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

\[
M \xrightarrow{X^n} \text{Encoder} \xrightarrow{p(y|x, s)} Y^n \xrightarrow{} \text{Decoder} \xrightarrow{\hat{M}}
\]

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].
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)
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]

\[ p(y|x, s) \]

Encoder

Decoder

\[ \hat{M} \]
Noncausal state information

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

![Diagram showing the communication process](image)

**Theorem**

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

**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

\[
\begin{array}{ccc|ccc|c}
S  & p(s) & X  & p(y|x, s) & Y & \\
0  & \frac{p}{2} & 0  & 0 & & \\
1  & \frac{p}{2} & 1  & 1 & 0 & stuck at 0 \\
2  & 1-p & 0  & 0 & 1 & stuck at 1 \\
\end{array}
\]
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

| $S$ | $p(s)$ | $X$ | $p(y|x, s)$ | $Y$ |
|-----|--------|-----|-------------|-----|
| 0   | $\frac{p}{2}$ | 0   | 0           | stuck at 0 |
| 1   | $\frac{p}{2}$ | 1   | 1           | stuck at 1 |
| 2   | $1 - p$     | 1   | 1           |             |

- 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

| S | p(s) | X | p(y|x, s) | Y | 
|---|------|---|----------|---|
| 0 | $p/2$ | 0 |          | 0 | stuck at 0
| 0 | $p/2$ | 1 |          | 1 | stuck at 1
| 2 | $1-p$ | 1 |          | 1 |

- 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/Laneman 07]
MAC with noncausal state information

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

\[
\text{MAC with Action-Dependent State}
\]

Uninformed Encoder

Informed Encoder

MAC

Decoder

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]

\[ C = \max \left[ I(U; Y) - I(U; S|A) \right] \]

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.
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 $p(s|a)$. 

Lior Dikstein, Haim Permuter and Shlomo (Shitz) Shamai
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**

![Diagram of a MAC with Action-Dependent State Information at One Encoder]

Lior Dikstein, Haim Permuter and Shlomo (Shitz) Shamai
Main Results
Main Results

Uninformed Encoder  \( X_1^n(M_1) \)

MAC  \( p(y|x_1, x_2, s) \)

Decoder  \((\hat{M}_1, \hat{M}_2)\)

Informed Encoder  \( X_2^n(M_1, M_2) \)

\( M_1 \)
\( M_2 \)

\( A^n(M_1, M_2) \)
\( S^n \)

\( p(s|a) \)

Lior Dikstein, Haim Permuter and Shlomo (Shitz) Shamai  MAC with Action-Dependent State
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
\[ |U| \leq |A||S||X_1||X_2| + 1. \]
Intuition

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

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

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

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

Notice that we can express the capacity region as:

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

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

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

Notice that we can express the capacity region as:

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

- The informed encoder transmits information using the action sequence $A$. 
Taking $\tilde{U} = (A, U)$, the following region is equivalent

\[
R_2 \leq I(A, U; Y | X_1) - I(U; S | X_1, A)
\]

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

Notice that we can express the capacity region as:

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

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

- The informed encoder transmits information using the action sequence $A$.
- This is used at the decoder to decode a second transmission, hence the conditioning.
Taking $\tilde{U} = (A, U)$, the following region is equivalent

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

Notice that we can express the capacity region as:

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

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

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

\[
R_1 + R_2 \leq I(X_1; Y) + I(A; Y|X_1) + I(U; Y|X_1, A) - I(U; S|X_1, A)
\]
Another presentation for the capacity region can be achieved by applying the chain rule and the Markov $X_1 - A - S$:

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

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

The corner points $(R_1, R_2)$:

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

$$\left( I(X_1; Y) \quad , \quad I(A; Y|X_1) + I(U; Y|X_1, A) - I(U; S|X_1, A) \right)$$
Corner Points

\[ I(A; Y|X_1) + I(U; Y|X_1, A) - I(U; S|X_1, A) \]

\[ I(X_1; Y) + I(X_1, U; Y|A) - I(X_1, U; S|A) \]
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.
Malfunction of the Action Encode.

We cannot choose an action that effects the formation of the states.

The Informed Encoder still knows the states noncausally.
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 noncausally.
- 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:

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

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


Lior Dikstein, Haim Permuter and Shlomo (Shitz) Shamai 
MAC with Action-Dependent State
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:

- 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 $(U^n, X_1^n, A^n, S^n) \in T_{\epsilon}^{(n)}(U, X_1, A, S)$. 
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:
\[
(U^n(\hat{M}_1, \hat{M}_2, k), X_1^n(\hat{M}_1), A^n(\hat{M}_1, \hat{M}_2), Y^n) \in \mathcal{T}_\epsilon^n(U, X_1, A, Y).
\]
Achievability Outline: Codebook Generation
Encoder 1 generates a codebook of $2^{nR_1}$ codewords
$\sim p(x_1)$. 
Encoder 1 generates a codebook of $2^{nR_1}$ codewords
\[ \sim p(x_1). \]
Encoder 2 generates $2^{n(R_1 + R_2)}$ action sequences $A^n(m_1, m_2) \sim p(a|x_1)$. 
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 $(m_1, m_2)$. 
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 $u^n(1), \ldots, u^n(2^{n\tilde{R}})$ according to $\sim p(u|a, x_1)$. 
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 $u^n(1), \ldots, u^n(2^{n\tilde{R}})$ 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

