Dirty Paper Coding

Gwanmo Ku

Adaptive Signal Processing and Information Theory Research Group

Nov. 7, 2012

Outline

Introduction

Writing on Dirty Paper

- System Model & Channel Capacity
- Achievable Rate Region
- Encoding & Decoding
- Achievability Proof

Example of Dirty Paper Coding

Outline

Introduction

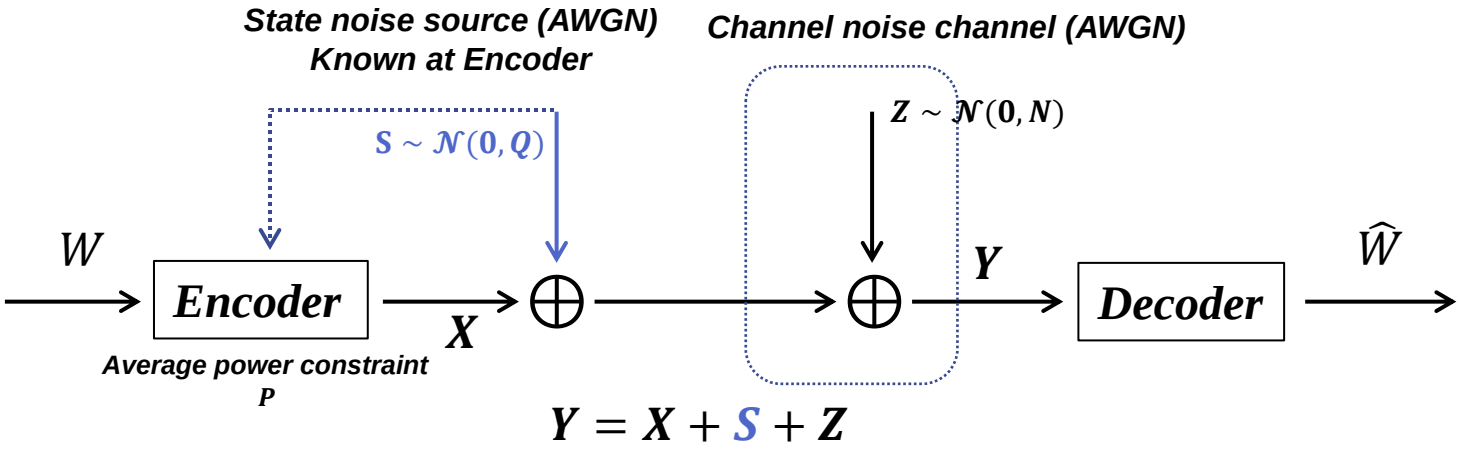
Writing on Dirty Paper

- System Model & Channel Capacity
- Achievable Rate Region
- Encoding & Decoding
- Achievability Proof

Example of Dirty Paper Coding

Writing on Dirty Paper (Costa, 83)

