Multi-user information theory and an example: the two-way relay channel

Natasha Devroye
Assistant Professor
University of Illinois at Chicago
http://www.ece.uic.edu/~devroye

Multi-user information theory and an example: the two-way relay channel

Monologues vs. Dialogues

Monologue = one-way = uni-directional

Dialogue = two-way = bi-directional

Two-way communication applications - wired

Video conferencing

Telesurgery

Data synchronization

Two-way communication applications - wireless

Battlefield telesurgery

Rural telesurgery

Video conferencing

Monologues vs. Dialogues

\[ W_{12} \quad 1 \quad \rightarrow \quad 2 \quad \tilde{W}_{12} \]

Monologue

\[ \frac{W_{12}}{W_{21}} \quad 1 \quad \leftarrow \quad 2 \quad \frac{\tilde{W}_{12}}{W_{21}} \]

Dialogue

1 Dialogue ≠ or = 2 Monologues?

It depends....

we will use information theory to find out.
Outline

- Information theory - what, why, when

- Two-way channel - channel coding

- Wireless channels and networks

- Two-way cellular-like networks

- Two-way relay channels - canonical example of wireless network coding

Overall - much is still unknown

Information theory - what, why, when


What is information?

How fast can we communicate?

How much can we compress information?

Source vs. channel coding

Source

Encoder

Source encoder

Channel

Decoder

Channel decoder

Destination

Remove redundancy

Controlled adding of redundancy

Noise

Decode signals, detect/correct errors

Restore source

Outline

- Information theory - what, why, when

- Two-way channel - channel coding

- Wireless channels and networks

- Two-way cellular-like networks

- Two-way relay channels - canonical example of wireless network coding

Information theory’s claims to fame

FADING CHANNEL

Reliable communication possible $\Leftrightarrow H<C$

Source coding

Source $= \text{random variable}$

Ultimate data compression limit is the source’s entropy $H$.

Minimum $R$ needed is $H(X) = -\sum_x p(x) \log_2(p(x))$

Channel coding

- Channel is $\text{independent distributions}$

- Ultimate transmission rate is the channel capacity $C$

Minimum $R$ needed is described by the Rate-Distortion function $R(D)$

$R(D) = \min_{p(\hat{x}|x)} I(X; \hat{X})$

s.t. $d(\hat{X}^n, X^n) \leq D$

Source coding = data compression

$X^n \rightarrow p(x)$

$X^n \rightarrow f(x^n) \in (1, 2, \ldots, 2^R)$

$R$ (bit/source symbol)

$\hat{X}^n$

$\hat{X}^n \rightarrow$ Destination

$X^n$
Channel capacity: a cute example

Source vs. channel coding

- Information channel capacity:
  \[ C = \max_{p(x)} I(X; Y) \]

- Operational channel capacity:
  Highest rate (bits/channel use) that can communicate at reliably

- Channel coding theorem says: information capacity = operational capacity

Communication system model

What is the capacity of this channel?

Channel capacity: a cute example

Source \rightarrow \text{Encoder} \rightarrow \text{Channel} \rightarrow \text{Decoder} \rightarrow \text{Destination}

\[ X^n \rightarrow P(y|x) \rightarrow Y^n \]

Source \rightarrow A, B, C, D \rightarrow \text{Encoder} \rightarrow \text{Channel} \rightarrow \text{Decoder} \rightarrow \text{Destination}

What is the capacity of this channel?
Channel capacity: a cute example

Source A, B, C, D → Encoder A → Channel

Decoder

MA → AB.

ABA. → AAA
AZ
BBA?

Non-overlapping outputs!

Inputs

Outputs

Capacity in general

- Main idea was to reduce the rate (from a 27-letter input per channel use to a 9-letter input per channel use) so as to produce non-overlapping outputs!

Capacity: key ideas

- choose input set of codewords so they are “non-confusable” at the output
- number of these that we can chose will determine the channel’s capacity
- number that we can choose will depend on the distribution \( p(y|x) \) which characterizes the channel

Mathematical description of capacity

- Can achieve reliable communication for all transmission rates \( R \):
  \[ R < C \]
  \[ R > C \]

- BUT, probability of decoding error always bounded away from zero if

One-way channel capacity

\[ C = \max_{p(x)} I(X; Y) \]

\[ I(X; Y) = \sum_{x,y} p(x,y) \log \left( \frac{p(x,y)}{p(x)p(y)} \right) \]

A few examples
Entropy of a random variable

- (A) entropy is the measure of average uncertainty in the random variable
- (B) entropy is the average number of bits needed to describe the random variable
- (C) entropy is measured in bits?
- (D) \( H(X) = - \sum_x p(x) \log_2(p(x)) \)
- (E) entropy of a deterministic value is 0

Entropy of a random variable

- (A) entropy is the measure of average uncertainty in the random variable
- (B) entropy is the average number of bits needed to describe the random variable
- (C) entropy is measured in bits?
- (D) \( H(X) = - \sum_x p(x) \log_2(p(x)) \)
- (E) entropy of a deterministic value is 0

What is capacity / mutual information?

\[ C = \max_{p(x)} I(X; Y) \]
Entropy of a uniform distribution

- Let X be uniformly distributed over 8 outcomes. What is the entropy of X?
  \[ H(X) = \sum_{x=1}^{8} p(x) \log_2(p(x)) = -\sum_{x=1}^{8} \frac{1}{8} \log_2 \left( \frac{1}{8} \right) = \log_2(8) = 3 \text{ (bits)} \]
- This is the number of bits needed to describe X!

- By extension, for a discrete random variable taking on K outcomes, the \textbf{maximal entropy} is attained by a uniform distribution and is equal to the number of bits needed to describe K:
  \[ H(X) = \log_2(K) \]

Entropy of a continuous random variable

- \textbf{Entropy:}
  \[ H(X) = -\sum_x p(x) \log_2(p(x)) \]

- \textbf{Differential entropy:}
  \[ h(X) = -\int f(x) \log(x) \, dx \]

Entropy maximization

- \textbf{Uniform} distribution maximizes entropy for a given # outcomes
  \[ \max_{X:|X|=K} H(X) = \log_2(K) \]

- \textbf{Gaussian} maximizes entropy for a given covariance constraint
  \[ \max_{E[X|X] = K} h(X) = \frac{1}{2} \log \left( (2\pi e)^n |K| \right) \]

Entropy of a non-uniform distribution

- Suppose X represents the outcome of a horse race with 8 horses, which win with probabilities \( \left( \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \frac{1}{32}, \frac{1}{64} \right) \)
  \[ H(X) = \frac{1}{2} \log_2 \left( \frac{1}{2} \right) - \frac{1}{2} \log_2 \left( \frac{1}{4} \right) - \frac{1}{8} \log_2 \left( \frac{1}{8} \right) - \frac{1}{16} \log_2 \left( \frac{1}{16} \right) - \frac{1}{32} \log_2 \left( \frac{1}{32} \right) - \frac{1}{64} \log_2 \left( \frac{1}{64} \right) \]
- 8 outcomes, 3 bits? But on average can represent with 2 bits!

\[ \{000,001,010,011,100,101,110,111\} \]
3 bits

\[ \{010,110,1110,11110,11111,1011110,1111110,1111111\} \]
2 bits (on average)

