Lectures on

Network Information Theory

Abbas El Gamal

Stanford University

Allerton 2009

The Early Years

- I started a course on multiple user (network) information theory at Stanford in 1982 and taught it 3 times

The Early Years

- I started a course on multiple user (network) information theory at Stanford in 1982 and taught it 3 times
- The course had some of today's big names in our field:





Syllabus Circa 1983

COURSE OUTLINE

1. Review of basic entropy and mutual information relations.
2. The Asymptotic Equipartition Property, joint typicality.
3. Shannon's source and channel coding theorem.
4. Multiple access channel.
5. The Slepian-Wolf source coding theorem.
6. The multiple access channel with correlated sources.
7. The broadcast channel.
8. Basic rate distortion theory.
9. Source coding with side information.
10. Channel with states.
11. The One Bit Theorem.

Syllabus Circa 1983

- I also gave a lecture on feedback

Syllabus Circa 1983

- I also gave a lecture on feedback
- Some results that were known then and are considered important today were absent:

  Interference channel: Strong interference; Han–Kobayashi

  Relay channel: cutset bound; decode–forward; compress–forward

  Multiple descriptions: El Gamal–Cover; Ozarow; Ahlswede

  Secrecy: Shannon; Wyner; Csiszár–Körner

Syllabus Circa 1983

- I also gave a lecture on feedback
- Some results that were known then and are considered important today were absent:

  Interference channel: Strong interference; Han–Kobayashi

  Relay channel: cutset bound; decode–forward; compress–forward

  Multiple descriptions: El Gamal–Cover; Ozarow; Ahlswede

  Secrecy: Shannon; Wyner; Csiszár–Körner

- There was no theoretical or practical interest in these results then

The Dog Years of NIT

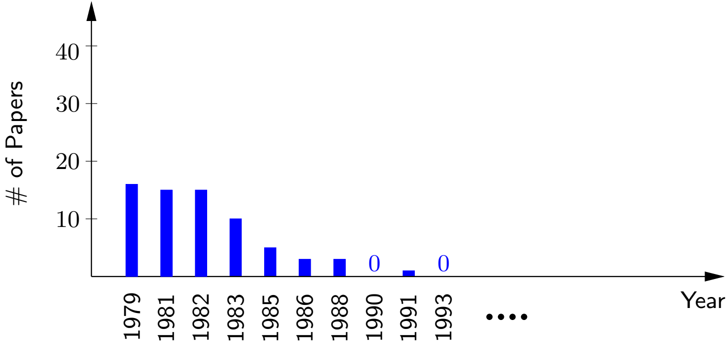
- By the mid 80s interest in NIT was all but gone

The Dog Years of NIT

- By the mid 80s interest in NIT was all but gone
- Theory was stuck and many basic problems remained open
- It seemed that the theory will have no applications

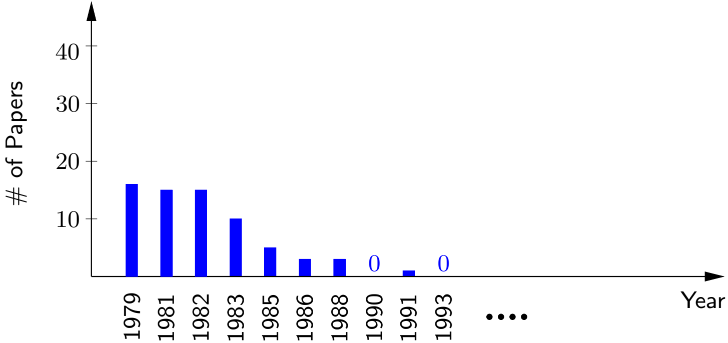
The Dog Years of NIT

- By the mid 80s interest in NIT was all but gone
- Theory was stuck and many basic problems remained open
- It seemed that the theory will have no applications
- By early 90s, the number of ISIT papers on NIT $\to 0$:



The Dog Years of NIT

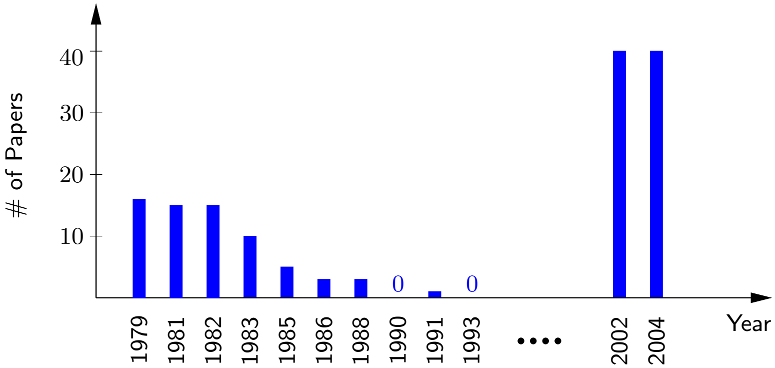
- By the mid 80s interest in NIT was all but gone
- Theory was stuck and many basic problems remained open
- It seemed that the theory will have no applications
- By early 90s, the number of ISIT papers on NIT $\to 0$:



- I stopped teaching the course and moved on to other things

The Resurgence

- By late 90s, the Internet and wireless communication began to revive interest in NIT; and by early 2000s, the field was in full swing



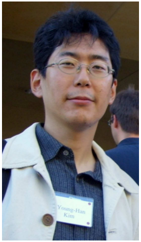
The Resurgence

- By late 90s, the Internet and wireless communication began to revive interest in NIT; and by early 2000s, the field was in full swing
- I started teaching the course again in 2002

The Resurgence

- By late 90s, the Internet and wireless communication began to revive interest in NIT; and by early 2000s, the field was in full swing
- I started teaching the course again in 2002
- The course had some of today's rising stars:



Syllabus Circa 2002

Rough Schedule

Sept 26	Introduction, review
Oct 1	Review
Oct 3	Review, MAC
Oct 8	No lecture
Oct 10	MAC, correlated sources
Oct 15	Correlated sources
Oct 17	Broadcast
Oct 22	Broadcast, interference
Oct 24	Interference, relay
Oct 29	Relay, feedback
Oct 31	Rate distortion
Nov 5	Rate distortion
Nov 7	Side information, multiple descriptions
Nov 12	Channel with state (Project selections)
Nov 14	Channel with state

Technology from Early 80s to 2002

- Chip technology: Scaled by a factor of 2^{11} (Moore's law)

Technology from Early 80s to 2002

- Chip technology: Scaled by a factor of 2^{11} (Moore's law)
- Computing: From VAX780 to PCs and laptops

Technology from Early 80s to 2002

- Chip technology: Scaled by a factor of 2^{11} (Moore's law)
- Computing: From VAX780 to PCs and laptops
- Communication: From 1200 Baud modems and wired phones to DSL, cellular, and 802.xx

Technology from Early 80s to 2002

- Chip technology: Scaled by a factor of 2^{11} (Moore's law)
- Computing: From VAX780 to PCs and laptops
- Communication: From 1200 Baud modems and wired phones to DSL, cellular, and 802.xx
- Networks: From ARPANET to the Internet

Technology from Early 80s to 2002

- Chip technology: Scaled by a factor of 2^{11} (Moore's law)
- Computing: From VAX780 to PCs and laptops
- Communication: From 1200 Baud modems and wired phones to DSL, cellular, and 802.xx
- Networks: From ARPANET to the Internet
- Multi-media: From film cameras and Sony Walkman to digital cameras and iPod

What's Wrong with This Picture?

What's Wrong with This Picture?

- Theory does not advance as fast as technology

What's Wrong with This Picture?

- Theory does not advance as fast as technology

- Nothing happened between early 80s and 2002

What's Wrong with This Picture?

- Theory does not advance as fast as technology

- Nothing happened between early 80s and 2002

- I didn't know what was going on

What's Wrong with This Picture?

- Theory does not advance as fast as technology

- Nothing happened between early 80s and 2002

- I didn't know what was going on

Answer: All of the above

What Happened Since Mid 80s?

- Some progress on old open problems (mainly Gaussian)

What Happened Since Mid 80s?

- Some progress on old open problems (mainly Gaussian)
- Work on new models: Fading channels; MIMO; secrecy, ...

What Happened Since Mid 80s?

- Some progress on old open problems (mainly Gaussian)
- Work on new models: Fading channels; MIMO; secrecy, . . .
- New directions in network capacity:

  Network coding

  Scaling laws

  Deterministic/high SNR approximations (within xx bits)

What Happened Since Mid 80s?

