Unreliable and Resource-Constrained Decoding

by

Lav R. Varshney

B.S., Electrical and Computer Engineering
Cornell University, 2004

S.M., Electrical Engineering and Computer Science
Massachusetts Institute of Technology, 2006

Electrical Engineer
Massachusetts Institute of Technology, 2008

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2010

Author .....
Department of Electrical Engineering and Computer Science
March 25, 2010

Certified by .....
Vivek K Goyal
Esther and Harold Edgerton Career Development
Associate Professor of Electrical Engineering and Computer Science
Thesis Supervisor

Certified by .....
Sanjoy K. Mitter
Professor of Electrical Engineering and Engineering Systems
Thesis Supervisor

Accepted by .....
Terry P. Orlando
Graduate Officer

Unreliable and Resource-Constrained Decoding

by

Lav R. Varshney

Submitted to the Department of Electrical Engineering and Computer Science
on March 25, 2010, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Electrical Engineering and Computer Science

Abstract

Traditional information theory and communication theory assume that decoders are noiseless and operate without transient or permanent faults. Decoders are also traditionally assumed to be unconstrained in physical resources like materiel, memory, and energy. This thesis studies how constraining reliability and resources in the decoder limits the performance of communication systems. Five communication problems are investigated. Broadly speaking these are communication using decoders that are wiring cost-limited, that are memory-limited, that are noisy, that fail catastrophically, and that simultaneously harvest information and energy. For each of these problems, fundamental trade-offs between communication system performance and reliability or resource consumption are established.

For decoding repetition codes using consensus decoding circuits, the optimal trade-off between decoding speed and quadratic wiring cost is defined and established. Designing optimal circuits is shown to be NP-complete, but is carried out for small circuit size. The natural relaxation to the integer circuit design problem is shown to be a reverse convex program. Random circuit topologies are also investigated.

Uncoded transmission is investigated when a population of heterogeneous sources must be categorized due to decoder memory constraints. Quantizers that are optimal for mean Bayes risk error, a novel fidelity criterion, are designed. Human decision making in segregated populations is also studied with this framework. The ratio between the costs of false alarms and missed detections is also shown to fundamentally affect the essential nature of discrimination.

The effect of noise on iterative message-passing decoders for low-density parity-check (LDPC) codes is studied. Concentration of decoding performance around its average is shown to hold. Density evolution equations for noisy decoders are derived. Decoding thresholds degrade smoothly as decoder noise increases, and in certain cases, arbitrarily small final error probability is achievable despite decoder noisiness. Precise information storage capacity results for reliable memory systems constructed from unreliable components are also provided.

Limits to communicating over systems that fail at random times are established. Communication with arbitrarily small probability of error is not possible, but schemes that optimize transmission volume communicated at fixed maximum message error probabilities are determined. System state feedback is shown not to improve performance.

For optimal communication with decoders that simultaneously harvest information and energy, a coding theorem that establishes the fundamental trade-off between the rates at which energy and reliable information can be transmitted over a single line is proven. The capacity-power function is computed for several channels; it is non-increasing and concave.

Thesis Supervisor: Vivek K Goyal
Title: Esther and Harold Edgerton Career Development
Associate Professor of Electrical Engineering and Computer Science

Thesis Supervisor: Sanjoy K. Mitter
Title: Professor of Electrical Engineering and Engineering Systems

Acknowledgments

It has been a great pleasure to have studied under Sanjoy Mitter and Vivek Goyal for the past several years and it is a high honor that I will be known as their student for time hereafter. Not only did they provide a warm and supportive environment in which to develop expertise, but also encouraged 'walking around'—both physically and metaphysically—so as to expose me to a wide range of epistemic and working cultures; this thesis is a product of both a rarefied home and a collection of ideas encountered elsewhere. Moreover, they have each been tireless advocates on my behalf. The contributions that Vivek and Sanjoy have made to this thesis and to my general education are very much appreciated.

I am grateful for the encouragement and suggestions provided by Dave Forney and Vladimir Stojanovic, the other members of my doctoral committee.

Research, at least the way I enjoy doing it, is an interactive endeavor. As such, the content of this thesis has been influenced by many besides my doctoral committee. The work in Chapter 3 benefited from conversations with José Moura, Mitya Chklovskii, Soummya Kar, João Xavier, Alex Olshevsky, and Pablo Parrilo. The work in Chapter 4 was carried out in large part with Kush Varshney and was improved through interactions with Alan Willsky, Daron Acemoglu, and Sendhil Mullainathan. The central ideas of Chapter 5 were developed through several enlightening discussions with Rüdiger Urbanke and Emre Telatar. The initial ideas for the problems studied in Chapters 6 and 7 came to mind when sitting at a synthetic biology bench with Team Biogurt and when taking an STS class with David Mindell, respectively. Figure 7-2 was created with assistance from Justin Dauwels.

Although works stemming from collaborations with Julius Kusuma, Vinith Misra, Mitya Chklovskii, Ha Nguyen, and Ram Srinivasan do not appear in this thesis, the interactions have also greatly enhanced my research experience.

An important part of graduate school for me was talking to fellow students. Besides those already mentioned, I would like to single out Adam Zelinski, John Sun, Demba Ba, Ahmed Kirmani, Dan Weller, Aniruddha Bhargava, and Joong Bum Rhim in STIR; Mukul Agarwal, Peter Jones, Barış Nakiboğlu, Mesrob Ohannessian, and Shashi Borade in LIDS; Urs Niesen, Charles Swannack, and Ashish Khisti in SIA; Pat Kreidl, Sujay Sanghavi, Emily Fox, Walter Sun, Vincent Tan, and Matt Johnson in SSG; Dinkar Vasudevan, Etienne Perron, and Shrinivas Kudekar at EPFL; Krish Eswaran and Bobak Nazer at Berkeley; and Quan Wen at CSHL.

Rachel Cohen and Eric Strattman provided unmatched administrative assistance.

Friends and family have provided immeasurable joy throughout this journey. Kush, who is not only my brother but also my friend, collaborator, and colleague, has influenced me and this thesis in myriad ways. My parents have given me strength and supported me every step of the way. ✉

- Financial support provided by the National Science Foundation Graduate Research Fellowship Program, Grant 0325774, Grant 0836720, and Grant 0729069.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Communication, computation, and decision systems process information, yet are constrained by their physical manifestations. There are costs in constructing, maintaining, operating, and disposing of physical systems. Moreover construction and operation are subject to the vagaries of noise. The traditional approach to studying information systems uses systems-theoretic concepts, distinguished by concern with mathematical properties rather than physical attributes. Motivated by several practical physical considerations, this thesis reopens and reexamines some of the traditional black boxes of systems-theoretic mathematical abstraction. Focus is placed on communication under resource constraints and in the presence of noise, not only in the channel but also elsewhere in the system.

The basic goal of communication is to transmit a message to a distant point through a noisy channel so that it may be recovered with acceptable fidelity while operating under resource constraints. The definition of fidelity depends on the desired end-to-end utility that a communication system is to provide. Constraints are specific to the physical resources that may be brought into use.

A communication system is said to solve a communication problem if it simultaneously meets fidelity and resource requirements when communicating a sufficiently large message. Much of information theory delimits which problems are solvable and which are not. Communication systems that operate at the edge of solvability are said to be optimal; communication theory is concerned with designing systems that are optimal or near-optimal. This thesis determines limits of solvability and finds optimal communication systems for several communication problems. In broad terms, these problems (with appropriate fidelity criteria) are:

- communication under wiring cost-limited consensus decoding,
- communication for a population of sources under memory-limited decoding,
- communication under noisy message-passing decoding,
- communication with decoders that fail catastrophically at random times, and
- communication with decoders that harvest both information and energy.

A point-to-point communication system typically has five parts: an information source, an encoder, a channel, a decoder, and an information destination, as depicted

Figure 1-1. Schematic diagram of a general point-to-point communication system, following [1, Figure 1].

in Figure 1-1. The channel is almost always thought to be noisy [1]. Physical properties of any of the five parts may govern whether a communication problem is solvable, but as noted, this thesis deals with limits imposed by decoder complexity, reliability, or energy.

The statistical problem of detection is to gain information from a mixture of the wanted signal and the unwanted noise, whether or not the source message is encoded with an error-control code. Modeling messages and signals as chance variables, denote the source message W, encoded signal X, received signal Y, and received message $\hat{W}$. The decoder tries to compute something useful from the received realization y. For example, the decoder may be tasked with producing a decoded signal $\hat{x}$ to minimize average symbol error probability with respect to X, or with producing a received message $\hat{w}$ to minimize a fidelity criterion such as Bayes risk or maximum error probability with respect to W.

By introducing redundancy to protect against channel noise, channel coding may increase the distance between signals in signal space and lead to better fidelity. A central goal of channel coding is to design an encoder-decoder pair that allows reliable communication over noisy channels at information rates close to the capacity promised by Shannon's noisy channel coding theorem [1]. In the quest for channel capacity, computational complexity of decoding has traditionally provided practical obstacles [2–4]. For example, a code chosen at random may achieve capacity, but require exponential decoding complexity.

Besides notions of complexity in an abstract computational model, decoder reliability or resource constraints may provide direct limits to practical channel coding. It has been suggested in other fields, such as cryptography when discussing adversary abilities [5], thermodynamics when discussing the abilities of Maxwell's demon [6, 7], and sensor networks when discussing the abilities of sensor nodes [8], that physical resources are more fundamental as constraints than computational ones. Results presented in this thesis demonstrate ways in which constrained decoder materiel and reliability may limit the performance of communication systems.

As listed above, several specific problems are considered under the general theme of unreliable and resource-constrained decoding. The basic motivations for these problems are briefly given here; fuller descriptions are provided in Section 1.2.

When manufacturing physical decoding circuits, one might be concerned about material costs in construction. Electronic circuits are built from costly logic gates, circuit components, wires, and heat sinks. Neural circuits are made from costly neural tissue. As an example, several works suggest that interconnect complexity, a form of wiring cost, is rather restrictive for implementing message-passing decoders for sparse graph codes [9]. Trading decoding performance for circuit simplicity would

help reduce the capital costs in building decoders.

Costs of provisioning information storage capacity on circuits may also be particularly limiting since memory is a physically-limited resource that impacts information processing performance. Indeed, memory registers consume much area on decoding circuits [10]. Reducing communication requirements to reduce information storage capacity at the decoder may expand the class of communication problems that are solvable.

In addition to infrastructure costs like wiring and memory, unreliability in operating decoding circuits is also constraining. Traditional information and coding theory implicitly assume that noise is localized in the channel and that the decoder operates without error. When computations are performed on faulty hardware, noise will manifest in the decoder. In fact Hamming's original development of parity-check codes [11] was motivated by applications in computing rather than in communication. Two kinds of decoder unreliability are studied in the thesis: transient and permanent.

Low-power decoding circuits are subject to transient noise for the same reasons as communication channels including thermal agitation, cosmic radiation, and interference from other signals. Techniques for modifying circuit operation such as voltage scaling can significantly reduce decoder power consumption but also reduce signal-to-noise ratios within the circuit. Nanoscale decoding circuits may additionally be subject to random quantum effects.

Catastrophic decoder failures may occur due to energy exhaustion, component failures [12], or adversarial actions. It is often impossible to fix decoders after failure [13] and so they cause one kind of permanent fault faced by communication systems. Unreliable manufacturing processes that produce devices different than designed lead to other kinds of permanent faults [14, 15].

Beyond their effect on decoder reliability, one might wonder whether energy constraints are also directly limiting. Surprisingly, there is no fundamental thermodynamic requirement for noiseless decoding to consume energy [16], however current technologies and any practical future technology will consume energy [17–20]. Given that energy is consumed by (noisy) decoders and energy limitations exacerbate decoder operation noise, it is of interest to determine the limits of embedding energy for decoding in the transmitted signal itself.

Section 1.1 reviews the relationship between engineering theory and engineering practice, describing how models mediate between the physical world and the world of mathematical abstraction. Section 1.2 summarizes the main contributions of the thesis.

1.1 Models and Fundamental Limits in Engineering Theory

Some have forcefully argued that "what every engineer needs is a good set of limit theorems" [21], say, limits to how much communication systems can be improved if all of the ingenuity of the engineer were brought to bear on the problem. Limit theorems, however, are deduced within engineering systems theories rather than within physical

theories.[1] They are properties of what Galileo called "machines in the abstract," as opposed to "machines in the concrete," characterized in structural and functional terms rather than in material ones.

An intellectual technology to mediate between the physical world and the world of mathematics is needed, a technology to convert large classes of machines in the concrete into a single machine in the abstract that is ripe for mathematical analysis. Systems-theoretic models—often represented in block diagram form—are precisely such a technology. Models convert inductive problems into deductive ones [23] by defining closed universes for deductive reasoning [24], where notions of fundamental limits and optimal systems are definable.

The abstraction of a physical engineering problem into a mathematical one requires that some properties of the 'real thing' are captured by the model. The process of abstraction must be so good that the black boxes of the model need not be reopened and reexamined.

Traditional models of communication systems take components to be noiseless, reliable, and unconstrained [25], but 'real' physical communication systems have noisy, unreliable, and constrained components. In moving from Hartley's early work [26] to the basic information theory problem [1, 27], Shannon introduced noise in the communication channel as well as transmit power constraints. In separate lines of work, fault-tolerant computing theory introduced noise into models of computational devices [14, 21, 28–33]. There has been little or no prior work on models of communication systems where the decoder is a faulty computer, because the decoder box of Figure 1-1 is thought to be reliable.

As given in the 2008 update of the International Technology Roadmap for Semiconductors (ITRS)[2] and elsewhere, recent technological trends imply that the traditional black boxes of information-theoretic modeling are inadequate for addressing future challenges. Decoder costs and reliability will be governing constraints.

This thesis takes steps in bridging the gap from noiseless and unconstrained models to noisy and resource-limited reality. Investigating novel communication system models leads to improved understanding of fundamental limits on the engineering problem of communication.

■ 1.2 Outline and Contributions

The central goal of the thesis is to move from the communication system of Figure 1-1 to the communication system of Figure 1-2, which has unreliable and resource-constrained decoding. Chapter 2 introduces noise into the channel by closing switch ② and reviews background material; switch ② remains closed throughout the rest of the thesis. Switches ③–⑦ are transitorily closed in Chapters 3–7, respectively, to study various extensions of the basic communication problem. Closing switch ③ leads to an explicit consideration of material cost limitations. Closing both switches labeled

[1] Physical theories are theories that detail natural laws, whereas engineering systems theories are theories of what can and cannot be built, a question that implies myriad engineering possibilities [22].

[2] The overall objective of the ITRS is to present the consensus of the semiconductor industry on the best current estimate of research and development needs for the next fifteen years.

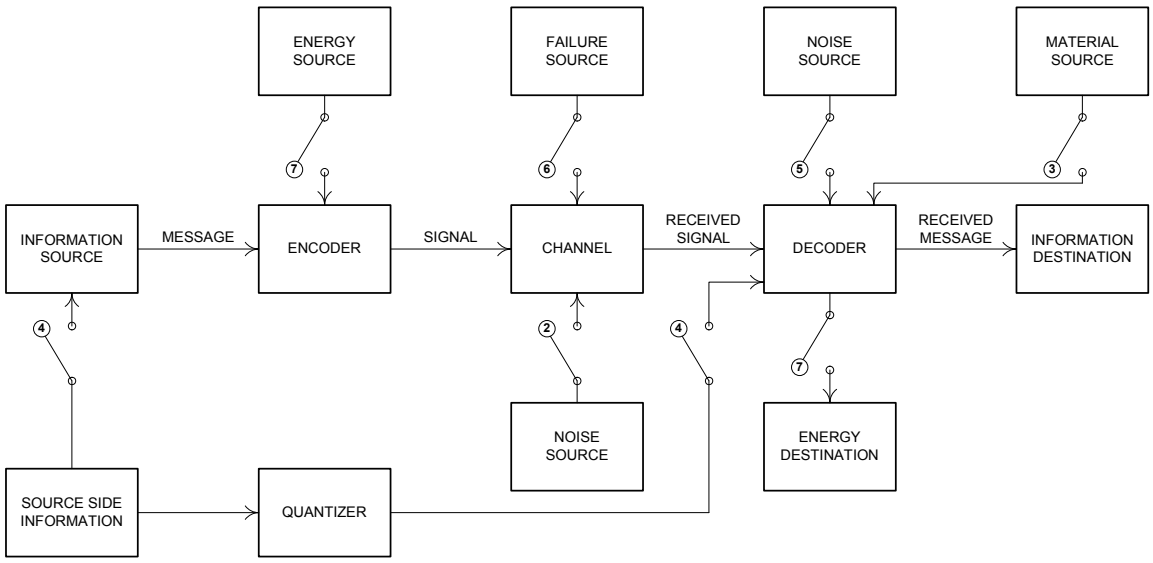

Figure 1-2. Schematic diagram of a point-to-point communication system with unreliable and resource-constrained decoding.

④ leads to limited memory for source side information. When switch ⑤ is closed, there is transient noise in decoding. Switch ⑥ is closed to introduce the possibility of catastrophic channel failure, generalizing decoder failure. Both switches labeled ⑦ are closed to consider energy transmission alongside information transmission. Chapter 8 concludes the thesis, briefly considering closing several of the switches simultaneously.

Chapter 2: Background

The background chapter reviews relevant results in information theory, coding theory, and statistical signal processing. In particular, the problem of reliable communication is described both for systems that use uncoded transmission and for systems that transmit coded signals. Performance criteria including maximal error probability, Bayes risk, and average error probability are defined. Some codes that achieve optimal performance under wholly unconstrained decoding are given. Finally, versions of the noisy channel coding theorem are stated [1].

Chapter 3: Infrastructure Costs—Wires

Integrated circuits are built from gates and the wires to connect them. The performances of these circuits depend on the electrical properties of the wiring such as resistance, conductance, and inductance [34]. Load for driving gates increases with wire capacitance; signal delay increases with wire resistance, capacitance, and inductance; and signal noise increases with inductive and capacitive coupling between wires. These electrical properties depend not only on the wire material, but also on geometric properties such as length [35]. The cost of wires, which may be much more restrictive in neuronal circuits [36] than in integrated circuits, also increases with length. The cost of a wire is often quadratic in its length [37, 38]. As such, it is desirable to have circuit topology and placement such that wires are short in length.

This chapter discusses the problem of communication using repetition codes with distributed consensus decoders [39] that have limited wiring length. Trade-offs between algebraic notions of decoding speed and algebraic notions of wiring cost are established. It is shown that separated topology design and node placement yields optimal circuit design. The design problem is shown to be NP-complete, but is carried out for small circuit size. A natural relaxation of the design problem is shown to be a reverse convex minimization problem. The performances of some random ensembles of decoders are studied.

The results of this chapter provide physical limits on the performance of decoders and may also have conceptual implications for the design and layout of future decoding circuits.

Chapter 4: Infrastructure Costs—Memory

Though memory is outwardly an informational resource, it is ultimately limited by physical resource constraints [40]. Limited decoder memory can reduce the fidelity of communication within informational communities where many different kinds of sources may want to communicate with the same destination. In such heterogeneous informational communities that additionally use uncoded communication, each possible source may yield a different optimal decoding metric. Since universal decoding methods [41] cannot be used with uncoded communication due to the lack of block length asymptotics, and since memory constraints may only allow a small finite set of decoding metrics to be used, mismatched decoding [42] may result. The size of the finite set determines the level of mismatch. This chapter considers communication for heterogeneous communities under memory-limited decoding.

In particular, when the source is selected at random from a population of sources, it is assumed that both the encoder and the decoder have access to perfect source state information, but that the decoder has limited adaptability due to the memory constraint. As a consequence, only a finite set of detection rules are used; the rule is almost surely mismatched if the family of sources is parameterized by an absolutely continuous random variable. Choosing the best subset of decoding metrics is a quantization problem, equivalent to quantizing prior probabilities for hypothesis testing over a population of objects. Nearest neighbor and centroid conditions are derived using mean Bayes risk error as a distortion measure for quantization. A high-resolution approximation to the distortion-rate function is obtained. Implications for human decision-making and the information economics of social discrimination are also presented.

Since any practical decoder has limited adaptability and universal decoders may not always be appropriate, results in this chapter give fundamental limits for heterogeneous uncoded transmission and hypothesis testing.

Chapter 5: Operation Reliability—Transient Faults

In communication theory, it is traditionally assumed that decoding algorithms perform without fault. Noise, however, provides a fundamental limit to computation systems just as it does to communication systems. This chapter determines limits in

processing noisy signals with noisy circuits. Performance analysis of noisy message-passing decoders for sparse graph codes shows that arbitrarily small error probability in communication is possible for some decoders at certain noise levels and is not for others.

Extending [43], concentration of decoder performance around its ensemble average performance is demonstrated even when noise is introduced into message-passing and local computation. Given this concentration result, density evolution equations for faulty iterative decoders are derived. In one model, computation of nonlinear estimation thresholds shows that performance degrades smoothly as decoder noise increases, however arbitrarily small probability of error is not achievable. In another model, probability of error may be driven to zero, with a decoding threshold that decreases smoothly with decoder noise. Thus, iterative decoding is robust to noise in the decoder. Comments on system optimization are also provided. In addition to the communication problem, the problem of constructing reliable memories from unreliable components is discussed.

The results in this chapter provide insights into the fundamental limits of processing unreliable signals with unreliable circuits.

Chapter 6: Operation Reliability—Permanent Faults

Separately from transient faults, decoders may also suffer permanent catastrophic failure due to energy exhaustion or component deterioration. Since it is slightly more general to model decoder failure as channel failure, this chapter establishes information-theoretic limits for channels that may fail at random times and presents optimal coding and decoding schemes.

Channels that fail at random times are finite-state semi-Markov channels. Communication over these channels with arbitrarily small probability of error is not possible. Making use of results in finite block length channel coding, sequences of block lengths that optimize transmission volume communicated at fixed maximum message error probabilities are determined. A dynamic programming formulation shows that channel state feedback does not improve performance.

Since any communication system will eventually fail when viewed at a suitably long timescale, the results of this chapter are of interest in many settings.

Chapter 7: Operation Costs—Energy

The penultimate chapter starts by describing the notion of reversible computing in the thermodynamics of computation [16], showing that a decoder need not dissipate energy if it is noiseless and has no memory constraints. As argued in Chapters 4–6, however, decoders are unreliable and memory-limited and must dissipate energy. To counter energy shortages, a transmitter might embed energy in the communication signal to help the receiver read the message.

The remainder of the chapter establishes the fundamental trade-off between the rates at which energy and reliable information can be transmitted over a single noisy line. A capacity-power function is defined and a coding theorem is given. The capacity-power function is a non-increasing concave $\cap$ function. Capacity-power

functions for several channels are computed. A suboptimal approach of time-sharing between energy reception and information reception, which requires simpler decoding circuitry, is also discussed.

This chapter reverses the view of a signal as separate from the communication medium over which it is transmitted, an abstraction established during the pre-history of information theory [44]. This new model of energy and information flowing together may be particularly important for future distributed energy/information systems.

Chapter 8: Conclusion

The conclusion chapter first recapitulates the main ideas and results presented in the thesis. Then some directions for extending the thesis are given. Finally, the chapter discusses analyzing and optimizing communication systems that are simultaneously subject to several of the constraints and sources of unreliability presented in the preceding chapters.

Bibliographical Note

Parts of Chapter 3 appear in the paper:

- L. R. Varshney, "Distributed Inference Networks with Costly Wires," to appear in *Proceedings of the 2010 American Control Conference*, June 2010.

Parts of Chapter 4 appear in papers:

- K. R. Varshney and L. R. Varshney, "Quantization of Prior Probabilities for Hypothesis Testing," *IEEE Transactions on Signal Processing*, vol. 56, pp. 4553–4562, Oct. 2008.

- K. R. Varshney and L. R. Varshney, "Minimum Mean Bayes Risk Error Quantization of Prior Probabilities," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 3445–3448, Apr. 2008.

and in the journal manuscript:

- L. R. Varshney and K. R. Varshney, "Decision Making with Quantized Priors Leads to Discrimination."

Parts of Chapter 5 appear in the paper:

- L. R. Varshney, "Performance of LDPC Codes Under Noisy Message-Passing Decoding," in *Proceedings of the IEEE Information Theory Workshop*, pp. 178–183, Sept. 2007.

and in the journal manuscript:

- L. R. Varshney, "Performance of LDPC Codes Under Faulty Iterative Decoding," arXiv:0806.1215 [cs.IT].

Parts of Chapter 6 appear in the paper:

- L. R. Varshney, S. K. Mitter, and V. K. Goyal, "Channels That Die," in *Proceedings of the Forty-Seventh Annual Allerton Conference on Communication, Control, and Computing*, Sept. 2009.

Parts of Chapter 7 appear in the paper:

- L. R. Varshney, "Transporting Information and Energy Simultaneously," in *Proceedings of the IEEE International Symposium on Information Theory*, pp. 1612–1616, July 2008.

Chapter 2

Background

The central task of communication decoders is to convert received signals into received messages. This background chapter first discusses the general problem of reliable communication and describes several performance criteria that are used to judge communication quality. Next, codes that improve communication quality when used with reliable, unconstrained decoders are reviewed. Finally, implementation of decoders is briefly discussed and the noisy channel coding theorem is presented.

Several texts on statistical signal processing [45, 46], coding theory [2, 47–49], and information theory [50–54] cover the material in this chapter much more deeply and completely, from several different perspectives.

2.1 The Problem of Reliable Communication

The goal of reliable communication is to transmit a message to a distant point through a noisy channel so that it may be recovered with high fidelity. As drawn in the schematic diagram Figure 1-1, there are five parts to a communication system.

The source message is modeled in probabilistic terms as a sequence of chance variables, $W_1^k = (W_1, W_2, \ldots, W_k)$ which are drawn from a common alphabet $\mathcal{W}$ according to a probability distribution $p_{W_1^k}(\cdot)$, defined for all k.[1]

The encoder converts the source message into a block of n channel symbols and is specified as a sequence of transition probability assignments $p_{X_1^n|W_1^k}(\cdot|\cdot)$ defined for all n. The output of the encoder is a sequence of channel input chance variables, $X_1^n = (X_1, X_2, \ldots, X_n)$ which are drawn from a common alphabet $\mathcal{X}$. Often the encoder is a sequence of deterministic maps, rather than a sequence of non-degenerate transition probability assignments.

The channel is defined mathematically as a sequence of transition probability assignments between the channel input space and the channel output space, $p_{Y_1^n|X_1^n}(\cdot|\cdot)$, for all n. The channel output chance variables, $Y_1^n = (Y_1, Y_2, \ldots, Y_n)$ are drawn from the common alphabet $\mathcal{Y}$.

There are two kinds of decoders that are employed for different but related purposes. One decodes messages whereas the other decodes signals. A message decoder is specified as a sequence of transition probability assignments $p_{\hat{W}_1^k|Y_1^n}(\cdot|\cdot)$ for all n.

[1] There are communication scenarios that assume semi-infinite sequences of source symbols, to be encoded into a semi-infinite sequences of channel symbols using tree codes [52, Chapter 10]. This thesis restricts attention to the block coding setting.

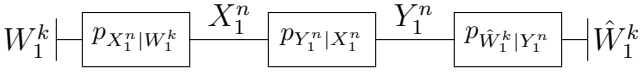

Figure 2-1. Factor graph of communication system with message decoding.

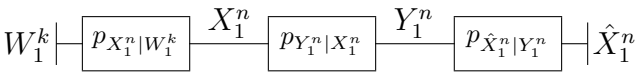

Figure 2-2. Factor graph of communication system with signal decoding.

The information destination receives a sequence of reconstruction chance variables, $\hat{W}_1^k = (\hat{W}_1, \hat{W}_2, \ldots, \hat{W}_k)$ which are drawn from the common alphabet $\hat{\mathcal{W}}$. On the other hand, a signal decoder tries to recover the transmitted codeword symbols rather than the original message. It is specified as a sequence of transition probability assignments $p_{\hat{X}_1^n|Y_1^n}(\cdot|\cdot)$ for all n. The information destination for a signal decoder receives a sequence of chance variables, $\hat{X}_1^n = (\hat{X}_1, \hat{X}_2, \ldots, \hat{X}_n)$ which are drawn from a common alphabet $\hat{\mathcal{X}}$. Just like encoders, decoders may be deterministic.

As described, the source message W_1^k, encoded signal X_1^n, received signal Y_1^n, and received message $\hat{W}_1^k$ (or decoded signal $\hat{X}_1^n$) chance variables obey the Markov information pattern implicit in Figure 1-1: $W_1^k \leftrightarrow X_1^n \leftrightarrow Y_1^n \leftrightarrow (\hat{W}_1^k, \hat{X}_1^n)$. Adopting a factorization view of communication, as in [55, Section 2.5] and elsewhere, one can convert directed block diagrams of point-to-point communication systems into Forney-style factor graphs [56, Section III.D], Figures 2-1 and 2-2. The half-edges and edges represent variables whereas the boxes represent probabilistic constraints. The undirected factor graph gives a behavioral view of a communication system [57], rather than an input-output view.

Among the three factor constraints in a communication system, the encoder and decoder are usually thought to be under design and the channel is thought to be fixed. In several sensing settings, the encoder is fixed as an identity map [58–60], leading to uncoded transmission. In other contexts, the channel might be designable [61]. In all systems considered in the thesis, the decoder is open to design under resource and reliability limitations. The decoder is designed to use the factor constraints and the received signal to determine what was transmitted.

■ 2.2 Bayes Risk and Error Probabilities

To define criteria for judging communication performance, first consider systems that have just one source symbol, $k = 1$, and employ message decoding rather than signal decoding. Restriction to $k = 1$ does not lead to much loss of generality, since a super-alphabet of size $|\mathcal{W}|^k$ can be defined for any k.

One might define a general distortion function between W and $\hat{W}$ and measure its expected value, maximum value, or probability of exceeding a fixed threshold.

Indeed, this is the basis for rate distortion-theoretic investigations of communication [55, 62, 63]. In this thesis, specific notions of fidelity are considered.

Bayes risk and various error probabilities are typically defined for cases when $\hat{\mathcal{W}} = \mathcal{W}$. For pairs of letters $(w, \hat{w}) \in \mathcal{W} \times \mathcal{W}$, there is a non-negative Bayes cost c_{ij} incurred when $(w = i, \hat{w} = j)$ is realized. The Bayes risk J of a communication system is the expected value of the Bayes cost:

$$J = E_{W,\hat{W}}[c] = \sum_{i=1}^{|\mathcal{W}|} \sum_{j=1}^{|\mathcal{W}|} c_{ij} \Pr[W = i, \hat{W} = j].$$

The average probability of message error, $P_e^{\text{avg}} = \Pr[\hat{W} \neq W]$, is a special case of Bayes risk when

$$c_{ij} = \begin{cases} 0 & i = j \\ 1 & i \neq j. \end{cases}$$

The average message error probability may also be written in terms of conditional message error probabilities $\lambda_w = \Pr[\hat{W} \neq w | W = w]$:

$$P_e^{\text{avg}} = E_W \left[\Pr[\hat{W} \neq w | W = w] \right] = E_W[\lambda_w].$$

The maximum message error probability is a worst-case alternative to average error probability:

$$P_e^{\max} = \max_{w \in \mathcal{W}} \lambda_w.$$

Clearly $P_e^{\text{avg}} \leq P_e^{\max}$.

Some communication systems perform signal decoding to produce an estimate $\hat{X}_1^n$ of the transmitted signal X_1^n and then use a surjective mapping to produce $\hat{W}$. For optimal (and reasonable non-optimal) communication systems of this type, the average message error probability is also given by

$$P_e^{\text{avg}} = E_{X_1^n} \left[\Pr[\hat{X}_1^n \neq x_1^n | X_1^n = x_1^n] \right].$$

Without need for the surjective mapping, average symbol error probability may be defined:

$$P_e^{\text{sym}} = \frac{1}{n} \sum_{i=1}^{n} \Pr[\hat{X}_i \neq x_i | X_i = x_i].$$

Clearly $P_e^{\text{sym}} \leq P_e^{\text{avg}} \leq P_e^{\max}$.

When there are several source message symbols, $k > 1$, a stringent maximax error requirement for message decoding may be defined:

$$P_e^{\text{maximax}} = \max_{i \in \{1,2,\ldots,k\}} P_e^{\max}(i),$$

where $P_e^{\max}(i)$ is the maximum message error probability for message i.

Definitions of error probability directly generalize to cases when $\mathcal{W} \subset \hat{\mathcal{W}}$, however

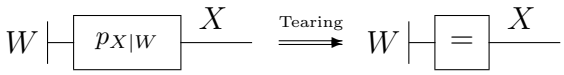

Figure 2-3. Factor graph of uncoded transmission.

there are alternatives to the natural generalization. One widely adopted set of definitions stem from *incomplete* or *errors-and-erasures* decoding [2,64], where $\hat{\mathcal{W}} = \mathcal{W} \cup \ominus$ and the special output symbol $\ominus$ indicates a declared message erasure. A declared message erasure is not counted as an error.

■ 2.3 Codes and Optimal Decoders

This section examines codes by 'tearing' [57] the encoder factor constraint in Figures 2-1 and 2-2 into specific code constraints. Discussion of decoding procedures that optimize the performance criteria developed in Section 2.2 are interleaved throughout the development.

Uncoded Transmission

The starting point of the discussion is uncoded transmission, where $\mathcal{W} = \mathcal{X}$, $n = 1$, and $p_{X|W}$ is an identity map. The factor graph for the encoder is shown in Figure 2-3, where $\boxed{=}$ indicates an equality constraint.

The decoder that minimizes Bayes risk for such an uncoded transmission system follows from optimal hypothesis testing [46]. To develop the main idea, it is easiest to restrict attention to the binary case, $\mathcal{W} = \{0, 1\}$.

There is a message prior probability, p_W, which is alternatively denoted by $p_W(0) = P_0$ and $p_W(1) = 1 - P_0$. The channel properties yield likelihood functions for the two possible messages: $p_{Y|W}(y|0)$ and $p_{Y|W}(y|1)$. The decoder is specified as a deterministic decision rule, $f_D(\cdot)$, that maps every possible received signal Y to one of the two messages, i.e., $f_D : \mathcal{Y} \mapsto \{0, 1\}$, thereby partitioning $\mathcal{Y}$ into two disjoint decision regions. More generally $f_D(\cdot)$ is the transition probability assignment $p_{\hat{W}|Y}$, with $\hat{\mathcal{W}} = \mathcal{W}$, but it can be shown that deterministic tests are always sufficient to achieve optimal performance [46, Section 2.6.1].

For fixed Bayes costs, c_{00}, c_{01}, c_{10}, and c_{11}, the decision rule is chosen to be the one that minimizes the Bayes risk:

$$f_D(\cdot) = \arg\min_{f(\cdot)} J(f),$$

where $J(f)$ is the Bayes risk of the communication system under decision rule f. The optimal decision rule f_D is the likelihood ratio test:

$$\frac{p_{Y|W}(y|1)}{p_{Y|W}(y|0)} \underset{f_D(y)=0}{\overset{f_D(y)=1}{\gtrless}} \frac{P_0(c_{10} - c_{00})}{(1 - P_0)(c_{01} - c_{11})}. \tag{2.1}$$

When the two quantities compared in decision-making are equal, either choice yields the same Bayes risk. The decision rule that optimizes for average error probability reduces to

$$\frac{p_{Y|W}(y|w=1)}{p_{Y|W}(y|w=0)} \underset{f_D(y)=0}{\overset{f_D(y)=1}{\gtrless}} \frac{P_0}{1-P_0}.$$

When $|\mathcal{W}|$ is greater than two, the same basic ideas hold. The optimal decoding rule can be structured as a procedure of eliminating one possible message at a time through a sequence of $|\mathcal{W}|-1$ binary comparisons similar in form to likelihood ratio tests [46, Section 2.8]. A simple example of optimal decoding is as follows.

Example 2.1. *Consider $\mathcal{W}=\{0,1\}$ with equiprobable W and a unit-variance additive white Gaussian noise communication channel used for uncoded transmission. The likelihood functions are:*

$$p_{Y|W}(y|w=0)=\frac{e^{-\frac{1}{2}y^2}}{\sqrt{2\pi}}$$

and

$$p_{Y|W}(y|w=1)=\frac{e^{-\frac{1}{2}(y-1)^2}}{\sqrt{2\pi}}.$$

The decoder is charged with minimizing Bayes risk with Bayes costs $c_{00}=c_{11}=0$ and $c_{01}=c_{10}=1$. Bayes risk with these costs is the average message error probability. Now the optimum decision rule is:

$$y \underset{f_D(y)=0}{\overset{f_D(y)=1}{\gtrless}} \tfrac{1}{2}.$$

The performance of this optimum rule is computed first in terms of the false alarm probability $P_e^I=\Pr[f_D(Y)=1|W=0]$ and the missed detection probability $P_e^{II}=\Pr[f_D(Y)=0|W=1]$.

$$P_e^I=\int_{1/2}^{\infty}\frac{e^{-\frac{1}{2}y^2}}{\sqrt{2\pi}}dy$$

and

$$P_e^{II}=1-\int_{1/2}^{\infty}\frac{e^{-\frac{1}{2}(y-1)^2}}{\sqrt{2\pi}}dy.$$

These error probabilities can be expressed in terms of the Q-function

$$Q(z)=1-Q(-z)=\frac{1}{\sqrt{2\pi}}\int_z^{\infty}e^{-\frac{1}{2}\zeta^2}d\zeta \tag{2.2}$$

as

$$P_e^I=Q\left(\tfrac{1}{2}\right)$$

and

$$P_e^{II}=1-Q\left(-\tfrac{1}{2}\right)=Q\left(\tfrac{1}{2}\right).$$

Thus the best Bayes risk possible in this communication problem is:

$$\begin{aligned} J &= p_W(w=0)\left[c_{00}(1-P_e^{II}) + c_{01}P_e^{II}\right] + p_W(w=1)\left[c_{11}(1-P_e^{I}) + c_{10}P_e^{I}\right] \\ &= \tfrac{1}{2}P_e^{II} + \tfrac{1}{2}P_e^{I} = Q\left(\tfrac{1}{2}\right). \end{aligned}$$

One may further note that in this example, $J = P_e^{\text{avg}} = P_e^{\max}$.

Repetition Codes

In uncoded transmission, error-control codes are not employed. The encoding process, however, may allow channel errors to be more easily detected and corrected by the decoder. In particular, embedding signals in higher dimensional spaces may increase distances between them. The simplest kind of code is a repetition code.

In a repetition code of length n, $p_{X_i|W}$ is an identity map for each $i = 1, 2, \ldots, n$, as depicted in the factor graph shown in Figure 2-4. The same basic optimal Bayesian hypothesis testing decoding procedure may be used as in the uncoded case. For some channels with $\mathcal{Y} = \mathcal{X}$, the optimal decoding rule reduces to a majority vote over the channel output coordinates.

In general, there can be several factor graphs that correspond to a given code. For example, a repetition code of length 8 can be drawn in tail-biting trellis form or tree form, as shown in Figure 2-5, or other forms. As will become apparent in later chapters, the choice of factor graph with an associated natural decoding algorithm can significantly impact decoding performance and decoder complexity. General factor graph synthesis is an open problem [65].

The following example shows that repetition coding can reduce error probability.

Example 2.2. *Consider the same source and channel as in Example 2.1 but using a repetition code of length* n*. The optimal decision rule is:*

$$\frac{1}{n}\sum_{i=1}^{n} y_i \underset{f_D(y)=0}{\overset{f_D(y)=1}{\gtrless}} \frac{1}{2}.$$

The error probabilities of this scheme are:

$$P_e^I = P_e^{II} = Q\left(\frac{n/2}{\sqrt{n}}\right) = Q\left(\frac{\sqrt{n}}{2}\right).$$

Therefore the average and maximum message error probabilities are:

$$P_e^{\max} = P_e^{\text{avg}} = Q\left(\frac{\sqrt{n}}{2}\right).$$

Since $Q(z)$ *is the tail probability of a standard Gaussian distribution, it decreases in* z *and goes to zero as* $z \to \infty$*. Therefore,* $P_e^{\max} \to 0$ *and* $P_e^{\text{avg}} \to 0$ *as* $n \to \infty$*.*

Notice that the optimal decision rule compares the average value of the channel output realizations to a fixed threshold. Such an optimal decision rule can be

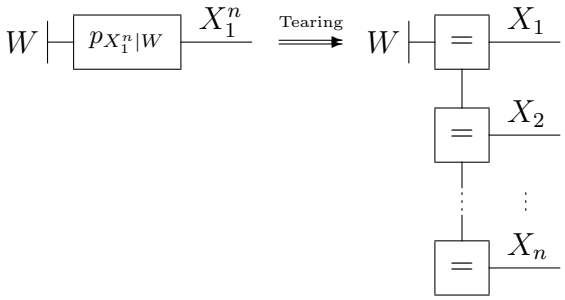

Figure 2-4. Factor graph of repetition coding.

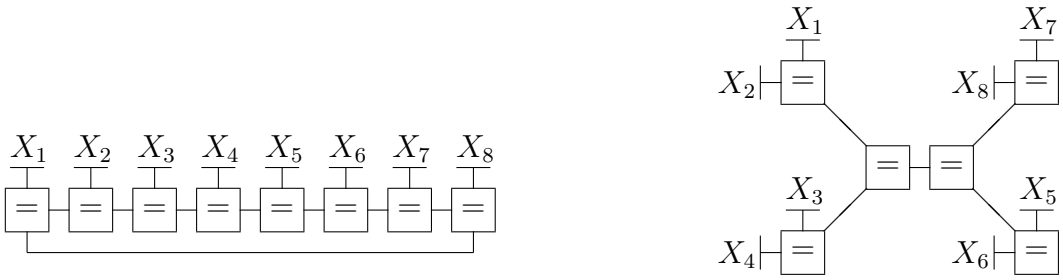

Figure 2-5. Two distinct factor graphs of repetition coding.

implemented using consensus decoding, as in Chapter 3.

Increasing the length of a repetition code drives the error probability to zero for communication systems with almost any kind of practical channel under optimal decoding, not just in Example 2.2.

Linear Codes

In general, codes are subsets of $\mathcal{X}^n$. The difficulty in code design is in choosing good subsets. Repetition codes are the simplest of codes, however there are more complicated ones that may be more useful.

An example of a code is a binary Hamming code of length $n = 7$ [11], shown in Table 2.1. The code is a subset of 16 sequences from the 128 sequences possible in $\{0, 1\}^7$. In channel coding, there are two other standard notions of quality besides fidelity: block length and number of messages represented. These three main communication system parameters are:

- block length n,
- number of messages M, and
- fidelity J, P_e^{avg}, $P_e^{\max}$, or P_e^{sym}.

For the Hamming code, the block length is $n = 7$ and the number of messages is $M = 16$.

0000000
0001011
0010110
0011101
0100111
0101100
0110001
0111010
1000101
1001110
1010011
1011000
1100010
1101001
1110100
1111111

Table 2.1. An ($n = 7$) binary Hamming code.

Many good codes are linear codes [2,66], drawn from a channel input alphabet $\mathcal{X}$ that is a finite field $\mathbb{F}_{|\mathcal{X}|}$. Linear codes satisfy certain linear constraints that may be expressed in terms of parity-check matrices or generator matrices and form a subspace of $\mathcal{X}^n$. In Forney-style factor graphs, the constraints may be expressed graphically using parity-check constraints, $\boxed{+}$, and equality constraints. Parity-check constraints enforce even parity among the variables corresponding to connected edges. Parity-check constraints are also called check nodes and through some misappropriation of terminology, equality constraints are also called variable nodes.

A linear code that maps k source symbols into n channel symbols can be completely described by any set of k linearly independent codewords of length n. Arranging these basis codewords into a $k \times n$ matrix yields the generator matrix G. The generator matrix is often useful for encoding.

A linear code may also be described in terms of a parity-check matrix. Since the code is a linear subspace of dimension k in an ambient space of dimension n, it has an orthogonal linear subspace of dimension $n - k$, which is called the dual code. The generator matrix for the dual code is an $(n - k) \times n$ matrix, which is called the parity-check matrix H for the original code. The parity-check matrix is often useful for decoding.

The $(n = 7, M = 16)$ binary Hamming code above is linear and has generator matrix

$$G = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}$$

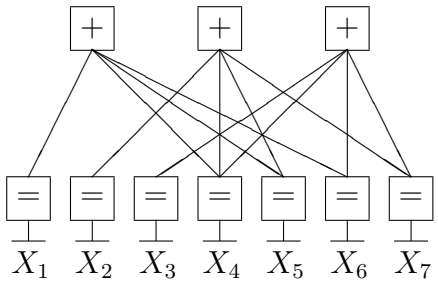

Figure 2-6. Factor graph of ($n = 7$) binary Hamming code.

and parity-check matrix

$$H = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}.$$

The parity-check matrix can be used to draw the factor graph shown in Figure 2-6, again tearing the general encoding constraints into specific code constraints.

Optimal decoding for linear codes is still specified by optimal Bayesian hypothesis testing for an alphabet of size $M = |\mathcal{W}|^k$. If all messages are equiprobable, then maximum likelihood (ML) decoding is optimal.

Low-Density Parity-Check Codes

A particular subclass of linear codes are the low-density parity-check (LDPC) codes [43, 49, 67]. As evident in name, the parity-check matrices of these codes are sparse and they equivalently have sparse factor graphs. In fact, they are typically defined through their factor graphs. Recall, however, that although a factor graph determines a code, the opposite is not true [65].

The standard ensemble of $(d_{\mathsf{v}}, d_{\mathsf{c}})$-regular binary LDPC codes of block length n, $\mathcal{C}^n(d_{\mathsf{v}}, d_{\mathsf{c}})$, is defined by a uniform measure on the set of labeled bipartite factor graphs with variable node degree d_{v} and check node degree d_{c}. There are n variable nodes corresponding to the codeword letters and $nd_{\mathsf{v}}/d_{\mathsf{c}}$ check nodes corresponding to the parity-check constraints.

The designed codebook size of the code is $M = 2^{n(1-d_{\mathsf{v}}/d_{\mathsf{c}})}$, though the actual codebook size might be higher since not all checks may be independent; the true size converges to the design size for large n [49, Lemma 3.22].

One may also consider irregular codes, $\mathcal{C}^n(\lambda, \rho)$ characterized by the degree distribution pair (λ, ρ). Generating functions of the degree distributions, $\lambda(\zeta)$ and $\rho(\zeta)$, are functions defined to be $\lambda(\zeta) = \sum_{i=2}^{\infty} \lambda_i \zeta^{i-1}$ and $\rho(\zeta) = \sum_{i=2}^{\infty} \rho_i \zeta^{i-1}$, where λ_i and ρ_i specify the fraction of edges that connect to nodes with degree i. The design size is $2^{n\left(1-\int_0^1 \rho(\zeta)d\zeta / \int_0^1 \lambda(\zeta)d\zeta\right)}$.

A particular LDPC code is typically chosen at random from an ensemble of LDPC codes. As for other linear codes, optimal Bayesian decoding procedures or ML decoding procedures may be developed.

Perfect Codes

Another class of codes that arise in communication are perfect codes [47, Chapter 7]. A perfect code is one for which there are equal-radius spheres centered at the codewords that are disjoint and that completely fill $\mathcal{X}^n$. The repetition codes, the binary $(n = 7, M = 16)$ Hamming code, and the binary linear $(n = 23, M = 4096)$ Golay code [68] are the only linear perfect codes. There are a few more nonlinear perfect codes [69, 70].

The optimal Bayesian decoding procedure for perfect codes is often very simple.

Modifying Codes

Since the design of good codes is difficult, modifying existing codes is often useful. For linear codes, there are six basic ways of modifying a code. Letting $r = n - k$, these are:

- *Extending.* Fix k, increase n, increase r. Corresponds to adding $\boxed{=}$ and $\boxed{+}$ vertices and new edges to connect them to existing vertices in a factor graph.
- *Puncturing.* Fix k, decrease r, decrease n. Corresponds to deleting $\boxed{=}$ and $\boxed{+}$ vertices and all edges connecting them to remaining vertices in a factor graph.
- *Lengthening.* Fix r, increase n, increase k. Corresponds to adding $\boxed{=}$ vertices and new edges to connect them to existing $\boxed{+}$ vertices in a factor graph.
- *Shortening.* Fix r, decrease n, decrease k. Corresponds to deleting $\boxed{=}$ vertices and all edges connecting them to $\boxed{+}$ vertices in a factor graph.
- *Augmenting.* Fix n, increase k, decrease r. Corresponds to deleting $\boxed{+}$ vertices and all edges connecting them to $\boxed{=}$ vertices in a factor graph.
- *Expurgating.* Fix n, decrease k, increase r. Corresponds to adding $\boxed{+}$ vertices and new edges to connect them to existing $\boxed{=}$ vertices in a factor graph.

Augmenting and expurgating are opposite operations, as are lengthening–shortening and extending–puncturing. Lengthening is dual to extending, whereas puncturing is dual to shortening.

Many of these code modification procedures also apply to nonlinear codes. For example, augmenting adds extra codewords to a general codebook, whereas expurgating removes codewords from a codebook.

Optimal Decoding Performance

The optimal decoding procedure for communication systems that use codes is given by optimal Bayesian hypothesis testing, but computing the precise error probabilities of systems that use complicated coding schemes is often difficult.

The optimal error probability can be computed in the special case of perfect, linear codes transmitted over discrete memoryless, symmetric channels, with codewords

chosen equiprobably. In this setting, the optimal decoding rule reduces to choosing the codeword closest to the received output sequence, measured using Hamming distance [53, Problem 2.13]. Identical computations also hold for linear codes that are not perfect, but that use errors-and-erasures decoding based on decoding spheres.

To compute error probability, it is convenient to determine the distances between various codewords. For linear codes, the multiset of distances from a given codeword to all other codewords is identical to the multiset of distances from any other given codeword to all other codewords, including the all-zero codeword (which is part of any linear code). Thence, considering the profile of the codeword Hamming weights is sufficient. Let A_ℓ denote the number of codewords of weight ℓ in a linear code. The $(n+1)$-dimensional vector with components A_ℓ is called the weight distribution of the code. For the binary Hamming code, this vector is

$$[1, 0, 0, 7, 7, 0, 0, 1]\,.$$

For the binary Golay code, this vector is

$$[1, 0, 0, 0, 0, 0, 0, 253, 506, 0, 0, 1288, 1288, 0, 0, 506, 253, 0, 0, 0, 0, 0, 0, 1]\,,$$

as computed exhaustively [2, Table 5.3].

Consider channels that make independent errors with probability ε in each component and transmit correctly with probability $1-\varepsilon$. Each of the $|\mathcal{X}|-1$ wrong symbols occur with probability $\varepsilon/(|\mathcal{X}|-1)$. Each pattern of h errors has probability:

$$p(h) = \left(\frac{\varepsilon}{|\mathcal{X}|-1}\right)^h (1-\varepsilon)^{n-h}$$

The sphere radius of a perfect code is denoted ρ. Overloading notation, the decoding radius of errors-and-erasures decoding for other linear codes is also fixed at ρ such that $2\rho+1$ is less than the minimum distance of the code.

The probability of correct message reception is:

$$\sum_{v=0}^{\rho} \binom{n}{v} \varepsilon^v (1-\varepsilon)^{n-v},$$

computed by finding the probability of a codeword only being perturbed to within its sphere of radius ρ.

Let $N_\ell^h(s)$ be the number of error patterns of weight h that are at distance s from a codeword of weight ℓ. Using type-counting combinatorics, this is:

$$N_\ell^h(s) = \sum_{\substack{0\le i\le n\\ 0\le j\le n\\ i+2j+h=s+1}} \binom{n-\ell}{j+h-\ell}\binom{\ell}{i}\binom{\ell-i}{j}(|\mathcal{X}|-1)^{j+h-1}(|\mathcal{X}|-2)^i.$$

Since decoding is limited to ρ apparent errors, the probability of error is

$$P_e^{\max} = P_e^{\text{avg}} = \sum_{h=0}^{n} \left(\frac{\varepsilon}{|\mathcal{X}| - 1} \right)^h (1-\varepsilon)^{n-h} \sum_{s=0}^{\rho} \sum_{\ell=1}^{n} A_\ell N_\ell^h(s).$$

For sphere-based errors-and-erasures decoding, there is a further erasure probability. For complete decoding of perfect codes, on the other hand, error and correct reception events partition the space.

Example 2.3. *Consider equiprobably choosing a codeword from the* $(n = 7, M = 16)$ *binary Hamming code and transmitting it over a binary symmetric channel with crossover probability* ε*. If the all-zero codeword is transmitted, the probability of correct reception under optimal decoding is:*

$$(1-\varepsilon)^7 + 7\varepsilon(1-\varepsilon)^6,$$

and so the probability of error is

$$P_e^{\max} = P_e^{\text{avg}} = 1 - (1-\varepsilon)^7 - 7\varepsilon(1-\varepsilon)^6.$$

This can alternatively be verified using the more extensive formula involving A_ℓ*.*

Implementing Decoders

This section has described codes and optimal decoders for them. For short algebraic codes, optimal decoders may be implementable [71, 72], but often implementation of optimal decoders and analysis of optimal performance is very difficult. To ease these difficulties, upper and lower bounds as well as suboptimal decoders have been developed. The various suboptimal decoders include joint typicality decoding used for proving theorems in information theory, iterative message-passing decoding for low-density parity-check codes [67] and optimization-based decoding for low-density parity-check codes [73,74]. Many of these decoding algorithms are based on the factor graphs of the encoding constraints. Moreover, suboptimal decoding methods are often sufficient to prove the information-theoretic coding theorems stated next.

■ 2.4 Information-Theoretic Limits

Information-theoretic limits on the fundamental trade-offs among the three block coding parameters (n, M, P_e), as a function of channel noise properties, are given in this section. For simplicity, the section is restricted to discrete, memoryless systems with message decoding and $k = 1$. Many of the stated results can be generalized to less restrictive scenarios, see e.g. [75, 76].

A discrete memoryless channel (DMC) is characterized by a finite input alphabet $\mathcal{X}$, a finite output alphabet $\mathcal{Y}$, and the transition probability assignment $p_{Y|X}(y|x)$. Since the channel is memoryless, its block length n extension is an n-fold product distribution. An (n, M) block code, t, for a DMC is defined by a deterministic

encoding function which maps messages to channel inputs:

$$f_E^t : \{1, \ldots, M\} \mapsto \mathcal{X}^n$$

and a decoding function which maps channel outputs to messages:

$$f_D^t : \mathcal{Y}^n \mapsto \{1, \ldots, M\}.$$

The set $\{1, \ldots, M\}$ is the message set $\mathcal{W}$. The sequences $f_E^t(1), \ldots, f_E^t(M)$ in $\mathcal{X}^n$ are codewords and the set of these is the codebook.

Channel inputs are described by chance variables X_1^n, where $X_1^n = f_E^t(W)$. The corresponding outputs are the chance variables Y_1^n. Recovered messages are $\hat{W}$ and satisfy $\hat{W} = f_D^t(Y_1^n)$.

The maximum probability of message error, $P_e^{\max}$, is used as the fidelity criterion here. For any n, it is desirable to have small $P_e^{\max}$ (and thereby small P_e^{avg} and P_e^{sym}) as well as to have large M. To determine the fundamental trade-off among $(n, M, P_e^{\max})$, define the maximal number of messages achievable for a given block length and target error probability as:

$$M^*(n, \eta) = \max\{M : \exists \text{ a system } (t, f_E^t, f_D^t) \text{ of length } n \text{ with } P_e^{\max} < \eta\}.$$

Although exact expressions for the maximal code size M^* are only known in very special cases (e.g. when perfect codes are optimal), it can be bounded and approximated using information-theoretic quantities like channel capacity and channel dispersion.

Let the mutual information between the input and output of DMC $p_{Y|X}$ with input distribution p_X (and output distribution $p_Y(y) = \sum_{x \in \mathcal{X}} p_{Y|X}(y|x) p_X(x)$) be:

$$I(p_X, p_{Y|X}) = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p_{Y|X}(y|x) p_X(x) \log \frac{p_{Y|X}(y|x)}{p_Y(y)}.$$

The channel capacity C is defined to be

$$C = \max_{p_X} I(p_X, p_{Y|X})$$

and the set of capacity-achieving input distributions is

$$\Pi = \{p_X : I(p_X, p_{Y|X}) = C\}.$$

The channel dispersion is defined to be

$$V = \min_{p_X \in \Pi} \operatorname{var}\left[\log \frac{p_{Y|X}(y|x)}{p_Y(y)}\right].$$

The following theorem bounds the asymptotic behavior of $\log M^*$ in terms of n and $P_e^{\max}$ by using an expression arising from the Berry-Esseen form of the central

limit theorem.

Theorem 2.1 ([77]). *For a DMC with capacity C and dispersion V, and $0 < \eta \leq 1/2$,*

$$\log M^*(n, \eta) = nC - \sqrt{nV}Q^{-1}(\eta) + O(\log n),$$

where $Q^{-1}(\cdot)$ is the inverse of the Q-function (2.2) *and standard asymptotic notation is used [78].*

It is often of interest to consider the information rate communicated through a communication system rather than (n, M). The definition of channel capacity also has an operational interpretation as the best achievable information rate that can be communicated through a given DMC. The operational interpretation is an asymptotic one, for $n \to \infty$.

Given a target maximum probability of error value η, $0 < \eta \leq 1$, a rate R is said to be η-achievable for the DMC $p_{Y|X}$ if for any $\delta > 0$ and sufficiently large n, there exist communication systems (t, f_E^t, f_D^t) with code parameters (n, M) such that

$$\tfrac{1}{n} \log M > R - \delta$$

and

$$P_e^{\max} < \eta.$$

The supremum of η-achievable rates is called the operational η-capacity of the channel.

Furthermore, a rate is said to be achievable if it is η-achievable for all $0 < \eta \leq 1$. The supremum of achievable rates is called the operational capacity of the channel.

Theorem 2.2. *For a DMC with capacity C and $0 < \eta \leq 1$, the operational η-capacity is equal to:*

$$\frac{C}{\log |\mathcal{X}| - h_2(\eta) - \eta \log(|\mathcal{X}| - 1)},$$

where $h_2(\cdot)$ is the binary entropy function.

Theorem 2.3. *For a DMC with capacity C, the operational capacity is also C.*

One can note that this result follows directly by considering Theorem 2.1 in the $n \to \infty$ regime. There are several ways of proving the achievability and converse parts of the information-theoretic limits stated in Theorems 2.1–2.3.

One way of finding upper bounds to M^* is through a sphere-packing argument. The basic idea is that achieving η-reliability will require decoding spheres of a minimal radius. Dividing the volume of the entire signaling alphabet $|\mathcal{X}|^n$ by the volume of the decoding sphere yields a bound on M^*.

Showing achievability may follow a random coding argument. In particular, the random codebook $T = t$ is chosen from an ensemble of codebooks and the choice is revealed to both the encoder and the decoder. The expected error performance of the communication system (T, f_E^T, f_D^T) with respect to $p_T(t)$ is measured. For capacity, this error probability is shown to be arbitrarily small with increasing block length n.

The achievability proof then essentially uses the mean value theorem to argue for the existence of at least one particular t that is achievable, further using code expurgation. Achievability arguments for η-capacity additionally require rate-distortion theory and the separation principle.

The information-theoretic limits given in this section determine the performance parameters of optimal codes. No restrictions arising from physically implemented decoding were made. The remainder of the thesis considers unreliable decoders and resource-constrained decoders. Under such decoding, information-theoretic converses remain valid but achievability results do not. Optimal performance may be far from the converse bounds given here.

Chapter 3

Infrastructure Costs—Wires

Communication decoders are physical constructs and the materials from which they are built can be costly. Computational elements such as logic gates, memory cells for local information storage, and the wires to connect together the various parts of a decoder are all costly. Restrictions on computational elements for general computation problems are considered in circuit complexity theory [79, Chapter 10] (see also Section 5.6 of this thesis) whereas restrictions on memory for general computation problems are considered in space complexity theory [79, Chapter 8] (see also Chapter 4 of this thesis). The study of wiring restrictions for computation problems is similarly of general interest.

More concretely, the wires used to connect computational elements in decoding circuits are costly [34,80–84]. For example, Thies [85] noted that "wiring has long been identified as an eventual limiter of integrated circuit performance, but the National Technology Roadmap for Semiconductors (NTRS) now estimates that without radical material, design, or architectural innovations, this point will be reached [soon]."

Whereas most previous systems-theoretic works that consider information processing circuits have ignored their spatial aspects, cf. [86, 87], this chapter considers trade-offs between functionality and spatially-driven wiring costs. To explore this decoder design question with specificity, focus is placed on the trade-off between decoding speed and wiring costs incurred in constructing consensus circuits for optimally decoding binary repetition codes. The goal is to choose circuit structures that yield fast decoding but that require short wires. A functionality-cost function that captures the optimal trade-off is defined in the sequel.

As suggested in Example 2.2, the optimal decoding rule for repetition codes transmitted over symmetric channels may involve comparing the average value of the several channel output values to a fixed constant. This average value can be computed in a distributed circuit using consensus algorithms [39, 88, 89].[1] Algebraic properties of circuits determine convergence speeds of consensus algorithms run on them; Appendix 3.A reviews certain relevant aspects of algebraic graph theory including definitions of eigenratio and algebraic connectivity of graphs. The eigenratio of a circuit graph determines the convergence speed of the distributed decoding algorithm studied here [39, 90]. The algebraic connectivity of a circuit graph has also been used

[1] Alternative implementations of optimal decoding for tree-structured factor graphs are possible using either the sum-product algorithm or parallel fusion. Reasons for using consensus are provided in Section 3.1.

to describe convergence speeds of distributed consensus algorithms [88, 91–93]. This is the functionality of the circuit.

As suggested in Figure 2-5, different factor graphs may be used to represent the equality constraints in repetition codes. In fact, any connected graph of equality constraint vertices generates a repetition code. The connectivity pattern of the factor graph chosen to represent a repetition code implies a natural consensus decoding circuit, where messages are passed along edges in the circuit graph. The choice of representation therefore determines the eigenratio and algebraic connectivity of the circuit graph.

A primary aspect of the problem formulation in this chapter is embedding an abstract graph structure in physical Euclidean space [94]. Although there are several theories of drawing graphs in space,[2] an adjacency model where the geometry respects the binary relations of adjacency/nonadjacency between vertices [100, 101] is used here. In particular, this chapter considers graph drawing in Euclidean space to minimize wiring cost that is quadratic in edge length [38, 102], a method which also has certain aesthetic appeal [103,104]. It is shown that any given graph has an optimal cost-minimizing placement in Euclidean space, so the choice of code representation also determines the wiring cost of the decoding circuit.

The remainder of the chapter is organized as follows. Section 3.1 describes consensus decoding of repetition codes, including deriving decoding graphs from factor graphs, and error performance of decoding. Section 3.2 determines functionality in terms of the convergence speed of consensus decoding on graphs. Section 3.3 explicates the placement of decoding graphs in physical space to make decoding circuits as well as the wiring costs of decoding circuits under optimal placement. Section 3.4 defines functionality-cost trade-offs for decoding circuits. Section 3.5.1 gives an NP-completeness result for decoding circuit structure design; Section 3.5.2 gives properties of the circuit design problem and lists small optimal circuits; and Section 3.5.3 provides a natural reverse convex relaxation that may be useful for designing large decoding circuits. The performance of random decoding circuits is presented in Section 3.6. Section 3.7 provides further discussion of results.

■ 3.1 Consensus Decoding of Repetition Codes

The communication problem investigated in this chapter is that of optimally decoding binary repetition codes transmitted over binary-input, memoryless channels with output alphabet $\mathcal{Y} \subseteq \mathbb{R}$.

Recall from Section 2.3 that the decision rule $f_D(\cdot)$ that optimizes average error

[2] Topological graph theory is concerned with drawing a graph on a surface so that no two edges cross [95]; surfaces of interest are the closed orientable surfaces: sphere, torus, double torus, triple torus,.... Alternatively, the metric theory of embeddings [96,97] and the theory of graph separators [98] are concerned with drawing a graph in another graph while minimizing the distortion in the graph distances. For finite graphs there is a simple semidefinite program which computes their Euclidean distortion [99].

probability for a repetition code with $\mathcal{W} = \mathcal{X} = \{\pm 1\}$ is:

$$\frac{p_{Y_1^n|W}(y_1^n|w=1)}{p_{Y_1^n|W}(y_1^n|w=-1)} \underset{f_D(y_1^n)=-1}{\overset{f_D(y_1^n)=1}{\gtrless}} \frac{P_0}{1-P_0},$$

where $P_0 = \Pr[W = -1]$. When transmitting a binary repetition code over a memoryless channel,

$$p_{Y_1^n|W}(y_1^n|w) = \prod_{i=1}^{n} p_{Y_i|W}(y_i|w).$$

By considering log-likelihoods ratios:

$$\ell_i = \log \frac{p_{Y_i|W}(y|1)}{p_{Y_i|W}(y|-1)},$$

and applying some algebraic manipulations, the optimal decoding rule $f_D(\cdot)$ reduces to comparing the average value of local log-likelihood ratios $\bar{\ell} = \frac{1}{n}\sum_{i=1}^{n} \ell_i$ to the fixed threshold

$$\frac{1}{n} \log \frac{P_0}{1-P_0}.$$

A memoryless channel is binary-input, symmetric if it satisfies

$$p_{Y_i|X_i}(y_i = y|x_i = 1) = p_{Y_i|X_i}(y_i = -y|x_i = -1).$$

When transmitting a repetition code over such a channel, $f_D(\cdot)$ compares the average value of the channel outputs $\frac{1}{n}\sum_{i=1}^{n} y_i$ themselves to a fixed threshold.

For the special case given in Example 2.2 of Gaussian channels and $P_0 = 1/2$, but with $\mathcal{X} = \{\pm 1\}$ instead of $\mathcal{X} = \{0, 1\}$, the optimal decision rule is to test whether the average value of the $(y_i)_{i=1}^n$ is positive or negative. This is also true for arbitrary channel noise variance σ^2 (rather than fixing $\sigma^2 = 1$).

The optimal decoding rule as derived thus far is a mathematical formula, but no physical implementation is specified. One approach is to compute the ℓ_i near where the decoder receives its inputs y_i, and then collect, average, and threshold them at some other point in the circuit, using a so-called parallel fusion circuit topology. Such an implementation is not robust to failure of the fusion center and does not provide the final solution at each point in the circuit.

If the factor graph used to describe the repetition code is a tree, e.g. the second graph shown in Figure 2-5 but not the first, and the number of vertices is known everywhere, the sum-product algorithm for the factor graph [56] yields an exact distributed way to optimally decode, with the final solution available everywhere. A loopy circuit has redundancy to protect against wire failure, but the sum-product algorithm operating on a loopy factor graph does not yield optimal solutions. For both optimality and robustness to circuit failures, the algorithm could be implemented on a subcircuit that is a tree, however the circuit would need active monitoring and reconfiguration to change the operative subcircuit in response to failures. Robustness

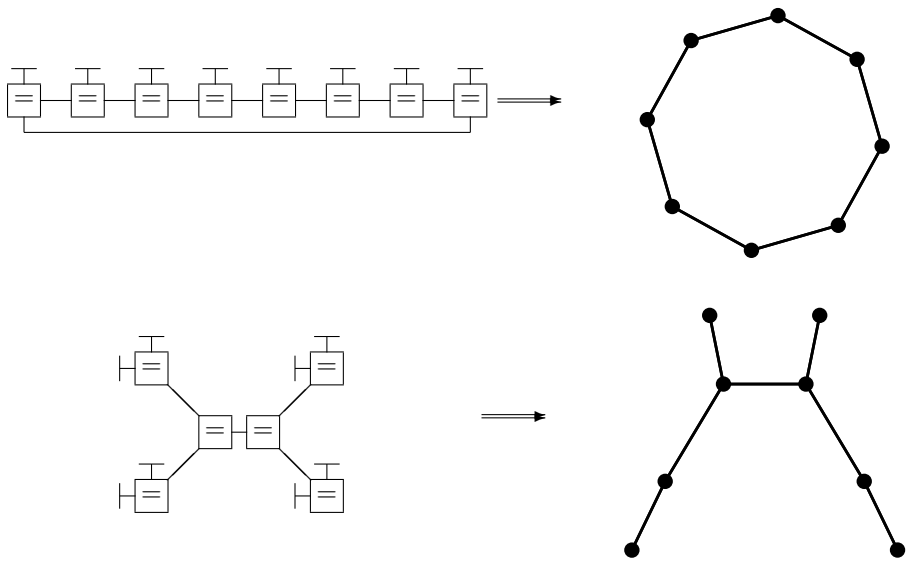

Figure 3-1. Circuit graphs for consensus decoding of repetition codes with tail-biting trellis factor graph and tree-structured factor graph.

of sum-product decoding to circuit failures is further discussed in [105].

For general circuit topologies with any degree sequence and the possibility of loops, distributed inference algorithms built on distributed averaging algorithms may be used. In distributed inference, all nodes in a circuit receive and agree upon the result of the final computation [106–108], contrary to settings with a fusion center that performs final processing [109–112]. Distributed inference is very robust to local circuit element failure.

An iterative consensus decoding algorithm is naturally built on any factor graph for the repetition code, e.g. ones shown in Figure 2-5. There is a direct mapping between the factor graph and a circuit graph that implements the decoding algorithm, as shown in Figure 3-1. Conversion is accomplished by changing half-edges in the factor graph into vertices, $\bullet$, in the circuit graph, connecting them according to the constraints, and then merging the constraints until no extraneous connections remain. Vertices in the circuit graph provide input/output functionality as well as perform local computations, whereas edges are used to pass messages.

Each vertex i has a state variable $\sigma_i(t) \in \mathbb{R}$ which is updated during each iteration t of the algorithm. The initial state is

$$\vec{\sigma}(t=0) = [\ell_1, \ldots, \ell_n]^T .$$

Vertices iteratively pass messages to their neighbors in the circuit graph based on weighted averaging of the current state and keep what they pass as their state variable. In particular,

$$\vec{\sigma}(t) = W\vec{\sigma}(t-1) = W^t\vec{\sigma}(0) \text{ for } t \geq 1,$$

where W is the weight matrix of the decoding algorithm and has the same sparsity

pattern as the adjacency matrix A of the decoding circuit graph. The weight matrix is chosen so that all edges have equal weight β:

$$W = I - \beta L,$$

where I is the identity matrix and L is the graph Laplacian of the decoding circuit graph.

At each time step t, a thresholding operation is performed at each vertex i:

$$\sigma_i(t) \overset{f_{D_i}(y_1^n)=1}{\underset{f_{D_i}(y_1^n)=-1}{\gtrless}} \frac{1}{n} \log \frac{P_0}{1-P_0}$$

where $f_{D_i}(\cdot)$ is the decoding function for symbol i.

Let the Laplacian spectral radius be $\lambda_n(L)$. If the decoding circuit graph is connected and β is bounded as $0 < \beta < 2/\lambda_n(L)$, it can be shown that $\lim_{t\to\infty} \vec{\sigma}(t)$ exists and satisfies

$$\lim_{t\to\infty} \sigma_i(t) = \frac{1}{n}\sum_{i=1}^{n} \sigma_i(0) = \frac{1}{n}\sum_{i=1}^{n} \ell_i = \bar{\ell},$$

for all $i = 1, \ldots, n$ [90]. Consequently the limiting performance of the decoding algorithm at every vertex, $P_e^{\text{avg}}(t)$ as $t \to \infty$, is the optimal performance of the parallel fusion decoding algorithm [39]. Since the state variables σ_i at every vertex surely converge to the same value, the final decoded message $\hat{W} \in \{\pm 1\}$ can be taken from any vertex.

■ 3.2 Convergence Speeds of Decoding Circuits

It is of central interest to study how quickly the decoding algorithm drives the error probability to its optimal value: the faster the better. To do so, the convergence speed of $\vec{\sigma}$ to $\bar{\ell}$ is studied. Convergence speed of decoding for weight matrix W is measured using the convergence factor

$$r(W) = \sup_{\vec{\sigma}(t)\neq\bar{\ell}} \frac{\|\vec{\sigma}(t+1) - \bar{\ell}\|_2}{\|\vec{\sigma}(t) - \bar{\ell}\|_2}$$

and associated convergence time

$$\tau(W) = \frac{1}{\log(1/r(W))}.$$

The goal is to minimize $r(W(\beta, L))$ through the choice of the weight matrix parameter β and the structure of the circuit graph, L.

The optimal weight

$$\beta^* = \frac{2}{\lambda_2(L) + \lambda_n(L)}$$

is determined by the algebraic connectivity $\lambda_2(L)$ and the Laplacian spectral radius $\lambda_n(L)$ for any decoding circuit graph structure L [90, (23)]. The convergence factor with optimal weight β^* is

$$r(L) = \frac{1 - \rho(L)}{1 + \rho(L)},$$

where $\rho(L) = \lambda_2(L)/\lambda_n(L)$ is the eigenratio [39, (6)]. Minimizing $r(L)$ is equivalent to maximizing $\rho(L)$. Decoding algorithms on circuit graphs with large eigenratio converge quickly.

The algebraic connectivity of a circuit graph $\lambda_2(L)$ has been found to determine the speed of convergence of several consensus algorithms related to the one presented. See e.g. [88, 91–93] and references therein. Therefore, these related consensus algorithms converge quickly on circuit graphs with large algebraic connectivity.

Kar et al. assert that $\rho(L) = \lambda_2(L)/\lambda_n(L)$ is more sensitive to variations in $\lambda_2(L)$ than variations in $\lambda_n(L)$. Hence they develop topology constructions that optimize $\lambda_2(L)$ [39]. Rad et al. demonstrate that graphs optimized for $\rho(L)$ also have large $\lambda_2(L)$ [93]. Some comparisons as to whether decoding circuits with large $\lambda_2(L)$ also have large $\rho(L)$, once wiring cost constraints are imposed, are presented in later sections.

A few previous works have considered optimizing $\lambda_2(L)$ or $\rho(L)$ through the choice of graph topology [39, 113–115]. Ramanujan graphs [116], a particular class of expander graphs, have been found to have extremal properties [39]; see also [117].

■ 3.3 Wiring Costs of Decoding Circuits

A physical decoding circuit is determined not only by the adjacency matrix (or equivalently the Laplacian matrix) of the circuit graph, but also by the physical placement of the vertices in Euclidean space. Through placement, the logical notions of vertices and edges in circuit graphs are transformed to the physical notions of nodes and wires in circuits. This section argues that wiring cost that is quadratic in length may be appropriate and then reviews results from graph drawing that demonstrate how to layout a circuit to minimize wiring cost [102].

Recall from Chapter 1 that quadratic wire cost is motivated by electrical and material properties that vary quadratically with wire length W. An example of such a property is the delay introduced by metal interconnects. The latency of an RC-limited interconnect can be expressed as:

$$\frac{\rho\varepsilon}{HT}W^2,$$

where $\rho\varepsilon$ is the resistivity-permittivity factor of the metal, H is the metal height, T is the insulator thickness, and W is the interconnect length [37, (3)]. As noted there [37], "the W^2 factor represents system-level opportunities to improve latency through the use of new microarchitectures that serve to 'keep interconnects short.'"

As another example, the maximum information rate that can be carried by a

simple electrical interconnection is

$$B_0 A \frac{1}{W^2},$$

where A is the cross-sectional area of the conductor, W is the interconnect length, and B_0 is essentially a constant [85, (6)]. Neural wiring has similar quadratic cost [38].

If nodes are positioned on the real line at $s = (s_1, s_2, \ldots, s_n)$, $s_i \in \mathbb{R}$, then a wire connecting nodes i and j will have cost $(s_i - s_j)^2$. If nodes are in the plane at $(s, u) = ((s_1, u_1), \ldots, (s_n, u_n)), (s_i, u_i) \in \mathbb{R}^2$, then a wire connecting nodes i and j will have cost $(s_i - s_j)^2 + (u_i - u_j)^2$. Similarly in three-dimensional Euclidean space. The total wiring costs $W_1, W_2, \ldots, W_d, \ldots$ in d-dimensional Euclidean space are the sums of the costs of individual wires.

The problem of drawing a graph to minimize quadratic wiring cost reduces to a problem in algebraic graph theory [38,102]. Consider the wiring cost for a circuit with adjacency matrix A and nodes on the real line at s. To establish the sum-of-squares property of the graph Laplacian, the wiring cost W_1 may be expressed as follows.

$$\begin{aligned}
W_1 &= \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} (s_i - s_j)^2 A_{ij} \\
&= \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} (s_i^2 - 2 s_i s_j + s_j^2) A_{ij} \\
&= \frac{1}{2} \left(\sum_{i=1}^{n} s_i^2 \sum_{j=1}^{n} A_{ij} - 2 \sum_{i=1}^{n} \sum_{j=1}^{n} s_i s_j A_{ij} + \sum_{j=1}^{n} s_j^2 \sum_{i=1}^{n} A_{ij} \right) \\
&= \sum_{i=1}^{n} s_i^2 \sum_{j=1}^{n} A_{ij} - \sum_{j=1}^{n} \sum_{i:i \neq j} s_i s_j A_{ij} \\
&= s^T (D - A) s = s^T L s.
\end{aligned}$$

Two non-triviality constraints should be imposed on the wiring cost-minimizing placement. First, all nodes are required to not be right on top of each other. Second, a normalization that $s^T s = 1$ is imposed, so that nodes are not arbitrarily close to being right on top of each other. This provides a measurement scale to the problem. Under these non-triviality constraints, the Courant-Fischer theorem implies that the optimal placement s should be the unit eigenvector associated with $\lambda_2(L)$. The wiring cost incurred for optimal placement in one-dimensional Euclidean space is $W_1 = \lambda_2(L)$.

If the network is to be drawn in two-dimensional Euclidean space with horizontal placement s and vertical placement u, it follows from simple geometry that the wiring cost is $W_2 = s^T L s + u^T L u$. If a further non-triviality constraint that the horizontal placement must be orthogonal to the vertical placement is imposed,[3] then the optimal

[3] If the circuit is to simultaneously sense and decode the output of a channel that is a band-limited field, a reason for imposing this non-triviality constraint may be related to sampling. Stable sampling requires the entire space to have reasonable sampling density [118].

placement has s be the eigenvector associated with $\lambda_2(L)$ and u be the eigenvector associated with $\lambda_3(L)$. The cost incurred is $W_2 = \lambda_2(L) + \lambda_3(L)$. In three dimensions, the cost is $W_3 = s^T L s + u^T L u + v^T L v$ and under optimal non-trivial placement, the cost is $W_3 = \lambda_2(L) + \lambda_3(L) + \lambda_4(L)$.

One can note the following separation principle:

Theorem 3.1. *There is an optimal non-trivial placement of circuit nodes in Euclidean space for any given choice of circuit graph topology A.*

Proof. The result follows directly, since for each A there is an optimal eigenplacement that incurs cost

$$W_d(A) = \sum_{i=2}^{d+1} \lambda_i(L(A))$$

for embedding in Euclidean space $\mathbb{R}^d$ for any $d = 1, \ldots, n-1$. □

In practice a circuit graph will be embedded in $\mathbb{R}^2$ or $\mathbb{R}^3$.

■ 3.4 Functionality-Cost Trade-off

Having defined circuit functionality in terms of either $\rho(L)$ or $\lambda_2(L)$ and having defined circuit cost in terms of optimized wiring costs in Euclidean space W_d, the trade-off between functionality and cost is studied. Large $\lambda_2(L)$ is desirable to enhance functionality whereas small $\lambda_2(L)$ is desirable to reduce costs. If the optimization problem were written in Lagrangian form, one could say that the cost term promotes circuits where spatially close nodes are connected, whereas the functionality term promotes long-range wires; in some sense, the cost-functionality trade-off involves optimizing the number of long-range connections.

Following Theorem 3.1, the circuit design problems posed here optimize circuit graph topology and involve algebraic graph theory. Fix the number of vertices of the graph Γ at n (the block length of the repetition code) and denote the Laplacian as $L(\Gamma)$. Also fix a dimension d, $1 \leq d < n$, as the dimension of the Euclidean space in which the circuit is to be built, e.g. $d = 2$, or $d = 3$. Then the set of circuit graphs Γ that meet a wiring cost constraint W_d under optimal placement is

$$\mathcal{G}(W_d) = \left\{ \Gamma : \sum_{i=2}^{d+1} \lambda_i(L(\Gamma)) \leq W_d \right\}. \tag{3.1}$$

The design problems to be solved are as follows.

Problem 3.1 (EIGENRATIO)**.** *Find the following functionality-cost function:*

$$b_\rho(W_d) = \max_{G \in \mathcal{G}(W_d)} \rho(L(G)) = \max_{G \in \mathcal{G}(W_d)} \frac{\lambda_2(L(G))}{\lambda_n(L(G))}. \tag{3.2}$$

Also find optimizing graphs

$$G^*_\rho(W_d) = \arg\max_{G \in \mathcal{G}(W_d)} \rho(L(G)) = \arg\max_{G \in \mathcal{G}(W_d)} \frac{\lambda_2(L(G))}{\lambda_n(L(G))}. \tag{3.3}$$

Problem 3.2 (Algebraic Connectivity)**.** *Find the following functionality-cost function:*

$$b_\lambda(W_d) = \max_{G \in \mathcal{G}(W_d)} \lambda_2(L(G)). \tag{3.4}$$

Also find optimizing graphs

$$G^*_\lambda(W_d) = \arg\max_{G \in \mathcal{G}(W_d)} \lambda_2(L(G)). \tag{3.5}$$

When $d = 1$, the objective function and the constraint coincide. Primary interest here, however, is in $d = 2$ or $d = 3$.

One might also consider the opposite problems of finding $W_d(b_\rho)$ and $W_d(b_\lambda)$. There may be slight differences, just like there are slight differences between the cost-distortion function and the distortion-cost function in information theory [55].

Problems 3.1 and 3.2 formalize optimization of decoding functionality under infrastructure cost constraints. There has been some previous work looking at trade-offs between functionality and cost in physical networks. For spatial distribution networks, the relationship between the lengths of paths from each node to the root node and the sum of the lengths of all paths in the network is discussed in [87]. Elsewhere, it has been suggested that neuronal networks are not exclusively optimized for minimal global wiring, but also for factors including minimization of computational processing steps [119].[4] A trade-off between algorithm performance and communication cost has also been discussed as a network design problem [91]. Ghosh and Boyd briefly discuss optimizing $\lambda_2(L)$ when costly links may be added to a network [113].

■ 3.5 Optimal Decoding Circuits with Costly Wires

The goal of this section is to design optimal consensus decoding circuits under wiring cost constraints by solving Problems 3.1 and 3.2.

■ 3.5.1 Optimization is NP-Hard

The decoding circuit design problems as defined are optimizations over graph Laplacians, which are discrete objects. A Laplacian matrix is symmetric, positive semidefinite, each row sums to zero, and its off-diagonal elements are either zero or minus one. Conversely, if L is any $n \times n$ matrix that satisfies these conditions, then it is the Laplacian of some graph on n nodes [121]. The set of Laplacian matrices is

$$\mathcal{L} = \{L \in \mathbb{R}^{n \times n} : L = L^T, L \succeq 0, L\vec{1} = 0, L_{ij} \in \{0, -1\} \text{ for } i \neq j\}.$$

[4] Note however that when discussing 'processing,' average path lengths rather than measures that take actual computations into account are used; see also [36, 120].

Since $\mathcal{L}$ is a discrete space, the optimization problems are integer programs, which are often difficult to solve. The difficulty of solving the algebraic connectivity with wiring costs problem, Problem 3.2, may be analyzed using computational complexity theory [122].

Before proceeding, define the decision version of the optimal algebraic connectivity optimization problem without any wiring cost constraints.

Problem 3.3. Maximum Algebraic Connectivity Augmentation

- *Given an undirected graph $G = (V, E)$, a non-negative integer k, and a non-negative threshold θ,*
- *Seek a subset $A \subseteq E^c$ of size $|A| \leq k$ such that the graph $H = (V, E \cup A)$ satisfies $\lambda_2(H) \geq \theta$.*

Now impose wiring costs on this decision problem to obtain the problem of interest.

Problem 3.4. Maximum Algebraic Connectivity Augmentation with Wiring Costs

- *Given an undirected graph $G = (V, E)$, a non-negative integer k, a non-negative threshold θ, and a non-negative wiring cost W_d,*
- *Seek a subset $A \subseteq E^c$ of size $|A| \leq k$ such that the graph $H = (V, E \cup A)$ satisfies $\lambda_2(H) \geq \theta$ and $\sum_{i=2}^{d+1} \lambda_i(H) \leq W_d$.*

The first thing to note is that when given a solution to Problem 3.4, it may be verified in polynomial time.

Theorem 3.2. *Problem 3.4 is in class NP.*

Proof. Eigenvalues of a matrix of size $n \times n$ can be computed in polynomial time $O(n^3)$ with Gaussian elimination, so as to verify the algebraic connectivity and wiring cost requirements. □

Although a solution to Problem 3.4 can be verified in polynomial time, finding a solution may be difficult. Some lemmas needed to prove the computational complexity of finding a solution are given after the statement and proof of this main result.

Theorem 3.3. *Problem 3.4 is NP-complete.*

Proof. Problem 3.4 is in class NP, by Theorem 3.2. Moreover, one can restrict Problem 3.4 to Problem 3.3 by only allowing instances having $W_d = nd$. This is the largest wiring cost possible for a graph with n vertices (achieved by a complete graph, Lemma 3.2). Since Problem 3.3 is a special case of Problem 3.4 and it is NP-complete (Lemma 3.1), the result follows from the special-case reduction [122]. □

Lemma 3.1 ([123])**.** *Problem 3.3 is NP-complete.*

Proof. Constructing a reduction from 3-colorability by creating a graph that is three disjoint copies of the graph G and making use of algebraic connectivity properties of the complete tripartite graph yields the desired result. □

Lemma 3.2. *The complete graph on n vertices, $n > 2$, has the following Laplacian eigenspectrum: $\lambda_1 = 0$, $\lambda_2 = \cdots = \lambda_n = n$.*

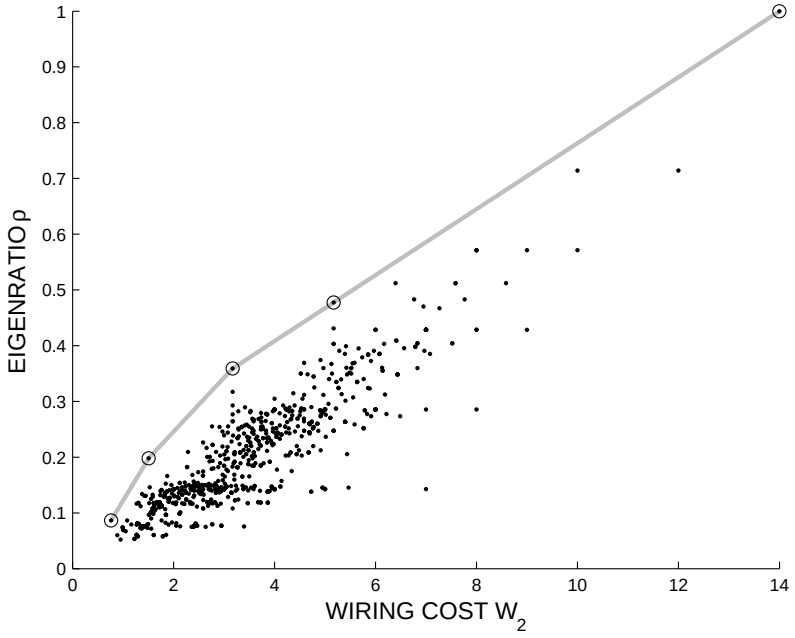

Figure 3-2. Eigenratio as a function of quadratic wiring cost under optimal non-trivial placement in $\mathbb{R}^2$ for all connected graphs on $n = 7$ vertices. The upper convex hull of these achievable functionality-cost points is denoted by the gray line. Points on the upper convex hull are circled.

■ 3.5.2 Small Optimal Decoding Circuits

Although designing decoding circuits optimally is computationally difficult, optimal circuits of small size can be determined. This design procedure is carried out here and exact solutions are found.

One may list all connected unlabeled graphs on n nodes [124, A001349] and calculate their cost and functionality values [125]. Since cospectral graphs exist, the number of graphs that need to be checked is smaller than the number of connected unlabeled graphs; the number of unlabeled graphs with distinct Laplacian spectra are given in [126, Table 1]. Besides using the sense of optimality defined in Problems 3.1 and 3.2, a stronger sense of optimality—seeking points on the upper convex hull of functionality-cost pairs—is considered here.

First, optimal decoding circuits of size $n = 7$ are computed. Figure 3-2 shows the eigenratio $\rho(L)$ as a function of optimized 2-dimensional wiring cost $W_2 = \lambda_2(L) + \lambda_3(L)$ for all possible connected circuit graphs. The upper convex hull of achievable functionality-cost points, the boundary of optimality, is also shown. Figure 3-3 shows the eigenratio as a function of optimized 3-dimensional wiring cost $W_3 = \lambda_2(L) + \lambda_3(L) + \lambda_4(L)$.

Considering algebraic connectivity as the notion of functionality, Figure 3-4 shows $\lambda_2(L)$ as a function of optimized 2-dimensional wiring cost W_2; the upper convex hull is the straight line $d\lambda_2(L) = W_d$, in this case $2\lambda_2 = W_2$. Figure 3-5 shows $\lambda_2(L)$ as a function of optimized 3-dimensional wiring cost W_3. Again, the achievable straight line upper bound $d\lambda_2(L) = W_d$, here $3\lambda_2(L) = W_3$, is evident. In fact, this upper

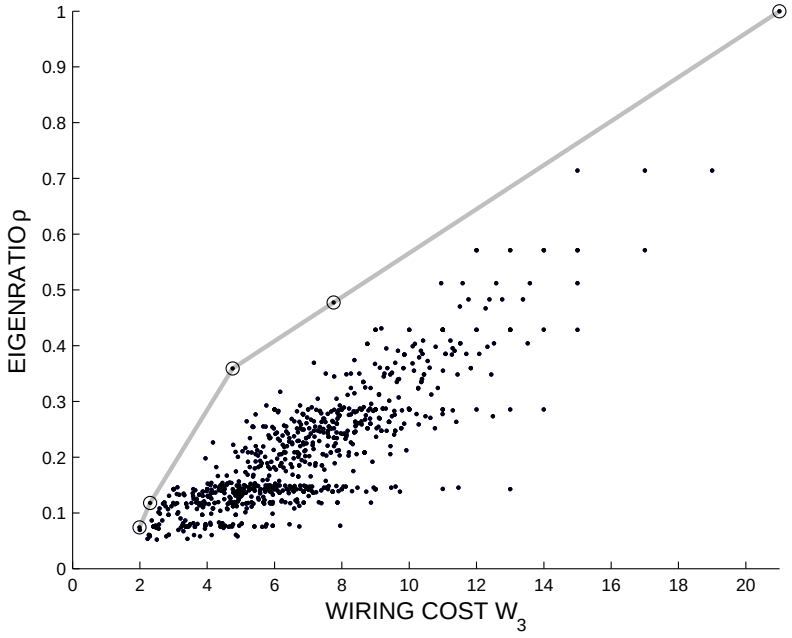

Figure 3-3. Eigenratio as a function of quadratic wiring cost under optimal non-trivial placement in $\mathbb{R}^3$ for all connected graphs on $n = 7$ vertices. The upper convex hull of these achievable functionality-cost points is denoted by the gray line. Points on the upper convex hull are circled.

bound always holds and is always achievable.

Proposition 3.1. *Solutions to the algebraic connectivity problem, Problem 3.2, satisfy*

$$b_\lambda(W_d) \leq \frac{W_d}{d},$$

and there exists at least one circuit that achieves the bound with equality for any admissible n and d.

Proof. Choose an admissible pair n and d. The bound

$$d\lambda_2(L) \leq \sum_{i=2}^{d+1} \lambda_i(L)$$

follows directly from the ordering $\lambda_2(L) \leq \lambda_3(L) \leq \cdots \leq \lambda_n(L)$, and therefore optimal graphs must obey

$$b_\lambda(W_d) \leq \frac{W_d}{d} = \frac{1}{d} \sum_{i=2}^{d+1} \lambda_i(L).$$

The bound is achievable with equality for complete graphs, which have $d\lambda_2(L) = W_d = nd$, as follows from Lemma 3.2. □

The five optimal decoding circuits in $\mathbb{R}^2$, in the sense of upper convex hull for the eigenratio, are shown in Figure 3-6; they are drawn in a minimum wiring cost

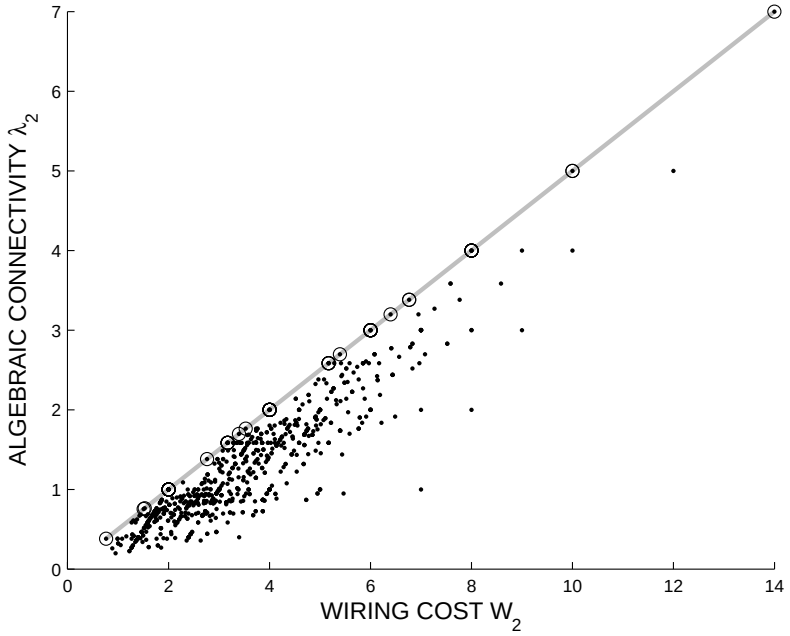

Figure 3-4. Algebraic connectivity as a function of quadratic wiring cost under optimal non-trivial placement in $\mathbb{R}^2$ for all connected graphs on $n = 7$ vertices. The upper convex hull of these achievable functionality-cost points is denoted by the gray line. Points on the upper convex hull are circled.

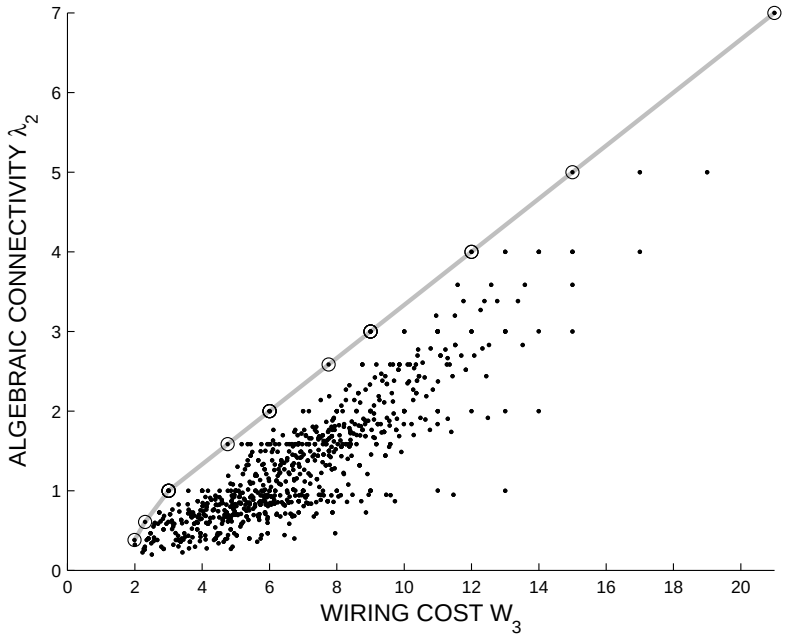

Figure 3-5. Algebraic connectivity as a function of quadratic wiring cost under optimal non-trivial placement in $\mathbb{R}^3$ for all connected graphs on $n = 7$ vertices. The upper convex hull of these achievable functionality-cost points is denoted by the gray line. Points on the upper convex hull are circled.

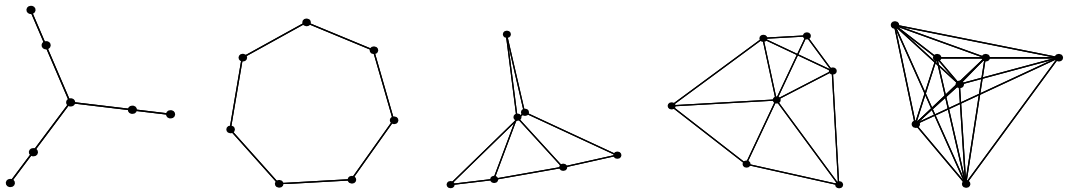

Figure 3-6. All circuits that are on the upper convex hull of achievable eigenratio-cost points in $\mathbb{R}^2$. Circuits drawn in a minimal wiring cost configuration (note that this may not be unique, even up to translation/rotation if some eigenvalues have multiplicity greater than 1).

n	$k(n)$	$K_\lambda(n)$	$K_\rho(n)$	$K_{\lambda\cap\rho}(n)$
2	1	1	1	1
3	2	2	2	2
4	6	3	3	3
5	21	8	3	3
6	112	22	4	4
7	853	62	5	5
8	11117	231	6	6

Table 3.1. Number of optimal decoding circuits in $\mathbb{R}^2$

configuration in order of increasing wiring cost. Three of these five decoding circuit graphs are also optimal in $\mathbb{R}^3$.

Although the optimal circuits have some aesthetic appeal, they do not have any extremal properties in terms of symmetry, as measured by circuit graph automorphism group order.

The optimal decoding circuits for the eigenratio problem in $\mathbb{R}^2$ are a strict subset of the optimal circuits for the algebraic connectivity problem in $\mathbb{R}^2$. There are 62 non-isomorphic circuit graphs that lie on the upper convex hull; some of these circuits are shown in Figures 3-6 and 3-7. The fact that the optimal circuits for Problem 3.1 are a strict subset of the optimal circuits for Problem 3.2 for $n = 7$ is an example of assertions that optimizing $\lambda_2(L)$ leads to optimal $\rho(L)$ [39, 93].

The solution containment property, that decoding circuits that achieve performance on the upper convex hull of $b_\rho(W_d)$ are a subset of the decoding circuits that achieve performance on the upper convex hull of $b_\lambda(W_d)$, holds for small n circuits but does not hold in general. Tables 3.1 and 3.2 list the number of connected graphs $k(n)$ [124, A001349], how many are optimal for algebraic connectivity $K_\lambda(n)$, how many are optimal for eigenratio $K_\rho(n)$, and how many optimal graphs for eigenratio are also optimal for algebraic connectivity $K_{\lambda\cap\rho}(n)$. Among the 14 rows in the tables, the containment property holds for 13 of them. It does not hold for decoding circuits in $\mathbb{R}^3$ of size $n = 8$, since $K_{\lambda\cap\rho}(8) < K_\rho(n)$.

Figures 3-8 and 3-9 show results for $n = 8$ in $\mathbb{R}^2$; there are many similarities to the $n = 7$ case. Observe that there are many more functionality-cost points for the larger dimension, as also evident in Table 3.1. The counterexample to the

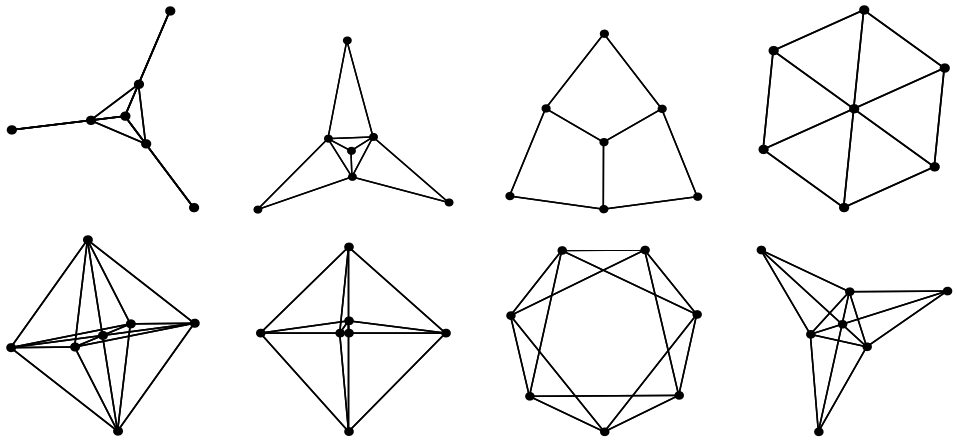

Figure 3-7. Some other circuits that are on the upper convex hull of achievable algebraic connectivity-cost points in $\mathbb{R}^2$. Circuits drawn in a minimal wiring cost configuration (note that this may not be unique, even up to translation/rotation if some eigenvalues have multiplicity greater than 1).

n	$k(n)$	$K_\lambda(n)$	$K_\rho(n)$	$K_{\lambda\cap\rho}(n)$
2	1	1	1	1
3	2	2	2	2
4	6	3	3	3
5	21	2	2	2
6	112	10	9	9
7	853	21	5	5
8	11117	61	6	4

Table 3.2. Number of optimal decoding circuits in $\mathbb{R}^3$

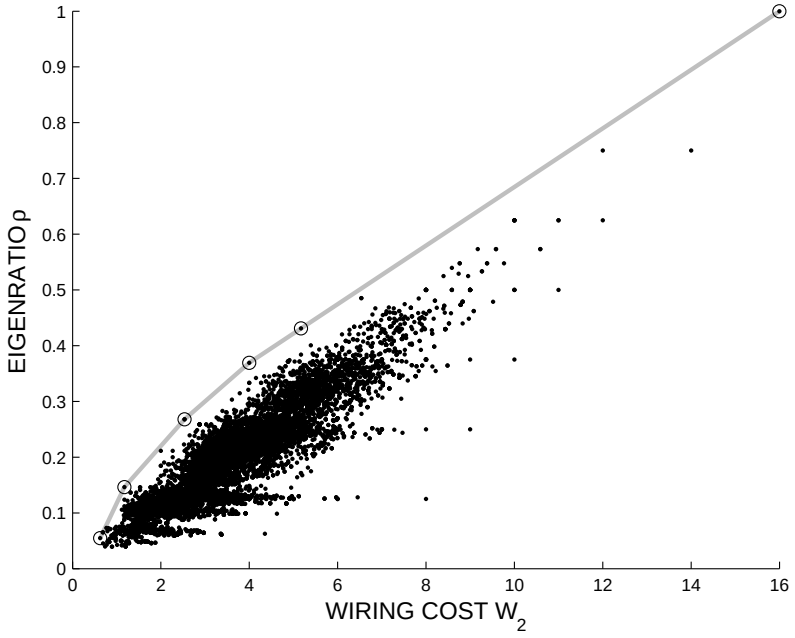

Figure 3-8. Eigenratio as a function of quadratic wiring cost under optimal non-trivial placement in $\mathbb{R}^2$ for all connected graphs on $n = 8$ vertices. The upper convex hull of these achievable functionality-cost points is denoted by the gray line. Points on the upper convex hull are circled.

containment property for $n = 8$ in $\mathbb{R}^3$ is displayed in Figures 3-10 and 3-11. These figures demonstrate that the general principle that circuits optimal for the eigenratio-cost trade-off are good for the algebraic connectivity-cost trade-off remains reasonably valid if not precisely so.

■ 3.5.3 Natural Relaxation is Reverse Convex Minimization

Although the previous section presented some general results, the focus was on exact solutions for small circuit size. Exhaustive numerical solutions may be determined for larger numbers of nodes, but the task of listing graphs on larger n becomes computationally difficult. Given that even simple decoding networks-on-chip [127], neuronal networks of small organisms [36, 128], and small wired sensor networks have $n \gg 10$, there is value in studying the natural relaxation to Problem 3.2, which may be useful for designing large decoding circuits.

Let $\mathcal{CL}$ be the convex hull of the set of all Laplacian matrices, $\mathcal{L}$. $\mathcal{CL}$ is the set of symmetric, positive semidefinite matrices, with zero row sums, and off-diagonal elements between minus one and zero.

$$\mathcal{CL} = \{L \in \mathbb{R}^{n \times n} : L = L^T, L \succeq 0, L\vec{1} = 0, -1 \leq L_{ij} \leq 0 \text{ for } i \neq j\}.$$

The Courant-Fischer idea essentially yields concavity properties of $\lambda_2(L)$.

Proposition 3.2. *The algebraic connectivity λ_2 is a concave function of L on $\mathcal{CL}$.*

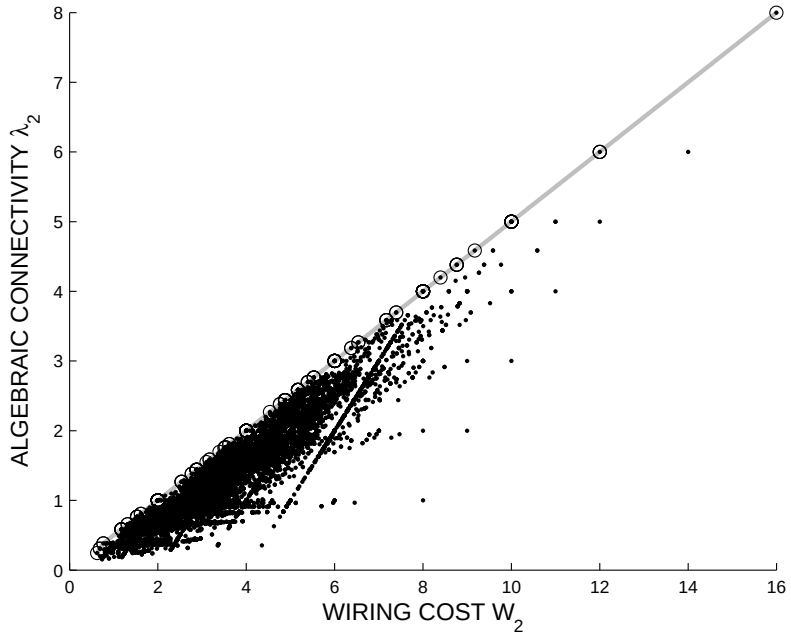

Figure 3-9. Algebraic connectivity as a function of quadratic wiring cost under optimal non-trivial placement in $\mathbb{R}^2$ for all connected graphs on $n = 8$ vertices. The upper convex hull of these achievable functionality-cost points is denoted by the gray line. Points on the upper convex hull are circled.

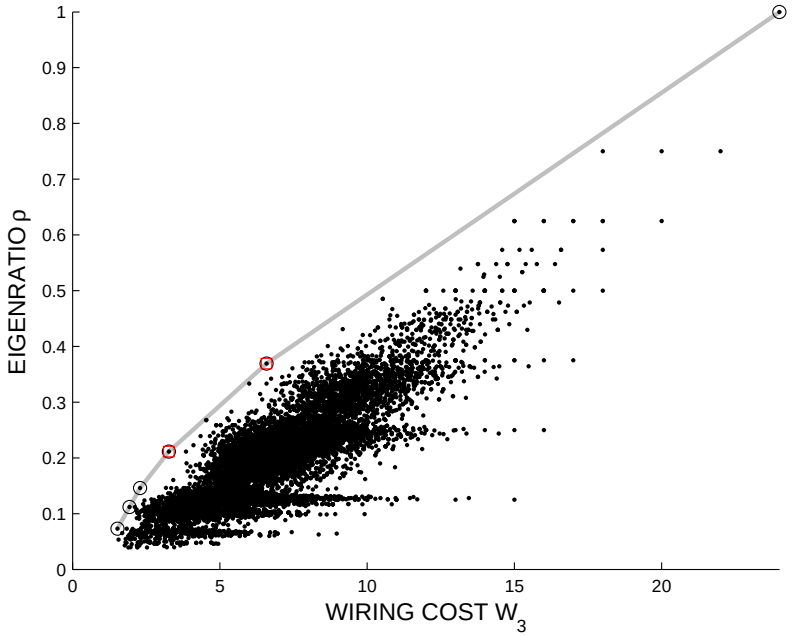

Figure 3-10. Eigenratio as a function of quadratic wiring cost under optimal non-trivial placement in $\mathbb{R}^3$ for all connected graphs on $n = 8$ vertices. The upper convex hull of these achievable functionality-cost points is denoted by the gray line. Points on the upper convex hull are circled; two of these points are further marked by red squares.

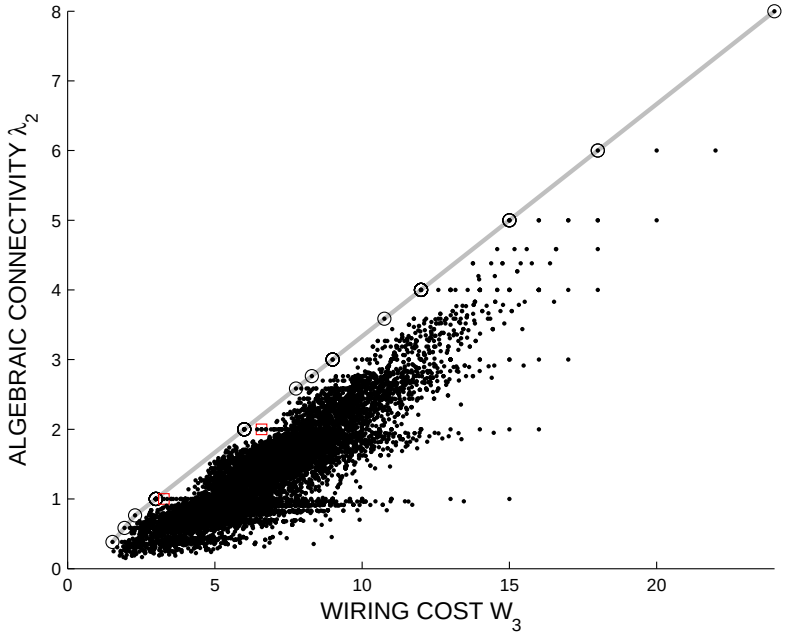

Figure 3-11. Algebraic connectivity as a function of quadratic wiring cost under optimal non-trivial placement in $\mathbb{R}^3$ for all connected graphs on $n = 8$ vertices. The upper convex hull of these achievable functionality-cost points is denoted by the gray line. Points on the upper convex hull are circled. Red squares give the performance of the circuits marked in Figure 3-10.

Proof. Each $L \in \mathcal{CL}$ is positive semidefinite and has $\lambda_1(L) = 0$, with corresponding eigenvector $\vec{1}$. Thus $\lambda_2(L)$ may be expressed as

$$\lambda_2(L) = \inf\{s^T L s : \|s\|_2 = 1 \text{ and } \vec{1}^T s = 0\}.$$

For each $s \in \mathbb{R}^n$ that satisfies $\|s\|_2 = 1$ and $\vec{1}^T s = 0$, $s^T L s$ is a linear (and therefore also concave) function of L. The formula shows that $\lambda_2(L)$ is the infimum of a family of concave functions in L, and is therefore also a concave function of L. □

A generalization of Proposition 3.2 also holds and is proven similarly.

Proposition 3.3. *The sum of the k smallest eigenvalues of L,*

$$g(L) = \sum_{i=2}^{k} \lambda_i(L)$$

is a concave function of L.

Proof. The sum of the smallest eigenvalues may be expressed in variational form as

$$\begin{aligned} g(L) &= \sum_{i=2}^{k} \lambda_i(L) = \sum_{i=1}^{k} \lambda_i(L) \\ &= \inf\{\text{trace}\left[S^T L S\right] : S \in \mathbb{R}^{n \times k} \text{ and } S^T S = I\}. \end{aligned}$$

Note that $S^T LS$ is a linear (and therefore also concave) function of L. The formula shows that $g(L)$ is the infimum of a family of concave functions in L, and is therefore also a concave function of L. □

Recall the algebraic connectivity problem, Problem 3.2, but now consider matrices in $\mathcal{CL}$ rather than in $\mathcal{L}$. This relaxed problem is a reverse convex program, defined as follows.

Definition 3.1 ([129]). *The optimization problem*

$$\begin{aligned} &\min\ f(z) \\ &\textit{subject to}\ z \in D \backslash C \end{aligned}$$

is a reverse convex program *when $f(z)$ is a convex function, D is a closed convex set, C is an open convex set, and D and C are given by explicit convex inequalities.*

Proposition 3.4. *Considering $L \in \mathcal{CL}$, the optimization problem:*

$$\begin{aligned} &\max\ \lambda_2(L) \\ &\textit{subject to}\ \sum_{i=2}^{d+1} \lambda_i(L) \leq W_d \end{aligned}$$

is a reverse convex program.

Proof. The function $\lambda_2(L)$ is concave, by Proposition 3.2. Maximizing a concave function is equivalent to minimizing a convex function, so the first condition is satisfied. The set $\mathcal{CL}$ is a closed convex set given by an explicit convex inequality. The wiring cost constraint excludes the set

$$\left\{ L \in \mathcal{CL} : \sum_{i=2}^{d+1} \lambda_i(L) > W_d \right\},$$

which by Proposition 3.3, is an open convex set given by an explicit convex inequality. Thus the second condition is satisfied. □

One may note that without the wiring cost constraint, the natural relaxation would have been a convex program [113]. There are several standard techniques for solving reverse convex programs [129–131], which may be used for the relaxed version of the cost-limited algebraic connectivity optimization problem. A heuristic such as rounding may be used to convert a solution for the relaxed problem into a solution for the algebraic connectivity problem, Problem 3.2.

Good decoding circuits for the wiring cost-limited eigenratio problem can then be found from circuits that have large algebraic connectivity.

It remains to implement and test this relaxation-based circuit design method.

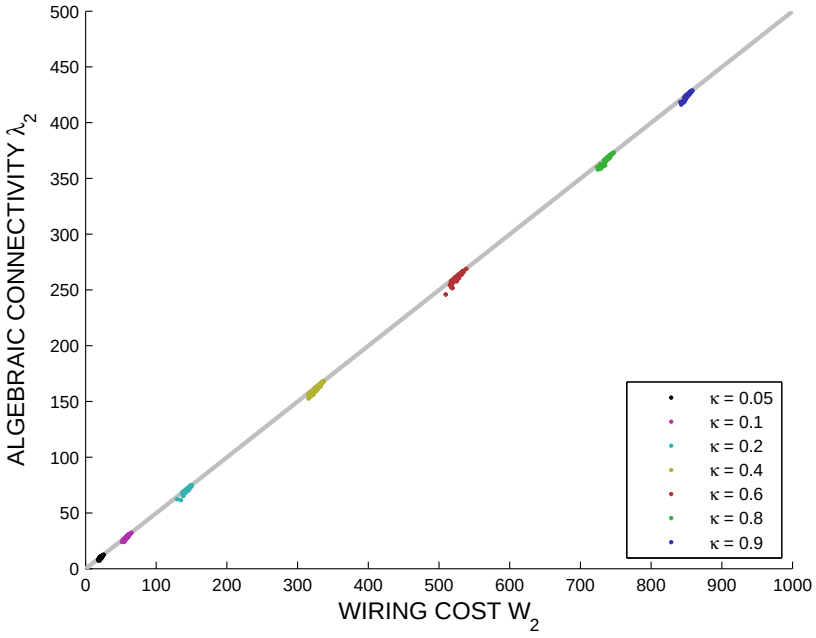

Figure 3-12. Algebraic connectivity as a function of quadratic wiring cost under optimal non-trivial placement in $\mathbb{R}^2$ for 100 random graphs each on $n = 500$ vertices for several different values of κ. An upper bound is denoted by the gray line.

3.6 Random Decoding Circuits

As an alternative to explicitly designing decoding circuits, one might wonder whether decoding circuits chosen at random perform well. Random circuits may be easier to manufacture at nanoscale [132]. Olfati-Saber studied the algebraic connectivity of graphs created by randomly rewiring regular lattices [133], along the lines of the Watts-Strogatz small-world graph ensemble [134]. The Erdös-Rényi random graph ensemble is obtained when all edges are randomly rewired, so that any edge is an i.i.d. Bernoulli random variable. Through simulations, it was found that increasing the level of random rewiring can greatly increase the algebraic connectivity while not increasing the Laplacian spectral radius too much. Thus the random rewiring procedure can also increase the eigenratio by a large amount.

Some simulations of Erdös-Rényi random graphs were carried out to test whether the random ensemble yields decoding circuits that provide good trade-offs between convergence speed and wiring cost. Figures 3-12 and 3-13 display the algebraic connectivities and wiring costs of 100 random circuits each of size $n = 500$ for several different edge existence probabilities κ. The upper bound from Proposition 3.1 is shown for comparison. As evident, random circuits perform well. Figures 3-14 and 3-15 display the eigenratios and wiring costs of these random circuits.

Beyond the simulation results, one might wonder whether one can use random matrix theory [135, 136] to say something more. The following concentration results that hold asymptotically almost surely can be proven.

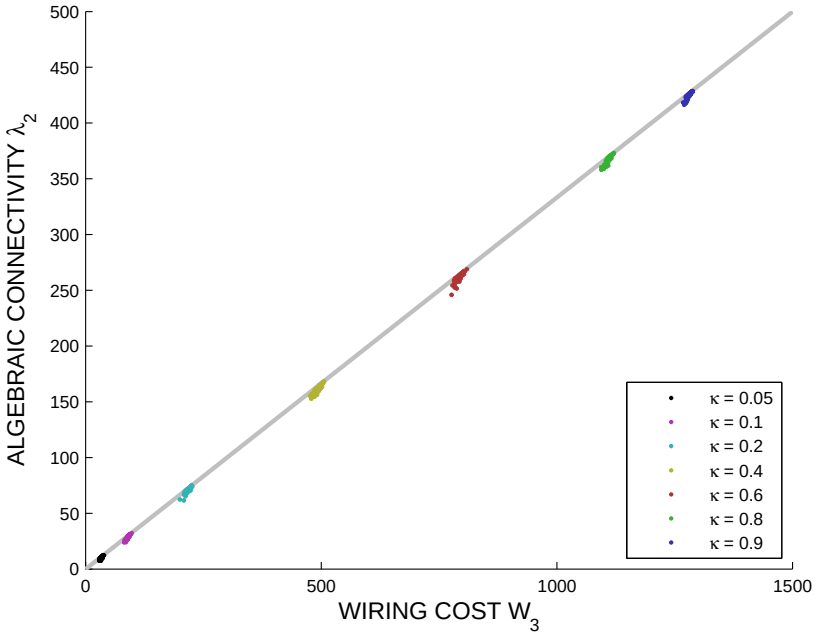

Figure 3-13. Algebraic connectivity as a function of quadratic wiring cost under optimal non-trivial placement in $\mathbb{R}^3$ for 100 random graphs each on $n = 500$ vertices for several different values of κ. An upper bound is denoted by the gray line.

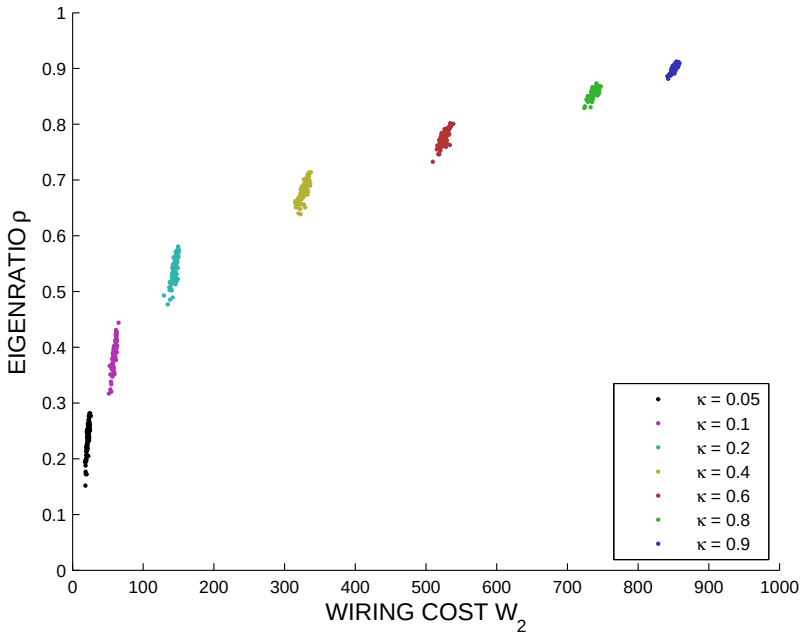

Figure 3-14. Eigenratio as a function of quadratic wiring cost under optimal non-trivial placement in $\mathbb{R}^2$ for 100 random graphs each on $n = 500$ vertices for several different values of κ.

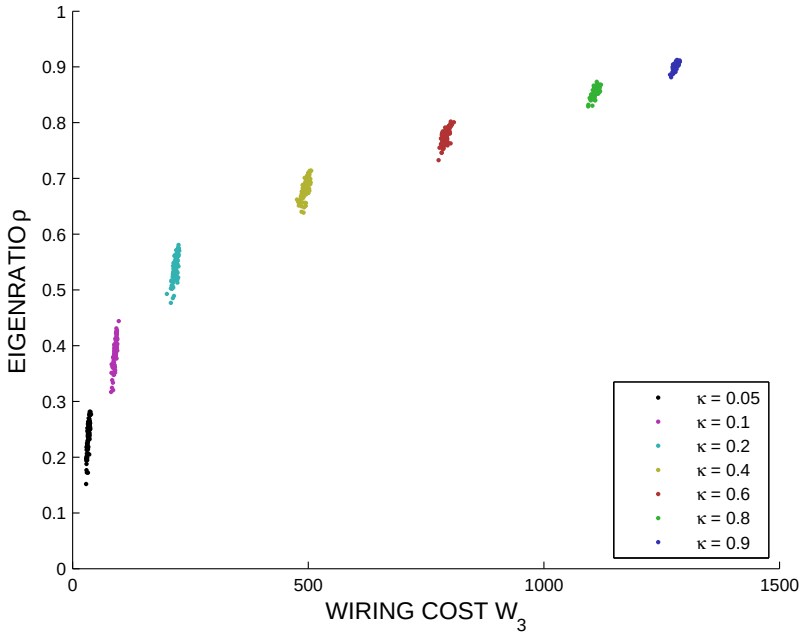

Figure 3-15. Eigenratio as a function of quadratic wiring cost under optimal non-trivial placement in $\mathbb{R}^3$ for 100 random graphs each on $n = 500$ vertices for several different values of κ.

Theorem 3.4 ([137, 138])**.** *The algebraic connectivity $\lambda_2(L)$ and Laplacian spectral radius $\lambda_n(L)$ of an Erdös-Rényi random graph with n vertices and probability κ of an edge being present asymptotically almost surely satisfy the following. For any $\epsilon > 0$,*

$$\kappa n - f_\epsilon^+(n) < \lambda_2(L) < \kappa n - f_\varepsilon^-(n)$$

and

$$\kappa n + f_\varepsilon^-(n) < \lambda_n(L) < \kappa n + f_\epsilon^+(n),$$

where

$$f_\epsilon^+(n) = \sqrt{(2+\epsilon)\kappa(1-\kappa)n\log n}$$

and

$$f_\epsilon^-(n) = \sqrt{(2-\epsilon)\kappa(1-\kappa)n\log n}.$$

This further implies that:

Corollary 3.1. *The Laplacian eigenvalues $\lambda_i(L)$ of an Erdös-Rényi random graph with n vertices and probability κ of an edge being present asymptotically almost surely satisfy the following. For any $\epsilon > 0$,*

$$\kappa n - f_\epsilon^+(n) < \lambda_2(L) \leq \lambda_3(L) \leq \cdots \leq \lambda_n(L) < \kappa n + f_\epsilon^+(n).$$

Proof. Follows directly from Theorem 3.4 by noting that $f_\epsilon^-(n) < f_\epsilon^+(n)$ and the ordering $\lambda_2(L) \leq \lambda_3(L) \leq \cdots \leq \lambda_n(L)$. □

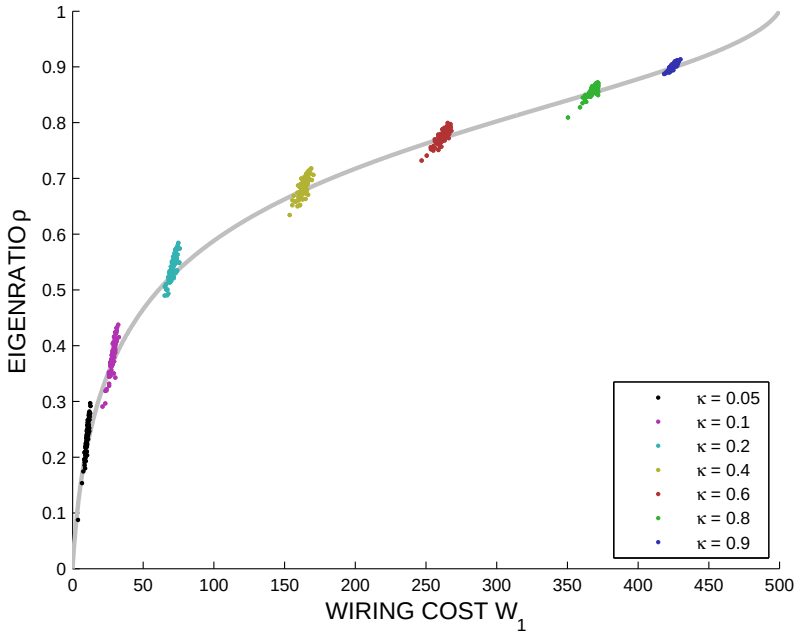

Figure 3-16. Eigenratio as a function of quadratic wiring cost under optimal non-trivial placement in $\mathbb{R}$ for 100 random graphs each on $n = 500$ vertices for several different values of κ. The asymptotic approximation from random matrix theory is shown as a gray line.

These concentration results imply that for large n, a decoding circuit constructed according to a graph selected from the Erdös-Rényi ensemble will almost surely behave like any other. Moreover, the decoding speed parameters will be approximately given by relatively simple functions of n and κ:

$$\lambda_2(L) \approx \kappa n - \sqrt{2\kappa(1-\kappa)n \log n}$$

and

$$\rho(L) \approx \frac{\kappa n - \sqrt{2\kappa(1-\kappa)n \log n}}{\kappa n + \sqrt{2\kappa(1-\kappa)n \log n}}.$$

As shown in Figure 3-16 for $n = 500$ and decoding circuits in $\mathbb{R}$ (where $W_1 = \lambda_2(L)$), these asymptotic approximations do predict the properties of randomly designed decoding circuits well.

■ 3.7 Discussion

This chapter used the framework of consensus decoding of binary repetition codes to convey the concept that material costs of decoding circuits might limit functionality. Moving beyond repetition codes, one might consider the effects of wiring cost limitations on general factor graph synthesis [48, Chapter 11], [65].

Problem 3.1 and Problem 3.2 formalized trade-offs between functionality, as measured by convergence speed of decoding, and wiring cost, as measured under optimal

placement. Using an enumerative strategy for the NP-complete problem, some optimal circuits for small n were found. An approach based on reverse convex program relaxations was suggested as a principled method for designing larger circuits that are close to optimal. Random design of decoding circuits was also investigated, using the Erdös-Rényi graph ensemble. It would be of further interest to test other random graph ensembles, such as exponential random graphs [139–142].

The basic mathematical formulation of the problem considered in this chapter may also inform engineering general distributed systems [88], such as sensor networks. These systems are often tasked to perform distributed detection, which is equivalent to decoding repetition codes. Like wiring, free-space radio propagation has quadratic attenuation cost in the length of the links. Moreover, nodes may be positioned in space to optimize performance.

The results given in this chapter may not only have conceptual implications for synthesis of engineered systems, but also for analysis of natural information-processing systems [36, 38, 119, 120]. One may further note that the eigenratio determines the speed of synchronization of networks of coupled oscillators [143, 144] and that the algebraic connectivity is central to rubber elasticity theory in polymer science [145].

■ 3.A Review of Algebraic Graph Theory

The eigenvalues and eigenvectors of certain matrices associated with graphs are often of central importance in understanding the properties of dynamical systems defined on those graphs.

Let $G = (V, E)$ be a graph with vertex set V of cardinality n and edge set $E \subseteq V \times V$. Let A be the adjacency matrix of the graph. Let d_j indicate the degree of vertex j and let D be the degree matrix of the graph, which takes value d_j along the diagonal and value 0 otherwise. The Laplacian matrix of a graph, L, satisfies $L = D - A$. Explicitly, its entries are

$$L_{ij} = \begin{cases} -1, & (i,j) \in E \\ d_i, & i = j \\ 0, & \text{otherwise.} \end{cases}$$

The Laplacian matrix is sometimes called the Kirchhoff matrix. The eigenvalues of L are denoted $\lambda_1(L) \leq \lambda_2(L) \leq \cdots \leq \lambda_n(L)$. Since L is symmetric, all of its eigenvalues are real and eigenvectors corresponding to different eigenvalues are orthogonal.

Let K be the $n \times |E|$ incidence matrix, where the columns are indexed by the edges and the rows are indexed by the vertices. After fixing an orientation for each edge, and for each column, the matrix is constructed as follows. For each edge, place $+1$ in the row corresponding to the positive end and -1 in the row corresponding to the negative end; all other entries are zero. Evidently, $L = KK^T$. If λ is an eigenvalue of L with associated eigenvector s,

$$\lambda \|s\|^2 = \langle \lambda s, s \rangle = \langle KK^T s, s \rangle = \langle K^T s, K^T s \rangle = \|K^T s\|^2 \geq 0.$$

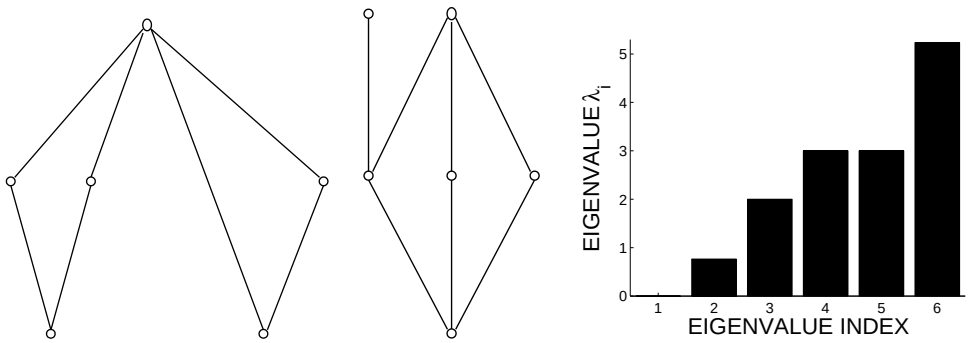

Figure 3-17. A pair of non-isomorphic graphs with identical Laplacian spectra.

Thus the Laplacian matrix of every graph is positive semidefinite and its eigenvalues are non-negative. Furthermore, since the row sums of L are all zero, the all-ones vector $\vec{1}$ is an eigenvector with eigenvalue $\lambda_1 = 0$. The multiplicity of the eigenvalue 0 determines the number of connected components in the graph; if $\lambda_2(L) > 0$, then the graph is a single connected component.

The celebrated Courant-Fischer min-max theorems endow the eigenvalues with certain optimality properties. In particular,

$$\lambda_2(L) = \min_{\{s \in \mathbb{R}^n:\ s \neq \vec{0} \text{ and } s \perp \vec{1}\}} \frac{\langle s, Ls \rangle}{\langle s, s \rangle},$$

where $\vec{0}$ is the all-zero vector. This second smallest eigenvalue $\lambda_2(L)$ is called the *algebraic connectivity* of the graph. The largest eigenvalue satisfies

$$\lambda_n(L) = \max_{\{s \in \mathbb{R}^n:\ s \neq \vec{0} \text{ and } s \perp \vec{1}\}} \frac{\langle s, Ls \rangle}{\langle s, s \rangle}.$$

The largest eigenvalue $\lambda_n(L)$ is called the *spectral radius* of the graph Laplacian. The *eigenratio*, $\rho(L) = \lambda_2(L)/\lambda_n(L)$, may also be defined.

Several non-isomorphic graphs may share the same Laplacian eigenspectrum, however characterizing such cospectral graphs is difficult [126,146]. An example of a pair of cospectral graphs is shown in Figure 3-17.

There is no precise characterization of how the Laplacian spectrum of a graph changes when edges are added or deleted. There are bounds [147–149], including ones on the relationship between $\lambda_2(L)$ and $\lambda_n(L)$ [150].

Chapter 4

Infrastructure Costs—Memory

The previous chapter was concerned with the costs of wiring in decoding circuits. This chapter focuses on another resource that might be constrained in constructing decoding circuits: memory. Though memory is outwardly an informational resource, it is ultimately limited by physical resource constraints [40, 151–153]. This chapter considers settings where memory constraints may limit the number of decoding rules that can be stored by a decoder for situations where any member of a family of sources might be communicating. It is assumed that both the encoder and the decoder have access to perfect source state information, but that the decoder has limited adaptability due to the memory constraint. As a consequence, only a finite set of decoding rules may be used and the rule will almost surely be mismatched [42] if the family of sources is parameterized by an absolutely continuous random variable.

To explore trade-offs between information storage at the decoder and information transmission performance of the overall communication system, focus is placed on uncoded transmission with Bayes risk optimal decoding. In particular, a class of sources $\{p_W^{(\theta)}(w)\}_{\Theta\in\mathcal{T}}$ is parameterized by the random variable Θ, which itself is distributed according to $p_\Theta(\theta)$. That is to say, a source $p_W^{(\theta)}(w)$ is selected by drawing its identity θ at random according to $p_\Theta(\theta)$. The memory constraint forces quantization of Θ into K quantization cells that partition $\mathcal{T}$.

For human decision-makers, one would expect K to be small since classical experiments in cognitive science demonstrate that people have limited memory [157]. For electronic decoders, one would expect K to be large but finite since there is typically a fair amount of storage capacity [152]; high-rate quantization theory would therefore be particularly relevant.

When $K = 1$, the decoder knows nothing about which source is feeding the channel except the distributional characterization $p_\Theta(\theta)$, similar to the weighted formulation of universal information theory [154]. When $K > 1$, knowledge of the quantization cell may yield better decoding. Quantization of source identity Θ is similar to quantization of channel state information for systems with rate-constrained feedback [155, 156]. For $K \to \infty$, the precise source is known to the decoder and the problem reduces to straightforward uncoded transmission, as described in Section 2.3. The goal here is to choose a quantization scheme for Θ with $1 < K < \infty$ cells that optimizes the Bayes risk in decoding with respect to the class of sources and the communication channel. The formulation is also extended to the setting where there are numerous classes of sources that must be quantized.

The remainder of the chapter is organized as follows. Section 4.1 frames the problem of quantizing prior probabilities for Bayesian hypothesis testing (uncoded transmission). Optimal low-rate quantization is discussed in Section 4.2; designs using dynamic programming and using the Lloyd-Max algorithm are derived using mean Bayes risk error as a distortion measure for quantization. A high-resolution approximation to the distortion-rate function is obtained in Section 4.3. Examples are given in Section 4.4. An extension to several classes of sources is provided in Section 4.5 and implications for human decision-making and the information economics of social discrimination are presented. Section 4.6 gives some further discussion of results.

■ 4.1 Population Decoding

Consider the uncoded transmission scenario with M possible messages, $\{0, \ldots, M-1\}$. Message m has probability p_m of being transmitted, $p_m = p_W(m)$, and the overall prior probability vector is $\theta = [p_0, \ldots, p_{M-1}]^T$, with non-negative entries and $\sum_{m=0}^{M-1} p_m = 1$, which is known. The message is transmitted over a channel $p_{Y|X}(y|x) = p_{Y|W}(y|w)$, and the decoder is to recover w from y. Optimal Bayesian hypothesis testing, as described in Section 2.3, minimizes Bayes risk [158].

Now consider a family of sources $\{p_W^{(\theta)}(w)\}_{\Theta \in \mathcal{T}}$, each as above, with its own prior probability vector drawn from the distribution $p_\Theta(\theta)$ supported on the $(M-1)$-dimensional probability simplex. The parameterized family of sources, $(\{p_W^{(\theta)}(w)\}_\Theta, p_\Theta(\theta))$, is also called a population. If the prior probability vector of each source were used perfectly by the decoder, then there would be no difference from standard Bayesian hypothesis testing.

The setting considered here, however, limits the decoder to work with at most K different prior probability vectors. A two-stage architecture with separation is imposed. When there are more than K sources in the population, the decoder first maps the true prior probability vector of the source being used to one of the K available vectors and then proceeds to perform the optimal Bayesian hypothesis test, treating that vector as the prior probabilities of the source.

An example of such a constrained decoding scenario is that of human decision making, e.g. in sports officiating. A referee decides whether a player has committed a foul using his or her noisy observation as well as prior experience. Players commit fouls with different frequencies; some players are dirtier or more aggressive than others. This normalized frequency is the prior probability for the 'foul committed' message. Hence, there is a distribution of prior probabilities over the population of players. If the referee tunes the prior probability to the particular player whose action is to be decoded, performance improves. Human decision makers, however, are limited in their information-processing capacity and only use a small number of categories [159]. The referee is thus limited and categorizes players into a small number of dirtiness levels, with associated representative prior probabilities, exactly the scenario described above.

Previous work that combines decoding and quantization looks at quantization of channel outputs, not source probabilities, and also only approximates the Bayes risk

function instead of working with it directly, e.g. [160–162] and references therein and thereto. In such work, there is a further communication constraint between the channel and the decoder, but the decision maker has unconstrained processing capability. This work deals with the opposite case, where there is no further communication constraint between the channel and the decoder, but the decoder is constrained.

The mapping from prior probability vectors in the population to one of K representative probability vectors will be optimized using quantization theory, but first mean Bayes risk error (MBRE) is defined as the fidelity criterion for quantization of Θ. A brief look at imperfect priors appears in [163, Sec. 2.E], but optimal quantization was not considered. Further results show that small deviations from the true prior yields small deviations in the Bayes risk [164, 165].

■ 4.1.1 Mismatch and Bayes Risk Error

This section defines mismatched decoding and how its performance is measured using MBRE. It is restricted to the binary case, $\mathcal{W} = \{0, 1\}$, where the channel implies likelihood functions $p_{Y|W}(y|0)$ and $p_{Y|W}(y|1)$.

For a given source, let the prior probabilities be $\theta = \Pr[W = 0]$ and $1 - \theta = \Pr[W = 1]$. A decoding function f_D is designed to uniquely map each y to either 0 or 1 so as to minimize Bayes risk J, an expectation over the non-negative Bayes cost functions c_{ij} (of which at least one is positive). Section 2.3 shows that f_D should be the likelihood ratio test. Recall that the two kinds of errors:

$$P_e^{\mathrm{I}} = \Pr[f_D(Y) = 1|W = 0],$$
$$P_e^{\mathrm{II}} = \Pr[f_D(Y) = 0|W = 1]$$

may be used to express the Bayes risk as

$$J = (c_{10} - c_{00})\theta P_e^{\mathrm{I}} + (c_{01} - c_{11})(1 - \theta)P_e^{\mathrm{II}} + c_{00}\theta + c_{11}(1 - \theta). \tag{4.1}$$

It is often of interest to assign no cost to correct decisions, i.e. $c_{00} = c_{11} = 0$, which is assumed in the sequel. The Bayes risk simplifies to

$$J(\theta) = c_{10}\theta P_e^{\mathrm{I}}(\theta) + c_{01}(1 - \theta)P_e^{\mathrm{II}}(\theta). \tag{4.2}$$

In (4.2), the dependence of the Bayes risk and error probabilities on θ has been explicitly noted. The error probabilities depend on θ through $f_D(\cdot)$. The function $J(\theta)$ is zero at the points $\theta = 0$ and $\theta = 1$ and is positive-valued, strictly concave, and continuous in the interval $(0, 1)$ [158, 166, 167].

When the true prior probability is θ, but $f_D(y)$ is designed according to the likelihood ratio test (2.1) using some other value a substituted for θ, there is mismatch, and the Bayes risk is

$$J(\theta, a) = c_{10}\theta P_e^{\mathrm{I}}(a) + c_{01}(1 - \theta)P_e^{\mathrm{II}}(a). \tag{4.3}$$

The mismatched Bayes risk $J(\theta, a)$ is a linear function of θ with slope $(c_{10}P_e^{\mathrm{I}}(a) -$

$c_{01}P_e^{\mathrm{II}}(a)$) and intercept $c_{01}P_e^{\mathrm{II}}(a)$. Note that $J(\theta, a)$ is tangent to $J(\theta)$ at a and that $J(\theta, \theta) = J(\theta)$.

Let Bayes risk error $d(\theta, a)$ be the difference between the mismatched Bayes risk function $J(\theta, a)$ and the matched Bayes risk function $J(\theta, \theta)$:

$$\begin{aligned} d(\theta, a) &= J(\theta, a) - J(\theta, \theta) \\ &= c_{10}\theta P_e^{\mathrm{I}}(a) + c_{01}(1-\theta)P_e^{\mathrm{II}}(a) - c_{10}\theta P_e^{\mathrm{I}}(\theta) - c_{01}(1-\theta)P_e^{\mathrm{II}}(\theta). \end{aligned} \tag{4.4}$$

Several useful properties of $d(\theta, a)$ as a function of θ and as a function of a may be proven.

Theorem 4.1. *The Bayes risk error $d(\theta, a)$ is non-negative and only equal to zero when $\theta = a$. As a function of $\theta \in (0, 1)$, it is continuous and strictly convex for all a.*

Proof. Since $J(\theta)$ is a continuous and strictly concave function, and lines $J(\theta, a)$ are tangent to $J(\theta)$, $J(\theta, a) \geq J(\theta)$ for all θ and a, with equality when $\theta = a$. Consequently, $d(\theta, a)$ is non-negative and only equal to zero when $\theta = a$. Moreover, $d(\theta, a)$ is continuous and strictly convex in $\theta \in (0, 1)$ for all a because it is the difference of a continuous linear function and a continuous strictly concave function. □

Lemma 4.1. *Let C be an arbitrary constant and let β and γ be arbitrary positive constants. Let $G(a) = \beta P_e^I(a) + \gamma P_e^{II}(a) + C$. Then for any deterministic likelihood ratio test $f_D(\cdot)$, as a function of $a \in (0, 1)$ for all θ, $G(a)$ has exactly one stationary point, which is a minimum.*

Proof. Consider the parameterized curve $(P_e^{\mathrm{I}}, P_e^{\mathrm{II}})$ traced out as a is varied; this is a flipped version of the receiver operating characteristic (ROC). The flipped ROC is a strictly convex function for deterministic likelihood ratio tests. At its endpoints, it takes values $(P_e^{\mathrm{I}} = 0, P_e^{\mathrm{II}} = 1)$ when $a = 1$ and $(P_e^{\mathrm{I}} = 1, P_e^{\mathrm{II}} = 0)$ when $a = 0$ [158], and therefore has average slope -1. By the mean value theorem and strict convexity, there exists a unique point on the flipped ROC at which

$$\frac{dP_e^{\mathrm{II}}}{dP_e^{\mathrm{I}}} = -1.$$

To the left of that point: $-\infty < dP_e^{\mathrm{II}}/dP_e^{\mathrm{I}} < -1$, and to the right of that point: $-1 < dP_e^{\mathrm{II}}/dP_e^{\mathrm{I}} < 0$.

For deterministic likelihood ratio tests, $\beta \frac{d}{da} P_e^{\mathrm{I}}(a) < 0$ and $\gamma \frac{d}{da} P_e^{\mathrm{II}}(a) > 0$ for all $a \in (0, 1)$ and positive constants β and γ [158]. Therefore, if

$$\frac{\gamma dP_e^{\mathrm{II}}}{\beta dP_e^{\mathrm{I}}} < -1,$$

i.e.

$$\frac{\gamma dP_e^{\mathrm{II}}}{da} \frac{da}{\beta dP_e^{\mathrm{I}}} < -1,$$

then

$$\gamma \frac{dP_e^{\mathrm{II}}}{da} > -\beta \frac{dP_e^{\mathrm{I}}}{da}$$

and

$$\beta\frac{dP_e^{\mathrm{I}}}{da}+\gamma\frac{dP_e^{\mathrm{II}}}{da}>0.$$

In the same manner, if

$$\frac{\gamma dP_e^{\mathrm{II}}}{\beta dP_e^{\mathrm{I}}}>-1,$$

then

$$\beta\frac{dP_e^{\mathrm{I}}}{da}+\gamma\frac{dP_e^{\mathrm{II}}}{da}<0.$$

Combining the above, the function $\beta P_e^{\mathrm{I}}(a)+\gamma P_e^{\mathrm{II}}(a)$ has exactly one stationary point in $(0,1)$, which occurs when the slope of the flipped ROC is $-\beta/\gamma$. Denote this stationary point as a_s. For $0<a<a_s$, $-1<dP_e^{\mathrm{II}}/dP_e^{\mathrm{I}}<0$ and the slope of $\beta P_e^{\mathrm{I}}(a)+\gamma P_e^{\mathrm{II}}(a)$ is negative; for $a_s<a<1$, $-\infty<dP_e^{\mathrm{II}}/dP_e^{\mathrm{I}}<-1$ and the slope of $\beta P_e^{\mathrm{I}}(a)+\gamma P_e^{\mathrm{II}}(a)$ is positive. Therefore, a_s is a minimum. □

Theorem 4.2. *For any deterministic likelihood ratio test $f_D(\cdot)$ and for any fixed θ, as a function of $a\in(0,1)$ the Bayes risk error $d(\theta,a)$ has exactly one stationary point, which is a minimum.*

Proof. As a function of a, the Bayes risk error $d(\theta,a)$ is directly seen to be of the form $\beta P_e^{\mathrm{I}}(a)+\gamma P_e^{\mathrm{II}}(a)+C$. Hence, by Lemma 4.1 it has exactly one stationary point, which is a minimum. □

Definition 4.1. *A continuous function $f:\mathbb{R}\mapsto\mathbb{R}$ is defined to be* quasiconvex *when one of the following hold:*

- *f is non-decreasing,*
- *f is non-increasing, or*
- *there is a point c in the domain of f such that for $t\le c$ (and t in the domain of f), $f(t)$ is non-increasing and also for $t\ge c$ (and t in the domain of f), $f(t)$ is non-decreasing.*

An alternate definition is through a Jensen-like inequality. The function f is quasiconvex if and only if:

$$f(\lambda x+(1-\lambda)y)\le\max(f(x),f(y))$$

for any x,y in the domain of f and any $0\le\lambda\le1$.

Definition 4.2. *The function f is defined to be* strictly quasiconvex *when*

$$f(\lambda x+(1-\lambda)y)<\max(f(x),f(y))$$

for any x,y in the domain of f and any $0\le\lambda\le1$.

Corollary 4.1. *For any deterministic likelihood ratio test $f_D(\cdot)$, as a function of $a\in(0,1)$ for all θ, the Bayes risk error $d(\theta,a)$ is strictly quasiconvex.*

Proof. Follows directly from Theorem 4.2. □

Theorem 4.3. *Let the function*

$$g(a) = \int_{b_1}^{b_2} d(\theta, a) p_\Theta(\theta) d\theta$$

for some arbitrary $b_1, b_2 \in (0,1)$. The function $g(a)$ has exactly one stationary point, which is a minimum. Moreover $g(a)$ is strictly quasiconvex.

Proof. The function $g(a)$ can be expressed as:

$$\begin{aligned}
g(a) &= \int_{b_1}^{b_2} d(\theta, a) p_\Theta(\theta) d\theta \\
&= \int_{b_1}^{b_2} \left[J(\theta, a) - J(\theta, \theta) \right] p_\Theta(\theta) d\theta \\
&\overset{(a)}{=} C + \int_{b_1}^{b_2} J(\theta, a) p_\Theta(\theta) d\theta \\
&= C + \int_{b_1}^{b_2} \left[c_{10} \theta P_e^{\mathrm{I}}(a) + c_{01}(1-\theta) P_e^{\mathrm{I}}(a) \right] p_\Theta(\theta) d\theta \\
&= C + c_{10} P_e^{\mathrm{I}}(a) \int_{b_1}^{b_2} \theta p_\Theta(\theta) d\theta + c_{01} P_e^{\mathrm{I}}(a) \int_{b_1}^{b_2} (1-\theta) p_\Theta(\theta) d\theta \\
&\overset{(b)}{=} C + \beta P_e^{\mathrm{I}}(a) + \gamma P_e^{\mathrm{I}}(a),
\end{aligned}$$

where the quantity C is defined to be $C = -\int J(\theta, \theta) p_\Theta(\theta) d\theta$ in step (a) and the quantities β and γ are defined to be $\beta = c_{10} \int \theta p_\Theta(\theta) d\theta$ and $\gamma = c_{01} \int (1-\theta) p_\Theta(\theta) d\theta$ in step (b).

Then the result follows from Lemma 4.1. □

As will become evident, these properties of $d(\theta, a)$ and $g(a)$ are useful to design optimal quantizers.

The third derivative of $d(\theta, a)$ with respect to θ is

$$- c_{10} \theta \tfrac{d^3 P_e^{\mathrm{I}}(\theta)}{d\theta^3} - 3 c_{10} \tfrac{d^2 P_e^{\mathrm{I}}(\theta)}{d\theta^2} - c_{01}(1-\theta) \tfrac{d^3 P_e^{\mathrm{II}}(\theta)}{d\theta^3} + 3 c_{01} \tfrac{d^2 P_e^{\mathrm{II}}(\theta)}{d\theta^2}, \tag{4.5}$$

when the constituent derivatives exist. When the third derivative exists and is continuous, $d(\theta, a)$ is locally quadratic, which is useful to develop high-rate quantization theory for Bayes risk error fidelity [168].

To design quantizers of Θ, the MBRE $E[d(\theta, a)]$ will be minimized.

■ 4.2 Optimal Quantization Design

Optimal fixed-rate quantizers for $p_\Theta(\theta)$ under Bayes risk error distortion are now derived. A K-point quantizer partitions the interval $[0,1]$ (the two-dimensional probability simplex) into K regions $\mathcal{R}_1, \mathcal{R}_2, \ldots, \mathcal{R}_K$. There is a representation point a_k

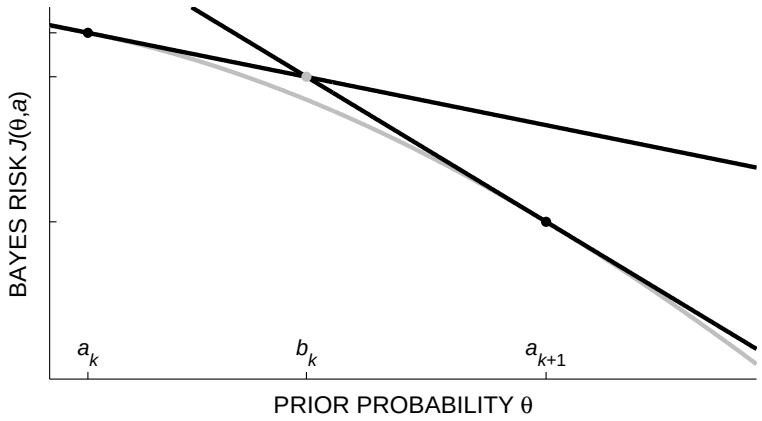

Figure 4-1. The intersection of the lines $J(\theta, a_k)$, tangent to $J(\theta)$ at a_k, and $J(\theta, a_{k+1})$, tangent to $J(\theta)$ at a_{k+1}, is the optimal interval boundary b_k.

to which elements of each of these quantization regions $\mathcal{R}_k$ are mapped. For regular quantizers, the regions are subintervals $\mathcal{R}_1 = [0, b_1]$, $\mathcal{R}_2 = (b_1, b_2]$, $\ldots$, $\mathcal{R}_K = (b_{K-1}, 1]$ and the representation points a_k are in $\mathcal{R}_k$.[1] A quantizer can be viewed as a nonlinear function $v_K(\cdot)$ such that $v_K(\theta) = a_k$ for $\theta \in \mathcal{R}_k$. For a given K, the goal is to find the quantizer that minimizes the MBRE:

$$D = E[d(\Theta, v_K(\Theta))] = \int d(\theta, v_K(\theta)) p_\Theta(\theta) d\theta. \tag{4.6}$$

■ 4.2.1 Local Optimality Conditions

There is no closed-form solution to optimal quantizer design, but an optimal quantizer must satisfy the nearest neighbor condition, the centroid condition, and the zero probability of boundary condition [169–172]. These conditions for MBRE are now developed.

Nearest Neighbor Condition

With the representation points $\{a_k\}$ fixed, an expression for the interval boundaries $\{b_k\}$ is derived. Given any $\theta \in [a_k, a_{k+1}]$, if $J(\theta, a_k) < J(\theta, a_{k+1})$ then Bayes risk error is minimized if θ is represented by a_k, and if $J(\theta, a_k) > J(\theta, a_{k+1})$ then Bayes risk error is minimized if θ is represented by a_{k+1}. The boundary point $b_k \in [a_k, a_{k+1}]$ is the abscissa of the point at which the lines $J(\theta, a_k)$ and $J(\theta, a_{k+1})$ intersect. The idea is illustrated graphically in Figure 4-1.

By manipulating the slopes and intercepts of $J(\theta, a_k)$ and $J(\theta, a_{k+1})$, the point of

[1] Due to the strict convexity of $d(\theta, a)$ in θ for all a shown in Theorem 4.1, quantizers that satisfy the necessary conditions for MBRE optimality are regular, see [169, Lemma 6.2.1]. Therefore, only regular quantizers are considered.

intersection is found to be:

$$b_k = \frac{c_{01}\left(P_e^{\mathrm{II}}(a_{k+1}) - P_e^{\mathrm{II}}(a_k)\right)}{c_{01}\left(P_e^{\mathrm{II}}(a_{k+1}) - P_e^{\mathrm{II}}(a_k)\right) - c_{10}\left(P_e^{\mathrm{I}}(a_{k+1}) - P_e^{\mathrm{I}}(a_k)\right)}. \tag{4.7}$$

Centroid Condition

With the quantization regions fixed, the MBRE is to be minimized over the $\{a_k\}$. Here, the MBRE is expressed as the sum of integrals over quantization regions:

$$D = \sum_{k=1}^{K} \int_{\mathcal{R}_k} \left(J(\theta, a_k) - J(\theta, \theta)\right) p_\Theta(\theta) d\theta. \tag{4.8}$$

Because the regions are fixed, the minimization may be performed for each interval separately.

Define $I_k^{\mathrm{I}} = \int_{\mathcal{R}_k} \theta p_\Theta(\theta) d\theta$ and $I_k^{\mathrm{II}} = \int_{\mathcal{R}_k} (1-\theta) p_\Theta(\theta) d\theta$, which are conditional means. Then:

$$a_k = \arg\min_a \left\{c_{10} I_k^{\mathrm{I}} P_e^{\mathrm{I}}(a) + c_{01} I_k^{\mathrm{II}} P_e^{\mathrm{II}}(a)\right\} = \arg\min_a g(a), \tag{4.9}$$

where $g(a)$ is as in Theorem 4.3. Since $g(a)$ has exactly one stationary point, which is a minimum, (4.9) is uniquely minimized by setting its derivative equal to zero. Thus, a_k is the solution to:

$$c_{10} I_k^{\mathrm{I}} \left.\frac{dP_e^{\mathrm{I}}(a)}{da}\right|_{a_k} + c_{01} I_k^{\mathrm{II}} \left.\frac{dP_e^{\mathrm{II}}(a)}{da}\right|_{a_k} = 0. \tag{4.10}$$

Commonly, differentiation of the two error probabilities is tractable; they are themselves integrals of the likelihood functions and the differentiation is with respect to some function of the limits of integration.

Zero Probability Boundary Condition

The third necessary condition for quantizer optimality arises when dealing with probability distribution functions that contain a discrete component. Suppose there is a quantizer that satisfies the nearest neighbor and centroid conditions, has two adjacent representation points $a_1 < a_2$ so that the corresponding boundary point b_1 separating $\mathcal{R}_1$ and $\mathcal{R}_2$ is the suitably defined midpoint, and that b_1 is an element of $\mathcal{R}_1$. Now suppose that this value of b_1 has positive probability mass. We modify the quantizer by reassigning b_1 to $\mathcal{R}_2$, which clearly does not affect the incurred distortion. The centroid of $\mathcal{R}_1$ however is now shifted so that a_1 is no longer optimal and the distortion can be reduced by replacing a_1 with the new centroid. This exercise demonstrates the necessity of requiring that the random variable Θ must have zero probability of occurring at a boundary between quantization regions. When $p_\Theta(\theta)$ is absolutely continuous, the zero probability of boundary condition is automatically satisfied.

Necessity and Sufficiency

Theorem 4.4 ([169–172]). *The nearest neighbor condition, the centroid condition, and the zero probability of boundary condition are necessary for a quantizer to be optimal.*

If additional conditions are met, then the necessary conditions for optimality are also sufficient for local optimality.[2]

Theorem 4.5. *If the following conditions hold for a source Θ and a distortion function $d(\theta, a)$:*

1. *$p_\Theta(\theta)$ is positive and continuous in $(0, 1)$;*
2. *$\int_0^1 d(\theta, a)p_\Theta(\theta)d\theta$ is finite for all a; and*
3. *$d(\theta, a)$ is zero only for $\theta = a$, is continuous in θ for all a, and is continuous and quasiconvex in a,*

then the nearest neighbor condition, centroid condition, and zero probability of boundary conditions are sufficient to guarantee local optimality of a quantizer.

Proof. Minor extension of results in [173], by relaxing the requirement of convexity of $d(\theta, a)$ in a to quasiconvexity of $d(\theta, a)$ in a. □

Note that the first and second conditions of Theorem 4.5 are met by common distributions such as the uniform distribution and the family of beta distributions [174]. The third condition is satisfied by Bayes risk error, as given in Theorem 4.1 and Corollary 4.1.

■ 4.2.2 Design using the Lloyd-Max Algorithm

Alternating between the nearest neighbor and centroid conditions, the iterative Lloyd-Max algorithm can be applied to find minimum MBRE quantizers [169, 171, 172]. The algorithm is widely used because of its simplicity, effectiveness, and convergence properties [175].

Algorithm 4.1 (Lloyd-Max).

1. *Choose an arbitrary set of initial representation points $\{a_k\}$.*
2. *For each $k = 1, \ldots, K$ set $\mathcal{R}_k = \{\theta \mid d(\theta, a_k) \leq d(\theta, a_j)$ for all $j \neq k\}$.*

[2] By local optimality, it is meant that the $\{a_k\}$ and $\{b_k\}$ minimize the objective function (4.6) among feasible representation and boundary points near them. More precisely, a K-level quantizer v_K is locally optimal if for some $\delta > 0$,

$$\int_0^1 d(\theta, v_K(\theta))p_\Theta(\theta)d\theta \leq \int_0^1 d(\theta, u_K(\theta))p_\Theta(\theta)d\theta$$

for every K-level quantizer u_K with boundaries β_k and representation points α_k satisfying $|\beta_k - b_k| < \delta$ for $k = 1, \ldots, K-1$ and $|\alpha_k - a_k| < \delta$ for $k = 1, \ldots, K$.

3. *For each* $k = 1, \ldots, K$ *set* $a_k = \arg\min_a E[d(\Theta, a) \mid \Theta \in \mathcal{R}_k]$.

4. *Repeat steps 2 and 3 until change in average distortion is negligible.*

5. *Revise* $\{a_k\}$ *and* $\{\mathcal{R}_k\}$ *to satisfy the zero probability of boundary condition.*

The average distortion decreases or remains the same after each execution of steps 2 and 3. Under the conditions of Theorem 4.5, the algorithm is guaranteed to converge to a local optimum [171]. The algorithm may be run several times with different initializations to find the global optimum.

Further conditions on $d(\theta, a)$ and $p_\Theta(\theta)$ are given in [173] for there to be a unique locally optimal quantizer, i.e. the global optimum. If these further conditions for unique local optimality hold, then the algorithm is guaranteed to find the globally minimum MBRE quantizer.

Design of globally optimal quantizers is considered again in Section 4.2.3 using a different optimization algorithm, but first it is demonstrated that increasing the number of quantization levels K does not worsen quantization performance.

Monotonic Convergence in K

Theorem 4.6. *Let*

$$D^*(K) = \sum_{k=1}^{K} \int_{\mathcal{R}_k^*} d(\theta, a_k^*) p_\Theta(\theta) d\theta \tag{4.11}$$

denote the MBRE for an optimal K*-point quantizer. This function* $D^*(K)$ *monotonically converges as* K *increases.*

Proof. The MBRE-optimal K-point quantizer is the solution to the following problem:

$$\begin{aligned} \min & \sum_{k=1}^{K} \int_{b_{k-1}}^{b_k} d(\theta, a_k) p_\Theta(\theta) d\theta \\ \text{subject to } & b_0 = 0 \\ & b_K = 1 \\ & b_{k-1} \le a_k, \ k = 1, \ldots, K \\ & a_k \le b_k, \ k = 1, \ldots, K. \end{aligned} \tag{4.12}$$

Adding the additional constraint $b_{K-1} = 1$ to (4.12) forced $a_K = 1$ and degeneracy of the Kth quantization region. The optimization problem for the K-point quantizer (4.12) with the additional constraint is equivalent to the optimization problem for the $(K-1)$-point quantizer. Thus, the $(K-1)$-point design problem and the K-point design problem have the same objective function, but the $(K-1)$-point problem has an additional constraint. Therefore, $D^*(K-1) \ge D^*(K)$.

Since $d(\theta, v_K(\theta)) \ge 0$, $D = E[d(\Theta, v_K(\Theta))] \ge 0$. The sequence $D^*(K)$ is nonincreasing and bounded from below and hence converges. □

Mean Bayes risk error cannot get worse when more quantization levels are employed. In typical settings, as in Section 4.4, performance always improves with an increase in the number of quantization levels.

■ 4.2.3 Design using Dynamic Programming

Design of absolutely optimal quantizers for many different fidelity criteria is possible through dynamic programming for discrete-valued Θ [176]. This section shows the applicability of this design procedure for the MBRE criterion. The same method can also be used to design approximately optimal quantizers for absolutely continuous Θ through discretization, but care is warranted [177].

Designing the v_K for a fixed number of levels K that minimizes D involves determining both the representation points $\{a_k\}$ and the boundaries $\{b_k\}$, but due to the centroid condition, there is an optimal $\{a_k\}$ for any given $\{b_k\}$ and so only the boundaries $\{b_k\}$ are designed here. Recall that there is no loss of optimality by only considering regular quantizers.

Let ${}_1D(\beta_1, \beta_2)$ be the minimum value of the expected distortion when one representation point is placed in the interval $(\beta_1, \beta_2) \subseteq (0, 1)$:

$$ {}_1D(\beta_1, \beta_2) = \min_a \int_{\beta_1}^{\beta_2} d(\theta, a) p_\Theta(\theta) d\theta. $$

Notice that this is a centroid defined using function $g(a)$ from Theorem 4.3. Also let ${}_\kappa D(\beta_1, \beta_2)$ be that function when $\kappa \geq 2$ points are placed in the interval (β_1, β_2):

$$ \begin{aligned} {}_\kappa D(\beta_1, \beta_2) &= \min_{\{a_k\}_{k=1}^{\kappa}, \{b_k\}_{k=1}^{\kappa-1} : \beta_1 < b_1 < \cdots < b_{\kappa-1} < \beta_2} \sum_{k=1}^{\kappa} \int_{b_{k-1}}^{b_k} d(\theta, a) p_\Theta(\theta) d\theta \\ &= \min_{\{b_k\}_{k=1}^{\kappa-1} : \beta_1 < b_1 < \cdots < b_{\kappa-1} < \beta_2} \sum_{k=1}^{\kappa} {}_1D(b_{k-1}, b_k), \end{aligned} $$

where $b_0 = \beta_1$ and $b_\kappa = \beta_2$.

Notice that $D^*(K)$ in (4.11) is the same as

$$ {}_K D(0, 1) = \min_{\{b_k\}_{k=1}^{K-1} : 0 < b_1 < \cdots < b_{K-1} < 1} \sum_{k=1}^{K} {}_1D(b_{k-1}, b_k), $$

where $b_0 = 0$ and $b_K = 1$. Let $b_1^*, b_2^*, \ldots, b_{K-1}^*$ be the optimizing boundary points for ${}_K D(b_0^* = 0, b_K^* = 1)$. By a *reductio ad absurdum* argument, it follows that $b_1^*, b_2^*, \ldots, b_{K-2}^*$ must be the optimizing boundary points for ${}_{K-1}(b_0^*, b_{K-1}^*)$. Thus for $K > 1$,

$$ {}_K D(b_0^*, b_K^*) = \min_{b_{K-1} : b_0^* < b_{K-1} < b_K^*} \left[{}_K D(b_0^*, b_{K-1}) + {}_1D(b_{K-1}, b_K^*) \right]. $$

Similarly for any k, $1 < k \leq K$,

$$ {}_kD(b_0, \beta) = \min_{b:b_0<b<\beta} \left[{}_{k-1}D(b_0, b) + {}_1D(b, \beta) \right]. \tag{4.13} $$

Moreover let the optimizing value be

$$ b^*_{k-1}(b_0, \beta) = \arg\min_{b:b_0<b<\beta} \left[{}_{k-1}D(b_0, b) + {}_1D(b, \beta) \right]. \tag{4.14} $$

The recursive partitioning structure suggested by the development above can be formalized as a dynamic programming algorithm for finding the optimal quantizer. The first two steps are preparatory, but are explicitly noted.

Algorithm 4.2 ([176])**.**

1. *Compute the values of* ${}_1D(\beta_1, \beta_2)$ *for all (discrete)* β_1 *and* β_2 *in* (b_0, b_N)*.*
2. *For each* $k = 2, \ldots, K$*, compute* ${}_kD(b_0, \beta)$ *and* $b^*_{k-1}(b_0, \beta)$ *for all (discrete)* β *in* (b_0, b_N) *using* (4.13) *and* (4.14)*.*
3. *Set* $B_K = b_k$*.*
4. *For each* $k = K, K-1, \ldots, 2$*, set* $B_{k-1} = b^*_{k-1}(b_0, B_k)$*.*
5. *Set* $B_0 = b_0$*.*
6. *For each* $k = 1, \ldots, K$ *set* $A_k = \arg\min_a E[d(\Theta, a) \mid \Theta \in (B_{k-1}, B_k)]$*.*

Theorem 4.7 ([176])**.** *The boundaries* $\{B_k\}_{k=0}^K$ *and representation points* $\{A_k\}_{k=1}^K$ *returned by Algorithm 4.2 represent the optimal quantizer.*

As noted by Sharma, step 1 of the algorithm involves minimizing the function $g(a)$ from Theorem 4.3 and may be computationally intense [176]. Since $g(a)$ is strictly quasiconvex, however, the computational burden can be greatly reduced using the Fibonacci Search method. Fibonacci Search is a line search procedure for minimizing a strictly quasiconvex function over a bounded interval [178, Chapter 8].

■ 4.3 High-Rate Quantization Theory

High-rate quantization theory [175] may also be applied to the study of minimum MBRE quantization. Note that the source alphabet $[0, 1]$ is bounded and that only fixed-rate quantization is studied. The distortion function for the MBRE criterion has a positive second derivative in θ (due to strict convexity) and for many families of likelihood functions, it has a continuous third derivative, see (4.5). Thus, it is locally quadratic in the sense of Li et al. [168] and in a manner similar to many perceptual, non-difference distortion functions, the high-rate quantization theory is well-developed.

Let

$$B(\theta) = -\tfrac{1}{2}c_{10}\theta\tfrac{d^2P_e^{\mathrm{I}}(\theta)}{d\theta^2} - c_{10}\tfrac{dP_e^{\mathrm{I}}(\theta)}{d\theta} - \tfrac{1}{2}c_{01}(1-\theta)\tfrac{d^2P_e^{\mathrm{II}}(\theta)}{d\theta^2} + c_{01}\tfrac{dP_e^{\mathrm{II}}(\theta)}{d\theta}. \tag{4.15}$$

Then at high rate, i.e. K large, $d(\theta, a_k)$ is approximated by the following second order Taylor expansion [168]:

$$d(\theta, a_k) \approx B(\theta)|_{\theta=a_k}(\theta - a_k)^2, \quad \theta \in \mathcal{R}_k. \tag{4.16}$$

Assuming that p_Θ is sufficiently smooth and substituting (4.16) into the objective of (4.12), the MBRE is approximated by:

$$D \approx \sum_{k=1}^{K} p_\Theta(a_k)B(a_k)\int_{\mathcal{R}_k}(\theta - a_k)^2\,d\theta. \tag{4.17}$$

At high rate, a quantizer is well-described by a quantizer point density function $\lambda(\theta)$. Integrating a quantizer point density over an interval yields the fraction of the $\{a_k\}$ that are in that interval.

The MBRE is greater than and approximately equal to the following lower bound, derived in [168] by relationships involving normalized moments of inertia of intervals $\mathcal{R}_k$ and by Hölder's inequality:

$$D_L = \tfrac{1}{12K^2}\int_0^1 B(\theta)p_\Theta(\theta)\lambda(\theta)^{-2}d\theta, \tag{4.18}$$

where the optimal quantizer point density is

$$\lambda(\theta) = \frac{(B(\theta)p_\Theta(\theta))^{1/3}}{\int_0^1 (B(\theta)p_\Theta(\theta))^{1/3}\,d\theta}. \tag{4.19}$$

Substituting (4.19) into (4.18) yields

$$D_L = \tfrac{1}{12K^2}\|B(\theta)p_\Theta(\theta)\|_{1/3}. \tag{4.20}$$

■ 4.4 Optimal Quantizers

This section presents examples of optimal quantizers. Consider the following scalar source and channel model:

$$Y = w + N, \quad w \in \{0, \mu\}, \tag{4.21}$$

where N is a zero-mean, Gaussian random variable with variance σ^2. The likelihoods are:

$$\begin{aligned} p_{Y|W}(y|0) &= \mathcal{N}(y;0,\sigma^2) = \tfrac{1}{\sigma\sqrt{2\pi}} e^{-y^2/2\sigma^2}, \\ p_{Y|W}(y|\mu) &= \mathcal{N}(y;\mu,\sigma^2) = \tfrac{1}{\sigma\sqrt{2\pi}} e^{-(y-\mu)^2/2\sigma^2}. \end{aligned} \tag{4.22}$$

The two error probabilities are:

$$\begin{aligned} P_e^{\mathrm{I}}(\theta) &= Q\left(\tfrac{\mu}{2\sigma} + \tfrac{\sigma}{\mu}\ln\left(\tfrac{c_{10}\theta}{c_{01}(1-\theta)}\right)\right), \\ P_e^{\mathrm{II}}(\theta) &= Q\left(\tfrac{\mu}{2\sigma} - \tfrac{\sigma}{\mu}\ln\left(\tfrac{c_{10}\theta}{c_{01}(1-\theta)}\right)\right). \end{aligned} \tag{4.23}$$

To use the centroid condition, the derivatives of the error probabilities are determined as follows.

$$\left.\tfrac{dP_e^{\mathrm{I}}(\theta)}{d\theta}\right|_{a_k} = -\tfrac{1}{\sqrt{2\pi}}\tfrac{\sigma}{\mu}\tfrac{1}{a_k(1-a_k)} e^{-\frac{1}{2}\left(\frac{\mu}{2\sigma}+\frac{\sigma}{\mu}\ln\left(\frac{c_{10}a_k}{c_{01}(1-a_k)}\right)\right)^2}, \tag{4.24}$$

$$\left.\tfrac{dP_e^{\mathrm{II}}(\theta)}{d\theta}\right|_{a_k} = +\tfrac{1}{\sqrt{2\pi}}\tfrac{\sigma}{\mu}\tfrac{1}{a_k(1-a_k)} e^{-\frac{1}{2}\left(\frac{\mu}{2\sigma}-\frac{\sigma}{\mu}\ln\left(\frac{c_{10}a_k}{c_{01}(1-a_k)}\right)\right)^2}. \tag{4.25}$$

By substituting these derivatives into (4.10) and simplifying, the following expression is obtained for the representation points:

$$a_k = \frac{I_k^{\mathrm{I}}}{I_k^{\mathrm{I}} + I_k^{\mathrm{II}}}. \tag{4.26}$$

For high-rate analysis, the second derivatives of the error probabilities are needed. They are:

$$\tfrac{d^2P_e^{\mathrm{I}}(\theta)}{d\theta^2} = -\tfrac{1}{\sqrt{8\pi}}\tfrac{\sigma}{\mu}\tfrac{1}{\theta^2(1-\theta)^2} e^{-\frac{1}{8\mu^2\sigma^2}\left(\mu^2+2\sigma^2\ln\left(\frac{c_{10}\theta}{c_{01}(1-\theta)}\right)\right)^2}\left[-3+4\theta-\tfrac{2\sigma^2}{\mu^2}\ln\left(\tfrac{c_{10}\theta}{c_{01}(1-\theta)}\right)\right] \tag{4.27}$$

and

$$\tfrac{d^2P_e^{\mathrm{II}}(\theta)}{d\theta^2} = +\tfrac{1}{\sqrt{8\pi}}\tfrac{\sigma}{\mu}\tfrac{1}{\theta^2(1-\theta)^2} e^{-\frac{1}{8\mu^2\sigma^2}\left(\mu^2-2\sigma^2\ln\left(\frac{c_{10}\theta}{c_{01}(1-\theta)}\right)\right)^2}\left[-1+4\theta-\tfrac{2\sigma^2}{\mu^2}\ln\left(\tfrac{c_{10}\theta}{c_{01}(1-\theta)}\right)\right]. \tag{4.28}$$

By inspection, note that the third derivatives are continuous. Substituting the first derivatives (4.24)–(4.25) and second derivatives (4.27)–(4.28) into (4.15), an expression for $B(\theta)$ can be obtained.

Examples with different distributions $p_\Theta(\theta)$ are presented below. All of the examples use scalar signals with additive Gaussian noise, $\mu = 1$, $\sigma = 1$ (4.21). As a point of reference, a comparison is made to quantizers designed under mean absolute error (MAE) [179], i.e. $d(\theta, a) = |\theta - a|$, an objective that does not account for hypothesis testing.[3]

[3]As shown by Kassam [179], minimizing the MAE criterion also minimizes the absolute distance between the cumulative distribution function of the source and the induced cumulative distribution

In the high-rate comparisons, the optimal point density for MAE [181]:

$$\lambda(\theta) = \frac{p_\Theta(\theta)^{1/2}}{\int_0^1 p_\Theta(\theta)^{1/2} d\theta}$$

is substituted into the high-rate distortion approximation for the MBRE criterion (4.18). Taking $R = \log_2(K)$, there is a constant gap between the rates using the MBRE point density and the MAE point density for all distortion values. This constant rate gap in the high-resolution regime is:

$$R_{\text{MBRE}}(D_L) - R_{\text{MAE}}(D_L) = \frac{1}{2}\log_2\left(\frac{\|p_\Theta(\theta)B(\theta)\|_{1/3}}{\|p_\Theta(\theta)\|_{1/2}\int_0^1 B(\theta)d\theta}\right).$$

The closer the ratio inside the logarithm is to one, the closer the MBRE- and MAE-optimal quantizers.

Uniformly Distributed P_0

First consider the setting in which all prior probabilities are equally likely. The MBRE of the MBRE-optimal quantizer and a quantizer designed to minimize MAE with respect to $p_\Theta(\theta)$ are plotted in Figure 4-2. (The optimal MAE quantizer for the uniform distribution is the uniform quantizer.) The plot shows MBRE as a function of K; the solid line with circle markers is the MBRE-optimal quantizer and the dotted line with asterisk markers is the MAE-optimal quantizer. D_L, the high-rate approximation to the distortion-rate function is plotted in Figure 4-3.

The performance of both quantizers is similar, but the MBRE-optimal quantizer always performs better or equally. For $K = 1, 2$, the two quantizers are identical, as seen in Figure 4-4. The plots in Figure 4-4 show $J(\theta, v_K(\theta))$ as solid and dotted lines for the MBRE- and MAE-optimal quantizers respectively; the markers are the representation points. The gray line is $J(\theta)$, the Bayes risk with unquantized prior probabilities. For $K = 3, 4$, the representation points for the MBRE-optimal quantizer are closer to $\theta = \frac{1}{2}$ than the uniform quantizer. Equivalently, the area under the point density function $\lambda(\theta)$ shown in Figure 4-5 is concentrated in the center. Each increment of K is associated with a large reduction in Bayes risk. There is a very large performance improvement from $K = 1$ to $K = 2$.

In Figure 4-6, Figure 4-7, Figure 4-8, and Figure 4-9, similar plots to those above are given for the case when the Bayes costs c_{10} and c_{01} are unequal. The unequal costs skew the Bayes risk function and consequently the representation point locations and point density function. The difference in performance between the MBRE-optimal

function of the quantized output. Since the induced distribution from quantization is used as the population prior distribution for hypothesis testing, requiring this induced distribution to be close to the true unquantized distribution is reasonable. If distance between probability distributions is to be minimized according to the Kullback-Leibler discrimination between the true and induced distributions (which is defined in terms of likelihood ratios), an application of Pinsker's inequality shows that a small absolute difference is required [180]. Although a reasonable criterion, MAE is suboptimal for hypothesis testing performance as seen in the examples.

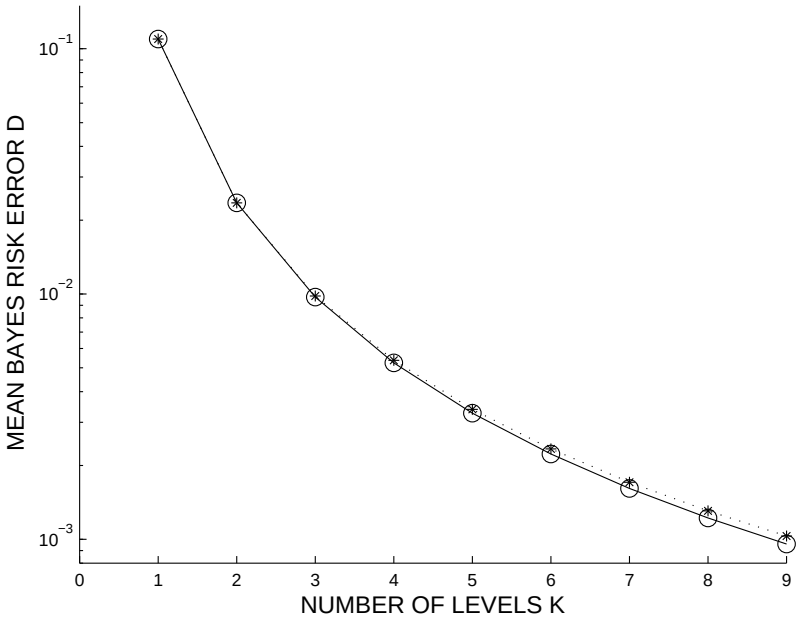

Figure 4-2. MBRE for uniformly distributed Θ and Bayes costs $c_{10} = c_{01} = 1$ plotted on a logarithmic scale as a function of the number of quantization levels K; the solid line with circle markers is the MBRE-optimal quantizer and the dotted line with asterisk markers is the MAE-optimal uniform quantizer.

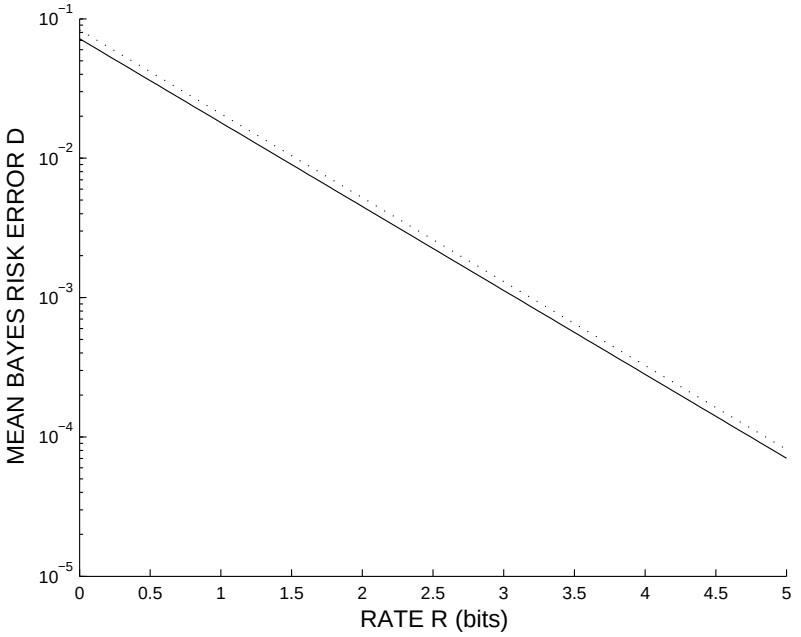

Figure 4-3. High-rate approximation of distortion-rate function D_L for uniformly distributed Θ and Bayes costs $c_{10} = c_{01} = 1$; the solid line is the MBRE-optimal quantizer and the dotted line is the MAE-optimal uniform quantizer.

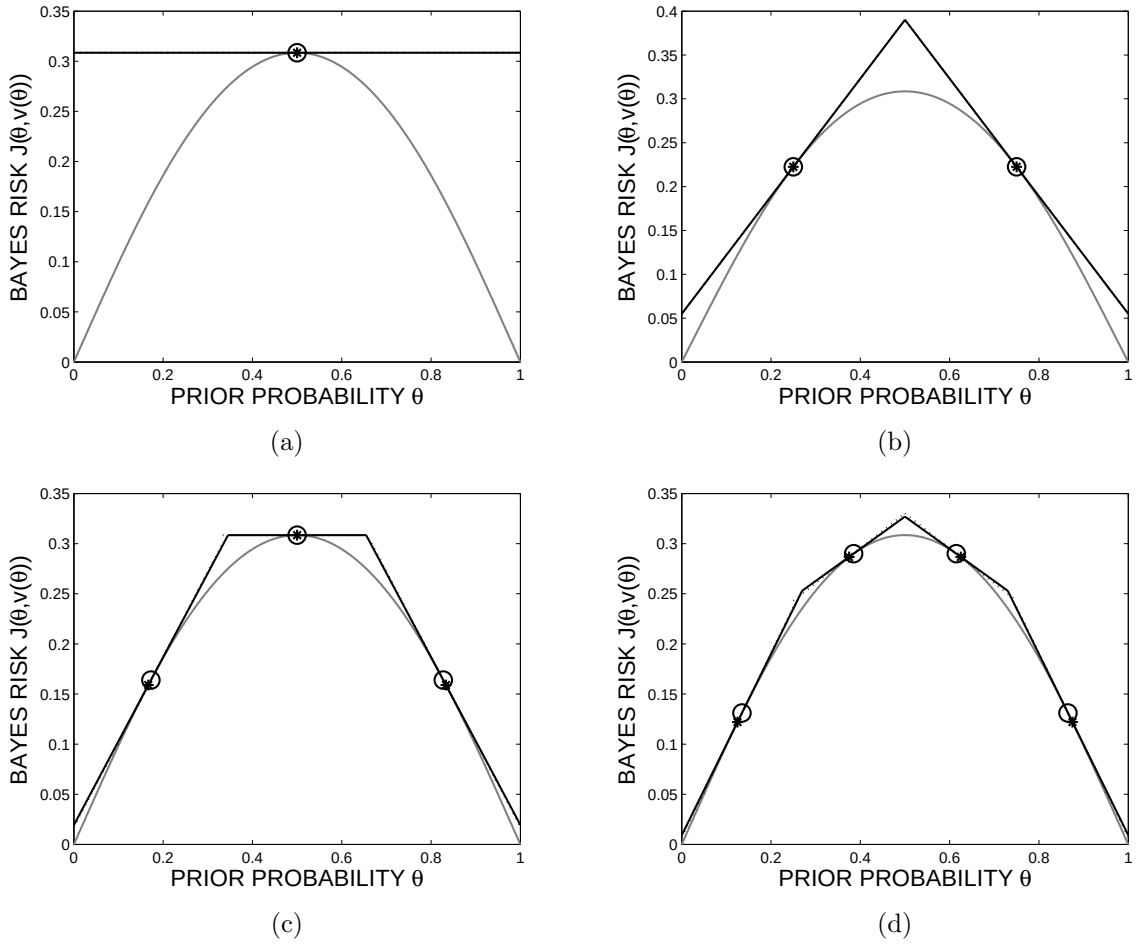

Figure 4-4. Quantizers for uniformly distributed Θ and Bayes costs $c_{10} = c_{01} = 1$. $J(\theta, v_K(\theta))$ is plotted for (a) $K = 1$, (b) $K = 2$, (c) $K = 3$, and (d) $K = 4$; the markers, circle and asterisk for the MBRE-optimal and MAE-optimal quantizers respectively, are the representation points $\{a_k\}$. The gray line is the unquantized Bayes risk $J(\theta)$.

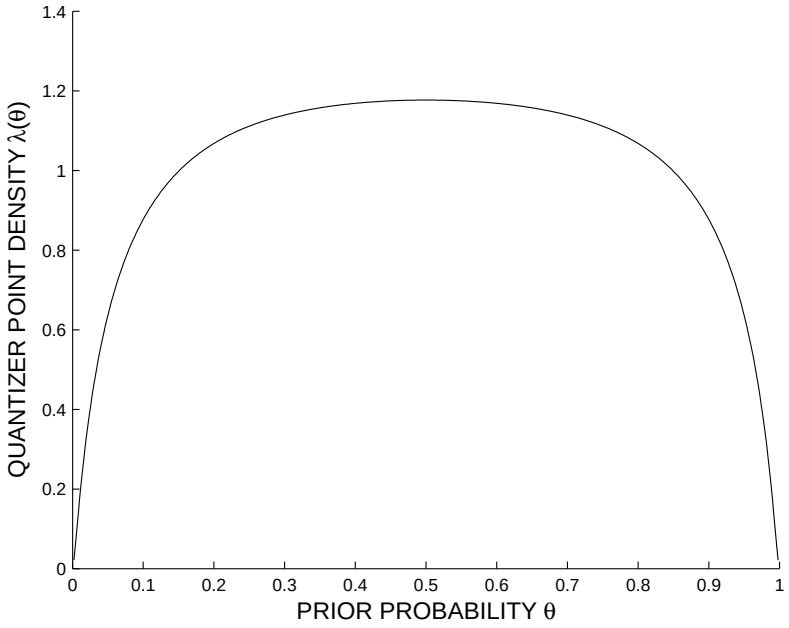

Figure 4-5. Optimal MBRE point density for uniformly distributed Θ and Bayes costs $c_{10} = c_{01} = 1$.

and MAE-optimal quantizers is greater in this example because the MAE-criterion cannot incorporate the Bayes costs, which factor into MBRE calculation.

Beta Distributed Θ

Now, consider a non-uniform distribution for Θ, in particular the Beta$(5, 2)$ distribution. The probability density function is shown in Figure 4-10. The MBRE of the MBRE-optimal and MAE-optimal quantizers are in Figure 4-11. Here, there are also large improvements in performance with an increase in K. The high-rate approximation to the distortion-rate function for this example is given in Figure 4-12.

The representation points $\{a_k\}$ are most densely distributed where $\lambda(\theta)$, plotted in Figure 4-13, has mass. In particular, more representation points are in the right half of the domain than in the left, as seen in Figure 4-14.

■ 4.5 Implications on Human Decision Making

The previous sections formulated the minimum MBRE quantization problem and discussed how to find the optimal MBRE quantizer. Having established the mathematical foundations of hypothesis testing with quantized priors, implications of such resource-constrained decoding on civic life may be explored.

Consider the particular setting for human decision making mentioned in Section 4.1: a referee determining whether a player has committed a foul or not using both his or her noisy observation and prior experience. The fraction of plays in which

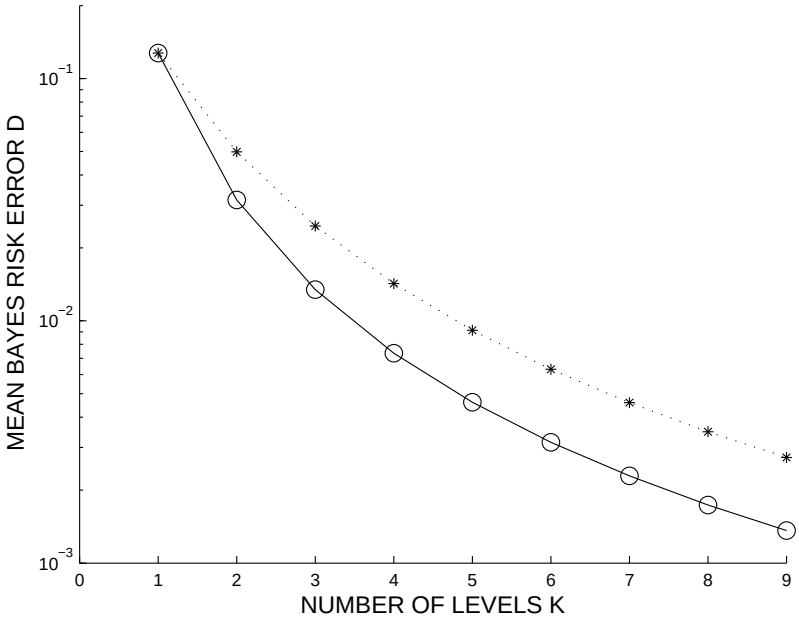

Figure 4-6. MBRE for uniformly distributed Θ and Bayes costs $c_{10} = 1, c_{01} = 4$ plotted on a logarithmic scale as a function of the number of quantization levels K; the solid line with circle markers is the MBRE-optimal quantizer and the dotted line with asterisk markers is the MAE-optimal uniform quantizer.

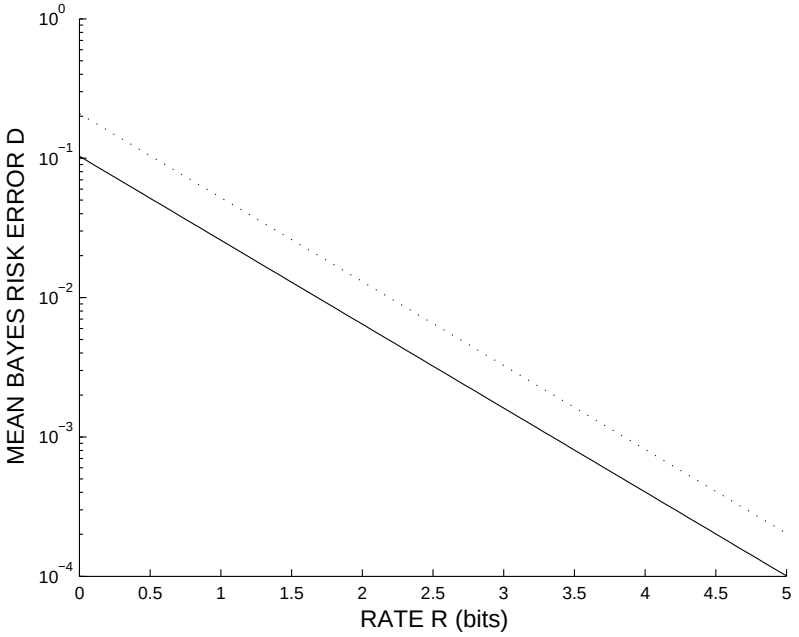

Figure 4-7. High-rate approximation of distortion-rate function D_L for uniformly distributed Θ and Bayes costs $c_{10} = 1, c_{01} = 4$; the solid line is the MBRE-optimal quantizer and the dotted line is the MAE-optimal uniform quantizer.

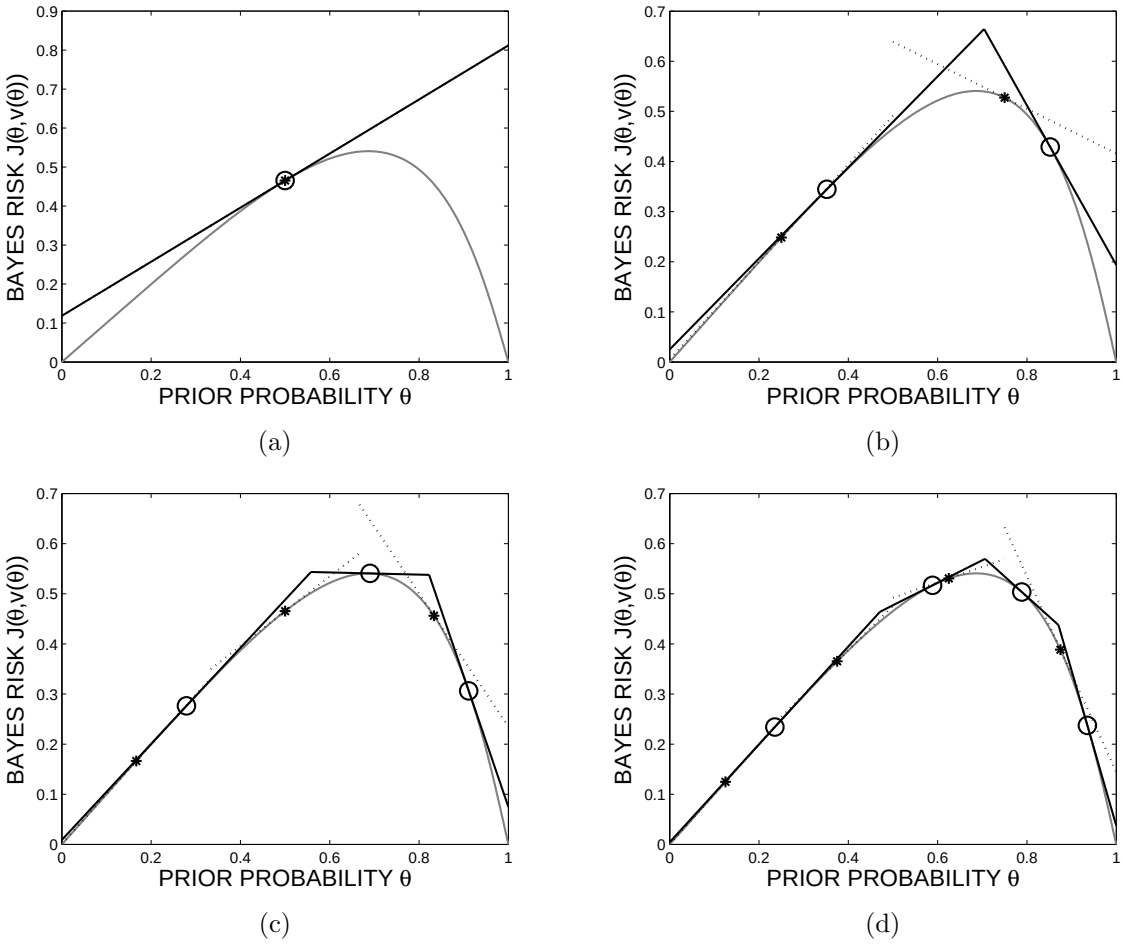

Figure 4-8. Quantizers for uniformly distributed Θ and Bayes costs $c_{10} = 1, c_{01} = 4$. $J(\theta, v_K(\theta))$ is plotted for (a) $K = 1$, (b) $K = 2$, (c) $K = 3$, and (d) $K = 4$; the markers, circle and asterisk for the MBRE-optimal and MAE-optimal quantizers respectively, are the representation points $\{a_k\}$. The gray line is the unquantized Bayes risk $J(\theta)$.

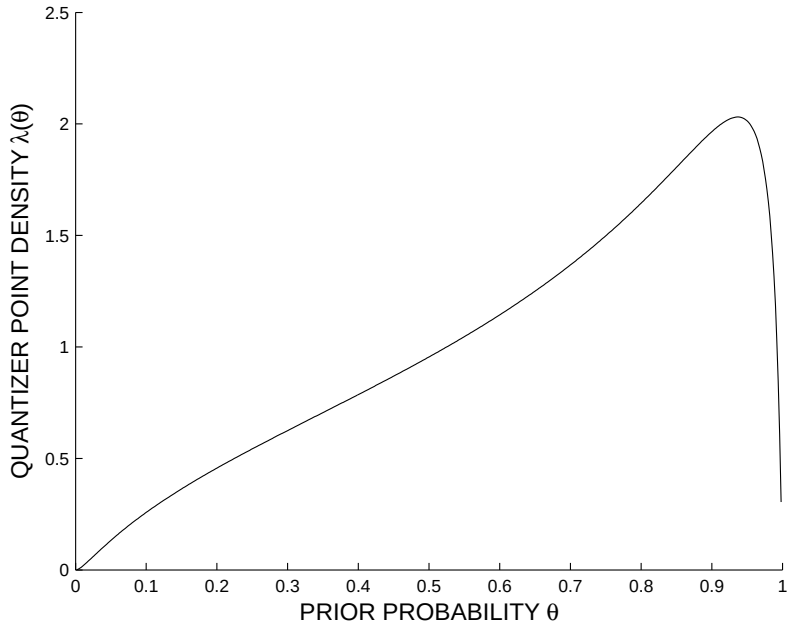

Figure 4-9. Optimal MBRE point density for uniformly distributed Θ and Bayes costs $c_{10} = 1, c_{01} = 4$.

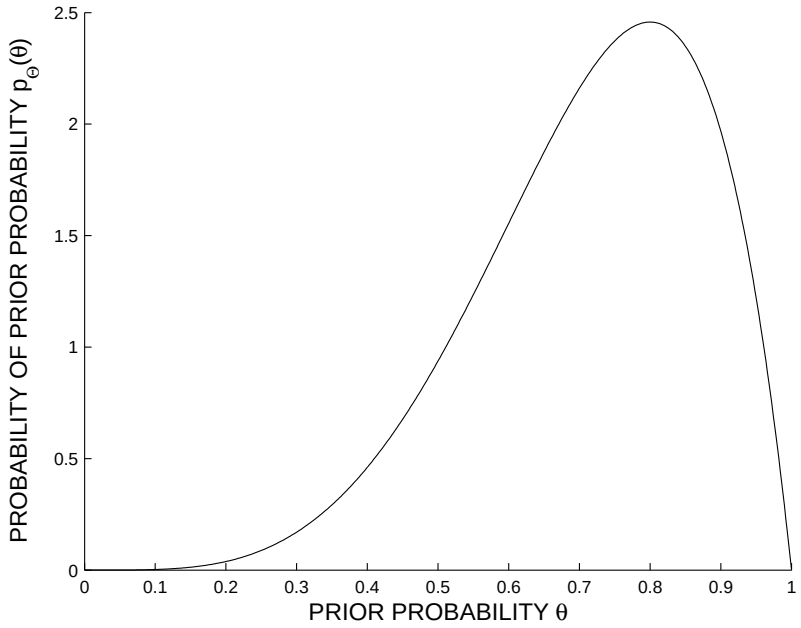

Figure 4-10. The probability density function $p_\Theta(\theta)$ for the Beta$(5, 2)$ distribution.

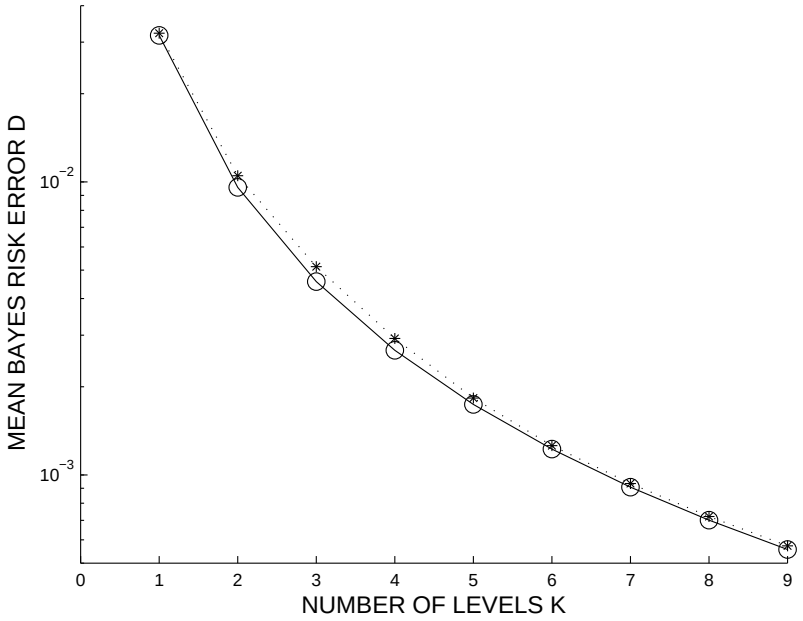

Figure 4-11. MBRE for Beta$(5, 2)$ distributed Θ and Bayes costs $c_{10} = c_{01} = 1$ plotted on a logarithmic scale as a function of the number of quantization levels K; the solid line with circle markers is the MBRE-optimal quantizer and the dotted line with asterisk markers is the MAE-optimal uniform quantizer.

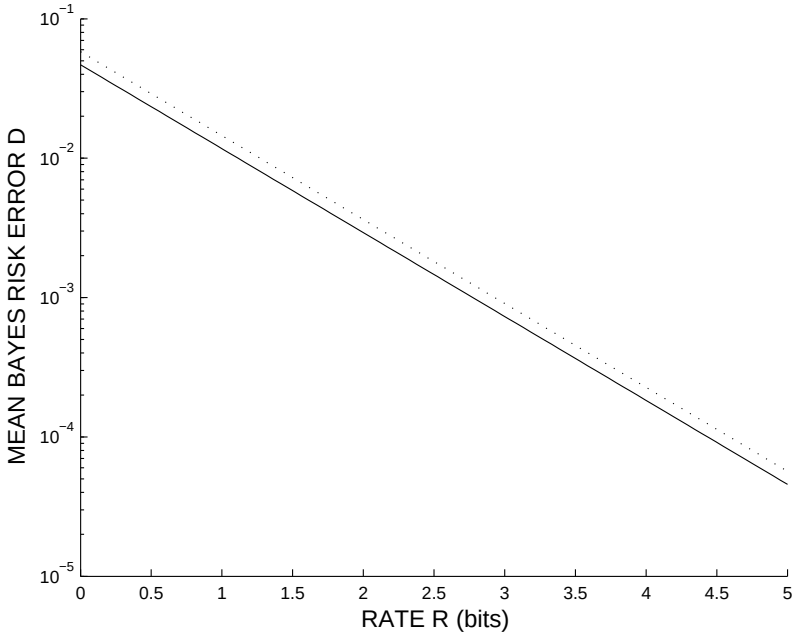

Figure 4-12. High-rate approximation of distortion-rate function D_L for Beta$(5, 2)$ distributed Θ and Bayes costs $c_{10} = c_{01} = 1$; the solid line is the MBRE-optimal quantizer and the dotted line is the MAE-optimal uniform quantizer.

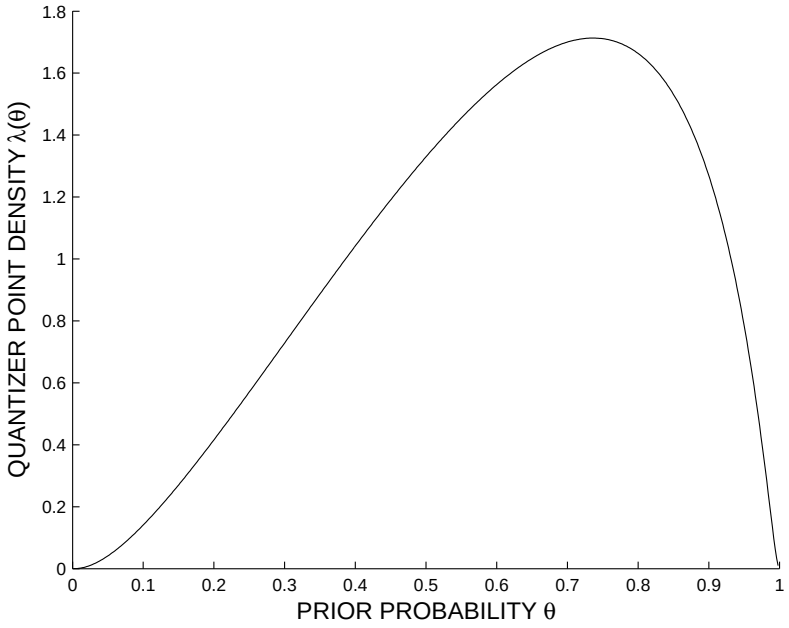

Figure 4-13. Optimal MBRE point density for Beta$(5, 2)$ distributed Θ and Bayes costs $c_{10} = c_{01} = 1$.

a player commits a foul is that player's prior probability. There is a distribution of prior probabilities over the population of players. As also mentioned in Section 4.1, human decision makers categorize into a small number of categories due to limitations in information processing capacity [159]. Decisions by humans may be modeled via quantization of the distribution of prior probabilities and the use of the quantization level centroid of the category in which a player falls as the prior probability when decoding that player's action. This is a form of bounded rationality [182–185], which steps away from full rationality where people act optimally [186–189].

A referee will do a better job with more categories rather than fewer. A police officer confronting an individual with whom he or she has prior experience will make a better decision if he or she has the mental categories 'law-abiding,' 'delinquent,' 'criminal,' and 'nefarious,' rather than just 'good' and 'bad.' Similarly, a doctor will have a smaller probability of error when interpreting a diagnostic test if he or she knows the prior probability of a positive result for many categorizations of patients rather than just one for the entire population at large. Additional examples could be given for a variety of decision-making tasks. Such implications are not surprising; however, fairly interesting implications arise when additional features are added to the decision-making scenario. In particular, the setting when the quantization of distinct populations is done separately.

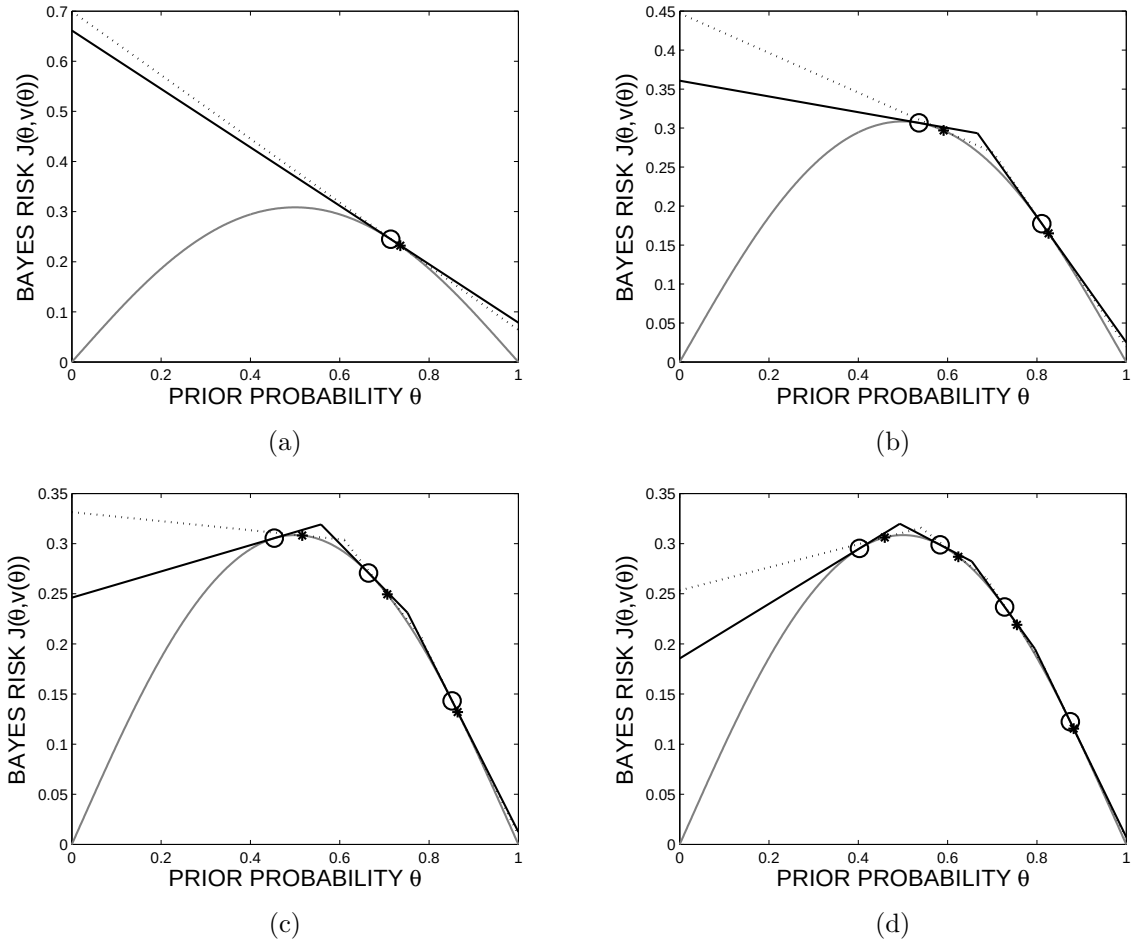

Figure 4-14. Quantizers for Beta$(5, 2)$ distributed Θ and Bayes costs $c_{10} = 1, c_{01} = 4$. $J(\theta, v_K(\theta))$ is plotted for (a) $K = 1$, (b) $K = 2$, (c) $K = 3$, and (d) $K = 4$; the markers, circle and asterisk for the MBRE-optimal and MAE-optimal quantizers respectively, are the representation points $\{a_k\}$. The gray line is the unquantized Bayes risk $J(\theta)$.

4.5.1 Multiple Populations

Suppose a decision maker must deal with subpopulations that are distinguished according to a socially observable part of identity like race [190]. For clarity and ease of connection to empirical studies, we restrict to two groups and use 'black' and 'white' to denote them. The rational memory-constrained decision maker should ignore the dimension of race altogether and simply quantize along the θ dimension due to the theorem of irrelevance [191, Theorem 8.4.2].

Although there is some debate in the social cognition literature [192], it is thought that race and gender categorization is essentially automatic, particularly when a human actor lacks the motivation, time, or cognitive capacity to think deeply. Therefore social cognition constraints prevent the decision maker from ignoring non-θ dimensions. Automaticity of racial categorization results in two quantizers designed separately for the two populations. The total quota on representation points, K_t, is split into some number of points for whites and some number for blacks, denoted $K_t = K_w + K_b$. The separate quantizers may then be denoted $v_{K_w}(\cdot)$ and $v_{K_b}(\cdot)$. It is assumed that the white population $(\{p_W(w)\}_{\Theta\in\mathcal{T}}, p_\Theta(\theta))$ and the black population $(\{p_B(b)\}_{\Psi\in\mathcal{T}}, p_\Psi(\psi))$ are identical.[4]

The definition of mean Bayes risk error for two populations, from the perspective of the decision maker, extends as:

$$\begin{aligned} D^{(2)} &= \tfrac{w_0}{w_0+b_0}\left\{E[J(\Theta, v_{K_w}(\Theta))] - E[J(\Theta)]\right\} + \tfrac{b_0}{w_0+b_0}\left\{E[J(\Psi, v_{K_b}(\Psi))] - E[J(\Psi)]\right\} \\ &= \tfrac{w_0}{w_0+b_0}E[J(\Theta, v_{K_w}(\Theta))] + \tfrac{b_0}{w_0+b_0}E[J(\Theta, v_{K_b}(\Theta))] - E[J(\Theta)], \end{aligned}$$

where w_0 and b_0 are the number of whites and blacks relevant to the decision maker.[5] Under Radner's notion of costly rationality [185], the goal is to minimize this extended Bayes risk error by finding the optimal quantizers $v_{K_w}(\cdot)$ and $v_{K_b}(\cdot)$ and the optimal allocation of representation points K_w and K_b.

Since the two populations are identical, $v_{K_w}(\cdot)$ and $v_{K_b}(\cdot)$ should be the quantizers that were to be designed in Section 4.2. Thus the problem reduces to minimizing expected Bayes risk error over all $K_t - 1$ possible allocations of K_w and K_b. Although there are sophisticated algorithms for optimal allocation of levels [199], just measuring the performance of all allocations and choosing the best one suffices.

Fryer and Jackson have previously suggested that it is better to allocate more representation points to a majority population than to a minority population [197].

[4] There might eventually be an adverse selection effect, as happens in insurance markets [193] or used car markets with 'lemons' [194], with $p_\Theta(\theta)$ or $p_\Psi(\psi)$ subject to modification through behavior, education, or investment. As noted in the previous information-based discrimination literature [195–197], a downward spiral of underinvestment in positive behavior or human capital by minority groups may occur when the minority groups are coarsely categorized. Thus two populations may not be identical, but this possibility is ignored here.

[5] One might assume that w_0 and b_0 are simply the number of whites and blacks in the general population, however these numbers are actually based on the social interaction pattern of the decision maker. Due to segregation in social interaction, see e.g. [198] and references therein, there is greater intra-population interaction than inter-population interaction. The decision maker has more training data from intra-population interaction.

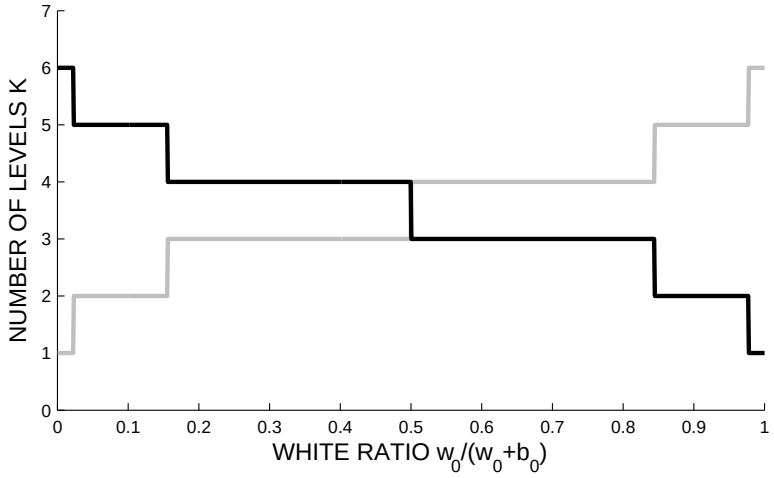

Figure 4-15. Optimal allocation of quantizer sizes to the white population and black population for $K_t = 7$ as a function of the proportion of whites. The gray line is K_w and the black line is K_b. The distribution of the prior probability is Beta(5,2), the channel is additive white Gaussian noise with unit signal to noise ratio, and the Bayes costs are $c_{10} = c_{01} = 1$.

Optimal allocation in this model yields the same result when the notion of majority and minority are with respect to the decision maker's interaction pattern. If w_0 is larger than b_0, it is better to allocate more representation points to whites whereas if b_0 is larger than w_0, it is better to allocate more representation points to blacks. An example of optimal allocation is shown in Figure 4-15.

The model may easily be extended to consider increased numbers of racial groups N. An exhaustive search over representation point allocations for various integer partitions of K_t: $K_t = K_1 + K_2 + \cdots + K_N$ may be used; this only involves the design of a linear number of optimal quantizers and a very small optimization problem. One may also consider the effects of a higher-dimensional socially-observable attribute space; identity is not just race. In fact some social dimensions may be consciously and explicitly correlated in order to further define identity [190]. The gains of vector quantization over scalar quantization are enhanced when there is dependence among dimensions [169].

■ 4.5.2 Social Discrimination

Due to social segregation, there is greater intra-population interaction than inter-population interaction. Whites interact more with whites whereas blacks interact more with blacks. Moreover, social life and economic life are so inextricably intertwined that human decision makers are unable to completely discount social experience in determining how to deploy their limited decision making resources [197]; the spillover may be due to the use of common brain regions for both kinds of interaction [200].

In the mathematical model, one would therefore expect the $w_0/(w_0 + b_0)$ ratio of a white decision maker to be greater than the $w_0/(w_0 + b_0)$ ratio of a black decision

maker. This fact, together with optimal level allocation and Theorem 4.6, implies that a white decision maker would perform worse than a black decision maker when dealing with blacks and vice versa.

This prediction is born out experimentally. A large body of literature in face recognition shows exactly the predicted own race bias effect, observed colloquially as "they [other-race persons] all look alike." In particular, both parts of the Bayes risk, P_e^{I} and P_e^{II} increase when trying to recognize members of the opposite population [201]. Own race bias in face recognition has been verified with laboratory experiments.

Econometric studies also provide a source for comparison to the proposed decision making model. The addition of police officers of a given race is associated with an increase in the number of arrests of suspects of a different race but has little impact on same-race arrests. The effect is more pronounced for minor offenses where the prior probability presumably plays a bigger role than the measurement [202]. There are similar own-race bias effects in the decision by police to search a vehicle during a traffic stop [203] and in the decision of National Basketball Association (NBA) referees to call a foul [204]. The rate of searching and the rate of foul calling are greater when the decision maker is of a different race than the driver and player, respectively.

A major difficulty in interpreting these studies, however, is that the ground truth is not known. Higher rates of arrest or foul calls, for example, may be explained by either a greater P_e^{I} or smaller P_e^{II}. It is possible that a greater probability of missed fouls would actually decrease the number of fouls called. This motivates a closer look at the Bayes risk by teasing it apart into its constituent parts and examining the Bayes costs in detail.

Using basketball fouls as a running example, the measurable quantity is the probability that a foul is called. This rate of fouls is:

$$\Pr[f_D^K(Y) = 1] = 1 - \theta + \theta P_e^{\mathrm{I}}(v_K(\theta)) - (1-\theta)P_e^{\mathrm{II}}(v_K(\theta)). \tag{4.29}$$

Looking at the average performance of a referee over the populations of black and white players, compare the expected foul rates on whites and blacks:

$$\begin{aligned}\Delta(c_{10}, c_{01}) &= E\left[\Pr[f_D^{K_b}(Y) = 1] - \Pr[f_D^{K_w}(Y) = 1]\right] \\ &= E[\Theta P_e^{\mathrm{I}}(v_{K_b}^*(\Theta)) - (1-\Theta)P_e^{\mathrm{II}}(v_{K_b}^*(\Theta)) - \Theta P_e^{\mathrm{I}}(v_{K_w}^*(\Theta)) + (1-\Theta)P_e^{\mathrm{II}}(v_{K_w}^*(\Theta))].\end{aligned} \tag{4.30}$$

If this discrimination quantity Δ is greater than zero, then the referee is calling more fouls on blacks. If Δ is less than zero, then the referee is calling more fouls on whites.

The dependence of Δ on c_{10} and c_{01} has been explicitly notated on the left side of (4.30) and is implicit in the two types of error probabilities on the right side of (4.30). The value of Δ also depends on the unquantized prior distribution $p_\Theta(\theta)$, the values of K_w and K_b, and the channel. Fixing these determines the regions in the c_{10}–c_{01} plane where a referee would call more fouls on blacks and where a referee would call more fouls on whites. This is shown in Figure 4-16. For uniform $p_\Theta(\theta)$, the two regions are divided by the line $c_{01} = c_{10}$.

For any population and channel, there is one half-plane where a referee would call

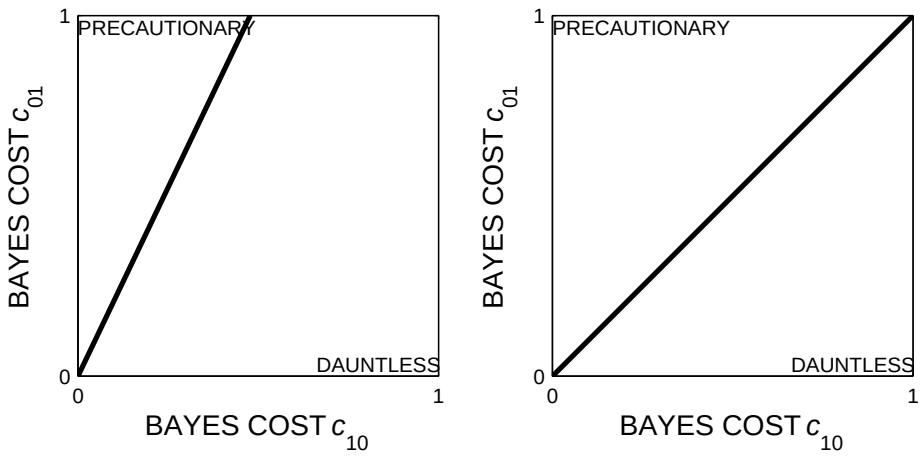

Figure 4-16. Dividing line between Bayes cost region in which referee will call more fouls on blacks and region in which referee will call more fouls on whites. A referee with $K_b < K_w$ will call more fouls on blacks in the upper left region and more fouls on whites in the lower right region, which correspond to precautionary and dauntless respectively. For the left panel, the distribution of the prior probability is Beta(5,2), the channel is additive white Gaussian noise with unit signal to noise ratio, and the level allocation is $K_b = 3$, $K_w = 4$. For the right panel, the prior probability distribution is uniform.

more fouls on black players and the other half-plane where the referee would call more fouls on white players. To reiterate, just because the Bayes risk for foul-calling on black players is greater than that for white players, it does not automatically imply that the foul call rate for blacks is higher. The high Bayes risk could well be the result of a preponderance of missed foul calls.

This result may be interpreted in terms of precautionary [205] and dauntless [206] decision making. The precautionary principle corresponds to a Bayes cost assignment with $c_{01} > c_{10}$, whereas the dauntless principle corresponds to a Bayes cost assignment with $c_{01} < c_{10}$. Thus, a referee with $K_w > K_b$ that calls more fouls on black players is precautionary and more fouls on white players is dauntless. A referee with $K_w < K_b$ that calls more fouls on black players is dauntless and more fouls on white players is precautionary.

Econometric studies give differences of differences to show racial bias. The first 'difference' is the difference in foul call rate between black players and white players, Δ. The second 'difference' is the difference in Δ between white referees and black referees. Denoting the foul call rate difference of a white referee by Δ_W and the foul call rate difference of a black referee by Δ_B, the difference of differences is $\Delta_W - \Delta_B$.

Figure 4-17 plots the difference of differences as a function of the ratio c_{01}/c_{10} for two different population distributions, the Beta(5,2) distribution and the uniform distribution. The right side of the plot is the precautionary regime, where white referees would call more fouls on black players than black referees. For the particular examples, if $c_{01}/c_{10} = 10$, then the white referee has a foul call rate 0.0132 greater than the black referee on black players for the Beta(5,2) distribution and 0.0142 greater for the uniform distribution.

The left side of the plot is the dauntless regime, where white referees would call

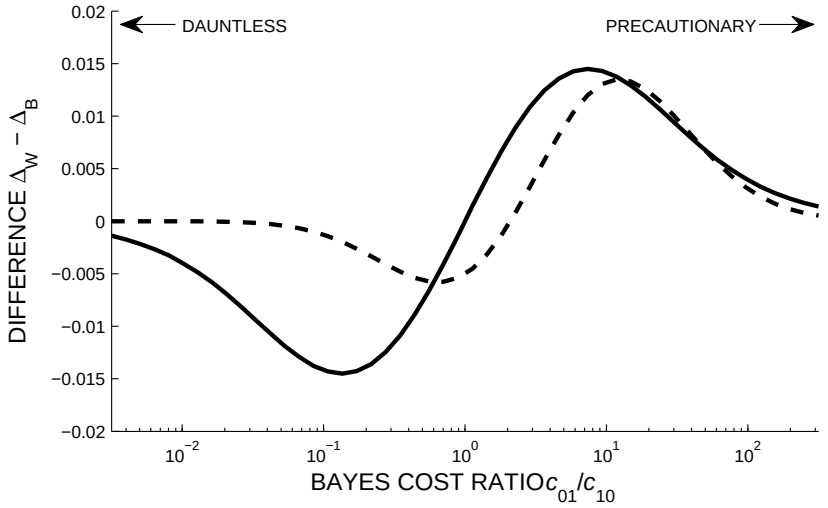

Figure 4-17. Difference of differences in foul calling as a function of the Bayes cost ratio. The white referee has $K_w = 4$, $K_b = 3$ and the black referee has $K_w = 3$, $K_b = 4$. For the dashed line, the distribution of the prior probability is Beta(5,2) and the channel is additive white Gaussian noise with unit signal to noise ratio. For the solid line, the prior probability distribution is uniform.

fewer fouls on black players than black referees. For the particular examples, if $c_{01}/c_{10} = 0.1$, then the white referee has a foul call rate 0.0013 less than the black referee on black players for the Beta(5,2) distribution and 0.0142 less for the uniform distribution. In these examples, the white referee has $K_w = 4$, $K_b = 3$, and the black referee has $K_w = 3$, $K_b = 4$.[6]

■ 4.5.3 The Price of Segregation

In the memory-constrained decision making setting with separate populations, decision makers of different races exhibited different biases because they had different K_w and K_b allocations due to different $w_0/(w_0 + b_0)$ ratios. This ratio was not the actual fraction of whites whose actions would be decoded by the decision maker, but was determined in part by the decision maker's segregated social life. If decision makers of all races had a bias that matched the true white fraction, then the phenomenon of racial bias would be rational. Since different decision making biases cannot simultaneously be optimal, divergent racial bias in decision making is an irrational phenomenon.

Fixing limitations of human memory, automaticity of racial classification, and intertwining of social and economic life, it follows that social segregation causes mismatch between social and economic lives and is therefore the root cause of irrational racial bias. To draw connections to econometric studies where ground truth is not known, the previous section used differences of differences. In analogy with notions of

[6]Note that there is no requirement for the white referee to have $K_w > K_b$ and the black referee to have $K_w < K_b$. It is only required that the K_w of the white referee be greater than the K_w of the black referee (assuming the same K_t). A qualitatively similar plot to Figure 4-17 is produced, e.g. when the white referee has $K_w = 5$, $K_b = 2$, and the black referee has $K_w = 4$, $K_b = 3$.

welfare loss in economic theory like deadweight loss, the social cost of monopoly [207, Chapter 4], and the price of anarchy [208], a *price of segregation* is defined here as a way to measure the deleterious effect of segregation. Note that segregation mismatch is distinct from mismatched decoding arising from memory constraints.

Let π_{true} be the probability that a member of the white population $(\{p_W(w)\}_\Theta, p_\Theta(\theta))$ is selected for decoding rather than a member of the black population $(\{p_B(b)\}_\Psi, p_\Psi(\psi))$. A particular decision maker that leads a segregated life, on the other hand, will have a white ratio $\pi_{\text{seg}} = w_0/(w_0 + b_0)$. The mean Bayes risk error, from the perspective of society, under the true white fraction is

$$D^{(2)}(\pi_{\text{true}}) = \pi_{\text{true}} E[J(\Theta, v_{K_w(\pi_{\text{true}})}(\Theta))] + (1 - \pi_{\text{true}}) E[J(\Theta, v_{K_b(\pi_{\text{true}})}(\Theta))] - E[J(\Theta)]$$

whereas the MBRE, from the perspective of society, under the segregated white fraction is

$$D^{(2)}(\pi_{\text{seg}}) = \pi_{\text{true}} E[J(\Theta, v_{K_w(\pi_{\text{seg}})}(\Theta))] + (1 - \pi_{\text{true}}) E[J(\Theta, v_{K_b(\pi_{\text{seg}})}(\Theta))] - E[J(\Theta)].$$

The difference between these two is the price of segregation:

$$\begin{aligned}
\Pi &= D^{(2)}(\pi_{\text{true}}) - D^{(2)}(\pi_{\text{seg}}) \\
&= \pi_{\text{true}} \left\{ E[J(\Theta, v_{K_w(\pi_{\text{true}})}(\Theta))] - E[J(\Theta, v_{K_w(\pi_{\text{seg}})}(\Theta))] \right\} \\
&\quad + (1 - \pi_{\text{true}}) \left\{ E[J(\Theta, v_{K_t - K_w(\pi_{\text{true}})}(\Theta))] - E[J(\Theta, v_{K_t - K_w(\pi_{\text{seg}})}(\Theta))] \right\}
\end{aligned}$$

The price of segregation Π depends strongly on the discontinuous, integer-valued $K_w(\cdot)$ function, and is also discontinuous. The price of segregation is a non-decreasing function of the level of segregation mismatch $|\pi_{\text{true}} - \pi_{\text{seg}}|$. An example of the price of segregation for a particular system and several different values of π_{true} is shown in Figure 4-18. Notice that if the level of mismatch is small, there may be no price of segregation. Figure 4-19 shows a similar plot but for a larger value of K_t. The range of mismatch that leads to zero Π is smaller, but the incurred Π in this non-zero regime may also be smaller.

The model predicts that greater homogeneity of social interaction among people would mitigate the price of segregation by driving the π_{seg} for all decision makers closer to π_{true}. This draws a connection to intergroup contact theory [209, 210]. One branch of contact theory suggests that mixing between members of different groups reduces prejudice since it allows individuals the chance to see previously unnoticed similarities and counter-stereotypic characteristics and behaviors in one another [210, 211], a conclusion similar to the model predictions.

Perhaps unexpectedly, social interaction is not linear in the overall ratio of subgroup populations [198], so a public policy to reduce segregation would be difficult to formulate. Discrimination appears to be a permanent artifact of memory-constrained decoding and automaticity of racial categorization.

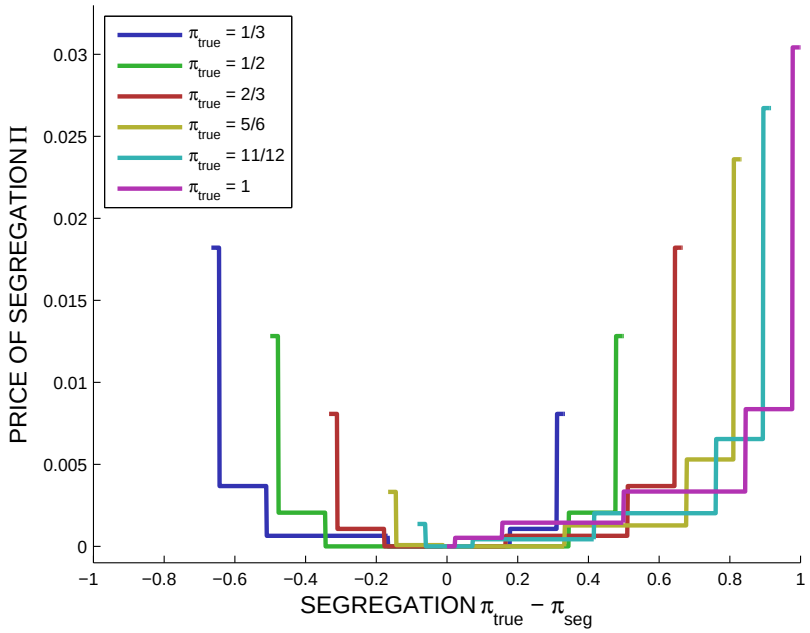

Figure 4-18. The price of segregation Π as a function of the level of segregation mismatch $\pi_{\text{true}} - \pi_{\text{seg}}$ for several values of π_{true}. The distribution of the prior probability is Beta(5,2), the channel is additive white Gaussian noise with unit signal to noise ratio, and the Bayes costs are $c_{10} = c_{01} = 1$, and $K_t = 7$.

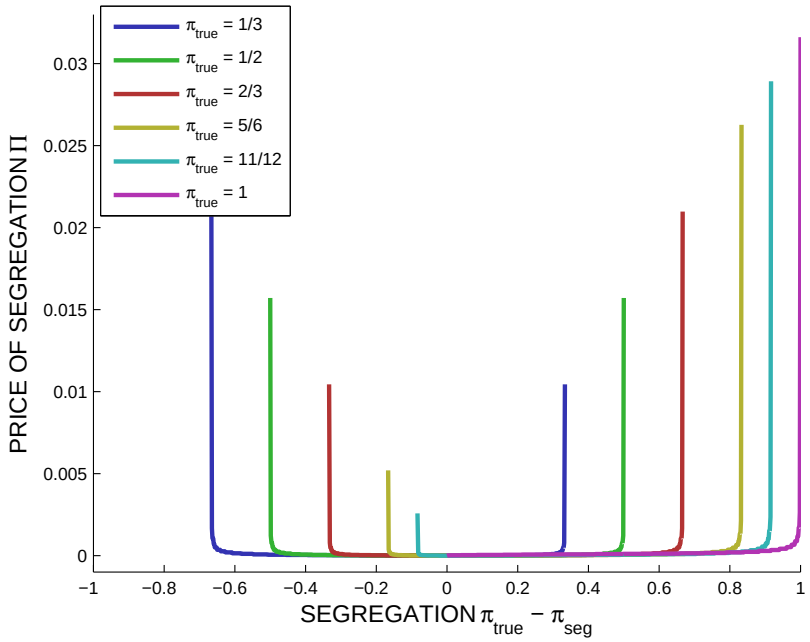

Figure 4-19. The price of segregation Π as a function of the level of segregation mismatch $\pi_{\text{true}} - \pi_{\text{seg}}$ for several values of π_{true}. The distribution of the prior probability is Beta(5,2), the channel is additive white Gaussian noise with unit signal to noise ratio, and the Bayes costs are $c_{10} = c_{01} = 1$, and $K_t = 37$.

4.6 Discussion

By analyzing binary uncoded communication, this chapter demonstrated that memory constraints in decoding circuits might reduce the performance of communication systems. Quantizing the space of decoding rules led to mismatched decoding and resultant performance loss. It would be of interest to extend this work by determining optimal methods of quantizing decoding rules for coded communication using properties of general mismatched decoding [42].

The setting of Bayesian hypothesis testing (uncoded transmission) with a population of sources, but where the decoder may only use a quantized representation of the true source in designing a decision rule was considered. Optimal quantizers for this purpose were determined using a new fidelity criterion built on the Bayes risk function. For larger message sets, $M > 2$, vector quantization rather than scalar quantization would be required, but determining the Lloyd-Max conditions and high-rate theory is no different conceptually due to the geometry of the Bayes risk function. Theories of quantized prior hypothesis testing, when combined with theories of social cognition and facts about social segregation, also led to a generative model of discriminative behavior in human affairs.

As an extension, one might consider a restricted class of quantizers rather than considering optimal quantization. Such restriction may model further cognitive constraints on human decision makers or information processing constraints in decoding circuits. In particular, Fryer and Jackson had suggested a heuristic algorithm for quantizer design based on splitting groups [197], which is a rediscovery of a tree-structured vector quantizer (TSVQ) design algorithm [212, Figure 20].

Memory constraints limited the adaptability of the decoder here, but limited online adaptability in information processing is a general problem that is only partially solved by universal information theory [41, 154]. A quantization-theoretic approach with novel, task-driven fidelity criteria may shed light on general trade-offs between adaptation and universality.

Chapter 5

Operation Reliability—Transient Faults

In the lifetime of a communication system, construction is followed by operation. The previous two chapters considered limitations in provisioning material for constructing decoders. This chapter and the next consider unreliability in operating decoders.

Traditional communication theory assumes that decoding algorithms perform without error, however noise is inherent in computation circuits just as in communication channels. The goal of this chapter is to investigate limits of communication systems with noisy decoders; there are dual motivations. The first is the eminently practical motivation of determining how well channel codes work when decoders are faulty. The second is the deeper motivation of determining fundamental limits for processing unreliable signals with unreliable computational devices. The motivations are intertwined. As noted by Pierce [21], "The down-to-earth problem of making a computer work, in fact, becomes tangled with this difficult philosophical problem: 'What is possible and what is impossible when unreliable circuits are used to process unreliable information?'"

A first step in understanding these issues is to analyze the average symbol error P_e^{sym} of a particular class of codes and decoding techniques: iterative message-passing decoding algorithms for low-density parity-check (LDPC) codes. Focus is placed on LDPC codes because they have emerged as a class of codes that have performance at or near the Shannon limit [213, 214] and yet are sufficiently structured as to have decoders with circuit implementations [215–217].[1] Recall that these codes were described in Section 2.3.

When the code is represented as a factor graph, decoding algorithm computations occur at vertices and decoding algorithm communication is carried out over edges. Correspondence between the factor graph and the algorithm is not only a tool for exposition but also the way decoding circuits are implemented [215–217], with vertices as circuit nodes and edges as circuit wires, similar to Figure 3-1. Here, there will be transient local computation and message-passing errors, whether the decoder is analog or digital.

To facilitate analysis, concentration of decoding performance around its average is

[1] One may also consider the effect of encoder complexity [218], however encoder noise need not be explicitly considered, since it may be incorporated into channel noise, using the noise combining argument suggested by Figure 5-2.

shown to hold when noise is introduced into message-passing and local computation. Density evolution equations for simple faulty iterative decoders are derived. In one model, computing nonlinear estimation thresholds shows that performance degrades smoothly as decoder noise increases, but arbitrarily small probability of error is not achievable. Probability of error may be driven to zero, however, in another system model; the decoding threshold again decreases smoothly with decoder noise. As an application of the methods developed, an achievability result for reliable memory systems constructed from unreliable components is provided.

The remainder of the chapter is organized as follows. Section 5.1 further reviews motivations and related work. Section 5.2 formalizes notation and Section 5.3 gives concentration results that allow the density evolution method of analysis, generalizing results in [43]. A noisy version of the Gallager A decoder for processing the output of a binary symmetric channel is analyzed in Section 5.4, where it is shown that Shannon reliability is unattainable. In Section 5.5, a noisy decoder for additive white Gaussian noise (AWGN) channels is analyzed. For this model, the probability of error may be driven to zero and the decoding threshold degrades smoothly as a function of decoder noise. As an application of the results of Section 5.4, Section 5.6 precisely characterizes the information storage capacity of a memory built from unreliable components. Section 5.7 provides further discussion.

5.1 Transient Circuit Faults: Causes and Consequences

Technological Trends

Although always present [11, 219], recent technological trends in digital circuit design bring practical motivations to the fore [35, 220, 221]. The 2008 update of the ITRS points out that for complementary metal-oxide-silicon (CMOS) technology, increasing power densities, decreasing supply voltages, and decreasing sizes have increased sensitivity to cosmic radiation, electromagnetic interference, and thermal fluctuations. The ITRS further says that an ongoing shift in the manufacturing paradigm will dramatically reduce costs but will lead to more transient failures of signals, logic values, devices, and interconnects. Device technologies beyond CMOS, such as single-electron tunnelling technology [222], carbon-based nanoelectronics [223], and chemically assembled electronic nanocomputers [224], are also projected to enter production, but they all display erratic, random device behavior [225, 226].

Analog computations are always subject to noise [227, 228]. Similar issues arise when performing real-valued computations on digital computers since quantization, whether fixed-point or floating-point, is often well-modeled as bounded, additive stochastic noise [229].

Fault-Tolerant Computing

Fault-tolerant computing theory [32, 33] has provided limits for processing reliable signals (inputs) with unreliable circuits [14, 21, 28–31]. This work brings it together with processing unreliable communication signals.

A fault is a physical defect, imperfection, or flaw that occurs within some hardware or software component. An error is the informational manifestation of a fault. A permanent fault exists indefinitely until corrective action is taken, whereas a transient fault appears and disappears in a short period of time. Noisy circuits in which interconnection patterns of components are trees are called formulas [230, 231].

In an error model, the effects of faults are given directly in the informational universe. For example, the basic Von Neumann model of noisy circuits [28] models transient faults in logic gates and wires as message and node computation noise that is both spatially and temporally independent. The Von Neumann model is used here. Error models of permanent faults [232, 233], of miswired circuit interconnection [14, 234], or of mismatched decoding metrics [235] have been considered elsewhere. Such permanent errors in decoding circuits may be interpreted as either changing the factor graph used for decoding or as introducing new potentials into the factor graph; the code used by the encoder and the code used by the decoder are different. An altogether different permanent error model is considered in Chapter 6.

There are several design philosophies to combat faults. Fault avoidance seeks to make physical components more reliable. Fault masking seeks to prevent faults from introducing errors. Fault tolerance is the ability of a system to continue performing its function in the presence of faults. This chapter is primarily concerned with fault tolerance, but Section 5.6 considers fault masking.

Related Work

Empirical characterizations of message-passing decoders have demonstrated that probability of error performance does not change much when messages are quantized at high resolution [67]. Even algorithms that are coarsely quantized versions of optimal belief propagation show little degradation in performance [43, 236–241]. It should be emphasized, however, that fault-free, quantized decoders differ significantly from decoders that make random errors.[2] The difference is similar to that between control systems with finite-capacity noiseless channels and control systems with noisy channels of equal capacity [245]. Seemingly the only previous work on message-passing algorithms with random errors is [246], which deals with problems in distributed inference.[3]

Noisy LDPC decoders were previously analyzed in the context of designing reliable memories from unreliable components [247, 248] (revisited in Section 5.6), using Gallager's original methods [67]. Several LPDC code analysis tools have since been developed, including simulation [249], expander graph arguments [250, 251], EXIT charts [252,253], and density evolution [43,254,255]. This work generalizes asymptotic characterizations developed by Richardson and Urbanke for noiseless decoders [43], showing that density evolution is applicable to faulty decoders. Expander graph

[2]Randomized algorithms [242] and stochastic computation [243] (used for decoding in [244]) make use of randomness to increase functionality, but the randomness is deployed in a controlled manner.

[3]If the graphical model of the code and the graph of noisy communication links in a distributed system coincide, then the distributed inference problem and the message-passing decoding problem can be made to coincide.

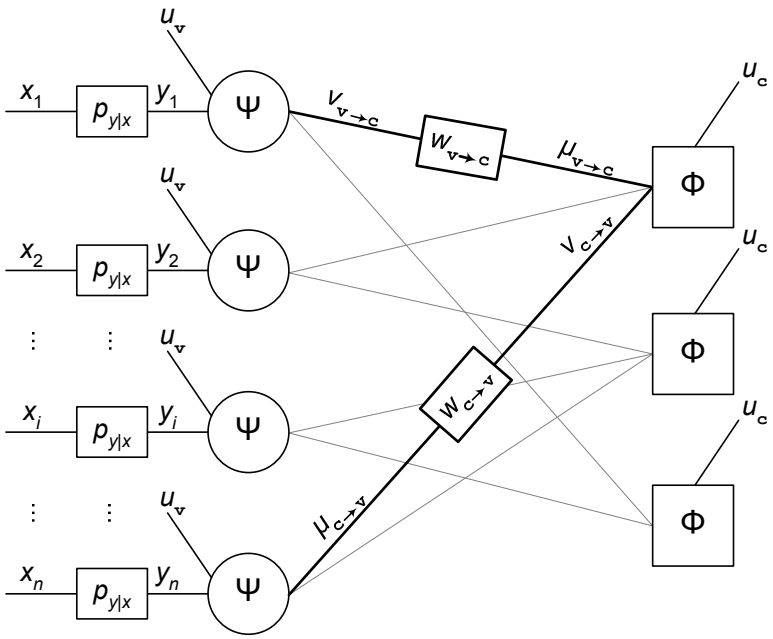

Figure 5-1. Schematic diagram of a factor graph-based implementation of a noisy decoder circuit. Only one variable-to-check message and one check-to-variable message are highlighted. Other wires, shown in gray, will also carry noisy messages.

arguments have also been extended to the case of noisy decoding in a paper [256] that appeared concurrently with the first presentation of this work [257]. Note that previous works have not even considered the possibility that Shannon reliability is achievable with noisy decoding.

■ 5.2 Message-Passing Decoders

This section establishes the basic notation of message-passing decoders for LDPC codes, which were themselves described in Section 2.3. Many of the notational conventions are depicted schematically in Figure 5-1 using a factor graph-based decoder implementation. The decoder circuit is formed from the Forney-style factor graph of the code by conversion to a Tanner-style factor graph with true variable node and check node vertices [258], followed by a direct mapping of vertices to nodes and edges to wires.

In a communication system, a codeword is selected by the encoder and is sent through the noisy channel. Channel input and output letters are $X \in \mathcal{X}$ and $Y \in \mathcal{Y}$. Since this chapter is restricted to binary codes, $\mathcal{X}$ is taken as $\{\pm 1\}$. The receiver contains a noisy message-passing decoder, which is used to process the channel output codeword to produce $\hat{X}$. Throughout this chapter, average probability of symbol error $P_e \triangleq P_e^{\text{sym}}$ is used as the performance criterion.

The message-passing decoder works in iterative stages and the iteration time is indexed by $\ell = 0, 1, \ldots$. Within the decoder, at time $\ell = 0$, each variable node has a realization of Y, y_i. A message-passing decoder exchanges messages between nodes

along wires. First each variable node sends a message to a neighboring check node over a noisy messaging wire. Generically, sent messages are denoted as $\nu_{\mathtt{v}\to\mathtt{c}}$, message wire noise realizations as $w_{\mathtt{v}\to\mathtt{c}}$,[4] and received messages as $\mu_{\mathtt{v}\to\mathtt{c}}$: assume without loss of generality that $\nu_{\mathtt{v}\to\mathtt{c}}$, $w_{\mathtt{v}\to\mathtt{c}}$, and $\mu_{\mathtt{v}\to\mathtt{c}}$ are drawn from a common messaging alphabet $\mathcal{M}$.

Each check node processes received messages and sends back a message to each neighboring variable node over a noisy message wire. The noisiness of the check node processing is generically denoted by an input random variable $U_{\mathtt{c}} \in \mathcal{U}$. The check node computation is denoted $\Phi^{(\ell)} : \mathcal{M}^{d_{\mathtt{c}}-1} \times \mathcal{U} \mapsto \mathcal{M}$. The notations $\nu_{\mathtt{c}\to\mathtt{v}}$, $\mu_{\mathtt{c}\to\mathtt{v}}$, and $w_{\mathtt{c}\to\mathtt{v}}$ are used for signaling from check node to variable node; again without loss of generality assume that $\nu_{\mathtt{c}\to\mathtt{v}}, w_{\mathtt{c}\to\mathtt{v}}, \mu_{\mathtt{c}\to\mathtt{v}} \in \mathcal{M}$.

Each variable node now processes its y_i and the messages it receives to produce new messages. The new messages are produced through possibly noisy processing, where the noise input is generically denoted $U_{\mathtt{v}} \in \mathcal{U}$. The variable node computation is denoted $\Psi^{(\ell)} : \mathcal{Y} \times \mathcal{M}^{d_{\mathtt{v}}-1} \times \mathcal{U} \mapsto \mathcal{M}$. Local computations and message-passing continue iteratively.

Message passing leads to computation graphs that describe how signals flow through a decoding system [49, Section 3.7.1]. Computation graphs have decoding neighborhoods, which involve nodes/wires that have communicated with one another. For a given node $\dot{n}$, its depth d neighborhood $\mathcal{N}_{\dot{n}}^{d}$ is the induced subgraph consisting of all nodes reached and wires traversed by paths of length at most d starting from $\dot{n}$ (including $\dot{n}$). The directed neighborhood of depth d of a wire $\mathtt{v} \to \mathtt{c}$, denoted by $\mathcal{N}_{\mathtt{v}\to\mathtt{c}}^{d}$, is defined as the induced subgraph containing all wires and nodes on paths starting from the same place as $\mathtt{v} \to \mathtt{c}$ but different from $\mathtt{v} \to \mathtt{c}$. Equivalently for a wire $\mathtt{c} \to \mathtt{v}$, $\mathcal{N}_{\mathtt{c}\to\mathtt{v}}^{d}$ is the induced subgraph containing all wires and nodes on paths starting from the same place as $\mathtt{c} \to \mathtt{v}$ but different from $\mathtt{c} \to \mathtt{v}$. If the induced subgraph (corresponding to a neighborhood) is a tree then the neighborhood is tree-like, otherwise it is not tree-like. The neighborhood is tree-like if and only if all involved nodes are distinct. Messages are statistically independent in tree-like neighborhoods.

Note that only extrinsic information is used in node computations. Also note that in the sequel, all decoder noises ($U_{\mathtt{c}}$, $U_{\mathtt{v}}$, $W_{\mathtt{v}\to\mathtt{c}}$, and $W_{\mathtt{c}\to\mathtt{v}}$) will be assumed to be independent of each other, as in the Von Neumann error model of faulty computing.

■ 5.3 Density Evolution Concentration Results

Considering the great successes achieved by analyzing the noiseless decoder performance of ensembles of codes [43, 49, 254] rather than of particular codes [67], the same approach is pursued for noisy decoders. The first mathematical contribution of this chapter is to extend the method of analysis promulgated in [43] to the case of decoders with random noise.

Several facts that simplify performance analysis are proven. First, under certain

[4]Note that this chapter is concerned purely with symbol decoding rather than message decoding. Hence, the notation w for wire noise should not be confused with the notation w for messages in other chapters.

symmetry conditions with wide applicability, the probability of error does not depend on which codeword is transmitted. Second, the individual performances of codes in an ensemble are, with high probability, the same as the average performance of the ensemble. Finally, this average behavior converges to the behavior of a code defined on a cycle-free graph. Performance analysis then reduces to determining average performance on an infinite tree: a noisy formula is analyzed in place of general noisy circuits.

For brevity, only regular LDPC codes are considered in this section, however the results can be generalized to irregular LDPC codes. In particular, replacing node degrees by maximum node degrees, the proofs stand *mutatis mutandis*. Similarly, only binary LDPC codes are considered; generalizations to non-binary alphabets also follow, as in [259].

5.3.1 Restriction to All-One Codeword

If certain symmetry conditions are satisfied by the system, then the probability of error conditioned on which codeword was chosen for transmission does not depend on the codeword. It is assumed throughout this section that messages in the decoder are in log-likelihood format, i.e. that the sign of the message indicates the bit estimate and the magnitude may indicate the level of confidence. Note, however, that it is not obvious that this is the best format for noisy message-passing [49, Appendix B.1] or that the symmetry conditions can be restated for messages in other formats. The several symmetry conditions are:

Definition 5.1 (Channel Symmetry)**.** *A memoryless channel is binary-input output-symmetric if it satisfies*

$$p(Y_i = y|X_i = 1) = p(Y_i = -y|X_i = -1)$$

for all channel usage times $i = 1, \ldots, n$.

Definition 5.2 (Check Node Symmetry)**.** *A check node message map is symmetric if it satisfies*

$$\Phi^{(\ell)}(b_1\mu_1, \ldots, b_{d_c-1}\mu_{d_c-1}, b_{d_c}u) = \Phi^{(\ell)}(\mu_1, \ldots, \mu_{d_c-1}, u)\left(\prod_{i=1}^{d_c} b_i\right)$$

for any ± 1 *sequence* $(b_1, \ldots, b_{d_c})$. *That is to say, the signs of the messages and the noise factor out of the map.*

Definition 5.3 (Variable Node Symmetry)**.** *A variable node message map is symmetric if it satisfies*

$$\Psi^{(0)}(-\mu_0, -u) = -\Psi^{(0)}(\mu_0, u)$$

and

$$\Psi^{(\ell)}(-\mu_0, -\mu_1, \ldots, -\mu_{d_v-1}, -u) = -\Psi^{(\ell)}(\mu_0, \mu_1, \ldots, \mu_{d_v-1}, u),$$

for $\ell \geq 1$. That is to say, the initial message from the variable node only depends on the received value and internal noise and there is sign inversion invariance for all messages.

Definition 5.4 (Message Wire Symmetry)**.** *Consider any message wire to be a mapping $\Xi : \mathcal{M} \times \mathcal{M} \to \mathcal{M}$. Then a message wire is symmetric if*

$$\mu = \Xi(\nu, u) = -\Xi(-\nu, -u),$$

where μ is any message received at a node when the message sent from the opposite node is ν and u is message wire noise.

Theorem 5.1 (Conditional Independence of Error)**.** *For a given binary linear code and a given noisy message-passing algorithm, let $P_e^{(\ell)}(\mathbf{x})$ denote the conditional probability of error after the ℓth decoding iteration, assuming that codeword $\mathbf{x}$ was sent. If the channel and the decoder satisfy the symmetry conditions given in Definitions 5.1–5.4, then $P_e^{(\ell)}(\mathbf{x})$ does not depend on $\mathbf{x}$.*

Proof. Minor modification of [43, Lemma 1] or [49, Lemma 4.92]. Appendix 5.A gives details. □

Suppose a system meets these symmetry conditions. Since probability of error is independent of the transmitted codeword and since all LDPC codes have the all-one codeword in the codebook, one may assume without loss of generality that this codeword is sent. Doing so removes the randomness associated with transmitted codeword selection.

■ 5.3.2 Concentration around Ensemble Average

The next simplification follows by seeing that the average performance of the ensemble of codes rather than the performance of a particular code may be studied, since all codes in the ensemble perform similarly. The performances of almost all LDPC codes closely match the average performance of the ensemble from which they are drawn. The average is over the instance of the code, the realization of the channel noise, and the realizations of the two forms of decoder noise. To simplify things, assume that the number of decoder iterations is fixed at some finite $\ell = \ell$ (with some reuse/abuse of notation). Let Z be the number of incorrect values held among all $d_{\mathrm{v}}n$ variable node-incident edges at the end of the ℓth iteration (for a particular code, channel noise realization, and decoder noise realization) and let $\mathrm{E}[Z]$ be the expected value of Z. By constructing a martingale through sequentially revealing all of the random elements and then using the Hoeffding-Azuma inequality, it can be shown that:

Theorem 5.2 (Concentration Around Expected Value)**.** *There exists a positive constant $\beta = \beta(d_{\mathrm{v}}, d_{\mathrm{c}}, \ell)$ such that for any $\epsilon > 0$,*

$$\Pr\left[|Z - \mathrm{E}[Z]| > n d_{\mathrm{v}} \epsilon / 2\right] \leq 2e^{-\beta \epsilon^2 n}.$$

Proof. Follows the basic ideas of the proofs of [43, Theorem 2] or [49, Theorem 4.94]. Appendix 5.B gives details. □

If the number of incorrect values Z concentrates, then so does P_e.

■ 5.3.3 Convergence to the Cycle-Free Case

The previous theorem showed that the noisy decoding algorithm behaves essentially deterministically for large n. As now shown, this ensemble average performance converges to the performance of an associated tree ensemble, which will allow the assumption of independent messages.

For a given edge whose directed neighborhood of depth 2ℓ is tree-like, let p be the expected number of incorrect messages received along this edge (after message noise) at the ℓth iteration, averaged over all graphs, inputs and decoder noise realizations of both types.

Theorem 5.3 (Convergence to Cycle-Free Case)**.** *There exists a positive constant $\gamma = \gamma(d_{\mathtt{v}}, d_{\mathtt{c}}, \ell)$ such that for any $\epsilon > 0$ and $n > 2\gamma/\epsilon$,*

$$|\,\mathrm{E}[Z] - nd_{\mathtt{v}}p| < nd_{\mathtt{v}}\epsilon/2.$$

The proof is identical to the proof of [43, Theorem 2]. The basic idea is that the computation tree created by unwrapping the code graph to a particular depth [260] almost surely has no repeated nodes.

The concentration and convergence results directly imply concentration around the average performance of a tree ensemble:

Theorem 5.4 (Concentration Around Cycle-Free Case)**.** *There exist positive constants $\beta = \beta(d_{\mathtt{v}}, d_{\mathtt{c}}, \ell)$ and $\gamma = \gamma(d_{\mathtt{v}}, d_{\mathtt{c}}, \ell)$ such that for any $\epsilon > 0$ and $n > 2\gamma/\epsilon$,*

$$\Pr\left[|Z - nd_{\mathtt{v}}p| > nd_{\mathtt{v}}\epsilon\right] \leq 2e^{-\beta\epsilon^2 n}.$$

Proof. Follows directly from Theorems 5.2 and 5.3. □

■ 5.3.4 Density Evolution

With the conditional independence and concentration results, all randomness is removed from explicit consideration and all messages are independent. The problem reduces to density evolution, the analysis of a discrete-time dynamical system [255]. The dynamical system state variable of most interest is the probability of symbol error, P_e.

Denote the probability of symbol error of a code $g \in \mathcal{C}^n$ after ℓ iterations of decoding by $P_e^{(\ell)}(g, \varepsilon, \alpha)$, where ε is a channel noise parameter (such as noise power or crossover probability) and α is a decoder noise parameter (such as logic gate error probability). Then density evolution computes

$$\lim_{n\to\infty} \mathrm{E}\left[P_e^{(\ell)}(g, \varepsilon, \alpha)\right],$$

where the expectation is over the choice of the code and the various noise realizations. The main interest is in the long-term behavior of the probability of error after performing many iterations. The long-term behavior of a generic dynamical system may be a limit cycle or a chaotic attractor, however density evolution usually converges to a stable fixed point. It should also be noted, that although monotonicity (either increasing or decreasing) of

$$\lim_{n\to\infty} \mathrm{E}\left[P_e^{(\ell)}(g,\varepsilon,\alpha)\right]$$

with respect to iteration number ℓ need not hold, it often does.

If there is a stable fixed point, the limiting performance corresponds to

$$\eta^* = \lim_{\ell\to\infty}\lim_{n\to\infty} \mathrm{E}\left[P_e^{(\ell)}(g,\varepsilon,\alpha)\right].$$

Certain sets of parameters (g,ε,α) lead to "good" performance, in the sense of small η^*, whereas other sets of parameters lead to "bad" performance with large η^*. The goal of density evolution analysis is to determine the boundary between these good and bad sets.

Though it is natural to expect the performance of an algorithm to improve as the quality of its input improves and as more resources are allocated to it, this may not be so. For many decoders, however, there is a monotonicity property that limiting behavior η^* improves as channel noise ε decreases and as decoder noise α decreases. Moreover, just as in other nonlinear estimation systems for dimensionality-expanding signals [60,261,262], there is a threshold phenomenon such that the limiting probability of error may change precipitously with the values of ε and α. The error probability properties of random codes around channel capacity under noiseless maximum likelihood decoding provide a prime example of the nonlinear estimation threshold.

In traditional coding theory, there is no parameter α, and the goal is often to determine the range of ε for which η^* is zero. The boundary is often called the decoding threshold and may be denoted $\varepsilon^*(0)$. A decoding threshold for optimal codes under optimal decoding may be computed from the rate of the code g and the capacity of the channel, given in Theorem 2.3. Since this Shannon limit threshold is for optimal codes and decoders, it is clearly an upper bound to $\varepsilon^*(0)$ for any given code and decoder. If the target error probability η^* is non-zero, then the Shannon limit threshold is derived from the η-capacity, given in Theorem 2.2.

In the case of faulty decoders, the Shannon limits also provide upper bounds on the ε-boundary for the set of (ε,α) that achieve good performance. One might hope for a Shannon theoretic characterization of the entire (ε,α)-boundary, but such results are not extant. Alternately, in the next sections, sets of (ε,α) that can achieve η^*-reliability for particular LDPC codes $g \in \mathcal{C}^n$ are characterized using the density evolution method developed in this section.

■ 5.4 Noisy Gallager A Decoder

Section 5.3 showed that density evolution equations determine the performance of almost all codes in the large block length regime. Here the density evolution equation

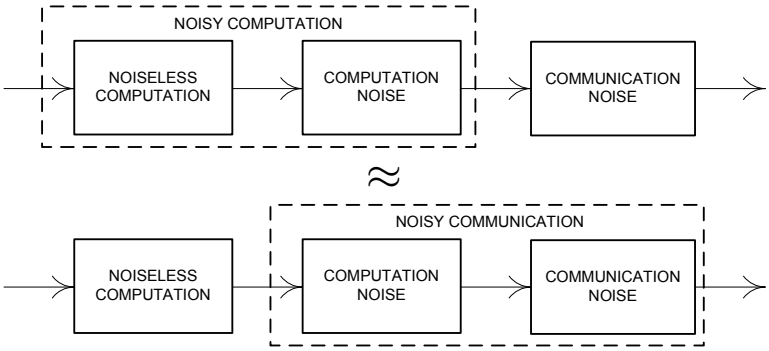

Figure 5-2. Local computation noise may be incorporated into message-passing noise without essential loss of generality.

for a simple noisy message-passing decoder, a noisy version of Gallager's decoding algorithm A [67, 263], is derived. The algorithm has message alphabet $\mathcal{M} = \{\pm 1\}$, with messages simply indicating the estimated sign of a bit.

At a check node, the outgoing message along edge $\vec{e}$ is the product of all incoming messages excluding the one incoming on $\vec{e}$, i.e. the check node map Φ is the XOR operation. At a variable node, the outgoing message is the original received code symbol unless all incoming messages give the opposite conclusion. That is,

$$\Psi = \begin{cases} -y, & \text{if } \mu_1 = \cdots = \mu_{d_\text{v}-1} = -y, \\ y, & \text{otherwise.} \end{cases}$$

Although this simple decoding algorithm cannot match the performance of belief propagation due to its restricted messaging alphabet $\mathcal{M}$, it is of interest since it is of extremely low complexity and can be analyzed analytically [263]. The density evolution equation leads to an analytic characterization of the set of (ε, α) pairs, which parameterize the noisiness of the communication system. There is no essential loss of generality by combining computation noise and message-passing noise into a single form of noise, as demonstrated schematically in Figure 5-2 and proven in [264, Lemma 3.1]. This noise combining is performed in the sequel to reduce the number of decoder noise parameters and allow a clean examination of the central phenomenon.

■ 5.4.1 Density Evolution Equation

Consider decoding the LDPC-coded output of a binary symmetric channel (BSC) with crossover probability ε. Each message in the Gallager algorithm A is passed through an independent and identical BSC with crossover probability α. The density evolution equation is developed for general irregular LDPC ensembles.

The state variable of density evolution, s_ℓ, is taken to be the probability of symbol error $P_e^{(\ell)}(\varepsilon, \alpha)$ at the variable nodes.

The original received message is in error with probability ε, thus

$$P_e^{(0)}(\varepsilon, \alpha) = s_0 = \varepsilon.$$

The initial variable-to-check message is in error with probability $(1-\varepsilon)\alpha+\varepsilon(1-\alpha)$, since it is passed through a BSC(α). For further iterations, ℓ, the probability of error, $P_e^{(\ell)}(\varepsilon,\alpha)$, is found by induction. Assume $P_e^{(i)}(\varepsilon,\alpha)=s_i$ for $0\le i\le\ell$. Now consider the error probability of a check-to-variable message in the $(\ell+1)$th iteration. A check-to-variable message emitted by a check node of degree d_{c} along a particular edge is the product of all the $(d_{\mathsf{c}}-1)$ incoming messages along all other edges. By assumption, each such message is in error with probability s_ℓ and all messages are independent. These messages are passed through BSC(α) before being received, so the probability of being received in error is

$$s_\ell(1-\alpha)+(1-s_\ell)\alpha=\alpha+s_\ell-2\alpha s_\ell.$$

Due to the XOR operation, the outgoing message will be in error if an odd number of these received messages are in error. The probability of this event, averaged over the degree distribution yields the probability

$$\frac{1-\rho\left[1-2(\alpha+s_\ell-2\alpha s_\ell)\right]}{2}.$$

Now consider $P_e^{(\ell+1)}(\varepsilon,\alpha)$, the error probability at the variable node in the $(\ell+1)$th iteration. Consider an edge which is connected to a variable node of degree d_{v}. The outgoing variable-to-check message along this edge is in error in the $(\ell+1)$th iteration if the original received value is in error and not all incoming messages are received correctly or if the originally received value is correct but all incoming messages are in error. The first event has probability

$$\varepsilon\left(1-\left[1-(1-\alpha)\left(\frac{1-\rho\left[1-2(\alpha+s_\ell-2\alpha s_\ell)\right]}{2}\right)\right.\right.$$
$$\left.\left.-\alpha\left(\frac{1+\rho\left[1-2(\alpha+s_\ell-2\alpha s_\ell)\right]}{2}\right)\right]^{d_{\mathsf{v}}-1}\right).$$

The second event has probability

$$(1-\varepsilon)\left(\left[(1-\alpha)\left(\frac{1-\rho\left[1-2(\alpha+s_\ell-2\alpha s_\ell)\right]}{2}\right)\right.\right.$$
$$\left.\left.+\alpha\left(\frac{1+\rho\left[1-2(\alpha+s_\ell-2\alpha s_\ell)\right]}{2}\right)\right]^{d_{\mathsf{v}}-1}\right).$$

Averaging over the degree distribution and adding the two terms together yields the density evolution equation in recursive form:

$$s_{\ell+1}=\varepsilon-\varepsilon q_\alpha^+(s_\ell)+(1-\varepsilon)q_\alpha^-(s_\ell). \tag{5.1}$$

The expressions

$$q_\alpha^+(\check{s}) = \lambda\left[\frac{1+\rho(\omega_\alpha(\check{s}))-2\alpha\rho(\omega_\alpha(\check{s}))}{2}\right],$$

$$q_\alpha^-(\check{s}) = \lambda\left[\frac{1-\rho(\omega_\alpha(\check{s}))+2\alpha\rho(\omega_\alpha(\check{s}))}{2}\right],$$

and $\omega_\alpha(\check{s}) = (2\alpha-1)(2\check{s}-1)$ are used to define the density evolution recursion.

■ 5.4.2 Performance Evaluation

With the density evolution equation established, the performance of the coding-decoding system with particular values of quality parameters ε and α may be determined. Taking the symbol error probability as the state variable, stable fixed points of the deterministic, discrete-time, dynamical system are to be found. Usually one would want the probability of error to converge to zero, but since this might not be possible, a weaker performance criterion may be needed. To start, consider partially noiseless cases.

Noisy Channel, Noiseless Decoder

For the noiseless decoder case, i.e. $\alpha = 0$, it has been known that there are thresholds on ε, below which the probability of error goes to zero as ℓ increases, and above which the probability of error goes to some large value. These can be found analytically for the Gallager A algorithm [263].

Noiseless Channel, Noisy Decoder

For the noisy Gallager A system under consideration, the probability of error does not go to zero as ℓ goes to infinity for any $\alpha > 0$. This can be seen by considering the case of the perfect original channel, $\varepsilon = 0$, and any $\alpha > 0$. The density evolution equation reduces to

$$s_{\ell+1} = q_\alpha^-(s_\ell), \tag{5.2}$$

with $s_0 = 0$. The recursion does not have a fixed point at zero, and since error probability is bounded below by zero, it must increase. The derivative is

$$\frac{\partial}{\partial s} q_\alpha^-(s) = \lambda'\left[\frac{1-\rho(\omega_\alpha(s))+2\alpha\rho(\omega_\alpha(s))}{2}\right]\rho'(\omega_\alpha(s))(2\alpha-1)^2,$$

which is greater than zero for $0 \le s \le \frac{1}{2}$ and $0 \le \alpha \le \frac{1}{2}$; thus the error evolution forms a monotonically increasing sequence. Since the sequence is monotone increasing starting from zero, and there is no fixed point at zero, it follows that this converges to a real solution of $s = q_\alpha^-(s)$. In fact it converges to the smallest real solution of $s = q_\alpha^-(s)$ since this smallest fixed point cannot be jumped, due to monotonicity. In particular, consider a fixed point z such that $s_\ell(s_0) \le z$ for some $\ell \ge 0$. Then $s_{\ell+1}(s_0) = q_\alpha^-(s_\ell(s_0)) \le q_\alpha^-(z) = z$ due to the monotonicity property of q_α^-. This implies that the final value $s_\infty \le z$.

The same phenomenon must also happen if the starting s_0 is positive, however the value to which the density evolution converges is a non-zero fixed point solution of the original equation (5.1), not of (5.2), and is a function of both α and ε. Intuitively, for somewhat large initial values of ε, the noisy decoder decreases the probability of error in the first few iterations, just like the noiseless one, but when the error probability becomes close to the internal decoder error, the probability of error settles at that level. This is summarized in the following proposition.

Proposition 5.1. *For any LDPC ensemble decoded using the noisy Gallager A system defined in Section 5.4, final error probability $\eta^* > 0$ when decoder noise level $\alpha > 0$ for any channel noise level ε.* $\square$

The fact that probability of error cannot asymptotically be driven to zero with the noisy Gallager decoder is expected yet is seemingly displeasing. In a practical scenario, however, the ability to drive P_e to a very small number is also desirable. As such, a performance objective of achieving P_e less than η is defined and the worst channel (ordered by ε) for which a decoder with noise level α can achieve that objective is determined. The channel parameter

$$\varepsilon^*(\eta, \alpha) = \sup\{\varepsilon \in [0, \tfrac{1}{2}] \mid \lim_{\ell \to \infty} P_e^{(\ell)}(g, \varepsilon, \alpha) < \eta\}$$

is called the threshold. For a large interval of η values, there is a single threshold value below which η-reliable communication is possible and above which it is not. Alternatively, one can determine the probability of error to which a system with particular α and ε can be driven, $\eta^*(\alpha, \varepsilon) = \lim_{\ell \to \infty} P_e^{(\ell)}$, and see whether this value is small.

In order to find the threshold in the case of $\alpha > 0$ and $\varepsilon > 0$, the real fixed point solutions of density evolution recursion (5.1) need to be found. The real solutions of the polynomial equation in s,

$$\varepsilon - \varepsilon q_\alpha^+(s) + (1 - \varepsilon) q_\alpha^-(s) - s = 0$$

are denoted $0 < r_1(\alpha, \varepsilon) \leq r_2(\alpha, \varepsilon) \leq r_3(\alpha, \varepsilon) \leq \cdots$.[5] The final probability of error η^* is determined by the r_i, since these are fixed points of the recursion (5.1).

The real solutions of the polynomial equation in s,

$$\frac{s - q_\alpha^-(s)}{1 - q_\alpha^+(s) - q_\alpha^-(s)} - s = 0, \tag{5.3}$$

are denoted $0 < \tau_1(\alpha) \leq \tau_2(\alpha) \leq \cdots$.[5] The threshold ε^* as well as the region in the $\alpha-\varepsilon$ plane where the decoder improves performance over no decoding are determined by the τ_i, since (5.3) is obtained by solving recursion (5.1) for ε and setting equal to zero. For particular ensembles of LDPC codes, these values can be computed analytically.

[5] The number of real solutions can be determined through Descartes' rule of signs or a similar tool [265].

For these particular ensembles, it can be determined whether the fixed points are stable or unstable. Moreover, various monotonicity results can be established to show that fixed points cannot be jumped.

Analytical expressions for the $r_i(\alpha, \varepsilon)$ and $\tau_i(\alpha)$ are determined for the (3,6) regular LDPC code by solving the appropriate polynomial equations and numerical evaluations of the r_i expressions are shown as thin lines in Figure 5-3 as functions of ε for fixed α. The point where $r_1(\alpha, \varepsilon) = \varepsilon$ is $\tau_1(\alpha)$ and the point where $r_2(\alpha, \varepsilon) = \varepsilon$ is $\tau_2(\alpha)$. In Figure 5-3, these are points where the thin lines cross.

By analyzing the dynamical system equation (5.1) for the (3,6) code in detail, it can be shown that $r_1(\alpha, \varepsilon)$ and $r_3(\alpha, \varepsilon)$ are stable fixed points of density evolution. Contrarily, $r_2(\alpha, \varepsilon)$ is an unstable fixed point, which determines the boundary between the regions of attraction for the two stable fixed points. Since $r_1(\alpha, \varepsilon)$ and $r_3(\alpha, \varepsilon)$ are stable fixed points, the final error probability η^* will take on one of these two values, depending on the starting point of the recursion, ε. The thick line in Figure 5-3 shows the final error probability η^* as a function of initial error probability ε. One may note that $\eta^* = r_1$ is the desirable small error probability, whereas $\eta^* = r_3$ is the undesirable large error probability and that τ_2 delimits these two regimes.

The $\tau(\alpha)$ points determine when it is beneficial to use the decoder, in the sense that $\eta^* < \varepsilon$. By varying α (as if in a sequence of plots like Figure 5-3), an $\alpha - \varepsilon$ region where the decoder is beneficial is demarcated; this is shown in Figure 5-4. The function $\tau_2(\alpha)$ is the η-reliability decoding threshold for large ranges of η.

Notice that the previously known special case, the decoding threshold of the noiseless decoder, can be recovered from these results. The decoding threshold for the noiseless decoder is denoted ε^*_{BRU} and is equal to the following expression [263].

$$\varepsilon^*_{BRU} = \frac{1 - \sqrt{\sigma}}{2},$$

where

$$\sigma = -\frac{1}{4} + \frac{\sqrt{-\frac{5}{12} - b}}{2} + \frac{\sqrt{-\frac{5}{6} + \frac{11}{4\sqrt{-5/12 - b}}}}{2}$$

and

$$b = \frac{8}{3}\left(\frac{2}{83 + 3\sqrt{993}}\right)^{\frac{1}{3}} - \frac{1}{3}\left(\frac{83 + 3\sqrt{993}}{2}\right)^{\frac{1}{3}}.$$

This value is recovered from noisy decoder results by noting that $\eta^*(\alpha = 0, \varepsilon) = 0$ for $\varepsilon \in [0, \varepsilon^*_{BRU}]$, which are the ordinate intercepts of the region in Figure 5-4.

To provide a better sense of the performance of the noisy Gallager A algorithm, Table 5.1 lists some values of α, ε, and η^* (numerical evaluations are listed and an example of an analytical expression is given in Appendix 5.C). As can be seen from these results, particularly from the τ_2 curve in Figure 5-4, the error probability performance of the system degrades gracefully as noise is added to the decoder.

Returning to threshold characterization, an analytical expression for the threshold

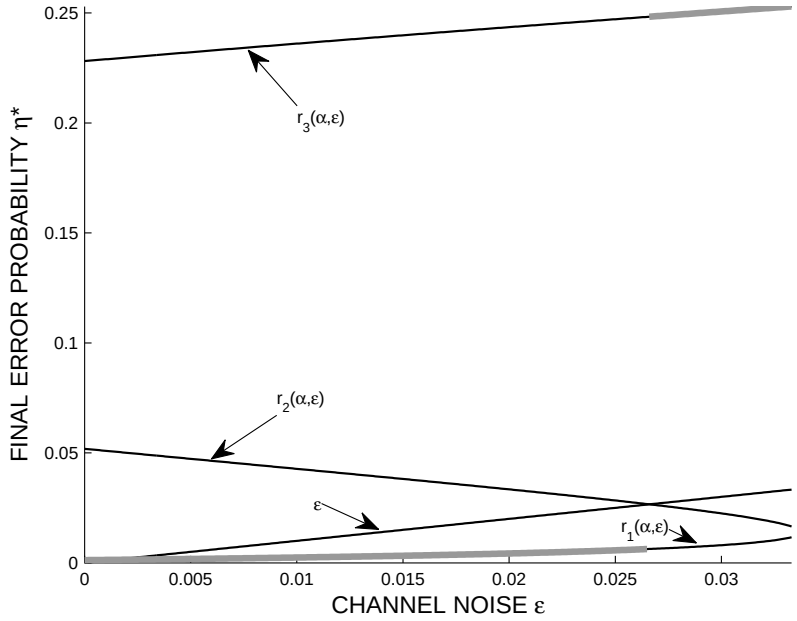

Figure 5-3. Thick line shows final error probability, η^*, after decoding a $\mathcal{C}^\infty(3,6)$ code with the noisy Gallager A algorithm, $\alpha = 0.005$. This is determined by the fixed points of density evolution, $r_i(\alpha, \varepsilon)$, shown with thin lines.

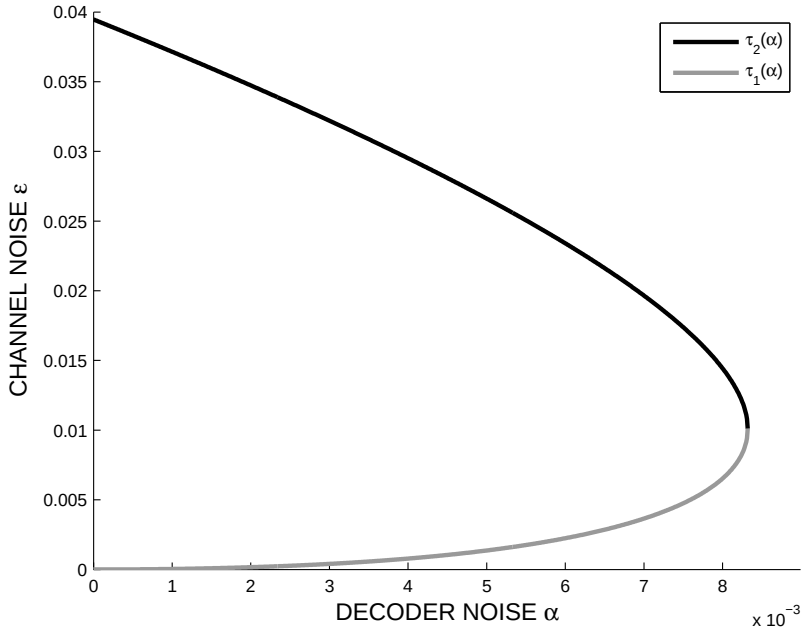

Figure 5-4. Decoding a $\mathcal{C}^\infty(3,6)$ code with the noisy Gallager A algorithm. Region where it is beneficial to use decoder is below τ_2 and above τ_1.

α	$\varepsilon^*(0.1,\alpha)$	$\eta^*(\alpha,\varepsilon^*)$	$\eta^*(\alpha,0.01)$
0	0.0394636562	0	0
1×10^{-10}	0.0394636560	7.8228×10^{-11}	1.3333×10^{-11}
1×10^{-8}	0.0394636335	7.8228×10^{-9}	1.3333×10^{-9}
1×10^{-6}	0.0394613836	7.8234×10^{-7}	1.3338×10^{-7}
1×10^{-4}	0.0392359948	7.8866×10^{-5}	1.3812×10^{-5}
3×10^{-4}	0.0387781564	2.4050×10^{-4}	4.4357×10^{-5}
1×10^{-3}	0.0371477336	8.4989×10^{-4}	1.8392×10^{-4}
3×10^{-3}	0.0321984070	3.0536×10^{-3}	9.2572×10^{-4}
5×10^{-3}	0.0266099758	6.3032×10^{-3}	2.4230×10^{-3}

Table 5.1. Performance of Noisy Gallager A algorithm for (3,6) code

within the region to use decoder is:

$$\varepsilon^*(\eta,\alpha)=\frac{\eta-q_\alpha^-(\eta)}{1-q_\alpha^+(\eta)-q_\alpha^-(\eta)},$$

which is the solution to the polynomial equation in $\check{\epsilon}$,

$$\check{\epsilon}-\check{\epsilon}q_\alpha^+(\eta)+(1-\check{\epsilon})q_\alpha^-(\eta)-\eta=0.$$

The threshold is drawn for several values of η in Figure 5-5. A threshold line determines the equivalence of channel noise and decoder noise with respect to final probability of error. If for example, the binary symmetric channels in the system are a result of hard-detected AWGN channels, such a line may be used to derive the equivalent channel noise power for decoder noise power or vice versa. Threshold lines therefore provide guidelines for power allocation in communication systems.

■ 5.4.3 Code Optimization

At this point, the symbol error performance of a system has simply been measured; no attempt has been made to optimize a code for a particular decoder and set of parameters. For fault-free decoding, it has been demonstrated that irregular code ensembles can perform much better than regular code ensembles like the (3,6) LDPC considered above [263, 266]. One might hope for similar improvements when LDPC code design takes decoder noise into account. The space of system parameters to be considered for noisy decoders is much larger than for noiseless decoders.

As a first step, consider the ensemble of rate 1/2 LDPC codes that were optimized by Bazzi et al. for the fault-free Gallager A decoding algorithm [263]. The left degree distribution is

$$\lambda(\zeta)=a\zeta^2+(1-a)\zeta^3$$

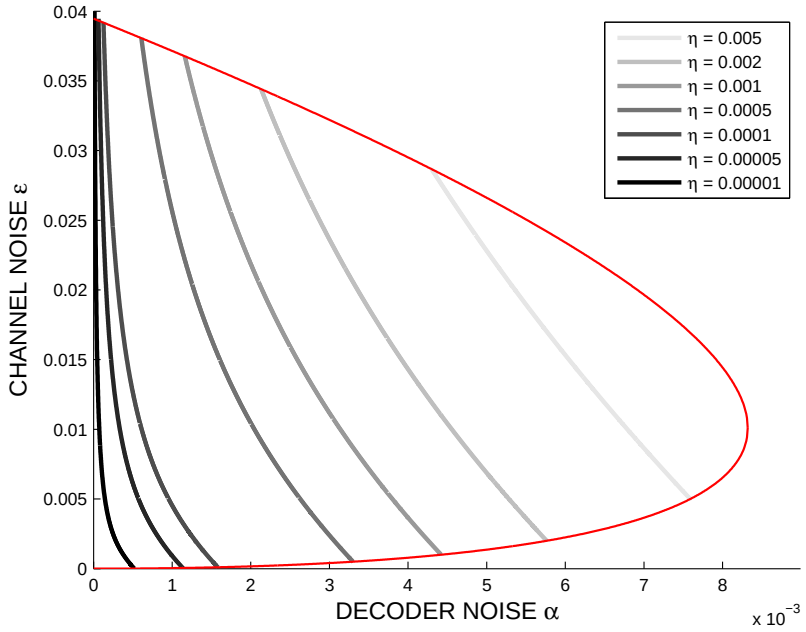

Figure 5-5. η-thresholds (gray lines) for decoding a $\mathcal{C}^{\infty}(3,6)$ code with the noisy Gallager A algorithm within the region to use decoder (delimited with red line).

and the right degree distribution is

$$\rho(\zeta) = \frac{7a}{3}\zeta^6 + \frac{3-7a}{3}\zeta^7,$$

where the optimal a is specified analytically. Numerically, $a_{\text{opt}} = 0.1115\ldots$. Measuring the performance of this code with the noisy Gallager A decoder yields the region to use decoder shown in Figure 5-6; the region to use decoder for the (3,6) code is shown for comparison. By essentially any criterion of performance, this optimized code is better than the (3,6) code.

Are there other codes that can perform better on the faulty decoder than the code optimized for the fault-free decoder? To see whether this is possible, arbitrarily restrict to the family of ensembles that were found to contain the optimal degree distribution for the fault-free decoder and take $a = 1/10$. Also let $\alpha = 1/500$ be fixed. The numerical value of the threshold $\varepsilon^*_{1/10}(1/10, \alpha) = 0.048239$, whereas the numerical value of the threshold $\varepsilon^*_{a_{\text{opt}}}(1/10, \alpha) = 0.047857$. In this sense, the $a = 1/10$ code is better than the $a = a_{\text{opt}}$ code. In fact, as seen in Figure 5-6, the region to use decoder for this $a = 1/10$ code contains the region to use decoder for the a_{opt} code.

On the other hand, the final error probability when operating at threshold for the $a = 1/10$ code $\eta^*_{1/10}(\alpha, \varepsilon^*_{1/10}(1/10, \alpha)) = 0.01869$, whereas the final error probability when operating at threshold for the $a = a_{\text{opt}}$ code is $\eta^*_{a_{\text{opt}}}(\alpha, \varepsilon^*_{a_{\text{opt}}}(1/10, \alpha)) = 0.01766$. So in this sense, the $a = a_{\text{opt}}$ code is better than the $a = 1/10$ code.

This phenomenon arises due to the fact that highly optimized ensembles usually lead to more simultaneous critical points including fixed points, see e.g. [49, Example

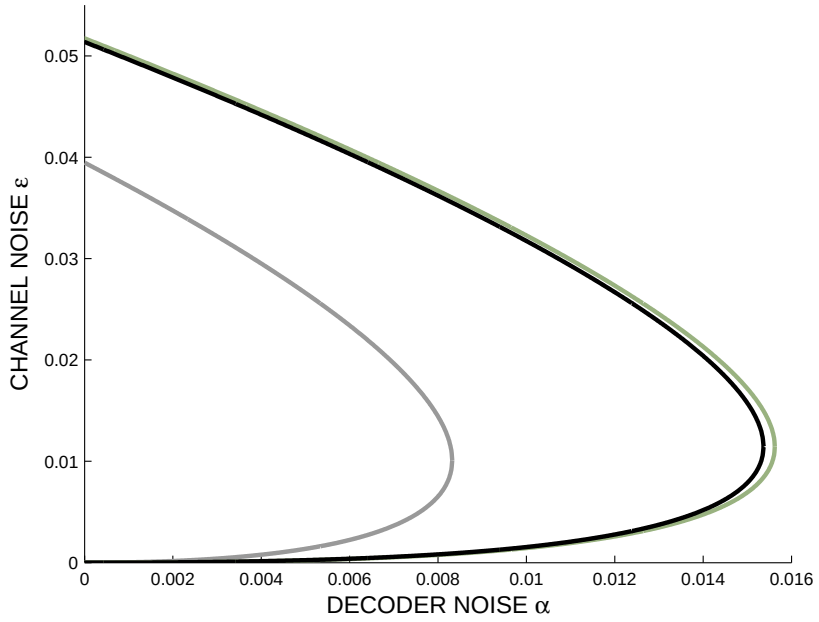

Figure 5-6. Region to use decoder for Bazzi et al.'s optimized rate 1/2 LDPC code with noisy Gallager A decoding (black) is contained within the region to use decoder for a rate 1/2 LDPC code in Bazzi et al.'s optimal family of codes with $a = 1/10$ (green) and contains the region to use decoder for the $\mathcal{C}^{\infty}(3, 6)$ code (gray).

3.64]. Since there are potentially more stable fixed points, they can capture the decoding process at greater values of final error probability.

If both threshold and final symbol error probability are performance criteria, there is no total order on codes and therefore there may be no notion of an optimal code.

■ 5.5 Noisy Gaussian Decoder

One might want to analyze a noisy version of the belief propagation decoder applied to the output of a continuous-alphabet channel, but density evolution for belief propagation is difficult to analyze even in the noiseless decoder case. This section considers a decoder inspired by [267] that has one-dimensional state variables and messages, rather than infinite-dimensional ones as for belief propagation. In reference to [267], the decoder is called a noisy Gaussian approximation decoder, but it should be thought of as physically implemented rather than simply as an approximation to aid performance analysis. The specific node computations carried out by the decoder, Φ and Ψ, as well as the noise model are defined below.

Section 5.4 had considered decoding the output of a BSC with a decoder that was constructed with BSC components and Proposition 5.1 had shown that probability of symbol error could never be driven to zero. Here, the probability of symbol error does in fact go to zero.

Consider a binary input AWGN channel with variance ε^2. The output is decoded

using a noisy Gaussian approximation decoder. For simplicity, only regular LDPC codes are considered. The messages that are passed in this decoder are real-valued, $\mathcal{M} = \mathbb{R} \cup \{\pm\infty\}$, and are in log-likelihood format. A positive-valued message indicates belief that a symbol is $+1$ whereas a negative-valued message indicates belief that a symbol is -1. A message of magnitude 0 indicates complete uncertainty whereas a message of infinite magnitude indicates complete confidence in a symbol value.

Letting $\nu(y)$ denote the log-likelihood ratios computed from the channel output symbols, the variable-to-check messages in the zeroth iteration are

$$\nu_{\mathtt{v}\to\mathtt{c}} = \nu(y) = \log \frac{p(y|x=1)}{p(y|x=-1)}.$$

The check node takes the received versions of these messages, $\mu_{\mathtt{v}\to\mathtt{c}}$, as input. The node implements a mapping Φ whose output, $\nu_{\mathtt{c}\to\mathtt{v}}$, satisfies:

$$\operatorname{etanh}(\nu_{\mathtt{c}\to\mathtt{v}}) = \prod_{i=1}^{d_{\mathtt{c}}-1} \operatorname{etanh}(\mu_{\mathtt{v}\to\mathtt{c}_i}),$$

where the product is taken over messages on all incoming edges except the one on which the message will be outgoing, and

$$\operatorname{etanh}(\check{v}) = \frac{1}{\sqrt{4\pi\check{v}}} \int_{\mathbb{R}} \tanh\frac{v}{2} e^{-\frac{(v-\check{v})^2}{4v}} dv.$$

The check node mapping is motivated by Gaussian likelihood computations. For the sequel, it is useful to define a slightly different function

$$\phi(\check{v}) = \begin{cases} 1 - \operatorname{etanh}(\check{v}), & \check{v} > 0 \\ 1, & \check{v} = 0 \end{cases}$$

which can be approximated as

$$\phi(\check{v}) \approx e^{a\check{v}^c + b},$$

with $a = -0.4527$, $b = 0.0218$, $c = 0.86$ [267].

For iterations $\ell \geq 1$, the variable node takes the received versions of the $\mathtt{c} \to \mathtt{v}$ messages, $\mu_{\mathtt{c}\to\mathtt{v}}$, as inputs. The mapping Ψ yields output $\nu_{\mathtt{v}\to\mathtt{c}}$ given by

$$\nu_{\mathtt{v}\to\mathtt{c}} = \nu(y) + \sum_{i=1}^{d_{\mathtt{v}}-1} \mu_{\mathtt{c}\to\mathtt{v}_i},$$

where the sum is taken over received messages from the neighboring check nodes except the one to which this message is outgoing. Again, the operation of the variable node is motivated by Gaussian likelihood computations.

As in Section 5.4, local computation noise is combined into message-passing noise (Figure 5-2). To model quantization [229] or random phenomena, consider each message passed in the decoder to be corrupted by signal-independent additive white noise

which is bounded as $-\alpha/2 \leq w \leq \alpha/2$. This class of noise models includes uniform noise and truncated Gaussian noise, among others. If the noise is symmetric, then Theorem 5.1 applies. Following the Von Neumann error model, each noise realization w is assumed to be independent.

5.5.1 Density Evolution Equation

The definition of the computation rules and the noise model may be used to derive the density evolution equation. The one-dimensional state variable chosen to be tracked is s, the approximate log-likelihood value at a variable node. If the all-one codeword was transmitted, then the value s going to $+\infty$ is equivalent to P_e going to 0.

To bound decoding performance under any noise model in the class of additive bounded noise, consider (non-stochastic) worst-case noise. Assuming that the all-one codeword was sent, all messages should be as positive as possible to move towards the correct decoded codeword (log-likelihoods of $+\infty$ indicate perfect confidence in a symbol being 1). Consequently, the worst bounded noise that may be imposed is to subtract $\alpha/2$ from all messages that are passed; this requires knowledge of the transmitted codeword being all-one. If another codeword is transmitted, then certain messages would have $\alpha/2$ added instead of subtracted.

Such a worst-case noise model does not meet the last condition (Definition 5.4) of Theorem 5.1, but transmission of the all-one codeword is assumed nonetheless. An adversary with knowledge of the transmitted codeword imposing worst-case noise on the decoder would still yield a probability of symbol error conditioned on the transmitted codeword that does not depend on the identity of the codeword if the symmetry conditions specified in Definitions 5.1–5.3 hold. This follows since the worst-case degradation causes the same loss in performance for any transmitted codeword.

Note that the adversary is restricted to selecting each noise realization independently. More complicated and devious error patterns in space or in time are not possible in the Von Neumann error model. Moreover, the performance criterion is probability of symbol error rather than probability of message error, so complicated error patterns would provide no great benefit.

Since the noise is conditionally deterministic given the transmitted codeword, derivation of the density evolution equation is much simplified. An induction argument is used, and the base case is

$$s_0 = \nu(y) = \tfrac{2}{\varepsilon^2},$$

where ε^2 is the channel noise power. The smaller the channel noise power, the closer to $+\infty$ the state variable starts. This follows from the log-likelihood computation for an AWGN communication channel with input alphabet $\mathcal{X} = \{\pm 1\}$.

The inductive assumption in the induction argument is $s_{\ell-1}$. This message is communicated over message-passing noise to get

$$\mu^{(\ell)}_{\mathsf{v}\to\mathsf{c}} = s_{\ell-1} - \tfrac{\alpha}{2}.$$

Next the check node computation is made:

$$\nu_{\mathsf{c}\to\mathsf{v}}^{(\ell)} = \phi^{-1}\left(1 - [1 - \phi(s_{\ell-1} - \tfrac{\alpha}{2})]^{d_{\mathsf{c}}-1}\right).$$

By the inductive assumption, all messages will be equivalent; that is why the product is a $(d_{\mathsf{c}} - 1)$-fold product of the same quantity. This value is communicated over message-passing noise to get

$$\mu_{\mathsf{c}\to\mathsf{v}}^{(\ell)} = \phi^{-1}\left(1 - [1 - \phi(s_{\ell-1} - \tfrac{\alpha}{2})]^{d_{\mathsf{c}}-1}\right) - \tfrac{\alpha}{2}.$$

Finally the variable-node computation yields

$$\nu(y) + (d_{\mathsf{v}} - 1)\left\{\phi^{-1}\left(1 - [1 - \phi(s_{\ell-1} - \tfrac{\alpha}{2})]^{d_{\mathsf{c}}-1}\right) - \tfrac{\alpha}{2}\right\}.$$

Again, all messages will be equivalent so the sum is a $(d_{\mathsf{v}} - 1)$-fold sum of the same quantity. Thus the density evolution equation is

$$s_\ell = \tfrac{2}{\varepsilon^2} - \tfrac{(d_{\mathsf{v}}-1)\alpha}{2} + (d_{\mathsf{v}} - 1)\left\{\phi^{-1}\left(1 - [1 - \phi(s_{\ell-1} - \tfrac{\alpha}{2})]^{d_{\mathsf{c}}-1}\right)\right\}. \tag{5.4}$$

■ 5.5.2 Performance Evaluation

One might wonder whether there are sets of noise parameters $\alpha > 0$ and $\varepsilon > 0$ such that $s_\ell \to +\infty$. Indeed there are, and there is a threshold phenomenon, just like Chung et al. showed for $\alpha = 0$ [267].

Proposition 5.2. *For LDPC ensembles decoded using the noisy Gaussian approximation system defined in Section 5.5, there exit positive decoding threshold values $\varepsilon^*(\alpha)$ such that final symbol error probability $\eta^* = 0$ for all binary-input AWGN channels with noise levels $\varepsilon < \varepsilon^*(\alpha)$.*

Proof. Substituting $s = +\infty$ into (5.4) demonstrates that it is a stable fixed point. It may further be verified that the dynamical system proceeds toward that fixed point if $\varepsilon < \varepsilon^*(\alpha)$. □

Unlike Section 5.4 where the $\varepsilon^*(\eta, \alpha)$ thresholds could be evaluated analytically, only numerical evaluations of these $\varepsilon^*(\alpha)$ thresholds are possible. This is shown in Figure 5-7 for the (3,6) regular LDPC ensemble. As can be observed, the threshold decreases smoothly as the decoder noise level increases.

The basic reason for the disparity between Propositions 5.1 and 5.2 is that here, the noise is bounded whereas the messages are unbounded. Thus once the messages grow large, the noise has essentially no effect. To use a term from [268], once the decoder reaches the *breakout value*, noise cannot stop the decoder from achieving Shannon reliability.

Perhaps a peak amplitude constraint on messages would provide a more realistic computation model, but the equivalent of Proposition 5.2 may not hold. Quantified data processing inequalities may provide insight into what forms of noise and message constraints are truly limiting [230, 231].

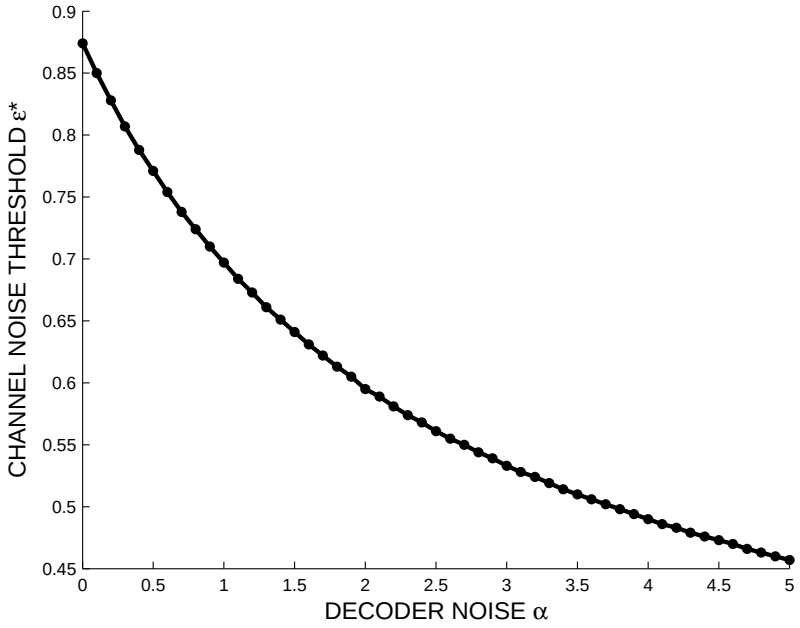

Figure 5-7. Thresholds for decoding a $\mathcal{C}^{\infty}(3,6)$ code with the noisy Gaussian approximation algorithm. Notice that the ordinate intercept is $\varepsilon^{*}_{CRU} = 0.8747$, [267].

■ 5.6 Constructing Reliable Memories from Unreliable Components

Complexity and reliability are primary limitations on practical decoding. By considering the design of fault masking techniques for memory systems, a communication problem beyond Figure 1-2, both complexity and reliability may be explicitly constrained. Indeed, the problem of constructing reliable information storage devices from unreliable components is central to fault-tolerant computing, and determining the information storage capacity of such devices is a long-standing open problem [269]. This problem is related to problems in distributed information storage [270] and is intimately tied to the performance of codes under faulty decoding. The analysis techniques developed thus far may be used directly.

In particular, one may construct a memory architecture with noisy registers and a noisy LDPC correcting network, as depicted in Figure 5-8. At each time step, the correcting network decodes the register contents and restores them. The correcting network prevents the codeword stored in the registers from wandering too far away. Taylor and others have shown that there exist non-zero levels of component noisiness such that the LDPC-based construction achieves non-zero storage capacity [247, 248, 256]. Results as in Section 5.4 may be used to precisely characterize storage capacity.

Before proceeding with an achievability result, requisite definitions and the problem statement are given [247].

Definition 5.5. *An* elementary operation *is any Boolean function of two binary operands.*

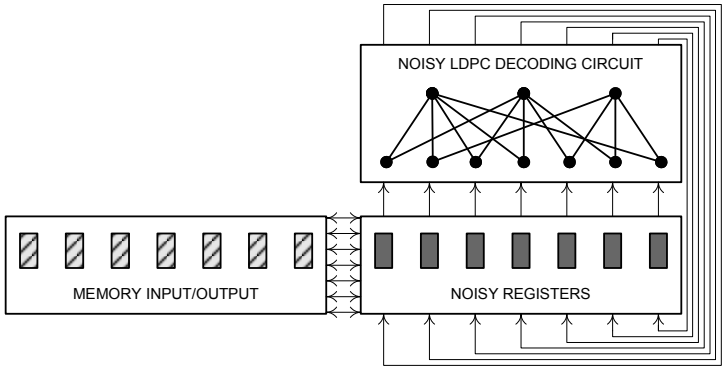

Figure 5-8. Architecture of memory system with noisy registers and noisy LDPC correcting network.

Definition 5.6. *A system is considered to be constructed from* components*, which are devices that either perform one elementary operation or store one bit.*

Definition 5.7. *The* complexity χ *of a system is the number of components within the system.*

Definition 5.8. *A memory system that stores k information bits is said to have an* information storage capability *of k.*

Definition 5.9. *Consider a sequence of memories $\{M_i\}$, ordered according to their information storage capability i (bits). The sequence $\{M_i\}$ is* stable *if it satisfies the following:*

1. *For any k, M_k must have 2^k allowed inputs denoted $\{I_{k_i}\}$, $1 \leq i \leq 2^k$.*
2. *A class of states, $C(I_{k_i})$, is associated with each input I_{k_i} of M_k. The classes $C(I_{k_i})$ and $C(I_{k_j})$ must be disjoint for all $i \neq j$ and all k.*
3. *The complexity of M_k, $\chi(M_k)$, must be bounded by θk, where* redundancy θ *is fixed for all k.*
4. *At $\ell = 0$, let one of the inputs from $\{I_{k_i}\}$ be stored in each memory M_k in the sequence of memories $\{M_i\}$, with no further inputs in times $\ell > 0$. Let I_{k_i} denote the particular input stored in memory M_k. Let $\lambda_{k_i}(T)$ denote the probability that the state of M_k does not belong to $C(I_{k_i})$ at $\ell = T$ and further let $P_k^{\max}(T) = \max_i \lambda_{k_i}(T)$. Then for any $T > 0$ and $\delta > 0$, there must exist a k such that $P_k^{\max}(T) < \delta$.*

The demarcation of classes of states is equivalent to demarcating decoding regions.

Definition 5.10. *The* storage capacity*, $\mathfrak{C}$, of memory is a number such that there exist stable memory sequences for all memory redundancy values θ greater than $1/\mathfrak{C}$.*

Note that unlike channel capacity for the communication problem, there is no informational definition of storage capacity that is known to go with the operational definition.

The basic problem then is to determine storage capacity, which is a measure of the circuit complexity required to achieve arbitrarily reliable information storage. Due to the definition of memory stability, the circuit complexity of a memory system with positive storage capacity must be linear in block length. It is not obvious that positive storage capacity is possible. Luckily systems with message-passing correcting networks for LDPC codes do have linear complexity. Thus the remainder of the chapter considers whether stable memories can be constructed using the architecture depicted in Figure 5-8 and whether the storage capacity can be quantified.

Although Proposition 5.1 shows that Shannon reliability is not achievable for any noisy Gallager A decoder, the definition of stable information storage does not require this. By only requiring maintenance within a decoding region, the definition implies that either the contents of the memory may be read-out in coded form or equivalently that there is a noiseless output device that yields decoded information; call this noiseless output device the *silver decoder*.

Taylor had previously proven that there exist stable memories with positive storage capacity [247].

Theorem 5.5 ([247])**.** *There exists a stable sequence of memories with positive storage capacity, where every component in every memory has a fixed non-zero probability of error.*

The goal here is to provide precise quantitative achievability results for storage capacity.

Consider the construction of a memory with noisy registers as storage elements according to the architecture of Figure 5-8. There are memory registers connected to the outside world through noisy input/output pins. These registers are also connected to a noisy Gallager A LDPC decoder (as described in Section 5.4), which takes the register values as inputs and stores its computational results back into the registers. To find the storage capacity of this construction, first compute the complexity (presupposing that the construction will yield a stable sequence of memories).

The Gallager A check node operation is a $(d_{\mathsf{c}}-1)$-input XOR gate, which may be constructed from $d_{\mathsf{c}}-2$ two-input XOR gates, Figure 5-9. A variable node determines whether its $d_{\mathsf{v}}-1$ inputs are all the same and then compares to the original received value. Let $D_{d_{\mathsf{v}}}$ denote the complexity of this logic. The output of the comparison to the original received value is the value of the consensus view. One construction to implement the consensus logic, as depicted in Figure 5-9, is to OR together the outputs of a $(d_{\mathsf{v}}-1)$-input AND gate and a $(d_{\mathsf{v}}-1)$-input AND gate with inverted inputs. This is then XORed with the stored value. Such a circuit can be implemented with $2(d_{\mathsf{v}}-2)+2$ components, so $D_{d_{\mathsf{v}}} = 2d_{\mathsf{v}}-2$. The storage is carried out in n registers. The total complexity of the memory M_k, $\chi(M_k)_{\mathcal{C}^n(d_{\mathsf{v}},d_{\mathsf{c}})}$, is

$$\chi(M_k)_{\mathcal{C}^n(d_{\mathsf{v}},d_{\mathsf{c}})} = n(1+2d_{\mathsf{v}}-2+d_{\mathsf{v}}(d_{\mathsf{c}}-2)) = n(d_{\mathsf{v}}d_{\mathsf{c}}-1).$$

The information storage capability is n times the rate of the code, R. The complexity of an irredundant memory with the same storage capability is $\chi_{\mathrm{irr}_n} = Rn$.

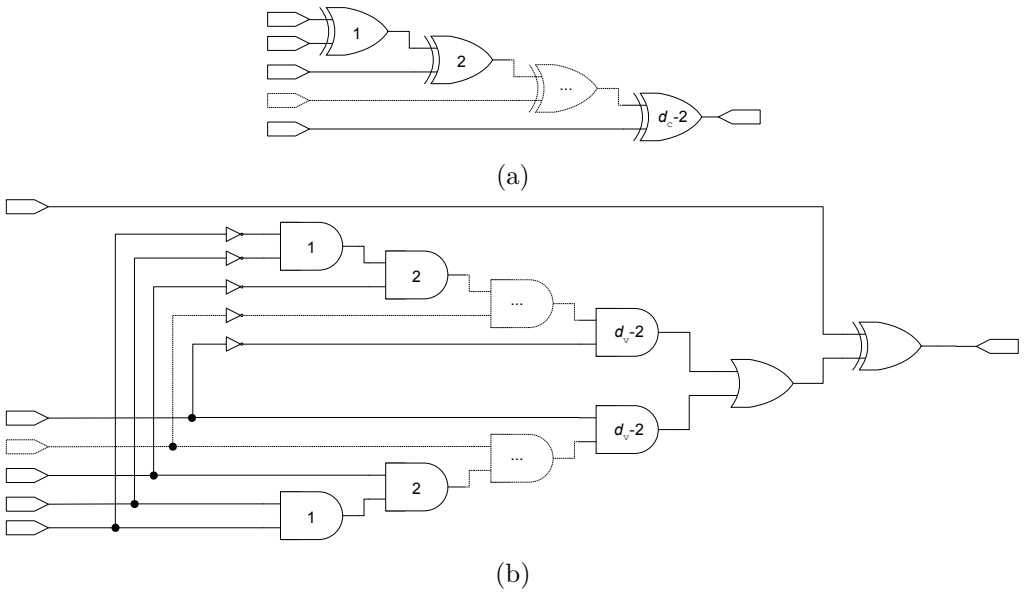

Figure 5-9. Gate-level implementation of nodes for a regular LDPC Gallager A correcting network. (a) check node. (b) variable node.

Hence, the redundancy is

$$\frac{\chi(M_k)_{\mathcal{C}^n(d_\mathsf{v},d_\mathsf{c})}}{\chi_{\mathrm{irr}_n}} = \frac{n(d_\mathsf{v}d_\mathsf{c}-1)}{Rn} \leq \frac{(d_\mathsf{v}d_\mathsf{c}-1)}{1-d_\mathsf{v}/d_\mathsf{c}}$$

which is a constant. By [49, Lemma 3.22], the inequality almost holds with equality with high probability for large n. For the $(3,6)$ regular LDPC code, the redundancy value is 34, so $\mathfrak{C} = 1/34$, if the construction does in fact yield stable memories.

The conditions under which the memory is stable depends on the silver decoder. Since silver decoder complexity does not enter, the silver decoder should be thought of as a maximum likelihood decoder. The Gallager lower bound to the ML decoding threshold for the $(3,6)$ regular LDPC code is $\varepsilon^*_{GLB} = 0.0914755$ [271, Table II]. Recall from Figure 5-4 that the decoding threshold for Gallager A decoding is $\varepsilon^*_{BRU} = 0.0394636562$.

If the probability of symbol error for the correcting network in the memory stays within the decoding threshold of the silver decoder, then stability follows. Thus the question reduces to determining the sets of component noisiness levels (α, ε) for which the decoding circuit achieves $(\eta = \varepsilon^*_{ML})$-reliability.

Consider a memory system where bits are stored in registers with probability α_r of flipping at each time step. An LDPC codeword is stored in these registers; the probability of incorrect storage at the first time step is ε. At each iteration, the variable node value from the correcting network is placed in the register. This stored value is used in the subsequent Gallager A variable node computation rather than a received value from the input pins. Suppose that the component noise values in the correcting network may be parameterized as in Section 5.4. Then a slight modification

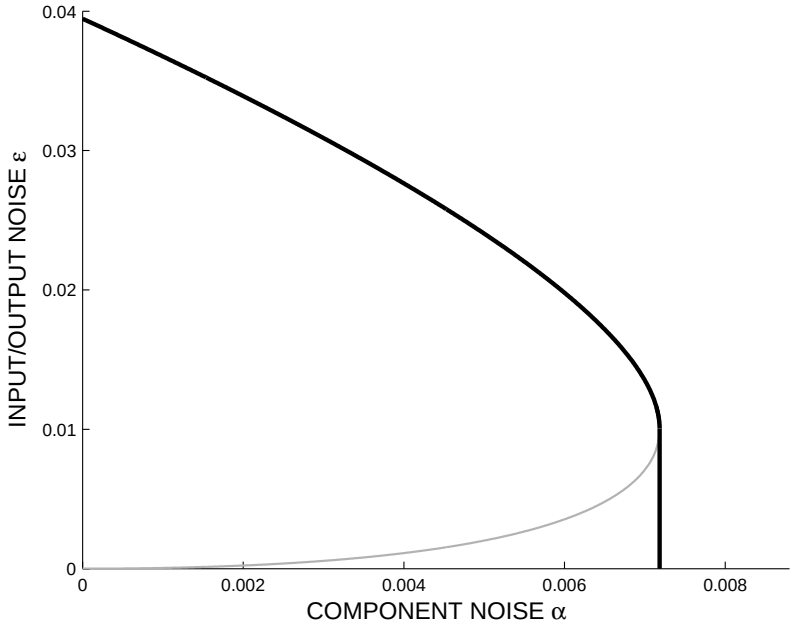

Figure 5-10. For a memory system constructed with noisy registers and a $(3, 6)$ LDPC Gallager A correcting network, the region $\mathfrak{R}$ is delimited by the black line, a region comprising the "region to use decoder" (thin gray line) and its hypograph.

of the analysis in Section 5.4 yields a density evolution equation

$$s_{\ell+1} = \varepsilon_2 - \varepsilon_2 q_\alpha^+(s_\ell) + (1 - \varepsilon_2) q_\alpha^-(s_\ell),$$

where $\varepsilon_2 = s_\ell(1 - \alpha_r) + \alpha_r(1 - s_\ell)$. There is a "region to use decoder" for this system, just as in Section 5.4. If $\alpha_r = \alpha$, this region is shown in Figure 5-10, and is slightly smaller than the region in Figure 5-4. Denote this region and its hypograph as $\mathfrak{R}$. It follows that $(\eta = \varepsilon^*_{BRU})$-reliability is achieved for $\mathfrak{R}$. Since ε^*_{BRU}-reliability is achievable, ε^*_{GLB}-reliability is achievable by monotonicity. Thus the construction yields stable memories.

Theorem 5.6. *Let $\mathfrak{R}$ be a set of memory component noise parameters (α, ε) within the region to use decoder of Figure 5-10 or its hypograph. Then a sequence of noisy memories constructed from $\mathfrak{R}$-components into a Gallager A correcting network for a $(3, 6)$ LDPC code have a storage capacity lower bounded as $\mathfrak{C} \geq 1/34$.*

By specifying a particular construction and using density evolution as the method of analysis, Theorem 5.6 gave a precise quantitative achievability result for storage capacity, expanding beyond the existence result in Theorem 5.5. Concentration around the cycle-free case, Theorem 5.4, which was used implicitly in this section also considerably reduced the circuit complexity that had previously been thought necessary for constructing reliable memories from unreliable components [247], by not requiring several copies of the correcting network.

Theorem 5.6 can directly be extended to use code ensembles better than the $(3, 6)$ regular code, but the question of an optimal architecture for memory systems remains open due to the lack of converse arguments. This theorem and the style of analysis used to prove it, however, give precise achievability results that lower bound the storage capacity.

5.7 Discussion

Loeliger et al. [215] had observed that decoders are robust to nonidealities and noise in physical implementations, however they had noted that "the quantitative analysis of these effects is a challenging theoretical problem." This chapter has taken steps to address this challenge by characterizing robustness to decoder noise.

The extension of the density evolution method to the case of faulty decoders allows a simplified means of asymptotic performance characterization. Results from this method show that in certain cases Shannon reliability is not achievable (Proposition 5.1), whereas in other cases it is achievable (Proposition 5.2). In either case, however, the degradation of a suitably defined decoding threshold is smooth with increasing decoder noise, whether in circuit nodes or circuit wires.

One may further consider a concatenated coding scheme where the inner code is an LDPC code with a noisy inner decoder, but the outer code has a silver decoder. Shannon reliable communication would be possible in many settings, even where results like Proposition 5.1 hold.

No attempt was made to apply fault masking methods to develop decoding algorithms with improved performance in the presence of noise. One approach might be to use coding within the decoder so as to reduce the decoder noise level α. Of course, the within-decoder code would need to be decoded. There are also more direct circuit-oriented techniques that may be applied [272, 273]. Following the concept of concatenated codes, concatenated decoders may also be promising. The basic idea of using a first (noiseless) decoder to correct many errors and then a second (noiseless) decoder to clean things up was already present in [254], but it may be extended to faulty decoders.

Reducing power consumption in decoder circuits has been an active area of research [19, 274–280], however power reduction often has the effect of increasing noise in the decoder. The trade-off developed between the quality of the communication channel and the quality of the decoder may provide guidelines for allocating resources in communication system design.

Analysis of other decoding algorithms with other error models may yield results similar to those obtained here. For greater generality, one might move beyond simple LDPC codes and consider arbitrary codes decoded with very general iterative decoding circuits [19] with suitable error models. An even more general model of computation such as a Turing machine or beyond [281] does not seem to have an obvious, appropriate error model.

Even just a bit of imagination provides numerous models of channel noise and transient circuit faults that may be investigated to provide further insights into the

fundamental limits of noisy communication and computing.

5.A Proof of Theorem 5.1

Let $\mathbf{x} \in \mathcal{C}^n$ be a codeword and let $\mathbf{Y}$ denote the corresponding channel output $\mathbf{Y} = \mathbf{xZ}$ (where the notation means pointwise multiplication on length n vectors). Note that $\mathbf{Z}$ is equal to the channel output observation when $\mathbf{x}$ is all-one. The goal is to show that messages sent during the decoding process for cases when the received codeword is either $\mathbf{xZ}$ or $\mathbf{x}$ correspond.

Let $\dot{n}_i$ be an arbitrary variable node and let $\dot{n}_j$ be one of its neighboring check nodes. Let $\nu_{ij}^{(\ell)}(\mathbf{y})$ and $\mu_{ij}^{(\ell)}(\mathbf{y})$ denote the variable-to-check message from $\dot{n}_i$ to $\dot{n}_j$ at the respective terminals in iteration ℓ, assuming received value $\mathbf{y}$. Similarly, let $\nu_{ji}^{(\ell)}(\mathbf{y})$ and $\mu_{ji}^{(\ell)}(\mathbf{y})$ be the check-to-variable message from $\dot{n}_j$ to $\dot{n}_i$ at the respective terminal in iteration ℓ assuming received value $\mathbf{y}$.

By Definition 5.1, the channel is memoryless binary-input output-symmetric and it may be modeled multiplicatively as

$$Y_t = x_t Z_t, \tag{5.5}$$

where $\{Z_t\}$ is a sequence of i.i.d. random variables and t is the channel usage time. The validity of the multiplicative model is shown in [43, p. 605] and [49, p. 184].

By the multiplicative model (5.5), $\nu_{ij}^{(0)}(\mathbf{y}) = \nu_{ij}^{(0)}(\mathbf{xz})$. Recalling that $x_i \in \{\pm 1\}$, by the variable node symmetry condition (Definition 5.3) which includes computation noise $u_{\dot{n}_i}^{(0)}$, it follows that $\nu_{ij}^{(0)}(\mathbf{y}) = \nu_{ij}^{(0)}(\mathbf{xz}) = x_i \nu_{ij}^{(0)}(\mathbf{z})$.

Now consider the wire noise $w_{ij}^{(0)}$ on the message from $\dot{n}_i$ to $\dot{n}_j$. It is symmetric (Definition 5.4) and so $\nu_{ij}^{(0)}(\mathbf{y}) = x_i \nu_{ij}^{(0)}(\mathbf{z})$ implies $\mu_{ij}^{(0)}(\mathbf{y}) = x_i \mu_{ij}^{(0)}(\mathbf{z})$.

Assume that $\mu_{ij}^{(0)}(\mathbf{y}) = x_i \mu_{ij}^{(0)}(\mathbf{z})$ for all (i, j) pairs and some $\ell \geq 0$. Let $\mathcal{N}_{\dot{n}_j}$ be the set of all variable nodes that are connected to check node $\dot{n}_j$. Since $\mathbf{x}$ is a codeword, it satisfies the parity checks, and so $\prod_{k \in \mathcal{N}_{\dot{n}_j}} = 1$. Then from the check node symmetry condition (Definition 5.2), $\nu_{ji}^{(\ell+1)}(\mathbf{y}) = x_i \nu_{ji}^{(\ell+1)}(\mathbf{z})$. Further, by the wire noise symmetry condition (Definition 5.4), $\mu_{ji}^{(\ell+1)}(\mathbf{y}) = x_i \mu_{ji}^{(\ell+1)}(\mathbf{z})$. By invoking the variable node symmetry condition (Definition 5.3) again, it follows that $\nu_{ij}^{(\ell+1)}(\mathbf{y}) = x_i \nu_{ij}^{(\ell+1)}(\mathbf{z})$ for all (i, j) pairs. Thus by induction, all messages to and from variable node $\dot{n}_i$ when $\mathbf{y}$ is received are equal to the product of x_i and the corresponding message when $\mathbf{z}$ is received.

Both decoders proceed in one-to-one correspondence and commit exactly the same number of errors.

5.B Proof of Theorem 5.2

Prior to giving the proof of Theorem 5.2, a review of some definitions from probability theory [282] and the Hoeffding-Azuma inequality are provided.

Consider a measurable space $(\Omega, \mathcal{F})$ consisting of a sample space Ω and a σ-algebra $\mathcal{F}$ of subsets of Ω that contains the whole space and is closed under complementation and countable unions. A random variable is an $\mathcal{F}$-measurable function on Ω. If there is a collection $(Z_\gamma | \gamma \in C)$ of random variables $Z_\gamma : \Omega \to \mathbb{R}$, then

$$\mathcal{Z} = \sigma(Z_\gamma | \gamma \in C)$$

is defined to be the smallest σ-algebra $\mathcal{Z}$ on Ω such that each map $(Z_\gamma | \gamma \in C)$ is $\mathcal{Z}$-measurable.

Definition 5.11 (Filtration)**.** *Let $\{\mathcal{F}_i\}$ be a sequence of σ-algebras with respect to the same sample space Ω. These $\mathcal{F}_i$ are said to form a filtration if $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots$ are ordered by refinement in the sense that each subset of Ω in $\mathcal{F}_i$ is also in $\mathcal{F}_j$ for $i \leq j$. Also $\mathcal{F}_0 = \{\emptyset, \Omega\}$.*

Usually, $\{\mathcal{F}_i\}$ is the natural filtration $\mathcal{F}_i = \sigma(Z_0, Z_1, \ldots, Z_i)$ of some sequence of random variables $(Z_0, Z_1, \ldots)$, and then the knowledge about ω known at step i consists of the values $Z_0(\omega), Z_1(\omega), \ldots, Z_i(\omega)$.

For a probability triple $(\Omega, \mathcal{F}, \mathbb{P})$, a version of the conditional expectation of a random variable Z given a σ-algebra $\mathcal{F}$ is a random variable denoted $\mathrm{E}[Z|\mathcal{F}]$. Two versions of conditional expectation agree almost surely, but measure zero departures are not considered subsequently; one version is fixed as canonical. Conditional expectation given a measurable event $\mathfrak{E}$ is denoted $\mathrm{E}\left[Z|\sigma(\mathfrak{E})\right]$ and conditional expectation given a random variable W is denoted $\mathrm{E}[Z|\sigma(W)]$.

Definition 5.12 (Martingale)**.** *Let $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots$ be a filtration on Ω and let $Z_0, Z_1, \ldots$ be a sequence of random variables on Ω such that Z_i is $\mathcal{F}_i$-measurable. Then $Z_0, Z_1, \ldots$ is a martingale with respect to the filtration $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots$ if $\mathrm{E}[Z_i|\mathcal{F}_{i-1}] = Z_{i-1}$.*

A generic way to construct a martingale is Doob's construction.

Definition 5.13 (Doob Martingale)**.** *Let $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots$ be a filtration on Ω and let Z be a random variable on Ω. Then the sequence of random variables $Z_0, Z_1, \ldots$ such that $Z_i = \mathrm{E}[Z|\mathcal{F}_i]$ is a Doob martingale.*

Lemma 5.1 (Hoeffding-Azuma Inequality [43,283,284])**.** *Let $Z_0, Z_1, \ldots$ be a martingale with respect to the filtration $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots$ such that for each $i > 0$, the following bounded difference condition is satisfied*

$$|Z_i - Z_{i-1}| \leq \alpha_i,\ \alpha_i \in [0, \infty).$$

Then for all $n > 0$ and any $\xi > 0$,

$$\Pr\left[|Z_n - Z_0| \geq \xi\right] \leq 2 \exp\left(-\frac{\xi^2}{2\sum_{k=1}^{n} \alpha_k^2}\right).$$

Now to the proof of Theorem 5.2; as noted before, it is an extension of [43, Theorem 2] or [49, Theorem 4.94]. The basic idea is to construct a Doob martingale about the object of interest by revealing various randomly determined aspects in a filtration-refining manner. The first set of steps is used to reveal which code was chosen from the ensemble of codes; the $nd_{\mathtt{v}}$ edges in the bipartite graph are ordered in some arbitrary manner and exposed one by one. Then the n channel noise realizations are revealed. At this point the exact graph and the exact channel noise realizations encountered have been revealed. Now the decoder noise realizations must be revealed. There are n variable nodes, so the computation noise in each of them is revealed one by one. There are $nd_{\mathtt{v}}$ edges over which variable-to-check communication noise is manifested. Then there are $nd_{\mathtt{v}}/d_{\mathtt{c}}$ check nodes with computation noise, and finally there are $nd_{\mathtt{v}}$ check-to-variable communication noises for one iteration of the algorithm. The decoder noise realizations are revealed for each iteration. At the beginning of the revelation process, the average (over choice of code, channel noise realization, and decoder noise realization) is known; after the $m = (d_{\mathtt{v}} + 2\ell d_{\mathtt{v}} + 1 + \ell + \ell d_{\mathtt{v}}/d_{\mathtt{c}})n$ revelation steps, the exact system used is known.

Recall that Z denotes the number of incorrect values held at the end of the ℓth iteration for a particular $(g, y, w, u) \in \Omega$. Since g is a graph in the set of labeled bipartite factor graphs with variable node degree $d_{\mathtt{v}}$ and check node degree $d_{\mathtt{c}}$, $\mathcal{G}^n(d_{\mathtt{v}}, d_{\mathtt{c}})$; y is a particular input to the decoder, $y \in \mathcal{Y}^n$; w is a particular realization of the message-passing noise, $w \in \mathcal{M}^{2\ell d_{\mathtt{v}} n}$; and u is a particular realization of the local computation noise, $u \in \mathcal{U}^{(\ell+\ell d_{\mathtt{v}}/d_{\mathtt{c}})n}$, the sample space is $\Omega = \mathcal{G}^n(d_{\mathtt{v}}, d_{\mathtt{c}}) \times \mathcal{Y}^n \times \mathcal{M}^{2\ell d_{\mathtt{v}} n} \times \mathcal{U}^{(\ell+\ell d_{\mathtt{v}}/d_{\mathtt{c}})n}$.

In order to define random variables, first define the following exposure procedure. Suppose realizations of random quantities are exposed sequentially. First expose the $d_{\mathtt{v}}n$ edges of the graph one at a time by exposing which check node is connected to which variable node. A connection point on a node is called a socket. At step $i \leq d_{\mathtt{v}}n$ expose the particular check node socket which is connected to the ith variable node socket. Next, in the following n steps, expose the received values y_i one at a time. Finally in the remaining $(2d_{\mathtt{v}} + 1 + d_{\mathtt{v}}/d_{\mathtt{c}})\ell n$ steps, expose the decoder noise values u_i and w_i that were encountered in all iterations up to iteration ℓ.

Let $\equiv_i$, $0 \leq i \leq m$, be a sequence of equivalence relations on the sample space Ω ordered by refinement. Refinement means that $(g', y', w', u') \equiv_i (g'', y'', w'', u'')$ implies $(g', y', w', u') \equiv_{i-1} (g'', y'', w'', u'')$. The equivalence relations define equivalence classes such that $(g', y', w', u') \equiv_i (g'', y'', w'', u'')$ if and only if the realizations of random quantities revealed in the first i steps for both pairs is the same.

Now, define a sequence of random variables $Z_0, Z_1, \ldots, Z_m$. Let the random variable Z_0 be $Z_0 = \mathrm{E}[Z]$, where the expectation is over the code choice, channel noise, and decoder noise. The remaining random variables Z_i are constructed as conditional expectations given the measurable equivalence events $(g', y', w', u') \equiv_i (g, y, w, u)$:

$$Z_i(g, y, w, u) = \mathrm{E}\left[Z(g', y', w', u') | \sigma((g', y', w', u') \equiv_i (g, y, w, u))\right].$$

Note that $Z_m = Z$ and that by construction $Z_0, Z_1, \ldots, Z_m$ is a Doob martingale. The filtration is understood to be the natural filtration of the random variables $Z_0, Z_1, \ldots, Z_m$.

To use the Hoeffding-Azuma inequality to give bounds on

$$\Pr\left[|Z - \mathrm{E}[Z]| > nd_{\mathtt{v}}\epsilon/2\right] = \Pr\left[|Z_m - Z_0| > nd_{\mathtt{v}}\epsilon/2\right],$$

bounded difference conditions

$$|Z_{i+1}(g, y, w, u) - Z_i(g, y, w, u)| \leq \alpha_i,\ i = 0, \ldots, m-1$$

need to be proved for suitable constants α_i that may depend on $d_{\mathtt{v}}$, $d_{\mathtt{c}}$, and ℓ.

For the steps where bipartite graph edges are exposed, it was shown in [43, p. 614] that

$$|Z_{i+1}(g, y, w, u) - Z_i(g, y, w, u)| \leq 8(d_{\mathtt{v}}d_{\mathtt{c}})^{\ell},\ 0 \leq i < nd_{\mathtt{v}}.$$

It was further shown in [43, p. 615] that for the steps when the channel outputs are revealed that

$$|Z_{i+1}(g, y, w, u) - Z_i(g, y, w, u)| \leq 2(d_{\mathtt{v}}d_{\mathtt{c}})^{\ell},\ nd_{\mathtt{v}} \leq i < n(1 + d_{\mathtt{v}}). \tag{5.6}$$

It remains to show that the inequality is also fulfilled for steps when decoder noise realizations are revealed. The bounding procedure is nearly identical to that which yields (5.6). When a node noise realization u is revealed, clearly only something whose directed neighborhood includes the node at which the noise u causes perturbations can be affected. Similarly, when an edge noise realization w is revealed, only something whose directed neighborhood includes the edge on which the noise w causes perturbations can be affected. In [43, p. 603], it is shown that the size of the directed neighborhood of depth 2ℓ of the node $\dot{n}(u)$ associated with noise u is bounded as $|\mathcal{N}^{2\ell}_{\dot{n}(u)}| \leq 2(d_{\mathtt{v}}d_{\mathtt{c}})^{\ell}$ and similarly the size of the directed neighborhood of length 2ℓ of the edge $\vec{e}(w)$ associated with noise w is bounded as $|\mathcal{N}^{2\ell}_{\vec{e}(w)}| \leq 2(d_{\mathtt{v}}d_{\mathtt{c}})^{\ell}$. Since the maximum depth that can be affected by a noise perturbation is 2ℓ, a weak uniform bound for the remaining exposure steps is

$$|Z_{i+1}(g, y, w, u) - Z_i(g, y, w, u)| \leq 2(d_{\mathtt{v}}d_{\mathtt{c}})^{\ell},\ n(1 + d_{\mathtt{v}})d_{\mathtt{v}} \leq i < m.$$

Since bounded difference constants α_i have been provided for all i, the theorem follows from application of the Hoeffding-Azuma inequality to the martingale.

One may compute a particular value of β to use as follows. The bounded difference sum is

$$\begin{aligned}\sum_{k=1}^{m} \alpha_k^2 &= 64nd_{\mathtt{v}}(d_{\mathtt{v}}d_{\mathtt{c}})^{2\ell} + 4n(d_{\mathtt{v}}d_{\mathtt{c}})^{2\ell} + 4[2\ell d_{\mathtt{v}}n + n\ell + n\ell d_{\mathtt{v}}/d_{\mathtt{c}}](d_{\mathtt{v}}d_{\mathtt{c}})^{2\ell} \\ &= n\left\{64d_{\mathtt{v}} + 4 + 8d_{\mathtt{v}}\ell + \ell + \tfrac{d_{\mathtt{v}}\ell}{d_{\mathtt{c}}}\right\} {d_{\mathtt{v}}}^{2\ell}{d_{\mathtt{c}}}^{2\ell}.\end{aligned}$$

Setting constants in the theorem and in the Hoeffding-Azuma inequality equal yields

$$\begin{aligned}\tfrac{1}{\beta} &= 512{d_{\mathsf{v}}}^{2\ell-1}{d_{\mathsf{c}}}^{2\ell} + 32{d_{\mathsf{v}}}^{2\ell-2}{d_{\mathsf{c}}}^{2\ell} + 64\ell{d_{\mathsf{v}}}^{2\ell-1}{d_{\mathsf{c}}}^{2\ell} + 8\ell{d_{\mathsf{v}}}^{2\ell-1}{d_{\mathsf{c}}}^{2\ell-1} + 8\ell{d_{\mathsf{v}}}^{2\ell-2}{d_{\mathsf{c}}}^{2\ell} \\ &\leq (544 + 80\ell){d_{\mathsf{v}}}^{2\ell-1}{d_{\mathsf{c}}}^{2\ell}.\end{aligned}$$

Thus $\frac{1}{\beta}$ can be taken as $(544 + 80\ell){d_{\mathsf{v}}}^{2\ell-1}{d_{\mathsf{c}}}^{2\ell}$.

■ 5.C An Analytical Expression

An analytical expression for $\varepsilon^*(\eta = 1/10, \alpha = 5 \times 10^{-3})$ is

$$\tfrac{1}{2} - \tfrac{1}{2}\sqrt{1 + 4c_7},$$

where c_7 is the second root of the polynomial in $\check{\varepsilon}$

$$c_1 + c_2\check{\varepsilon} + c_3\check{\varepsilon}^2 + c_4\check{\varepsilon}^3 + c_5\check{\varepsilon}^4 + c_6\check{\varepsilon}^5,$$

and constants $(c_1, \ldots, c_6)$ are defined as follows.

$$\begin{aligned}c_1 &= 36\alpha^2 - 360\alpha^3 + 1860\alpha^4 - 6240\alpha^5 + 14752\alpha^6 - 25344\alpha^7 + 31680\alpha^8 \\ &\quad - 28160\alpha^9 + 16896\alpha^{10} - 6144\alpha^{11} + 1024\alpha^{12} \\ &= \frac{3424572914129280658801}{400000000000000000000000}\end{aligned}$$

$$\begin{aligned}c_2 &= 1 - 72\alpha + 1080\alpha^2 - 8160\alpha^3 + 38640\alpha^4 - 125952\alpha^5 + 295424\alpha^6 \\ &\quad - 506880\alpha^7 + 633600\alpha^8 - 563200\alpha^9 + 337920\alpha^{10} - 122880\alpha^{11} + 20480\alpha^{12} \\ &= \frac{133200752195329280658801}{200000000000000000000000}\end{aligned}$$

$$\begin{aligned}c_3 &= 32 - 864\alpha + 10080\alpha^2 - 69120\alpha^3 + 314880\alpha^4 - 1012224\alpha^5 + 2364928\alpha^6 \\ &\quad - 4055040\alpha^7 + 5068800\alpha^8 - 4505600\alpha^9 + 2703360\alpha^{10} - 983040\alpha^{11} + 163840\alpha^{12} \\ &= \frac{698088841835929280658801}{250000000000000000000000}\end{aligned}$$

$$\begin{aligned}c_4 &= 160 - 3840\alpha + 42240\alpha^2 - 281600\alpha^3 + 1267200\alpha^4 - 4055040\alpha^5 + 9461760\alpha^6 \\ &\quad - 16220160\alpha^7 + 20275200\alpha^8 - 18022400\alpha^9 + 10813440\alpha^{10} - 3932160\alpha^{11} \\ &\quad + 655360\alpha^{12} \\ &= \frac{886384871716129280658801}{625000000000000000000}\end{aligned}$$

$$
\begin{aligned}
c_5 &= 320 - 7680\alpha + 84480\alpha^2 - 563200\alpha^3 + 2534400\alpha^4 - 8110080\alpha^5 + 18923520\alpha^6 \\
&\quad - 32440320\alpha^7 + 40550400\alpha^8 - 36044800\alpha^9 + 21626880\alpha^{10} - 7864320\alpha^{11} \\
&\quad + 1310720\alpha^{12} \\
&= \frac{88638487171612928065880 1}{3125000000000000000 0}
\end{aligned}
$$

$$
\begin{aligned}
c_6 &= 256 - 6144\alpha + 67584\alpha^2 - 450560\alpha^3 + 2027520\alpha^4 - 6488064\alpha^5 + 15138816\alpha^6 \\
&\quad - 25952256\alpha^7 + 32440320\alpha^8 - 28835840\alpha^9 + 17301504\alpha^{10} - 6291456\alpha^{11} \\
&\quad + 1048576\alpha^{12} \\
&= \frac{88638487171612928065880 1}{3906250000000000000 0}
\end{aligned}
$$

As given in Table 5.1, the numerical value of $\varepsilon^*(\eta = 1/10, \alpha = 5 \times 10^{-3})$ is 0.0266099758.

Similarly complicated analytical expressions are available for the other entries of Table 5.1 and the values used to create Figures 5-3, 5-4, and 5-5.

Chapter 6

Operation Reliability—Permanent Faults

Physical systems such as decoding circuits have a tendency to fail at random times [12], due to device failure, energy exhaustion, or adversarial attacks. Such a permanent circuit fault is qualitatively different from the kind of transient faults studied in Chapter 5 because there is no possibility of future recovery. To explore the impact of permanent decoder failures on communication systems, this chapter defines and determines information-theoretic limits on the maximum number of bits that can be communicated to meet a fixed maximax message error requirement.

In designing resource-limited communication systems, there is often a trade-off between decoder survivability and channel reliability before failure, similar to the trade-off between channel and decoder resources in Figure 5-5. Results in this chapter formalize a resource allocation equivalence relationship between survivability and reliability.

Since catastrophic decoder failure causes the entire communication system to fail, the failure can be modeled elsewhere in the system without much change to the mathematical formulation. This chapter therefore models system failure as communication channel failure: channels that die. As noted by Jacobs [13], "a communication channel... might be inoperative because of an amplifier failure, a broken or cut telephone wire, ..." The notion of outage in wireless communication [285, 286] is similar to channel death, except that outage is not a permanent condition. Likewise for lost letters in postal channels [287]. Therefore the block length asymptotics that are useful to study those channel models are not useful for channels that die. Recent work that has similar motivations as this chapter provides the outage capacity of a wireless channel [288].

Given that block length asymptotics are not useful, limits on channel coding with finite block length [289], which have seen renewed interest [290–294], are central to the development.[1] Indeed, channels that die bring the notion of finite block length to the fore and provide a concrete reason to step back from infinity.[2]

[1]Recall from Chapter 2 that tree codes are beyond the scope of this thesis. A reformulation of communicating over channels that die using tree codes [52, Chapter 10] with early termination [295] would, however, be interesting. In fact, communicating over channels that die using convolutional codes with sequential decoding would be very natural, but would require performance criteria different from the ones developed here.

[2]The phrase "back from infinity" is borrowed from J. Ziv's 1997 Shannon Lecture.

As it turns out, the central trade-off in communicating over channels that die is in the lengths of codeword blocks. Longer blocks improve communication performance as classically known, whereas shorter blocks have a smaller probability of being prematurely terminated due to channel death. Dynamic programming can be used to find an optimal ordered integer partition for the sequence of block lengths. Solving the dynamic program demonstrates that channel state feedback does not improve performance.

Optimization of codeword block lengths is reminiscent of frame size control in wireless networks [296–299], however such techniques are used in conjunction with automatic repeat request protocols and are motivated by amortizing protocol information. Moreover, those results demonstrate the benefit of adapting to either channel state or decision feedback; in contrast, adaptation provides no benefit for channels that die.

The remainder of the chapter is organized as follows. Section 6.1 defines discrete channels that die and shows that these channels have zero capacity. Section 6.2 states the communication system model and also fixes novel performance criteria. Section 6.3 shows that Shannon reliability in communication is not achievable, strengthening the result of zero capacity, as well as provides the optimal communication scheme and determines its performance. Section 6.4 optimizes performance for several death distributions and provides some design examples. Section 6.5 provides further discussion.

■ 6.1 Channel Model of Permanent Circuit Faults

Consider a channel with finite input alphabet $\mathcal{X}$ and finite output alphabet $\mathcal{Y}$. It has an alive state $s = a$ when it acts like a discrete memoryless channel (DMC)[3] and a dead state $s = d$ when it erases the input, producing the special symbol '?'. The dead state of the channel is meant as a model of catastrophic decoder failure. Assume throughout the chapter that the DMC from the alive state has zero-error capacity [301] equal to zero.[4]

If the channel acts like a binary symmetric channel (BSC) with crossover probability $0 < \varepsilon < 1$ in the alive state, with $\mathcal{X} = \{0, 1\}$, and $\mathcal{Y} = \{0, 1, ?\}$, then the transmission matrix in the alive state is

$$p_{Y|X}(y|x, s = a) = p_a(y|x) = \begin{bmatrix} 1-\varepsilon & \varepsilon & 0 \\ \varepsilon & 1-\varepsilon & 0 \end{bmatrix}, \tag{6.1}$$

and the transmission matrix in the dead state is

$$p_{Y|X}(y|x, s = d) = p_d(y|x) = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}. \tag{6.2}$$

[3]Results can be extended to cases where the channel acts like other channels [76,300] in the alive state.

[4]If the channel is noiseless in the alive state, the problem is similar to settings where fountain codes [302] are used in the point-to-point case and growth codes [303] are used in the network case.

The channel starts in state $s = a$ and then transitions to $s = d$ at some random time T, where it remains for all time thereafter. That is, the channel is in state a for times $n = 1, 2, \ldots, T$ and in state d for times $n = T + 1, T + 2, \ldots$. The death time distribution is denoted $p_T(t)$. It is assumed throughout the chapter that there exists a finite $t^\dagger$ such that $p_T(t^\dagger) > 0$.

6.1.1 Finite-State Semi-Markov Channel

Channels that die can be classified as finite-state channels (FSCs) [51, Section 4.6].

Proposition 6.1. *A channel that dies $(\mathcal{X}, p_a(y|x), p_d(y|x), p_T(t), \mathcal{Y})$ is a finite-state channel.*

Proof. Follows by definition, since the channel has two states. □

Channels that die have semi-Markovian [304, Section 4.8], [305, Section 5.7] properties.

Definition 6.1. *A semi-Markov process changes state according to a Markov chain but takes a random amount of time between changes. More specifically, it is a stochastic process with states from a discrete alphabet $\mathcal{S}$, such that whenever it enters state s, $s \in \mathcal{S}$:*

- *The next state it will enter is state r with probability that depends only on $s, r \in \mathcal{S}$.*
- *Given that the next state to be entered is state r, the time until the transition from s to r occurs has distribution that depends only on $s, r \in \mathcal{S}$.*

Definition 6.2. *The Markovian sequence of states of a semi-Markov process is called the embedded Markov chain of the semi-Markov process.*

Definition 6.3. *A semi-Markov process is irreducible if its embedded Markov chain is irreducible.*

Proposition 6.2. *A channel that dies $(\mathcal{X}, p_a(y|x), p_d(y|x), p_T(t), \mathcal{Y})$ has a channel state sequence that is a non-irreducible semi-Markov process.*

Proof. When in state a, the next state is d with probability 1 and given that the next state is to be d, the time until the transition from a to d has distribution $p_T(t)$. When in state d, the next state is d with probability 1. Thus, the channel state sequence is a semi-Markov process.

The semi-Markov state process is not irreducible because the a state of the embedded Markov chain is transient. □

Note that when T is a geometric random variable, the channel state process forms a Markov chain, with transient state a and recurrent, absorbing state d.

There are further special classes of FSCs.

Definition 6.4. *An FSC is a finite-state semi-Markov channel (FSSMC) if its state sequence forms a semi-Markov process.*

Definition 6.5. *An FSC is a finite-state Markov channel (FSMC) if its state sequence forms a Markov chain.*

Proposition 6.3. *A channel that dies* $(\mathcal{X}, p_a(y|x), p_d(y|x), p_T(t), \mathcal{Y})$ *is an FSSMC and is an FSMC when* T *is geometric.*

Proof. Follows from Propositions 6.1 and 6.2. □

FSMCs have been widely studied in the literature [51, 306, 307], particularly the child's toy/panic button channel of Gallager [306, p. 26], [51, p. 103] and the Gilbert–Elliott channel and its extensions [308, 309]. Contrarily, FSSMCs seem to not have been specifically studied in information theory. There are a few works [310–312] that give semi-Markov channel models for wireless communications systems but do not provide information-theoretic characterizations.

■ 6.1.2 Capacity is Zero

A channel that dies has capacity (see Theorem 2.3) equal to zero. To show this, first notice that if the initial state of a channel that dies were not fixed, then it would be an indecomposable FSC [51, Section 4.6], where the effect of the initial state dies away.

Proposition 6.4. *If the initial state of a channel that dies* $(\mathcal{X}, p_a(y|x), p_d(y|x), p_T(t), \mathcal{Y})$ *is not fixed, then it is an indecomposable FSC.*

Proof. The embedded Markov chain for a channel that dies has a unique absorbing state d. □

Indecomposable FSCs have the property that the upper capacity, defined in [51, (4.6.6)], and lower capacity, defined in [51, (4.6.3)], are identical [51, Theorem 4.6.4]. This can be used to show that the capacity of a channel that dies is zero.

Proposition 6.5. *The capacity of a channel that dies* $(\mathcal{X}, p_a(y|x), p_d(y|x), p_T(t), \mathcal{Y})$ *is zero.*

Proof. Although the initial state $s_1 = a$ here, temporarily suppose that s_1 may be either a or d. Then the channel is indecomposable by Proposition 6.4.

The lower capacity $\underline{C}$ equals the upper capacity $\overline{C}$, for indecomposable channels by [51, Theorem 4.6.4]. The information rate of a memoryless $p_d(y|x)$ 'dead' channel is clearly zero for any input distribution, so the lower capacity $\underline{C} = 0$. Thus the capacity for a channel that dies with initial alive state is $C = \overline{C} = 0$. □

6.2 Communication System Operation

To information theoretically characterize a channel that dies, a communication system that contains the channel is described.

There is an information stream (like i.i.d. equiprobable bits), which can be grouped into a sequence of k messages, $(W_1, W_2, \ldots, W_k)$. Each message W_i is drawn from a message set $\mathcal{W}_i = \{1, 2, \ldots, M_i\}$. Each message W_i is encoded into a channel input codeword $X_1^{n_i}(W_i)$ and these codewords $(X_1^{n_1}(W_1), X_1^{n_2}(W_2), \ldots, X_1^{n_k}(W_k))$ are transmitted in sequence over the channel. A noisy version of this codeword sequence is received, $Y_1^{n_1+n_2+\cdots+n_k}(W_1, W_2, \ldots, W_k)$. The receiver then guesses the sequence of messages using an appropriate decoding rule f_D, to produce $(\hat{W}_1, \hat{W}_2, \ldots, \hat{W}_k) = f_D(Y_1^{n_1+n_2+\cdots+n_k})$. The $\hat{W}_i$s are drawn from alphabets $\mathcal{W}_i^{\ominus} = \mathcal{W}_i \cup \ominus$, where the $\ominus$ message indicates the decoder declaring an erasure. The decoder makes an error on message i if $\hat{W}_i \neq W_i$ and $\hat{W}_i \neq \ominus$.

These system definitions may be formalized, following Chapter 2.

Definition 6.6. *For a channel that dies $(\mathcal{X}, p_a(y|x), p_d(y|x), p_T(t), \mathcal{Y})$, an (M_i, n_i) individual message code consists of:*

- *An individual message set $\{1, 2, \ldots, M_i\}$, and*
- *An individual message encoding function $f_E^{(i)} : \{1, 2, \ldots, M_i\} \mapsto \mathcal{X}^{n_i}$.*

The individual message index set $\{1, 2, \ldots, M_i\}$ is denoted $\mathcal{W}_i$, and the set of individual message codewords $\{f_E^{(i)}(1), f_E^{(i)}(2), \ldots, f_E^{(i)}(M_i)\}$ is called the individual message codebook.

Definition 6.7. *An $(M_i, n_i)_{i=1}^k$ code for a channel that dies $(\mathcal{X}, p_a(y|x), p_d(y|x), p_T(t), \mathcal{Y})$ is a sequence of k individual message codes, $(M_i, n_i)_{i=1}^k$, in the sense of comprising:*

- *A sequence of individual message index sets $\mathcal{W}_1, \ldots, \mathcal{W}_k$,*
- *A sequence of individual message encoding functions $f_E = (f_E^{(1)}, \ldots, f_E^{(k)})$, and*
- *A decoding function $f_D : \mathcal{Y}^{\sum_{i=1}^k n_i} \mapsto \mathcal{W}_1^{\ominus} \times \cdots \times \mathcal{W}_k^{\ominus}$.*

There is no essential loss of generality by assuming that the decoding function f_D is decomposed into a sequence of individual message decoding functions $f_D = (f_D^{(1)}, f_D^{(2)}, \ldots, f_D^{(k)})$ where $f_D^{(i)} : \mathcal{Y}^{n_i} \mapsto \mathcal{W}_i^{\ominus}$ when individual messages are chosen independently, due to this independence and the conditional memorylessness of the channel. To define performance measures, assume that the decoder operates on an individual message basis. That is, when applying the communication system, let $\hat{W}_1 = f_D^{(1)}(Y_1^{n_1})$, $\hat{W}_2 = f_D^{(2)}(Y_{n_1+1}^{n_1+n_2})$, and so on.

For the sequel, make a further assumption on the operation of the decoder.

Assumption 6.1. *If all n_i channel output symbols used by individual message decoder $f_D^{(i)}$ are not ?, then the range of $f_D^{(i)}$ is $\mathcal{W}_i$. If any of the n_i channel output symbols used by individual message decoder $f_D^{(i)}$ are ?, then $f_D^{(i)}$ maps to $\ominus$.*

This assumption corresponds to the physical properties of a communication system where the decoder fails catastrophically. Once the decoder fails, it cannot perform any decoding operations, and so the ? symbols in the channel model of system failure must be ignored.

6.2.1 Performance Measures

Slightly modifying definitions of message error probabilities from Chapter 2 so as not to penalize declared erasures $\ominus$ leads to the following.

Definition 6.8. *For all $1 \leq w \leq M_i$, let*

$$\lambda_w(i) = \Pr[\hat{W}_i \neq w | W_i = w, \hat{W}_i \neq \ominus]$$

be the conditional message probability of error given that the ith individual message is w.

Definition 6.9. *The maximal probability of error for an (M_i, n_i) individual message code is*

$$P_e^{\max}(i) = \max_{w \in \mathcal{W}_i} \lambda_w(i).$$

Definition 6.10. *The maximal probability of error for an $(M_i, n_i)_{i=1}^k$ code is*

$$P_e^{\text{maximax}} = \max_{i \in \{1,\ldots,k\}} P_e^{\max}(i).$$

Performance criteria weaker than traditional in information theory are defined, since the capacity of a channel that dies is zero (Proposition 6.5). In particular, define formal notions of how much information is transmitted using a code and how long it takes.

Definition 6.11. *The transmission time of an $(M_i, n_i)_{i=1}^k$ code is $N = \sum_{i=1}^k n_i$.*

Definition 6.12. *The expected transmission volume of an $(M_i, n_i)_{i=1}^k$ code is*

$$V = \mathrm{E}_T \left\{ \sum_{i \in \{1,\ldots,k | \hat{W}_i \neq \ominus\}} \log M_i \right\}.$$

Notice that although declared erasures do not lead to errors, they do not contribute transmission volume either.

The several performance criteria for a code may be combined together.

Definition 6.13. *Given $0 \leq \eta < 1$, a pair of numbers (N_0, V_0) (where N_0 is a positive integer and V_0 is non-negative) is said to be an achievable transmission time-volume at η-reliability if there exists, for some k, an $(M_i, n_i)_{i=1}^k$ code for the channel that dies $(\mathcal{X}, p_a(y|x), p_d(y|x), p_T(t), \mathcal{Y})$ such that*

$$P_e^{\text{maximax}} \leq \eta,$$

$$N \leq N_0,$$

and

$$V \geq V_0.$$

Moreover, (N_0, V_0) is said to be an achievable transmission time-volume at Shannon reliability if it is an achievable transmission time-volume at η-reliability for all $0 < \eta < 1$.

6.3 Limits on Communication

Having defined the notion of achievable transmission time-volume at various levels of reliability, the goal of this section is to demarcate what is achievable.

6.3.1 Shannon Reliability is Not Achievable

Not only is the capacity of a channel that dies zero, but there is no $V > 0$ such that (N, V) is an achievable transmission time-volume at Shannon reliability. A coding scheme that always declares erasures would achieve zero error probability (and therefore Shannon reliability) but would not provide positive transmission volume; this is also not allowed under Assumption 6.1.

Lemmata are stated and proved after the proof of the main proposition. For brevity, the proof is limited to the alive-BSC case, but can be extended to general alive-DMCs by choosing the two most distant letters in $\mathcal{Y}$ for constructing the repetition code, among other things.

Proposition 6.6. *For a channel that dies $(\mathcal{X}, p_a(y|x), p_d(y|x), p_T(t), \mathcal{Y})$, there is no $V > 0$ such that (N, V) is an achievable transmission time-volume at Shannon reliability.*

Proof. From the error probability point of view, transmitting longer codes is not harder than transmitting shorter codes (Lemma 6.1) and transmitting smaller codes is not harder than transmitting larger codes (Lemma 6.2). Hence, the desired result follows from showing that even the longest and smallest code that has positive expected transmission volume cannot achieve Shannon reliability.

Clearly the longest and smallest code uses a single individual message code of length $n_1 \to \infty$ and size $M_1 = 2$. Among such codes, transmitting the binary repetition code is not harder than transmitting any other code (Lemma 6.3). Hence showing that the binary repetition code cannot achieve Shannon reliability yields the desired result.

Consider transmitting a single $(M_1 = 2, n_1)$ individual message code that is simply a binary repetition code over a channel that dies $(\mathcal{X}, p_a(y|x), p_d(y|x), p_T(t), \mathcal{Y})$.

Let $\mathcal{W}_1 = \{00000\ldots, 11111\ldots\}$, where the two codewords are of length n_1. Assume that the all-zeros codeword and the all-ones codeword are each transmitted with probability 1/2 and measure average probability of error, since average error probability lower bounds $P_e^{\max}(1)$ [51, Problem 5.32]. The transmission time $N = n_1$ and let $N \to \infty$. The expected transmission volume is $\log 2 > 0$.

Under equiprobable signaling over a BSC, the minimum error probability decoder is the maximum likelihood decoder, which in turn is the minimum distance decoder [53, Problem 2.13].

The scenario corresponds to binary hypothesis testing over a BSC(ε) with T observations (since after the channel dies, the output symbols do not help with hypothesis testing). Let $\underline{\varepsilon} = \min(\varepsilon, 1-\varepsilon)$ and $\overline{\varepsilon} = \max(\varepsilon, 1-\varepsilon)$. Given the realization $T = t$, the error probability under minimum distance decoding is

$$P_e^{\text{maximax}}(t) = \sum_{i=\lceil t/2 \rceil}^{t} \binom{t}{i} \underline{\varepsilon}^i \overline{\varepsilon}^{t-i} > K(t) > 0, \tag{6.3}$$

where $K(t)$ is a fixed constant. Since $P_e^{\text{maximax}}(t) > K(t) > 0$ for any finite t, it follows that there is a fixed constant

$$K = \min_{t:\ t<\infty \text{ and } p_T(t)>0} K(t)$$

such that $P_e^{\text{maximax}}(t) > K > 0$ for any $p_T(t)$ that satisfies the property that there is a finite $t^\dagger$ such that $p_T(t^\dagger) > 0$.

Thus Shannon reliability is not achievable. □

Lemma 6.1. *When transmitting over the alive state's memoryless channel $p_a(y|x)$, let the maximal probability of error $P_e^{\max}(i)$ for an optimal (M_i, n_i) individual message code and minimum probability of error individual decoder $f_D^{(i)}$ be $P_e^{\max}(i; n_i)$. Then $P_e^{\max}(i; n_i + 1) \le P_e^{\max}(i; n_i)$.*

Proof. Consider the optimal block-length-n_i individual message code/decoder, which achieves $P_e^{\max}(i; n_i)$. Use it to construct an $n_i + 1$ individual message code that appends a dummy symbol to each codeword and an associated decoder that operates by ignoring this last symbol. The error performance of this (suboptimal) code/decoder is clearly $P_e^{\max}(i; n_i)$, and so the optimal performance can only be better: $P_e^{\max}(i; n_i + 1) \le P_e^{\max}(i; n_i)$. □

Lemma 6.2. *When transmitting over the alive state's memoryless channel $p_a(y|x)$, let the maximal probability of error $P_e^{\max}(i)$ for an optimal (M_i, n_i) individual message code and minimum probability of error individual decoder $f_D^{(i)}$ be $P_e^{\max}(i; M_i)$. Then $P_e^{\max}(i; M_i) \le P_e^{\max}(i; M_i + 1)$.*

Proof. Follows from sphere-packing principles. □

Lemma 6.3. *When transmitting over the alive state's memoryless channel $p_a(y|x)$, the optimal $(M_i = 2, n_i)$ individual message code can be taken as a binary repetition code.*

Proof. Under minimum distance decoding (which yields the minimum error probability [53, Problem 2.13]) for a code transmitted over a BSC, increasing the distance between codewords can only reduce error probability. The repetition code has maximum Hamming distance between codewords. □

Notice that Proposition 6.6 also directly implies Proposition 6.5.

6.3.2 Finite Block Length Channel Coding

Before developing an optimal scheme for η-reliable communication over a channel that dies, finite block length channel coding is reviewed.

Under the definitions above, traditional channel coding results [289–294] provide information about individual message codes, determining the achievable trios $(n_i, M_i, P_e^{\max}(i))$. In particular, the largest possible $\log M_i$ for a given n_i and $P_e^{\max}(i)$ is denoted $\log M^*(n_i, P_e^{\max}(i))$.

The purpose of this work is not to improve upper and lower bounds on finite block length channel coding, but to use existing results to study channels that die. In fact, for the sequel, simply assume that the function $\log M^*(n_i, P_e^{\max}(i))$ is known, as are codes/decoders that achieve this value. In principle, optimal individual message codes may be found through exhaustive search [290, 313]. Although algebraic notions of code quality do not directly imply error probability quality [314], perfect codes such as the Hamming or Golay codes may also be optimal in certain limited cases.

Omitting the $O(\log n)$ term in Theorem 2.1 yields an approximation to the function $\log M^*(n_i, P_e^{\max}(i))$, which is called Strassen's normal approximation [77]. Recent results comparing upper and lower bounds around Strassen's approximation have demonstrated that the approximation is quite good [292].

In the sequel, assume that optimal $\log M^*(n_i, \eta)$-achieving individual message codes are known. Exact upper and lower bounds to $\log M^*(n_i, \eta)$ can be substituted to make results precise. For numerical demonstrations, further assume that optimal codes have performance given by Strassen's approximation.

For the BSC(ε), Strassen's approximation is:

$$\log M^* \approx n_i(1 - h_2(\varepsilon)) - \sqrt{n_i \varepsilon (1-\varepsilon)} Q^{-1}(\eta) \log_2 \tfrac{\varepsilon}{1-\varepsilon}. \tag{6.4}$$

This BSC expression first appeared in [315].

To provide intuition, the approximate $\log M^*(n_i, \eta)$ function for a BSC(ε) is plotted in Figure 6-1. Notice that $\log M^*$ is zero for small n_i since no code can achieve the target error probability η. Also notice that $\log M^*$ is a monotonically increasing function of n_i. Moreover, notice in Figure 6-2 that even when normalized, the approximate $(\log M^*)/n_i$ is a monotonically increasing function of n_i. Therefore under Strassen's approximation, longer blocks provide more 'bang for the buck.' The curve in Figure 6-2 asymptotically approaches capacity.

6.3.3 η-reliable Communication

A coding scheme that achieves positive expected transmission volume at η-reliability is now described. Survival probability of the channel plays a key role in measuring performance.

Definition 6.14. *For a channel that dies* $(\mathcal{X}, p_a(y|x), p_d(y|x), p_T(t), \mathcal{Y})$*, the survival*

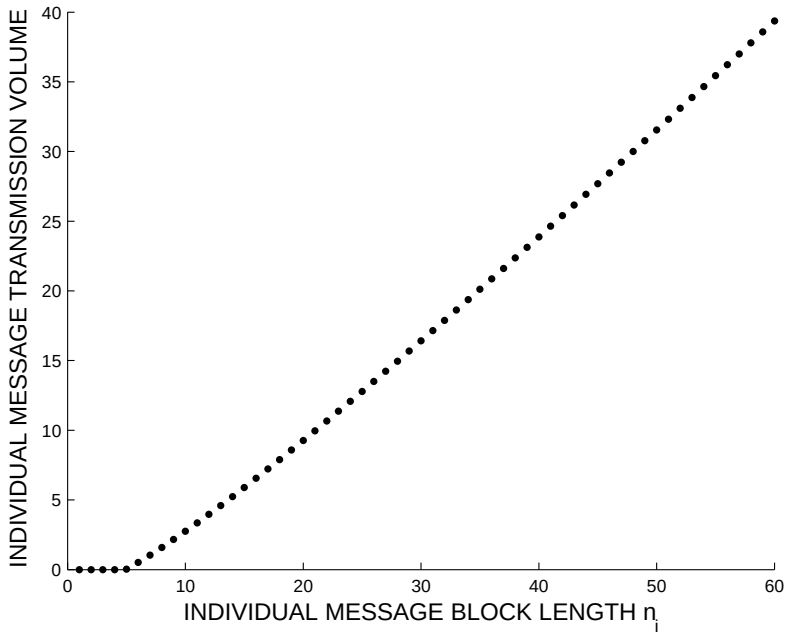

Figure 6-1. Maximal codebook size for given block length, the expression (6.4), for $\varepsilon = 0.01$ and $\eta = 0.001$.

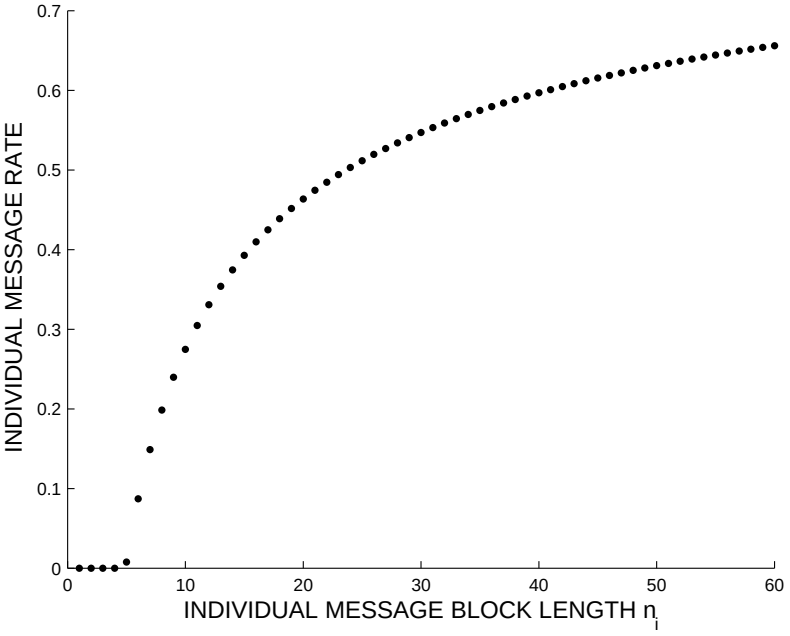

Figure 6-2. Normalized maximal codebook size for given block length, $(\log M^*(n_i, \eta))/n_i$, for $\varepsilon = 0.01$ and $\eta = 0.001$. The capacity of a BSC(ε) is $1 - h_2(\varepsilon) = 0.92$.

function is $\Pr[T > t]$*, is denoted* $R_T(t)$*, and satisfies*

$$R_T(t) = \Pr[T > t] = 1 - \sum_{\tau=1}^{t} p_T(\tau) = 1 - F_T(t),$$

where F_T *is the cumulative distribution function.*

$R_T(t)$ is a non-increasing function.

Proposition 6.7. *The transmission time-volume*

$$\left(N = \sum_{i=1}^{k} n_i, V = \sum_{i=1}^{k} R_T(e_i) \log M^*(n_i, \eta) \right)$$

is achievable at η*-reliability for any sequence* $(n_i)_{i=1}^k$ *of individual message codeword lengths, and* $e_0 = 0, e_1 = n_1, e_2 = n_1 + n_2, \ldots, e_k = \sum_{i=1}^k n_i$*.*

Proof.

Code Design: A target error probability η and a sequence $(n_i)_{i=1}^k$ of individual message codeword lengths are fixed. Construct a length-k sequence of (M_i, n_i) individual message codes and individual decoding functions $(\mathcal{W}_i, f_E^{(i)}, f_D^{(i)})$ that achieve optimal performance. The size of $\mathcal{W}_i$ is $|\mathcal{W}_i| = \log M^*(n_i, \eta)$. Note that individual decoding functions $f_D^{(i)}$ have range $\mathcal{W}_i$ rather than $\mathcal{W}_i^{\ominus}$.

Encoding: A codeword $W_1 = w_1$ is selected uniformly at random from the codebook $\mathcal{W}_1$. The mapping of this codeword into n_1 channel input letters, $X_{e_0+1}^{e_1} = f_1(w_1)$, is transmitted in channel usage times $n = e_0 + 1, e_0 + 2, \ldots, e_1$.

Then a codeword $W_2 = w_2$ is selected uniformly at random from the codebook $\mathcal{W}_2$. The mapping of this codeword into n_2 channel input letters, $X_{e_1+1}^{e_2} = f_2(w_2)$, is transmitted in channel usage times $n = e_1 + 1, e_1 + 2, \ldots, e_2$.

This procedure continues until the last individual message code in the code is transmitted. That is, a codeword $W_k = w_k$ is selected uniformly at random from the codebook $\mathcal{W}_k$. The mapping of this codeword into n_k channel input letters, $X_{e_{k-1}+1}^{e_k} = f_k(w_k)$, is transmitted in channel usage times $n = e_{k-1}+1, e_{k-1}+2, \ldots, e_k$.

Refer to channel usage times $n \in \{e_{i-1}+1, e_{i-1}+2, \ldots, e_i\}$ as the ith transmission epoch.

Decoding: For decoding, the channel output symbols for each epoch are processed separately. If any of the channel output symbols in an epoch are erasure symbols ?, then a decoding erasure $\ominus$ is declared for the message in that epoch, i.e. $\hat{W}_i = \ominus$. Otherwise, the individual message decoding function $f_D^{(i)} : \mathcal{Y}^{n_i} \to \mathcal{W}_i$ is applied to obtain $\hat{W}_i = f_D^{(i)}(Y_{e_{i-1}+1}^{e_i})$.

Performance Analysis: Having defined the communication scheme, the error probability, transmission time, and expected transmission volume are measured.

The decoder will either produce an erasure $\ominus$ or use an individual message decoder $f_D^{(i)}$. When $f_D^{(i)}$ is used, the maximal error probability of individual message code error is bounded as $P_e^{\max}(i) < \eta$ by construction. Since declared erasures $\ominus$ do not lead to

error, and since all $P_e^{\max}(i) < \eta$, it follows that

$$P_e^{\text{maximax}} < \eta.$$

The transmission time is simply $N = \sum n_i$.

Recall the definition of expected transmission volume:

$$\mathrm{E}\left\{ \sum_{i \in \{1,\ldots,k | \hat{W}_i \neq \ominus\}} \log M_i \right\} = \sum_{i \in \{1,\ldots,k | \hat{W}_i \neq \ominus\}} \mathrm{E}\left\{\log M_i\right\}$$

and the fact that the channel produces the erasure symbol ? for all channel usage times after death, $n > T$, but not before. Combining this with the length of an optimal code, $\log M^*(n_i, \eta)$, leads to the expression

$$\sum_{i=1}^{k} \Pr[T > e_i] \log M^*(n_i, \eta) = \sum_{i=1}^{k} R_T(e_i) \log M^*(n_i, \eta),$$

since all individual message codewords that are received in their entirety before the channel dies are decoded using $f_D^{(i)}$ whereas any individual message codewords that are even partially cut off are declared $\ominus$. □

Proposition 6.7 is valid for any choice of $(n_i)_{i=1}^k$. Since $(\log M^*)/n_i$ is approximately monotonically increasing, it is better to use individual message codes that are as long as possible. With longer individual message codes, however, there is a greater chance of many channel usages being wasted if the channel dies in the middle of transmission. The basic trade-off is captured in picking the set of values $\{n_1, n_2, \ldots, n_k\}$. For fixed and finite N, this involves picking an ordered integer partition $n_1 + n_2 + \cdots + n_k = N$. This choice is optimized in Section 6.4.

■ 6.3.4 Converse Arguments

Since there are simply operational expressions and no informational expressions in the development, and since optimal individual message codes and individual message decoders are assumed to be used, it may seem as though converse arguments are not required. This would indeed follow, if the following two things were true, which follow from Assumption 6.1. First, that there is no benefit in trying to decode the last partially erased message block. Second, that there is no benefit to errors-and-erasures decoding [64] by the $f_D^{(i)}$ for codewords that are received before channel death. Under Assumption 6.1, Proposition 6.7 gives the best performance possible.

One might wonder whether Assumption 6.1 is needed. That there would be no benefit in trying to decode the last partially erased block follows from the conjecture that an optimal individual message code would have no latent redundancy that could be exploited to achieve a $P_e^{\max}(i = \text{last}) < \eta$, but this is a property of the actual optimal code.

On the other hand the effect of errors-and-erasures decoding by the individual message decoders it unclear. As given in a Neyman-Pearson style result by Forney [64], decoding regions determined by likelihood ratio tests should be used to optimally trade (average) error probability for erasure probability by varying the threshold. It is unclear how the choice of threshold would affect the expected transmission volume

$$\sum_{i=1}^{k}(1-\xi_i)R_T(e_i)\log M^*(n_i,\xi_i,\eta),$$

where ξ_i would be the specified erasure probability for individual message i, and $M^*(n_i,\xi_i,\eta)$ would be the maximum individual message codebook size under erasure probability ξ_i and maximum error probability η. Error bounds for errors-and-erasures decoding [64, Theorem 2] can certainly be converted into bounds on $\log M^*(n_i,\xi_i,\eta)$. It is unknown, however, whether there is a good Strassen-style approximation to this quantity.

6.4 Optimizing the Communication Scheme

Section 6.3.3 had not optimized the lengths of the individual message codes, as done here. For fixed η and N, maximize the expected transmission volume V over the choice of the ordered integer partition $n_1+n_2+\cdots+n_k=N$:

$$\max_{(n_i)_{i=1}^k:\sum n_i=N}\sum_{i=1}^{k}R_T(e_i)\log M^*(n_i,\eta). \tag{6.5}$$

For finite N, this optimization can be carried out by an exhaustive search over all 2^{N-1} ordered integer partitions. If the death distribution $p_T(t)$ has finite support, there is no loss of generality in considering only finite N. Since exhaustive search has exponential complexity, however, there is value in trying to use a simplified algorithm. A dynamic programming formulation for the finite horizon case is developed in Section 6.4.3. The next subsection develops a greedy algorithm which is applicable to both the finite and infinite horizon cases and yields the optimal solution for certain problems.

6.4.1 A Greedy Algorithm

To try to solve the optimization problem (6.5), consider a greedy algorithm that optimizes block lengths n_i one by one.

Algorithm 6.1.

1. *Maximize $R_T(n_1)\log M^*(n_1,\eta)$ through the choice of n_1 independently of any other n_i.*

2. *Maximize $R_T(e_2)\log M^*(n_2,\eta)$ after fixing n_1, but independently of later n_i.*

3. *Maximize $R_T(e_3)\log M^*(n_3,\eta)$ after fixing e_2, but independently of later n_i.*

4. *Continue in the same manner for all subsequent n_i.*

Sometimes the algorithm produces the correct solution.

Proposition 6.8. *The solution produced by the greedy algorithm, (n_i), is locally optimal if*

$$\frac{R_T(e_i)\log M^*(n_i,\eta) - R_T(e_i-1)\log M^*(n_i-1,\eta)}{R_T(e_{i+1})\left[\log M^*(n_{i+1}+1,\eta) - \log M^*(n_{i+1},\eta)\right]} \geq 1 \tag{6.6}$$

for each i.

Proof. The solution of the greedy algorithm partitions time using a set of epoch boundaries (e_i). The proof proceeds by testing whether local perturbation of an arbitrary epoch boundary can improve performance. There are two possible perturbations: a shift to the left or a shift to the right.

First consider shifting an arbitrary epoch boundary e_i to the right by one. This makes the left epoch longer and the right epoch shorter. Lengthening the left epoch does not improve performance due to the greedy optimization of the algorithm. Shortening the right epoch does not improve performance since $R_T(e_i)$ remains unchanged whereas $\log M^*(n_i,\eta)$ does not increase since $\log M^*$ is a non-decreasing function of n_i.

Now consider shifting an arbitrary epoch boundary e_i to the left by one. This makes the left epoch shorter and the right epoch longer. Reducing the left epoch will not improve performance due to greediness, but enlarging the right epoch might improve performance, so the gain and loss must be balanced.

The loss in performance (a positive quantity) for the left epoch is

$$\Delta_l = R_T(e_i)\log M^*(n_i,\eta) - R_T(e_i-1)\log M^*(n_i-1,\eta)$$

whereas the gain in performance (a positive quantity) for the right epoch is

$$\Delta_r = R_T(e_{i+1})\left[\log M^*(n_{i+1}+1,\eta) - \log M^*(n_{i+1},\eta)\right].$$

If $\Delta_l \geq \Delta_r$, then perturbation will not improve performance. The condition may be rearranged as

$$\frac{R_T(e_i)\log M^*(n_i,\eta) - R_T(e_i-1)\log M^*(n_i-1,\eta)}{R_T(e_{i+1})\left[\log M^*(n_{i+1}+1,\eta) - \log M^*(n_{i+1},\eta)\right]} \geq 1$$

This is the condition (6.6), so the left-perturbation does not improve performance. Hence, the solution produced by the greedy algorithm is locally optimal. □

Proposition 6.9. *The solution produced by the greedy algorithm, (n_i), is globally optimal if*

$$\frac{R_T(e_i)\log M^*(n_i,\eta) - R_T(e_i-K_i)\log M^*(n_i-K_i,\eta)}{R_T(e_{i+1})\left[\log M^*(n_{i+1}+K_i,\eta) - \log M^*(n_{i+1},\eta)\right]} \geq 1 \tag{6.7}$$

for each i, and any non-negative integers $K_i \leq n_i$.

Proof. The result follows by repeating the argument for local optimality in Proposition 6.8 for shifts of any admissible size K_i. □

There is an easily checked special case of global optimality condition (6.7) under the Strassen approximation.

Lemma 6.4. *The function $\log M_S^*(z,\eta) - \log M_S^*(z-K,\eta)$ is a non-decreasing function of z for any K, where*

$$\log M_S^*(z,\eta) = zC - \sqrt{z\rho}Q^{-1}(\eta) \tag{6.8}$$

is Strassen's approximation.

Proof. Essentially follows from the fact that $\sqrt{z}$ is a concave $\cap$ function in z. More specifically $\sqrt{z}$ satisfies

$$-\sqrt{z} + \sqrt{z-K} \leq -\sqrt{z+1} + \sqrt{z+1-K}$$

for $K \leq z < \infty$. This implies:

$$-\sqrt{z}\sqrt{\rho}Q^{-1}(\eta) + \sqrt{z-K}\sqrt{\rho}Q^{-1}(\eta) \leq -\sqrt{z+1}\sqrt{\rho}Q^{-1}(\eta) + \sqrt{z+1-K}\sqrt{\rho}Q^{-1}(\eta).$$

Adding the positive constant KC to both sides, in the form $zC - zC + KC$ on the left and in the form $(z+1)C - (z+1)C + KC$ on the right yields

$$\begin{aligned} &zC - \sqrt{z\rho}Q^{-1}(\eta) - (z-K)C + \sqrt{z-K}\sqrt{\rho}Q^{-1}(\eta) \\ &\quad \leq (z+1)C - \sqrt{z+1}\sqrt{\rho}Q^{-1}(\eta) - (z+1-K)C + \sqrt{z+1-K}\sqrt{\rho}Q^{-1}(\eta) \end{aligned}$$

and so

$$\left[\log M_S^*(z,\eta) - \log M_S^*(z-K,\eta)\right] \leq \left[\log M_S^*(z+1,\eta) - \log M_S^*(z+1-K,\eta)\right].$$

□

Proposition 6.10. *If the solution produced by the greedy algorithm using Strassen's approximation* (6.8) *satisfies $n_1 \geq n_2 \geq \cdots \geq n_k$, then condition* (6.7) *for global optimality is satisfied.*

Proof. Since $R_T(\cdot)$ is a non-increasing survival function,

$$R_T(e_i - K) \geq R_T(e_{i+1}), \tag{6.9}$$

for the non-negative integer K. Since the function $[\log M_S^*(z,\eta) - \log M_S^*(z-K,\eta)]$ is a non-decreasing function of z by Lemma 6.4, and since the n_i are in non-increasing order,

$$\log M_S^*(n_i,\eta) - \log M_S^*(n_i-K,\eta) \geq \log M_S^*(n_{i+1}+K,\eta) - \log M_S^*(n_{i+1},\eta). \tag{6.10}$$

Taking products of (6.9) and (6.10), and rearranging yields the condition:

$$\frac{R_T(e_i - K)\left[\log M_S^*(n_i, \eta) - \log M_S^*(n_i - K, \eta)\right]}{R_T(e_{i+1})\left[\log M_S^*(n_{i+1} + K, \eta) - \log M_S^*(n_{i+1}, \eta)\right]} \geq 1.$$

Since $R_T(\cdot)$ is a non-increasing survival function,

$$R_T(e_i - K) \geq R_T(e_i) \geq R_T(e_{i+1}).$$

Therefore the global optimality condition (6.7) is also satisfied, by substituting $R_T(e_i)$ for $R_T(e_i - K)$ in one place. □

■ 6.4.2 Geometric Death Distribution

A common failure mode for systems that do not age is a geometric death time T [12]:

$$p_T(t) = \alpha(1 - \alpha)^{(t-1)},$$

and

$$R_T(t) = (1 - \alpha)^t,$$

where α is the death time parameter.

Proposition 6.11. *When T is geometric, then the solution to* (6.5) *under Strassen's approximation yields equal epoch sizes. This optimal size is given by*

$$\arg\max_{\nu} R_T(\nu) \log M^*(\nu, \eta).$$

Proof. Begin by showing that Algorithm 6.1 will produce a solution with equal epoch sizes. Recall that the survival function of a geometric random variable with parameter $0 < \alpha \leq 1$ is $R_T(t) = (1 - \alpha)^t$. Therefore the first step of the algorithm will choose n_1 as

$$n_1 = \arg\max_{\nu} (1 - \alpha)^{\nu} \log M^*(\nu, \eta).$$

The second step of the algorithm will choose

$$\begin{aligned} n_2 &= \arg\max_{\nu} (1 - \alpha)^{n_1} (1 - \alpha)^{\nu} \log M^*(\nu, \eta) \\ &= \arg\max_{\nu} (1 - \alpha)^{\nu} \log M^*(\nu, \eta), \end{aligned}$$

which is the same as n_1. In general,

$$\begin{aligned} n_i &= \arg\max_{\nu} (1 - \alpha)^{e_{i-1}} (1 - \alpha)^{\nu} \log M^*(\nu, \eta) \\ &= \arg\max_{\nu} (1 - \alpha)^{\nu} \log M^*(\nu, \eta), \end{aligned}$$

so $n_1 = n_2 = \cdots$.

Such a solution satisfies $n_1 \geq n_2 \geq \cdots$ and so it is optimal by Proposition 6.10. □

Notice that the geometric death time distribution forms a boundary case for Proposition 6.10.

The optimal epoch size for geometric death under Strassen's approximation can be found analytically. This is done for when the alive state corresponds to a BSC(ε). Let $C = 1 - h_2(\varepsilon)$, $K = \sqrt{\varepsilon(1-\varepsilon)}Q^{-1}(\eta)\log_2\frac{\varepsilon}{1-\varepsilon}$, $\bar{\alpha} = 1-\alpha$ and $\ell = \log\bar{\alpha}$. The goal is then to solve

$$\arg\max_{\nu} \bar{\alpha}^{\nu}\left[\nu C - \sqrt{\nu}K\right].$$

Applying the differentiation operator $\frac{d}{d\nu}$ to the expression yields:

$$\frac{d}{d\nu}\bar{\alpha}^{\nu}\left[\nu C - \sqrt{\nu}K\right] = \bar{\alpha}^{\nu}C + \bar{\alpha}^{\nu}\nu C\ell - \frac{\bar{\alpha}^{\nu}K}{2\sqrt{\nu}} - \bar{\alpha}^{\nu}\sqrt{\nu}K\ell.$$

Finding the appropriate root yields:

$$\begin{aligned}\nu^*_{\text{real}} =& \tfrac{-2C^2+K^2\ell}{3C^2\ell} \\ &- \frac{\left(-16C^4\ell^2+16C^2K^2\ell^3-16K^4\ell^4\right)}{\left(24C^2\ell^2\left(8C^6\ell^3+15C^4K^2\ell^4-12C^2K^4\ell^5+8K^6\ell^6+3\sqrt{3}\sqrt{16C^{10}K^2\ell^7-13C^8K^4\ell^8+8C^6K^6\ell^9}\right)^{1/3}\right)} \\ &+ \tfrac{1}{6C^2\ell^2}\left(8C^6\ell^3+15C^4K^2\ell^4-12C^2K^4\ell^5+8K^6\ell^6+3\sqrt{3}\sqrt{16C^{10}K^2\ell^7-13C^8K^4\ell^8+8C^6K^6\ell^9}\right)^{1/3}\end{aligned}$$

The solution ν^* is then given by choosing the best between $\lceil\nu^*_{\text{real}}\rceil$ and $\lfloor\nu^*_{\text{real}}\rfloor$ so as to meet the integer constraint. For fixed crossover ε and target error probability η, the solution is plotted as a function of α in Figure 6-3. The less likely the channel is to die early, the longer the optimal epoch length.

As an alternative computation, rather than fixing η, one might fix the number of bits to be communicated and find the best level of reliability that is possible. Figure 6-4 shows the best $P_e^{\text{maximax}} = \eta$ that is possible when communicating 5 bits over a BSC(ε)-geometric(α) channel that dies.

■ 6.4.3 Dynamic Programming

The greedy algorithm of Section 6.4.1 solves (6.5) under certain conditions. For finite N, a dynamic program (DP) may be used to solve (6.5) under any conditions. To develop the DP formulation [316], assume that channel state feedback (whether the channel output is ? or whether it is either 0 or 1) is available to the transmitter, however solving the DP will show that channel state feedback is not required.

System Dynamics:

$$\begin{bmatrix}\zeta_n \\ \omega_n\end{bmatrix} = \begin{bmatrix}(\zeta_{n-1}+1)\hat{s}_{n-1} \\ \omega_{n-1}\kappa_{n-1}\end{bmatrix}, \tag{6.11}$$

for $n = 1, 2, \ldots, N+1$. The following state variables, disturbances, and controls are used:

- $\zeta_n \in \mathbb{Z}^*$ is a state variable that counts the location in the current transmission epoch,

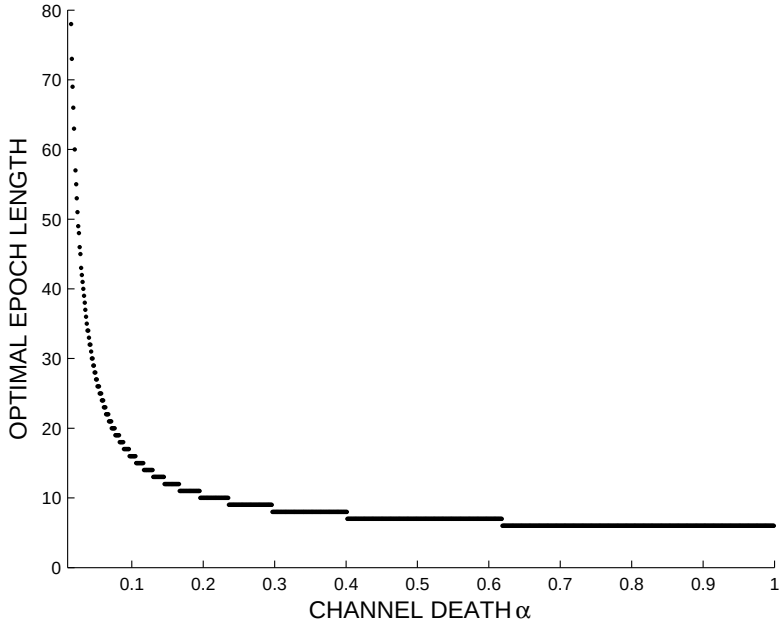

Figure 6-3. Optimal epoch lengths under Strassen's approximation for an (ε, α) BSC-geometric channel that dies for $\varepsilon = 0.01$ and $\eta = 0.001$.

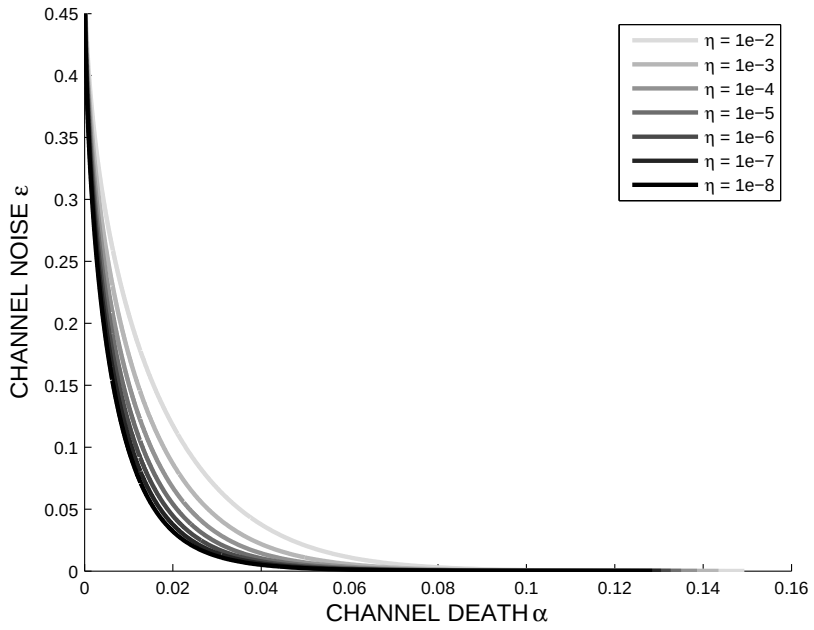

Figure 6-4. Achievable η-reliability in sending 5 bits over (ε, α) BSC-geometric channel that dies.

- $\omega_n \in \{0,1\}$ is a state variable that indicates whether the channel is alive (1) or dead (0),
- $\kappa_n \in \{0,1\} \sim \text{Bern}\left(R_T(n)\right)$ is a disturbance that kills (0) or preserves (1) the channel in the next time step, and
- $\hat{s}_n \in \{0,1\}$ is a control input that starts (0) or continues (1) a transmission epoch in the next time step.

Initial State: Since the channel starts alive (note that $R_T(1) = 1$) and since the first transmission epoch starts at the beginning of time,

$$\begin{bmatrix} \zeta_1 \\ \omega_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{6.12}$$

Additive Cost: Transmission volume $\log M^*(\zeta_n + 1, \eta)$ is credited if the channel is alive (i.e. $\omega_n = 1$) and the transmission epoch is to be restarted in the next time step (i.e. $1 - \hat{s}_n = 1$). This implies a cost function

$$c_n(\zeta_n, \omega_n, \hat{s}_n) = -(1 - \hat{s}_n)\omega_n \log M^*(\zeta_n + 1, \eta). \tag{6.13}$$

This is negative so that smaller is better.

Terminal Cost: There is no terminal cost, $c_{N+1} = 0$.

Cost-to-go: From time n to time $N + 1$ is:

$$\mathrm{E}_{\vec{\kappa}}\left\{\sum_{i=n}^{N} c_i(\zeta_i, \omega_i, \hat{s}_i)\right\} = -\,\mathrm{E}_{\vec{\kappa}}\left\{\sum_{i=n}^{N}(1 - \hat{s}_i)\omega_i \log M^*(\zeta_i + 1, \eta)\right\}.$$

Notice that the state variable ζ_n which counts epoch time is known to the transmitter and is determinable by the receiver through transmitter simulation. The state variable ω_n indicates the channel state and is known to the receiver by observing the channel output. It may be communicated to the transmitter through channel state feedback. It follows directly that:

Proposition 6.12. *A communication scheme that follows the dynamics* (6.11) *and additive cost* (6.13) *achieves the transmission time-volume*

$$\left(N, V = -\,\mathrm{E}\left[\sum_{n=1}^{N} c_n\right]\right)$$

at η-reliability.

DP may be used to find the optimal control policy $(\hat{s}_n)$.

Proposition 6.13. *The optimal $-V$ for the initial state* (6.12)*, dynamics* (6.11)*, additive cost* (6.13)*, and no terminal cost is equal to the cost of the solution produced by the dynamic programming algorithm.*

Proof. The system described by initial state (6.12), dynamics (6.11), and additive cost (6.13) is in the form of the *basic problem* [316, Sec. 1.2]. Thus the result follows from [316, Prop. 1.3.1] □

The DP optimization computations are now carried out; standard J notation[5] is used for cost [316]. The base case at time $N+1$ is

$$J_{N+1}(\zeta_{N+1}, \omega_{N+1}) = c_{N+1} = 0.$$

In proceeding backwards from time N to time 1:

$$J_n(\zeta_n, \omega_n) = \min_{\hat{s}_n \in \{0,1\}} \mathrm{E}_{\kappa_n} \left\{ c_n(\zeta_n, \omega_n, \hat{s}_n) + J_{n+1}\left(f_n(\zeta_n, \omega_n, \hat{s}_n, \kappa_n)\right) \right\},$$

for $n = 1, 2, \ldots, N$, where

$$f_n(\zeta_n, \omega_n, \hat{s}_n, \kappa_n) = \begin{bmatrix} \zeta_{n+1} \\ \omega_{n+1} \end{bmatrix} = \begin{bmatrix} (\zeta_n + 1)\hat{s}_n \\ \omega_n \kappa_n \end{bmatrix}.$$

Substituting the additive cost function yields:

$$\begin{aligned} J_n(\zeta_n, \omega_n) &= \min_{\hat{s}_n \in \{0,1\}} -\mathrm{E}_{\kappa_n} \left\{ (1 - \hat{s}_n)\omega_n \log M^*(\zeta_n + 1, \eta) \right\} + \mathrm{E}_{\kappa_n}\{J_{n+1}\} \qquad (6.14) \\ &= \min_{\hat{s}_n \in \{0,1\}} -(1 - \hat{s}_n) R_T(n) \log M^*(\zeta_n + 1, \eta) + \mathrm{E}_{\kappa_n}\{J_{n+1}\}. \end{aligned}$$

Notice that the state variable ω_n dropped out of the first term when the expectation operation was taken with respect to the disturbance κ_n. This is true for each stage in the DP.

Proposition 6.14. *For a channel that dies* $(\mathcal{X}, p_a(y|x), p_d(y|x), p_T(t), \mathcal{Y})$*, channel state feedback does not improve performance.*

Proof. By repeating the expectation calculation in (6.14) for each stage n in the stage-by-stage DP algorithm, it is verified that state variable ω does not enter into the stage optimization problem. Hence the transmitter does not require channel state feedback to determine the optimal signaling strategy. □

■ 6.4.4 A Dynamic Programming Example

A short example of applying dynamic programming optimization is provided to develop some intuition on the choice of epoch lengths. Consider the channel that dies with BSC($\varepsilon = 0.01$) alive state and $p_T(t)$ that is uniform over a finite horizon of length 40 (disallowing death in the first time step):

$$p_T(t) = \begin{cases} 1/39, & t = 2, \ldots, 40, \\ 0 & \text{otherwise.} \end{cases}$$

[5] Not to be confused with Bayes risk.

The goal is to communicate with η-reliability, $\eta = 0.001$.

Since the death distribution has finite support, there is no benefit to transmitting after death is guaranteed. Suppose some sequence of n_is is arbitrarily chosen: $(n_1 = 13, n_2 = 13, n_3 = 13, n_4 = 1)$. This has expected transmission volume (under the Strassen approximation)

$$\begin{aligned} V &= \sum_{i=1}^{4} R_T(e_i) \log M^*(n_i, \eta) \\ &\overset{(a)}{=} \log M^*(13, 0.001) \sum_{i=1}^{3} R_T(e_i) \\ &= \log M^*(13, 0.001)[R_T(13) + R_T(26) + R_T(39)] \\ &= 4.600[9/13 + 14/39 + 1/39] = 4.954 \text{ bits.} \end{aligned}$$

where (a) removes the fourth epoch since uncoded transmission cannot achieve η-reliability.

The result from optimizing the ordered integer partition (under the Strassen approximation) using the DP algorithm is $(n_1 = 20, n_2 = 12, n_3 = 6, n_4 = 2)$.[6] Notice that since the solution is in order, the greedy algorithm would also have succeeded. The expected transmission volume for this strategy (under the Strassen approximation) is

$$\begin{aligned} V &= R_T(20) \log M^*(20, 0.001) + R_T(32) \log M^*(12, 0.001) + R_T(38) \log M^*(6, 0.001) \\ &= (20/39) \cdot 9.2683 + (8/39) \cdot 3.9694 + (2/39) \cdot 0.5223 \\ &= 5.594 \text{ bits.} \end{aligned}$$

Optimization improves performance.

■ 6.4.5 A Precise Solution

It has been assumed that optimal finite block length codes are known and used. Moreover, the Strassen approximation has been used for certain computations. It is, however, also of interest to determine precisely which code should be used over a channel that dies. This subsection gives an example where a sequence of length-23 binary Golay codes [68] are optimal. Similar examples may be developed for other perfect codes (described in Chapter 2).

Before presenting the example, the sphere-packing upper bound on $\log M^*(n_i, \eta)$ for a BSC(ε) is derived. Recall the notion of decoding radius from Chapter 2 and let $\rho(\varepsilon, \eta)$ be the largest integer such that

$$\sum_{s=0}^{\rho} \binom{n_i}{s} \varepsilon^s (1-\varepsilon)^{n_i - s} \leq 1 - \eta.$$

[6] Equivalently $(n_1 = 20, n_2 = 12, n_3 = 6, n_4 = 1, n_5 = 1)$, since the last two channel usages are wasted (see Figure 6-1) to hedge against channel death.

The sphere-packing bound follows from counting how many decoding regions of radius ρ could conceivably fit in the Hamming space 2^{n_i} disjointly. Let $D_{s,m}$ be the number of channel output sequences that are decoded into message w_m and have distance s from the mth codeword. By the nature of Hamming space,

$$D_{s,m} \leq \binom{n_i}{s}$$

and due to the volume constraint,

$$\sum_{m=1}^{M} \sum_{s=0}^{\rho} D_{s,m} \leq 2^{n_i}.$$

Hence, the maximal codebook size $M^*(n_i, \eta)$ is upper-bounded as

$$\begin{aligned} M^*(n_i, \eta) &\leq \frac{2^{n_i}}{\sum_{s=0}^{\rho} D_{s,m}} \\ &\leq \frac{2^{n_i}}{\sum_{s=0}^{\rho(\varepsilon,\eta)} \binom{n_i}{s}}. \end{aligned}$$

Thus the sphere-packing upper bound on $\log M^*(n_i, \eta)$ is

$$\log M^*(n_i, \eta) \leq n_i - \log \left[\sum_{s=0}^{\rho(\varepsilon,\eta)} \binom{n_i}{s} \right] \triangleq \log M_{sp}(n_i, \eta).$$

As implied in Chapter 2, perfect codes such as the binary Golay code of length 23 can sometimes achieve the sphere-packing bound with equality.

Consider an (ε, α) BSC-geometric channel that dies, with $\varepsilon = 0.01$ and $\alpha = 0.05$. The target error probability is fixed at $\eta = 2.9 \times 10^{-6}$. For these values of ε and η, the decoding radius $\rho(\varepsilon, \eta) = 1$ for $2 \leq n_i \leq 3$. It is $\rho(\varepsilon, \eta) = 2$ for $4 \leq n_i \leq 10$; $\rho(\varepsilon, \eta) = 3$ for $11 \leq n_i \leq 23$; $\rho(\varepsilon, \eta) = 4$ for $24 \leq n_i \leq 40$; and so on.

Moreover, one can note that the $(n = 23, M = 4096)$ binary Golay code has a decoding radius of 3; thus it meets the BSC sphere-packing bound

$$M_{sp}(23, 2.9 \times 10^{-6}) = \frac{2^{23}}{1 + 23 + 253 + 1771} = 4096$$

with equality.

Now to bring channel death into the picture. If one proceeds greedily, following Algorithm 6.1, but using the sphere-packing bound $\log M_{sp}(n_i, \eta)$ rather than the optimal $\log M^*(n_i, \eta)$,

$$n_1(\varepsilon = 0.01, \alpha = 0.05, \eta = 2.9 \times 10^{-6}) = \arg\max_{\nu} \bar{\alpha}^{\nu} \log_2 \frac{2^{\nu}}{\sum_{s=0}^{\rho(\varepsilon,\eta)}} = 23.$$

By the memorylessness argument of Proposition 6.11, it follows that running Algorithm 6.1 with the sphere-packing bound will yield $23 = n_1 = n_2 = \cdots$.

It remains to show that Algorithm 6.1 actually gives the true solution. Had Strassen's approximation been used rather than the sphere-packing bound, the result would follow directly from Proposition 6.11. Instead, the global optimality condition (6.7) can be verified exhaustively for all 23 possible shift sizes K for the first epoch:

$$\frac{\bar{\alpha}^{23} \log M_{sp}(23, \eta) - \bar{\alpha}^{23-K} \log M_{sp}(23-K, \eta)}{\bar{\alpha}^{46} \log M_{sp}(23+K) - \bar{\alpha}^{46} \log M_{sp}(23, \eta)} \geq 1.$$

Then the same exhaustive verification is performed for all 23 possible shifts for the second epoch:

$$\begin{aligned}
\frac{\bar{\alpha}^{46} \log M_{sp}(23, \eta) - \bar{\alpha}^{46-K} \log M_{sp}(23-K, \eta)}{\bar{\alpha}^{69} \log M_{sp}(23+K) - \bar{\alpha}^{69} \log M_{sp}(23, \eta)} &\geq 1 \\
\frac{\bar{\alpha}^{23} \left[\bar{\alpha}^{23} \log M_{sp}(23, \eta) - \bar{\alpha}^{23-K} \log M_{sp}(23-K, \eta)\right]}{\bar{\alpha}^{23} \left[\bar{\alpha}^{46} \log M_{sp}(23+K) - \bar{\alpha}^{46} \log M_{sp}(23, \eta)\right]} &\geq 1 \\
\frac{\bar{\alpha}^{23} \log M_{sp}(23, \eta) - \bar{\alpha}^{23-K} \log M_{sp}(23-K, \eta)}{\bar{\alpha}^{46} \log M_{sp}(23+K) - \bar{\alpha}^{46} \log M_{sp}(23, \eta)} &\geq 1.
\end{aligned}$$

The exhaustive verification can be carried out indefinitely to show that using the length-23 binary Golay code for every epoch is optimal.

■ 6.5 Discussion

Channels that die arise in communication systems embedded in networks that may run out of energy [317], synthetic systems embedded in biological cells that may die [318], systems embedded in spacecraft that may enter black holes [319], or systems embedded in oceans with undersea cables that may be cut [320]. One might even consider communication systems where users may permanently lose attention [321].

Information-theoretic limits of communicating over channels that die, in the sense of maximizing expected transmission volume at a given level of error probability, were determined. The results presented in this chapter therefore provide insight into the fundamental limits of communication systems that might randomly undergo catastrophic failure.

Implementable optimization algorithms were also given, so as to find approximately or precisely optimal coding methods. These methods work with arbitrary death distributions, even empirically measured ones. Further, rather than considering the $\log M^*(n_i, \eta)$ function for optimal finite block length codes, the code optimization procedures would work just as well if a set of finite block length codes was provided. Such a limited set of codes might be selected for decoding complexity or other practical reasons.

Chapter 7

Operation Costs—Energy

Operating communication systems is costly, in part, because they consume energy. Energy constraints therefore limit the performance of extant decoding technologies [18, 19]. Moreover unreliability in decoding operation as discussed in Chapters 5 and 6 may be due to low-power operation [322] or energy exhaustion [317]. Mitigating decoding restrictions due to energy limits is therefore of paramount importance.

There are several approaches to countering energy limitations in decoding circuits. The most obvious is to make circuits more energy efficient. In fact, there is no fundamental thermodynamic reason for energy to be dissipated in the decoding process, since energy is not required to perform mathematical work [323, Chapter 5]. Decoders that are reversible computational devices would lie in this extreme regime and would not dissipate any energy [16,324]. Physical mechanisms proposed for reversible computing such as ballistic computers, externally clocked Brownian machines, and fully Brownian machines [325, 326], however require that the system have no faults (contrary to the model in Chapter 5), have infinite storage capability (contrary to the model in Chapter 4), and operate arbitrarily slowly (contrary to the goals in Chapter 3).

If energy resources are asymmetric within the communication system, such that the transmitter is relatively unconstrained whereas the receiver is severely constrained [327,328], an alternate approach to countering energy limitations presents itself. One might embed energy for decoding in the transmitted signal itself, allowing the receiver to harvest both energy and information simultaneously [329].

The earliest telegraph, telephone, and crystal radio receivers had no external power sources [44], providing historical examples of systems where energy for future information processing is included in the communication signal itself. Modern examples include retinal prostheses that receive power and data from light entering the eye [330], implanted brain-machine interfaces that receive configuration signals and energy through inductive coupling [331], RFID systems where the energy provided through the forward channel is used to transmit over the backward channel [332], and mudpulse telemetry systems in the oil industry where energy and information are provisioned to remote instruments over a single line [333].

With this method of mitigating decoding energy constraints in mind, it is of interest to determine the fundamental trade-off between the rates at which energy and reliable information can be transmitted over a single noisy line. A capacity-power function is defined here and an appropriate coding theorem is proven to endow

it with operational significance. In particular, it characterizes communication systems that simultaneously meet two goals:

1. large received energy per unit time (large received power), and
2. large information per unit time.

Notice that unlike traditional transmitter power constraints, where small transmitted power is desired, here large received power is desired. One previous study has looked at maximum received power constraints [334].

The remainder of the chapter is organized as follows. Section 7.1 reviews the energy that might be required for decoding, first considering reversible computing and then considering practical computing. Section 7.2 considers transmitting energy and information simultaneously, finding the capacity-power function. Properties of the capacity-power function are given in Section 7.3 and examples are computed in Section 7.4. Finally Section 7.5 further discusses the results of the chapter.

7.1 Energy Requirements for Decoding

This section reviews how much energy might be required to decode an error-control codeword observed through a noisy communication channel. Some might call it an examination of the energy complexity [8] of decoding.

7.1.1 Reversible Decoders

As described in Chapter 2, a message decoder computes an estimate of the transmitted message using the channel output sequence according to the probabilistic kernel $p_{\hat{W}_1^k|Y_1^n}(\cdot|\cdot)$. If the decoder is deterministic, it implements a decoding function $f_D : \mathcal{Y}^n \mapsto \hat{\mathcal{W}}^k$. A signal decoder computes an estimate of the transmitted signal according to $p_{\hat{X}_1^n|Y_1^n}(\cdot|\cdot)$ which reduces to $f_D : \mathcal{Y}^n \mapsto \mathcal{X}^n$ if deterministic.

The decoders are designed to dissipate the irrelevant uncertainty associated with channel noise, but in doing so, they erase the information that was in the noise signal. Since they are not informationally lossless (logically reversible), by Landauer's principle [16,335], they must dissipate energy irrespective of physical implementation. It can however be shown that any deterministic computation can be reprogrammed as a sequence of informationally lossless computation steps, if the computation is allowed to save a copy of its input [324, 326]. Thus, deterministic decoders $f_D(\cdot)$ can be made logically reversible.

When a logically irreversible operation is implemented using logically reversible gates like the Fredkin gate, extra inputs that are fixed constants might need to be supplied to the circuit, and the circuit might produce undesired garbage outputs which cannot be erased [326]. Indeed, storing garbage may require large amounts of extra memory.

It is generally accepted that any logically reversible informational transformation can, in principle, be performed by a physical mechanism that operates in a thermodynamically reversible manner. The physical mechanisms proposed for carrying out

reversible computing are rather strange and include such devices as ballistic computers, externally clocked Brownian machines, and fully Brownian machines. These physical mechanisms are thermodynamically reversible but require that there are no hardware errors [20,324]. Electronic circuits that are almost reversible have also been proposed [17].

In addition to requiring no computing errors, the Brownian models require operation in the limit of zero speed, so as to eliminate friction. Ballistic computers, on the other hand, can operate at non-zero speed and remain thermodynamically reversible but are much more sensitive to noise than Brownian computers [325].

Thus deterministic decoding can, in principle, be carried out without loss of energy if the decoder is not limited by memory constraints or unreliability.

7.1.2 Traditional Decoders

Unlike reversible decoders, energy is consumed by any practically realizable decoder. A typical power budget analysis for wireless communications systems is given in [336], and similarly for neural systems in [337, 338]. The specific energy consumption of decoding is investigated in [339].

Sahai and Grover have developed lower bounds to energy complexity of iterative message-passing decoders [19], which abstract and generalize the kind of decoder studied in Chapter 5. Their energy model assumes that each computational node consumes a fixed amount of energy per iteration; passing messages over wires does not require energy. Combining lower bounds on circuit complexity, similar to the upper bounds developed in Section 5.6, and lower bounds on error probability as a function of number of iterations for any code yields a lower bound on error probability as a function of decoding energy. It is shown that systems with iterative message-passing decoders may require an infinite number of iterations to achieve Shannon reliability and therefore infinite energy. Even more negative results are given for other decoding architectures.

One might perform the same kind of analysis relating error probability and decoder power consumption for still other kinds of decoding architectures. As an example, consider the consensus decoders of Chapter 3. For a given circuit topology, the error probability can be related to the number of iterations in a straightforward manner [340, Corollary 6]. If energy is consumed due to attenuation over lengthy wires, it can be characterized using the wiring length analysis of Chapter 3. Since the operations of the computational nodes are very simple, the energy consumed by them is also easily determined.

The fact that decoders do consume energy advocates the alternative strategy mentioned at the beginning of the chapter: sending energy along with information.

7.2 Transmitting Energy and Information Simultaneously

Recall that when sending energy along with information, a characterization of communication systems that simultaneously meet two goals:

1. large received power, and

2. large information rate

is to be found.

In order to achieve the first goal, one would want the most energetic symbol received all the time, whereas to achieve the second goal, one would want to use the unconstrained capacity-achieving input distribution. This intuition is formalized for discrete memoryless channels, as follows.

Recall from Chapter 2 that a DMC is characterized by the input alphabet $\mathcal{X}$, the output alphabet $\mathcal{Y}$, and the transition probability assignment $p_{Y|X}(y|x)$. Furthermore, each output letter $y \in \mathcal{Y}$ has an energy $b(y)$, a nonnegative real number. The n-fold extension of the energy function $b(\cdot)$ is also defined on $\mathcal{Y}^n$ for all n. Throughout the chapter, it is assumed that

$$b(y_1^n) = \sum_{i=1}^{n} b(y_i).$$

The average received energy is

$$\begin{aligned} E\left[b(Y_1^n)\right] &= \sum_{y_1^n \in \mathcal{Y}^n} b(y_1^n) p_{Y_1^n}(y_1^n) \\ &= \sum_{y_1^n \in \mathcal{Y}^n} b(y_1^n) \sum_{x_1^n \in \mathcal{X}^n} p_{X_1^n}(x_1^n) p_{Y_1^n|X_1^n}(y_1^n|x_1^n). \end{aligned} \tag{7.1}$$

When normalized by block length n, this is the average received power. When using a deterministic encoder $f_E : \mathcal{W} \mapsto \mathcal{X}^n$, the average received energy for a specific message w with codeword $x_1^n(w)$ is

$$E\left[b(Y_1^n(x_1^n(w)))\right] = \sum_{y_1^n \in \mathcal{Y}^n} b(y_1^n) p_{Y_1^n|X_1^n}(y_1^n|x_1^n(w)).$$

If all messages $w \in \mathcal{W}$ yield the same average received energy $E\left[b(Y_1^n(x_1^n(w)))\right]$, the overall average received energy $E\left[b(Y_1^n)\right]$ is also the same under any input distribution.

As in Chapter 2, the information rate of an (n, M) block code is $\frac{1}{n}\log M$. Here it must be chosen to achieve arbitrarily small maximum message error probability $P_e^{\max}$.

The goal is to find a reliable code that maximizes information rate under a minimum received power constraint; a formal definition is as follows.

Definition 7.1. *Given $0 < \epsilon \leq 1$, a non-negative number R is an ϵ-achievable rate for the channel $p_{Y|X}$ with constraint (b, B) if for every $\delta > 0$ and every sufficiently large n there exists an (n, M)-block code with maximum error probability $P_e^{\max} < \epsilon$ of rate exceeding $R - \delta$ for which $E[b(Y_1^n(x_1^n(w)))] < B$ for each message $w \in \mathcal{W}$. R is an achievable rate if it is ϵ-achievable for all $0 < \epsilon < 1$. The supremum of achievable rates is called the capacity of the channel under constraint (b, B) and is denoted $C_O(B)$.*

7.2.1 Time-sharing Approaches

One simple approach to transmitting energy and information simultaneously is to interleave the energy signal and the information-bearing signal. The encoder and decoder simultaneously commutate between the information source–information destination line and the energy source–energy destination line, both depicted in the extensive schematic diagram of a communication system (Figure 1-2). When in information mode, the energy in the signal is not exploited and when in energy mode, the information in the signal is ignored. This suboptimal time-sharing might arise due to limitations on receiver circuits [Y. Ramadass, personal communication].

If the system is in information mode for τ fraction of the time, it follows directly from Theorem 2.3 that the best a communication scheme could do is to achieve information rate of τC and received power $(1-\tau)B_{\max}$, where $B_{\max}$ is the maximum element of the vector $b^T p_{Y|X}$ computed from the vector-matrix product of the column vector of the $b(y)$ denoted b and the channel transition matrix $p_{Y|X}$.

In the setting where the interleaving schedule cannot be coordinated between the encoder and decoder, the decoder might randomly switch between the information destination and the energy destination. This would cause random puncturing of the code.[1] Randomly puncturing a capacity-achieving random code just yields a shorter capacity-achieving random code. Thus, under these conditions, an information rate of τC and received power of $(1-\tau)B_{p_Y^*}$ is achievable, where

$$B_{p_Y^*} = \sum_{y \in \mathcal{Y}} b(y) \sum_{x \in \mathcal{X}} p_X^*(x) p_{Y|X}(y|x)$$

is the power under the capacity-achieving input distribution p_X^*.

Rather than random puncturing, one might also use controlled puncturing [341, Chapter 2]. This would involve exploiting the information in the transmitted signal from the beginning of time until the codeword is decoded to the desired error probability, and then switching to harvesting energy afterward. Analysis of such sequential decoding procedures [342] is beyond the scope of this thesis.

7.2.2 Optimizing Energy and Information Transmission

A receiver circuit that simultaneously exploits all of the energy and information in the signal may perform better than one that is restricted to time-sharing. An optimization problem that precisely captures the trade-off between energy and information for a general receiver is developed as follows. The goal is set to maximize information rate under a minimum received power constraint.

For each n, the nth capacity-power function $C_n(B)$ of the channel is defined as

$$C_n(B) = \max_{X_1^n : E\left[b(Y_1^n)\right] \geq nB} I(X_1^n; Y_1^n).$$

An input random vector X_1^n is a test source; one that satisfies $E\left[b(Y_1^n)\right] \geq nB$ is

[1] Recall that the notion of puncturing a code was described in Chapter 2.

B-admissible. The maximization is over all n-dimensional B-admissible test sources. The set of probability distributions $p_{X_1^n}(x_1^n)$ corresponding to the set of B-admissible test sources is a closed subset of $\mathbb{R}^{|\mathcal{X}|^n}$ and is bounded since $\sum p(x_1^n) = 1$. Since the set is closed and bounded, it is compact. Mutual information is a continuous function of the input distribution and since continuous, real-valued functions defined on compact subsets of metric spaces achieve their supremums (see Theorem 7.13), defining the optimization as a maximum is not problematic. The nth capacity-power functions are only defined for $0 \leq B \leq B_{\max}$.

The capacity-power function of the channel is defined as

$$C(B) = \sup_n \frac{1}{n} C_n(B). \tag{7.2}$$

A coding theorem can be proven that endows this informational definition with operational significance.

Theorem 7.1. *$C_O(B) = C(B)$.*

Proof. Follows by reversing the output constraint inequality in the solution to [343, P20 on p. 117], which uses a maximal code argument. See also [334], which suggests a random coding argument for the proof. □

Note that Theorem 7.1 can be generalized to settings where the input and output alphabets $\mathcal{X}$ and $\mathcal{Y}$ are continuous, using the quantization-based techniques of [51, Section 7.3].

■ 7.3 Properties of the Capacity-Power Function

The coding theorem provided operational significance to the capacity-power function. Some properties of this function may also be developed.

It is immediate that $C_n(B)$ is non-increasing, since the feasible set in the optimization becomes smaller as B increases. The function also has convexity properties.

Theorem 7.2. *$C_n(B)$ is a concave function of B for $0 \leq B \leq B_{\max}$.*

Proof. Let $\alpha_1, \alpha_2 \geq 0$ with $\alpha_1 + \alpha_2 = 1$. The inequality to be proven is that for $0 \leq B_1, B_2 \leq B_{\max}$,

$$C_n(\alpha_1 B_1 + \alpha_2 B_2) \geq \alpha_1 C_n(B_1) + \alpha_2 C_n(B_2).$$

Let $X^{(1)}$ and $X^{(2)}$ be n-dimensional test sources distributed according to $p^{(1)}(x_1^n)$ and $p^{(2)}(x_1^n)$ that achieve $C_n(B_1)$ and $C_n(B_2)$ respectively. Denote the corresponding channel outputs as $Y^{(1)}$ and $Y^{(2)}$. By definition, $E[b(Y^{(i)})] \geq nB_i$ and $I(X^{(i)}; Y^{(i)}) = C_n(B_i)$ for $i = 1, 2$. Define another source X distributed according to $p(x_1^n) = \alpha_1 p^{(1)}(x_1^n) + \alpha_2 p^{(2)}(x_1^n)$ with corresponding output Y. The various input distributions $p^{(1)}$, $p^{(2)}$, and p are written as column vectors, as is the vector of energies b.

Recall that the channel transition assignment is written as the matrix $p_{Y|X}$. Then

$$\begin{aligned} E[b(Y)] &= b^T p_{Y|X} p = b^T p_{Y|X}[\alpha_1 p^{(1)} + \alpha_2 p^{(2)}] \\ &= \alpha_1 b^T p_{Y|X} p^{(1)} + \alpha_2 b^T p_{Y|X} p^{(2)} \\ &= \alpha_1 E[b(Y^{(1)})] + \alpha_2 E[b(Y^{(2)})] \\ &\geq n(\alpha_1 B_1 + \alpha_2 B_2), \end{aligned} \tag{7.3}$$

where b and $p_{Y|X}$ have been suitably extended. Thus, X is $(\alpha_1 B_1 + \alpha_2 B_2)$-admissible. Now, by definition of $C_n(\cdot)$, $I(X;Y) \leq C_n(\alpha_1 B_1 + \alpha_2 B_2)$. However, since $I(X;Y)$ is a concave function of the input probability,

$$\begin{aligned} I(X;Y) &\geq \alpha_1 I(X^{(1)};Y^{(1)}) + \alpha_2 I(X^{(2)};Y^{(2)}) \\ &= \alpha_1 C_n(B_1) + \alpha_2 C_n(B_2). \end{aligned}$$

Linking the two inequalities yields the desired result:

$$C_n(\alpha_1 B_1 + \alpha_2 B_2) \geq I(X;Y) \geq \alpha_1 C_n(B_1) + \alpha_2 C_n(B_2).$$

□

The concavity property of the capacity-power functions demonstrates that time-sharing approaches are never better than full exploitation of both energy and information.

A single-letterization property also holds.

Theorem 7.3. *For any DMC, $C_n(B) = nC_1(B)$ for all $n = 1, 2, \ldots$ and $0 \leq nB \leq nB_{\max}$.*

Proof. Let X_1^n be a B-admissible test source with corresponding output Y_1^n that achieves $C_n(B)$, so $E[b(Y)] \geq nB$ and $I(X_1^n;Y_1^n) = C_n(B)$. Since the channel is memoryless, $I(X_1^n;Y_1^n) \leq \sum_{i=1}^n I(X_i;Y_i)$. Let $B_i = E[b(Y_i)]$, then $\sum_{i=1}^n B_i = \sum_{i=1}^n E[b(Y_i)] = E[b(Y)] \geq nB$. By the definition of $C_1(B_i)$, $I(X_i;Y_i) \leq C_1(B_i)$. Now since $C_1(B)$ is a concave function of B, by Jensen's inequality,

$$\frac{1}{n}\sum_{i=1}^n C_1(B_i) \leq C_1\left(\frac{1}{n}\sum_{i=1}^n B_i\right) = C_1\left(\frac{1}{n}E[b(Y)]\right).$$

But since $\frac{1}{n}E[b(Y)] \geq B$ and $C_1(B)$ is a non-increasing function of B,

$$\frac{1}{n}\sum_{i=1}^n C_1(B_i) \leq C_1\left(\frac{1}{n}E[b(Y)]\right) \leq C_1(B),$$

that is,

$$\sum_{i=1}^n C_1(B_i) \leq nC_1(B).$$

Combining yields $C_n(B) \leq nC_1(B)$.

For the reverse, let X be a scalar test source with corresponding output Y that achieves $C_1(B)$. That is, $E[b(Y)] \geq B$ and $I(X;Y) = C_1(B)$. Now let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables drawn according to p_X with corresponding outputs $Y_1, \ldots, Y_n$. Then

$$E[b(Y_1^n)] = \sum_{i=1}^{n} E[b(Y_i)] \geq nB.$$

Moreover by memorylessness,

$$I(X_1^n; Y_1^n) = \sum_{i=1}^{n} I(X_i; Y_i) = nC_1(B).$$

Thus, $C_n(B) \geq nC_1(B)$. Since $C_n(B) \geq nC_1(B)$ and $C_n(B) \leq nC_1(B)$, $C_n(B) = nC_1(B)$. □

The theorem implies that single-letterization, $C(B) = C_1(B)$, is valid to establish the fundamental trade-off between energy and information.

7.4 Some Optimal Trade-offs

It is instructive to compute the capacity-power function for several channels, thereby establishing the optimal trade-offs in transmitting energy and information over them. Closed form expressions of capacity-power for some binary channels are given first, followed by properties and expressions for Gaussian channels.

7.4.1 Binary Channels

Here, three binary channels with output alphabet energy function $b(0) = 0$ and $b(1) = 1$ are considered. Such an energy function corresponds to discrete particles and packets, among other commodities.

Consider a noiseless binary channel. The optimization problem is solved by the maximum entropy method, hence the capacity-achieving input distribution is in the exponential family of distributions. It is easy to show that the capacity-power function is

$$C(B) = \begin{cases} \log(2), & 0 \leq B \leq \frac{1}{2} \\ h_2(B), & \frac{1}{2} \leq B \leq 1. \end{cases}$$

The capacity-power functions for other discrete noiseless channels are similarly easy to work out using maximum entropy methods.

Consider a binary symmetric channel with crossover probability $\varepsilon \leq 1/2$. It can be shown that the capacity-power function is

$$C(B) = \begin{cases} \log(2) - h_2(\varepsilon), & 0 \leq B \leq \frac{1}{2} \\ h_2(B) - h_2(\varepsilon), & \frac{1}{2} \leq B \leq 1 - \varepsilon, \end{cases}$$

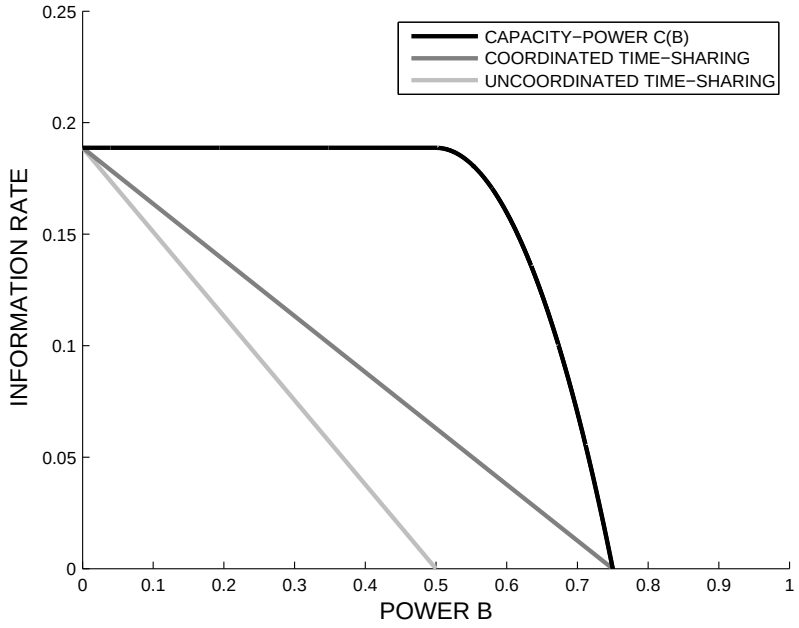

Figure 7-1. Capacity-power function for a binary symmetric channel with crossover probability $1/4$ and $b(0) = 0, b(1) = 1$. Performances of time-sharing schemes are shown for comparison.

as shown in Figure 7-1, where several properties of general capacity-power functions are evident. Recall that for the unconstrained problem, equiprobable inputs are capacity-achieving, which yield output power $\frac{1}{2}$. For $B > \frac{1}{2}$, the distribution must be perturbed so that the symbol 1 is transmitted more frequently. The maximum power receivable through this channel is $1-\varepsilon$, when 1 is always transmitted. The information rates and received power levels achievable through coordinated and uncoordinated time-sharing approaches are shown for comparison.

A third worked example is the Z-channel. It has a transmission matrix

$$p_{Y|X} = \begin{bmatrix} 1 & 0 \\ \varepsilon & 1-\varepsilon \end{bmatrix}$$

with 1 to 0 crossover probability ε. Its unconstrained capacity expression and associated capacity-achieving input distribution are as follows [344].

$$C(0) = \log\left(1 - \varepsilon^{\frac{1}{1-\varepsilon}} + \varepsilon^{\frac{\varepsilon}{1-\varepsilon}}\right),$$

achieved using input distribution with probability of the 1 symbol

$$\pi^* = \frac{\varepsilon^{\frac{\varepsilon}{1-\varepsilon}}}{1 + (1-\varepsilon)\varepsilon^{\frac{\varepsilon}{1-\varepsilon}}}.$$

Using this result and the need to switch over to using more the more energetic symbol

for more stringent power constraints, it can be shown that the capacity-power function is

$$C(B) = \begin{cases} \log\left(1 - \varepsilon^{\frac{1}{1-\varepsilon}} + \varepsilon^{\frac{\varepsilon}{1-\varepsilon}}\right), 0 \le B \le (1-\varepsilon)\pi^* \\ h_2(B) - \frac{B}{1-\varepsilon}h_2(\varepsilon), (1-\varepsilon)\pi^* \le B \le 1-\varepsilon. \end{cases}$$

A Z-channel models quantal synaptic failure [345] and other "stochastic leaky pipes" where the commodity may be lost en route.

■ 7.4.2 A Gaussian Channel

Now consider a memoryless, discrete-time AWGN channel. Continuous additive noise systems have the interesting property that for the goal of received power, noise power is actually helpful, whereas for the goal of information, noise power is hurtful. In discrete-alphabet channels, such an interpretation is not obvious.

Rather than working with the output power constraint directly, it is convenient to think of the output energy function $b(y)$ as inducing costs on the input alphabet $\mathcal{X}$:

$$\rho(x) = \int p_{Y|X}(y|x)b(y)dy.$$

By construction, this cost function preserves the constraint:

$$\begin{aligned} E[\rho(X)] &= \int \rho(x)dF_X(x) = \int dF_X(x) \int p_{Y|X}(y|x)b(y)dy \\ &= \int\int p_{Y|X}(y|x)b(y)dF_X(x)dy \\ &= E[b(Y)], \end{aligned}$$

where $F_X(x)$ is the input cumulative distribution function. Basically, $\rho(x)$ is the expected output energy provided by input letter x. For the AWGN channel $\mathcal{N}(0, \sigma_N^2)$ with $b(y) = y^2$,

$$\rho(x) = \int_{-\infty}^{\infty} \frac{y^2}{\sigma_N\sqrt{2\pi}} \exp\left\{-\frac{(y-x)^2}{2\sigma_N^2}\right\} dy = x^2 + \sigma_N^2.$$

That is, the output power is just the sum of the input power and the noise power.

Since the coding theorem, Theorem 7.1, and the single-letterization theorem, Theorem 7.3, extend directly to memoryless, continuous alphabet channels,

$$C(B) = \sup_{X:E[\rho(X)]\ge B} I(X;Y). \tag{7.4}$$

When working with real-valued alphabets, some sort of transmitter constraint must be imposed so as to disallow arbitrarily powerful signals. Hard amplitude constraints that model both rail limitations in power circuits and regulatory restrictions [346, Chapter 4] are suitable.

In particular, consider the AWGN channel, $\mathcal{N}(0, \sigma_N^2)$, with input alphabet $\mathcal{X} =$

$[-A, A] \subset \mathbb{R}$, and energy function $b(y) = y^2$. Denote the capacity-power function as $C(B; A)$.

The channel input cumulative distribution function is denoted $F_X(x)$, the channel transition probability density is denoted $p_{Y|X}(y|x)$, and the channel output density under input F (which always exists) is denoted

$$p_Y(y; F) = \int p_{Y|X}(y|x) dF(x).$$

Two informational functionals that arise are

$$i(x; F) = \int p_{Y|X}(y|x) \log \frac{p_{Y|X}(y|x)}{p_Y(y; F)} dy,$$

which is variously known as the marginal information density [347], the Bayesian surprise [348], or without name [349, Eq. 1]; and

$$h(x; F) = -\int p_{Y|X}(y|x) \log p_Y(y; F) dy,$$

which is known as the marginal entropy density.

Following lockstep with Smith [347, 350], it is shown that the capacity-power achieving input distribution consists of a finite number of mass points.

Let $\mathcal{F}_A$ be the set of input probability distribution functions $F_X(x)$ having all points of increase on the finite interval $[-A, A]$. This space has the following properties:

Lemma 7.1. *$\mathcal{F}_A$ is convex and compact in the Lévy metric.*

Proof. Let $F_1(x)$ and $F_2(x)$ be arbitrary elements of $\mathcal{F}_A$ and $\theta \in [0, 1]$. Define $F_\theta(x) = (1 - \theta)F_1(x) + \theta F_2(x)$. Both F_1 and F_2 are non-decreasing, right-continuous real functions that are zero for all $x < -A$ and one for all $x > A$. Clearly all of these properties are preserved in $F_\theta(x)$ and so $\mathcal{F}_A$ is convex.

The Lévy distance is a metric on the set of cumulative distribution functions [351, 3(b) on p. 215].

Helly's weak compactness theorem [351, p. 179] guarantees that every sequence of distribution functions has a weakly convergent subsequence and so $\mathcal{F}_A$ is weakly compact. It can further be shown that weak convergence is equivalent to complete convergence when restricting to distribution functions supported on a finite interval [351, Section 11.2] and so $\mathcal{F}_A$ is also completely compact. Moreover, complete convergence is equivalent to convergence in the Lévy distance [351, 3(c) on p. 215] and so $\mathcal{F}_A$ is also compact in the Lévy distance. □

Since the channel is fixed, the mutual information $I(X; Y) = I(F_X, p_{Y|X})$ can be written as a function of only the input distribution, $I(F_X)$. Although not always true [352, XVII], the mutual information functional has the following properties for an AWGN channel.

Lemma 7.2. *Mutual information $I : \mathcal{F}_A \to \mathbb{R}$ is a strictly concave, continuous, weakly differentiable functional with weak derivative*

$$I'_{F_1}(F_2) = \int_{-A}^{A} i(x; F_1) dF_2(x) - I(F_1).$$

Proof. Concavity of mutual information is shown in [350, Lemma on p. 25] by generalizing the convexity property of entropy. Strict concavity is shown by excluding the possibility of equality. The equality condition would only arise if the output densities $p(y; F_1)$ and $p(y; \theta F_1 + (1-\theta)F_2)$ were equal pointwise for some distinct F_1 and F_2, but due to the smoothness of the Gaussian noise, $p(y; F_1) = p(y; \theta F_1 + (1-\theta)F_2)$ pointwise implies that the Lévy distance between F_1 and $\theta F_1 + (1-\theta)F_2$ is zero.

Continuity of mutual information is shown in [350, Lemma on p. 27] and follows essentially from the Helly-Bray theorem [351, p. 182].

Weak differentiability of mutual information is shown in [350, Lemma on p. 29]. □

The average squared value of the channel input under the input distribution F_X is denoted as

$$\sigma_F^2 \triangleq \int_{-A}^{A} x^2 dF_X(x) = E[x^2].$$

Recall the received power constraint

$$B \leq E[\rho(X)] = E[\sigma_N^2 + x^2] = \sigma_N^2 + \sigma_F^2,$$

which is equivalent to $B - \sigma_N^2 - \sigma_F^2 \leq 0$. Now define the functional $J : \mathcal{F}_A \to \mathbb{R}$ as

$$J(F_X) \triangleq B - \sigma_N^2 - \int_{-A}^{A} x^2 dF_X(x).$$

Lemma 7.3. *J is a concave, continuous, weakly differentiable functional with weak derivative*

$$J'_{F_1}(F_2) = J(F_2) - J(F_1).$$

Proof. Clearly J is linear in F_X (see (7.3) for basic argument). Moreover, J is bounded as $B - \sigma_N^2 - A^2 \leq J \leq B - \sigma_N^2$. Since J is linear and bounded, it is concave, continuous, and weakly differentiable. The weak derivative for a linear functional is simply the difference between the functional values of the endpoints. □

Now return to the information-theoretic optimization problem to be solved. Two optimization theorems establishing properties of Lagrangian optimization and convex optimization, given in Appendix 7.A, will be used.

Theorem 7.4. *There exists a Lagrange multiplier $\lambda \geq 0$ such that*

$$C(B; A) = \sup_{F_X \in \mathcal{F}_A} [I(F_X) - \lambda J(F_X)].$$

Proof. The result follows from Theorem 7.11 in the appendix, since I is a concave functional (Lemma 7.2), since J is a concave functional (Lemma 7.3), since capacity is finite whenever $A < \infty$ and $\sigma_N^2 > 0$, and since there is obviously an $F_X \in \mathcal{F}_A$ such that $J(F_X) < 0$. □

Theorem 7.5. *There exists a unique capacity-energy achieving input X_0 with distribution function F_0 such that*

$$C(B;A) = \max_{F_X \in \mathcal{F}_A} [I(F_X) - \lambda J(F_X)] = I(F_0) - \lambda J(F_0),$$

with constant $\lambda \geq 0$. Moreover, a necessary and sufficient condition for F_0 to achieve capacity-power is

$$I'_{F_0}(F_X) - \lambda J'_{F_0}(F_X) \leq 0 \text{ for all } F_X \in \mathcal{F}_A. \tag{7.5}$$

Proof. Since I and J are both continuous and weakly differentiable (Lemmas 7.2, 7.3), so is $I - \lambda J$. Since I is strictly concave (Lemma 7.2) and J is concave (7.3), $I - \lambda J$ is strictly concave. Furthermore $\mathcal{F}_A$ is a convex, compact space (Lemma 7.1). Therefore Theorem 7.13 in the appendix applies and yields the desired result. □

For this function $I - \lambda J$, the optimality condition (7.5) may be rewritten as

$$\int_{-A}^{A} [i(x;F_0) + \lambda x^2] dF_X(x) \leq I(F_0) + \lambda \int x^2 dF_0(x)$$

for all $F_X \in \mathcal{F}_A$, using the known weak derivative expressions. If $\int x^2 dF_0(x) > B - \sigma_N^2$, then the moment constraint is trivial and the constant λ is zero, thus the optimality condition can be written as

$$\int_{-A}^{A} [i(x;F_0) + \lambda x^2] dF_X(x) \leq I(F_0) + \lambda[B - \sigma_N^2]. \tag{7.6}$$

The optimality condition (7.6) may be rejiggered to a condition on the input alphabet.

Theorem 7.6. *Let F_0 be an arbitrary distribution function in $\mathcal{F}_A$ satisfying the received power constraint. Let E_0 denote the points of increase of F_0 on $[-A, A]$. Then F_0 is optimal if and only if, for some $\lambda \geq 0$,*

$$\begin{aligned} &i(x;F_0) \leq I(F_0) + \lambda[B - \sigma_N^2 - x^2] \text{ for all } x \in [-A, A], \\ &i(x;F_0) = I(F_0) + \lambda[B - \sigma_N^2 - x^2] \text{ for all } x \in E_0. \end{aligned} \tag{7.7}$$

Proof. If both conditions (7.7) hold for some $\lambda \geq 0$, F_0 must be optimal and λ is the one from Theorem 7.5. This is because integrating both sides of the conditions (7.7) by an arbitrary F yields satisfaction of condition (7.5).

For the converse, assume that F_0 is optimal but that the inequality condition is not satisfied. Then there is some $x_1 \in [-A, A]$ and some $\lambda \geq 0$ such that $i(x_1;F_0) >$

$I(F_0) + \lambda[B - \sigma_N^2 - x_1^2]$. Let $F_1(x)$ be the unit step $\mathbf{1}(x - x_1) \in \mathcal{F}_A$; but then

$$\int_{-A}^{A} [i(x; F_0) + \lambda x^2] dF_1(x) = i(x_1; F_0) + \lambda x_1^2 > I(F_0) + \lambda[B - \sigma_N^2].$$

This violates (7.6), thus the inequality condition must be valid with λ from Theorem 7.5.

Now assume that F_0 is optimal but that the equality condition is not satisfied, i.e. there is set $E^\dagger \subset E_0$ such that the following is true:

$$\int_{E^\dagger} dF_0(x) = \delta > 0 \text{ and } \int_{E_0 - E^\dagger} dF_0(x) = 1 - \delta,$$

$$i(x; F_0) + \lambda x^2 < I(F_0) + \lambda[B - \sigma_N^2] \text{ for all } x \in E^\dagger,$$

and

$$i(x; F_0) + \lambda x^2 = I(F_0) + \lambda[B - \sigma_N^2] \text{ for all } x \in E_0 - E^\dagger.$$

Then,

$$\begin{aligned} 0 &= \int_{E_0} [i(x; F_0) + \lambda x^2] dF_0(x) - I(F_0) - \lambda[B - \sigma_N^2] \\ &= \int_{E^\dagger} [i(x; F_0) + \lambda x^2] dF_0(x) + \int_{E_0 - E^\dagger} [i(x; F_0) + \lambda x^2] dF_0(x) - I(F_0) - \lambda[B - \sigma_N^2] \\ &< \delta[I(F_0) + \lambda(B - \sigma_N^2)] + (1 - \delta)[I(F_0) + \lambda(B - \sigma_N^2)] - I(F_0) - \lambda[B - \sigma_N^2] = 0, \end{aligned}$$

a contradiction. Thus the equality condition must be valid. □

At a point like this in the development with the appropriate Karush-Kuhn-Tucker (KKT) conditions evident, one might try to develop measure-matching conditions like Gastpar et al. [349] for undetermined $b(\cdot)$, but this path is not pursued here.

To show that the input distribution is supported on a finite number of mass points requires Smith's *reductio ab absurdum* argument. The proof uses optimality conditions from Theorem 7.6 to derive a contradiction using the analytic extension property of the marginal entropy density $h(x; F)$.

Two facts from analysis will be needed:

Theorem 7.7 (Bolzano-Weierstrass Theorem [353, 8.13 on p. 76])**.** *Every bounded infinite subset of $\mathbb{R}^n$ has a limit point.*

Theorem 7.8 (Identity Theorem for Analytic Functions [354, p. 87])**.** *If two functions are analytic in a region $\mathcal{R}$ and if they coincide in a neighborhood of a point $z_0 \in \mathcal{R}$, or coincide along a path segment terminating in z_0, or coincide for an infinite number of distinct points with the limit point z_0, then the two functions are equal everywhere in $\mathcal{R}$.*

Moreover, the characteristic function of Gaussian noise also takes Gaussian form and is therefore non-zero on $(-\infty, \infty)$. This implies that

$$h(x; F_0) = -\int p_{Y|X}(y|x) \log p_Y(y; F_0) dy$$

has an extension to the complex plane $h(z; F_0)$ that is well-defined and analytic.

Now establish the two following lemmas, which are useful for contradiction.

Lemma 7.4. *If E_0 is not finite, then $h(x; F_0) = I(F_0) + D + \lambda[B - \sigma_N^2 - x^2]$ for all $x \in \mathbb{R}$, where $D = -\int p_{Y|X}(z) \log p_{Y|X}(z) dz$ is the differential entropy of the channel noise.*

Proof. Since E_0 is assumed to be an infinite set of points on $[-A, A]$, it has a limit point by the Bolzano-Weierstrass Theorem. The function $h(z; F_0)$ and the constant $I(F_0) + D + \lambda[B - \sigma_N^2 - z^2]$ are analytic on the complex plane and agree on an infinite set of points E_0 in the region $[-A, A]$. Since E_0 has a limit point in the region, it follows from the Identity Theorem that $h(z; F_0) = I(F_0) + D + \lambda[B - \sigma_N^2 - z^2]$ on the complex plane, and in particular

$$h(x; F_0) = I(F_0) + D + \lambda[B - \sigma_N^2 - x^2] \text{ for all } x \in \mathbb{R}.$$

□

Lemma 7.5. *If $h(x; F_0) = I(F_0) + D + \lambda[B - \sigma_N^2 - x^2]$ for all $x \in \mathbb{R}$, then the channel output distribution $p_Y(y; F_0)$ is distributed according to a Gaussian distribution with variance greater than σ_N^2.*

Proof. The proof is a computational exercise and is omitted. Parallel computations are provided in [350, pp. 63–64]. □

Now the statement and proof of the discreteness result.

Theorem 7.9. *E_0 is a finite set of points.*

Proof. By Theorem 7.6, $i(x; F_0) = I(F_0) + \lambda[B - \sigma_N^2 - x^2]$ for all $x \in E_0$, which is equivalent to

$$h(x; F_0) = I(F_0) + D + \lambda[B - \sigma_N^2 - x^2] \text{ for all } x \in E_0.$$

Now suppose that E_0 is *not* a finite set of points. Lemmas 7.4 and 7.5 then imply that the output distribution $p_Y(y; F_0)$ is a Gaussian distribution. Since no input distribution supported on $[-A, A]$ can yield an output distribution that is Gaussian with variance greater than σ_N^2, there is a contradiction.

Hence, E_0 is a finite set of points. □

Since the capacity-power achieving input distribution is supported on a finite set of mass points, a finite numerical optimization algorithm may be used to determine the capacity-power function [347, 355, 356]. Furthermore, since the optimal signaling

alphabet is discrete, practical signaling constellations may be used without shaping loss.

Considering the AWGN channel $\mathcal{N}(0,1)$, the capacity-power point $C(B=0;A=1.5)$ is found. The unconstrained capacity achieving input density is $p(x)=\frac{1}{2}\delta(x+1.5)+\frac{1}{2}\delta(x-1.5)$. The rate achieved for such a binary-input AWGN channel is [49, Example 4.39]:

$$C(0;1.5)=\int_{-1}^{+1}\frac{2/3}{\sqrt{2\pi}(1-y^2)}e^{-\frac{(1-(4/9)\tanh^{-1}(y))^2}{8/9}}\log(1+y)dy.$$

The output power achieved is $E[Y^2]=3.25$. In fact, this is the maximum output power possible over this channel, since $E[Y^2]=E[X^2]+\sigma_N^2$, and $E[X^2]$ cannot be improved over operating at the edges $\{-A,A\}$. Thus,

$$C(B;1.5)=C(0;1.5),\ 0\leq B\leq B_{\max}=3.25.$$

For these particular channel parameters, there actually is no trade-off between information and power: antipodal signaling should be used all the time. This is not a general phenomenon, however. This is not true for the same noise, but for say $A\geq 1.7$ rather than $A=1.5$ [350].

As an example, a particle-based numerical optimization procedure [356] is used to determine the capacity-power function for $A=5$, $C(B;5)$. This is shown in Figure 7-2. The unconstrained capacity-achieving input distribution is supported on 5 mass points, whereas the maximum power delivery is achieved with antipodal signaling at the edges.

■ 7.5 Discussion

Some have argued that the greatest inventions of civilization either transform, store, and transmit energy or they transform, store, and transmit information [357]. Although quite reasonable, many engineering systems actually deal with both energy and information: signals must be embodied in energy or matter.

In particular, the problem of communication is usually cast as one of transmitting a message generated at one point to another point. During the pre-history of information theory, a primary accomplishment was the abstraction of the message to be communicated from the communication medium. As noted, "electricity in the wires became merely a carrier of messages, not a source of power, and hence opened the door to new ways of thinking about communications" [44]. As Norbert Wiener said, "Information is information, not matter or energy. No materialism which does not admit this can survive at the present day" [358, p. 132].

Understanding signals independently from their physical manifestations led to modern communication theory, but it also blocked other possible directions. Separating messages from media arguably even led to the division of electrical engineering into two distinct subfields, electric power engineering and communication engineering. The separation of messages and media however is not always warranted.

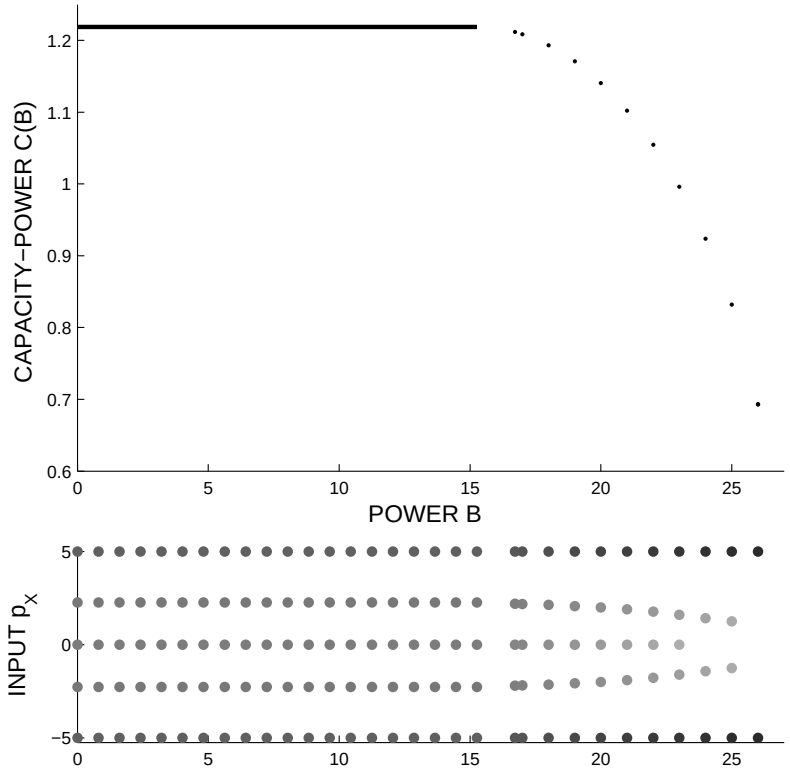

Figure 7-2. Capacity-power function for an AWGN channel with unit noise power, amplitude constraint $[-5, 5]$, and $b(y) = y^2$. The capacity-power-achieving input distribution, supported on a finite set of points, is also shown. Darker markers indicate greater probability mass.

This chapter argued that, notwithstanding reversible decoding [16], there may be benefits to transmitting power for information processing along with the information-bearing signal so as to overcome energy limitations in the receiver. The idea of broadcasting power has a long history [359].

The fundamental trade-off between the rate of transporting energy[2] and the rate of simultaneously transmitting information by modulating in energy has been defined, established, and computed for several channels.

Additional engineering inspiration for studying the simultaneous transmission of power and information is provided by powerline communication. Powerline communication has received significant attention [346, 361, 362], but the literature has focused on the informational aspect under the constraint that modulation schemes not severely degrade power delivery. This need not be the case in future engineering systems, where power and energy constraints are both limiting.

■ 7.A Review of Optimization Theory

This appendix reviews some results from optimization theory and gives some alternative versions that apply more directly to this chapter.

The first theorem deals with Lagrange multipliers.

Theorem 7.10 (Theorem 1 in Section 8.3 of [363])**.** *Let $\mathcal{A}$ be a linear vector space, $\mathcal{Z}$ a normed space, Ω a convex subset of $\mathcal{A}$, and $\mathcal{P}$ the positive cone in $\mathcal{Z}$. Assume that $\mathcal{P}$ contains an interior point. Let f be a real-valued convex functional on Ω and G a convex mapping from Ω into $\mathcal{Z}$. Assume the existence of a point $a_1 \in \Omega$ for which $G(a_1) < \vec{0}$. Let*

$$D' = \inf_{a \in \Omega, G(a) \leq \vec{0}} f(a) \tag{7.8}$$

and assume D' is finite.

Then there is a $z_0^* \geq \vec{0}$ in the normed dual of $\mathcal{Z}$, $\mathcal{Z}^*$, such that

$$D' = \inf_{a \in \Omega} \left[f(a) + \langle G(a), z_0^* \rangle \right]. \tag{7.9}$$

Moreover, if the infimum is achieved in (7.8) by an $a_0 \in \Omega$ such that $G(a_0) \leq \vec{0}$, it is achieved by a_0 in (7.9) and $\langle G(a_0), z_0^* \rangle = 0$.

A simplification of this result (as well as a conversion from an infimization problem to a supremization problem) is as follows.

Theorem 7.11. *Let $\mathcal{A}$ be a linear vector space and Ω a convex subset of $\mathcal{A}$. Let f and g be real-valued concave functionals on Ω. Assume the existence of a point $a_1 \in \Omega$ for which $g(a_1) < 0$. Let*

$$D' = \sup_{a \in \Omega, g(a) \leq 0} f(a) \tag{7.10}$$

[2] Although the chapter focuses on energy, the results apply equally well to other commodities like water, railroad cars, or packets in communication networks (whose timing is modulated [360]).

and assume D' is finite.

Then there is a $\lambda \geq 0$ in $\mathbb{R}$, such that

$$D' = \sup_{a \in \Omega} \left[f(a) - \lambda g(a) \right]. \tag{7.11}$$

Moreover, if the supremum is achieved in (7.10) *by an $a_0 \in \Omega$ such that $g(a_0) \leq 0$, it is achieved by a_0 in* (7.11) *and $\lambda g(a_0) = 0$.*

Proof. Take $\mathcal{Z}$ in Theorem 7.10 to be $\mathbb{R}$ with the usual norm. Then $\mathcal{P} = \{x \in \mathbb{R} : x \geq 0\}$ which clearly has an interior point. Moreover, $\mathcal{Z}^* = \mathbb{R}$.

The conversion from an infimization problem to a supremization problem is carried out by negation and the requirement of concave functionals rather than convex. □

The next result, a form of the Karush-Kuhn-Tucker (KKT) conditions for convex optimization, requires notions of differentiability.

Definition 7.2. *Let $\mathcal{A}$ be a vector space, and f a real-valued functional defined on domain $\Omega \subset \mathcal{A}$. Fix an $a_0 \in \Omega$ and let h be arbitrary in $\mathcal{A}$. If there exists a map $\delta f : \Omega \mapsto \mathbb{R}$ such that*

$$\delta f(a_0; h) = \lim_{\alpha \to 0} \frac{f\left[a_0 + \alpha h\right] - f(a_0)}{\alpha},$$

then f is said to be Gateaux differentiable at a_0 with increment h *and $\delta f(a_0; h)$ is the* Gateaux derivative at a_0 with increment h.

If f is Gateaux differentiable at a_0 with increment h for all $h \in \mathcal{A}$, then f is said to be Gateaux differentiable at a_0.

A slightly different definition of differentiability is useful for convex optimization, where the direction of the derivative is required to come from within the convex domain.

Definition 7.3. *Let $\mathcal{A}$ be a vector space, and f a real-valued functional defined on domain $\Omega \subset \mathcal{A}$, where Ω is a convex set. Fix an $a_0 \in \Omega$ and let $\theta \in [0, 1]$. If there exists a map $f'_{a_0} : \Omega \mapsto \mathbb{R}$ such that*

$$f'_{a_0}(a) = \lim_{\theta \downarrow 0} \frac{f\left[(1-\theta)a_0 + \theta a\right] - f(a_0)}{\theta}$$

for all $a \in \Omega$, then f is said to be weakly differentiable in Ω at a_0 *and f'_{a_0} is the* weak derivative *in Ω at a_0.*

If f is weakly differentiable in Ω at a_0 for all a_0 in Ω, then f is said to be weakly differentiable.

Notice the relationship between the weak derivative and the Gateaux derivative: $f'_{a_0}(a) = \delta f(a_0; a - a_0)$.

The following is a general result on differential methods for finding extrema.

Theorem 7.12 (Theorem 2 in Section 7.4 of [363])**.** *Let $\mathcal{A}$ be a vector space and Ω a convex subset. Let f be a real-valued functional on $\mathcal{A}$. Suppose that a_0 minimizes f on Ω and that f is Gateaux differentiable at a_0.*

Then $\delta f(a_0; a - a_0) \geq 0$ for all $a \in \Omega$.

This can be used to develop the KKT conditions for convex optimization problems, for which the existence and uniqueness of solutions can also be proven.

Theorem 7.13. *Let $\mathcal{A}$ be a normed vector space and Ω a compact and convex subset. Let f be a continuous, weakly differentiable, strictly concave real-valued functional on Ω. Let*

$$D = \sup_{a \in \Omega} f(a).$$

Then the two following properties hold:

1. *$D = \max f(a) = f(a_0)$ for some unique $a_0 \in \Omega$, and*
2. *A necessary and sufficient condition for $f(a_0) = D$ is for $f'_{a_0}(a) \leq 0$ for all $a \in \Omega$.*

Proof. Since f is a continuous function and Ω is a compact set, the range of f is a compact subset of $\mathbb{R}$ [353, Theorem 16.5]. Since the range is a compact set, it follows from the extreme value theorem that the supremum is achieved and so there is some maximizing $a_0 \in \Omega$. Uniqueness is a consequence of strict concavity, as follows. Suppose the contrary, that there are two maximizers a_1 and a_2 that achieve $f(a_1) = f(a_2) = \max f(a)$. By strict concavity,

$$f\left((1-\theta)a_1 + \theta a_2\right) > (1-\theta)f(a_1) + \theta f(a_2) = f(a_1) = \max f(a),$$

which is a contradiction. This proves the first property.

The necessity part of the second property follows directly from Theorem 7.12 by the relation $f'_{a_0}(a) = \delta f(a_0; a - a_0)$. The sufficiency part holds since there is a unique maximum. □

Chapter 8

Conclusion

Designing a communication system is a messy and complicated problem. Finite resources restrain how well communication systems can be constructed, operated, and maintained. Moreover, communication systems are subjected to the vagaries of noise in channels, encoders, and decoders. The general theme of this thesis has been to understand the effect of limited resources and limited reliability on the performance of single-user communication systems, concentrating on limitations in decoding.

By developing new mathematical models that bring key properties of physical communication systems into focus, this thesis sought insights into the basic trade-offs between the performance of a system and the resources/reliability allocated to constructing, operating, and maintaining it. In Chapter 5, the performance degradation caused by imposing extra system constraints was precisely quantified and computed, in accordance with Radner's notion of truly bounded rationality where the system designer is unaware of the degradation [185]. In other chapters, the optimal relationship between performance and resource/reliability was established quantitatively and computed, in accordance with Radner's notion of costly rationality where the system designer can optimize under knowledge of the constraint [185].

The central argument of the thesis has been that since communication systems are physically manifest, an expansion of communication theory to cover constraints imposed by their physical form leads to meaningful insights on the engineering science question: what is a good communication system?

Specific results obtained from studying these new models of communication systems show that even under optimal placement of computational nodes, decreasing the wire length of a consensus decoding circuit reduces decoding speed. Further, even when optimally quantizing sources into categories, there is a loss in optimal Bayes risk performance when performing one-shot decoding for a class of sources. The results of the thesis also establish that forms of transient and permanent noise in decoders can rule out the possibility of communicating over noisy channels with arbitrarily small error probability, but surprisingly there are system parameters where arbitrarily small error probability is possible. When arbitrarily small error probability is impossible but the goal is to communicate with some non-zero error probability, the degradation in system performance is smooth. When optimally using a system that fails catastrophically, some number of bits can be communicated at a fixed error probability. One way to mitigate decoder unreliability is to provide extra energy. With this method in mind it is shown that with a minimal received power requirement, there

are regimes where information can be communicated at the unconstrained channel capacity rate, but the rate must be reduced for more stringent power requirements even under optimal transmission.

The main results of the thesis are further detailed in Section 8.1. These results correspond to closing single switches in Figure 1-2, however one might wonder what happens when more than one switch is closed at the same time. Some problems that may be of interest for future study along the direction of simultaneously closing several switches are described in Section 8.3. Although the attentive reader has surely envisaged areas where the work presented in the individual chapters may be extended, some of these areas of possible future work primarily inspired by the contents of a single chapter are also explicitly mentioned in Section 8.2.

8.1 Recapitulation

There are several kinds of classifications of the chapters. As evident from the titles, Chapters 3 and 4 are primarily concerned with construction costs; Chapters 5 and 6 are concerned with operational reliability; and Chapter 7 is concerned with operational costs. The systems in Chapters 3 and 4 use either uncoded transmission or repetition codes; the systems in Chapter 5 use low-density parity-check codes; and the systems in Chapters 6 and 7 use optimal codes. Chapter 5 is primarily concerned with measuring performance, whereas the other chapters are concerned with optimizing performance.

The primary contributions of Chapters 3–7 are now summarized.

Wiring Costs

The general question of how material limitations in building decoding circuits might degrade their performance was investigated through a specific formulation of circuit design where quadratic wiring costs limit the convergence speed of consensus decoding.

It was shown that there is no loss of optimality in first designing the circuit topology and then optimally placing it in Euclidean space using exact spectral graph layout methods. The mathematical form of this topology design problem was shown to be the optimization of algebraic connectivity or eigenratio under a constraint on the sum of the smallest eigenvalues of the graph Laplacian. The decision version of the optimal topology design problem was proven to be NP-complete and the natural relaxation of the optimal topology design problem was proven to be a reverse convex minimization problem, thereby allowing the use of standardized optimization algorithms for circuit design. Through enumerative methods, several exactly optimal circuits were designed and exhibited. It was demonstrated that optimizing algebraic connectivity under wiring cost constraints also optimizes the eigenratio either exactly or approximately. Finally, circuits with random topologies and optimal placement were demonstrated to have a good trade-off between convergence speed and wiring cost.

Memory Costs

The general question of how limitations on the adaptability of decoding circuits might degrade their performance was investigated through a specific formulation where memory constraints limit the number of distinct rules that can be used to decode the uncoded messages produced by a source drawn at random from a population of sources. The sources must be categorized due to the limitation.

The best approach to decoder design was suggested as first categorizing the population of sources and then using a prototype source from each category to design its optimal likelihood ratio test, the decoding rule to be used for the entire category. The mathematical form of the population categorization problem was shown to be quantization of the population under the mean Bayes risk error (MBRE) criterion, a novel fidelity criterion. Unlike existing (but different) work on quantization for hypothesis testing, it was shown that minimum MBRE quantizers can be designed, either using dynamic programming for discrete populations or the Lloyd-Max algorithm for continuous populations. Several optimal quantizers were designed and exhibited. The high-rate approximation to the trade-off between memory size and performance was also established.

In the context of human decision making, the basic formulation was extended to settings where there are multiple populations. A novel model of human decision making that generates racial discrimination without a taste for discrimination was formulated. It combines the memory constraint, the automaticity of racial categorization, and facts about racial segregation. It was shown that the attitude of a decision maker, in the sense of Bayes costs, has a large impact on whether in-race or out-of-race members of the population have better outcomes. Results from econometric studies imply that for consistency with the model, most human decision makers are precautionary. Social welfare loss arising from the costly rationality of human decision makers is primarily due to the mismatch between social life and economic life arising from social segregation. This welfare loss was defined as the price of segregation and was quantified.

Transient Faults

The general question of how transient faults in operating decoding circuits might degrade their performance was investigated through a specific formulation of performance analysis of iterative message-passing decoders for low-density parity-check codes with noisy computational elements and noisy wires.

The validity of the density evolution method of performance analysis was extended to the setting where the decoder is noisy by proving the conditional independence of error probability relative to the choice of codeword, the concentration of error probability around ensemble average performance, and convergence to the cycle-free case.

Density evolution equations were derived and analyzed for a noisy version of the Gallager A decoder. It was shown that arbitrarily small communication error probability is not possible. An exact analytic expression for the final error probability as a function of channel noise and decoder noise levels was determined for this de-

coder. The set of noise parameters for which it is beneficial to use the decoder was demarcated. An equivalence between channel noise power and decoder noise power was defined so as to facilitate resource allocation in communication system design and provide guidelines for voltage scaling. It was shown that if decoding threshold and final error probability are both important metrics for assessing system performance of the noisy Gallager A decoder, then there is no total order on codes and therefore no notion of an optimal code. Density evolution equations were also derived and analyzed for a noisy Gaussian decoder. Arbitrarily small communication error probability is possible since the noise eventually becomes negligible. A decoding threshold that is a function of channel noise and decoder noise levels was determined numerically for the noisy Gaussian decoder.

New achievability results for the problem of constructing reliable memories from noisy logic gates and noisy wires were proven. The problem formulation was concerned with both circuit complexity and circuit reliability.

Permanent Faults

The general question of how permanent faults in operating decoding circuits might degrade their performance was investigated through a specific formulation of communication system design where the system fails catastrophically at a random time.

A channel model of catastrophic system failure, namely a channel that dies, was developed. Channels that die were shown to be finite-state semi-Markov channels. They were also shown to be indecomposable, to have zero capacity, and more strongly to not even be able to communicate a single binary message with arbitrarily small probability of error. A method of communication over channels that die using block codes of specified lengths was developed and proven to be optimal. Greedy algorithms and dynamic programming methods were developed to optimize the expected transmission volume at a given level of final message error probability for any death time distribution through the choice of the block code lengths. For a geometric death time distribution and an alive state that is a binary symmetric channel, it was shown that the same code should be used repeatedly. In particular, an example was given where the binary Golay code should be used in sequence.

Energy Considerations

Many engineering systems need to receive both power and information from an external source. The fundamental trade-off between transmitting power and information over a single noisy line was investigated.

A capacity-power function that specifies the optimal trade-off between transmitting power and transmitting information was defined and a coding theorem was proven to endow it with operational significance. The capacity-power function was shown to be non-decreasing and concave. The capacity-power function was computed for several channels. A notable example is the additive white Gaussian noise channel with hard amplitude constraints, which has an optimizing input distribution that is supported on a finite number of mass points.

8.2 Future Directions

Although understanding the trade-off between wiring cost and convergence speed of consensus circuits may be useful for scientific understanding or engineering design, it is just one instance of the general trade-off between infrastructure costs and performance in information processing. It would be of interest to study other related problems. As an example, one might consider the factor graph synthesis problem for belief propagation decoders under wiring-cost minimizing placement. For the specific problem solved in Chapter 3, it would be of interest to perform large-scale circuit design using the reverse convex minimization relaxation.

Besides wiring costs, another limitation on decoding circuits may be the reliability of the wiring connectivity pattern [364], due to process variation in circuit manufacturing [14, 15, 365]. It would be of interest to investigate the effects of miswiring on the performance of information processing circuits.

Quantization of prior probabilities (source side information) due to memory constraints, as in Chapter 4, led to limited adaptability in settings with heterogenous sources. A related scenario might require quantizing populations of channels rather than sources, leading to mismatched channel decoding [42]. It would be of interest to determine optimal quantizers for channels under a fidelity criterion derived from the performance of mismatched decoding. Quantization of channel state information for wireless systems with rate-constrained feedback [155] may have strong connections. It would also be of interest to look at settings where there is no distribution on the population of sources/channels, bringing ϵ-entropy to the fore.

In studying noisy message-passing decoders for LDPC codes in Chapter 5, the primary approach was performance analysis rather than determination of optimality bounds and so determining the performance of optimal systems would be of interest. To do so would require developing converse arguments for communication with noisy receivers. It might be possible to use notions of entropy generation and dissipation from the thermodynamics of computation [7, 16, 326] to study noisy computational devices, and then to tie together entropy and error probability [366]. Converse arguments for the problem of constructing reliable memories from unreliable components would require quantifying how much entropy a noisy gate can dissipate.

Chapter 5 exhibited two examples: one where arbitrarily small error probability (Shannon reliability) is achievable and one where it is not. It would be of interest to demarcate the boundaries of the two regimes. For example, one might wonder whether the fact that in the Shannon reliable case the wire noise is bounded and messages are possibly unbounded is the differentiating factor, or whether it is the fact that the wires have positive zero-error capacity, or whether it is something else altogether. Finite block length analysis of noisy decoding is also of interest.

For systems that fail catastrophically, rather than thinking of death time as independent of the signaling scheme X_1^n, taking inspiration from synthetic biology [318] one might consider channels that die because they lose fitness as a consequence of operation: T would be dependent on X_1^n. This would be similar to Gallager's panic button/child's toy channel, and would have intersymbol interference [51, 306]. There would also be strong connections to channels that heat up [367] and communication

with a dynamic cost [368, Chapter 3].

Inspired by communication terminals that randomly lie within communication range, one might also consider a channel that is born at a random time and then dies at a random time. One would suspect that channel state feedback would be beneficial. Bokodes [369] are a technology that imply a two-dimensional version of this problem and that explicitly disallow feedback. Networks of birth-death channels are also of interest and would have connections to percolation-style work [13].

For the problem of transmitting energy and information simultaneously, a closer look at applications in prosthetics [330, 331], RFID [332], and powerline communication [346] might suggest further system constraints. It is also of interest to determine the performance of specific coding schemes for the energy/information problem and finite block length bounds similar to Theorem 2.1 for comparison.

■ 8.3 Building, Maintaining, and Operating a Communication System

As depicted in Figure 1-2, there are all kinds of things that affect a communication system. The main results of the thesis consider these effects in isolation, but it is also of interest to consider them together. A few problems that suggest themselves are now described.

Chapter 4 was concerned with decoding a single observation under memory constraints whereas Chapter 3 was concerned with combining the results of several observations using consensus under wiring cost constraints. It might be of interest to decode several observations using consensus, but where the source is drawn from a population and each local computation node has finite memory. To use the terminology of human decision making, this would be an n-referees problem under wiring cost constraints.

One might consider communication systems where the decoder is subject to both transient faults and catastrophic failure: channels that die with noisy decoders. As was noted in Section 6.5, the code optimization procedures developed in Chapter 6 would work just as well if a set of finite block length codes was provided and their performance determined. In fact, the code optimization procedures would also work if $\log M(n, \eta)$ were determined for a set of codes with noisy decoders. Finite block length analysis of noisy decoding, as suggested in Section 8.2, would provide the required results.

Consensus algorithms have been studied in the setting where all wires have independent and identically distributed noise. Error probability first decreases as the algorithm iterates, but then the error probability increases; an optimal number of iterations can be determined [340]. It would be of interest to analyze these noisy decoders by relating the total cost, counting both wiring cost in construction and energy cost in operation, to the error probability and convergence speed performance.

Finally, one might study the ultimate thermodynamic trade-offs between error and energy dissipation in decoding circuits. Bennett had previously considered a particular biochemical proofreading system that models DNA replication and had analyzed the trade-off between the desired final error probability and the amount of

energy that must be dissipated at each time step in the correcting system, for a given level of noise within the decoder [370]. The speed of operation was also considered together with final error probability and energy consumption [20]. It would be of interest to study the kinds of decoders described in the thesis using these kinds of thermodynamic tools.

Thermodynamic limitations on energy and error performance, combined with information/coding theoretic limitations on error performance may provide fundamental physical and informational limits on unreliable and resource-constrained systems. As Shannon himself said [371, p. 52], "I think the connection between information theory and thermodynamics will hold up in the long run, but it has not been fully explored and understood. There is more there than we know at present."

Bibliography

[1] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423, 623–656, July/Oct. 1948. 16, 18, 19

[2] R. E. Blahut, *Theory and Practice of Error Control Codes.* Reading, MA: Addison-Wesley Publishing Company, 1983. 16, 25, 28, 32, 35

[3] A. R. Calderbank, "The art of signaling: Fifty years of coding theory," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2561–2595, Oct. 1998. doi: 10.1109/18.720549 16

[4] D. J. Costello, Jr. and G. D. Forney, Jr., "Channel coding: The road to channel capacity," *Proc. IEEE*, vol. 95, no. 6, pp. 1150–1177, Jun. 2007. doi: 10.1109/JPROC.2007.895188 16

[5] U. M. Maurer, "Conditionally-perfect secrecy and a provably-secure randomized cipher," *J. Cryptol.*, vol. 5, no. 1, pp. 53–66, Jan. 1992. doi: 10.1007/BF00191321 16

[6] O. Penrose, *Foundations of Statistical Mechanics: A Deductive Treatment.* Oxford: Pergamon Press, 1970. 16

[7] S. K. Mitter and N. J. Newton, "Information and entropy flow in the Kalman-Bucy filter," *J. Stat. Phys.*, vol. 118, no. 1-2, pp. 145–176, Jan. 2005. doi: 10.1007/s10955-004-8781-9 16, 181

[8] F. Zhao, "The physical side of computing," *Commun. ACM*, vol. 51, no. 7, p. 98, Jul. 2008. doi: 10.1145/1364782.1364803 16, 158

[9] A. Darabiha, A. C. Carusone, and F. R. Kschischang, "Block-interlaced LDPC decoders with reduced interconnect complexity," *IEEE Trans. Circuits Syst. II*, vol. 55, no. 1, pp. 74–78, Jan. 2008. doi: 10.1109/TCSII.2007.905328 16

[10] H. Sankar and K. R. Narayanan, "Memory-efficient sum–product decoding of LDPC codes," *IEEE Trans. Commun.*, vol. 52, no. 8, pp. 1225–1230, Aug. 2004. doi: 10.1109/TCOMM.2004.833016 17

[11] R. W. Hamming, "Error detecting and error correcting codes," *Bell Syst. Tech. J.*, vol. 26, no. 2, pp. 147–160, Apr. 1950. 17, 31, 100

[12] D. J. Davis, "An analysis of some failure data," *J. Am. Stat. Assoc.*, vol. 47, no. 258, pp. 113–150, Jun. 1952. 17, 133, 148

[13] I. M. Jacobs, "Connectivity in probabilistic graphs: An abstract study of reliable communications in systems containing unreliable components," Sc.D. thesis, Massachusetts Institute of Technology, Cambridge, MA, Aug. 1959. 17, 133, 182

[14] S. Winograd and J. D. Cowan, *Reliable Computation in the Presence of Noise*. Cambridge, MA: MIT Press, 1963. 17, 18, 100, 101, 181

[15] A. M. Pappu, X. Zhang, A. V. Harrison, and A. B. Apsel, "Process-invariant current source design: Methodology and examples," *IEEE J. Solid-State Circuits*, vol. 42, no. 10, pp. 2293–2302, Oct. 2007. doi: 10.1109/JSSC.2007.905240 17, 181

[16] R. Landauer, "Computation, measurement, communication and energy dissipation," in *Selected Topics in Signal Processing*, S. Haykin, Ed. Englewood Cliffs, NJ: Prentice Hall, 1989, pp. 18–47. 17, 21, 157, 158, 174, 181

[17] M. P. Frank, "Reversibility for efficient computing," Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA, Jun. 1999. doi: 1721.1/9464 17, 159

[18] A. Darabiha, A. C. Carusone, and F. R. Kschischang, "Power reduction techniques for LDPC decoders," *IEEE J. Solid-State Circuits*, vol. 43, no. 8, pp. 1835–1845, Aug. 2008. doi: 10.1109/JSSC.2008.925402 17, 157

[19] A. Sahai and P. Grover, "The price of certainty: "waterslide curves" and the gap to capacity," EECS Department, University of California, Berkeley, Tech. Rep. UCB/EECS-2008-1, Jan. 2008. 17, 125, 157, 159

[20] C. H. Bennett and M. Donkor, "Thermodynamics of error correction: Speed-error-dissipation tradeoff in copying," in *Proc. IEEE Inf. Theory Workshop (ITW'08)*, May 2008, p. 1. doi: 10.1109/ITW.2008.4578608 17, 159, 183

[21] W. H. Pierce, *Failure-Tolerant Computer Design*. New York: Academic Press, 1965. 17, 18, 99, 100

[22] S. Y. Auyang, *Engineering—An Endless Frontier*. Cambridge, MA: Harvard University Press, 2004. 18

[23] G. Harman and S. Kulkarni, *Reliable Reasoning: Induction and Statistical Learning Theory*. Cambridge, MA: MIT Press, 2007. 18

[24] R. Netz, *The Shaping of Deduction in Greek Mathematics: A Study in Cognitive History*. Cambridge: Cambridge University Press, 1999. 18

[25] S. Rajagopalan and L. Schulman, "A coding theorem for distributed computation," in *Proc. 26th Annu. ACM Symp. Theory Comput. (STOC'94)*, May 1994, pp. 790–799. doi: 10.1145/195058.195462 18

[26] R. V. L. Hartley, "Transmission of information," *Bell Syst. Tech. J.*, vol. 7, pp. 535–563, Jul. 1928. 18

[27] C. E. Shannon, "Communication in the presence of noise," *Proc. IRE*, vol. 37, no. 1, pp. 10–21, Jan. 1949. 18

[28] J. von Neumann, "Probabilistic logics and the synthesis of reliable organisms from unreliable components," in *Automata Studies*, C. E. Shannon and J. McCarthy, Eds. Princeton: Princeton University Press, 1956, pp. 43–98. 18, 100, 101

[29] C. N. Hadjicostis, *Coding Approaches to Fault Tolerance in Combinational and Dynamic Systems*. Boston: Kluwer Academic Publishers, 2002. 18, 100

[30] P. Gács and A. Gál, "Lower bounds for the complexity of reliable Boolean circuits with noisy gates," *IEEE Trans. Inf. Theory*, vol. 40, no. 2, pp. 579–583, Mar. 1994. doi: 10.1109/18.312190 18, 100

[31] P. Elias, "Computation in the presence of noise," *IBM J. Res. Develop.*, vol. 2, no. 4, pp. 346–353, Oct. 1958. 18, 100

[32] B. W. Johnson, *Design and Analysis of Fault-Tolerant Digital Systems*. Reading, MA: Addison-Wesley Publishing Company, 1989. 18, 100

[33] D. K. Pradhan, *Fault-Tolerant Computer System Design*. Upper Saddle River, NJ: Prentice Hall, 1996. 18, 100

[34] J. A. Davis, R. Venkatesan, A. Kaloyeros, M. Beylansky, S. J. Souri, K. Banerjee, K. C. Saraswat, A. Rahman, R. Reif, and J. D. Meindl, "Interconnect limits on gigascale integration (GSI) in the 21st century," *Proc. IEEE*, vol. 89, no. 3, pp. 305–324, Mar. 2001. doi: 10.1109/5.915376 19, 41

[35] R. Ho, K. W. Mai, and M. A. Horowitz, "The future of wires," *Proc. IEEE*, vol. 89, no. 4, pp. 490–504, Apr. 2001. doi: 10.1109/5.920580 19, 100

[36] B. L. Chen, D. H. Hall, and D. B. Chklovskii, "Wiring optimization can relate neuronal structure and function," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 103, no. 12, pp. 4723–4728, Mar. 2006. doi: 10.1073/pnas.0506806103 19, 49, 56, 64

[37] J. D. Meindl, J. A. Davis, P. Zarkesh-Ha, C. S. Patel, K. P. Martin, and P. A. Kohl, "Interconnect opportunities for gigascale integration," *IBM J. Res. Develop.*, vol. 46, no. 2/3, pp. 245–263, Mar.-May 2002. 19, 46

[38] D. B. Chklovskii, "Exact solution for the optimal neuronal layout problem," *Neural Comput.*, vol. 16, no. 10, pp. 2067–2078, Oct. 2004. doi: 10.1162/0899766041732422 19, 42, 47, 64

[39] S. Kar, S. Aldosari, and J. M. F. Moura, "Topology for distributed inference on graphs," *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2609–2613, Jun. 2008. doi: 10.1109/TSP.2008.923536 20, 41, 45, 46, 54

[40] L. R. Varshney, P. J. Sjöström, and D. B. Chklovskii, "Optimal information storage in noisy synapses under resource constraints," *Neuron*, vol. 52, no. 3, pp. 409–423, Nov. 2006. doi: 10.1016/j.neuron.2006.10.017 20, 67

[41] V. D. Goppa, "Nonprobabilistic mutual information without memory," *Probl. Control Inf. Theory*, vol. 4, no. 2, pp. 97–102, 1975. 20, 98

[42] A. Ganti, A. Lapidoth, and İ. E. Telatar, "Mismatched decoding revisited: General alphabets, channels with memory, and the wide-band limit," *IEEE Trans. Inf. Theory*, vol. 46, no. 7, pp. 2315–2328, Nov. 2000. doi: 10.1109/18.887846 20, 67, 98, 181

[43] T. J. Richardson and R. L. Urbanke, "The capacity of low-density parity-check codes under message-passing decoding," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 599–618, Feb. 2001. doi: 10.1109/18.910577 21, 33, 100, 101, 103, 105, 106, 126, 127, 128, 129

[44] D. A. Mindell, "Opening Black's box: Rethinking feedback's myth of origin," *Technol. Cult.*, vol. 41, no. 3, pp. 405–434, Jul. 2000. 22, 157, 172

[45] H. L. Van Trees, *Detection, Estimation, and Modulation Theory.* John Wiley & Sons, 1968. 25

[46] A. S. Willsky, G. W. Wornell, and J. H. Shapiro, *Stochastic Processes, Detection and Estimation: 6.432 Course Notes.* Massachusetts Institute of Technology, Fall 2002. 25, 28, 29

[47] J. H. van Lint, *Introduction to Coding Theory.* Springer, 1998. 25, 34

[48] G. D. Forney, Jr., *6.451 Principles of Digital Communication II Course Notes.* Massachusetts Institute of Technology, Spring 2005. 25, 63

[49] T. Richardson and R. Urbanke, *Modern Coding Theory.* Cambridge: Cambridge University Press, 2008. 25, 33, 103, 104, 105, 106, 115, 123, 126, 128, 172

[50] R. M. Fano, *Transmission of Information: A Statistical Theory of Communications.* Cambridge, MA: MIT Press, 1961. 25

[51] R. G. Gallager, *Information Theory and Reliable Communication.* New York: John Wiley & Sons, 1968. 25, 135, 136, 139, 162, 181

[52] F. Jelinek, *Probabilistic Information Theory: Discrete and Memoryless Models.* New York: McGraw-Hill Book Company, 1968. 25, 133

[53] R. J. McEliece, *The Theory of Information and Coding*. Cambridge: Cambridge University Press, 2002. 25, 35, 140

[54] R. W. Yeung, *A First Course in Information Theory*. New York: Kluwer Academic/Plenum Publishers, 2002. 25

[55] L. R. Varshney, "Optimal information storage: Nonsequential sources and neural channels," S.M. thesis, Massachusetts Institute of Technology, Cambridge, MA, Jun. 2006. doi: 1721.1/37851 26, 27, 49

[56] G. D. Forney, Jr., "Codes on graphs: Normal realizations," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 520–548, Feb. 2001. doi: 10.1109/18.910573 26, 43

[57] J. C. Willems, "The behavioral approach to open and interconnected systems," *IEEE Control Syst. Mag.*, vol. 27, no. 6, pp. 46–99, Dec. 2007. doi: 10.1109/MCS.2007.906923 26, 28

[58] N. Wiener, *The Extrapolation, Interpolation and Smoothing of Stationary Time Series*. New York: John Wiley & Sons, 1949. 26

[59] J. A. O'Sullivan, R. E. Blahut, and D. L. Snyder, "Information-theoretic image formation," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2094–2123, Oct. 1998. doi: 10.1109/18.720533 26

[60] P. M. Woodward, *Probability and Information Theory, With Applications to Radar*. New York: McGraw-Hill, 1953. 26, 107

[61] T. Berger, "Living information theory," *IEEE Inf. Theory Soc. Newsletter*, vol. 53, no. 1, pp. 1/6–19, Mar. 2003 [2002], 2002 Shannon Lecture. 26

[62] ——, *Rate Distortion Theory: A Mathematical Basis for Data Compression*. Englewood Cliffs, NJ: Prentice-Hall, 1971. 27

[63] R. M. Gray, *Entropy and Information Theory*. New York: Springer-Verlag, 1990. 27

[64] G. D. Forney, Jr., "Exponential error bounds for erasure, list, and decision feedback schemes," *IEEE Trans. Inf. Theory*, vol. IT-14, no. 2, pp. 206–220, Mar. 1968. 28, 144, 145

[65] T. R. Halford and K. M. Chugg, "The extraction and complexity limits of graphical models for linear codes," *IEEE Trans. Inf. Theory*, vol. 54, no. 9, pp. 3884–3906, Sep. 2008. doi: 10.1109/TIT.2008.928271 30, 33, 63

[66] P. Elias, "Coding for noisy channels," in *IRE Nat. Conv. Rec., Part 4*, 1955, pp. 37–46. 32

[67] R. G. Gallager, *Low-Density Parity-Check Codes*. Cambridge, MA: MIT Press, 1963. 33, 36, 101, 103, 108

[68] M. J. E. Golay, "Notes on digital coding," *Proc. IRE*, vol. 37, no. 6, p. 657, Jun. 1949. 34, 153

[69] A. Tietäväinen, "A short proof for the nonexistence of unknown perfect codes over $GF(q), q > 2$," *Ann. Acad. Sci. Fenn., Ser. A*, no. 580, pp. 1–6, 1974. 34

[70] J. H. van Lint, "A survey of perfect codes," *Rocky Mt. J. Math.*, vol. 5, no. 2, pp. 199–224, Spring 1975. doi: 10.1216/RMJ-1975-5-2-199 34

[71] T. Kasami, "A decoding procedure for multiple-error-correcting cyclic codes," *IEEE Trans. Inf. Theory*, vol. IT-10, no. 2, pp. 134–139, Apr. 1964. 36

[72] J. E. Meggitt, "Error correcting codes and their implementation for data transmission systems," *IRE Trans. Inf. Theory*, vol. IT-7, no. 4, pp. 234–244, Oct. 1961. doi: 10.1109/TIT.1961.1057659 36

[73] J. Feldman, M. J. Wainwright, and D. R. Karger, "Using linear programming to decode binary linear codes," *IEEE Trans. Inf. Theory*, vol. 51, no. 3, pp. 954–972, Mar. 2005. doi: 10.1109/TIT.2004.842696 36

[74] K. Yang, J. Feldman, and X. Wang, "Nonlinear programming approaches to decoding low-density parity-check codes," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 8, pp. 1603–1613, Aug. 2006. doi: 10.1109/JSAC.2006.879405 36

[75] S. Verdú and T. S. Han, "A general formula for channel capacity," *IEEE Trans. Inf. Theory*, vol. 40, no. 4, pp. 1147–1157, Jul. 1994. doi: 10.1109/18.335960 36

[76] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Dispersion of the Gilbert-Elliott channel," in *Proc. 2009 IEEE Int. Symp. Inf. Theory*, Jul. 2009, pp. 2209–2213. doi: 10.1109/ISIT.2009.5205838 36, 134

[77] V. Strassen, "Asymptotische abschätzungen in Shannons informationstheorie," in *Transactions of the 3rd Prague Conference on Information Theory, Statistical Decision Functions, Random Processes*. Prague: Pub. House of the Czechoslovak Academy of Sciences, 1962, pp. 689–723. 38, 141

[78] D. E. Knuth, "Big omicron and big omega and big theta," *SIGACT News*, vol. 8, no. 2, pp. 18–24, Apr.-June 1976. doi: 10.1145/1008328.1008329 38

[79] M. Sipser, *Introduction to the Theory of Computation*. Boston: PWS Publishing Company, 1997. 41

[80] M. A. Sivilotti, "Wiring considerations in analog VLSI systems, with application to field-programmable networks," Ph.D. thesis, California Institute of Technology, Pasadena, 1991. 41

[81] D. B. Chklovskii, T. Schikorski, and C. F. Stevens, "Wiring optimization in cortical circuits," *Neuron*, vol. 34, no. 3, pp. 341–347, Apr. 2002. doi: 10.1016/S0896-6273(02)00679-7 41

[82] J. A. Davis, V. K. De, and J. D. Meindl, "A stochastic wire-length distribution for gigascale integration (GSI)—Part I: Derivation and validation," *IEEE Trans. Electron Devices*, vol. 45, no. 3, pp. 580–589, Mar. 1998. doi: 10.1109/16.661219 41

[83] ——, "A stochastic wire-length distribution for gigascale integration (GSI)—Part II: Applications to clock frequency, power dissipation, and chip size estimation," *IEEE Trans. Electron Devices*, vol. 45, no. 3, pp. 590–597, Mar. 1998. doi: 10.1109/16.661220 41

[84] D. Stroobandt, *A Priori Wire Length Estimates for Digital Design*. Boston: Kluwer Academic Publishers, 2001. 41

[85] T. N. Theis, "The future of interconnection technology," *IBM J. Res. Develop.*, vol. 44, no. 3, pp. 379–390, May 2000. 41, 47

[86] M. T. Gastner and M. E. J. Newman, "The spatial structure of networks," *Eur. Phys. J. B*, vol. 49, no. 2, pp. 247–252, Jan. 2006. doi: 10.1140/epjb/e2006-00046-8 41

[87] ——, "Shape and efficiency in spatial distribution networks," *J. Stat. Mech.*, vol. 2006, no. 1, pp. 247–252, Jan. 2006. doi: 10.1088/1742-5468/2006/01/P01015 41, 49

[88] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proc. IEEE*, vol. 95, no. 1, pp. 215–233, Jan. 2007. doi: 10.1109/JPROC.2006.887293 41, 42, 46, 64

[89] P. Denantes, F. Bénézit, P. Thiran, and M. Vetterli, "Which distributed averaging algorithm should I choose for my sensor network?" in *Proc. 27th IEEE Conf. Computer Commun. (INFOCOM 2008)*, Apr. 2008, pp. 986–994. doi: 10.1109/INFOCOM.2008.152 41

[90] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," *Syst. Control Lett.*, vol. 53, no. 1, pp. 65–78, Sep. 2004. doi: 10.1016/j.sysconle.2004.02.022 41, 45, 46

[91] R. Olfati-Saber and R. M. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1520–1533, Sep. 2004. doi: 10.1109/TAC.2004.834113 42, 46, 49

[92] A. Olshevsky and J. N. Tsitsiklis, "Convergence rates in distributed consensus and averaging," in *Proc. 45th IEEE Conf. Decision Control*, Dec. 2006, pp. 3387–3392. doi: 10.1109/CDC.2006.376899 42, 46

[93] A. A. Rad, M. Jalili, and M. Hasler, "Efficient rewirings for enhancing synchronizability of dynamical networks," *Chaos*, vol. 18, no. 3, p. 037104, Sep. 2008. doi: 10.1063/1.2967738 42, 46, 54

[94] J. Diaz, J. Petit, and M. Serna, "A survey of graph layout problems," *ACM Comput. Surv.*, vol. 34, no. 3, pp. 313–356, Sep. 2002. doi: 10.1145/568522.568523 42

[95] J. L. Gross, *Topological Graph Theory.* New York: John Wiley & Sons, 1987. 42

[96] M. M. Deza and M. Laurent, *Geometry of Cuts and Metrics.* Berlin: Springer, 1997. 42

[97] N. Linial, E. London, and Y. Rabinovich, "The geometry of graphs and some of its algorithmic applications," *Combinatorica*, vol. 15, no. 2, pp. 215–245, Jun. 1995. doi: 10.1007/BF01200757 42

[98] A. L. Rosenberg and L. S. Heath, *Graph Separators, with Applications.* New York: Kluwer Academic / Plenum Publishers, 2001. 42

[99] S. Khot and A. Naor, "Nonembeddability theorems via Fourier analysis," *Math. Ann.*, vol. 334, no. 4, pp. 821–852, Apr. 2006. doi: 10.1007/s00208-005-0745-0 42

[100] P. Frankl and H. Maehara, "The Johnson-Lindenstrauss lemma and the sphericity of some graphs," *J. Comb. Theory, Ser. A*, vol. 44, no. 3, pp. 355–362, Jun. 1987. 42

[101] L. Lovász, "On the Shannon capacity of a graph," *IEEE Trans. Inf. Theory*, vol. IT-25, no. 1, pp. 1–7, Jan. 1979. 42

[102] K. M. Hall, "An r-dimensional quadratic placement algorithm," *Manage. Sci.*, vol. 17, no. 3, pp. 219–229, Nov. 1970. doi: 10.1287/mnsc.17.3.219 42, 46, 47

[103] Y. Koren, "Drawing graphs by eigenvectors: Theory and practice," *Comput. Math. Appl.*, vol. 49, no. 11-12, pp. 1867–1888, Jun. 2005. doi: 10.1016/j.camwa.2004.08.015 42

[104] A. J. Seary and W. D. Richards, "Spectral methods for analyzing and visualizing networks: An introduction," in *Dynamic Social Network Modeling and Analysis: Workshop Summary and Papers*, R. Breiger, K. Carley, and P. Pattison, Eds. Washington, DC: The National Academies Press, 2003, pp. 209–228. 42

[105] R. Olfati-Saber, E. Franco, E. Frazzoli, and J. S. Shamma, "Belief consensus and distributed hypothesis testing in sensor networks," in *Networked Embedded Sensing and Control*, ser. Lecture Notes in Control and Information Sciences, P. J. Antsaklis and P. Tabuada, Eds. Berlin: Springer, 2005, vol. 331, pp. 169–182. doi: 10.1007/11533382 44

[106] V. Borkar and P. P. Varaiya, "Asymptotic agreement in distributed estimation," *IEEE Trans. Autom. Control*, vol. AC-27, no. 3, pp. 650–655, Jun. 1982. 44

[107] J. N. Tsitsiklis and M. Athans, "Convergence and asymptotic agreement in distributed decision problems," *IEEE Trans. Autom. Control*, vol. AC-29, no. 1, pp. 42–50, Jan. 1984. 44

[108] P. F. Swaszek and P. Willett, "Parley as an approach to distributed detection," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 31, no. 1, pp. 447–457, Jan. 1995. doi: 10.1109/7.366326 44

[109] P. K. Varshney, *Distributed Detection and Data Fusion*. New York: Springer-Verlag, 1997. 44

[110] T. Berger, Z. Zhang, and H. Viswanathan, "The CEO problem," *IEEE Trans. Inf. Theory*, vol. 42, no. 3, pp. 887–902, May 1996. doi: 10.1109/18.490552 44

[111] T. S. Han and S.-I. Amari, "Statistical inference under multiterminal data compression," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2300–2324, Oct. 1998. doi: 10.1109/18.720540 44

[112] V. Misra, V. K. Goyal, and L. R. Varshney, "High-resolution functional quantization," in *Proc. Data Compression Conf. (DCC 2008)*, Mar. 2008, pp. 113–122. doi: 10.1109/DCC.2008.100 44

[113] A. Ghosh and S. Boyd, "Growing well-connected graphs," in *Proc. 45th IEEE Conf. Decision Control*, Dec. 2006, pp. 6605–6611. doi: 10.1109/CDC.2006.377282 46, 49, 59

[114] A. Hagberg and D. A. Schult, "Rewiring networks for synchronization," *Chaos*, vol. 18, no. 3, p. 037105, Sep. 2008. doi: 10.1063/1.2975842 46

[115] M. Jalili and A. A. Rad, "Comment on "rewiring networks for synchronization"," *Chaos*, vol. 19, no. 2, p. 028101, May 2009. doi: 10.1063/1.3130929 46

[116] A. Lubotzky, R. Phillips, and P. Sarnak, "Ramanujan graphs," *Combinatorica*, vol. 8, no. 3, pp. 261–277, Sep. 1988. doi: 10.1007/BF02126799 46

[117] S. Belhaiza, N. M. M. de Abreu, P. Hansen, and C. S. Oliveira, "Variable neighborhood search for extremal graphs. XI. bounds on algebraic connectivity," in *Graph Theory and Combinatorial Optimization*, D. Avis, A. Hertz, and O. Marcotte, Eds. New York: Springer, 2005, pp. 1–16. 46

[118] H. J. Landau, "Sampling, data transmission, and the Nyquist rate," *Proc. IEEE*, vol. 55, no. 10, pp. 1701–1706, Oct. 1967. 47

[119] M. Kaiser and C. C. Hilgetag, "Nonoptimal component placement, but short processing paths, due to long-distance projections in neural systems," *PLoS Comput. Biol.*, vol. 2, no. 7, p. e95, Jul. 2006. doi: 10.1371/journal.pcbi.0020095 49, 64

[120] V. A. Klyachko and C. F. Stevens, "Connectivity optimization and the positioning of cortical areas," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 100, no. 13, pp. 7937–7941, Jun. 2003. doi: 10.1073/pnas.0932745100 49, 64

[121] A. Ghosh and S. Boyd, "Upper bounds on algebraic connectivity via convex optimization," *Linear Algebra Appl.*, vol. 418, no. 2-3, pp. 693–707, Oct. 2006. doi: 10.1016/j.laa.2006.03.006 49

[122] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*. San Francisco: W. H. Freeman and Company, 1979. 50

[123] D. Mosk-Aoyama, "Maximum algebraic connectivity augmentation is NP-hard," *Oper. Res. Lett.*, vol. 36, no. 6, pp. 677–679, Nov. 2008. doi: 10.1016/j.orl.2008.09.001 50

[124] N. J. A. Sloane, "The on-line encyclopedia of integer sequences." [Online]. Available: http://www.research.att.com/~njas/sequences 51, 54

[125] L. A. Goldberg, *Efficient Algorithms for Listing Combinatorial Structures*. Cambridge: Cambridge University Press, 1993. 51

[126] W. H. Haemers and E. Spence, "Enumeration of cospectral graphs," *Eur. J. Comb.*, vol. 25, no. 2, pp. 199–211, Feb. 2004. doi: 10.1016/S0195-6698(03)00100-8 51, 65

[127] R. Marculescu and P. Bogdan, "The chip is the network: Toward a science of network-on-chip design," *Found. Trends Electron. Design Autom.*, vol. 2, no. 4, pp. 371–461, 2009. doi: 10.1561/1000000011 56

[128] L. R. Varshney, B. L. Chen, E. Paniagua, D. H. Hall, and D. B. Chklovskii, "Structural properties of the *Caenorhabditis elegans* neuronal network," Jul. 2009, arXiv:0907.2373 [q-bio]. 56

[129] H. Konno, P. T. Thach, and H. Tuy, *Optimization on Low Rank Nonconvex Structures*. Dordrecht: Kluwer Academic Publishers, 1996. 59

[130] H. Tuy, *Convex Analysis and Global Optimization*. Dordrecht: Kluwer Academic Publishers, 1998. 59

[131] J.-B. Hiriart-Urruty, "Conditions for global optimality 2," *J. Global Optim.*, vol. 13, no. 4, pp. 349–367, Dec. 1998. doi: 10.1023/A:1008365206132 59

[132] M. R. Stan, P. D. Franzon, S. C. Goldstein, J. C. Lach, and M. M. Ziegler, "Molecular electronics: From devices and interconnect to circuits and architecture," *Proc. IEEE*, vol. 91, no. 11, pp. 1940–1957, Nov. 2003. doi: 10.1109/JPROC.2003.818327 60

[133] R. Olfati-Saber, "Ultrafast consensus in small-world networks," in *Proc. Am. Contr. Conf. (ACC 2005)*, Jun. 2005, pp. 2371–2378. doi: 10.1109/ACC.2005.1470321 60

[134] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, no. 6684, pp. 440–442, Jun. 1998. doi: 10.1038/30918 60

[135] Z. Füredi and J. Kómlos, "The eigenvalues of random symmetric matrices," *Combinatorica*, vol. 1, no. 3, pp. 233–241, Sep. 1981. doi: 10.1007/BF02579329 60

[136] F. Chung, L. Lu, and V. Vu, "Spectra of random graphs with given expected degrees," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 100, no. 11, pp. 6313–6318, May 2003. doi: 10.1073/pnas.0937490100 60

[137] M. Juvan and B. Mohar, "Laplace eigenvalues and bandwidth-type invariants of graphs," *J. Graph Theory*, vol. 17, no. 3, pp. 393–407, Jul. 1993. doi: 10.1002/jgt.3190170313 62

[138] B. Mohar and S. Poljak, "Eigenvalues in combinatorial optimization," in *Combinatorial and Graph-Theoretic Problems in Linear Algebra*, R. A. Brualdi, S. Friedland, and V. Klee, Eds. Springer-Verlag, 1993, vol. 50, pp. 107–151. 62

[139] G. Robins, P. Pattison, Y. Kalish, and D. Lusher, "An introduction to exponential random graph (p^*) models for social networks," *Soc. Networks*, vol. 29, no. 2, pp. 173–191, May 2007. doi: 10.1016/j.socnet.2006.08.002 64

[140] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon, "Network motifs: Simple building blocks of complex networks," *Science*, vol. 298, no. 5594, pp. 824–827, Oct. 2002. doi: 10.1126/science.298.5594.824 64

[141] N. Przulj, "Biological network comparison using graphlet degree distribution," *Bioinformatics*, vol. 23, no. 2, pp. e177–e183, Jan. 2007. doi: 10.1093/bioinformatics/btl301 64

[142] P. Mahadevan, D. Krioukov, K. Fall, and A. Vahdat, "Systematic topology analysis and generation using degree correlations," in *Proc. 2006 Conf. Appl. Technol. Archit. Protocols Comput. Commun. (SIGCOMM'06)*, Sep. 2006, pp. 135–146. doi: 10.1145/1151659.1159930 64

[143] M. Barahona and L. M. Pecora, "Synchronization in small-world systems," *Phys. Rev. Lett.*, vol. 89, no. 5, p. 054101, Jul. 2002. doi: 10.1103/PhysRevLett.89.054101 64

[144] L. Donetti, P. I. Hurtado, and M. A. Muñoz, "Entangled networks, synchronization, and optimal network topology," *Phys. Rev. Lett.*, vol. 95, no. 18, p. 188701, Oct. 2005. doi: 10.1103/PhysRevLett.95.188701 64

[145] L. Y. Shy and B. E. Eichinger, "Large computer simulations on elastic networks: Small eigenvalues and eigenvalue spectra of the Kirchhoff matrix," *J. Chem. Phys.*, vol. 90, no. 9, pp. 5179–5189, May 1989. doi: 10.1063/1.456561 64

[146] E. R. van Dam and W. H. Haemers, "Which graphs are determined by their spectrum?" *Linear Algebra Appl.*, vol. 373, pp. 241–272, Nov. 2003. doi: 10.1016/S0024-3795(03)00483-X 65

[147] F. M. Atay and T. Biyikoglu, "Graph operations and synchronization of complex networks," *Phys. Rev. E*, vol. 72, no. 1, p. 016217, Jul. 2005. doi: 10.1103/PhysRevE.72.016217 65

[148] N. M. M. de Abreu, "Old and new results on algebraic connectivity of graphs," *Linear Algebra Appl.*, vol. 423, no. 1, pp. 53–73, May 2007. doi: 10.1016/j.laa.2006.08.017 65

[149] C. Maas, "Transportation in graphs and the admittance spectrum," *Discrete Appl. Math.*, vol. 16, no. 1, pp. 31–49, Jan. 1987. doi: 10.1016/0166-218X(87)90052-7 65

[150] F. Goldberg, "Bounding the gap between extremal Laplacian eigenvalues of graphs," *Linear Algebra Appl.*, vol. 416, no. 1, pp. 68–74, Jul. 2006. doi: 10.1016/j.laa.2005.07.007 65

[151] J. M. Mulder, N. T. Quach, and M. J. Flynn, "An area model for on-chip memories and its application," *IEEE J. Solid-State Circuits*, vol. 26, no. 2, pp. 98–106, Feb. 1991. doi: 10.1109/4.68123 67

[152] P. R. Panda, N. Dutt, and A. Nicolau, *Memory Issues in Embedded Systems-On-Chip: Optimizations and Exploration.* Norwell, MA: Kluwer Academic Publishers, 1999. 67

[153] B. Jacob, S. W. Ng, and D. T. Wang, *Memory Systems: Cache, DRAM, Disk.* Burlington, MA: Morgan Kaufmann Publishers, 2008. 67

[154] D. L. Neuhoff, R. M. Gray, and L. D. Davisson, "Fixed rate universal block source coding with a fidelity criterion," *IEEE Trans. Inf. Theory*, vol. IT-21, no. 5, pp. 511–523, Sep. 1975. 67, 98

[155] A. Narula, M. J. Lopez, M. D. Trott, and G. W. Wornell, "Efficient use of side information in multiple-antenna data transmission over fading channels," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1423–1436, Oct. 1998. doi: 10.1109/49.730451 67, 181

[156] D. J. Love, R. W. Heath, Jr., V. K. N. Lau, D. Gesbert, B. D. Rao, and M. Andrews, "An overview of limited feedback in wireless communication systems," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 8, pp. 1341–1365, Oct. 2008. doi: 10.1109/JSAC.2008.081002 67

[157] G. A. Miller, "The magical number seven, plus or minus two: Some limits on our capacity for processing information," *Psychol. Rev.*, vol. 63, pp. 81–97, 1956. 67

[158] A. S. Willsky, G. W. Wornell, and J. H. Shapiro, *Stochastic Processes, Detection and Estimation 6.432 Course Notes*. Cambridge, MA: Dept. Elect. Eng. Comput. Sci., Mass. Inst. Tech., Fall 2003. 68, 69, 70

[159] G. A. Miller, "Human memory and the storage of information," *IRE Trans. Inf. Theory*, vol. IT-2, no. 3, pp. 129–137, Sep. 1956. doi: 10.1109/TIT.1956.1056815 68, 89

[160] S. A. Kassam, "Optimum quantization for signal detection," *IEEE Trans. Commun.*, vol. COM-25, no. 5, pp. 479–484, May 1977. 69

[161] H. V. Poor and J. B. Thomas, "Applications of Ali–Silvey distance measures in the design of generalized quantizers for binary decision systems," *IEEE Trans. Commun.*, vol. COM-25, no. 9, pp. 893–900, Sep. 1977. 69

[162] R. Gupta and A. O. Hero, III, "High-rate vector quantization for detection," *IEEE Trans. Inf. Theory*, vol. 49, no. 8, pp. 1951–1969, Aug. 2003. doi: 10.1109/TIT.2003.814482 69

[163] C. Hildreth, "Bayesian statisticians and remote clients," *Econometrica*, vol. 31, no. 3, pp. 422–438, Jul. 1963. 69

[164] R. E. Kihlstrom, "The use of approximate prior distributions in a Bayesian decision model," *Econometrica*, vol. 39, no. 6, pp. 899–910, Nov. 1971. 69

[165] D. C. Gilliland and M. K. Helmers, "On continuity of the Bayes response," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 4, pp. 506–508, Jul. 1978. 69

[166] R. A. Wijsman, "Continuity of the Bayes risk," *Ann. Math. Stat.*, vol. 41, no. 3, pp. 1083–1085, Jun. 1970. doi: 10.1214/aoms/1177696987 69

[167] M. H. DeGroot, *Optimal Statistical Decisions*. Hoboken, NJ: Wiley-Interscience, 2004. 69

[168] J. Li, N. Chaddha, and R. M. Gray, "Asymptotic performance of vector quantizers with a perceptual distortion measure," *IEEE Trans. Inf. Theory*, vol. 45, no. 4, pp. 1082–1091, May 1999. doi: 10.1109/18.761252 72, 78, 79

[169] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Boston: Kluwer Academic Publishers, 1992. 73, 75, 92

[170] H. Steinhaus, "Sur la division des corp materiels en parties," *Bull. Acad. Polonaise Sci.*, vol. C1.III, no. IV, pp. 801–804, 1956. 73, 75

[171] S. P. Lloyd, "Least squares quantization in PCM," Jul. 1957, unpublished Bell Laboratories Technical Note. 73, 75, 76

[172] J. Max, "Quantizing for minimum distortion," *IRE Trans. Inf. Theory*, vol. IT-6, no. 1, pp. 7–12, Mar. 1960. doi: 10.1109/TIT.1960.1057548 73, 75

[173] A. V. Trushkin, "Sufficient conditions for uniqueness of a locally optimal quantizer for a class of convex error weighting functions," *IEEE Trans. Inf. Theory*, vol. IT-28, no. 2, pp. 187–198, Mar. 1982. 75, 76

[174] T. L. Fine, *Probability and Probabilistic Reasoning for Electrical Engineering*. Upper Saddle River, NJ: Prentice Hall, 2006. 75

[175] R. M. Gray and D. L. Neuhoff, "Quantization," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2325–2383, Oct. 1998. doi: 10.1109/18.720541 75, 78

[176] D. K. Sharma, "Design of absolutely optimal quantizers for a wide class of distortion measures," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 6, pp. 693–702, Nov. 1978. 77, 78

[177] K. Rose, "Deterministic annealing for clustering, compression, classification, regression, and related optimization problems," *Proc. IEEE*, vol. 86, no. 11, pp. 2210–2239, Nov. 1998. doi: 10.1109/5.726788 77

[178] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear Programming: Theory and Algorithms*. Hoboken, NJ: Wiley-Interscience, 2006. 78

[179] S. A. Kassam, "Quantization based on the mean-absolute-error criterion," *IEEE Trans. Commun.*, vol. COM-26, no. 2, pp. 267–270, Feb. 1978. 80

[180] F. Topsøe, "Some inequalities for information divergence and related measures of discrimination," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1602–1609, Jul. 2000. doi: 10.1109/18.850703 81

[181] R. M. Gray and A. H. Gray, Jr., "Asymptotically optimal quantizers," *IEEE Trans. Inf. Theory*, vol. IT-23, no. 1, pp. 143–144, Jan. 1977. 81

[182] J. Conlisk, "Why bounded rationality?" *J. Econ. Lit.*, vol. 34, no. 2, pp. 669–700, Jun. 1996. 89

[183] A. Rubinstein, *Modeling Bounded Rationality*. Cambridge, MA: MIT Press, 1998. 89

[184] P. Weirich, *Realistic Decision Theory: Rules for Nonideal Agents in Nonideal Circumstances*. Oxford: Oxford University Press, 2004. 89

[185] R. Radner, "Costly and bounded rationality in individual and team decision-making," in *Understanding Industrial and Corporate Change*, G. Dosi, D. J. Teece, and J. Chytry, Eds. Oxford: Oxford University Press, 2005, pp. 3–35. 89, 91, 177

[186] F. P. Ramsey, "Truth and probability," in *The Foundations of Mathematics and Other Logical Essays*, R. B. Braithwaite, Ed. New York: Harcourt, Brace and Company, 1931, pp. 156–198. 89

[187] B. de Finetti, "La prévision: Ses lois logiques, ses sources subjectives," *Ann. Inst. Henri Poincaré*, vol. 7, no. 1, pp. 1–68, 1937. 89

[188] J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press, 1944. 89

[189] L. J. Savage, *The Foundations of Statistics*. New York: Wiley, 1954. 89

[190] G. A. Akerlof and R. E. Kranton, "Economics and identity," *Quart. J. Econ.*, vol. 115, no. 3, pp. 715–753, Aug. 2000. doi: 10.1162/003355300554881 91, 92

[191] R. G. Gallager, *Principles of Digital Communication*. Cambridge: Cambridge University Press, 2008. 91

[192] C. N. Macrae and G. V. Bodenhausen, "Social cognition: Thinking categorically about others," *Annu. Rev. Psychol.*, vol. 51, pp. 93–120, Feb. 2000. doi: 10.1146/annurev.psych.51.1.93 91

[193] J. E. Stiglitz, "Information and the change in the paradigm in economics," Dec. 2001, nobel Prize Lecture. 91

[194] G. A. Akerlof, "The market for "lemons": Quality uncertainty and the market mechanism," *Quart. J. Econ.*, vol. 84, no. 3, pp. 488–500, Aug. 1970. 91

[195] K. Arrow, "The theory of discrimination," in *Discrimination in Labor Markets*, O. Ashenfelter and A. Rees, Eds. Princeton, NJ: Princeton University Press, 1973, pp. 3–33. 91

[196] S. Coate and G. C. Loury, "Will affirmative-action policies eliminate negative stereotypes?" *Am. Econ. Rev.*, vol. 83, no. 5, pp. 1220–1240, Dec. 1993. 91

[197] R. Fryer and M. O. Jackson, "A categorical model of cognition and biased decision-making," *B. E. J. Theor. Econ.*, vol. 8, no. 1, Jan. 2008. 91, 92, 98

[198] F. Echenique and R. G. Fryer, Jr., "A measure of segregation based on social interactions," *Quart. J. Econ.*, vol. 122, no. 2, pp. 441–485, May 2007. doi: 10.1162/qjec.122.2.441 91, 96

[199] Y. Shoham and A. Gersho, "Efficient bit allocation for an arbitrary set of quantizers," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 36, no. 9, pp. 1445–1453, Sep. 1988. doi: 10.1109/29.90373 91

[200] J. N. Wood, "Social cognition and the prefrontal cortex," *Behav. Cogn. Neurosci. Rev.*, vol. 2, no. 2, pp. 97–114, Jun. 2003. doi: 10.1177/1534582303253625 92

[201] C. A. Meissner and J. C. Brigham, "Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review," *Psychol. Pub. Pol. L.*, vol. 7, no. 1, pp. 3–35, Jan. 2001. 93

[202] J. J. Donohue, III and S. D. Levitt, "The impact of race on policing and arrests," *J. Law Econ.*, vol. 44, no. 2, pp. 367–394, Oct. 2001. doi: 10.1086/322810 93

[203] K. L. Antonovics and B. G. Knight, "A new look at racial profiling: Evidence from the Boston Police Department," *Rev. Econ. Stat.*, vol. 91, no. 1, pp. 163–177, Feb. 2009. doi: 10.1162/rest.91.1.163 93

[204] J. Price and J. Wolfers, "Racial discrimination among NBA referees," NBER, Working Paper 13206, Jun. 2007. 93

[205] N. A. Ashford and C. S. Miller, "Low-level chemical exposures: A challenge for science and policy," *Environ. Sci. Technol.*, vol. 32, no. 21, pp. 508–509, Nov. 1998. 94

[206] H. M. Sapolsky, "The politics of risk," *Daedalus*, vol. 19, no. 4, pp. 83–96, Fall 1990. 94

[207] W. K. Viscusi, J. E. Jr., Harrington, and J. M. Vernon, *Economics of Regulation and Antitrust*. Cambridge, MA: MIT Press, 2005. 96

[208] C. H. Papadimitriou, "Algorithms, games, and the internet," in *Proc. 33rd Annu. ACM Symp. Theory Comput. (STOC'01)*, Jul. 2001, pp. 749–753. doi: 10.1145/380752.380883 96

[209] G. W. Allport, *The Nature of Prejudice*. Reading, MA: Addison-Wesley, 1954. 96

[210] C. Van Laar, S. Levin, S. Sinclair, and J. Sidanius, "The effect of university roommate contact on ethnic attitudes and behavior," *J. Exp. Soc. Psychol.*, vol. 41, no. 4, pp. 329–345, Jul. 2005. doi: 10.1016/j.jesp.2004.08.002 96

[211] M. Rothbart and O. P. John, "Social categorization and behavioral episodes: A cognitive analysis of the effects of intergroup contact," *J. Soc. Issues*, vol. 41, no. 3, pp. 81–104, Fall 1985. 96

[212] J. Makhoul, S. Roucos, and H. Gish, "Vector quantization in speech coding," *Proc. IEEE*, vol. 73, no. 11, pp. 1551–1588, Nov. 1985. 98

[213] H. D. Pfister, I. Sason, and R. Urbanke, "Capacity-achieving ensembles for the binary erasure channel with bounded complexity," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2352–2379, Jul. 2005. doi: 10.1109/TIT.2005.850079 99

[214] S.-Y. Chung, G. D. Forney, Jr., T. J. Richardson, and R. Urbanke, "On the design of low-density parity-check codes within 0.0045 dB of the Shannon limit," *IEEE Commun. Lett.*, vol. 5, no. 2, pp. 58–60, Feb. 2001. doi: 10.1109/4234.905935 99

[215] H.-A. Loeliger, F. Lustenberger, M. Helfenstein, and F. Tarköy, "Probability propagation and decoding in analog VLSI," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 837–843, Feb. 2001. doi: 10.1109/18.910594 99, 125

[216] A. J. Blanksby and C. J. Howland, "A 690-mW 1-Gb/s 1024-b, rate-1/2 low-density parity-check code decoder," *IEEE J. Solid-State Circuits*, vol. 37, no. 3, pp. 404–412, Mar. 2002. doi: 10.1109/4.987093 99

[217] T. Zhang and K. K. Parhi, "An FPGA implementation of $(3, 6)$-regular low-density parity-check code decoder," *EURASIP J. Appl. Signal Process.*, vol. 2003, no. 6, pp. 530–542, 2003. doi: 10.1155/S1110865703212105 99

[218] L. M. J. Bazzi and S. K. Mitter, "Endcoding complexity versus minimum distance," *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 2103–2112, Jun. 2005. doi: 10.1109/TIT.2005.847727 99

[219] P. Larsson and C. Svensson, "Noise in digital dynamic CMOS circuits," *IEEE J. Solid-State Circuits*, vol. 29, no. 6, pp. 655–662, Jun. 1994. doi: 10.1109/4.293110 100

[220] C. Zhao, X. Bai, and S. Dey, "Evaluating transient error effects in digital nanometer circuits," *IEEE Trans. Rel.*, vol. 56, no. 3, pp. 381–391, Sep. 2007. doi: 10.1109/TR.2007.903288 100

[221] T. Rejimon, K. Lingasubramanian, and S. Bhanja, "Probabilistic error modeling for nano-domain logic circuits," *IEEE Trans. VLSI Syst.*, vol. 17, no. 1, pp. 55–65, Jan. 2009. doi: 10.1109/TVLSI.2008.2003167 100

[222] K. K. Likharev, "Single-electron devices and their applications," *Proc. IEEE*, vol. 87, no. 4, pp. 606–632, Apr. 1999. doi: 10.1109/5.752518 100

[223] A. Bachtold, P. Hadley, T. Nakanishi, and C. Dekker, "Logic circuits with carbon nanotube transistors," *Science*, vol. 294, no. 5545, pp. 1317–1320, Nov. 2001. doi: 10.1126/science.1065824 100

[224] C. P. Collier, E. W. Wong, M. Belohradský, F. M. Raymo, J. F. Stoddart, P. J. Kuekes, R. S. Williams, and J. R. Heath, "Electronically configurable molecular-based logic gates," *Science*, vol. 285, no. 5426, pp. 391–394, Jul. 1999. doi: 10.1126/science.285.5426.391 100

[225] J. Han and P. Jonker, "A defect- and fault-tolerant architecture for nanocomputers," *Nanotechnology*, vol. 14, no. 2, pp. 224–230, Feb. 2003. doi: 10.1088/0957-4484/14/2/324 100

[226] L. Anghel and M. Nicolaidis, "Defects tolerant logic gates for unreliable future nanotechnologies," in *Computational and Ambient Intelligence*, ser. Lecture Notes in Computer Science, F. Sandoval, A. Prieto, J. Cabestany, and M. Graña, Eds. Berlin: Springer, 2007, vol. 4507, pp. 422–429. doi: 10.1007/978-3-540-73007-1 100

[227] V. Bush, "The differential analyzer. A new machine for solving differential equations," *J. Franklin Inst.*, vol. 212, no. 4, pp. 447–488, Oct. 1931. doi: 10.1016/S0016-0032(31)90616-9 100

[228] R. Sarpeshkar, "Analog versus digital: Extrapolating from electronics to neurobiology," *Neural Comput.*, vol. 10, no. 7, pp. 1601–1638, Oct. 1998. doi: 10.1162/089976698300017052 100

[229] B. Widrow and I. Kollár, *Quantization Noise: Roundoff Error in Digital Computation, Signal Processing, Control, and Communications*. Cambridge: Cambridge University Press, 2008. 100, 117

[230] N. Pippenger, "Reliable computation by formulas in the presence of noise," *IEEE Trans. Inf. Theory*, vol. 34, no. 2, pp. 194–197, Mar. 1988. doi: 10.1109/18.2628 101, 119

[231] W. S. Evans and L. J. Schulman, "Signal propagation and noisy circuits," *IEEE Trans. Inf. Theory*, vol. 45, no. 7, pp. 2367–2373, Nov. 1999. doi: 10.1109/18.796377 101, 119

[232] C. Crick and A. Pfeffer, "Loopy belief propagation as a basis for communication in sensor networks," in *Proc. 19th Annu. Conf. Uncertainty in Artificial Intelligence (UAI'03)*, Aug. 2003, pp. 151–158. 101

[233] O. W. Yeung and K. M. Chugg, "On the error tolerance of iterative decoder circuitry," in *Proc. 2008 Inf. Theory Appl. Workshop*, Jan. 2008. 101

[234] L. Wei, "Robustness of LDPC codes and internal noisy systems," in *Proc. 41st Annu. Allerton Conf. Commun. Control Comput.*, Oct. 2003, pp. 1665–1674. 101

[235] H. Saeedi and A. H. Banihashemi, "Performance of belief propagation for decoding LDPC codes in the presence of channel estimation error," *IEEE Trans. Commun.*, vol. 55, no. 1, pp. 83–89, Jan. 2007. doi: 10.1109/TCOMM.2006.887488 101

[236] L. Ping and W. K. Leung, "Decoding low density parity check codes with finite quantization bits," *IEEE Commun. Lett.*, vol. 4, no. 2, pp. 62–64, Feb. 2000. doi: 10.1109/4234.824757 101

[237] R. Singhal, G. S. Choi, and R. N. Mahapatra, "Quantized LDPC decoder design for binary symmetric channels," in *Proc. IEEE Int. Symp.*

Circuits Syst. (ISCAS 2005), vol. 6, May 2005, pp. 5782–5785. doi: 10.1109/ISCAS.2005.1465952 101

[238] N. Miladinovic and M. P. C. Fossorier, "Improved bit-flipping decoding of low-density parity-check codes," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1594–1606, Apr. 2005. doi: 10.1109/TIT.2005.844095 101

[239] J. Chen, A. Dholakia, E. Eleftheriou, M. P. C. Fossorier, and X.-Y. Hu, "Reduced-complexity decoding of LDPC codes," *IEEE Trans. Commun.*, vol. 53, no. 8, pp. 1288–1299, Aug. 2005. doi: 10.1109/TCOMM.2005.852852 101

[240] J. Zhao, F. Zarkeshvari, and A. H. Banihashemi, "On implementation of min-sum algorithm and its modifications for decoding low-density parity-check (LDPC) codes," *IEEE Trans. Commun.*, vol. 53, no. 4, pp. 549–554, Apr. 2005. doi: 10.1109/TCOMM.2004.836563 101

[241] T. Yu, R.-S. Lin, B. Super, and B. Tang, "Efficient message representations for belief propagation," in *Proc. 11th IEEE Int. Conf. Computer Vision*, Oct. 2007, pp. 1–8. doi: 10.1109/ICCV.2007.4408905 101

[242] R. Motwani and P. Raghavan, *Randomized Algorithms*. Cambridge: Cambridge University Press, 1995. 101

[243] S. T. Ribeiro, "Random-pulse machines," *IEEE Trans. Electron. Comput.*, vol. EC-16, no. 3, pp. 261–276, Jun. 1967. doi: 10.1109/PGEC.1967.264661 101

[244] S. S. Tehrani, W. J. Gross, and S. Mannor, "Stochastic decoding of LDPC codes," *IEEE Commun. Lett.*, vol. 10, no. 10, pp. 716–718, Oct. 2006. doi: 10.1109/LCOMM.2006.060570 101

[245] A. Sahai and S. Mitter, "The necessity and sufficiency of anytime capacity for stabilization of a linear system over a noisy communication link—Part I: Scalar systems," *IEEE Trans. Inf. Theory*, vol. 52, no. 8, pp. 3369–3395, Aug. 2006. doi: 10.1109/TIT.2006.878169 101

[246] A. T. Ihler, J. W. Fisher, III, and A. S. Willsky, "Loopy belief propagation: Convergence and effects of message errors," *J. Mach. Learn. Res.*, vol. 6, pp. 905–936, May 2005. 101

[247] M. G. Taylor, "Reliable information storage in memories designed from unreliable components," *Bell Syst. Tech. J.*, vol. 47, no. 10, pp. 2299–2337, Dec. 1968. 101, 120, 122, 124

[248] A. V. Kuznetsov, "Information storage in a memory assembled from unreliable components," *Probl. Inf. Transm.*, vol. 9, no. 3, pp. 100–114, July-Sept. 1973. 101, 120

[249] D. J. C. MacKay and R. M. Neal, "Near Shannon limit performance of low density parity check codes," *IEE Electron. Lett.*, vol. 33, no. 6, pp. 457–458, Mar. 1997. doi: 10.1049/el:19961141 101

[250] M. Sipser and D. A. Spielman, "Expander codes," *IEEE Trans. Inf. Theory*, vol. 42, no. 6, pp. 1710–1722, Nov. 1996. doi: 10.1109/18.556667 101

[251] D. Burshtein and G. Miller, "Expander graph arguments for message-passing algorithms," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 782–790, Feb. 2001. doi: 10.1109/18.910588 101

[252] S. ten Brink, "Convergence of iterative decoding," *IEE Electron. Lett.*, vol. 35, no. 10, pp. 806–808, May 1999. doi: 10.1049/el:19990555 101

[253] M. Ardakani and F. R. Kschischang, "A more accurate one-dimensional analysis and design of irregular LDPC codes," *IEEE Trans. Commun.*, vol. 52, no. 12, pp. 2106–2114, Dec. 2004. doi: 10.1109/TCOMM.2004.838718 101

[254] M. G. Luby, M. Mitzenmacher, M. A. Shokrollahi, and D. A. Spielman, "Improved low-density parity-check codes using irregular graphs," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 585–598, Feb. 2001. doi: 10.1109/18.910576 101, 103, 125

[255] T. Richardson and R. Urbanke, "Fixed points and stability of density evolution," *Commun. Inf. Syst.*, vol. 4, no. 1, pp. 103–116, Sep. 2004. 101, 106

[256] S. K. Chilappagari and B. Vasic, "Fault tolerant memories based on expander graphs," in *Proc. IEEE Inf. Theory Workshop (ITW'07)*, Sep. 2007, pp. 126–131. doi: 10.1109/ITW.2007.4313061 102, 120

[257] L. R. Varshney, "Performance of LDPC codes under noisy message-passing decoding," in *Proc. IEEE Inf. Theory Workshop (ITW'07)*, Sep. 2007, pp. 178–183. doi: 10.1109/ITW.2007.4313070 102

[258] H.-A. Loeliger, "An introduction to factor graphs," *IEEE Signal Process. Mag.*, vol. 21, no. 1, pp. 28–41, Jan. 2004. doi: 10.1109/MSP.2004.1267047 102

[259] V. Rathi and R. Urbanke, "Density evolution, thresholds and the stability condition for non-binary LDPC codes," *IEE Proc. Commun.*, vol. 152, no. 6, pp. 1069–1074, Dec. 2005. doi: 10.1049/ip-com:20050230 104

[260] Y. Weiss, "Correctness of local probability propagation in graphical models with loops," *Neural Comput.*, vol. 12, no. 1, pp. 1–41, Jan. 2000. doi: 10.1162/089976600300015880 106

[261] D. Slepian, "The threshold effect in modulation systems that expand bandwidth," *IRE Trans. Inf. Theory*, vol. IT-8, no. 5, pp. 122–127, Sep. 1962. doi: 10.1109/TIT.1962.1057759 107

[262] J. M. Wozencraft and I. M. Jacobs, *Principles of Communication Engineering*. New York: John Wiley & Sons, 1965. 107

[263] L. Bazzi, T. J. Richardson, and R. L. Urbanke, "Exact thresholds and optimal codes for the binary-symmetric channel and Gallager's decoding algorithm A," *IEEE Trans. Inf. Theory*, vol. 50, no. 9, pp. 2010–2021, Sep. 2004. doi: 10.1109/TIT.2004.833352 108, 110, 112, 114

[264] R. L. Dobrushin and S. I. Ortyukov, "Lower bound for the redundancy of self-correcting arrangements of unreliable functional elements," *Probl. Inf. Transm.*, vol. 13, no. 1, pp. 82–89, Jan.-Mar. 1977. 108

[265] L. Yang, "Recent advances on determining the number of real roots of parametric polynomials," *J. Symbol. Comput.*, vol. 28, no. 1-2, pp. 225–242, Jul. 1999. doi: 10.1006/jsco.1998.0274 111

[266] T. J. Richardson, M. A. Shokrollahi, and R. L. Urbanke, "Design of capacity-approaching irregular low-density parity-check codes," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 619–637, Feb. 2001. doi: 10.1109/18.910578 114

[267] S.-Y. Chung, T. J. Richardson, and R. L. Urbanke, "Analysis of sum-product decoding of low-density parity-check codes using a Gaussian approximation," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 657–670, Feb. 2001. doi: 10.1109/18.910580 116, 117, 119, 120

[268] M. Lentmaier, D. V. Truhachev, K. S. Zigangirov, and D. J. Costello, Jr., "An analysis of the block error probability performance of iterative decoding," *IEEE Trans. Inf. Theory*, vol. 51, no. 11, pp. 3834–3855, Nov. 2005. doi: 10.1109/TIT.2005.856942 119

[269] S. K. Chilappagari, B. Vasic, and M. Marcellin, "Can the storage capacity of memories built from unreliable components be determined?" in *Proc. 2008 Inf. Theory Appl. Workshop*, Jan.-Feb. 2008, pp. 41–43. doi: 10.1109/ITA.2008.4601020 120

[270] A. G. Dimakis and K. Ramchandran, "Network coding for distributed storage in wireless networks," in *Networked Sensing Information and Control*, V. Saligrama, Ed. New York: Springer, 2008, pp. 115–136. 120

[271] A. Montanari, "Tight bounds for LDPC and LDGM codes under MAP decoding," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3221–3246, Sep. 2005. doi: 10.1109/TIT.2005.853320 123

[272] J. G. Tryon, "Quadded logic," in *Redundancy Techniques for Computing Systems*, R. H. Wilcox and W. C. Mann, Eds. Washington: Spartan Books, 1962, pp. 205–228. 125

[273] J. B. Gao, Y. Qi, and J. A. B. Fortes, "Bifurcations and fundamental error bounds for fault-tolerant computations," *IEEE Trans. Nanotechnol.*, vol. 4, no. 4, pp. 395–402, Jul. 2005. doi: 10.1109/TNANO.2005.851289 125

[274] S. H. Nawab, A. V. Oppenheim, A. P. Chandrakasan, J. M. Winograd, and J. T. Ludwig, "Approximate signal processing," *J. VLSI Signal Process.*, vol. 15, no. 1-2, pp. 177–200, Jan. 1997. doi: 10.1023/A:1007986707921 125

[275] A. Sinha, A. Wang, and A. Chandrakasan, "Energy scalable system design," *IEEE Trans. VLSI Syst.*, vol. 10, no. 2, pp. 135–145, Apr. 2002. doi: 10.1109/92.994990 125

[276] M. Bhardwaj, R. Min, and A. P. Chandrakasan, "Quantifying and enhancing power awareness of VLSI systems," *IEEE Trans. VLSI Syst.*, vol. 9, no. 6, pp. 757–772, Dec. 2001. doi: 10.1109/92.974890 125

[277] R. Hegde and N. R. Shanbhag, "Toward achieving energy efficiency in presence of deep submicron noise," *IEEE Trans. VLSI Syst.*, vol. 8, no. 4, pp. 379–391, Aug. 2000. doi: 10.1109/92.863617 125

[278] L. Wang and N. R. Shanbhag, "Energy-efficiency bounds for deep submicron VLSI systems in the presence of noise," *IEEE Trans. VLSI Syst.*, vol. 11, no. 2, pp. 254–269, Apr. 2003. doi: 10.1109/TVLSI.2003.810783 125

[279] G. Masera, M. Mazza, G. Piccinini, F. Viglione, and M. Zamboni, "Architectural strategies for low-power VLSI turbo decoders," *IEEE Trans. VLSI Syst.*, vol. 10, no. 3, pp. 279–285, Jun. 2002. doi: 10.1109/TVLSI.2002.1043330 125

[280] M. M. Mansour and N. R. Shanbhag, "A 640-Mb/s 2048-bit programmable LDPC decoder chip," *IEEE J. Solid-State Circuits*, vol. 41, no. 3, pp. 684–698, Mar. 2006. doi: 10.1109/JSSC.2005.864133 125

[281] B. J. Copeland, "Hypercomputation," *Minds Mach.*, vol. 12, no. 4, pp. 461–502, Nov. 2002. doi: 10.1023/A:1021105915386 125

[282] D. Williams, *Probability with Martingales*. Cambridge: Cambridge University Press, 1991. 126

[283] K. Azuma, "Weighted sums of certain dependent random variables," *Tohoku Math. J.*, vol. 19, no. 3, pp. 357–367, 1967. 127

[284] W. Hoeffding, "Probability inequalities for sums of bounded random variables," *J. Am. Stat. Assoc.*, vol. 58, no. 301, pp. 13–30, Mar. 1963. 127

[285] L. H. Ozarow, S. Shamai, and A. D. Wyner, "Information theoretic considerations for cellular mobile radio," *IEEE Trans. Veh. Technol.*, vol. 43, no. 2, pp. 359–378, May 1994. doi: 10.1109/25.293655 133

[286] A. Goldsmith, *Wireless Communications*. New York: Cambridge University Press, 2005. 133

[287] J. K. Wolf, A. D. Wyner, and J. Ziv, "The channel capacity of the postal channel," *Inf. Control*, vol. 16, no. 2, pp. 167–172, Apr. 1970. doi: 10.1016/S0019-9958(70)90097-5 133

[288] M. Zeng, R. Zhang, and S. Cui, "On the outage capacity of a dying channel," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM 2008)*, Dec. 2008. doi: 10.1109/GLOCOM.2008.ECP.198 133

[289] D. Slepian, "Bounds on communication," *Bell Syst. Tech. J.*, vol. 42, pp. 681–707, May 1963. 133, 141

[290] S. J. MacMullan and O. M. Collins, "A comparison of known codes, random codes, and the best codes," *IEEE Trans. Inf. Theory*, vol. 44, no. 7, pp. 3009–3022, Nov. 1998. doi: 10.1109/18.737529 133, 141

[291] J. N. Laneman, "On the distribution of mutual information," in *Proc. Inf. Theory Appl. Inaugural Workshop*, Feb. 2006. 133, 141

[292] Y. Polyanskiy, H. V. Poor, and S. Verdú, "New channel coding achievability bounds," in *Proc. 2008 IEEE Int. Symp. Inf. Theory*, Jul. 2008, pp. 1763–1767. doi: 10.1109/ISIT.2008.4595291 133, 141

[293] D. Buckingham and M. C. Valenti, "The information-outage probability of finite-length codes over AWGN channels," in *Proc. 42nd Annu. Conf. Inf. Sci. Syst. (CISS 2008)*, Mar. 2008, pp. 390–395. doi: 10.1109/CISS.2008.4558558 133, 141

[294] G. Wiechman and I. Sason, "An improved sphere-packing bound for finite-length codes over symmetric memoryless channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 5, pp. 1962–1990, May 2008. doi: 10.1109/TIT.2008.920216 133, 141

[295] G. D. Forney, Jr., "Convolutional codes II. Maximum-likelihood decoding," *Inf. Control*, vol. 25, no. 3, pp. 222–266, Jul. 1974. doi: 10.1016/S0019-9958(74)90870-5 133

[296] S. Hara, A. Ogino, M. Araki, M. Okada, and N. Morinaga, "Throughput performance of SAW–ARQ protocol with adaptive packet length in mobile packet data transmission," *IEEE Trans. Veh. Technol.*, vol. 45, no. 3, pp. 561–569, Aug. 1996. doi: 10.1109/25.533771 134

[297] E. Modiano, "An adaptive algorithm for optimizing the packet size used in wireless ARQ protocols," *Wireless Netw.*, vol. 5, no. 4, pp. 279–286, Jul. 1999. doi: 10.1023/A:1019111430288 134

[298] P. Lettieri and M. B. Srivastava, "Adaptive frame length control for improving wireless link throughput, range, and energy efficiency," in *Proc. 17th Annu. Joint Conf. IEEE Computer Commun. Soc. (INFOCOM'98)*, vol. 2, Mar. 1998, pp. 564–571. doi: 10.1109/INFCOM.1998.665076 134

[299] S. Ci, H. Sharif, and K. Nuli, "Study of an adaptive frame size predictor to enhance energy conservation in wireless sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 283–292, Feb. 2005. doi: 10.1109/JSAC.2004.839400 134

[300] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Dispersion of Gaussian channels," in *Proc. 2009 IEEE Int. Symp. Inf. Theory*, Jul. 2009, pp. 2204–2208. doi: 10.1109/ISIT.2009.5205834 134

[301] C. E. Shannon, "The zero error capacity of a noisy channel," *IRE Trans. Inf. Theory*, vol. IT-2, no. 3, pp. 8–19, Sep. 1956. doi: 10.1109/TIT.1956.1056798 134

[302] S. Sanghavi, "Intermediate performance of rateless codes," in *Proc. IEEE Inf. Theory Workshop (ITW'07)*, Sep. 2007, pp. 478–482. doi: 10.1109/ITW.2007.4313121 134

[303] A. Kamra, V. Misra, J. Feldman, and D. Rubenstein, "Growth codes: Maximizing sensor network data persistence," in *Proc. 2006 Conf. Appl. Technol. Archit. Protocols Comput. Commun. (SIGCOMM'06)*, Sep. 2006, pp. 255–266. doi: 10.1145/1159913.1159943 134

[304] S. M. Ross, *Stochastic Processes*. John Wiley & Sons, 1996. 135

[305] R. G. Gallager, *Discrete Stochastic Processes*. Boston: Kluwer Academic Publishers, 1996. 135

[306] R. Gallager, *Information Theory and Reliable Communication*, ser. International Centre for Mechanical Sciences, Courses and Lectures. Vienna: Springer-Verlag, 1972, no. 30. 136, 181

[307] S. Tatikonda and S. Mitter, "The capacity of channels with feedback," *IEEE Trans. Inf. Theory*, vol. 55, no. 1, pp. 323–349, Jan. 2009. doi: 10.1109/TIT.2008.2008147 136

[308] M. Mushkin and I. Bar-David, "Capacity and coding for the Gilbert–Elliott channels," *IEEE Trans. Inf. Theory*, vol. 35, no. 6, pp. 1211–1290, Nov. 1989. doi: 10.1109/18.45284 136

[309] A. J. Goldsmith and P. P. Varaiya, "Capacity, mutual information, and coding for finite-state Markov channels," *IEEE Trans. Inf. Theory*, vol. 42, no. 3, pp. 868–886, May 1996. doi: 10.1109/18.490551 136

[310] L. E. