Lecture 14: Polar codes

Course instructor: Alexey Frolov

al.frolov@skoltech.ru

Teaching Assistant: Stanislav Kruglik

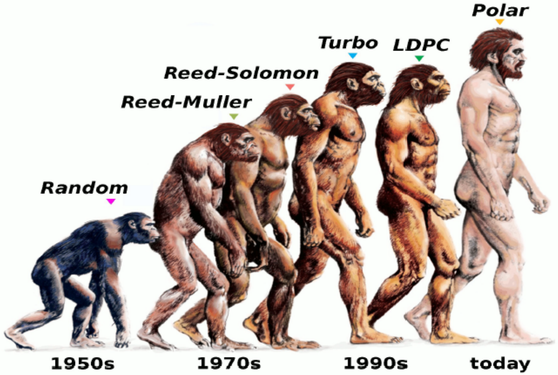
stanislav.kruglik@skolkovotech.ru

March 7, 2017

Outline

1. Polarization
2. Encoding
3. Decoding

Introduction



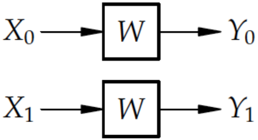
- Polar codes are the ***"next big thing"*** in coding theory since the advent of turbo codes and LDPC codes in the late 1990s.
- In terms of performance/complexity trade-off, the only competition to polar codes today are ***spatially-coupled LDPC codes***.

Outline

1. Polarization
2. Encoding
3. Decoding

Polarization

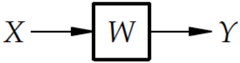
A new class of error-correcting codes, invented in 2009 by Erdal Arıkan, based on the universal phenomenon of ***channel polarization.***



Two independent uses of the channel W

Polarization

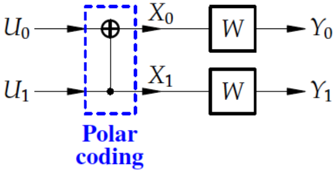
A new class of error-correcting codes, invented in 2009 by Erdal Arıkan, based on the universal phenomenon of ***channel polarization.***



Binary-input symmetric DMC

Polarization

A new class of error-correcting codes, invented in 2009 by Erdal Arıkan, based on the universal phenomenon of ***channel polarization.***



Polar code of length $n = 2$

Polarization

As the mapping is invertible, we have

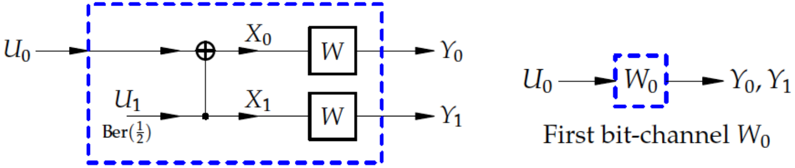
$$\begin{aligned} I(U_0, U_1; Y_0, Y_1) &= I(X_0, X_1; Y_0, Y_1) = I(X_0; Y_0) + I(X_1; Y_1) \\ &= 2I(X_1; Y_1) = 2I(W). \end{aligned}$$

The chain rule decomposition

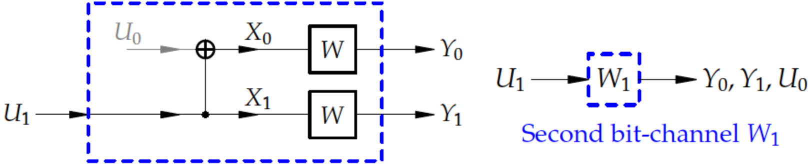
$$\begin{aligned} I(U_0, U_1; Y_0, Y_1) &= I(U_0; Y_0, Y_1) + I(U_1; Y_0, Y_1 | U_0) \\ &= 2I(W). \end{aligned}$$

Consider “virtual” sub-channels!

"Virtual" sub-channels



First bit-channel W_0



Second bit-channel W_1

“Virtual” sub-channels

Two channels are created

$$(W, W) \to (W_0, W_1)$$

W_0 is worse than W, W_1 is better than W. Will be denoted also as W^- and W^+.

Example

Polarization is easy to analyze when W is a BEC.

