MAC with Action-Dependent State Information at One Encoder

Lior Dikstein, Haim Permuter and Shlomo (Shitz) Shamai

Ben Gurion university, Technion

November, 2012

This work has been supported by the CORNET Consortium Israel Ministry for Industry and Commerce

Outline

- Motivation and history
- Problem setting
- Main results
- Achievability and converse outline
- The Gaussian channel
    - The action-dependent MAC
    - The action-dependent point-to-point channel
- Rate distortion dual
- Summary

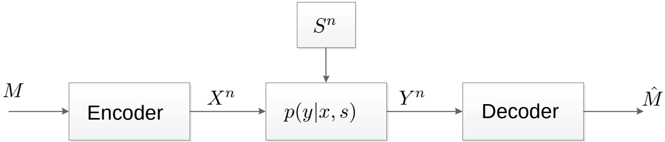
Channels with state information

Channels with state information

- Channels with state information model a communication situation where the channel is time variant:

Channels with state information

- Channels with state information model a communication situation where the channel is time variant:



  the channel is memoryless without feedback:
  $p(y^n|x^n,s^n,m) = \prod_{i=1}^{n} p(y_i|x_i,s_i)$
- Capacity of a channels where the states are available causally to the encoder [Shannon58].

Channels with state information

STATE-DEPENDENT channels characterize a significant collection of communication models

Channels with state information

STATE-DEPENDENT channels characterize a significant collection of communication models

- Interference in a wireless network

Channels with state information

STATE-DEPENDENT channels characterize a significant collection of communication models

- Interference in a wireless network
- Uncertainty about channel

Channels with state information

STATE-DEPENDENT channels characterize a significant collection of communication models

- Interference in a wireless network
- Uncertainty about channel
- Jamming (arbitrarily varying channel)

Channels with state information

STATE-DEPENDENT channels characterize a significant collection of communication models

- Interference in a wireless network
- Uncertainty about channel
- Jamming (arbitrarily varying channel)
- Channel fading

Channels with state information

STATE-DEPENDENT channels characterize a significant collection of communication models

- Interference in a wireless network
- Uncertainty about channel
- Jamming (arbitrarily varying channel)
- Channel fading
- Write-Once-Memory (WOM)

Channels with state information

STATE-DEPENDENT channels characterize a significant collection of communication models

- Interference in a wireless network
- Uncertainty about channel
- Jamming (arbitrarily varying channel)
- Channel fading
- Write-Once-Memory (WOM)
- Memory with defects

Channels with state information

STATE-DEPENDENT channels characterize a significant collection of communication models

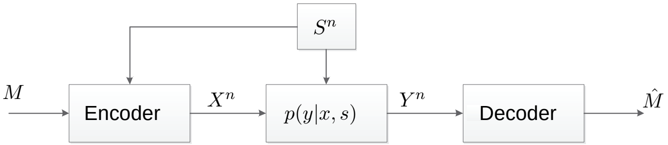
- Interference in a wireless network
- Uncertainty about channel
- Jamming (arbitrarily varying channel)
- Channel fading
- Write-Once-Memory (WOM)
- Memory with defects
- Feedback from the receiver

Noncausal state information

- Channels with noncausal side information at the encoder [Gelfand & Pinsker 80]

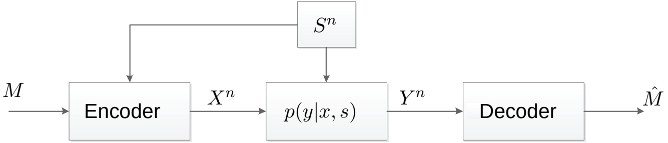
Noncausal state information

- Channels with noncausal side information at the encoder [Gelfand & Pinsker 80]



Noncausal state information

- Channels with noncausal side information at the encoder [Gelfand & Pinsker 80]



Theorem

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

for some joint distribution
$p(s,u,x,y) = p(s)p(u|s)p(x|u,s)p(y|x,s)$.

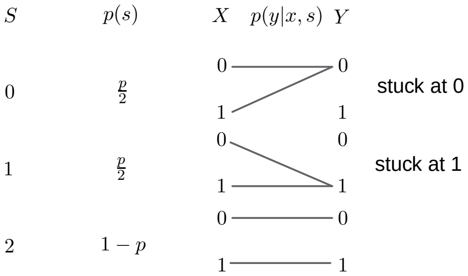
Channels with state information

One application for such a model is the Write-once memory such as a ROM or a CD-ROM.

Channels with state information

One application for such a model is the Write-once memory such as a ROM or a CD-ROM.

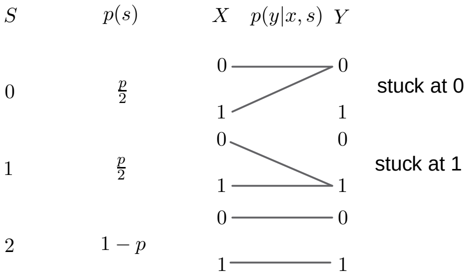
- Models a memory with stuck-at faults



Channels with state information

One application for such a model is the Write-once memory such as a ROM or a CD-ROM.

- Models a memory with stuck-at faults

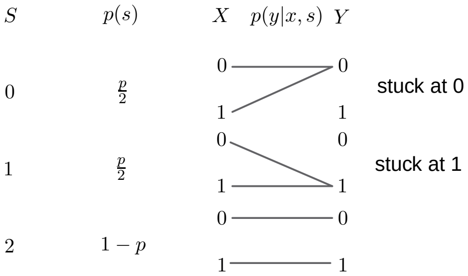


- The writer (encoder) who knows the locations of the faults (by first reading the memory)

Channels with state information

One application for such a model is the Write-once memory such as a ROM or a CD-ROM.

- Models a memory with stuck-at faults



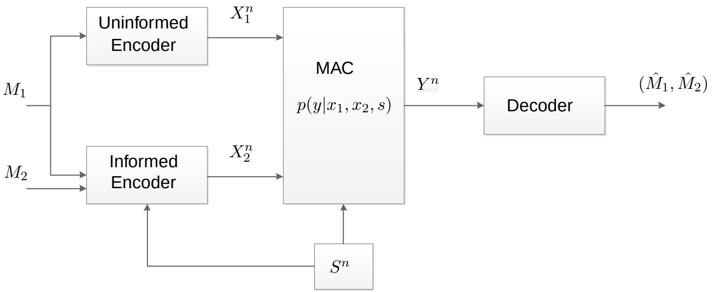
- The writer (encoder) who knows the locations of the faults (by first reading the memory)
- It wishes to reliably store information in a way that does not require the reader (decoder) to know the locations of the faults

MAC with noncausal state information

- MAC with states available at one encoder [Somekh-Baruch,Shamai & Verdú 07] [Kotagiri/Laneman07]

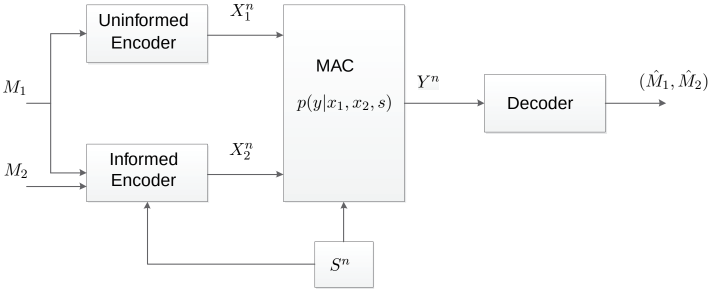
MAC with noncausal state information

- MAC with states available at one encoder [Somekh-Baruch,Shamai & Verdú 07] [Kotagiri/Laneman07]



MAC with noncausal state information

- MAC with states available at one encoder [Somekh-Baruch,Shamai & Verdú 07] [Kotagiri/Laneman07]



Theorem

$$R_2 \leq I(U;Y|X_1) - I(U;S|X_1)$$
$$R_1 + R_2 \leq I(U,X_1;Y) - I(U,X_1;S)$$

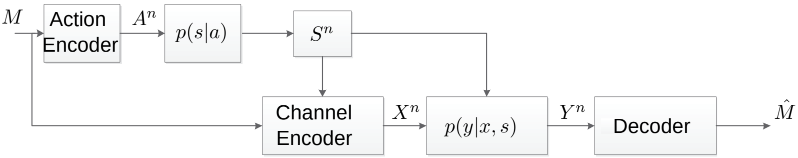
for some joint distribution $p(s,x_1,u,x_2,y) = p(s)p(x_1)p(u,x_2|s,x_1)p(y|s,x_1,x_2)$.

Action-dependent states

- Channels with Action-Dependent States [Wiessman10]

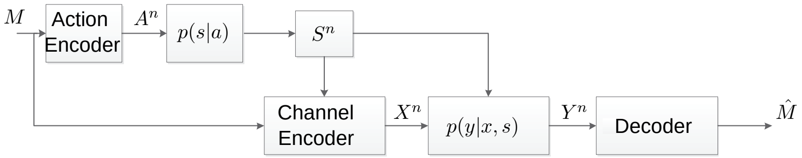
Action-dependent states

- Channels with Action-Dependent States [Wiessman10]



Action-dependent states

- Channels with Action-Dependent States [Wiessman10]

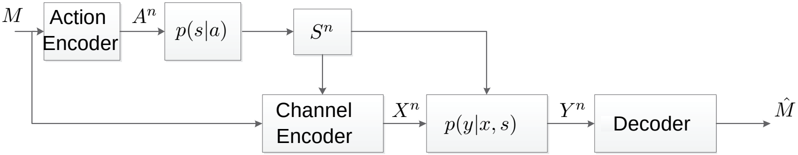


Theorem

$$\begin{aligned} C &= \max\left[I(U;Y) - I(U;S|A)\right] \\ &= \max\left[I(A,U;Y) - I(U;S|A)\right] \end{aligned}$$

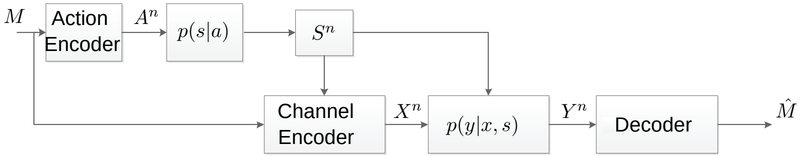
for some joint distribution
$p(a,s,u,x,y) = p(a)p(s|a)p(u|s,a)1_{x=f(u,s)}p(y|x,s)$.

Motivation



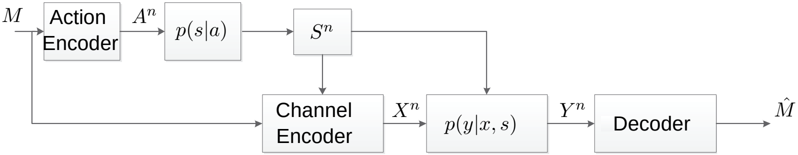
- One interpretation of the action can be a noisy public relay.

Motivation



- One interpretation of the action can be a noisy public relay.
- Provide a function of the message to the transmitter: $A(M)$ and get S, via the memoryless noisy transformation $p(s|a)$.

Motivation



- One interpretation of the action can be a noisy public relay.
- Provide a function of the message to the transmitter: $A(M)$ and get S, via the memoryless noisy transformation $p(s|a)$.
- The relay outputs are public, and monitored before hand, thus S is known at transmitter.

