

Multiterminal source coding with complementary delivery

Akisato Kimura^{*1,*2}, Tomohiko Uyematsu^{*1}

^{*1} Department of Communications and Integrated Systems,
Graduate School of Science and Engineering,
Tokyo Institute of Technology

^{*2} Media Information Laboratory,
NTT Communication Science Laboratories,
NTT Corporation

Abstract

- Investigate a coding problem for correlated information sources, where
 - messages from two correlated sources are jointly encoded
 - each decoder has access to one of two messages to reproduce the other message
- Clarify the rate-distortion function for this problem

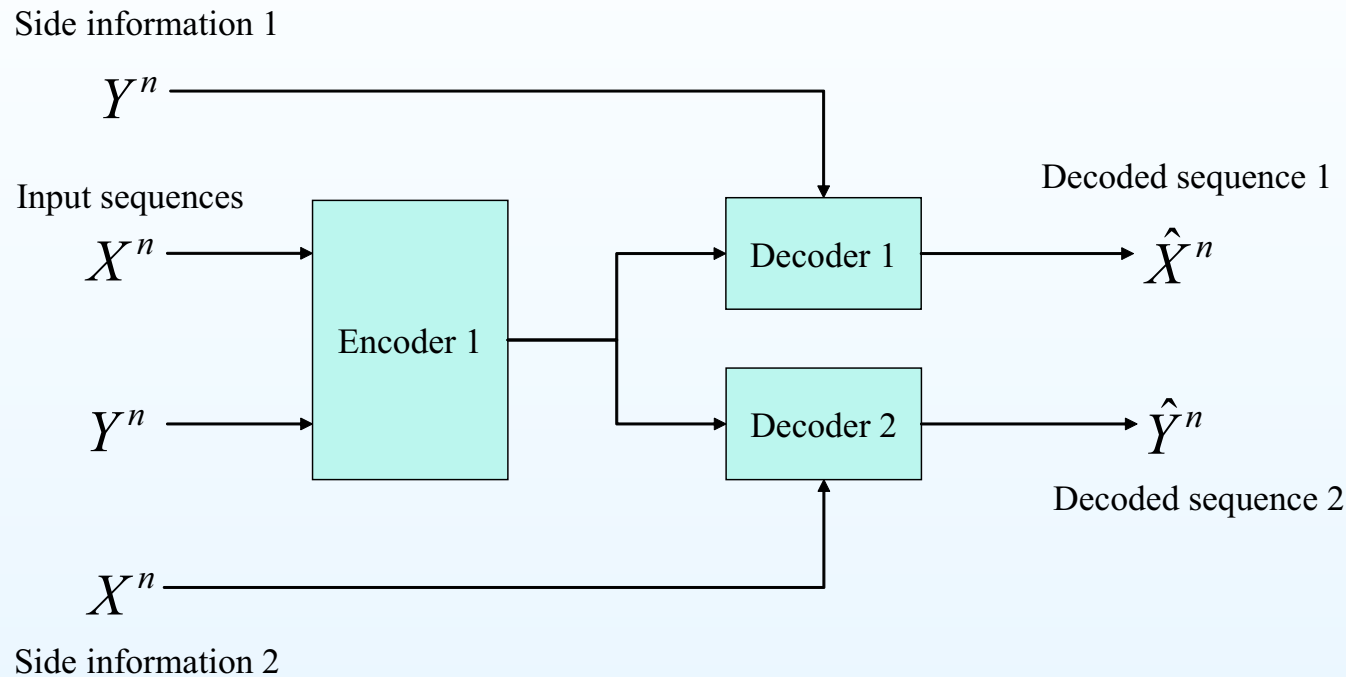
(Out of the focus of this talk)

- Clarify inner and outer bounds of the rate-distortion region for an extended coding problem, where
 - only an encoded sequence of a message is available as side information at each decoder.