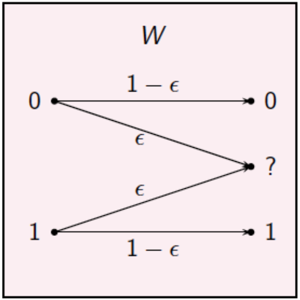
If W is a BEC(ϵ), then so are W^- and W^+, with erasure probabilities

$$\epsilon^- \triangleq 2\epsilon - \epsilon^2$$

and

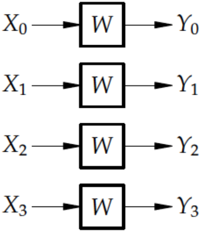
$$\epsilon^+ \triangleq \epsilon^2$$

respectively.



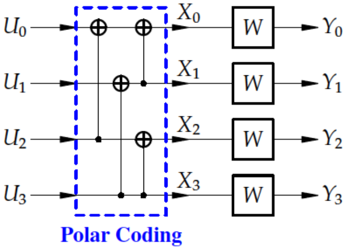
$n = 4$

A new class of error-correcting codes, invented in 2009 by Erdal Arıkan, based on the universal phenomenon of ***channel polarization.***



Four independent uses of the channel W

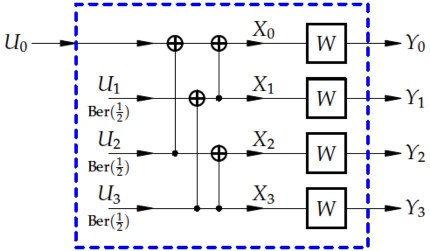
A new class of error-correcting codes, invented in 2009 by Erdal Arıkan, based on the universal phenomenon of ***channel polarization.***

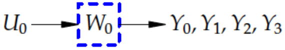


Polar code of length $n = 4$

$n = 4$

A new class of error-correcting codes, invented in 2009 by Erdal Arıkan, based on the universal phenomenon of ***channel polarization.***

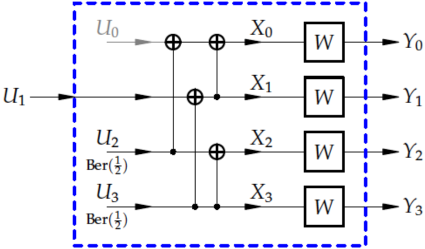


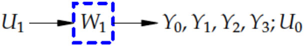


First bit-channel W_0

$n = 4$

A new class of error-correcting codes, invented in 2009 by Erdal Arıkan, based on the universal phenomenon of ***channel polarization.***

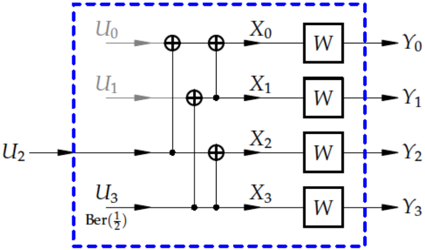


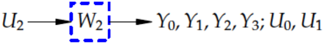


Second bit-channel W_1

$n = 4$

A new class of error-correcting codes, invented in 2009 by Erdal Arıkan, based on the universal phenomenon of ***channel polarization.***

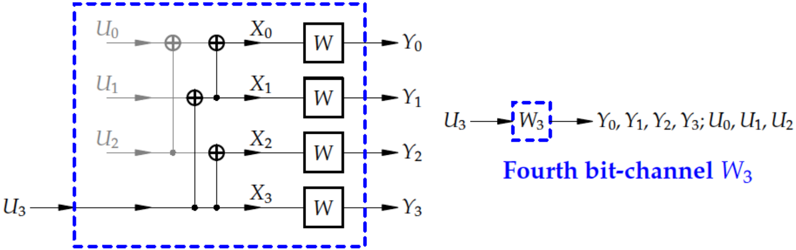




Third bit-channel W_2

$n = 4$

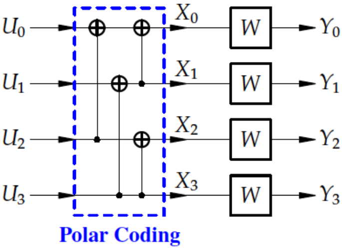
A new class of error-correcting codes, invented in 2009 by Erdal Arıkan, based on the universal phenomenon of ***channel polarization.***

