we conclude that there are exactly four equivalence classes of triples over* $\mathbb{F}_{25}$.

Chapter 5

Conclusions and Discussion

In this thesis, we provide applications of algebraic geometric codes to stopping sets and to polar codes. We also connect open problems concern triples of rational places on the Hermitian curve to polar coding and three-point Hermitian kernels. However, there are still many open questions and problems to be considered in all three of these topics.

For stopping sets, it remains to be shown if stopping sets of sizes $m-1$ and m exist for algebraic geometry codes over hyperelliptic curves of genus 2 as well as the existence of stopping sets consisting of groups of noncollinear points of Hermitian codes. In addition, stopping sets of algebraic geometric codes from other curves remain to be studied.

There are also still many open questions in polar coding. For example, further investigation of Hermitian triples leads to greater insight to three-point Heritian kernels and shortening. In addition, shortening may increase the exponent while decreasing the size of the kernel; however, shortening may also decrease the exponent. A hybrid approach might allow one to balance these competing goals. There are many other open problems not discussed in this thesis, such as finite block-length analysis

and other decoding techniques. It would be interesting to see how one might use the underlying algebraic structure of the kernel in decoding.

Bibliography

[1] E. Arikan, Channel polarization: a method for constructing capacity-achieving codes for symmetric binary-input memoryless channels, IEEE Trans. Inform. Theory 55 (2009), no. 7, 3051–3073.

[2] E. Arikan and E. Telatar, On the rate of channel polarization, IEEE ISIT, Seoul, South Korea, 28 June - 3 July 2005, 1493 – 1495.

[3] E. Ballico and A. Ravagnani, A zero-dimensional cohomological approach to Hermitian codes, J. Pure Appl. Algebra 219 (2015), no. 4, 1031–1044.

[4] C. Carvalho and F. Torres, On Goppa codes and Weierstrass gaps at several points, Des. Codes and Cryptog. 32 (2005), 211–225.

[5] C. Chen and I. Duursma, Geometric Reed-Solomon codes of length 64 and 65 over $\mathbb{F}_8$, IEEE Trans. Inform. Theory 49 (2003), no 5, 1351–1353.

[6] M. Daberkow, C. Fieker, J. Klüners, M. Pohst, K. Roegner and K. Wildanger, *KANT V4*, in J. Symbolic Comp. 24 (1997), 267–283.

[7] T. Høholdt, J. H. van Lint, and R. Pellikaan, Algebraic-geometric codes, in *Handbook of Coding Theory* 1 (1998), V. Pless, W. C. Huffman, and R. A. Brualdi, Eds., Elsevier, Amsterdam, 871–961.

[8] M. Homma and S. J. Kim, The Complete Determination of the Minimum Distance of Two-Point Codes on a Hermitian Curve, Des. Codes and Cryptog. 40 (2006), 5–24.

[9] M. Homma and S. J. Kim, Goppa Codes with Weierstrass Pairs, J. Pure Appl. Algebra 162 (2001), no. 2, 273–290.

[10] V. D. Goppa. Algebraic-geometric codes. Math. USSR-Izv., 21:7591, 1983.

[11] V. D. Goppa. Geometry and Codes. Kluwer, 1988.

[12] S. Korada, E. Şaşoğlu, and R. Urbanke, Polar codes: characterization of exponent, bounds, and constructions, IEEE Trans. Inform. Theory 56 (2010), no. 12, 6253–6264.

[13] J. Little, private communication, 2002.

[14] N. Goela, S. Korada, and M. Gastpar, On LP decoding of polar codes, IEEE Inform. Theory Workshop, Dublin, Ireland, 30 Aug - 3 Sept 2010, 1–5.

[15] H. Maharaj, G. Matthews, and G. Pirsic, Riemann-Roch spaces of the Hermitian function field with applications to algebraic-geometric codes and low-discrepancy sequences, J. Pure Appl. Algebra 195 (2005), no. 3, 261–280.

[16] J. Marshall, On the number of Weierstrass semigroups of triples on the Hermitian curve, M.S. thesis, Clemson University, 2007.

[17] E. Martinez-Moro, Advances in algebraic geometry codes, World Scientific, Vol. 5., 2008.

[18] G. L. Matthews, Weierstrass pairs and minimum distance of Goppa codes, Des. Codes and Cryptog. 22 (2001), 107–121.

[19] R. Mori and T. Tanaka, Channel Polarization on q-ary discrete memoryless channels by arbitrary kernels, IEEE ISIT, Austin, Texas, 13 June - 18 June 2010, 894 – 898.

[20] R. Mori and T. Tanaka, Non-binary Polar codes using Reed-Solomon codes and algebraic geometry codes, IEEE Inform. Theory Workshop, Dublin, Ireland, 30 Aug - 3 Sept 2010, 1–5.

[21] R. Mori and T. Tanaka, Source and channel polarization over finite fields and Reed-Solomon matrices, IEEE Trans. Inform. Theory 60 (2014), no 5, 2720–2736.

[22] W. Park and A. Barg, Polar codes for q-ary channels, $q = 2^r$, IEEE Trans. Inform. Theory 59 (2013), no 2, 955–969.

[23] E. Şaşoğlu, Polar Coding Theorems for Discrete Systems, Ph.D. dissertation, Ecole Polytechnique Federale de Lausanne, 2011.

[24] E. Şaşoğlu, E. Telatar, and E. Arkan, Polarization for arbitrary discrete memoryless channels, IEEE Inform. Theory Workshop, Taormina, Italy, Oct. 2009, 144–148.

[25] William A. Stein et al. Sage Mathematics Software (Version 5.11), The Sage Development Team, 2014, http://www.sagemath.org.

[26] H. Stichtenoth, *Algebraic Function Fields and Codes*, Springer, Berlin, 1993.

[27] I. Tal and A. Vardy, How to construct polar codes, IEEE Trans. Inform. Theory 59 (2013), no. 10, 6562–6582.

[28] I. Tal and A. Vardy, List decoding of polar codes, IEEE ISIT, Saint-Petersburg, Russia, 31 July - 5 August 2011, 1–5.

[29] K. Yang and P. V. Kumar, On the true minimum distance of Hermitian codes, Coding Theory and algebraic-geometric, Proceedings, Luminy, 1991, Lecture Notes in Mathematics 1518 (1992), Springer-Verlag, 99–107.

[30] C.-P. Xing, On Automorphism Groups of the Hermitian Codes, IEEE Trans. Inform. Theory 41 (1995), no. 6, 1629–1635.

[31] J. Zhang, F.W. Fu, and D. Wan, Stopping sets of algebraic geometry codes, IEEE Trans. Inform. Theory 60 (2013), no. 3, 1488–1495.