Multiterminal Source Coding with an Entropy-Based Distortion Measure

Thomas Courtade and Rick Wesel

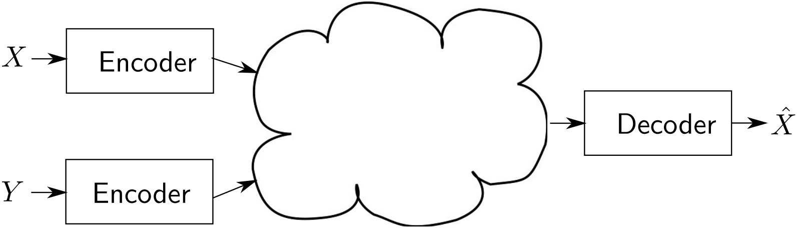
Department of Electrical Engineering
University of California, Los Angeles

4 August, 2011

IEEE International Symposium on Information Theory
Saint Petersburg, Russia

Motivation

The lossless one-helper problem

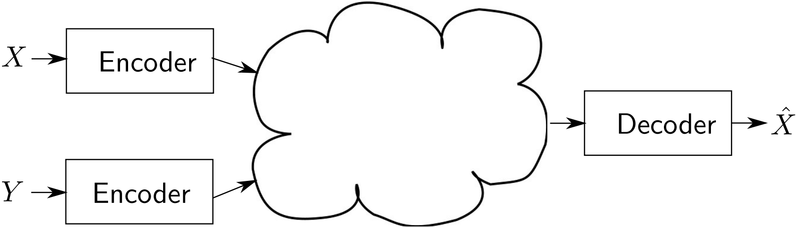


Question

What is the achievable rate region for a lossless one-helper network with a single source?

Motivation

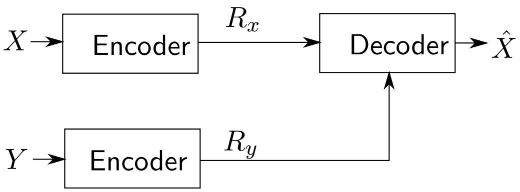
The lossless one-helper problem



Question

What is the achievable rate region for a lossless one-helper network with a single source?

- The answer to this question appears to be out of reach for now.

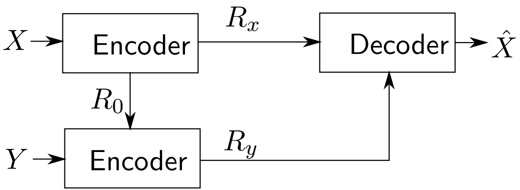


Theorem (Ahlswede-Körner-Wyner 1975)

$$\mathcal{R} = \left\{ (R_x, R_y) : \begin{array}{l} R_x \geq H(X|U) \\ R_y \geq I(Y;U) \\ \textit{for some distribution} \\ p(x,y,u) = p(x,y)p(u|y), \\ \textit{where } |\mathcal{U}| \leq |\mathcal{Y}| + 2 \end{array} \right\}$$

Motivation

The lossless one-helper problem

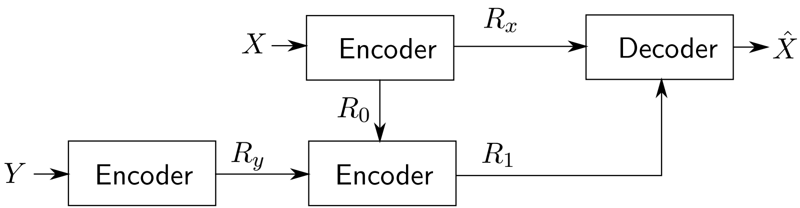


Theorem (Kaspi-Berger 1982)

$$\mathcal{R} = \left\{ (R_0, R_x, R_y) : \begin{array}{l} \exists p(x,y,v,u) = p(x,y)p(u|x)p(v|y,u) \textit{ such that} \\ R_0 \geq I(X;U|Y), \\ R_x \geq H(X|V,U), \\ R_x + R_y \geq H(X) + I(Y;V|U,X) \end{array} \right\}$$