- **Costa's Idea**

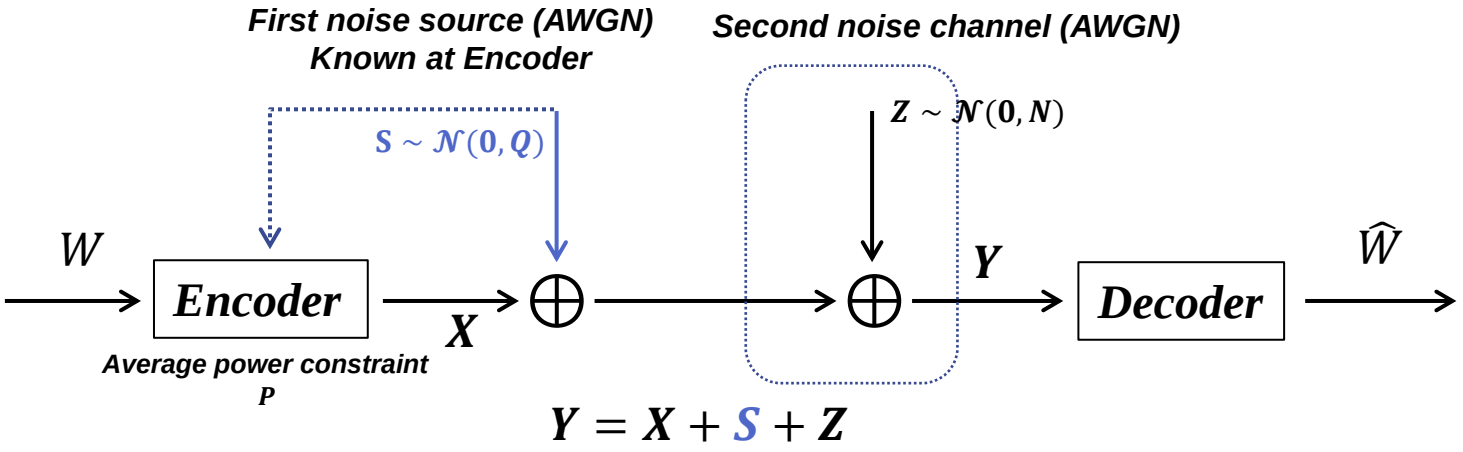


Capacity

$$\frac{1}{2}\ln\left(1+\frac{P}{N}\right) \quad \Leftarrow \quad \frac{1}{2}\ln\left(1+\frac{P}{N+Q}\right)$$

State S does not effect the capacity

Writing on Dirty Paper (Costa, 83)



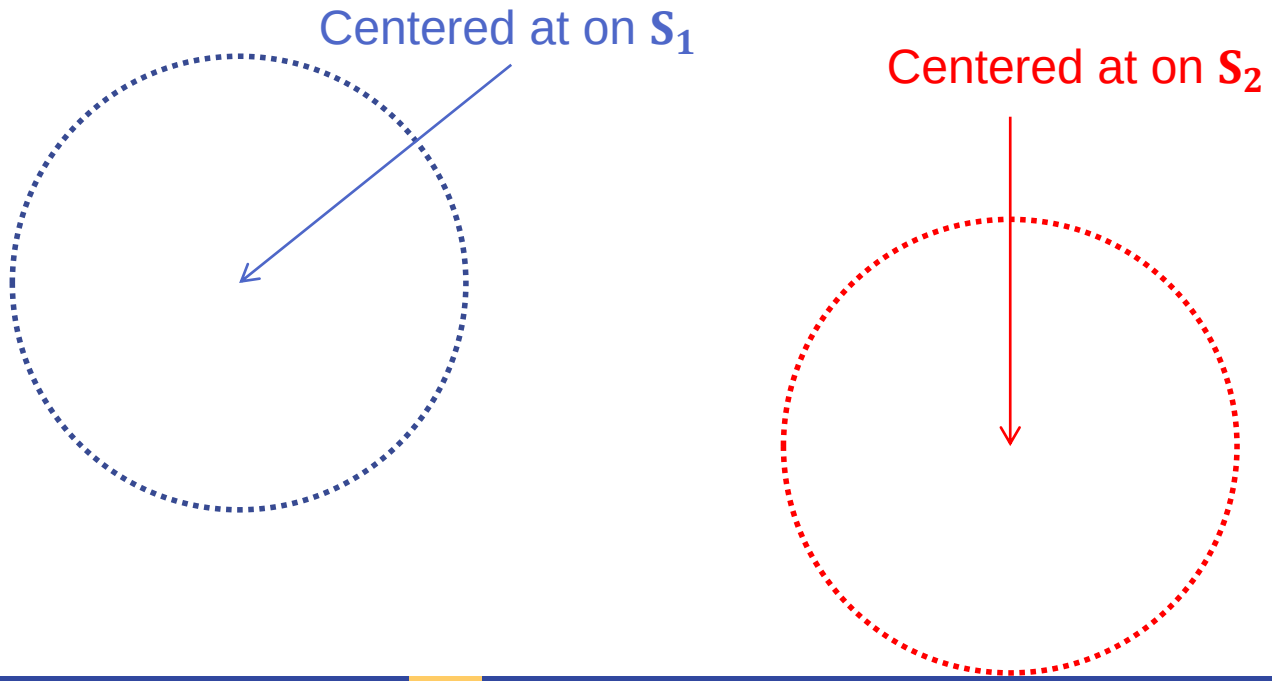
- ✓ **S** *: a piece of paper covered with normally distributed dots spot*
- ✓ **W** *: a message on it wit a limited amount of ink (**P**)*
- ✓ **How much information can we reliably send?**

$$\frac{1}{2}\ln\left(1+\frac{P}{N}\right) \quad \Leftarrow \quad \frac{1}{2}\ln\left(1+\frac{P}{N+Q}\right)$$

Dirt spots do not effect the capacity

Writing on Dirty Paper (Costa, 83)

- **Practical Importance**
  - ✓ **Diversity without capacity loss**
  - ✓ **Layering by dirt spots (Watermarking)**
  - ✓ **Extension to MIMO case for Diversity**



