NEW COMMUNICATION STRATEGIES FOR
BROADCAST AND INTERFERENCE NETWORKS

S. Sandeep Pradhan
(Joint work with Arun Padakandla and Aria Sahebi)

University of Michigan, Ann Arbor
Distributed Information Coding

- Proliferation of wireless data and sensor network applications
- Supported by distributed information processing
- Information-theoretic perspective
1: Distributed Field Gathering
2: Broadcast and Interference Networks

- Mobile Transmitter 1
- Mobile Transmitter 2
- Mobile Receiver 1
- Mobile Receiver 2
Information and Coding Theory: Tradition

Information Theory:

- Develop efficient communication strategies
- No constraints on memory/computation for encoding/decoding
- Obtain performance limits that are independent of technology
Information and Coding Theory: Tradition

Information Theory:
- Develop efficient communication strategies
- No constraints on memory/computation for encoding/decoding
- Obtain performance limits that are independent of technology

Coding Theory:
- Approach these limits using algebraic codes (Ex: linear codes)
- Fast encoding and decoding algorithms
- Objective: practical implementability of optimal communication systems
Information theory: Orders of magnitude

- Subatomic scale: $10^{-23} - 10^{-15}$ Physicists
- Atomic scale: $10^{-15} - 10^{-6}$ Chemists
Information theory: Orders of magnitude

- Subatomic scale: $10^{-23} - 10^{-15}$ Physicists
- Atomic scale: $10^{-15} - 10^{-6}$ Chemists
- Human scale: $10^{-6} - 10^6$ Biologists
- Astronomical scale: $10^6 - 10^{27}$ Astronomers
Information theory: Orders of magnitude

- Subatomic scale: $10^{-23} - 10^{-15}$ Physicists
- Atomic scale: $10^{-15} - 10^{-6}$ Chemists
- Human scale: $10^{-6} - 10^6$ Biologists
- Astronomical scale: $10^6 - 10^{27}$ Astronomers
- Information-theory scale: $10^n$, $n$ sufficiently large.
Probability versus Algebra

Information Theory Tools: based on probability

- Finding the optimal communication system directly is difficult
Information Theory Tools: based on probability

- Finding the optimal communication system directly is difficult
- Random Coding:
  - Build a collection of communication systems (ensemble)
  - Put a probability distribution on them
  - Show good average performance
  - Craft ensembles using probability
Information Theory Tools: based on probability

- Finding the optimal communication system directly is difficult
- Random Coding:
  - Build a collection of communication systems (ensemble)
  - Put a probability distribution on them
  - Show good average performance
  - Craft ensembles using probability

Coding Theory Tools: Abstract algebra (groups, fields)

- Exploit algebraic structure to develop algorithms of polynomial complexity for encoding/decoding
- Study a very small ensemble at a time.
Random Coding in Networks

- Prob. distribution on a collection of codebooks (ensemble)
- Extensions of Shannon ensembles
**Random Coding in Networks**

- Prob. distribution on a collection of codebooks (ensemble)
- Extensions of Shannon ensembles
- Lot of bad codebooks in the ensemble
- Average performance significantly affected by these bad codes
- Do not achieve optimality in general
- Many problems have remained open for decades.
It turns out that algebraic structure can be used to weed out bad codes
Coding theory to the rescue?

- It turns out that algebraic structure can be used to weed out bad codes.
- Gain barely noticeable in point-to-point communication.
  - Improvement in second order performance (error exponents).
Coding theory to the rescue?

- It turns out that algebraic structure can be used to weed out bad codes.
- Gain barely noticeable in point-to-point communication.
  - Improvement in second order performance (error exponents).
- Gains significant in multi-terminal communication.
It turns out that algebraic structure can be used to weed out bad codes.