- Some progress on old open problems (mainly Gaussian)
- Work on new models: Fading channels; MIMO; secrecy, . . .
- New directions in network capacity:

  Network coding

  Scaling laws

  Deterministic/high SNR approximations (within xx bits)
- Attempts to consummate marriage (or at least dating) between IT and networking

Lectures on NIT: 2009

- Developed jointly with Young-Han Kim of UCSD
- Incorporate many of the recent results

Lectures on NIT: 2009

- Developed jointly with Young-Han Kim of UCSD
- Incorporate many of the recent results
- Attempt to organize the field in a “top-down” way

Lectures on NIT: 2009

- Developed jointly with Young-Han Kim of UCSD
- Incorporate many of the recent results
- Attempt to organize the field in a "top-down" way
- Balance introduction of new techniques and new models

Lectures on NIT: 2009

- Developed jointly with Young-Han Kim of UCSD
- Incorporate many of the recent results
- Attempt to organize the field in a "top-down" way
- Balance introduction of new techniques and new models
- Unify, simplify, and formalize achievability proofs

Lectures on NIT: 2009

- Developed jointly with Young-Han Kim of UCSD
- Incorporate many of the recent results
- Attempt to organize the field in a "top-down" way
- Balance introduction of new techniques and new models
- Unify, simplify, and formalize achievability proofs
- Emphasize extension to networks

Lectures on NIT: 2009

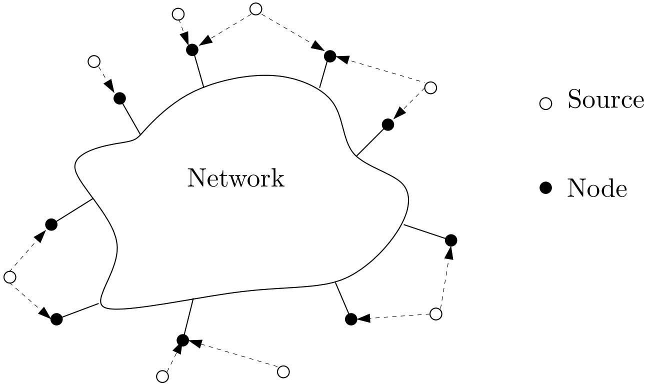
- Developed jointly with Young-Han Kim of UCSD
- Incorporate many of the recent results
- Attempt to organize the field in a “top-down” way
- Balance introduction of new techniques and new models
- Unify, simplify, and formalize achievability proofs
- Emphasize extension to networks
- Use clean and unified notation and terminology

Outline

1. The First Lecture
2. Achievability for DM Sources and Channels
3. Gaussian Sources and Channels
4. Converse
5. Extension to Networks
6. Conclusion

Network Information Flow

- Consider a general networked information processing system:



- Sources: data, speech, music, images, video, sensor data
- Nodes: handsets, base stations, servers, sensor nodes
- Network: wired, wireless, or hybrid

Network Information Flow

- Each node observes some sources, wishes to obtain descriptions of other sources, or to compute function/make decision based on them

Network Information Flow

- Each node observes some sources, wishes to obtain descriptions of other sources, or to compute function/make decision based on them
- To achieve the goal, the nodes communicate and perform local computing

Network Information Flow

- Each node observes some sources, wishes to obtain descriptions of other sources, or to compute function/make decision based on them
- To achieve the goal, the nodes communicate and perform local computing
- Information flow questions:

  What are the necessary and sufficient conditions on information flow?

  What are the optimal schemes/techniques needed to achieve them?

Network Information Flow

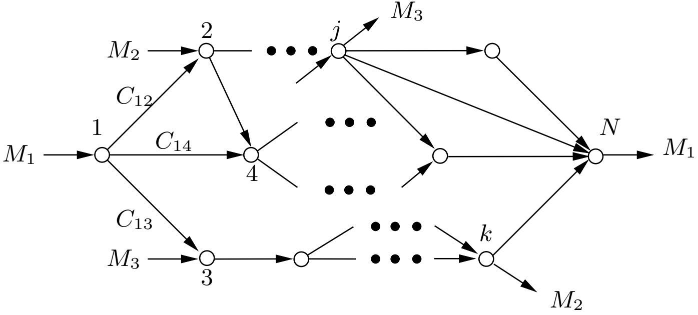
- Each node observes some sources, wishes to obtain descriptions of other sources, or to compute function/make decision based on them
- To achieve the goal, the nodes communicate and perform local computing
- Information flow questions:

  What are the necessary and sufficient conditions on information flow?

  What are the optimal schemes/techniques needed to achieve them?
- The difficulty in answering these questions depends on:
  - Source and network models
  - Information processing goals
  - Computational capabilities of the nodes

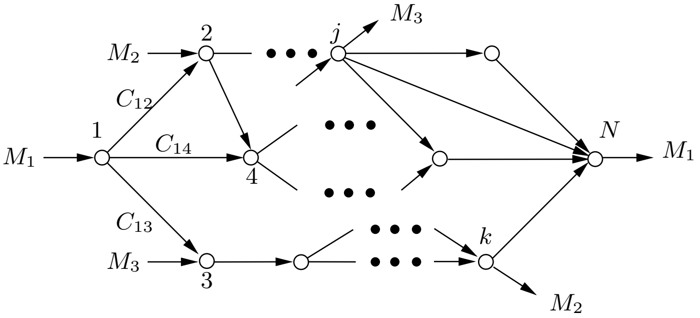
Example: Multi-Commodity Flow

- If the sources are commodities with demands (rates in bits/sec); the nodes are connected by noiseless rate-constrained links; each intermediate node forwards the bits it receives; the goal is to send each commodity to a destination node; the problem reduces to the multi-commodity flow with known conditions on optimal flow



Example: Multi-Commodity Flow

- If the sources are commodities with demands (rates in bits/sec); the nodes are connected by noiseless rate-constrained links; each intermediate node forwards the bits it receives; the goal is to send each commodity to a destination node; the problem reduces to the multi-commodity flow with known conditions on optimal flow



- For single commodity, these conditions reduce to the celebrated max-flow min-cut theorem

Network Information Theory

- This simple networked information processing system model does not capture many important aspects of real-world systems:
  - Real-world information sources have redundancies, time and space correlations, time variations

Network Information Theory

- This simple networked information processing system model does not capture many important aspects of real-world systems:
  - Real-world information sources have redundancies, time and space correlations, time variations
  - Real-world networks may suffer from noise, interference, node/link failures, delay, time variation

Network Information Theory

- This simple networked information processing system model does not capture many important aspects of real-world systems:
  - Real-world information sources have redundancies, time and space correlations, time variations
  - Real-world networks may suffer from noise, interference, node/link failures, delay, time variation
  - Real-world networks may allow for broadcasting

Network Information Theory

- This simple networked information processing system model does not capture many important aspects of real-world systems:
  - Real-world information sources have redundancies, time and space correlations, time variations
  - Real-world networks may suffer from noise, interference, node/link failures, delay, time variation
  - Real-world networks may allow for broadcasting
  - Real-world communication nodes may allow for more complex node operations than forwarding

Network Information Theory

- This simple networked information processing system model does not capture many important aspects of real-world systems:
  - Real-world information sources have redundancies, time and space correlations, time variations
  - Real-world networks may suffer from noise, interference, node/link failures, delay, time variation
  - Real-world networks may allow for broadcasting
  - Real-world communication nodes may allow for more complex node operations than forwarding
  - The goal in many information processing systems is to partially recover the sources or to compute/make a decision

Network Information Theory

- This simple networked information processing system model does not capture many important aspects of real-world systems:
  - Real-world information sources have redundancies, time and space correlations, time variations
  - Real-world networks may suffer from noise, interference, node/link failures, delay, time variation
  - Real-world networks may allow for broadcasting
  - Real-world communication nodes may allow for more complex node operations than forwarding
  - The goal in many information processing systems is to partially recover the sources or to compute/make a decision
- Network information theory aims to answer the information flow questions while capturing essential elements of real-world networks in the probabilistic framework of Shannon's information theory

State of the Theory

- Focus has been on compression and communication for discrete memoryless (DM) and Gaussian sources and channels

State of the Theory

- Focus has been on compression and communication for discrete memoryless (DM) and Gaussian sources and channels
- Most results are for separate source–channel settings

State of the Theory

- Focus has been on compression and communication for discrete memoryless (DM) and Gaussian sources and channels
- Most results are for separate source–channel settings
- Computable characterizations of capacity/optimal rate regions known for few cases. For other cases, only inner and outer bounds are known

State of the Theory

- Focus has been on compression and communication for discrete memoryless (DM) and Gaussian sources and channels
- Most results are for separate source–channel settings
- Computable characterizations of capacity/optimal rate regions known for few cases. For other cases, only inner and outer bounds are known
- Some results on joint source–channel coding, communication for computing, secrecy, and in intersection with networking

State of the Theory

- Focus has been on compression and communication for discrete memoryless (DM) and Gaussian sources and channels
- Most results are for separate source–channel settings
- Computable characterizations of capacity/optimal rate regions known for few cases. For other cases, only inner and outer bounds are known
- Some results on joint source–channel coding, communication for computing, secrecy, and in intersection with networking
- Coding techniques developed, e.g., superposition, successive cancellation, Slepian–Wolf, Wyner–Ziv, successive refinement, dirty paper coding, network coding are starting to impact real-world networks

