The Multiple Access Channel with Causal and Strictly Causal Side Information at the Encoders

Amos Lapidoth and Yossef Steinberg

Outline

Outline

- Problem Formulation: The MAC with strictly causal and causal common SI

Outline

- Problem Formulation: The MAC with strictly causal and causal common SI
- An achievable region for the strictly causal model

Outline

- Problem Formulation: The MAC with strictly causal and causal common SI
- An achievable region for the strictly causal model
- Example

Outline

- Problem Formulation: The MAC with strictly causal and causal common SI
- An achievable region for the strictly causal model
- Example
- An achievable region for the causal model

Outline

- Problem Formulation: The MAC with strictly causal and causal common SI
- An achievable region for the strictly causal model
- Example
- An achievable region for the causal model
- The naïve approach

Outline

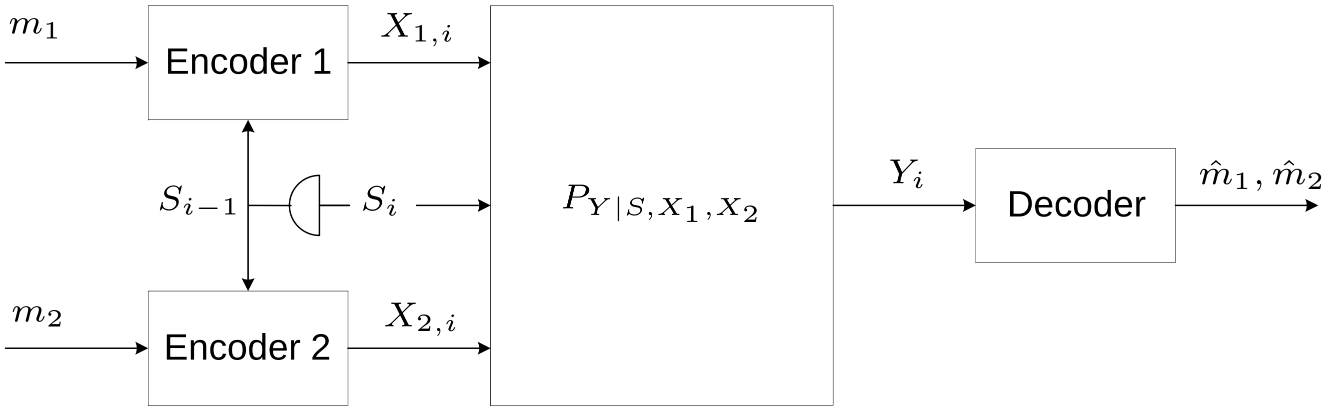
- Problem Formulation: The MAC with strictly causal and causal common SI
- An achievable region for the strictly causal model
- Example
- An achievable region for the causal model
- The naïve approach
- Example

Outline

- Problem Formulation: The MAC with strictly causal and causal common SI
- An achievable region for the strictly causal model
- Example
- An achievable region for the causal model
- The naïve approach
- Example
- Two independent states

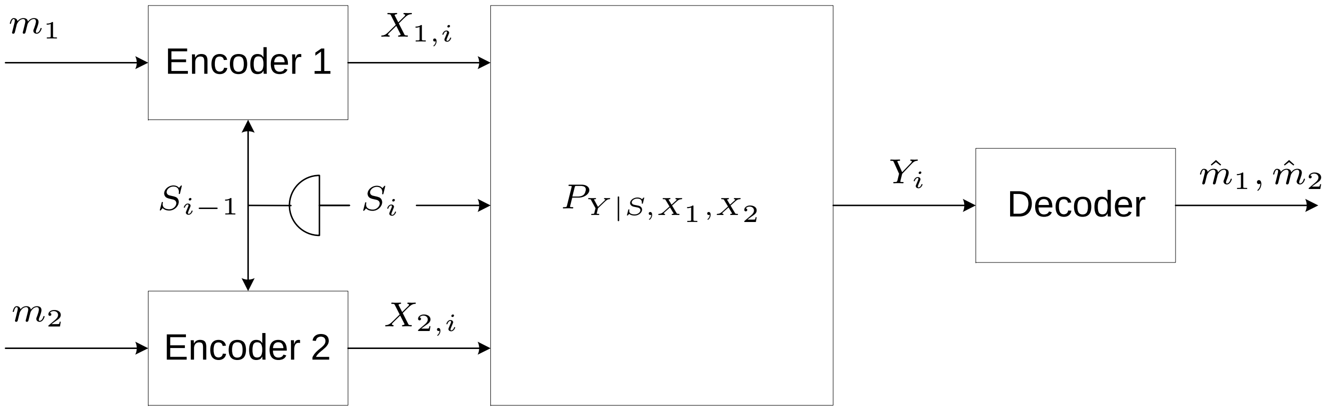
Problem Formulation

MAC with strictly causal side information (SI):



Problem Formulation

MAC with strictly causal side information (SI):

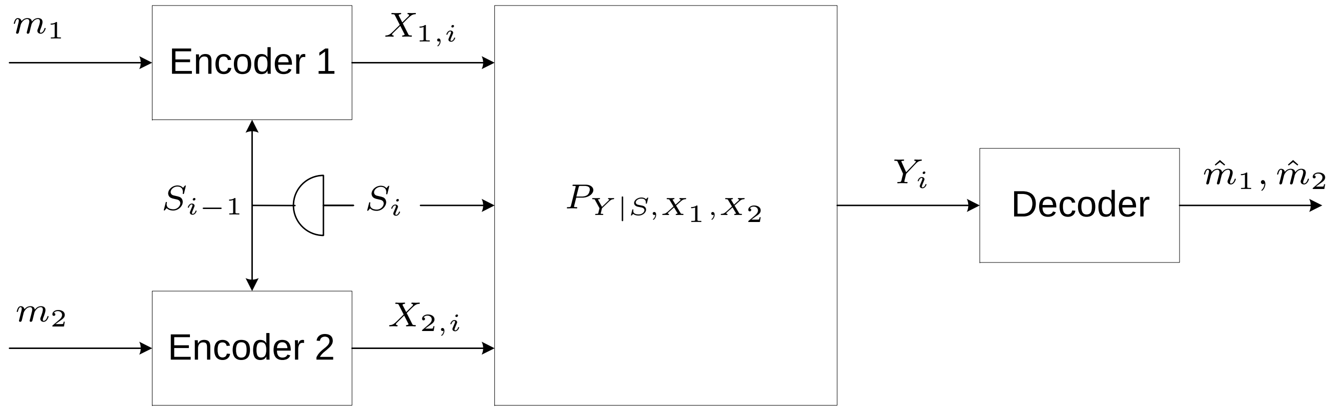


- One state sequence S^n, available to the encoders in a strictly causal manner:

$$X_{1,i} = f_{1,i}(m_1, S^{i-1}), \quad X_{2,i} = f_{2,i}(m_2, S^{i-1}), \quad i = 1, \ldots, n$$

Problem Formulation

MAC with strictly causal side information (SI):



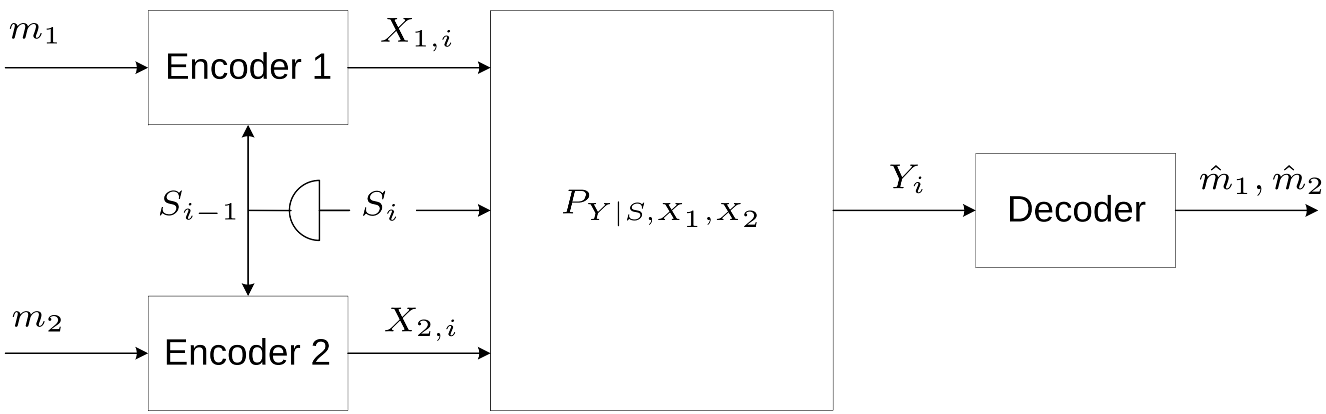
- One state sequence S^n, available to the encoders in a strictly causal manner:

$$X_{1,i} = f_{1,i}(m_1, S^{i-1}), \quad X_{2,i} = f_{2,i}(m_2, S^{i-1}), \quad i = 1, \ldots, n$$

$$(\hat{m}_1, \hat{m}_2) = g(Y^n)$$

Problem Formulation

MAC with strictly causal side information (SI):