- Gain barely noticeable in point-to-point communication
  - Improvement in second order performance (error exponents)

- Gains significant in multi-terminal communication

- Time for Question?
Broadcast Networks

Base

Station

Mobile Receiver 1

Mobile Receiver 2
Start with Binary symmetric channel

\[ N \sim Be(\delta), \text{ and } + \text{ is addition modulo 2} \]

\[ \text{Capacity } = \max_{P(X)} I(X; Y) = 1 - h(\delta). \]
Output is within a ball around a transmitted codeword

Maximum likelyhood decoding
Output is within a ball around a transmitted codeword

Maximum likelyhood decoding

Time for Question?
Twitter and Eddington Number

- Suppose you want to tweet on a BSC:
- 140 characters
Suppose you want to tweet on a BSC:

- 140 characters
- Entropy of tweets = 1.9 bits/character, ⇒ 266 bits.
- Suppose $\delta = 0.11$, then $C = 0.5$ bits/channel use
Twitter and Eddington Number

- Suppose you want to tweet on a BSC:
  - 140 characters
  - Entropy of tweets = 1.9 bits/character, $\Rightarrow 266 \text{ bits}$.
  - Suppose $\delta = 0.11$, then $C = 0.5 \text{ bits/channel use}$
  - A tweet can be sent by using BSC 532 times.
  - Number of possible tweets $= 2^{266}$
Twitter and Eddington Number

- Suppose you want to tweet on a BSC:
- 140 characters
- Entropy of tweets $= 1.9$ bits/character, $\Rightarrow 266$ bits.
- Suppose $\delta = 0.11$, then $C = 0.5$ bits/channel use
- A tweet can be sent by using BSC 532 times.
- Number of possible tweets $= 2^{266}$
- Equals the number of protons in the observable universe
- Named after Arthur Eddington.
BSC WITH COST CONSTRAINT

\[ \frac{1}{n} \mathbb{E} w_H(X^n) \leq q \]

i.e., a codeword has at most \( q \) fractions of 1's
**BSC with cost constraint**

1. \( \frac{1}{n} \mathbb{E} w_H(X^n) \leq q \)
2. i.e., a codeword has at most \( q \) fractions of 1's
3. Capacity-cost function

\[
C(q) = \max_{Ew_H(X) \leq q} I(X; Y) = H(Y) - H(Y|X) = h(q \delta) - h(\delta)
\]

4. \( q \delta = (1 - q)\delta + q(1 - \delta) \)
BSC with cost constraint

- \( \frac{1}{n} \mathbb{E} w_H(X^n) \leq q \)
- i.e., a codeword has at most \( q \) fractions of 1’s
- Capacity-cost function

\[
C(q) = \max_{\mathbb{E} w_H(X) \leq q} I(X; Y) = H(Y) - H(Y|X) = h(q \delta) - h(\delta)
\]

- \( q \delta = (1 - q)\delta + q(1 - \delta) \)
- \( X \sim Be(q) \)
Picture of an optimal code

- Big circle: the set of all words with $q$ fraction of 1's
BSC WITH COST CONSTRAINT AND INTERFERENCE

\[ S \sim Be(0.5) \text{ and } N \sim Be(\delta) \]

\[ S \text{ is non-causally observable only at encoder} \]
BSC WITH COST CONSTRAINT AND INTERFERENCE

- $S \sim \text{Be}(0.5)$ and $N \sim \text{Be}(\delta)$
- $S$ is non-causally observable only at encoder
- $\frac{1}{n} \mathbb{E} w_H(X^n) \leq q$
Applications

Digital watermarking, data hiding, covert communication

- Blind watermarking
Applications

Digital watermarking, data hiding, covert communication

- Blind watermarking
- You want big govt. but you don't trust it too much
Q1: What is the communication strategy?
Q1: What is the communication strategy?

A1. Try cancelling it
Q1: What is the communication strategy?

A1. Try cancelling it

   You cannot, you do not have enough number of ones.
Q1: What is the communication strategy?

A1. Try cancelling it
   - You cannot, you do not have enough number of ones.

A2. Ride on the interference
Q1: What is the communication strategy?

A1. Try cancelling it
   - You cannot, you do not have enough number of ones.

A2. Ride on the interference
   - Nudge the interference with channel input toward a codeword
   - But, you have got just $q$ fraction of ones.
BSC with cost constraint and interference

Q1: What is the communication strategy?

A1. Try cancelling it
   You cannot, you do not have enough number of ones.

A2. Ride on the interference
   Nudge the interference with channel input toward a codeword
   But, you have got just $q$ fraction of ones.
   Gelfand-Pinsker: Nudge toward a codeword from a set
Q1: What is the communication strategy?

