

Prestige Lecture Series on Science of Information



Information Theory Today

Sergio Verdú

Princeton University

.

Department of Computer Science

Purdue University

October 2, 2006

699 months ago...

A Mathematical Theory of Communication

By C. E. SHANNON

INTRODUCTION

THE recent development of various methods of modulation such as PCM and PPM which exchange bandwidth for signal-to-noise ratio has intensified the interest in a general theory of communication. A basis for such a theory is contained in the important papers of Nyquist¹ and Hartley² on this subject. In the present paper we will extend the theory to include a number of new factors, in particular the effect of noise in the channel, and the savings possible due to the statistical structure of the original message and due to the nature of the final destination of the information.

The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point. Frequently the messages have *meaning*; that is they refer to or are

Published in THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 27, pp. 379-423, 623-656, July, October, 1948
Copyright 1948 by AMERICAN TELEPHONE AND TELEGRAPH CO.
Printed in U. S. A.

Information Theory as a Design Driver

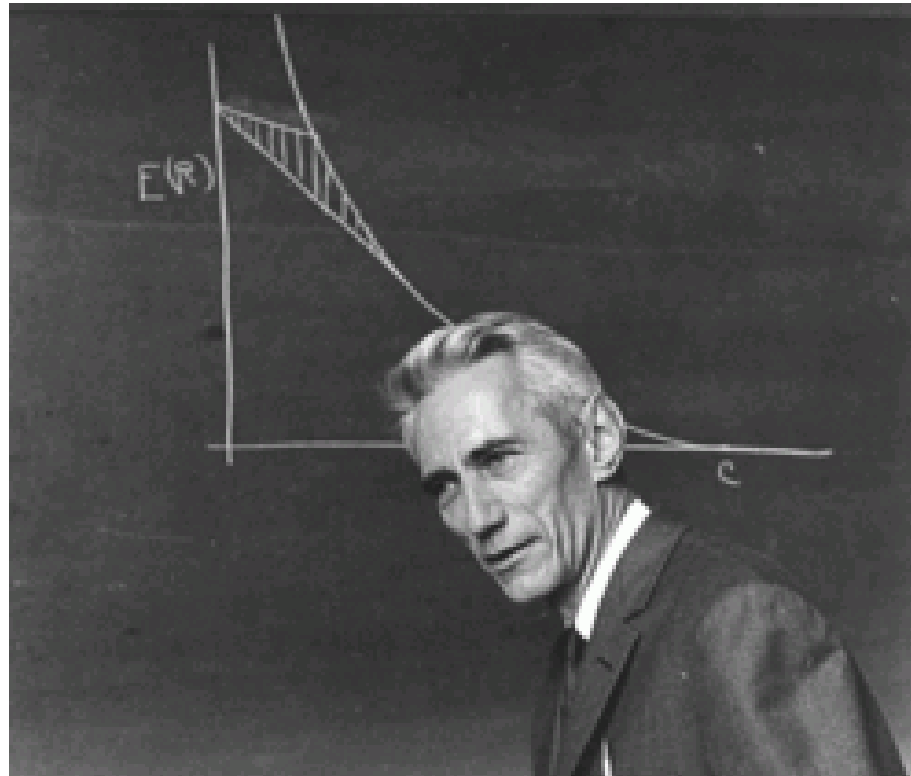
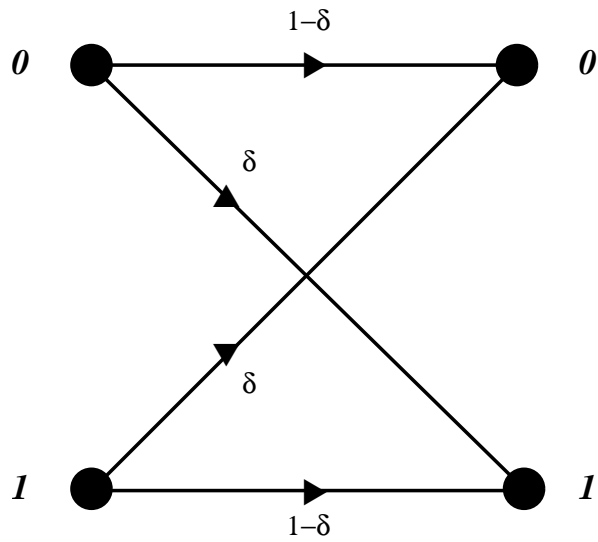
- Sparse-graph codes
- Universal data compression
- Voiceband modems
- Discrete multitone modulation
- CDMA
- Multiuser detection
- Multiantenna
- Space-time codes
- Opportunistic signaling
- Discrete denoising
- Cryptography

Open Problems: Single-User Channels

Open Problems: Single-User Channels

- Reliability Function

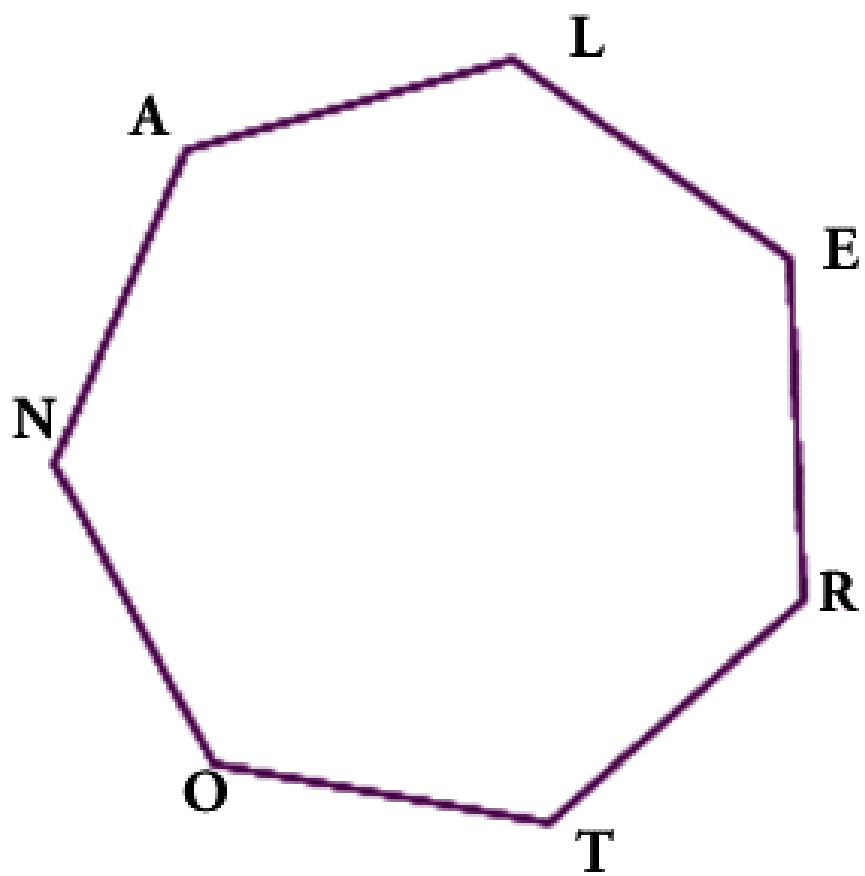
Reliability function



Open Problems: Single-User Channels

- Reliability Function
- Zero-error Capacity

Zero-Error Capacity



Open Problems: Single-User Channels

- Reliability Function
- Zero-error Capacity
- Delay – Error Probability Tradeoff

