Polar-Coded Non-Coherent Communication[1]

Mustafa Cemil Coşkun
mustafa.coskun@tum.de

Institute for Communications Engineering (LNT)
Technical University of Munich (TUM)

Short Packet Transmission for Wireless Communications
Paris, France – November 24, 2021

[1]supported by the German Research Foundation (DFG) under Grant KR 3517/9-1, and by the Munich Aerospace under the grant "Efficient Coding and Modulation for Satellite Links with Severe Delay Constraints"

Outline

Outline

Polar Codes

IEEE TRANSACTIONS ON INFORMATION THEORY, VOL. 55, NO. 7, JULY 2009 3051

Channel Polarization: A Method for Constructing Capacity-Achieving Codes for Symmetric Binary-Input Memoryless Channels

Erdal Arıkan, *Senior Member, IEEE*

***Abstract*—A method is proposed, called channel polarization, to construct code sequences that achieve the symmetric capacity $I(W)$ of any given binary-input discrete memoryless channel (B-DMC) W. The symmetric capacity is the highest rate achievable subject to using the input letters of the channel with equal probability. Channel polarization refers to the fact that it is pos-**

A. Preliminaries

We write $W : \mathcal{X} \to \mathcal{Y}$ to denote a generic B-DMC with input alphabet $\mathcal{X}$, output alphabet $\mathcal{Y}$, and transition probabilities $W(y|x)$, $x \in \mathcal{X}$, $y \in \mathcal{Y}$. The input alphabet $\mathcal{X}$ will always be $\{0, 1\}$, the output alphabet and the transition probabilities may

- They are capacity-achieving on binary memoryless symmetric (BMS) channels with low encoding/decoding complexity [Arı09].

Polar Codes

IEEE TRANSACTIONS ON INFORMATION THEORY, VOL. 55, NO. 7, JULY 2009 3051

Channel Polarization: A Method for Constructing Capacity-Achieving Codes for Symmetric Binary-Input Memoryless Channels

Erdal Arıkan, *Senior Member, IEEE*

***Abstract*—A method is proposed, called channel polarization, to construct code sequences that achieve the symmetric capacity $I(W)$ of any given binary-input discrete memoryless channel (B-DMC) W. The symmetric capacity is the highest rate achievable subject to using the input letters of the channel with equal probability. Channel polarization refers to the fact that it is pos-**

A. Preliminaries

We write $W : \mathcal{X} \rightarrow \mathcal{Y}$ to denote a generic B-DMC with input alphabet $\mathcal{X}$, output alphabet $\mathcal{Y}$, and transition probabilities $W(y|x)$, $x \in \mathcal{X}$, $y \in \mathcal{Y}$. The input alphabet $\mathcal{X}$ will always be $\{0, 1\}$, the output alphabet and the transition probabilities may

- They are capacity-achieving on binary memoryless symmetric (BMS) channels with low encoding/decoding complexity [Arı09].
- But successive cancellation (SC) decoding performs poorly for small blocks due to imperfect polarization.

Channel Polarization

Among all, there are channels for which it is easy to communicate optimally:

Channel Polarization

Among all, there are channels for which it is easy to communicate optimally:

- Noiseless channels: The output Y determines the input X (i.e., $I(X;Y) \approx 1$).
- Useless channels: The output Y is independent from the input X (i.e., $I(X;Y) \approx 0$).

Channel Polarization

Among all, there are channels for which it is easy to communicate optimally:

- Noiseless channels: The output Y determines the input X (i.e., $I(X;Y) \approx 1$).
- Useless channels: The output Y is independent from the input X (i.e., $I(X;Y) \approx 0$).

Channel polarization is a technique to convert any BMS channel to a mixture of easy channels + vanishing fraction of mediocre channels, asymptotically in the block length.

Channel Polarization

Among all, there are channels for which it is easy to communicate optimally:

- Noiseless channels: The output Y determines the input X (i.e., $I(X;Y) \approx 1$).
- Useless channels: The output Y is independent from the input X (i.e., $I(X;Y) \approx 0$).

Channel polarization is a technique to convert any BMS channel to a mixture of easy channels + vanishing fraction of mediocre channels, asymptotically in the block length.

- Transmit uniformly distributed information bits over the noiseless channels and set the inputs of the rest to zeros (the receiver knows their positions).

Channel Polarization

Among all, there are channels for which it is easy to communicate optimally:

- Noiseless channels: The output Y determines the input X (i.e., $I(X;Y) \approx 1$).
- Useless channels: The output Y is independent from the input X (i.e., $I(X;Y) \approx 0$).

Channel polarization is a technique to convert any BMS channel to a mixture of easy channels + vanishing fraction of mediocre channels, asymptotically in the block length.

- Transmit uniformly distributed information bits over the noiseless channels and set the inputs of the rest to zeros (the receiver knows their positions).
- The technique is lossless in terms of mutual information (required to achieve the capacity).

Channel Polarization

Among all, there are channels for which it is easy to communicate optimally:

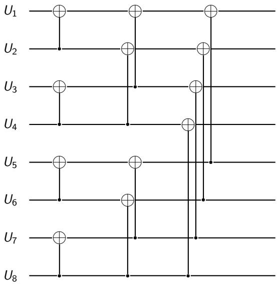
- Noiseless channels: The output Y determines the input X (i.e., $I(X;Y) \approx 1$).
- Useless channels: The output Y is independent from the input X (i.e., $I(X;Y) \approx 0$).

Channel polarization is a technique to convert any BMS channel to a mixture of easy channels + vanishing fraction of mediocre channels, asymptotically in the block length.

- Transmit uniformly distributed information bits over the noiseless channels and set the inputs of the rest to zeros (the receiver knows their positions).
- The technique is lossless in terms of mutual information (required to achieve the capacity).
- The technique is of low complexity (there exists an encoder-decoder pair, realizing the technique with $\mathcal{O}(N \log N)$ complexity, where N is the block length).

Encoding

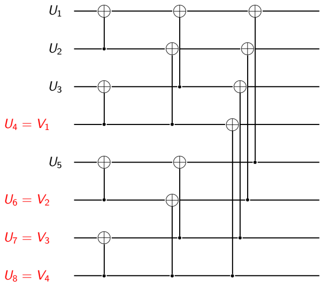
Let V_1^K denote the random information bits to be encoded:



Encoding

Let V_1^K denote the random information bits to be encoded:

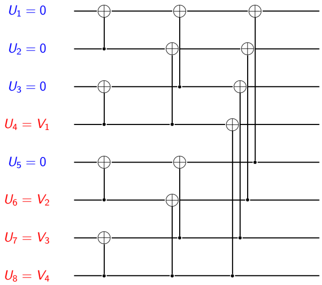
1. For a given set $\mathcal{A}$, map V_1^K onto $U_{\mathcal{A}}$.



Encoding

Let V_1^K denote the random information bits to be encoded:

1. For a given set $\mathcal{A}$, map V_1^K onto $U_{\mathcal{A}}$.
2. Set the remaining elements to 0, i.e., $U_{\mathcal{F}} = 0$ (frozen bits).



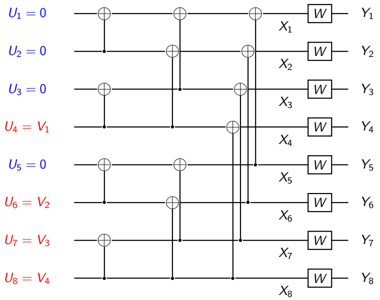
Encoding

Let V_1^K denote the random information bits to be encoded:

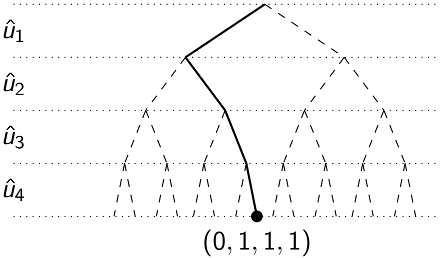
1. For a given set $\mathcal{A}$, map V_1^K onto $U_{\mathcal{A}}$.
2. Set the remaining elements to 0, i.e., $U_{\mathcal{F}} = 0$ (frozen bits).
3. Apply polar transform of length$-N$, i.e.,

$$X_1^N = U_1^N \mathbf{G}_2^{\otimes \log_2 N}$$

and transmit X_1^N over the channel after suitable modulation (the figure assumes w.l.o.g. a binary-input channel).



