Polar-Coded Modulation
- A Tutorial -

Mathis Seidl and Johannes B. Huber

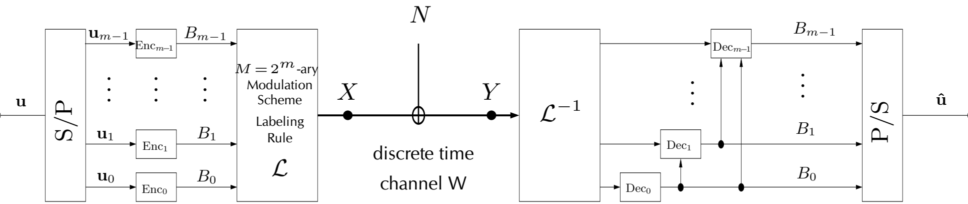
FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
TECHNISCHE FAKULTÄT

LEHRSTUHL FÜR INFORMATIONSÜBERTRAGUNG

Institute for Information Transmission
Friedrich-Alexander-Universität Erlangen-Nürnberg

Outline

- Coded Modulation by means of Multilevel Coding (MLC)
- Sequential Binary Partitions
- Concatenation of SBPs and Interpretation of Polar Codes as concatenated SBPs
- Concatenation of MLC and Polar Codes forming one Polar Code
- BICM and Polar Codes
- Conclusions

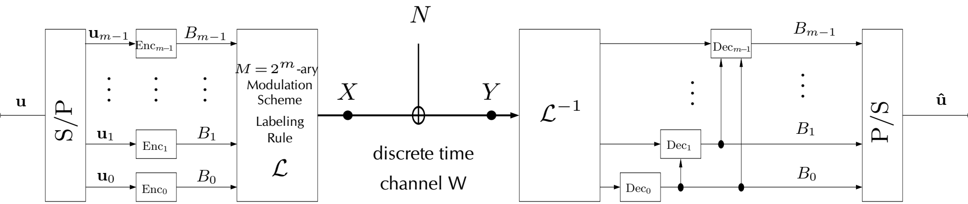


Symmetric channel capacity:

$$I(X;Y) = I\left(B_0, \ldots, B_{m-1}; Y\right) = \sum_{i=0}^{m-1} I\left(B_i; Y | B_0, \ldots, B_{i-1}\right) =: \sum_{i=0}^{m-1} I(\mathsf{B}^{(i)})$$

with $I(\mathsf{B}^{(i)})$: i-th level / i-th (symmetric) binary channel capacity

Multilevel Coding with Multistage Decoding



Symmetric channel capacity:

$$I(X;Y) = I\left(B_0,\ldots,B_{m-1};Y\right) = \sum_{i=0}^{m-1} I\left(B_i;Y|B_0,\ldots,B_{i-1}\right) =: \sum_{i=0}^{m-1} I(\mathsf{B}^{(i)})$$

with $I(\mathsf{B}^{(i)})$: i-th level / i-th (symmetric) binary channel capacity

- If rates R_i of the codes for equivalent binary channels B_i are chosen to $I(\mathsf{B}_i)$, channel capacity $I(X;Y)$ is achievable by means of capacity-achieving binary codes together with successive decoding
- Although the $I(\mathsf{B}_i)$ strongly depend on the chosen labeling rule $\mathcal{L}$, „capacity achievability" is irrespective of the choice of $\mathcal{L}$

Sequential Binary Partition

Let $\mathsf{W} : \mathcal{X} \mapsto \mathcal{Y}$ be a discrete, memoryless channel (DMC) with $|\mathcal{X}| = M = 2^m$ and mutual information $I(X;Y)$.

Sequential Binary Partition of order m of W (m-SBP of W):

SBP: Transformation $\quad \varphi : \mathsf{W} \mapsto \{\mathsf{B}^{(0)}, \ldots, \mathsf{B}^{(m-1)}\}$

$\mathsf{B}^{(i)} : \{0,1\} \mapsto \mathcal{Y} \times \{0,1\}^i$ B-DMC, binary channels

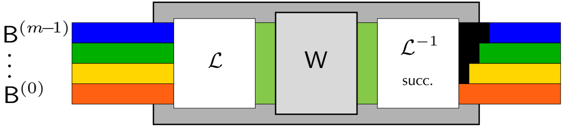
$I(\mathsf{B}^{(i)}) := I(B_i; Y | B_0, \ldots, B_{i-1})$ symmetric capacity

such that

$$\sum_{i=0}^{m-1} I(\mathsf{B}^{(i)}) = I(X;Y)$$ (Information is preserved)

Labeling: $\quad \mathcal{L} : B_0 \ldots B_{m-1} \mapsto X \quad$ ($(2^m)!$ possibilities)

Sequential Binary Partition



Properties of m-SBPs:

- average capacity per binary channel:

$$M_\varphi(\mathsf{W}) := \frac{1}{m}\sum_{i=0}^{m-1} I(\mathsf{B}^{(i)}) = \frac{1}{m} I(X;Y)$$

  - independent of labeling rule

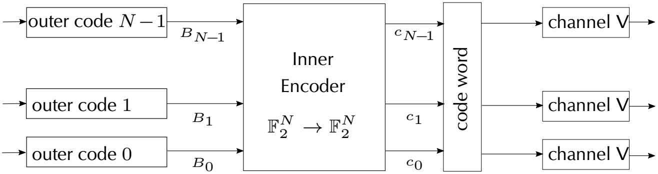
- variance of the binary channel capacities:

$$V_\varphi(\mathsf{W}) := \frac{1}{m}\sum_{i=0}^{m-1} I(\mathsf{B}^{(i)})^2 - M_\varphi(\mathsf{W})^2$$

  - strongly depending on labeling rule!

Sequential Binary Partition

Concatenated Channel Coding interpreted as Multilevel Coding:



- In old days of channel coding well known as „generalized concatenated codes"
- For $m = N$ and one-to-one mapping, a simple change of mapping and a transform of identical binary channels V into **differing equivalent channels** $\mathsf{B}^{(i)}$ ($i = 0, \ldots, N-1$) in a capacity-preserving way:

$$\sum_{i=0}^{N-1} I(\mathsf{B}^{(i)}) = N \cdot I(\mathsf{V})$$

with $I(\mathsf{B}^{(i)}) := I(B_i; \boldsymbol{Y} | B_0, \ldots, B_{i-1})$.

Sequential Binary Partitions

- Polar code as a „special“ „generalized“ concatenated code:
  - Inner encoder: $N \times N$ binary generator matrix

$$\boldsymbol{G}_N = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}^{\otimes \log_2(N)} \in \mathbb{F}_2^{N \times N}$$

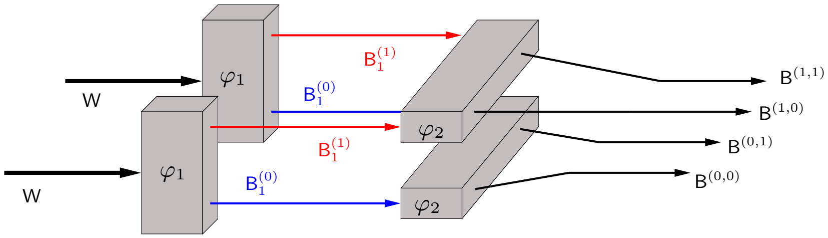
  - Outer encoders with rates

$$R_i = \begin{cases} 0 & \text{for frozen} \\ 1 & \text{for used} \end{cases} \text{ symbols}$$

- Polar-Coded Modulation:
  - Inner encoder: Labeling of coded modulation
  - Outer encoders: Polar codes with rates $R_i \leq I(\mathsf{B}^{(i)})$

Concatenation of two SBPs:

- φ_1: $\mathsf{W} \mapsto \{\mathsf{B}_1^{(0)}, \ldots, \mathsf{B}_1^{(k_1-1)}\}$
- φ_2: $\mathsf{B}^{k_2} \mapsto \{\mathsf{B}_2^{(0)}, \ldots, \mathsf{B}_2^{(k_2-1)}\}$ (Vector channel of k_2 B-DMCs B as input)



Concatenation of SBPs

Concatenation of two SBPs:

- φ_1: $\mathsf{W} \mapsto \{\mathsf{B}_1^{(0)}, \ldots, \mathsf{B}_1^{(k_1-1)}\}$
- φ_2: $\mathsf{B}^{k_2} \mapsto \{\mathsf{B}_2^{(0)}, \ldots, \mathsf{B}_2^{(k_2-1)}\}$ (Vector channel of k_2 B-DMCs B as input)

Concatenation is again an SBP (of order $k_1 k_2$):

$$\varphi_1 \circ \varphi_2 : \mathsf{W}^{k_2} \mapsto \{\mathsf{B}^{(0,0)}, \ldots, \mathsf{B}^{(0,k_2-1)}, \ldots, \mathsf{B}^{(k_1-1,0)}, \ldots, \mathsf{B}^{(k_1-1,k_2-1)}\}$$

$$\mathsf{B}^{(i,j)} : \{0,1\} \mapsto \mathcal{Y}^{k_2} \times \{0,1\}^{k_2 i + j}$$ B-DMC, binary channels

$$I(\mathsf{B}^{(i,j)}) := I(B_{i,j}; \boldsymbol{Y} | \boldsymbol{B}_0 \ldots \boldsymbol{B}_{i-1}, B_{i,0}, \ldots, B_{i,j-1})$$

Linear indexing: $(i,j) \mapsto (k_2 i + j)$

Concatenation of SBPs

Concatenation of two SBPs:

- φ_1: $\mathsf{W} \mapsto \{\mathsf{B}_1^{(0)}, \ldots, \mathsf{B}_1^{(k_1-1)}\}$
- φ_2: $\mathsf{B}^{k_2} \mapsto \{\mathsf{B}_2^{(0)}, \ldots, \mathsf{B}_2^{(k_2-1)}\}$ (Vector channel of k_2 B-DMCs B as input)

Concatenation is again an SBP (of order $k_1 k_2$):

$$\varphi_1 \circ \varphi_2 : \mathsf{W}^{k_2} \mapsto \{\mathsf{B}^{(0,0)}, \ldots, \mathsf{B}^{(0,k_2-1)}, \ldots, \mathsf{B}^{(k_1-1,0)}, \ldots, \mathsf{B}^{(k_1-1,k_2-1)}\}$$

$$\mathsf{B}^{(i,j)} : \{0,1\} \mapsto \mathcal{Y}^{k_2} \times \{0,1\}^{k_2 i+j}$$ B-DMC, binary channels

$$I(\mathsf{B}^{(i,j)}) := I(B_{i,j}; \boldsymbol{Y} | \boldsymbol{B}_0 \ldots \boldsymbol{B}_{i-1}, B_{i,0}, \ldots, B_{i,j-1})$$

Linear indexing: $(i,j) \mapsto (k_2 i + j)$

- **average** capacity per binary channel does not change:

$$M_{\varphi_1 \circ \varphi_2}(\mathsf{W}^{k_2}) = M_{\varphi_1}(\mathsf{W}) = \frac{1}{k_1} I(X;Y)$$

Variance of concatenated SBPs

Theorem 1: variance increases:

$$\boxed{V_{\varphi_1 \circ \varphi_2}(\mathsf{W}^{k_2}) = V_{\varphi_1}(\mathsf{W}) + \frac{1}{k_1} \sum_{i=0}^{k_1-1} V_{\varphi_2}(\mathsf{B}_1^{(i)})}$$

Theorem 1: variance increases:

$$\boxed{V_{\varphi_1 \circ \varphi_2}(\mathsf{W}^{k_2}) = V_{\varphi_1}(\mathsf{W}) + \frac{1}{k_1} \sum_{i=0}^{k_1-1} V_{\varphi_2}(\mathsf{B}_1^{(i)})}$$

Proof: By definition:

$$V_{\varphi_1 \circ \varphi_2}(\mathsf{W}^{k_2}) = \frac{1}{k_1 k_2} \sum_{i=0}^{k_1-1} \sum_{j=0}^{k_2-1} I(\mathsf{B}^{(i,j)})^2 - M_{\varphi_1 \circ \varphi_2}(\mathsf{W}^{k_2})^2$$

Variance of concatenated SBPs

Theorem 1: variance increases:

$$\boxed{V_{\varphi_1 \circ \varphi_2}(\mathsf{W}^{k_2}) = V_{\varphi_1}(\mathsf{W}) + \frac{1}{k_1} \sum_{i=0}^{k_1-1} V_{\varphi_2}(\mathsf{B}_1^{(i)})}$$

Proof: By definition:

$$V_{\varphi_1 \circ \varphi_2}(\mathsf{W}^{k_2}) = \frac{1}{k_1 k_2} \sum_{i=0}^{k_1-1} \sum_{j=0}^{k_2-1} I(\mathsf{B}^{(i,j)})^2 - M_{\varphi_1 \circ \varphi_2}(\mathsf{W}^{k_2})^2$$

Adding and subtracting the term $\frac{1}{k_1} \sum_{i=0}^{k_1-1} I(\mathsf{B}_1^{(i)})^2$:

$$V_{\varphi_1 \circ \varphi_2}(\mathsf{W}^{k_2}) = \frac{1}{k_1} \sum_{i=0}^{k_1-1} I(\mathsf{B}_1^{(i)})^2 - M_{\varphi_1 \circ \varphi_2}(\mathsf{W}^{k_2})^2$$

$$+ \frac{1}{k_1 k_2} \sum_{i=0}^{k_1-1} \sum_{j=0}^{k_2-1} I(\mathsf{B}^{(i,j)})^2 - \frac{1}{k_1} \sum_{i=0}^{k_1-1} I(\mathsf{B}_1^{(i)})^2$$

Variance of concatenated SBPs

Theorem 1: variance increases:

$$\boxed{V_{\varphi_1 \circ \varphi_2}(\mathsf{W}^{k_2}) = V_{\varphi_1}(\mathsf{W}) + \frac{1}{k_1} \sum_{i=0}^{k_1-1} V_{\varphi_2}(\mathsf{B}_1^{(i)})}$$

Proof: By definition:

$$V_{\varphi_1 \circ \varphi_2}(\mathsf{W}^{k_2}) = \frac{1}{k_1 k_2} \sum_{i=0}^{k_1-1} \sum_{j=0}^{k_2-1} I(\mathsf{B}^{(i,j)})^2 - M_{\varphi_1 \circ \varphi_2}(\mathsf{W}^{k_2})^2$$

Adding and subtracting the term $\frac{1}{k_1} \sum_{i=0}^{k_1-1} I(\mathsf{B}_1^{(i)})^2$:

$$V_{\varphi_1 \circ \varphi_2}(\mathsf{W}^{k_2}) = \frac{1}{k_1} \sum_{i=0}^{k_1-1} I(\mathsf{B}_1^{(i)})^2 - M_{\varphi_1}(\mathsf{W})^2$$

$$+ \frac{1}{k_1 k_2} \sum_{i=0}^{k_1-1} \sum_{j=0}^{k_2-1} I(\mathsf{B}^{(i,j)})^2 - \frac{1}{k_1} \sum_{i=0}^{k_1-1} I(\mathsf{B}_1^{(i)})^2$$

Variance of concatenated SBPs

Theorem 1: variance increases:

$$\boxed{V_{\varphi_1 \circ \varphi_2}(\mathsf{W}^{k_2}) = V_{\varphi_1}(\mathsf{W}) + \frac{1}{k_1} \sum_{i=0}^{k_1-1} V_{\varphi_2}(\mathsf{B}_1^{(i)})}$$

