Strong Polarization for Shortened and Punctured Polar Codes

Boaz Shuval, Ido Tal

Technion

- Question: do shortened/punctured polar codes have the same error exponent as seminal polar codes?
- Answers:
  - ChatGPT 3: No (and we did not understand the explanation)
  - ChatGPT 4: Yes (and we did not understand the explanation)
  - Co-pilot: Yes, see Shuval & Tal's recent paper

Big picture first

- Seminal polar codes have probability of error $\approx 2^{-\sqrt{N}}$, where $N = 2^n$
- Polar codes can be either shortened or punctured to lengths M that are not powers of 2
- We analyze:
  - the shortening method of Wang and Liu, and
  - the puncturing method of Niu, Chen, and Lin
- In both cases, the probability of error is $\approx 2^{-\sqrt{M}}$
- No restriction on M
- We are *not* assuming a symmetric channel *nor* a symmetric input

Theorem

Let $\mathbf{X}$ *be a random vector of length* M *with i.i.d. entries, each sampled from an input distribution* $p(x)$*. Let* $\mathbf{Y}$ *be the result of passing* $\mathbf{X}$ *through a BM channel* $W(y|x)$*. Let* $\mathbf{U}$ *of length* M *be the result of transforming* $\mathbf{X}$ *via either the shortening transform or the puncturing transform. Fix* $0 < \beta < 1/2$*. Then,*

$$\lim_{M\to\infty} \frac{1}{M} \left|\left\{ i : Z(U_i|U^{i-1}, \mathbf{Y}) < 2^{-M^\beta} \right\}\right| = 1 - H(X|Y),$$

$$\lim_{M\to\infty} \frac{1}{M} \left|\left\{ i : K(U_i|U^{i-1}) < 2^{-M^\beta} \right\}\right| = H(X).$$

Reminder: Bhattacharyya parameter and total variation

$$Z(X|Y) = \sum_y p(Y=y) \cdot \sqrt{P(X=0|Y=y)P(X=1|Y=y)}$$

$$K(X|Y) = \sum_y p(Y=y) \cdot \left|P(X=0|Y=y) - P(X=1|Y=y)\right|$$

Shortening and puncturing

Shortening a general code $\mathcal{C}$:

- Pick an index set $\mathcal{S}$
- Subcode: $\mathbf{c} \in \mathcal{C}$ such that

$$i \in \mathcal{S} \Longrightarrow c_i = 0$$

- Do not transmit indices in $\mathcal{S}$

Puncturing a general code $\mathcal{C}$:

- Pick an index set $\mathcal{P}$
- Use all $\mathbf{c} \in \mathcal{C} \ldots$
- Do not transmit indices in $\mathcal{P}$

For polar codes (Wang and Liu)

$$\mathcal{S} = \{\overleftarrow{N-1}, \overleftarrow{N-2}, \ldots, \overleftarrow{N-(N-M)}\}$$

For polar codes (Niu, Chen, and Lin):

$$\mathcal{P} = \{\overleftarrow{0}, \overleftarrow{1}, \ldots, \overleftarrow{N-M-1}\}$$

Notation for the polar transform

For a binary vector $\mathbf{x} = \begin{bmatrix} x_0 & x_1 & \cdots & x_{N-1} \end{bmatrix}$ of length $N = 2^n$

$$\begin{bmatrix} x_0 & x_1 & \cdots & x_{N-1} \end{bmatrix}^{[0]} = \begin{bmatrix} x_0 \oplus x_1 & x_2 \oplus x_3 & \cdots & x_{N-2} \oplus x_{N-1} \end{bmatrix}$$

and

$$\begin{bmatrix} x_0 & x_1 & \cdots & x_{N-1} \end{bmatrix}^{[1]} = \begin{bmatrix} x_1 & x_3 & \cdots & x_{N-1} \end{bmatrix},$$

Notation for the polar transform

For a binary vector $\mathbf{x} = \begin{bmatrix} x_0 & x_1 & \cdots & x_{N-1} \end{bmatrix}$ of length $N = 2^n$

$$\begin{bmatrix} x_0 & x_1 & \cdots & x_{N-1} \end{bmatrix}^{[0]}$$
$$= \begin{bmatrix} x_0 \oplus x_1 & x_2 \oplus x_3 & \cdots & x_{N-2} \oplus x_{N-1} \end{bmatrix}$$

and

$$\begin{bmatrix} x_0 & x_1 & \cdots & x_{N-1} \end{bmatrix}^{[1]}$$
$$= \begin{bmatrix} x_0 \triangleright x_1 & x_2 \triangleright x_3 & \cdots & x_{N-2} \triangleright x_{N-1} \end{bmatrix},$$

where

$$\alpha \triangleright \beta \triangleq \beta$$

Notation for the polar transform

- Let $N = 2^n$
- Polar transform:

$$\mathbf{x} = \begin{bmatrix} x_0 & x_1 & \cdots & x_{N-1} \end{bmatrix} \Longrightarrow \mathbf{u} = \begin{bmatrix} u_0 & u_1 & \cdots & u_{N-1} \end{bmatrix}$$

- Definition: for an index

$$i = (b_{n-1}, b_{n-2}, \ldots, b_0)_2 = \sum_{j=0}^{n-1} b_j 2^j$$

we have

$$u_i = \mathbf{x}^{[\overleftarrow{\mathbf{b}}]} = \left(\cdots \left(\left(\mathbf{x}^{[b_{n-1}]} \right)^{[b_{n-2}]} \right) \cdots \right)^{[b_0]}$$

Notation for shortening and puncturing

Recall that

$$\alpha \triangleright \beta \triangleq \beta$$

We now generalize the $\alpha \oplus \beta$ and $\alpha \triangleright \beta$ operations to

$$\alpha, \beta \in \{0, 1, \mathtt{s}, \mathtt{p}\}$$

$\oplus$	0	1	s	p
0	0	1	0	$\emptyset$
1	1	0	1	$\emptyset$
s	$\emptyset$	$\emptyset$	s	$\emptyset$
p	p	p	p	p

$\triangleright$	0	1	s	p
0	0	1	s	$\emptyset$
1	0	1	s	$\emptyset$
s	$\emptyset$	$\emptyset$	s	$\emptyset$
p	0	1	s	p

.