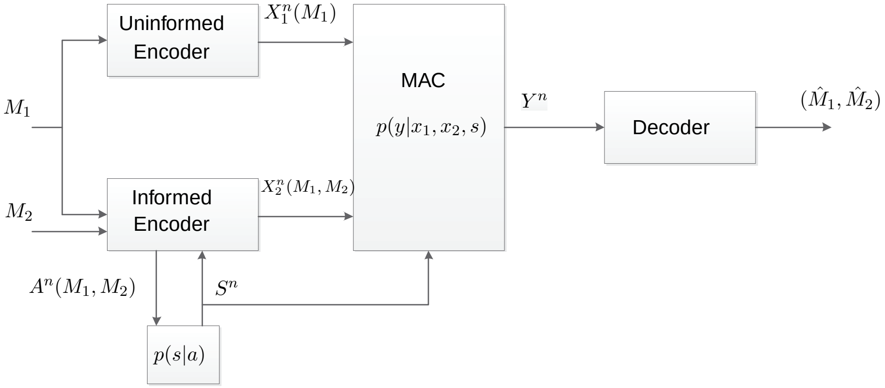
Problem setting

Problem setting

- MAC with Action-Dependent State Information at One Encoder

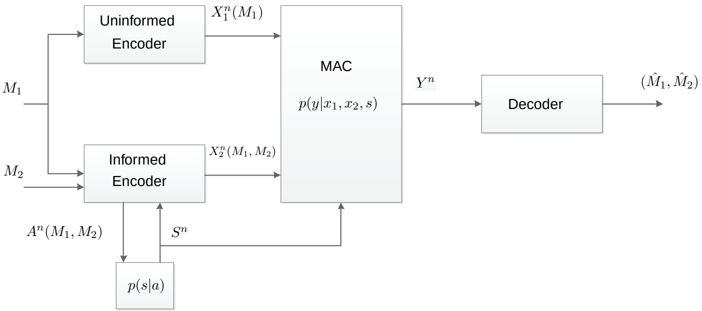
Problem setting

- MAC with Action-Dependent State Information at One Encoder

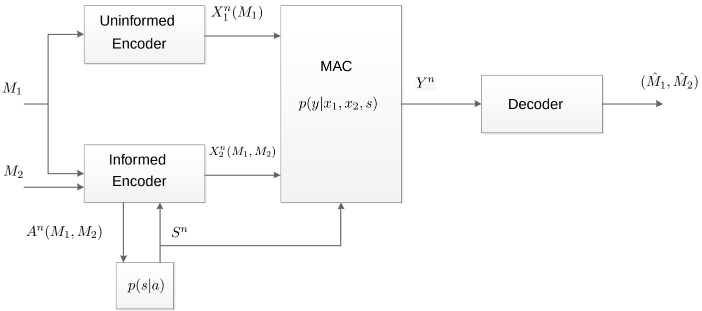


Main Results

Main Results



Main Results



Theorem

$$R_2 \le I(U;Y|X_1) - I(U;S|A,X_1)$$
$$R_1 + R_2 \le I(U,X_1;Y) - I(U,X_1;S|A)$$

for some joint distribution
$p(x_1)p(a|x_1)p(s|a)p(u|s,a,x_1)p(x_2|x_1,s,u)p(y|s,x_1,x_2)$ *and*
$|\mathcal{U}| \le |\mathcal{A}||\mathcal{S}||\mathcal{X}_1||\mathcal{X}_2| + 1$.

Intuition

Taking $\tilde{U} = (A, U)$, the following region is equivalent

$$\begin{aligned} R_2 &\leq I(A, U; Y|X_1) - I(U; S|X_1, A) \\ R_1 + R_2 &\leq I(X_1, A, U; Y) - I(X_1, U; S|A) \end{aligned}$$

Intuition

Taking $\tilde{U} = (A, U)$, the following region is equivalent

$$\begin{aligned} R_2 &\leq I(A, U; Y|X_1) - I(U; S|X_1, A) \\ R_1 + R_2 &\leq I(X_1, A, U; Y) - I(X_1, U; S|A) \end{aligned}$$

Notice that we can express the capacity region as:

$$\begin{aligned} R_2 &\leq I(A; Y|X_1) + I(U; Y|X_1, A) - I(U; S|X_1, A) \\ R_1 + R_2 &\leq I(A; Y) + I(X_1, U; Y|A) - I(X_1, U; S|A). \end{aligned}$$

Intuition

Taking $\tilde{U} = (A, U)$, the following region is equivalent

$$\begin{aligned} R_2 &\leq I(A, U; Y|X_1) - I(U; S|X_1, A) \\ R_1 + R_2 &\leq I(X_1, A, U; Y) - I(X_1, U; S|A) \end{aligned}$$

Notice that we can express the capacity region as:

$$\begin{aligned} R_2 &\leq I(A; Y|X_1) + I(U; Y|X_1, A) - I(U; S|X_1, A) \\ R_1 + R_2 &\leq I(A; Y) + I(X_1, U; Y|A) - I(X_1, U; S|A). \end{aligned}$$

- The informed encoder transmits information using the action sequence A.

Intuition

Taking $\tilde{U} = (A, U)$, the following region is equivalent

$$\begin{aligned} R_2 &\leq I(A, U; Y|X_1) - I(U; S|X_1, A) \\ R_1 + R_2 &\leq I(X_1, A, U; Y) - I(X_1, U; S|A) \end{aligned}$$

Notice that we can express the capacity region as:

$$\begin{aligned} R_2 &\leq I(A; Y|X_1) + I(U; Y|X_1, A) - I(U; S|X_1, A) \\ R_1 + R_2 &\leq I(A; Y) + I(X_1, U; Y|A) - I(X_1, U; S|A). \end{aligned}$$

- The informed encoder transmits information using the action sequence A.
- This is used at the decoder to decode a second transmission, hence the conditioning.

Intuition

Taking $\tilde{U} = (A, U)$, the following region is equivalent

$$\begin{aligned} R_2 &\leq I(A, U; Y|X_1) - I(U; S|X_1, A) \\ R_1 + R_2 &\leq I(X_1, A, U; Y) - I(X_1, U; S|A) \end{aligned}$$

Notice that we can express the capacity region as:

$$\begin{aligned} R_2 &\leq I(A; Y|X_1) + I(U; Y|X_1, A) - I(U; S|X_1, A) \\ R_1 + R_2 &\leq I(A; Y) + I(X_1, U; Y|A) - I(X_1, U; S|A). \end{aligned}$$

- The informed encoder transmits information using the action sequence A.
- This is used at the decoder to decode a second transmission, hence the conditioning.
- By Gel'fand-Pinsker given A: (U, X_1) can be decoded.

Corner Points

Another presentation for the capacity region can be achieved by applying the chain rule and the Markov $X_1 - A - S$:

$$\begin{aligned} R_2 &\leq I(A;Y|X_1) + I(U;Y|X_1,A) - I(U;S|X_1,A) \\ R_1 + R_2 &\leq I(X_1;Y) + I(A;Y|X_1) + I(U;Y|X_1,A) - I(U;S|X_1,A) \end{aligned}$$

Corner Points

Another presentation for the capacity region can be achieved by applying the chain rule and the Markov $X_1 - A - S$:

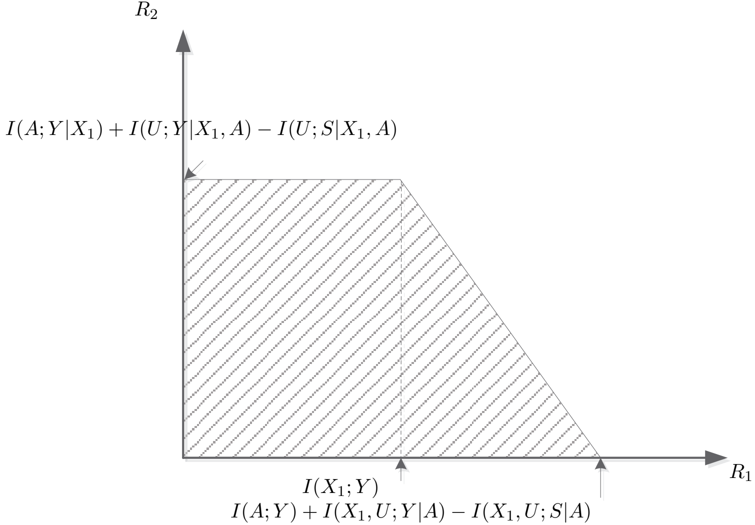
$$\begin{aligned} R_2 &\le I(A;Y|X_1) + I(U;Y|X_1,A) - I(U;S|X_1,A) \\ R_1 + R_2 &\le I(X_1;Y) + I(A;Y|X_1) + I(U;Y|X_1,A) - I(U;S|X_1,A) \end{aligned}$$

The corner points (R_1, R_2):

$$\Big(I(X_1;Y) + I(A;Y|X_1) + I(Y;U|A,X_1) - I(S;U|A,X_1) \quad , \quad 0\Big)$$

$$\Big(I(X_1;Y) \quad , \quad I(A;Y|X_1) + I(U;Y|X_1,A) - I(U;S|X_1,A)\Big)$$

Corner Points



Special Case

Special Case

- Malfunction of the Action Encode.

Special Case

- Malfunction of the Action Encode.
- We cannot choose an action that effects the formation of the states.

Special Case

- Malfunction of the Action Encode.
- We cannot choose an action that effects the formation of the states.
- The Informed Encoder still knows the states noncausaly.

Special Case

- Malfunction of the Action Encode.
- We cannot choose an action that effects the formation of the states.
- The Informed Encoder still knows the states noncausaly.
- The following expressions $I(U;S|A)$ and $I(X_1,U;S|A)$, become $I(U;S)$ and $I(X_1,U;S)$ respectively.
- We have the capacity:

$$
\begin{aligned}
R_2 &\leq I(U;Y|X_1) - I(U;S|X_1) \\
R_1 + R_2 &\leq I(U,X_1;Y) - I(U,X_1;S)
\end{aligned}
$$

Achievability Outline:

The main idea is based on a three-part coding scheme:

1. The uninformed encoder transmits X_1 at rate $I(X_1;Y)$.

Achievability Outline:

The main idea is based on a three-part coding scheme:

1. The uninformed encoder transmits X_1 at rate $I(X_1;Y)$.
2. - The informed encoder chooses an action A^n. As a result a state S^n is generated.

Achievability Outline:

The main idea is based on a three-part coding scheme:

1. The uninformed encoder transmits X_1 at rate $I(X_1; Y)$.

2.
   - The informed encoder chooses an action A^n. As a result a state S^n is generated.
   - The action sequence is sent at rate $I(A; Y|X_1)$.

Achievability Outline:

The main idea is based on a three-part coding scheme:

1. The uninformed encoder transmits X_1 at rate $I(X_1;Y)$.
2. - The informed encoder chooses an action A^n. As a result a state S^n is generated.
   - The action sequence is sent at rate $I(A;Y|X_1)$.
3. The informed encoder transmits using a Gel'fand-Pinsker scheme at rate $I(U;Y|A,X_1) - I(U;S|A,X_1)$.

Achievability Outline:

Achievability Outline:

- Choose the codeword $X_1(M_1)$ from Encoder 1's codebook of size 2^{nR_1}.
- Choose an action sequence $A^n(M_1, M_2)$.

Achievability Outline:

- Choose the codeword $X_1(M_1)$ from Encoder 1's codebook of size 2^{nR_1}.
- Choose an action sequence $A^n(M_1, M_2)$.
- As a result, a state S^n is generated.

Achievability Outline:

- Choose the codeword $X_1(M_1)$ from Encoder 1's codebook of size 2^{nR_1}.
- Choose an action sequence $A^n(M_1, M_2)$.
- As a result, a state S^n is generated.
- Encoder 2 chooses a codeword $U^n(k)$ from bin (M_1, M_2) such that $(U^n, X_1^n, A^n, S^n) \in \mathcal{T}_\epsilon^{(n)}(U, X_1, A, S)$.