Introduction

- Coding problems for correlated information sources were originally investigated by Slepian and Wolf [SW73]
- Corresponding rate-distortion problems and various coding problems have been considered.
- The main focus in the 1970's was coding problems with
 - **Separate encoding** : each message is separately encoded
 - **Joint decoding** : some of codewords are sent to a decoder and decoded at once
- Since the 1980's, coding problems that involve
 - **Joint encoding** : messages from several sources are encoded at once
 - **Separate decoding** : each message is separately decodedhave been explored.

Main contribution: complementary delivery



- We clarify the rate-distortion function and some interesting properties for the above coding problem.
 - The minimum achievable rate for a lossless configuration have been clarified. (Willems, Wolf & Wyner [WWW02])

Preliminaries (1/2)

- \mathcal{X} : finite alphabet, $|\mathcal{X}|$: cardinality of \mathcal{X} , $\mathcal{I}_M = \{1, 2, \dots, M\}$
 - In what follows, only finite alphabets will be considered.
- $x^n = (x_1, x_2, \dots, x_n)$: member of \mathcal{X}^n ,
 $x_i^j = (x_i, x_{i+1}, \dots, x_j)$: substring of x^n ($i \leq j$)
 - When the dimension is clear from the content, vectors will be denoted by boldface letters, i.e., $\mathbf{x} \in \mathcal{X}^n$
- $\mathcal{M}(\mathcal{X})$: the set of all probability distributions on \mathcal{X} ,
 $\mathcal{M}(\mathcal{X}|P_Y)$: the set of all probability distributions on \mathcal{X} given a distribution $P_Y \in \mathcal{M}(\mathcal{Y})$
- X : a discrete memoryless source taking values in \mathcal{X} with a generic distribution $P_X \in \mathcal{M}(\mathcal{X})$

Preliminaries (2/2)

- $H(X)$: entropy of X ,
- $H(X|Y)$: conditional entropy of X given Y ,
- $I(X; Y)$: mutual information between X and Y ,
- $I(X; Y|Z)$: conditional mutual information between X and Y given Z
- $\hat{\mathcal{X}}$: reconstruction alphabet,
- $\Delta_X : \mathcal{X} \times \hat{\mathcal{X}} \rightarrow [0, \infty)$: single-letter distortion function,
- $\Delta_X^n(\mathbf{x}, \hat{\mathbf{x}})$: vector distortion function

$$\Delta_X^n(\mathbf{x}, \hat{\mathbf{x}}) = \frac{1}{n} \sum_{k=1}^n \Delta_X(x_k, \hat{x}_k)$$

Problem formulation (1/2)

Definition 1. (Lossy Fully-informed Complementary Delivering (FCD) code)

A set $(\varphi^n, \hat{\varphi}_{(1)}^n, \hat{\varphi}_{(2)}^n)$ of an encoder and decoders is a lossy FCD code

$(n, M_n, \rho_n^{(X)}, \rho_n^{(Y)})$ for the source (X, Y) if and only if

$$\varphi^n : \mathcal{X}^n \times \mathcal{Y}^n \rightarrow \mathcal{I}_{M_n}$$

$$\hat{\varphi}_{(1)}^n : \mathcal{I}_{M_n} \times \mathcal{Y}^n \rightarrow \hat{\mathcal{X}}^n, \quad \hat{\varphi}_{(2)}^n : \mathcal{I}_{M_n} \times \mathcal{X}^n \rightarrow \hat{\mathcal{Y}}^n,$$

$$\rho_n^{(X)} = E \left[\Delta_X^n(X^n, \varphi_{(1)}^n(A_n, Y^n)) \right],$$

$$\rho_n^{(Y)} = E \left[\Delta_Y^n(Y^n, \varphi_{(2)}^n(A_n, X^n)) \right],$$

where $A_n \stackrel{\text{def.}}{=} \varphi^n(X^n, Y^n)$.

