Polar Coding
Part 2 - Construction and Performance

Erdal Arıkan

Electrical-Electronics Engineering Department,
Bilkent University, Ankara, Turkey

Algorithmic Coding Theory Workshop
June 13 - 17, 2016
ICERM, Providence, RI
Table of Contents

Channel polarization

Polar coding

Polar codes for selected applications
Channel polarization

Polar coding

Polar codes for selected applications
The channel

Let $W : X \rightarrow Y$ be a binary-input discrete memoryless channel

- input alphabet: $\mathcal{X} = \{0, 1\}$,
- output alphabet: $\mathcal{Y}$,
- transition probabilities:

$$W(y|x), \quad x \in \mathcal{X}, y \in \mathcal{Y}$$
Symmetry assumption

Assume that the channel has “input-output symmetry.”

Examples:

- **BSC(\(\epsilon\))**
  - Input: 0
  - Output: 0 with probability 1 - \(\epsilon\)
  - Output: 1 with probability \(\epsilon\)
  - Input: 1
  - Output: 1 with probability 1 - \(\epsilon\)
  - Output: 0 with probability \(\epsilon\)

- **BEC(\(\epsilon\))**
  - Input: 0
  - Output: 0 with probability 1 - \(\epsilon\)
  - Output: ? with probability \(\epsilon\)
  - Input: 1
  - Output: 1 with probability 1 - \(\epsilon\)
  - Output: ? with probability \(\epsilon\)
For channels with input-output symmetry, the capacity is given by

\[ C(W) \triangleq I(X; Y), \quad \text{with } X \sim \text{unif. } \{0, 1\} \]

Use base-2 logarithms:

\[ 0 \leq C(W) \leq 1 \]
The main idea

- Channel coding problem trivial for two types of channels
  - Perfect: \( C(W) = 1 \)
  - Useless: \( C(W) = 0 \)
- Transform ordinary \( W \) into such extreme channels
The method: aggregate and redistribute capacity

Original channels (uniform)

\[ W \]

Vector channel

\[ W_{\text{vec}} \]

New channels (polarized)

\[ W_1 \]

\[ W_{N-1} \]

\[ W_N \]

Combine → Split
Combining

- Begin with $N$ copies of $W$,
- use a 1-1 mapping

$$G_N : \{0,1\}^N \rightarrow \{0,1\}^N$$

- to create a vector channel

$$W_{vec} : U^N \rightarrow Y^N$$
Conservation of capacity

Combining operation is lossless:
- Take $U_1, \ldots, U_N$ i.i.d. unif. $\{0, 1\}$
- then, $X_1, \ldots, X_N$ i.i.d. unif. $\{0, 1\}$
- and

\[
C(W_{\text{vec}}) = I(U^N; Y^N) \\
= I(X^N; Y^N) \\
= NC(W)
\]
Splitting

\[ C(W_{vec}) = I(U^N; Y^N) \]
\[ = \sum_{i=1}^{N} I(U_i; Y^N, U^{i-1}) \]
\[ = \sum_{i=1}^{N} C(W_i) \]

Define bit-channels

\[ W_i : U_i \rightarrow (Y^N, U^{i-1}) \]
Polarization is commonplace

- Polarization is the rule not the exception
- A random permutation $G_N : \{0, 1\}^N \rightarrow \{0, 1\}^N$

is a good polarizer with high probability

- Equivalent to Shannon’s random coding approach

Channel polarization

The method
Random polarizers: stepwise, isotropic

Isotropy: any redistribution order is as good as any other.
The complexity issue

- Random polarizers lack structure, too complex to implement
- Need a low-complexity polarizer
- May sacrifice stepwise, isotropic properties of random polarizers in return for less complexity
Basic module for a low-complexity scheme

Combine two copies of $W$

and split to create two bit-channels

$W_1 : U_1 \rightarrow (Y_1, Y_2)$

$W_2 : U_2 \rightarrow (Y_1, Y_2, U_1)$
The first bit-channel $W_1$

$W_1 : U_1 \rightarrow (Y_1, Y_2)$

$C(W_1) = I(U_1; Y_1, Y_2)$
The second bit-channel $W_2$

$W_2 : U_2 \rightarrow (Y_1, Y_2, U_1)$

$C(W_2) = I(U_2; Y_1, Y_2, U_1)$
Capacity conserved but redistributed unevenly

Conservation:

\[ C(W_1) + C(W_2) = 2C(W) \]

Extremization:

\[ C(W_1) \leq C(W) \leq C(W_2) \]

with equality iff \( C(W) \) equals 0 or 1.
Extremality of BEC

\[ H(U_1|Y_1Y_2) \leq H(X_1|Y_1) + H(X_2|Y_2) \]
\[ - H(X_1|Y_1)H(X_2|Y_2) \]

with equality iff \( W \) is a BEC.
Extremality of BSC (Mrs. Gerber’s lemma)