A1. Try cancelling it
   - You cannot, you do not have enough number of ones.

A2. Ride on the interference
   - Nudge the interference with channel input toward a codeword
   - But, you have got just $q$ fraction of ones.
   - Gelfand-Pinsker: Nudge toward a codeword from a set

Q2. How large should the set be?

Rate of the set: $1 - h(q)$. 
Picture of an optimal set of code words
All these codewords are assigned for a message.
Picture of an Optimal Set of Codewords

- All these codewords are assigned for a message.
- Select a codeword to which you can nudge the interference.
- ..by spending just $q$ fraction of ones $\Rightarrow U = X + S$
All these codewords are assigned for a message

Select a codeword to which you can nudge the interference...

..by spending just $q$ fraction of ones $\Rightarrow U = X + S$

New effective channel: $Y = U + N$ with capacity $1 - h(\delta)$
Precoding for Interference

- Rate of the composite codebook: $1 - h(\delta)$
- Rate of a sub-code-book: $1 - h(q)$
- Transmission rate: difference $= h(q) - h(\delta)$
- Capacity in general case [Gelfand-Pinsker ’80]
Precoding for Interference

- Rate of the composite codebook: $1 - h(\delta)$
- Rate of a sub-code-book: $1 - h(q)$
- Transmission rate: difference $= h(q) - h(\delta)$
- Capacity in general case [Gelfand-Pinsker ’80]

$$C(q) = \max_{P(U,X|S):Ew_H(X) \leq q} I(U; Y) - I(U; S)$$
Bottomline: Rate loss as compared to no interference
Broadcast Channel: Cover ’72

- Channel with one input and multiple outputs
- Same signal should contain info. meant for both receivers
- Capacity region still not known in general
Channel with one input and multiple outputs

Same signal should contain info. meant for both receivers

Capacity region still not known in general

Time for questions?
Marton’s Coding Strategy: Two receivers

- Create a signal that carry information for the second receiver
Marton’s Coding Strategy: Two Receivers

- Create a signal that carry information for the second receiver
- This signal acts as interference for the signal of the first
- How to tackle (self) interference?
Marton’s Coding Strategy: Two receivers

- Create a signal that carry information for the second receiver
- This signal acts as interference for the signal of the first
- How to tackle (self) interference?
  - Make the first receiver decode a large portion of interference
  - This portion is given by a (univariate) function
Marton’s Coding Strategy: Two Receivers

- Create a signal that carry information for the second receiver
- This signal acts as interference for the signal of the first
- How to tackle (self) interference?
  - Make the first receiver decode a large portion of interference
  - This portion is given by a (univariate) function
  - The rest is precoded for using Gelfand-Pinsker strategy
Marton’s Coding Strategy: Two Receivers

- Create a signal that carry information for the second receiver
- This signal acts as interference for the signal of the first
- How to tackle (self) interference?
  - Make the first receiver decode a large portion of interference
  - This portion is given by a (univariate) function
  - The rest is precoded for using Gelfand-Pinsker strategy
- This strategy is optimal for many special cases
- We do not know whether it is optimal in general
**Example: so-called non-degraded channel**

- $N_1 \sim Be(\delta)$, and $N_2 \sim Be(\epsilon)$, and no constraint on $X_2$
- Hamming weight constraint on $X_1$: $\frac{1}{n} \mathbb{E} w_H(X_1^n) \leq q$
**Example: So-called non-degraded channel**

- $N_1 \sim \text{Be}(\delta)$, and $N_2 \sim \text{Be}(\epsilon)$, and no constraint on $X_2$
- Hamming weight constraint on $X_1$: $\frac{1}{n} \mathbb{E} w_H(X_1^n) \leq q$
- $\text{Fix } R_2 = 1 - h(\epsilon)$, and assume $\delta < \epsilon$
**Example: so-called non-degraded channel**

- $N_1 \sim Be(\delta)$, and $N_2 \sim Be(\epsilon)$, and no constraint on $X_2$
- Hamming weight constraint on $X_1$: $\frac{1}{n} \mathbb{E} w_H(X_1^n) \leq q$
- Fix $R_2 = 1 - h(\epsilon)$, and assume $\delta < \epsilon$
- When $q \cdot \delta \leq \epsilon$, Rec. 1 can decode interference completely
Example: So-called non-degraded channel