Problem formulation (2/2)

Definition 2. (Lossy FCD-achievable rate)

R is a lossy FCD-achievable rate of the source (X, Y) for a given distortion pair (D_1, D_2) if and only if there exists a sequence of lossy FCD codes

$\left\{ \left(n, M_n, \rho_n^{(X)}, \rho_n^{(Y)} \right) \right\}_{n=1}^{\infty}$ for the source (X, Y) such that

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log M_n \leq R,$$

$$\limsup_{n \rightarrow \infty} \rho_n^{(X)} \leq D_1, \quad \limsup_{n \rightarrow \infty} \rho_n^{(Y)} \leq D_2.$$

Definition 3. (Inf lossy FCD-achievable rate)

$$R_F(X, Y | D_1, D_2)$$

$$= \inf \left\{ R : R \text{ is a lossy FCD-achievable rate of } (X, Y) \text{ for } (D_1, D_2) \right\}.$$

Main results

Theorem 1. (Coding theorem of lossy FCD code)

$$R_F(X, Y | D_1, D_2) = \min \left[\max \{ I(X; U | Y), I(Y; U | X) \} \right],$$

where the auxiliary random variable U takes a value over an alphabet \mathcal{U} satisfying $|\mathcal{U}| \leq |\mathcal{X} \times \mathcal{Y}| + 2$, and the minimum is taken over all $P_{U|XY} \in \mathcal{M}(\mathcal{U} | P_{XY})$ such that there exist functions $\phi_{(1)} : \mathcal{U} \times \mathcal{Y} \rightarrow \hat{\mathcal{X}}$ and $\phi_{(2)} : \mathcal{U} \times \mathcal{X} \rightarrow \hat{\mathcal{Y}}$ which satisfy

$$D_1 \geq E \left[\Delta_X(X, \phi_{(1)}(U, Y)) \right],$$

$$D_2 \geq E \left[\Delta_Y(Y, \phi_{(2)}(U, X)) \right].$$

- This theorem can be easily extended to any finite number of correlated sources.

Properties (1/3)

Corollary 1.

(Compatibility with the previous result for a lossless configuration)

If $\Delta_X(x, \hat{x}) = 0 \Leftrightarrow x = \hat{x}$ and $\Delta_Y(y, \hat{y}) = 0 \Leftrightarrow y = \hat{y}$, then

$$\begin{aligned} R_F(X, Y) &\stackrel{\text{def.}}{=} R_1(X, Y | D_1 = 0, D_2 = 0) \\ &= \max\{H(X|Y), H(Y|X)\}, \end{aligned}$$

which coincides with the result reported by Willems, Wolf and Wyner [WWW02].

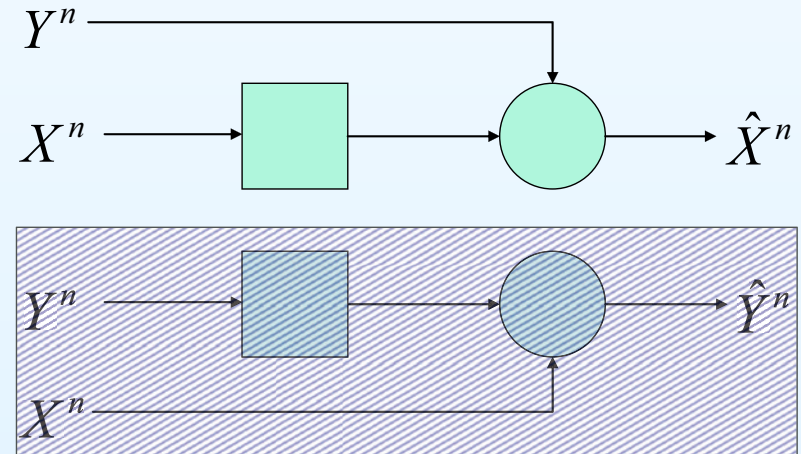
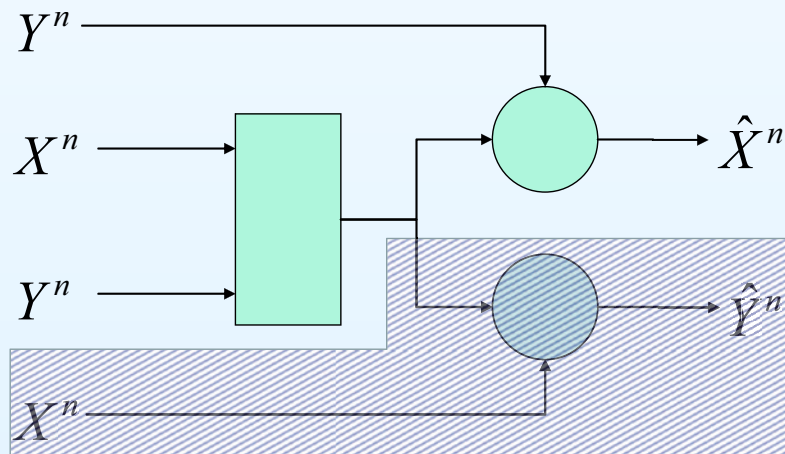
Properties (2/3)

