An Explicit Optimal Scheme for Distributed Lossy Compression

Min Ye and Alexander Barg

University of Maryland, College Park

Institute for Systems Research

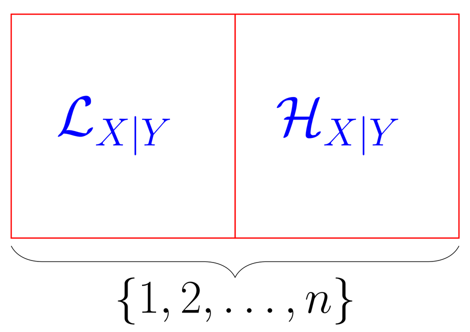
Abstract

We show that polar codes can be used to achieve the rate-distortion functions in the problem of hierarchical lossy compression also known as the successive refinement problem.

Introduction

Polar coding was introduced by Arikan in the seminal paper [Arıkan '09].

Let $n = 2^m, G_n = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}^{\otimes m}$. Arıkan showed that given a binary-input channel W, there is a sequence of linear codes, whose generator matrices are appropriately chosen from the rows of G_n, achieving the symmetric capacity of W. Later it was proved that polar codes work equally well for source coding [Arıkan '10].



- $X, Y \sim P_{X,Y}$
- Let $U^n = X^n G_n$. $U_{\mathcal{L}_{X|Y}}$ can be determined by previous bits together with Y^n, while $U_{\mathcal{H}_{X|Y}}$ is uniformly random.
- For channel transmission, put information in $U_{\mathcal{L}_{X|Y}}$; for source coding, only need to record $U_{\mathcal{H}_{X|Y}}$.
- $$\lim_{n\to\infty} \frac{1}{n} |\mathcal{H}_{X|Y}| = H(X|Y).$$
  Achieve optimal rates for both channel and source coding.

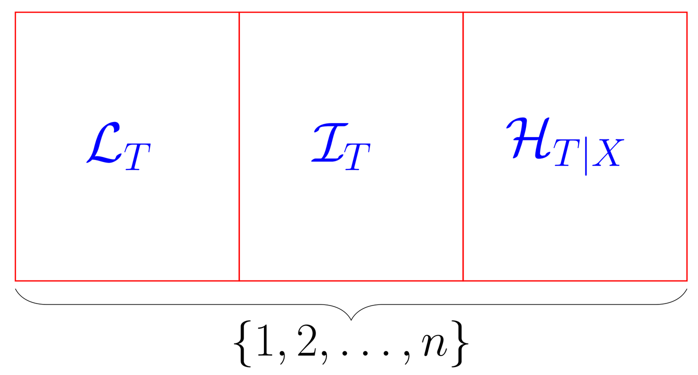
Polar Codes for Lossy Source Coding

Rate-Distortion:

- Source $X \sim P_X$ over a finite alphabet $\mathcal{X}$
- Distortion function $d : \mathcal{X} \times \{0, 1\} \to [0, \infty)$
- Rate distortion function $R(D) = \min_{P_{T|X}} I(X;T)$, where $P_{T|X}$ is such that $E_{XT}(d(X,T)) \le D$.

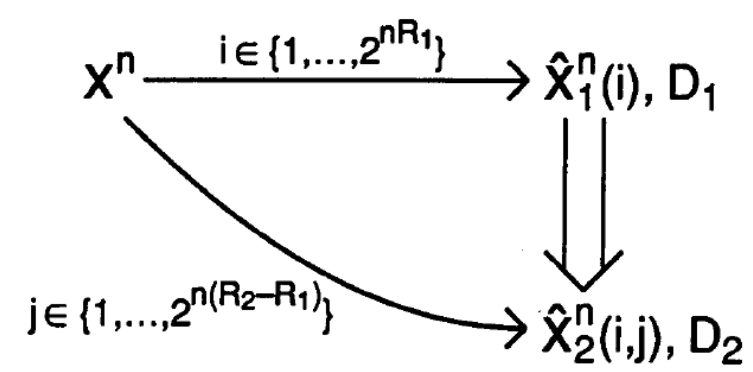
Polar Coding Scheme [Honda and Yamamoto '13]

- Objective: to approximate the distribution $P_{T^n X^n}$.
- Let $U^n = T^n G_n$. $P_{T^n X^n}$ induces a joint distribution $P_{U^n X^n}$. It is equivalent to approximating $P_{U^n X^n}$.
- Divide U^n into three parts.



