**Noname manuscript No.**
(will be inserted by the editor)

# Exponents of polar codes using algebraic geometric code kernels

**Sarah E. Anderson · Gretchen L. Matthews**

Received: date / Accepted: date

**Abstract** Reed-Solomon and BCH codes were considered as kernels of polar codes by Mori and Tanaka (IEEE Inform. Theory Workshop, Dublin, Ireland, 30 Aug. - 3 Sept. 2010, 1–5) and Korada, Şaşoğlu, and Urbanke (IEEE Trans. Inform. Theory vol. 56, no. 12, 6253–6264 (2010)) to create polar codes with large exponents. Mori and Tanaka showed that Reed-Solomon codes over the finite field $\mathbb{F}_q$ with $q$ elements give the best possible exponent among all codes of length $l \leq q$. They also stated that a Hermitian code over $\mathbb{F}_{2^r}$ with $r \geq 4$, a simple algebraic geometric code, gives a larger exponent than the Reed-Solomon matrix over the same field. In this paper, we expand on these ideas by employing more general algebraic geometric codes to produce kernels of polar codes. Lower bounds on the exponents are given for kernels from general algebraic geometric codes, Hermitian codes, and Suzuki codes. We demonstrate that both Hermitian and Suzuki kernels have larger exponents than Reed-Solomon codes over the same field, for $q \geq 3$; however, the larger exponents are at the expense of larger kernel matrices. Comparing kernels of the same size, though over different fields, we see that Reed-Solomon kernels have larger exponents than both Hermitian and Suzuki kernels. These results indicate a tradeoff between the exponent, kernel matrix size, and field size.

**Keywords** Polar codes · kernel · exponent · maximal function fields · algebraic geometric codes

S. Anderson
Department of Mathematical Sciences
Clemson University
Clemson, South Carolina 29634–0975
E-mail: sarah5@g.clemson.edu

G. Matthews
Department of Mathematical Sciences
Clemson University
Clemson, South Carolina 29634–0975
E-mail: gmatthe@g.clemson.edu

**Mathematics Subject Classification (2000)** 14G50 · 94B27

## 1 Introduction

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

$$G_2 := \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

for encoding a block of $2^n$ channels; $G_2$ is called the kernel matrix. As the number of channels grows, each channel becomes either a noiseless channel or a pure-noise channel. Korada, Şaşoğlu, and Urbanke considered larger binary matrices as kernels and characterized the speed of polarization by introducing a quantity called the exponent [4]. Şaşoğlu also explored the polarization phenomenon for nonbinary alphabets [10], and polar codes were generalized to arbitrary discrete memoryless channels by Şaşoğlu, Telatar, and Arikan [11]. Mori and Tanaka generalized the arguments of Korada et. al. to general finite fields [6] 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 [7]. 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. [4], suggests algebraic geometric codes are good candidates for constructing kernel matrices.

In this paper, we expand on these ideas by employing more general algebraic geometric codes to produce kernels of polar codes. We consider a $q$-ary discrete memoryless channel $W : \mathcal{X} \to \mathcal{Y}$ with input alphabet the finite field with $q$ elements, where $q$ is a prime power. A summary of notation is provided at the end of this section. Section 2 provides background on polar coding in this setting where a $q$-ary matrix serves as the kernel. Prerequisite material on algebraic geometric (AG) codes is found in Section 3. In Section 4, we construct kernels from AG codes, and consider examples from the Hermitian and Suzuki function fields. Lower bounds on the exponents are given for kernels from general algebraic geometric codes, Hermitian codes, and Suzuki codes. In addition, we demonstrate that the kernel obtained by shortening a one-point AG code yields a kernel associated with a multipoint AG code. Finally, in Section 5, we consider the best achievable probability of block error for polar coding over $W$ using an arbitrary kernel $G$ under successive cancellation (SC) decoding.