Successive Cancellation Decoding

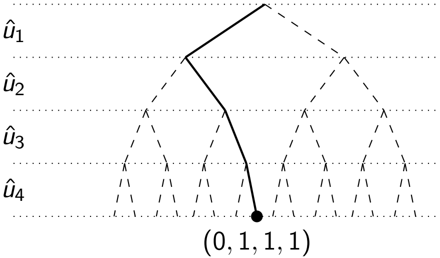


$$\hat{u}_i = \begin{cases} u_i & \text{if } i \in \mathcal{F} \\ f_i\left(y_1^N, \hat{u}_1^{i-1}\right) & \text{if } i \in \mathcal{A}. \end{cases}$$

$$f_i\left(y_1^N, \hat{u}_1^{i-1}\right) \triangleq \begin{cases} 0 & \text{if } p_{Y^N,U^{i-1}|U_i}(y_1^N, \hat{u}_1^{i-1}|0) \geq p_{Y^N,U^{i-1}|U_i}(y_1^N, \hat{u}_1^{i-1}|1) \\ 1 & \text{otherwise} \end{cases}$$

where a frame error occurs if $\hat{u}_i \neq u_i$ for any $i \in \mathcal{A}$

Successive Cancellation Decoding

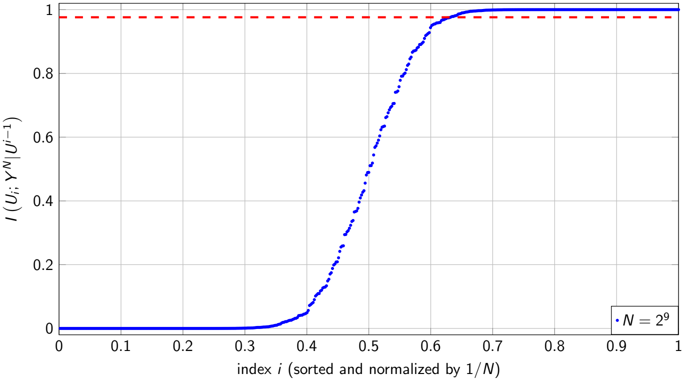


$$\hat{u}_i = \begin{cases} u_i & \text{if } i \in \mathcal{F} \\ f_i\left(y_1^N, \hat{u}_1^{i-1}\right) & \text{if } i \in \mathcal{A}. \end{cases}$$

$$f_i\left(y_1^N, \hat{u}_1^{i-1}\right) \triangleq \begin{cases} 0 & \text{if } p_{Y^N,U^{i-1}|U_i}(y_1^N, \hat{u}_1^{i-1}|0) \geq p_{Y^N,U^{i-1}|U_i}(y_1^N, \hat{u}_1^{i-1}|1) \\ 1 & \text{otherwise} \end{cases}$$

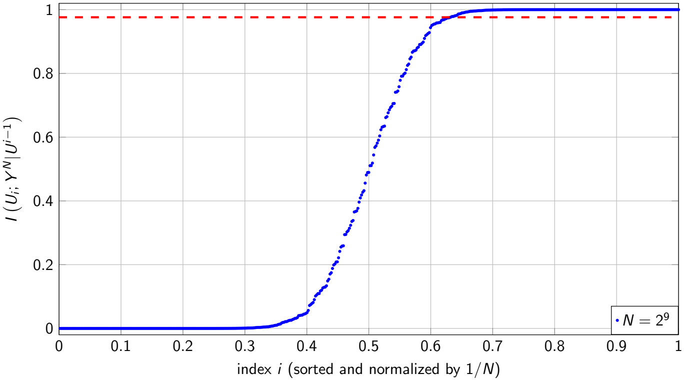
where a frame error occurs if $\hat{u}_i \neq u_i$ for any $i \in \mathcal{A} \longleftrightarrow$ imperfect channel polarization at finite-length regime!

Imperfect Channel Polarization at Finite-Length Regime



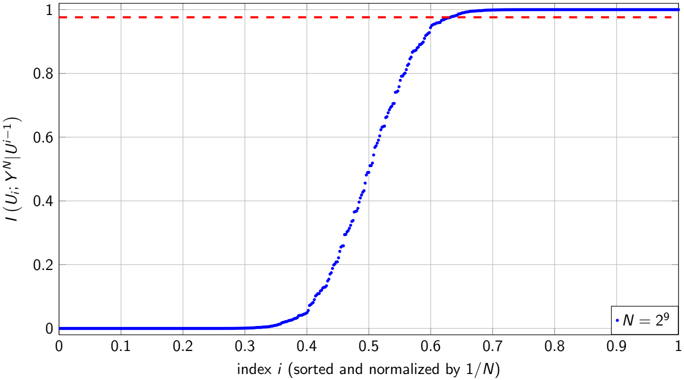
- Choose set $\mathcal{A}$ to contain the most reliable K indices.

Imperfect Channel Polarization at Finite-Length Regime



- Choose set $\mathcal{A}$ to contain the most reliable K indices.
- Any error made by SC decoding cannot be corrected

Imperfect Channel Polarization at Finite-Length Regime



- Choose set $\mathcal{A}$ to contain the most reliable K indices.
- Any error made by SC decoding cannot be corrected $\rightarrow$ use successive cancellation list (SCL) decoding to make use of the frozen bits in reliable positions for error-correction!

Successive List Cancellation Decoding

IEEE TRANSACTIONS ON INFORMATION THEORY, VOL. 61, NO. 5, MAY 2015 2213

List Decoding of Polar Codes

Ido Tal, *Member, IEEE* and Alexander Vardy, *Fellow, IEEE*

***Abstract*—We describe a successive-cancellation list decoder for polar codes, which is a generalization of the classic successive-cancellation decoder of Arıkan. In the proposed list decoder, L decoding paths are considered concurrently at each decoding stage, where L is an integer parameter. At the end of the decoding process, the most likely among the L paths is selected as the single codeword at the decoder output. Simulations show that the resulting performance is very close to that of maximum-likelihood decoding, even for moderate values of L. Alternatively, if a genie is allowed to pick the transmitted codeword from the list, the results are comparable with the performance of current state-of-the-art LDPC codes. We show that such a genie can be easily implemented using simple CRC precoding. The specific list-decoding algorithm that achieves this performance doubles the number of decoding paths for each information bit, and then uses a pruning procedure to discard all but the L most likely paths. However, straightforward implementation of this**

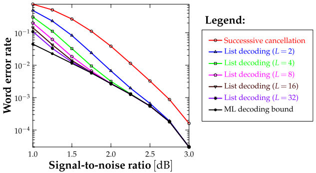


Fig. 1. List-decoding performance for a polar code of length $n = 2048$ and rate $R = 0.5$ on the BPSK-modulated Gaussian channel. The code was constructed using the methods of [15], with optimization for $E_b/N_0 = 2$ dB.

- SC list (SCL) decoding with CRC and large list-size performs very well and approaches maximum-likelihood (ML) decoding performance [TV15].

Successive List Cancellation Decoding

IEEE TRANSACTIONS ON INFORMATION THEORY, VOL. 61, NO. 5, MAY 2015 2213

List Decoding of Polar Codes

Ido Tal, *Member, IEEE* and Alexander Vardy, *Fellow, IEEE*

***Abstract*— We describe a successive-cancellation list decoder for polar codes, which is a generalization of the classic successive-cancellation decoder of Arıkan. In the proposed list decoder, L decoding paths are considered concurrently at each decoding stage, where L is an integer parameter. At the end of the decoding process, the most likely among the L paths is selected as the single codeword at the decoder output. Simulations show that the resulting performance is very close to that of maximum-likelihood decoding, even for moderate values of L. Alternatively, if a genie is allowed to pick the transmitted codeword from the list, the results are comparable with the performance of current state-of-the-art LDPC codes. We show that such a genie can be easily implemented using simple CRC precoding. The specific list-decoding algorithm that achieves this performance doubles the number of decoding paths for each information bit, and then uses a pruning procedure to discard all but the L most likely paths. However, straightforward implementation of this**

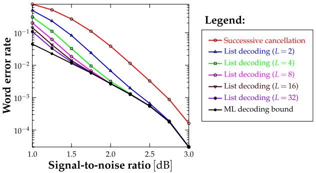


Fig. 1. List-decoding performance for a polar code of length $n = 2048$ and rate $R = 0.5$ on the BPSK-modulated Gaussian channel. The code was constructed using the methods of [15], with optimization for $E_b/N_0 = 2$ dB.

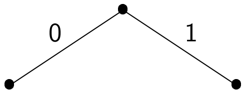
- SC list (SCL) decoding with CRC and large list-size performs very well and approaches maximum-likelihood (ML) decoding performance [TV15].
- It can also be used to decode other codes (e.g., Reed–Muller codes, PAC codes, etc.).

Successive Cancellation List Decoding

Key idea: Each time a decision is needed on $\hat{u}_i$, both options, i.e., $\hat{u}_i = 0$ and $\hat{u}_i = 1$, are stored. This doubles the number of partial input sequences (paths) at each decoding stage.

Successive Cancellation List Decoding

Key idea: Each time a decision is needed on $\hat{u}_i$, both options, i.e., $\hat{u}_i = 0$ and $\hat{u}_i = 1$, are stored. This doubles the number of partial input sequences (paths) at each decoding stage.

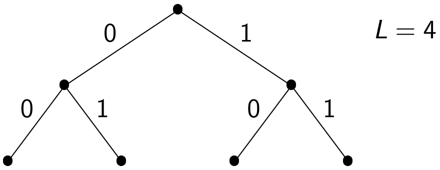


$L = 4$

- When the number of paths exceeds a predefined list size L, discard the least likely paths.