Corollary 2. (Compatibility with the previous result for a special case)

$$R_F(X, Y | D_1 = d_1, D_2) = R_{C1}(Y | X, D_2),$$

$$R_F(X, Y | D_1, D_2 = d_2) = R_{C1}(X | Y, D_1)$$

if $d_1 \geq \bar{\Delta}_X = \max \Delta_X(x, \hat{x})$ and $d_2 \geq \bar{\Delta}_Y = \max \Delta_Y(y, \hat{y})$, where $R_{C1}(X | Y, D)$ is the conditional rate-distortion function [Ber71].



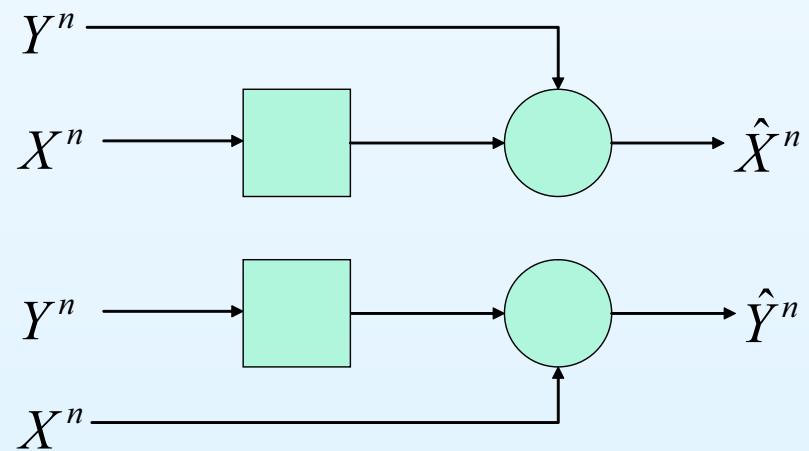
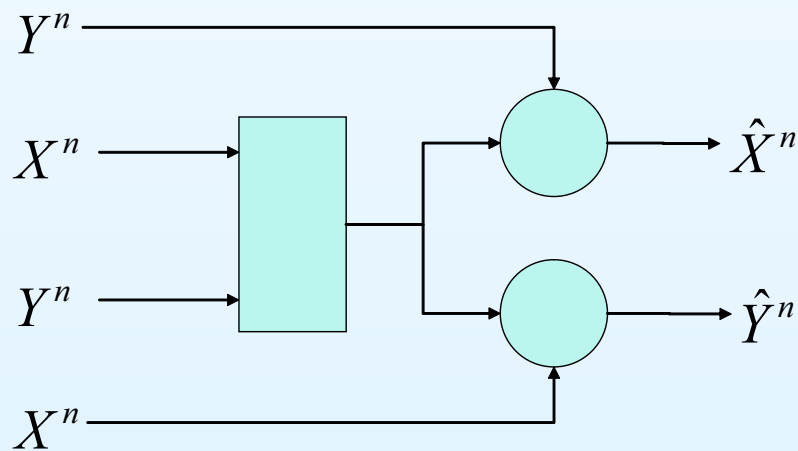
Properties (3/3)