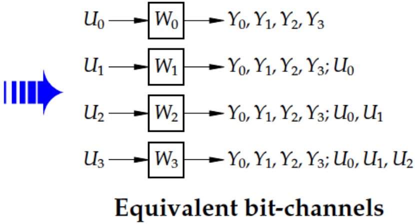


Fourth bit-channel W_3

$n = 4$

A new class of error-correcting codes, invented in 2009 by Erdal Arıkan, based on the universal phenomenon of ***channel polarization.***



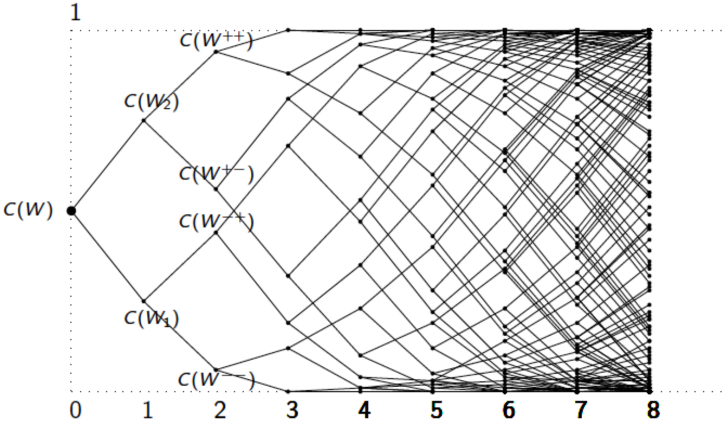


Polarization Theorem [Arıkan 2009]

As $n \to \infty$, the bit-channels polarize: they become either good (nearly noiseless) or bad (nearly useless). The fraction of good channels approaches capacity(W).

Polar codes: send information bits over the good channels, freeze the input to the bad channels to *a priori* known values (say, zeros).

General n



Outline

1 Polarization

2 Encoding

3 Decoding

Kroneker product

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

Kroneker product

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

Kroneker product

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

Kroneker product

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

Kroneker product

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

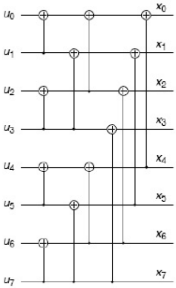
Kroneker product

length $N = 2^m$, $m \in \mathbb{N}$

generator matrix: rows of $G_2^{\otimes m}$

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

$$\bar{x} = [u_0\ u_1\ u_2\ u_3\ u_4\ u_5\ u_6\ u_7] G_2^{\otimes 3}$$



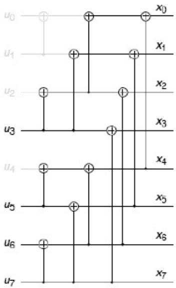
Kroneker product

length $N = 2^m$, $m \in \mathbb{N}$

generator matrix: rows of $G_2^{\otimes m}$

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

$$\bar{x} = [0\ 0\ 0\ u_3\ 0\ u_5\ u_6\ u_7]G_2^{\otimes 3}$$



How to choose the rows?

Reed-Muller codes

length $N = 2^m$, $m \in \mathbb{N}$

generator matrix: rows of $G_2^{\otimes m}$

How to choose the rows?

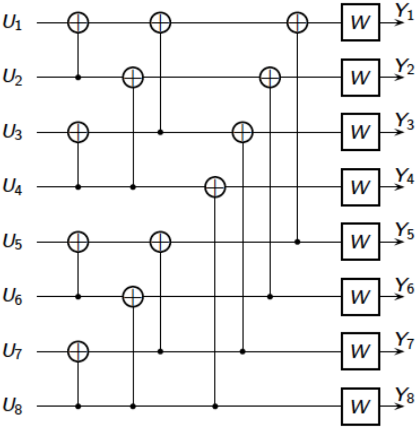
$$G_2^{\otimes 3} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