- \( N_1 \sim Be(\delta) \), and \( N_2 \sim Be(\epsilon) \), and no constraint on \( X_2 \)
- Hamming weight constraint on \( X_1 \): \( \frac{1}{n} \mathbb{E} w_H(X_1^n) \leq q \)
- Fix \( R_2 = 1 - h(\epsilon) \), and assume \( \delta < \epsilon \)
- When \( q \times \delta \leq \epsilon \), Rec. 1 can decode interference completely
  - a.k.a no interference \( \Rightarrow R_1 = h(q \times \delta) - h(\delta) \)
Example: So-called non-degraded channel

- $N_1 \sim Be(\delta)$, and $N_2 \sim Be(\epsilon)$, and no constraint on $X_2$
- Hamming weight constraint on $X_1$: $\frac{1}{n}Ew_H(X_1^n) \leq q$
- Fix $R_2 = 1 - h(\epsilon)$, and assume $\delta < \epsilon$
- When $q \cdot \delta \leq \epsilon$, Rec. 1 can decode interference completely
  - a.k.a no interference $\Rightarrow R_1 = h(q \cdot \delta) - h(\delta)$
- Otherwise, precode for $X_2$: $\Rightarrow R_1 = h(q) - h(\delta)$
Decide a \textit{univariate} function of interference & precode for the rest.
Broadcast with more receivers

- Marton’s strategy can be easily extended
- Consider 3 receiver case: At receiver 1:
Broadcast with more receivers

- Marton’s strategy can be easily extended
- Consider 3 receiver case: At receiver 1:
  - (self) interference of signals of Rec. 2 and Rec. 3
Broadcast with more receivers

- Marton’s strategy can be easily extended

- Consider 3 receiver case: At receiver 1:
  - (self) interference of signals of Rec. 2 and Rec. 3
  - Decode a univariate function of signal meant for Rec. 2...
  - .. and a univariate function of signal meant for Rec. 3.
Broadcast with more receivers

- Marton’s strategy can be easily extended
- Consider 3 receiver case: At receiver 1:
  - (self) interference of signals of Rec. 2 and Rec. 3
  - Decode a univariate function of signal meant for Rec. 2...
  - .. and a univariate function of signal meant for Rec. 3.
  - Precode for the rest
Broadcast with more receivers

- Marton’s strategy can be easily extended
- Consider 3 receiver case: At receiver 1:
  - (self) interference of signals of Rec. 2 and Rec. 3
  - Decode a univariate function of signal meant for Rec. 2...
  - .. and a univariate function of signal meant for Rec. 3.
  - Precode for the rest
- All these being done using random codes
- No need for linear or algebraic codes till now
Marton’s strategy can be easily extended.

Consider 3 receiver case: At receiver 1:

- (self) interference of signals of Rec. 2 and Rec. 3
- Decode a univariate function of signal meant for Rec. 2...
- .. and a univariate function of signal meant for Rec. 3.
- Precode for the rest

All these being done using random codes

No need for linear or algebraic codes till now

We can show that such a strategy is strictly suboptimal
**New Strategy**

- Decode a bivariate function of the signals meant for other two
NEW STRATEGY

- Decode a bivariate function of the signals meant for other two
- It turns out that to exploit this we need linear codes
**New Strategy**

- Decode a bivariate function of the signals meant for other two.
- It turns out that to exploit this we need linear codes.

\[ X_1 + X_2 + X_3 \rightarrow Y, N_1 \]
\[ X_2 + N_2 \rightarrow Z, N_2 \]
\[ X_3 + N_3 \rightarrow A, N_3 \]

- \( N_2, N_3 \sim Be(\epsilon) \), and no constraints on \( X_2 \) and \( X_3 \).
- \( N_1 \sim Be(\delta) \) and the usual: \( \frac{1}{n} \mathbb{E} w_H(X_1^n) \leq q \).
New Strategy

- Decode a bivariate function of the signals meant for other two
- It turns out that to exploit this we need linear codes

\[ X_1 + X_2 + X_3 \]