- For $i \in \mathcal{L}_T$, U_i can be determined by U^{i-1}.
- For $i \in \mathcal{H}_{T|X}$, U_i is independent of U^{i-1} and X^n.
- For $i \in \mathcal{I}_T$, U_i is independent of U^{i-1} but can be determined by U^{i-1} and X^n.
- Set $U_{\mathcal{H}_{T|X}}$ to be uniformly random independently of X^n. Make it known to both the encoder and the decoder.
- The encoder has access to X^n, thus it knows the value of $U_{\mathcal{I}_T}$ and $U_{\mathcal{L}_T}$.
- Transmit $U_{\mathcal{I}_T}$ to the decoder.
  $$\lim_{n\to\infty} \frac{1}{n} |\mathcal{I}_T| = I(X;T) = R(D).$$
- The decoder knows both $U_{\mathcal{H}_{T|X}}$ and $U_{\mathcal{I}_T}$.
- Thus it can recover $U_{\mathcal{L}_T}$ by looking at the previous bits.

Successive Refinement of Information



The source X is said to be successively refinable with distortions D_1 and D_2, $D_2 \le D_1$, if the pair of rate values $(R(D_1), R(D_2))$ is achievable.

Theorem 1 (Koshelev '80, Equitz and Cover '91)**.** *Let X be a source and let T, W be two binary random variables. The source is successively refinable with distortions D_1 and D_2 ($D_2 \le D_1$) if and only if there exists a conditional distribution $P_{TW|X}$ with*

$$E_{XT} d(X,T) \le D_1, \quad E_{XW} d(X,W) \le D_2,$$

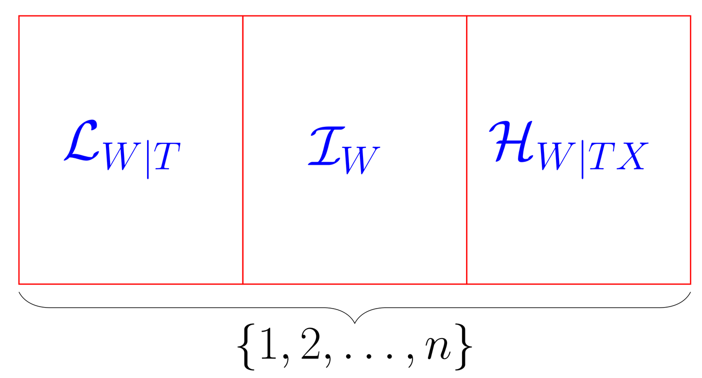
$$I(X;T) = R(D_1), \quad I(X;W) = R(D_2), \tag{1}$$

and such that X, W, T satisfy the Markov condition

$$X \to W \to T. \tag{2}$$

Polar Coding Scheme for Successive Refinement

- Objective: to approximate the distribution $P_{W^n T^n X^n}$.
- Let $U^n = T^n G_n$ and $V^n = W^n G_n$. $P_{W^n T^n X^n}$ induces the joint distribution $P_{V^n U^n X^n}$. It is equivalent to approximate $P_{V^n U^n X^n}$.
- We have shown how to approximate $P_{U^n X^n}$. That is the construction of the coarse layer coding scheme.
- Now we only need to approximate $P_{V^n | U^n X^n}$. Divide V^n into three parts.



- The refinement layer can reconstruct T^n using the encoded sequence at the coarse layer. No need to record $V_{\mathcal{L}_{W|T}}$.
- $V_{\mathcal{H}_{W|TX}}$ is approximately independent of previous bits and $T^n X^n$. No need to record either. Set $V_{\mathcal{H}_{W|TX}}$ to be uniformly random and make these bits known to both encoder and decoder.
- The encoder only needs to record $V_{\mathcal{I}_W}$. Due to the Markov condition,
  $$\frac{|\mathcal{I}_W|}{n} \to R(D_2) - R(D_1).$$

References

[1] M. Ye and A. Barg. "Polar codes for distributed hierarchical source coding." Advances in Mathematics of Communications, 9(1), 87-103, 2015.

[2] V. N. Koshelev, "Hierarchical coding of discrete sources." Probl. Inform. Trans., 16 (1980), 11-19.

[3] W. Equitz and T. Cover, "Successive refinement of information." IEEE Trans. Inform. Theory, 37 (1991), 269-275.