Polar Coding for Multi-Terminal Information Theory

Erdal Arıkan

Bilkent University, Ankara, Turkey

26 Nov. 2018
Information Theory Workshop 2018 (ITW2018)
Guangzhou, PRC

Goal

- Discuss polarization as a method for solving coding problems in information theory
- Pose some open problem areas

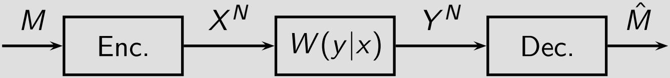
Goal

- Discuss polarization as a method for solving coding problems in information theory
- Pose some open problem areas

Outline

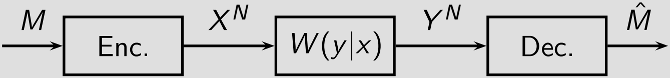
- Basic polar coding
- Polar coding for multi-terminal scenarios

Channel coding



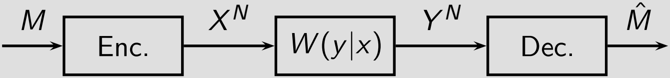
- Reliability measure $P_e = P(\hat{M} \neq M)$
- Encoder maps message M to a codeword X^N of length N
- There are 2^{NR} messages, $H(M) = NR$
- Channel W is memoryless with capacity $C(W)$
- Shannon showed that P_e can be made arbitrarily small by selecting N large enough if (and only if) $R < C(W)$
- Shannon's method was non-constructive

Channel coding



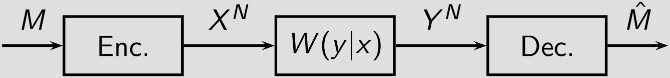
- Reliability measure $P_e = P(\hat{M} \neq M)$
- Encoder maps message M to a codeword X^N of length N
- There are 2^{NR} messages, $H(M) = NR$
- Channel W is memoryless with capacity $C(W)$
- Shannon showed that P_e can be made arbitrarily small by selecting N large enough if (and only if) $R < C(W)$
- Shannon's method was non-constructive

Channel coding



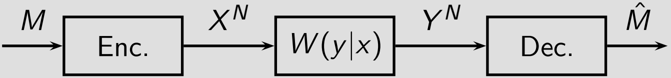
- Reliability measure $P_e = P(\hat{M} \neq M)$
- Encoder maps message M to a codeword X^N of length N
- There are 2^{NR} messages, $H(M) = NR$
- Channel W is memoryless with capacity $C(W)$
- Shannon showed that P_e can be made arbitrarily small by selecting N large enough if (and only if) $R < C(W)$
- Shannon's method was non-constructive

Channel coding



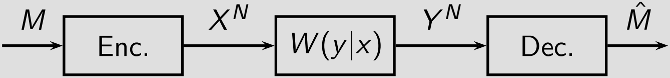
- Reliability measure $P_e = P(\hat{M} \neq M)$
- Encoder maps message M to a codeword X^N of length N
- There are 2^{NR} messages, $H(M) = NR$
- Channel W is memoryless with capacity $C(W)$
- Shannon showed that P_e can be made arbitrarily small by selecting N large enough if (and only if) $R < C(W)$
- Shannon's method was non-constructive

Channel coding



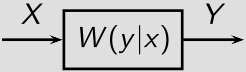
- Reliability measure $P_e = P(\hat{M} \neq M)$
- Encoder maps message M to a codeword X^N of length N
- There are 2^{NR} messages, $H(M) = NR$
- Channel W is memoryless with capacity $C(W)$
- Shannon showed that P_e can be made arbitrarily small by selecting N large enough if (and only if) $R < C(W)$
- Shannon's method was non-constructive

Channel coding



- Reliability measure $P_e = P(\hat{M} \neq M)$
- Encoder maps message M to a codeword X^N of length N
- There are 2^{NR} messages, $H(M) = NR$
- Channel W is memoryless with capacity $C(W)$
- Shannon showed that P_e can be made arbitrarily small by selecting N large enough if (and only if) $R < C(W)$
- Shannon's method was non-constructive

Channel capacity



- $(X, Y) \sim Q(x)W(y|x)$ an input-output pair for W
- Capacity

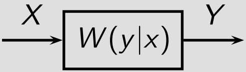
$$C(W) = \max_{Q} I(X;Y),$$

- Symmetric capacity

$$I(W) = I(X;Y),$$

with Q uniform

Channel capacity



- $(X, Y) \sim Q(x)W(y|x)$ an input-output pair for W
- Capacity

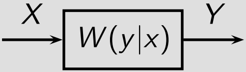
$$C(W) = \max_Q I(X; Y),$$

- Symmetric capacity

$$I(W) = I(X; Y),$$

with Q uniform

Channel capacity



- $(X, Y) \sim Q(x)W(y|x)$ an input-output pair for W
- Capacity

$$C(W) = \max_{Q} I(X; Y),$$

- Symmetric capacity

$$I(W) = I(X, Y),$$

with Q uniform

Polar coding

Polar coding is a block coding method that can achieve symmetric channel capacity $I(W)$ with

- an explicit construction
- encoding and decoding methods of complexity $N \log N$
- reliability $P_e \lesssim 2^{-\sqrt{N}}$

Polar coding

Polar coding is a block coding method that can achieve symmetric channel capacity $I(W)$ with

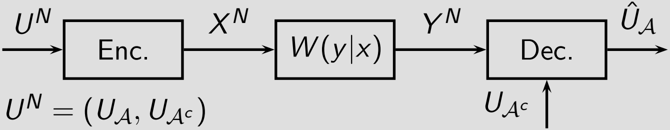
- an explicit construction
- encoding and decoding methods of complexity $N \log N$
- reliability $P_e \lesssim 2^{-\sqrt{N}}$

Polar coding

Polar coding is a block coding method that can achieve symmetric channel capacity $I(W)$ with

- an explicit construction
- encoding and decoding methods of complexity $N \log N$
- reliability $P_e \lesssim 2^{-\sqrt{N}}$

Basic polar coding method

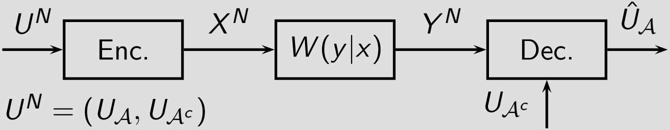


- W is a channel with input alphabet $\{0, 1\}$
- Block length is $N = 2^n$ for some $n \geq 1$
- U^N and X^N are vectors of length N over $\{0, 1\}$
- U^N consists of a message part $U_{\mathcal{A}}$ and a frozen part $U_{\mathcal{A}^c}$
- Encoder computes $X^N = U^N G_N$ where

$$G_N = F^{\otimes n}, \qquad \text{with } F = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

- Decoder estimates $U_{\mathcal{A}}$ with knowledge of $U_{\mathcal{A}^c}$

Basic polar coding method

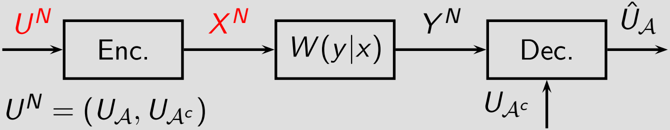


- W is a channel with input alphabet $\{0, 1\}$
- Block length is $N = 2^n$ for some $n \geq 1$
- U^N and X^N are vectors of length N over $\{0, 1\}$
- U^N consists of a message part $U_{\mathcal{A}}$ and a frozen part $U_{\mathcal{A}^c}$
- Encoder computes $X^N = U^N G_N$ where

$$G_N = F^{\otimes n}, \qquad \text{with } F = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

- Decoder estimates $U_{\mathcal{A}}$ with knowledge of $U_{\mathcal{A}^c}$

Basic polar coding method

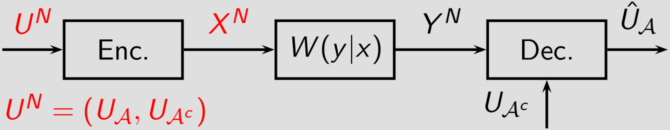


- W is a channel with input alphabet $\{0, 1\}$
- Block length is $N = 2^n$ for some $n \geq 1$
- U^N and X^N are vectors of length N over $\{0, 1\}$
- U^N consists of a message part $U_{\mathcal{A}}$ and a frozen part $U_{\mathcal{A}^c}$
- Encoder computes $X^N = U^N G_N$ where

$$G_N = F^{\otimes n}, \qquad \text{with } F = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

- Decoder estimates $U_{\mathcal{A}}$ with knowledge of $U_{\mathcal{A}^c}$

Basic polar coding method

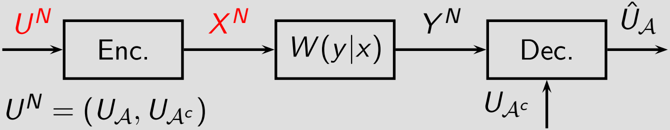


- W is a channel with input alphabet $\{0,1\}$
- Block length is $N = 2^n$ for some $n \geq 1$
- U^N and X^N are vectors of length N over $\{0,1\}$
- U^N consists of a message part $U_{\mathcal{A}}$ and a frozen part $U_{\mathcal{A}^c}$
- Encoder computes $X^N = U^N G_N$ where

$$G_N = F^{\otimes n}, \qquad \text{with } F = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

- Decoder estimates $U_{\mathcal{A}}$ with knowledge of $U_{\mathcal{A}^c}$

Basic polar coding method

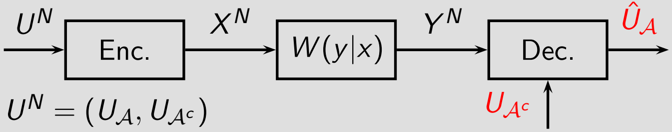


- W is a channel with input alphabet $\{0, 1\}$
- Block length is $N = 2^n$ for some $n \geq 1$
- U^N and X^N are vectors of length N over $\{0, 1\}$
- U^N consists of a message part $U_{\mathcal{A}}$ and a frozen part $U_{\mathcal{A}^c}$
- Encoder computes $X^N = U^N G_N$ where

$$G_N = F^{\otimes n}, \qquad \text{with } F = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

- Decoder estimates $U_{\mathcal{A}}$ with knowledge of $U_{\mathcal{A}^c}$

Basic polar coding method



- W is a channel with input alphabet $\{0,1\}$
- Block length is $N = 2^n$ for some $n \geq 1$
- U^N and X^N are vectors of length N over $\{0,1\}$
- U^N consists of a message part $U_{\mathcal{A}}$ and a frozen part $U_{\mathcal{A}^c}$
- Encoder computes $X^N = U^N G_N$ where

$$G_N = F^{\otimes n}, \qquad \text{with } F = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

- Decoder estimates $U_{\mathcal{A}}$ with knowledge of $U_{\mathcal{A}^c}$

Polar coding

Polar coding is a block coding method that can achieve symmetric channel capacity $I(W)$ with

- an explicit construction
- encoding and decoding methods of complexity $N \log N$
- reliability $P_e \lesssim 2^{-\sqrt{N}}$

Polar coding

Polar coding is a block coding method that can achieve symmetric channel capacity $I(W)$ with

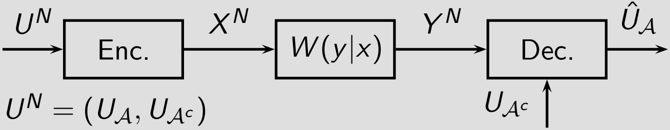
- an explicit construction
- encoding and decoding methods of complexity $N \log N$
- reliability $P_e \lesssim 2^{-\sqrt{N}}$

Polar coding

Polar coding is a block coding method that can achieve symmetric channel capacity $I(W)$ with

- an explicit construction
- encoding and decoding methods of complexity $N \log N$
- reliability $P_e \lesssim 2^{-\sqrt{N}}$

Polarization



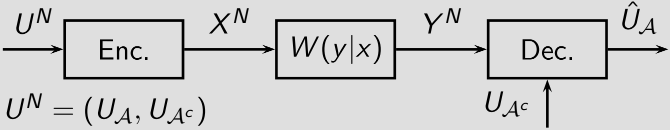
- Polarization refers to the possibility of selecting a set $\mathcal{A}$ of cardinality $|\mathcal{A}| = NI(W) - o(N)$ with $o(N)/N \to 0$ so that

$$H(U_i|Y^N, U^{i-1}) = O(2^{-N^{0.49}}), \quad \text{for all } i \in \mathcal{A},$$

  where $U^{i-1} = (U_1, \ldots, U_{i-1})$.

- It follows that polar codes achieve symmetric capacity $I(W)$ under successive cancellation (SC) decoding

Polarization



- Polarization refers to the possibility of selecting a set $\mathcal{A}$ of cardinality $|\mathcal{A}| = NI(W) - o(N)$ with $o(N)/N \to 0$ so that

$$H(U_i|Y^N, U^{i-1}) = O(2^{-N^{0.49}}), \quad \text{for all } i \in \mathcal{A},$$

  where $U^{i-1} = (U_1, \ldots, U_{i-1})$.
- It follows that polar codes achieve symmetric capacity $I(W)$ under successive cancellation (SC) decoding

Limitations of basic polar coding

Before polar codes can be considered as a candidate for solving all types of coding problems, a number of issues need to be addressed.

- Universality:
  A polar code good for one channel may not be good for another channel of the same or even higher capacity
- Capacity vs. symmetric capacity:
  Polar codes achieve the symmetric capacity $I(W)$ not the true capacity $C(W)$
- Channels with non-binary inputs:
  How to generalize the polar code construction to channels with non-binary input alphabets

Limitations of basic polar coding

Before polar codes can be considered as a candidate for solving all types of coding problems, a number of issues need to be addressed.

- Universality:
  A polar code good for one channel may not be good for another channel of the same or even higher capacity
- Capacity vs. symmetric capacity:
  Polar codes achieve the symmetric capacity $I(W)$ not the true capacity $C(W)$
- Channels with non-binary inputs:
  How to generalize the polar code construction to channels with non-binary input alphabets

Limitations of basic polar coding

Before polar codes can be considered as a candidate for solving all types of coding problems, a number of issues need to be addressed.

- Universality:
  A polar code good for one channel may not be good for another channel of the same or even higher capacity
- Capacity vs. symmetric capacity:
  Polar codes achieve the symmetric capacity $I(W)$ not the true capacity $C(W)$
- Channels with non-binary inputs:
  How to generalize the polar code construction to channels with non-binary input alphabets

Universality issue

- If W is degraded wrt W', then a polar code good for W is also good for W'.
- But in general, a polar code good for W may be very bad for W' even if $I(W') > I(W)$.
- Şaşoğlu (2010) shows that if a polar code is designed for a Binary Symmetric Channel (BSC) W, then it is at least as good for any other channel W' with $I(W') \geq I(W)$ under ML decoding.
- So, the non-universal nature of polar coding is an artifact of the successive cancellation decoding method.
- Hassani and Urbanke (2014), Ye and Barg (2014), and Şaşoğlu and Wang (2016) show that it is possible to extend polar codes so that they retain low-complexity encoding and decoding properties while achieving symmetric capacity.

Universality issue

- If W is degraded wrt W', then a polar code good for W is also good for W'.
- But in general, a polar code good for W may be very bad for W' even if $I(W') > I(W)$.
- Şaşoğlu (2010) shows that if a polar code is designed for a Binary Symmetric Channel (BSC) W, then it is at least as good for any other channel W' with $I(W') \geq I(W)$ under ML decoding.
- So, the non-universal nature of polar coding is an artifact of the successive cancellation decoding method.
- Hassani and Urbanke (2014), Ye and Barg (2014), and Şaşoğlu and Wang (2016) show that it is possible to extend polar codes so that they retain low-complexity encoding and decoding properties while achieving symmetric capacity.

Universality issue

- If W is degraded wrt W', then a polar code good for W is also good for W'.
- But in general, a polar code good for W may be very bad for W' even if $I(W') > I(W)$.
- Şaşoğlu (2010) shows that if a polar code is designed for a Binary Symmetric Channel (BSC) W, then it is at least as good for any other channel W' with $I(W') \geq I(W)$ under ML decoding.
- So, the non-universal nature of polar coding is an artifact of the successive cancellation decoding method.
- Hassani and Urbanke (2014), Ye and Barg (2014), and Şaşoğlu and Wang (2016) show that it is possible to extend polar codes so that they retain low-complexity encoding and decoding properties while achieving symmetric capacity.

Universality issue

- If W is degraded wrt W', then a polar code good for W is also good for W'.
- But in general, a polar code good for W may be very bad for W' even if $I(W') > I(W)$.
- Şaşoğlu (2010) shows that if a polar code is designed for a Binary Symmetric Channel (BSC) W, then it is at least as good for any other channel W' with $I(W') \geq I(W)$ under ML decoding.
- So, the non-universal nature of polar coding is an artifact of the successive cancellation decoding method.
- Hassani and Urbanke (2014), Ye and Barg (2014), and Şaşoğlu and Wang (2016) show that it is possible to extend polar codes so that they retain low-complexity encoding and decoding properties while achieving symmetric capacity.

Universality issue

- If W is degraded wrt W', then a polar code good for W is also good for W'.
- But in general, a polar code good for W may be very bad for W' even if $I(W') > I(W)$.
- Şaşoğlu (2010) shows that if a polar code is designed for a Binary Symmetric Channel (BSC) W, then it is at least as good for any other channel W' with $I(W') \geq I(W)$ under ML decoding.
- So, the non-universal nature of polar coding is an artifact of the successive cancellation decoding method.
- Hassani and Urbanke (2014), Ye and Barg (2014), and Şaşoğlu and Wang (2016) show that it is possible to extend polar codes so that they retain low-complexity encoding and decoding properties while achieving symmetric capacity.

Symmetric capacity issue

- Polar codes in their basic form achieve symmetric capacity $I(W)$ not the true capacity $C(W)$.
- All linear codes suffer from this defect and some general remedies exist.
- Korada (2009) proposed constructing polar codes over an extended alphabet and mapping extended alphabet to the binary alphabet to effect the desired distribution.
- Honda and Yamamoto (2012) combined source polarization and channel polarization and provided a solution fully within the framework of polar coding.

Symmetric capacity issue

- Polar codes in their basic form achieve symmetric capacity $I(W)$ not the true capacity $C(W)$.
- All linear codes suffer from this defect and some general remedies exist.
- Korada (2009) proposed constructing polar codes over an extended alphabet and mapping extended alphabet to the binary alphabet to effect the desired distribution.
- Honda and Yamamoto (2012) combined source polarization and channel polarization and provided a solution fully within the framework of polar coding.

Symmetric capacity issue

- Polar codes in their basic form achieve symmetric capacity $I(W)$ not the true capacity $C(W)$.
- All linear codes suffer from this defect and some general remedies exist.
- Korada (2009) proposed constructing polar codes over an extended alphabet and mapping extended alphabet to the binary alphabet to effect the desired distribution.
- Honda and Yamamoto (2012) combined source polarization and channel polarization and provided a solution fully within the framework of polar coding.

Symmetric capacity issue

- Polar codes in their basic form achieve symmetric capacity $I(W)$ not the true capacity $C(W)$.
- All linear codes suffer from this defect and some general remedies exist.
- Korada (2009) proposed constructing polar codes over an extended alphabet and mapping extended alphabet to the binary alphabet to effect the desired distribution.
- Honda and Yamamoto (2012) combined source polarization and channel polarization and provided a solution fully within the framework of polar coding.

Non-binary polar coding issue

- Şaşoğlu et al. (2009) showed that the standard 2-by-2 kernel F achieves polarization for q-ary channels for any prime q, but may fail if q is composite.
- Mori and Tanaka (2010) showed that for $q = p^m$, $p > 2$, the kernel

$$F = \begin{bmatrix} 1 & 0 \\ \gamma & 1 \end{bmatrix}, \quad \gamma \neq 0, 1,$$

  polarizes channels with q-ary inputs.
- For channels for which the input alphabet size is not a prime power, there exist multi-level polarization approaches as discussed in Şaşoğlu et al. (2009), Park and Barg (2011,2012, 2013), and Sahebi and Pradhan (2011).

Non-binary polar coding issue

- Şaşoğlu et al. (2009) showed that the standard 2-by-2 kernel F achieves polarization for q-ary channels for any prime q, but may fail if q is composite.
- Mori and Tanaka (2010) showed that for $q = p^m$, $p > 2$, the kernel

$$F = \begin{bmatrix} 1 & 0 \\ \gamma & 1 \end{bmatrix}, \quad \gamma \neq 0, 1,$$

polarizes channels with q-ary inputs.
- For channels for which the input alphabet size is not a prime power, there exist multi-level polarization approaches as discussed in Şaşoğlu et al. (2009), Park and Barg (2011,2012, 2013), and Sahebi and Pradhan (2011).

Non-binary polar coding issue

- Şaşoğlu et al. (2009) showed that the standard 2-by-2 kernel F achieves polarization for q-ary channels for any prime q, but may fail if q is composite.
- Mori and Tanaka (2010) showed that for $q = p^m$, $p > 2$, the kernel

$$F = \begin{bmatrix} 1 & 0 \\ \gamma & 1 \end{bmatrix}, \quad \gamma \neq 0, 1,$$

  polarizes channels with q-ary inputs.
- For channels for which the input alphabet size is not a prime power, there exist multi-level polarization approaches as discussed in Şaşoğlu et al. (2009), Park and Barg (2011,2012, 2013), and Sahebi and Pradhan (2011).

Polarization theory and methods

Topic	References
Basic polarization	[28], [4], [1], [2]
General binary kernels	[29], [29], [30], [31], [32]
Non-binary kernels	[33], [30], [34], [35], [36], [37], [38]
Multi-level polarization	[39], [40], [41], [42],[43]
Universal polar codes	[44], [45], [46], [47]
Rate of polarization	[48], [49], [29]
Finite-length performance	[50], [51], [52],[53] ,[54], [55]
Code construction	[56], [57], [58], [59], [51]
Input distribution opt.	[2], [33], [60]

Recap of first part

Given a target rate R and an integer $q \geq 2$, there exist q-ary polar codes with length N and rate $> R$ such that, on any channel W with $C(W) > R$,

- encoding and decoding complexities are bounded by $O(N \log N)$,
- the probability of error is bounded by $O(e^{-N^{\beta}})$ for some $\beta > 0$ that depends on the channel W.

Recap of first part

Given a target rate R and an integer $q \geq 2$, there exist q-ary polar codes with length N and rate $> R$ such that, on any channel W with $C(W) > R$,

- encoding and decoding complexities are bounded by $O(N \log N)$,
- the probability of error is bounded by $O(e^{-N^\beta})$ for some $\beta > 0$ that depends on the channel W.

Outline

- Basic polar coding
- Polar coding for multi-terminal scenarios

A new approach

- The references cited use standard low-complexity properties of polar coding
- In each case, difficulty lies mainly in resolving complications arising from the need for universality, non-uniform input distributions, non-binary alphabets, etc.
- To explore the limits of polarization approach, it makes sense to use stronger polarizers unconstrained by complexity considerations.
- The primary question of interest is to see if polarization without complexity limits is as strong as random-coding/typical-set approach for proving achievability results.

A new approach

- The references cited use standard low-complexity properties of polar coding
- In each case, difficulty lies mainly in resolving complications arising from the need for universality, non-uniform input distributions, non-binary alphabets, etc.
- To explore the limits of polarization approach, it makes sense to use stronger polarizers unconstrained by complexity considerations.
- The primary question of interest is to see if polarization without complexity limits is as strong as random-coding/typical-set approach for proving achievability results.

A new approach

- The references cited use standard low-complexity properties of polar coding
- In each case, difficulty lies mainly in resolving complications arising from the need for universality, non-uniform input distributions, non-binary alphabets, etc.
- To explore the limits of polarization approach, it makes sense to use stronger polarizers unconstrained by complexity considerations.
- The primary question of interest is to see if polarization without complexity limits is as strong as random-coding/typical-set approach for proving achievability results.

A new approach

- The references cited use standard low-complexity properties of polar coding
- In each case, difficulty lies mainly in resolving complications arising from the need for universality, non-uniform input distributions, non-binary alphabets, etc.
- To explore the limits of polarization approach, it makes sense to use stronger polarizers unconstrained by complexity considerations.
- The primary question of interest is to see if polarization without complexity limits is as strong as random-coding/typical-set approach for proving achievability results.

Random polarizers

Definition

A random polarizer is a one-to-one transformation $G_N : \mathcal{X}^N \to \mathcal{X}^N$ chosen at random (in equiprobable manner) from the class of all one-to-one functions from $\mathcal{X}^n$ to $\mathcal{X}^n$.

Polarization profile of a single random variable X

- Let X^N consist of N iid copies of a random variable over a finite alphabet $\mathcal{X}$.
- Let G_N be a random polarizer
- Let $U^N = G_N(X^N)$ be the transform of X^N
- The polarization profile of X^N is defined as the plot of $H(U_i|U^{i-1})$ vs. the normalized index $\hat{i} = i \log|\mathcal{X}|/N$.
- Polarization profile looks as follows:

Polarization profile of a single random variable X

- Let X^N consist of N iid copies of a random variable over a finite alphabet $\mathcal{X}$.
- Let G_N be a random polarizer
- Let $U^N = G_N(X^N)$ be the transform of X^N
- The polarization profile of X^N is defined as the plot of $H(U_i|U^{i-1})$ vs. the normalized index $\hat{i} = i \log|\mathcal{X}|/N$.
- Polarization profile looks as follows:

Polarization profile of a single random variable X

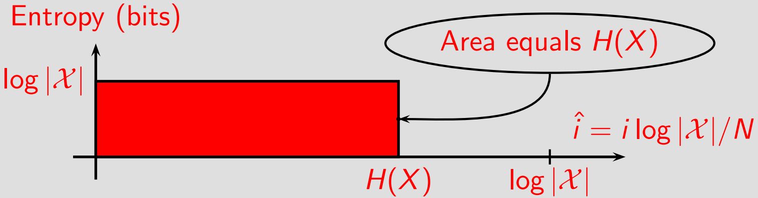
- Let X^N consist of N iid copies of a random variable over a finite alphabet $\mathcal{X}$.
- Let G_N be a random polarizer
- Let $U^N = G_N(X^N)$ be the transform of X^N
- The polarization profile of X^N is defined as the plot of $H(U_i|U^{i-1})$ vs. the normalized index $\hat{i} = i \log|\mathcal{X}|/N$.
- Polarization profile looks as follows:

Polarization profile of a single random variable X

- Let X^N consist of N iid copies of a random variable over a finite alphabet $\mathcal{X}$.
- Let G_N be a random polarizer
- Let $U^N = G_N(X^N)$ be the transform of X^N
- The polarization profile of X^N is defined as the plot of $H(U_i|U^{i-1})$ vs. the normalized index $\hat{i} = i \log|\mathcal{X}|/N$.
- Polarization profile looks as follows:

Polarization profile of a single random variable X

- Let X^N consist of N iid copies of a random variable over a finite alphabet $\mathcal{X}$.
- Let G_N be a random polarizer
- Let $U^N = G_N(X^N)$ be the transform of X^N
- The polarization profile of X^N is defined as the plot of $H(U_i|U^{i-1})$ vs. the normalized index $\hat{i} = i \log|\mathcal{X}|/N$.
- Polarization profile looks as follows:



Polarization profile of X given Y

- Let (X^N, Y^N) consist of iid copies of (X, Y) over $\mathcal{X} \times \mathcal{Y}$.
- Let $G_N : \mathcal{X}^N \to \mathcal{X}^N$ be a random polarizer
- Let $U^N = G_N(X^N)$ be the transform of X^N
- The polarization profile of X^N given Y^N is defined as the plot of $H(U_i|Y^N, U^{i-1})$ vs. the normalized index $\hat{i} = i \log |\mathcal{X}|/N$.
- Typical polarization profile:

Polarization profile of X given Y

- Let (X^N, Y^N) consist of iid copies of (X, Y) over $\mathcal{X} \times \mathcal{Y}$.
- Let $G_N : \mathcal{X}^N \to \mathcal{X}^N$ be a random polarizer
- Let $U^N = G_N(X^N)$ be the transform of X^N
- The polarization profile of X^N given Y^N is defined as the plot of $H(U_i|Y^N, U^{i-1})$ vs. the normalized index $\hat{i} = i \log|\mathcal{X}|/N$.
- Typical polarization profile:

Polarization profile of X given Y

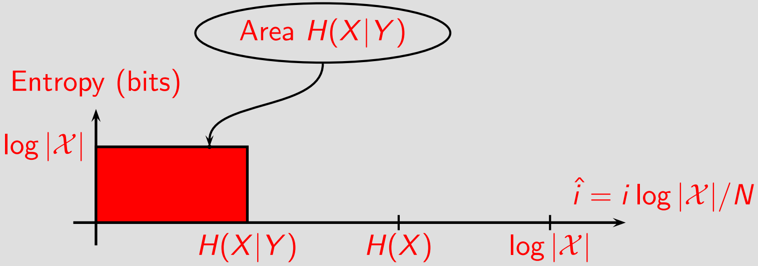
- Let (X^N, Y^N) consist of iid copies of (X, Y) over $\mathcal{X} \times \mathcal{Y}$.
- Let $G_N : \mathcal{X}^N \to \mathcal{X}^N$ be a random polarizer
- Let $U^N = G_N(X^N)$ be the transform of X^N
- The polarization profile of X^N given Y^N is defined as the plot of $H(U_i|Y^N, U^{i-1})$ vs. the normalized index $\hat{i} = i \log|\mathcal{X}|/N$.
- Typical polarization profile:

Polarization profile of X given Y

- Let (X^N, Y^N) consist of iid copies of (X, Y) over $\mathcal{X} \times \mathcal{Y}$.
- Let $G_N : \mathcal{X}^N \to \mathcal{X}^N$ be a random polarizer
- Let $U^N = G_N(X^N)$ be the transform of X^N
- The polarization profile of X^N given Y^N is defined as the plot of $H(U_i|Y^N, U^{i-1})$ vs. the normalized index $\hat{i} = i \log|\mathcal{X}|/N$.
- Typical polarization profile:

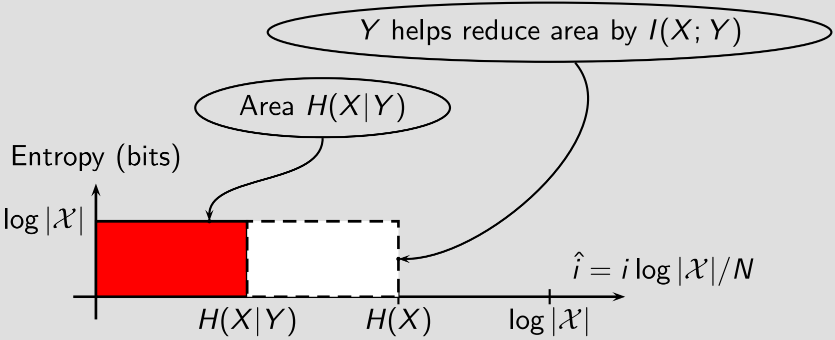
Polarization profile of X given Y

- Let (X^N, Y^N) consist of iid copies of (X, Y) over $\mathcal{X} \times \mathcal{Y}$.
- Let $G_N : \mathcal{X}^N \to \mathcal{X}^N$ be a random polarizer
- Let $U^N = G_N(X^N)$ be the transform of X^N
- The polarization profile of X^N given Y^N is defined as the plot of $H(U_i|Y^N, U^{i-1})$ vs. the normalized index $\hat{i} = i \log |\mathcal{X}|/N$.
- Typical polarization profile:

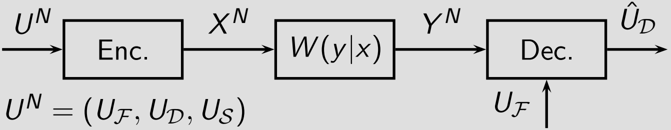


Polarization profile of X given Y

- Let (X^N, Y^N) consist of iid copies of (X, Y) over $\mathcal{X} \times \mathcal{Y}$.
- Let $G_N : \mathcal{X}^N \to \mathcal{X}^N$ be a random polarizer
- Let $U^N = G_N(X^N)$ be the transform of X^N
- The polarization profile of X^N given Y^N is defined as the plot of $H(U_i|Y^N, U^{i-1})$ vs. the normalized index $\hat{i} = i \log|\mathcal{X}|/N$.
- Typical polarization profile:

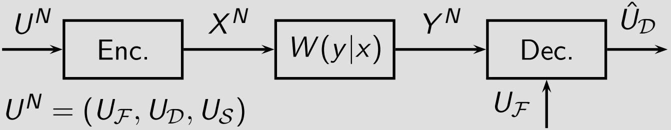


Achievability of channel capacity by random polarization



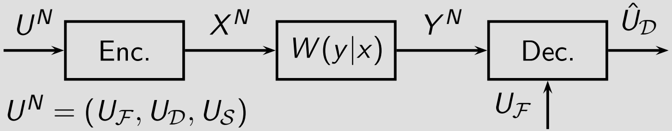
- W is a channel with an arbitrary finite input alphabet $\mathcal{X}$
- Capacity-achieving input distribution is Q
- $\mathcal{F}$: index set of symbols available to the decoder
- $\mathcal{D}$: index set of data symbols
- $\mathcal{S}$: index set of symbols computed by the encoder from $(\mathcal{F}, \mathcal{D})$
- Decoder estimates $U_{\mathcal{D}}$ given $(U_{\mathcal{F}}, Y^N)$
- Direct part of channel-coding theorem follows by choosing $|\mathcal{F}| = NH(X|Y) + o(N)$ and $|\mathcal{D}| = NI(X;Y) - o(N)$.

Achievability of channel capacity by random polarization



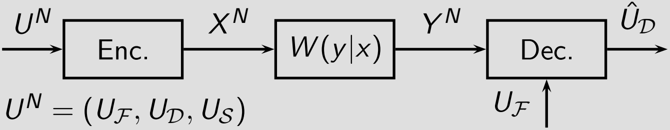
- W is a channel with an arbitrary finite input alphabet $\mathcal{X}$
- Capacity-achieving input distribution is Q
- $\mathcal{F}$: index set of symbols available to the decoder
- $\mathcal{D}$: index set of data symbols
- $\mathcal{S}$: index set of symbols computed by the encoder from $(\mathcal{F}, \mathcal{D})$
- Decoder estimates $U_{\mathcal{D}}$ given $(U_{\mathcal{F}}, Y^N)$
- Direct part of channel-coding theorem follows by choosing $|\mathcal{F}| = NH(X|Y) + o(N)$ and $|\mathcal{D}| = NI(X;Y) - o(N)$.

Achievability of channel capacity by random polarization



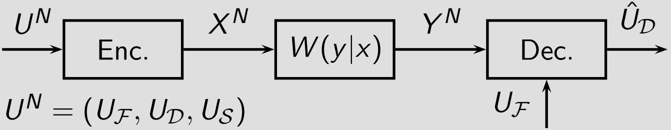
- W is a channel with an arbitrary finite input alphabet $\mathcal{X}$
- Capacity-achieving input distribution is Q
- $\mathcal{F}$: index set of symbols available to the decoder
- $\mathcal{D}$: index set of data symbols
- $\mathcal{S}$: index set of symbols computed by the encoder from $(\mathcal{F}, \mathcal{D})$
- Decoder estimates $U_{\mathcal{D}}$ given $(U_{\mathcal{F}}, Y^N)$
- Direct part of channel-coding theorem follows by choosing $|\mathcal{F}| = NH(X|Y) + o(N)$ and $|\mathcal{D}| = NI(X;Y) - o(N)$.

