PACKING AND COVERING LEMMA
+
COMMUNICATION FOR COMPUTING

Department of Electrical Engineering of Arizona State University

Bahman Moraffah

OUTLINE

- Packing lemma
- Covering lemma
- Communication for computing

PACKING LEMMA

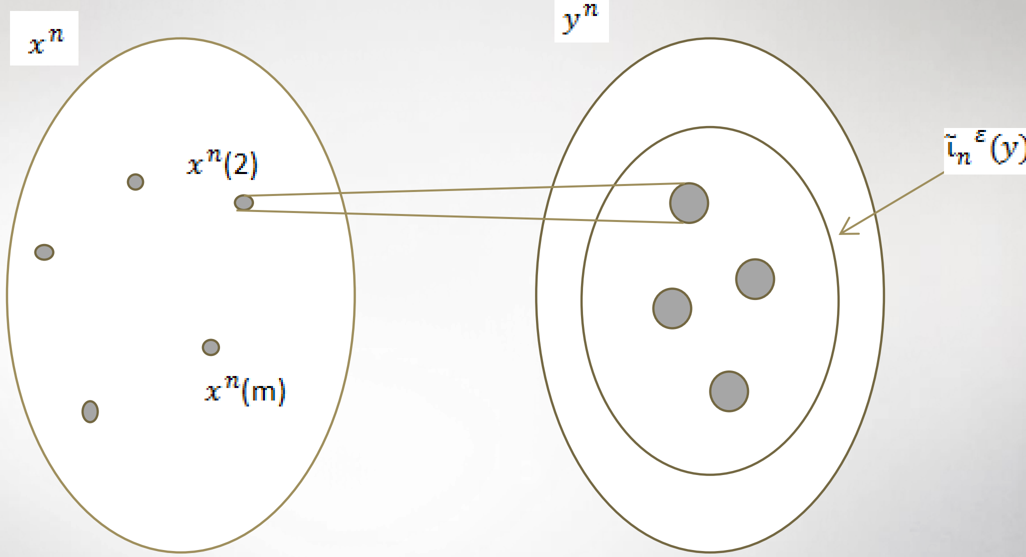
- Fix p(x) and channel p(y|x)
- Now, according to p(x), construct 2^{nR} code words, i.i.d. with length n :

$$X^n(\mathrm{m}) \sim \prod_{i=1}^{n} p(\mathrm{x_i})$$

- Suppose $\hat{X}(1)$ is our message and $\tilde{Y}^n$ is our output.

OBJECTIVE

- ❖ $\tilde{Y}^n, X^n(m)$ m $\in [2,\ 2^{nR}]$ are NOT Jointly Typical!



$$R < I(x, y) \Rightarrow \frac{2^{n(H(y|x)+R)}}{2^{nH(y)}} \to 0$$

SIMPLIFIED PACKING LEMMA

Probability that $\tilde{Y}^n$ and X^n(m) m $\in [2,\ 2^{nR}]$ are joint typical goes to zero if R < I(X; Y)

❖ X^n(m) and $\tilde{Y}^n$ are independent and uniformly chosen.

GENERALIZE PACKING LEMMA

We generalize packing in 3 steps:

1. $\tilde{Y}^n$ with arbitrary distribution.
2. Dependency of code words to each other.
3. Structured Code Book.

$\tilde{Y}^n$ WITH ARBITRARY DISTRIBUTION

- ❖ Suppose $\tilde{Y}^n$ is not necessarily the output of channel related to message 1, and it has an arbitrary distribution.
- ❖ We suppose $\tilde{Y}^n, X^n(m)$ *are independent, we have*:

R < I(X, Y)

$$\Rightarrow \lim_{n\to\infty} \mathrm{P}(\exists m \; 1 \le m \le 2^{nR} : \left(X^n(m), \tilde{Y}^n\right) \in \tilde{\imath}_{\varepsilon}{}^{n}(x, y)) = 0$$

DEPENDENCY OF CODE WORDS TO EACH OTHER

- By looking at proof of previous slide we understand that:
  - $\tilde{Y}^n$ and X^n(m) m $\in [1,\ 2^{nR}]$ are independent.
  - Marginal distribution of X^n(m) should be i.i.d.
- So, we don't need independency of code words and independency of each code words and $\tilde{Y}^n$ suffice. To have the same result.

STRUCTURED CODE BOOK

- It means we have a relationship or closeness among code words .
- A common way is :
  - Let p(U, X).
  - Generate $\widetilde{U}^n$ i.i.d distribution of p(U)
  - Send $\widetilde{U}^n$ through channel p(X|U) 2^{nR} times.
  - Have 2^{nR} code words according to

$$X^n(\mathrm{m}) \sim \prod_{i=1}^{n} p_{X|U}(x_i|\,\hat{u}_i)$$

OBJECTIVE

- Suppose $\hat{X}(1)$ *is our message and* $\widetilde{Y}^n$ is our output.