Open Problems: Single-User Channels

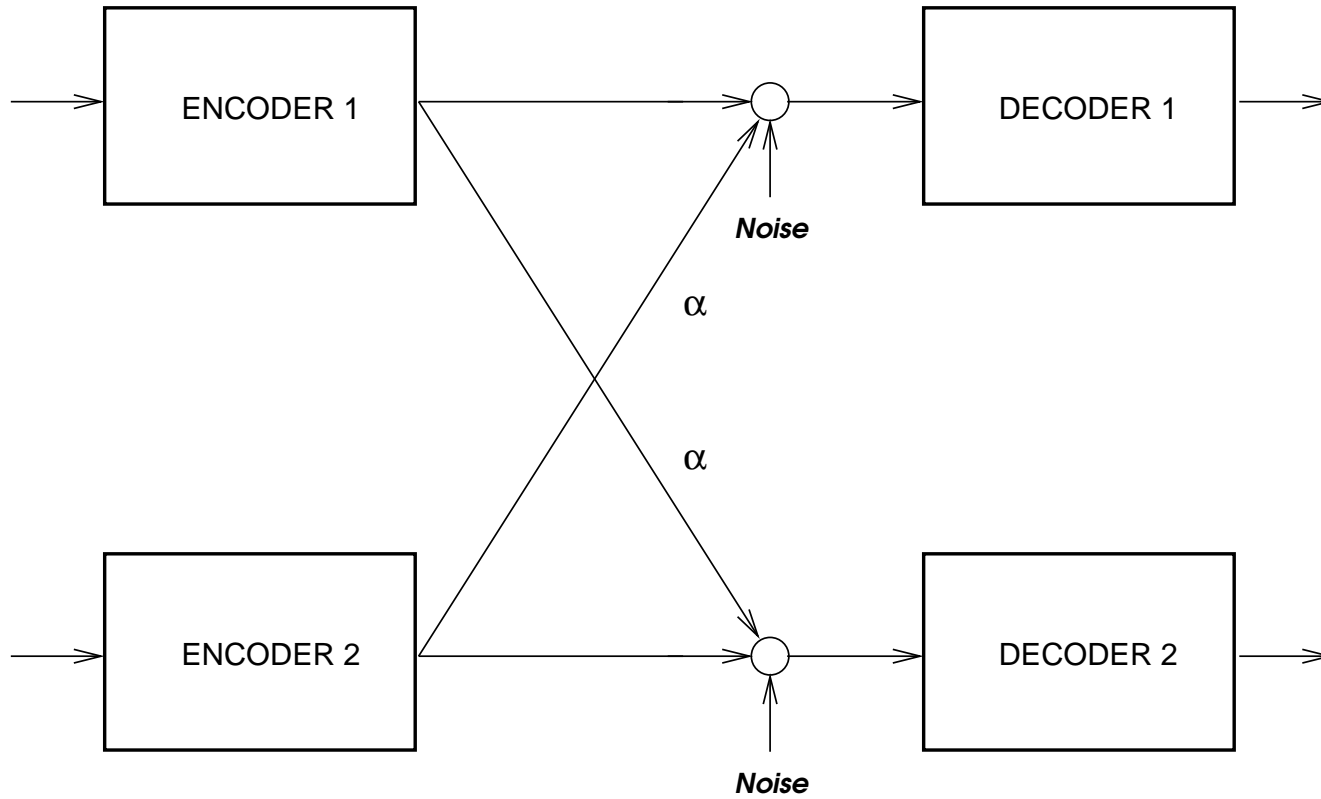
- Reliability Function
- Zero-error Capacity
- Delay – Error Probability Tradeoff
- Feedback
 - Partial/Noisy feedback
 - Constructive schemes
 - Gaussian channels with memory
- Deletions, Synchronization

