Lecture #13 Joint Source–Channel Coding

(Reading: NIT 3.9, 14.1)

- Communication of a DMS over a DMC
- Communication of a 2-DMS over a DM-MAC
  - A joint source–channel coding scheme
  - Common part of a 2-DMS
  - A more general joint source–channel coding scheme

Communication of a DMS over a DMC

- DMS U with entropy $H(U)$
- DMC $p(y|x)$ with capacity C
- A $(|\mathcal{U}|^k, n)$ joint source–channel code of rate $r = k/n$ symbols/transmission:
  - Encoder: $x^n(u^k) \in \mathcal{X}^n$
  - Decoder: $\hat{u}^k(y^n) \in \hat{\mathcal{U}}^k$
- r achievable if $\exists$ $(|\mathcal{U}|^{rn}, n)$ codes such that $\lim_{n\to\infty} \mathrm{P}\{\hat{U}^{rn} \neq U^{rn}\} = 0$
- What is the necessary and sufficient condition for r to be achievable?

Joint source–channel coding

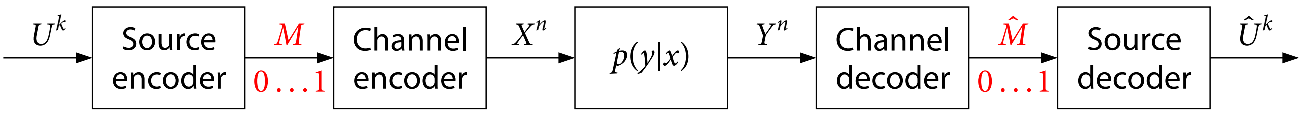
Source–channel separation theorem (Shannon 1959)

- If $rH(U) < C$, then r is achievable
- If r is achievable, then $rH(U) \le C$

- Proof of the converse:

$$\begin{aligned} H(U) &\le \frac{1}{k} I(U^k; \hat{U}^k) \\ &\le \frac{1}{k} I(U^k; Y^n) \\ &\le \frac{1}{k} \sum_{i=1}^{n} I(X_i; Y_i) \le \frac{1}{r} C \end{aligned}$$

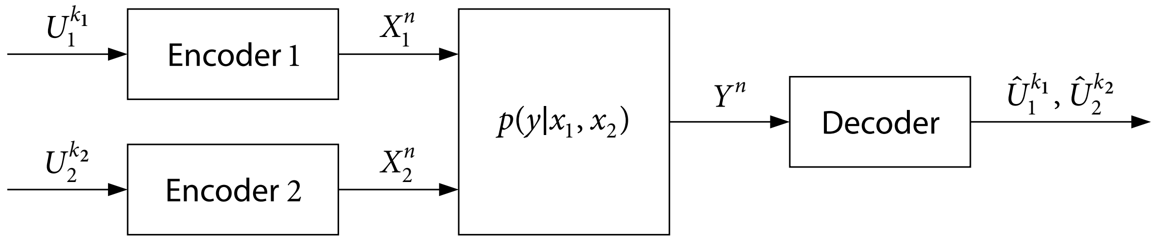
Proof of achievability



- Separate source and channel coding is asymptotically optimal
- $\mathrm{P}\{\hat{U}^{rn} \ne U^{rn}\} \le \mathrm{P}\{\hat{U}^{rn} \ne U^{rn}, \hat{M} = M\} + \mathrm{P}\{\hat{M} \ne M\}$
- Source coding: $\mathrm{P}\{\hat{U}^{rn} \ne U^{rn}, \hat{M} = M\} \to 0$ if $R > H(U)$ bits/symbol
- Channel coding: $\mathrm{P}\{\hat{M} \ne M\} \to 0$ if $rR < C$
- Combined together, $\mathrm{P}\{\hat{U}^{rn} \ne U^{rn}\} \to 0$ if $rH(U) < C$