State of the Theory

- Focus has been on compression and communication for discrete memoryless (DM) and Gaussian sources and channels
- Most results are for separate source–channel settings
- Computable characterizations of capacity/optimal rate regions known for few cases. For other cases, only inner and outer bounds are known
- Some results on joint source–channel coding, communication for computing, secrecy, and in intersection with networking
- Coding techniques developed, e.g., superposition, successive cancellation, Slepian–Wolf, Wyner–Ziv, successive refinement, dirty paper coding, network coding are starting to impact real-world networks
- However, many basic problems remain open and a complete theory is yet to be developed

Outline of Lectures

- Lectures aim to provide broad coverage of the models, fundamental results, proof techniques, and open problems in NIT

Outline of Lectures

- Lectures aim to provide broad coverage of the models, fundamental results, proof techniques, and open problems in NIT
- Include both teaching material and advanced results

Outline of Lectures

- Lectures aim to provide broad coverage of the models, fundamental results, proof techniques, and open problems in NIT
- Include both teaching material and advanced results
- Divided into four parts:

  Part I: Background

  Part II: Single-hop Networks

  Part III: Multi-hop Networks

  Part IV: Extensions

Outline of Lectures

- Lectures aim to provide broad coverage of the models, fundamental results, proof techniques, and open problems in NIT
- Include both teaching material and advanced results
- Divided into four parts:

  Part I: Background

  Part II: Single-hop Networks

  Part III: Multi-hop Networks

  Part IV: Extensions
- Global appendices for general techniques and background, e.g., bounding cardinalities of auxiliary random variables and Fourier–Motzkin elimination

Part: I Background

- Purpose: Introduce notation and basic techniques used throughout; point out some differences between point-to-point and multiple user communication

Part: I Background

- Purpose: Introduce notation and basic techniques used throughout; point out some differences between point-to-point and multiple user communication
- Entropy, differential entropy, and mutual information

Part: I Background

- Purpose: Introduce notation and basic techniques used throughout; point out some differences between point-to-point and multiple user communication
- Entropy, differential entropy, and mutual information
- Strong typicality: Orlitsky–Roche definition; properties

Part: I Background

- Purpose: Introduce notation and basic techniques used throughout; point out some differences between point-to-point and multiple user communication
- Entropy, differential entropy, and mutual information
- Strong typicality: Orlitsky–Roche definition; properties
- Key achievability lemmas:
  - Typical average lemma
  - Joint typicality lemma
  - Packing lemma
  - Covering lemma
  - Conditional typicality lemma

Part: I Background

- Purpose: Introduce notation and basic techniques used throughout; point out some differences between point-to-point and multiple user communication
- Entropy, differential entropy, and mutual information
- Strong typicality: Orlitsky–Roche definition; properties
- Key achievability lemmas:
  - Typical average lemma
  - Joint typicality lemma
  - Packing lemma
  - Covering lemma
  - Conditional typicality lemma
- Shannon's point-to-point communication theorems: Random coding; joint typicality encoding/decoding

Part II: Single-hop Networks

- Single round one-way communication

Part II: Single-hop Networks

- Single round one-way communication

  Independent messages over noisy channels:

  Correlated sources over noiseless (wireline) channels:

  Correlated sources over DM channels:

Part II: Single-hop Networks

- Single round one-way communication

  Independent messages over noisy channels:

  - Multiple access channels: time sharing; successive cancellation

  Correlated sources over noiseless (wireline) channels:

  Correlated sources over DM channels:

Part II: Single-hop Networks

- Single round one-way communication

  Independent messages over noisy channels:

  - Multiple access channels: time sharing; successive cancellation
  - Degraded broadcast channels: superposition coding

  Correlated sources over noiseless (wireline) channels:

  Correlated sources over DM channels:

Part II: Single-hop Networks

- Single round one-way communication

  Independent messages over noisy channels:

  - Multiple access channels: time sharing; successive cancellation
  - Degraded broadcast channels: superposition coding
  - Interference channels: strong interference; Han–Kobayashi

  Correlated sources over noiseless (wireline) channels:

  Correlated sources over DM channels:

Part II: Single-hop Networks

- Single round one-way communication

  Independent messages over noisy channels:
  - Multiple access channels: time sharing; successive cancellation
  - Degraded broadcast channels: superposition coding
  - Interference channels: strong interference; Han–Kobayashi
  - Channels with state: Gelfand–Pinsker; writing on dirty paper

  Correlated sources over noiseless (wireline) channels:

  Correlated sources over DM channels:

Part II: Single-hop Networks

- Single round one-way communication

  Independent messages over noisy channels:

  - Multiple access channels: time sharing; successive cancellation
  - Degraded broadcast channels: superposition coding
  - Interference channels: strong interference; Han–Kobayashi
  - Channels with state: Gelfand–Pinsker; writing on dirty paper
  - Fading channels: alternative performance measures (outage capacity)

  Correlated sources over noiseless (wireline) channels:

  Correlated sources over DM channels:

Part II: Single-hop Networks

- Single round one-way communication

  Independent messages over noisy channels:

  - Multiple access channels: time sharing; successive cancellation
  - Degraded broadcast channels: superposition coding
  - Interference channels: strong interference; Han–Kobayashi
  - Channels with state: Gelfand–Pinsker; writing on dirty paper
  - Fading channels: alternative performance measures (outage capacity)
  - General broadcast channels: Marton coding; mutual covering

  Correlated sources over noiseless (wireline) channels:

  Correlated sources over DM channels:

Part II: Single-hop Networks

- Single round one-way communication

  Independent messages over noisy channels:

  - Multiple access channels: time sharing; successive cancellation
  - Degraded broadcast channels: superposition coding
  - Interference channels: strong interference; Han–Kobayashi
  - Channels with state: Gelfand–Pinsker; writing on dirty paper
  - Fading channels: alternative performance measures (outage capacity)
  - General broadcast channels: Marton coding; mutual covering
  - Vector Gaussian channels: dirty paper coding; MAC–BC duality

  Correlated sources over noiseless (wireline) channels:

  Correlated sources over DM channels:

Part II: Single-hop Networks

- Single round one-way communication

  Independent messages over noisy channels:

  - Multiple access channels: time sharing; successive cancellation
  - Degraded broadcast channels: superposition coding
  - Interference channels: strong interference; Han–Kobayashi
  - Channels with state: Gelfand–Pinsker; writing on dirty paper
  - Fading channels: alternative performance measures (outage capacity)
  - General broadcast channels: Marton coding; mutual covering
  - Vector Gaussian channels: dirty paper coding; MAC–BC duality

  Correlated sources over noiseless (wireline) channels:

  - Distributed lossless source coding: Slepian–Wolf; random binning

  Correlated sources over DM channels:

Part II: Single-hop Networks

- Single round one-way communication

  Independent messages over noisy channels:

  - Multiple access channels: time sharing; successive cancellation
  - Degraded broadcast channels: superposition coding
  - Interference channels: strong interference; Han–Kobayashi
  - Channels with state: Gelfand–Pinsker; writing on dirty paper
  - Fading channels: alternative performance measures (outage capacity)
  - General broadcast channels: Marton coding; mutual covering
  - Vector Gaussian channels: dirty paper coding; MAC–BC duality

  Correlated sources over noiseless (wireline) channels:

  - Distributed lossless source coding: Slepian–Wolf; random binning
  - Source coding with side information: Wyner–Ziv

  Correlated sources over DM channels:

Part II: Single-hop Networks

- Single round one-way communication

  Independent messages over noisy channels:

  - Multiple access channels: time sharing; successive cancellation
  - Degraded broadcast channels: superposition coding
  - Interference channels: strong interference; Han–Kobayashi
  - Channels with state: Gelfand–Pinsker; writing on dirty paper
  - Fading channels: alternative performance measures (outage capacity)
  - General broadcast channels: Marton coding; mutual covering
  - Vector Gaussian channels: dirty paper coding; MAC–BC duality

  Correlated sources over noiseless (wireline) channels:

  - Distributed lossless source coding: Slepian–Wolf; random binning
  - Source coding with side information: Wyner–Ziv
  - Distributed lossy source coding: Berger–Tung; quadratic Gaussian

  Correlated sources over DM channels:

Part II: Single-hop Networks

- Single round one-way communication

  Independent messages over noisy channels:
  - Multiple access channels: time sharing; successive cancellation
  - Degraded broadcast channels: superposition coding
  - Interference channels: strong interference; Han–Kobayashi
  - Channels with state: Gelfand–Pinsker; writing on dirty paper
  - Fading channels: alternative performance measures (outage capacity)
  - General broadcast channels: Marton coding; mutual covering
  - Vector Gaussian channels: dirty paper coding; MAC–BC duality

  Correlated sources over noiseless (wireline) channels:
  - Distributed lossless source coding: Slepian–Wolf; random binning
  - Source coding with side information: Wyner–Ziv
  - Distributed lossy source coding: Berger–Tung; quadratic Gaussian
  - Multiple descriptions: El Gamal–Cover; successive refinement