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

$2^{nR_2}$ Bins

$2^{n(R - R_1 - R_2)}$ codewords $u^n \sim p(u|a, x_1)$
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 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 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 have to show that for any \((2^n R_1, 2^n R_2, n)\) code with
\[ P_{\text{error}} \to 0 \text{ as } n \to \infty \] 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,
  \[
  \sum_{i=1}^n I(X_{i+1}^n; Y_i|Y_{i-1}^i) = \sum_{i=1}^n I(Y_{i-1}^i; X_i|X_{i+1}^n)
  \]
- We identify our auxiliary random variable,
  \[
  U_i = (X_{i-1}^{i-1}, X_{i+1}^n, S_{i+1}^n, Y_{i-1}^i, A^n, M_1, M_2).
  \]
- We use a time-sharing random variable \(Q\) uniformly distributed in \(\{1, 2, \ldots, n\}\).
Main Results

![Diagram of a MAC with Action-Dependent State]

**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

\[
|U| \leq |A||S||X_1||X_2| + 1.
\]

Lior Dikstein, Haim Permuter and Shlomo (Shitz) Shamai

MAC with Action-Dependent State
Gaussian Channel-Channel Model
Gaussian Channel-Channel Model

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

$$
Y_i = X_{1,i}(M_1) + X_{2,i}(M_1, M_2, S^n) + S_i + Z_i \\
= X_{1,i}(M_1) + X_{2,i}(M_1, M_2, S^n) + A_i(M_1, M_2) + W_i + Z_i
$$
Gaussian Channel-Channel Model

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

$$Y_i = X_{1,i}(M_1) + X_{2,i}(M_1, M_2, S^n) + S_i + Z_i$$

$$= X_{1,i}(M_1) + X_{2,i}(M_1, M_2, S^n) + A_i(M_1, M_2) + W_i + Z_i$$

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

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

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

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

- $Z^n$ and $W^n$ are independent, $W^n$ is i.i.d.$\sim N(0, Q)$ and $Z^n$ is i.i.d.$\sim N(0, N)$. 

Lior Dikstein, Haim Permuter and Shlomo (Shitz) Shamai
Gaussian Channel-Channel Model

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

$$Y_i = X_{1,i}(M_1) + X_{2,i}(M_1, M_2, S^n) + S_i + Z_i$$

$$= X_{1,i}(M_1) + X_{2,i}(M_1, M_2, S^n) + A_i(M_1, M_2) + W_i + Z_i$$

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

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

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

and

$$\frac{1}{n} \sum_{i=1}^{n} (A_i)^2 \leq P_A.$$
Theorem
\begin{align*}
R_2 & \leq \frac{1}{2} \log \frac{(N + P_2 + P_A + Q - P_2 \rho_{12}^2 - P_A \rho_{1A}^2 + 2\sqrt{P_2 P_A} \rho_{2A} - 2\sqrt{P_2 P_A} \rho_{12} \rho_{1A} + 2\sqrt{P_2 Q} \rho_{2W})}{N((\rho_{1A}^2 - 1)(N + Q + P_2 \rho_{2W}^2 + 2\sqrt{P_2 Q} \rho_{2W}) - P_2 \Delta)} \\
+ \frac{1}{2} \log (N(\rho_{1A}^2 - 1) - P_2 \Delta)
\end{align*}

\begin{align*}
R_1 + R_2 & \leq \frac{1}{2} \log \frac{(N + P_1 + P_2 + P_A + Q + 2\sqrt{P_1 P_2} \rho_{12} + 2\sqrt{P_1 P_A} \rho_{1A} + 2\sqrt{P_2 P_A} \rho_{2A} + 2\sqrt{P_2 Q} \rho_{2W})}{N((\rho_{1A}^2 - 1)(N + Q + P_2 \rho_{2W}^2 + 2\sqrt{P_2 Q} \rho_{2W}) - P_2 \Delta)} \\
+ \frac{1}{2} \log (N(\rho_{1A}^2 - 1) - P_2 \Delta)
\end{align*}

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

Rate Region

$P_A = 0$

$P_A = 1$

$P_A = 2$

$P_A = 3$

$P_A = 4$

$P_A = 5$
We state two lemmas that show that our region is upper-bounded by:

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

\[
R_1 + R_2 \leq I(U, X_1; Y) - I(U, X_1; S|A)
\]
\[
\leq I(A, X_1; Y) + h(X_2|X_1, A, W) - h(X_2 - \hat{X}_2^{\text{lin}}(X_1, A, W, X_2 + Z))
\]
We state two lemmas that show that our region is upper-bounded by:

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

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

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:

\[
R_2 \leq I(U; Y|X_1) - I(U; S|A, X_1) \\
\leq I(A; Y|X_1) + h(X_2|X_1, A, W) - h(X_2 - \hat{X}_2^{\text{lin}}(X_1, A, W, X_2 + Z)) \\
= \frac{1}{2} \log \left( \frac{\sigma_Y^2}{\sigma_W^2} \right)
\]

\[
R_1 + R_2 \leq I(U, X_1; Y) - I(U, X_1; S|A) \\
\leq I(A, X_1; Y) + h(X_2|X_1, A, W) - h(X_2 - \hat{X}_2^{\text{lin}}(X_1, A, W, X_2 + Z)) \\
= \frac{1}{2} \log \left( \frac{\sigma_Y^2}{\sigma_W^2} \right)
\]

- We show that it suffices to consider only jointly Gaussian random variables.
We state two lemmas that show that our region is upper-bounded by:

\[
R_2 \leq I(U; Y|X_1) - I(U; S|A, X_1) \\
\leq I(A; Y|X_1) + h(X_2|X_1, A, W) - h(X_2 - \hat{X}_2^{\text{lin}}(X_1, A, W, X_2 + Z)) \\
= \frac{1}{2} \log \left( \frac{\sigma_Y^2|X_1 \sigma_W^2|Y, X_1, A}{QN} \right)
\]

\[
R_1 + R_2 \leq I(U, X_1; Y) - I(U, X_1; S|A) \\
\leq I(A, X_1; Y) + h(X_2|X_1, A, W) - h(X_2 - \hat{X}_2^{\text{lin}}(X_1, A, W, X_2 + Z)) \\
= \frac{1}{2} \log \left( \frac{\sigma_Y^2 \sigma_W^2|Y, X_1, A}{QN} \right)
\]

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

Now we define \( E[X_1^2] \triangleq \sigma_X^2, E[X_2^2] \triangleq \sigma_X^2, E[A^2] \triangleq \sigma_A^2 \) and calculate the expression.
Proof Outline-Converse

\[ R_2 \leq I(U; Y|X_1) - I(U; S|A, X_1) \]
\[ \leq I(A; Y|X_1) + h(X_2|X_1, A, W) - h(X_2 - \hat{X}_2^{\text{lin}}(X_1, A, W, X_2 + Z)) \]
\[ = \frac{1}{2} \log \left( \frac{\sigma_Y^2 | X_1 \sigma_W^2 | Y, X_1, A}{QN} \right) \]
\[ = \frac{1}{2} \log \left( \frac{N + \sigma_X^2 + \sigma_A^2 + Q - \sigma_X^2 \rho_{12}^2 - \sigma_A^2 \rho_{1A}^2 - 2 \sqrt{\sigma_X^2 \sigma_A^2 \rho_{2A}^2} - 2 \sqrt{\sigma_X^2 \sigma_A^2 \rho_{12} \rho_{1A}} + 2 \sqrt{\sigma_X^2 Q \rho_{2W}}} {N \left( (\rho_{1A}^2 - 1)(N + Q + \sigma_X^2 \rho_{2W}^2 + 2 \sqrt{\sigma_X^2 Q \rho_{2W}}) - \sigma_X^2 \Delta \right)} \right) \]
\[ + \frac{1}{2} \log \left( N(\rho_{1A}^2 - 1) - \sigma_X^2 \Delta \right) \]

such that

\[ \sigma_X^2 \leq P_1 \quad \sigma_X^2 \leq P_2 \quad \sigma_A^2 \leq P_A. \]
Proof Outline-Converse

\[ R_1 + R_2 \leq I(U, X_1; Y) - I(U, X_1; S|A) \]
\[ \leq I(A, X_1; Y) + h(X_2|X_1, A, W) - h(X_2 - \hat{X}_2^{\text{lin}}(X_1, A, W, X_2 + Z)) \]
\[ = \frac{1}{2} \log \left( \frac{\sigma_Y^2 \sigma_W^2 | Y, X_1, A}{QN} \right) \]
\[ = \frac{1}{2} \log \left( N + \sigma_{X_1}^2 + \sigma_{X_2}^2 + \sigma_A^2 + Q + 2 \sqrt{\sigma_{X_1}^2 \sigma_{X_2}^2 \rho_{12}} + 2 \sqrt{\sigma_{X_1}^2 \sigma_A^2 \rho_{1A}} + 2 \sqrt{\sigma_{X_2}^2 \sigma_A^2 \rho_{2A}} + 2 \sqrt{\sigma_{X_2}^2 Q \rho_{2W}} \right) \]
\[ \times N \left( (\rho_{1A}^2 - 1)(N + Q + \sigma_{X_2}^2 \rho_{2W}^2 + 2 \sqrt{\sigma_{X_2}^2 Q \rho_{2W}}) - \sigma_{X_2}^2 \Delta \right) \]
\[ + \frac{1}{2} \log \left( N(\rho_{1A}^2 - 1) - \sigma_{X_2}^2 \Delta \right) \]