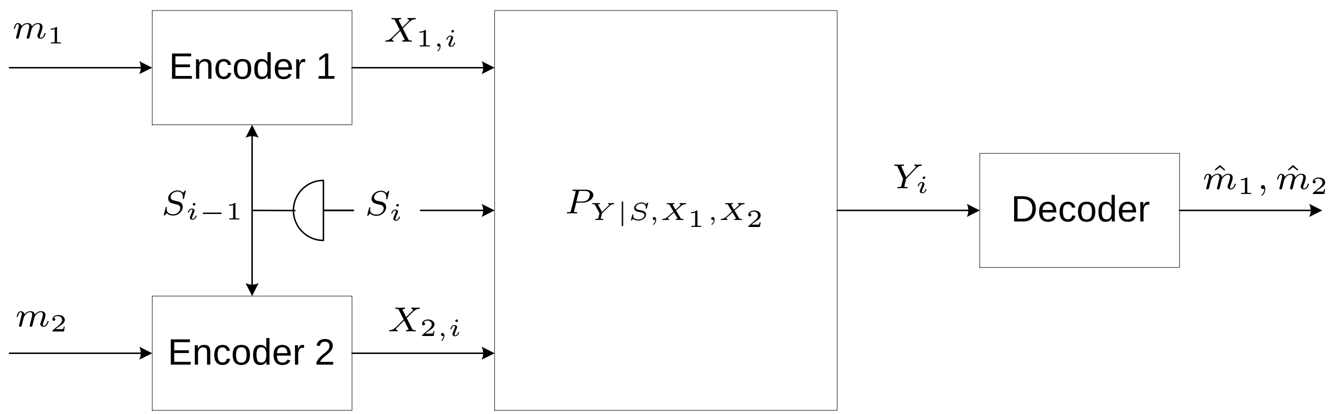
- One state sequence S^n, available to the encoders in a strictly causal manner:

$$X_{1,i} = f_{1,i}(m_1, S^{i-1}), \quad X_{2,i} = f_{2,i}(m_2, S^{i-1}), \quad i = 1, \ldots, n$$
$$(\hat{m}_1, \hat{m}_2) = g(Y^n)$$

- Transmission is subject to input constraints $\frac{1}{n}\sum_{i=1}^{n} \phi_k(X_{k,i}) \le \Gamma_k, \quad k = 1, 2.$

Problem Formulation

MAC with strictly causal side information (SI):



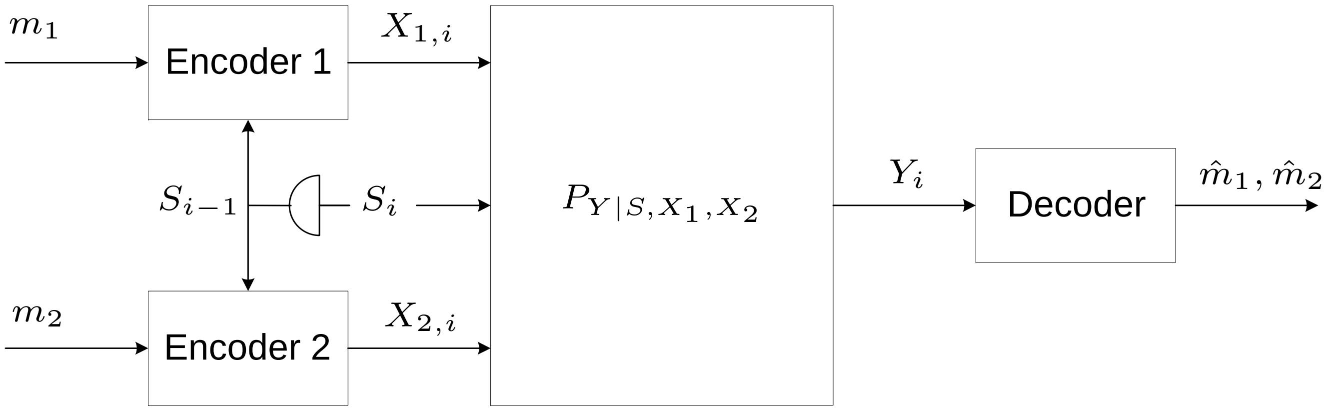
- One state sequence S^n, available to the encoders in a strictly causal manner:

$$X_{1,i} = f_{1,i}(m_1, S^{i-1}), \quad X_{2,i} = f_{2,i}(m_2, S^{i-1}), \quad i = 1, \ldots, n$$
$$(\hat{m}_1, \hat{m}_2) = g(Y^n)$$

- Transmission is subject to input constraints $\frac{1}{n}\sum_{i=1}^{n} \phi_k(X_{k,i}) \le \Gamma_k, \quad k = 1, 2.$
- Memoryless, time invariant channel and state $P_{Y|S,X_1,X_2}, P_S$.

Problem Formulation

MAC with strictly causal side information (SI):

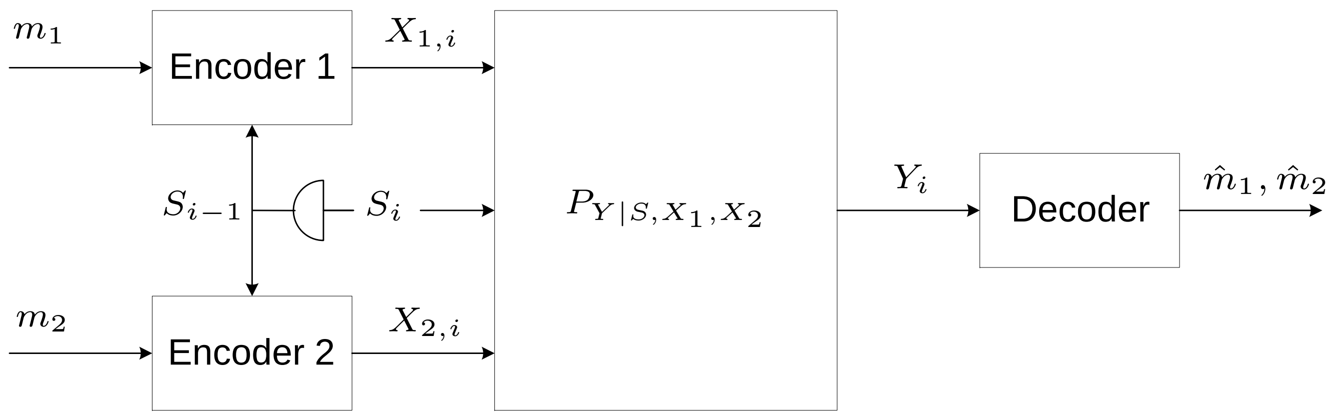


We are interested in $\mathcal{C}_{\text{s-c}}$, the region of all achievable rate and cost pairs

$$(R_1, R_2, \Gamma_1, \Gamma_2).$$

Problem Formulation

MAC with strictly causal side information (SI):



We are interested in $\mathcal{C}_{\text{s-c}}$, the region of all achievable rate and cost pairs

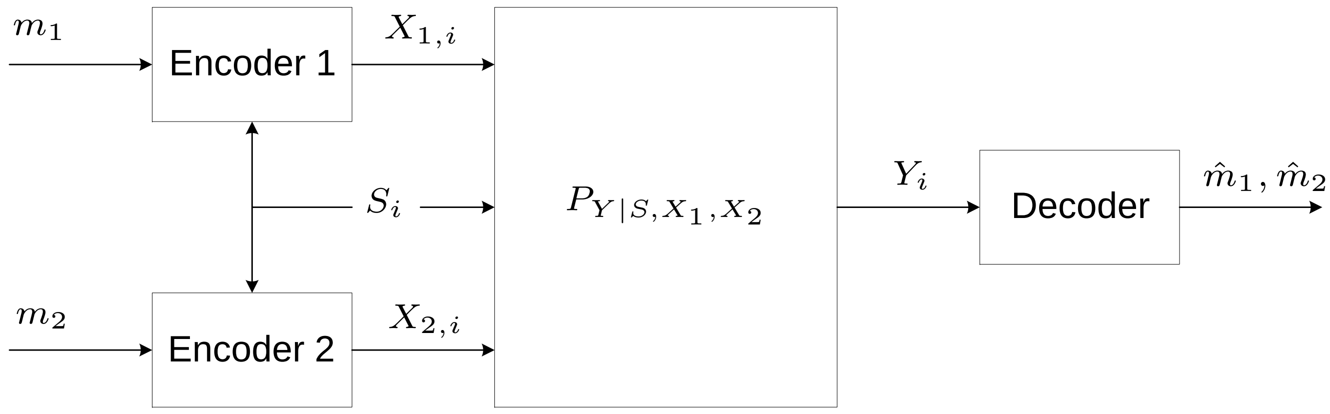
$$(R_1, R_2, \Gamma_1, \Gamma_2).$$

$\mathcal{C}_{\text{s-c}}(\Gamma_1, \Gamma_2)$ – the collection of all rate pairs (R_1, R_2) such that

$$(R_1, R_2, \Gamma_1, \Gamma_2) \in \mathcal{C}_{\text{s-c}}.$$

Problem Formulation

MAC with causal SI:



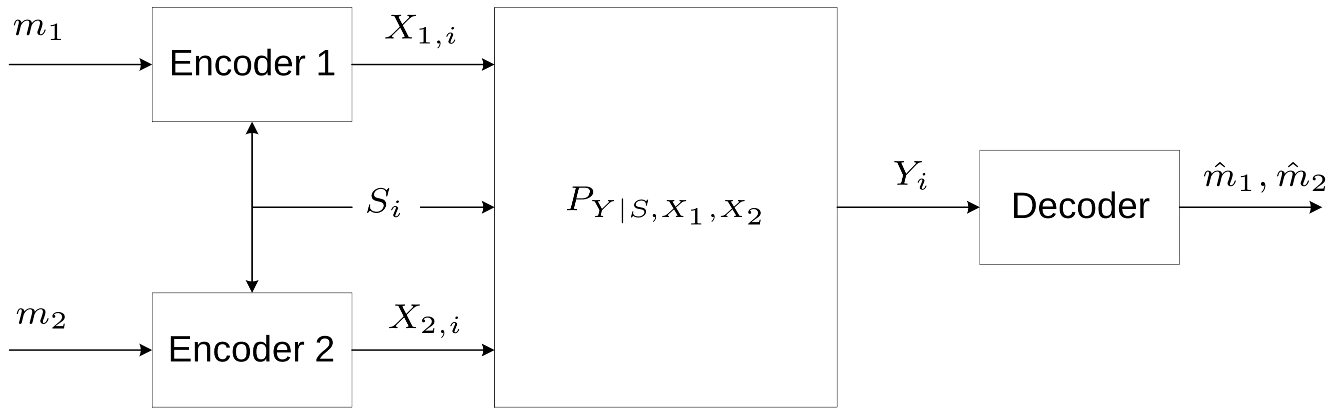
- One state sequence S^n, available to the encoders in a causal manner:

$$X_{1,i} = f_{1,i}(m_1, S^i), \quad X_{2,i} = f_{2,i}(m_2, S^i), \quad i = 1, \ldots, n$$
$$(\hat{m}_1, \hat{m}_2) = g(Y^n)$$

- Transmission is subject to input constraints $\frac{1}{n}\sum_{i=1}^{n} \phi_k(X_{k,i}) \leq \Gamma_k, \quad k = 1, 2.$
- Memoryless, time invariant channel and state $P_{Y|S,X_1,X_2}$, P_S.

Problem Formulation

MAC with causal SI:



We are interested in $\mathcal{C}_{\text{cau}}$, the region of all achievable rate and cost pairs

$$(R_1, R_2, \Gamma_1, \Gamma_2).$$

$\mathcal{C}_{\text{cau}}(\Gamma_1, \Gamma_2)$ – the collection of all rate pairs (R_1, R_2) such that

$$(R_1, R_2, \Gamma_1, \Gamma_2) \in \mathcal{C}_{\text{cau}}.$$

Single user and BC with SC SI

- Strictly causal SI does not increase the capacity of the single user channel

Single user and BC with SC SI

- Strictly causal SI does not increase the capacity of the single user channel

$$
\begin{aligned}
nR - n\epsilon_n &\leq I(M;Y^n) = \sum_{i=1}^{n} I(M;Y_i|Y^{i-1}) \\
&\leq \sum_{i=1}^{n} I(M,Y^{i-1};Y_i) \\
&\leq \sum_{i=1}^{n} I(M,Y^{i-1},X_i;Y_i) \\
&= \sum_{i=1}^{n} I(X_i;Y_i) \\
&\leq \max_{P_X} I(X;Y) = nC
\end{aligned}
$$

where C is the capacity without SI.

Single user and BC with SC SI

- Strictly causal SI does not increase the capacity of the single user channel (a reminiscent of the situation in feedback capacity)

Single user and BC with SC SI

- Strictly causal SI does not increase the capacity of the single user channel (a reminiscent of the situation in feedback capacity)
- Transmission of the state (or compressed version thereof) to the other side is sub optimal: waste of precious rate, without increase in capacity.

Single user and BC with SC SI

- An example by Dueck (1980): A non degraded additive noise BC with feedback.

  The noise is common to the two channels.

Single user and BC with SC SI

- An example by Dueck (1980): A non degraded additive noise BC with feedback.

  The noise is common to the two channels.

  - The encoder transmits the noise to the two users, uncompressed.

Single user and BC with SC SI

- An example by Dueck (1980): A non degraded additive noise BC with feedback. The noise is common to the two channels.
  - The encoder transmits the noise to the two users, uncompressed.
  - Knowledge of the additive noise at the decoder facilitates decoding of the messages.

Single user and BC with SC SI

- An example by Dueck (1980): A non degraded additive noise BC with feedback. The noise is common to the two channels.
  - The encoder transmits the noise to the two users, uncompressed.
  - Knowledge of the additive noise at the decoder facilitates decoding of the messages.
  - Although precious rate is spent on transmitting the noise, the net effect is an increase in the capacity region.

Single user and BC with SC SI

- An example by Dueck (1980): A non degraded additive noise BC with feedback. The noise is common to the two channels.
  - The encoder transmits the noise to the two users, uncompressed.
  - Knowledge of the additive noise at the decoder facilitates decoding of the messages.
  - Although precious rate is spent on transmitting the noise, the net effect is an increase in the capacity region.
  - Yields gains in capacity also when only lossy transmission of the noise is possible.

Single user and BC with SC SI

- An example by Dueck (1980): A non degraded additive noise BC with feedback. The noise is common to the two channels.
  - The encoder transmits the noise to the two users, uncompressed.
  - Knowledge of the additive noise at the decoder facilitates decoding of the messages.
  - Although precious rate is spent on transmitting the noise, the net effect is an increase in the capacity region.
  - Yields gains in capacity also when only lossy transmission of the noise is possible.
- In the MAC: The two users can *cooperate* in transmitting the noise (state) to the decoder.

MAC with SC SI

$\mathcal{R}_{\text{s-c}}$ - the CH of all $(R_1, R_2, \Gamma_1, \Gamma_2)$ satisfying

$$
\begin{aligned}
R_1 &\leq I(X_1;Y|X_2,U,V) \\
R_2 &\leq I(X_2;Y|X_1,U,V) \\
R_1 + R_2 &\leq I(X_1,X_2;Y|U,V) \\
R_1 + R_2 &\leq I(X_1,X_2,V;Y) - I(V;S) \\
\Gamma_k &\geq \mathsf{E}[\phi_k(X_k)], \qquad k = 1,2
\end{aligned}
$$

for some joint distribution

$$P_{U,V,X_1,X_2,S,Y} = P_S P_{X_1|U} P_{X_2|U} P_U P_{V|S} P_{Y|S,X_1,X_2}.$$

MAC with SC SI

$\mathcal{R}_{\text{s-c}}$ - the CH of all $(R_1, R_2, \Gamma_1, \Gamma_2)$ satisfying

$$\begin{aligned} R_1 &\leq I(X_1; Y | X_2, U, V) \\ R_2 &\leq I(X_2; Y | X_1, U, V) \\ R_1 + R_2 &\leq I(X_1, X_2; Y | U, V) \\ R_1 + R_2 &\leq I(X_1, X_2, V; Y) - I(V; S) \\ \Gamma_k &\geq \mathsf{E}[\phi_k(X_k)], \qquad k = 1, 2 \end{aligned}$$

for some joint distribution

$$P_{U,V,X_1,X_2,S,Y} = P_S P_{X_1|U} P_{X_2|U} P_U P_{V|S} P_{Y|S,X_1,X_2}.$$

$$X_1 - U - X_2$$
$$(X_1, U, X_2) \perp (V, S)$$

MAC with SC SI

$\mathcal{R}_{\text{s-c}}$ - the CH of all $(R_1, R_2, \Gamma_1, \Gamma_2)$ satisfying

$$\begin{aligned} R_1 &\leq I(X_1; Y|X_2, U, V) \\ R_2 &\leq I(X_2; Y|X_1, U, V) \\ R_1 + R_2 &\leq I(X_1, X_2; Y|U, V) \\ R_1 + R_2 &\leq I(X_1, X_2, V; Y) - I(V; S) \\ \Gamma_k &\geq \mathsf{E}[\phi_k(X_k)], \qquad k = 1, 2 \end{aligned}$$

for some joint distribution

$$P_{U,V,X_1,X_2,S,Y} = P_S P_{X_1|U} P_{X_2|U} P_U P_{V|S} P_{Y|S,X_1,X_2}.$$

$$X_1 - U - X_2$$

$$(X_1, U, X_2) \perp (V, S)$$

$$V - S - Y$$

Main result

$\mathcal{R}_{\text{s-c}}$ - the CH of all $(R_1, R_2, \Gamma_1, \Gamma_2)$ satisfying

$$\begin{aligned} R_1 &\leq I(X_1; Y | X_2, U, V) \\ R_2 &\leq I(X_2; Y | X_1, U, V) \\ R_1 + R_2 &\leq I(X_1, X_2; Y | U, V) \\ R_1 + R_2 &\leq I(X_1, X_2, V; Y) - I(V; S) \\ \Gamma_k &\geq \mathsf{E}[\phi_k(X_k)], \qquad k = 1, 2 \end{aligned}$$

for some joint distribution

$$P_{U,V,X_1,X_2,S,Y} = P_S P_{X_1|U} P_{X_2|U} P_U P_{V|S} P_{Y|S,X_1,X_2}.$$

Theorem 1 (Strictly-Causal SI) $\mathcal{R}_{\text{s-c}} \subseteq \mathcal{C}_{\text{s-c}}$