**Notation:** Given a prime power $q$, $\mathbb{F}_q$ denotes the finite field with $q$ elements, and $\mathbb{F}_q^n$ denotes 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 $u$. For $1 \leq i \leq j \leq n$, it is often convenient to write $u_i^j := (u_i, \ldots, u_j) \in \mathbb{F}_q^{j-i+1}$.

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 $Row_i A$ 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 $Col_j A$. We use the notation $\otimes$ to denote the Kronecker product; that is, the Kronecker product of two matrices $A$ and $B$ with entries in a field $\mathbb{F}$ is written $A \otimes B$, and $A^{\otimes n} := \underbrace{A \otimes \cdots \otimes A}_{n}$.

## 2 Background

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

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

The rate and reliability are given by the capacity

$$I(W) = \sum_{y \in \mathcal{Y}} \sum_{x \in \mathcal{X}} \frac{1}{q} W(y|x) \log_q \left( \frac{W(y|x)}{\frac{1}{q} \sum_{x' \in \mathcal{X}} W(y|x')} \right) \tag{1}$$

and the Bhattacharyya parameter

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

of $W$.

Let $G$ be a $l \times l$ matrix over $\mathbb{F}_q$. A block of $N = l^n$ channels is produced from $W$ by combining and splitting channels as we describe now. 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

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

where $B_N$ is an $N \times N$ permutation matrix that sends $u_1^N$ to

$$(u_1, u_{l+1}, \ldots, u_{N-(l-1)}, u_2, u_{l+2}, \ldots, u_{N-(l-2)}, \ldots, u_l, u_{2l}, \ldots, u_N).$$

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

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

As we discuss below, the properties of the channels $W_N^{(i)}$ depend on the matrix $G$, which is called the **kernel**, and on $q$.

Take $q = 2$ and $G = G_2$ as in Section 1. This yields Arikan's original construction. Here, for fixed $\delta > 0$, 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 and is demonstrated by Arikan [1].

It is important to note that not all matrices polarize a given channel. 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 1** *[6, 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 1** *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 1 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. □

**Theorem 2** *Let $q$ be a prime power and $\mathcal{X}$ be a finite field of order $q$. 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 1, 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 1. 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 $Row_k G$ 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 $Row_k G$ using the rows below $k$, the only entries affected are entries to the right of the diagonal since rows of $G$ below $Row_k G$ 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 $Row_k(G_{kk}^{-1}G)$, then $L$ satisfies Lemma 1. Therefore, $G$ polarizes $W$ as $L$ does. □

A natural question to consider is if Theorem 2 provides a characterization of those matrices which polarize $q$-ary DMCs. When $q$ is a prime, this is the case, as demonstrated in Theorem 3 below. When $q$ is a prime power that is not prime, whether or not a matrix polarizes a channel is channel dependent; Example 1 illustrates this. A necessary and sufficient condition for a matrix over a finite field of size $q$, where $q$ is any prime power, to polarize any $q$-ary DMC is given in [8].

**Theorem 3** *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 [4, Lemma 1], we see that $L$ does not polarize $W$. Hence $G$ does not polarize $W$.

*Example 1* 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 $\alpha$ is a primitive element of $\mathbb{F}_4$. According to Equation (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 2. Calculating capacities as in Equation (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 $\alpha$ 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 [11].

The rate of polarization of a kernel is known as the exponent and its definition is below. In preparation, we define a random variable $W'_n$ that is uniformly distributed over $\left\{W_N^{(i)}\right\}$, $1 \le i \le N$. Consider a sequence of independent and identically distributed (i.i.d.) random variables $\{B_n | n \ge 1\}$, where $B_n$ is uniformly distributed over the set $\{1, \ldots, l\}$. Let $Z_n := Z(W'_n)$ where the channels $W'_i$, $i \ge 0$, are defined recursively by

$$W'_0 = W, W'_1 = W_l^{(B_1)}, \text{ and } W'_{n+1} = (W'_n)_l^{(B_{n+1})}.$$

**Definition 1** [6] 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 $G$ (or of the polar code with kernel $G$) is the value $E(G)$ such that

- for any fixed $\beta < E(G)$,

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

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

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

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 exist a sequence $\{\mathcal{A}_N\}$ of sets $\mathcal{A}_N \subseteq \{1, \ldots, N\}$ such that $|\mathcal{A}_N| \geq NR$ and

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

Information is then sent across the channels indexed by elements of $\mathcal{A}_N$.