Successive Cancellation List Decoding

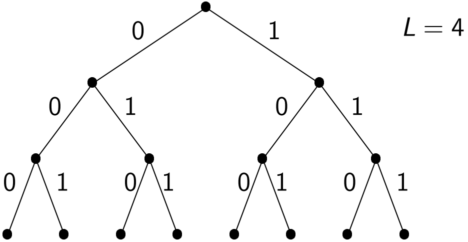
Key idea: Each time a decision is needed on $\hat{u}_i$, both options, i.e., $\hat{u}_i = 0$ and $\hat{u}_i = 1$, are stored. This doubles the number of partial input sequences (paths) at each decoding stage.



- When the number of paths exceeds a predefined list size L, discard the least likely paths.

Successive Cancellation List Decoding

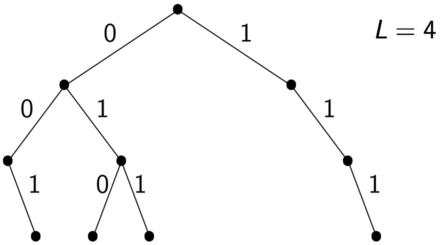
Key idea: Each time a decision is needed on $\hat{u}_i$, both options, i.e., $\hat{u}_i = 0$ and $\hat{u}_i = 1$, are stored. This doubles the number of partial input sequences (paths) at each decoding stage.



- When the number of paths exceeds a predefined list size L, discard the least likely paths.

Successive Cancellation List Decoding

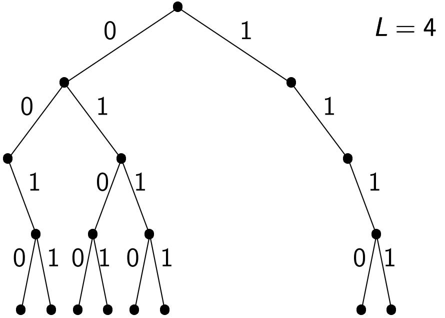
Key idea: Each time a decision is needed on $\hat{u}_i$, both options, i.e., $\hat{u}_i = 0$ and $\hat{u}_i = 1$, are stored. This doubles the number of partial input sequences (paths) at each decoding stage.



- When the number of paths exceeds a predefined list size L, discard the least likely paths.

Successive Cancellation List Decoding

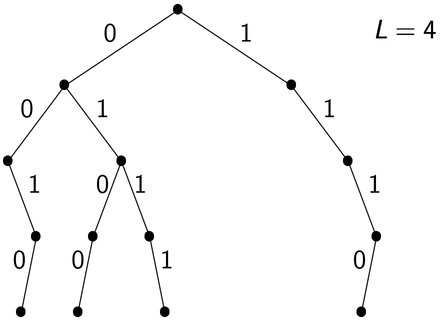
Key idea: Each time a decision is needed on $\hat{u}_i$, both options, i.e., $\hat{u}_i = 0$ and $\hat{u}_i = 1$, are stored. This doubles the number of partial input sequences (paths) at each decoding stage.



- When the number of paths exceeds a predefined list size L, discard the least likely paths.

Successive Cancellation List Decoding

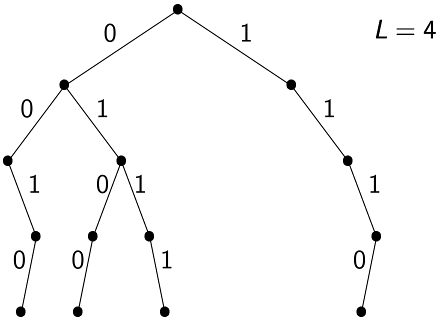
Key idea: Each time a decision is needed on $\hat{u}_i$, both options, i.e., $\hat{u}_i = 0$ and $\hat{u}_i = 1$, are stored. This doubles the number of partial input sequences (paths) at each decoding stage.



- When the number of paths exceeds a predefined list size L, discard the least likely paths.

Successive Cancellation List Decoding

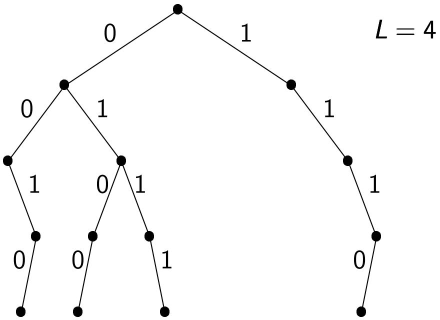
Key idea: Each time a decision is needed on $\hat{u}_i$, both options, i.e., $\hat{u}_i = 0$ and $\hat{u}_i = 1$, are stored. This doubles the number of partial input sequences (paths) at each decoding stage.



- When the number of paths exceeds a predefined list size L, discard the least likely paths.
- After N-th stage, $\hat{u}_1^N = \arg\max_{u_1^N \in \mathcal{L}_N} p_{Y_1^N, U_1^{N-1}|U_N}(y_1^N, u_1^{N-1}|u_N) = \arg\max_{u_1^N \in \mathcal{L}_N} \Pr(y_1^N|u_1^N)$.

Successive Cancellation List Decoding

Key idea: Each time a decision is needed on $\hat{u}_i$, both options, i.e., $\hat{u}_i = 0$ and $\hat{u}_i = 1$, are stored. This doubles the number of partial input sequences (paths) at each decoding stage.



- When the number of paths exceeds a predefined list size L, discard the least likely paths.
- After N-th stage, $\hat{u}_1^N = \arg\max_{u_1^N \in \mathcal{L}_N} p_{Y_1^N, U_1^{N-1}|U_N}(y_1^N, u_1^{N-1}|u_N) = \arg\max_{u_1^N \in \mathcal{L}_N} \Pr(y_1^N|u_1^N)$.
- Very similar ideas were applied to RM codes (see, e.g., [Sto02, DS06]).

Outline

Motivating Question

- Is it possible to use frozen bits not only for error-correction but also for estimating the channel state when communicating over unknown fading channels?

Motivating Question

- Is it possible to use frozen bits not only for error-correction but also for estimating the channel state when communicating over unknown fading channels?
    - Frozen bits are known at the receiver (similar to pilot symbols), which...

Motivating Question

- Is it possible to use frozen bits not only for error-correction but also for estimating the channel state when communicating over unknown fading channels?
  - Frozen bits are known at the receiver (similar to pilot symbols), which...
  - ... cannot be treated as pilot symbols because of the polar transform before transmitting :(

Motivating Question

- Is it possible to use frozen bits not only for error-correction but also for estimating the channel state when communicating over unknown fading channels?
  - Frozen bits are known at the receiver (similar to pilot symbols), which...
  - ... cannot be treated as pilot symbols because of the polar transform before transmitting :(
- We propose a pilot-free two-stage polar-coded transmission (PCT) scheme to jointly estimate the channel state and data...

Motivating Question

- Is it possible to use frozen bits not only for error-correction but also for estimating the channel state when communicating over unknown fading channels?
  - Frozen bits are known at the receiver (similar to pilot symbols), which...
  - ... cannot be treated as pilot symbols because of the polar transform before transmitting :(
- We propose a pilot-free two-stage polar-coded transmission (PCT) scheme to jointly estimate the channel state and data...
  - ... where, first, SCL decoding and the polar code constraints are used to estimate the channel state.

Motivating Question

- Is it possible to use frozen bits not only for error-correction but also for estimating the channel state when communicating over unknown fading channels?
  - Frozen bits are known at the receiver (similar to pilot symbols), which...
  - ... cannot be treated as pilot symbols because of the polar transform before transmitting :(
- We propose a pilot-free two-stage polar-coded transmission (PCT) scheme to jointly estimate the channel state and data...
  - ... where, first, SCL decoding and the polar code constraints are used to estimate the channel state. Then, mismatched SCL decoding proceeds with this estimate.

Motivating Question

- Is it possible to use frozen bits not only for error-correction but also for estimating the channel state when communicating over unknown fading channels?
  - Frozen bits are known at the receiver (similar to pilot symbols), which...
  - ... cannot be treated as pilot symbols because of the polar transform before transmitting :(
- We propose a pilot-free two-stage polar-coded transmission (PCT) scheme to jointly estimate the channel state and data...
  - ... where, first, SCL decoding and the polar code constraints are used to estimate the channel state. Then, mismatched SCL decoding proceeds with this estimate.

Based on a joint work with Peihong Yuan and Gerhard Kramer (TUM) [YCK21]


Related Works

- A related method to estimate CSI uses the parity-check constraints of a low-density parity-check (LDPC) code [IHG10]; however, ...

Related Works

- A related method to estimate CSI uses the parity-check constraints of a low-density parity-check (LDPC) code [IHG10]; however, ...
  - ✓ ... SCL decoding of polar codes naturally provides soft estimates of frozen bits (no modification required for the decoder).

Related Works

- A related method to estimate CSI uses the parity-check constraints of a low-density parity-check (LDPC) code [IHG10]; however, ...
  - ✓ ... SCL decoding of polar codes naturally provides soft estimates of frozen bits (no modification required for the decoder).
  - ✓ ... polar codes are usually used with a high-rate outer code [TV15, 5G21] that can resolve CSI ambiguities, e.g., the phase ambiguity due to all-one codeword when using QPSK and Gray labeling [IHG10].