- \( N_2, N_3 \sim Be(\epsilon) \), and no constraints on \( X_2 \) and \( X_3 \)
- \( N_1 \sim Be(\delta) \) and the usual : \( \frac{1}{n} \mathbb{E} w_{H}(X_1^n) \leq q \)
- Let \( R_2 = R_3 = 1 - h(\epsilon) \), the incorrigible brutes!
- Let \( \delta < \epsilon \)
Deficiency of random codes

- \( \delta = 0.1 \) and \( \epsilon = 0.2 \)

\[
\begin{align*}
0.2781 & \quad 0.5310 & \quad 0.5562 \\
1 - h(\epsilon) & \quad 1 - h(\delta) & \quad 2(1 - h(\epsilon))
\end{align*}
\]
Deficiency of random codes

- $\delta = 0.1$ and $\epsilon = 0.2$

- Marton wishes to decode “full” interference: $(X_2, X_3)$:
  - $1 - h(q \ast \delta) > 2(1 - h(\epsilon))$
DEFICIENCY OF RANDOM CODES

- \( \delta = 0.1 \) and \( \epsilon = 0.2 \)

\[
\begin{align*}
0.2781 & \quad 0.5310 & \quad 0.5562 \\
1-h(\epsilon) & \quad 1-h(\delta) & \quad 2(1-h(\epsilon))
\end{align*}
\]

- Marton wishes to decode “full” interference: \((X_2, X_3)\):
  - \( 1 - h(q \ast \delta) > 2(1 - h(\epsilon)) \)
  - a.k.a never going to happen
  - Marton ends up doing precoding incurring rate loss
DEFICIENCY OF RANDOM CODES

- \( \delta = 0.1 \) and \( \epsilon = 0.2 \)

\[
\begin{array}{ccc}
0.2781 & & 0.5310 & 0.5562 \\
1-h(\epsilon) & & 1-h(\delta) & 2(1-h(\epsilon))
\end{array}
\]

- Marton wishes to decode “full” interference: \((X_2, X_3)\):
  - \( 1 - h(q \ast \delta) > 2(1 - h(\epsilon)) \)
  - a.k.a never going to happen
  - Marton ends up doing precoding incurring rate loss

- New Approach: Try decoding actual interference: \(X_2 + X_3\)
Deficiency of random codes

- \( \delta = 0.1 \) and \( \epsilon = 0.2 \)

\[
\begin{align*}
0.2781 & \quad 0.5310 & \quad 0.5562 \\
1 - h(\epsilon) & \quad 1 - h(\delta) & \quad 2(1 - h(\epsilon))
\end{align*}
\]

- Marton wishes to decode “full” interference: \((X_2, X_3)\):
  - \(1 - h(q \ast \delta) > 2(1 - h(\epsilon))\)
  - a.k.a never going to happen
  - Marton ends up doing precoding incurring rate loss

- New Approach: Try decoding *actual* interference: \(X_2 + X_3\)
  - Benefit if the range of \(X_2 + X_3\) is \(\ll\) range of \((X_2, X_3)\)
DEFICIENCY OF RANDOM CODES

- $\delta = 0.1$ and $\epsilon = 0.2$

- Marton wishes to decode “full” interference: $(X_2, X_3)$:
  - $1 - h(q \ast \delta) > 2(1 - h(\epsilon))$
  - a.k.a never going to happen
  - Marton ends up doing precoding incurring rate loss

- New Approach: Try decoding actual interference: $X_2 + X_3$
  - Benefit if the range of $X_2 + X_3$ is $\ll$ range of $(X_2, X_3)$
  - If $X_2$ and $X_3$ are “random”, this won’t happen
Picture of sum of two random sets

\[ \begin{array}{c}
\text{Set 1} \\
\text{Set 2} \\
\hline
\text{Sum}
\end{array} \]
Picture of sum of two cosets of a linear code
Exploits of Linear Codes

- The “incorrigible brutes” can have their capacities
Exploits of Linear Codes

- The “incorrigible brutes” can have their capacities
- We just need their codebooks to behave “algebraic”
- We know that linear codes achieve the capacity of BSC
Exploits of Linear Codes

- The “incorrigible brutes” can have their capacities
- We just need their codebooks to behave “algebraic”
- We know that linear codes achieve the capacity of BSC
- Rate of $X_2 = rate of X_3 = rate of X_2 + X_3 = 1 - h(\epsilon)$
- Since $\delta < \epsilon$, we have for small $q$: $q \cdot \delta < \epsilon$
Exploits of Linear Codes