Let $\mathcal{H}^{-1} : [0, \frac{1}{2}] \rightarrow [0, \frac{1}{2}]$ be the inverse of the binary entropy function $\mathcal{H}(p) = -p \log(p) - (1 - p) \log(1 - p)$, $0 \leq p \leq \frac{1}{2}$.

$$H(U_1 | Y_1 Y_2) \geq \mathcal{H}(\mathcal{H}^{-1}(H(X_1 | Y_1)) * \mathcal{H}^{-1}(H(X_2 | Y_2)))$$

with equality iff $W$ is a BSC.
The two channels created by the basic transform

\[(W, W) \rightarrow (W_1, W_2)\]

will be denoted also as

\[W^- = W_1 \text{ and } W^+ = W_2\]

Likewise, we write \(W^{--}\), \(W^{+-}\) for descendants of \(W^-\); and \(W^{+-}\), \(W^{++}\) for descendants of \(W^+\).
For the size-4 construction
... duplicate the basic transform
... obtain a pair of $W^-$ and $W^+$ each
... apply basic transform on each pair
... decode in the indicated order

\[ U_1 \rightarrow W^- \rightarrow U_3 \rightarrow W^+ \rightarrow U_2 \rightarrow W^- \rightarrow U_4 \rightarrow W^+ \]
... obtain the four new bit-channels

\[ U_1 \quad W^{--} \]
\[ U_3 \quad W^{+-} \]
\[ U_2 \quad W^{-+} \]
\[ U_4 \quad W^{++} \]
Overall size-4 construction

Channel polarization

Recursive method
“Rewire” for standard-form size-4 construction

Channel polarization

Recursive method
Size 8 construction
Polarization of a BEC $W$

Polarization is easy to analyze when $W$ is a BEC.

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

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

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

respectively.

![Diagram showing the recursive method](image-url)
The first bit channel $W^-$ is a BEC.

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

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

and

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

respectively.
The second bit channel $W^+$ is a BEC.

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

$$\epsilon^- \overset{\Delta}{=} 2\epsilon - \epsilon^2$$

and

$$\epsilon^+ \overset{\Delta}{=} \epsilon^2$$

respectively.

![Diagram showing channel polarization and recursive method](image)
Polarization for BEC($\frac{1}{2}$): $N = 16$

Channel polarization

Recursive method
Polarization for BEC($\frac{1}{2}$): $N = 32$
Polarization for BEC($\frac{1}{2}$): $N = 64$
Polarization for BEC\(^{\frac{1}{2}}\): \( N = 128 \)

![Graph showing capacity of bit channels vs bit channel index for BEC(\( \frac{1}{2} \)) with \( N = 128 \)].
Polarization for BEC($\frac{1}{2}$): $N = 256$
Polarization for BEC($\frac{1}{2}$): $N = 512$
Polarization for BEC($\frac{1}{2}$): $N = 1024$
Theorem (Polarization, A. 2007)

The bit-channel capacities \( \{ C(W_i) \} \) polarize: for any \( \delta \in (0, 1) \), as the construction size \( N \) grows

\[
\left[ \frac{\text{no. channels with } C(W_i) > 1 - \delta}{N} \right] \longrightarrow C(W)
\]

and

\[
\left[ \frac{\text{no. channels with } C(W_i) < \delta}{N} \right] \longrightarrow 1 - C(W)
\]

Theorem (Rate of polarization, A. and Telatar (2008))

Above theorem holds with \( \delta \approx 2^{-\sqrt{N}} \).
Channel polarization

Polar coding

Polar codes for selected applications
Polar coding

- Code construction
- Encoding
- Decoding
- Performance
Polar code example: $W = \text{BEC}(\frac{1}{2})$, $N = 8$, rate $1/2$

<table>
<thead>
<tr>
<th>$I(W_i)$</th>
<th>Rank</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0039</td>
<td>8</td>
<td>frozen</td>
</tr>
<tr>
<td>0.1211</td>
<td>7</td>
<td>frozen</td>
</tr>
<tr>
<td>0.1914</td>
<td>6</td>
<td>frozen</td>
</tr>
<tr>
<td>0.6836</td>
<td>4</td>
<td>data</td>
</tr>
<tr>
<td>0.3164</td>
<td>5</td>
<td>frozen</td>
</tr>
<tr>
<td>0.8086</td>
<td>3</td>
<td>data</td>
</tr>
<tr>
<td>0.8789</td>
<td>2</td>
<td>data</td>
</tr>
<tr>
<td>0.9961</td>
<td>1</td>
<td>data</td>
</tr>
</tbody>
</table>