The exponent of a matrix can be found using partial distances, a method introduced Korada, Şaşoğlu, and Urbanke for the binary case [4] and explored for larger alphabets by Mori and Tanaka [6].

**Definition 2** 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 2** *[4, 5] If $G$ is an $l \times l$ matrix, then the exponent of the polar code with kernel $G$ is*

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

Given an $l \times l$ matrix $G$ as in Definition 2, let $C_i := \langle g_{l-i+1}, \ldots, g_l \rangle$ for $0 \leq i \leq l$ and suppose $C_i \neq C_{i+1}$ for all $i$. Then

$$C_1 \subseteq C_2 \subseteq \cdots \subseteq C_l.$$

Notice that $D_i = d\left(g_i, C_{l-i}\right)$. Because $D_i = \min\{d(g_i, c) : c \in C_{l-i}\}$,

$$D_i \geq d\left(C_{l-i+1}\right).$$

It follows that

$$E(G) \geq \frac{1}{l} \sum_{i=1}^{l} \log_l\left(d\left(C_{l-i+1}\right)\right),$$

providing a lower bound on the exponent based on the minimum distances of the codes $C_i$. As the next result demonstrates, it can happen that $D_i > d\left(C_{l-i+1}\right)$; an instance of this is given in Example 2.

**Proposition 1** *Consider an $l \times l$ matrix $G = [g_1^T, \ldots, g_l^T]^T$ over $\mathbb{F}_q$, and let $C_i := \langle g_{l-i+1}, \ldots, g_l \rangle$ for $1 \leq i \leq l$. Assume $C_i \neq C_{i+1}$ for all $i$, $1 \leq i \leq l-1$. If there exists $i$ such that $d(C_{l-i}) = d(C_{l-i+1})$ and $C_{l-i}$ and $C_{l-I+1}$ have precisely the same minimum weight words, then*

$$D_i > d\left(C_{l-i+1}\right).$$

*Proof* By definition , $D_i \geq d\left(C_{l-i+1}\right)$. Let $d = d\left(C_{l-i+1}\right)$, and suppose $D_i = d$. Then there exists $c \in C_{l-i}$ with $d(g_i, c) = d$. Then

$$w := g_i - c \in C_{l-i+1}.$$

Because $wt(w) = d$ and the codes $C_{l-i}$ and $C_{l-i+1}$ have the same minimum weight words,

$$w \in C_{l-i}.$$

Then

$$g_i = c + w \in C_{i-1},$$

which is a contradiction as $\dim \langle g_{i+1}, \ldots, g_l \rangle < \dim \langle g_i, \ldots, g_l \rangle$ . Hence, $D_i > d$. □

According to Lemma 2, Arikan's original kernel matrix $G_2$ has exponent $E(G_2) = \frac{1}{2}$. In [4], a matrix obtained from a generator matrix for a shortened BCH code is found to have exponent greater than $\frac{1}{2}$. 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 section, algebraic geometric codes lend themselves naturally to this construction.