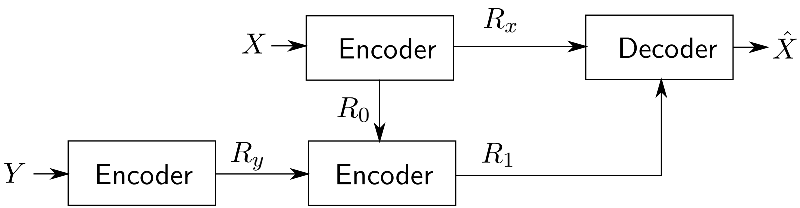
Motivation

The lossless one-helper problem

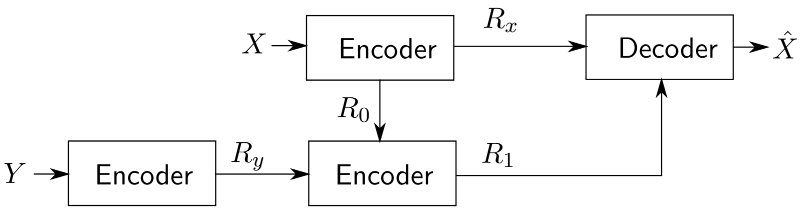


Motivation

The lossless one-helper problem



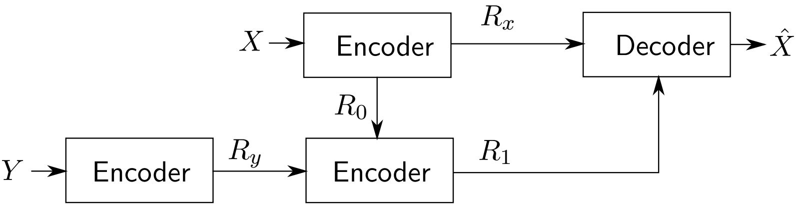
- Achievable rate region appears to be unknown.



- Achievable rate region appears to be unknown.
- The encoder without a source is problematic.

Motivation

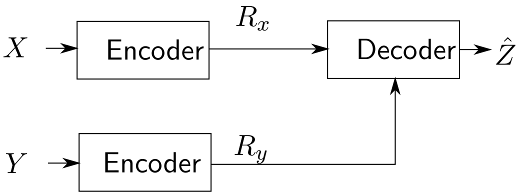
The lossless one-helper problem



- Achievable rate region appears to be unknown.
- The encoder without a source is problematic.
- Intuitively, it should send some lossy estimate of X.

Multiterminal Source Coding

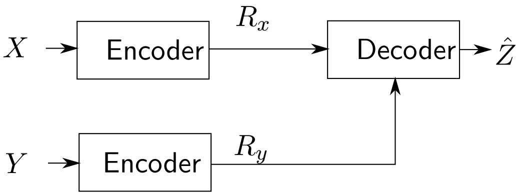
with an entropy-based distortion measure



- We study two cases of the MTSC problem:
  1. Joint distortion constraint: $\mathbb{E}\left[d(X, Y, \hat{Z})\right] \leq D$,
  2. Distortion constraint only on X: $\mathbb{E}\left[d(X, \hat{Z})\right] \leq D$.

Multiterminal Source Coding

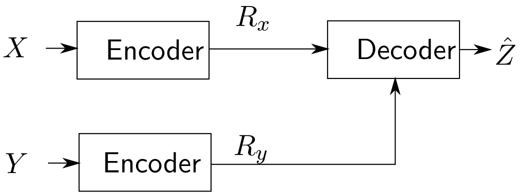
with an entropy-based distortion measure



- We study two cases of the MTSC problem:
  1. Joint distortion constraint: $\mathbb{E}\left[d(X,Y,\hat{Z})\right] \leq D$,
  2. Distortion constraint only on X: $\mathbb{E}\left[d(X,\hat{Z})\right] \leq D$.
- We consider a particular choice of $\hat{\mathcal{Z}}$ and $d(\cdot)$ (Case 1):
  - $\hat{\mathcal{Z}} = \mathbb{R}_+^{|\mathcal{X}\times\mathcal{Y}|}$ (i.e., the set of functions from $\mathcal{X} \times \mathcal{Y}$ to $\mathbb{R}_+$).
  - $d(x,y,\hat{z}) = \log\left(\frac{1}{\hat{z}(x,y)}\right)$.

Multiterminal Source Coding

with an entropy-based distortion measure



- We study two cases of the MTSC problem:
  1. Joint distortion constraint: $\mathbb{E}\left[d(X,Y,\hat{Z})\right] \leq D$,
  2. Distortion constraint only on X: $\mathbb{E}\left[d(X,\hat{Z})\right] \leq D$.
- We consider a particular choice of $\hat{\mathcal{Z}}$ and $d(\cdot)$ (Case 2):
  - $\hat{\mathcal{Z}} = \mathbb{R}_+^{|\mathcal{X}|}$ (i.e., the set of functions from $\mathcal{X}$ to $\mathbb{R}_+$).
  - $d(x,\hat{z}) = \log\left(\frac{1}{\hat{z}(x)}\right)$.

Examples

An entropy-based distortion measure

- To make things interesting, we assume $\sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}} \hat{z}(x,y) \leq M$.