Achievability of channel capacity by random polarization



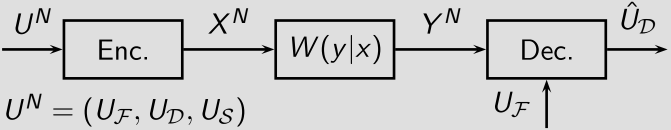
- W is a channel with an arbitrary finite input alphabet $\mathcal{X}$
- Capacity-achieving input distribution is Q
- $\mathcal{F}$: index set of symbols available to the decoder
- $\mathcal{D}$: index set of data symbols
- $\mathcal{S}$: index set of symbols computed by the encoder from $(\mathcal{F}, \mathcal{D})$
- Decoder estimates $U_{\mathcal{D}}$ given $(U_{\mathcal{F}}, Y^N)$
- Direct part of channel-coding theorem follows by choosing $|\mathcal{F}| = NH(X|Y) + o(N)$ and $|\mathcal{D}| = NI(X;Y) - o(N)$.

Achievability of channel capacity by random polarization



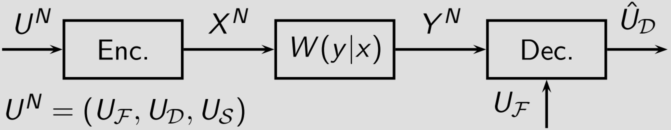
- W is a channel with an arbitrary finite input alphabet $\mathcal{X}$
- Capacity-achieving input distribution is Q
- $\mathcal{F}$: index set of symbols available to the decoder
- $\mathcal{D}$: index set of data symbols
- $\mathcal{S}$: index set of symbols computed by the encoder from $(\mathcal{F}, \mathcal{D})$
- Decoder estimates $U_{\mathcal{D}}$ given $(U_{\mathcal{F}}, Y^N)$
- Direct part of channel-coding theorem follows by choosing $|\mathcal{F}| = NH(X|Y) + o(N)$ and $|\mathcal{D}| = NI(X;Y) - o(N)$.

Achievability of channel capacity by random polarization



- W is a channel with an arbitrary finite input alphabet $\mathcal{X}$
- Capacity-achieving input distribution is Q
- $\mathcal{F}$: index set of symbols available to the decoder
- $\mathcal{D}$: index set of data symbols
- $\mathcal{S}$: index set of symbols computed by the encoder from $(\mathcal{F}, \mathcal{D})$
- Decoder estimates $U_{\mathcal{D}}$ given $(U_{\mathcal{F}}, Y^N)$
- Direct part of channel-coding theorem follows by choosing $|\mathcal{F}| = NH(X|Y) + o(N)$ and $|\mathcal{D}| = NI(X;Y) - o(N)$.

Achievability of channel capacity by random polarization



- W is a channel with an arbitrary finite input alphabet $\mathcal{X}$
- Capacity-achieving input distribution is Q
- $\mathcal{F}$: index set of symbols available to the decoder
- $\mathcal{D}$: index set of data symbols
- $\mathcal{S}$: index set of symbols computed by the encoder from $(\mathcal{F}, \mathcal{D})$
- Decoder estimates $U_{\mathcal{D}}$ given $(U_{\mathcal{F}}, Y^N)$
- Direct part of channel-coding theorem follows by choosing $|\mathcal{F}| = NH(X|Y) + o(N)$ and $|\mathcal{D}| = NI(X;Y) - o(N)$.

Random polarization for two random variables (X_1, X_2)

- Let (X_1^N, X_2^N) consist of iid copies of the pair (X_1, X_2) with an arbitrary distribution over $\mathcal{X}_1 \times \mathcal{X}_2$.
- Let $G_N^{(1)}$ and $G_N^{(2)}$ be independent random polarizers for X_1^N and X_2^N, respectively.
- Let $\overline{X}_1 = G_N^{(1)}(X_1^N)$ and $\overline{X}_2 = G_N^{(2)}(X_2^N)$ be the transforms.
- The polarization profile of (X_1^N, X_2^N) is defined as the plot of

$$H(\overline{X}_{1,i+1}|\overline{X}_1^i, \overline{X}_2^j) \quad and \quad H(\overline{X}_{2,j+1}|\overline{X}_1^i, \overline{X}_2^j)$$

as a function of $\hat{i} = i \log|\mathcal{X}_1|/N$ and $\hat{j} = j \log|\mathcal{X}_2|/N$.

Random polarization for two random variables (X_1, X_2)

- Let (X_1^N, X_2^N) consist of iid copies of the pair (X_1, X_2) with an arbitrary distribution over $\mathcal{X}_1 \times \mathcal{X}_2$.
- Let $G_N^{(1)}$ and $G_N^{(2)}$ be independent random polarizers for X_1^N and X_2^N, respectively.
- Let $\overline{X}_1 = G_N^{(1)}(X_1^N)$ and $\overline{X}_2 = G_N^{(2)}(X_2^N)$ be the transforms.
- The polarization profile of (X_1^N, X_2^N) is defined as the plot of

$$H(\overline{X}_{1,i+1}|\overline{X}_1^i, \overline{X}_2^j) \quad and \quad H(\overline{X}_{2,j+1}|\overline{X}_1^i, \overline{X}_2^j)$$

as a function of $\hat{i} = i \log|\mathcal{X}_1|/N$ and $\hat{j} = j \log|\mathcal{X}_2|/N$.

Random polarization for two random variables (X_1, X_2)

- Let (X_1^N, X_2^N) consist of iid copies of the pair (X_1, X_2) with an arbitrary distribution over $\mathcal{X}_1 \times \mathcal{X}_2$.
- Let $G_N^{(1)}$ and $G_N^{(2)}$ be independent random polarizers for X_1^N and X_2^N, respectively.
- Let $\overline{X}_1 = G_N^{(1)}(X_1^N)$ and $\overline{X}_2 = G_N^{(2)}(X_2^N)$ be the transforms.
- The polarization profile of (X_1^N, X_2^N) is defined as the plot of

$$H(\overline{X}_{1,i+1}|\overline{X}_1^i, \overline{X}_2^j) \quad and \quad H(\overline{X}_{2,j+1}|\overline{X}_1^i, \overline{X}_2^j)$$

as a function of $\hat{i} = i \log|\mathcal{X}_1|/N$ and $\hat{j} = j \log|\mathcal{X}_2|/N$.

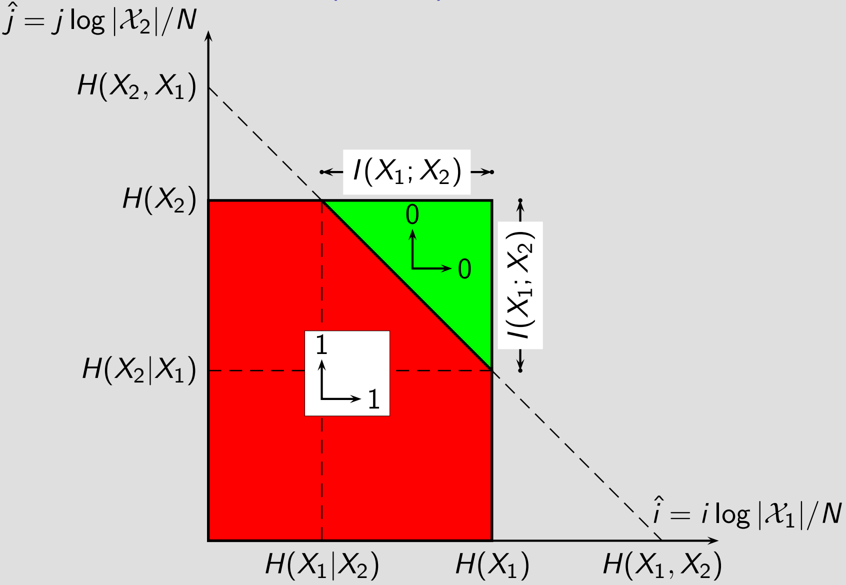
Random polarization for two random variables (X_1, X_2)

- Let (X_1^N, X_2^N) consist of iid copies of the pair (X_1, X_2) with an arbitrary distribution over $\mathcal{X}_1 \times \mathcal{X}_2$.
- Let $G_N^{(1)}$ and $G_N^{(2)}$ be independent random polarizers for X_1^N and X_2^N, respectively.
- Let $\overline{X}_1 = G_N^{(1)}(X_1^N)$ and $\overline{X}_2 = G_N^{(2)}(X_2^N)$ be the transforms.
- The polarization profile of (X_1^N, X_2^N) is defined as the plot of

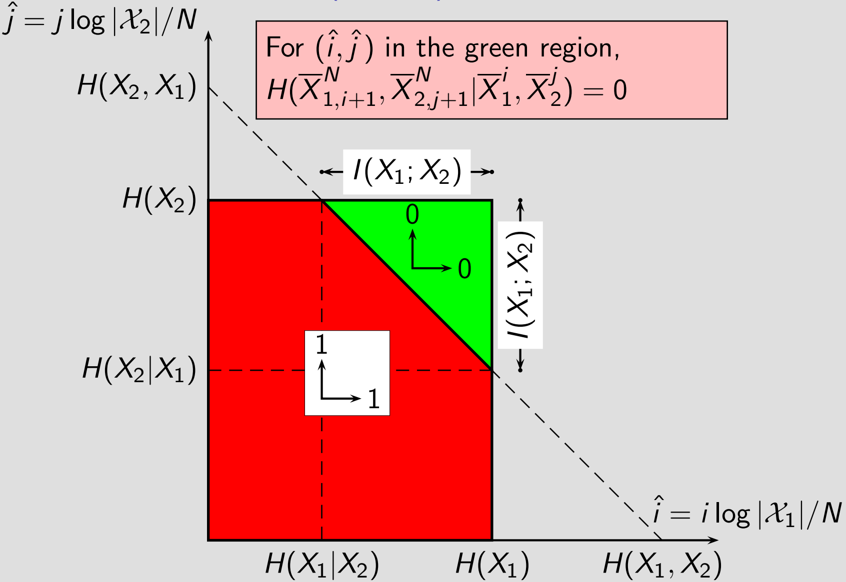
$$H(\overline{X}_{1,i+1}|\overline{X}_1^i, \overline{X}_2^j) \quad and \quad H(\overline{X}_{2,j+1}|\overline{X}_1^i, \overline{X}_2^j)$$

as a function of $\hat{i} = i \log |\mathcal{X}_1|/N$ and $\hat{j} = j \log |\mathcal{X}_2|/N$.