Entropy of a Gaussian random variable

- \textbf{Differential entropies of Gaussian distributions:}
  \[ h \left( \mathcal{N}(0, \sigma^2) \right) = \frac{1}{2} \log \left( 2\pi e \sigma^2 \right) \]
  \[ h \left( \mathcal{N}_a(\mu, K) \right) = \frac{1}{2} \log \left( (2\pi e)^n |K| \right) \]

Mutual information between 2 random variables:

\[ I(X;Y) = \sum p(x,y) \log \left( \frac{p(x,y)}{p(x)p(y)} \right) \]
\[ = H(X) - H(X|Y) \]
\[ = H(Y) - H(Y|X) \]
Mutual information between 2 random variables:

\[ I(X;Y) = \sum p(x, y) \log \left( \frac{p(x, y)}{p(x)p(y)} \right) \]

\[ = H(X) - H(X|Y) \]

\[ = H(Y) - H(Y|X) \]

(A) \( I(X;Y) \) is the reduction in the uncertainty about \( X \) due to knowledge of \( Y \)

(B) if \( X, Y \) are independent \( I(X;Y) = 0 \)

(C) if \( X=Y \) then \( I(X;Y) = H(X) \)

(D) \( I(X;Y) \) is non-negative

Mutual information between 2 random variables:

\[ I(X;Y) = \sum p(x, y) \log \left( \frac{p(x, y)}{p(x)p(y)} \right) \]

\[ = H(X) - H(X|Y) \]

\[ = H(Y) - H(Y|X) \]

(A) \( I(X;Y) \) is the reduction in the uncertainty about \( X \) due to knowledge of \( Y \)

(B) if \( X, Y \) are independent \( I(X;Y) = 0 \)

(C) if \( X=Y \) then \( I(X;Y) = H(X) \)

(D) \( I(X;Y) \) is non-negative

Mutual information between 2 random variables:

\[ I(X;Y) = \sum p(x, y) \log \left( \frac{p(x, y)}{p(x)p(y)} \right) \]

\[ = H(X) - H(X|Y) \]

\[ = H(Y) - H(Y|X) \]

(A) \( I(X;Y) \) is the reduction in the uncertainty about \( X \) due to knowledge of \( Y \)

(B) if \( X, Y \) are independent \( I(X;Y) = 0 \)

(C) if \( X=Y \) then \( I(X;Y) = H(X) \)

(D) \( I(X;Y) \) is non-negative

Properties of mutual information

Channel capacity

\[ C = \max_{p(x)} I(X;Y) \]

\[ I(X;Y) = \sum_{x,y} p(x, y) \log \left( \frac{p(x, y)}{p(x)p(y)} \right) \]
Channel capacity
1 bit/channel use

\[ C = \max_{p(x)} I(X; Y) \]
\[ = \max_{p(x)} H(X) - H(X|Y) \]
\[ = \max_{p(x)} H(X) - \log_2(3) \]
\[ = \log_2(27) - \log_2(3) = \log_2(9) \]

Discrete memoryless channel capacity

What if

X and Y are not bits, but real numbers?
AWGN channel capacity

\[ C = \max_{p(x)} I(X;Y) \]

\[ = \max_{p(x)} h(X) - h(Y|X) \]

\[ = \max_{p(x)} h(Y) - h(X|Y) \]

\[ = \max_{p(x)} h(X + N) - h(Y|X) \]

\[ = \max_{p(x)} \log(2^{2|h|^2P / (N/2)}) - \frac{1}{2} \log(2\pi ePN) \]

\[ = \frac{1}{2} \log \left( \frac{h^2P}{N/2} \right) \]

\[ = \frac{1}{2} \log \left( \frac{h^2P}{N/2} \right) \]

Source vs. channel coding

**Source**

- Remove redundancy

**Channel**

- Controlled adding of redundancy

**Decode signals, detect/correct errors**

**Restore source**

Use?

- algebraic codes
- convolutional codes
- iterative codes (LDPC, turbo)

Claude Shannon — Born on the planet Earth (Sid III) in the year 1916 A.D. Generally regarded as the father of the Information Age, he formulated the notion of channel capacity in 1948 A.D. Within several decades, mathematicians and engineers had found practical ways to communicate reliably at data rates within 1% of the Shannon limit...
Use?

- Benchmark for performance of practical systems
- Guideline in designing systems - what’s worth shooting for?
- Theoretical insights can lead to practical insights

Outline

- Information theory - what, why, when
- Source coding, channel coding, entropy and mutual information, capacity, Gaussian noise channel
- Two-way channel - channel coding
- Wireless channels and networks
- Two-way cellular-like networks
- Two-way relay channels - canonical example of wireless network coding

One-way channel capacity

$$C = \max_{p(x)} I(X; Y)$$ bits/channel use

$$I(X; Y) = \sum_{x,y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)}$$ symmetric in its arguments!

Two-way channel (historical)

So now what?

Unsolved

Fundamental

-One-way channel capacity - notation

$$X^n \xrightarrow{p(y|x)} Y^n$$

$$X \xrightarrow{p(y|x)} Y$$

$$1^{R_{12}} Y$$

-One-way channel capacity - notation

$$X^n \xrightarrow{p(y|x)} Y^n$$

$$X \xrightarrow{p(y|x)} Y$$

$$1^{R_{12}} Y$$

-One-way channel capacity - notation

$$X^n \xrightarrow{p(y|x)} Y^n$$

$$X \xrightarrow{p(y|x)} Y$$

$$1^{R_{12}} Y$$

-One-way channel capacity - notation

$$X^n \xrightarrow{p(y|x)} Y^n$$

$$X \xrightarrow{p(y|x)} Y$$

$$1^{R_{12}} Y$$

-One-way channel capacity - notation

$$X^n \xrightarrow{p(y|x)} Y^n$$

$$X \xrightarrow{p(y|x)} Y$$

$$1^{R_{12}} Y$$
Two-way channel capacity

One-way

\[ X \xrightarrow{p(y|x)} Y \]

Two-way

\[ X_1 \xrightarrow{p(y_1|y_2,x_1,x_2)} Y_1 \]

\[ X_2 \xrightarrow{Y_2} \]

Two-way communication channels

Claude E. Shannon
Massachusetts Institute of Technology
Cambridge, Massachusetts

1. Introduction

A two-way communication channel is shown schematically in figure 1. Here \( x_i \) is an input letter to the channel at terminal 1 and \( y_i \) an output while \( x_2 \) is an input at terminal 2 and \( y_2 \) the corresponding output. Once each second, say, new inputs \( x_i \) and \( x_2 \) may be chosen from corresponding input alphabets and put into the channel; outputs \( y_i \) and \( y_2 \) may then be observed. These outputs will be related statistically to the inputs and perhaps historically to previous inputs and outputs if the channel has memory. The problem is to communicate in both directions through the channel as effectively as possible. Particularly, we wish to determine what pairs of signalling rates \( R_1 \) and \( R_2 \) for the two directions can be approached with arbitrarily small error probabilities.

Before making these notions precise, we give some simple examples. In figure 2 the two-way channel decomposes into two independent one-way noiseless binary channels.

This work was supported in part by the U.S. Army (Signal Corps), the U.S. Air Force (Office of Scientific Research, Air Research and Development Command), and the U.S. Navy (Office of Naval Research).