Intuition:

- $\mathtt{s}$ is another name for 0
- $\mathtt{p}$ signifies a bit with arbitrary value

Two definitions of the polar shortening transform

$\oplus$	0	1	s	p
0	0	1	0	$\emptyset$
1	1	0	1	$\emptyset$
s	$\emptyset$	$\emptyset$	s	$\emptyset$
p	p	p	p	p

$\triangleright$	0	1	s	p
0	0	1	s	$\emptyset$
1	0	1	s	$\emptyset$
s	$\emptyset$	$\emptyset$	s	$\emptyset$
p	0	1	s	p

.

Suppose $M = 5$, and so $N = 2^{\lceil \log_2 M \rceil} = 8$

$$\mathcal{S} = \{\overleftarrow{N-1}, \overleftarrow{N-2}, \dots, \overleftarrow{N-(N-M)}\} = \{\overleftarrow{7}, \overleftarrow{6}, \overleftarrow{5}\} = \{7, 3, 5\}$$

First definition:

$$\mathbf{x} = \begin{bmatrix} 0 & 1 & 1 & & 0 & & 1 & \end{bmatrix}$$

$$\bar{\mathbf{x}} = \begin{bmatrix} 0 & 1 & 1 & \mathtt{s} & 0 & \mathtt{s} & 1 & \mathtt{s} \end{bmatrix}$$

$$\bar{\mathbf{x}}^{[0]} = \begin{bmatrix} 1 & 1 & 0 & 1 \end{bmatrix} \quad \bar{\mathbf{x}}^{[1]} = \begin{bmatrix} 1 & \mathtt{s} & \mathtt{s} & \mathtt{s} \end{bmatrix}$$

$$\bar{\mathbf{x}}^{[00]} = \begin{bmatrix} 0 & 1 \end{bmatrix} \quad \bar{\mathbf{x}}^{[01]} = \begin{bmatrix} 1 & 1 \end{bmatrix} \quad \bar{\mathbf{x}}^{[10]} = \begin{bmatrix} 1 & \mathtt{s} \end{bmatrix} \quad \bar{\mathbf{x}}^{[11]} = \begin{bmatrix} \mathtt{s} & \mathtt{s} \end{bmatrix}$$

$$\bar{\mathbf{u}} = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & \mathtt{s} & \mathtt{s} & \mathtt{s} \end{bmatrix}$$

$$\mathbf{u} = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & & & \end{bmatrix}$$

Two definitions of the polar shortening transform

$\oplus$	0	1	s	p
0	0	1	0	$\emptyset$
1	1	0	1	$\emptyset$
s	$\emptyset$	$\emptyset$	s	$\emptyset$
p	p	p	p	p

$\triangleright$	0	1	s	p
0	0	1	s	$\emptyset$
1	0	1	s	$\emptyset$
s	$\emptyset$	$\emptyset$	s	$\emptyset$
p	0	1	s	p

.

Suppose $M = 5$, and so $N = 2^{\lceil \log_2 M \rceil} = 8$

$$\mathcal{S} = \{\overleftarrow{N-1}, \overleftarrow{N-2}, \ldots, \overleftarrow{N-(N-M)}\} = \{\overleftarrow{7}, \overleftarrow{6}, \overleftarrow{5}\} = \{7, 3, 5\}$$

Second definition:

$$\mathbf{x} = \begin{bmatrix} 0 & 1 & 1 & & 0 & & 1 & \end{bmatrix}$$

$$\bar{\mathbf{x}} = \begin{bmatrix} 0 & 1 & 1 & {\color{red}0} & 0 & {\color{red}0} & 1 & {\color{red}0} \end{bmatrix}$$

$$\bar{\mathbf{x}}^{[0]} = \begin{bmatrix} 1 & 1 & 0 & 1 \end{bmatrix} \quad \bar{\mathbf{x}}^{[1]} = \begin{bmatrix} 1 & {\color{red}0} & {\color{red}0} & {\color{red}0} \end{bmatrix}$$

$$\bar{\mathbf{x}}^{[00]} = \begin{bmatrix} 0 & 1 \end{bmatrix} \quad \bar{\mathbf{x}}^{[01]} = \begin{bmatrix} 1 & 1 \end{bmatrix} \quad \bar{\mathbf{x}}^{[10]} = \begin{bmatrix} 1 & {\color{red}0} \end{bmatrix} \quad \bar{\mathbf{x}}^{[11]} = \begin{bmatrix} {\color{red}0} & {\color{red}0} \end{bmatrix}$$

$$\bar{\mathbf{u}} = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & {\color{red}0} & {\color{red}0} & {\color{red}0} \end{bmatrix}$$

$$\mathbf{u} = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & & & \end{bmatrix}$$

Two definitions of the polar puncturing transform

$\oplus$	0	1	s	p
0	0	1	0	$\emptyset$
1	1	0	1	$\emptyset$
s	$\emptyset$	$\emptyset$	s	$\emptyset$
p	p	p	p	p

$\triangleright$	0	1	s	p
0	0	1	s	$\emptyset$
1	0	1	s	$\emptyset$
s	$\emptyset$	$\emptyset$	s	$\emptyset$
p	0	1	s	p

.

