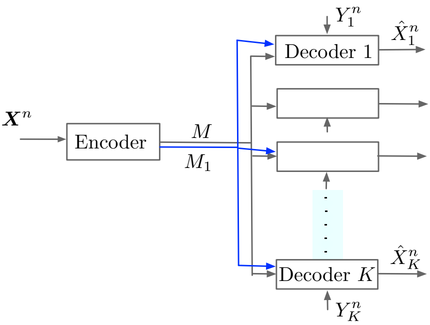
Variations of Source Coding with Side-Information at the Decoder(s)

Michèle Wigger
Telecom ParisTech
michele.wigger@telecom-paristech.fr

Chalmers University, 31 May 2013

joint work with T. Laich, A. Lapidoth, A. Malär, T. Oechtering, R. Timo

Source Coding with Side-Information



- A central processor stores data
- Different users want to reconstruct data
- Each user has SI about the data

- Users have SI because:
  - they can measure correlated data (e.g., correlated temperature measurements)
  - they have previously obtained descriptions of related data (previous queries)

Things we shall Address

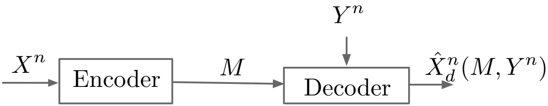
- Minimum description rate
- Encoder wishes to "control" decoder's reconstruction, even without knowing SI
- Usefulness of Encoder-SI

Part I:

Constraints on the Decoder's Reconstruction

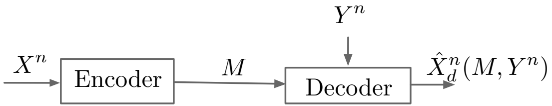
(Single Decoder)

Lossless Source Coding with SI; Single Decoder



- $\{(X_i, Y_i)\}$ IID $\sim P_{XY}$ over $\mathcal{X} \times \mathcal{Y}$
- Message $M \in \{1, \ldots, \lfloor 2^{nR} \rfloor\}$
- Side information Y^n known at decoder only!
- Decoder's source-reconstruction $\hat{X}_{\mathsf{d}}^n(M, Y^n)$ takes value in $\hat{\mathcal{X}}^n$
- Rate R achievable, if $\quad \overline{\lim_{n \to \infty}} \Pr\left[X^n \neq \hat{X}_{\mathsf{d}}^n\right] = 0$

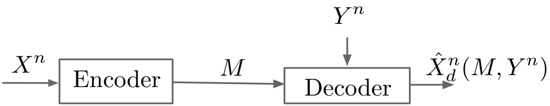
Lossless Source Coding with SI; Single Decoder



Slepian-Wolf '73: Infimum over achievable rates: $R^\star = H(X|Y)$

- Send bin-index of X^n to the decoder, which reconstructs X^n with Y^n
- "Coding language": Send syndrome of X^n of a code where each coset forms a good channel code for $X^n \to Y^n$

Lossy Source Coding with SI; Single Decoder

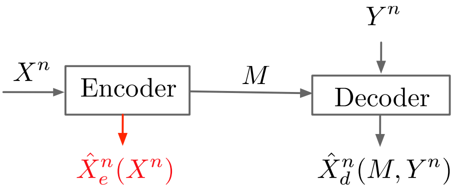


- Per-symbol distortion: $d_\mathsf{d} : \mathcal{X} \times \hat{\mathcal{X}} \to \mathbb{R}_0^+$
- (R, D_d) achievable if $\overline{\lim}_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \mathsf{E}\Big[d_\mathsf{d}(X_i, \hat{X}_{\mathsf{d},i})\Big] \le D_\mathsf{d}$

Wyner-Ziv'76: $R^\star_{\mathsf{WZ}}(D_\mathsf{d}) = \min_{\substack{Z, \hat{X}_\mathsf{d}(Z,Y) \text{ s.t.} \\ Z -\!\circ\!- X -\!\circ\!- Y \\ \mathsf{E}[d_\mathsf{d}(X, \hat{X}_\mathsf{d})] \le D_\mathsf{d}}} I(X; Z|Y)$ where $|\mathcal{Z}| = |\mathcal{X}| + 1$

$\to$ Encoder ignorant of Y^n $\Rightarrow$ cannot compute $\hat{X}_\mathsf{d}^n(M, Y^n)$!

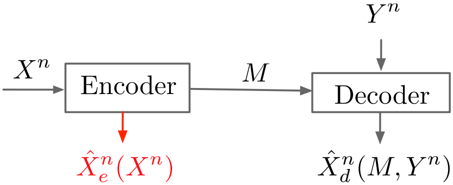
Wyner-Ziv with Common-Reconstruction Constraint (Steinberg'09)



- (R, D_{d}) achievable if:
  1. $\overline{\lim}_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \mathsf{E}\left[d_{\mathsf{d}}(X_i, \hat{X}_{\mathsf{d},i})\right] \leq D_{\mathsf{d}}$
  2. $\overline{\lim}_{n\to\infty} \mathsf{Pr}\left[\hat{X}_{\mathsf{e}}^n \neq \hat{X}_{\mathsf{d}}^n\right] = 0$

Steinberg'09: $$R_{\mathsf{CR}}(D_{\mathsf{d}}) = \min_{\substack{Z, \hat{X}_{\mathsf{d}}(Z) \text{ s.t.} \\ Z -\circ- X -\circ- Y \\ \mathsf{E}\left[d_{\mathsf{d}}(X, \hat{X}_{\mathsf{d}})\right] \leq D_{\mathsf{d}}}} I(X; Z|Y)$$

Wyner-Ziv with Lossy Common-Reconstruction Constraint



- Encoder-side distortion-function $d_\mathsf{e} \colon \hat{\mathcal{X}} \times \hat{\mathcal{X}} \to \mathbb{R}_0^+$
- $(R, D_\mathsf{d}, D_\mathsf{e})$ achievable if:
  1. $\overline{\lim_{n\to\infty}} \frac{1}{n} \sum_{i=1}^{n} \mathsf{E}\left[d_\mathsf{d}(X_i, \hat{X}_{\mathsf{d},i})\right] \leq D_\mathsf{d}$
  2. $\overline{\lim_{n\to\infty}} \frac{1}{n} \sum_{i=1}^{n} \mathsf{E}\left[d_\mathsf{e}(\hat{X}_{\mathsf{d},i}, \hat{X}_{\mathsf{e},i})\right] \leq D_\mathsf{e}$