Two-way channel capacity region

One-way **Capacity**

\[ R = \max_{p(x,y)} I(X;Y) \]

Two-way **Capacity Region**

\[ R_{12} \]

\[ R_{21} \]

Two-way channel capacity

One-way

\[ X \xrightarrow{p(y|x)} Y \]

Two-way

\[ X_1 \xrightarrow{p(y_1|y_2,x_1,x_2)} Y_1 \]

\[ X_2 \xrightarrow{Y_2} \]

Two-way communication channels

Claude E. Shannon
Massachusetts Institute of Technology
Cambridge, Massachusetts

1. Introduction

A two-way communication channel is shown schematically in figure 1. Here \( x_i \) is an input letter to the channel at terminal 1 and \( y_i \) an output while \( x_2 \) is an input at terminal 2 and \( y_2 \) the corresponding output. Once each second, say, new inputs \( x_i \) and \( x_2 \) may be chosen from corresponding input alphabets and put into the channel; outputs \( y_i \) and \( y_2 \) may then be observed. These outputs will be related statistically to the inputs and perhaps historically to previous inputs and outputs if the channel has memory. The problem is to communicate in both directions through the channel as effectively as possible. Particularly, we wish to determine what pairs of signalling rates \( R_1 \) and \( R_2 \) for the two directions can be approached with arbitrarily small error probabilities.

Before making these notions precise, we give some simple examples. In figure 2 the two-way channel decomposes into two independent one-way noiseless binary channels.

This work was supported in part by the U.S. Army (Signal Corps), the U.S. Air Force (Office of Scientific Research, Air Research and Development Command), and the U.S. Navy (Office of Naval Research).

Two-way channel capacity region

One-way **Capacity**

\[ R = \max_{p(x,y)} I(X;Y) \]

Two-way **Capacity Region**

\[ R_{12} \]

\[ R_{21} \]

Two-way channel capacity

One-way

\[ X \xrightarrow{p(y|x)} Y \]

Two-way

\[ X_1 \xrightarrow{p(y_1|y_2,x_1,x_2)} Y_1 \]

\[ X_2 \xrightarrow{Y_2} \]

Two-way communication channels

Claude E. Shannon
Massachusetts Institute of Technology
Cambridge, Massachusetts

1. Introduction

A two-way communication channel is shown schematically in figure 1. Here \( x_i \) is an input letter to the channel at terminal 1 and \( y_i \) an output while \( x_2 \) is an input at terminal 2 and \( y_2 \) the corresponding output. Once each second, say, new inputs \( x_i \) and \( x_2 \) may be chosen from corresponding input alphabets and put into the channel; outputs \( y_i \) and \( y_2 \) may then be observed. These outputs will be related statistically to the inputs and perhaps historically to previous inputs and outputs if the channel has memory. The problem is to communicate in both directions through the channel as effectively as possible. Particularly, we wish to determine what pairs of signalling rates \( R_1 \) and \( R_2 \) for the two directions can be approached with arbitrarily small error probabilities.

Before making these notions precise, we give some simple examples. In figure 2 the two-way channel decomposes into two independent one-way noiseless binary channels.

This work was supported in part by the U.S. Army (Signal Corps), the U.S. Air Force (Office of Scientific Research, Air Research and Development Command), and the U.S. Navy (Office of Naval Research).

When is

Two-way

\[ W_{12} \]

\[ R_{12} \]

\[ W_{21} \]

Two one-ways

\[ W_{12} \]

\[ R_{12} \]

\[ W_{21} \]

\[ \tilde{W}_{12} \]

\[ \tilde{W}_{21} \]

equal to

Two one-ways

\[ W_{12} \]

\[ R_{12} \]

\[ W_{21} \]

\[ \tilde{W}_{12} \]

\[ \tilde{W}_{21} \]

Two-way channel capacity region

One-way **Capacity**

\[ R = \max_{p(x,y)} I(X;Y) \]

Two-way **Capacity Region**

\[ R_{12} \]

\[ R_{21} \]

When is

Two-way

\[ W_{12} \]

\[ R_{12} \]

\[ W_{21} \]

equal to

Two one-ways

\[ W_{12} \]

\[ R_{12} \]

\[ W_{21} \]

\[ \tilde{W}_{12} \]

\[ \tilde{W}_{21} \]

Models for two-way adaptation

One-way: no adaptation possible

\[ x_1^n(w_{12}) \]

\[ x_2^n(w_{21}) \]

Two-way: no adaptation

\[ x_1^n(w_{12}, y_1^{n-1}) \]

\[ x_2^n(w_{21}, y_2^{n-1}) \]

Two-way: full adaptation

\[ x_1^n(w_{12}) \]

\[ x_2^n(w_{21}) \]

\[ x_1^n(w_{12}) \]

\[ x_2^n(w_{21}) \]

\[ \tilde{W}_{12} \]

\[ \tilde{W}_{21} \]

(Aside: \( I(X;Y) \) and \( I(X;Y|Z) \))

\[ I(X;Y) = \sum_{x \in X} \sum_{y \in Y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)} \]

\[ = E_{p(x,y)} \log \frac{p(X,Y)}{p(X)p(Y)} \]

\[ I(X;Y|Z) = \sum_{x \in X \cap Y \cap Z} \sum_{y \in Y \cap Z} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \]

\[ = E_{p(x,y,z)} \log \frac{p(X,Y|Z)}{p(X|Z)p(Y|Z)} \]
(Aside: $I(X;Y)$ and $I(X;Y|Z)$)

Bernoulli(1/2)

\[ I(X;Y) = H(Y) - H(Y|X) \]
\[ = \log_2(3)/4 + \log_2(2)/2 + \log_2(4)/4 - \log_2(2) \]
\[ = 0.5 \text{ bits} \]

\[ I(X;Y|Z) = H(Y|Z) - H(Y|X,Z) \]
\[ = \log_2(2) - 0 \]
\[ = 1.0 \text{ bit} \]

When is capacity known

- Parallel two-way channel
- Mod-2 adder
- Two-way restricted channel
- Two-way “push-to-talk” channel
- Two-way Gaussian noise channel (full & half duplex, restricted & unrestricted)

When is capacity unknown

- General unrestricted discrete memoryless channels
- Binary multiplier channel (BMC)

Capacity: binary mod-2 adder channel

\[ y_1 = y_2 = x_1 + x_2 \text{ (mod 2)} \]

How to achieve capacity region?

General results

**Inner bound**

\[ R_1 \leq I(X_1;Y_2|X_2) \]
\[ R_2 \leq I(X_2;Y_1|X_1) \]

where $X_1$ and $X_2$ follow the joint distribution $p(x_1, x_2) = p(x_1)p(x_2)$. **Not in general equal!**

**Outer bound**

\[ R_1 \leq I(X_1;Y_2|X_2) \]
\[ R_2 \leq I(X_2;Y_1|X_1) \]

where the joint distribution of random variables $X_1$ and $X_2$ is $p(x_1, x_2)$.

Figure 2

\[ R_{12} \leq C_{K_1} \]
\[ R_{21} \leq C_{K_2} \]

Achieving mod 2 adder channel capacity

**Receiver 1:**

\[ \hat{y}_1 = \hat{y}_2 = x_1 + x_2 \text{ (mod 2)} \]

**Receiver 2:**

\[ \hat{y}_1 = \hat{y}_2 = x_1 + x_2 \text{ (mod 2)} \]

\[ \hat{x}_1 = \hat{x}_2 = x_1 + x_2 \text{ (mod 2)} \]

**EXPLOIT TWO-WAY!**
Capacity: restricted channel

\[ R_1 \leq I(X_1; Y_2 | X_2) \]
\[ R_2 \leq I(X_2; Y_1 | X_1) \]

where \( X_1 \) and \( X_2 \) follow the joint distribution \( p(x_1, x_2) = p(x_1) p(x_2) \).