Suppose $M = 5$, and so $N = 2^{\lceil \log_2 M \rceil} = 8$

$$\mathcal{P} = \{\overleftarrow{0}, \overleftarrow{1}, \ldots, \overleftarrow{N-M-1}\} = \{\overleftarrow{0}, \overleftarrow{1}, \overleftarrow{2}\} = \{0, 4, 2\}$$

First definition:

$$\mathbf{x} = \begin{bmatrix} & 0 & & 1 & & 1 & 0 & 1 \end{bmatrix}$$

$$\tilde{\mathbf{x}} = \begin{bmatrix} \mathtt{p} & 0 & \mathtt{p} & 1 & \mathtt{p} & 1 & 0 & 1 \end{bmatrix}$$

$$\tilde{\mathbf{x}}^{[0]} = \begin{bmatrix} \mathtt{p} & \mathtt{p} & \mathtt{p} & 1 \end{bmatrix} \quad \tilde{\mathbf{x}}^{[1]} = \begin{bmatrix} 0 & 1 & 1 & 1 \end{bmatrix}$$

$$\tilde{\mathbf{x}}^{[00]} = \begin{bmatrix} \mathtt{p} & \mathtt{p} \end{bmatrix} \quad \tilde{\mathbf{x}}^{[01]} = \begin{bmatrix} \mathtt{p} & 1 \end{bmatrix} \quad \bar{\mathbf{x}}^{[10]} = \begin{bmatrix} 1 & 0 \end{bmatrix} \quad \tilde{\mathbf{x}}^{[11]} = \begin{bmatrix} 1 & 1 \end{bmatrix}$$

$$\tilde{\mathbf{u}} = \begin{bmatrix} \mathtt{p} & \mathtt{p} & \mathtt{p} & 1 & 1 & 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{u} = \begin{bmatrix} & & & 1 & 1 & 0 & 0 & 1 \end{bmatrix}$$

Two definitions of the polar puncturing transform

$\oplus$	0	1	s	p
0	0	1	0	$\emptyset$
1	1	0	1	$\emptyset$
s	$\emptyset$	$\emptyset$	s	$\emptyset$
p	p	p	p	p

$\triangleright$	0	1	s	p
0	0	1	s	$\emptyset$
1	0	1	s	$\emptyset$
s	$\emptyset$	$\emptyset$	s	$\emptyset$
p	0	1	s	p

.

Suppose $M = 5$, and so $N = 2^{\lceil \log_2 M \rceil} = 8$

$$\mathcal{P} = \{\overleftarrow{0}, \overleftarrow{1}, \ldots, \overleftarrow{N-M-1}\} = \{\overleftarrow{0}, \overleftarrow{1}, \overleftarrow{2}\} = \{0, 4, 2\}$$

Second definition:

$$\mathbf{x} = \begin{bmatrix} & 0 & & 1 & & 1 & 0 & 1 \end{bmatrix}$$

$$\tilde{\mathbf{x}} = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 1 & 0 & 1 \end{bmatrix}$$

$$\tilde{\mathbf{x}}^{[0]} = \begin{bmatrix} 1 & 1 & 0 & 1 \end{bmatrix} \quad \tilde{\mathbf{x}}^{[1]} = \begin{bmatrix} 0 & 1 & 1 & 1 \end{bmatrix}$$

$$\tilde{\mathbf{x}}^{[00]} = \begin{bmatrix} 0 & 1 \end{bmatrix} \quad \tilde{\mathbf{x}}^{[01]} = \begin{bmatrix} 1 & 1 \end{bmatrix} \quad \bar{\mathbf{x}}^{[10]} = \begin{bmatrix} 1 & 0 \end{bmatrix} \quad \tilde{\mathbf{x}}^{[11]} = \begin{bmatrix} 1 & 1 \end{bmatrix}$$

$$\tilde{\mathbf{u}} = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{u} = \begin{bmatrix} & & & 1 & 1 & 0 & 0 & 1 \end{bmatrix}$$

Second definition, for now

- We now think of shortening/puncturing using the second definition

$$\bar{\mathbf{x}} = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$
$$\tilde{\mathbf{x}} = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 1 & 0 & 1 \end{bmatrix}$$

- The first definition will come into play later...

Distributions

- Denote the probability distribution of "regular" input-output as

$$W(x;y) = P(X = x, Y = y)$$

- What about shortening/puncturing?
- Shortening:
  - Input is forced to be 0
  - No corresponding output

$$\mathtt{S}(x;y) = \begin{cases} 1, & x = 0, y = ? \\ 0, & \text{otherwise} \end{cases}$$

- Puncturing:
  - Input is arbitrary
  - No corresponding output

$$\mathtt{P}(x;y) = \begin{cases} \frac{1}{2}, & x \in \{0,1\}, y = ? \\ 0, & \text{otherwise} \end{cases}$$

The '$-$' and '$+$' operations on joint distributions

- Denote
$$\mathcal{X} = \{0, 1\}$$
- Let $A(x_0; y_0)$ be a joint distribution on $(x_0, y_0) \in \mathcal{X} \times \mathcal{Y}_0$
- Let $B(x_0; y_1)$ be a joint distribution on $(x_1, y_1) \in \mathcal{X} \times \mathcal{Y}_1$
- The '$-$' operation:
$$(A \boxast B)(u_0; y_0, y_1) = \sum_{x_1 \in \mathcal{X}} A(u_0 \oplus x_1; y_0) B(x_1; y_1)$$
- The '$+$' operation:
$$(A \circledast B)(u_1; u_0, y_0, y_1) = A(u_0 \oplus u_1; y_0) B(u_1; y_1)$$

Degrading and the symmetric setting

- For two joint distributions $A(x_0; y_0)$ and $B(x_0; y_1)$, denote

$$A \overset{\text{d}}{\sqsubseteq} B$$

  if A is (stochastically) degraded with respect to B
- That is, if there exists $Q(y_0|y_1)$ over $\mathcal{Y}_0 \times \mathcal{Y}_1$ such that

$$A(x_0; y_0) = \sum_{y_1} B(x_0; y_1) Q(y_0|y_1)$$

- In the special case where $A(x_0; y_0)$ corresponds to a symmetric input distribution and a symmetric channel,

$$\underbrace{A \boxast A}_{A^-} \overset{\text{d}}{\sqsubseteq} A \overset{\text{d}}{\sqsubseteq} \underbrace{A \circledast A}_{A^+}$$

- Goal: generalize "$\overset{\text{d}}{\sqsubseteq}$" to some "$\sqsubseteq$" so that for general A, B

$$A \boxast B \sqsubseteq A \sqsubseteq A \circledast B \,, \quad A \boxast B \sqsubseteq B \sqsubseteq A \circledast B$$