Main Result

We can write $\mathcal{R}_{\text{s-c}}$ as

$$\begin{aligned}
R_1 &\leq I(X_1;Y|X_2,U,V) \\
R_2 &\leq I(X_2;Y|X_1,U,V) \\
R_1+R_2 &\leq I(X_1,X_2;Y|U,V) \\
R_0+R_1+R_2 &\leq I(X_1,X_2;Y|V) \\
R_0 &\geq I(V;S)-I(V;Y). \\
\Gamma_k &\geq \mathsf{E}[\phi_k(X_k)], \qquad k=1,2
\end{aligned}$$

Main Result

We can write $\mathcal{R}_{\text{s-c}}$ as

$$\begin{aligned} R_1 &\leq I(X_1; Y | X_2, U, V) \\ R_2 &\leq I(X_2; Y | X_1, U, V) \\ R_1 + R_2 &\leq I(X_1, X_2; Y | U, V) \\ R_0 + R_1 + R_2 &\leq I(X_1, X_2; Y | V) \\ R_0 &\geq I(V; S) - I(V; Y). \\ \Gamma_k &\geq \mathsf{E}[\phi_k(X_k)], \qquad k = 1, 2 \end{aligned}$$

A block Markov scheme:

Main Result

We can write $\mathcal{R}_{\text{s-c}}$ as

$$
\begin{aligned}
R_1 &\leq I(X_1;Y|X_2,U,V) \\
R_2 &\leq I(X_2;Y|X_1,U,V) \\
R_1+R_2 &\leq I(X_1,X_2;Y|U,V) \\
R_0+R_1+R_2 &\leq I(X_1,X_2;Y|V) \\
R_0 &\geq I(V;S)-I(V;Y). \\
\Gamma_k &\geq \mathsf{E}[\phi_k(X_k)], \qquad k=1,2
\end{aligned}
$$

A block Markov scheme:

- The state sequence S^n is compressed by a Wyner-Ziv scheme, with coding random variable V, and decoder side information Y^n.

Main Result

We can write $\mathcal{R}_{\text{s-c}}$ as

$$\begin{aligned} R_1 &\leq I(X_1;Y|X_2,U,V) \\ R_2 &\leq I(X_2;Y|X_1,U,V) \\ R_1+R_2 &\leq I(X_1,X_2;Y|U,V) \\ R_0+R_1+R_2 &\leq I(X_1,X_2;Y|V) \\ R_0 &\geq I(V;S)-I(V;Y). \\ \Gamma_k &\geq \mathsf{E}[\phi_k(X_k)], \qquad k=1,2 \end{aligned}$$

A block Markov scheme:

- The state sequence S^n is compressed by a Wyner-Ziv scheme, with coding random variable V, and decoder side information Y^n.

$$V - S - Y$$

Main Result

We can write $\mathcal{R}_{\text{s-c}}$ as

$$
\begin{aligned}
R_1 &\leq I(X_1;Y|X_2,U,V) \\
R_2 &\leq I(X_2;Y|X_1,U,V) \\
R_1+R_2 &\leq I(X_1,X_2;Y|U,V) \\
R_0+R_1+R_2 &\leq I(X_1,X_2;Y|V) \\
R_0 &\geq I(V;S)-I(V;Y). \\
\Gamma_k &\geq \mathsf{E}[\phi_k(X_k)], \qquad k=1,2
\end{aligned}
$$

A block Markov scheme:

- The state sequence S^n is compressed by a Wyner-Ziv scheme, with coding random variable V, and decoder side information Y^n.

$$V - S - Y$$

- The compressed state is transmitted to the decoder in the *next transmission block* as a *common message*, together with the independent messages m_1, m_2.

Main Result

We can write $\mathcal{R}_{\text{s-c}}$ as

$$\begin{aligned} R_1 &\leq I(X_1;Y|X_2,U,V) \\ R_2 &\leq I(X_2;Y|X_1,U,V) \\ R_1+R_2 &\leq I(X_1,X_2;Y|U,V) \\ R_0+R_1+R_2 &\leq I(X_1,X_2;Y|V) \\ R_0 &\geq I(V;S)-I(V;Y). \\ \Gamma_k &\geq \mathsf{E}[\phi_k(X_k)], \qquad k=1,2 \end{aligned}$$

A block Markov scheme:

- The state sequence S^n is compressed by a Wyner-Ziv scheme, with coding random variable V, and decoder side information Y^n.

$$V - S - Y$$

- The compressed state is transmitted to the decoder in the *next transmission block* as a *common message*, together with the independent messages m_1, m_2.

$$X_1 - U - X_2, \quad \text{independent of} \quad (V,S).$$

Main Result

We can write $\mathcal{R}_{\text{s-c}}$ as

$$\begin{aligned} R_1 &\le I(X_1;Y|X_2,U,V) \\ R_2 &\le I(X_2;Y|X_1,U,V) \\ R_1+R_2 &\le I(X_1,X_2;Y|U,V) \\ R_0+R_1+R_2 &\le I(X_1,X_2;Y|V) \\ R_0 &\ge I(V;S)-I(V;Y). \\ \Gamma_k &\ge \mathsf{E}[\phi_k(X_k)], \qquad k=1,2 \end{aligned}$$

A block Markov scheme:

- The state sequence S^n is compressed by a Wyner-Ziv scheme, with coding random variable V, and decoder side information Y^n.

$$V - S - Y$$

- The compressed state is transmitted to the decoder in the *next transmission block* as a *common message*, together with the independent messages m_1, m_2.

$$X_1 - U - X_2, \quad \text{independent of} \quad (V,S).$$

- The two codes are decoupled.

The coding scheme

The total transmission time is divided into $B+1$ blocks, each of length n.

The coding scheme

The total transmission time is divided into $B+1$ blocks, each of length n.

- First block - User 1 and User 2 transmit messages at rate R_1 and R_2.

The coding scheme

The total transmission time is divided into $B+1$ blocks, each of length n.

- First block - User 1 and User 2 transmit messages at rate R_1 and R_2.
- Block $b \in [2:B]$: the users cooperatively transmit a common message at rate R_0, and superimpose on it their private messages at rates R_1, R_2.

The coding scheme

The total transmission time is divided into $B+1$ blocks, each of length n.

- First block - User 1 and User 2 transmit messages at rate R_1 and R_2.
- Block $b \in [2:B]$: the users cooperatively transmit a common message at rate R_0, and superimpose on it their private messages at rates R_1, R_2.
  - The common message consists of a Wyner-Ziv codeword of the state in previous block, $b-1$.

The coding scheme

The total transmission time is divided into $B+1$ blocks, each of length n.

- First block - User 1 and User 2 transmit messages at rate R_1 and R_2.
- Block $b \in [2:B]$: the users cooperatively transmit a common message at rate R_0, and superimpose on it their private messages at rates R_1, R_2.
  - The common message consists of a Wyner-Ziv codeword of the state in previous block, $b-1$.
  - The channel output at block $b-1$ serves as the decoder's SI.

The coding scheme

The total transmission time is divided into $B+1$ blocks, each of length n.

- First block - User 1 and User 2 transmit messages at rate R_1 and R_2.
- Block $b \in [2:B]$: the users cooperatively transmit a common message at rate R_0, and superimpose on it their private messages at rates R_1, R_2.
  - The common message consists of a Wyner-Ziv codeword of the state in previous block, $b-1$.
  - The channel output at block $b-1$ serves as the decoder's SI.
  - The Wyner-Ziv codeword is independent of the state during its transmission.

The coding scheme

The total transmission time is divided into $B+1$ blocks, each of length n.

- First block - User 1 and User 2 transmit messages at rate R_1 and R_2.
- Block $b \in [2:B]$: the users cooperatively transmit a common message at rate R_0, and superimpose on it their private messages at rates R_1, R_2.
  - The common message consists of a Wyner-Ziv codeword of the state in previous block, $b-1$.
  - The channel output at block $b-1$ serves as the decoder's SI.
  - The Wyner-Ziv codeword is independent of the state during its transmission.
- Block $B+1$: The users do not transmit private information. Transmit only the common message, consisting of the Wyner-Ziv codeword of the state in block B.

The coding scheme

The total transmission time is divided into $B+1$ blocks, each of length n.