- The “incorrigible brutes” can have their capacities
- We just need their codebooks to behave “algebraic”
- We know that linear codes achieve the capacity of BSC
- Rate of $X_2 = rate of X_3 = rate of X_2 + X_3 = 1 - h(\epsilon)$
- Since $\delta < \epsilon$, we have for small $q$: $q \cdot \delta < \epsilon$
- Hence $1 - h(q \cdot \delta) > 1 - h(\epsilon)$
- Rec. 1 can decode the actual interference and subtract it off
The “incorrigible brutes” can have their capacities

We just need their codebooks to behave “algebraic”

We know that linear codes achieve the capacity of BSC

rate of $X_2 = rate\ of\ X_3 = rate\ of\ X_2 + X_3 = 1 - h(\epsilon)$

Since $\delta < \epsilon$, we have for small $q$: $q \ast \delta < \epsilon$

Hence $1 - h(q \ast \delta) > 1 - h(\epsilon)$

Rec. 1 can decode the actual interference and subtract it off

Then decodes her message at rate $h(q \ast \delta) - h(\delta)$

$R_1 = h(q \ast \delta) - h(\delta)$, $R_2 = R_3 = 1 - h(\epsilon)$
We have banked on
Symmetry and addition saved the world

We have banked on

1. Channels of Rec. 2 and 3 are symmetric
   - so uniform input distribution achieves capacity
Symmetry and addition saved the world

We have banked on

- Channels of Rec. 2 and 3 are symmetric
  - so uniform input distribution achieves capacity
- Interference in the broadcast channel is additive
Symmetry and addition saved the world

We have banked on

- Channels of Rec. 2 and 3 are symmetric
  - so uniform input distribution achieves capacity
- Interference in the broadcast channel is additive

But Shannon theory is all about not getting bogged down in an example

- Objective is to develop a theory for general case
Caution: Even in point-to-point communication

In general, linear codes do not achieve Shannon capacity of an arbitrary discrete memoryless channel
However?

- Caution: Even in point-to-point communication
  - In general, linear codes do not achieve Shannon capacity of an arbitrary discrete memoryless channel
- What hope do we have in using them for network communication for the arbitrary discrete memoryless case?
Algebraic structure in codes may be necessary in a fundamental way.
Algebraic structure in codes may be necessary in a fundamental way.

Algebraic structure alone is not sufficient.

A right mix of algebraic structure along with non-linearity.

Nested algebraic code appears to be a universal structure.
Given: Channel I/P = X, O/P = Y, with $p_Y|X$, and cost function $w(x)$

Find: maximum transmission rate $R$ for a target cost $W$. 

Noisy Channel Coding in Point-to-point case
Noisy Channel Coding in Point-to-point case

- Given: Channel I/P = X, O/P = Y, with $p_{Y|X}$, and cost function $w(x)$
- Find: maximum transmission rate $R$ for a target cost $W$.
- Answer: Shannon Capacity-Cost function (Shannon ’49)

$$C(W) = \max_{p_X : E[w] \leq W} I(X; Y)$$
Picture of a near-optimal channel code

Obtained from Shannon ensemble

- Box = \( X^n \)
- Red dot = codeword
- \( C \) = code book
Obtained from Shannon ensemble

- Box = \( \mathcal{X}^n \)
- Red dot = codeword
- \( C \) = code book
- \( C \) has Packing Property
- \( C \) has Shaping Property
Picture of a near-optimal channel code

Obtained from Shannon ensemble

- Box = $\mathcal{X}^n$
- Red dot = codeword
- $\mathcal{C}$ = code book
- $\mathcal{C}$ has Packing Property
- $\mathcal{C}$ has Shaping Property
- Shape Region = Typical set
- Size of code = $I(X; Y)$
- Codeword density =
  
  $$I(X; Y) - H(X) = -H(X|Y)$$

Broadcast and interference
New Result: An optimal linear code

- Let $|\mathcal{X}| = p$, prime no.
- $C_1 =$ code book
- $C_1$ has Packing Property
- Size of code
  \[= \log |\mathcal{X}| - H(X|Y)\]
NEW RESULT: AN OPTIMAL LINEAR CODE