Result: Rate-Distortions Function for Discrete Finite Alphabets

Theorem (Lapidoth/Malär/Wigger'11)

$$R_{\mathsf{lossyCR}}(D_{\mathsf{d}}, D_{\mathsf{e}}) = \min_{\substack{Z, \hat{X}_{\mathsf{d}}(Z,Y), \hat{X}_{\mathsf{e}}(Z,X) \text{ s.t.} \\ Z -\!\circ\!- X -\!\circ\!- Y \\ \mathsf{E}\left[d_{\mathsf{d}}(X,\hat{X}_{\mathsf{d}})\right] \leq D_{\mathsf{d}} \\ \mathsf{E}\left[d_{\mathsf{e}}(\hat{X}_{\mathsf{d}},\hat{X}_{\mathsf{e}})\right] \leq D_{\mathsf{e}}}} I(X;Z|Y)$$

where $|\mathcal{Z}| = |\mathcal{X}| + 3$ *suffices*

Corollary

When $d_{\mathsf{e}}(\hat{x}_{\mathsf{d}}, \hat{x}_{\mathsf{e}}, x) = I\{\hat{x}_{\mathsf{e}} \neq \hat{x}_{\mathsf{d}}\}$, *then* $R_{\mathsf{lossyCR}}(D_{\mathsf{d}}, 0) = R_{\mathsf{CR}}(D_{\mathsf{d}})$

Gaussian Sources and Quadratic Distortions

- $X \sim \mathcal{N}(0, \sigma_X^2)$
- $Y = X + U$, where $U \sim \mathcal{N}(0, \sigma_U^2)$ independent of X
- $d_{\mathsf{d}}(x, \hat{x}_{\mathsf{d}}) = (x - \hat{x}_{\mathsf{d}})^2$
- $d_{\mathsf{e}}(\hat{x}_{\mathsf{d}}, \hat{x}_{\mathsf{e}}) = (\hat{x}_{\mathsf{d}} - \hat{x}_{\mathsf{e}})^2$

Result: Rate-Distortions Function of Quadratic-Gaussian Setup

Theorem (Lapidoth/Malär/Wigger'11)

$$R_{\mathsf{lossyCR}}(D_{\mathsf{d}}, D_{\mathsf{e}}) = \begin{cases} \left[\frac{1}{2}\log\left(\frac{\sigma_X^2\sigma_U^2}{(\sigma_X^2+\sigma_U^2)D_{\mathrm{d}}}\right)\right]^+, & \text{if } \sqrt{D_{\mathsf{e}}\sigma_U^2} \geq \min\left\{D_{\mathsf{d}}, \frac{\sigma_X^2\sigma_U^2}{\sigma_X^2+\sigma_U^2}\right\} \\ \left[\frac{1}{2}\log\left(\frac{\sigma_X^2}{\sigma_X^2+\sigma_U^2}\,\frac{\sigma_U^2+D_{\mathrm{d}}-2\sqrt{\sigma_U^2 D_{\mathrm{e}}}}{D_{\mathrm{d}}-D_{\mathrm{e}}}\right)\right]^+, & \text{else.} \end{cases}$$

Corollary

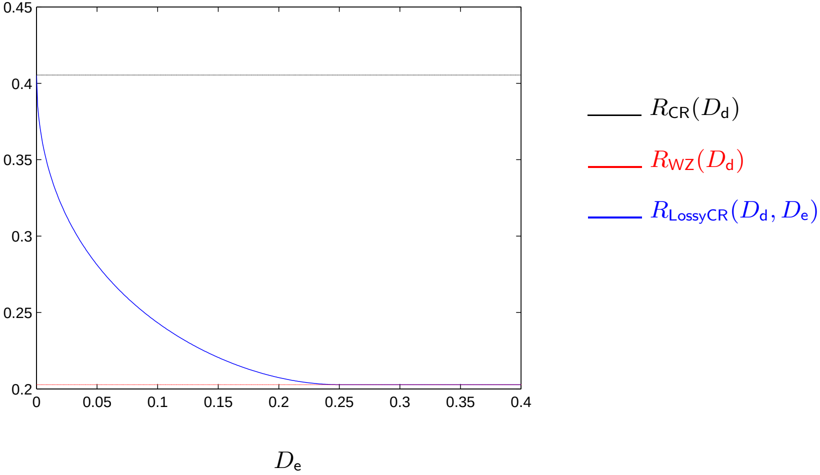
- *If* $\sqrt{D_{\mathsf{e}}\sigma_U^2} \geq \min\left\{D_{\mathsf{d}}, \frac{\sigma_X^2\sigma_U^2}{\sigma_X^2+\sigma_U^2}\right\}$ *or* $\left(1-\sqrt{\frac{D_{\mathsf{e}}}{\sigma_U^2}}\right)^2\sigma_X^2 \leq D_{\mathsf{d}} - D_{\mathsf{e}}$,

  then: $$R_{\mathsf{lossyCR}} = R_{\mathsf{WZ}} = R_{\mathsf{SI}}$$

- *If* $D_{\mathsf{e}} = 0$, *then* $R_{\mathsf{lossyCR}} = R_{\mathsf{CR}}$

Plots for Quadratic-Gaussian Setup

$$\sigma_X^2 = 3; \quad \sigma_U^2 = 1; \quad D_{\mathsf{d}} = 0.5$$



More General Reconstruction Constraints

- $K \geq 1$ distortion-functions $\quad d_k \colon \hat{\mathcal{X}} \times \hat{\mathcal{X}} \times X \to \mathbb{R}_0^+, \quad k \in \{1, \ldots, K\}$
- $(R, D_1, \ldots, D_K)$ achievable if:

$$\varlimsup_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \mathsf{E}\left[d_k(X_i, \hat{X}_{\mathsf{d},i}, \hat{X}_{\mathsf{e},i})\right] \leq D_k, \qquad k \in \{1, \ldots, K\}$$

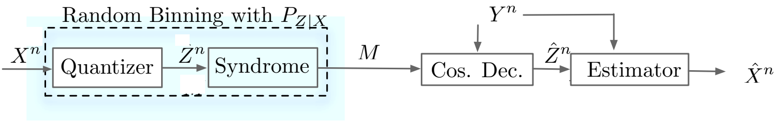
Theorem (Lapidoth/Malär/Wigger'11)

$$R^{\star}_{\mathsf{lossyCR}}(D_1, \ldots, D_K) = \min_{\substack{T,Z,\hat{X}_{\mathsf{d}}(Z,Y),\hat{X}_{\mathsf{e}}(Z,X,T) \text{ s.t.} \\ (T,Z) -\!\circ\!- X -\!\circ\!- Y \\ \mathsf{E}[d_k(X,\hat{X}_{\mathsf{d}},\hat{X}_{\mathsf{e}})] \leq D_k,\ k\in\{1,\ldots,K\}}} I(X;Z|Y)$$

where $|\mathcal{Z}| = |\mathcal{X}||\mathcal{T}| + K + 1$ *and* $|\mathcal{T}| = K$ *suffices*