  Correlated sources over DM channels:

Part II: Single-hop Networks

- Single round one-way communication

  Independent messages over noisy channels:

  - Multiple access channels: time sharing; successive cancellation
  - Degraded broadcast channels: superposition coding
  - Interference channels: strong interference; Han–Kobayashi
  - Channels with state: Gelfand–Pinsker; writing on dirty paper
  - Fading channels: alternative performance measures (outage capacity)
  - General broadcast channels: Marton coding; mutual covering
  - Vector Gaussian channels: dirty paper coding; MAC–BC duality

  Correlated sources over noiseless (wireline) channels:

  - Distributed lossless source coding: Slepian–Wolf; random binning
  - Source coding with side information: Wyner–Ziv
  - Distributed lossy source coding: Berger–Tung; quadratic Gaussian
  - Multiple descriptions: El Gamal–Cover; successive refinement

  Correlated sources over DM channels:

  Separation does not hold in general; common information; sufficient conditions for MAC, BC

Part III: Multi-hop Networks

- Relaying and multiple communication rounds

Part III: Multi-hop Networks

- Relaying and multiple communication rounds

  Independent messages over noiseless networks:

  Independent messages over noisy networks:

  Correlated sources over noiseless (wireline) channels:

Part III: Multi-hop Networks

- Relaying and multiple communication rounds

Independent messages over noiseless networks:

Max-flow min-cut theorem; network coding

Independent messages over noisy networks:

Correlated sources over noiseless (wireline) channels:

Part III: Multi-hop Networks

- Relaying and multiple communication rounds

Independent messages over noiseless networks:

Max-flow min-cut theorem; network coding

Independent messages over noisy networks:

- Relay channel: cutset bound; decode–forward; compress–forward

Correlated sources over noiseless (wireline) channels:

Part III: Multi-hop Networks

- Relaying and multiple communication rounds

  Independent messages over noiseless networks:

  Max-flow min-cut theorem; network coding

  Independent messages over noisy networks:

  - Relay channel: cutset bound; decode–forward; compress–forward
  - Interactive communication: feedback capacity; iterative refinement

  Correlated sources over noiseless (wireline) channels:

Part III: Multi-hop Networks

- Relaying and multiple communication rounds

Independent messages over noiseless networks:

Max-flow min-cut theorem; network coding

Independent messages over noisy networks:

- Relay channel: cutset bound; decode–forward; compress–forward
- Interactive communication: feedback capacity; iterative refinement
- DM networks: cutset bound; decode–forward; compress–forward

Correlated sources over noiseless (wireline) channels:

Part III: Multi-hop Networks

- Relaying and multiple communication rounds

  Independent messages over noiseless networks:

  Max-flow min-cut theorem; network coding

  Independent messages over noisy networks:

  - Relay channel: cutset bound; decode–forward; compress–forward
  - Interactive communication: feedback capacity; iterative refinement
  - DM networks: cutset bound; decode–forward; compress–forward
  - Gaussian networks: scaling laws; high SNR approximations

  Correlated sources over noiseless (wireline) channels:

Part III: Multi-hop Networks

- Relaying and multiple communication rounds

Independent messages over noiseless networks:

Max-flow min-cut theorem; network coding

Independent messages over noisy networks:

- Relay channel: cutset bound; decode–forward; compress–forward
- Interactive communication: feedback capacity; iterative refinement
- DM networks: cutset bound; decode–forward; compress–forward
- Gaussian networks: scaling laws; high SNR approximations

Correlated sources over noiseless (wireline) channels:

Multiple descriptions networks; interactive source coding

Part IV: Extensions

- Extensions of the theory to other settings

Part IV: Extensions

- Extensions of the theory to other settings

  Communication for computing:

  Distributed coding for computing: Orlitsky–Roche; μ-sum problem; distributed consensus

Part IV: Extensions

- Extensions of the theory to other settings

  Communication for computing:

  Distributed coding for computing: Orlitsky–Roche; μ-sum problem; distributed consensus

  Information theoretic secrecy:

  Wiretap channels; key generation from common randomness

Part IV: Extensions

- Extensions of the theory to other settings

  Communication for computing:

  Distributed coding for computing: Orlitsky–Roche; μ-sum problem; distributed consensus

  Information theoretic secrecy:

  Wiretap channels; key generation from common randomness

  Asynchronous communication:

  Random arrivals; asynchronous MAC

Balancing Introduction of Models and Techniques: Broadcast Channel

Degraded broadcast channels:

Channels with state:

Fading channels

General broadcast channels:

Gaussian vector channels:

Balancing Introduction of Models and Techniques: Broadcast Channel

Degraded broadcast channels:

- Superposition coding inner bound
- Degraded broadcast channels
- AWGN broadcast channels
- Less noisy and more capable broadcast channels

Channels with state:

Fading channels

General broadcast channels:

Gaussian vector channels:

Balancing Introduction of Models and Techniques: Broadcast Channel

Degraded broadcast channels:

Channels with state:

- Compound channel
- Arbitrarily varying channel
- Channels with random state
- Causal state information available at encoder
- Noncausal state information available at the encoder
- Writing on dirty paper
- Partial state information

Fading channels

General broadcast channels:

Gaussian vector channels:

Balancing Introduction of Models and Techniques: Broadcast Channel

Degraded broadcast channels:

Channels with state:

Fading channels

General broadcast channels:

- DM-BC with degraded message sets
- 3-Receiver multilevel DM-BC with degraded message sets
- Marton inner bound
- Relationship to Gelfand–Pinsker
- Nair–El Gamal outer bound
- Inner bound for more than 2 receivers

Gaussian vector channels:

Balancing Introduction of Models and Techniques: Broadcast Channel

Degraded broadcast channels:

Channels with state:

Fading channels

General broadcast channels:

Gaussian vector channels:

- Gaussian vector channel
- Gaussian vector fading channel
- Gaussian vector multiple access channel
- Spectral Gaussian broadcast channel
- Vector writing on dirty paper
- Gaussian vector broadcast channel

Typicality

- Let (u^n, x^n, y^n) be a triple of sequences with elements drawn from finite alphabets $(\mathcal{U}, \mathcal{X}, \mathcal{Y})$. Define their joint type as

$$\pi(u, x, y | u^n, x^n, y^n) = \frac{|\{i : (u_i, x_i, y_i) = (u, x, y)\}|}{n}$$

for $(u, x, y) \in \mathcal{U} \times \mathcal{X} \times \mathcal{Y}$

- Let $(U, X, Y) \sim p(u, x, y)$. The set $\mathcal{T}_\epsilon^{(n)}(U, X, Y)$ of ϵ-typical n-sequences is defined as

$$\{(u^n, x^n, y^n) : |\pi(u, x, y | u^n, x^n, y^n) - p(u, x, y)| \le \epsilon \cdot p(u, x, y) \text{ for all } (u, x, y) \in \mathcal{U} \times \mathcal{X} \times \mathcal{Y}\}$$

Typicality

- Let (u^n, x^n, y^n) be a triple of sequences with elements drawn from finite alphabets $(\mathcal{U}, \mathcal{X}, \mathcal{Y})$. Define their joint type as

$$\pi(u, x, y | u^n, x^n, y^n) = \frac{|\{i : (u_i, x_i, y_i) = (u, x, y)\}|}{n}$$

  for $(u, x, y) \in \mathcal{U} \times \mathcal{X} \times \mathcal{Y}$

- Let $(U, X, Y) \sim p(u, x, y)$. The set $\mathcal{T}_\epsilon^{(n)}(U, X, Y)$ of ϵ-typical n-sequences is defined as

$$\{(u^n, x^n, y^n) : |\pi(u, x, y|u^n, x^n, y^n) - p(u, x, y)| \leq \epsilon \cdot p(u, x, y) \text{ for all } (u, x, y) \in \mathcal{U} \times \mathcal{X} \times \mathcal{Y}\}$$

- Typical average lemma: Let $x^n \in \mathcal{T}_\epsilon^{(n)}(X)$. Then for any $g(x) \geq 0$,

$$(1 - \epsilon) \, \mathsf{E}(g(X)) \leq (1/n) \textstyle\sum_{i=1}^n g(x_i) \leq (1 + \epsilon) \, \mathsf{E}(g(X))$$

Joint Typicality Lemma

- Let $(U, X, Y) \sim p(u, x, y)$.
  1. Let $(u^n, x^n) \in \mathcal{T}_\epsilon^{(n)}(U, X)$ and $\tilde{Y}^n \sim \prod_{i=1}^n p_{Y|U}(\tilde{y}_i|u_i)$. Then

$$\mathsf{P}\{(u^n, x^n, \tilde{Y}^n) \in \mathcal{T}_\epsilon^{(n)}(U, X, Y)\} \doteq 2^{-nI(X;Y|U)}$$