$Y_1 \rightarrow W \rightarrow Y_2 \rightarrow W \rightarrow Y_3 \rightarrow W \rightarrow Y_4 \rightarrow W \rightarrow Y_5 \rightarrow W \rightarrow Y_6 \rightarrow W \rightarrow Y_7 \rightarrow W \rightarrow Y_8 \rightarrow W$
Polar code example: $W = \text{BEC}(\frac{1}{2})$, $N = 8$, rate $1/2$

<table>
<thead>
<tr>
<th>$I(W_i)$</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0039</td>
<td>8</td>
</tr>
<tr>
<td>0.1211</td>
<td>7</td>
</tr>
<tr>
<td>0.1914</td>
<td>6</td>
</tr>
<tr>
<td>0.6836</td>
<td>4</td>
</tr>
<tr>
<td>0.3164</td>
<td>5</td>
</tr>
<tr>
<td>0.8086</td>
<td>3</td>
</tr>
<tr>
<td>0.8789</td>
<td>2</td>
</tr>
<tr>
<td>0.9961</td>
<td>1</td>
</tr>
</tbody>
</table>

Diagram showing the encoding process with nodes $U_i$ and $W_i$ connected to outputs $Y_i$. The diagram illustrates the decoding process for the given polar code parameters.
Encoding complexity

Theorem

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

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 $O(N \log N)$. 

Encoding: an example

Polar coding
Encoding: an example

Polar coding

Encoding
Encoding: an example

Polar coding  Encoding
Encoding: an example

```
frozen 0 0 1 1
frozen 0 0 1 1
frozen 0 1 1 0
free  1 1 1 0
frozen 0 1 0 0
free  1 1 0 0
free  0 1 1 1
free  1 1 1 1
```

```
Y1
Y2
Y3
Y4
Y5
Y6
Y7
Y8
```
Successive Cancellation Decoding (SCD)

Theorem

The complexity of successive cancellation decoding for polar codes is $O(N \log N)$.

Proof: Given below.
SCD: Exploit the $x = \lfloor a \lfloor a + b \rfloor$ structure
First phase: treat $a$ as noise, decode $(u_1, u_2, u_3, u_4)$
End of first phase
Second phase: Treat $\hat{b}$ as known, decode $(u_5, u_6, u_7, u_8)$
First phase in detail

\[ u_1 \quad b_1 \quad x_1 \quad w \quad y_1 \]
\[ u_2 \quad b_2 \quad x_2 \quad w \quad y_2 \]
\[ u_3 \quad b_3 \quad x_3 \quad w \quad y_3 \]
\[ u_4 \quad b_4 \quad x_4 \quad w \quad y_4 \]
\[ \text{noise } a_1 \quad x_5 \quad w \quad y_5 \]
\[ \text{noise } a_2 \quad x_6 \quad w \quad y_6 \]
\[ \text{noise } a_3 \quad x_7 \quad w \quad y_7 \]
\[ \text{noise } a_4 \quad x_8 \quad w \quad y_8 \]
Equivalent channel model

\[ b_1 \rightarrow x_1 \rightarrow y_1 \]
\[ b_2 \rightarrow x_2 \rightarrow y_2 \]
\[ b_3 \rightarrow x_3 \rightarrow y_3 \]
\[ b_4 \rightarrow x_4 \rightarrow y_4 \]
\[ \text{noise } a_1 \rightarrow x_5 \rightarrow y_5 \]
\[ \text{noise } a_2 \rightarrow x_6 \rightarrow y_6 \]
\[ \text{noise } a_3 \rightarrow x_7 \rightarrow y_7 \]
\[ \text{noise } a_4 \rightarrow x_8 \rightarrow y_8 \]
First copy of $W^-$
Second copy of $W^-$
Third copy of $W^-$
Fourth copy of $W^-$
Decoding on $W^-$

$u_1 \rightarrow b_1 \rightarrow w^- \rightarrow (y_1, y_5)$

$u_2 \rightarrow b_2 \rightarrow w^- \rightarrow (y_2, y_6)$

$u_3 \rightarrow b_3 \rightarrow w^- \rightarrow (y_3, y_7)$

$u_4 \rightarrow b_4 \rightarrow w^- \rightarrow (y_4, y_8)$
\[ b = |t|t + w | \]

\[ \begin{align*}
  u_1 & \quad w_1 & \quad b_1 & \quad w^- & \quad (y_1, y_5) \\
  u_2 & \quad w_2 & \quad b_2 & \quad w^- & \quad (y_2, y_6) \\
  u_3 & \quad t_1 & \quad b_3 & \quad w^- & \quad (y_3, y_7) \\
  u_4 & \quad t_2 & \quad b_4 & \quad w^- & \quad (y_4, y_8)
\end{align*} \]
Decoding on $W^{--}$

\[ u_1 \xrightarrow{w_1} w^{--} (y_1, y_3, y_5, y_7) \]

\[ u_2 \xrightarrow{w_2} w^{--} (y_2, y_4, y_6, y_8) \]
Decoding on $W$---

\[ u_1 \xrightarrow{\text{W}---} (y_1, y_2, \ldots, y_8) \]