[Shannon ’61]

Capacity: Gaussian noise channel

\[ Y_1 = aX_1 + bX_2 + N_1 \sim \mathcal{N}(0, \sigma_1^2) \]
\[ Y_2 = cX_1 + dX_2 + N_2 \sim \mathcal{N}(0, \sigma_2^2) \]

No dependence on “a” or “d”  [Han ’84]

When is capacity known

- Parallel two-way channel
- Mod-2 adder
- Two-way restricted channel
- Two-way “push-to-talk” channel
- Two-way Gaussian noise channel (full & half duplex, restricted & unrestricted)

When is capacity unknown

- General unrestricted discrete memoryless channels
- Binary multiplier channel (BMC)

Two-way: half duplex

\[ Y_1 = aX_1 + bX_2 + N_1 \]
\[ Y_2 = cX_1 + dX_2 + N_2 \]

\[ R_1 \leq (1/2) \log(1 + c^2 P_1 / \sigma_2^2) \]
\[ R_2 \leq (1/2) \log(1 + b^2 P_2 / \sigma_1^2) \]

- Two parallel channels!!
  - Achieved by Gaussian inputs
  - “Feedback” does not help here

Capacity unknown: Binary Multiplier Channel

\[ y_1 = y_2 = x_1 x_2, \text{ where } x_1, x_2 \in \{0, 1\}. \]

<table>
<thead>
<tr>
<th>lower/upper bounds</th>
<th>Rate (bits)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shannon’s lower bound</td>
<td>0.01695</td>
</tr>
<tr>
<td>Hagelbarger’s lower bound</td>
<td>0.503</td>
</tr>
<tr>
<td>Schalkwijk’s lower bound</td>
<td>0.63056</td>
</tr>
<tr>
<td>Shannon’s upper bound</td>
<td>0.056424</td>
</tr>
<tr>
<td>Zheng’s upper bound</td>
<td>0.63901</td>
</tr>
</tbody>
</table>

General techniques for two-way channels

Adaptation

Adaptive codewords

![Adaptive codewords](image)

The space over which we can code (x’s) is enormous!

Adaptive codewords:

\[ x_1^{(L)}(w_1), y_2^{(L)}(w_2) \]

Non-adaptive codewords:

\[ x_1 = \{0, 1\}, \quad y_1 = x_1, x_2 \]

Code of user 1:

\[
\begin{align*}
u_{11} &= 0 \\
u_{12} &= 1
\end{align*}
\]

Code of user 2:

\[
\begin{align*}
u_{21} &= 0 \\
u_{22} &= 1
\end{align*}
\]

Can we take this adaptation into account?

CAUSAL adaptation
From mutual information to directed information

Need to extend the symmetric \( I(X^N; Y^N) \) to account for causally adaptive codewords

Marko/Massey's "Directed Information"

\[
I(X^N \rightarrow Y^N) := \sum_{n=1}^{N} I(X^n; Y_n | Y^{n-1})
\]

Kramer's "Causally conditioned Directed Information"

\[
I(X^N \rightarrow Y^N | Z^N) := \sum_{n=1}^{N} I(X^n; Y_n | Y^{n-1}, Z^n)
\]

Usage: capacity of two-way channels

\[
X_1 \in \{0, 1\}, \quad Y_1 = X_1 \oplus X_2
\]

Adaptive codewords \( A_3 \)

\[
L = N = 3 \text{ channel uses}
\]

\[
Y_2 = Y_1 \mod 2, \quad X_2 \in \{0, 1, 2\}
\]

Adaptive codewords \( A_3 \)

Take away points - AWGN two-way channel

- If have half-duplex constraint and memoryless channels, time-share
- If have full-duplex - obtain two parallel clean channels

For applications - full duplex gains a lot!

Aside - could this be what we need?

\[
I(X^N; Y^N) = I(Y^N; X^N)
\]

BUT

\[
I(X^n \rightarrow Y^n) \neq I(Y^n \rightarrow X^n)
\]

Usage: capacity of two-way channels

The capacity region of two-way channel is given by the limit as \( L \rightarrow \infty \) of the regions

\[
R_1 = I(A_1^L \rightarrow Y_2^L | X_1^L)
\]

\[
R_2 = I(A_2^L \rightarrow Y_1^L | X_1^L)
\]

where \( A_1^L, A_2^L \) are adaptive independent codewords.

Difficulty: space of codewords hard to compute!

Take away points - Discrete memoryless two-way channel

- If have half-duplex constraint ("push-to-talk"), time-share
- If have parallel two-way channels, mod-2 adder
- If have restricted channel

\[
R_1 \leq I(X_1; Y_1 | X_2)
\]

\[
R_2 \leq I(X_2; Y_2 | X_1)
\]

where \( X_1 \) and \( X_2 \) follow the joint distribution \( p(x_1, x_2) = p(x_1)p(x_2) \).

In general may need adaptive codewords

In general OPEN PROBLEM
Outline

• Information theory - what, why, when
Source coding, channel coding, entropy and mutual information, capacity, Gaussian noise channel

• Two-way channel - channel coding
Adaptive codewords, capacity in Gaussian noise = two parallel channels

• Wireless channels and networks

• Two-way relay channels - canonical example of wireless network coding

NO! Motivation 1: two-way channels

Two-way source coding = what skipping

• Two-way (lossy) source coding using block-coding and protocols of K rounds

• Interactive communication framework / communication complexity

Point-to-point

• Channel capacity ✓

• How to approach it for memoryless Gaussian noise channels ✓

Is that the end of the story?

NO! Motivation 2 - networks
Two-way networks!

- Multi-user two-way

Capacity regions

- Outer bound
- Capacity region
- Achievable region

\[
R_1 \leq I(X_1;Y|X_2) \\
R_2 \leq I(X_2;Y|X_1) \\
R_1 + R_2 \leq I(X_1, X_2; Y)
\]

- Prove that error is bounded away from 0 when \( \oplus \) not satisfied
- Find a more capable channel whose capacity is known
- Be creative!

Achievable rate region

- Propose a coding scheme (random codes!)
- Prove that as long as \( \oplus \) holds, reliable communication possible

\[
R_1 \leq I(X_1;Y|X_2) \\
R_2 \leq I(X_2;Y|X_1) \\
R_1 + R_2 \leq I(X_1, X_2; Y)
\]

Capacity regions

- Limit of communication, NOT how to achieve it in practice necessarily!
- However, benchmark and guidance in practical designs
Three key multi-user channels

- Broadcast channel
- Relay channel
- Multiple-access channel

Multiple-access channel (MAC)

- Capacity region is the closure of the convex hull of all rate pairs (R1, R2) satisfying

\[
R_1 \leq I(X_1; Y | X_2) \\
R_2 \leq I(X_2; Y | X_1) \\
R_1 + R_2 \leq I(X_1, X_2; Y)
\]

for some distribution

\[p(x_1, x_2, y) = p(x_1)p(x_2)p(y|x_1, x_2)\]

Multiple access channel

\[
\frac{1}{2} \log(1 + P_2) \\
\frac{1}{2} \log\left(\frac{1+P_2}{1+P_1}\right) \\
\frac{1}{2} \log\left(\frac{1+P_1+P_2}{1+P_1}\right) \\
\frac{1}{2} \log(1 + P_1)
\]

