Polar codes with large exponent using AG code kernels

Sarah Anderson

Department of Mathematical Sciences
Clemson University

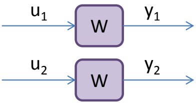
April 14, 2013

Polar codes acheive symmetric capacity of certain channels.

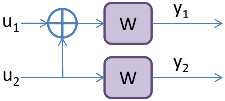
Erdal Arikan introduced polar codes in 2009. Polar codes are a channel dependent construction of symmetric capacity achieving codes for binary DMCs inspired by the chain rule for mutual information, which states

$$\begin{aligned} NI(W) &= I(\mathcal{X}_1^N; \mathcal{Y}_1^N) \\ &= \sum_{i=1}^{N} I(U_i; \mathcal{Y}_1^N U_1^{i-1}). \end{aligned}$$

Classically, a message is encoded and each bit is sent across W.

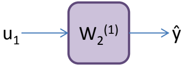


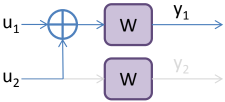
In polar coding, sums of bits are sent across W.



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

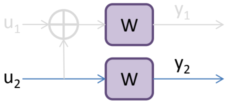
This results in upgraded and degraded channels.

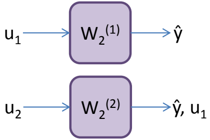




$$G_2 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

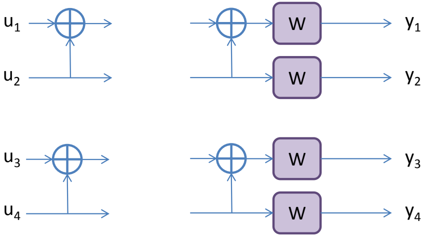
This results in upgraded and degraded channels.





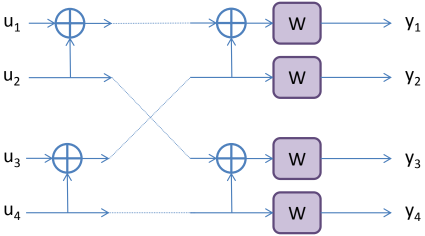
$$G_2 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

The 4-bit diagram has 4 embedded copies of the 2-bit diagram represented by a permutation of $G_2^{\otimes 2}$.



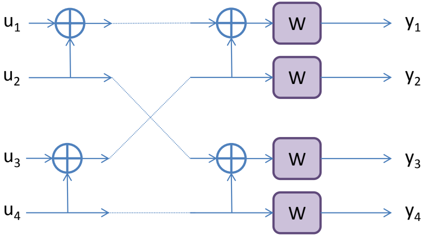
$$G_2^{\otimes 2} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

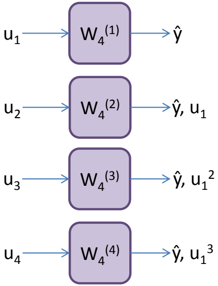
Bit-reversals are represented by switching columns.



$$G_2^{\otimes 2} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

This results in upgraded and degraded channels.





$$B_4 G_2^{\otimes 2} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

The channel W_N is defined recursively.

We define $W_i : \mathcal{X}^i \to \mathcal{Y}^i$, $1 \leq i \leq N = 2^n$, as

$$W_1 = W,$$
$$W_2(y_1^2|u_1^2) = W(y_1|u_1 \oplus u_2)W(y_2|u_2),$$

and

$$W_N(y_1^N|u_1^N) = W^N(y|u(B_N G_2^{\otimes n})),$$

where $u \in \mathcal{X}^N$ and $y \in \mathcal{Y}^N$ and W^N denotes N independent uses of W.

The channels $W_N^{(i)}$ are defined based on the chain rule for mutual information.

For $1 \leq i \leq N$, the binary channels

$$W_N^{(i)} : \mathcal{X} \to \mathcal{Y}^N \times \mathcal{X}^{i-1}$$

are defined by the transition probabilities

$$W_N^{(i)}(y_1^N, u_1^{i-1} | u_i) = \sum_{u_{i+1}^N \in \mathcal{X}^{N-i}} \frac{1}{2^{N-1}} W_N(y_1^N | u_1^N).$$

The fraction of better channels goes to $I(W)$.

Theorem (Arikan, 2009)

For any binary DMC W, the channels $W_N^{(i)}$ polarize in the sense that, for any fixed $\delta \in (0,1)$, as N goes to infinity, the fraction of indices $i \in 1, \ldots, N$ for which

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

goes to

$$I(W).$$

Polar codes for q-ary DMC were first studied by Mori and Tanaka.