The 'input permutation' relation

- We say that A has undergone an input permutation, resulting in A' if there exists a function $f : \mathcal{Y}_0 \to \mathcal{X}$ such that
$$A'(x_0; y_0) = A(x_0 \oplus f(y_0); y_0)$$
- We denote this by
$$A' \overset{\mathrm{p}}{\sqsubseteq} A$$
- Note that
$$Z(A') = Z(A),$$
$$K(A') = K(A'),$$
$$H(A') = H(A)$$

Since
$$Z(A) = \sum_y \sqrt{A(0; y) \cdot A(1; y)}$$
$$Z(A') = \sum_y \sqrt{A(0 \oplus f(y); y) \cdot A(1 \oplus f(y); y)}$$

The 'inferior' relation

- We define that $A \sqsubseteq B$ if we can identify a finite sequence of 'degradation' and 'input permutation' relations that will lead to A from B
- In other words, there exists $0 < t < \infty$, a sequence of joint distributions $C_1, C_2, \ldots, C_{t-1}$, and a sequence $\mathrm{r}_1, \mathrm{r}_2, \ldots, \mathrm{r}_t \in \{\mathrm{d}, \mathrm{p}\}$ such that

$$A \overset{\mathrm{r}_1}{\sqsubseteq} C_1 \overset{\mathrm{r}_2}{\sqsubseteq} C_2 \overset{\mathrm{r}_3}{\sqsubseteq} \cdots \overset{\mathrm{r}_{t-1}}{\sqsubseteq} C_{t-1} \overset{\mathrm{r}_t}{\sqsubseteq} B$$

Key properties of the 'inferior' relation

$A \sqsubseteq B$ if there exists $0 < t < \infty$, a sequence of joint distributions $C_1, C_2, \ldots, C_{t-1}$, and a sequence $\mathrm{r}_1, \mathrm{r}_2, \ldots, \mathrm{r}_t \in \{\mathrm{d}, \mathrm{p}\}$ such that

$$A \overset{\mathrm{r}_1}{\sqsubseteq} C_1 \overset{\mathrm{r}_2}{\sqsubseteq} C_2 \overset{\mathrm{r}_3}{\sqsubseteq} \cdots \overset{\mathrm{r}_{t-1}}{\sqsubseteq} C_{t-1} \overset{\mathrm{r}_t}{\sqsubseteq} B$$

Key properties:

- Transitivity:

$$A \sqsubseteq B \quad \text{and} \quad B \sqsubseteq C \Longrightarrow A \sqsubseteq C$$

- Z, K, and H monotonicity:

$$A \sqsubseteq B \Longrightarrow Z(A) \geq Z(B),\ K(A) \leq K(B),\ H(A) \geq H(B)$$

- Preservation by polar operations:

$$A' \sqsubseteq A \quad \text{and} \quad B' \sqsubseteq B \Longrightarrow$$
$$A' \boxast B' \sqsubseteq A \boxast B \quad \text{and} \quad A' \circledast B' \sqsubseteq A \circledast B.$$

- The two extremes: For any A,

$$\mathtt{P} \sqsubseteq A \sqsubseteq \mathtt{S}$$

Look familiar?

- If $A \sqsubseteq B$ and $B \sqsubseteq A$ then we will treat A and B as equivalent
- The following holds, up to equivalence:

$\boxast$	B	$\mathtt{S}$	$\mathtt{P}$
A	$A \boxast B$	A	$\mathtt{P}$
$\mathtt{S}$	B	$\mathtt{S}$	$\mathtt{P}$
$\mathtt{P}$	$\mathtt{P}$	$\mathtt{P}$	$\mathtt{P}$

$\circledast$	B	$\mathtt{S}$	$\mathtt{P}$
A	$A \circledast B$	$\mathtt{S}$	A
$\mathtt{S}$	$\mathtt{S}$	$\mathtt{S}$	$\mathtt{S}$
$\mathtt{P}$	B	$\mathtt{S}$	$\mathtt{P}$

- Look familiar?

Look familiar?

- If $A \sqsubseteq B$ and $B \sqsubseteq A$ then we will treat A and B as equivalent
- The following holds, up to equivalence:

$\boxast$	B	$\mathtt{S}$	$\mathtt{P}$
A	$A \boxast B$	A	$\mathtt{P}$
$\mathtt{S}$	B	$\mathtt{S}$	$\mathtt{P}$
$\mathtt{P}$	$\mathtt{P}$	$\mathtt{P}$	$\mathtt{P}$

$\circledast$	B	$\mathtt{S}$	$\mathtt{P}$
A	$A \circledast B$	$\mathtt{S}$	A
$\mathtt{S}$	$\mathtt{S}$	$\mathtt{S}$	$\mathtt{S}$
$\mathtt{P}$	B	$\mathtt{S}$	$\mathtt{P}$

- Look familiar?
- Yes! For $a, b \in \{0, 1\}$,

$\oplus$	b	$\mathtt{s}$	$\mathtt{p}$
a	$a \oplus b$	a	$\mathtt{p}$
$\mathtt{s}$	b	$\mathtt{s}$	$\mathtt{p}$
$\mathtt{p}$	$\mathtt{p}$	$\mathtt{p}$	$\mathtt{p}$

$\triangleright$	b	$\mathtt{s}$	$\mathtt{p}$
a	$a \triangleright b$	$\mathtt{s}$	a
$\mathtt{s}$	$\mathtt{s}$	$\mathtt{s}$	$\mathtt{s}$
$\mathtt{p}$	b	$\mathtt{s}$	$\mathtt{p}$

The advantages of good bookkeeping

$$\mathbf{x} = \begin{bmatrix} 0 & 1 & 1 & & 0 & & 1 & \end{bmatrix}$$
$$\bar{\mathbf{x}} = \begin{bmatrix} 0 & 1 & 1 & \mathtt{s} & 0 & \mathtt{s} & 1 & \mathtt{s} \end{bmatrix}$$
$$\bar{\mathbf{u}} = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & \mathtt{s} & \mathtt{s} & \mathtt{s} \end{bmatrix}$$
$$\mathbf{u} = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & & & \end{bmatrix}$$