Related Works

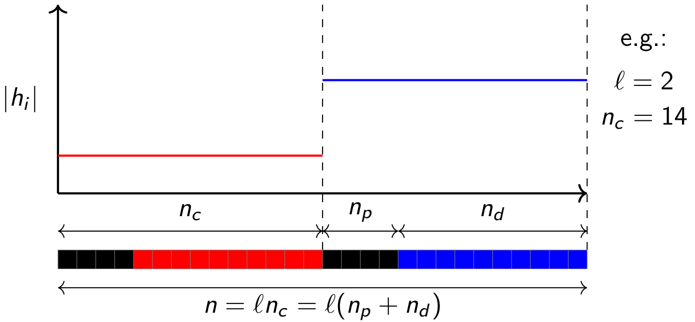
- A related method to estimate CSI uses the parity-check constraints of a low-density parity-check (LDPC) code [IHG10]; however, ...
  - ✓ ... SCL decoding of polar codes naturally provides soft estimates of frozen bits (no modification required for the decoder).
  - ✓ ... polar codes are usually used with a high-rate outer code [TV15, 5G21] that can resolve CSI ambiguities, e.g., the phase ambiguity due to all-one codeword when using QPSK and Gray labeling [IHG10].
  - ✓ ... polar codes concatenated with outer CRC codes are very competitive in short block length regime [CDJ+19], where pilot symbols cost a large overhead [ODS+19].

System Model (PAT)

- The input-output relationship of the channel is given by

$$\boldsymbol{y}_i = h_i \boldsymbol{x}_i + \boldsymbol{n}_i \qquad \text{for} \qquad i = 1, \ldots, \ell$$

where $\boldsymbol{x}_i = \left[\boldsymbol{x}_i^{(p)}, \boldsymbol{x}_i^{(d)}\right] \in \mathcal{X}^{n_c}$, $\boldsymbol{y}_i \in \mathbb{C}^{n_c}$, $H_i \sim P_H$ and $\boldsymbol{N}_i \sim \mathcal{CN}(0, \sigma^2 \boldsymbol{I}_{n_c})$.

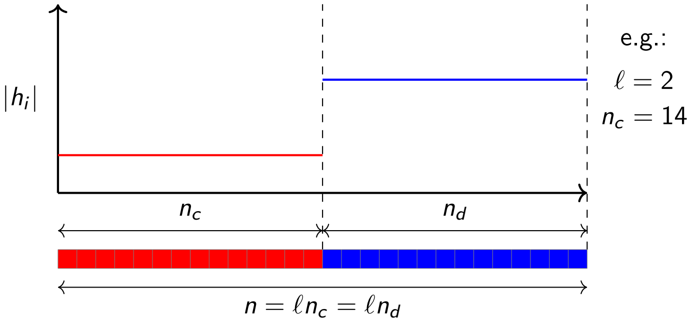


System Model (PCT)

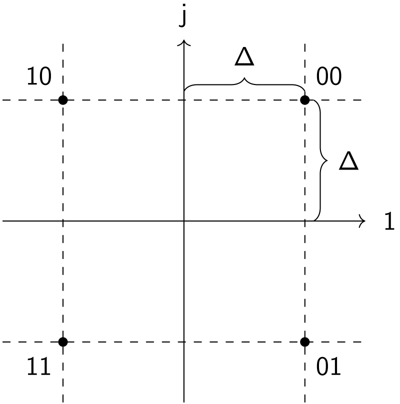
- The input-output relationship of the channel is given by

$$\boldsymbol{y}_i = h_i \boldsymbol{x}_i + \boldsymbol{n}_i \qquad \text{for} \qquad i = 1, \ldots, \ell \tag{1}$$

where $\boldsymbol{x}_i = \boldsymbol{x}_i^{(d)} \in \mathcal{X}^{n_c}$, $\boldsymbol{y}_i \in \mathbb{C}^{n_c}$, $H_i \sim P_H$ and $\boldsymbol{N}_i \sim \mathcal{CN}(0, \sigma^2 \boldsymbol{I}_{n_c})$.



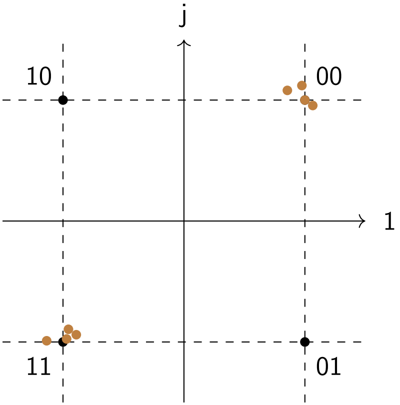
Mini Example



- $\ell = 1$
- Repetition code, $R = 0.5$

$$w_1, w_2, \ldots \mapsto w_1, w_1, w_2, w_2, \ldots$$

Mini Example

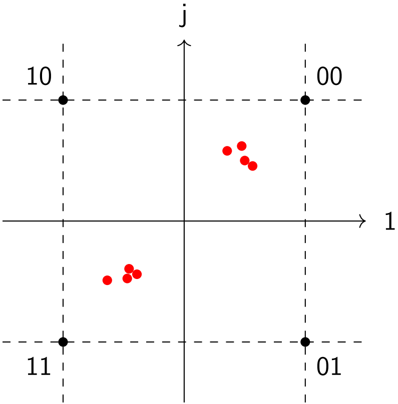


$h = 1$

- $\ell = 1$
- Repetition code, $R = 0.5$

$$w_1, w_2, \ldots \mapsto w_1, w_1, w_2, w_2, \ldots$$

Mini Example

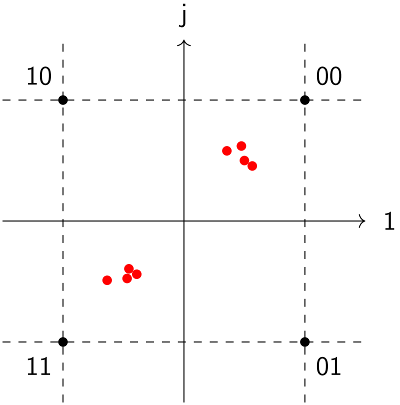


$h = ?$

- $\ell = 1$
- Repetition code, $R = 0.5$

$$w_1, w_2, \ldots \mapsto w_1, w_1, w_2, w_2, \ldots$$

Mini Example



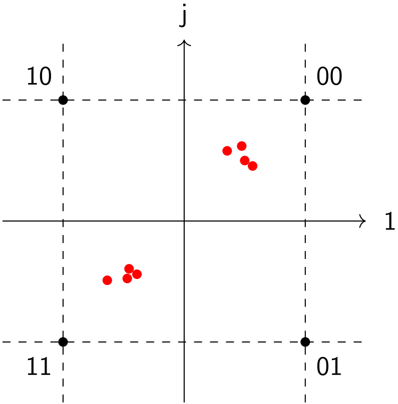
$h = ?$

- $\ell = 1$
- Repetition code, $R = 0.5$

$$w_1, w_2, \ldots \mapsto w_1, w_1, w_2, w_2, \ldots$$

- $h \approx 0.5$ or $h \approx -0.5$

Mini Example



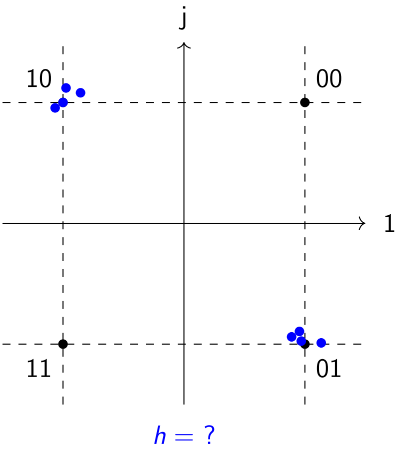
$h = ?$

- $\ell = 1$
- Repetition code, $R = 0.5$

$$w_1, w_2, \ldots \mapsto w_1, w_1, w_2, w_2, \ldots$$

- $h \approx 0.5$ or $h \approx -0.5$
  - all one is a valid codeword
  - symmetric mapping
  - $h \cdot \boldsymbol{x} = -h \cdot -\boldsymbol{x}$

Mini Example

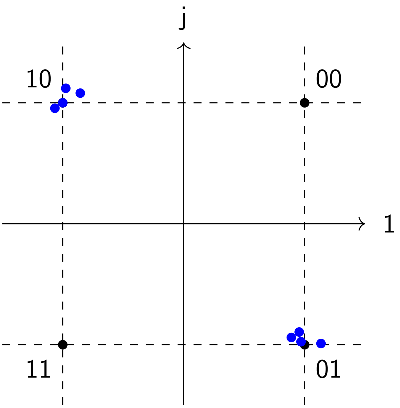


- $\ell = 1$
- Repetition code, $R = 0.5$

$$w_1, w_2, \ldots \mapsto w_1, w_1, w_2, w_2, \ldots$$

- $h \approx 0.5$ or $h \approx -0.5$
  - all one is a valid codeword
  - symmetric mapping
  - $h \cdot \boldsymbol{x} = -h \cdot -\boldsymbol{x}$