Comparison of Coding Schemes

- Wyner-Ziv coding:



- Steinberg's coding: no estimator since encoder cannot reproduce it $\rightarrow \hat{X}^n = \hat{Z}^n$

- Our coding scheme: constrained estimator

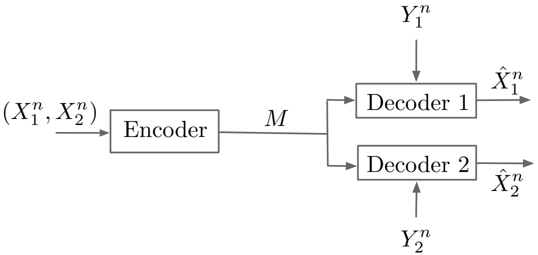
Challenges in the Converses

- Identification of "single-letter" variable $\hat{X}_{\mathsf{e}}$
- Gaussian case: proving optimality of $(\hat{X}_{\mathsf{d}}, \hat{X}_{\mathsf{e}})$ jointly Gaussian with (X, Y)
- Cardinality constraint on $\mathcal{U}$ in last setup $\rightarrow$ wish to recover our original result

Part II:

New Results on the Lossless Kaspi/Heegard-Berger Problem

The Lossless Heegard-Berger Problem with Two Sources



- $\{(X_{1,i}, X_{2,i}, Y_{1,i}, Y_{2,i})\}$ IID $\sim P_{X_1 X_2 Y_1 Y_2}$
- Decoder 1 wishes to learn X_1^n and Decoder 2 X_2^n
- Message $M \in \{1, \ldots, \lfloor 2^{nR} \rfloor\}$
- SI Y_1^n and Y_2^n known at the two decoders only!
- Rate R achievable, if $\quad \overline{\lim_{n\to\infty}} \Pr\left[X_1^n \neq \hat{X}_1^n \text{ and } X_2^n \neq \hat{X}_2^n\right] = 0$

Known Minimum Description Rates

- complementary SI $Y_1^n = X_2^n$ and $Y_2^n = X_1^n$ (Sgarro'77)

$$R^\star = \max\left\{H(X_1|Y_1), H(X_2|Y_2)\right\}$$

$\rightarrow$ send $X_1^n \oplus X_2^n$

- equal sources $X_1^n = X_2^n = X^n$ (Sgarro'77)

$$R^\star = \max_{i\in\{1,2\}} H(X|Y_i)$$

$\rightarrow$ send "random bin index" of X^n

- physically-degraded SI $(X_1^n, X_2^n) -\!\circ\!- Y_1^n -\!\circ\!- Y_2^n$ (Kaspi'94, Heegard/Berger'85)

$$R^\star = H(X_2|Y_2) + H(X_1|Y_1X_2)$$

$\rightarrow$ describe X_2^n to both decoders; describe X_1^n to decoder 1 which knows X_2^n, Y_1^n

Bounds on Minimum Description Rate

- Achievability:

$$R^\star \le \min_W \Big\{ \max \{ I(W;X_1X_2|Y_1), I(W;X_1X_2|Y_2) \} + H(X_1|WY_1) + H(X_2|WY_2) \Big\}$$

$\rightarrow$ send a "quantization" W^n to both decoders; then send X_1^n to Decoder 1 which knows W^n, Y_1^n and send X_2^n to Decoder 2 which knows W^n, Y_2^n

- Converses:
  - Reveal SI Y_2^n to Decoder 1 $\Rightarrow$ physically degraded setup

$$R^\star \ge H(X_2|Y_2) + H(X_1|X_2Y_1Y_2)$$

  - Single-decoder lower bound:

$$R^\star \ge \max_{i \in \{1,2\}} H(X_i|Y_i)$$

Conditionally Less-Noisy SI

Definition

Y_1 is conditionally less noisy than Y_2 given X_2, $(Y_1 \succeq Y_2|X_2)$, if

$$I(U;Y_1|X_2) \geq I(U;Y_2|X_2)$$

for all $U \multimap (X_1, X_2) \multimap (Y_1, Y_2)$.

- If the SI is physically degraded $(X_1, X_2) \multimap Y_1 \multimap Y_2$
- If $X_1 \multimap X_2 \multimap Y_2$

$$\left.\begin{array}{l} (X_1, X_2) \multimap Y_1 \multimap Y_2 \\ X_1 \multimap X_2 \multimap Y_2 \end{array}\right\} \Longrightarrow (Y_1 \succeq Y_2|X_2)$$

Result: Minimum Description Rate for Conditionally Less-Noisy SI

Timo/Oechtering/Wigger'12

Lemma (New Converse)

If $(Y_1 \succeq Y_2 | X_2)$, *then*

$$R^\star \geq H(X_2|Y_2) + H(X_1|X_2Y_1)$$

Theorem (Converse tight when also $H(X_2|Y_1) \leq H(X_2|Y_2)$)

If $(Y_1 \succeq Y_2 | X_2)$ *and* $H(X_2|Y_1) \leq H(X_2|Y_2)$, *then*

$$R^\star = H(X_2|Y_2) + H(X_1|X_2Y_1)$$

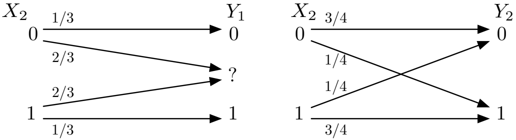
(achievability presented 2 slides ago, $W = X_2$)

Example: SI Conditionally Less-Noisy but not Physically Degraded

- X_2, Z independent Bernoulli-$1/2$ and-$1/3$

- $X_1 = X_2 \oplus Z$

- SI Y_1 and Y_2 defined by channels



- $H(X_2|Y_1) = 2/3 \quad < \quad H(X_2|Y_2) = H_b(1/4) \approx 0.8113$

- Minimum description rate:

$$R^{\star} = H_b(1/4) + H_b(1/3).$$

Converse based on Entropy-Characterization Problem

- Converse for physically degraded SI does not apply/cannot be extended

- Converse for conditionally less-noisy SI relies on:

Lemma (Entropy-Characterization Lemma)

Assume

$$(R^n, {S_1}^n, S_2^n, T^n, L^n) \text{ IID } \sim (R, S_1, S_2, T, L)$$

and

$$J \multimap\!\!- (R^n, L^n) \multimap\!\!- (S_1^n, S_2^n, T^n).$$

There exists a W *with cardinality constraint* $|\mathcal{W}| \leq |\mathcal{R}||\mathcal{L}|$ *such that*

$$I(J; S_2^n|L^n) - I(J; S_1^n|L^n) = n\big(I(W; S_2|L) - I(W; S_1|L)\big)$$

and $W \multimap\!\!- (R, L) \multimap\!\!- (S_1, S_2, T)$.