Proof: By definition:

$$V_{\varphi_1 \circ \varphi_2}(\mathsf{W}^{k_2}) = \frac{1}{k_1 k_2} \sum_{i=0}^{k_1-1} \sum_{j=0}^{k_2-1} I(\mathsf{B}^{(i,j)})^2 - M_{\varphi_1 \circ \varphi_2}(\mathsf{W}^{k_2})^2$$

Adding and subtracting the term $\frac{1}{k_1} \sum_{i=0}^{k_1-1} I(\mathsf{B}_1^{(i)})^2$:

$$V_{\varphi_1 \circ \varphi_2}(\mathsf{W}^{k_2}) = \frac{1}{k_1} \sum_{i=0}^{k_1-1} I(\mathsf{B}_1^{(i)})^2 - M_{\varphi_1}(\mathsf{W})^2 + \frac{1}{k_1 k_2} \sum_{i=0}^{k_1-1} \sum_{j=0}^{k_2-1} I(\mathsf{B}^{(i,j)})^2 - \frac{1}{k_1} \sum_{i=0}^{k_1-1} I(\mathsf{B}_1^{(i)})^2$$

Variance of concatenated SBPs

Theorem 1: variance increases:

$$\boxed{V_{\varphi_1\circ\varphi_2}(\mathsf{W}^{k_2}) = V_{\varphi_1}(\mathsf{W}) + \frac{1}{k_1}\sum_{i=0}^{k_1-1} V_{\varphi_2}(\mathsf{B}_1^{(i)})}$$

Proof: By definition:

$$V_{\varphi_1\circ\varphi_2}(\mathsf{W}^{k_2}) = \frac{1}{k_1 k_2}\sum_{i=0}^{k_1-1}\sum_{j=0}^{k_2-1} I(\mathsf{B}^{(i,j)})^2 - M_{\varphi_1\circ\varphi_2}(\mathsf{W}^{k_2})^2$$

Adding and subtracting the term $\frac{1}{k_1}\sum_{i=0}^{k_1-1} I(\mathsf{B}_1^{(i)})^2$:

$$V_{\varphi_1\circ\varphi_2}(\mathsf{W}^{k_2}) = \frac{1}{k_1}\sum_{i=0}^{k_1-1} I(\mathsf{B}_1^{(i)})^2 - M_{\varphi_1}(\mathsf{W})^2$$

$$+ \frac{1}{k_1}\sum_{i=0}^{k_1-1}\left(\frac{1}{k_2}\sum_{j=0}^{k_2-1} I(\mathsf{B}^{(i,j)})^2 - I(\mathsf{B}_1^{(i)})^2\right)$$

Variance of concatenated SBPs

Theorem 1: variance increases:

$$\boxed{V_{\varphi_1 \circ \varphi_2}(\mathsf{W}^{k_2}) = V_{\varphi_1}(\mathsf{W}) + \frac{1}{k_1} \sum_{i=0}^{k_1-1} V_{\varphi_2}(\mathsf{B}_1^{(i)})}$$

Proof: By definition:

$$V_{\varphi_1 \circ \varphi_2}(\mathsf{W}^{k_2}) = \frac{1}{k_1 k_2} \sum_{i=0}^{k_1-1} \sum_{j=0}^{k_2-1} I(\mathsf{B}^{(i,j)})^2 - M_{\varphi_1 \circ \varphi_2}(\mathsf{W}^{k_2})^2$$

Adding and subtracting the term $\frac{1}{k_1} \sum_{i=0}^{k_1-1} I(\mathsf{B}_1^{(i)})^2$:

$$V_{\varphi_1 \circ \varphi_2}(\mathsf{W}^{k_2}) = \frac{1}{k_1} \sum_{i=0}^{k_1-1} I(\mathsf{B}_1^{(i)})^2 - M_{\varphi_1}(\mathsf{W})^2$$

$$+ \frac{1}{k_1} \sum_{i=0}^{k_1-1} \left(\frac{1}{k_2} \sum_{j=0}^{k_2-1} I(\mathsf{B}^{(i,j)})^2 - M_{\varphi_2}(\mathsf{B}_1^{(i)})^2 \right)$$

Variance of concatenated SBPs

Theorem 1: variance increases:

$$\boxed{V_{\varphi_1\circ\varphi_2}(\mathsf{W}^{k_2}) = V_{\varphi_1}(\mathsf{W}) + \frac{1}{k_1}\sum_{i=0}^{k_1-1} V_{\varphi_2}(\mathsf{B}_1^{(i)})}$$

Proof: By definition:

$$V_{\varphi_1\circ\varphi_2}(\mathsf{W}^{k_2}) = \frac{1}{k_1 k_2}\sum_{i=0}^{k_1-1}\sum_{j=0}^{k_2-1} I(\mathsf{B}^{(i,j)})^2 - M_{\varphi_1\circ\varphi_2}(\mathsf{W}^{k_2})^2$$

Adding and subtracting the term $\frac{1}{k_1}\sum_{i=0}^{k_1-1} I(\mathsf{B}_1^{(i)})^2$:

$$\begin{aligned} V_{\varphi_1\circ\varphi_2}(\mathsf{W}^{k_2}) &= \frac{1}{k_1}\sum_{i=0}^{k_1-1} I(\mathsf{B}_1^{(i)})^2 - M_{\varphi_1}(\mathsf{W})^2 \\ &+ \frac{1}{k_1}\sum_{i=0}^{k_1-1}\left(\frac{1}{k_2}\sum_{j=0}^{k_2-1} I(\mathsf{B}^{(i,j)})^2 - M_{\varphi_2}(\mathsf{B}_1^{(i)})^2\right) \end{aligned}$$

Variance of concatenated SBPs

Theorem 1: variance increases:

$$\boxed{V_{\varphi_1 \circ \varphi_2}(\mathsf{W}^{k_2}) = V_{\varphi_1}(\mathsf{W}) + \frac{1}{k_1} \sum_{i=0}^{k_1-1} V_{\varphi_2}(\mathsf{B}_1^{(i)})}$$

Proof: By definition:

$$V_{\varphi_1 \circ \varphi_2}(\mathsf{W}^{k_2}) = \frac{1}{k_1 k_2} \sum_{i=0}^{k_1-1} \sum_{j=0}^{k_2-1} I(\mathsf{B}^{(i,j)})^2 - M_{\varphi_1 \circ \varphi_2}(\mathsf{W}^{k_2})^2$$

Adding and subtracting the term $\frac{1}{k_1} \sum_{i=0}^{k_1-1} I(\mathsf{B}_1^{(i)})^2$:

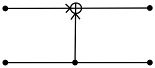
$$V_{\varphi_1 \circ \varphi_2}(\mathsf{W}^{k_2}) = \frac{1}{k_1} \sum_{i=0}^{k_1-1} I(\mathsf{B}_1^{(i)})^2 - M_{\varphi_1}(\mathsf{W})^2 + \frac{1}{k_1} \sum_{i=0}^{k_1-1} \left(\frac{1}{k_2} \sum_{j=0}^{k_2-1} I(\mathsf{B}^{(i,j)})^2 - M_{\varphi_2}(\mathsf{B}_1^{(i)})^2 \right)$$

q.e.d.