Examples

An entropy-based distortion measure

- To make things interesting, we assume $\sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}} \hat{z}(x,y) \leq M$.
- Without loss of generality:

$$\sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}} \hat{z}(x,y) = 1 \quad \Rightarrow d(x,y,\hat{z}) = D\left(1_{\{(x',y')=(x,y)\}} \| \hat{z}(x',y')\right).$$

Examples

An entropy-based distortion measure

- To make things interesting, we assume $\sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}} \hat{z}(x,y) \leq M$.
- Without loss of generality:

$$\sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}} \hat{z}(x,y) = 1 \quad \Rightarrow d(x,y,\hat{z}) = D\left(1_{\{(x',y')=(x,y)\}} \| \hat{z}(x',y')\right).$$

- We obtain the erasure distortion measure if we restrict $\hat{z}(x,y)$ to functions of the form:

$$\hat{z}(x,y) = \begin{cases} 1 & \text{if } (x,y) = (x',y') \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad \hat{z}(x,y) = \frac{1}{|\mathcal{X}\times\mathcal{Y}|}\ \forall x,y.$$

Examples

An entropy-based distortion measure

- To make things interesting, we assume $\sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}} \hat{z}(x,y) \leq M$.
- Without loss of generality:

$$\sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}} \hat{z}(x,y) = 1 \quad \Rightarrow d(x,y,\hat{z}) = D\left(1_{\{(x',y')=(x,y)\}} \| \hat{z}(x',y')\right).$$

- We obtain the erasure distortion measure if we restrict $\hat{z}(x,y)$ to functions of the form:

$$\hat{z}(x,y) = \begin{cases} 1 & \text{if } (x,y)=(x',y') \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad \hat{z}(x,y) = \frac{1}{|\mathcal{X}\times\mathcal{Y}|} \ \forall x,y.$$

- If $(R_x, R_y) = (0,0)$, setting $\hat{z}(x,y) = p(x,y)$ results in distortion $H(X,Y)$.

- To make things interesting, we assume $\sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}} \hat{z}(x,y) \leq M$.
- Without loss of generality:

$$\sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}} \hat{z}(x,y) = 1 \quad \Rightarrow d(x,y,\hat{z}) = D\left(1_{\{(x',y')=(x,y)\}} \| \hat{z}(x',y')\right).$$

- We obtain the erasure distortion measure if we restrict $\hat{z}(x,y)$ to functions of the form:

$$\hat{z}(x,y) = \begin{cases} 1 & \text{if } (x,y) = (x',y') \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad \hat{z}(x,y) = \frac{1}{|\mathcal{X}\times\mathcal{Y}|} \; \forall x,y.$$

- If $(R_x, R_y) = (0,0)$, setting $\hat{z}(x,y) = p(x,y)$ results in distortion $H(X,Y)$.
- If $(R_x, R_y) = (0, H(Y))$, setting $\hat{z}(x,y) = p(x|y)$ results in distortion $H(X|Y)$.

Theorem (Wyner-Ziv 1976)

Let (X, Y) be drawn i.i.d. and let $d(x, \hat{z})$ be given. The rate distortion function with side information is

$$R_Y(D) = \min_{p(w|x)} \min_{f} I(X; W|Y)$$

where the minimization is over all functions $f : \mathcal{Y} \times \mathcal{W} \to \hat{\mathcal{Z}}$ and conditional distributions $p(w|x)$ such that $\mathbb{E}\left[d(X, f(Y, W))\right] \leq D$.

Examples

Rate distortion with side information

Theorem (Wyner-Ziv 1976)

Let (X, Y) be drawn i.i.d. and let $d(x, \hat{z})$ be given. The rate distortion function with side information is

$$R_Y(D) = \min_{p(w|x)} \min_{f} I(X; W|Y)$$

where the minimization is over all functions $f : \mathcal{Y} \times \mathcal{W} \to \hat{\mathcal{Z}}$ and conditional distributions $p(w|x)$ such that $\mathbb{E}\left[d(X, f(Y, W))\right] \leq D$.

Corollary

For our choice of $\hat{\mathcal{Z}}$ and $d(\cdot)$, the rate distortion function with side information is:

$$R_Y(D) = H(X|Y) - D.$$

Examples

Rate distortion with side information

Proof of Corollary:

$$D \geq \mathbb{E}\left[d(X, f(Y,W))\right] = \mathbb{E}\left[\log\left(\frac{1}{f(Y,W)[X]}\right)\right]$$
$$= D(p(x|y,w)\|f(y,w)[x]) + H(X|Y,W) \geq H(X|Y,W).$$

Therefore:

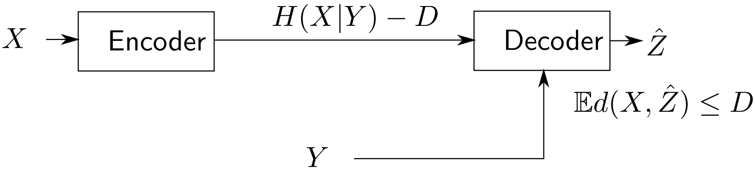
$$R_Y(D) = \min_{p(w|x)} \min_{f} \{H(X|Y) - H(X|Y,W)\} \geq H(X|Y) - D.$$

