One-Shot Wyner–Ziv Compression of a Uniform Source

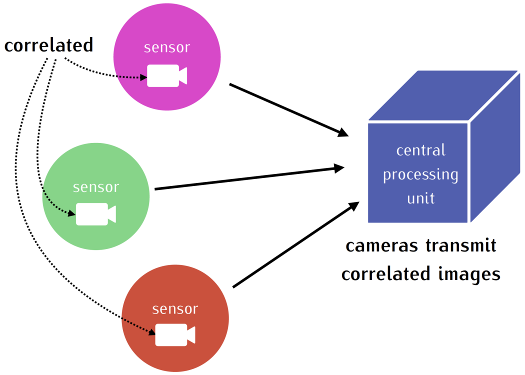
Oğuzhan Kubilay Ülger, Elza Erkip

New York University

2024 IEEE International Symposium on Information Theory

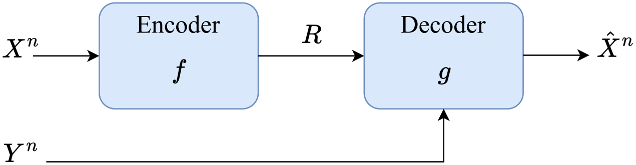
7/10/2024

Distributed Compression



- Distributed sensor network such as a distributed camera array.
- Correlated observations transmitted to a central processing unit.
- How can the sensors exploit this correlation to send their data efficiently?

Compression with Side Information



Rate Distortion Function

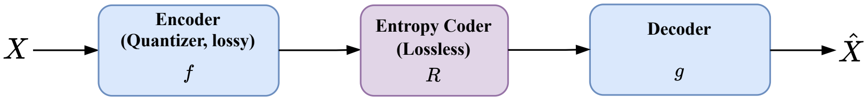
$$R(\Delta) = \min I(X;U) - I(Y;U)$$

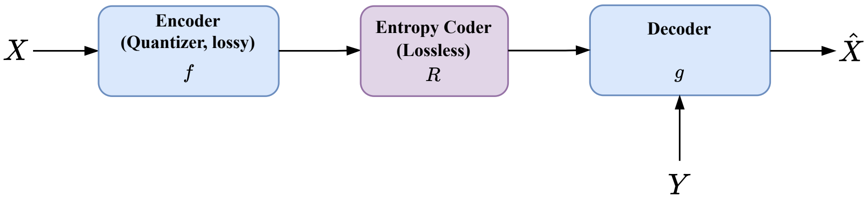
where minimization is over all $p(u|x)$ and $g(U,Y) = \hat{X}$ such that $\mathbb{E}[d(X,\hat{X})] \leq D$.

Special case by Wyner and Ziv (1976):

- Compress X^n under some distortion constraint.
- Correlated side information is available only at the decoder:
- They characterize the rate-distortion trade-off in the asymptotic setting.
- Asymptotic methods do not directly apply to one-shot!

One-Shot Compression with Side Information

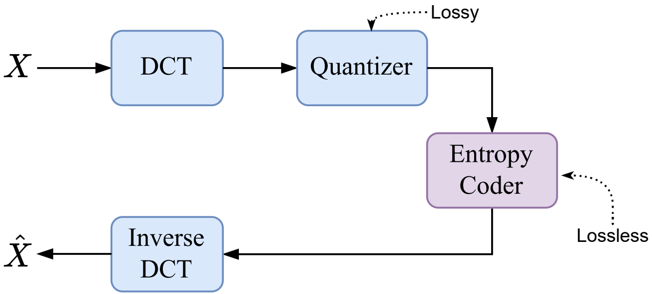




- Popular approach for **point-to-point** (no side information) setting.
- **Quantizer** (encoder) followed by an **entropy coder** (lossless, variable length such

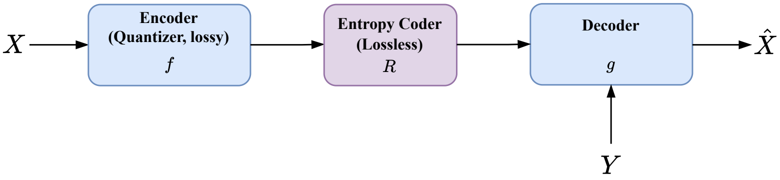
Example: JPEG

- Classical image compression: JPEG



- Also popular method for learned compressors.
- In our problem, we have to include the side information!