## 3 Algebraic geometric codes

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

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

where

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

is the Riemann-Roch space of $A$. If $|supp(A)| = 1$, then $C(D, A)$ is called a one-point code; otherwise, $C(D, A)$ is known as a multipoint code. 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 $C(D, A)$. The AG code $C(D, A)$ is an $[n, k, d]$ code where $d \geq n - deg(A)$,

$$k = l(A) - l(A - D),$$

and $l(A) := dim(\mathcal{L}(A))$. If $2g-1 \leq deg(A) < n$, then $k = deg(A) + 1 - g$. Moreover, $C(D, A)$ satisfies a Singleton-like bound

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

An especially useful property of AG codes is their "nested" structure. If $A = \sum_{Q_i \in \mathbb{P}_F} a_i Q_i$ and $B = \sum_{Q_i \in \mathbb{P}_F} b_i Q_i$ are divisors on $F$, where $\mathbb{P}_F$ is the set of all places of $F/\mathbb{F}_q$, then $A \leq B$ if $a_i \leq b_i$ for all $i$. Given divisors $A$ and $B$ with $supp(A) \cap supp(D) = \emptyset = supp(B) \cap supp(D)$,

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

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

## 4 Construction of kernels using AG codes

### 4.1 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 $n \geq 2g$. Construct a sequence of divisors

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

so that the supports of $D := P_1 + \cdots + P_n$ and the $A_j$, $1 \leq j \leq n$, are disjoint and

$$C(D, A_1) \subsetneqq C(D, A_2) \subsetneqq \cdots \subsetneqq C(D, A_n) = \mathbb{F}_q^n. \tag{2}$$

Let $G$ be 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 (2) 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

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

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 $2g-1$. Finally, notice that for all positive integers $\alpha$,

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

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 (2). 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 divisors satisfying (2) for all $i$.

**Theorem 4** *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*

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

*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

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

This nested structure of codes allows us to bound the partial distances $D_i$ of $G$ by

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

This bound combined with Lemma 2 yields the desired result. □

*Remark 1* 1. Theorem 4 provides a lower bound on the exponent, as the partial distances associated with these matrices are not necessarily nondecreasing even though the minimum distances of the associated codes are. 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 space. This structure may prove useful in further studies, such as into shortened AG code kernels.

2. One may use multi-level code construction with the AG code kernels constructed above; however, a manipulation of the kernel will ensure that these kernels polarize according to Theorem 2. 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 $\alpha$ 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 2. The exponent of $G'$ may be bounded as well. Indeed, note that the proof of Theorem 4 applies except for the term $d_n$. Even so, $D_n \geq 2$ since $G'_{nj} = \alpha$ and $G'_{nn} = 1$.

An immediate corollary of Theorem 4 is the exponent of a kernel based on a Reed-Solomon code; this is computed by Mori and Tanaka [6]. 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 $\mathbb{F}_q$.

**Corollary 1** *The exponent of a Reed-Solomon kernel $G_{RS}$ over $\mathbb{F}_q$ is*

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

*Proof* This follows directly from Theorem 4 using the fact that $F$ has genus $g = 0$. □

Another consequence of Theorem 4 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 $F$ of degree one is $q + 1 + 2g\sqrt{q}$.

**Theorem 5** *Let $F/\mathbb{F}_q$ be a maximal function field of genus $g$, and let $G$ be a generator matrix of an AG code on $F$ of length $n = q + 2g\sqrt{q}$ constructed as in (2). Then*

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

*Proof* Because $F/\mathbb{F}_q$ is a maximal function field,

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

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,

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

Then

$$\begin{aligned} E(G) &\geq \frac{1}{n\ln(n)} \left( \ln((n-g)!) + \sum_{i=n-g+1}^{n} \ln(d_1) \right) \\ &\geq \frac{\ln((n-g)!)}{n\ln(n)} \\ &\geq \frac{1}{q^{3/2}\ln(q^{3/2})} \ln\left( \left( q^{3/2} - \frac{3q}{2} + \frac{q^{1/2}}{2} \right)! \right) \\ &= \frac{1}{q^{3/2}\ln(q^{3/2})} \sum_{i=2}^{q^{3/2} - \frac{3q}{2} + \frac{q^{1/2}}{2}} \ln(i) \\ &\geq \frac{1}{q^{3/2}\ln(q^{3/2})} \int_1^{q^{3/2} - \frac{3q}{2} + \frac{q^{1/2}}{2}} \ln(x)dx \\ &= \frac{q^{3/2} - \frac{3q}{2} + \frac{q^{1/2}}{2}}{q^{3/2}\ln(q^{3/2})} \ln\left( q^{3/2} - \frac{3q}{2} + \frac{q^{1/2}}{2} \right) - \frac{q^{3/2} - \frac{3q}{2} + \frac{q^{1/2}}{2} - 1}{q^{3/2}\ln(q^{3/2})} \\ &= \left(1 - \frac{3}{2q^{1/2}} + \frac{1}{2q}\right) \frac{\ln\left(q^{3/2} - \frac{3q}{2} + \frac{q^{1/2}}{2}\right)}{\ln(q^{3/2})} - \frac{1}{\ln(q^{3/2})} + \frac{3}{2q^{1/2}\ln(q^{3/2})} \\ &\qquad - \frac{1}{2q\ln(q^{3/2})} + \frac{1}{q^{3/2}\ln(q^{3/2})}. \end{aligned}$$