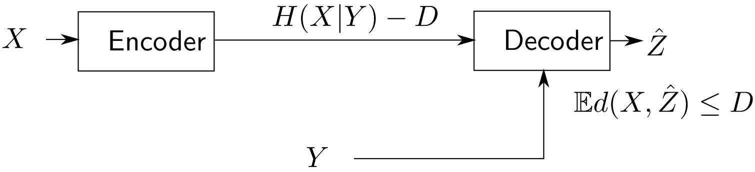
Taking $f(y,w)[x] = p(x|y,w)$ and $W = \begin{cases} X & \text{with probability } 1 - \frac{D}{H(X|Y)} \\ \emptyset & \text{with probability } \frac{D}{H(X|Y)} \end{cases}$

achieves equality throughout $\Rightarrow R_Y(D) = H(X|Y) - D$.

Examples

Rate distortion with side information

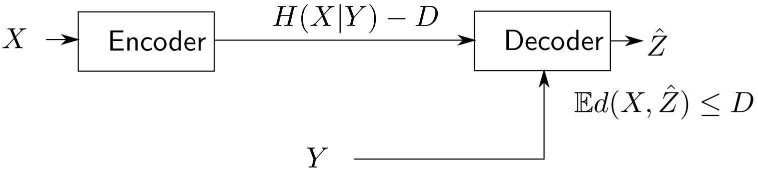




- Intuition: every "bit" of distortion we tolerate saves one bit of rate.

Examples

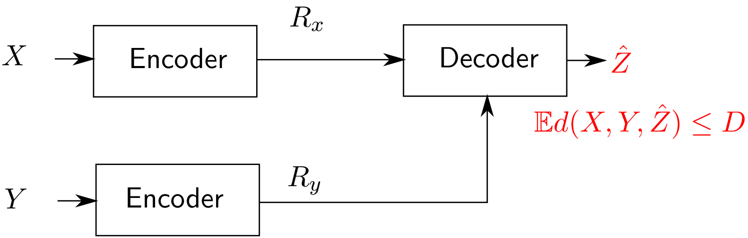
Rate distortion with side information



- Intuition: every "bit" of distortion we tolerate saves one bit of rate.
- What if Y is rate limited? Does a similar theme prevail?

d-Lossy Coding of Correlated Sources

Joint distortion criterion



Theorem

$$\mathcal{R} = \left\{ (R_x, R_y, D) : \begin{array}{l} R_x \geq H(X|Y) - D \\ R_y \geq H(Y|X) - D \\ R_x + R_y \geq H(X,Y) - D \end{array} \right\}$$

d-Lossy Coding of Correlated Sources

Joint distortion criterion

- We obtain a D-bit enlargement of the achievable rate region in all directions.

- We obtain a D-bit enlargement of the achievable rate region in all directions.
- The proof relies on the WZ corollary and lemmas of the form:

Lemma (Distortion Preimage)

Define $\mathcal{A}(\hat{z}^n) = \{(x^n, y^n) : d(x^n, y^n, \hat{z}^n) \leq D + \epsilon\}$. *Then*

$$|\mathcal{A}(\hat{z}^n)| \leq 2^{n(D+2\epsilon)}.$$

d-Lossy Coding of Correlated Sources

Joint distortion criterion

- We obtain a D-bit enlargement of the achievable rate region in all directions.
- The proof relies on the WZ corollary and lemmas of the form:

Lemma (Distortion Preimage)

Define $\mathcal{A}(\hat{z}^n) = \{(x^n, y^n) : d(x^n, y^n, \hat{z}^n) \leq D + \epsilon\}$. *Then*

$$|\mathcal{A}(\hat{z}^n)| \leq 2^{n(D+2\epsilon)}.$$

- Intuitively, if we have a reconstruction $\hat{z}^n$ then nD additional bits of information about (x^n, y^n) are required to determine (x^n, y^n) completely.

- We obtain a D-bit enlargement of the achievable rate region in all directions.
- The proof relies on the WZ corollary and lemmas of the form:

Lemma (Distortion Preimage)

Define $\mathcal{A}(\hat{z}^n) = \{(x^n, y^n) : d(x^n, y^n, \hat{z}^n) \leq D + \epsilon\}$. *Then*

$$|\mathcal{A}(\hat{z}^n)| \leq 2^{n(D+2\epsilon)}.$$

- Intuitively, if we have a reconstruction $\hat{z}^n$ then nD additional bits of information about (x^n, y^n) are required to determine (x^n, y^n) completely.
- This hints at a relationship with the Slepian-Wolf region.