Open Problems: Multiuser Channels

Open Problems: Multiuser Channels

- Interference Channels

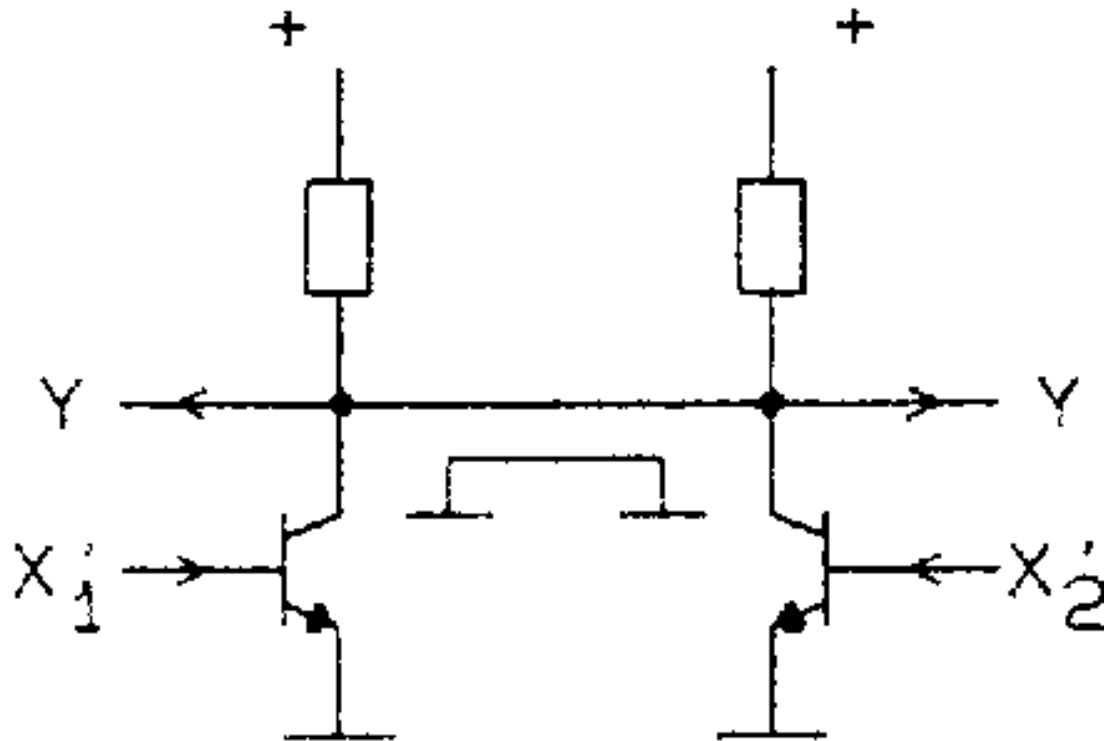
Interference Channels



Open Problems: Multiuser Channels

- Interference Channels
- Two-way Channels

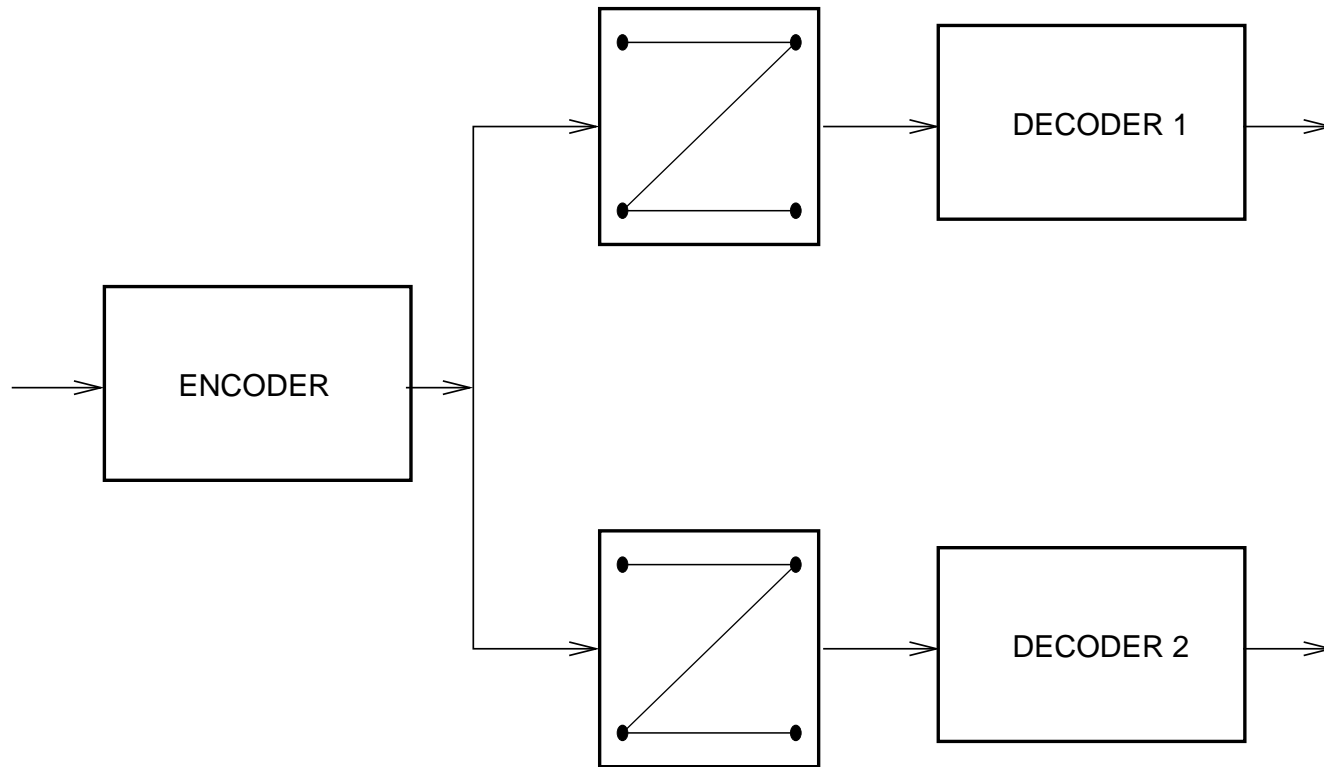
Two-Way Channels



Open Problems: Multiuser Channels

- Interference Channels
- Two-way Channels
- Broadcast Channels

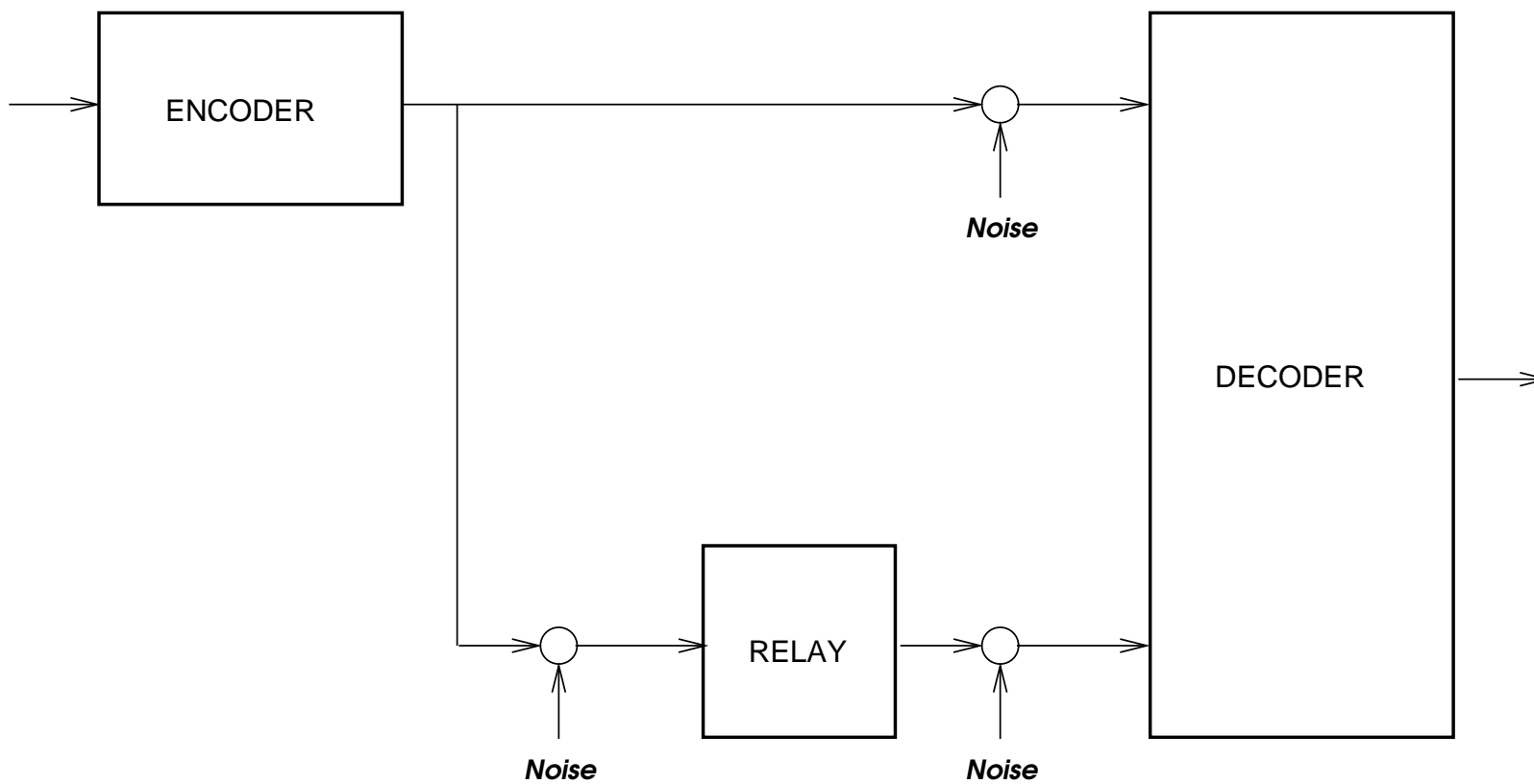
Broadcast Channels



Open Problems: Multiuser Channels

- Interference Channels
- Two-way Channels
- Broadcast Channels
- Relay Channels

Relay Channels



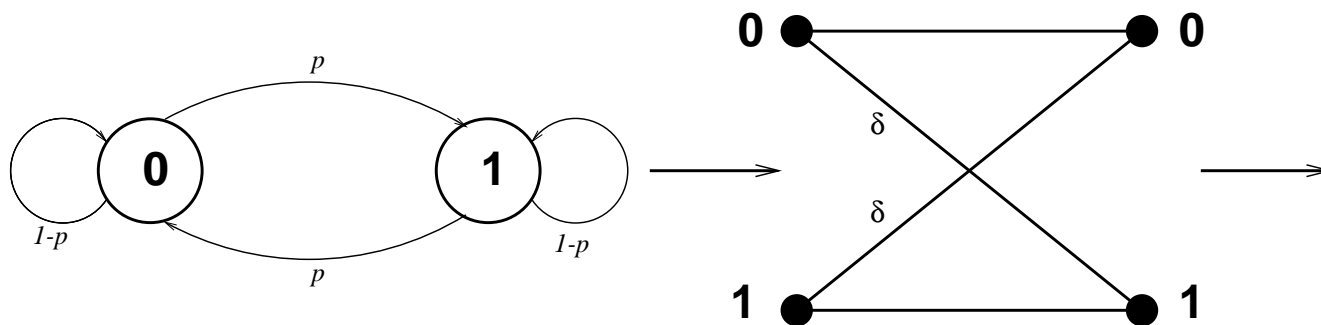
Open Problems: Multiuser Channels

- Interference Channels
- Two-way Channels
- Broadcast Channels
- Relay Channels
- Compression-Transmission

Open Problems: Lossless Data Compression

- Joint Source/Channel Coding
- Two-dimensional sources
- Implementing Slepian-Wolf:
 - Backup hard-disks with dialup modems?
- \Leftarrow Artificial Intelligence
- Entropy Rate of Sources with Memory

Entropy Rate of Sources with Memory



Open Problems: Lossy Data Compression

- Theory \leftrightarrow \leftrightarrow Practice