choose rows of largest weight

$$\bar{x} = [0\ 0\ 0\ u_3\ 0\ u_5\ u_6\ u_7]G_2^{\otimes 3}$$

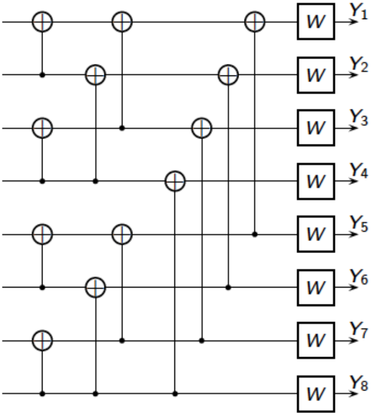
$W = BEC(0.5),\ N = 8,\ R = 1/2$

$I(W_i)$
0.0039
0.1211
0.1914
0.6836
0.3164
0.8086
0.8789
0.9961



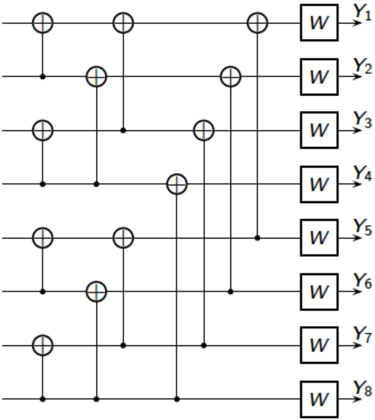
$W = BEC(0.5)$, $N = 8$, $R = 1/2$

$I(W_i)$	Rank		
0.0039	8	frozen	U_1
0.1211	7	frozen	U_2
0.1914	6	frozen	U_3
0.6836	4	data	U_4
0.3164	5	frozen	U_5
0.8086	3	data	U_6
0.8789	2	data	U_7
0.9961	1	data	U_8



$W = BEC(0.5),\ N = 8,\ R = 1/2$

$I(W_i)$	Rank		
0.0039	8	frozen	0
0.1211	7	frozen	0
0.1914	6	frozen	0
0.6836	4	data	U_4
0.3164	5	frozen	0
0.8086	3	data	U_6
0.8789	2	data	U_7
0.9961	1	data	U_8



$W = BEC(0.5)$, $N = 8$, $R = 1/2$

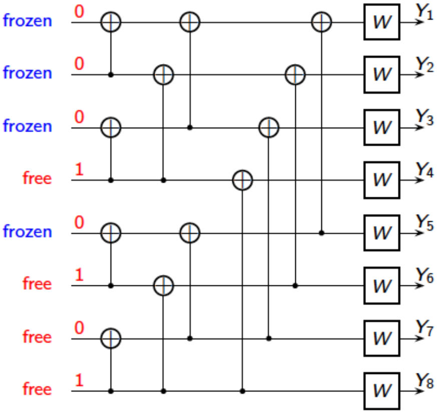
Theorem

Encoding complexity for polar coding is $\mathcal{O}(N \log N)$.

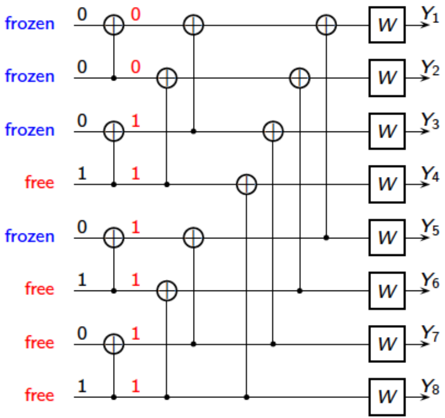
Proof:

- Polar coding transform can be represented as a graph with $N[1 + \log(N)]$ variables.
- The graph has $(1 + \log(N))$ levels with N variables at each level.
- Computation begins at the source level and can be carried out level by level.
- Space complexity $O(N)$, time complexity $O(N \log N)$.