Therefore,

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

□

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

### 4.2 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; $F$ is known as the Hermitian function field. 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 [7]. Applying Theorem 4 and the exact distances of one-point Hermitian codes [15] 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 (2).

**Corollary 2** *The exponent of a Hermitian kernel $G_H$ over $\mathbb{F}_{q^2}$ is bounded below by*

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

*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 Section 4.1 is

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

where $g = \frac{q(q-1)}{2}$ and $[a]^t$ denotes the multiset $\{a,\ldots,a\}$ of cardinality $t$ [15]. Hence, $E(G)$ 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

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

Thus,

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

□

As the next example demonstrates, the bound in Corollary 2 only provides a lower bound on the exponent, not its exact value, for $q \geq 4$.

*Example 2* Let $q \geq 4$, and consider a kernel $G_H$ constructed as above. Doing so, one may take $\alpha_{n-5} = q^3+q^2-3q-3$ and $\alpha_{n-4} = q^3+q^2-3q-2$, where $n = q^3$. Then

$$C\left(D, \left(q^3+q^2-3q-3\right)P_\infty\right) \subsetneqq C\left(D, \left(q^3+q^2-3q-2\right)P_\infty\right);$$

indeed,

$$\dim C\left(D, \left(q^3+q^2-3q-3\right)P_\infty\right) = q^3-5$$

whereas

$$\dim C\left(D, \left(q^3+q^2-3q-2\right)P_\infty\right) = q^3-4;$$

see [12] for details. Both codes have minimum distance $d = 3$ [15]. According to [5], the two codes have the same minimum weight words. Taking $\alpha_{n-5} = q^3+q^2-3q-3$ and $\alpha_{n-4} = q^3+q^2-3q-2$ and constructing the kernel matrix $G_H$ as above, we see that

$$D_5 > d\left(C\left(D, \alpha_{n-4}P_\infty\right)\right).$$

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, provided $q \neq 2$. Indeed, the proof follows immediately from Theorem 4, Corollary 1, and Corollary 2.

**Proposition 2** *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 $q \geq 3$.*

| | 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 1** Lower bounds on exponents of Reed-Solomon and Hermitian kernels over $\mathbb{F}_{q^m}$

*Remark 2* It should be observed that the kernel matrices in Proposition 2 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 2.

### 4.3 Kernels from Suzuki codes

In this subsection, we see that the asymptotic behavior of the exponent in Theorem 5 is not restricted to maximal function fields. 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 with defining equation

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

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 4 then yields the following result.

**Corollary 3** *Let $G_{Suz}$ be a Suzuki kernel over $\mathbb{F}_q$ where $q = 2^{2r+1}$. Then*

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

The exact minimum distances of Suzuki one-point codes over $\mathbb{F}_8$ are known according to the work of Chen and Duursma [2]. We further explore this function field in the example below.