Open Problems: Lossy Data Compression

- Theory \leftrightarrow \leftrightarrow Practice
- Constructive Schemes
 - Memoryless Sources
 - Universal Lossy Data Compression

Open Problems: Lossy Data Compression

- Theory \leftrightarrow \leftrightarrow Practice
- Constructive Schemes
 - Memoryless Sources
 - Universal Lossy Data Compression
- Multi-source Fundamental Limits

Multi-source Fundamental Limits

DIGITAL · STEREO
453 478-2 GH

PY
925



a PolyGram company

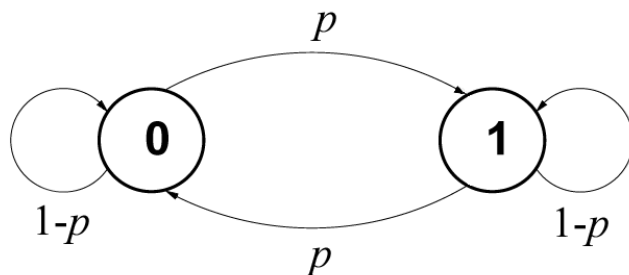
4D DDD
AUDIO RECORDING

A new dimension in clarity and realism
Klarheit und Klangtreue in neuer
Dimension
Une nouvelle dimension dans la clarté
et le naturel
La nuova dimensione di un suono
limpido e naturale

Open Problems: Lossy Data Compression

- Theory \leftrightarrow \leftrightarrow Practice
- Constructive Schemes
 - Memoryless Sources
 - Universal Lossy Data Compression
- Multi-source Fundamental Limits
- Rate-Distortion Functions

Binary Markov chain; Bit Error Rate



$$0 \leq p \leq \frac{1}{2}$$

$$R(D) = \begin{cases} h(p) - h(D) & \text{for } 0 \leq D \leq D_? \\ \text{UNKNOWN} & \text{otherwise.} \end{cases}$$