- First block - User 1 and User 2 transmit messages at rate R_1 and R_2.
- Block $b \in [2:B]$: the users cooperatively transmit a common message at rate R_0, and superimpose on it their private messages at rates R_1, R_2.
  - The common message consists of a Wyner-Ziv codeword of the state in previous block, $b-1$.
  - The channel output at block $b-1$ serves as the decoder's SI.
  - The Wyner-Ziv codeword is independent of the state during its transmission.
- Block $B+1$: The users do not transmit private information. Transmit only the common message, consisting of the Wyner-Ziv codeword of the state in block B.

- Backward decoding: In block $B+1$, the decoder decodes the state of block B, using the output of block B as side information.

The coding scheme

The total transmission time is divided into $B+1$ blocks, each of length n.

- First block - User 1 and User 2 transmit messages at rate R_1 and R_2.
- Block $b \in [2:B]$: the users cooperatively transmit a common message at rate R_0, and superimpose on it their private messages at rates R_1, R_2.
  - The common message consists of a Wyner-Ziv codeword of the state in previous block, $b-1$.
  - The channel output at block $b-1$ serves as the decoder's SI.
  - The Wyner-Ziv codeword is independent of the state during its transmission.
- Block $B+1$: The users do not transmit private information. Transmit only the common message, consisting of the Wyner-Ziv codeword of the state in block B.

- Backward decoding: In block $B+1$, the decoder decodes the state of block B, using the output of block B as side information.
- The decoded state of block B is used to decode the messages sent at block B: private messages, and compressed state of block $B-1$....

Example

The Gaussian MAC where the state comprises the channel noise

$$Y = X_1 + X_2 + S, \qquad S \sim \mathcal{N}\left(0, \sigma_s^2\right)$$

$$\mathsf{E}\left[X_1^2\right] \leq \Gamma_1, \qquad \mathsf{E}\left[X_2^2\right] \leq \Gamma_2.$$

Example

The Gaussian MAC where the state comprises the channel noise

$$Y = X_1 + X_2 + S, \qquad S \sim \mathcal{N}\left(0, \sigma_s^2\right)$$

$$\mathsf{E}\left[X_1^2\right] \leq \Gamma_1, \qquad \mathsf{E}\left[X_2^2\right] \leq \Gamma_2.$$

$\mathcal{C}_{\text{s-c}}(\Gamma_1, \Gamma_2)$ is the collection of all rate-pairs (R_1, R_2) satisfying

$$R_1 + R_2 \leq \frac{1}{2} \log \left(1 + \frac{(\Gamma_1^{\frac{1}{2}} + \Gamma_2^{\frac{1}{2}})^2}{\sigma_s^2} \right).$$

Example

The Gaussian MAC where the state comprises the channel noise

$$Y = X_1 + X_2 + S, \qquad S \sim \mathcal{N}\left(0, \sigma_s^2\right)$$
$$\mathsf{E}\left[X_1^2\right] \le \Gamma_1, \qquad \mathsf{E}\left[X_2^2\right] \le \Gamma_2.$$

$\mathcal{C}_{\text{s-c}}(\Gamma_1, \Gamma_2)$ is the collection of all rate-pairs (R_1, R_2) satisfying

$$R_1 + R_2 \le \frac{1}{2} \log\left(1 + \frac{(\Gamma_1^{\frac{1}{2}} + \Gamma_2^{\frac{1}{2}})^2}{\sigma_s^2}\right).$$

I.e., the full cooperation line can be achieved.

Example

The Gaussian MAC where the state comprises the channel noise

$$Y = X_1 + X_2 + S, \qquad S \sim \mathcal{N}\left(0, \sigma_s^2\right)$$

$$\mathsf{E}\left[X_1^2\right] \leq \Gamma_1, \qquad \mathsf{E}\left[X_2^2\right] \leq \Gamma_2.$$

$\mathcal{C}_{\text{s-c}}(\Gamma_1, \Gamma_2)$ is the collection of all rate-pairs (R_1, R_2) satisfying

$$R_1 + R_2 \leq \frac{1}{2} \log \left(1 + \frac{(\Gamma_1^{\frac{1}{2}} + \Gamma_2^{\frac{1}{2}})^2}{\sigma_s^2} \right).$$

I.e., the full cooperation line can be achieved.

Proof:

Example

The Gaussian MAC where the state comprises the channel noise

$$Y = X_1 + X_2 + S, \qquad S \sim \mathcal{N}\left(0, \sigma_s^2\right)$$

$$\mathsf{E}\left[X_1^2\right] \le \Gamma_1, \qquad \mathsf{E}\left[X_2^2\right] \le \Gamma_2.$$

$\mathcal{C}_{\text{s-c}}(\Gamma_1, \Gamma_2)$ is the collection of all rate-pairs (R_1, R_2) satisfying

$$R_1 + R_2 \le \frac{1}{2} \log\left(1 + \frac{(\Gamma_1^{\frac{1}{2}} + \Gamma_2^{\frac{1}{2}})^2}{\sigma_s^2}\right).$$

I.e., the full cooperation line can be achieved.

Proof:

Converse. Since strictly causal SI does not increase the capacity of the single user channel, full cooperation is an upper bound.

Example

The Gaussian MAC where the state comprises the channel noise

$$Y = X_1 + X_2 + S, \qquad S \sim \mathcal{N}\left(0, \sigma_s^2\right)$$

$$\mathsf{E}\left[X_1^2\right] \leq \Gamma_1, \qquad \mathsf{E}\left[X_2^2\right] \leq \Gamma_2.$$

$\mathcal{C}_{\text{s-c}}(\Gamma_1, \Gamma_2)$ is the collection of all rate-pairs (R_1, R_2) satisfying

$$R_1 + R_2 \leq \frac{1}{2} \log \left(1 + \frac{(\Gamma_1^{\frac{1}{2}} + \Gamma_2^{\frac{1}{2}})^2}{\sigma_s^2}\right).$$

I.e., the full cooperation line can be achieved.

Proof:

Converse. Since strictly causal SI does not increase the capacity of the single user channel, full cooperation is an upper bound.

Direct part. Two methods:

Example

The Gaussian MAC where the state comprises the channel noise

$$Y = X_1 + X_2 + S, \qquad S \sim \mathcal{N}\left(0, \sigma_s^2\right)$$
$$\mathsf{E}\left[X_1^2\right] \le \Gamma_1, \qquad \mathsf{E}\left[X_2^2\right] \le \Gamma_2.$$

$\mathcal{C}_{\text{s-c}}(\Gamma_1, \Gamma_2)$ is the collection of all rate-pairs (R_1, R_2) satisfying

$$R_1 + R_2 \le \frac{1}{2}\log\left(1 + \frac{(\Gamma_1^{\frac{1}{2}} + \Gamma_2^{\frac{1}{2}})^2}{\sigma_s^2}\right).$$

I.e., the full cooperation line can be achieved.

Proof:

Converse. Since strictly causal SI does not increase the capacity of the single user channel, full cooperation is an upper bound.

Direct part. Two methods:

- Good choice of random variables in our achievability region $\mathcal{R}_{\text{s-c}}$. (In some cases, $\mathcal{R}_{\text{s-c}}$ is tight.)

Example

The Gaussian MAC where the state comprises the channel noise

$$Y = X_1 + X_2 + S, \qquad S \sim \mathcal{N}\left(0, \sigma_s^2\right)$$
$$\mathsf{E}\left[X_1^2\right] \le \Gamma_1, \qquad \mathsf{E}\left[X_2^2\right] \le \Gamma_2.$$

$\mathcal{C}_{\text{s-c}}(\Gamma_1, \Gamma_2)$ is the collection of all rate-pairs (R_1, R_2) satisfying

$$R_1 + R_2 \le \frac{1}{2}\log\left(1 + \frac{(\Gamma_1^{\frac{1}{2}} + \Gamma_2^{\frac{1}{2}})^2}{\sigma_s^2}\right).$$

I.e., the full cooperation line can be achieved.

Proof:

Converse. Since strictly causal SI does not increase the capacity of the single user channel, full cooperation is an upper bound.

Direct part. Two methods:

- Good choice of random variables in our achievability region $\mathcal{R}_{\text{s-c}}$. (In some cases, $\mathcal{R}_{\text{s-c}}$ is tight.)
- A Schalkwijk-Kailath algorithm

MAC with causal SI

The region we had for the strictly causal case is still achievable

$$
\begin{aligned}
R_1 &\leq I(X_1; Y | X_2, U, V) \\
R_2 &\leq I(X_2; Y | X_1, U, V) \\
R_1 + R_2 &\leq I(X_1, X_2; Y | U, V) \\
R_1 + R_2 &\leq I(X_1, X_2, V; Y) - I(V; S) \\
\Gamma_k &\geq \mathsf{E}[\phi_k(X_k)], \qquad k = 1, 2
\end{aligned}
$$

with the Markov conditions

$$
\begin{aligned}
& X_1 - U - X_2 \\
& (X_1, U, X_2) \perp (V, S) \\
& V - S - Y
\end{aligned}
$$

MAC with causal SI

The region we had for the strictly causal case is still achievable

$$\begin{aligned}
R_1 &\leq I(X_1; Y|X_2, U, V) \\
R_2 &\leq I(X_2; Y|X_1, U, V) \\
R_1 + R_2 &\leq I(X_1, X_2; Y|U, V) \\
R_1 + R_2 &\leq I(X_1, X_2, V; Y) - I(V; S) \\
\Gamma_k &\geq \mathsf{E}[\phi_k(X_k)], \qquad k = 1, 2
\end{aligned}$$

with the Markov conditions

$$\begin{aligned}
&X_1 - U - X_2 \\
&(X_1, U, X_2) \perp (V, S) \\
&V - S - Y
\end{aligned}$$

But now, X_1, X_2 can depend on S.

MAC with causal SI

The region we had for the strictly causal case is still achievable

$$\begin{aligned} R_1 &\leq I(X_1; Y | X_2, U, V) \\ R_2 &\leq I(X_2; Y | X_1, U, V) \\ R_1 + R_2 &\leq I(X_1, X_2; Y | U, V) \\ R_1 + R_2 &\leq I(X_1, X_2, V; Y) - I(V; S) \\ \Gamma_k &\geq \mathsf{E}[\phi_k(X_k)], \qquad k = 1, 2 \end{aligned}$$

with the Markov conditions

$$X_1 - U - X_2$$
$$(X_1, U, X_2) \perp (V, S)$$
$$V - S - Y$$

But now, X_1, X_2 can depend on S.

$\Rightarrow$ Use Shannon strategies on top of our block Markov scheme.

MAC with causal SI

The region we had for the strictly causal case is still achievable

$$\begin{aligned} R_1 &\leq I(X_1;Y|X_2,U,V) \\ R_2 &\leq I(X_2;Y|X_1,U,V) \\ R_1+R_2 &\leq I(X_1,X_2;Y|U,V) \\ R_1+R_2 &\leq I(X_1,X_2,V;Y)-I(V;S) \\ \Gamma_k &\geq \mathsf{E}[\phi_k(X_k)], \qquad k=1,2 \end{aligned}$$

with the Markov conditions

$$\begin{aligned} & X_1 - U - X_2 \\ & (X_1,U,X_2) \perp (V,S) \\ & V - S - Y \end{aligned}$$

But now, X_1, X_2 can depend on S.

Replace (X_1,X_2) by (U_1,U_2) independent of S, and let

$$P_{X_1|U,U_1,S}, \quad P_{X_2|U,U_2,S}$$

Main result

$\mathcal{R}_{\text{cau}}$ - the CH of all $(R_1, R_2, \Gamma_1, \Gamma_2)$ satisfying

$$
\begin{aligned}
R_1 &\le I(U_1; Y | U_2, U, V) \\
R_2 &\le I(U_2; Y | U_1, U, V) \\
R_1 + R_2 &\le I(U_1, U_2; Y | U, V) \\
R_1 + R_2 &\le I(U_1, U, U_2, V; Y) - I(V; S) \\
\Gamma_k &\ge \mathsf{E}[\phi_k(X_k)], \qquad k = 1, 2
\end{aligned}
$$

for some joint distribution

$$
\begin{aligned}
P_{U,U_1,U_2,V,X_1,X_2,S,Y} = {} & P_U P_{U_1|U} P_{U_2|U} P_{V|S} P_S \cdot \\
& P_{X_1|U,U_1,S} P_{X_2|U,U_2,S} P_{Y|S,X_1,X_2}.
\end{aligned}
$$

Main result

$\mathcal{R}_{\text{cau}}$ - the CH of all $(R_1, R_2, \Gamma_1, \Gamma_2)$ satisfying

$$
\begin{aligned}
R_1 &\leq I(U_1; Y | U_2, U, V) \\
R_2 &\leq I(U_2; Y | U_1, U, V) \\
R_1 + R_2 &\leq I(U_1, U_2; Y | U, V) \\
R_1 + R_2 &\leq I(U_1, U, U_2, V; Y) - I(V; S) \\
\Gamma_k &\geq \mathsf{E}[\phi_k(X_k)] , \qquad k = 1, 2
\end{aligned}
$$

for some joint distribution

$$
\begin{aligned}
P_{U,U_1,U_2,V,X_1,X_2,S,Y} = {} & P_U P_{U_1|U} P_{U_2|U} P_{V|S} P_S \cdot \\
& P_{X_1|U,U_1,S} P_{X_2|U,U_2,S} P_{Y|S,X_1,X_2} .
\end{aligned}
$$

$$
\begin{aligned}
U_1 - U - U_2 \\
V - S - Y
\end{aligned}
$$

Main result

$\mathcal{R}_{\text{cau}}$ - the CH of all $(R_1, R_2, \Gamma_1, \Gamma_2)$ satisfying

$$\begin{aligned} R_1 &\leq I(U_1; Y | U_2, U, V) \\ R_2 &\leq I(U_2; Y | U_1, U, V) \\ R_1 + R_2 &\leq I(U_1, U_2; Y | U, V) \\ R_1 + R_2 &\leq I(U_1, U, U_2, V; Y) - I(V; S) \\ \Gamma_k &\geq \mathsf{E}[\phi_k(X_k)], \qquad k = 1, 2 \end{aligned}$$

for some joint distribution

$$\begin{aligned} P_{U,U_1,U_2,V,X_1,X_2,S,Y} = {} & P_U P_{U_1|U} P_{U_2|U} P_{V|S} P_S \cdot \\ & P_{X_1|U,U_1,S} P_{X_2|U,U_2,S} P_{Y|S,X_1,X_2}. \end{aligned}$$

$$\begin{aligned} & U_1 - U - U_2 \\ & V - S - Y \\ & (U_1, U, U_2) \perp (V, S) \end{aligned}$$

Main result

$\mathcal{R}_{\text{cau}}$ - the CH of all $(R_1, R_2, \Gamma_1, \Gamma_2)$ satisfying

$$\begin{aligned} R_1 &\le I(U_1; Y | U_2, U, V) \\ R_2 &\le I(U_2; Y | U_1, U, V) \\ R_1 + R_2 &\le I(U_1, U_2; Y | U, V) \\ R_1 + R_2 &\le I(U_1, U, U_2, V; Y) - I(V; S) \\ \Gamma_k &\ge \mathsf{E}[\phi_k(X_k)], \qquad k = 1, 2 \end{aligned}$$

for some joint distribution

$$\begin{aligned} P_{U,U_1,U_2,V,X_1,X_2,S,Y} = {} & P_U P_{U_1|U} P_{U_2|U} P_{V|S} P_S \cdot \\ & P_{X_1|U,U_1,S} P_{X_2|U,U_2,S} P_{Y|S,X_1,X_2}. \end{aligned}$$

Theorem 2 (Causal SI) $\mathcal{R}_{\text{cau}} \subseteq \mathcal{C}_{\text{cau}}$

The naïve approach

The naïve approach – using Shannon strategies, without block Markov coding of the state.

The naïve approach

The naïve approach – using Shannon strategies, without block Markov coding of the state. It leads to the region of all (R_1, R_2) satisfying

$$
\begin{aligned}
R_1 &\leq I(T_1; Y|T_2, Q) \\
R_2 &\leq I(T_2; Y|T_1, Q) \\
R_1 + R_2 &\leq I(T_1, T_2; Y|Q)
\end{aligned}
$$

for some joint distribution $P_Q P_{T_1|Q} P_{T_2|Q} P_{Y|T_1,T_2}$.

The naïve approach

The naïve approach – using Shannon strategies, without block Markov coding of the state. It leads to the region of all (R_1, R_2) satisfying

$$\begin{aligned} R_1 &\le I(T_1; Y|T_2, Q) \\ R_2 &\le I(T_2; Y|T_1, Q) \\ R_1 + R_2 &\le I(T_1, T_2; Y|Q) \end{aligned}$$

for some joint distribution $P_Q P_{T_1|Q} P_{T_2|Q} P_{Y|T_1,T_2}$. Here

$T_k,\, k = 1, 2$ are random Shannon strategies:

$$T_k \in \mathcal{T}_k, \quad \text{the set of mappings} \quad t_k : \mathcal{S} \to \mathcal{X}_k$$

The naïve approach

The naïve approach – using Shannon strategies, without block Markov coding of the state. It leads to the region of all (R_1, R_2) satisfying

$$R_1 \leq I(T_1; Y | T_2, Q)$$

$$R_2 \leq I(T_2; Y | T_1, Q)$$

$$R_1 + R_2 \leq I(T_1, T_2; Y | Q)$$

for some joint distribution $P_Q P_{T_1|Q} P_{T_2|Q} P_{Y|T_1,T_2}$. Here

T_k, $k = 1, 2$ are random Shannon strategies:

$$T_k \in \mathcal{T}_k, \quad \text{the set of mappings} \quad t_k : \mathcal{S} \to \mathcal{X}_k$$

Q is a time sharing random variable,

The naïve approach

The naïve approach – using Shannon strategies, without block Markov coding of the state. It leads to the region of all (R_1, R_2) satisfying

$$\begin{aligned} R_1 &\leq I(T_1; Y | T_2, Q) \\ R_2 &\leq I(T_2; Y | T_1, Q) \\ R_1 + R_2 &\leq I(T_1, T_2; Y | Q) \end{aligned}$$

for some joint distribution $P_Q P_{T_1|Q} P_{T_2|Q} P_{Y|T_1,T_2}$. Here

T_k, $k = 1, 2$ are random Shannon strategies:

$$T_k \in \mathcal{T}_k, \quad \text{the set of mappings} \quad t_k : \mathcal{S} \to \mathcal{X}_k$$

Q is a time sharing random variable, and

$$P_{Y|T_1,T_2}(y|t_1, t_2) = \sum_{s \in \mathcal{S}} P_S(s) P_{Y|S,X_1,X_2}\big(y|s, t_1(s), t_2(s)\big).$$

The naïve approach

The naïve approach – using Shannon strategies, without block Markov coding of the state. It leads to the region of all (R_1, R_2) satisfying

$$R_1 \leq I(T_1; Y|T_2, Q)$$
$$R_2 \leq I(T_2; Y|T_1, Q)$$
$$R_1 + R_2 \leq I(T_1, T_2; Y|Q)$$

for some joint distribution $P_Q P_{T_1|Q} P_{T_2|Q} P_{Y|T_1,T_2}$.

We denote this region as $\mathcal{R}^{\text{naïve}}$.

The naïve approach

- $\mathcal{R}_{\text{cau}}$ contains the region of the naïve approach, since we can always choose degenerate V.

- In some cases, the inclusion is strict.

Example

The noiseless binary MAC with input selector:

$$\mathcal{X}_1 = \mathcal{X}_2 = \mathcal{Y} = \{0,1\}, \quad \mathcal{S} = \{1,2\}, \quad P_S(S=2) = p > 0.5$$

Example

The noiseless binary MAC with input selector:

$$\mathcal{X}_1 = \mathcal{X}_2 = \mathcal{Y} = \{0, 1\}, \quad \mathcal{S} = \{1, 2\}, \quad P_S(S = 2) = p > 0.5$$

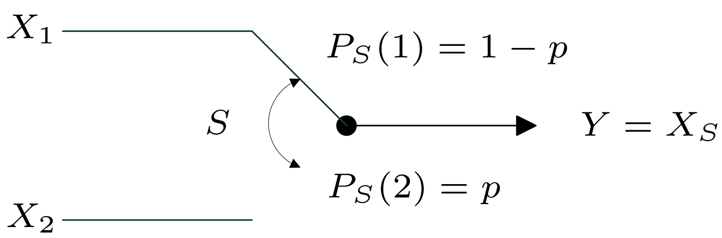
$$Y = X_S$$

Example

The noiseless binary MAC with input selector:

$$\mathcal{X}_1 = \mathcal{X}_2 = \mathcal{Y} = \{0, 1\}, \quad \mathcal{S} = \{1, 2\}, \quad P_S(S = 2) = p > 0.5$$

$$Y = X_S$$

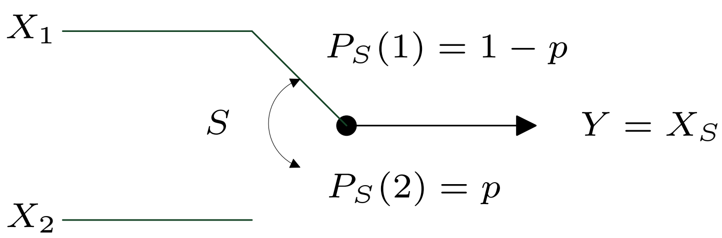


Example

The noiseless binary MAC with input selector:

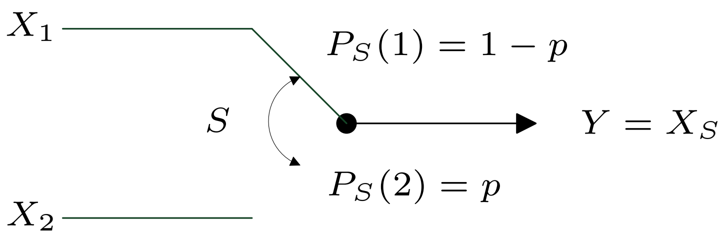
$$\mathcal{X}_1 = \mathcal{X}_2 = \mathcal{Y} = \{0,1\}, \quad \mathcal{S} = \{1,2\}, \quad P_S(S=2) = p > 0.5$$

$$Y = X_S$$

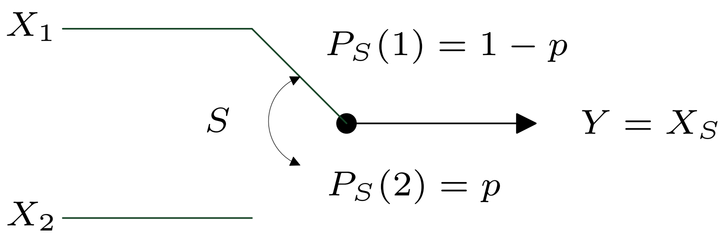


- If the decoder knows S, user 1 can transmit at rate $1-p$.
- Hence, $1-p$ is an upper bound on the transmission rate of user 1 in our model.

Example



Example

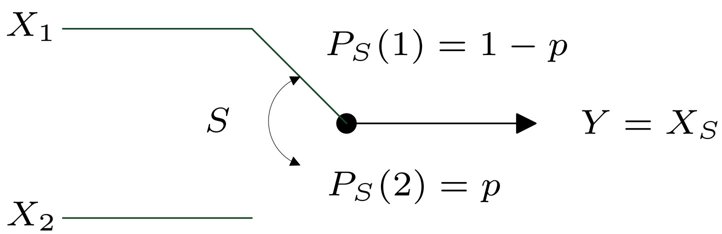


With a proper choice of random variables in $\mathcal{R}_{\text{cau}}$

$$(R_1, R_2) = \Big(\min\big\{1-p, 1-H_b(p)\big\},\, 0\Big) \in \mathcal{R}_{\text{cau}}$$

(Observe – achieves the maximal rate of user 1 for $p \geq H_b(p)$.)

Example



With a proper choice of random variables in $\mathcal{R}_{\text{cau}}$

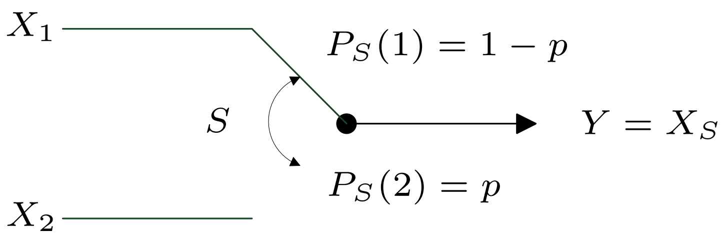
$$(R_1, R_2) = \Big(\min\big\{1-p, 1-H_b(p)\big\},\ 0\Big) \in \mathcal{R}_{\text{cau}}$$

(Observe – achieves the maximal rate of user 1 for $p \geq H_b(p)$.)

The maximal rate of user 1 in the naïve approach is

$$R_{2,\text{max}}^{\text{naïve}} = \log_2\left(1 + (1-p)p^{\frac{p}{1-p}}\right) \quad \text{bits}$$

Example



With a proper choice of random variables in $\mathcal{R}_{\text{cau}}$

$$(R_1, R_2) = \Big(\min\big\{1 - p, 1 - H_b(p)\big\},\, 0\Big) \in \mathcal{R}_{\text{cau}}$$

(Observe – achieves the maximal rate of user 1 for $p \geq H_b(p)$.)

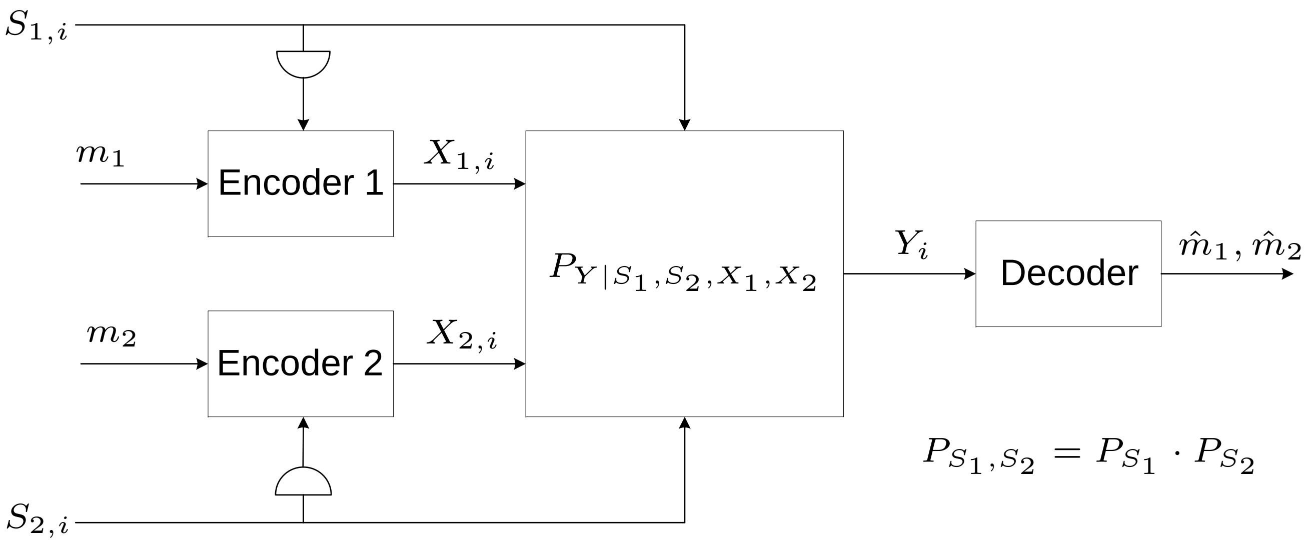
The maximal rate of user 1 in the naïve approach is

$$R_{2,\text{max}}^{\text{naïve}} = \log_2\left(1 + (1 - p)p^{\frac{p}{1-p}}\right) \quad \text{bits}$$