SBPs: Examples

Examples of SBPs:

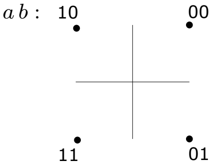
- M-ary transmission ($M = 2^m$) with binary labeling: m-SBP φ
- Arıkan's polarizing construction: 2-SBP π :



Polar Code as SBP

- Subsequent binary bipolar PAM signalling using orthogonal pulses, i.e.,

$$\int_{-\infty}^{\infty} g(t) \cdot g^{*}(t - kT)\,\mathrm{d}t = \delta_{0,k} \qquad \text{(1. Nyquist condition)}$$



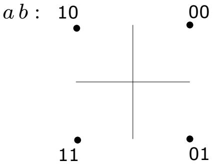
individual decisions

„Gray“ Mapping

Polar Code as SBP

- Subsequent binary bipolar PAM signalling using orthogonal pulses, i.e.,

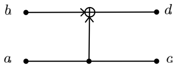
$$\int_{-\infty}^{\infty} g(t) \cdot g^*(t - kT)\, \mathrm{d}t = \delta_{0,k} \qquad \text{(1. Nyquist condition)}$$

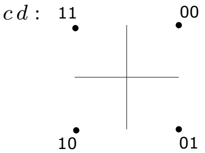


individual decisions

„Gray“ Mapping

- Application of 2-SBP π :



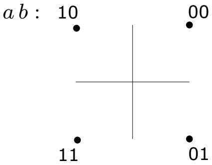


Mapping by Set Partitioning

Polar Code as SBP

- Subsequent binary bipolar PAM signalling using orthogonal pulses, i.e.,

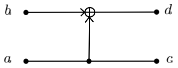
$$\int_{-\infty}^{\infty} g(t) \cdot g^*(t - kT)\, \mathrm{d}t = \delta_{0,k} \qquad \text{(1. Nyquist condition)}$$

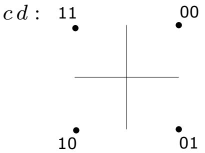


individual decisions

„Gray“ Mapping

- Application of 2-SBP π :



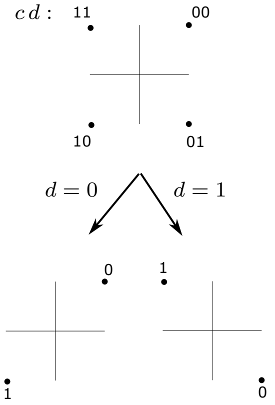


Mapping by Set Partitioning

$\Rightarrow$ 2-SBP π: Transform of mapping from „Gray“ to „Set Partitioning“

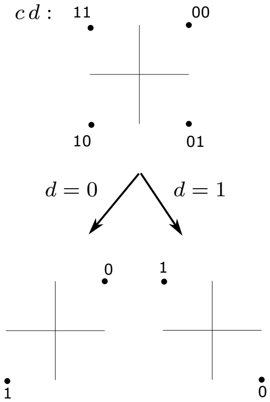
Polar Code as Coded Modulation

- Set Partitioning of

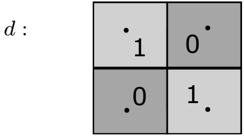


Polar Code as Coded Modulation

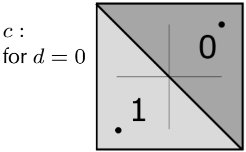
- Set Partitioning of



- Decision Regions:



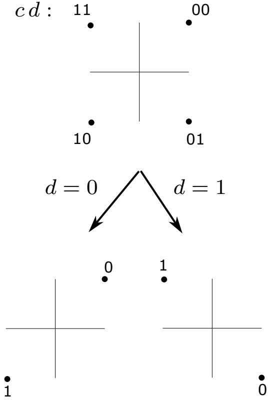
loss to BPSK due to
2 nearest neighbours



3 dB gain to BPSK

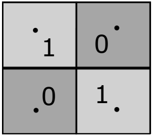
Polar Code as Coded Modulation

- Set Partitioning of



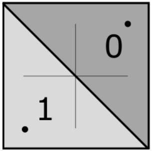
- Decision Regions:

d :



loss to BPSK due to
2 nearest neighbours

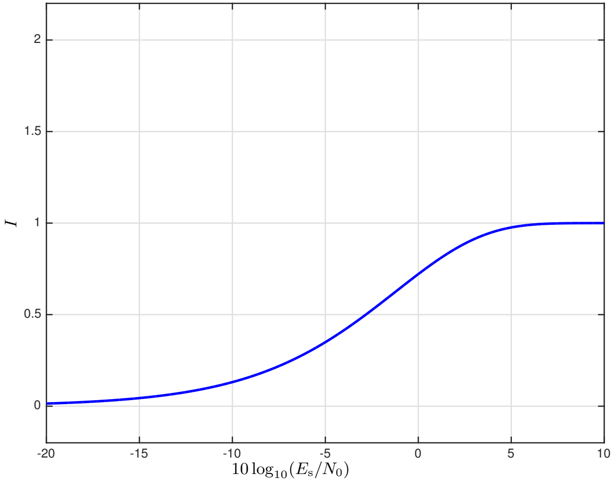
c :
for $d = 0$



3 dB gain to BPSK

- length-N Polar Code: A special „Anti-Gray“ mapping for the vertices of a hypercube in $\mathbb{R}^N$, a method of **Coded Modulation**

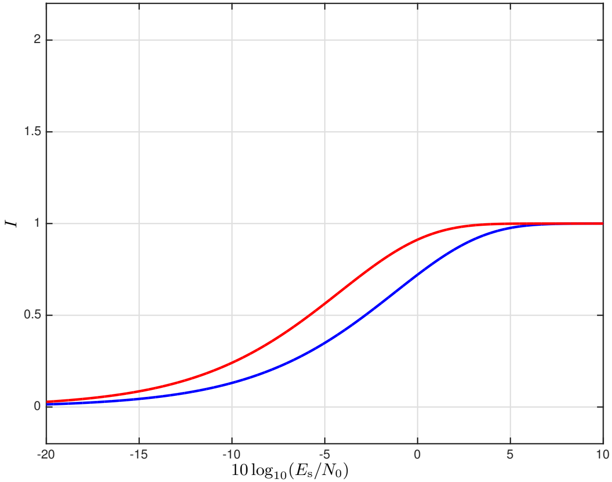
Polar Code as Coded Modulation



Legend:

- — BPSK

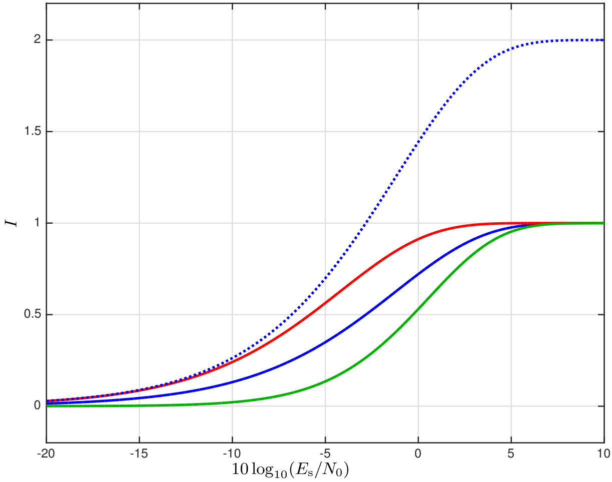
Polar Code as Coded Modulation



Legend:

- BPSK
- higher level c

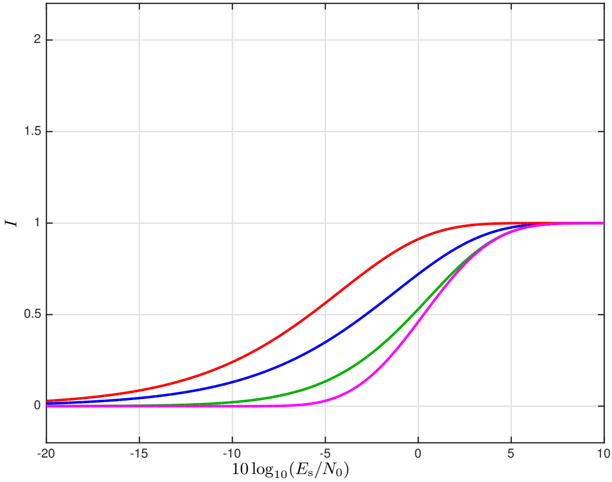
Polar Code as Coded Modulation



Legend:

- BPSK
- higher level c
- lower level d
- BPSK, two uses

Polar Code as Coded Modulation

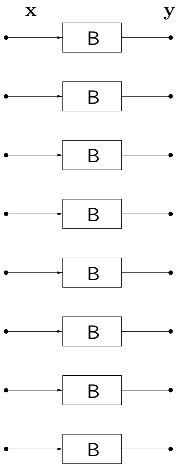


Legend:

- — BPSK
- — higher level c
- — lower level d

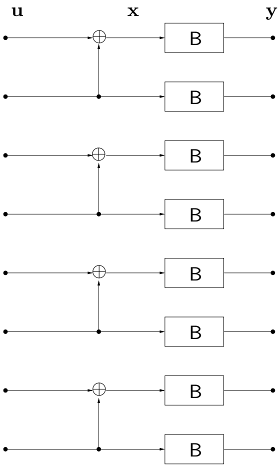
- — BPSK-mod channel (lowest level in $M \gg 4$-ary ASK)

Polar Codes as SBPs



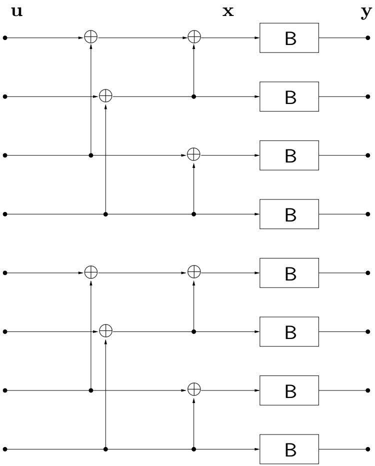
- Polar codes as a particular instance of concatenated linear SBPs

Polar Codes as SBPs



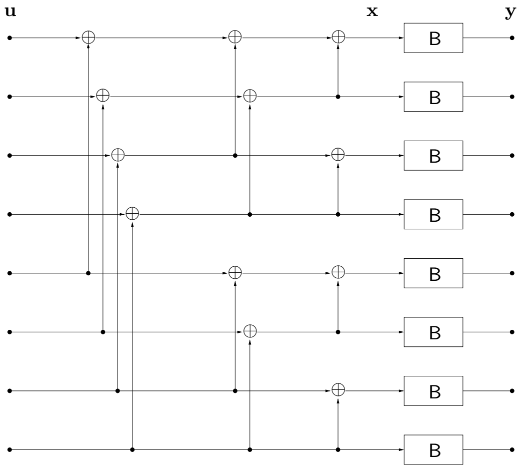
- Polar codes as a particular instance of concatenated linear SBPs

Polar Codes as SBPs



- Polar codes as a particular instance of concatenated linear SBPs

Polar Codes as SBPs



- Polar codes as a particular instance of concatenated linear SBPs

Concatenation of π

- n-fold concatenation of π leads to N-SBP π^n $(N = 2^n)$

$$\pi^n : \mathsf{B}^N \mapsto \{\mathsf{B}_N^{(0)}, \ldots, \mathsf{B}_N^{(N-1)}\}$$

with binary channels

$$\mathsf{B}_N^{(i)} : \{0,1\} \to \mathcal{Y}^N \times \{0,1\}^i \quad , \quad i = 0, \ldots, N-1$$

$$I(\mathsf{B}_N^{(i)}) := I(B_i; Y_0, \ldots, Y_{N-1} | B_0, \ldots B_{i-1})$$

The chain rule of mutual information assures

$$\sum_{i=0}^{N-1} I(\mathsf{B}_N^{(i)}) = N \cdot I(\mathsf{B}) = N \cdot M_{\pi^n}(\mathsf{B})$$

capacity-preserving change of mapping

Polarization and Variance

Variance:

$$V_{\pi^n}(\mathsf{B}) = \frac{1}{N}\sum_{i=0}^{N-1} I(\mathsf{B}_N^{(i)})^2 - I(\mathsf{B})^2$$

Polarization and Variance

Variance:

$$V_{\pi^n}(\mathsf{B}) = \frac{1}{N} \sum_{i=0}^{N-1} I(\mathsf{B}_N^{(i)})^2 - I(\mathsf{B})^2$$

- maximum possible variance corresponds to state of perfect polarization:

$$V_{\pi^n}(\mathsf{B}) \leq I(\mathsf{B}) \cdot (1 - I(\mathsf{B}))$$

with equality iff $I(\mathsf{B}_N^{(i)}) \in \{0, 1\} \quad \forall i$

Polarization and Variance

Variance:

$$V_{\pi^n}(\mathsf{B}) = \frac{1}{N}\sum_{i=0}^{N-1} I(\mathsf{B}_N^{(i)})^2 - I(\mathsf{B})^2$$

- maximum possible variance corresponds to state of perfect polarization:

$$V_{\pi^n}(\mathsf{B}) \leq I(\mathsf{B}) \cdot (1 - I(\mathsf{B}))$$

with equality iff $I(\mathsf{B}_N^{(i)}) \in \{0, 1\} \quad \forall i$

Proof:

$$V_{\pi^n}(\mathsf{B}) = \frac{1}{N}\sum_{i=0}^{N-1} I(\mathsf{B}_N^{(i)})^2 - I(\mathsf{B})^2$$

$$\leq \frac{1}{N}\sum_{i=0}^{N-1} I(\mathsf{B}_N^{(i)})^1 - I(\mathsf{B})^2 = I(\mathsf{B}) \cdot (1 - I(\mathsf{B}))$$

Polarization and Variance

Variance:

$$V_{\pi^n}(\mathsf{B}) = \frac{1}{N}\sum_{i=0}^{N-1} I(\mathsf{B}_N^{(i)})^2 - I(\mathsf{B})^2$$

- maximum possible variance corresponds to state of perfect polarization:

  $$V_{\pi^n}(\mathsf{B}) \leq I(\mathsf{B}) \cdot (1 - I(\mathsf{B}))$$

  with equality iff $I(\mathsf{B}_N^{(i)}) \in \{0, 1\} \quad \forall i$

- $V_{\pi^n}(\mathsf{B})$ increases with each step of polarization (Theorem 1):

  $$V_{\pi^{n+1}}(\mathsf{B}) = V_{\pi^n}(\mathsf{B}) + \frac{1}{N}\sum_{i=0}^{N-1} V_{\pi}(\mathsf{B}_N^{(i)})$$

Polarization and Variance

Variance:

$$V_{\pi^n}(\mathsf{B}) = \frac{1}{N}\sum_{i=0}^{N-1} I(\mathsf{B}_N^{(i)})^2 - I(\mathsf{B})^2$$

- maximum possible variance corresponds to state of perfect polarization:

  $$V_{\pi^n}(\mathsf{B}) \le I(\mathsf{B}) \cdot (1 - I(\mathsf{B}))$$

  with equality iff $I(\mathsf{B}_N^{(i)}) \in \{0, 1\} \quad \forall i$

- $V_{\pi^n}(\mathsf{B})$ increases with each step of polarization (Theorem 1):

  $$V_{\pi^{n+1}}(\mathsf{B}) = V_{\pi^n}(\mathsf{B}) + \frac{1}{N}\sum_{i=0}^{N-1} V_\pi(\mathsf{B}_N^{(i)})$$

- Byproduct: a simple proof of capacity-achieving property of Polar Codes

Polarization and Variance

Variance:

$$V_{\pi^n}(\mathsf{B}) = \frac{1}{N}\sum_{i=0}^{N-1} I(\mathsf{B}_N^{(i)})^2 - I(\mathsf{B})^2$$