- Proof of lemma by Kramer's telescoping identity or Csiszar's sum-identity

Extensions

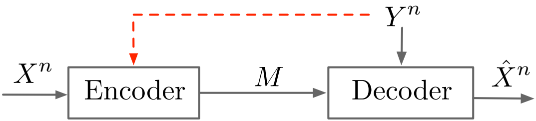
We can extend our result on minimum description length to:

- $K \geq 2$ decoders
- Partially lossy case $\rightarrow$ one decoder needs only a lossy reconstruction of its source
- Successive Refinement $\rightarrow$ one decoder obtains an additional private message

Part III:

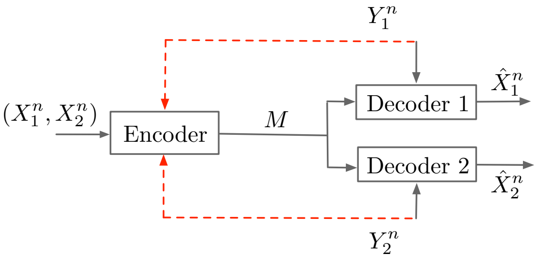
Utility of Encoder-SI (Lossless Kaspi/Heegard-Berger Problem)

Utility of Encoder-SI: Single Decoder



- Here: $M(X^n, Y^n)$
- $R^\star_{\text{cogn}}$: minimum achievable rate with encoder-SI ($R^\star_{\text{ign}}$ without encoder-SI)
- $R^\star_{\text{cogn}}$ with source X^n equals $R^\star_{\text{ign}}$ with modified source (X^n, Y^n)
- Lossless case: $R^\star_{\text{cogn}} = R^\star_{\text{ign}} = H(X|Y) \rightarrow$ encoder-SI useless!
- Lossy case: $R^\star_{\text{cogn}} \leq R^\star_{\text{ign}} = R^\star_{\text{WZ}} \rightarrow$ encoder-SI can help!
  (not in Quadratic-Gaussian setup)

Heegard-Berger Setup with 2 Decoders, 2 Sources, Encoder-SI



- With encoder-SI: $M(X_1^n, X_2^n, Y_1^n, Y_2^n) \to$ minimum description rate R_{cogn}
- Without encoder-SI: $M(X_1^n, X_2^n) \to$ minimum description rate R_{ign}
- R_{cogn} with sources X_1^n and X_2^n equals R_{ign} with modified sources (X_1^n, Y_1^n) and (X_2^n, Y_2^n)

Utility of Encoder-SI?

$$\text{Can} \quad R^\star_{\text{cogn}} < R^\star_{\text{ign}} \quad ?$$

- Yes, for lossy case, e.g., for Gaussian sources and physically degraded Gaussian SI (Kaspi'94, Perron/Diggavi/Telatar'06)

- No, for lossless case for (Sgarro'77):
  - equal sources: $X_1^n = X_2^n$
  - complementary SI: $Y_1^n = X_2^n$ and $Y_2^n = X_1^n$

  $$\Bigg\} \Longrightarrow R^\star_{\text{cogn}} = R^\star_{\text{ign}}$$

- General lossless case?

Result: More Lossless Setups where Encoder-SI NOT Useful

Theorem (Laich/Wigger'13)

Encoder-SI useless when:

- *physically degraded SI:* $(X_1, X_2) -\!\!\circ\!\!- Y_1 -\!\!\circ\!\!- Y_2$
- $X_1 -\!\!\circ\!\!- (X_2, Y_1) -\!\!\circ\!\!- Y_2$ *and* $H(X_2|Y_1) \leq H(X_2|Y_2)$
  (subset of conditionally less-noisy SI)
- *"Noisy Complementary SI":* $X_2 -\!\!\circ\!\!- (X_1, Y_1) -\!\!\circ\!\!- Y_2$ *and* $X_1 -\!\!\circ\!\!- (X_2, Y_2) -\!\!\circ\!\!- Y_1$

Ex: $\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim \text{DSBS}(p), \quad Y_1 = \begin{cases} X_2 & E_1 = 0 \\ ? & E_1 = 1 \end{cases}, \quad Y_2 = \begin{cases} X_1 & E_2 = 0 \\ ? & E_2 = 1 \end{cases}$

Result: Encoder-SI Useful for following example!

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim \text{DSBS}(p); \qquad \tilde{Y}_k = \begin{cases} X_1, X_2 & \text{if } E_k = 0 \\ ? & \text{if } E_k = 1 \end{cases}; \qquad Y_k = (\tilde{Y}_k, E_1, E_2)$$

where

$$\Pr[E_1 = 1, E_2 = 0] = q; \qquad \Pr[E_1 = 0, E_2 = 1] = 1 - q; \qquad q \leq 1/3$$

Encoder-SI strictly decreases minimum description rate!

$$R^\star_{\text{cogn}} = H(X_2|Y_2) + H(X_1|X_2Y_1Y_2) = 1 - q$$
$$R^\star_{\text{ign}} = H(X_2|Y_2) + H(X_1|X_2Y_1) = 1 - q + qH_b(p)$$

With Enc-SI:

1. describe (X_2^n, Y_2^n) to both decoders $\rightarrow$ describing also Y_2^n needs no extra rate (!) since $H(X_2|Y_2) = H(X_2Y_2|Y_2) \geq H(X_2Y_2|Y_1)$;

2. describe X_1^n to Decoder 1 which knows X_2^n, Y_1^n, Y_2^n

Result: Encoder-SI Useful for following example!