Mini Example



$h = ?$

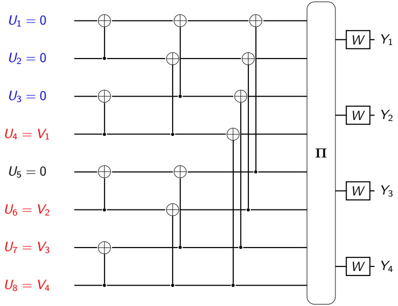
- $\ell = 1$
- Repetition code, $R = 0.5$

$$w_1, w_2, \ldots \mapsto w_1, w_1, w_2, w_2, \ldots$$

- $h \approx 0.5$ or $h \approx -0.5$
  - all one is a valid codeword
  - symmetric mapping
  - $h \cdot \boldsymbol{x} = -h \cdot -\boldsymbol{x}$
- $h \approx e^{\frac{\pi}{2}}$ or $h \approx e^{-\frac{\pi}{2}}$

Mini Example

- $\ell = 1$, $\mathcal{A} = \{4, 6, 7, 8\}$ and $\mathcal{F} = \{1, 2, 3, 5\}$

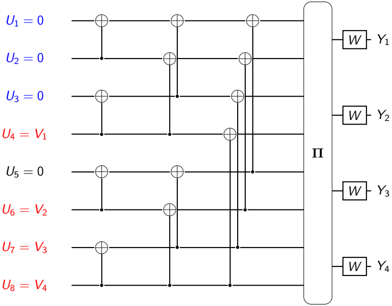


Mini Example

- $\ell = 1$, $\mathcal{A} = \{4, 6, 7, 8\}$ and $\mathcal{F} = \{1, 2, 3, 5\}$
- A channel estimate is obtained as

$$\hat{h} = \arg\max_{h} p_{Y_1^4|U_1^3,H_1}(y_1^4|000, h)$$

where $p_{Y_1^4|U_1^3,H_1}(y_1^4|000, h)$ can be efficiently computed via SCL decoding with $L_e = 1$ and $\beta = 3$ for any h.



Mini Example

- $\ell = 1$, $\mathcal{A} = \{4,6,7,8\}$ and $\mathcal{F} = \{1,2,3,5\}$
- A channel estimate is obtained as

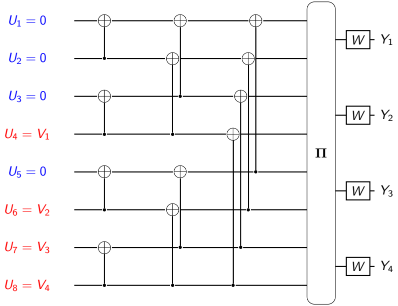
$$\hat{h} = \arg\max_{h} p_{Y_1^4|U_1^3,H_1}(y_1^4|000, h)$$

where $p_{Y_1^4|U_1^3,H_1}(y_1^4|000, h)$ can be efficiently computed via SCL decoding with $L_e = 1$ and $\beta = 3$ for any h.

- A more accurate estimate is

$$\begin{aligned}\hat{h} &= \arg\max_{h} p_{Y_1^4|U_1^3,U_5,H_1}(y_1^4|0000, h) \\ &= \arg\max_{h} \left[p_{Y_1^4,U_4|U_1^3,U_5,H_1}(y_1^4, 0|0000, h) \right. \\ &\qquad\qquad \left. + p_{Y_1^4,U_4|U_1^3,U_5,H_1}(y_1^4, 1|0000, h)\right]\end{aligned}$$

where the cost function requires SCL with $L_e = 2$ and $\beta = 5$ for any h.



General Case

- Let $h_i = r_i e^{j\theta_i}$ where $r_i \in [0, \infty)$ and $\theta_i \in [0, 2\pi)$, $i \in [\ell]$.

General Case

- Let $h_i = r_i e^{j\theta_i}$ where $r_i \in [0, \infty)$ and $\theta_i \in [0, 2\pi)$, $i \in [\ell]$.
- Let β be a number of input bits, and let $\mathcal{A}^{(\beta)} = \mathcal{A} \cap [\beta]$ and $\mathcal{F}^{(\beta)} = \mathcal{F} \cap [\beta]$ be sets of information and frozen indices among the first β input bits u_1^β.

General Case

- Let $h_i = r_i e^{j\theta_i}$ where $r_i \in [0, \infty)$ and $\theta_i \in [0, 2\pi)$, $i \in [\ell]$.
- Let β be a number of input bits, and let $\mathcal{A}^{(\beta)} = \mathcal{A} \cap [\beta]$ and $\mathcal{F}^{(\beta)} = \mathcal{F} \cap [\beta]$ be sets of information and frozen indices among the first β input bits u_1^β.

1. Estimate the amplitudes $r_i = |h_i|$ as

$$\hat{r}_i = \left(\sqrt{2}\Delta\right)^{-1} \sqrt{\frac{1}{n_c}\|\boldsymbol{y}_i\|^2 - 2\sigma^2}, \quad i = 1, \ldots, \ell.$$

General Case

- Let $h_i = r_i e^{j\theta_i}$ where $r_i \in [0, \infty)$ and $\theta_i \in [0, 2\pi)$, $i \in [\ell]$.
- Let β be a number of input bits, and let $\mathcal{A}^{(\beta)} = \mathcal{A} \cap [\beta]$ and $\mathcal{F}^{(\beta)} = \mathcal{F} \cap [\beta]$ be sets of information and frozen indices among the first β input bits u_1^{β}.

1. Estimate the amplitudes $r_i = |h_i|$ as

$$\hat{r}_i = \left(\sqrt{2}\Delta\right)^{-1} \sqrt{\frac{1}{n_c}\|\boldsymbol{y}_i\|^2 - 2\sigma^2}, \quad i = 1, \ldots, \ell.$$

2. Use the polar code constraints to estimate the phase as

$$\left\{\hat{\theta}_1, \ldots, \hat{\theta}_\ell\right\} = \underset{\{\theta_1,\ldots,\theta_\ell\}}{\arg\max}\; p_{Y_1^n | U_{\mathcal{F}^{(\beta)}}, H_1^\ell}\left(y_1^n \middle| 0, \hat{h}_1^\ell\right)$$

$$= \underset{\{\theta_1,\ldots,\theta_\ell\}}{\arg\max} \sum_{u_{\mathcal{A}^{(\beta)}}} p_{Y_1^n, U_{\mathcal{A}^{(\beta)}} | U_{\mathcal{F}^{(\beta)}}, H_1^B}\left(y_1^n, u_{\mathcal{A}^{(\beta)}} \middle| 0, \hat{h}_1^\ell\right)$$

Complexity

- The search space grows exponentially in the number of diversity branches ℓ. Although there can be other ways to reduce the complexity, the following observation halves the search space.

Corollary

Polar-coded modulations with the QPSK and Gray labeling over the channel (1) satisfy

$$p_{Y_1^n|U_{\mathcal{F}(\beta)},H_1^\ell}\left(y_1^n\left|0,h_1^\ell\right.\right)=p_{Y_1^n|U_{\mathcal{F}(\beta)},H_1^\ell}\left(y_1^n\left|0,-h_1^\ell\right.\right)$$

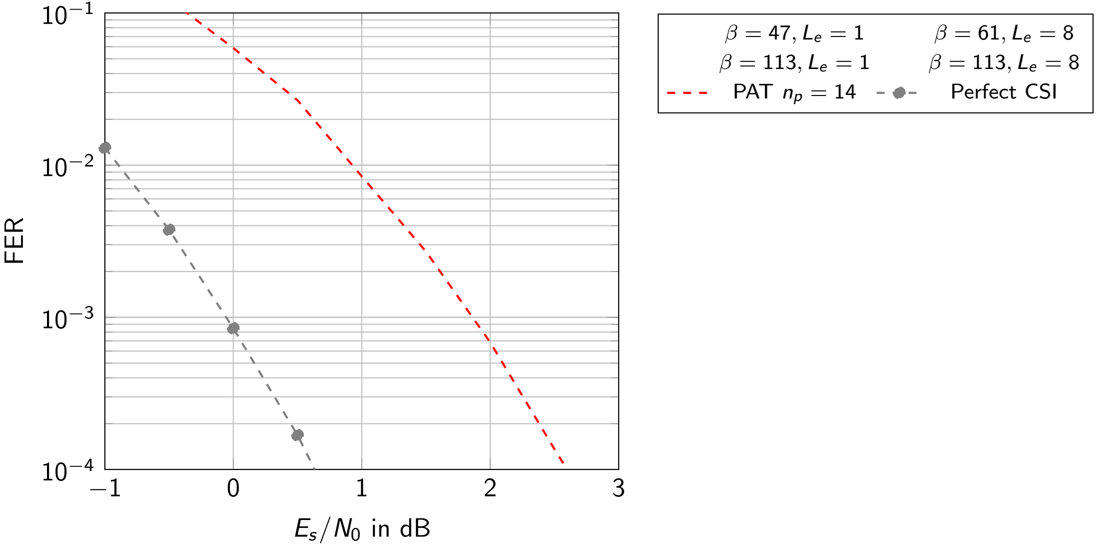
for all y_1^n and h_1^ℓ.