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

\[
\Lambda = \begin{pmatrix}
\sigma_{X_1}^2 & \sigma_{12} & \sigma_{1A} & 0 & 0 \\
\sigma_{12} & \sigma_{X_2}^2 & \sigma_{2A} & \sigma_{2W} & 0 \\
\sigma_{1A} & \sigma_{2A} & \sigma_A^2 & 0 & 0 \\
0 & \sigma_{2W} & 0 & Q & 0 \\
0 & 0 & 0 & 0 & N
\end{pmatrix},
\]

satisfies the nonnegative-definiteness condition

\[
\det (\Lambda) = \sigma_{1A}^2 \sigma_{2W}^2 N \sigma_{X_1}^2 \sigma_A^2 + 2\sigma_{12} \sigma_{1A} \sigma_{2A} N Q - \sigma_{2A}^2 N \sigma_{X_1}^2 Q - \sigma_{12}^2 N \sigma_A^2 Q + N \sigma_{X_1}^2 \sigma_{X_2}^2 \sigma_A^2
\]

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

\[
1 - \rho_{12}^2 - \rho_{1A}^2 - \rho_{2A}^2 - \rho_{2W}^2 + \rho_{1A}^2 \rho_{2W}^2 + 2\rho_{1A} \rho_{2A} \rho_{12} \geq 0
\]
We show that replacing $\sigma^2_{X_1}, \sigma^2_{X_2}, \sigma^2_A$ with $P_1, P_2$ and $P_A$ respectively, further increases the region.
We show that replacing $\sigma^2_{X_1}, \sigma^2_{X_2}, \sigma^2_{A}$ with $P_1, P_2$ and $P_A$ respectively, further increases the region.

Substituting $\sigma^2_{X_1}, \sigma^2_{X_2}, \sigma^2_{A}$ with $P_1, P_2$ and $P_A$, we obtain the capacity region of the theorem.
Proof Outline-Converse

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

- Substituting $\sigma_{X_1}^2$, $\sigma_{X_2}^2$, $\sigma_A^2$ 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 $\sigma^2_{X_1}, \sigma^2_{X_2}, \sigma^2_A$ with $P_1, P_2$ and $P_A$ respectively, further increases the region.

- Substituting $\sigma^2_{X_1}, \sigma^2_{X_2}, \sigma^2_A$ 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.
We choose specific distributions of our r.v.
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.

\[
U = X_1 + X_2 + \beta S \\
= X_1 + X_2 + \beta(A + W).
\]
Substituting $U = X_1 + X_2 + \beta(A + W)$ in the capacity region:

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

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

- The capacity for the Gaussian action-dependent point-to-point channel, was left open in [Wiessman10].
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.
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 $R_1 = 0$. 
The capacity for the Gaussian action-dependent point-to-point channel, was left open in [Wiessman10].

We find the capacity for 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.
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
An information-theoretic duality between our Action-MAC and the "Successive Refinement with Actions" [Chia, Asnani & Weissman 11].
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.
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.
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:
  - A rate distortion dual for the action dependent point-to-point channel.
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

\[ T_1(X^n) \in \{1, 2, ..., 2^{nR_1}\} \]

\[ T_2(X^n) \in \{1, 2, ..., 2^{nR_2}\} \]

Encoder

Uninformed Decoder

\( \hat{X}_1^n \)

Informed Decoder

\( \hat{X}_2^n \)

Vender

\[ p(s|a, x) \]

Lior Dikstein, Haim Permuter and Shlomo (Shitz) Shamai

MAC with Action-Dependent State
The "Successive Refinement with Actions" model

Theorem

\[ R_1 \geq I(X; \hat{X}_1) \]
\[ R_1 + R_2 \geq I(X; \hat{X}_1) + I(A; X|\hat{X}_1) + I(X; U|X, A, \hat{X}_1) \]

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
Channel Coding $\leftrightarrow$ Rate Distortion

Encoder inputs / Decoder outputs: $\leftrightarrow$ Decoder inputs / Encoder outputs:

$M_1 \in \{1, 2, \ldots, 2^{nR_1}\} \leftrightarrow T_1 \in \{1, 2, \ldots, 2^{nR_1}\}$

$M_2 \in \{1, 2, \ldots, 2^{nR_2}\} \leftrightarrow T_2 \in \{1, 2, \ldots, 2^{nR_2}\}$
Duality Transformation Principles