Polarization profile of (X_1, X_2)

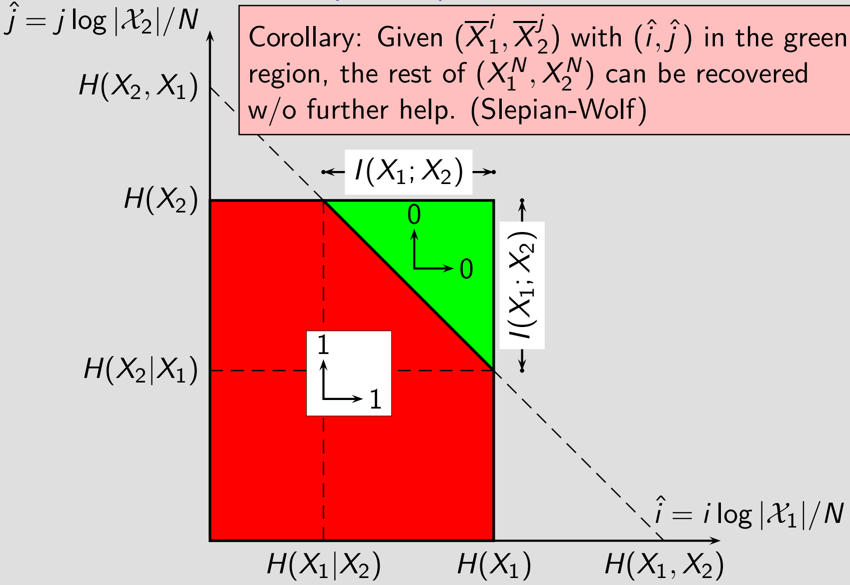


Polarization profile of (X_1, X_2)

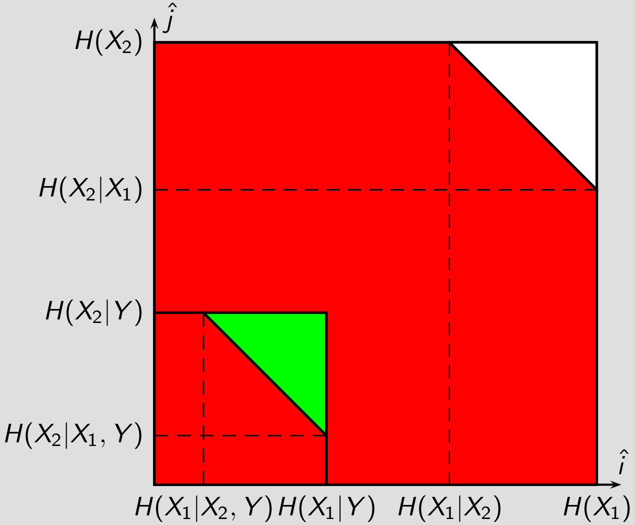


Polarization profile of (X_1, X_2)

Corollary: Given $(\overline{X}_1^i, \overline{X}_2^j)$ with $(\hat{i}, \hat{j})$ in the green region, the rest of (X_1^N, X_2^N) can be recovered w/o further help. (Slepian-Wolf)

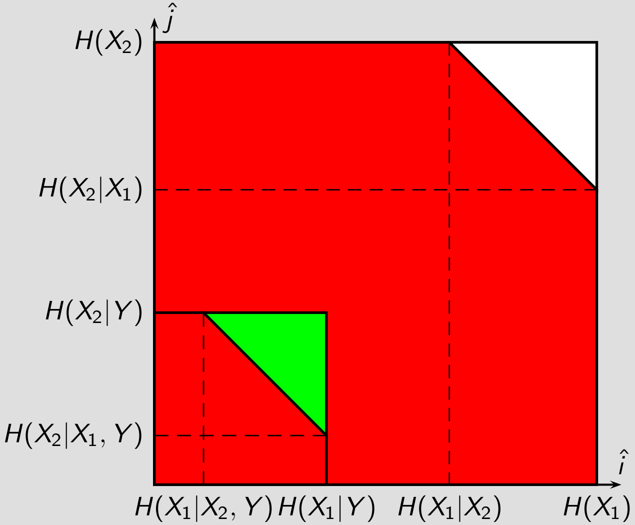


Polarization profile of (X_1, X_2) given Y



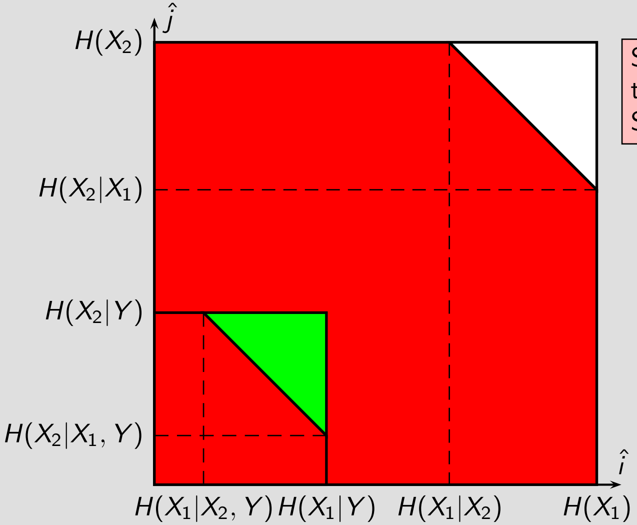
Polarization profile of (X_1, X_2) given Y

Starting from anywhere in the green region, a decoder can reach the top right corner.



Polarization profile of (X_1, X_2) given Y

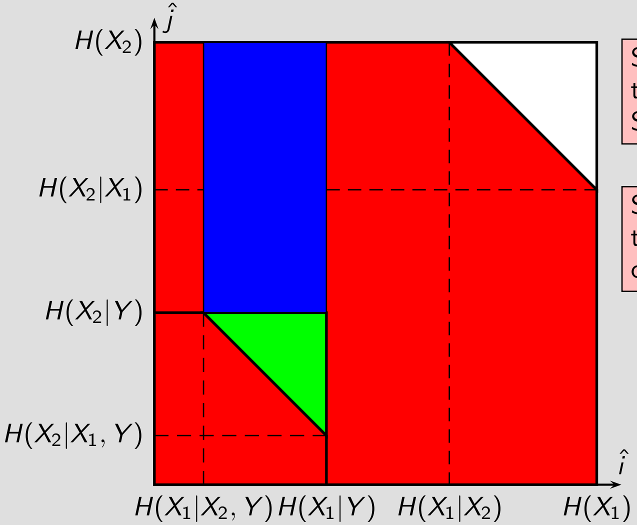
Starting from anywhere in the green region, a decoder can reach the top right corner.



Step 1: Green is traversed by Slepian-Wolf.

Polarization profile of (X_1, X_2) given Y

Starting from anywhere in the green region, a decoder can reach the top right corner.

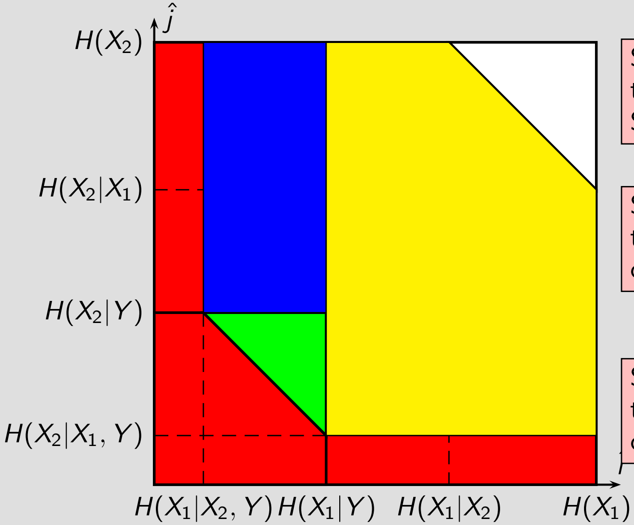


Step 1: Green is traversed by Slepian-Wolf.