*Example 3* 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 $\alpha$ be a primitive element of $\mathbb{F}_8$.

| | q = 8 | | q = 32 | |
|---|---|---|---|---|
| Kernel | Exponent | Size of Kernel | Exponent | Size of Kernel |
| Reed-Solomon | 0.63747 | $8 \times 8$ | 0.73540 | $32 \times 32$ |
| Suzuki | 0.65555 | $64 \times 64$ | 0.73635 | $1024 \times 1024$ |

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

A one-point code over this function field, called a Suzuki one-point code, is of length 64. The Suzuki one-point codes

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

satisfy condition (2). 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

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

according to [2]. Hence, Theorem 4 implies $E(G_{Suz}(8)) \geq 0.65555$.

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 exponent comes at the price of a larger kernel size. When comparing kernels of similar size (though over different fields), Reed-Solomon kernels give larger exponent than Suzuki kernels.

**Proposition 3** *Let $q = 2^{2r+1}$. 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 $r \geq 1$.*

*Remark 3* It should be observed kernel matrices in Proposition 3 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 3.

The limiting behavior of the exponent in Theorem 5 is not restricted to maximal function fields. In fact, kernels from Suzuki one-point codes display similar asymptotics.

**Theorem 6** *Let $G_{Suz}$ be a Suzuki kernel over $\mathbb{F}_q$ where $q = 2^{2r+1}$. Then*

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

*Proof* Let $G$ be a Suzuki kernel over $\mathbb{F}_q$ where $q = 2^{2r+1}$. 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 $q \to \infty$. □

### 4.4 Shortening an AG code kernel

The method of shortening can be used to create smaller kernels with large exponent. In [4], Korada, Şaşoğlu, and Urbanke used repeated shortening of a BCH code to create the smallest binary kernel with exponent exceeding $\frac{1}{2}$, that of Arikan's original kernel $G_2$.

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 $Row_i G$ to all the rows with a nonzero element in $Col_j G$. Finally, remove $Col_j G$ and $Row_i G$ 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 7** *Let $\alpha_1 \leq \cdots \leq \alpha_n$ be integers such that*

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

*and $G$ be a generator matrix of $C(D, \alpha_n P)$ constructed according to (2). Suppose $G'$ is the matrix obtained by shortening applied to the $j^{th}$ column of $G$. Then*

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

*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

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

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,

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

is a sequence of codes satisfying (2). □

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

*Example 4* Consider a generator matrix $G_H(4)$ for the Hermitian code $C(D, 9P_\infty)$ over $\mathbb{F}_4$ constructed according to (2), where $\alpha$ is a primitive element of $\mathbb{F}_4$ satisfying $\alpha^2 + \alpha + 1 = 0$:

$$G_H(4) = \begin{array}{c} \\ X^3Y \\ X^2Y \\ X^3 \\ XY \\ X^2 \\ Y \\ X \\ 1 \end{array} \begin{array}{c} \begin{array}{cccccccc} (0,0) & (0,1) & (1,\alpha) & (1,\alpha^2) & (\alpha,\alpha) & (\alpha,\alpha^2) & (\alpha^2,\alpha) & (\alpha^2,\alpha^2) \end{array} \\ \begin{pmatrix} 0 & 0 & \alpha & \alpha^2 & \alpha & \alpha^2 & \alpha & \alpha^2 \\ 0 & 0 & \alpha & \alpha^2 & 1 & \alpha & \alpha^2 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & \alpha & \alpha^2 & \alpha^2 & 1 & 1 & \alpha \\ 0 & 0 & 1 & 1 & \alpha^2 & \alpha^2 & \alpha & \alpha \\ 0 & 1 & \alpha & \alpha^2 & \alpha & \alpha^2 & \alpha & \alpha^2 \\ 0 & 0 & 1 & 1 & \alpha & \alpha & \alpha^2 & \alpha^2 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} \end{array}.$$

The columns of $G_H(4)$ are indexed by $(\alpha, \beta)$ 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 $\mathcal{L}(9P_\infty)$. Specifically, let

$$\alpha_i = \begin{cases} 0 & \text{if } i = 1 \\ i & \text{if } 2 \le i \le 7 \\ 9 & \text{if } i = 8. \end{cases}$$

For $1 \le i \le 7$, 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 $G_H(4)$. Since the last row of $G_H(4)$ is the only row with a nonzero entry in the first column, we will remove the last row and the first column of $G_H(4)$. 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 $C(D - P_{0,0}, 9P_\infty - P_{0,0})$.