Multiple-access channel (MAC)

\[p(x_1, x_2) = p(x_1)p(x_2)\]

\[p(y|x_1, x_2)\]

- Introduced by Shannon in 1961
- Capacity known for discrete and Gaussian noise channels
  - Capacity [Ahlswede ’71, Liao ’72]
  - MIMO [Telatar ’99]
  - Fading [Gallager ’94, Shamai-Wyner ’97, Tse-Hanly ’98]

Gaussian MAC

Optimal multiple access

- TDM/FDM multiple access

\[
R_1 = \frac{1}{2} \log \left(1 + \frac{P_1}{N_0}\right) \\
R_2 = \frac{1}{2} \log \left(1 + \frac{P_2}{N_0}\right) \\
R_1 + R_2 = \frac{1}{2} \log \left(1 + P_1 + P_2\right)
\]

Multiple access channels in practice

- GSM
- CDMA
- WiFi
Broadcast channel

- Introduced by Cover in 1972
- Capacity known special cases:
  - Degraded broadcast channels [Bergmans ’73,74, Gallager ’74]
  - General BC with degraded message sets [Kramer + Marton ’77]
  - Gaussian MIMO broadcast channel [Weingarten, Steinberg, Shamai ’06]
- Best achievable rate region [Marton ’79]

Best achievable rate region: Marton’s region

- Capacity region is the closure of the convex hull of all rate pairs \((R_1, R_2)\) satisfying
  
  \[
  0 \leq R_1 \leq I(U_1; Y_1)
  
  0 \leq R_2 \leq I(U_2; Y_2)
  
  R_1 + R_2 \leq I(U_1; Y_1) + I(U_2; Y_2) - I(U_1; U_2)
  
  \]

  for some distribution

  \[
  p(u_1, u_2, x, y_1, y_2) = p(u_1, u_2)p(x|u_1, u_2)p(y_1, y_2|x)
  \]

Relay channel

- Introduced by Van der Meulen in 1968
- Capacity known special cases:
  - Physically degraded relay channels [Cover, El Gamal ’79]
- 3 types of forwarding:
  - Decode-forward (DF), Compress-forward (CF), Amplify-forward (AF)

Gaussian broadcast channel

\[
R_1 < \frac{1}{2} \log_2 \left( 1 + \frac{(1 - \alpha)P}{\sigma_1^2} \right) \quad \frac{1}{2} \log_2 \left( 1 + \frac{\alpha P}{(1 - \alpha)P + \sigma_2^2} \right)
\]

Relay channel: outer bounds

- Cut set outer bound intuitive + useful!
- Point-to-point

\[
R_{out} = \bigcup_{p(x)} I(X; Y)
\]

- Relay channel

\[
R_{out} = \bigcup_{p(x, x_1)} \min \{ I(X; Y, Y| X_1), I(X, X_1; Y) \}
\]
Relay channel: Decode + forward (DF)

\[ R \leq I(X; Y|X_1) \]
\[ R \leq I(X_1; Y) + I(X; Y|X_1) = I(X, X_1; Y) \]
\[ R \leq \max_{p(x,x_1)} \min\{I(X; Y|X_1), I(X, X_1; Y)\} \]

- Exploit broadcast transmission of source
- Source + relay transmit simultaneously (full duplex)
- Create joint codebooks at source + relay

Outline

- Information theory - what, why, when
  - Source coding, channel coding, entropy and mutual information, capacity, Gaussian noise channel
- Two-way channel - channel coding
  - Adaptive codewords, capacity in Gaussian noise + two parallel channels
- Wireless channels and networks
  - MAC, relay, BC channels
  - Two-way cellular-like networks
  - Two-way relay channels - canonical example of wireless network coding

Two-way MAC/BC channels

- account for feedback / data transmission tradeoff in a MAC/BC channel

MAC with feedback

No feedback

With feedback

MAC with feedback results

- Discrete Memoryless MAC
  - capacity region with feedback not known
  - with perfect feedback [Cover+Leung '81], [Bross+Lapidoth '05]
  - with generalized feedback [Carleial '82], [Kim '78]
  - noisy feedback bounds [Gastpar+Kramer '06]

- Gaussian MAC
  - perfect feedback - capacity is known [Ozarow '84]
  - noisy feedback bounds [Gastpar+Kramer '06]
Feedback allows the sources to correlate

- $R_1 = R_2 = R$ for $P_f > P$
- What is minimum $P_f$ to have a rate gain?
- Cooperative info = $\log(1 + P_f) - (1 - q) \log(1 + P_f)$

Two-way framework

- half-duplex nodes
- MAC phase - time $nT$
- BC phase - time $(1 - n)T$
  
  - what if BC phase used only for feedback?
  - power $P_f = P_i$ is used to feedback

Two-way is necessary!

- If have data in both directions, then the joint encoding of up-down links outperforms separate treatment of up-down links!
- combine feedback and new information!
Outline

- Information theory - what, why, when
  Source coding, channel coding, entropy and mutual information, capacity, Gaussian noise channel
- Two-way channel - channel coding
  Adaptive codewords, capacity in Gaussian noise = two parallel channels
- Wireless channels and networks
  MAC, relay, BC channels
- Two-way cellular-like networks
  Tradeoff between forward information and feedback
- Two-way relay channels - canonical example of wireless network coding
  - Single flow, single relay
  - Multiple flows, single relay
  - Single flow, multiple relays

Two-way relay channel

Motivation

Telesurgery
**Full-duplex**


**Half-duplex**


**All through the relay**


**Relaying type**

- **Decompress and forward**
- **Amplify and forward**

**Multiple terminals**

- **Temporal “phases”: who transmits when**

**Are 4 phases needed? NO!**

**Two-way relay channel: half-duplex**

Nodes can either transmit or receive, but not both.

**Better protocol**

![Diagram](image-url)
Message-level network coding

One possible approach we could take:

- Transmission Strategy:
  - Phase 1: Node a sends $w_a$.
  - Phase 2: Node b sends $w_b$.
  - Phase 3: Node r (the relay) sends $w_a \oplus w_b$.

- Decoding
  - Node a computes $w_a \oplus (w_a \oplus w_b) = w_b$, and
  - Node b computes $w_b \oplus (w_a \oplus w_b) = w_a$.

- Is this the best strategy?

Key exploits

- "own message side information" at nodes used to cancel out own message
- "overheard side information" available to nodes when not transmitting
- broadcast nature of wireless channels: relay broadcasts one thing, both nodes hear it.

(i) DT: Direct Transmission

(ii) MABC: Multiple Access Broadcast Channel
Relaying schemes

Amplify and forward (AF)

- The relay sends a scaled version of the signal it receives.
- Very little computation is needed.

\[ x_r = w_a Y_a^{(1)} \]

Decode and forward (DF)

- The relay decodes both \( w_a \) and \( w_b \).
- Much computation, and transmitter codebooks are needed at the relay.

\[ x_r = w_a \oplus w_b \]
### Achievable rate regions: one example

#### Mixed Forward (MF)

- The relay decodes $w_a$ and compresses $w_b$, combines them into a new message $w_r$ according to a bijective function, which it encodes and transmits.