For $0 \le i \le M$,

$$Z(U_i|U^{i-1}, \mathbf{Y}) = Z(\bar{U}_i|\bar{U}^{i-1}, \bar{\mathbf{Y}})$$

$$K(U_i|U^{i-1}, \mathbf{Y}) = K(\bar{U}_i|\bar{U}^{i-1}, \bar{\mathbf{Y}})$$

$$\mathbf{x} = \begin{bmatrix} & 0 & & 1 & & 1 & 0 & 1 \end{bmatrix}$$
$$\tilde{\mathbf{x}} = \begin{bmatrix} \mathtt{p} & 0 & \mathtt{p} & 1 & \mathtt{p} & 1 & 0 & 1 \end{bmatrix}$$
$$\tilde{\mathbf{u}} = \begin{bmatrix} \mathtt{p} & \mathtt{p} & \mathtt{p} & 1 & 1 & 0 & 0 & 1 \end{bmatrix}$$
$$\mathbf{u} = \begin{bmatrix} & & & 1 & 1 & 0 & 0 & 1 \end{bmatrix}$$

For $0 \le i \le M$,

$$Z(U_i|U^{i-1}, \mathbf{Y}) = Z(\tilde{U}_{i+N-M}|\tilde{U}^{i+N-M-1}, \tilde{\mathbf{Y}})$$

$$K(U_i|U^{i-1}, \mathbf{Y}) = K(\tilde{U}_{i+N-M}|\tilde{U}^{i+N-M-1}, \tilde{\mathbf{Y}})$$

The advantages of good bookkeeping

$$\mathbf{x} = \begin{bmatrix} 0 & 1 & 1 & & 0 & & 1 & \end{bmatrix} \qquad \mathbf{x} = \begin{bmatrix} & 0 & & 1 & & 1 & 0 & 1 \end{bmatrix}$$

$$\bar{\mathbf{x}} = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \qquad \tilde{\mathbf{x}} = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 1 & 0 & 1 \end{bmatrix}$$

$$\bar{\mathbf{u}} = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \end{bmatrix} \qquad \tilde{\mathbf{u}} = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{u} = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & & & \end{bmatrix} \qquad \mathbf{u} = \begin{bmatrix} & & & 1 & 1 & 0 & 0 & 1 \end{bmatrix}$$

For $0 \leq i \leq M$,

$$Z(U_i | U^{i-1}, \mathbf{Y}) = Z(\bar{U}_i | \bar{U}^{i-1}, \bar{\mathbf{Y}})$$

$$K(U_i | U^{i-1}, \mathbf{Y}) = K(\bar{U}_i | \bar{U}^{i-1}, \bar{\mathbf{Y}})$$

For $0 \leq i \leq M$,

$$Z(U_i | U^{i-1}, \mathbf{Y}) = Z(\tilde{U}_{i+N-M} | \tilde{U}^{i+N-M-1}, \tilde{\mathbf{Y}})$$

$$K(U_i | U^{i-1}, \mathbf{Y}) = K(\tilde{U}_{i+N-M} | \tilde{U}^{i+N-M-1}, \tilde{\mathbf{Y}})$$

Main Theorem, reworded

Theorem

Let $W(x; y)$ be a joint distribution over $\mathcal{X} \times \mathcal{Y}$. Let $\mathbf{X}, \mathbf{Y}$ be a pair of random vectors of length M, with each (X_i, Y_i) sampled independently from W. Let $\mathbf{U}$ of length M be the result of transforming $\mathbf{X}$ via either the shortening transform or the puncturing transform. Fix $0 < \beta < 1/2$ and $\epsilon > 0$. Then, there exists M_0 such that for all $M \geq M_0$,

$$\frac{1}{M}\left|\left\{i : Z(U_i|U^{i-1}, \mathbf{Y}) < 2^{-M^\beta}\right\}\right| > 1 - H(X|Y) - \epsilon,$$

$$\frac{1}{M}\left|\left\{i : K(U_i|U^{i-1}, \mathbf{Y}) < 2^{-M^\beta}\right\}\right| > H(X|Y) - \epsilon.$$

A halfway lemma

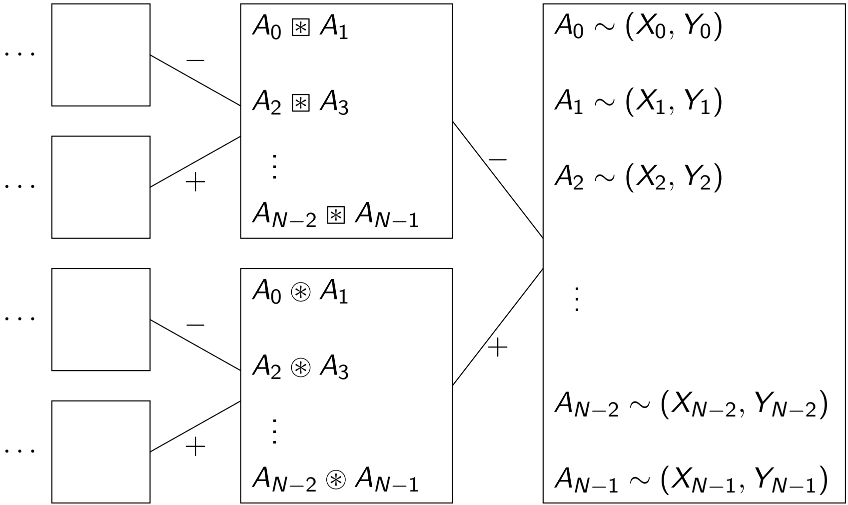
Lemma

Let $W(x;y)$, $\mathbf{X}$, $\mathbf{Y}$, and $\mathbf{U}$ be as in the main theorem. Fix $0 < \beta' < 1/2$ and $\epsilon' > 0$. Fix integers $t > 0$ and $a \in \{2^{t-1} + 1, 2^{t-1} + 2, \ldots, 2^t\}$. There exists n_0 such that for all $n \geq n_0$, if $M = a \cdot 2^{n-t}$, then for $N = 2^n$,