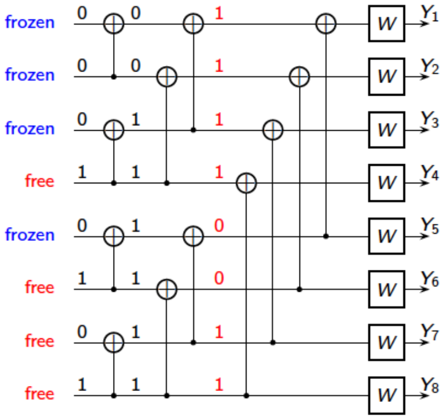
$W = BEC(0.5)$, $N = 8$, $R = 1/2$



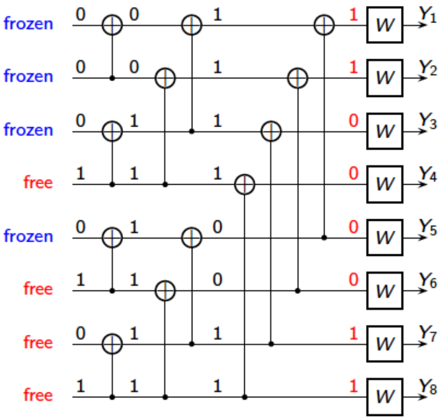
$W = BEC(0.5)$, $N = 8$, $R = 1/2$



$W = BEC(0.5)$, $N = 8$, $R = 1/2$



$W = BEC(0.5)$, $N = 8$, $R = 1/2$



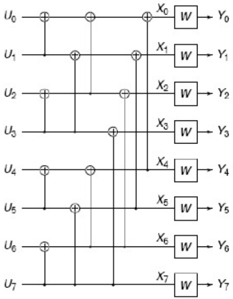
Outline

1 Polarization

2 Encoding

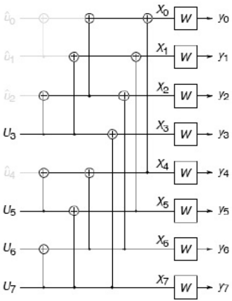
3 Decoding

Decoding



U_0
U_1
U_2
U_3
U_4
U_5
U_6
U_7
X_0
X_1
X_2
X_3
X_4
X_5
X_6
X_7
W
W
W
W
W
W
W
W
Y_0
Y_1
Y_2
Y_3
Y_4
Y_5
Y_6
Y_7

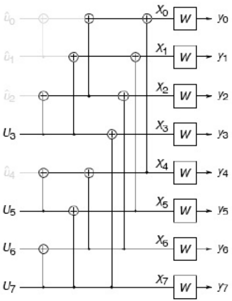
$F = \{0, 1, 2, 4\}$


$\hat{u}_0$
$\hat{u}_1$
$\hat{u}_2$
u_3
$\hat{u}_4$
u_5
u_6
u_7
x_0
x_1
x_2
x_3
x_4
x_5
x_6
x_7
W
y_0
y_1
y_2
y_3
y_4
y_5
y_6
y_7

Decoding

$F = \{0, 1, 2, 4\}$

From 0 till $N - 1$



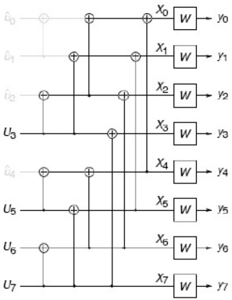
Decoding

$F = \{0, 1, 2, 4\}$

From 0 till $N - 1$

if $i \in F$, $\hat{u}_i = 0$
if $i \in F^c$,

$$\hat{u}_i = \begin{cases} 0, & \text{if } \frac{P(0|\hat{u}_0^{i-1}, \bar{y})}{P(1|\hat{u}_0^{i-1}, \bar{y})} > 1 \\ 1, & \text{otherwise} \end{cases}$$



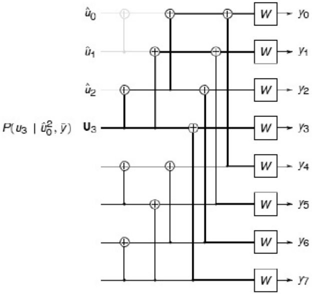
Decoding

$F = \{0, 1, 2, 4\}$

From 0 till $N-1$

if $i \in F$, $\hat{u}_i = 0$
if $i \in F^c$,

$$\hat{u}_i = \begin{cases} 0, & \text{if } \frac{P(0|\hat{u}_0^{i-1}, \bar{y})}{P(1|\hat{u}_0^{i-1}, \bar{y})} > 1 \\ 1, & \text{otherwise} \end{cases}$$



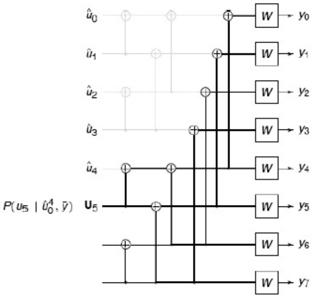
Decoding

$F = \{0, 1, 2, 4\}$

From 0 till $N - 1$

if $i \in F$, $\hat{u}_i = 0$
if $i \in F^c$,

$$\hat{u}_i = \begin{cases} 0, & \text{if } \frac{P(0|\hat{u}_0^{i-1}, \bar{y})}{P(1|\hat{u}_0^{i-1}, \bar{y})} > 1 \\ 1, & \text{otherwise} \end{cases}$$