Achievability Outline:

- Choose the codeword $X_1(M_1)$ from Encoder 1's codebook of size 2^{nR_1}.
- Choose an action sequence $A^n(M_1, M_2)$.
- As a result, a state S^n is generated.
- Encoder 2 chooses a codeword $U^n(k)$ from bin (M_1, M_2) such that $(U^n, X_1^n, A^n, S^n) \in \mathcal{T}_\epsilon^{(n)}(U, X_1, A, S)$.
- The decoder looks for the smallest value of $(\hat{M}_1, \hat{M}_2)$ for which exists a $\hat{k}$ such that: $\big(U^n(\hat{M}_1, \hat{M}_2, k), X_1^n(\hat{M}_1), A^n(\hat{M}_1, \hat{M}_2), Y^n\big) \in \mathcal{T}_\epsilon^{(n)}(U, X_1, A, Y)$.

Achievability Outline: Codebook Generation

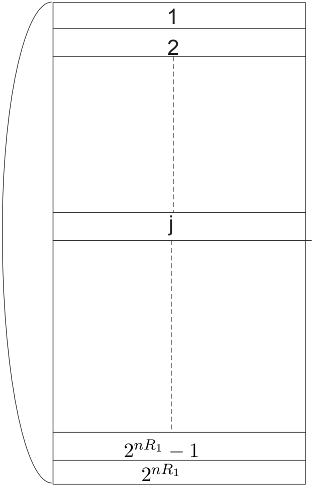
Achievability Outline: Codebook Generation

- Encoder 1 generates a codebook of 2^{nR_1} codewords $\sim p(x_1)$.

Achievability Outline: Codebook Generation

- Encoder 1 generates a codebook of 2^{nR_1} codewords $\sim p(x_1)$.

2^{nR_1} codewords $x_1^n \sim p(x_1)$



Achievability Outline: Codebook Generation

- Encoder 2 generates $2^{n(R_1+R_2)}$ action sequences $A^n(m_1,m_2) \sim p(a|x_1)$.

Achievability Outline: Codebook Generation

- Encoder 2 generates $2^{n(R_1+R_2)}$ action sequences $A^n(m_1, m_2) \sim p(a|x_1)$.
- Generate $2^{n(R_1+R_2)}$ bins, one for each set of messages (m_1, m_2).

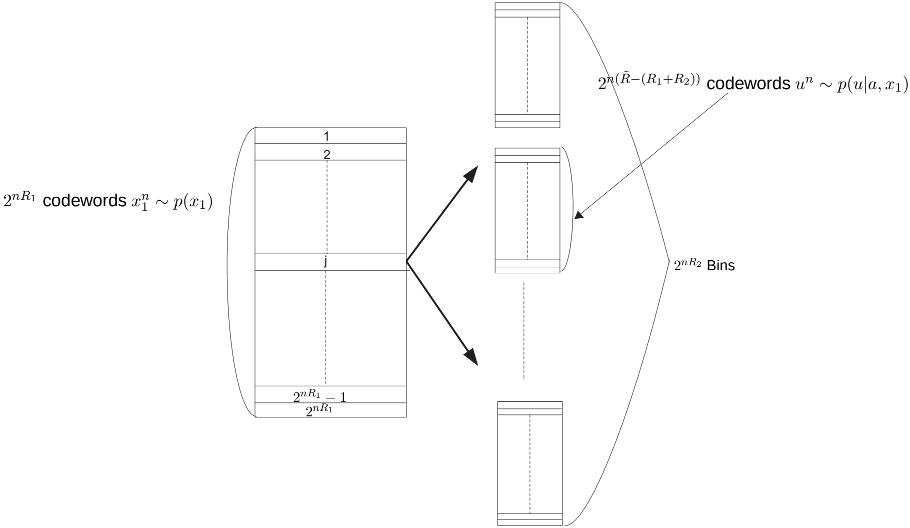
Achievability Outline: Codebook Generation

- Encoder 2 generates $2^{n(R_1+R_2)}$ action sequences $A^n(m_1, m_2) \sim p(a|x_1)$.
- Generate $2^{n(R_1+R_2)}$ bins, one for each set of messages (m_1, m_2).
- Generate randomly $2^{n\tilde{R}}$ codewords $u^n(1), ..., u^n(2^{n\tilde{R}})$ according to $\sim p(u|a, x_1)$.

Achievability Outline: Codebook Generation

- Encoder 2 generates $2^{n(R_1+R_2)}$ action sequences $A^n(m_1, m_2) \sim p(a|x_1)$.
- Generate $2^{n(R_1+R_2)}$ bins, one for each set of messages (m_1, m_2).
- Generate randomly $2^{n\tilde{R}}$ codewords $u^n(1), ..., u^n(2^{n\tilde{R}})$ according to $\sim p(u|a, x_1)$.
- Distribute the codewords uniformly to the bins, giving us a subcodebook $c(m_1, m_2)$ for each message set of $2^{n(\tilde{R}-(R_1+R_2))}$ codewords.

Achievability Outline: Codebook Generation



Converse outline

We have to show that for any $(2^{nR_1}, 2^{nR_2}, n)$ code with $P_{\text{error}} \to 0$ as $n \to \infty$ we must have

$$R_2 \leq I(U;Y|X_1) - I(U;S|A,X_1)$$
$$R_1 + R_2 \leq I(U,X_1;Y) - I(U,X_1;S|A)$$

Converse outline

We have to show that for any $(2^{nR_1}, 2^{nR_2}, n)$ code with $P_{\text{error}} \to 0$ as $n \to \infty$ we must have

$$\begin{aligned} R_2 &\le I(U;Y|X_1) - I(U;S|A,X_1) \\ R_1 + R_2 &\le I(U,X_1;Y) - I(U,X_1;S|A) \end{aligned}$$

- We use Fano's inequality in the form of $H(M_1, M_2|Y^n) \le n(R_1 + R_2)P_e^{(n)} + H(P_e^{(n)})$.

Converse outline

We have to show that for any $(2^{nR_1}, 2^{nR_2}, n)$ code with $P_{\text{error}} \to 0$ as $n \to \infty$ we must have

$$R_2 \leq I(U;Y|X_1) - I(U;S|A,X_1)$$
$$R_1 + R_2 \leq I(U,X_1;Y) - I(U,X_1;S|A)$$

- We use Fano's inequality in the form of $H(M_1, M_2|Y^n) \leq n(R_1 + R_2)P_e^{(n)} + H(P_e^{(n)})$.
- We use the Csiszar sum identity, $\sum_{i=1}^{n} I(X_{i+1}^n; Y_i|Y^{i-1}) = \sum_{i=1}^{n} I(Y^{i-1}; X_i|X_{i+1}^n)$

Converse outline

We have to show that for any $(2^{nR_1}, 2^{nR_2}, n)$ code with $P_{\text{error}} \to 0$ as $n \to \infty$ we must have

$$R_2 \leq I(U;Y|X_1) - I(U;S|A,X_1)$$
$$R_1 + R_2 \leq I(U,X_1;Y) - I(U,X_1;S|A)$$

- We use Fano's inequality in the form of $H(M_1,M_2|Y^n) \leq n(R_1+R_2)P_e^{(n)} + H(P_e^{(n)})$.
- We use the Csiszar sum identity, $\sum_{i=1}^{n} I(X_{i+1}^n;Y_i|Y^{i-1}) = \sum_{i=1}^{n} I(Y^{i-1};X_i|X_{i+1}^n)$
- We identify our auxiliary random variable, $U_i = (X_1^{i-1}, X_{i+1}^n, S_{i+1}^n, Y^{i-1}, A^n, M_1, M_2)$.

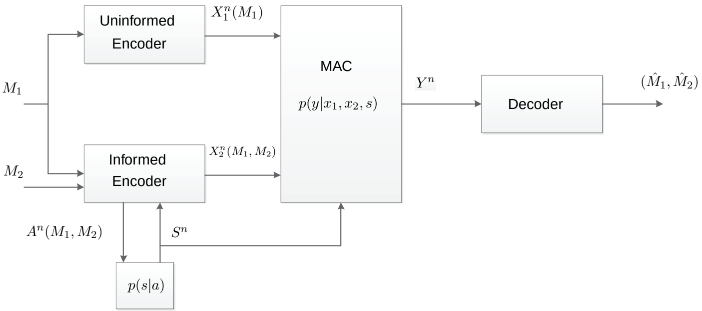
Converse outline

We have to show that for any $(2^{nR_1}, 2^{nR_2}, n)$ code with $P_{\text{error}} \to 0$ as $n \to \infty$ we must have

$$R_2 \leq I(U;Y|X_1) - I(U;S|A,X_1)$$
$$R_1 + R_2 \leq I(U,X_1;Y) - I(U,X_1;S|A)$$

- We use Fano's inequality in the form of $H(M_1,M_2|Y^n) \leq n(R_1+R_2)P_e^{(n)} + H(P_e^{(n)})$.
- We use the Csiszar sum identity, $\sum_{i=1}^n I(X_{i+1}^n;Y_i|Y^{i-1}) = \sum_{i=1}^n I(Y^{i-1};X_i|X_{i+1}^n)$
- We identify our auxiliary random variable, $U_i = (X_1^{i-1}, X_{i+1}^n, S_{i+1}^n, Y^{i-1}, A^n, M_1, M_2)$.
- We use a time-sharing random variable Q uniformly distributed in $\{1,2,...,n\}$.

Main Results



Theorem

$$R_2 \leq I(U;Y|X_1) - I(U;S|A,X_1)$$
$$R_1 + R_2 \leq I(U,X_1;Y) - I(U,X_1;S|A)$$

for some joint distribution $p(x_1)p(a|x_1)p(s|a)p(u|s,a,x_1)p(x_2|x_1,s,u)p(y|s,x_1,x_2)$ *and* $|\mathcal{U}| \leq |\mathcal{A}||\mathcal{S}||\mathcal{X}_1||\mathcal{X}_2| + 1$.

Gaussian Channel-Channel Model

Gaussian Channel-Channel Model

- The channel probability is defined by the following relations between X_1, X_2, S and Y:

$$\begin{aligned} Y_i &= X_{1,i}(M_1) + X_{2,i}(M_1, M_2, S^n) + S_i + Z_i \\ &= X_{1,i}(M_1) + X_{2,i}(M_1, M_2, S^n) + A_i(M_1, M_2) + W_i + Z_i \end{aligned}$$

Gaussian Channel-Channel Model

- The channel probability is defined by the following relations between X_1, X_2, S and Y:

$$
\begin{aligned}
Y_i &= X_{1,i}(M_1) + X_{2,i}(M_1, M_2, S^n) + S_i + Z_i \\
&= X_{1,i}(M_1) + X_{2,i}(M_1, M_2, S^n) + A_i(M_1, M_2) + W_i + Z_i
\end{aligned}
$$

- $S^n = A^n(M_1, M_2) + W^n$.

Gaussian Channel-Channel Model

- The channel probability is defined by the following relations between X_1, X_2, S and Y:

$$\begin{aligned} Y_i &= X_{1,i}(M_1) + X_{2,i}(M_1, M_2, S^n) + S_i + Z_i \\ &= X_{1,i}(M_1) + X_{2,i}(M_1, M_2, S^n) + A_i(M_1, M_2) + W_i + Z_i \end{aligned}$$

- $S^n = A^n(M_1, M_2) + W^n$.
- Z^n and W^n are independent, W^n is i.i.d.$\sim N(0, Q)$ and Z^n is i.i.d.$\sim N(0, N)$.