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim \mathsf{DSBS}(p); \qquad \tilde{Y}_k = \begin{cases} X_1, X_2 & \text{if } E_k = 0 \\ ? & \text{if } E_k = 1 \end{cases}; \qquad Y_k = (\tilde{Y}_k, E_1, E_2)$$

where

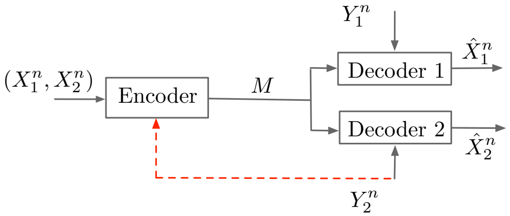
$$\mathsf{Pr}[E_1 = 1, E_2 = 0] = q; \qquad \mathsf{Pr}[E_1 = 0, E_2 = 1] = 1 - q; \qquad q \leq 1/3$$

Encoder-SI strictly decreases minimum description rate!

$$R^\star_{\mathsf{cogn}} = H(X_2|Y_2) + H(X_1|X_2 Y_1 Y_2) = 1 - q$$
$$R^\star_{\mathsf{ign}} = H(X_2|Y_2) + H(X_1|X_2 Y_1) = 1 - q + q H_b(p)$$

- in our scheme Decoder 1 learns (X_1^n, X_2^n)
- scheme & rates apply also to degraded source sets where Dec. 1 needs (X_1^n, X_2^n)

Result: Even Partial (One-Sided) Encoder-SI can be Strictly Useful



$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim \text{DSBS}(p); \qquad Y_1 = \begin{cases} X_1 & \text{if } E_1 = 0 \\ ? & \text{if } E_1 = 1 \end{cases}; \qquad Y_2 = \begin{cases} X_2, E_1 & \text{if } E_2 = 0 \\ ?, E_1 & \text{if } E_2 = 1 \end{cases}$$

where

$$q_2 \triangleq \Pr[E_2 = 1]; \qquad q_1 \triangleq \Pr[E_1 = 1]; \qquad \Pr[E_1 = E_2 = 1] = q_e$$

$$R^\star_{\text{partial-cogn}} = q_2 + H_b(p) + \max\left\{0, (q_1 - q_2)(1 - H_b(p)) - (q_2 - q_e)H_b(p)\right\}$$
$$R^\star_{\text{ign}} = q_2 + H_b(p) + \max\left\{0, (q_1 - q_2)(1 - H_b(p))\right\}.$$

→ if $q_2 > q_1 > q_e$, then $R^\star_{\text{partial-cogn}} < R^\star_{\text{ign}}$

Summary

Wyner-Ziv Problem with *Lossy* Encoder-Decoder Reconstruction-Constraints

- Rate-distortions function (single-letter) for discrete case
- Rate-distortions function for quadratic-Gaussian case

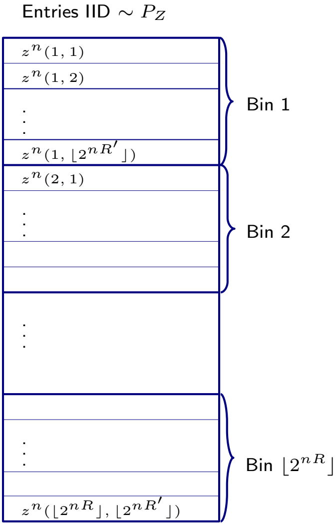
Minimum Description Rate for Lossless Heegard-Berger problem

- Solution for *conditionally less noisy* SI
- Also for partially lossy problem or successive refinement problem
- Converse with new "entropy-characterization lemma"

Utility of Encoder SI for 2-sources HB problem

- Encoder-SI strictly useful! Also with degraded source sets or partial encoder-SI
- Intuition: sometimes can describe SI Y_1^n to Decoder 2 for free!

Wyner-Ziv's Scheme



- Encoding:
  - Choose M, K s.t.

    $$(Z^n(M,K), X^n) \in \mathcal{T}_\epsilon^{(n)}(P_{ZX})$$

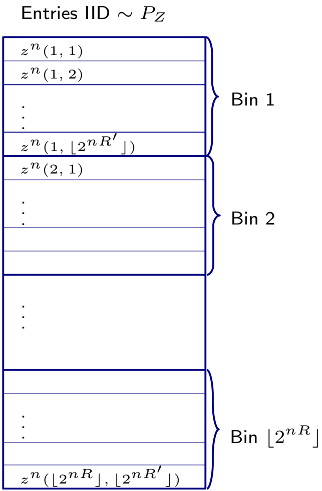
  - Message M is bin-index!
- Decoding:
  - Binning phase: Look for $\hat{K}$ s.t.

    $$(Z^n(M,\hat{K}), Y^n) \in \mathcal{T}_\epsilon^{(n)}(P_{ZY})$$

  - Estimation phase:
    $\hat{X}_{\mathsf{d},i} = \phi(Z_i(M,\hat{K}), Y_i)$

With high prob: $Z^n(M,K) = Z^n(M,\hat{K})$

Steinberg's Scheme



- Encoding:
  - Choose M, K s.t.

    $$(Z^n(M,K), X^n) \in \mathcal{T}_\epsilon^{(n)}(P_{ZX})$$

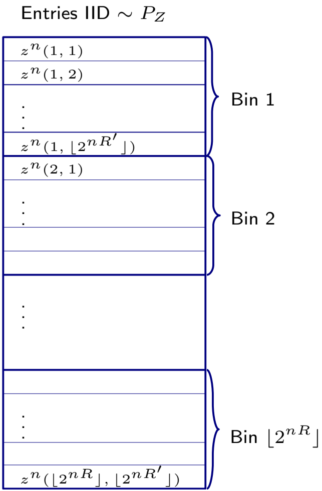
  - Message M is bin-index!
  - $\hat{X}_\mathsf{e}^n = Z^n(M,K)$
- Decoding:
  - Binning phase: Look for $\hat{K}$ s.t.

    $$(Z^n(M,\hat{K}), Y^n) \in \mathcal{T}_\epsilon^{(n)}(P_{ZY})$$

  - "Estimation phase": $\hat{X}_\mathsf{d}^n = Z^n(M,\hat{K})$

Estimation phase independent of Y^n!

Our Scheme



- Encoding:
  - Choose M, K s.t.

    $$(Z^n(M,K), X^n) \in \mathcal{T}_\epsilon^{(n)}(P_{ZX})$$

  - Message M is bin-index!
  - $\hat{X}_{\mathsf{e}}^n = \psi(Z_i(M,K), X_i)$
- Decoding:
  - Binning phase: Look for $\hat{K}$ s.t.

    $$(Z^n(M,\hat{K}), Y^n) \in \mathcal{T}_\epsilon^{(n)}(P_{ZY})$$

  - Estimation phase: $\hat{X}_{\mathsf{d},i} = \phi(Z_i(M,\hat{K}), Y_i)$

Estimation phase can *moderately* depend on Y^n!