Outline

Introduction

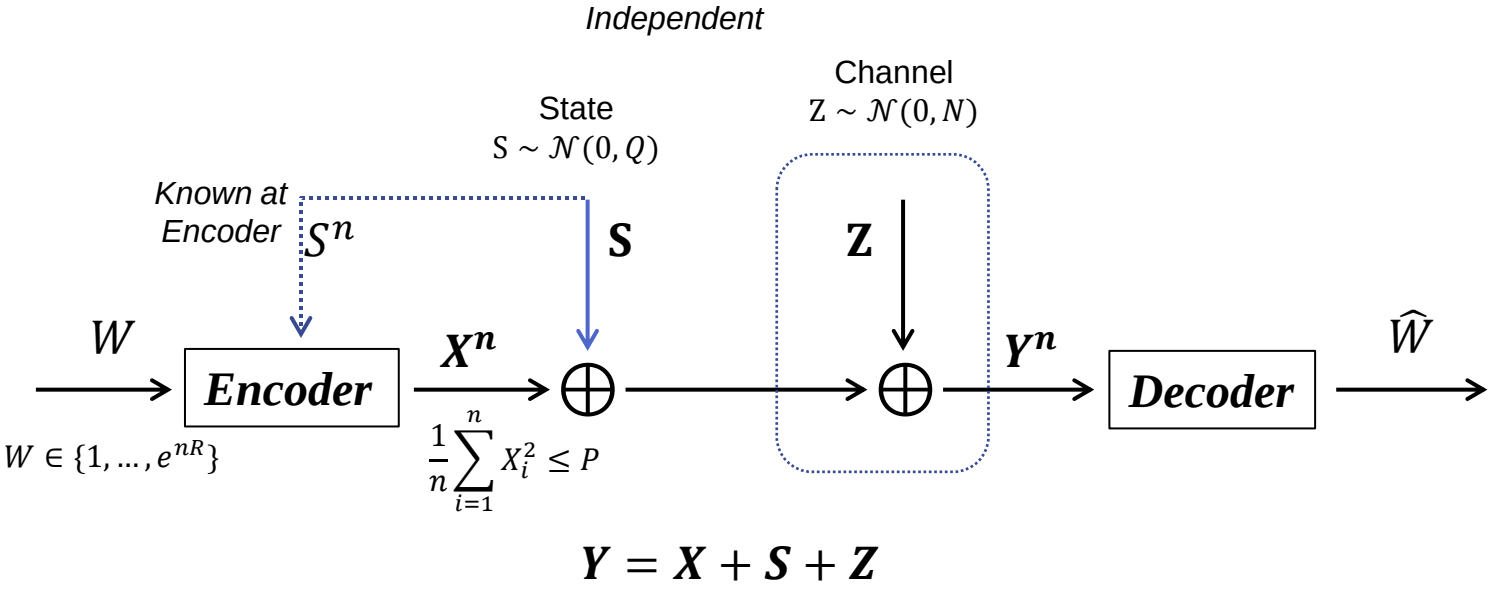
Writing on Dirty Paper

- System Model & Channel Capacity
- Achievable Rate Region
- Encoding & Decoding
- Achievability Proof

Example of Dirty Paper Coding

Writing on Dirty Paper

- **System Model**



$$Y = X + S + Z$$

✓ **Capacity**

$$\frac{1}{2}\ln\left(1+\frac{P}{N}\right) \quad \Leftarrow \quad \frac{1}{2}\ln\left(1+\frac{P}{N+Q}\right)$$

State S does not effect the capacity

Capacity Region

- **Refer to Gelfand Pinsker Theorem**
  - ✓ **Capacity of the with DM state (S) known only at Encoder**

$$C_{SI-E} = \max_{p(u,x|s)} [I(U;Y) - I(U;S)]$$

  - ✓ **Extension to DMC with continuous alphabet**

$$C_{SI-E} = \sup_{p(u,x|s)} [I(U_d;Y_p) - I(U_d;S_q)]$$

  - ✓**Proof :** Refer to Gelfand Pinsker Paper or El Gamal' book (Ch 7.6)

- **Capacity Region of DPC**
  - ✓ **Normally Distributed State S**
  - ✓ **Input power constraint P**
  - ✓ **AWGN channel with variance of Q**

$$\sup_{p(u,x|s)} [I(U;Y) - I(U;S)]$$

$$\frac{1}{2}\log\left(1+\frac{P}{N}\right)$$

Encoding

- **Encoding Scheme**
  - ✓ **Generate $e^{n[I(U;Y)-\epsilon]}$ iid sequences U**

    According to Uniform Distribution over the set of typical U
  - ✓ **Distribute these sequences uniformly over e^{nR} bins**

    Each sequence u with bin index $i(u)$
  - ✓ **Given state vector S and the message W**

    Look in bin W for a sequence U s.t. joint typical (U, S)
  - ✓ **Choose X s.t. jointly typical (X, U, S), send it**