Gaussian Channel-Channel Model

- The channel probability is defined by the following relations between X_1, X_2, S and Y:

$$
\begin{aligned}
Y_i &= X_{1,i}(M_1) + X_{2,i}(M_1, M_2, S^n) + S_i + Z_i \\
&= X_{1,i}(M_1) + X_{2,i}(M_1, M_2, S^n) + A_i(M_1, M_2) + W_i + Z_i
\end{aligned}
$$

- $S^n = A^n(M_1, M_2) + W^n$.
- Z^n and W^n are independent, W^n is i.i.d.$\sim N(0, Q)$ and Z^n is i.i.d.$\sim N(0, N)$.
- We have the following power constraints:

$$
\frac{1}{n}\sum_{i=1}^{n}(X_{1i})^2 \le P_1 \qquad \frac{1}{n}\sum_{i=1}^{n}(X_{2i})^2 \le P_2
$$

$$
\text{and} \quad \frac{1}{n}\sum_{i=1}^{n}(A_i)^2 \le P_A.
$$

Results-Gaussian Action MAC

Theorem

Results-Gaussian Action MAC

Theorem

$$R_2 \leq \frac{1}{2}\log \frac{\Big(N + P_2 + P_A + Q - P_2\rho_{12}^2 - P_A\rho_{1A}^2 + 2\sqrt{P_2P_A}\rho_{2A} - 2\sqrt{P_2P_A}\rho_{12}\rho_{1A} + 2\sqrt{P_2Q}\rho_{2W}}{N\Big((\rho_{1A}^2 - 1)(N + Q + P_2\rho_{2W}^2 + 2\sqrt{P_2Q}\rho_{2W}) - P_2\Delta\Big)}$$

$$+ \frac{1}{2}\log\Big(N(\rho_{1A}^2 - 1) - P_2\Delta\Big)$$

$$R_1 + R_2 \leq \frac{1}{2}\log \frac{\Big(N + P_1 + P_2 + P_A + Q + 2\sqrt{P_1P_2}\rho_{12} + 2\sqrt{P_1P_A}\rho_{1A} + 2\sqrt{P_2P_A}\rho_{2A} + 2\sqrt{P_2Q}\rho_{2W}\Big)}{N\Big((\rho_{1A}^2 - 1)(N + Q + P_2\rho_{2W}^2 + 2\sqrt{P_2Q}\rho_{2W}) - P_2\Delta\Big)},$$

$$+ \frac{1}{2}\log\Big(N(\rho_{1A}^2 - 1) - P_2\Delta\Big)$$

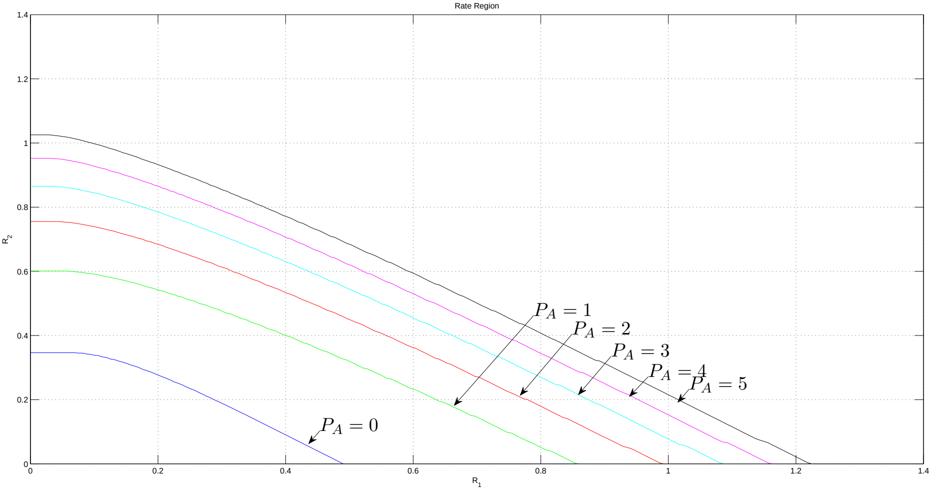
for some $\rho_{12} \in [-1, 1]$, $\rho_{1A} \in [-1, 1]$, $\rho_{2A} \in [-1, 1]$, $\rho_{2W} \in [-1, 1]$ *where*

$$\Delta = 1 - \rho_{12}^2 - \rho_{1A}^2 - \rho_{2A}^2 - \rho_{2W}^2 + \rho_{1A}^2\rho_{2W}^2 + 2\rho_{1A}\rho_{2A}\rho_{12},$$

such that

$$\Delta \geq 0.$$

Capacity Region-Gaussian Action MAC



Proof Outline-Converse

Proof Outline-Converse

- We state two lemmas that show that our region is upper-bounded by:

$$
\begin{aligned}
R_2 &\leq I(U;Y|X_1) - I(U;S|A,X_1) \\
&\leq I(A;Y|X_1) + h(X_2|X_1,A,W) - h(X_2 - \hat{X}_2^{\text{lin}}(X_1,A,W,X_2+Z))
\end{aligned}
$$

$$
\begin{aligned}
R_1 + R_2 &\leq I(U,X_1;Y) - I(U,X_1;S|A) \\
&\leq I(A,X_1;Y) + h(X_2|X_1,A,W) - h(X_2 - \hat{X}_2^{\text{lin}}(X_1,A,W,X_2+Z))
\end{aligned}
$$

Proof Outline-Converse

- We state two lemmas that show that our region is upper-bounded by:

$$\begin{aligned} R_2 &\leq I(U;Y|X_1) - I(U;S|A,X_1) \\ &\leq I(A;Y|X_1) + h(X_2|X_1,A,W) - h(X_2 - \hat{X}_2^{\text{lin}}(X_1,A,W,X_2+Z)) \end{aligned}$$

$$\begin{aligned} R_1 + R_2 &\leq I(U,X_1;Y) - I(U,X_1;S|A) \\ &\leq I(A,X_1;Y) + h(X_2|X_1,A,W) - h(X_2 - \hat{X}_2^{\text{lin}}(X_1,A,W,X_2+Z)) \end{aligned}$$

- We show that it suffices to consider only jointly Gaussian random variables.

Proof Outline-Converse

- We state two lemmas that show that our region is upper-bounded by:

$$
\begin{aligned}
R_2 &\leq I(U;Y|X_1) - I(U;S|A,X_1) \\
&\leq I(A;Y|X_1) + h(X_2|X_1,A,W) - h(X_2 - \hat{X}_2^{\text{lin}}(X_1,A,W,X_2+Z)) \\
&= \frac{1}{2}\log\big(\frac{\sigma^2_{Y|X_1}\sigma^2_{W|Y,X_1,A}}{QN}\big) \\
R_1 + R_2 &\leq I(U,X_1;Y) - I(U,X_1;S|A) \\
&\leq I(A,X_1;Y) + h(X_2|X_1,A,W) - h(X_2 - \hat{X}_2^{\text{lin}}(X_1,A,W,X_2+Z)) \\
&= \frac{1}{2}\log\big(\frac{\sigma^2_{Y}\sigma^2_{W|Y,X_1,A}}{QN}\big)
\end{aligned}
$$

- We show that it suffices to consider only jointly Gaussian random variables.

Proof Outline-Converse

- We state two lemmas that show that our region is upper-bounded by:

$$
\begin{aligned}
R_2 &\leq I(U;Y|X_1) - I(U;S|A,X_1) \\
&\leq I(A;Y|X_1) + h(X_2|X_1,A,W) - h(X_2 - \hat{X}_2^{\text{lin}}(X_1,A,W,X_2+Z)) \\
&= \frac{1}{2}\log\left(\frac{\sigma^2_{Y|X_1}\sigma^2_{W|Y,X_1,A}}{QN}\right) \\
R_1 + R_2 &\leq I(U,X_1;Y) - I(U,X_1;S|A) \\
&\leq I(A,X_1;Y) + h(X_2|X_1,A,W) - h(X_2 - \hat{X}_2^{\text{lin}}(X_1,A,W,X_2+Z)) \\
&= \frac{1}{2}\log\left(\frac{\sigma^2_{Y}\sigma^2_{W|Y,X_1,A}}{QN}\right)
\end{aligned}
$$

- We show that it suffices to consider only jointly Gaussian random variables.
- Now we define $E[X_1^2] \triangleq \sigma^2_{X_1}$, $E[X_2^2] \triangleq \sigma^2_{X_2}$, $E[A^2] \triangleq \sigma^2_A$ and calculate the expression.

Proof Outline-Converse

$$
\begin{aligned}
R_2 &\le I(U;Y|X_1) - I(U;S|A,X_1)\\
&\le I(A;Y|X_1) + h(X_2|X_1,A,W) - h(X_2 - \hat{X}_2^{\text{lin}}(X_1,A,W,X_2+Z))\\
&= \frac{1}{2}\log\left(\frac{\sigma^2_{Y|X_1}\sigma^2_{W|Y,X_1,A}}{QN}\right)\\
&= \frac{1}{2}\log\frac{\left(N+\sigma^2_{X_2}+\sigma^2_A+Q-\sigma^2_{X_2}\rho^2_{12}-\sigma^2_A\rho^2_{1A}+2\sqrt{\sigma^2_{X_2}\sigma^2_A}\rho_{2A}-2\sqrt{\sigma^2_{X_2}\sigma^2_A}\rho_{12}\rho_{1A}+2\sqrt{\sigma^2_{X_2}Q}\rho_{2W}\right)}{N\left((\rho^2_{1A}-1)(N+Q+\sigma^2_{X_2}\rho^2_{2W}+2\sqrt{\sigma^2_{X_2}Q}\rho_{2W})-\sigma^2_{X_2}\Delta\right)}\\
&\quad+\frac{1}{2}\log\left(N(\rho^2_{1A}-1)-\sigma^2_{X_2}\Delta\right)
\end{aligned}
$$

such that

$$
\sigma^2_{X_1} \le P_1 \quad \sigma^2_{X_2} \le P_2 \quad \sigma^2_A \le P_A.
$$

Proof Outline-Converse

$$
\begin{aligned}
R_1 + R_2 &\leq I(U, X_1; Y) - I(U, X_1; S|A) \\
&\leq I(A, X_1; Y) + h(X_2|X_1, A, W) - h(X_2 - \hat{X}_2^{\text{lin}}(X_1, A, W, X_2 + Z)) \\
&= \frac{1}{2} \log \big(\frac{\sigma_Y^2 \sigma_{W|Y,X_1,A}^2}{QN}\big) \\
&= \frac{1}{2} \log \frac{\Big(N + \sigma_{X_1}^2 + \sigma_{X_2}^2 + \sigma_A^2 + Q + 2\sqrt{\sigma_{X_1}^2 \sigma_{X_2}^2}\rho_{12} + 2\sqrt{\sigma_{X_1}^2 \sigma_A^2}\rho_{1A} + 2\sqrt{\sigma_{X_2}^2 \sigma_A^2}\rho_{2A} + 2\sqrt{\sigma_{X_2}^2 Q}\rho_{2W}\Big)}{N\Big((\rho_{1A}^2 - 1)(N + Q + \sigma_{X_2}^2 \rho_{2W}^2 + 2\sqrt{\sigma_{X_2}^2 Q}\rho_{2W}) - \sigma_{X_2}^2 \Delta\Big)} \\
&\quad + \frac{1}{2} \log \Big(N(\rho_{1A}^2 - 1) - \sigma_{X_2}^2 \Delta\Big)
\end{aligned}
$$

such that

$$
\sigma_{X_1}^2 \leq P_1 \quad \sigma_{X_2}^2 \leq P_2 \quad \sigma_A^2 \leq P_A
$$