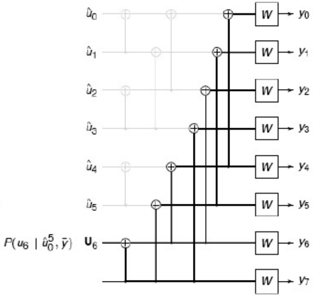
Decoding

$F = \{0, 1, 2, 4\}$

From 0 till $N-1$

if $i \in F$, $\hat{u}_i = 0$
if $i \in F^c$,

$$\hat{u}_i = \begin{cases} 0, & \text{if } \frac{P(0|\hat{u}_0^{i-1}, \bar{y})}{P(1|\hat{u}_0^{i-1}, \bar{y})} > 1 \\ 1, & \text{otherwise} \end{cases}$$



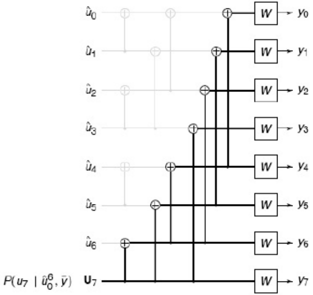
Decoding

$F = \{0, 1, 2, 4\}$

From 0 till $N-1$

if $i \in F$, $\hat{u}_i = 0$
if $i \in F^c$,

$$\hat{u}_i = \begin{cases} 0, & \text{if } \frac{P(0|\hat{u}_0^{i-1},\bar{y})}{P(1|\hat{u}_0^{i-1},\bar{y})} > 1 \\ 1, & \text{otherwise} \end{cases}$$



Decoding

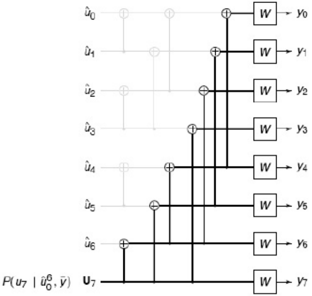
$F = \{0, 1, 2, 4\}$

From 0 till $N - 1$

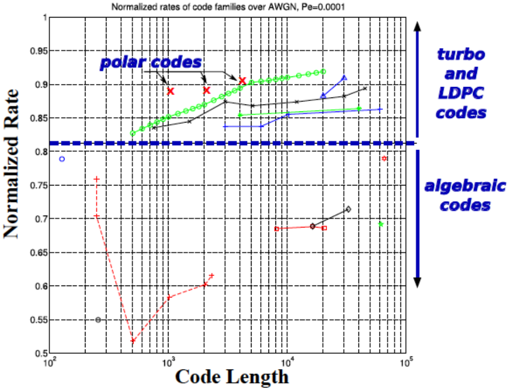
if $i \in F$, $\hat{u}_i = 0$
if $i \in F^c$,

$$\hat{u}_i = \begin{cases} 0, & \text{if } \frac{P(0|\hat{u}_0^{i-1},\bar{y})}{P(1|\hat{u}_0^{i-1},\bar{y})} > 1 \\ 1, & \text{otherwise} \end{cases}$$

Complexity $O(N \log N)$

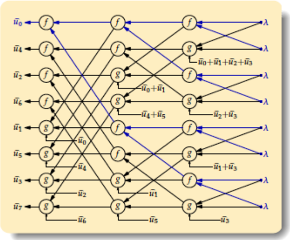


Comparison to another coding schemes



- Polar codes achieve ***better performance*** than the best-known LDPC codes with lower complexity, even at very short code lengths.

Summary



- Polar codes provably **achieve capacity** on general symmetric channels.
- Polar codes have **low encoding and decoding complexity** — both $O(n \log n)$.
- Polar codes **are explicit:** there is no random ensemble to choose from; they also do not suffer from an **error-floor.**
- Polar codes have **beautiful structure** that resembles the fast Fourier transform. This makes them **well suited for VLSI** implementation.

  G. Sarkis, P. Giard, A. Vardy, C. Thibeault, and **W.J. Gross**, Fast polar decoders: Algorithm and implementation, *IEEE Journal on Selected Areas in Communications*, **32**, pp. 946–957, May 2014.

- Polar codes work in **many important scenarios** other than point-to-point channel coding: wiretap channels, broadcast channels, multi-user channels, Wyner-Ziv coding, source coding, and more...

Thank you for your attention!