Gradient

↗ Constructive

↘ Combinatorics

↗ Applied

↘ Continuous Time

↗ Multiuser

↘ Ergodic Theory

↗ Universal Methods

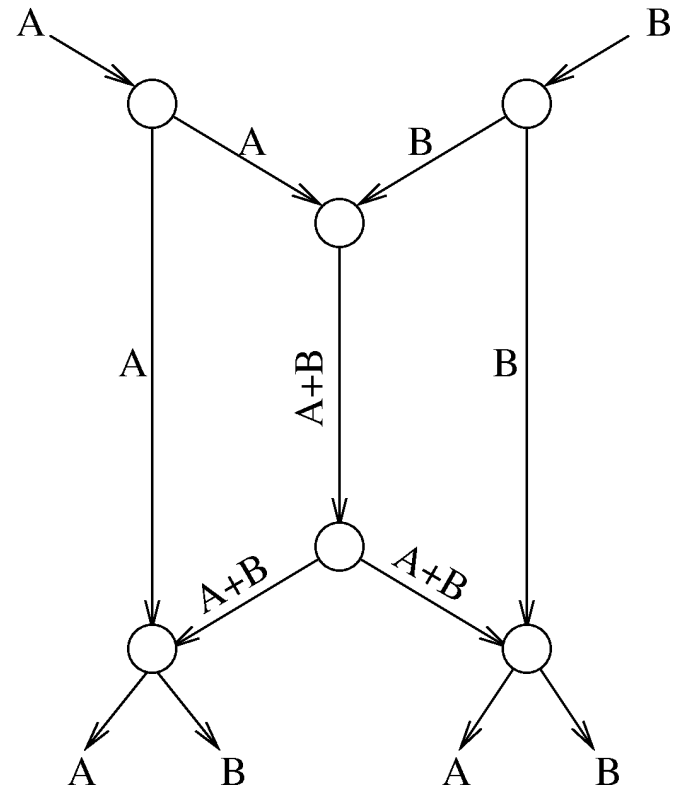
↘ Error Exponents

↗ Intersections

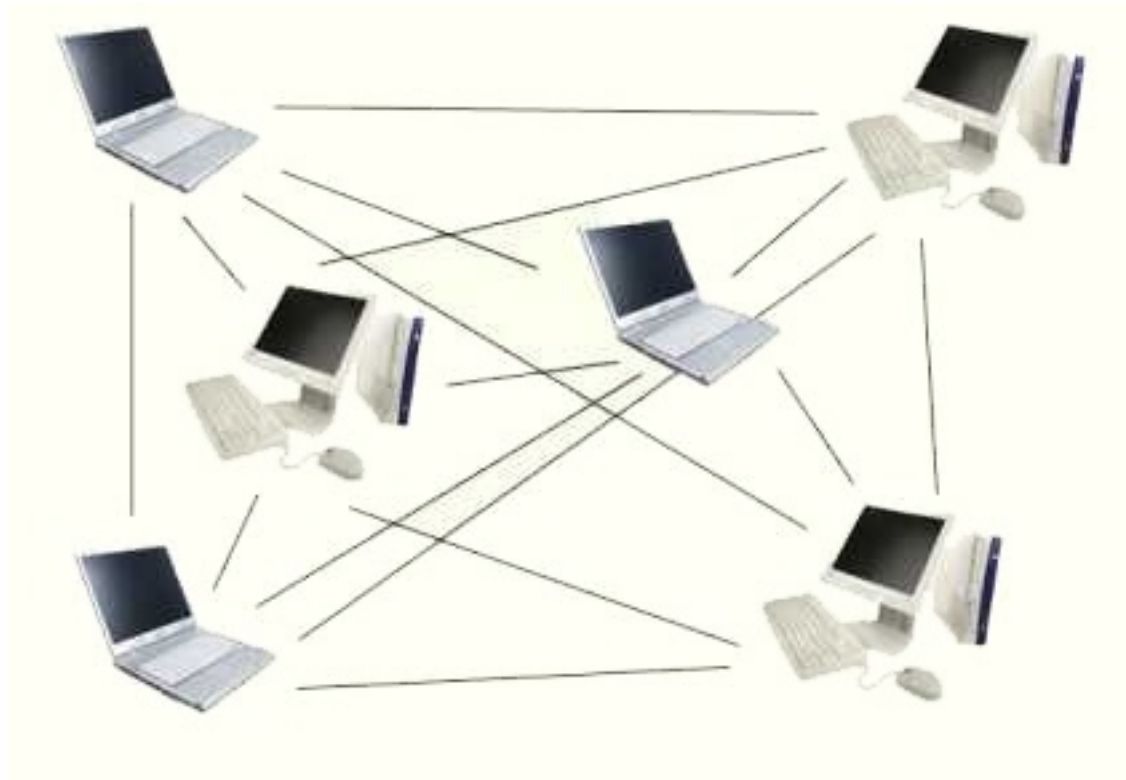
Intersections

- Networks
 - Network coding
 - Scaling laws

Network Coding



Scaling Laws



Intersections

- Networks

- Network coding
- Scaling laws

- Signal Processing

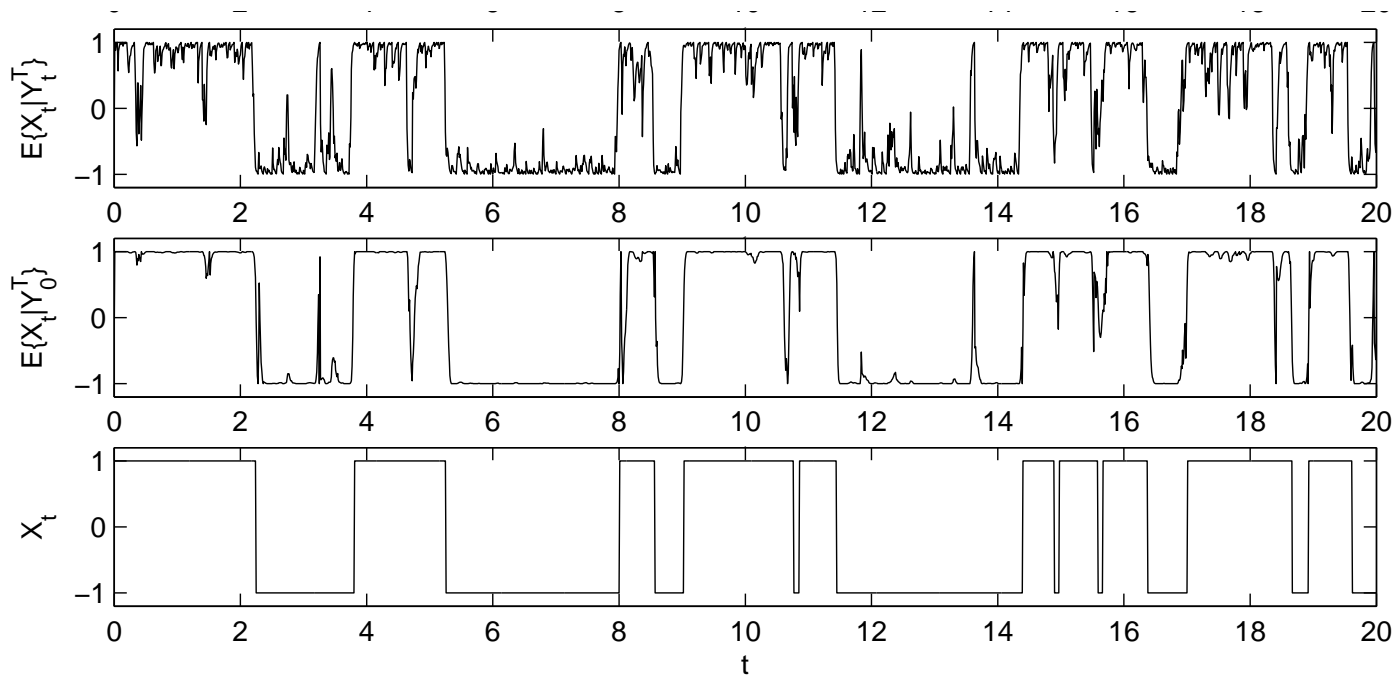
- Estimation theory
- Discrete denoising
- Finite-alphabet