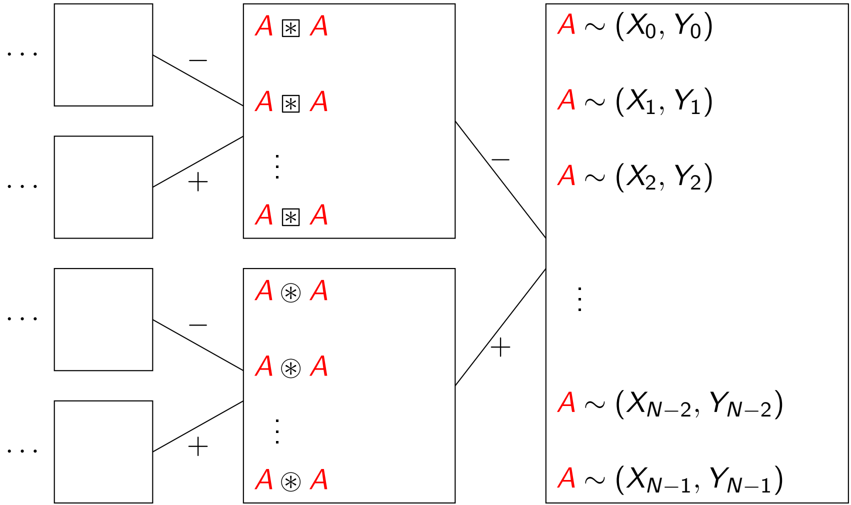
$$\frac{1}{M}\left|\left\{i : Z(U_i|U^{i-1}, \mathbf{Y}) < 2^{-N^{\beta'}}\right\}\right| > 1 - H(X|Y) - \epsilon',$$

$$\frac{1}{M}\left|\left\{i : K(U_i|U^{i-1}, \mathbf{Y}) < 2^{-N^{\beta'}}\right\}\right| > H(X|Y) - \epsilon'.$$