- maximum possible variance corresponds to state of perfect polarization:

  $$V_{\pi^n}(\mathsf{B}) \leq I(\mathsf{B}) \cdot (1 - I(\mathsf{B}))$$

  with equality iff $I(\mathsf{B}_N^{(i)}) \in \{0, 1\} \quad \forall i$

- $V_{\pi^n}(\mathsf{B})$ increases with each step of polarization (Theorem 1):

  $$V_{\pi^{n+1}}(\mathsf{B}) = V_{\pi^n}(\mathsf{B}) + \frac{1}{N}\sum_{i=0}^{N-1} V_\pi(\mathsf{B}_N^{(i)})$$

$\Rightarrow$ The sequence $V_{\pi^n}(\mathsf{B})$ converges for $n \to \infty$, which implies

$$\lim_{n\to\infty} \left(V_{\pi^{n+1}}(\mathsf{B}) - V_{\pi^n}(\mathsf{B})\right) = \lim_{n\to\infty} \frac{1}{2^n}\sum_{i=0}^{2^n-1} V_\pi(\mathsf{B}_N^{(i)}) = 0$$

Polarization and Variance

$$\lim_{n\to\infty} \frac{1}{2^n} \sum_{i=0}^{2^n-1} V_\pi(\mathsf{B}_N^{(i)}) = 0$$

Polarization and Variance

$$\lim_{n\to\infty} \frac{1}{2^n} \sum_{i=0}^{2^n-1} V_\pi(\mathsf{B}_N^{(i)}) = 0$$

- $V_\pi(\mathsf{B}_N^{(i)}) = 0$ only possible if $I(\mathsf{B}_N^{(i)})$ is either 0 or 1 , because serial and parallel information combining for two identical channels has to result in two identical channels which only is possible for input MI either $I = 0$ or $I = 1$ (bounds on information combining!)

Polarization and Variance

$$\lim_{n\to\infty} \frac{1}{2^n} \sum_{i=0}^{2^n-1} V_\pi(\mathsf{B}_N^{(i)}) = 0$$

- $V_\pi(\mathsf{B}_N^{(i)}) = 0$ only possible if $I(\mathsf{B}_N^{(i)})$ is either 0 or 1

$\Rightarrow$ for $n \to \infty$ and $\epsilon > 0$, the fraction of binary channels $\mathsf{B}_N^{(i)}$ with capacity $I(\mathsf{B}_N^{(i)}) \in [\epsilon, 1-\epsilon]$ has to vanish

Polarization and Variance

$$\lim_{n\to\infty} \frac{1}{2^n} \sum_{i=0}^{2^n-1} V_\pi(\mathsf{B}_N^{(i)}) = 0$$

- $V_\pi(\mathsf{B}_N^{(i)}) = 0$ only possible if $I(\mathsf{B}_N^{(i)})$ is either 0 or 1

$\Rightarrow$ for $n \to \infty$ and $\epsilon > 0$, the fraction of binary channels $\mathsf{B}_N^{(i)}$ with capacity $I(\mathsf{B}_N^{(i)}) \in [\epsilon, 1-\epsilon]$ has to vanish

$\Rightarrow$ convergence to state of perfect polarization:

$$\lim_{n\to\infty} V_{\pi^n}(\mathsf{B}) = I(\mathsf{B}) \cdot (1 - I(\mathsf{B}))$$

and all transforms capacity-preserving

Polarization and Variance

$$\lim_{n\to\infty} \frac{1}{2^n} \sum_{i=0}^{2^n-1} V_\pi(\mathsf{B}_N^{(i)}) = 0$$

- $V_\pi(\mathsf{B}_N^{(i)}) = 0$ only possible if $I(\mathsf{B}_N^{(i)})$ is either 0 or 1

$\Rightarrow$ for $n \to \infty$ and $\epsilon > 0$, the fraction of binary channels $\mathsf{B}_N^{(i)}$ with capacity $I(\mathsf{B}_N^{(i)}) \in [\epsilon, 1-\epsilon]$ has to vanish

$\Rightarrow$ convergence to state of perfect polarization:

$$\lim_{n\to\infty} V_{\pi^n}(\mathsf{B}) = I(\mathsf{B}) \cdot (1 - I(\mathsf{B}))$$

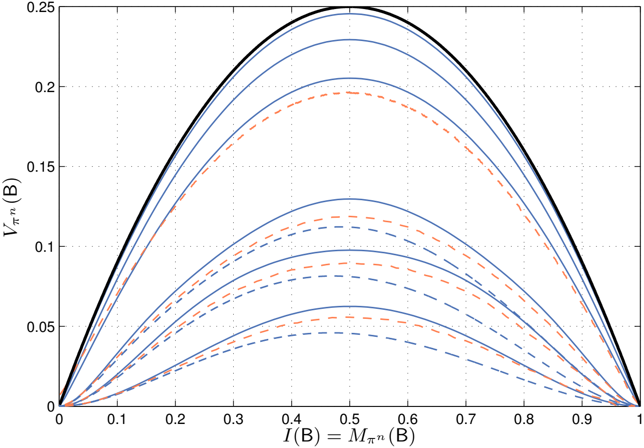
and all transforms capacity-preserving

$\Rightarrow$ capacity-achieving property of polar codes for symmetric binary channels

Variance for Binary Polar Codes

Polar Codes: variance of binary channel capacities

$$N = 2^n,\ n = 1, 2, 4, 8, 12, 20$$



Legend:

- — BEC
- - - BSC
- - - BI-AWGN

Multilevel Polar Coding

Polarization: MLC

Variance for MLC with Polar Codes:

$$V_{\varphi \circ \pi^n}(\mathsf{W}) = V_{\varphi}(\mathsf{W}) + \frac{1}{m} \sum_{i=0}^{m-1} V_{\pi^n}(\mathsf{B}^{(i)})$$

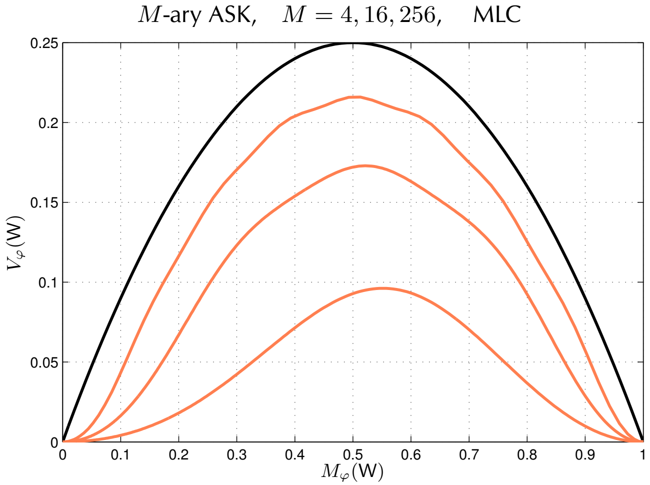
- Variance $V_{\varphi}(\mathsf{W})$ of binary capacities $I(\mathsf{B}^{(i)})$ influences overall polarization

$\Rightarrow$ choose labeling that maximizes $V_{\varphi \circ \pi^n}(\mathsf{W})$

- **Theorem 2:** MLC with Polar Codes as component codes results again in a Polar Code

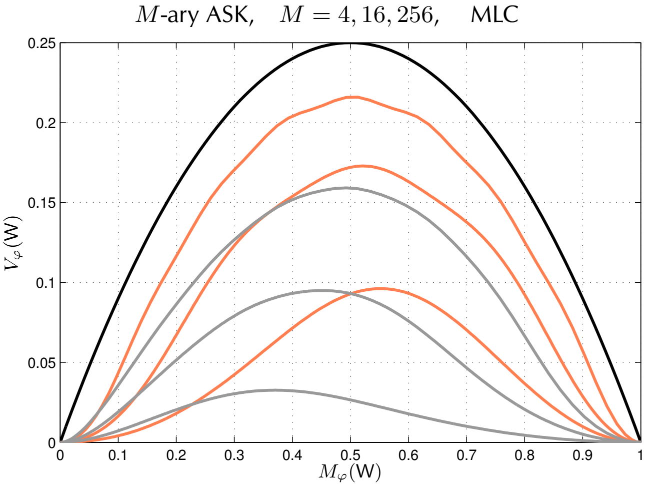
  Proof: concatenation of SBPs: $\varphi \circ \pi^n$