- Basis for digital communication: Bits as "universal" source–channel interface

Communication of a 2-DMS over a DM-MAC



- 2-DMS (U_1, U_2) and a DM-MAC $p(y|x_1, x_2)$
- A $(|\mathcal{U}_1|^{k_1}, |\mathcal{U}_2|^{k_2}, n)$ joint source–channel code of rate pair $(r_1 = k_1/n, r_2 = k_2/n)$:
  - Encoder $j = 1, 2$: $x_j^n(u_j^{k_j})$
  - Decoder: $(\hat{u}_1^{k_1}(y^n), \hat{u}_2^{k_2}(y^n))$
- Probability of error: $P_e^{(n)} = \mathrm{P}\{(\hat{U}_1^{k_1}, \hat{U}_2^{k_2}) \neq (U_1^{k_1}, U_2^{k_2})\}$
- Lossless communication if $\exists$ $(|\mathcal{U}_1|^{k_1}, |\mathcal{U}_2|^{k_2}, n)$ codes with $\lim_{n\to\infty} P_e^{(n)} = 0$
- What is the necessary and sufficient condition for lossless communication?
- For simplicity, assume that $r_1 = r_2 = 1$ symbol/transmission

A separate source and channel coding scheme

- $\mathscr{C}$ of the DM-MAC is the set of (R_1, R_2) such that

$$\begin{aligned} R_1 &\le I(X_1; Y|X_2, Q), \\ R_2 &\le I(X_2; Y|X_1, Q), \\ R_1 + R_2 &\le I(X_1, X_2; Y|Q) \end{aligned}$$

  for some $p(q)p(x_1|q)p(x_2|q)$

- $\mathscr{R}^*$ for distributed lossless compression is the set of (R_1, R_2) such that

$$\begin{aligned} R_1 &\ge H(U_1|U_2), \\ R_2 &\ge H(U_2|U_1), \\ R_1 + R_2 &\ge H(U_1, U_2) \end{aligned}$$

- Hence, separation achieves lossless communication if $\mathscr{C} \cap \mathscr{R}^*$ is not empty:

$$\begin{aligned} H(U_1|U_2) &< I(X_1; Y|X_2, Q), \\ H(U_2|U_1) &< I(X_2; Y|X_1, Q), \\ H(U_1, U_2) &< I(X_1, X_2; Y|Q) \end{aligned}$$

  for some $p(q)p(x_1|q)p(x_2|q)$

Examples

- MAC with orthogonal components $p(y|x_1, x_2) = p(y_1|x_1)p(y_2|x_2)$:
  Lossless communication is possible iff

$$\begin{aligned} H(U_1|U_2) &\le C_1, \\ H(U_2|U_1) &\le C_2, \\ H(U_1, U_2) &\le C_1 + C_2 \end{aligned}$$

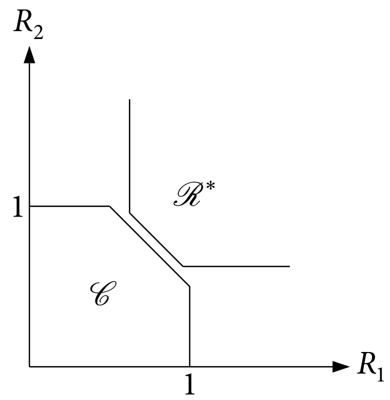
- Independent sources $(U_1, U_2) \sim p(u_1)p(u_2)$:
  Lossless communication is possible iff

$$\begin{aligned} H(U_1) &\le I(X_1; Y|X_2, Q), \\ H(U_2) &\le I(X_2; Y|X_1, Q), \\ H(U_1) + H(U_2) &\le I(X_1, X_2; Y|Q) \end{aligned}$$

  for some $p(q)p(x_1|q)p(x_2|q)$

Counterexample (Cover–El Gamal–Salehi 1980)

- Let (U_1, U_2) be a 2-DMS with $p_{U_1,U_2}(0,0) = p_{U_1,U_2}(0,1) = p_{U_1,U_2}(1,1) = 1/3$
- Send (U_1, U_2) over a binary erasure MAC ($X_1, X_2 \in \{0,1\}$, $Y = X_1 + X_2$)
- $\mathscr{R}^* \cap \mathscr{C} = \emptyset$



- Cannot send (U_1, U_2) losslessly using this separate source and channel scheme
- Now consider uncoded transmission: $X_{1i} = U_{1i}$, $X_{2i} = U_{2i}$, $i \in [1:n]$
- Hence, joint source–channel coding achieves error-free transmission!
- Separation does not hold in general for sending sources over multiuser channels