Decoding

- **Decoding Scheme**
  - ✓ **Look for unique sequence U s.t. jointly typical (U, Y)**
  - ✓ **Declare Error**

    More than one or no such sequence exist
  - ✓ **Set the estimate $\widehat{W}$**

    Equal to the index of bin containing the obtained sequence U
  - ✓ **If $R < I(U;Y) - I(U;S) - \epsilon - \delta$**

    The probability of error averaged over all codes

    decrease exponentially to zero as $n \to \infty$

Achievability Proof

- $\boldsymbol{R \leq \frac{1}{2}\log\left(1+\frac{P}{N}\right)}$

Step 1 : Set $U = X + \alpha S$

$$
\begin{aligned}
I(U;Y) &= h(Y) - h(Y|U) \\
&= h(X+S+Z) - h(X+S+Z|X+\alpha S) \\
&= h(X+S+Z) + h(X+\alpha S) - h(X+\alpha S, X+S+Z) \\
&= \frac{1}{2}\ln[(2\pi e)^n (P+Q+N)(P+\alpha^2 Q) \\
&\quad - \frac{1}{2}\ln\{(2\pi e)^n [(P+Q+N)(P+\alpha^2 Q) - (P+\alpha Q)^2]\} \\
&= \frac{1}{2}\ln\frac{(P+Q+N)(P+\alpha^2 Q)}{PQ(1-\alpha)^2 + N(P+\alpha^2 Q)}
\end{aligned}
$$

Achievability Proof

$$
\begin{aligned}
I(U;S) &= h(U) - h(U|S) \\
&= h(X + \alpha S) - h(X + \alpha S|S) \\
&= h(X + \alpha S) + h(X) \\
&= \frac{1}{2}\ln[(2\,\pi e)^n(P + \alpha^2 Q)] - \frac{1}{2}\ln[(2\,\pi e)^n P] \\
&= \frac{1}{2}\ln\frac{(P + \alpha^2 Q)}{P}
\end{aligned}
$$

Let $R(\alpha) = I(U;Y) - I(U;S)$,

$$
R(\alpha) = \frac{1}{2}\ln\frac{P(P + Q + N)}{PQ(1-\alpha)^2 + N(P + \alpha^2 Q)}
$$

Achievability Proof

$$R(\alpha) = \frac{1}{2}\ln\frac{P(P+Q+N)}{PQ(1-\alpha)^2 + N(P+\alpha^2 Q)}$$

Find maximum $R(\alpha)$

Fixed $P(P+Q+N)$

minimize $PQ(1-\alpha)^2 + N(P+\alpha^2 Q)$

$$\alpha^* = \arg\min_{\alpha}[PQ(1-\alpha)^2 + N(P+\alpha^2 Q)] = \frac{-2PQ}{-2Q(N+P)} = \frac{P}{P+N}$$