- The basic idea is to show a correspondence with the Slepian-Wolf region as follows:

Claim 1

If (R_x, R_y, D) is an achievable RD point, then $(R_x + \theta D, R_y + (1-\theta)D)$ is an achievable Slepian-Wolf rate pair for some $\theta \in [0, 1]$.

- Proved using the distortion preimage lemmas combined with a decomposition of the distortion measure and a random binning argument.

d-Lossy Coding of Correlated Sources

Joint distortion criterion: Proof sketch

- The basic idea is to show a correspondence with the Slepian-Wolf region as follows:

Claim 1

If (R_x, R_y, D) is an achievable RD point, then $(R_x + \theta D, R_y + (1 - \theta)D)$ is an achievable Slepian-Wolf rate pair for some $\theta \in [0, 1]$.

Claim 2

If $(R_x + \theta D, R_y + (1 - \theta)D)$ is an achievable Slepian-Wolf rate pair for some $\theta \in [0, 1]$, then (R_x, R_y, D) is an achievable RD point.

- Proved via the WZ corollary and timesharing.

d-Lossy Coding of Correlated Sources

Joint distortion criterion: Proof sketch

- The basic idea is to show a correspondence with the Slepian-Wolf region as follows:

Claim 1

If (R_x, R_y, D) is an achievable RD point, then $(R_x + \theta D, R_y + (1-\theta)D)$ is an achievable Slepian-Wolf rate pair for some $\theta \in [0,1]$.

Claim 2

If $(R_x + \theta D, R_y + (1-\theta)D)$ is an achievable Slepian-Wolf rate pair for some $\theta \in [0,1]$, then (R_x, R_y, D) is an achievable RD point.

- Proved via the WZ corollary and timesharing.
- Combining the two claims with the known form of the Slepian-Wolf region gives the expression for the achievable rate distortion region.

Decomposing the Entropy Measure $d(\cdot)$

Joint distortion criterion: Proof sketch

- Given a rate distortion code, $\hat{z}$ can be thought of as a probability mass function on $\mathcal{X} \times \mathcal{Y}$, we can decompose it uniquely as: $\hat{z}(x, y) = \hat{z}(x)\hat{z}(y|x)$.

Decomposing the Entropy Measure $d(\cdot)$

Joint distortion criterion: Proof sketch

- Given a rate distortion code, $\hat{z}$ can be thought of as a probability mass function on $\mathcal{X} \times \mathcal{Y}$, we can decompose it uniquely as: $\hat{z}(x, y) = \hat{z}(x)\hat{z}(y|x)$.
- This allows the definition of the marginal and conditional distortions D_x and $D_{y|x}$:

$$D \geq \mathbb{E}\left[d(X^n, Y^n, \hat{Z}^n)\right]$$
$$= \underbrace{\mathbb{E}\left[d_x(X^n, \hat{Z}^n)\right]}_{D_x} + \underbrace{\mathbb{E}\left[d_{y|x}(X^n, Y^n, \hat{Z}^n)\right]}_{D_{y|x}}.$$

Decomposing the Entropy Measure $d(\cdot)$

Joint distortion criterion: Proof sketch

- Given a rate distortion code, $\hat{z}$ can be thought of as a probability mass function on $\mathcal{X} \times \mathcal{Y}$, we can decompose it uniquely as: $\hat{z}(x, y) = \hat{z}(x)\hat{z}(y|x)$.
- This allows the definition of the marginal and conditional distortions D_x and $D_{y|x}$:

$$D \geq \mathbb{E}\left[d(X^n, Y^n, \hat{Z}^n)\right]$$
$$= \underbrace{\mathbb{E}\left[d_x(X^n, \hat{Z}^n)\right]}_{D_x} + \underbrace{\mathbb{E}\left[d_{y|x}(X^n, Y^n, \hat{Z}^n)\right]}_{D_{y|x}}.$$

- This prompts the following lemma:

Lemma (Marginal and Conditional Distortion Preimages)

Define $\mathcal{A}_x(\hat{z}^n) = \{(x^n) : d_x(x^n, \hat{z}^n) \leq D_x + \epsilon\}$ *and* $\mathcal{A}_{y|x}(x^n, \hat{z}^n) = \left\{(y^n) : d_{y|x}(x^n, y^n, \hat{z}^n) \leq D_{y|x} + \epsilon\right\}$*. Then*

$$|\mathcal{A}_x(\hat{z}^n)| \leq 2^{n(D_x + 2\epsilon)} \text{ and } |\mathcal{A}_{y|x}(x^n, \hat{z}^n)| \leq 2^{n(D_{y|x} + 2\epsilon)}.$$

Proving Claim 1

Joint distortion criterion: Proof sketch

Claim 1

If (R_x, R_y, D) is an achievable RD point, then $(R_x + \theta D, R_y + (1-\theta)D)$ is an achievable Slepian-Wolf rate pair for some $\theta \in [0,1]$.