Outline

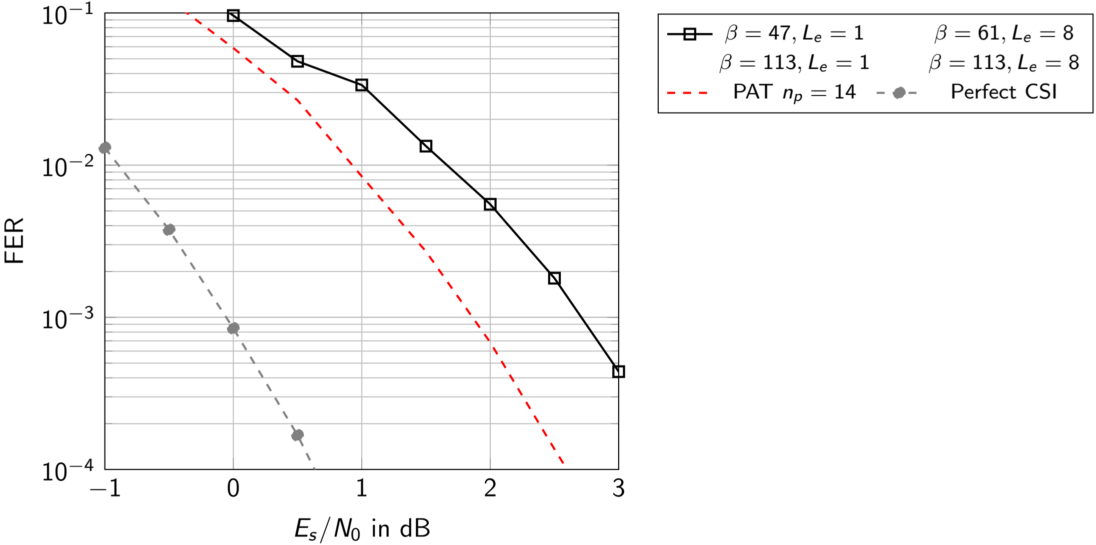
Example: Single Block ($\ell = 1$, $n_c = 64$)

- $y_i = e^{j\theta} x_i + z_i, \quad i = 1, \ldots, n_c = 64$ where $\theta \sim \mathcal{U}[0, 2\pi)$
- No CSIT/CSIR (including the amplitude)
- $(128, 38)$ 5G polar code with 6 bits CRC, $R = 0.5$ bpcu, Pilot-free
- Random interleaver
- QPSK (Gray)

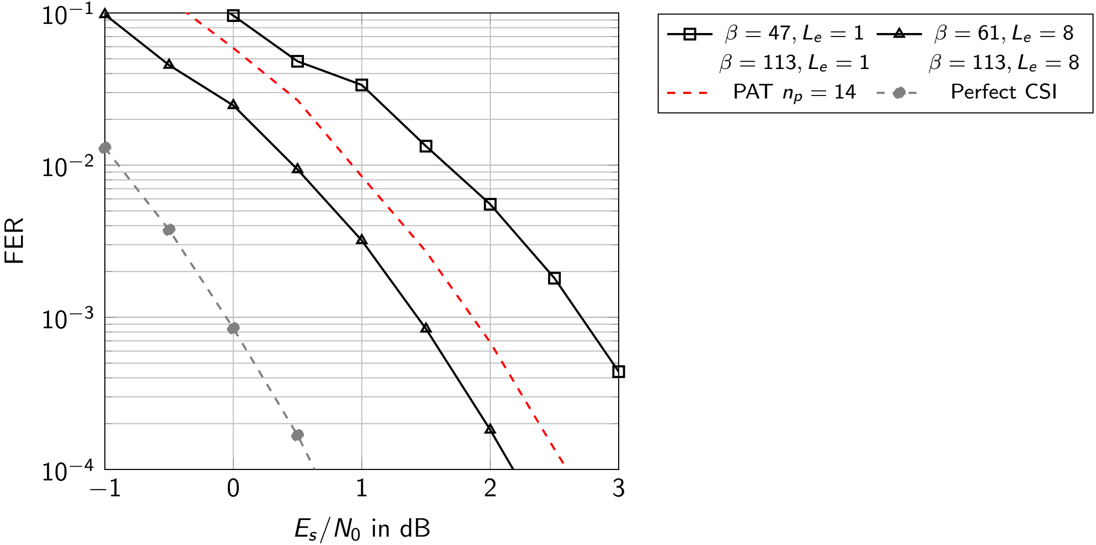
Example: Single Block ($\ell = 1$, $n_c = 64$)



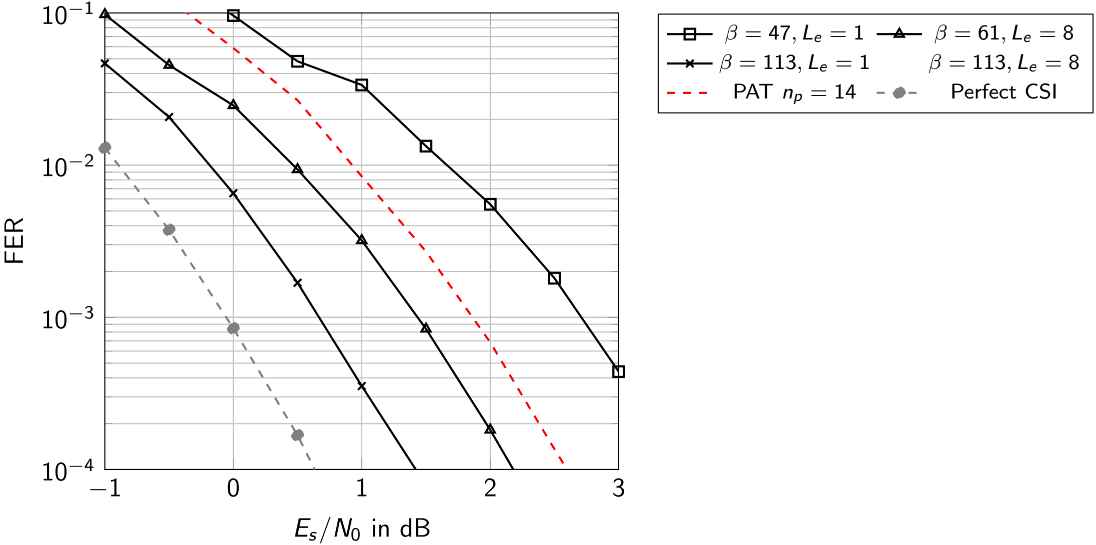
Example: Single Block ($\ell = 1,\ n_c = 64$)



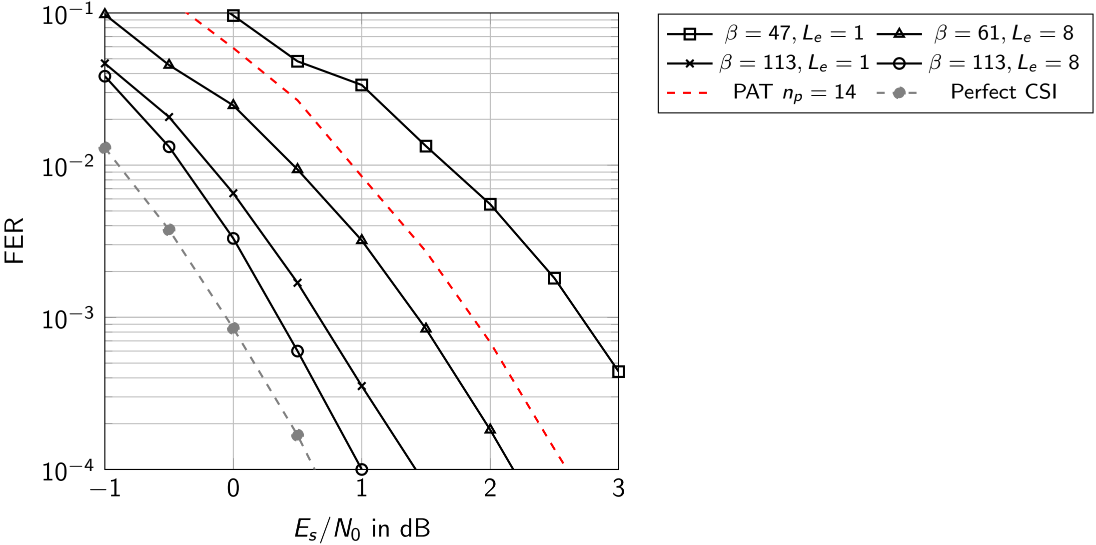
Example: Single Block ($\ell = 1$, $n_c = 64$)



Example: Single Block ($\ell = 1$, $n_c = 64$)



Example: Single Block ($\ell = 1$, $n_c = 64$)



Complexity

Table 1: Number of Visited Nodes per Frame at $E_s/N_0 = 1$ dB

Method	FER	Visited Nodes
PAT $(n_p = 14,\ L = 8)$	8.43×10^{-3}	631
PAT $(n_p = 14,\ L = 32)$	3.16×10^{-3}	2223
PCT $(\beta = 47,\ L_e = 1,\ L = 8)$	3.36×10^{-2}	1383
PCT $(\beta = 61,\ L_e = 8,\ L = 8)$	3.20×10^{-3}	2151
PCT $(\beta = 113,\ L_e = 1,\ L = 8)$	3.50×10^{-4}	2439
PCT $(\beta = 113,\ L_e = 8,\ L = 8)$	1.00×10^{-4}	8807
Perfect CSI $(L = 8)$	2.40×10^{-5}	631