Achievability in Quadratic-Gaussian Case

- Previous achievability fails (strong typicality!)
- New achievability: similar, but with coding over spheres

Converse for Discrete Case

Converse: $$R_{\text{lossyCR}}(D_{\text{d}}, D_{\text{e}}) \geq \bar{R}(D_{\text{d}}, D_{\text{e}}) \triangleq \min_{\substack{Z, \hat{X}_{\text{d}}(Z,Y), \hat{X}_{\text{e}}(Z,X) \\ \text{s.t.:}\ Z -\!\!\circ\!\!- X -\!\!\circ\!\!- Y \\ \mathsf{E}\left[d_{\text{d}}(X,\hat{X}_{\text{d}})\right] \leq D_{\text{d}} \\ \mathsf{E}\left[d_{\text{e}}(\hat{X}_{\text{d}},\hat{X}_{\text{e}})\right] \leq D_{\text{e}}}} I(X;Z|Y)$$

a) Relax source coding problem, i.e., relax 2. distortion constraint

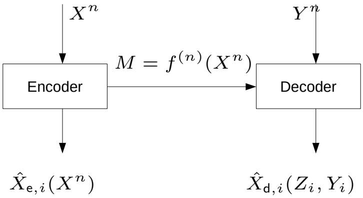
Then: $$R_{\text{lossyCR}}(D_{\text{d}}, D_{\text{e}}) \geq R_{\text{Relaxed}}(D_{\text{d}}, D_{\text{e}})$$

b) Converse to relaxed problem:

$$R_{\text{Relaxed}}(D_{\text{d}}, D_{\text{e}}) \geq \bar{R}(D_{\text{d}}, D_{\text{e}})$$

Converse Step a): Relax Source-Coding Problem

original problem

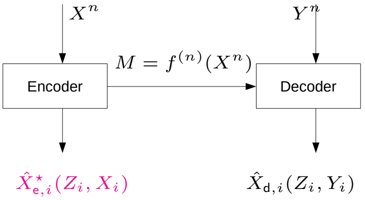


1. $\overline{\lim}_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \mathsf{E}\left[d_{\mathsf{d}}(X_i, \hat{X}_{\mathsf{d},i})\right] \leq D_{\mathsf{d}}$
2. $\overline{\lim}_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \mathsf{E}\left[d_{\mathsf{e}}(\hat{X}_{\mathsf{d},i}, \hat{X}_{\mathsf{e},i})\right] \leq D_{\mathsf{e}}$

- Define $Z_i \triangleq (M, Y^{i-1}, Y_{i+1}^{n})$

Converse Step a): Relax Source-Coding Problem

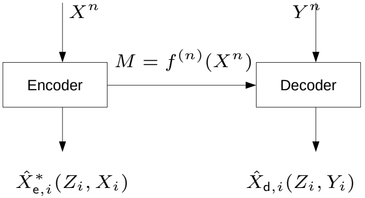
relaxed problem



1. $\overline{\lim_{n\to\infty}} \frac{1}{n} \sum_{i=1}^{n} \mathsf{E}\left[d_{\mathsf{d}}(X_i, \hat{X}_{\mathsf{d},i})\right] \le D_{\mathsf{d}}$
2. $\overline{\lim_{n\to\infty}} \frac{1}{n} \sum_{i=1}^{n} \mathsf{E}\left[d_{\mathsf{e}}(\hat{X}_{\mathsf{d},i}, \hat{X}^\star_{\mathsf{e},i})\right] \le D_{\mathsf{e}}$

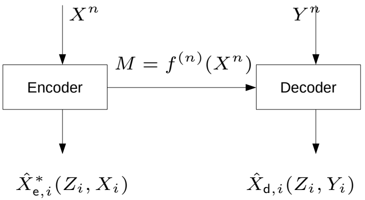
- Define $Z_i \triangleq (M, Y^{i-1}, Y^n_{i+1})$
- Because of $X^n \multimap (Z_i, X_i) \multimap (Z_i, Y_i)$: new constraint 2. weaker

Converse Step b): Converse to Relaxed Problem



1. $\overline{\lim_{n\to\infty}} \dfrac{1}{n} \sum_{i=1}^{n} \mathsf{E}\left[d_{\mathsf{d}}(X_i, \hat{X}_{\mathsf{d},i})\right] \le D_{\mathsf{d}}$
2. $\overline{\lim_{n\to\infty}} \dfrac{1}{n} \sum_{i=1}^{n} \mathsf{E}\left[d_{\mathsf{e}}(\hat{X}_{\mathsf{d},i}, \hat{X}^*_{\mathsf{e},i})\right] \le D_{\mathsf{e}}$

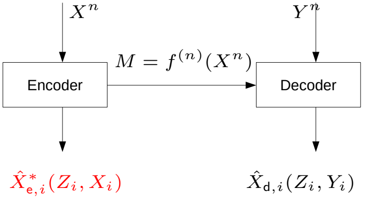
Converse Step b): Converse to Relaxed Problem



1. $\overline{\lim\limits_{n\to\infty}} \frac{1}{n}\sum_{i=1}^{n} \mathsf{E}\left[d_{\mathsf{d}}(X_i, \hat{X}_{\mathsf{d},i})\right] \le D_{\mathsf{d}}$
2. $\overline{\lim\limits_{n\to\infty}} \frac{1}{n}\sum_{i=1}^{n} \mathsf{E}\left[d_{\mathsf{e}}(\hat{X}_{\mathsf{d},i}, \hat{X}^*_{\mathsf{e},i})\right] \le D_{\mathsf{e}}$

- By definition $Z_i \triangleq (M, Y^{i-1}, Y^n_{i+1})$: $\quad Z_i - \circ - X_i - \circ - Y_i$
- $R_{\mathsf{Relaxed}} \ge \frac{1}{n} H(M) \ge \ldots \ge \frac{1}{n}\sum_{i=1}^{n} I(X_i; Z_i | Y_i)$

Converse Step b): Converse to Relaxed Problem



1. $\overline{\lim}_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \underbrace{\mathsf{E}\left[d_{\mathsf{d}}(X_i, \hat{X}_{\mathsf{d},i})\right]}_{D_{\mathsf{d},i}} \leq D_{\mathsf{d}}$

2. $\overline{\lim}_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \underbrace{\mathsf{E}\left[d_{\mathsf{e}}(\hat{X}_{\mathsf{d},i}, \hat{X}^*_{\mathsf{e},i})\right]}_{D_{\mathsf{e},i}} \leq D_{\mathsf{e}}$

- By definition $Z_i \triangleq (M, Y^{i-1}, Y_{i+1}^n)$: $\quad Z_i \multimap\!\!- X_i \multimap\!\!- Y_i$