- Let $|\mathcal{X}| = p$, prime no.
- $C_1 = \text{code book}$
- $C_1$ has Packing Property
- Size of code
  $$= \log |\mathcal{X}| - H(X|Y)$$
- Finite field is $\mathbb{Z}_p$
- Bounding Region $= \mathcal{X}^n$
- Density $= -H(X|Y)$

Broadcast and interference
**New Theorem: An Optimal Nested Linear Code**

- $C_1$ fine code (red & black)
- $C_2$ coarse code (black)
- $C_1$ has Packing property

Going beyond symmetry
**New Theorem:** An Optimal Nested Linear Code

- $C_1$ fine code (red & black)
- $C_2$ coarse code (black)
- $C_1$ has Packing property
- $C_2$ has Shaping property
- Size of $C_1 = \log |\mathcal{X}| - H(X|Y)$
- Size of $C_2 = \log |\mathcal{X}| - H(X)$

Going beyond symmetry
**New Theorem: An Optimal Nested Linear Code**

- $C_1$ fine code (red & black)
- $C_2$ coarse code (black)
- $C_1$ has Packing property
- $C_2$ has Shaping property
- Size of $C_1 = \log |\mathcal{X}| - H(X|Y)$
- Size of $C_2 = \log |\mathcal{X}| - H(X)$
- Code book $= C_1/C_2$
- Code book size $= I(X; Y)$
- Achieves $C(W)$

Going beyond symmetry
GOING BEYOND ADDITION

- $X_2 \lor X_3$ (logical OR function)
Going beyond addition

- $X_2 \lor X_3$ (logical OR function)
- What kind of glasses you wear so this looks like addition?
Going beyond addition

- $X_2 \vee X_3$ (logical OR function)
- What kind of glasses you wear so this looks like addition?
- Can be embedded in the addition table in $\mathbb{F}_3$

\[
\begin{array}{c|c|c}
0 & 1 & 2 \\
\hline
0 & 0 & 1 \\
1 & 1 & 2 \\
2 & 2 & 0 \\
\end{array}
\]
Going beyond addition

- $X_2 \lor X_3$ (logical OR function)
- What kind of glasses you wear so this looks like addition?
- Can be embedded in the addition table in $\mathbb{F}_3$

\[
\begin{array}{ccc}
0 & 0 & 1 & 2 \\
0 & & & \\
1 & 1 & 2 & 0 \\
2 & 2 & 0 & 1 \\
\end{array}
\]

- Map binary sources into $\mathbb{F}_3$, and use linear codes built on $\mathbb{F}_3$
- Can do better than traditional random coding
**Going beyond addition**

- $X_2 \lor X_3$ (logical OR function)
- What kind of glasses you wear so this looks like addition?
- Can be embedded in the addition table in $\mathbb{F}_3$

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

- Map binary sources into $\mathbb{F}_3$, and use linear codes built on $\mathbb{F}_3$
- Can do better than traditional random coding
- In general we ‘embed’ bivariate functions in groups
Groups - An Introduction

- $G$ - a finite abelian group of order $n$
- $G \cong \mathbb{Z}_{p_1^{e_1}} \times \mathbb{Z}_{p_2^{e_2}} \cdots \times \mathbb{Z}_{p_k^{e_k}}$
- $G$ isomorphic to direct product of possibly repeating primary cyclic groups

$$g \in G \iff g = (g_1, \ldots, g_k), \ g_i \in \mathbb{Z}_{p_i^{e_i}}$$

- Call $g_i$ as the $i$th digit of $g$
Groups - An Introduction

- $G$ - a finite abelian group of order $n$
- $G \cong \mathbb{Z}_{p_1^{e_1}} \times \mathbb{Z}_{p_2^{e_2}} \times \cdots \times \mathbb{Z}_{p_k^{e_k}}$
- $G$ isomorphic to direct product of possibly repeating primary cyclic groups

$$g \in G \iff g = (g_1, \ldots, g_k), \ g_i \in \mathbb{Z}_{p_i^{e_i}}$$

- Call $g_i$ as the $i$th digit of $g$
- Prove coding theorems for primary cyclic groups
Nested Group Codes