$8 + 8$ coarse-fine search for the optimization

$$\hat{\theta} = \arg\max_{\theta} p_{Y_1^n | U_{\mathcal{F}_\beta} H}\left(y_1^n \middle| 0, \hat{r}e^{j\theta}\right)$$

Example: Rayleigh Block-Fading Channel ($\ell = 2,\ n_c = 32$)

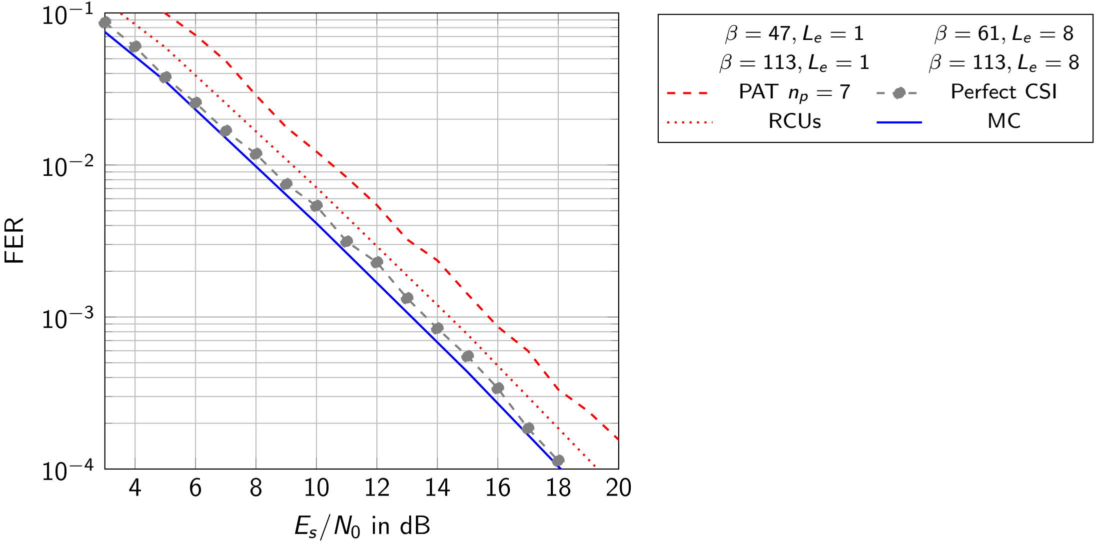
- $\boldsymbol{y}_i = h_i \boldsymbol{x}_i + \boldsymbol{n}_i, \quad i = 1, 2$ where $H_i \sim \mathcal{CN}(0, 1)$ and $\boldsymbol{N}_i \sim \mathcal{CN}(0, \sigma^2 \boldsymbol{I}_{n_c})$.
- No CSIT/CSIR
- $(128, 38)$ 5G polar code with 6 bits CRC, $R = 0.5$ bpcu, Pilot-free
- Random interleaver
- QPSK (Gray)
- Compared to RCUs [MG11] and MC [PPV10] bounds (see [LOD+20] for details)

Example: Rayleigh Block-Fading Channel ($\ell = 2,\ n_c = 32$)

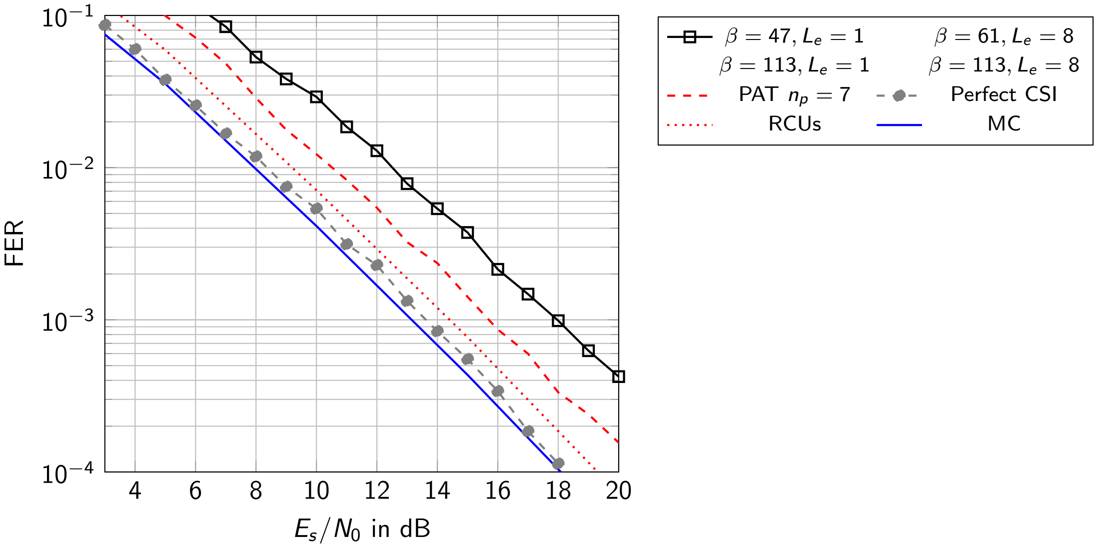
- $\boldsymbol{y}_i = h_i \boldsymbol{x}_i + \boldsymbol{n}_i, \quad i = 1, 2$ where $H_i \sim \mathcal{CN}(0, 1)$ and $\boldsymbol{N}_i \sim \mathcal{CN}(0, \sigma^2 \boldsymbol{I}_{n_c})$.
- No CSIT/CSIR
- $(128, 38)$ 5G polar code with 6 bits CRC, $R = 0.5$ bpcu, Pilot-free
- Random interleaver
- QPSK (Gray)
- Compared to RCUs [MG11] and MC [PPV10] bounds (see [LOD+20] for details)

(Courtesy of Dr. A. Lancho (Chalmers, MIT))

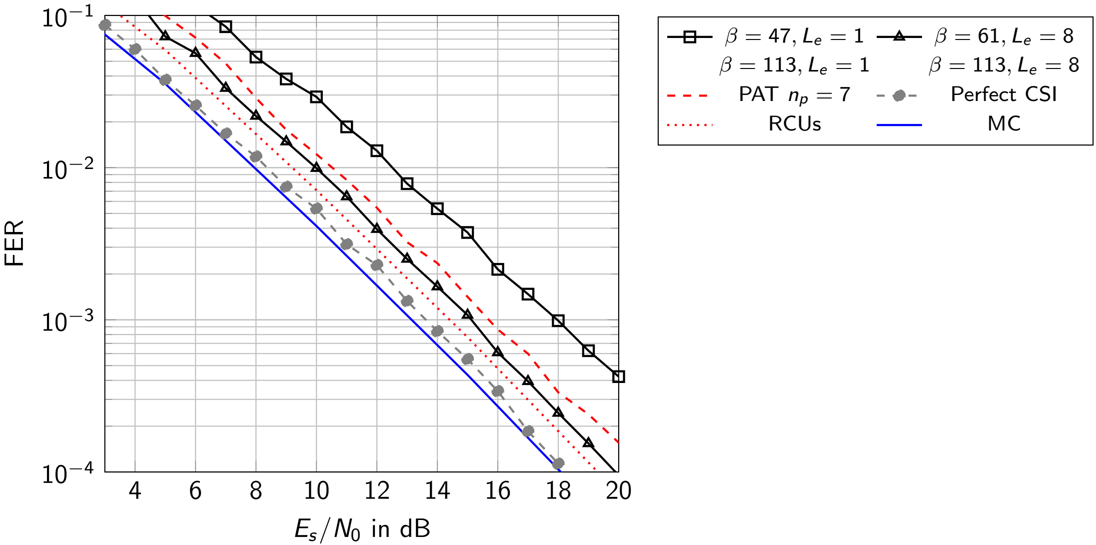
Example: Rayleigh Block-Fading Channel ($\ell = 2$, $n_c = 32$)