- $R_{\text{Relaxed}} \geq \frac{1}{n}H(M) \geq \ldots \geq \frac{1}{n}\sum_{i=1}^{n} I(X_i; Z_i|Y_i) \geq \frac{1}{n}\sum_{i=1}^{n} \bar{R}(D_{\mathsf{d},i}, D_{\mathsf{e},i})$

$$\geq \bar{R}\left(\frac{1}{n}\sum_{i=1}^{n} D_{\mathsf{d},i}, \frac{1}{n}\sum_{i=1}^{n} D_{\mathsf{e},i}\right) \geq \bar{R}(D_{\mathsf{d}}, D_{\mathsf{e}})$$

Converse in Quadratic-Gaussian Case

- $X \sim \mathcal{N}(0, \sigma_X^2)$
- $Y = X + U$, where $U \sim \mathcal{N}(0, \sigma_U^2)$ independent of X
- $d_{\mathsf{d}}(x, \hat{x}_{\mathsf{d}}) = (x - \hat{x}_{\mathsf{d}})^2$ and $d_{\mathsf{e}}(\hat{x}_{\mathsf{d}}, \hat{x}_{\mathsf{e}}) = (\hat{x}_{\mathsf{d}} - \hat{x}_{\mathsf{e}})^2$

$$R_{\mathsf{lossyCR}}(D_{\mathsf{d}}, D_{\mathsf{e}}) \geq \begin{cases} \left[\frac{1}{2}\log\left(\frac{\sigma_X^2\sigma_U^2}{(\sigma_X^2+\sigma_U^2)D_{\mathrm{d}}}\right)\right]^+, & \text{if } \sqrt{D_{\mathsf{e}}\sigma_U^2} \geq \min\left\{D_{\mathsf{d}}, \frac{\sigma_X^2\sigma_U^2}{\sigma_X^2+\sigma_U^2}\right\} \\ \left[\frac{1}{2}\log\left(\frac{\sigma_X^2}{\sigma_X^2+\sigma_U^2}\frac{\sigma_U^2+D_{\mathrm{d}}-2\sqrt{\sigma_U^2 D_{\mathrm{e}}}}{D_{\mathrm{d}}-D_{\mathrm{e}}}\right)\right]^+, & \text{else.} \end{cases}$$

Converse in Quadratic-Gaussian Case, First Step

Step 1:
$$R_{\mathsf{lossyCR}}(D_{\mathsf{d}}, D_{\mathsf{e}}) \geq \inf_{\substack{Z, \hat{X}_{\mathsf{d}}(Z,Y), \hat{X}_{\mathsf{e}}(Z,X) \\ \text{s.t.:} \quad Z -\!\circ\!- X -\!\circ\!- Y \\ \mathsf{E}\left[(X-\hat{X}_{\mathsf{d}})^2\right] \leq D_{\mathsf{d}} \\ \mathsf{E}\left[(\hat{X}_{\mathsf{d}}-\hat{X}_{\mathsf{e}})^2\right] \leq D_{\mathsf{e}}}} I(X;Z|Y) \qquad (1)$$

Step 2-: Evaluate RHS(1);

Converse in Quadratic-Gaussian Case, First Step

Step 1: $$R_{\mathsf{lossyCR}}(D_{\mathsf{d}}, D_{\mathsf{e}}) \geq h(X|Y) - \sup_{\substack{Z, \hat{X}_{\mathsf{d}}(Z,Y), \hat{X}_{\mathsf{e}}(Z,X) \\ \text{s.t.: } Z \multimap X \multimap Y \\ \mathsf{E}\left[(X - \hat{X}_{\mathsf{d}})^2\right] \leq D_{\mathsf{d}} \\ \mathsf{E}\left[(\hat{X}_{\mathsf{d}} - \hat{X}_{\mathsf{e}})^2\right] \leq D_{\mathsf{e}}}} h(X|YZ) \qquad (1)$$

Step 2-: Evaluate RHS(1); First Thoughts:

- Conditional Max-Entropy Theorem:
  Given $K_{XYZ\hat{X}_{\mathsf{d}}\hat{X}_{\mathsf{e}}}$ Gaussian tuple $(Z, \hat{X}_{\mathsf{d}}(Z,Y), \hat{X}_{\mathsf{e}}(Z,X))$ optimizes (1)
- Not $\forall$ $K_{XYZ\hat{X}_{\mathsf{d}}\hat{X}_{\mathsf{e}}}$ the Gaussian tuple is valid because $\hat{X}_{\mathsf{d}}(Z,Y)$ and $\hat{X}_{\mathsf{e}}(Z,X)$
- If we relax $\hat{X}_{\mathsf{d}}(Z,Y)$ and $\hat{X}_{\mathsf{e}}(Z,X)$ $\Rightarrow$ RHS(1)=0 (too low!)

Converse in Quadratic-Gaussian Case, Further Steps

Step 1:
$$R_{\mathsf{lossyCR}}(D_{\mathsf{d}}, D_{\mathsf{e}}) \geq h(X|Y) - \sup_{\substack{Z, \hat{X}_{\mathsf{d}}(Z,Y), \hat{X}_{\mathsf{e}}(Z,X) \\ \text{s.t.: } Z -\!\circ\!- X -\!\circ\!- Y \\ \mathsf{E}\left[(X-\hat{X}_{\mathsf{d}})^2\right] \leq D_{\mathsf{d}} \\ \mathsf{E}\left[(\hat{X}_{\mathsf{d}}-\hat{X}_{\mathsf{e}})^2\right] \leq D_{\mathsf{e}}}} h(X|YZ) \tag{1}$$

Step 2: RHS(1) lower bounded by:
$$h(X|Y) - \sup_{\substack{\hat{X}_{\mathsf{d}} \text{ s.t.:} \\ \mathsf{E}\left[(X-\hat{X}_{\mathsf{d}})^2\right] \leq D_{\mathsf{d}} \\ \left|\mathsf{E}\left[(X-\hat{X}_{\mathsf{d}})U\right]\right| \leq \sqrt{\sigma_U^2 D_{\mathsf{e}}}}} h(X - \hat{X}_{\mathsf{d}} | Y - \hat{X}_{\mathsf{d}}, \hat{X}_{\mathsf{d}}) \tag{2}$$

Step 3: (2) maximized by jointly Gaussian $(\hat{X}_{\mathsf{d}}, X, U)$ (cond. max-entropy thm)