Joint Typicality Lemma

- Let $(U, X, Y) \sim p(u, x, y)$.
  1. Let $(u^n, x^n) \in \mathcal{T}_\epsilon^{(n)}(U, X)$ and $\tilde{Y}^n \sim \prod_{i=1}^n p_{Y|U}(\tilde{y}_i|u_i)$. Then
  $$\mathsf{P}\{(u^n, x^n, \tilde{Y}^n) \in \mathcal{T}_\epsilon^{(n)}(U, X, Y)\} \doteq 2^{-nI(X;Y|U)}$$
  2. If $(\tilde{U}^n, \tilde{X}^n) \sim p(\tilde{u}^n, \tilde{x}^n)$ and $\tilde{Y}^n \sim \prod_{i=1}^n p_{Y|U}(\tilde{y}_i|\tilde{u}_i)$. Then
  $$\mathsf{P}\{(\tilde{U}^n, \tilde{X}^n, \tilde{Y}^n) \in \mathcal{T}_\epsilon^{(n)}(U, X, Y)\} \le 2^{-n(I(X;Y|U)-\delta(\epsilon))}$$

Packing Lemma

Let $(U, X, Y) \sim p(u, x, y)$ and $\tilde{U}^n \sim p(\tilde{u}^n)$. Let $X^n(m)$, $m \in \mathcal{A}$, where $|\mathcal{A}| \leq 2^{nR}$, be random sequences, each distributed according to $\prod_{i=1}^{n} p_{X|U}(x_i|\tilde{u}_i)$ with arbitrary dependence on the rest

Packing Lemma

Let $(U, X, Y) \sim p(u, x, y)$ and $\tilde{U}^n \sim p(\tilde{u}^n)$. Let $X^n(m)$, $m \in \mathcal{A}$, where $|\mathcal{A}| \le 2^{nR}$, be random sequences, each distributed according to $\prod_{i=1}^n p_{X|U}(x_i|\tilde{u}_i)$ with arbitrary dependence on the rest

Let $\tilde{Y}^n \in \mathcal{Y}^n$ be another random sequence, conditionally independent of each $X^n(m), m \in \mathcal{A}$, given $\tilde{U}^n$, and distributed according to an arbitrary pmf $p(\tilde{y}^n|\tilde{u}^n)$

Packing Lemma

Let $(U, X, Y) \sim p(u, x, y)$ and $\tilde{U}^n \sim p(\tilde{u}^n)$. Let $X^n(m)$, $m \in \mathcal{A}$, where $|\mathcal{A}| \le 2^{nR}$, be random sequences, each distributed according to $\prod_{i=1}^n p_{X|U}(x_i|\tilde{u}_i)$ with arbitrary dependence on the rest

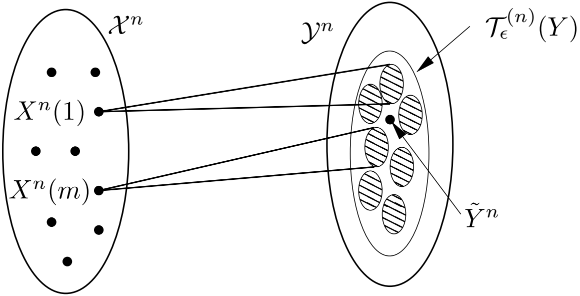
Let $\tilde{Y}^n \in \mathcal{Y}^n$ be another random sequence, conditionally independent of each $X^n(m), m \in \mathcal{A}$, given $\tilde{U}^n$, and distributed according to an arbitrary pmf $p(\tilde{y}^n|\tilde{u}^n)$

Then, there exists $\delta(\epsilon) \to 0$ as $\epsilon \to 0$ such that

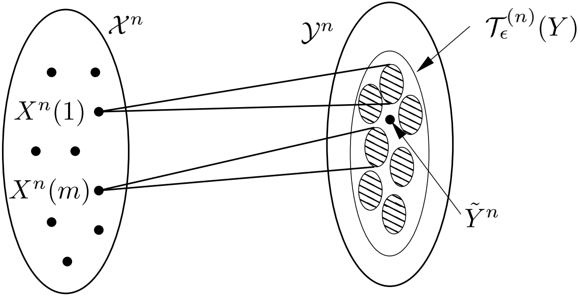
$$\mathsf{P}\{(\tilde{U}^n, X^n(m), \tilde{Y}^n) \in \mathcal{T}_\epsilon^{(n)} \text{ for some } m \in \mathcal{A}\} \to 0$$

as $n \to \infty$, if $R < I(X; Y|U) - \delta(\epsilon)$

The sequences $X^n(m)$, $m \in \mathcal{A}$, represent codewords. The $\tilde{Y}^n$ sequence represents the received sequence as a result of sending a codeword $\notin \mathcal{A}$



The sequences $X^n(m)$, $m \in \mathcal{A}$, represent codewords. The $\tilde{Y}^n$ sequence represents the received sequence as a result of sending a codeword $\notin \mathcal{A}$



The lemma shows that under any pmf on $\tilde{Y}^n$ the probability that some codeword in $\mathcal{A}$ is jointly typical with $\tilde{Y}^n \to 0$ as $n \to \infty$ if the rate of the code $R < I(X;Y|U)$

Covering Lemma

Let $(U, X, \hat{X}) \sim p(u, x, \hat{x})$. Let $(U^n, X^n) \sim p(u^n, x^n)$ be a pair of arbitrarily distributed random sequences such that $\mathsf{P}\{(U^n, X^n) \in \mathcal{T}_\epsilon^{(n)}(U, X)\} \to 1$ as $n \to \infty$

Covering Lemma

Let $(U, X, \hat{X}) \sim p(u, x, \hat{x})$. Let $(U^n, X^n) \sim p(u^n, x^n)$ be a pair of arbitrarily distributed random sequences such that $\mathsf{P}\{(U^n, X^n) \in \mathcal{T}_\epsilon^{(n)}(U, X)\} \to 1$ as $n \to \infty$

Let $\hat{X}^n(m), m \in \mathcal{A}$, where $|\mathcal{A}| \geq 2^{nR}$, be random sequences, conditionally independent of each other and of X^n given U^n, and distributed according to $\prod_{i=1}^n p_{\hat{X}|U}(\hat{x}_i|u_i)$

Covering Lemma

Let $(U, X, \hat{X}) \sim p(u, x, \hat{x})$. Let $(U^n, X^n) \sim p(u^n, x^n)$ be a pair of arbitrarily distributed random sequences such that $\mathsf{P}\{(U^n, X^n) \in \mathcal{T}_\epsilon^{(n)}(U, X)\} \to 1$ as $n \to \infty$

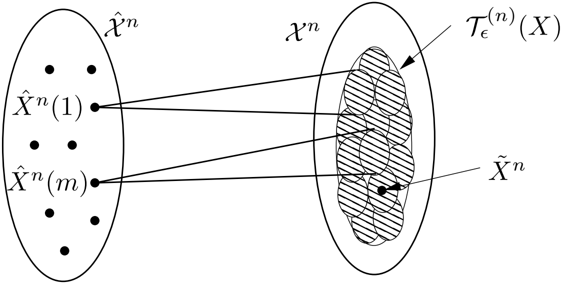
Let $\hat{X}^n(m), m \in \mathcal{A}$, where $|\mathcal{A}| \geq 2^{nR}$, be random sequences, conditionally independent of each other and of X^n given U^n, and distributed according to $\prod_{i=1}^n p_{\hat{X}|U}(\hat{x}_i|u_i)$

Then, there exists $\delta(\epsilon) \to 0$ as $\epsilon \to 0$ such that

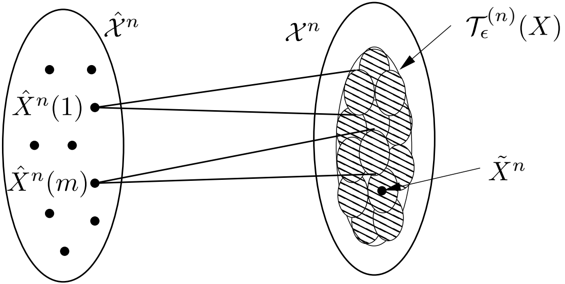
$$\mathsf{P}\{(U^n, X^n, \hat{X}^n(m)) \notin \mathcal{T}_\epsilon^{(n)} \text{ for all } m \in \mathcal{A}\} \to 0$$

as $n \to \infty$, if $R > I(X; \hat{X}|U) + \delta(\epsilon)$

The sequences $\hat{X}^n(m)$, $m \in \mathcal{A}$, represent reproduction sequences and X^n represents the source sequence



The sequences $\hat{X}^n(m)$, $m \in \mathcal{A}$, represent reproduction sequences and X^n represents the source sequence



The lemma shows that if $R > I(X; \hat{X}|U)$ then there is at least one reproduction sequence that is jointly typical with $\tilde{X}^n$

Conditional Typicality Lemma

- Let $(X,Y) \sim p(x,y)$, $x^n \in \mathcal{T}_{\epsilon'}^{(n)}(X)$, and $Y^n \sim \prod_{i=1}^n p_{Y|X}(y_i|x_i)$. Then, for every $\epsilon > \epsilon'$,

$$\mathsf{P}\{(x^n, Y^n) \in \mathcal{T}_{\epsilon}^{(n)}(X,Y)\} \to 1 \text{ as } n \to \infty$$