$$\boldsymbol{R(\alpha) \le R(\alpha^*)|_{\alpha^*=\frac{P}{P+N}}}$$

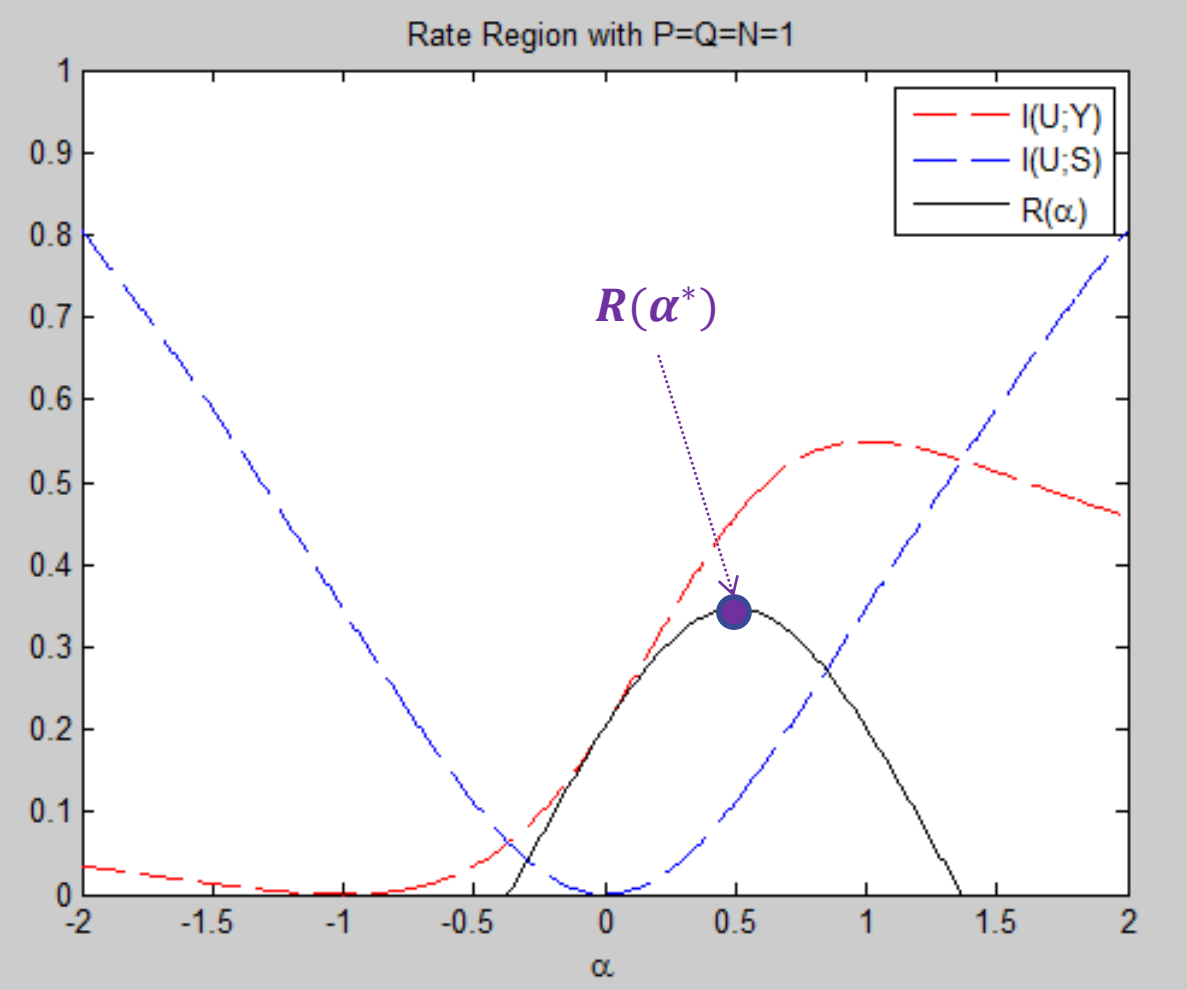
$$= \frac{1}{2}\ln[1+\frac{P}{N}]$$

Also,

$$I(U^*,Y) = \frac{1}{2}\ln[1+\frac{P(P+Q+N)}{N(P+N)}] \qquad I(U^*,S) = \frac{1}{2}\ln[1+\frac{PQ}{P+N}]$$

Achievability Proof

- **Rate Regions and Capacity**



Encoding & Decoding with α^*

- **Encoding Scheme**

✓ **Generate $e^{n[I(U^*;Y)-\epsilon]}$ sequences U**

With components independently drawn according to $\mathcal{N}(0, P + \alpha^{*2} Q)$

✓ **Place these sequences into $e^{nR} = e^{n(C^*-2\,\epsilon)}$ bins**

Each bin contain same number of sequences

✓ **Share Codebook between Encoder and Decoder**

✓ **Given a State Vector $S = S_0$ and a message $W = k$**

Look for jointly typical pair (U, S_0) among the U in bin k

$$|(U - \alpha^* S_0)' S_0| \leq \delta$$

Encoder calculate $\mathbf{X_0 = U_0 - \alpha^* S_0}$ **, send** $\mathbf{X_0}$

Encoding & Decoding

- **Decoding Scheme**

✓ **Receive $Y = Y_0$**

Look for a sequence U s.t jointly typical (U, Y_0)