Step 4: Evaluate (2) for jointly Gaussian $(\hat{X}_{\mathsf{d}}, X, U)$

Step 2-I: Apply $\hat{X}_{\mathsf{d}}(Z,Y)$ to transform Objective Function

- Because $\hat{X}_{\mathsf{d}}(Z,Y)$:

$$
\begin{aligned}
h(X|Y,Z) = h(X|Y,Z,\hat{X}_{\mathsf{d}}) &= h(X-\hat{X}_{\mathsf{d}}|Y-\hat{X}_{d},Z,\hat{X}_{\mathsf{d}}) \\
&\le h(X-\hat{X}_{\mathsf{d}}|X-\hat{X}_{\mathsf{d}}+U,\hat{X}_{\mathsf{d}})
\end{aligned}
$$

Step 2-I:

$$
R_{\mathsf{lossyCR}}(D_{\mathsf{d}},D_{\mathsf{e}}) \ge h(X|Y) - \sup_{\substack{Z,\hat{X}_{\mathsf{d}}(Z,Y),\hat{X}_{\mathsf{e}}(Z,X) \\ \text{s.t.: } Z - \circ - X - \circ - Y \\ \mathsf{E}\left[(X-\hat{X}_{\mathsf{d}})^2\right] \le D_{\mathsf{d}} \\ \mathsf{E}\left[(\hat{X}_{\mathsf{d}}-\hat{X}_{\mathsf{e}})^2\right] \le D_{\mathsf{e}}}} h(X-\hat{X}_{\mathsf{d}}|X-\hat{X}_{\mathsf{d}}+U,\hat{X}_{\mathsf{d}})
$$

Step 2-I: Apply $\hat{X}_{\mathsf{d}}(Z,Y)$ to transform Objective Function

- Because $\hat{X}_{\mathsf{d}}(Z,Y)$:

$$\begin{aligned} h(X|Y,Z) &= h(X|Y,Z,\hat{X}_{\mathsf{d}}) = h(X-\hat{X}_{\mathsf{d}}|Y-\hat{X}_{d},Z,\hat{X}_{\mathsf{d}}) \\ &\leq h(X-\hat{X}_{\mathsf{d}}|X-\hat{X}_{\mathsf{d}}+U,\hat{X}_{\mathsf{d}}) \end{aligned}$$

Step 2-I:

$$R_{\mathsf{lossyCR}}(D_{\mathsf{d}},D_{\mathsf{e}}) \geq h(X|Y) - \sup_{\substack{Z,\hat{X}_{\mathsf{d}}(Z,Y),\hat{X}_{\mathsf{e}}(Z,X) \\ \text{s.t.:}\ Z -\circ- X -\circ- Y \\ \mathsf{E}\left[(X-\hat{X}_{\mathsf{d}})^2\right]\leq D_{\mathsf{d}} \\ \mathsf{E}\left[(\hat{X}_{\mathsf{d}}-\hat{X}_{\mathsf{e}})^2\right]\leq D_{\mathsf{e}}}} h(X-\hat{X}_{\mathsf{d}}|X-\hat{X}_{\mathsf{d}}+U,\hat{X}_{\mathsf{d}})$$

- Relax function-constraint now $\rightarrow$ Wyner-Ziv result (too loose)

- First need to use $\hat{X}_{\mathsf{e}}(Z,X)$ to limit dependence of $\hat{X}_{\mathsf{d}}$ on U

Step 2-II: Apply $\hat{X}_{\mathsf{e}}(Z,X)$ to transform Constraints

- $Z \multimap\!\!- X \multimap\!\!- Y = X + U \Rightarrow (X,Z)$ ind. of U
- $\hat{X}_{\mathsf{e}}(Z,X)$ & Constraint $\mathsf{E}\left[(\hat{X}_{\mathsf{d}} - \hat{X}_{\mathsf{e}})^2\right] \leq D_{\mathsf{e}}$:

$$\left|\mathsf{E}\left[\hat{X}_{\mathsf{d}} \cdot U\right]\right| = \left|\mathsf{E}\left[(\hat{X}_{\mathsf{d}} - \hat{X}_{\mathsf{e}})U\right]\right| \leq \sqrt{\sigma_U^2 D_{\mathsf{e}}} \tag{3}$$

Step 2-II: relax constraints

$$R_{\mathsf{lossyCR}}(D_{\mathsf{d}}, D_{\mathsf{e}}) \geq h(X|Y) - \sup_{\substack{Z, \hat{X}_{\mathsf{d}}(Z,Y), \hat{X}_{\mathsf{e}}(Z,X) \\ \text{s.t.:} \quad (Z,X) \text{ ind. of } U \\ \mathsf{E}\left[(X-\hat{X}_{\mathsf{d}})^2\right] \leq D_{\mathsf{d}} \\ \left|\mathsf{E}\left[\hat{X}_{\mathsf{d}} U\right]\right| \leq \sqrt{\sigma_U^2 D_{\mathsf{e}}}}} h(X - \hat{X}_{\mathsf{d}} | X - \hat{X}_{\mathsf{d}} + U, \hat{X}_{\mathsf{d}})$$

Step 2-II: Apply $\hat{X}_{\mathsf{e}}(Z,X)$ to transform Constraints

- $Z - \circ - X - \circ - Y = X + U \Rightarrow (X,Z)$ ind. of U
- $\hat{X}_{\mathsf{e}}(Z,X)$ & Constraint $\mathsf{E}\left[(\hat{X}_{\mathsf{d}} - \hat{X}_{\mathsf{e}})^2\right] \leq D_{\mathsf{e}}$:

$$\left|\mathsf{E}\left[\hat{X}_{\mathsf{d}} \cdot U\right]\right| = \left|\mathsf{E}\left[(\hat{X}_{\mathsf{d}} - \hat{X}_{\mathsf{e}})U\right]\right| \leq \sqrt{\sigma_U^2 D_{\mathsf{e}}} \quad (3)$$

Step 2-II: relax constraints

$$R_{\mathsf{lossyCR}}(D_{\mathsf{d}}, D_{\mathsf{e}}) \geq h(X|Y) - \sup_{\substack{Z, \hat{X}_{\mathsf{d}}(Z,Y), \hat{X}_{\mathsf{e}}(Z,X) \\ \text{s.t.: } (Z,X) \text{ ind. of } U \\ \mathsf{E}\left[(X - \hat{X}_{\mathsf{d}})^2\right] \leq D_{\mathsf{d}} \\ \left|\mathsf{E}\left[\hat{X}_{\mathsf{d}} U\right]\right| \leq \sqrt{\sigma_U^2 D_{\mathsf{e}}}}} h(X - \hat{X}_{\mathsf{d}} | X - \hat{X}_{\mathsf{d}} + U, \hat{X}_{\mathsf{d}})$$

- Relax function constraints now