Polar codes

Another way to look at channel coding

\((X_1, \ldots, X_N)\) channel inputs.

\((Y_1, \ldots, Y_N)\) channel outputs

\[
X_i \xrightarrow{W} Y_i
\]

\(\rightsquigarrow \text{BMS}\)

\[W = \{W(y|x)\}_{y \in \{0,1\}, x \in \{0,1\}}\]

Generator matrix

\[X_1^N = U_1^M \hat{G} \quad \hat{G} \in \{0,1\}^{M \times N}\]

can complete to

\[
G = \begin{bmatrix} \hat{G} \\ \hat{G} \end{bmatrix}^{\uparrow M}_{\downarrow N-M}
\]

rank of \(G\) = \(N\)

\(\leftarrow M\)
\[ X_i^n = U_i^n G, \quad G \in \{0,1\}^{N \times N} \]

invertible transform.

\[ U^{n}_{M+1} = 0 \]

- Generate \( U^m_i \in \{0,1\}^m \) uniformly.
- Complete with 0's
- Transmit \( X_i^n = U_i^n G \)
- Decode \( U^m_i \) from \( Y_i^n \) and \( U^{n}_{M+1} = 0 \)

By symmetry of the channel, equivalent

- Generate \( U^m_i \in \{0,1\}^n \) uniformly
- Transmit \( X_i^n = U_i^n G \)
- Transmit \( U^{n}_{M+1} \) through a clean channel
- Decode \( U^m_i \) from \( Y_i^n \) and \( U^{n}_{M+1} \).
The split need not to be first $M$ vs last $N-M$.

\[ [N] = F \cup I \]
\[ \uparrow \]
\[ \text{frozen bits} \]
\[ \text{set to 0} \]
\[ \text{revealed} \]
\[ r = \frac{|lI|}{N} \]
\[ \text{rate.} \]

Sequential decoder. - successive decoding

- Computes conditional distr.

\[ \hat{0}_i = \sum_{i 
otin F} y_i \text{ if } y_i \in F \]

\[ \hat{0}_i = 0 \text{ if } y_i \notin F \]
Call log-likelihood ratio

\[ L_i(y_1^n, \hat{u}_{i-1}) = \log \left\{ \frac{P_{u_i|y_1^n, u_{i-1}}(0|y_1^n, \hat{u}_{i-1})}{P_{u_i|y_1^n, u_{i-1}}(1|y_1^n, \hat{u}_{i-1})} \right\} \]

Recursively set

\[ \hat{u}_i = \begin{cases} 
0 & \text{if } i \in F, \quad \text{if } i \in I, \quad L_i(y_1^n, \hat{u}_{i-1}) > 0 \\
1 & \text{if } i \in I, \quad L_i(y_1^n, \hat{u}_{i-1}) < 0
\end{cases} \]

Suppose I give you \( G \in 0,1 \)^{n \times n}. How should you choose \( F, I \) ?

suppose we correctly de code \( U_1 \cdots U_{i-1} \), then we are trying to estimate \( U_i \) from \( Y_1^n U_i \)
Equivalent channel

\[ U_i \rightarrow W_i \rightarrow \sum_{i=1}^{N} U_{i-1} \]

Quality of the channel.

- Conditional entropy

\[ H(W_i) = H(U_i | \sum_{i=1}^{N} U_{i-1}) \]

- Minimal error prob.

\[ P_e(W_i) = \sum_{\tilde{y}_i \in \hat{Y}} W_i(\tilde{y}_i | 10) \cdot 1_{W_i(\tilde{y}_i | 10) \leq W_i(\tilde{y}_i | 11)} \]

- Bhattacharya parameter.

\[ Z(W_i) = \sum_{\tilde{y}_i \in \hat{Y}} \sqrt{W_i(\tilde{y}_i | 10) W_i(\tilde{y}_i | 11)} \]
Information set

\[ I(\delta) = \left\{ i \in [N] : P_e(W_i) \leq \delta \right\} \]

Block error probability

\[
P_\delta = \sum_{i=1}^{n} P(\hat{Y}_i = U_i; \hat{Y}_i \neq U_i)
\leq \sum_{i=1}^{n} P_e(W_i) \leq N \delta.
\]

Therefore need to take \( \delta = \delta_N \, \delta = o\left(\frac{1}{N}\right) \)

Relationship btw different metrics.

\[ H(W_i) \rightarrow 0 \iff P_e(W_i) \rightarrow 0 \iff Z(W_i) \rightarrow 0. \]

easier to analyze.
In particular

**Lemma**

\[ P_c(W_i) \leq Z(W_i) \]

**Proof**

\[ P_c(W_i) = \sum_{\gamma \in \mathcal{Y}} W_i(\gamma) 1_{\frac{W_i(\gamma_0)}{W_i(\gamma_1)} \leq 1} \leq \sum_{\gamma \in \mathcal{Y}} W_i(\gamma_0) \sqrt{\frac{W_i(\gamma_1)}{W_i(\gamma_0)}} = Z(W_i) \]

In particular I can use the info set

\[ J(\delta) = \{ i \in [N] : Z(W_i) \leq \delta \} \]

\[ \delta = o\left(\frac{1}{N}\right) \]
Lemma
\[ \sum_{i=1}^{N} H(W_i) = H(X | Y) = N \cdot H(W) = N(1 - C_W) \]

Proof: Chain rule of entropy

We say that the transform polarizes the channel \( W \) if
\[ \frac{1}{N} \# \left\{ i : H(W_i) \in [\delta, 1-\delta] \right\} \to 0 \]

It follows that
\[ I = \frac{1}{N} \left| \mathbb{I}(\delta) \right| = 1 - \frac{1}{N} \left\{ i : H(W_i) \geq (1-\delta) \right\} \]
\[ + o_N(1) \]
\[ = C_{KW} + o_N(1) \]

\( \Rightarrow \) the code achieves capacity