ECSQ with Side Information

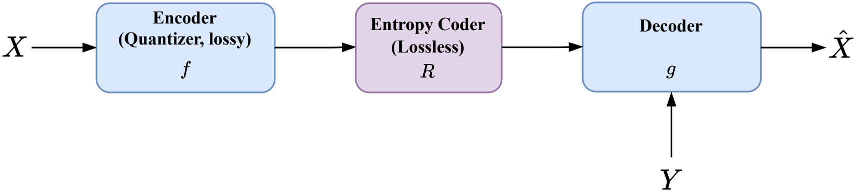


- Source X, side information Y and distortion metric $d(x, \hat{x})$.
- Encoder $f : \mathcal{X} \to \mathbb{N}$, decoder $g : \mathbb{N} \times \mathcal{Y} \to \hat{\mathcal{X}}$.
- The entropy and distortion of this encoder-decoder pair with $\hat{X} = g(f(X), Y)$:

$$H(f) = -\sum_i \mathbb{P}[f(X) = i] \log(\mathbb{P}[f(X) = i])$$

$$D_{SI}(f, g) = \mathbb{E}[d(X, \hat{X})]$$

Side Information Entropy–Distortion Functions

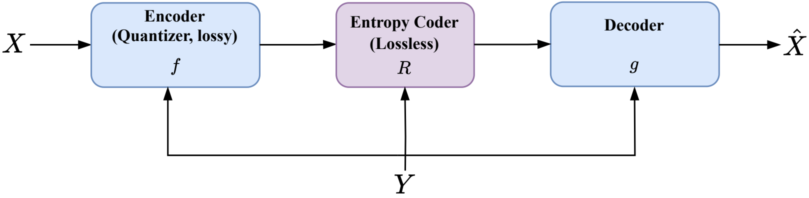


- **Entropy-distortion function with SI** is given by:

$$H_{SI}(\Delta) = \inf_{f} H(f)$$

- Infimum is taken over all encoders f such that there exists a decoder g with $D_{SI}(f, g) \leq \Delta$.

Conditional ECSQ

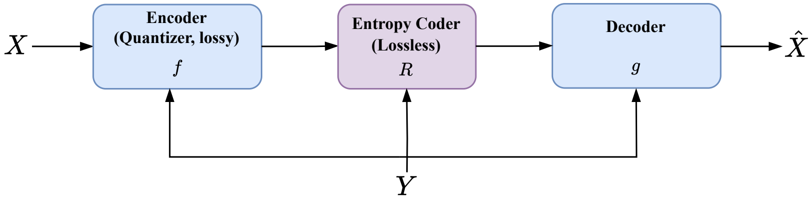


- Consider the case where Y is also available at the encoder. We refer to this case as **conditional ECSQ**.
- Encoder $f : \mathcal{X} \times \mathcal{Y} \to \mathbb{N}$, decoder $g : \mathbb{N} \times \mathcal{Y} \to \hat{\mathcal{X}}$.
- The conditional entropy and distortion of (f, g) with $\hat{X} = g(f(X,Y),Y)$:

$$H(f|Y) = \mathbb{E}\left[-\sum_i \mathbb{P}[f(X) = i|Y = y] \log(\mathbb{P}[f(X) = i|Y = y])\right]$$

$$D_C(f,g) = \mathbb{E}[d(X,\hat{X})]$$

Conditional Entropy–Distortion Function



- Similar to the previous case, we define the **conditional entropy-distortion function** as:

$$H_C(\Delta) = \inf_f H(f|Y)$$

- The infimum is taken over all encoders f such that there exists a decoder g with $D_C(f, g) \leq \Delta$.

Our Goal

We wish to characterize $H_{SI}(\Delta)$ and $H_C(\Delta)$.

- We consider a uniform source on the unit interval $[0, 1]$.

Motivation

- Recent interest in bench-marking performance of neural compressors on processes generated by uniform sources[1,2].
- Emergence of neural compressors in decoder-only side information settings[3].
- Even without SI, closed form expressions are only known for a limited source distributions.

---

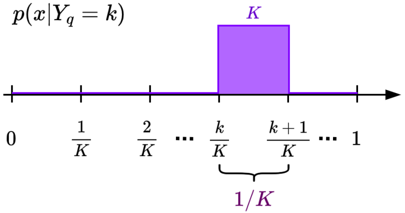
[1] A. B. Wagner and J. Balle, "Neural networks optimally compress the sawbridge," in 2021 Data Compression Conference (DCC). IEEE, 2021.

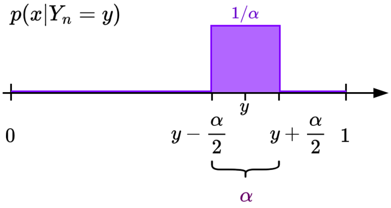
[2] S. Bhadane, A. B. Wagner, and J. Balle, "Do neural networks compress manifolds optimally?" in 2022 IEEE Information Theory Workshop (ITW). IEEE, 2022.