Corollary 3. (Relationships to the previous results)

$$\begin{aligned} & \max \{R_{C1}(X|Y, D_1), R_{C1}(Y|X, D_2)\} \\ & \leq R_F(X, Y|D_1, D_2) \leq R_{C1}(X|Y, D_1) + R_{C1}(Y|X, D_2), \end{aligned}$$

On the other hand, for lossless configuration

$$\max\{H(X|Y), H(Y|X)\} = R_F(X, Y) \leq H(X|Y) + H(Y|X).$$



Conclusions

- Investigated a coding problem for correlated information sources, where
 - messages from two correlated sources are jointly encoded
 - each decoder has access to one of two messages to reproduce the other message
- Clarified the rate-distortion function for the problem

Future work

- Extend the results to other classes of information sources
- Explicit construction of the codes for the coding system

Thank you

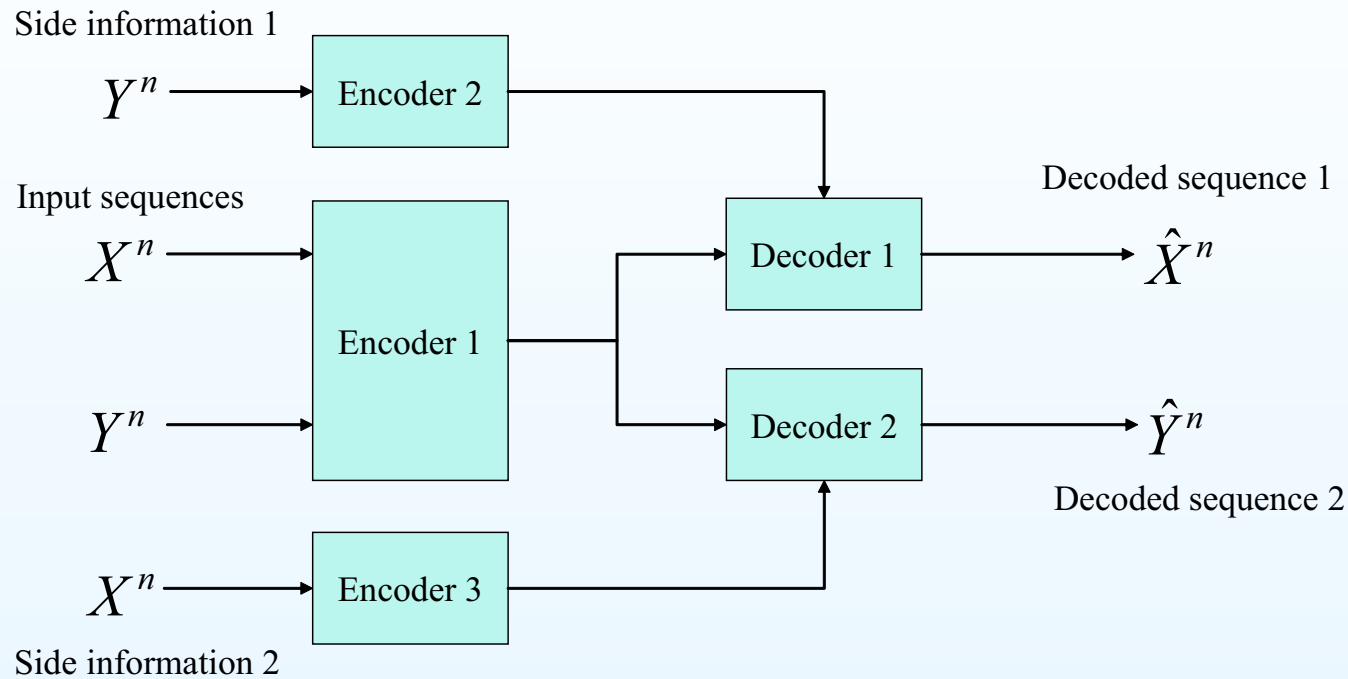
References

- [Ber71] T. Berger. *Rate distortion theory: A mathematical basis for data compression*. Prentice Hall, Englewood Cliffs, NJ, 1971.
- [CT91] T.M. Cover and J.A. Thomas. *Elements of information theory*. Wiley, New York, 1991.
- [EC82] A. A. ElGamal and T. M. Cover. Achievable rates for multiple descriptions. *IEEE Trans. Inform. Theory*, 28(6):851–857, November 1982.
- [GW74] R. M. Gray and A. D. Wyner. Source coding for a simple network. *Bell Syst. Tech. J.*, 53(9):1681–1721, November 1974.
- [HK80] T. S. Han and K. Kobayashi. A unified achievable rate region for a general class of multiterminal source coding systems. *IEEE Trans. Inform. Theory*, 26(3):277–288, 1980.
- [KM75] J. Körner and K. Marton. Images of a set via two channels and thier role in multi-user communication. *IEEE Trans. Inform. Theory*, 23(6):751–761, November 1975.
- [Sga77] A. Sgarro. Source coding with side information at several decoders. *IEEE Trans. Inform. Theory*, 23(2):179–182, March 1977.
- [SW73] D. Slepian and J. K. Wolf. Noiseless coding of correlated information sources. *IEEE Trans. Inform. Theory*, 19(4):471–480, July 1973.
- [Tun78] S. Y. Tung. *Multiterminal source coding*. PhD thesis, School of Electrical Engineering, Cornell University, May 1978.