Compute

\[ L--- \triangleq \frac{W---(y_1, \ldots, y_8 \mid u_1 = 0)}{W---(y_1, \ldots, y_8 \mid u_1 = 1)} \]

and set

\[ \hat{u}_1 = \begin{cases} u_1 & \text{if } u_1 \text{ is frozen} \\ 0 & \text{else if } L--- > 0 \\ 1 & \text{else} \end{cases} \]
Decoding on $W^{---+}$

- Known $\hat{u}_1$
- $u_2$

Diagram:

- $W^{--}$
- $(y_1, y_3, y_5, y_7)$
- $(y_2, y_4, y_6, y_8)$
Decoding on $W^{--+}$

Compute

$$L^{--+} \overset{\Delta}{=} \frac{W^{--+}(y_1, \ldots, y_8, \hat{u}_1 | u_2 = 0)}{W^{--+}(y_1, \ldots, y_8, \hat{u}_1 | u_2 = 1)}$$

and set

$$\hat{u}_2 = \begin{cases} u_2 & \text{if } u_2 \text{ is frozen} \\ 0 & \text{else if } L^{--+} > 0 \\ 1 & \text{else} \end{cases}$$
Complexity for successive cancelation decoding

- Let $C_N$ be the complexity of decoding a code of length $N$
- Decoding problem of size $N$ for $W$ reduced to two decoding problems of size $N/2$ for $W^-$ and $W^+$
- So
  \[ C_N = 2C_{N/2} + kN \]
  for some constant $k$
- This gives $C_N = \mathcal{O}(N \log N)$
Performance of polar codes

Probability of Error (A. and Telatar (2008))

For any binary-input symmetric channel $W$, the probability of frame error for polar coding at rate $R < C(W)$ and using codes of length $N$ is bounded as

$$P_e(N, R) \leq 2^{-N^{0.49}}$$

for sufficiently large $N$.

A more refined versions of this result has been given given by S. H. Hassani, R. Mori, T. Tanaka, and R. L. Urbanke (2011).
Construction complexity

Polar codes can be constructed in time $O(N \text{poly}(\log(N)))$.

This result has been developed in a sequence of papers by:

- R. Mori and T. Tanaka (2009)
- I. Tal and A. Vardy (2011)
Trifonov (2011) introduced a Gaussian approximation technique for constructing polar codes.

Dai et al. (2015) studied various refinements of Gaussian approximation for polar code construction.

These methods work extremely well although a satisfactory explanation of why they work is still missing.
Example of Gaussian approximation

Polar code construction and performance estimation by Gaussian approximation

![Graph showing performance of Polar coding and Shannon limits with Gaussian approximation.](image)

- **FER** vs. **$E_s/N_0$ (dB)**
- **Polar(65536,61440,8) - BPSK**
- **Ultimate Shannon limit**
- **BPSK Shannon limit**
- **Threshold SNR at target FER**
- **Gaussian approximation**

- Gap to ultimate capacity = 3.42
- Gap to BPSK capacity = 1.06
Given $W$, $N = 2^n$, and $R < I(W)$, a polar code can be constructed such that it has

- construction complexity $\mathcal{O}(N \text{poly}(\log(N)))$,
- encoding complexity $\approx N \log N$,
- successive-cancellation decoding complexity $\approx N \log N$,
- frame error probability $P_e(N, R) = o \left(2^{-\sqrt{N} + o(\sqrt{N})}\right)$.
Performance improvement for polar codes

- Concatenation to improve minimum distance
- List decoding to improve SC decoder performance
## Concatenation

<table>
<thead>
<tr>
<th>Method</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>Block turbo coding with polar constituents</td>
<td>AKMOP (2009)</td>
</tr>
<tr>
<td>Generalized concatenated coding with polar inner</td>
<td>AM (2009)</td>
</tr>
<tr>
<td>Reed-Solomon outer, polar inner</td>
<td>BJE (2010)</td>
</tr>
<tr>
<td>Polar outer, block inner</td>
<td>SH (2010)</td>
</tr>
<tr>
<td>Polar outer, LDPC inner</td>
<td>EP (ISIT’2011)</td>
</tr>
</tbody>
</table>

AKMOP: A., Kim, Markarian, Özgür, Poyraz
GCC: A., Markarian
BJE: Bakshi, Jaggi, and Effros
SH: Seidl and Huber
EP: Eslami and Pishro-Nik
Overview of decoders for polar codes