A joint source–channel coding scheme

Theorem 14.1 (Cover–El Gamal–Salehi 1980)

The 2-DMS (U_1, U_2) can be sent losslessly over a DM-MAC $p(y|x_1, x_2)$ if

$$\begin{aligned} H(U_1|U_2) &< I(X_1; Y|X_2, U_2, Q), \\ H(U_2|U_1) &< I(X_2; Y|X_1, U_1, Q), \\ H(U_1, U_2) &< I(X_1, X_2; Y|Q) \end{aligned}$$

for some $p(q, x_1, x_2|u_1, u_2) = p(q)p(x_1|u_1, q)p(x_2|u_2, q)$

- Special cases:
  - Separate source and channel coding: Set $p(x_1|u_1, q)p(x_2|u_2, q) = p(x_1|q)p(x_2|q)$
  - Counterexample: Set $Q = \emptyset$, $X_1 = U_1$, and $X_2 = U_2$

Proof of achievability

- Assume $|\mathcal{Q}| = 1$ (the general case follows by coded time sharing)
- Codebook generation: Fix $p(x_1|u_1)$ and $p(x_2|u_2)$
  - For each $u_1^n \in \mathcal{U}_1^n$, randomly and independently generate $x_1^n(u_1^n) \sim \prod_{i=1}^n p_{X_1|U_1}(x_{1i}|u_{1i})$
  - For each $u_2^n \in \mathcal{U}_2^n$, randomly and independently generate $x_2^n(u_2^n) \sim \prod_{i=1}^n p_{X_2|U_2}(x_{2i}|u_{2i})$
- Encoding:
  - Upon observing u_1^n, encoder 1 transmits $x_1^n(u_1^n)$
  - Upon observing u_2^n, encoder 2 transmits $x_2^n(u_2^n)$
  - No more than $2^{n(H(U_1,U_2)+\delta(\epsilon))}$ codeword pairs (x_1^n, x_2^n) can simultaneously occur w.h.p.
- Decoding:
  - Find the unique pair $(\hat{u}_1^n, \hat{u}_2^n)$ such that $(\hat{u}_1^n, \hat{u}_2^n, x_1^n(\hat{u}_1^n), x_2^n(\hat{u}_2^n), y^n) \in \mathcal{T}_\epsilon^{(n)}$

Analysis of the probability of error

- Consider the error events:

$$\begin{aligned}
\mathcal{E}_1 &= \{(U_1^n, U_2^n, X_1^n(U_1^n), X_2^n(U_2^n), Y^n) \notin \mathcal{T}_\epsilon^{(n)}\},\\
\mathcal{E}_2 &= \{(\tilde{u}_1^n, U_2^n, X_1^n(\tilde{u}_1^n), X_2^n(U_2^n), Y^n) \in \mathcal{T}_\epsilon^{(n)} \text{ for some } \tilde{u}_1^n \neq U_1^n\},\\
\mathcal{E}_3 &= \{(U_1^n, \tilde{u}_2^n, X_1^n(U_1^n), X_2^n(\tilde{u}_2^n), Y^n) \in \mathcal{T}_\epsilon^{(n)} \text{ for some } \tilde{u}_2^n \neq U_2^n\},\\
\mathcal{E}_4 &= \{(\tilde{u}_1^n, \tilde{u}_2^n, X_1^n(\tilde{u}_1^n), X_2^n(\tilde{u}_2^n), Y^n) \in \mathcal{T}_\epsilon^{(n)} \text{ for some } \tilde{u}_1^n \neq U_1^n, \tilde{u}_2^n \neq U_2^n\}.
\end{aligned}$$

- Then, the average probability of error

$$\mathsf{P}(\mathcal{E}) \le \mathsf{P}(\mathcal{E}_1) + \mathsf{P}(\mathcal{E}_2) + \mathsf{P}(\mathcal{E}_3) + \mathsf{P}(\mathcal{E}_4)$$

- Cannot use the packing lemma or joint typicality lemma to bound $\mathsf{P}(\mathcal{E}_j)$, $j = 2, 3, 4$

- Use basic properties of joint typicality (see NIT 14.1.2)