- [WWW02] A. D. Wyner, J. K. Wolf, and F. M. J. Willems. Communicating via a processing broadcast satellite. *IEEE Trans. Inform. Theory*, 48(6):1243–1249, June 2002.
- [Wyn75] A. D. Wyner. On source coding with side information at the decoder. *IEEE Trans. Inform. Theory*, 21(3):294–300, May 1975.
- [WZ76] A. D. Wyner and J. Ziv. The rate-distortion function for source coding with side information at the decoder. *IEEE Trans. Inform. Theory*, 22(1):1–10, January 1976.
- [Yam96] H. Yamamoto. Source coding theory for a triangular communication system. *IEEE Trans. Inform. Theory*, 42(3):848–853, May 1996.

Some materials will be available at <http://www.brl.ntt.co.jp/people/akisato/>

Appendix

Related coding problem



- We call this the lossy Partially-informed complementary delivering (PCD) coding problem
- $\mathcal{R}_P(X, Y | D_1, D_2)$: Inf lossy PCD-achievable rate region defined as well as $R_F(X, Y | D_1, D_2)$

Coding theorem of lossy PCD code

Theorem 2.

$$\begin{aligned} \mathcal{R}_P(X, Y | D_1, D_2) \subseteq \\ \{ (R_1, R_2, R_3) : R_1 \geq \max\{I(X; U|V), I(Y; U|W)\}, \\ R_2 \geq I(Y; V), R_3 \geq I(X; W), \\ V \rightarrow Y \rightarrow X, \quad W \rightarrow X \rightarrow Y \} \quad (\text{outer bound}) \end{aligned}$$

where the auxiliary random variables U , V and W take values over the alphabets \mathcal{U} , \mathcal{V} and \mathcal{W} , respectively, satisfying $|\mathcal{V}| \leq |\mathcal{Y}| + 2$, $|\mathcal{W}| \leq |\mathcal{X}| + 2$ and $|\mathcal{U}| \leq |\mathcal{X} \times \mathcal{Y}|(|\mathcal{Y}| + 2)(|\mathcal{X}| + 2) + 3$, and there exist functions $\phi_{(1)} : \mathcal{U} \times \mathcal{V} \rightarrow \hat{\mathcal{X}}$ and $\phi_{(2)} : \mathcal{U} \times \mathcal{W} \rightarrow \hat{\mathcal{Y}}$ satisfying

$$\begin{aligned} D_1 &\geq E [\Delta_X(X, \phi_{(1)}(U, V))], \\ D_2 &\geq E [\Delta_Y(Y, \phi_{(2)}(U, W))]. \end{aligned}$$

Coding theorem of lossy PCD code

Theorem 2. Contd.