We want to find conditions on R such that $\tilde{Y}^n$ and X^n(m) m $\in [2,\ 2^{nR}]$ would NOT be jointly typical.

- We know that X^n(m) →U→ $\widetilde{Y}^n$

CONT.

We generated $\widetilde{U}^n, X^n, \tilde{Y}^n\ i.i.d.$ from

$$q(u,x,y) = p(u)p(x|u)p(y|u)$$

So, the probability which is equal to p(u, x, y) is

$$2^{-nD(p(u,x,y)||q(u,x,y))} = 2^{-nI(X;Y|U)}$$

PACKING LEMMA

Consider p(u, x, y) on (U, X, Y). Suppose $\widetilde{Y}^n$, $\widetilde{U}^n$ *have an* arbitrary distribution p($\widetilde{U}^n$, $\widetilde{Y}^n$). $X^n(m)$ is 2^{nR} sequences such that

P($X^n(m)$ $= x^n \mid \widetilde{U}^n = \widetilde{u}^n$) = $\prod p_{X|U}(x_i|\widetilde{u}_i)$

We also have X^n(m) → $\widetilde{U}^n$ → $\widetilde{Y}^n$ then,

R < I(X; Y|U)

⇒ $\lim_{n\to\infty}$ P($\exists m\ 1 \le m \le 2^{nR}$: $\left(\widetilde{U}^n, \tilde{Y}^n, X(m)\right) \in$ ĩ$_\varepsilon{}^n(U,X,Y)$) = 0

COVERING LEMMA

In packing lemma we saw that

R < I(X; Y) ⇒ $P\{\exists m\ m \in [1, 2^{nR}]: \left(X(m), \tilde{Y}^n\right) \in \tilde{\iota}_{\varepsilon}{}^{n}(X, Y)\} \rightarrow 0$

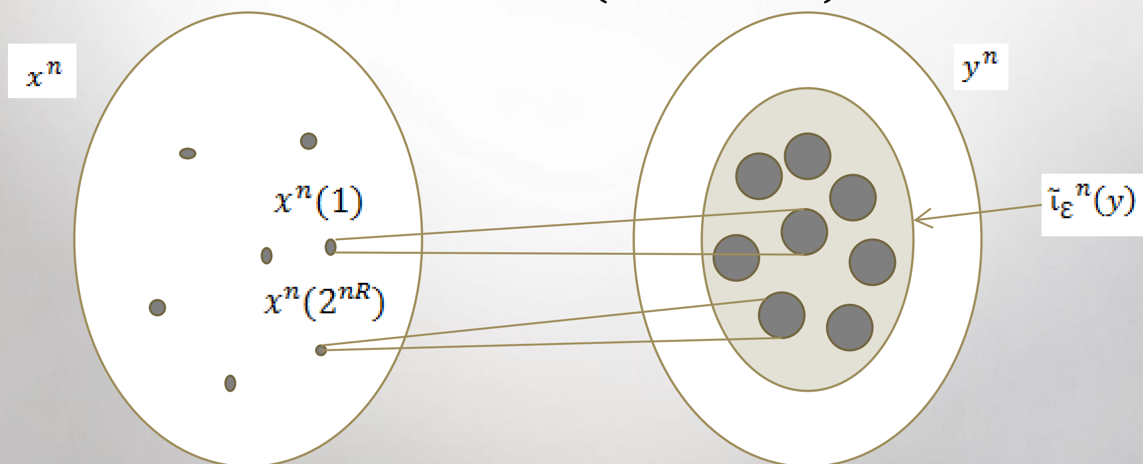
No, what if R > I(X; Y)?

Answer: Covering lemma

COVERING LEMMA

$\tilde{Y}^n$ is a typical sequence with distribution of p(y) then

$$R > I(X; Y) \Rightarrow P\{\exists m \in [1, 2^{nR}]: \left(X^n(m), \tilde{Y}^n\right) \in \tilde{\imath}_{\varepsilon}{}^{n}(X, Y) \ \rightarrow 1$$



COMPARISON

- ❖ If number of points is less than $2^{-n(I(X;Y)-\varepsilon)}$

$\Rightarrow$ *grey circles have no intersection.*

- ❖ If number of points is greater than $2^{-n(I(X;Y)+\varepsilon)} \Rightarrow$ *grey circles* cover $\tilde{\iota}_{\varepsilon}{}^{n}$.

✓ So, I(X; Y) causes a change in phase.

NOTE

When $R < I(X; Y)$, we don't have a good book for channel coding.

we use packing lemma in channel coding and covering lemma in source coding.