[3] E. Ozyilkan, J. Balle, and E. Erkip, "Neural distributed compressor discovers binning," IEEE Journal on Selected Areas in Information Theory, 2024.

Source and Side Information

- $X \sim \text{Unif}([0,1])$ and L_1 distance $d(x,\hat{x}) = |x - \hat{x}|$.
  - The results are valid for any r^{th} power distortion.
- Two different SI settings:
  1. **Quantized**: $Y_q = \lfloor KX \rfloor / K$ for some integer $K \geq 1$.
  2. **Noisy**: $Y_n = X + Z$ (mod 1) where Z is distributed uniformly on $[-\alpha/2, \alpha/2]$.





- Given Y, X is distributed uniformly on a smaller set. Size $1/K$ or α.

Source and Side Information

- Quantized SI, Y_q is a function of X so the encoder also has access to it.
  1. For Y_q we consider the conditional model.
  2. Y_q is a coarse quantized version of X.
- Noisy SI, Y_n is not available at the encoder because of the independent noise.

Goals

1. For Y_q: Characterize conditional entropy distortion function $H_C^q(\Delta)$
2. For Y_n: Characterize entropy-distortion function with side information $H_{SI}^n(\Delta)$

Optimal ECSQ for Uniform Source

Point-to-point entropy-distortion function[1]

For $X \sim \text{Unif}([0,1])$ and L_1 distortion metric $d(x,\hat{x}) = |x - \hat{x}|$. The entropy-distortion function is given by:

$$U(\Delta) = \begin{cases} -\lfloor 1/p \rfloor p \log p - q \log q, & 0 < \Delta < 1/4 \\ 0, & \Delta \geq 1/4 \end{cases}$$

where $q = (1 - \lfloor 1/p \rfloor p)$ and $p \in (0,1)$ is the unique solution to

$$\lfloor 1/p \rfloor p^2 + q^2 = 4\Delta.$$

- Achieved by quantizing X with biuniform intervals: $\lfloor 1/p \rfloor$ intervals of size p and one of size $q = (1 - \lfloor \frac{1}{p} \rfloor p)$

[1] A. Gyorgy and T. Linder, "Optimal entropy-constrained scalar quantization of a uniform source," IEEE Transactions on Information Theory, 2000.

Quantized Side Information: Characterization

Quantized SI: $Y_q = \lfloor KX \rfloor / K$, coarsely quantized X.

Quantized SI Conditional Entropy-Distortion Function

Consider $X \sim \text{Unif}([0,1])$ source and side information $Y_q = \lfloor KX \rfloor / K$ for some integer $K \geq 1$. The conditional entropy–distortion function is given by

$$H_C^q(\Delta) = \min_{\{\Delta_k\}_{k=1}^K} \frac{1}{K} \sum_{k=1}^{K} U(K\Delta_k)$$

$$\text{s.t.} \quad \frac{1}{K} \sum_{k=1}^{K} \Delta_k \leq \Delta,$$

$$\Delta_k \geq 0 \text{ for all } k \in \{1, \ldots, K\}$$

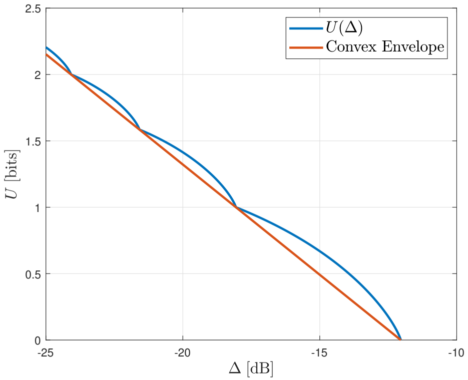
Quantized Side Information: Bounds

- The optimization problem is non-convex.
- We present simpler lower and upper bounds.

The lower and upper bounds to conditional entropy distortion function:

$$\breve{U}(K\Delta) \leq H_C^q(\Delta) \leq U(K\Delta)$$

where $\breve{U}(\cdot)$ is the convex envelope of $U(\cdot)$



- **Upper bound:** Set $\Delta_k = \Delta$ for all k.
- **Lower bound:** Lower bound the optimization problem with a convex one.

Noisy Side Information: Converse

Noisy SI: $Y_n = X + Z$ (mod 1) where Z is distributed uniformly on $[-\alpha/2, \alpha/2]$.