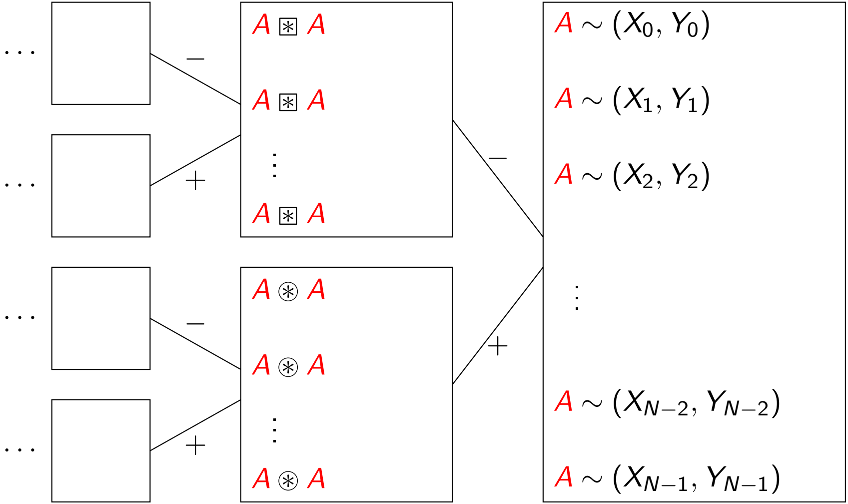
Proof of halfway lemma – part 1



Proof of halfway lemma – part 1

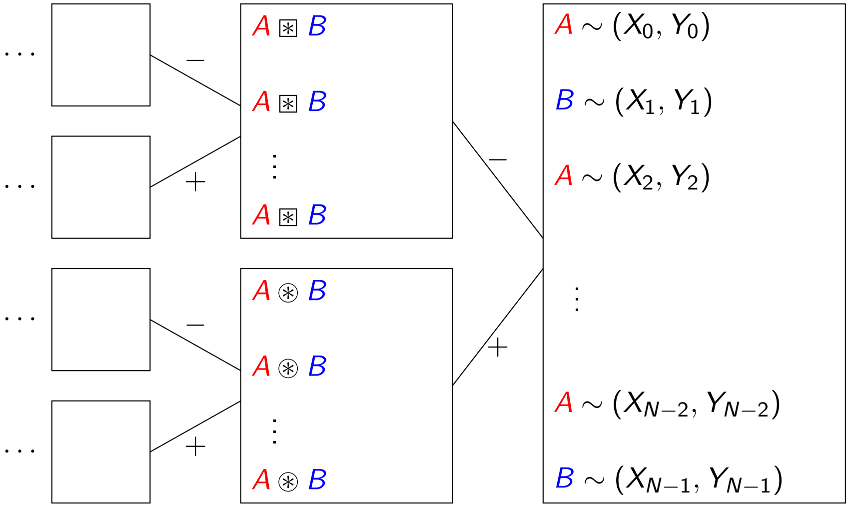


Proof of halfway lemma – part 1

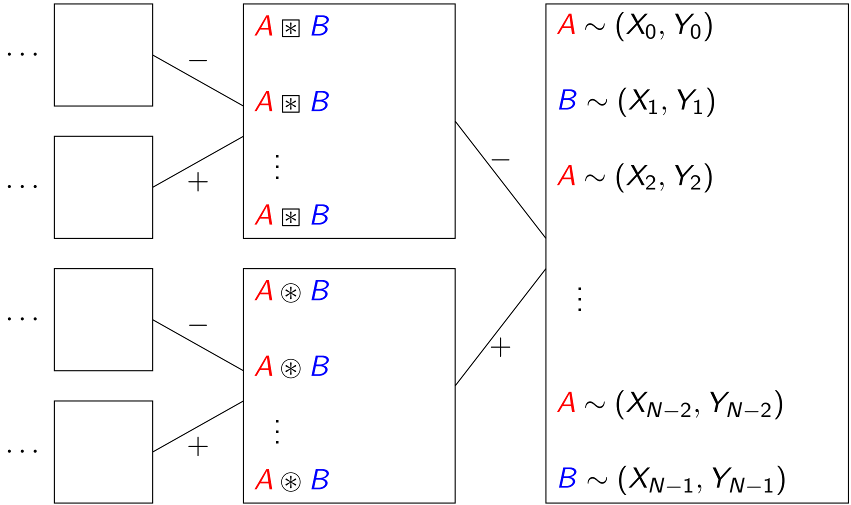


When all A_i are equal: Arıkan & Telatar '09 gives fast polarization

Proof of halfway lemma – part 1

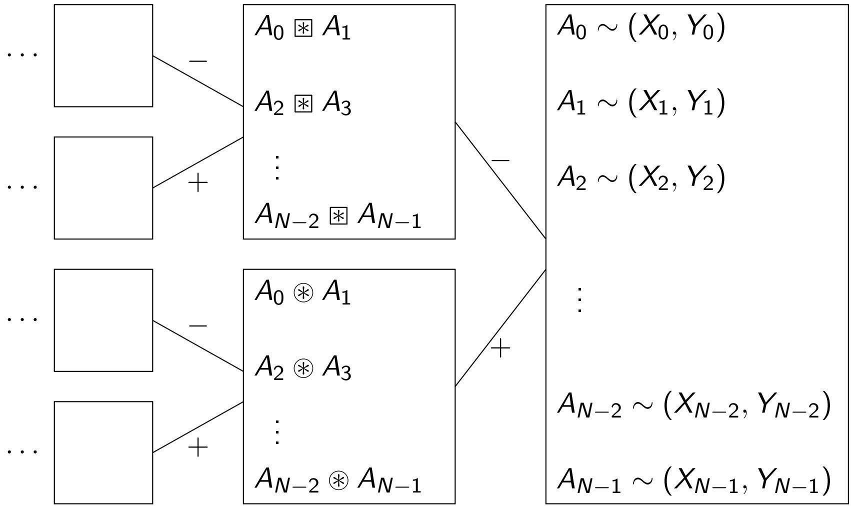


Proof of halfway lemma – part 1



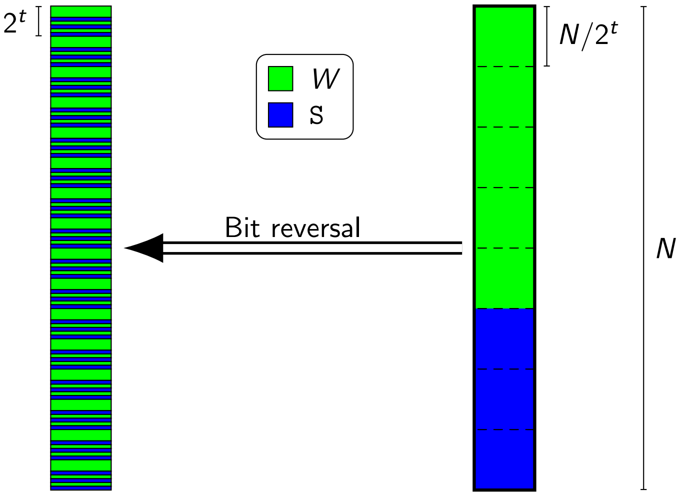
When A_i have period 2: Arıkan & Telatar, '09 applied after first transform gives fast polarization

Proof of halfway lemma – part 1



Generally: if the A_i have period 2^t, then we have fast polarization

Proof of halfway lemma – part 2



Proof of main theorem – key properties of "$\sqsubseteq$"

Recall key properties of "$\sqsubseteq$" relation:

- The two extremes: For any A,

$$\mathtt{P} \sqsubseteq A \sqsubseteq \mathtt{S}$$

- Preservation by polar operations:

$$A' \sqsubseteq A \quad \text{and} \quad B' \sqsubseteq B \Longrightarrow$$
$$A' \boxast B' \sqsubseteq A \boxast B \quad \text{and} \quad A' \circledast B' \sqsubseteq A \circledast B.$$

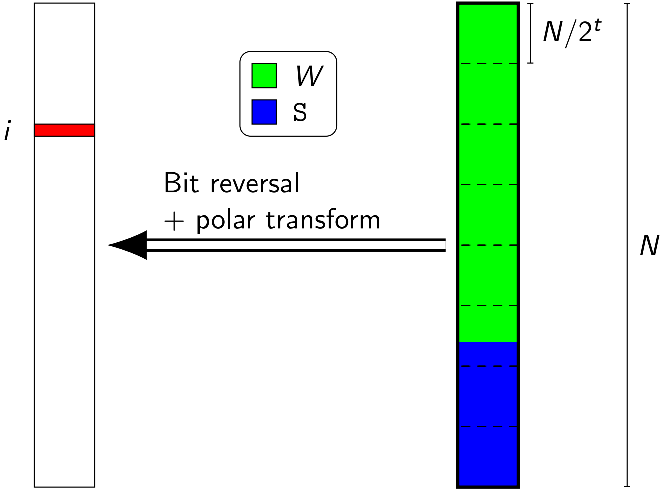
- Transitivity:

$$A \sqsubseteq B \quad \text{and} \quad B \sqsubseteq C \Longrightarrow A \sqsubseteq C$$

- Z, K, and H monotonicity:

$$A \sqsubseteq B \Longrightarrow Z(A) \geq Z(B),\ K(A) \leq K(B),\ H(A) \geq H(B)$$