Example: Variance ASK binary capacities



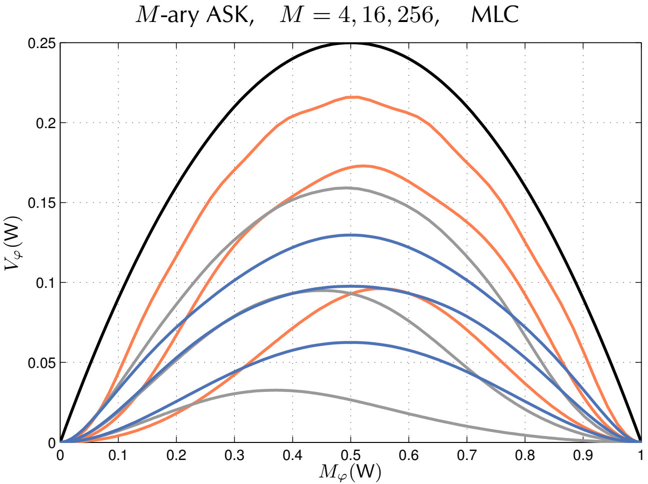
— Ungerboeck labeling,

Example: Variance ASK binary capacities



— Ungerboeck labeling, — Gray labeling

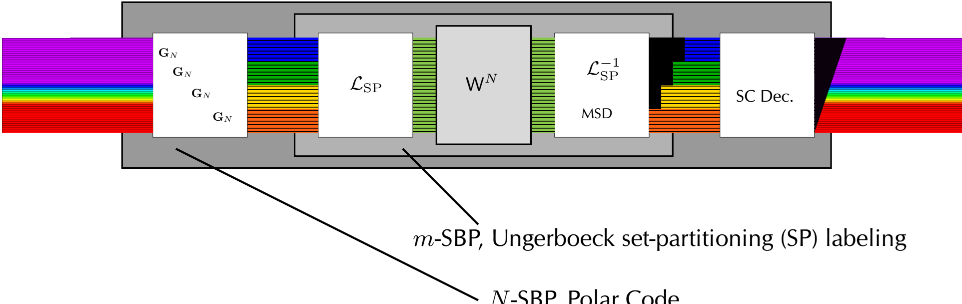
Example: Variance ASK binary capacities



— Ungerboeck labeling, — Gray labeling — Polar Code BEC ($N = 2, 4, 8$)

Multilevel Polar Code

Multilevel Polar Code, length mN



m-SBP, Ungerboeck set-partitioning (SP) labeling

N-SBP, Polar Code

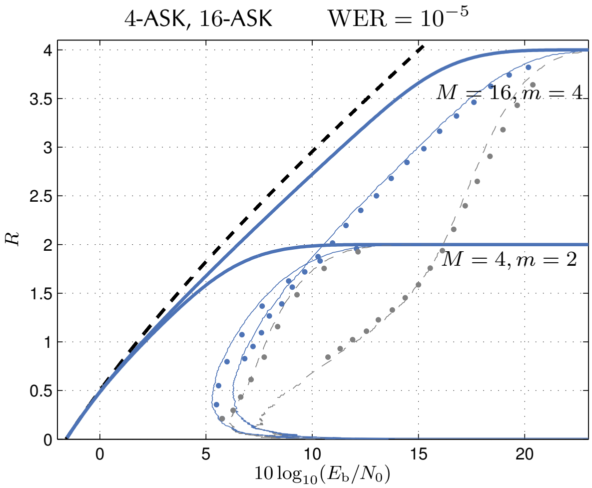
Performance - Remarks

- Only SC (Successive Cancellation) Decoding considered

Performance Analysis:

- Density Evolution (DE) with Gaussian Approximation (solid lines —)
- Full Simulation

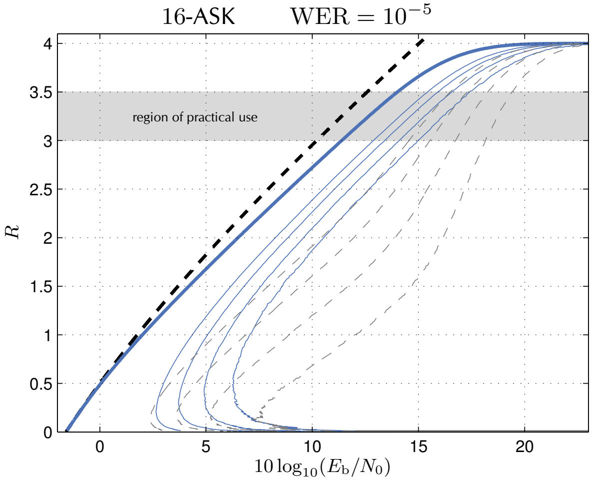
Accuracy of Gaussian DE (ASK)



Legend:

- Overall Blocklength: $mN = 512$
- – – Shannon bound (real constellations)
- — Const. constrained capacity
- — DE MLC-MSD, SP
- – – DE MLC-MSD, Gray
- $*$ $*$ Simulation

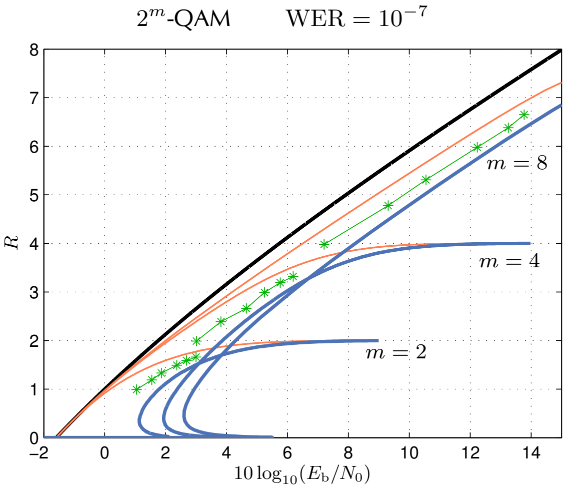
Performance: MLC - ASK



Legend:

- Overall Blocklength: $mN = 2^9, 2^{11}, 2^{13}, 2^{15}$
- – – Shannon bound (real constellations)
- — Const. constrained capacity
- — DE MLC-MSD, SP
- – – DE MLC-MSD, Gray

Comparison: Polar Codes - LDPC



Legend:

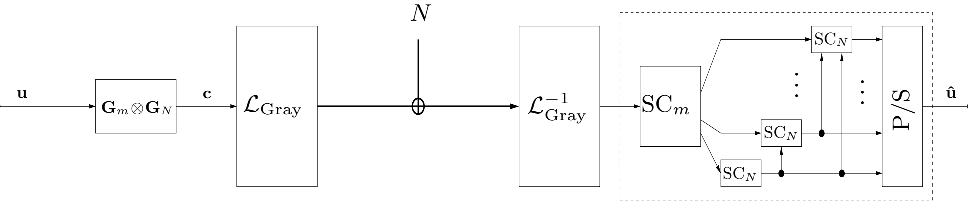
- Overall Blocklength:
  65.536 (Polar Code)
  64.800 (LDPC + BCH)
- — Shannon bound
- — Const. constrained capacity
- — DE MLC-MSD, SP
  4-QAM, 16-QAM, 256-QAM
- $*$ DVB-T2 LDPC + BCH
  4-QAM, 16-QAM, 256-QAM