![Mixed Forward (MF)](image)

- Theorem 1: The capacity region of the half-duplex bi-directional relay channel with the MABC protocol is the union of

$$R_a < \min \left\{ \Delta_1 I(X_a^{(1)}; Y_r^{(1)}|X_b^{(1)}), Q), \Delta_2 I(X_a^{(2)}; Y_r^{(2)}|Q) \right\}$$

$$R_b < \min \left\{ \Delta_1 I(X_b^{(1)}; Y_r^{(1)}|X_a^{(1)}), Q), \Delta_2 I(X_b^{(2)}; Y_r^{(2)}|Q) \right\}$$

$$R_a + R_b < \Delta_1 I(X_a^{(1)}, X_b^{(1)}; Y_r^{(1)}|Q)$$

over all joint distributions $p(q)p_r^{(1)}(x_a|q)p_r^{(1)}(x_b|q)p_r^{(2)}(x_r|q)$ with $|Q| \leq 5$.

#### Outer bounds: cut-set bound

If the rates $\{R^{(i)}\}$ are achievable with a protocol $P$ and $R_{\Sigma}(S \rightarrow S')$ denotes the total rate of independent information sent from set $S$ to set $S'$ then for all sets $S$:

$$R_{\Sigma}(S \rightarrow S') \leq \sum_i \Delta_i I(X_i^{(i)}; Y_r^{(i)}|X_i^{(i)}), Q).$$

---

### Achievable rate regions: an example

<table>
<thead>
<tr>
<th>Protocol</th>
<th>Side information</th>
<th>Phase</th>
<th>Interference</th>
</tr>
</thead>
<tbody>
<tr>
<td>MABC</td>
<td>not present</td>
<td>2</td>
<td>present</td>
</tr>
<tr>
<td>TDBC</td>
<td>present</td>
<td>3</td>
<td>not present</td>
</tr>
<tr>
<td>HBC</td>
<td>present</td>
<td>4</td>
<td>present</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Relaying</th>
<th>Complexity</th>
<th>Noise</th>
<th>Relay needs</th>
</tr>
</thead>
<tbody>
<tr>
<td>AF</td>
<td>very low</td>
<td>carried</td>
<td>full codebooks</td>
</tr>
<tr>
<td>DF</td>
<td>high</td>
<td>eliminated</td>
<td>p(y_a)</td>
</tr>
<tr>
<td>CF</td>
<td>low</td>
<td>distortion</td>
<td>a codebook, p(y_b)</td>
</tr>
</tbody>
</table>

#### Comparison of protocols

- The relay compresses/quantizes the received signal.
- Less computation than DF and transmitter codebooks are not needed at the relay.

![Comparison of protocols](image)

- **Theorem 1**: The capacity region of the half-duplex bi-directional relay channel with the MABC protocol is the union of

$$R_a < \min \left\{ \Delta_1 I(X_a^{(1)}; Y_r^{(1)}|X_b^{(1)}), Q), \Delta_2 I(X_a^{(2)}; Y_r^{(2)}|Q) \right\}$$

$$R_b < \min \left\{ \Delta_1 I(X_b^{(1)}; Y_r^{(1)}|X_a^{(1)}), Q), \Delta_2 I(X_b^{(2)}; Y_r^{(2)}|Q) \right\}$$

$$R_a + R_b < \Delta_1 I(X_a^{(1)}, X_b^{(1)}; Y_r^{(1)}|Q)$$

over all joint distributions $p(q)p_r^{(1)}(x_a|q)p_r^{(1)}(x_b|q)p_r^{(2)}(x_r|q)$ with $|Q| \leq 5$.

**Outer bounds: cut-set bound**

If the rates $\{R^{(i)}\}$ are achievable with a protocol $P$ and $R_{\Sigma}(S \rightarrow S')$ denotes the total rate of independent information sent from set $S$ to set $S'$ then for all sets $S$:

$$R_{\Sigma}(S \rightarrow S') \leq \sum_i \Delta_i I(X_i^{(i)}; Y_r^{(i)}|X_i^{(i)}), Q).$$

---

### Achievable rate regions: an example

- The relay compresses/quantizes the received signal.
- Less computation than DF and transmitter codebooks are not needed at the relay.

![Achievable rate regions: an example](image)
Consider two nodes which wish to exchange messages with the help of a single relay: one can think it difficult to evaluate expressions involved. Directed information-based capacity characterizations have played a crucial role in the development of capacity theorems for feedback channels \[\text{[\ldots]}\] and general two-way communications. Unfortunately, the mutual information is sufficient or optimal for two-way communications.

The capacity of the two-way channel \(C_{\text{XOR}}\) is given by one of the following formulas:

\[
C_{\text{XOR}} = C_{\text{XOR}}(a, b) = \min \left\{ I(X_a; Y_b), I(X_b; Y_a) \right\}
\]

This observation sparked the introduction of Decode and Forward (DF) coding, which provides low minimum distance. In the quadrature-phase direct transmission between the two terminal nodes over two time phases, network coding is used in the third phase. The decoding at the receiving terminal takes place in the fourth phase. It has been shown that the straightforward application of one-way metrics and techniques is not necessarily sufficient or optimal for two-way communications.

When two nodes wish to exchange messages using the help of a single relay, one can think it difficult to evaluate expressions involved. Directed information-based capacity characterizations have played a crucial role in the development of capacity theorems for feedback channels and general two-way communications. Unfortunately, the mutual information is sufficient or optimal for two-way communications.

**Recent developments**

- Capacity is known to within a constant \# of bits in Gaussian noise
  

- Constellation design for two-way relaying is considered in
  

**Relation to network coding?**

- bit-level / packet level network coding \(\rightarrow\) **Decode and Forward (DF)**

  - excellent systems-level demonstration of 2-way relaying gains (all layers, actual testbed)

- physical / analog network coding \(\rightarrow\) similar to **Amplify and Forward (AF)**

  - excellent systems-level demonstration of analog network coding (all layers, actual testbed)

---


Gaussian simulations

\[ h_{ar} = h_{br} = 1, \ h_{ab} = 0.2, \ N = 1, \text{ and } P = 50 \text{ dB}. \]

--

Simulations for the Gaussian noise channel

---

Gaussian simulations

--

Recent developments

- Capacity is known to within a constant \# of bits in Gaussian noise


- Constellation design for two-way relaying is considered in


**Relation to network coding?**

- bit-level / packet level network coding \(\rightarrow\) **Decode and Forward (DF)**

  - excellent systems-level demonstration of 2-way relaying gains (all layers, actual testbed)

- physical / analog network coding \(\rightarrow\) similar to **Amplify and Forward (AF)**

  - excellent systems-level demonstration of analog network coding (all layers, actual testbed)

---

Key exploits

• "own message side information" at nodes used to cancel out own message
• "overheard side information" available to nodes when not transmitting
• broadcast nature of wireless channels: relay broadcasts one thing, both nodes hear it.

Outline

• Information theory - what, why, when
• Two-way channel - channel coding
  Adaptive codewords, capacity in Gaussian noise = two parallel channels
• Wireless channels and networks
  MAC, relay, BC channels
• Two-way cellular-like networks
  Tradeoff between forward information and feedback
• Two-way relay channels - canonical example of wireless network coding
  • Single flow, single relay
  • Multiple flows, single relay
  • Single flow, multiple relays

Open questions

• codeword adaptation?

Motivation

Block 1

Block 2

Block 3

• Will it be used beyond academic demonstrations?