Proving Claim 1

Joint distortion criterion: Proof sketch

Claim 1

If (R_x, R_y, D) is an achievable RD point, then $(R_x + \theta D, R_y + (1-\theta)D)$ is an achievable Slepian-Wolf rate pair for some $\theta \in [0,1]$.

- Suppose we have a sequence of $(2^{nR_x}, 2^{nR_y}, n)$ rate-distortion codes achieving average distortion D.

Claim 1

If (R_x, R_y, D) is an achievable RD point, then $(R_x + \theta D, R_y + (1-\theta)D)$ is an achievable Slepian-Wolf rate pair for some $\theta \in [0,1]$.

- Suppose we have a sequence of $(2^{nR_x}, 2^{nR_y}, n)$ rate-distortion codes achieving average distortion D.
- With high probability, $\mathbb{E}\left[d_x(X^n, \hat{Z}^n)\right] \leq D_x + \epsilon$ and $\mathbb{E}\left[d_{y|x}(X^n, Y^n, \hat{Z}^n)\right] \leq D_{y|x} + \epsilon$. Also, $D_x + D_{y|x} \leq D + \epsilon$.

Claim 1

If (R_x, R_y, D) is an achievable RD point, then $(R_x + \theta D, R_y + (1-\theta)D)$ is an achievable Slepian-Wolf rate pair for some $\theta \in [0,1]$.

- Suppose we have a sequence of $(2^{nR_x}, 2^{nR_y}, n)$ rate-distortion codes achieving average distortion D.
- With high probability, $\mathbb{E}\left[d_x(X^n, \hat{Z}^n)\right] \le D_x + \epsilon$ and $\mathbb{E}\left[d_{y|x}(X^n, Y^n, \hat{Z}^n)\right] \le D_{y|x} + \epsilon$. Also, $D_x + D_{y|x} \le D + \epsilon$.
- Bin the x^n's into $2^{n(D_x+3\epsilon)}$ bins and send bin index along with rate-distortion codeword. This requires rate $R_x + D_x + 3\epsilon$ and allows the decoder to recover X^n w.h.p. by the MDPI lemma.

Proving Claim 1

Joint distortion criterion: Proof sketch

Claim 1

If (R_x, R_y, D) is an achievable RD point, then $(R_x + \theta D, R_y + (1-\theta)D)$ is an achievable Slepian-Wolf rate pair for some $\theta \in [0,1]$.

- Suppose we have a sequence of $(2^{nR_x}, 2^{nR_y}, n)$ rate-distortion codes achieving average distortion D.
- With high probability, $\mathbb{E}\left[d_x(X^n, \hat{Z}^n)\right] \leq D_x + \epsilon$ and $\mathbb{E}\left[d_{y|x}(X^n, Y^n, \hat{Z}^n)\right] \leq D_{y|x} + \epsilon$. Also, $D_x + D_{y|x} \leq D + \epsilon$.
- Bin the x^n's into $2^{n(D_x+3\epsilon)}$ bins and send bin index along with rate-distortion codeword. This requires rate $R_x + D_x + 3\epsilon$ and allows the decoder to recover X^n w.h.p. by the MDPI lemma.
- Bin the y^n's into $2^{n(D_{y|x}+3\epsilon)}$ bins and send bin index along with rate-distortion codeword. This requires rate $R_y + D_{y|x} + 3\epsilon$ and allows the decoder to recover Y^n w.h.p. by the CDPI lemma.

Proving Claim 2

Joint distortion criterion: Proof sketch

Claim 2

If $(R_x + \theta D, R_y + (1-\theta)D)$ is an achievable Slepian-Wolf rate pair for some $\theta \in [0,1]$, then (R_x, R_y, D) is an achievable RD point.

Claim 2

If $(R_x + \theta D, R_y + (1-\theta)D)$ is an achievable Slepian-Wolf rate pair for some $\theta \in [0,1]$, then (R_x, R_y, D) is an achievable RD point.

- Suppose $(\tilde{R}_x, \tilde{R}_y)$ is an achievable Slepian-Wolf rate pair.

Claim 2

If $(R_x + \theta D, R_y + (1-\theta)D)$ is an achievable Slepian-Wolf rate pair for some $\theta \in [0,1]$, then (R_x, R_y, D) is an achievable RD point.

- Suppose $(\tilde{R}_x, \tilde{R}_y)$ is an achievable Slepian-Wolf rate pair.
- Can assume $(\tilde{R}_x, \tilde{R}_y) = (1-\theta) \times (H(X), H(Y|X)) + \theta \times (H(X|Y), H(Y))$.

Proving Claim 2

Joint distortion criterion: Proof sketch

Claim 2

If $(R_x + \theta D, R_y + (1-\theta)D)$ is an achievable Slepian-Wolf rate pair for some $\theta \in [0,1]$, then (R_x, R_y, D) is an achievable RD point.