Proof Outline-Converse

The values of the covariances are such that the covariance matrix

$$\Lambda = \begin{pmatrix} \sigma^2_{X_1} & \sigma_{12} & \sigma_{1A} & 0 & 0 \\ \sigma_{12} & \sigma^2_{X_2} & \sigma_{2A} & \sigma_{2W} & 0 \\ \sigma_{1A} & \sigma_{2A} & \sigma^2_A & 0 & 0 \\ 0 & \sigma_{2W} & 0 & Q & 0 \\ 0 & 0 & 0 & 0 & N \end{pmatrix},$$

satisfies the nonnegative-definiteness condition

$$\det(\Lambda) = \sigma^2_{1A}\sigma^2_{2W}N\sigma^2_{X_1}\sigma^2_A + 2\sigma_{12}\sigma_{1A}\sigma_{2A}NQ - \sigma^2_{2A}N\sigma^2_{X_1}Q - \sigma^2_{12}N\sigma^2_A Q + N\sigma^2_{X_1}\sigma^2_{X_2}\sigma^2_A$$

or equivalently as a function of $\rho_{12}, \rho_{1A}, \rho_{2A}$ and ρ_{2W}

$$1 - \rho^2_{12} - \rho^2_{1A} - \rho^2_{2A} - \rho^2_{2W} + \rho^2_{1A}\rho^2_{2W} + 2\rho_{1A}\rho_{2A}\rho_{12} \geq 0$$

Proof Outline-Converse

- We show that replacing $\sigma^2_{X_1}$, $\sigma^2_{X_2}$, σ^2_A with P_1, P_2 and P_A respectively, further increases the region.

Proof Outline-Converse

- We show that replacing $\sigma_{X_1}^2$, $\sigma_{X_2}^2$, σ_A^2 with P_1, P_2 and P_A respectively, further increases the region.
- Substituting $\sigma_{X_1}^2$, $\sigma_{X_2}^2$, σ_A^2 with P_1, P_2 and P_A, we obtain the capacity region of the theorem.

Proof Outline-Converse

- We show that replacing $\sigma^2_{X_1}$, $\sigma^2_{X_2}$, σ^2_A with P_1, P_2 and P_A respectively, further increases the region.
- Substituting $\sigma^2_{X_1}$, $\sigma^2_{X_2}$, σ^2_A with P_1, P_2 and P_A, we obtain the capacity region of the theorem.
- To conclude, the upper bound is obtained as an optimization problem on $\rho_{12} \in [-1, 1]$, $\rho_{1A} \in [-1, 1]$, $\rho_{2A} \in [-1, 1]$ and $\rho_{2W} \in [-1, 1]$.

Proof Outline-Converse

- We show that replacing $\sigma^2_{X_1}$, $\sigma^2_{X_2}$, σ^2_A with P_1, P_2 and P_A respectively, further increases the region.
- Substituting $\sigma^2_{X_1}$, $\sigma^2_{X_2}$, σ^2_A with P_1, P_2 and P_A, we obtain the capacity region of the theorem.
- To conclude, the upper bound is obtained as an optimization problem on $\rho_{12} \in [-1, 1]$, $\rho_{1A} \in [-1, 1]$, $\rho_{2A} \in [-1, 1]$ and $\rho_{2W} \in [-1, 1]$.
- In the achievability part, we show that this bound is also achievable.

Proof Outline-Direct Part

- We choose specific distributions of our r.v.

Proof Outline-Direct Part

- We choose specific distributions of our r.v.
- We take (X_1, X_2, A, W, Y) to be jointly Gaussian.

Proof Outline-Direct Part

- We choose specific distributions of our r.v.
- We take (X_1, X_2, A, W, Y) to be jointly Gaussian.
- We choose random variables $X_1 \sim N(0, P_1)$, $X_2 \sim N(0, P_2)$, $A \sim N(0, P_A)$.

Proof Outline-Direct Part

- We choose specific distributions of our r.v.
- We take (X_1, X_2, A, W, Y) to be jointly Gaussian.
- We choose random variables $X_1 \sim N(0, P_1)$, $X_2 \sim N(0, P_2)$, $A \sim N(0, P_A)$.
- We choose the auxiliary r.v.

$$\begin{aligned} U &= X_1 + X_2 + \beta S \\ &= X_1 + X_2 + \beta(A + W). \end{aligned}$$

Proof Outline-Direct Part

- Substituting $U = X_1 + X_2 + \beta(A+W)$ in the capacity region:

$$\begin{aligned} R_2 &\leq I(U;Y|X_1) - I(U;S|A,X_1) \\ R_1 + R_2 &\leq I(U,X_1;Y) - I(U,X_1;S|A) \end{aligned}$$

Proof Outline-Direct Part

- Substituting $U = X_1 + X_2 + \beta(A + W)$ in the capacity region:

$$R_2 \leq I(U;Y|X_1) - I(U;S|A,X_1)$$
$$R_1 + R_2 \leq I(U,X_1;Y) - I(U,X_1;S|A)$$

we achieve the equalities of the upper bound

$$\begin{aligned} R_2 &\leq I(U;Y|X_1) - I(U;S|A,X_1) \\ &= I(A;Y|X_1) + h(X_2|X_1,A,W) - h(X_2 - \hat{X}_2^{\text{lin}}(X_1,A,W,X_2+Z)) \end{aligned}$$

$$\begin{aligned} R_1 + R_2 &\leq I(U,X_1;Y) - I(U,X_1;S|A) \\ &= I(A,X_1;Y) + h(X_2|X_1,A,W) - h(X_2 - \hat{X}_2^{\text{lin}}(X_1,A,W,X_2+Z)) \end{aligned}$$

Gaussian Channel-Remarks

- The capacity for the Gaussian action-dependent point-to-point channel, was left open in [Wiessman10].

Gaussian Channel-Remarks

- The capacity for the Gaussian action-dependent point-to-point channel, was left open in [Wiessman10].
- We find the capacity for the the action dependent point-to-point channel by taking similar steps.

Gaussian Channel-Remarks

- The capacity for the Gaussian action-dependent point-to-point channel, was left open in [Wiessman10].
- We find the capacity for the the action dependent point-to-point channel by taking similar steps.
- We can derive the capacity of the p.t.p channel from the Action-MAC by taking $R_1 = 0$.

Gaussian Channel-Remarks

- The capacity for the Gaussian action-dependent point-to-point channel, was left open in [Wiessman10].
- We find the capacity for the the action dependent point-to-point channel by taking similar steps.
- We can derive the capacity of the p.t.p channel from the Action-MAC by taking $R_1 = 0$.
- We give an alternative proof for the capacity of the point to point channel

Gaussian Channel-Remarks

- The capacity for the Gaussian action-dependent point-to-point channel, was left open in [Wiessman10].
- We find the capacity for the the action dependent point-to-point channel by taking similar steps.
- We can derive the capacity of the p.t.p channel from the Action-MAC by taking $R_1 = 0$.
- We give an alternative proof for the capacity of the point to point channel
- We obtain a one-to-one correspondence with the Gaussian GGP MAC [Somekh-Baruch,Shamai & Verdú 07]: with only a common message.

Duality Channel-Source Coding with Action

Duality Channel-Source Coding with Action

- An information-theoretic duality between our Action-MAC and the ”Successive Refinement with Actions” [Chia, Asnani & Weissman 11].

Duality Channel-Source Coding with Action

- An information-theoretic duality between our Action-MAC and the ”Successive Refinement with Actions” [Chia, Asnani & Weissman 11].
- For a given channel-coding problem, we obtain a rate-distortion problem and vice versa.

Duality Channel-Source Coding with Action

- An information-theoretic duality between our Action-MAC and the ”Successive Refinement with Actions” [Chia, Asnani & Weissman 11].
- For a given channel-coding problem, we obtain a rate-distortion problem and vice versa.
- The roles of the encoders and decoders are functionally interchangeable.

Duality Channel-Source Coding with Action

- An information-theoretic duality between our Action-MAC and the ”Successive Refinement with Actions” [Chia, Asnani & Weissman 11].
- For a given channel-coding problem, we obtain a rate-distortion problem and vice versa.
- The roles of the encoders and decoders are functionally interchangeable.
- The input-output joint distribution is equivalent with some renaming of variables.

Duality Channel-Source Coding with Action

- An information-theoretic duality between our Action-MAC and the ”Successive Refinement with Actions” [Chia, Asnani & Weissman 11].
- For a given channel-coding problem, we obtain a rate-distortion problem and vice versa.
- The roles of the encoders and decoders are functionally interchangeable.
- The input-output joint distribution is equivalent with some renaming of variables.
- Recognizing this duality, further dualities emerge:

Duality Channel-Source Coding with Action

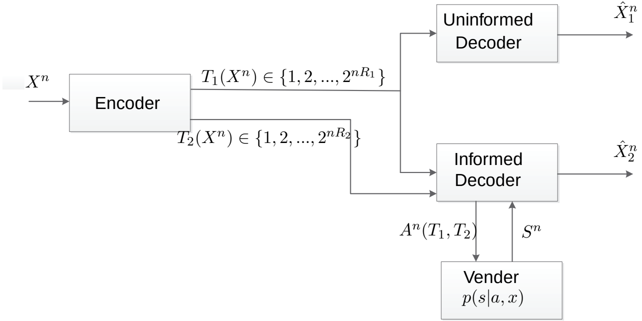
- An information-theoretic duality between our Action-MAC and the ”Successive Refinement with Actions” [Chia, Asnani & Weissman 11].
- For a given channel-coding problem, we obtain a rate-distortion problem and vice versa.
- The roles of the encoders and decoders are functionally interchangeable.
- The input-output joint distribution is equivalent with some renaming of variables.
- Recognizing this duality, further dualities emerge:
    1. A rate distortion dual for the action dependent point-to-point channel.

Duality Channel-Source Coding with Action

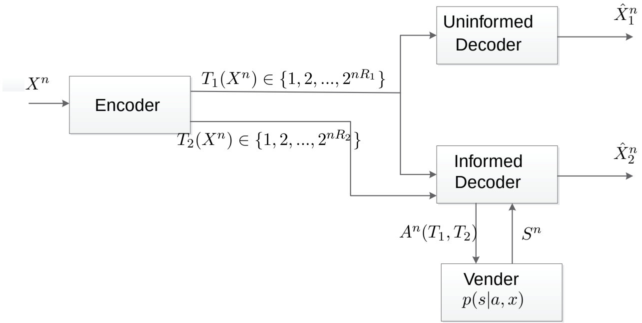
- An information-theoretic duality between our Action-MAC and the ”Successive Refinement with Actions” [Chia, Asnani & Weissman 11].
- For a given channel-coding problem, we obtain a rate-distortion problem and vice versa.
- The roles of the encoders and decoders are functionally interchangeable.
- The input-output joint distribution is equivalent with some renaming of variables.
- Recognizing this duality, further dualities emerge:
  1. A rate distortion dual for the action dependent point-to-point channel.
  2. A rate distortion dual for the GGP MAC.