✓ **Declare Error**

More than one or no such sequence

Find U_0

✓ **Set the estimate $\widehat{W}$**

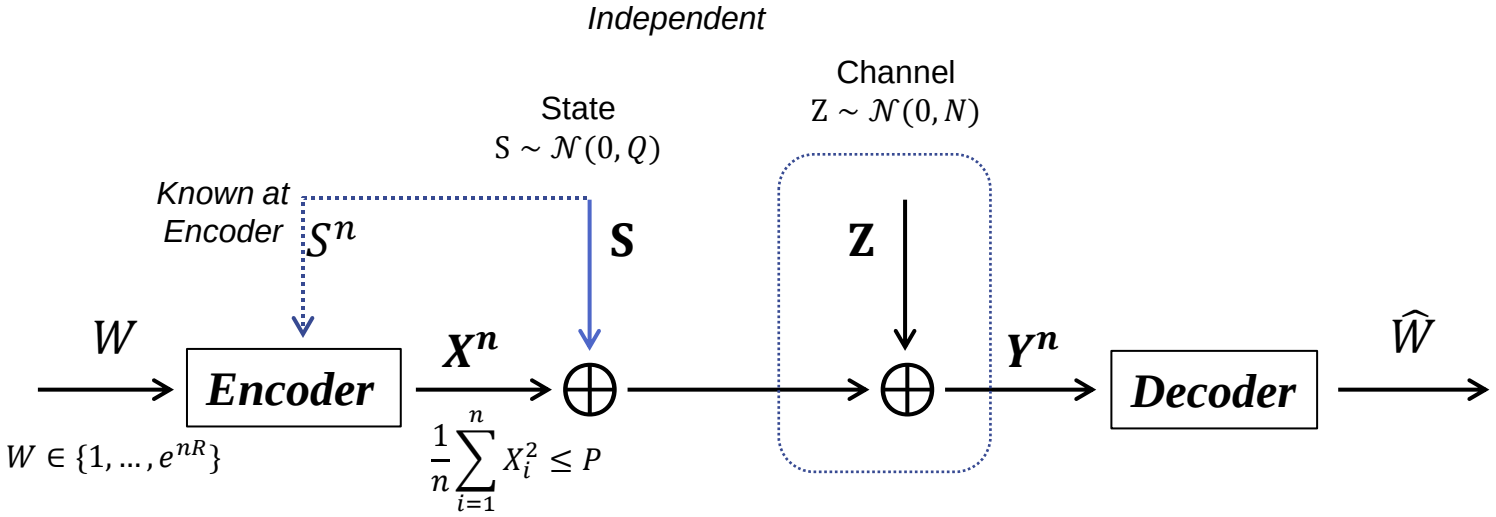
Equal to the index of bin containing this sequence

The probability of error averaged over random choice of code

decrease exponentially to zero as $n \to \infty$

Conclusion of DPC

- **System Model**



$$Y = X + S + Z$$

✓ **Capacity**

$$\frac{1}{2}\ln\left(1+\frac{P}{N}\right) \quad \Leftarrow \quad \frac{1}{2}\ln\left(1+\frac{P}{N+Q}\right)$$

State S does not effect the capacity

Outline

Introduction

Writing on Dirty Paper

- System Model & Channel Capacity
- Achievable Rate Region
- Encoding & Decoding
- Achievability Proof

Example of Dirty Paper Coding

Example of DPC

- **DPC for Gaussian BC**

$$Y_1 = X + S_1 + Z_1$$
$$Y_2 = X + S_2 + Z_2$$

$$R_1 \le C\left(\frac{\alpha P}{N_1}\right) \qquad \alpha \in [0, 1]$$
$$R_2 \le C\left(\frac{(1-\alpha)P}{\alpha P + N_2}\right)$$

✓ **Assumptions**

Two States S_1 and S_2 are only known at Encoder

$S_1 \sim \mathcal{N}(0, Q_1)$ $\quad$ $S_2 \sim \mathcal{N}(0, Q_2)$

$Z_1 \sim \mathcal{N}(0, N_1)$ $\quad$ $Z_2 \sim \mathcal{N}(0, N_2)$ $\quad$ $N_2 \ge N_1$

Average Power Constraint P on X

Example of DPC

- **Encoding & Decoding**

Split X into two independent parts X_1 and X_2

$$X_1 \sim \mathcal{N}(0, \alpha P) \text{ and } X_2 \sim \mathcal{N}(0, (1-\alpha)P)$$

For weaker receiver Y_2,

$$Y_2 = X_2 + S_2 + (X_1 + Z_2) \text{ with known state } S_2$$

Taking $U_2 = X_2 + \beta_2 S_2$ with $\beta_2 = \frac{(1-\alpha)P}{P+N_2}$

Then, achieve $\boldsymbol{R_2 \le I(U_2; Y_2) - I(U_2; S_2) \le C\left(\frac{(1-\alpha)P}{\alpha P + N_2}\right)}$

Example of DPC

For stronger receiver Y_1,

$Y_1 = X_1 + (X_2 + S_1) + Z_1$ with known state S_1

Taking $U_1 = X_1 + \beta_1(X_2 + S_1)$ with $\beta_1 = \frac{\alpha P}{\alpha P + N_1}$