Let $W : \mathcal{X} \to \mathcal{Y}$ be a q-ary DMC.

- Rate: The symmetric capacity is

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

- Reliability: The Bhattacharyya parameter is

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

where

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

for $x, x' \in \mathcal{X}$.

Polarization is not restricted to G_2.

Theorem (Korada,Şaşoğlu, and Urbanke, 2009)

For any binary channel W, G polarizes if and only if G is not upper triangular.

Polarization is not restricted to G_2.

Theorem (Korada,Şaşoğlu, and Urbanke, 2009)

For any binary channel W, G polarizes if and only if G is not upper triangular.

Theorem (Mori and Tanaka, 2010)

For any q-ary channel W, suppose G is linear kernel which is not diagonal. Let k be the index of the row with the largest number of non-zero elements. If there exists $j \in \{0, \dots, k-1\}$ such that G_{kj} is a primitive element, then G polarizes.

The rate of polarization of a kernel depends on partial distances, which are governed by nested vector spaces.

Each kernel matrix has a rate of polarization, $E(G)$, called the exponent of G. Let

$$G = \begin{bmatrix} \text{---} & g_1 & \text{---} \\ \text{---} & g_2 & \text{---} \\ & \vdots & \\ \text{---} & g_{\ell-1} & \text{---} \\ \text{---} & g_\ell & \text{---} \end{bmatrix} \in \mathbb{F}_q^{\ell \times \ell}.$$

The i^{th} partial distance of G is

$$D_i = d(g_i, \langle g_{i+1}, \ldots, g_\ell \rangle).$$

The rate of polarization of a kernel depend on partial distances, which are governed by nested vector spaces.

Each kernel matrix has a rate of polarization, $E(G)$, called the exponent of G. Let

$$G = \begin{bmatrix} — & g_1 & — \\ — & g_2 & — \\ & \vdots & \\ — & g_{\ell-1} & — \\ — & g_\ell & — \end{bmatrix} \in \mathbb{F}_q^{\ell \times \ell}.$$

The i^{th} partial distance of G is

$$D_i = d(g_i, \langle g_{i+1}, \ldots, g_\ell \rangle).$$

Note that

$$\langle g_\ell \rangle \subseteq \langle g_{\ell-1}, g_\ell \rangle \subseteq \ldots \subseteq \langle g_2, \ldots, g_\ell \rangle .$$

The rate of polarization of a kernel depend on partial distances, which are governed by nested vector spaces.

Each kernel matrix has a rate of polarization, $E(G)$, called the exponent of G. Let

$$G = \begin{bmatrix} \text{---} & g_1 & \text{---} \\ \text{---} & g_2 & \text{---} \\ & \vdots & \\ \text{---} & g_{\ell-1} & \text{---} \\ \text{---} & g_\ell & \text{---} \end{bmatrix} \in \mathbb{F}_q^{\ell \times \ell}.$$

The i^{th} partial distance of G is

$$D_i = d(g_i, \langle g_{i+1}, \ldots, g_\ell \rangle).$$

Definition

For any channel W and any $\ell \times \ell$ kernel matrix G with partial distances $\{D_i\}_{i=1}^{\ell}$,

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

The exponent provides a bound on the block error probability.

Theorem (Korada, Şaşoğlu, and Urbanke, 2009)

For any W with $0 < I(W) < 1$, an $\ell \times \ell$ kernel G has a rate of polarization $E(G)$ if and only if

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

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

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

$$\liminf_{n \to \infty} Pr[Z_n \leq 2^{-\ell^{n\beta}}] = 0.$$

Here, $Z_n = Z(W_n)$, and the W_i are defined recursively as

$$W_0 = W, \quad \text{and} \quad W_{n+1} = (W_n)_N^{(B_n+1)},$$

where $\{B_n \mid n \geq 1\}$ is a sequence of i.i.d random variables uniformly distributed over the set $\{1, \ldots, \ell\}$.

The exponent provides a bound on the block error probability.

Theorem

Consider polar coding over a q-ary DMC using kernel G at a fixed rate $0 < R < I(W)$ with block length $N = \ell^n$. Then

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

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

The nested structure of AG codes provides a systematic construction of nice kernels.

Let F be a function field over $\mathbb{F}_q$ of genus g. Consider divisors A and

$$D = P_1 + \ldots + P_n$$

with disjoint support, where P_i are places of F of degree 1. The Riemann-Roch space of A is

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

An algebraic geometry (AG) code, $C(D, A)$, is

$$C(D, A) = \{(f(P_1), \ldots, f(P_n)) \; : \; f \in \mathcal{L}(A)\}.$$

The nested structure of AG codes provides a systematic construction of nice kernels.