## 5 Probability of block error

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 [6], then the exponent helps bound the block error probability under successive cancellation (SC) decoding. Let $P_e$ be the best achievable probability of block error under SC decoding for polar coding over $W$ using kernel $G$. Using techniques similar to [1] and [9], the following result holds.

**Theorem 8** *Consider polar coding over a $q$-ary DMC using kernel $G$ at a fixed rate $0 < R < I(W)$ with block length $N = l^n$. 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 $\mathcal{E} = \cup_{i \in \mathcal{A}} \mathcal{B}_i$ 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

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

Thus,

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

Then

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

Hence,

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

□

## 6 Conclusions

AG codes can be used to create kernels for polar codes. Certain AG codes have known Riemann-Roch bases, so we have explicit kernels that may be described simply by basis functions and places. Both Hermitian and Suzuki kernels have larger exponents than Reed-Solomon codes over the same field, for $q \geq 3$; however, the larger exponents are at the expense of larger kernel matrices. Comparing kernels of the same size, though over different fields, we see that Reed-Solomon kernels have larger exponents than both Hermitian and Suzuki kernels. These results indicate a tradeoff between the exponent, kernel matrix size, and field size. Shortening may be used to create smaller kernels, but this might decrease the exponent; multipoint AG code kernels may do the same. A hybrid approach might allow one to balance these competing goals.

## References

1. E. Arikan, Channel polarization: a method for constructing capacity-achieving codes for symmetric binary-input memoryless channels, IEEE Trans. Inform. Theory **55**, no. 7, 3051–3073 (2009).
2. C. Chen and I. Duursma, Geometric Reed-Solomon codes of length 64 and 65 over $\mathbb{F}_8$, IEEE Trans. Inform. Theory **49**, no. 5, 1351–1353 (2003).
3. 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.
4. S. Korada, E. Şaşoğlu, and R. Urbanke, Polar codes: characterization of exponent, bounds, and constructions, IEEE Trans. Inform. Theory **56**, no.. 12, 6253–6264 (2010).
5. C. Marcolla, M. Pellegrini, and M. Sala, On the small weights codewords of some Hermitian codes, preprint.
6. 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.
7. 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.
8. R. Mori and T. Tanaka, Source and channel polarization over finite fields and Reed-Solomon matrices, IEEE Trans. Inform. Theory, to appear.
9. W. Park and A. Barg, Polar codes for $q$-ary channels, $q = 2^r$, IEEE Trans. Inform. Theory **59**, no. 2, 955–969 (2013).
10. E. Şaşoğlu, Polar Coding Theorems for Discrete Systems, Ph.D. dissertation, Ecole Polytechnique Federale de Lausanne, 2011.
11. E. Şaşoğlu, E. Telatar, and E. Arkan, Polarization for arbitrary discrete memoryless channels, in Proc. IEEE Information Theory Workshop, Taormina, Italy, Oct. 2009, pp. 144-148.
12. H. Stichtenoth, Algebraic Function Fields and Codes, Springer, Berlin, 1993.
13. I. Tal and A. Vardy, How to construct polar codes, IEEE Trans. Inform. Theory, **59**, no. 10, 6562–6582 (2013).
14. I. Tal and A. Vardy, List decoding of polar codes, IEEE ISIT, Saint-Petersburg, Russia, 31 July - 5 August 2011, 1–5.
15. K. Yang and P. Kumar, On the true minimum distance of Hermitian codes, Coding theory and algebraic geometry, Lecture no.tes in Mathematics, vol. 1518, 99–107. Springer: Berlin (1992).