- Group code over \( \mathbb{Z}_{p^r}^n \): \( \mathcal{C} < \mathbb{Z}_{p^r}^n \)
- \( \mathcal{C} = \text{Image}(\phi) \) for some homomorphism \( \phi: \mathbb{Z}_{p^r}^k \rightarrow \mathbb{Z}_{p^r}^n \)
Nested Group Codes

- Group code over $\mathbb{Z}_{p^r}^n$: $C < \mathbb{Z}_{p^r}^n$
- $C = \text{Image}(\phi)$ for some homomorphism $\phi: \mathbb{Z}_{p^r}^k \rightarrow \mathbb{Z}_{p^r}^n$
- $(C_1, C_2)$ nested if $C_2 \subset C_1$
Nested Group Codes

- Group code over $\mathbb{Z}_{p^r}^n$: $C < \mathbb{Z}_{p^r}^n$
- $C = \text{Image}(\phi)$ for some homomorphism $\phi: \mathbb{Z}_{p^r}^k \rightarrow \mathbb{Z}_{p^r}^n$
- $(C_1, C_2)$ nested if $C_2 \subset C_1$
- We need:
  - $C_1 < \mathbb{Z}_{p^r}^n$: “good” packing code
  - $C_2 < \mathbb{Z}_{p^r}^n$: “good” covering code
Good Group Packing Codes

- Good group channel code $C_2$ for the triple $(\mathcal{U}, \mathcal{V}, P_{UV})$
- Assume $\mathcal{U} = \mathbb{Z}_{pr}$ for some prime $p$ and exponent $r > 0$
Good Group Packing Codes

- Good group channel code $C_2$ for the triple $(U, V, P_{UV})$
- Assume $U = \mathbb{Z}_{pr}$ for some prime $p$ and exponent $r > 0$

**Lemma**

 Exists for large $n$ if

$$\frac{1}{n} \log |C_2| \leq \log p^r - \max_{0 \leq i < r} \left( \frac{r}{r-i} \right) \left( H(U|V) - H([U]_i|V) \right)$$
Good Group Packing Codes

- Good group channel code $C_2$ for the triple $(U, V, P_{UV})$
- Assume $U = \mathbb{Z}_{pr}$ for some prime $p$ and exponent $r > 0$

**Lemma**

Exists for large $n$ if

$$\frac{1}{n} \log |C_2| \leq \log p^r - \max_{0 \leq i < r} \left( \frac{r}{r-i} \right) (H(U|V) - H([U]_i|V))$$

- $[U]_i$ is a function of $U$ and depends on the group
- Extra penalty for imposing group structure beyond linearity
Good Group Packing Codes

- Good group channel code $C_2$ for the triple $(U, V, P_{UV})$
- Assume $U = \mathbb{Z}_{p^r}$ for some prime $p$ and exponent $r > 0$

Lemma

 Exists for large $n$ if

$$\frac{1}{n} \log |C_2| \leq \log p^r - \max_{0 \leq i < r} \left( \frac{r}{r-i} \right) (H(U|V) - H([U]_i|V))$$

- $[U]_i$ is a function of $U$ and depends on the group
- Extra penalty for imposing group structure beyond linearity
- Time for questions?
Encoders observe different components of a vector source

Central decoder receives quantized observations from the encoders

Given source distribution $p_{XYZ}$

Best known rate region - Berger-Tung Rate Region, ’77
Conclusions

- Presented a nested group codes based coding scheme
- Can recover known rate regions of broadcast channel
- Offers rate gains over random coding coding scheme
Conclusions

- Presented a nested group codes based coding scheme
- Can recover known rate regions of broadcast channel
- Offers rate gains over random coding scheme
- New bridge between probability and algebra, between information theory and coding theory
CONCLUSIONS

- Presented a nested group codes based coding scheme
- Can recover known rate regions of broadcast channel
- Offers rate gains over random coding coding scheme
- New bridge between probability and algebra, between information theory and coding theory
- It was thought that probability and algebra are nemesis
Conclusions

- Presented a nested group codes based coding scheme
- Can recover known rate regions of broadcast channel
- Offers rate gains over random coding coding scheme
- New bridge between probability and algebra, between information theory and coding theory
- It was thought that probability and algebra are nemesis
- Instead the match made in heaven