Information Theory \Leftrightarrow Estimation Theory

$$\frac{d}{d\text{snr}} I(\mathbf{X}; \sqrt{\text{snr}} \cdot \mathbf{H} \mathbf{X} + \mathbf{W}) = \frac{1}{2} \text{mmse}(\text{snr})$$

- Entropy power inequality
- Monotonicity of nonGaussianness
- Mercury-Waterfilling
- Continuous-time Nonlinear Filtering

Information Theory \Leftrightarrow Nonlinear Filtering



$$\text{cmmse}(\text{snr}) = \frac{1}{\text{snr}} \int_0^{\text{snr}} \text{mmse}(\gamma) d\gamma$$

Information Theory \Leftrightarrow Estimation Theory

$$E^2[|E[X|Z] - E[X]|] \leq 2A^2 I(X; Z)$$

for any (X, Z) s.t $X \in [-A, A]$.

Intersections

- Networks

- Network coding
- Scaling laws

- Signal Processing

- Estimation theory
- Discrete denoising
- Finite-alphabet

Text Denoising: Don Quixote de La Mancha

Noisy Text (21 errors, 5% error rate):

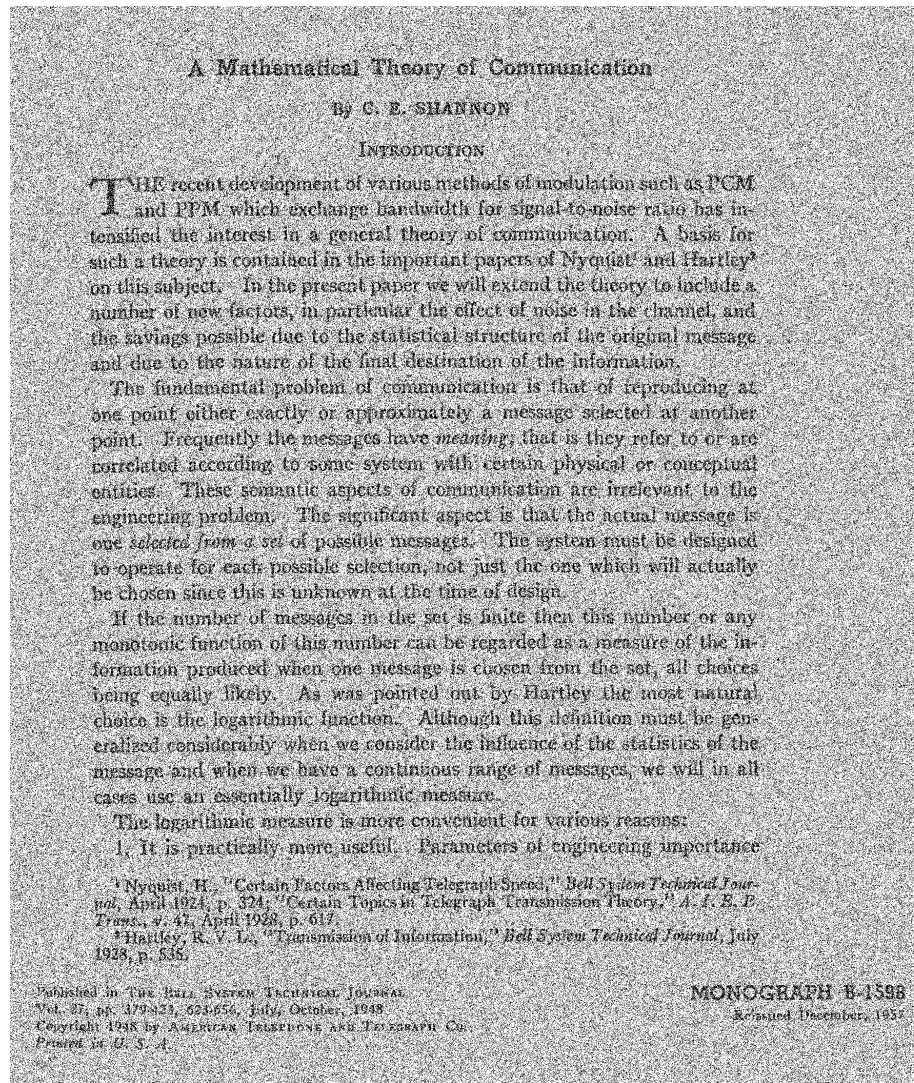
"Whar giants?" said Sancho Panza. "Those thou seest theee," snswered yis master, "with the long arms, and spne have tgem ndarly two leagues long." "Look, yIur worship," sair Sancho; "what we see there zre not gianrs but windmills, and what seem to be their arms are the sails that turned by the wind make rhe millstpne go." "Kt is easy to see," replied Don Quixote, "that thou art not used to this business of adventures; fhose are giantz; and if thou arf wfraod, away with thee out of this and betake thysepf to prayer while I engage them in fierce and unequal combat."

DUDE output, $k = 2$ (7 errors):

"What giants?" said Sancho Panza. "Those thou seest there," answered his master, "with the long arms, and spne have them nearly two leagues long." "Look, your worship," said Sancho; "what we see there are not giants but windmills, and what seem to be their arms are the sails that turned by the wind make the millstone go." "It is easy to see," replied Don Quixote, "that thou art not used to this business of adventures; fhose are giantz; and if thou arf wfraod, away with thee out of this and betake thyself to prayer while I engage them in fierce and unequal combat."