Channel Coding $\leftrightarrow$ Rate Distortion

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

- 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 ↔ Rate Distortion

- Encoder inputs / Decoder outputs: ↔ Decoder inputs / Encoder outputs:
  \( M_1 \in \{1, 2, \ldots, 2^{nR_1}\} \leftrightarrow T_1 \in \{1, 2, \ldots, 2^{nR_1}\} \)
  \( M_2 \in \{1, 2, \ldots, 2^{nR_2}\} \leftrightarrow T_2 \in \{1, 2, \ldots, 2^{nR_2}\} \)

- Encoder outputs / Channel input: ↔ Decoder output / Source reconstruction:
  \( X_1^n, X_2^n \leftrightarrow \hat{X}_1^n, \hat{X}_2^n \)

- Decoder input / Channel output: ↔ Encoder input / Source:
  \( Y^n \leftrightarrow X^n \)
Duality Transformation Principles

Channel Coding ↔ Rate Distortion

- Encoder inputs / Decoder outputs: ↔ Decoder inputs / Encoder outputs:
  \[ M_1 \in \{1, 2, \ldots, 2^{nR_1}\} \leftrightarrow T_1 \in \{1, 2, \ldots, 2^{nR_1}\} \]
  \[ M_2 \in \{1, 2, \ldots, 2^{nR_2}\} \leftrightarrow T_2 \in \{1, 2, \ldots, 2^{nR_2}\} \]

- Encoder outputs / Channel input: ↔ Decoder output / Source reconstruction:
  \[ X_1^n, X_2^n \leftrightarrow \hat{X}_1^n, \hat{X}_2^n \]

- Decoder input / Channel output: ↔ Encoder input / Source:
  \[ Y^n \leftrightarrow X^n \]

- Encoder functions: ↔ Decoder functions:
  \[ f_1 : \{1, 2, \ldots, 2^{nR_1}\} \to X_1^n \leftrightarrow g_1 : \{1, 2, \ldots, 2^{nR_1}\} \to \hat{X}_1^n \]
  \[ f_2 : \{1, 2, \ldots, 2^{nR_1}\} \times \{1, 2, \ldots, 2^{nR_1}\} \times S^n \to X_2^n \leftrightarrow \]
  \[ g_2 : \{1, 2, \ldots, 2^{nR_1}\} \times \{1, 2, \ldots, 2^{nR_1}\} \times S^n \to \hat{X}_2^n \]
Duality Transformation Principles

Channel Coding ↔ Rate Distortion

- Encoder inputs / Decoder outputs: ↔ Decoder inputs / Encoder outputs:
  \[ M_1 \in \{1, 2, \ldots, 2^{nR_1}\} \leftrightarrow T_1 \in \{1, 2, \ldots, 2^{nR_1}\} \]
  \[ M_2 \in \{1, 2, \ldots, 2^{nR_2}\} \leftrightarrow T_2 \in \{1, 2, \ldots, 2^{nR_2}\} \]

- Encoder outputs / Channel input: ↔ Decoder output / Source reconstruction:
  \[ X_1^n, X_2^n \leftrightarrow \hat{X}_1^n, \hat{X}_2^n \]

- Decoder input / Channel output: ↔ Encoder input / Source:
  \[ Y^n \leftrightarrow X^n \]

- Encoder functions: ↔ Decoder functions:
  \[ f_1 : \{1, 2, \ldots, 2^{nR_1}\} \to X_1^n \leftrightarrow g_1 : \{1, 2, \ldots, 2^{nR_1}\} \to \hat{X}_1^n \]
  \[ f_2 : \{1, 2, \ldots, 2^{nR_1}\} \times \{1, 2, \ldots, 2^{nR_1}\} \times S^n \to X_2^n \leftrightarrow \]
  \[ g_2 : \{1, 2, \ldots, 2^{nR_1}\} \times \{1, 2, \ldots, 2^{nR_1}\} \times S^n \to \hat{X}_2^n \]

- Action encoder: ↔ Action strategy:
  \[ f_A : \mathcal{M}_1 \times \mathcal{M}_2 \to \mathcal{A}^n \leftrightarrow f_A : \mathcal{T}_1 \times \mathcal{T}_2 \to \mathcal{A}^n \]
Duality Transformation Principles

Channel Coding ↔ Rate Distortion

- Encoder inputs / Decoder outputs: ↔ Decoder inputs / Encoder outputs:
  \[ M_1 \in \{1, 2, \ldots, 2^{nR_1} \} \leftrightarrow T_1 \in \{1, 2, \ldots, 2^{nR_1} \} \]
  \[ M_2 \in \{1, 2, \ldots, 2^{nR_2} \} \leftrightarrow T_2 \in \{1, 2, \ldots, 2^{nR_2} \} \]

- Encoder outputs / Channel input: ↔ Decoder output / Source reconstruction:
  \[ X_1^n, X_2^n \leftrightarrow \hat{X}_1^n, \hat{X}_2^n \]

- Decoder input / Channel output: ↔ Encoder input / Source:
  \[ Y^n \leftrightarrow X^n \]