- Successive cancellation decoding: A depth-first search method with complexity roughly $N \log N$
  - Sufficient to prove that polar codes achieve capacity
  - Equivalent to an earlier algorithm by Schnabl and Bossert (1995) for RM codes
  - Simple but not powerful enough to challenge LDPC and turbo codes in short to moderate lengths
- List decoding: A breadth-first search algorithm with limited branching (known as “beam search” in AI).
  - First proposed by Tal and Vardy (2011) for polar codes.
  - List decoding was used earlier by Dumer and Shabunov (2006) for RM codes
  - Complexity grows as $O(LN \log N)$ for a list size $L$. But hardware implementation becomes problematic as $L$ grows due to sorting and memory management.
- Sphere-decoding (“British Museum” search with branch and bound, starts decoding from the opposite side).
List decoder for polar codes

- First produce $L$ candidate decisions
- Pick the most likely word from the list
- Complexity $O(LN \log N)$
Polar code performance

Successive cancellation decoder

![Graph showing the performance of successive cancellation decoder with EsNo (dB) on the x-axis and FER on the y-axis. The graph indicates the performance of \(P(2048,1024)\), 4-QAM, L-1, CRC-0, SNR = 2.]

Polar coding

Performance
Improvement by list-decoding: List-32

![Graph showing the performance of Polar codes with different list-decoding sizes. The x-axis represents Es/No (dB) and the y-axis represents FER (Frames Error Rate). Two curves are shown: one for P(2048,1024), 4-QAM, L-1, CRC-0, SNR = 2, and another for P(2048,1024), 4-QAM, L-32, CRC-0, SNR = 2. The graph illustrates a clear improvement in performance with list-decoding size.]
Polar code performance

Improvement by list-decoding: List-1024

![Graph showing Polar code performance with list-decoding for different scenarios.]

- P(2048,1024), 4-QAM, L-1, CRC-0, SNR = 2
- P(2048,1024), 4-QAM, L-32, CRC-0, SNR = 2
- P(2048,1024), 4-QAM, L-1024, CRC-0, SNR = 2
Polar code performance

Comparison with ML bound

![Graph showing FER vs. EsNo for different polar coding configurations and an ML Bound comparison.](image-url)
Polar code performance

Introducing CRC improves performance at high SNR
Polar code performance

Comparison with dispersion bound

- $P(2048,1024)$, 4-QAM, L-1, CRC-0, SNR = 2
- $P(2048,1024)$, 4-QAM, L-32, CRC-0, SNR = 2
- $P(2048,1024)$, 4-QAM, L-1024, CRC-0, SNR = 2
- ML Bound for $P(2048,1024)$, 4-QAM
- $P(2048,1024)$, 4-QAM, L-32, CRC-16, SNR = 2
- Dispersion bound for (2048,1024)
Polar codes vs WiMAX Turbo Codes

Comparable performance obtained with List-32 + CRC
Polar codes vs WiMAX LDPC Codes

Better performance obtained with List-32 + CRC
Polar Codes vs DVB-S2 LDPC Codes

LDPC (16200,13320), Polar (16384,13421). Rates = 0.82. BPSK-AWGN channel.

![Graph showing Frame Error Rate (FER) vs Eb/N0 for Polar and DVB-S2 codes.](image)
Park (2014) gives the following performance comparison.

(Park’s result on LDPC conflicts with reference IEEE 802.11-10/0432r2. Whether there exists an error floor as shown needs to be confirmed independently.)

Summary of performance comparisons

- Successive cancellation decoder is simplest but inherently sequential which limits throughput.
- BP decoder improves throughput and with careful design performance.
- List decoder but significantly improves performance at low SNR.
- Adding CRC to list decoding improves performance significantly at high SNR with little extra complexity.
- Overall, polar codes under list-32 decoding with CRC offer performance comparable to codes used in present wireless standards.
Implementation performance metrics

Implementation performance is measured by

- Chip area (mm²)
- Throughput (Mbits/sec)
- Energy efficiency (nJ/bit)
- Hardware efficiency (Mb/s/mm²)
Successive cancellation decoder comparisons

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Decoder Type</strong></td>
<td>SC</td>
<td>SC</td>
<td>BP</td>
</tr>
<tr>
<td><strong>Block Length</strong></td>
<td>1024</td>
<td>1024</td>
<td>1024</td>
</tr>
<tr>
<td><strong>Technology</strong></td>
<td>90 nm</td>
<td>65 nm</td>
<td>65 nm</td>
</tr>
<tr>
<td><strong>Area [mm²]</strong></td>
<td>3.213</td>
<td>0.68</td>
<td>1.476</td>
</tr>
<tr>
<td><strong>Voltage [V]</strong></td>
<td>1.0</td>
<td>1.2</td>
<td>1.0</td>
</tr>
<tr>
<td><strong>Frequency [MHz]</strong></td>
<td>2.79</td>
<td>1010</td>
<td>300</td>
</tr>
<tr>
<td><strong>Power [mW]</strong></td>
<td>32.75</td>
<td>-</td>
<td>477.5</td>
</tr>
<tr>
<td><strong>Throughput [Mb/s]</strong></td>
<td>2860</td>
<td>497</td>
<td>4676</td>
</tr>
<tr>
<td><strong>Engy.-per-bit [pJ/b]</strong></td>
<td>11.45</td>
<td>-</td>
<td>102.1</td>
</tr>
<tr>
<td><strong>Hard. Eff. [Mb/s/mm²]</strong></td>
<td>890</td>
<td>730</td>
<td>3168</td>
</tr>
</tbody>
</table>