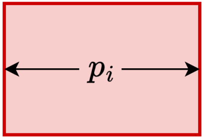
Lower Bound to the Entropy-Distortion Function with Noisy SI

For $X \sim$ Unif$(0,1)$ and $Y_n = X + Z$ (mod 1) with $Z \sim$ Unif$(\alpha/2, \alpha/2)$ the entropy distortion function is lower bounded as:

$$H_{SI}^n(\Delta) \geq \breve{U}(\Delta/\alpha)$$

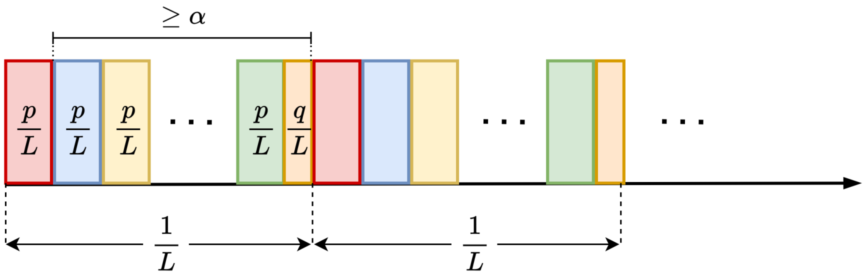
- Genie-aided lower-bound: We assume Y_n is also available at the encoder.
- Similar to the quantized case, we can obtain a lower bound in terms of the point-to-point entropy-distortion function.

Noisy Side Information: Achievability



Noisy Side Information: Achievability

- Similar to point-to-point case, we choose biuniform p_i: **All but one** encoding indices have size p, **remaining one** has size $q = 1 - \lfloor 1/p \rfloor p$.
- Each is split into L **subintervals**, placed at least α apart.



Noisy Side Information: Achievability

Upper Bound to the Entropy-Distortion Function with Noisy SI

For $X \sim \text{Unif}(0,1)$ and $Y_n = X + Z$ (mod 1) with $Z \sim \text{Unif}(\alpha/2, \alpha/2)$ the entropy distortion function is upper bounded as:

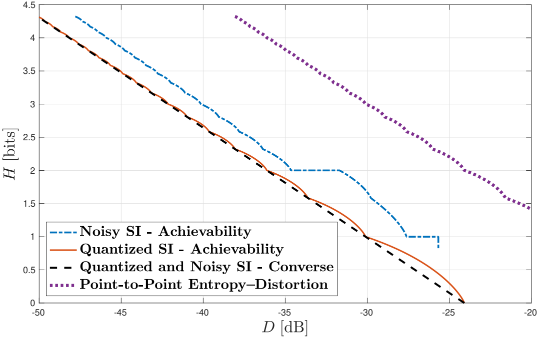
$$H_{SI}(\Delta) \leq -\lfloor 1/p \rfloor p \log p - q \log q$$

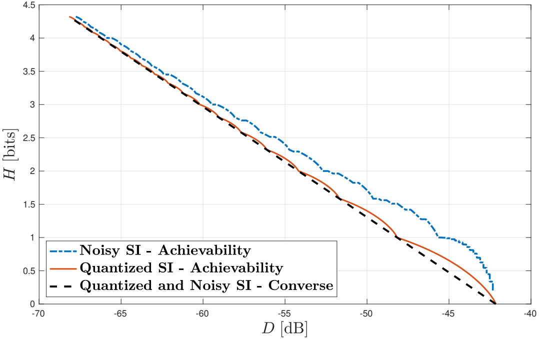
where $q = (1 - \lfloor \frac{1}{p} \rfloor p)$ and $p \in (0, 1-\alpha)$ is the solution to

$$4\Delta = \left\lfloor \frac{1}{p} \right\rfloor \left(p \left(\frac{p}{L} \wedge \alpha \right) - \frac{L}{4\alpha} \left(\frac{p}{L} \wedge \alpha \right)^3 \right) + \left(q \left(\frac{q}{L} \wedge \alpha \right) - \frac{L}{4\alpha} \left(\frac{q}{L} \wedge \alpha \right)^3 \right)$$

with $L = \lfloor (1-p)/\alpha \rfloor$ and $(a \wedge b) = \min(a, b)$.

Illustrations of the Results





Entropy–distortion trade-off for quantized and noisy SI

$K = 1/\alpha = 4$ $\qquad$ $K = 1/\alpha = 32$

Conclusions & Future Work

- We investigated a one-shot Wyner–Ziv problem, i.e. entropy-distortion trade-off with SI, for a uniform source with two SI models.
- We presented upper and lower bounds for the entropy–distortion functions and showed that they get tighter at higher rates.

Future work:

- Complex sources, general side information.
- High rate regime analysis where our achievabilities are close to optimal.