- Encoder functions: ↔ Decoder functions:
  \[ f_1 : \{1, 2, \ldots, 2^{nR_1} \} \rightarrow X_1^n \leftrightarrow g_1 : \{1, 2, \ldots, 2^{nR_1} \} \rightarrow \hat{X}_1^n \]
  \[ f_2 : \{1, 2, \ldots, 2^{nR_1} \} \times \{1, 2, \ldots, 2^{nR_1} \} \times S^n \rightarrow X_2^n \leftrightarrow \]
  \[ g_2 : \{1, 2, \ldots, 2^{nR_1} \} \times \{1, 2, \ldots, 2^{nR_1} \} \times S^n \rightarrow \hat{X}_2^n \]

- Action encoder: ↔ Action strategy:
  \[ f_A : M_1 \times M_2 \rightarrow A^n \leftrightarrow f_A : T_1 \times T_2 \rightarrow A^n \]

- \( U \) auxiliary random variable ↔ \( U \) auxiliary random variable
Channel Coding ↔ Rate Distortion

- Encoder inputs / Decoder outputs: ↔ Decoder inputs / Encoder outputs:
  \[ M_1 \in \{1, 2, \ldots, 2^n R_1 \} \leftrightarrow T_1 \in \{1, 2, \ldots, 2^n R_1 \} \]
  \[ M_2 \in \{1, 2, \ldots, 2^n R_2 \} \leftrightarrow T_2 \in \{1, 2, \ldots, 2^n R_2 \} \]

- Encoder outputs / Channel input: ↔ Decoder output / Source reconstruction:
  \[ X_1^n, X_2^n \leftrightarrow \hat{X}_1^n, \hat{X}_2^n \]

- Decoder input / Channel output: ↔ Encoder input / Source:
  \[ Y^n \leftrightarrow X^n \]

- Encoder functions: ↔ Decoder functions:
  \[ f_1 : \{1, 2, \ldots, 2^n R_1 \} \rightarrow X_1^n \leftrightarrow g_1 : \{1, 2, \ldots, 2^n R_1 \} \rightarrow \hat{X}_1^n \]
  \[ f_2 : \{1, 2, \ldots, 2^n R_1 \} \times \{1, 2, \ldots, 2^n R_1 \} \times S^n \rightarrow X_2^n \leftrightarrow \]
  \[ g_2 : \{1, 2, \ldots, 2^n R_1 \} \times \{1, 2, \ldots, 2^n R_1 \} \times S^n \rightarrow \hat{X}_2^n \]

- Action encoder: ↔ Action strategy:
  \[ f_A : M_1 \times M_2 \rightarrow A^n \leftrightarrow f_A : T_1 \times T_2 \rightarrow A^n \]

- \( U \) auxiliary random variable ↔ \( U \) auxiliary random variable

- \( S \) state information ↔ \( S \) side information
Channel Coding ↔ Rate Distortion

- Encoder inputs / Decoder outputs: ↔ Decoder inputs / Encoder outputs:
  \[ M_1 \in \{1, 2, \ldots, 2^{nR_1}\} \leftrightarrow T_1 \in \{1, 2, \ldots, 2^{nR_1}\} \]
  \[ M_2 \in \{1, 2, \ldots, 2^{nR_2}\} \leftrightarrow T_2 \in \{1, 2, \ldots, 2^{nR_2}\} \]

- Encoder outputs / Channel input: ↔ Decoder output / Source reconstruction:
  \[ X_1^n, X_2^n \leftrightarrow \hat{X}_1^n, \hat{X}_2^n \]

- Decoder input / Channel output: ↔ Encoder input / Source:
  \[ Y^n \leftrightarrow X^n \]

- Encoder functions: ↔ Decoder functions:
  \[ f_1 : \{1, 2, \ldots, 2^{nR_1}\} \rightarrow X_1^n \leftrightarrow g_1 : \{1, 2, \ldots, 2^{nR_1}\} \rightarrow \hat{X}_1^n \]
  \[ f_2 : \{1, 2, \ldots, 2^{nR_1}\} \times \{1, 2, \ldots, 2^{nR_1}\} \times S^n \rightarrow X_2^n \leftrightarrow \]
  \[ g_2 : \{1, 2, \ldots, 2^{nR_1}\} \times \{1, 2, \ldots, 2^{nR_1}\} \times S^n \rightarrow \hat{X}_2^n \]

- Action encoder: ↔ 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 ↔ \( U \) auxiliary random variable

- \( S \) state information ↔ \( S \) side information
Duality Transformation Principles

Channel Coding ↔ Rate Distortion

- Encoder inputs / Decoder outputs: ↔ Decoder inputs / Encoder outputs:
  \[ M_1 \in \{1, 2, \ldots, 2^{nR_1}\} \leftrightarrow T_1 \in \{1, 2, \ldots, 2^{nR_1}\} \]
  \[ M_2 \in \{1, 2, \ldots, 2^{nR_2}\} \leftrightarrow T_2 \in \{1, 2, \ldots, 2^{nR_2}\} \]