COVERING LEMMA

Because, we use this lemma in source coding we are going to change the notation:

(U, X, $\hat{X}$) ~ p(u, x, $\hat{x}$) and X^n, U^n have an arbitrary distribution such that:

$$P((U^n, X^n) \in \tilde{\imath}_{\varepsilon}{}^{n}(U,X)) \to 1 \text{ and } \hat{X}^n(\mathrm{m})\ m \in [1, 2^{nR}]$$

are generated independently by sending $\widetilde{U}^n$ through $p_{\hat{X}|U}$

CONT.

i.e.

$$p(\hat{X}^n(\mathrm{m}) = \hat{x}^n \,|U^n = u^n) = \prod_{i=1}^{n} p_{\hat{X}|U}(\hat{x}_i|u_i)$$

and

$$p(\hat{X}^n(1),\ldots,\hat{X}^n(2^{nR})|\ U^n) = \prod_{m=1}^{2^{nR}} p\ (\hat{X}^n(\mathrm{m})|u_i)$$

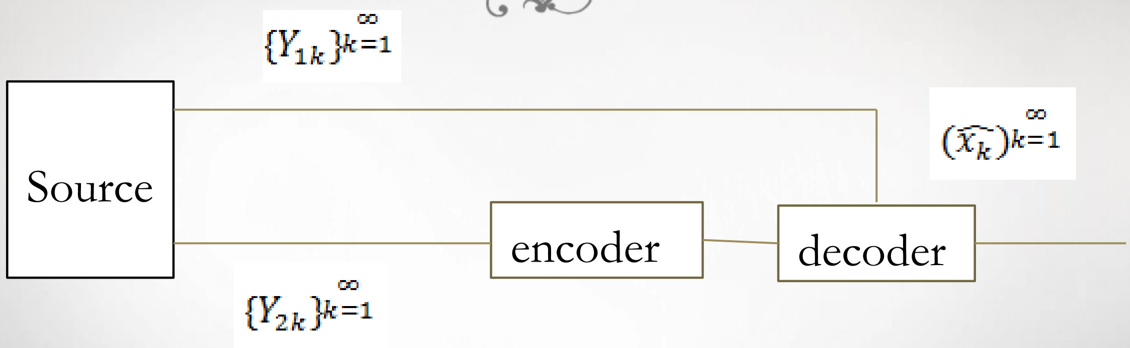
We also have

$$\hat{X}^n(1),\ldots,\hat{X}^n(2^{nR}) \to\ U^n \to X^n$$

Then

$$\mathrm{R} > \mathrm{I}(\mathrm{X};\hat{X}|U) \Rightarrow \mathrm{P}\{\exists m \in [1, 2^{nR}]: \big(U^n, X^n, \hat{X}^n(\mathrm{m})\big) \in\ \tilde{\imath}_{\varepsilon}{}^{n}\big(U, X, \hat{X}\big)\} \ \to 1$$

WYNER- ZIV THEORY FOR A GENERAL FUNCTION OF CORROLATED SOURCE



- ❖ In some cases, The user at decoder will be satisfied with estimating the value of $X = G(Y_1, Y_2)$ instead of Y_2.

- Let $\mathcal{Y}_1, \mathcal{Y}_2, \mathcal{X}, \hat{\mathcal{X}}$ finite sets.

- $(Y_{1k}, Y_{2k})_{k=1}^{\infty}$ a sequence of i.i.d pairs of dependent r.v. (Y_1, Y_2)

- G: $\mathcal{Y}_1 * \mathcal{Y}_2 \rightarrow \mathcal{X}$

- $\mathcal{X} * \hat{\mathcal{X}} \rightarrow [0, +\infty)$: distortion function

Objective : Decoder desires to reproduce X = G(Y_1, Y_2) with a distortion D.

Define : code(F_E, F_D) for G:

$$F_E : \mathcal{Y}_2{}^n \rightarrow \; I_M$$

$$F_D : \mathcal{Y}_1{}^n * I_M \rightarrow \hat{X}^n$$

$$\Delta = \mathrm{E}\left(\frac{1}{n}\sum D(X_k, \widehat{X_k})\right)$$

Det: (R, D) is said to be achievable for function of G if, for $\in > 0$ and n sufficiently large, there exists a code(F_E, F_D) which satisfies

$$M \leq 2^{n(R+\in)}$$

$$\Delta \leq d + \in$$

R = Set of all achievable (R, D)

$$R(d) = \min_{(R, D) \in \boldsymbol{R}} R$$

ThM: For d ≥ 0

$$R(d) = \min_{\widehat{Y_2} \in p(d)} I(Y_2;\ \widehat{Y_2} | Y_1)$$

p(d) = Set of all $\widehat{Y_2} \in \widehat{\mathcal{Y}_2}$ which:

- $\widehat{Y_2} \rightarrow Y_2 \rightarrow Y_1$
- $\exists\, f : \mathcal{Y}_1 * \widehat{\mathcal{Y}_2} \rightarrow \widehat{\mathcal{X}} : E\,(D(G(Y_1, Y_2), f(Y_1, \widehat{Y_2}))) \leq D$