- Suppose $(\tilde{R}_x, \tilde{R}_y)$ is an achievable Slepian-Wolf rate pair.
- Can assume $(\tilde{R}_x, \tilde{R}_y) = (1-\theta) \times (H(X), H(Y|X)) + \theta \times (H(X|Y), H(Y))$.
- By the WZ corollary, $(R_x, R_y, D) = (\tilde{R}_x - \theta D, \tilde{R}_y - (1-\theta)D, D)$ is an achievable RD point.

Proving Claim 2

Joint distortion criterion: Proof sketch

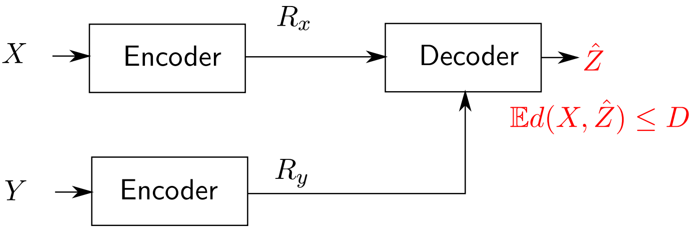
Claim 2

If $(R_x + \theta D, R_y + (1-\theta)D)$ is an achievable Slepian-Wolf rate pair for some $\theta \in [0,1]$, then (R_x, R_y, D) is an achievable RD point.

- Suppose $(\tilde{R}_x, \tilde{R}_y)$ is an achievable Slepian-Wolf rate pair.
- Can assume $(\tilde{R}_x, \tilde{R}_y) = (1-\theta) \times (H(X), H(Y|X)) + \theta \times (H(X|Y), H(Y))$.
- By the WZ corollary, $(R_x, R_y, D) = (\tilde{R}_x - \theta D, \tilde{R}_y - (1-\theta)D, D)$ is an achievable RD point.
- Making the substitution $(\tilde{R}_x, \tilde{R}_y) = (R_x + \theta D, R_y + (1-\theta)D)$ completes the proof.

d-Lossy Coding with Coded Side Information

Distortion constraint on X only



Theorem

$$\mathcal{R} = \left\{ (R_x, R_y, D) : \begin{array}{l} R_x \geq H(X|U) - D \\ R_y \geq I(Y;U) \\ \textit{for some distribution} \\ p(x,y,u) = p(x,y)p(u|y), \\ \textit{where } |\mathcal{U}| \leq |\mathcal{Y}| + 2 \end{array} \right\}$$

d-Lossy Coding with Coded Side Information

Distortion constraint on X only

- In this case, the achievable rate region is enlarged by D bits in the R_x-direction only. The idea is that Y is already at D_{max}, and thus any increase in D doesn't necessarily help decrease R_y.

Theorem

$$\mathcal{R} = \left\{ (R_x, R_y, D) : \begin{array}{l} R_x \geq H(X|U) - D \\ R_y \geq I(Y;U) \\ \textit{for some distribution} \\ p(x,y,u) = p(x,y)p(u|y), \\ \textit{where } |\mathcal{U}| \leq |\mathcal{Y}| + 2. \end{array} \right\}$$

d-Lossy Coding with Coded Side Information

Distortion constraint on X only

- In this case, the achievable rate region is enlarged by D bits in the R_x-direction only. The idea is that Y is already at D_{max}, and thus any increase in D doesn't necessarily help decrease R_y.
- The proof is similar in spirit to the previous case. In particular, we show a correspondence with the Ahlswede-Körner-Wyner region.

Theorem

$$\mathcal{R} = \left\{ (R_x, R_y, D) : \begin{array}{l} R_x \geq H(X|U) - D \\ R_y \geq I(Y;U) \\ \textit{for some distribution} \\ p(x,y,u) = p(x,y)p(u|y), \\ \textit{where } |\mathcal{U}| \leq |\mathcal{Y}| + 2. \end{array} \right\}$$

An Application

Gambling with Rate-Limited Side Information

- Suppose we bet on n i.i.d. horse races with outcomes X^n and correlated side information Y^n.

- Suppose we bet on n i.i.d. horse races with outcomes X^n and correlated side information Y^n.
- Our insider friend is only able to provide rate-limited side information $f(Y^n)$ at rate R_y.

- Suppose we bet on n i.i.d. horse races with outcomes X^n and correlated side information Y^n.
- Our insider friend is only able to provide rate-limited side information $f(Y^n)$ at rate R_y.
- Starting with 1 ruble, in the i^{th} race we bet a fraction $\hat{z}_i(x)$ of our wealth on horse x.