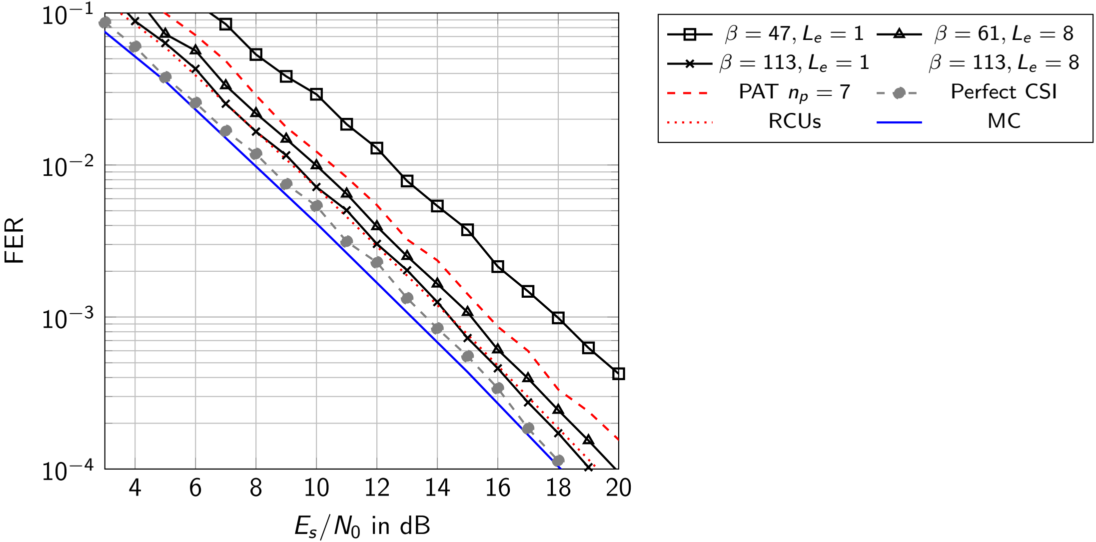
Example: Rayleigh Block-Fading Channel ($\ell = 2$, $n_c = 32$)



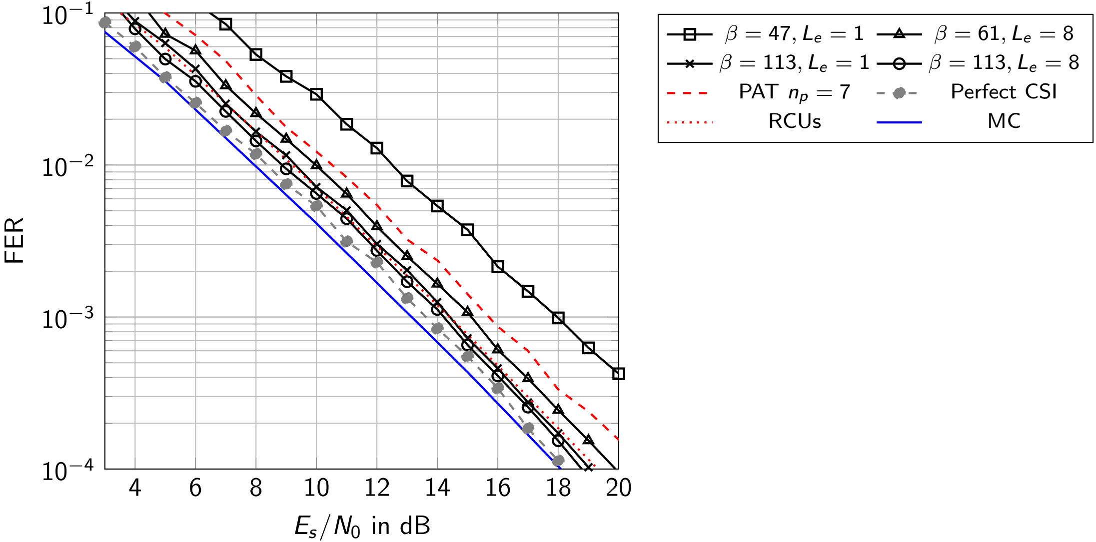
Example: Rayleigh Block-Fading Channel ($\ell = 2$, $n_c = 32$)



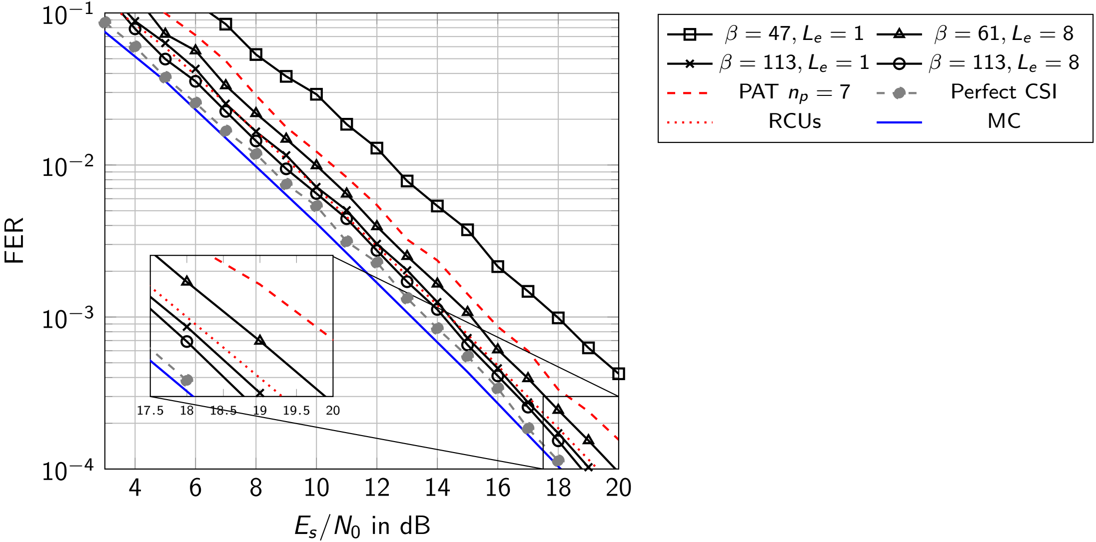
Example: Rayleigh Block-Fading Channel ($\ell = 2$, $n_c = 32$)



Example: Rayleigh Block-Fading Channel ($\ell = 2$, $n_c = 32$)



Example: Rayleigh Block-Fading Channel ($B = 2$, $n_c = 32$)



Outline

Conclusions and Outlook

- We presented pilot-free non-coherent transmission, providing gains up to 2 dB over classical PAT schemes.

Conclusions and Outlook

- We presented pilot-free non-coherent transmission, providing gains up to 2 dB over classical PAT schemes.
- Search space grows exponentially with diversity branches $\rightarrow$ combination with PAT is promising!

Conclusions and Outlook

- We presented pilot-free non-coherent transmission, providing gains up to 2 dB over classical PAT schemes.
- Search space grows exponentially with diversity branches $\rightarrow$ combination with PAT is promising!
- The scheme could be extended to operate over multiple antenna systems, as well as in combination with high-order modulation.

Conclusions and Outlook

- We presented pilot-free non-coherent transmission, providing gains up to 2 dB over classical PAT schemes.
- Search space grows exponentially with diversity branches $\rightarrow$ combination with PAT is promising!
- The scheme could be extended to operate over multiple antenna systems, as well as in combination with high-order modulation.
- A tailored code design should lead to interesting optimization problems, where...

Conclusions and Outlook

- We presented pilot-free non-coherent transmission, providing gains up to 2 dB over classical PAT schemes.
- Search space grows exponentially with diversity branches $\rightarrow$ combination with PAT is promising!
- The scheme could be extended to operate over multiple antenna systems, as well as in combination with high-order modulation.
- A tailored code design should lead to interesting optimization problems, where...
  - ... freezing reliable bit positions could improve the channel estimation, and this may be reflected in an overall performance gain.

References I

Technical specification group radio access network - nr - multiplexing and channel coding, March 2021.

E. Arıkan.
Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels.
IEEE Trans. Inf. Theory, 55(7):3051–3073, July 2009.

M. C. Coşkun, G. Durisi, T. Jerkovits, G. Liva, W. E. Ryan, B. Stein, and F. Steiner.
Efficient error-correcting codes in the short blocklength regime.
Elsevier Phys. Commun., 34:66–79, June 2019.

I. Dumer and K. Shabunov.
Soft-decision decoding of reed-muller codes: recursive lists.
IEEE Trans. Inf. Theory, 52(3):1260–1266, 2006.

R. Imad, S. Houcke, and M. Ghogho.
Blind estimation of the phase and carrier frequency offsets for LDPC-coded systems.
EURASIP J. Adv. Signal Process., 2010(1):1–13, 2010.

References II

A. Lancho, J. Östman, G. Durisi, T. Koch, and G. Vazquez-Vilar.
Saddlepoint approximations for short-packet wireless communications.
IEEE Trans. Wireless Commun., 19(7):4831–4846, 2020.

A. Martinez and A. Guillén i Fàbregas.
Saddlepoint approximation of random–coding bounds.
In *Inf. Theory Applic. Workshop (ITA)*, San Diego, CA, U.S.A., February 2011.

J. Östman, G. Durisi, E. G. Ström, M. C. Coşkun, and G. Liva.
Short packets over block-memoryless fading channels: Pilot-assisted or noncoherent transmission?
IEEE Trans. Commun., 67(2):1521–1536, Feb. 2019.

Y. Polyanskiy, H. V. Poor, and S. Verdú.
Channel coding rate in the finite blocklength regime.
IEEE Trans. Inf. Theory, 56(5):2307–2359, May 2010.

N. Stolte.
Rekursive Codes mit der Plotkin-Konstruktion und ihre Decodierung.
PhD thesis, TU Darmstadt, 2002.

References III

I. Tal and A. Vardy.
List decoding of polar codes.
IEEE Trans. Inf. Theory, 61(5):2213–2226, May 2015.

P. Yuan, M. C. Coşkun, and G. Kramer.
Polar-coded non-coherent communication.
IEEE Commun. Lett., 25(6):1786–1790, 2021.