Multiple terminals

Multiple terminals

Arbitrary (m) number of end users
To be presented at ISIT 2010:

Multiple terminals

End user

Node

End user

W_0.0

W_0.1

W_0.2

W_0.3

Base-station

Relay

W_1.0

W_1.1

W_1.2

W_1.3

W_2.0

W_2.1

W_2.2

W_2.3

W_3.0

W_3.1

W_3.2

W_3.3

Half-duplex nodes

Decode + forward relay

Compress + forward end user cooperation

Per-flow network coding of messages at relay

“Protocols” = time “phases”

Phase 1 = MAC phase

Phase 2 = BC phase

Protocol 1: FMABC (Full MAC then BC)

Phase 1 = MAC phase

Phase 2 = BC phase

Protocol 2: PMABC (Partial MAC then BC)

Phase 1

Phase 2

Phase 3

Phase 4 = BC phase

Protocol 3: FTDBC (Full Time Division then BC)

Phase 1

Phase 2

Phase 3

Phase 4

Phase 5
Which protocol is “better”?

1. Extended Marton’s region for broadcasting
2. Per-flow network coding
3. Random-binning to exploit side-information
4. Terminal node cooperation
Per-flow Network coding (N)

\[ w_r = w_1 \oplus w_2 \]

\[ w_{r,1} = w_{1,0} \oplus w_{0,1} \]
\[ w_{r,2} = w_{2,0} \oplus w_{0,2} \]

\[ x_r(w_{r,1}, w_{r,2}) \]

Example of per-flow Network coding

**FMABC**

\[ R_{0,1} \leq \Delta_1 I(U_1^{(0)}, Y_1^{(2)}) \]
\[ R_{0,0} \leq \Delta_2 I(U_2^{(0)}, Y_2^{(2)}) \]
\[ R_{1,0} + R_{2,0} + R_{0,1} \leq \Delta_3 I(U_0^{(0)}, Y_0^{(2)}) + I(U_2^{(2)}, Y_2^{(2)}) - I(U_0^{(0)}, Y_0^{(1)}) \]
\[ R_{1,0} + R_{2,0} + R_{0,1} \leq \Delta_2 I(U_2^{(2)}, Y_2^{(2)}) - I(U_1^{(2)}, Y_1^{(2)}) \]

**FMABC-N**

\[ R_{0,1} \leq \Delta_1 I(U_1^{(0)}, Y_1^{(2)}) \]
\[ R_{0,0} \leq \Delta_2 I(U_2^{(0)}, Y_2^{(2)}) \]
\[ R_{1,0} + R_{2,0} + R_{0,1} \leq \Delta_3 I(U_0^{(0)}, Y_0^{(2)}) + I(U_2^{(2)}, Y_2^{(2)}) - I(U_0^{(0)}, Y_0^{(1)}) \]
\[ R_{1,0} + R_{2,0} + R_{0,1} \leq \Delta_2 I(U_2^{(2)}, Y_2^{(2)}) - I(U_1^{(2)}, Y_1^{(2)}) \]

Random binning (R) for exploiting overheard information

\[ R_{1,0} \leq \Delta_1 I(X_1^{(1)}; Y_r^{(1)}) \]
\[ R_{1,0} \leq \Delta_1 I(X_1^{(1)}; Y_0^{(1)}) + \Delta_3 I(X_r^{(3)}, Y_0^{(3)}) \]
\[ R_{0,1} \leq \Delta_2 I(X_0^{(2)}, Y_1^{(2)}) \]

Random binning with network coding

\[ \Delta_1 \]
\[ \Delta_2 \]
\[ \Delta_3 \]

Random binning and Marton’s binning

**Which protocol is “better”?**

1. Extended Marton’s region for broadcasting
2. Per-flow network coding
3. Random-binning to exploit side-information
4. Terminal node cooperation

**FMABC**

\[ R_{0,1} \leq \Delta_1 I(U_1^{(0)}, Y_1^{(2)}) \]
\[ R_{0,0} \leq \Delta_2 I(U_2^{(0)}, Y_2^{(2)}) \]
\[ R_{1,0} + R_{2,0} + R_{0,1} \leq \Delta_3 I(U_0^{(0)}, Y_0^{(2)}) + I(U_2^{(2)}, Y_2^{(2)}) - I(U_0^{(0)}, Y_0^{(1)}) \]
\[ R_{1,0} + R_{2,0} + R_{0,1} \leq \Delta_2 I(U_2^{(2)}, Y_2^{(2)}) - I(U_1^{(2)}, Y_1^{(2)}) \]

**FMABC-N**

\[ R_{0,1} \leq \Delta_1 I(U_1^{(0)}, Y_1^{(2)}) \]
\[ R_{0,0} \leq \Delta_2 I(U_2^{(0)}, Y_2^{(2)}) \]
\[ R_{1,0} + R_{2,0} + R_{0,1} \leq \Delta_3 I(U_0^{(0)}, Y_0^{(2)}) + I(U_2^{(2)}, Y_2^{(2)}) - I(U_0^{(0)}, Y_0^{(1)}) \]
\[ R_{1,0} + R_{2,0} + R_{0,1} \leq \Delta_2 I(U_2^{(2)}, Y_2^{(2)}) - I(U_1^{(2)}, Y_1^{(2)}) \]
Which protocol is “better”?

1. Extended Marton’s region for broadcasting
2. Per-flow network coding
3. Random-binning to exploit side-information
4. Terminal node cooperation

Cooperation (C) between terminal nodes

PMABC - NRC

Transmit

Process

Compress

Decompress

Phase 1

Phase 2

Phase 3

slot k

slot k+1

Cooperation (C) between terminal nodes

PMABC - NRC

Transmit

Process

Compress

Decompress

Phase 1

Phase 2

Phase 3

Cooperation (C) between terminal nodes

PMABC - NRC

Transmit

Process

Compress

Decompress

Phase 1

Phase 2

Phase 3

Cooperation (C) between terminal nodes

PMABC - NRC

Transmit

Process

Compress

Decompress

Phase 1

Phase 2

Phase 3

Cooperation (C) between terminal nodes

PMABC - NRC

Transmit

Process

Compress

Decompress

Phase 1

Phase 2

Phase 3
In this section, we evaluate the achievable rate regions and bounds are variations of cut-set outer bounds along the line of Theorem 17. Furthermore, we can significantly improve the value of the minimum data rate for all choices of the joint distribution in for example the high or low SNR regimes, which is the subject of current investigation[20].

Theorem 20: There are no inclusion relationships between FMABC, PMABC, and asymmetric channel.

A. Achievable rate region comparisons

Simple - MB - MB+NR

Simulations in Gaussian noise

$$Y[k] = HX[k] + Z[k]$$

$$H_1 = \begin{bmatrix} 0 & 0.3 & 0.05 & 1 \\ 0.3 & 0 & 1.5 & 1 \\ 0.05 & 1.5 & 0 & 0.2 \\ 1 & 1 & 0.2 & 0 \end{bmatrix}$$

$$H_2 = \begin{bmatrix} 0 & 0.9 & 0.4 & 1 \\ 0 & 0 & 0.02 & 1 \\ 0.02 & 0 & 0.5 \\ 1 & 1 & 0.5 & 0 \end{bmatrix}$$

FIGURE 10. Comparison with $\Delta r_n = F_1 = F_2 = F_3 = 0.5$, $H = H_2$.

B. PTDBC protocol

The main outcome is that different protocols are optimal under different information flow in each data link. Without this limitation, the PTDBC protocol is outer bounded by the set of all points described next.

$$R_{0_1} + R_{0_2} \leq \Delta I(X_0^{(1)}, Y_0^{(2)}; X_1^{(1)}, X_0^{(1)})$$

$$R_{1_0} + R_{2_0} \leq \Delta I(X_0^{(2)}, Y_0^{(3)}; X_0^{(3)}, X_0^{(3)})$$

$$R_{2_0} \leq \Delta I(X_0^{(3)}, X_1^{(3)}, Y_0^{(3)}; X_0^{(3)})$$

$$R_{0_1} \leq \Delta I(X_0^{(3)}, Y_0^{(3)}; X_0^{(3)})$$

$$R_{0_2} \leq \Delta I(X_0^{(3)}, Y_0^{(3)}; X_0^{(3)})$$

Evaluate and optimize over

- phase durations
- correlation matrices of Marton binning RVs subject to power constraints
- compression parameters

N = Network coding
R = Random binning
C = Cooperation between terminals

<table>
<thead>
<tr>
<th>Protocol</th>
<th>Multiple Access</th>
<th>Marton’s Broadcast</th>
<th>Network coding</th>
<th>Random binning</th>
<th>User cooperation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple</td>
<td>X</td>
<td>X</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>FMABC</td>
<td>X</td>
<td>X</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>FMABC-N</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>PMABC-N</td>
<td>X</td>
<td>X</td>
<td>–</td>
<td>X</td>
<td>–</td>
</tr>
<tr>
<td>PMABC-NR</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>FTDBC</td>
<td>–</td>
<td>X</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>FTDBC-NR</td>
<td>–</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>–</td>
</tr>
<tr>
<td>FTDBC-NR</td>
<td>–</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>PTDBC</td>
<td>X</td>
<td>X</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>PTDBC-NR</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

Table 1

<table>
<thead>
<tr>
<th>Protocol</th>
<th>Multiple Access</th>
<th>Marton’s Broadcast</th>
<th>Network coding</th>
<th>Random binning</th>
<th>User cooperation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple</td>
<td>X</td>
<td>X</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>FMABC</td>
<td>X</td>
<td>X</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>FMABC-N</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>PMABC-N</td>
<td>X</td>
<td>X</td>
<td>–</td>
<td>X</td>
<td>–</td>
</tr>
<tr>
<td>PMABC-NR</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

FIGURE 9. Comparison with $R_0 = R_1 = R_2 = R_3 = 0.4$ dB, $H = H_2$.

Simulations in Gaussian noise

$$Y[k] = HX[k] + Z[k]$$

FIGURE 7. Comparison with $F_1 = F_2 = F_3 = 0.5$, $H = H_2$.

FIGURE 8. Comparison with $F_1 = F_2 = F_3 = 0.4$ dB, $H = H_2$.

FIGURE 11. Comparison with $F_1 = F_2 = F_3 = 0.4$ dB, $H = H_2$.

FIGURE 12. Comparison with $F_1 = F_2 = F_3 = 0.4$ dB, $H = H_2$.

Network coding + random binning

FMABC

PMABC

FTDBC
January 29, 2010

C. Cooperation coding gain

In Fig. 12, we fixed the data rate with and without cooperation. To show the cooperation coding gain, we plot the achievable rates in Fig. 13.

Multi-flow take-away points

- Most schemes use per-flow network coding
- Significantly more complex: protocols and opportunities abound, only starting to understand when to do what.
- Due to practical relevance - crucial to develop insight and thereafter demonstrations (there are none)

Other recent multi-flow developments

- In a recent multi-hop network model, no base-station or direct link is assumed. Capacity is obtained in high-SNR linear deterministic model; capacity to within 2 bits/sec/Hz for general Gaussian channel.
- A generalized multi-way cluster model is explored; bounds on the symmetric rates are obtained.
- The relay has multiple antennas and the delay-limited diversity-multiplexing tradeoff is explored, showing that the multiplexing gain may be improved (over fully connected interference case).

Multi-flow take-away points

- Most schemes use per-flow network coding
- Significantly more complex: protocols and opportunities abound. Only starting to understand when to do what.

Multi-flow take-away points

- Most schemes use per-flow network coding
- Significantly more complex: protocols and opportunities abound, only starting to understand when to do what.
- Two-way nature must be explicitly accounted for (side-information, ability to network code and broadcast) in order to see gains.
### Outline

- Information theory - what, why, when
  - Source coding, channel coding, entropy and mutual information, capacity, Gaussian noise channel
- Two-way channel - channel coding
  - Adaptive codewords, capacity in Gaussian noise - two parallel channels
- Wireless channels and networks
  - MAC, relay, BC channels
- Two-way cellular-like networks
  - Tradeoff between forward information and feedback
- Two-way relay channels - canonical example of wireless network coding
  - Single flow, single relay
  - Multiple flows, single relay
  - Single flow, multiple relays

### Assumptions

- Half-duplex nodes - protocols in time
- Decode and Forward (DF) and Amplify and Forward (AF) only
- What rates are achievable, what protocols are best?

### Protocol 2: (m,3) TDBC

**Duration Δ₁**
- Allows for direct transmissions from a to b (side-information)
  - If r₁ decodes w₁ and w₂ : x₁(w₁, w₂)
  - If r₁ decodes w₁ only : x₁(w₁)
  - If r₁ decodes w₂ only : x₂(w₂)
  - Otherwise, r₁ is silent

**Duration Δ₂**
- ... (diagram showing nodes and arrows)

**Duration Δ₃**
- ... (diagram showing nodes and arrows)

### Protocol 3: (m, m+2) Multi-hop Multi-relay

**Better!**
- Does not combine messages in 2 directions!

---

### Multiple relays

- **Protocol 1: (m,2) MABC**
  - Duration Δ₁
  - Duration Δ₂
  - If r₁ decodes w₁ and w₂ : x₁(w₁, w₂)
  - If r₁ decodes w₁ only : x₁(w₁)
  - If r₁ decodes w₂ only : x₂(w₂)
  - Otherwise, r₁ is silent.

- **Protocol 3: (m, m+2) MHMR**
  - ... (diagram showing nodes and arrows)
**Initialization**

**Before communication**

- $w_a(0)$, $w_a(1)$, $w_a(2)$
- $w_b(0)$, $w_b(1)$, $w_b(2)$

**After initialization**

- $w_a(0)$, $w_a(1)$, $w_a(2)$
- $w_b(0)$, $w_b(1)$, $w_b(2)$

**After termination**

- $w_a(0)$, $w_b(1)$, $w_b(2)$
- $w_a(0)$, $w_a(1)$, $w_a(2)$

**Performance in Gaussian noise**

- In low SNR: (2,4) DF MHMR
- In high SNR: (2,4) DF MHMR or (2,2) AF MABC
Multi-relay take-aways

- numerical results indicate that \((m,m+2)\) MHMR protocol with information flowing in 2 directions yields the largest rates
- for low # of hops, or at high SNR, AF relaying may do well, but rapidly degrades as # hops increases or SNR decreases
- fundamentally unsolved, in part due to complexity of space.

Outline

- Information theory - what, why, when
- Two-way channel - channel coding
- Wireless channels and networks
- Two-way cellular-like networks
- Two-way relay channels - canonical example of wireless network coding
  - Single flow, single relay
  - Multiple flows, single relay
  - Single flow, multiple relays

Future areas of two-way channels

- one-way information theory “fairly” well understood
- advances in processing power
- never ending desire for bandwidth and limited wireless spectrum

Two-way wireless networks
Questions?

Natasha Devroye
Assistant Professor
University of Illinois at Chicago
http://www.ece.uic.edu/~devroye