The ”Successive Refinement with Actions” model

The "Successive Refinement with Actions" model



The "Successive Refinement with Actions" model



Theorem

$$
\begin{aligned}
R_1 &\geq I(X; \hat{X}_1) \\
R_1 + R_2 &\geq I(X; \hat{X}_1) + I(A; X|\hat{X}_1) + I(X; U|X, A, \hat{X}_1)
\end{aligned}
$$

for some joint distribution $P(x, a, u, s, \hat{x}_1) = P(x)P(a, u, \hat{x}_1|x)P(s|x, a)$

Duality Transformation Principles

Duality Transformation Principles

Channel Coding $\leftrightarrow$ Rate Distortion

- Encoder inputs / Decoder outputs: $\leftrightarrow$ Decoder inputs / Encoder outputs:
  $M_1 \in \{1, 2, ..., 2^{nR_1}\} \leftrightarrow T_1 \in \{1, 2, ..., 2^{nR_1}\}$
  $M_2 \in \{1, 2, ..., 2^{nR_2}\} \leftrightarrow T_2 \in \{1, 2, ..., 2^{nR_2}\}$

Duality Transformation Principles

Channel Coding $\leftrightarrow$ Rate Distortion

- Encoder inputs / Decoder outputs: $\leftrightarrow$ Decoder inputs / Encoder outputs:
  $M_1 \in \{1, 2, ..., 2^{nR_1}\} \leftrightarrow T_1 \in \{1, 2, ..., 2^{nR_1}\}$
  $M_2 \in \{1, 2, ..., 2^{nR_2}\} \leftrightarrow T_2 \in \{1, 2, ..., 2^{nR_2}\}$
- Encoder outputs / Channel input: $\leftrightarrow$ Decoder output / Source reconstruction:
  $X_1^n, X_2^n \leftrightarrow \hat{X}_1^n, \hat{X}_2^n$

Duality Transformation Principles

Channel Coding $\leftrightarrow$ Rate Distortion

- Encoder inputs / Decoder outputs: $\leftrightarrow$ Decoder inputs / Encoder outputs:
  $M_1 \in \{1, 2, ..., 2^{nR_1}\} \leftrightarrow T_1 \in \{1, 2, ..., 2^{nR_1}\}$
  $M_2 \in \{1, 2, ..., 2^{nR_2}\} \leftrightarrow T_2 \in \{1, 2, ..., 2^{nR_2}\}$
- Encoder outputs / Channel input: $\leftrightarrow$ Decoder output / Source reconstruction:
  $X_1^n, X_2^n \leftrightarrow \hat{X}_1^n, \hat{X}_2^n$
- Decoder input / Channel output: $\leftrightarrow$ Encoder input / Source:
  $Y^n \leftrightarrow X^n$

Duality Transformation Principles

Channel Coding $\leftrightarrow$ Rate Distortion

- Encoder inputs / Decoder outputs: $\leftrightarrow$ Decoder inputs / Encoder outputs:
  $M_1 \in \{1,2,...,2^{nR_1}\} \leftrightarrow T_1 \in \{1,2,...,2^{nR_1}\}$
  $M_2 \in \{1,2,...,2^{nR_2}\} \leftrightarrow T_2 \in \{1,2,...,2^{nR_2}\}$
- Encoder outputs / Channel input: $\leftrightarrow$ Decoder output / Source reconstruction:
  $X_1^n, X_2^n \leftrightarrow \hat{X}_1^n, \hat{X}_2^n$
- Decoder input / Channel output: $\leftrightarrow$ Encoder input / Source:
  $Y^n \leftrightarrow X^n$
- Encoder functions: $\leftrightarrow$ Decoder functions:
  $f_1 : \{1,2,...,2^{nR_1}\} \rightarrow \mathcal{X}_1^n \leftrightarrow g_1 : \{1,2,...,2^{nR_1}\} \rightarrow \hat{\mathcal{X}}_1^n$
  $f_2 : \{1,2,...,2^{nR_1}\} \times \{1,2,...,2^{nR_1}\} \times \mathcal{S}^n \rightarrow \mathcal{X}_2^n \leftrightarrow$
  $g_2 : \{1,2,...,2^{nR_1}\} \times \{1,2,...,2^{nR_1}\} \times \mathcal{S}^n \rightarrow \hat{\mathcal{X}}_2^n$

Duality Transformation Principles

Channel Coding $\leftrightarrow$ Rate Distortion

- Encoder inputs / Decoder outputs: $\leftrightarrow$ Decoder inputs / Encoder outputs:
  $M_1 \in \{1, 2, ..., 2^{nR_1}\} \leftrightarrow T_1 \in \{1, 2, ..., 2^{nR_1}\}$
  $M_2 \in \{1, 2, ..., 2^{nR_2}\} \leftrightarrow T_2 \in \{1, 2, ..., 2^{nR_2}\}$
- Encoder outputs / Channel input: $\leftrightarrow$ Decoder output / Source reconstruction:
  $X_1^n, X_2^n \leftrightarrow \hat{X}_1^n, \hat{X}_2^n$
- Decoder input / Channel output: $\leftrightarrow$ Encoder input / Source:
  $Y^n \leftrightarrow X^n$
- Encoder functions: $\leftrightarrow$ Decoder functions:
  $f_1 : \{1, 2, ..., 2^{nR_1}\} \rightarrow \mathcal{X}_1^n \leftrightarrow g_1 : \{1, 2, ..., 2^{nR_1}\} \rightarrow \hat{\mathcal{X}}_1^n$
  $f_2 : \{1, 2, ..., 2^{nR_1}\} \times \{1, 2, ..., 2^{nR_1}\} \times \mathcal{S}^n \rightarrow \mathcal{X}_2^n \leftrightarrow$
  $g_2 : \{1, 2, ..., 2^{nR_1}\} \times \{1, 2, ..., 2^{nR_1}\} \times \mathcal{S}^n \rightarrow \hat{\mathcal{X}}_2^n$
- Action encoder: $\leftrightarrow$ Action strategy:
  $f_A : \mathcal{M}_1 \times \mathcal{M}_2 \rightarrow \mathcal{A}^n \leftrightarrow f_A : \mathcal{T}_1 \times \mathcal{T}_2 \rightarrow \mathcal{A}^n$

Duality Transformation Principles

Channel Coding $\leftrightarrow$ Rate Distortion

- Encoder inputs / Decoder outputs: $\leftrightarrow$ Decoder inputs / Encoder outputs:
  $M_1 \in \{1, 2, ..., 2^{nR_1}\} \leftrightarrow T_1 \in \{1, 2, ..., 2^{nR_1}\}$
  $M_2 \in \{1, 2, ..., 2^{nR_2}\} \leftrightarrow T_2 \in \{1, 2, ..., 2^{nR_2}\}$
- Encoder outputs / Channel input: $\leftrightarrow$ Decoder output / Source reconstruction:
  $X_1^n, X_2^n \leftrightarrow \hat{X}_1^n, \hat{X}_2^n$
- Decoder input / Channel output: $\leftrightarrow$ Encoder input / Source:
  $Y^n \leftrightarrow X^n$
- Encoder functions: $\leftrightarrow$ Decoder functions:
  $f_1 : \{1, 2, ..., 2^{nR_1}\} \rightarrow \mathcal{X}_1^n \leftrightarrow g_1 : \{1, 2, ..., 2^{nR_1}\} \rightarrow \hat{\mathcal{X}}_1^n$
  $f_2 : \{1, 2, ..., 2^{nR_1}\} \times \{1, 2, ..., 2^{nR_1}\} \times \mathcal{S}^n \rightarrow \mathcal{X}_2^n \leftrightarrow$
  $g_2 : \{1, 2, ..., 2^{nR_1}\} \times \{1, 2, ..., 2^{nR_1}\} \times \mathcal{S}^n \rightarrow \hat{\mathcal{X}}_2^n$
- Action encoder: $\leftrightarrow$ Action strategy:
  $f_A : \mathcal{M}_1 \times \mathcal{M}_2 \rightarrow \mathcal{A}^n \leftrightarrow f_A : \mathcal{T}_1 \times \mathcal{T}_2 \rightarrow \mathcal{A}^n$
- U auxiliary random variable $\leftrightarrow$ U auxiliary random variable

Duality Transformation Principles

Channel Coding $\leftrightarrow$ Rate Distortion

- Encoder inputs / Decoder outputs: $\leftrightarrow$ Decoder inputs / Encoder outputs:
  $M_1 \in \{1, 2, ..., 2^{nR_1}\} \leftrightarrow T_1 \in \{1, 2, ..., 2^{nR_1}\}$
  $M_2 \in \{1, 2, ..., 2^{nR_2}\} \leftrightarrow T_2 \in \{1, 2, ..., 2^{nR_2}\}$
- Encoder outputs / Channel input: $\leftrightarrow$ Decoder output / Source reconstruction:
  $X_1^n, X_2^n \leftrightarrow \hat{X}_1^n, \hat{X}_2^n$
- Decoder input / Channel output: $\leftrightarrow$ Encoder input / Source:
  $Y^n \leftrightarrow X^n$
- Encoder functions: $\leftrightarrow$ Decoder functions:
  $f_1 : \{1, 2, ..., 2^{nR_1}\} \rightarrow \mathcal{X}_1^n \leftrightarrow g_1 : \{1, 2, ..., 2^{nR_1}\} \rightarrow \hat{\mathcal{X}}_1^n$
  $f_2 : \{1, 2, ..., 2^{nR_1}\} \times \{1, 2, ..., 2^{nR_1}\} \times \mathcal{S}^n \rightarrow \mathcal{X}_2^n \leftrightarrow$
  $g_2 : \{1, 2, ..., 2^{nR_1}\} \times \{1, 2, ..., 2^{nR_1}\} \times \mathcal{S}^n \rightarrow \hat{\mathcal{X}}_2^n$
- Action encoder: $\leftrightarrow$ Action strategy:
  $f_A : \mathcal{M}_1 \times \mathcal{M}_2 \rightarrow \mathcal{A}^n \leftrightarrow f_A : \mathcal{T}_1 \times \mathcal{T}_2 \rightarrow \mathcal{A}^n$
- U auxiliary random variable $\leftrightarrow$ U auxiliary random variable
- S state information $\leftrightarrow$ S side information

Duality Transformation Principles

Channel Coding $\leftrightarrow$ Rate Distortion

- Encoder inputs / Decoder outputs: $\leftrightarrow$ Decoder inputs / Encoder outputs:
  $M_1 \in \{1, 2, ..., 2^{nR_1}\} \leftrightarrow T_1 \in \{1, 2, ..., 2^{nR_1}\}$
  $M_2 \in \{1, 2, ..., 2^{nR_2}\} \leftrightarrow T_2 \in \{1, 2, ..., 2^{nR_2}\}$
- Encoder outputs / Channel input: $\leftrightarrow$ Decoder output / Source reconstruction:
  $X_1^n, X_2^n \leftrightarrow \hat{X}_1^n, \hat{X}_2^n$
- Decoder input / Channel output: $\leftrightarrow$ Encoder input / Source:
  $Y^n \leftrightarrow X^n$
- Encoder functions: $\leftrightarrow$ Decoder functions:
  $f_1 : \{1, 2, ..., 2^{nR_1}\} \to \mathcal{X}_1^n \leftrightarrow g_1 : \{1, 2, ..., 2^{nR_1}\} \to \hat{\mathcal{X}}_1^n$
  $f_2 : \{1, 2, ..., 2^{nR_1}\} \times \{1, 2, ..., 2^{nR_1}\} \times \mathcal{S}^n \to \mathcal{X}_2^n \leftrightarrow$
  $g_2 : \{1, 2, ..., 2^{nR_1}\} \times \{1, 2, ..., 2^{nR_1}\} \times \mathcal{S}^n \to \hat{\mathcal{X}}_2^n$
- Action encoder: $\leftrightarrow$ Action strategy:
  $f_A : \mathcal{M}_1 \times \mathcal{M}_2 \to \mathcal{A}^n \leftrightarrow f_A : \mathcal{T}_1 \times \mathcal{T}_2 \to \mathcal{A}^n$
- U auxiliary random variable $\leftrightarrow$ U auxiliary random variable
- S state information $\leftrightarrow$ S side information

Duality Transformation Principles

Channel Coding $\leftrightarrow$ Rate Distortion

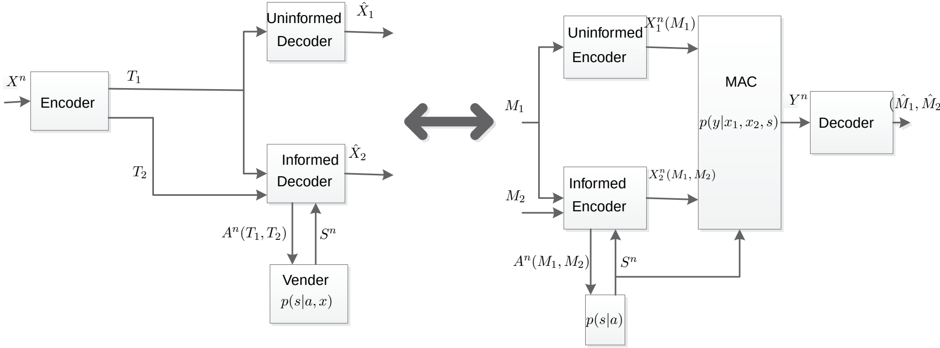
- Encoder inputs / Decoder outputs: $\leftrightarrow$ Decoder inputs / Encoder outputs:
  $M_1 \in \{1, 2, ..., 2^{nR_1}\} \leftrightarrow T_1 \in \{1, 2, ..., 2^{nR_1}\}$
  $M_2 \in \{1, 2, ..., 2^{nR_2}\} \leftrightarrow T_2 \in \{1, 2, ..., 2^{nR_2}\}$
- Encoder outputs / Channel input: $\leftrightarrow$ Decoder output / Source reconstruction:
  $X_1^n, X_2^n \leftrightarrow \hat{X}_1^n, \hat{X}_2^n$
- Decoder input / Channel output: $\leftrightarrow$ Encoder input / Source:
  $Y^n \leftrightarrow X^n$
- Encoder functions: $\leftrightarrow$ Decoder functions:
  $f_1 : \{1, 2, ..., 2^{nR_1}\} \to \mathcal{X}_1^n \leftrightarrow g_1 : \{1, 2, ..., 2^{nR_1}\} \to \hat{\mathcal{X}}_1^n$
  $f_2 : \{1, 2, ..., 2^{nR_1}\} \times \{1, 2, ..., 2^{nR_1}\} \times \mathcal{S}^n \to \mathcal{X}_2^n \leftrightarrow$
  $g_2 : \{1, 2, ..., 2^{nR_1}\} \times \{1, 2, ..., 2^{nR_1}\} \times \mathcal{S}^n \to \hat{\mathcal{X}}_2^n$
- Action encoder: $\leftrightarrow$ Action strategy:
  $f_A : \mathcal{M}_1 \times \mathcal{M}_2 \to \mathcal{A}^n \leftrightarrow f_A : \mathcal{T}_1 \times \mathcal{T}_2 \to \mathcal{A}^n$
- U auxiliary random variable $\leftrightarrow$ U auxiliary random variable
- S state information $\leftrightarrow$ S side information
- Markov $S - A - X_1 \leftrightarrow$ Markov $S - (A, X) - U, \hat{X}_1$

Duality Transformation Principles

"Successive Refinement with Actions"

MAC with action-dependent state Information

at One Encoder



Duality- Corner Points

Duality- Corner Points

The best way to ”see” the duality relationship is to consider the corner points for the rate regions:

Duality- Corner Points

The best way to ”see” the duality relationship is to consider the corner points for the rate regions:

- Recall the capacity region of the Action-MAC

$$\begin{aligned} R_2 &\leq I(A;Y|X_1) + I(Y;U|A,X_1) - I(S;U|A,X_1) \\ R_1 + R_2 &\leq I(X_1;Y) + I(A;Y|X_1) + I(Y;U|A,X_1) - I(S;U|A,X_1). \end{aligned} \quad (2)$$

Duality- Corner Points

The best way to ”see” the duality relationship is to consider the corner points for the rate regions:

- Recall the capacity region of the Action-MAC

$$\begin{aligned} R_2 &\leq I(A;Y|X_1) + I(Y;U|A,X_1) - I(S;U|A,X_1) \\ R_1 + R_2 &\leq I(X_1;Y) + I(A;Y|X_1) + I(Y;U|A,X_1) - I(S;U|A,X_1). \end{aligned} \quad (2)$$

- The corner points for this region are:

$$\Big(I(X_1;Y) + I(A;Y|X_1) + I(Y;U|A,X_1) - I(S;U|A,X_1) \quad , \quad 0 \Big)$$

$$\Big(I(X_1;Y) \quad , \quad I(A;Y|X_1) + I(U;Y|X_1,A) - I(U;S|X_1,A) \Big)$$

Duality- Corner Points

The best way to "see" the duality relationship is to consider the corner points for the rate regions:

- Recall the capacity region of the Action-MAC

$$\begin{aligned} R_2 &\leq I(A;Y|X_1) + I(Y;U|A,X_1) - I(S;U|A,X_1) \\ R_1 + R_2 &\leq I(X_1;Y) + I(A;Y|X_1) + I(Y;U|A,X_1) - I(S;U|A,X_1). \end{aligned} \quad (2)$$

- The corner points for this region are:

$$\Big(I(X_1;Y) + I(A;Y|X_1) + I(Y;U|A,X_1) - I(S;U|A,X_1) \quad , \quad 0\Big)$$

$$\Big(I(X_1;Y) \quad , \quad I(A;Y|X_1) + I(U;Y|X_1,A) - I(U;S|X_1,A)\Big)$$

- Recall the rate region for the "Successive Refinement with Actions"

$$\begin{aligned} R_1 &\geq I(X;\hat{X}_1) \\ R_1 + R_2 &\geq I(X;\hat{X}_1) + I(A;X|\hat{X}_1) + I(X;U|A,\hat{X}_1) - I(S;U|A,\hat{X}_1) \end{aligned} \quad (3)$$

Duality- Corner Points

The best way to ”see” the duality relationship is to consider the corner points for the rate regions:

- Recall the capacity region of the Action-MAC

$$\begin{aligned} R_2 &\leq I(A;Y|X_1) + I(Y;U|A,X_1) - I(S;U|A,X_1) \\ R_1 + R_2 &\leq I(X_1;Y) + I(A;Y|X_1) + I(Y;U|A,X_1) - I(S;U|A,X_1). \end{aligned} \quad (2)$$

- The corner points for this region are:

$$\Big(I(X_1;Y) + I(A;Y|X_1) + I(Y;U|A,X_1) - I(S;U|A,X_1) \quad , \quad 0 \Big)$$
$$\Big(I(X_1;Y) \quad , \quad I(A;Y|X_1) + I(U;Y|X_1,A) - I(U;S|X_1,A) \Big)$$

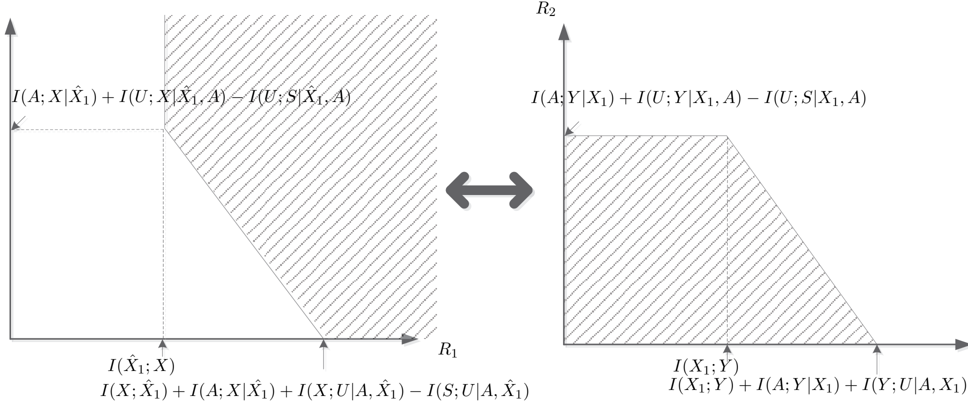
- Recall the rate region for the ”Successive Refinement with Actions”

$$\begin{aligned} R_1 &\geq I(X;\hat{X}_1) \\ R_1 + R_2 &\geq I(X;\hat{X}_1) + I(A;X|\hat{X}_1) + I(X;U|A,\hat{X}_1) - I(S;U|A,\hat{X}_1) \end{aligned} \quad (3)$$

- The corner points for this region are:

$$\Big(I(\hat{X}_1;X) + I(A;X|\hat{X}_1) + I(X;U|A,\hat{X}_1) - I(S;U|A,\hat{X}_1) \quad , \quad 0 \Big)$$
$$\Big(I(\hat{X}_1;X) \quad , \quad I(A;X|\hat{X}_1) + I(X;U|A,\hat{X}_1) - I(S;U|A,\hat{X}_1) \Big)$$

Duality Transformation Principles



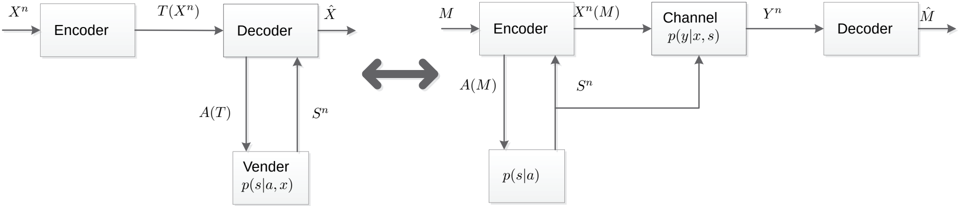
More Dualities

More Dualities

Duality between the action-dependent point-to-point channel and the source coding with side information ”Vending Machine” [Permuter & Weissman 11]

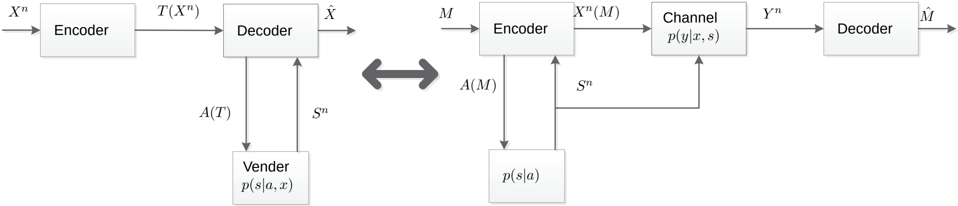
More Dualities

Duality between the action-dependent point-to-point channel and the source coding with side information ”Vending Machine” [Permuter & Weissman 11]



More Dualities

Duality between the action-dependent point-to-point channel and the source coding with side information "Vending Machine" [Permuter & Weissman 11]



$$R(D) = I(X;A) + I(X;U|A) - I(S;U|A) \qquad C = I(Y;A) + I(Y;U|A) - I(S;U|A)$$

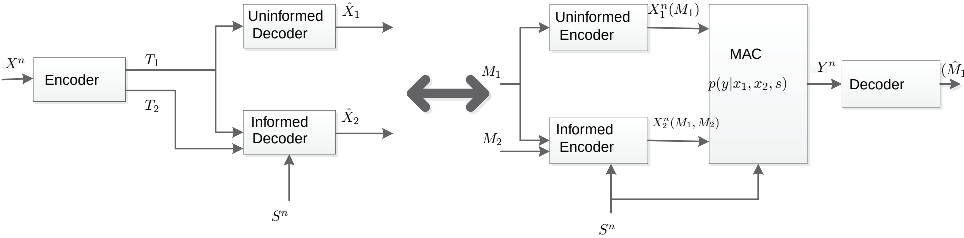
More Dualities

More Dualities

Duality between the GGP MAC and the Stienberg-Merhav rate distortion setting [Stienberg & Merhav 04]:

More Dualities

Duality between the GGP MAC and the Stienberg-Merhav rate distortion setting [Stienberg & Merhav 04]:



$$
\begin{aligned}
R_2 &\leq I(Y;U|X_1) - I(S;U|X_1) \\
R_1 + R_2 &\leq I(X_1;Y) + I(Y;U|X_1) - I(S;U|X_1).
\end{aligned} \quad (4)
$$

$$
\begin{aligned}
R_1 &\geq I(X;\hat{X}_1) \\
R_1 + R_2 &\geq I(X;\hat{X}_1) + I(X;U|\hat{X}_1) - I(S;U|\hat{X}_1).
\end{aligned} \quad (5)
$$

Summary

- We discussed state-dependent and action-dependent channel coding problems.