An Application

Gambling with Rate-Limited Side Information

- Suppose we bet on n i.i.d. horse races with outcomes X^n and correlated side information Y^n.
- Our insider friend is only able to provide rate-limited side information $f(Y^n)$ at rate R_y.
- Starting with 1 ruble, in the i^{th} race we bet a fraction $\hat{z}_i(x)$ of our wealth on horse x.
- The expected doubling rate is:

$$\mathbb{E}\frac{1}{n}\sum_{i=1}^{n}\log \hat{z}_i(x_i)o(x_i) = \mathbb{E}\log o(X) - \mathbb{E}d(X^n, \hat{Z}^n).$$

An Application

Gambling with Rate-Limited Side Information

- Suppose we bet on n i.i.d. horse races with outcomes X^n and correlated side information Y^n.
- Our insider friend is only able to provide rate-limited side information $f(Y^n)$ at rate R_y.
- Starting with 1 ruble, in the i^{th} race we bet a fraction $\hat{z}_i(x)$ of our wealth on horse x.
- The expected doubling rate is:

$$\mathbb{E}\frac{1}{n}\sum_{i=1}^{n}\log \hat{z}_i(x_i)o(x_i) = \mathbb{E}\log o(X) - \mathbb{E}d(X^n, \hat{Z}^n).$$

Theorem (Erkip 1996)

The optimum doubling rate is:

$$W^* = \mathbb{E}\log o(X) - D^*, \quad \textit{where}$$

$$D^* = \inf_{p(x,y,u)=p(u|y)p(x,y)} \left\{D : H(X|U) \le D, I(Y;U) \le R_y\right\}.$$

- Suppose we bet on n i.i.d. horse races with outcomes X^n and correlated side information Y^n.
- Our insider friend is only able to provide rate-limited side information $f(Y^n)$ at rate R_y, and an oracle provides side info $g(X^n)$ at rate R_x.
- Starting with 1 ruble, in the i^{th} race we bet a fraction $\hat{z}_i(x)$ of our wealth on horse x.
- The expected doubling rate is:

$$\mathbb{E}\frac{1}{n}\sum_{i=1}^{n}\log \hat{z}_i(x_i)o(x_i) = \mathbb{E}\log o(X) - \mathbb{E}d(X^n, \hat{Z}^n).$$

Theorem (One oracle bit adds more than one bit to doubling rate)

The optimum doubling rate is:

$$W^* = \mathbb{E}\log o(X) - D^*, \quad \textit{where}$$

$$D^* = \inf_{p(x,y,u)=p(u|y)p(x,y)} \left\{D : H(X|U) - R_x \leq D, I(Y;U) \leq R_y\right\}.$$

Concluding Remarks

- Two lossy source coding problems which are open in general can be solved for our choice of $d(\cdot)$:
  1. MTSC with a joint distortion constraint.
  2. MTSC with a single distortion constraint: $D_x = D$, $D_y = D_{max}$.

Concluding Remarks

- Two lossy source coding problems which are open in general can be solved for our choice of $d(\cdot)$:
  1. MTSC with a joint distortion constraint.
  2. MTSC with a single distortion constraint: $D_x = D$, $D_y = D_{max}$.
- It is possible that more cases can be solved.

Concluding Remarks

- Two lossy source coding problems which are open in general can be solved for our choice of $d(\cdot)$:
  1. MTSC with a joint distortion constraint.
  2. MTSC with a single distortion constraint: $D_x = D$, $D_y = D_{max}$.
- It is possible that more cases can be solved.
- The entropy-based distortion measure $d(\cdot)$ has shown up in the literature before (cf. image compression, the information bottleneck problem). This may point to some potential applications of the results presented today.

Concluding Remarks

- Two lossy source coding problems which are open in general can be solved for our choice of $d(\cdot)$:
  1. MTSC with a joint distortion constraint.
  2. MTSC with a single distortion constraint: $D_x = D$, $D_y = D_{max}$.
- It is possible that more cases can be solved.
- The entropy-based distortion measure $d(\cdot)$ has shown up in the literature before (cf. image compression, the information bottleneck problem). This may point to some potential applications of the results presented today.
- Due to the relationship between $d(\cdot)$ and relative entropy, there may be some additional applications in estimation/portfolio/gambling theory.

Concluding Remarks

- Two lossy source coding problems which are open in general can be solved for our choice of $d(\cdot)$:
  1. MTSC with a joint distortion constraint.
  2. MTSC with a single distortion constraint: $D_x = D$, $D_y = D_{max}$.
- It is possible that more cases can be solved.
- The entropy-based distortion measure $d(\cdot)$ has shown up in the literature before (cf. image compression, the information bottleneck problem). This may point to some potential applications of the results presented today.
- Due to the relationship between $d(\cdot)$ and relative entropy, there may be some additional applications in estimation/portfolio/gambling theory.
- Algorithmically satisfying from an engineering perspective.

Thank You!