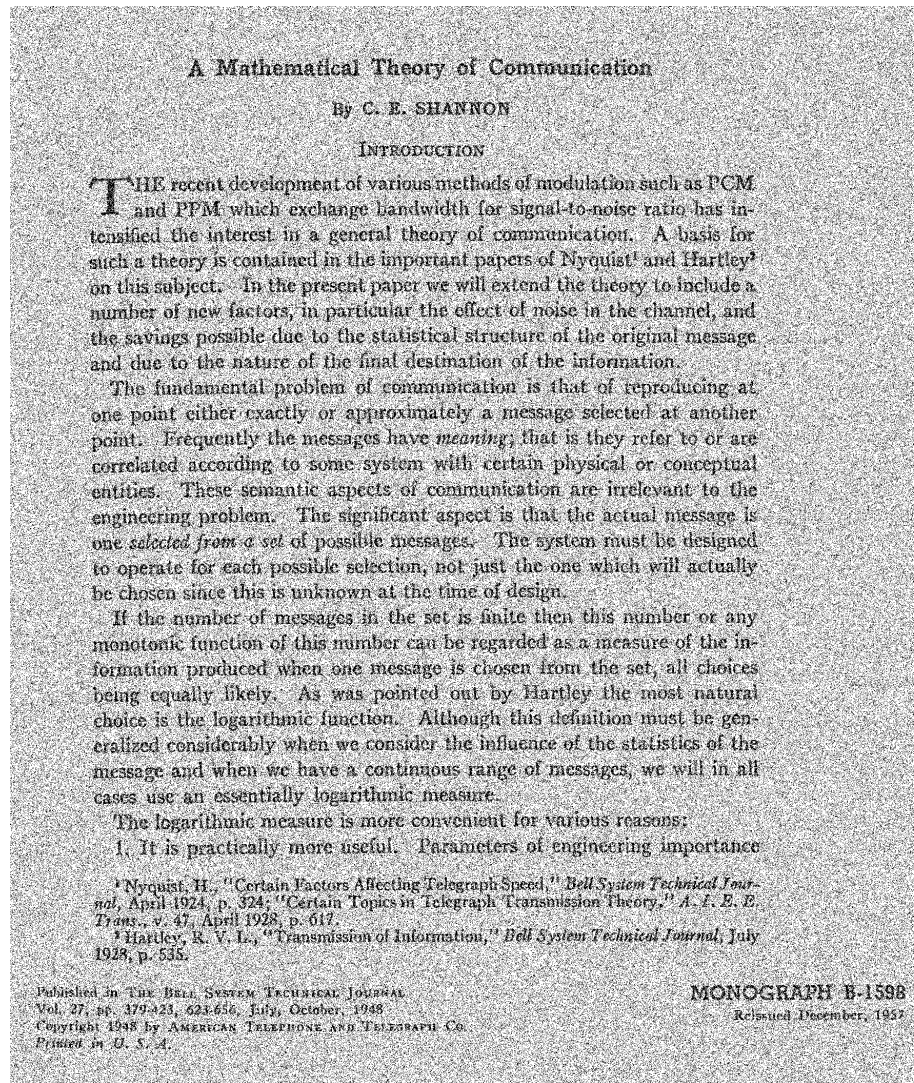
BSC systematic output; $C = 1 - h(0.25) = 0.19$

0.25



BP Decoder Output (RA; Rate = 0.25; $k = 4000$; 30 iter.)

0.21



Denoising+Decoding

0.0003

A Mathematical Theory of Communication

By C. E. SHANNON

INTRODUCTION

THE recent development of various methods of modulation such as PCM and PPM which exchange bandwidth for signal-to-noise ratio has intensified the interest in a general theory of communication. A basis for such a theory is contained in the important papers of Nyquist¹ and Hartley² on this subject. In the present paper we will extend the theory to include a number of new factors, in particular the effect of noise in the channel, and the savings possible due to the statistical structure of the original message and due to the nature of the final destination of the information.

The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point. Frequently the messages have *meaning*; that is they refer to or are correlated according to some system with certain physical or conceptual entities. These semantic aspects of communication are irrelevant to the engineering problem. The significant aspect is that the actual message is one *selected from a set* of possible messages. The system must be designed to operate for each possible selection, not just the one which will actually be chosen since this is unknown at the time of design.

If the number of messages in the set is finite then this number or any monotonic function of this number can be regarded as a measure of the information produced when one message is chosen from the set, all choices being equally likely. As was pointed out by Hartley the most natural choice is the logarithmic function. Although this definition must be generalized considerably when we consider the influence of the statistics of the message and when we have a continuous range of messages, we will in all cases use an essentially logarithmic measure.

The logarithmic measure is more convenient for various reasons:

1. It is practically more useful. Parameters of engineering importance

¹ Nyquist, H., "Certain Factors Affecting Telegraph Speed," *Bell System Technical Journal*, April 1924, p. 324; "Certain Topics in Telegraph Transmission Theory," *A. I. E. E. Trans.*, v. 47, April 1928, p. 617.

² Hartley, R. V. L., "Transmission of Information," *Bell System Technical Journal*, July 1928, p. 535.

Intersections

- Networks
 - Network coding
 - Scaling laws
- Control
 - Noisy [plant \longrightarrow controller] channel.
 - Control-oriented feedback communication schemes.
- Signal Processing
 - Estimation theory
 - Discrete denoising
 - Finite-alphabet

Intersections

- Networks
 - Network coding
 - Scaling laws
- Signal Processing
 - Estimation theory
 - Discrete denoising
 - Finite-alphabet
- Control
 - Noisy [plant \longrightarrow controller] channel.
 - Control-oriented feedback communication schemes.
- Computer Science
 - Analytic information theory
 - Interactive communication

Other Intersections

- Economics
- Quantum
- Bio
- Physics

Emerging Tools

- Optimization
- Statistical Physics
- Random Matrices

To Probe Further: Design Driver

- R. Gallager, *Low-Density Parity-Check Codes*, MIT Press, 1963.
- J. Ziv and A. Lempel, “A Universal algorithm for sequential data compression,” *IEEE Trans. Inform. Theory*, IT-24, pp. 337-343, May 1977
- G. Foschini, “Layered space-time architecture for wireless communication in a fading environment when using multiple antennas,” *Bell Labs Technical Journal 2*, Vol. 1, no. 2, pp 41-59, 1996
- U. Maurer, Information-Theoretic Cryptography, *Advances in Cryptology - CRYPTO '99*, Lecture Notes in Computer Science, Springer-Verlag, vol. 1666, pp. 47-64, Aug 1999.
- V. Tarokh V, N. Seshadri, and A. R. Calderbank, “Space-time Codes for High Data Rate Wireless Communication: Performance Criterion and

Code Construction,” *IEEE Trans. on Information Theory*, Vol. 44, No. 2, pp. 744-765, Mar. 1998.