¹ Throughput 730 Mb/s calculated by technology conversion metrics
² Performance at 4 dB SNR with average no of iterations 6.57
## BP decoder comparisons

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Decoding type and Scheduling</td>
<td></td>
<td>SCD with folded HPPSN</td>
<td>Specialized SC</td>
<td>BP Circular Unidirectional</td>
<td>BP Circular Unidirectional</td>
<td>BP All-ON, Fully Parallel</td>
</tr>
<tr>
<td>Block length</td>
<td></td>
<td>1024</td>
<td>16384</td>
<td>1024</td>
<td>1024</td>
<td>1024</td>
</tr>
<tr>
<td>Rate</td>
<td></td>
<td>0.9</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>Technology</td>
<td></td>
<td>CMOS</td>
<td>Altera Stratix 4</td>
<td>CMOS</td>
<td>CMOS</td>
<td>CMOS</td>
</tr>
<tr>
<td>Process nm</td>
<td></td>
<td>65</td>
<td>40</td>
<td>65</td>
<td>65</td>
<td>45</td>
</tr>
<tr>
<td>Core area mm²</td>
<td></td>
<td>0.068</td>
<td>1.48</td>
<td>1.48</td>
<td>12.46</td>
<td>1.65</td>
</tr>
<tr>
<td>Supply V</td>
<td></td>
<td>1.2</td>
<td>1.35</td>
<td>1</td>
<td>0.475</td>
<td>1</td>
</tr>
<tr>
<td>Frequency MHz</td>
<td></td>
<td>1010</td>
<td>106</td>
<td>300</td>
<td>50</td>
<td>606</td>
</tr>
<tr>
<td>Power mW</td>
<td></td>
<td>477.5</td>
<td>18.6</td>
<td>2056.5</td>
<td>328.4</td>
<td>328.4</td>
</tr>
<tr>
<td>Iterations</td>
<td></td>
<td>1</td>
<td>1</td>
<td>15</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Throughput Mb/s</td>
<td></td>
<td>497</td>
<td>1091</td>
<td>1024</td>
<td>171</td>
<td>2068</td>
</tr>
<tr>
<td>Energy efficiency pJ/b</td>
<td></td>
<td>102.1</td>
<td>23.8</td>
<td>110.5</td>
<td>19.3</td>
<td>19.3</td>
</tr>
<tr>
<td>Energy eff. per iter. pJ/b/iter</td>
<td></td>
<td>15.54</td>
<td>3.63</td>
<td>7.36</td>
<td>1.28</td>
<td>1.28</td>
</tr>
<tr>
<td>Area efficiency Mb/s/mm²</td>
<td></td>
<td>7306.78</td>
<td>693.77</td>
<td>99.80</td>
<td>166.01</td>
<td>1187.71</td>
</tr>
</tbody>
</table>

Normalized to 45 nm according to ITRS roadmap:

| Throughput Mb/s     |            | 613.4                           | 1263.8                          | 210.6          | 2068           | 1960                            |                                      |
| Energy efficiency pJ/b |            | 149.6                           | 34.9                            | 110.5          | 19.3           | 19.3                            |                                      |
| Area efficiency Mb/s/mm² |          | 18036.5                          | 1250.21                         | 179.85         | 166.01         | 1187.71                         |                                      |

* Throughput obtained by disabling the BP early-stopping rules for fair comparison.


Channel polarization

Polar coding

Polar codes for selected applications
Polar codes for selected applications

- 60 GHz wireless
- Optical access networks
- 5G
Millimeter Wave 60 GHz Communications

- 7 GHz of bandwidth available (57-64 GHz allocated in the US)
- Free-space path loss \((4\pi d/\lambda)^2\) is high at \(\lambda = 5\) mm but compensated by large antenna arrays.
- Propagation range limited severely by \(O_2\) absorption. Cells confined to rooms.

![Graph showing absorption in GHz](image-url)
Recent IEEE 802.11.ad Wi-Fi standard operates at 60 GHz ISM band and uses an LDPC code with block length 672 bits, rates 1/2, 5/8, 3/4, 13/16.

Two papers compare polar codes that study polar coding for 60 GHz applications:

Millimeter Wave 60 GHz Communications

Wei et al compare polar codes with the LDPC codes used in the standard using a nonlinear channel model

Wei, B. Li, and C. Zhao, “On the polar code for the 60 GHz millimeter-wave systems,” EURASIP, JWCN, 2015.
Millimeter Wave 60 GHz Communications

Wei et al compare polar codes with the LDPC codes used in the standard using a nonlinear channel model.