Suboptimality of the scheme

- Let $U_1 = U_2 = U$

- The scheme yields:
$$H(U) < \max_{p(x_1|u)p(x_2|u)} I(X_1, X_2; Y)$$
Cannot generate all joint pmfs on (X_1, X_2) in general

- But, since both senders observe same source, can use cooperative coding:

$$H(U) < \max_{p(x_1,x_2)} I(X_1, X_2; Y),$$

which can be less stringent than using the joint source-channel coding scheme

- In general the scheme can be improved when (U_1, U_2) have a common part

Common part of a 2-DMS

- Let $(U_1, U_2) \sim p(u_1, u_2)$. Arrange $p(u_1, u_2)$ in largest block diagonal form:

$u_2 \backslash u_1$				
	$u_0 = 1$	0	$\cdots$	0
	0	$u_0 = 2$	$\cdots$	0
	$\vdots$	$\vdots$	$\ddots$	$\vdots$
	0	0	$\cdots$	$u_0 = k$

- Common part between U_1 and U_2: $U_0 = u_0$ if (U_1, U_2) is in block $u_0 \in [1:k]$
- Formally (Gács–Körner 1973, Witsenhausen 1975):
  Let $g_j : \mathcal{U}_j \to [1:k]$, $j = 1, 2$, be functions with largest k such that:

  $$\mathrm{P}\{g_j(U_j) = u_0\} > 0 \text{ for every } u_0 \in [1:k],\ j = 1, 2, \text{ and } \mathrm{P}\{g_1(U_1) = g_2(U_2)\} = 1$$

  Then the common part is $U_0 = g_1(U_1) = g_2(U_2)$

Common part of a 2-DMS (U_1, U_2)

- Example:

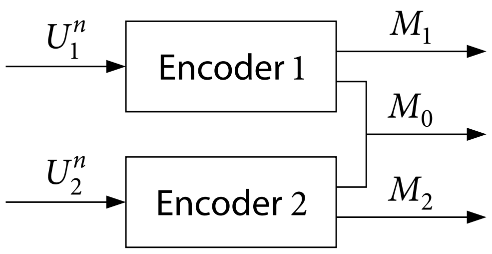
		$u_0 = 1$		$u_0 = 2$	
	$u_2 \backslash u_1$	1	2	3	4
	1	0.1	0.2	0	0
$u_0 = 1$	2	0.1	0.1	0	0
	3	0.1	0.1	0	0
$u_0 = 2$	4	0	0	0.2	0.1

  Here $k = 2$, $\mathrm{P}\{U_0 = 1\} = 0.7$, $\mathrm{P}\{U_0 = 2\} = 0.3$

- Now, consider a 2-DMS (U_1, U_2)
- Common part between U_1^n and U_2^n is U_0^n (up to relabeling)
- Hence, the DMS U_0 is the common part of the 2-DMS (U_1, U_2)
- Can modify our joint source–channel scheme to include common part

Three-index separate source and channel coding scheme

- Source coding:



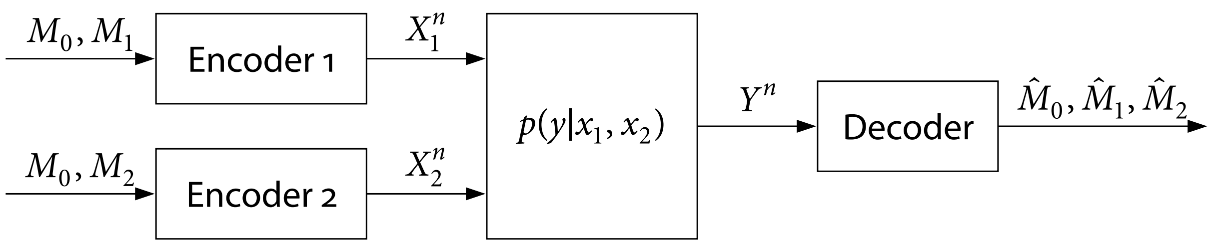
- M_0 can depend only on U_0^n
- Hence, the optimal rate region $\mathscr{R}^*$ is the set of (R_0, R_1, R_2) such that

$$\begin{aligned} R_1 &\geq H(U_1|U_2), \\ R_2 &\geq H(U_2|U_1), \\ R_1 + R_2 &\geq H(U_1, U_2|U_0), \\ R_0 + R_1 + R_2 &\geq H(U_1, U_2) \end{aligned}$$

- Lossless compression of U_0; S–W coding of (U_1, U_2) for every u_0^n

Three-index separate source and channel coding scheme

- Channel coding:



- The capacity region $\mathscr{C}$ with a common message is the set of (R_0, R_1, R_2) such that

$$\begin{aligned} R_1 &\leq I(X_1; Y|X_2, W), \\ R_2 &\leq I(X_2; Y|X_1, W), \\ R_1 + R_2 &\leq I(X_1, X_2; Y|W), \\ R_0 + R_1 + R_2 &\leq I(X_1, X_2; Y) \end{aligned}$$

  for some $p(w)p(x_1|w)p(x_2|w)$ with $|\mathcal{W}| \leq \min\{|\mathcal{X}_1| \cdot |\mathcal{X}_2| + 2,\ |\mathcal{Y}| + 3\}$ (Problem 5.19)

- Use superposition coding

Three-index separate source and channel coding scheme

- Hence, separation achieves lossless communication if $\mathscr{C} \cap \mathscr{R}^*$ is not empty:

$$
\begin{aligned}
H(U_1|U_2) &< I(X_1; Y|X_2, W),\\
H(U_2|U_1) &< I(X_2; Y|X_1, W),\\
H(U_1, U_2|U_0) &< I(X_1, X_2; Y|W),\\
H(U_1, U_2) &< I(X_1, X_2; Y)
\end{aligned}
$$

for some $p(w)p(x_1|w)p(x_2|w)$ with $|\mathcal{W}| \le \min\{|\mathcal{X}_1| \cdot |\mathcal{X}_2| + 2,\ |\mathcal{Y}| + 3\}$

A joint source–channel coding scheme with common part

- Combine three-index coding scheme and Theorem 14.1:

Theorem 14.2 (Cover–El Gamal–Salehi 1980)

A 2-DMS with common part can be sent losslessly over a DM-MAC if

$$
\begin{aligned}
H(U_1|U_2) &< I(X_1; Y|X_2, U_2, W),\\
H(U_2|U_1) &< I(X_2; Y|X_1, U_1, W),\\
H(U_1, U_2|U_0) &< I(X_1, X_2; Y|U_0, W),\\
H(U_1, U_2) &< I(X_1, X_2; Y)
\end{aligned}
$$

for some $p(w)p(x_1|u_1, w)p(x_2|u_2, w)$

- U_0 is represented by W, chosen to maximize cooperation between the senders
- Use the previous joint source–channel scheme for each w^n
- Optimal for the DM-MAC with common message (Slepian–Wolf 1973)
- Not optimal in general (Dueck 1981)

Summary

- Source channel separation holds for point to point communication
- Source–channel separation does not hold in general for multiuser channels
- Joint source–channel coding schemes that utilize the correlation between the sources for cooperative transmission
- Common part of a 2-DMS

References

Cover, T. M., El Gamal, A., and Salehi, M. (1980). Multiple access channels with arbitrarily correlated sources. *IEEE Trans. Inf. Theory*, 26(6), 648–657.

Dueck, G. (1981). A note on the multiple access channel with correlated sources. *IEEE Trans. Inf. Theory*, 27(2), 232–235.

Gács, P. and Körner, J. (1973). Common information is far less than mutual information. *Probl. Control Inf. Theory*, 2(2), 149–162.

Shannon, C. E. (1959). Coding theorems for a discrete source with a fidelity criterion. In *IRE Int. Conv. Rec.*, vol. 7, part 4, pp. 142–163. Reprint with changes (1960). In R. E. Machol (ed.) *Information and Decision Processes*, pp. 93–126. McGraw-Hill, New York.

Slepian, D. and Wolf, J. K. (1973). A coding theorem for multiple access channels with correlated sources. *Bell Syst. Tech. J.*, 52(7), 1037–1076.

Witsenhausen, H. S. (1975). On sequences of pairs of dependent random variables. *SIAM J. Appl. Math.*, 28(1), 100–113.