Step 2: Blue is traversed by single-user channel coding.

Polarization profile of (X_1, X_2) given Y

Starting from anywhere in the green region, a decoder can reach the top right corner.

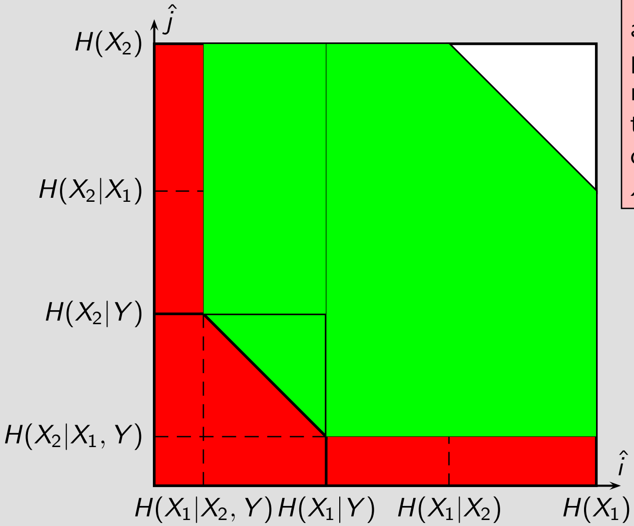


Step 1: Green is traversed by Slepian-Wolf.

Step 2: Blue is traversed by single-user channel coding.

Step 3: Yellow is traversed by single-user channel coding.

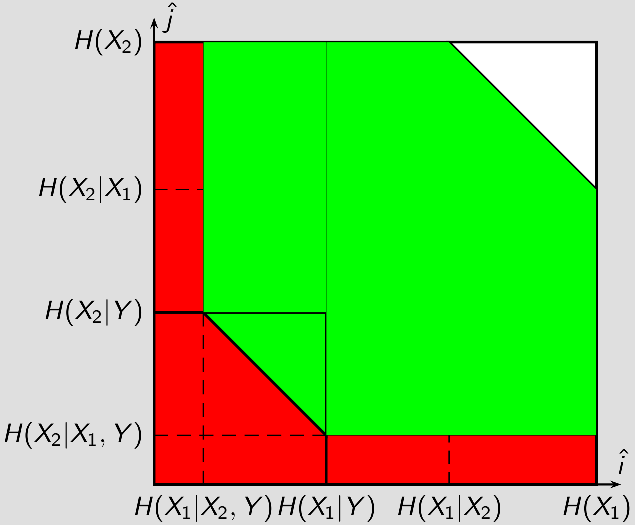
Polarization profile of (X_1, X_2) given Y



Top right corner is accessible from any point in the green region. The white triangle corresponds to common information in X_1 and X_2.

Polarization profile of (X_1, X_2) given Y

Achievable rates for a channel $(X_1, X_2) \to Y$ are obtained as
$R_1 < I(X_1; Y|X_2)$, $R_2 < I(X_2; Y|X_1)$, and $R_1 + R_2 < I(X_1, X_2; Y) - I(X_1; X_2)$



Summary

- We described a random polarization method for proving source and channel coding theorems.
- Some cases have been discussed for demonstration of the method.
- Questions remain whether the polarization method described here is as strong as the standard random-coding method.

Summary

- We described a random polarization method for proving source and channel coding theorems.
- Some cases have been discussed for demonstration of the method.
- Questions remain whether the polarization method described here is as strong as the standard random-coding method.

Summary

- We described a random polarization method for proving source and channel coding theorems.
- Some cases have been discussed for demonstration of the method.
- Questions remain whether the polarization method described here is as strong as the standard random-coding method.

Thank you!

References I

[1] N. Hussami, S. B. Korada, and R. Urbanke, "Performance of polar codes for channel and source coding," in *2009 IEEE International Symposium on Information Theory*, Jun. 2009, pp. 1488–1492.

[2] S. B. Korada, "Polar codes for channel and source coding," Ph.D. dissertation, ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE, 2009. [Online]. Available: http://biblion.epfl.ch/EPFL/theses/2009/4461/EPFL_TH4461.pdf

[3] S. B. Korada and R. Urbanke, "Polar codes for Slepian-Wolf, Wyner-Ziv, and Gelfand-Pinsker," in *2010 IEEE Information Theory Workshop (ITW)*. IEEE, Jan. 2010, pp. 1–5.

[4] E. Arıkan, "Source polarization," in *2010 IEEE International Symposium on Information Theory Proceedings (ISIT)*. IEEE, Jun. 2010, pp. 899–903.

[5] ——, "Polar coding for the Slepian-Wolf problem based on monotone chain rules," in *2012 IEEE International Symposium on Information Theory Proceedings*, Jul. 2012, pp. 566–570.

[6] S. Onay, "Successive cancellation decoding of polar codes for the two-user binary-input MAC," in *2013 IEEE International Symposium on Information Theory*. Istanbul, Turkey: IEEE, Jul. 2013, pp. 1122–1126. [Online]. Available: http://ieeexplore.ieee.org/document/6620401/

References II

[7] S. Salamatian, M. Médard, and E. Telatar, "A Successive Description property of Monotone-Chain Polar Codes for Slepian-Wolf coding," in *2015 IEEE International Symposium on Information Theory (ISIT)*, Jun. 2015, pp. 1522–1526.

[8] E. Sasoglu, E. Telatar, and E. Yeh, "Polar codes for the two-user multiple-access channel," *1006.4255*, Jun. 2010. [Online]. Available: http://arxiv.org/abs/1006.4255

[9] E. Abbe and E. Telatar, "MAC polar codes and matroids," in *Information Theory and Applications Workshop (ITA), 2010*. IEEE, Jan. 2010, pp. 1–8, 00021.

[10] R. Nasser and E. Telatar, "Fourier Analysis of MAC Polarization," *arXiv:1501.06076 [cs, math]*, Jan. 2015, 00000 arXiv: 1501.06076. [Online]. Available: http://arxiv.org/abs/1501.06076

[11] N. Goela, E. Abbe, and M. Gastpar, "Polar Codes for Broadcast Channels," *IEEE Transactions on Information Theory*, vol. 61, no. 2, pp. 758–782, Feb. 2015. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6975233

[12] M. Mondelli, S. H. Hassani, I. Sason, and R. Urbanke, "Achieving the Superposition and Binning Regions for Broadcast Channels Using Polar Codes," *arXiv:1401.6060 [cs, math]*, Jan. 2014. [Online]. Available: http://arxiv.org/abs/1401.6060

References III

[13] M. Mondelli, S. H. Hassani, I. Sason, and R. L. Urbanke, "Achieving Marton's Region for Broadcast Channels Using Polar Codes," *IEEE Transactions on Information Theory*, vol. 61, no. 2, pp. 783–800, Feb. 2015.

[14] M. Andersson, V. Rathi, R. Thobaben, J. Kliewer, and M. Skoglund, "Nested Polar Codes for Wiretap and Relay Channels," *IEEE Communications Letters*, vol. 14, no. 8, pp. 752–754, Aug. 2010. [Online]. Available: http://ieeexplore.ieee.org/document/5545658/

[15] R. Blasco-Serrano, R. Thobaben, V. Rathi, and M. Skoglund, "Polar codes for compress-and-forward in binary relay channels," in *2010 Conference Record of the Forty Fourth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*. IEEE, Nov. 2010, pp. 1743–1747, 00009.

[16] R. Blasco-Serrano, R. Thobaben, M. Andersson, V. Rathi, and M. Skoglund, "Polar Codes for Cooperative Relaying," *IEEE Transactions on Communications*, vol. 60, no. 11, pp. 3263–3273, Nov. 2012.

[17] M. Karzand, "Polar Codes for Degraded Relay Channels," in *International Zurich Seminar on Communications*. Citeseer, 2012, p. 59.

[18] L. Wang, "Polar coding for relay channels," in *2015 IEEE International Symposium on Information Theory (ISIT)*. Hong Kong, Hong Kong: IEEE, Jun. 2015, pp. 1532–1536. [Online]. Available: http://ieeexplore.ieee.org/document/7282712/

References IV

[19] L. Wang and E. Sasoglu, “Polar coding for interference networks,” *arXiv:1401.7293 [cs, math]*, Jan. 2014. [Online]. Available: http://arxiv.org/abs/1401.7293

[20] M. Mondelli, S. H. Hassani, and R. Urbanke, “A New Coding Paradigm for the Primitive Relay Channel,” *arXiv:1801.03153 [cs, math]*, Jan. 2018, arXiv: 1801.03153. [Online]. Available: http://arxiv.org/abs/1801.03153

[21] O. O. Koyluoglu and H. El Gamal, “Polar coding for secure transmission and key agreement,” in *2010 IEEE 21st International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*. IEEE, Sep. 2010, pp. 2698–2703.

[22] E. Hof and S. Shamai, “Secret and private rates on degraded wire-tap channels via polar coding,” in *2010 IEEE 26th Convention of Electrical and Electronics Engineers in Israel (IEEEI)*. IEEE, Nov. 2010, pp. 000 094–000 096, 00000.

[23] H. Mahdavifar and A. Vardy, “Achieving the Secrecy Capacity of Wiretap Channels Using Polar Codes,” *IEEE Transactions on Information Theory*, vol. 57, no. 10, pp. 6428–6443, Oct. 2011, 00177.

[24] Y. P. Wei and S. Ulukus, “Polar Coding for the General Wiretap Channel With Extensions to Multiuser Scenarios,” *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 2, pp. 278–291, Feb. 2016.

References V

[25] R. A. Chou and A. Yener, "Polar coding for the multiple access wiretap channel via rate-splitting and cooperative jamming," in *2016 IEEE International Symposium on Information Theory (ISIT)*. Barcelona, Spain: IEEE, Jul. 2016, pp. 983–987. [Online]. Available: http://ieeexplore.ieee.org/document/7541446/

[26] A. Bhatt, N. Ghaddar, and L. Wang, "Polar coding for multiple descriptions using monotone chain rules," in *2017 55th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. Monticello, IL, USA: IEEE, Oct. 2017, pp. 565–571. [Online]. Available: http://ieeexplore.ieee.org/document/8262787/

[27] T. C. Gulcu and A. Barg, "Interactive Function Computation via Polar Coding," *arXiv:1405.0894 [cs, math]*, May 2014. [Online]. Available: http://arxiv.org/abs/1405.0894

[28] E. Arıkan, "Channel Polarization: A Method for Constructing Capacity-Achieving Codes for Symmetric Binary-Input Memoryless Channels," *IEEE Transactions on Information Theory*, vol. 55, no. 7, pp. 3051–3073, Jul. 2009.

[29] S. B. Korada, E. Sasoglu, and R. Urbanke, "Polar codes: Characterization of exponent, bounds, and constructions," in *IEEE International Symposium on Information Theory, 2009. ISIT 2009*. IEEE, Jun. 2009, pp. 1483–1487.

References VI

[30] R. Mori and T. Tanaka, “Channel polarization on q-ary discrete memoryless channels by arbitrary kernels,” in *2010 IEEE International Symposium on Information Theory*, Jun. 2010, pp. 894–898.

[31] N. Presman, O. Shapira, S. Litsyn, T. Etzion, and A. Vardy, “Binary Polarization Kernels From Code Decompositions,” *IEEE Transactions on Information Theory*, vol. 61, no. 5, pp. 2227–2239, May 2015.

[32] H. Lin, S. Lin, and K. A. S. Abdel-Ghaffar, “Linear and Nonlinear Binary Kernels of Polar Codes of Small Dimensions With Maximum Exponents,” *IEEE Transactions on Information Theory*, vol. 61, no. 10, pp. 5253–5270, Oct. 2015.

[33] E. Şaşoğlu, E. Telatar, and E. Arikan, “Polarization for arbitrary discrete memoryless channels,” in *2009 IEEE Information Theory Workshop*, Oct. 2009, pp. 144–148.

[34] R. Mori and T. Tanaka, “Source and Channel Polarization Over Finite Fields and Reed–Solomon Matrices,” *IEEE Transactions on Information Theory*, vol. 60, no. 5, pp. 2720–2736, May 2014.

[35] ——, “Non-Binary Polar Codes using Reed-Solomon Codes and Algebraic Geometry Codes,” *1007.3661*, Jul. 2010. [Online]. Available: http://arxiv.org/abs/1007.3661

References VII

[36] ——, "Source and Channel Polarization over Finite Fields and Reed-Solomon Matrix," *arXiv:1211.5264*, Nov. 2012. [Online]. Available: http://arxiv.org/abs/1211.5264

[37] N. Presman, O. Shapira, and S. Litsyn, "Polar codes with mixed kernels," in *2011 IEEE International Symposium on Information Theory Proceedings*, Jul. 2011, pp. 6–10.

[38] F. Gabry, V. Bioglio, I. Land, and J. Belfiore, "Multi-kernel construction of polar codes," in *2017 IEEE International Conference on Communications Workshops (ICC Workshops)*, May 2017, pp. 761–765.

[39] W. Park and A. Barg, "Multilevel polarization for nonbinary codes and parallel channels," *1107.4965*, Jul. 2011. [Online]. Available: http://arxiv.org/abs/1107.4965

[40] ——, "Polar codes for q-ary channels, q=2^r," *arXiv:1107.4965*, Jul. 2011. [Online]. Available: http://arxiv.org/abs/1107.4965

[41] ——, "Polar codes for q-ary channels, q^2r," in *2012 IEEE International Symposium on Information Theory Proceedings (ISIT)*, Jul. 2012, pp. 2142–2146, 00000.

[42] ——, "Polar Codes for Q-Ary Channels,," *IEEE Transactions on Information Theory*, vol. 59, no. 2, pp. 955–969, Feb. 2013, 00028.

References VIII

[43] A. G. Sahebi and S. S. Pradhan, "Multilevel Polarization of Polar Codes Over Arbitrary Discrete Memoryless Channels," *arXiv:1107.1535 [cs, math]*, Jul. 2011. [Online]. Available: http://arxiv.org/abs/1107.1535

[44] Eren Sasoglu, "Polar Coding Theorems for Discrete Systems," Ph.D. dissertation, EPFL, Lausanne, 2011. [Online]. Available: doi:10.5075/epfl-thesis-5219

[45] S. H. Hassani and R. Urbanke, "Universal polar codes," in *2014 IEEE International Symposium on Information Theory*, Jun. 2014, pp. 1451–1455.

[46] E. Şaşoğlu and L. Wang, "Universal Polarization," *IEEE Transactions on Information Theory*, vol. 62, no. 6, pp. 2937–2946, Jun. 2016.

[47] M. Ye and A. Barg, "Universal Source Polarization and an Application to a Multi-User Problem," *arXiv:1408.6824 [cs, math]*, Aug. 2014, arXiv: 1408.6824. [Online]. Available: http://arxiv.org/abs/1408.6824

[48] E. Arıkan and E. Telatar, "On the rate of channel polarization," in *IEEE International Symposium on Information Theory, 2009. ISIT 2009.* IEEE, Jun. 2009, pp. 1493–1495.

[49] S. H. Hassani, R. Mori, T. Tanaka, and R. Urbanke, "Rate-Dependent Analysis of the Asymptotic Behavior of Channel Polarization," *arXiv:1110.0194*, Oct. 2011. [Online]. Available: http://arxiv.org/abs/1110.0194

References IX

[50] S. H. Hassani, K. Alishahi, and R. Urbanke, "Finite-Length Scaling of Polar Codes," *arXiv:1304.4778*, Apr. 2013. [Online]. Available: http://arxiv.org/abs/1304.4778

[51] V. Guruswami and P. Xia, "Polar Codes: Speed of Polarization and Polynomial Gap to Capacity," *IEEE Transactions on Information Theory*, vol. 61, no. 1, pp. 3–16, Jan. 2015, 00023.

[52] J. Błasiok, V. Guruswami, P. Nakkiran, A. Rudra, and M. Sudan, "General strong polarization," in *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing - STOC 2018*. Los Angeles, CA, USA: ACM Press, 2018, pp. 485–492. [Online]. Available: http://dl.acm.org/citation.cfm?doid=3188745.3188816

[53] ——, "General Strong Polarization," *arXiv:1802.02718 [cs, math]*, Feb. 2018, arXiv: 1802.02718. [Online]. Available: http://arxiv.org/abs/1802.02718

[54] D. Goldin and D. Burshtein, "Improved Bounds on the Finite Length Scaling of Polar Codes," *IEEE Transactions on Information Theory*, vol. 60, no. 11, pp. 6966–6978, Nov. 2014. [Online]. Available: http://ieeexplore.ieee.org/document/6905834/

[55] H. D. Pfister and R. Urbanke, "Near-Optimal Finite-Length Scaling for Polar Codes over Large Alphabets," *arXiv:1605.01997 [cs, math]*, May 2016, arXiv: 1605.01997. [Online]. Available: http://arxiv.org/abs/1605.01997

References X

[56] R. Mori and T. Tanaka, "Performance and construction of polar codes on symmetric binary-input memoryless channels," in *IEEE International Symposium on Information Theory, 2009. ISIT 2009*. IEEE, Jun. 2009, pp. 1496–1500.

[57] I. Tal and A. Vardy, "How to Construct Polar Codes," *1105.6164*, May 2011. [Online]. Available: http://arxiv.org/abs/1105.6164

[58] R. Pedarsani, S. H. Hassani, I. Tal, and I. E. Telatar, "On the construction of polar codes," in *Information Theory Proceedings (ISIT), 2011 IEEE International Symposium on*. IEEE, 2011, pp. 11–15, 00038. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6033724

[59] R. Pedarsani, "Polar Codes: Construction and Performance Analysis Ramtin Pedarsani," Ph.D. dissertation, EPFL, Lausanne, Jun. 2011, 00000.

[60] J. Honda and H. Yamamoto, "Polar coding without alphabet extension for asymmetric channels," in *2012 IEEE International Symposium on Information Theory Proceedings (ISIT)*, 2012, pp. 2147–2151.