Then, achieve

$$\boldsymbol{R_1 \leq C(\frac{\alpha P}{N_1})}$$

Converse Proof of Capacity Region of DPC

Gwanmo Ku

Adaptive Signal Processing and Information Theory Research Group

Nov. 9, 2012

Outline

Converse Proof

System Structure and Gel'fand Pinsker Theorem

Csiszar Sum Equality

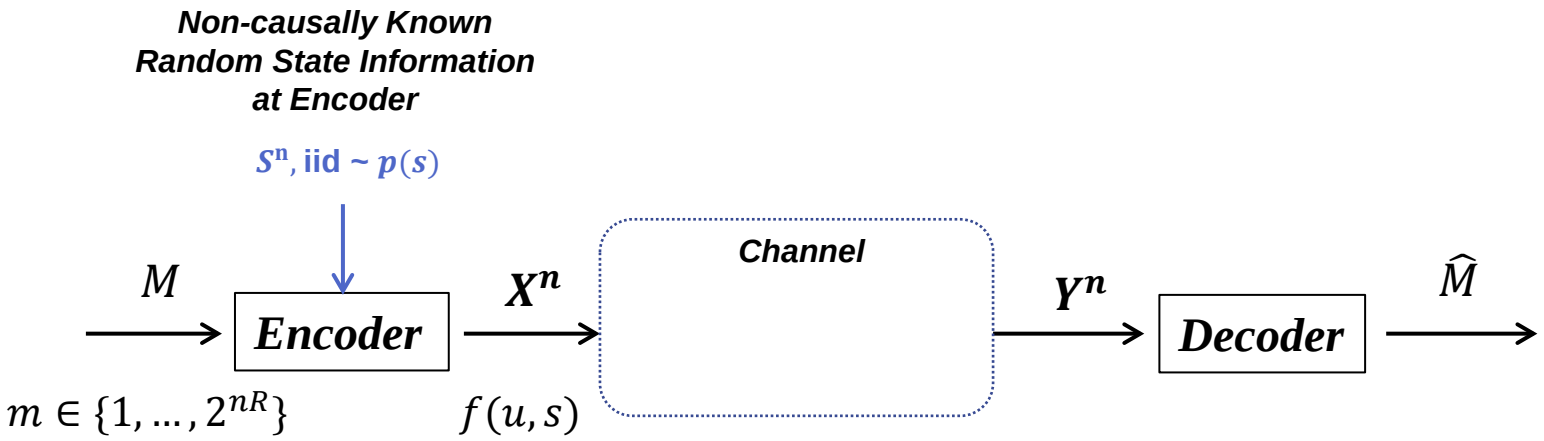
Applying Non-causality in Capacity

References

1. S. I. Gelf'and and M. S. Pinsker, "Coding for channel with random parameters", Problems of Control and Information Theory, Vol 9, No 1, pp 19-31, 1980
2. Abbas El Gamal and Young-han Kim, "Network Information Theory"
3. Thomas Cover, "Elements of Information Theory"

Non-causal State Information at Encoder

- **DMC with DM State**



✓ **Applying Gel'fand Pinsker Theorem**

$$C_{SI-E} = \max_{p(u|s),x=f(u,s)} [I(U;Y) - I(U;S)]$$

$$C_{Gel'fand} = \max_{p(u,s)} [I(U;Y) - I(U;S)]$$

Converse Proof

· Let R be an achievable rate, $\exists\ (2^{nR}, n)$ code with $P_e \to 0$ as $n \to \infty$.

· Find the auxiliary random variable U_i that forms

Markov Chain $\boldsymbol{U_i \to (X_i, S_i) \to Y_i}$,

$$nR = H(M) \quad \leftarrow +0$$

$$= H(M) - H(M|Y^n) + H(M|Y^n)$$

$$= I(M;Y^n) + H(M|Y^n)$$

← By Fano's lemma $H(M|Y^n) \le n\,\epsilon_n$

$$\le I(M;Y^n) + n\epsilon_n$$

$$= H(Y^n) - H(Y^n|M) + n\,\epsilon_n$$

Converse Proof

$$= H(Y^n) - H(Y^n|M) + n\,\epsilon_n$$

← By Chain Rule

$$= \sum_{i=1}^{n} H(Y_i|Y^{i-1}) - \sum_{i=1}^{n} H(Y_i|Y^{i-1}, M) + n\,\epsilon_n$$

← conditioning reduces entropy

$$\leq \sum_{i=1}^{n} H(Y_i) - \sum_{i=1}^{n} H(Y_i|Y^{i-1}, M) + n\,\epsilon_n$$

$$= \sum_{i=1}^{n} I(Y_i; Y^{i-1}, M) + n\,\epsilon_n$$

Converse Proof

$$= \sum_{i=1}^{n} I(Y_i; Y^{i-1}, M) \; + n\,\epsilon_n$$

Define n random variables $U_1, \ldots, U_n$ as follow

$$U_i = (M, Y_1, \ldots, Y_{i-1}, S_{i+1}, \ldots, S_n) = (M, Y^{i-1}, S_{i+1}^{n}) \cong (Q, U_Q)$$