$$\mathcal{R}_P(X, Y | D_1, D_2) \supseteq$$

$$\{ (R_1, R_2, R_3) :$$

$$R_1 \geq \max\{I(VX; U), I(WY; U)\} - \min\{I(V; U), I(W; U)\},$$

$$R_2 \geq I(Y; V), R_3 \geq I(X; W), \quad V \rightarrow Y \rightarrow X \rightarrow W \}, \quad (\text{inner bound}).$$

Sketch of proof (1/3)

[Converse]

Consist of 3 parts:

1. Derive a lower bound of the coding rate R
2. Evaluate average distortions, e.g. $E[\Delta_X^n(X^n, \hat{X}^n)]$
3. limit the size of the alphabet \mathcal{U}

1. $U_k = A_n X^{k-1} Y^{k-1}$, where $A_n = \varphi^n(X^n, Y^n)$

$$\rightarrow \begin{cases} R + \delta & \geq \frac{1}{n} \sum_{k=1}^n I(X_k; U_k | Y_k) \\ R + \delta & \geq \frac{1}{n} \sum_{k=1}^n I(Y_k; U_k | X_k) \end{cases}$$

Sketch of proof (2/3)

2. $Y_{k+1}^n(U_k, Y_k)$: r.v. selected to minimize $E[\Delta_X(X_k, \hat{X}_k)]$

$$\longrightarrow \phi_{(1)k}(U_k, Y_k) \stackrel{\text{def.}}{=} \hat{\varphi}_{(1)k}(A_n, Y^k * Y_{k+1}^n(U_k, Y_k))$$

$$\begin{aligned} D_1 + \gamma &\geq \frac{1}{n} \sum_{k=1}^n E [\Delta_X(X_k, \hat{\varphi}_{(1)k}(A_n, Y^n))] \\ &\geq \frac{1}{n} \sum_{k=1}^n E [\Delta_X(X_k, \phi_{(1)k}(U_k, Y_k))] \\ &= E [\Delta_X(X, \phi_{(1)}(U, Y))] \end{aligned}$$

3. **Introduce the support lemma.**

Need $|\mathcal{X} \times \mathcal{Y}| + 2$ constraints to preserve

$$P_{XY}(x, y) \forall (x, y) \in \mathcal{X} \times \mathcal{Y}, \quad \max\{I(X; U|Y), I(Y; U|X)\}, \\ E[\Delta_X(X, \phi_{(1)}(U, Y))], \quad \text{and } E[\Delta_Y(Y, \phi_{(2)}(U, X))].$$

sketch of Proof (3/3)

[Achievability]

Codeword selection

- Generate $\mathcal{A}_U = \{\mathbf{u}_i\}_{i=1}^{M_U}$ according to P_U .
- Divide \mathcal{A}_U into N_U bins, each containing $L_U = M_U/N_U$ members.

Encoding

- $\mathbf{u}(\mathbf{x}, \mathbf{y}) = \mathbf{u}_i$ s.t. $(\mathbf{u}_i, \mathbf{x}, \mathbf{y}) \in T_{UXY}(k_1\delta)$ (Strong typicality)
No \mathbf{u}_i we can find : encoding error.
→ $M_U \geq \exp\{n(I(XY; U) + \gamma_1)\}$ is necessary to prevent encoding errors.
- $\varphi^n(\mathbf{x}, \mathbf{y}) = j$ s.t. $\mathbf{u}(\mathbf{x}, \mathbf{y})$ belongs to j -th bin.

Decoding

- $\hat{\mathbf{u}}(\mathbf{y}) = \mathbf{u}$ s.t. $(\mathbf{u}, \mathbf{y}) \in T_{UY}(k_2\delta)$, \mathbf{u} belongs to j -th bin.
No or more than one \mathbf{u} we can find : decoding error.
→ $L_U \leq \exp\{n(I(Y; U) - \gamma_2)\}$ is necessary to avoid decoding errors.
- $\hat{\mathbf{x}} = (x_1, \dots, x_n)$, $\hat{x}_k = \phi_{(1)}(\hat{\mathbf{u}}_k(\mathbf{y}), y_k)$ $k \in \mathcal{I}_n$.