Summary

- We discussed state-dependent and action-dependent channel coding problems.
- We found the capacity region the action-dependent MAC.

Summary

- We discussed state-dependent and action-dependent channel coding problems.
- We found the capacity region the action-dependent MAC.
- We obtained the capacity of the Gaussian action-dependent MAC.

Summary

- We discussed state-dependent and action-dependent channel coding problems.
- We found the capacity region the action-dependent MAC.
- We obtained the capacity of the Gaussian action-dependent MAC.
- We found the capacity of the Gaussian p.t.p action-dependent channel, which was left open in [Wiessman10].

Summary

- We discussed state-dependent and action-dependent channel coding problems.
- We found the capacity region the action-dependent MAC.
- We obtained the capacity of the Gaussian action-dependent MAC.
- We found the capacity of the Gaussian p.t.p action-dependent channel, which was left open in [Wiessman10].
- We established rate distortion dualities of action dependent models.

Summary

- We discussed state-dependent and action-dependent channel coding problems.
- We found the capacity region the action-dependent MAC.
- We obtained the capacity of the Gaussian action-dependent MAC.
- We found the capacity of the Gaussian p.t.p action-dependent channel, which was left open in [Wiessman10].
- We established rate distortion dualities of action dependent models.

Thank you!

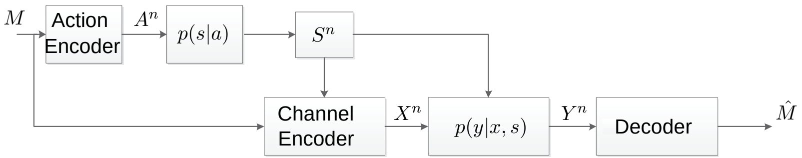
Gaussian Channel

Gaussian Channel

- To solve the Gaussian case for the MAC, we needed to first solve the case of the point-to-point channel, left open in [Wiessman10].

Gaussian Channel

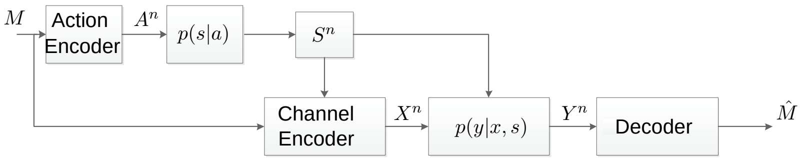
- To solve the Gaussian case for the MAC, we needed to first solve the case of the point-to-point channel, left open in [Wiessman10].



- The channel model:
  $Y^n = X^n(M,S^n)+S^n+Z^n = X^n(M,S^n)+A^n(M)+W^n+Z^n$
- $S^n = A^n(M) + W^n$.

Gaussian Channel

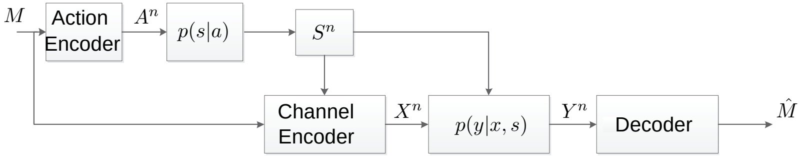
- To solve the Gaussian case for the MAC, we needed to first solve the case of the point-to-point channel, left open in [Wiessman10].



- The channel model:
  $Y^n = X^n(M,S^n)+S^n+Z^n = X^n(M,S^n)+A^n(M)+W^n+Z^n$
- $S^n = A^n(M) + W^n$.
- Z^n and W^n are independent, W^n is i.i.d.$\sim N(0,Q)$ and Z^n is i.i.d.$\sim N(0,N)$.

Gaussian Channel

- To solve the Gaussian case for the MAC, we needed to first solve the case of the point-to-point channel, left open in [Wiessman10].



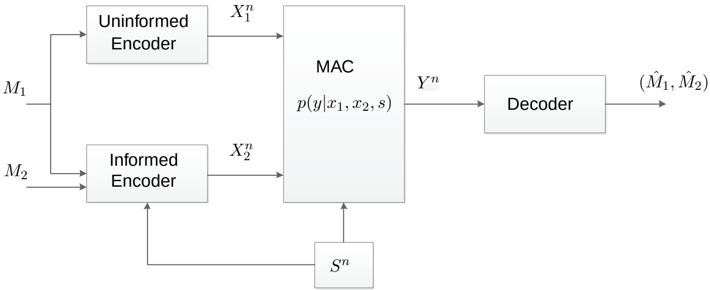
- The channel model:
  $Y^n = X^n(M,S^n)+S^n+Z^n = X^n(M,S^n)+A^n(M)+W^n+Z^n$
- $S^n = A^n(M) + W^n$.
- Z^n and W^n are independent, W^n is i.i.d.$\sim N(0,Q)$ and Z^n is i.i.d.$\sim N(0,N)$.
- We have the following power constraints:
  $\frac{1}{n}\sum_{i=1}^{n}(X_i)^2 \leq P_x$ and $\frac{1}{n}\sum_{i=1}^{n}(A_i)^2 \leq P_A$.

Gaussian Channel-Point-to-Point

- We look at the Gaussian MAC channel model (GGP channel) [Somekh-Baruch,Shamai & Verdú 07]

Gaussian Channel-Point-to-Point

- We look at the Gaussian MAC channel model (GGP channel) [Somekh-Baruch,Shamai & Verdú 07]



- The channel model is:
$Y^n = X_1(M_1)^n + X_2^n(M_1, M_2, W^n) + W^n + Z^n$.

Gaussian Channel-Point-to-Point

- The capacity of the GGP MAC [Somekh-Baruch,Shamai & Verdú 07]:

Gaussian Channel-Point-to-Point

- The capacity of the GGP MAC [Somekh-Baruch,Shamai & Verdú 07]:

$$\begin{aligned} R_2 &\leq \frac{1}{2}\log\left(1+\frac{P_2(1-\rho_{12}^2-\rho_{2S}^2)}{N}\right) \\ R_1+R_2 &\leq \frac{1}{2}\log\left(1+\frac{P_2(1-\rho_{12}^2-\rho_{2S}^2)}{N}\right) \\ &+ \frac{1}{2}\log\left(1+\frac{(\sqrt{P_1}+\sqrt{P_2})^2}{P_2(1-\rho_{12}^2-\rho_{2S}^2)+(\sigma_W+\rho_{2S}\sqrt{P_2})^2 N}\right), \end{aligned}$$

where

$$\rho_{12}=\frac{\sigma_{12}}{\sqrt{P_1P_2}},\quad \rho_{2W}=\frac{\sigma_{2W}}{\sqrt{P_2Q}}.$$

$$\rho_{12}^2+\rho_{2W}^2\leq 1.$$

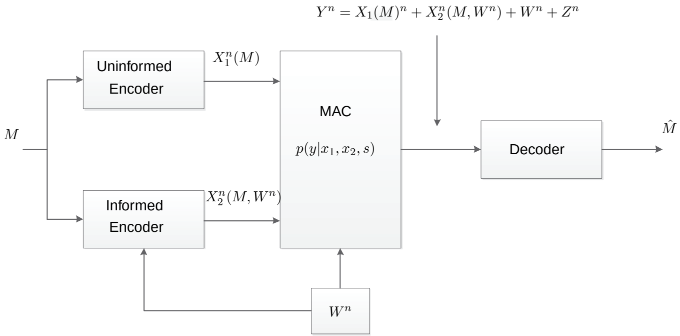
- How is this result relevant to the action-dependent Gaussian channel?

Gaussian Channel-Point-to-Point

- We found a one-to-one correspondence between the action-dependent Gaussain point-to-point channel and the GGP MAC.

Gaussian Channel-Point-to-Point

- We found a one-to-one correspondence between the action-dependent Gaussain point-to-point channel and the GGP MAC.
- This is done by looking at the GGP MAC with only a common message:



Gaussian Channel-Point-to-Point

- We can look at the block of ”Action Encoder” as the ”Uninformed Encoder” and the block of ”Channel Encoder” as the ”Informed Encoder”:

Action-dependent p-t-p channel	GGP channel with common message
A^n	X_1^n
X^n	X_2^n
$f_A : \mathcal{M} \rightarrow \mathcal{A}^n$	$f_{X_1} : \mathcal{M} \rightarrow \mathcal{X}_1^n$
$f_X : \mathcal{M} \times \mathcal{S}^n \rightarrow \mathcal{X}^n$	$f_{X_2} : \mathcal{M} \times \mathcal{S}^n \rightarrow \mathcal{X}_2^n$

Gaussian Channel-Point-to-Point

- We can look at the block of ”Action Encoder” as the ”Uninformed Encoder” and the block of ”Channel Encoder” as the ”Informed Encoder”:

Action-dependent p-t-p channel	GGP channel with common message
A^n	X_1^n
X^n	X_2^n
$f_A : \mathcal{M} \to \mathcal{A}^n$	$f_{X_1} : \mathcal{M} \to \mathcal{X}_1^n$
$f_X : \mathcal{M} \times \mathcal{S}^n \to \mathcal{X}^n$	$f_{X_2} : \mathcal{M} \times \mathcal{S}^n \to \mathcal{X}_2^n$

- Notice we don’t lose any of the properties of the settings.

Gaussian Channel-Point-to-Point

The capacity is achieved by substituting:

- $M_2 = 0$, thus $R_2 = 0$,
- $P_1 = P_A$,
- $P_2 = P_X$,
- $\rho_{12} = \rho_{XA}$ and $\rho_{2W} = \rho_{XW}$,

we have:

$$\begin{aligned} C &= \frac{1}{2}\log\left(1 + \frac{P_X(1-\rho_{XA}^2-\rho_{XW}^2)}{N}\right) \\ &+ \frac{1}{2}\log\left(1 + \frac{(\sqrt{P_A}+\rho_{XA}\sqrt{P_X})^2}{P_X(1-\rho_{XA}^2-\rho_{XW}^2)+(\sigma_W+\rho_{XW}\sqrt{P_X})^2+N}\right), \end{aligned}$$

such that

$$\rho_{XA}^2 + \rho_{XW}^2 \leq 1.$$

Similar results where obtained simultaneously and independently in [Choudhuri-Mitra,GLOBECOM'12].