Conditional Typicality Lemma

- Let $(X,Y) \sim p(x,y)$, $x^n \in \mathcal{T}_{\epsilon'}^{(n)}(X)$, and $Y^n \sim \prod_{i=1}^n p_{Y|X}(y_i|x_i)$. Then, for every $\epsilon > \epsilon'$,

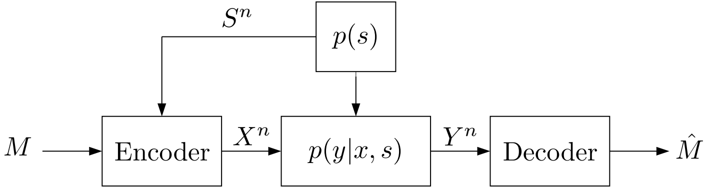
$$\mathsf{P}\{(x^n, Y^n) \in \mathcal{T}_{\epsilon}^{(n)}(X,Y)\} \to 1 \text{ as } n \to \infty$$

- Markov lemma is a special case: $U \to X \to Y$ form a Markov chain. If $(u^n, x^n) \in \mathcal{T}_{\epsilon'}^{(n)}(U,X)$ and $Y^n \sim \prod_{i=1}^n p_{Y|X}(y_i|x_i)$, then for every $\epsilon > \epsilon'$,

$$\mathsf{P}\{(u^n, x^n, Y^n) \in \mathcal{T}_{\epsilon}^{(n)}(U,X,Y)\} \to 1 \text{ as } n \to \infty$$

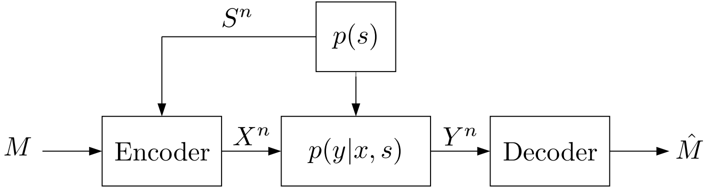
Gelfand–Pinsker

- Consider a DMC with DM state $(\mathcal{X} \times \mathcal{S}, p(y|x,s)p(s), \mathcal{Y})$
- The sender X who knows the state sequence S^n noncausally and wishes to send a message $M \in [1 : 2^{nR}]$ to the receiver Y



Gelfand–Pinsker

- Consider a DMC with DM state $(\mathcal{X} \times \mathcal{S}, p(y|x,s)p(s), \mathcal{Y})$
- The sender X who knows the state sequence S^n noncausally and wishes to send a message $M \in [1:2^{nR}]$ to the receiver Y



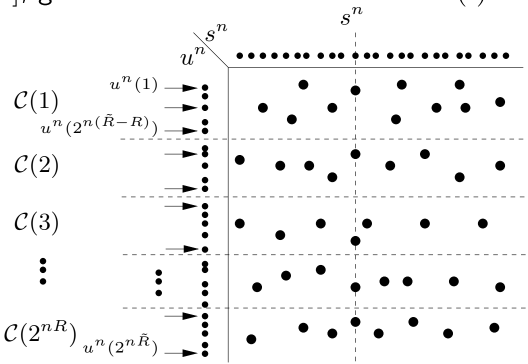
Gelfand–Pinsker Theorem

The capacity of a DMC with DM state available noncausally at the encoder is

$$C_{\mathrm{SI-E}} = \max_{p(u|s),\ x(u,s)} \left(I(U;Y) - I(U;S) \right)$$

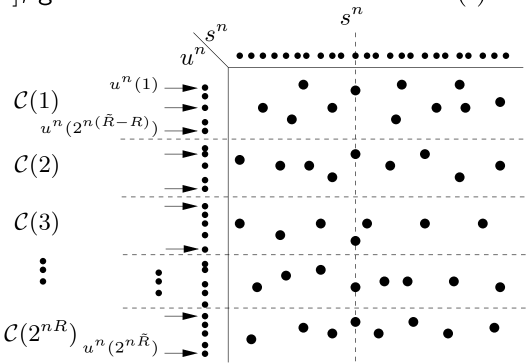
Outline of Achievability [Heegard, El Gamal]

- Fix $p(u|s)$, $x(u,s)$ that achieve capacity. For each message $m \in [1:2^{nR}]$, generate a subcode of $2^{n(\tilde{R}-R)}$ $u^n(l)$ sequences



Outline of Achievability [Heegard, El Gamal]

- Fix $p(u|s)$, $x(u,s)$ that achieve capacity. For each message $m \in [1:2^{nR}]$, generate a subcode of $2^{n(\tilde{R}-R)}$ $u^n(l)$ sequences



- To send m given s^n, find $u^n(l) \in \mathcal{C}(m)$ that is jointly typical with s^n and transmit $x_i = x(u_i(l), s_i)$ for $i \in [1:n]$
- The receiver finds a jointly typical $\hat{u}^n$ with y^n and declares the subcode index $\hat{m}$ of $\hat{u}^n$ to be the message sent

Analysis of the Probability of Error

- Assume $M = 1$ and let L be the index of the chosen U^n codeword for $M = 1$ and S^n
- We bound each probability of error event:
  - $\mathcal{E}_1 = \{(S^n, U^n(l)) \notin \mathcal{T}_{\epsilon'}^{(n)}$ for all $U^n(l) \in \mathcal{C}(1)\}$:
    By the covering lemma, $\mathsf{P}(\mathcal{E}_1) \to 0$ as $n \to \infty$ if $\tilde{R} - R > I(U;S)$

Analysis of the Probability of Error

- Assume $M = 1$ and let L be the index of the chosen U^n codeword for $M = 1$ and S^n
- We bound each probability of error event:
  - $\mathcal{E}_1 = \{(S^n, U^n(l)) \notin \mathcal{T}_{\epsilon'}^{(n)}$ for all $U^n(l) \in \mathcal{C}(1)\}$:
    By the covering lemma, $\mathsf{P}(\mathcal{E}_1) \to 0$ as $n \to \infty$ if $\tilde{R} - R > I(U;S)$
  - $\mathcal{E}_2 = \{(U^n(L), Y^n) \notin \mathcal{T}_{\epsilon}^{(n)}\}$:
    By the conditional typicality lemma, $\mathsf{P}(\mathcal{E}_1^c \cap \mathcal{E}_2) \to 0$ as $n \to \infty$

Analysis of the Probability of Error

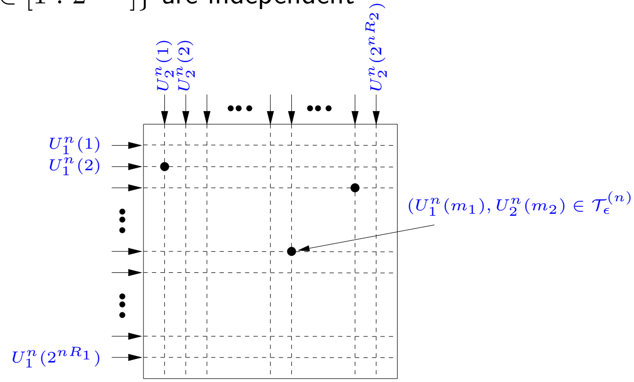
- Assume $M = 1$ and let L be the index of the chosen U^n codeword for $M = 1$ and S^n
- We bound each probability of error event:
  - $\mathcal{E}_1 = \{(S^n, U^n(l)) \notin \mathcal{T}_{\epsilon'}^{(n)}$ for all $U^n(l) \in \mathcal{C}(1)\}$:
    By the covering lemma, $\mathsf{P}(\mathcal{E}_1) \to 0$ as $n \to \infty$ if $\tilde{R} - R > I(U;S)$
  - $\mathcal{E}_2 = \{(U^n(L), Y^n) \notin \mathcal{T}_{\epsilon}^{(n)}\}$:
    By the conditional typicality lemma, $\mathsf{P}(\mathcal{E}_1^c \cap \mathcal{E}_2) \to 0$ as $n \to \infty$
  - $\mathcal{E}_3 = \{(U^n(\tilde{l}), Y^n) \in \mathcal{T}_{\epsilon}^{(n)}$ for some $U^n(\tilde{l}) \notin \mathcal{C}(1)\}$
    Since each $U^n(\tilde{l}) \notin \mathcal{C}(1)$ is independent of Y^n and generated according to $\prod_{i=1}^n p_U(u_i)$, by the packing lemma, $\mathsf{P}(\mathcal{E}_3) \to 0$ as $n \to \infty$ if $\tilde{R} < I(U;Y)$