Let F be a function field over $\mathbb{F}_q$ of genus g. Consider divisors A and

$$D = P_1 + \ldots + P_n$$

with disjoint support, where P_i are places of F of degree 1. The Riemann-Roch space of A is

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

An algebraic geometry (AG) code, $C(D, A)$, is

$$C(D, A) = \{(f(P_1), \ldots, f(P_n)) \; : \; f \in \mathcal{L}(A)\}.$$

AG codes have a “nested” structure such that given divisors A and B,

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

The nested structure of AG codes provides a systematic construction of nice kernels.

Construct a sequence of divisors

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

so that the supports of $D := P_1 + \cdots + P_n$ and A_j are disjoint and

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

The nested structure of AG codes provides a systematic construction of nice kernels.

Construct a sequence of divisors

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

so that the supports of $D := P_1 + \cdots + P_n$ and A_j are disjoint and

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

Let $\{f_1, \ldots, f_k\}$ is a basis for $\mathcal{L}(A_n)$, so

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

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

The nested structure of AG codes provides a systematic construction of nice kernels.

Construct a sequence of divisors

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

so that the supports of $D := P_1 + \cdots + P_n$ and A_j are disjoint and

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

Let $\{f_1, \ldots, f_k\}$ is a basis for $\mathcal{L}(A_n)$, so

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

is a generator matrix for $C(D, A_n)$. The matrix with rows $\mathrm{Row}_{k-i} G, \ldots, \mathrm{Row}_k G$ is a generator matrix for

$$C(D, A_i).$$

The partial distances of the kernel are bounded by the minimum distance of the nested codes.

Let $\{f_1, \ldots, f_k\}$ is a basis for $\mathcal{L}(A_n)$, so

$$G = \begin{bmatrix} f_k(P_1) & f_k(P_2) & \cdots & f_k(P_n) \\ f_{k-1}(P_1) & f_{k-1}(P_2) & \cdots & f_{k-1}(P_n) \\ \vdots & \vdots & \cdots & \vdots \\ f_2(P_1) & f_2(P_2) & \cdots & f_2(P_n) \\ f_1(P_1) & f_1(P_2) & \cdots & f_1(P_n) \end{bmatrix}$$

is a generator matrix for $C(D, A_n)$. This G will be the kernel matrix, so

$$D_i \geq d(C(D, A_{n-i})) := d_{n-i}.$$

Bounds on the minimum distance of nested codes give bounds on the exponent.

Theorem

The exponent of the polar code with kernel G constructed using an AG code of length n of a function field of genus g as above satisfies

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

Bounds on the minimum distance of nested codes give bounds on the exponent.

Theorem

The exponent of the polar code with kernel G constructed using an AG code of length n of a function field of genus g as above satisfies

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

Corollary (Mori and Tanaka, 2010)

If G_{RS} is a Reed-Solomon kernel over $\mathbb{F}_q$, then the exponent of G_{RS} is

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

Maximal function fields give kernels with exponents very close to 1.

Theorem

Let $F/\mathbb{F}_q$ be a maximal function field of genus g. Also, let G be a generator matrix of an AG code on F of length n constructed as before where $n = q + 2gq^{1/2}$. Then

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

The Hermitian function field is an example of a maximal function field.

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. A 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 $y - \beta$.

Bounds on the minimum distances can be used to bound the exponent of Hermitian kernels.

Corollary

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

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

where $a^{\underline{i}} := a(a-1)\dots(a-i+1)$.

Bounds on the minimum distances can be used to bound the exponent of Hermitian kernels.

Corollary

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

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

where $a^{\underline{i}} := a(a-1)\ldots(a-i+1)$.

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

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

Hermitian kernels usually produce larger exponents than Reed-Solomon kernels.

Proposition

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 for $q \geq 3$

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

Hermitian kernels usually produce larger exponents than Reed-Solomon kernels.

Proposition

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 for $q \geq 3$

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

Corollary

If G_H is a Hermitian kernel over $\mathbb{F}_{q^2}$, then

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

References

(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) 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.

(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** (2010), no. 12, 6253–6264.

(5) 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.

(6) 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.

(7) I. Tal and A. Vardy, How to construct polar codes, submitted to IEEE Trans. Inform. Theory, available online as arXiv:1105.6164v2, 2011.

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

(9) K. Yang and P. Kumar, On the true minimum distance of Hermitian codes, coding theory and algebraic geometry, Lecture Notes in Mathematics. Springer: Berlin, 1992, vol. 1518, 99–107.