Wei, B. Li, and C. Zhao, “On the polar code for the 60 GHz millimeter-wave systems,” EURASIP, JWCN, 2015.
Wei et al compare polar codes with the LDPC codes used in the standard using a nonlinear channel model

Wei, B. Li, and C. Zhao, “On the polar code for the 60 GHz millimeter-wave systems,” EURASIP, JWCN, 2015.
Polar codes vs IEEE 802.11ad LDPC codes

Park (2014) gives the following performance comparison.

(Park’s result on LDPC conflicts with reference IEEE 802.11-10/0432r2. Whether there exists an error floor as shown needs to be confirmed independently.)

In terms of implementation complexity and throughput, Park (2014) gives the following figures.

<table>
<thead>
<tr>
<th></th>
<th>LPDC</th>
<th>Polar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Throughput Gb/s</td>
<td>0.5</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>0.779</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.676</td>
</tr>
<tr>
<td>Energy efficiency (pJ/b)</td>
<td>21</td>
<td>61.7</td>
</tr>
<tr>
<td></td>
<td>89.5</td>
<td>23.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>102.1</td>
</tr>
<tr>
<td>Area efficiency (Gb/s/mm2)</td>
<td>0.31</td>
<td>3.75</td>
</tr>
<tr>
<td></td>
<td>5.63</td>
<td>0.528</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.168</td>
</tr>
</tbody>
</table>

Optical access/transport network

- 10-100 Gb/s at 1E-12 BER
- OTU4 (100 Gb/s Ethernet) and ITU G.975.1 standards use Reed-Solomon (RS) codes
- The challenge is to provide high reliability at low hardware complexity.
Polar codes for optical access/transport

There have been some studies of polar codes for optical transmission.


- Z. Wu and B. Lankl, “Polar codes for low-complexity forward error correction in optical access networks,” ITG-Fachbericht 248: Photonische Netze - 05, 06.05.2014, Leipzig. (Compares polar codes with G.975.1 RS codes.)


Comparison of polar codes with G.975.1 RS codes

Comparison of polar codes with G.975.1 RS codes

Comparison of polar codes with all codes in G.975.1

In a recent MS thesis, T. Ahmad compared polar codes with G.975.1 codes.

Polar codes are optimized for a target BER of $10^{-12}$

- BER\(_{in(\text{coded})}\) (with FEC overhead)
- BER\(_{in(\text{uncoded})}\) (without FEC overhead)
- Concatenated BCH super FEC
- Concatenated RS/BCH
- RS/Product code, quantization = 2 bits
- RS/Product code, quantization = 1 bit
- LDPC super FEC
- RS (255, 239)
- RS (2720, 2550)
- Polar (2040, 1912)
- Polar (32640, 30592)
- Polar (130560, 122368)
- Polar (261120, 244736)

Polar codes for selected applications

Optical access
The conclusion of Ahmad (2016) is that polar codes perform better than all G.975.1 FEC schemes.

<table>
<thead>
<tr>
<th>FEC Code</th>
<th>$BER_{in}$</th>
<th>NCG (dB)</th>
<th>CG (dB)</th>
<th>Q (dB)</th>
<th>$\frac{Eb}{No}$ (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RS (255, 239)</td>
<td>1.82E-04</td>
<td>5.62</td>
<td>5.90</td>
<td>11.04</td>
<td>8.31</td>
</tr>
<tr>
<td>LDPC super FEC code</td>
<td>1.33E-03</td>
<td>7.10</td>
<td>7.39</td>
<td>9.56</td>
<td>6.83</td>
</tr>
<tr>
<td>RS (2720, 2550)</td>
<td>1.26E-03</td>
<td>7.06</td>
<td>7.34</td>
<td>9.60</td>
<td>6.87</td>
</tr>
<tr>
<td>Conc. RS/CSOC code (24.5%OH)</td>
<td>5.80E-03</td>
<td>7.95</td>
<td>8.90</td>
<td>8.04</td>
<td>5.31</td>
</tr>
<tr>
<td>Concatenated BCH code</td>
<td>3.30E-03</td>
<td>7.98</td>
<td>8.26</td>
<td>8.68</td>
<td>5.95</td>
</tr>
<tr>
<td>Conc. RS/BCH code</td>
<td>2.26E-03</td>
<td>7.63</td>
<td>7.91</td>
<td>9.06</td>
<td>6.34</td>
</tr>
<tr>
<td>Conc. RS/Product code</td>
<td>4.60E-03</td>
<td>8.40</td>
<td>8.68</td>
<td>8.30</td>
<td>5.57</td>
</tr>
<tr>
<td>Polar (2040, 1912)</td>
<td>2.81E-04</td>
<td>5.91</td>
<td>6.19</td>
<td>10.75</td>
<td>8.02</td>
</tr>
<tr>
<td>Polar (32640, 30592)</td>
<td>2.60E-03</td>
<td>7.74</td>
<td>8.02</td>
<td>8.92</td>
<td>6.20</td>
</tr>
<tr>
<td>Polar (130560, 122368)</td>
<td>4.61E-03</td>
<td>8.35</td>
<td>8.63</td>
<td>8.31</td>
<td>5.58</td>
</tr>
<tr>
<td>Polar (261120, 244736)</td>
<td>5.72E-03</td>
<td>8.60</td>
<td>8.89</td>
<td>8.06</td>
<td>5.33</td>
</tr>
</tbody>
</table>
Ahmad’s study finds that polar codes fall short of beating 3G FEC proposed for optical transport.