- S. Verdú, *Multuser Detection*, Cambridge University Press, Cambridge UK, 1998
- T. Weissman, E. Ordentlich, G. Seroussi, S. Verdú and M. Weinberger, “Universal Discrete Denoising: Known Channel,” *IEEE Trans. Information Theory*, vol. 51, no. 1, pp. 5-28, Jan. 2005.
- P. Viswanath, D. Tse and R. Laroia, “Opportunistic Beamforming using Dumb Antennas,” *IEEE Trans. Information Theory*, vol. 48, pp. 1277-94, June 2002

To Probe Further: Network Coding

- R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, “Network Information Flow”, *IEEE Trans. on Information Theory*, IT-46, pp. 1204-1216, 2000.
- R. Yeung, S. Y. Li, N. Cai, and Z. Zhang, “Network Coding Theory,” *Foundations and Trends in Communications and Information Theory*, vol. 2, no 4-5, pp. 241-381, 2005

To Probe Further: Scaling Laws Ad-hoc Networks

- F. Xue and P. R. Kumar, “Scaling Laws for Ad Hoc Wireless Networks: An information Theoretic Approach,” *Foundations and Trends in Networking*, vol 1, no. 2, 2006.

To Probe Further: IT and Estimation Theory

- D. Guo, S. Shamai, and S. Verdú, “Mutual Information and Minimum Mean-Square Error in Gaussian Channels,” *IEEE Trans. on Information Theory*, vol. 51, pp. 1261-1283, Apr. 2005.
- G. D. Forney, Jr., “Shannon meets Wiener II: On MMSE estimation in successive decoding schemes,” in *Proceedings 42nd Annual Allerton Conference on Communication, Control, and Computing*, Monticello, IL, USA, 2004. <http://arxiv.org/pdf/cs.IT/0409011>.
- T. Tao, “Szemerédi’s Regularity Lemma Revisited,” *Contrib. Discrete Math*.

To Probe Further: Denoising

- T. Weissman, E. Ordentlich, G. Seroussi, S. Verdú and M. Weinberger, “Universal Discrete Denoising: Known Channel,” *IEEE Trans. Information Theory*, vol. 51, no. 1, pp. 5-28, Jan. 2005.
- E. Ordentlich, G. Seroussi, S. Verdú, K. Viswanathan, “Channel decoding of systematically encoded unknown redundant sources,” 2006.

To Probe Further: IT and Control

- S. Tatikonda and S. Mitter, “Control Over Noisy Channels”, *IEEE Trans. on Auto. Control*, Vol 49, Issue 7, pp. 1196-1201
- N. C. Martins and M. A. Dahleh, “Feedback Control in the Presence of Noisy Channels: Bode-Like Fundamental Limitations of Performance.” *IEEE Trans. on Auto. Control*, submitted.
- N. Elia, When Bode meets Shannon: Control-Oriented feedback communication schemes, *IEEE Trans. on Auto. Control*, Vol 49, No 9, pp. 1477, September 2004
- A. Sahai and S. Mitter, “The Necessity and Sufficiency of Anytime Capacity for Stabilization of a Linear System Over a Noisy Communication LinkPart I: Scalar Systems” *IEEE Trans. Information Theory*, pp. 3369- 3395, Aug 2006

- S. Yuksel, T. Basar Coding and Control over Discrete Noisy Forward and Feedback Channels, *Proc. IEEE CDC/ECC'05*

To Probe Further: Analytic Information Theory

- W. Szpankowski, *Average Case Analysis of Algorithms on Sequences*, Wiley, New York, 2001.
- W. Szpankowski, J. Kieffer and E-H. Yang, “Problems on Sequences: Information Theory and Computer Science Interface”, *IEEE Trans. Information Theory*, 50, 1385-1392, 2004

To Probe Further: Communication Complexity

- J. Korner and A. Orlitsky, “Zero-error Information Theory,” *IEEE Trans. Information Theory*, Oct. 1998
- E. Kushilevitz, N. Nisan, *Communication Complexity*, Cambridge University Press, 1996
- R. G. Gallager, “Finding Parity in a Simple Broadcast Network”, *IEEE Transactions on Information Theory*, Vol. 34, No. 2, March 1988.
- N. Goyal, G. Kindler and M. Saks, “Lower Bounds for the Noisy Broadcast Problem,” FOCS 2005.

To Probe Further: IT and Economics

- T. Cover and J. Thomas, *Elements of Information Theory, 2nd Edition*, Chapter 16, Wiley, 2006
- C. Sims, "Implications of Rational Inattention," *Journal of Monetary Economics*, 50(3), 665-690.
- C. A. Sims: "Rational Inattention: A Research Agenda", *7th Deutsche Bundesbank Spring Conference*, Berlin 2005.

To Probe Further: Quantum Information Theory

- C. H. Bennett, P. W. Shor, “Quantum Information Theory”, *IEEE Transactions on Information Theory*, Vol. 44, No. 6, Oct 1998.

To Probe Further: Information Theory and Physics

- Tom Siegfried, *The Bit and the Pendulum: From Quantum Computing to M Theory-The New Physics of Information*, Wiley 2000

To Probe Further: Bio and Information Theory

- T. Berger, “Living Information Theory,” *IEEE Information Theory Newsletter*, March 2003

To Probe Further: IT and Optimization

- M. Chiang, “Geometric programming for communication systems”, *Foundations and Trends in Communications and Information Theory*, vol. 2, no. 1-2, pp. 1-154, July 2005.

To Probe Further: IT and Statistical Physics

- N. Surlas. “Spin-glass models as error-correcting codes”. *Nature*, 339:693-694, 1989.
- T. Tanaka “A Statistical-Mechanics Approach to Large-System Analysis of CDMA Multiuser Detectors,” *IEEE Trans. Inform. Theory*, vol. 48, no. 11, pp. 2888-2910, Nov. 2002
- D. Guo and S. Verdú, “Randomly Spread CDMA: Asymptotics via Statistical Physics,” *IEEE Trans. Information Theory*, vol. 51, no. 6, pp. 1983- 2010, June 2005.
- H. Nishimori. *Statistical Physics of Spin Glasses and Information Processing:An Introduction*. Oxford University Press, Oxford, UK, 2001
- C. Measson, A. Montanari, T. Richardson, and R. Urbanke. Life above threshold:from list decoding to area theorem and MSE. *Proc. ITW*, San Antonio, TX, October 2004

To Probe Further: IT and Random Matrices

- A. M. Tulino and S. Verdú, “Random Matrices and Wireless Communications,” *Foundations and Trends in Communications and Information Theory*, vol. 1, no. 1, pp. 1-182, June 2004.