- Encoder outputs / Channel input: ↔ Decoder output / Source reconstruction:
  \[ X_1^n, X_2^n \leftrightarrow \hat{X}_1^n, \hat{X}_2^n \]

- Decoder input / Channel output: ↔ Encoder input / Source:
  \[ Y^n \leftrightarrow X^n \]

- Encoder functions: ↔ Decoder functions:
  \[ f_1 : \{1, 2, \ldots, 2^{nR_1}\} \rightarrow X_1^n \leftrightarrow g_1 : \{1, 2, \ldots, 2^{nR_1}\} \rightarrow \hat{X}_1^n \]
  \[ f_2 : \{1, 2, \ldots, 2^{nR_1}\} \times \{1, 2, \ldots, 2^{nR_1}\} \times S^n \rightarrow X_2^n \leftrightarrow \hat{X}_2^n \]

- Action encoder: ↔ 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 ↔ \( U \) auxiliary random variable

- \( S \) state information ↔ \( S \) side information

- Markov \( S - A - X_1 \) ↔ Markov \( S - (A, X) - U, \hat{X}_1 \)
"Successive Refinement with Actions"

MAC with action-dependent state information at one encoder

Duality Transformation Principles

Lior Dikstein, Haim Permuter and Shlomo (Shitz) Shamai

MAC with Action-Dependent State
The best way to "see" the duality relationship is to consider the corner points for the rate regions:
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

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

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

\[
\begin{align*}
R_2 & \leq I(A; Y|X_1) + I(Y; U|A, X_1) - I(S; U|A, X_1) \\
R_1 + R_2 & \leq I(X_1; Y) + I(A; Y|X_1) + I(Y; U|A, X_1) - I(S; U|A, X_1). \tag{2}
\end{align*}
\]

- The corner points for this region are:

\[
\left( I(X_1; Y) + I(A; Y|X_1) + I(Y; U|A, X_1) - I(S; U|A, X_1) , 0 \right)

\left( I(X_1; Y) , I(A; Y|X_1) + I(U; Y|X_1, A) - I(U; S|X_1, A) \right)
\]
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

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

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

- The corner points for this region are:

\[
\left( I(X_1; Y) + I(A; Y|X_1) + I(Y; U|A, X_1) - I(S; U|A, X_1) , \ 0 \right)
\]

\[
\left( I(X_1; Y), \ I(A; Y|X_1) + I(U; Y|X_1, A) - I(U; S|X_1, A) \right)
\]

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

\[
R_1 \geq I(X; \hat{X}_1)
\]

\[
R_1 + R_2 \geq I(X; \hat{X}_1) + I(A; X|\hat{X}_1) + I(X; U|A, \hat{X}_1) - I(S; U|A, \hat{X}_1).
\]
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**

\[
\begin{align*}
R_2 \leq & \ I(A; Y|X_1) + I(Y; U|A, X_1) - I(S; U|A, X_1) \\
R_1 + R_2 \leq & \ I(X_1; Y) + I(A; Y|X_1) + I(Y; U|A, X_1) - I(S; U|A, X_1).
\end{align*}
\]

(2)

- **The corner points for this region are:**

\[
\begin{align*}
(I(X_1; Y), & \ I(A; Y|X_1) + I(U; Y|X_1, A) - I(U; S|X_1, A)) \\
(I(X_1; Y) + I(A; Y|X_1) + I(Y; U|A, X_1) - I(S; U|A, X_1)), & \ 0
\end{align*}
\]

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

\[
\begin{align*}
R_1 \geq & \ I(X; \hat{X}_1) \\
R_1 + R_2 \geq & \ I(X; \hat{X}_1) + I(A; X|\hat{X}_1) + I(X; U|A, \hat{X}_1) - I(S; U|A, \hat{X}_1).
\end{align*}
\]

(3)

- **The corner points for this region are:**

\[
\begin{align*}
(I(\hat{X}_1; X), & \ I(A; X|\hat{X}_1) + I(X; U|A, \hat{X}_1) - I(S; U|A, \hat{X}_1)) \\
(I(\hat{X}_1; X) + I(A; X|\hat{X}_1) + I(X; U|A, \hat{X}_1) - I(S; U|A, \hat{X}_1)), & \ 0
\end{align*}
\]
Duality Transformation Principles

\[ I(A; X|\hat{X}_1) + I(U; X|\hat{X}_1, A) - I(U; S|\hat{X}_1, A) \]

\[ I(A; Y|X_1) + I(U; Y|X_1, A) - I(U; S|X_1, A) \]

\[ I(X_1; Y) + I(A; Y|X_1) + I(Y; U|A, X_1) \]

Lior Dikstein, Haim Permuter and Shlomo (Shitz) Shamai
MAC with Action-Dependent State
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]

\[ R(D) = I(X; A) + I(X; U|A) - I(S; U|A) \]

\[ C = I(Y; A) + I(Y; U|A) - I(S; U|A) \]
More Dualities
More Dualities

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

Lior Dikstein, Haim Permuter and Shlomo (Shitz) Shamai
MAC with Action-Dependent State
More Dualities