Bit-Interleaved Polar-Coded Modulation

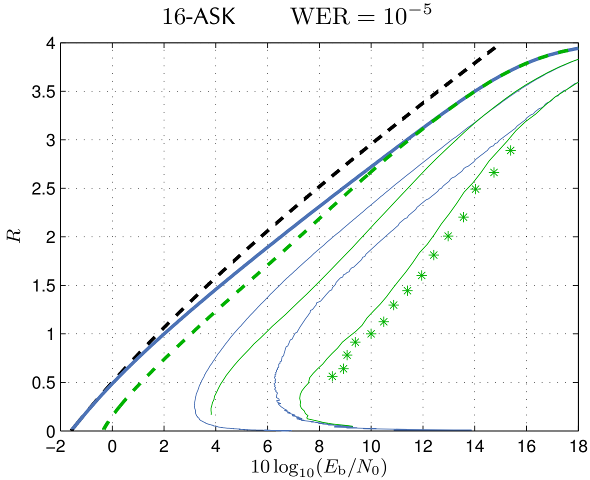
System Model: BICM

BICM with Parallel Decoding over M-ary constellation

- Gray labeling
- no interleaver considered here
- polar code of length mN ($m = \log_2(M)$)



Performance: BICM



Legend:

- Overall Blocklength:
  $mN = 2^9, 2^{14}$
  $N = 2^7, 2^{12}$
- – – Shannon bound (real constellations)
- — Const. constrained capacity
- – – BICM capacity
- — DE MLC-MSD, SP
- — DE BICM, Gray
- $*$ Simulation

Labeling Transforms

Is there a way to transform Gray into SP labeling?

Labeling Transforms

Is there a way to transform Gray into SP labeling?

- M-ASK / PSK constellations [Alvarado et al., 2012]:

$$\mathbf{M}_{\mathrm{SP}} \cdot \mathbf{T}_m = \mathbf{M}_{\mathrm{Gray}}$$

Labeling Transforms

Is there a way to transform Gray into SP labeling?

- M-ASK / PSK constellations [Alvarado et al., 2012]:

$$\mathbf{M}_{\mathrm{SP}} \cdot \mathbf{T}_m = \mathbf{M}_{\mathrm{Gray}}$$

$$\mathbf{M}_{\mathrm{SP}} = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & 1 \end{bmatrix} \quad , \quad \mathbf{M}_{\mathrm{Gray}} = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 \\ 1 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

$$\mathbf{T}_m = \begin{bmatrix} \color{blue}{1} & 0 & 0 & \dots & 0 & 0 \\ \color{blue}{1} & \color{blue}{1} & 0 & \dots & 0 & 0 \\ 0 & \color{blue}{1} & \color{blue}{1} & \dots & 0 & 0 \\ & \vdots & & \ddots & \vdots & \\ 0 & 0 & 0 & \dots & \color{blue}{1} & \color{blue}{1} \end{bmatrix}$$

Labeling Transforms

Is there a way to transform Gray into SP labeling?

- M-ASK / PSK constellations [Alvarado et al., 2012]:

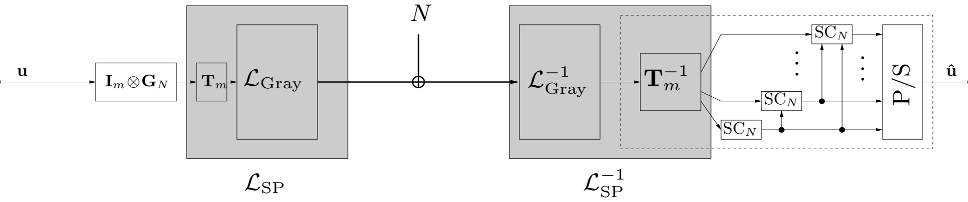
$$\mathbf{M}_{\mathrm{SP}} \cdot \mathbf{T}_m = \mathbf{M}_{\mathrm{Gray}}$$

- quadratic M^2-QAM: $\mathbf{M}_{\mathrm{SP}} \cdot (\mathbf{G}_2 \otimes \mathbf{T}_m) = \mathbf{M}_{\mathrm{Gray}}$

System Model: BICM

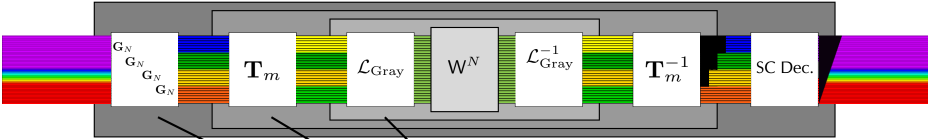
BICM with Parallel Decoding over M-ary constellation, modified polar code

- identical encoder like in the MLC approach
- quasi-identical decoder
- performance loss due to Parallel Decoding



BICM Polar Code

BICM Polar Code, length mN

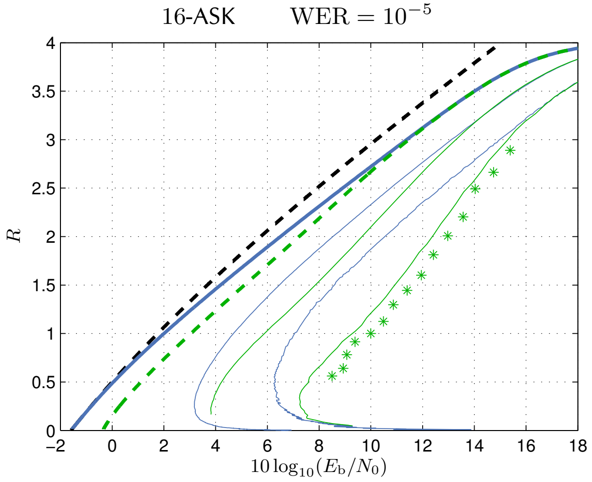


m-SBP, Labeling

Transformation SP $\mapsto$ Gray

N-SBP, Polar Code

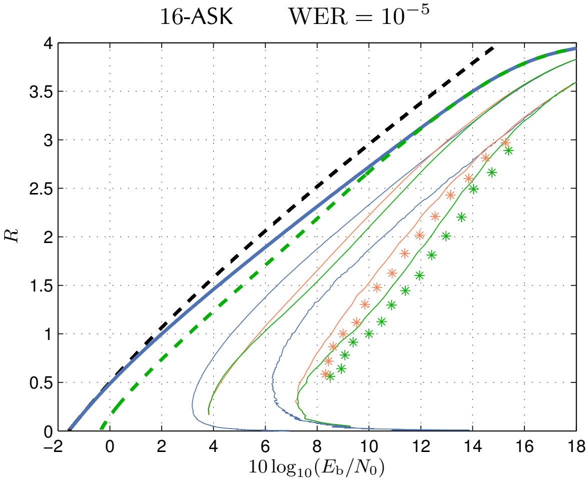
Comparison: MLC - BICM



Legend:

- Overall Blocklength: $mN = 2^9, 2^{14}$
- – – Shannon bound (real constellations)
- — Const. constrained capacity
- – – BICM capacity
- — DE MLC-MSD, SP
- — DE BICM, Gray
- — DE BICM mod. Polar Code
- *, * Simulation

Comparison: MLC - BICM



Legend:

- Overall Blocklength: $mN = 2^9, 2^{14}$
- – – Shannon bound (real constellations)
- — Const. constrained capacity
- – – BICM capacity
- — DE MLC-MSD, SP
- — DE BICM, Gray
- — DE BICM mod. Polar Code
- *, * Simulation

Conclusions

- Unified description of both Polar Coding and Multilevel Coding
- Polar Coding is a sort of Coded Modulation: Transform of natural „Gray“ mapping for the vertices of an N-dimensional hypercube for orthogonal channel uses into an „Anti-Gray“ mapping
- MLC/Polar may be seen as one single succ. decoded Polar Code
- MLC/Polar: Ungerboeck labeling should be used
- BICM/Polar (no interleaver!) behaves like a degraded version of MLC/Polar
- Easy to analyze via DE (with Gaussian approximation)

Thank you for your attention!