Analysis of the Probability of Error

- Assume $M = 1$ and let L be the index of the chosen U^n codeword for $M = 1$ and S^n
- We bound each probability of error event:
  - $\mathcal{E}_1 = \{(S^n, U^n(l)) \notin \mathcal{T}_{\epsilon'}^{(n)}$ for all $U^n(l) \in \mathcal{C}(1)\}$:
    By the covering lemma, $\mathsf{P}(\mathcal{E}_1) \to 0$ as $n \to \infty$ if $\tilde{R} - R > I(U;S)$
  - $\mathcal{E}_2 = \{(U^n(L), Y^n) \notin \mathcal{T}_{\epsilon}^{(n)}\}$:
    By the conditional typicality lemma, $\mathsf{P}(\mathcal{E}_1^c \cap \mathcal{E}_2) \to 0$ as $n \to \infty$
  - $\mathcal{E}_3 = \{(U^n(\tilde{l}), Y^n) \in \mathcal{T}_{\epsilon}^{(n)}$ for some $U^n(\tilde{l}) \notin \mathcal{C}(1)\}$
    Since each $U^n(\tilde{l}) \notin \mathcal{C}(1)$ is independent of Y^n and generated according to $\prod_{i=1}^n p_U(u_i)$, by the packing lemma, $\mathsf{P}(\mathcal{E}_3) \to 0$ as $n \to \infty$ if $\tilde{R} < I(U;Y)$
- Thus the probability or error $\to 0$ as $n \to \infty$ if $R < I(U;Y) - I(U;S)$

Mutual Covering Lemma [El Gamal, van der Meulen]

Let $(U_1, U_2) \sim p(u_1, u_2)$. For $j = 1, 2$, let $U_j^n(m_j), m_j \in [1 : 2^{nR_j}]$, be pairwise independent random sequences, each distributed according to $\prod_{i=1}^n p_{U_j}(u_{ji})$. Assume that $\{U_1^n(m_1) : m_1 \in [1 : 2^{nR_1}]\}$ and $\{U_2^n(m_2) : m_2 \in [1 : 2^{nR_2}]\}$ are independent



Mutual Covering Lemma [El Gamal, van der Meulen]

Let $(U_1, U_2) \sim p(u_1, u_2)$. For $j = 1, 2$, let $U_j^n(m_j), m_j \in [1 : 2^{nR_j}]$, be pairwise independent random sequences, each distributed according to $\prod_{i=1}^n p_{U_j}(u_{ji})$. Assume that $\{U_1^n(m_1) : m_1 \in [1 : 2^{nR_1}]\}$ and $\{U_2^n(m_2) : m_2 \in [1 : 2^{nR_2}]\}$ are independent

Mutual Covering Lemma [El Gamal, van der Meulen]

Let $(U_1, U_2) \sim p(u_1, u_2)$. For $j = 1, 2$, let $U_j^n(m_j), m_j \in [1 : 2^{nR_j}]$, be pairwise independent random sequences, each distributed according to $\prod_{i=1}^n p_{U_j}(u_{ji})$. Assume that $\{U_1^n(m_1) : m_1 \in [1 : 2^{nR_1}]\}$ and $\{U_2^n(m_2) : m_2 \in [1 : 2^{nR_2}]\}$ are independent

Then, there exists $\delta(\epsilon) \to 0$ as $\epsilon \to 0$ such that

$$\mathsf{P}\{(U_1^n(m_1), U_2^n(m_2)) \notin \mathcal{T}_\epsilon^{(n)} \text{ for all } (m_1, m_2)\} \to 0$$

as $n \to \infty$ if $R_1 + R_2 > I(U_1; U_2)$

Mutual Covering Lemma [El Gamal, van der Meulen]

Let $(U_1, U_2) \sim p(u_1, u_2)$. For $j = 1, 2$, let $U_j^n(m_j), m_j \in [1 : 2^{nR_j}]$, be pairwise independent random sequences, each distributed according to $\prod_{i=1}^n p_{U_j}(u_{ji})$. Assume that $\{U_1^n(m_1) : m_1 \in [1 : 2^{nR_1}]\}$ and $\{U_2^n(m_2) : m_2 \in [1 : 2^{nR_2}]\}$ are independent

Then, there exists $\delta(\epsilon) \to 0$ as $\epsilon \to 0$ such that

$$\mathsf{P}\{(U_1^n(m_1), U_2^n(m_2)) \notin \mathcal{T}_\epsilon^{(n)} \text{ for all } (m_1, m_2)\} \to 0$$

as $n \to \infty$ if $R_1 + R_2 > I(U_1; U_2)$

- Used in the proof of Marton inner bound for BC
- Can be extended to k variables. Extension used in the proof of El Gamal–Cover inner bound for multiple descriptions and for extending Marton inner bound to k receivers

Mutual Packing Lemma

Let $(U_1, U_2) \sim p(u_1, u_2)$. For $j = 1, 2$, let $U_j^n(m_j)$, $m_j \in [1 : 2^{nR_j}]$, be random sequences, each distributed according to $\prod_{i=1}^n p_{U_j}(u_{ji})$ with arbitrary dependence on the rest of the $U_j^n(m_j)$ sequences. Assume that $\{U_1^n(m_1) : m_1 \in [1 : 2^{nR_1}]\}$ and $\{U_2^n(m_2) : m_2 \in [1 : 2^{nR_2}]\}$ are independent

Mutual Packing Lemma

Let $(U_1, U_2) \sim p(u_1, u_2)$. For $j = 1, 2$, let $U_j^n(m_j)$, $m_j \in [1 : 2^{nR_j}]$, be random sequences, each distributed according to $\prod_{i=1}^n p_{U_j}(u_{ji})$ with arbitrary dependence on the rest of the $U_j^n(m_j)$ sequences. Assume that $\{U_1^n(m_1) : m_1 \in [1 : 2^{nR_1}]\}$ and $\{U_2^n(m_2) : m_2 \in [1 : 2^{nR_2}]\}$ are independent

Then, there exists $\delta(\epsilon) \to 0$ as $\epsilon \to 0$ such that

$$\mathsf{P}\{(U_1^n(m_1), U_2^n(m_2)) \in \mathcal{T}_\epsilon^{(n)} \text{ for some } (m_1, m_2)\} \to 0$$

as $n \to \infty$ if $R_1 + R_2 < I(U_1; U_2) - \delta(\epsilon)$

Mutual Packing Lemma

Let $(U_1, U_2) \sim p(u_1, u_2)$. For $j = 1, 2$, let $U_j^n(m_j)$, $m_j \in [1 : 2^{nR_j}]$, be random sequences, each distributed according to $\prod_{i=1}^n p_{U_j}(u_{ji})$ with arbitrary dependence on the rest of the $U_j^n(m_j)$ sequences. Assume that $\{U_1^n(m_1) : m_1 \in [1 : 2^{nR_1}]\}$ and $\{U_2^n(m_2) : m_2 \in [1 : 2^{nR_2}]\}$ are independent

Then, there exists $\delta(\epsilon) \to 0$ as $\epsilon \to 0$ such that

$$\mathsf{P}\{(U_1^n(m_1), U_2^n(m_2)) \in \mathcal{T}_\epsilon^{(n)} \text{ for some } (m_1, m_2)\} \to 0$$

as $n \to \infty$ if $R_1 + R_2 < I(U_1; U_2) - \delta(\epsilon)$

- Used in the proof of the Berger–Tung inner bound for distributed lossy source coding

Gaussian Sources and Channels

- Because Gaussian models are quite popular in wireless communication, we have complete coverage of all basic results

Gaussian Sources and Channels

- Achievability:
  1. Show that Gaussian optimizes mutual information expressions
  2. Prove achievability of optimized expressions via DM counterpart (with cost) by discretization and taking appropriate limits

  The second step is detailed only for AWGN channel and quadratic Gaussian source coding

Gaussian Sources and Channels

- Treatment of Gaussian is interspersed within each lecture,

Gaussian Sources and Channels

- Treatment of Gaussian is interspersed within each lecture, e.g., the interference channel lecture:
  - Inner and outer bounds on capacity region of DM-IC

Gaussian Sources and Channels

- Treatment of Gaussian is interspersed within each lecture, e.g., the interference channel lecture:
  - Inner and outer bounds on capacity region of DM-IC
  - Capacity region of DM-IC under strong interference

Gaussian Sources and Channels

- Treatment of Gaussian is interspersed within each lecture, e.g., the interference channel lecture:
  - Inner and outer bounds on capacity region of DM-IC
  - Capacity region of DM-IC under strong interference
  - AWGN-IC

Gaussian Sources and Channels

- Treatment of Gaussian is interspersed within each lecture, e.g., the interference channel lecture:
  - Inner and outer bounds on capacity region of DM-IC
  - Capacity region of DM-IC under strong interference
  - AWGN-IC
  - Capacity region of AWGN-IC under strong interference

Gaussian Sources and Channels

- Treatment of Gaussian is interspersed within each lecture, e.g., the interference channel lecture:
  - Inner and outer bounds on capacity region of DM-IC
  - Capacity region of DM-IC under strong interference
  - AWGN-IC
  - Capacity region of AWGN-IC under strong interference
  - Han–Kobayashi inner bound for DM-IC