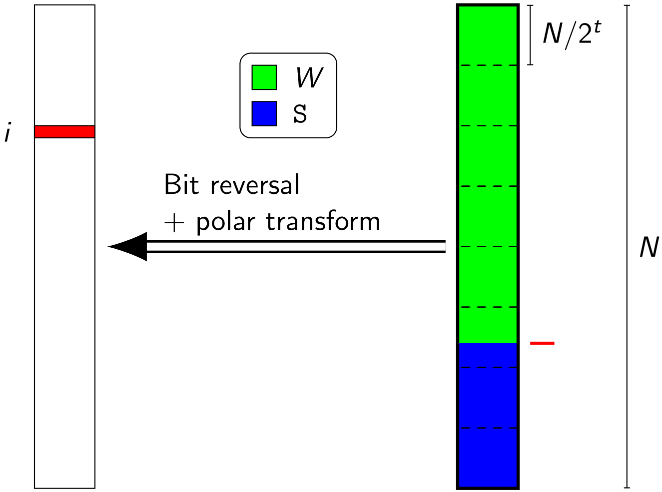
Proof of main theorem



$Z(\bar{U}_i|\bar{U}^{i-1}, \bar{\mathbf{Y}})$

$K(\bar{U}_i|\bar{U}^{i-1}, \bar{\mathbf{Y}})$

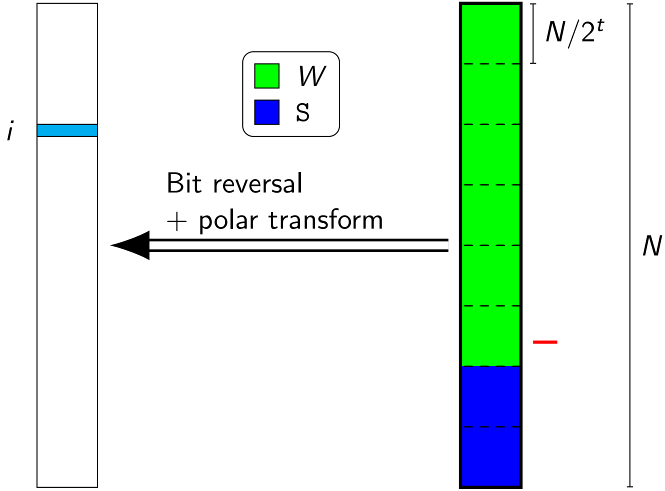
Proof of main theorem



$Z(\bar{U}_i|\bar{U}^{i-1}, \bar{\mathbf{Y}})$

$K(\bar{U}_i|\bar{U}^{i-1}, \bar{\mathbf{Y}})$

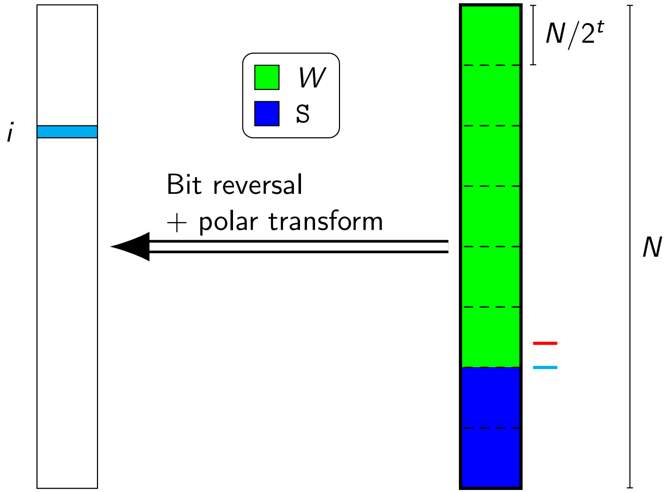
Proof of main theorem



$Z(\bar{U}_i | \bar{U}^{i-1}, \bar{\mathbf{Y}})$

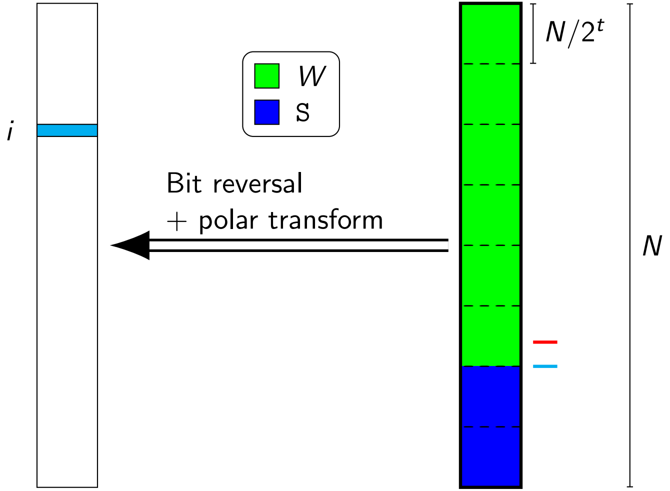
$K(\bar{U}_i | \bar{U}^{i-1}, \bar{\mathbf{Y}})$

Proof of main theorem



$$Z(\bar{U}_i|\bar{U}^{i-1},\bar{\mathbf{Y}})$$
$$K(\bar{U}_i|\bar{U}^{i-1},\bar{\mathbf{Y}})$$

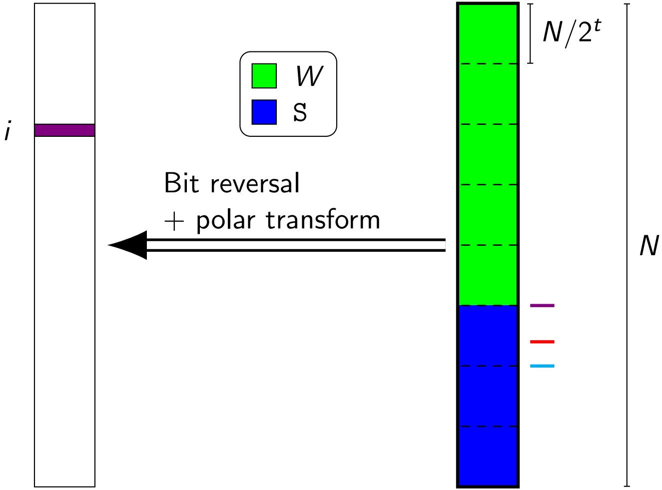
Proof of main theorem



$$Z(\bar{U}_i|\bar{U}^{i-1}, \bar{\mathbf{Y}}) \leq Z(\bar{U}_i|\bar{U}^{i-1}, \bar{\mathbf{Y}})$$

$$K(\bar{U}_i|\bar{U}^{i-1}, \bar{\mathbf{Y}})$$

Proof of main theorem



$$Z(\bar{U}_i|\bar{U}^{i-1},\bar{\mathbf{Y}}) \leq Z(\bar{U}_i|\bar{U}^{i-1},\bar{\mathbf{Y}})$$
$$K(\bar{U}_i|\bar{U}^{i-1},\bar{\mathbf{Y}}) \leq K(\bar{U}_i|\bar{U}^{i-1},\bar{\mathbf{Y}})$$