<table>
<thead>
<tr>
<th>FEC code</th>
<th>NCG (dB)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Polar (32640, 27200)</td>
<td>10.07</td>
<td>Ahmad (2016)</td>
</tr>
<tr>
<td>Polar (130560, 108800)</td>
<td>10.79</td>
<td>Ahmad (2016)</td>
</tr>
<tr>
<td>Polar (261120, 217600)</td>
<td>11.07</td>
<td>Ahmad (2016)</td>
</tr>
<tr>
<td>Polar (522240, 435200)</td>
<td>11.30</td>
<td>Ahmad (2016)</td>
</tr>
<tr>
<td>CC-LDPC (10032, 4, 24)</td>
<td>11.50</td>
<td>3G FEC, 12 iterations</td>
</tr>
<tr>
<td>QC-LDPC (18360, 15300)</td>
<td>11.30</td>
<td>3G FEC, 12 iterations</td>
</tr>
</tbody>
</table>
With list-decoding and CRC polar codes deliver comparable performance to LDPC and Turbo codes used in present wireless standards.

SoA in coding is already close to theoretical limits, leaving little margin for improvement.

Polar coding compared to SoA offers some advantages:

- Universal: the same hardware can be used with different code lengths, rates, channels
- Flexible: the code rate can be adjusted readily to any number between 0 and 1
- Versatile: can be used in multi-terminal coding scenarios
FEC for 5G

- What is 5G?
- What will be new in terms of FEC?
What is 5G?

Andrews et al.\(^3\) answer this question as follows.

- It will not be an incremental advance over 4G.
- Will be characterized by
  - Very high frequencies and massive bandwidths with very large number of antennas
  - Extreme base station and device connectivity
  - Universal connectivity between 5G new air interfaces, LTE, WiFi, etc.

\(^3\)Andrews et al., “What will 5G be?” JSAC 2014
Technical requirements for 5G

Again, according to Andrews et al., 5G will have to meet the following requirements (not all at once):

- Data rates compared to 4G
  - Aggregate: 1000 times more capacity/km2 compared to 4G
  - Cell-edge: 100 - 1000 Mb/s/user with 95% guarantee
  - Peak: 10s of Gb/s/user

- Round-trip latency: Some applications (tactile Internet, two-way gaming, virtual reality) will require 1 ms latency compared to 10-15 ms that 4G can provide

- Energy and cost: Link energy consumption should remain the same as data rates increase, meaning that a 100-times more energy-efficient link is required

- No of devices: 10,000 more low-rate devices for M2M communications, along with traditional high-rate users
Key technology ingredients for 5G

It is generally agreed that the 1000x aggregate data rate increase will be possible through a combination of three types of gains.

- Densification of network access nodes
- Increased bandwidth (move to mm waves)
- Increased spectral efficiency through new communication techniques:
  - advanced MIMO
  - improved multi-access
  - better interference management
  - improved coding and modulation schemes
Energy challenge

- Latest smartphone batteries have a power rating of about 10 Wh (2550 mAh at 3.85 V) or an energy of 36 kJ = 3.6E13 nJ
- Typical energy consumption at the decoder today is 1-10 nJ/bit
- At 3.6 nJ/bit and a data rate of 1 Gb/s, the FEC consumes all battery power in 20 mins!
- Technology challenge: Build FEC with energy consumption less than 10 pJ/bit
Throughput challenge

Typical backbone (fronthaul/backhaul) requirement for the near future:

- 1 Tb/s data rate
- Frame Error Rate (FER) better than 1E-15
- Latency: 10-200 $\mu$sec
- Total power consumption: 20 W max (20 pJ/bit)
Outlook for applied FEC research

Applied FEC research will remain an active area for the next decade

- Diverse set of applications, diverse set of requirements
- More sophisticated coding techniques to serve multi-user techniques (relaying, etc.) are in demand
- One-size-fits-all type of solution impossible
- A trade-off will emerge given the conflicting demands for energy-efficiency, low-latency, flexibility, and near-optimal performance
Thank you!