Gaussian Sources and Channels

- Treatment of Gaussian is interspersed within each lecture, e.g., the interference channel lecture:
  - Inner and outer bounds on capacity region of DM-IC
  - Capacity region of DM-IC under strong interference
  - AWGN-IC
  - Capacity region of AWGN-IC under strong interference
  - Han–Kobayashi inner bound for DM-IC
  - Capacity region of a Class of deterministic DM-IC

Gaussian Sources and Channels

- Treatment of Gaussian is interspersed within each lecture, e.g., the interference channel lecture:
  - Inner and outer bounds on capacity region of DM-IC
  - Capacity region of DM-IC under strong interference
  - AWGN-IC
  - Capacity region of AWGN-IC under strong interference
  - Han–Kobayashi inner bound for DM-IC
  - Capacity region of a Class of deterministic DM-IC
  - Capacity region of AWGN-IC within Half a Bit

Gaussian Sources and Channels

- Treatment of Gaussian is interspersed within each lecture, e.g., the interference channel lecture:
  - Inner and outer bounds on capacity region of DM-IC
  - Capacity region of DM-IC under strong interference
  - AWGN-IC
  - Capacity region of AWGN-IC under strong interference
  - Han–Kobayashi inner bound for DM-IC
  - Capacity region of a Class of deterministic DM-IC
  - Capacity region of AWGN-IC within Half a Bit
  - Sum-capacity of AWGN-IC under weak interference

Converse

- The lectures discuss only weak converses

Converse

- The lectures discuss only weak converses
- The tools are introduced gradually:
  - DMC: Fano's inequality; convexity (data processing inequality); Markovity (memoryless)
  - AWGN: Gaussian optimizes differential entropy under power constraint

Converse

- The lectures discuss only weak converses
- The tools are introduced gradually:
  - DMC: Fano's inequality; convexity (data processing inequality); Markovity (memoryless)
  - AWGN: Gaussian optimizes differential entropy under power constraint
  - MAC: Time sharing random variable

Converse

- The lectures discuss only weak converses
- The tools are introduced gradually:
  - DMC: Fano's inequality; convexity (data processing inequality); Markovity (memoryless)
  - AWGN: Gaussian optimizes differential entropy under power constraint
  - MAC: Time sharing random variable
  - Degraded BC: Gallager's identification of auxiliary random variable; bounding cardinality
  - Binary Symmetric BC: Mrs. Gerber's lemma
  - AWGN-BC: EPI
  - More capable/less noisy BC: Csiszár's sum identity

Converse

- The lectures discuss only weak converses
- The tools are introduced gradually:
  - DMC: Fano's inequality; convexity (data processing inequality); Markovity (memoryless)
  - AWGN: Gaussian optimizes differential entropy under power constraint
  - MAC: Time sharing random variable
  - Degraded BC: Gallager's identification of auxiliary random variable; bounding cardinality
  - Binary Symmetric BC: Mrs. Gerber's lemma
  - AWGN-BC: EPI
  - More capable/less noisy BC: Csiszár's sum identity
  - Strong interference: Extension of more capable from scalar to vectors
  - Deterministic IC: Genie
  - Weak interference: Gaussian worst noise

Converse

- The lectures discuss only weak converses
- The tools are introduced gradually:
  - DMC: Fano's inequality; convexity (data processing inequality); Markovity (memoryless)
  - AWGN: Gaussian optimizes differential entropy under power constraint
  - MAC: Time sharing random variable
  - Degraded BC: Gallager's identification of auxiliary random variable; bounding cardinality
  - Binary Symmetric BC: Mrs. Gerber's lemma
  - AWGN-BC: EPI
  - More capable/less noisy BC: Csiszár's sum identity
  - Strong interference: Extension of more capable from scalar to vectors
  - Deterministic IC: Genie
  - Weak interference: Gaussian worst noise
  - Vector Gaussian BC: MAC/BC duality; convex optimization
  - Quadratic Gaussian distributed coding: MMSE

Extension to Networks

- The lectures include extensions (or lack thereof) of results for ≤ 3 users to networks

Extension to Networks

- The lectures include extensions (or lack thereof) of results for ≤ 3 users to networks
- In some rare cases the results extend naturally to many users:
  - MAC
  - Degraded BC
  - MIMO BC
  - Slepian–Wolf

Extension to Networks

- The lectures include extensions (or lack thereof) of results for ≤ 3 users to networks
- In some rare cases the results extend naturally to many users:
  - MAC
  - Degraded BC
  - MIMO BC
  - Slepian–Wolf
- In most cases the results don't extend and naive extensions of results for ≤ 3 users can be improved using new coding techniques

Extension to Networks

- The lectures include extensions (or lack thereof) of results for ≤ 3 users to networks
- In some rare cases the results extend naturally to many users:
  - MAC
  - Degraded BC
  - MIMO BC
  - Slepian–Wolf
- In most cases the results don't extend and naive extensions of results for ≤ 3 users can be improved using new coding techniques
- The lectures provide several examples of such cases

Some Interesting Extensions

- Inner bound for DM-BC with degraded message sets for 3 receivers
- Marton for ≥ 3 receivers
- General BC inner bound construction for ≥ 3 receivers

Some Interesting Extensions

- Inner bound for DM-BC with degraded message sets for 3 receivers
- Marton for ≥ 3 receivers
- General BC inner bound construction for ≥ 3 receivers
- Network coding for multicast noiseless networks and special cases

Some Interesting Extensions

- Inner bound for DM-BC with degraded message sets for 3 receivers
- Marton for ≥ 3 receivers
- General BC inner bound construction for ≥ 3 receivers
- Network coding for multicast noiseless networks and special cases
- Decode–forward
- Compress–forward

Some Interesting Extensions

- Inner bound for DM-BC with degraded message sets for 3 receivers
- Marton for ≥ 3 receivers
- General BC inner bound construction for ≥ 3 receivers
- Network coding for multicast noiseless networks and special cases
- Decode–forward
- Compress–forward
- Slepian–Wolf over noiseless broadcast network (CFO problem)

Some Interesting Extensions

- Inner bound for DM-BC with degraded message sets for 3 receivers
- Marton for ≥ 3 receivers
- General BC inner bound construction for ≥ 3 receivers
- Network coding for multicast noiseless networks and special cases
- Decode–forward
- Compress–forward
- Slepian–Wolf over noiseless broadcast network (CFO problem)
- Wiretap channel with > 2 receivers; key generation for many sources

Some Interesting Extensions

- Inner bound for DM-BC with degraded message sets for 3 receivers
- Marton for ≥ 3 receivers
- General BC inner bound construction for ≥ 3 receivers

- Network coding for multicast noiseless networks and special cases

- Decode–forward
- Compress–forward

- Slepian–Wolf over noiseless broadcast network (CFO problem)

- Wiretap channel with > 2 receivers; key generation for many sources

- Several cutset bounds for various types of networks

Some Interesting Extensions

- Inner bound for DM-BC with degraded message sets for 3 receivers
- Marton for ≥ 3 receivers
- General BC inner bound construction for ≥ 3 receivers
- Network coding for multicast noiseless networks and special cases
- Decode–forward
- Compress–forward
- Slepian–Wolf over noiseless broadcast network (CFO problem)
- Wiretap channel with > 2 receivers; key generation for many sources
- Several cutset bounds for various types of networks
- Scaling laws and high SNR approximations for Gaussian networks

Conclusion

- Lectures on NIT:
  - Top-down organization
  - Balances introduction of new tools and models
  - Elementary tools and proof techniques for most material
  - Unified approach to achievability
  - Comprehensive coverage of key results
  - Extensions to networks

Conclusion

- Lectures on NIT:
  - Top-down organization
  - Balances introduction of new tools and models
  - Elementary tools and proof techniques for most material
  - Unified approach to achievability
  - Comprehensive coverage of key results
  - Extensions to networks

- Some of the basic material ready to be included in graduate comm course sequences (with introductory IT course as prereq)

Conclusion

- Lectures on NIT:
  - Top-down organization
  - Balances introduction of new tools and models
  - Elementary tools and proof techniques for most material
  - Unified approach to achievability
  - Comprehensive coverage of key results
  - Extensions to networks
- Some of the basic material ready to be included in graduate comm course sequences (with introductory IT course as prereq)
- We plan to make the teaching subset of the lectures available early next year

Acknowledgments

- Many people have contributed to the development of the lectures over the years:
  - Many of my graduate students
  - My course TAs
  - Students that took the class

Acknowledgments

- Many people have contributed to the development of the lectures over the years:
  - Many of my graduate students
  - My course TAs
  - Students that took the class

- Tom Cover has been an inspiring and encouraging figure throughout

Acknowledgments

- Many people have contributed to the development of the lectures over the years:
  - Many of my graduate students
  - My course TAs
  - Students that took the class

- Tom Cover has been an inspiring and encouraging figure throughout

- Partial financial support from NSF and DARPA ITMANET