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

\[
\begin{align*}
R_2 &\leq I(Y; U|X_1) - I(S; U|X_1) \\
R_1 + R_2 &\leq I(X_1; Y) + I(Y; U|X_1) - I(S; U|X_1).
\end{align*}
\]

(4)

\[
\begin{align*}
R_1 &\geq I(X; \hat{X}_1) \\
R_1 + R_2 &\geq I(X; \hat{X}_1) + I(X; U|\hat{X}_1) - I(S; U|\hat{X}_1).
\end{align*}
\]

(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 of the action-dependent MAC.
- We obtained the capacity of the Gaussian action-dependent MAC.
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 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

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].

\[ Y^n = X^n(M, S^n) + Z^n = X^n(M, S^n) + A^n(M) + W^n + Z^n \]

- The channel model:

\[ S^n = A^n(M) + W^n. \]
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:
\[ Y^n = X^n(M, S^n) + S^n + Z^n = X^n(M, S^n) + A^n(M) + W^n + Z^n \]
\[ S^n = A^n(M) + W^n. \]
\[ Z^n \text{ and } W^n \text{ are independent, } W^n \text{ is i.i.d.} \sim N(0, Q) \text{ and } Z^n \text{ is i.i.d.} \sim N(0, N). \]
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:

\[ Y^n = X^n(M, S^n) + S^n + Z^n = X^n(M, S^n) + A^n(M) + W^n + Z^n \]

\[ S^n = A^n(M) + W^n. \]

\[ Z^n \text{ and } W^n \text{ are independent, } W^n \text{ is i.i.d. } \sim N(0, Q) \text{ and } Z^n \text{ is i.i.d. } \sim N(0, N). \]

We have the following power constraints:

\[ \frac{1}{n} \sum_{i=1}^{n} (X_i)^2 \leq P_x \text{ and } \frac{1}{n} \sum_{i=1}^{n} (A_i)^2 \leq P_A. \]
We look at the Gaussian MAC channel model (GGP channel) [Somekh-Baruch, Shamai & Verdú 07]
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]:

\[
R_2 \leq \frac{1}{2} \log \left( 1 + \frac{P_2(1 - \rho_{12}^2 - \rho_{2S}^2)}{N} \right)
\]

\[
R_1 + R_2 \leq \frac{1}{2} \log \left( 1 + \frac{P_2(1 - \rho_{12}^2 - \rho_{2S}^2)}{N} \right) + \frac{1}{2} \log \left( 1 + \frac{(\sqrt{P_1} + \sqrt{P_2})^2}{P_2(1 - \rho_{12}^2 - \rho_{2S}^2) + (\sigma_W + \rho_{2S} \sqrt{P_2}^2 N)} \right),
\]

where

\[
\rho_{12} = \frac{\sigma_{12}}{\sqrt{P_1 P_2}}, \quad \rho_{2W} = \frac{\sigma_{2W}}{\sqrt{P_2 Q}}.
\]

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

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 Gaussian point-to-point channel and the GGP MAC.
Gaussian Channel-Point-to-Point

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

\[
Y^n = X_1(M)^n + X_2^n(M, W^n) + W^n + Z^n
\]
We can look at the block of "Action Encoder" as the "Uninformed Encoder" and the block of "Channel Encoder" as the "Informed Encoder":

<table>
<thead>
<tr>
<th>Action-dependent p-t-p channel</th>
<th>GGP channel with common message</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A^n$</td>
<td>$X^n_1$</td>
</tr>
<tr>
<td>$X^n$</td>
<td>$X^n_2$</td>
</tr>
<tr>
<td>$f_A : \mathcal{M} \rightarrow A^n$</td>
<td>$f_{X_1} : \mathcal{M} \rightarrow X^n_1$</td>
</tr>
<tr>
<td>$f_X : \mathcal{M} \times S^n \rightarrow X^n$</td>
<td>$f_{X_2} : \mathcal{M} \times S^n \rightarrow X^n_2$</td>
</tr>
</tbody>
</table>
We can look at the block of ”Action Encoder” as the ”Uninformed Encoder” and the block of ”Channel Encoder” as the ”Informed Encoder”:

<table>
<thead>
<tr>
<th>Action-independent p-t-p channel</th>
<th>GGP channel with common message</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A^n$</td>
<td>$X_1^n$</td>
</tr>
<tr>
<td>$X^n$</td>
<td>$X_2^n$</td>
</tr>
<tr>
<td>$f_A : M \to A^n$</td>
<td>$f_{X_1} : M \to X_1^n$</td>
</tr>
<tr>
<td>$f_X : M \times S^n \to X^n$</td>
<td>$f_{X_2} : M \times S^n \to X_2^n$</td>
</tr>
</tbody>
</table>

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:

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

such that

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