For sufficiently large value of p,

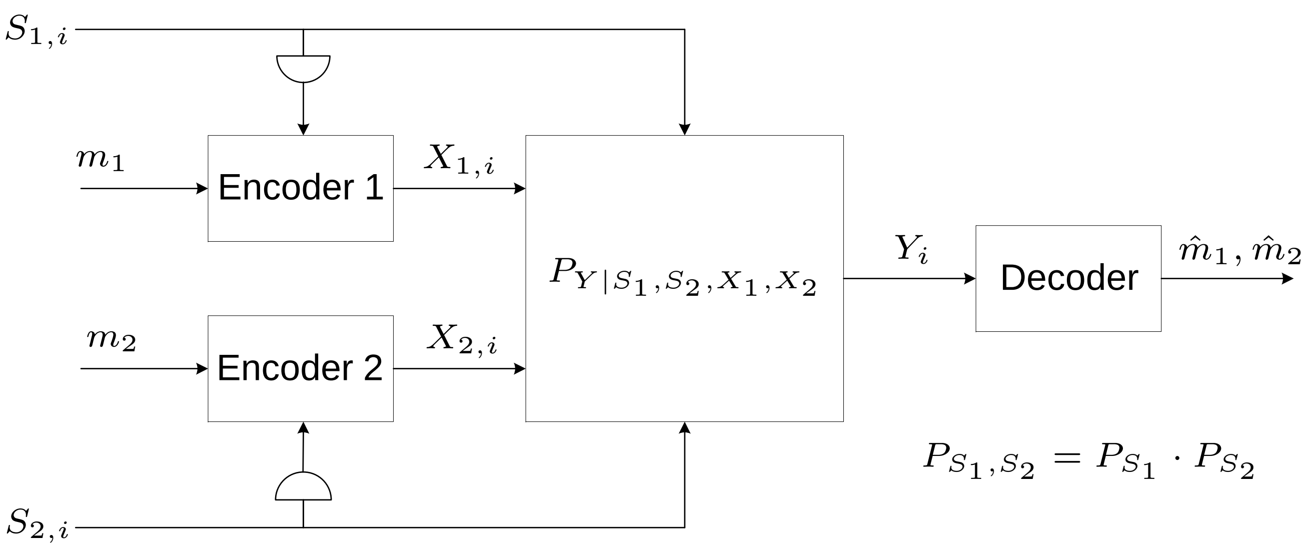
$$R_{2,\text{max}}^{\text{naïve}} < \min\big\{1 - p, 1 - H_b(p)\big\}$$

MAC with independent SI streams



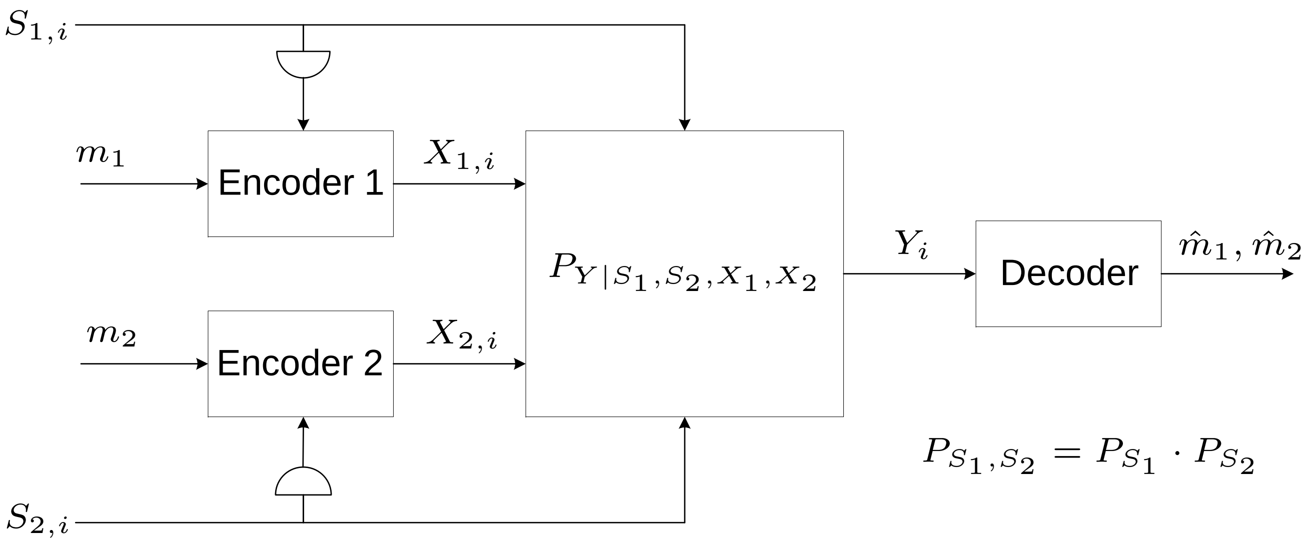
$$P_{S_1,S_2} = P_{S_1} \cdot P_{S_2}$$

MAC with independent SI streams



- *Cooperation* in the compression and transmission of the state is not possible.

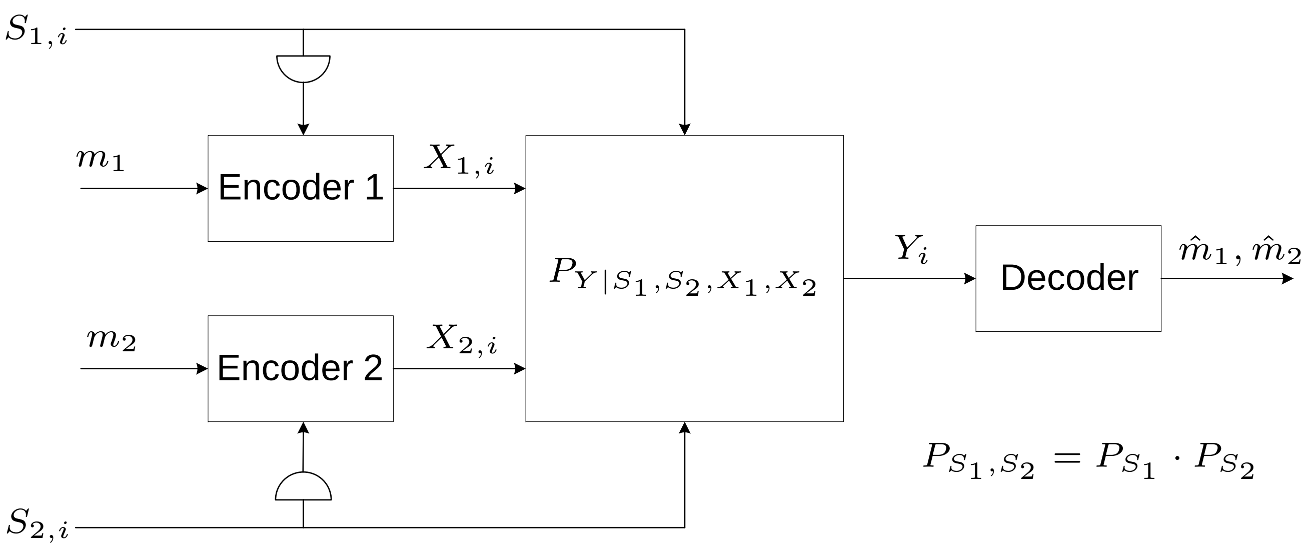
MAC with independent SI streams



$$P_{S_1,S_2} = P_{S_1} \cdot P_{S_2}$$

- *Cooperation* in the compression and transmission of the state is not possible.
- Yet, compression and transmission of the states to the decoder is beneficial, and enlarges the capacity region of the MAC.

MAC with independent SI streams



- *Cooperation* in the compression and transmission of the state is not possible.
- Yet, compression and transmission of the states to the decoder is beneficial, and enlarges the capacity region of the MAC.
- Utilize distributed Wyner-Ziv compression and block Markov coding (ISIT 2010).

Summary

- Derived achievable region for the MAC with common strictly causal SI, based on block Markov encoding of the state.
- Strictly causal SI enlarges the capacity region of the MAC.
- Extended the results to causal SI
- The new region for causal Si is strictly better that the region obtained by the naïve approach.

- Strictly causal SI is beneficial even when the states available at the encoders are independent (ISIT 2010).

Thank You!