Where $Q = (M, Y^{i-1})$ and $U_Q = S_{i+1}^{n}$

When Q is independent $Y_Q = Y_i$, then

$$I(Y_Q; Q) = I(Y_Q; U_Q, Q) - I(Y_Q; U_Q|Q)$$

$$I(Y_Q; U_Q|Q) = I(Y_Q; U_Q, Q)$$

By Eq. 15.281 in Elements of Information Theory ,Thomas Cover

$$\sum_{i=1}^{n} I(Y_i; Y^{i-1}, M) \; = \sum_{i=1}^{n} [I(M, Y^{i-1}, S_{i+1}^{n}; Y_i) - I(Y_i; S_{i+1}^{n}|M, Y^{i-1})]$$

Converse Proof

$$\leq \sum_{i=1}^{n} [I(M, Y^{i-1}, S_{i+1}^{n}; Y_i) - \boldsymbol{I(Y_i; S_{i+1}^{n} | M, Y^{i-1})}] + n\,\epsilon_n$$

← By Csiszar Sum Equality

$$= \sum_{i=1}^{n} [I(M, Y^{i-1}, S_{i+1}^{n}; Y_i) - \boldsymbol{I(Y^{i-1}; S_i | S_{i+1}^{n}, M)}] + n\,\epsilon_n$$

Csiszar Sum Equality

For Two random vectors X^n *and* Y^n *with joint probability distribution, then*

$$\sum_{i=1}^{n} \boldsymbol{I(Y_i; X_{i+1}^{n} | Y^{i-1})} = \sum_{i=1}^{n} \boldsymbol{I(Y^{i-1}; X_i | X_{i+1}^{n})}$$

Csiszar Sum Equality Proof

$$\sum_{i=1}^{n} I\left(Y_i; X_{i+1}^{n} \middle| Y^{i-1}\right) = \sum_{i=1}^{n} \sum_{j=i+1}^{n} I(X_j; Y_i | Y^{i-1}, X_{j+1}^{n})$$

Chain rule

Switching of summation

$$= \sum_{j=2}^{n} \sum_{i=1}^{j-1} I(X_j; Y_i | Y^{i-1}, X_{j+1}^{n})$$

Chain rule

$$= \sum_{j=2}^{n} I(X_j; Y^{j-1} | X_{j+1}^{n})$$

$$= \sum_{j=1}^{n} I(X_j; Y^{j-1} | X_{j+1}^{n})$$

Also define $I(X_1; Y^0 | X_2^n) = 0$

Switching of Index

$$= \sum_{i=1}^{n} I(X_i; Y^{i-1} | X_{i+1}^{n})$$

Converse Proof

$$= \sum_{i=1}^{n} [I(M, Y^{i-1}, S_{i+1}^{n}; Y_i) - I(Y^{i-1}; S_i | S_{i+1}^{n}, M)] + n\,\epsilon_n$$

Independency between (M, S_{i+1}^{n}) *and* S_i

$$= \sum_{i=1}^{n} [I(M, Y^{i-1}, S_{i+1}^{n}; Y_i) - I(S_{i+1}^{n}, M, Y^{i-1}; S_i)] + n\,\epsilon_n$$

Use $U_i = (M, Y^{i-1}, S_{i+1}^{n})$

$$nR \leq \sum_{i=1}^{n} [I(U_i; Y_i) - I(U_i; S_i)] + n\,\epsilon_n$$

$$\leq n \max_{p(u,x|s)} [I(U; Y) - I(U; S)] + n\,\epsilon_n$$

Converse Proof

$$R \leq \max_{p(u,x|s)} [I(U;Y) - I(U;S)] + \epsilon_n$$

Also show that it suffices to maximize over $p(u|s)$ and function $x = f(u,s)$

$$p(u,x|s) = p(u|s)p(x|u,s)$$

Since $x = f(u,s)$ is deterministic function of u and s,

$$p(x|u,s) = 0,1$$

For fixed $p(u|s)$,

Fixed

$$R \leq \max_{p(u|s),x=f(u,s)} [I(U;Y) - I(U;S)] + \epsilon_n$$

Only over $I(U;Y)$ Convex of $p(y|u)$ fixing $p(u|s)$

Converse Proof

Since mutual information $I(U;Y)$ is a convex function of $p(y|u)$ for a fixed $p(u|s)$

$$p(y|u) = \sum_{x,s} p(s|u)p(x|s,u)p(y|x,s)$$ linear in $p(x|u,s)$

$I(U;Y)$ is also convex of $p(x|u,s)$

Maximum is attained at an extreme points of the set of pmf $p(x|u,s)$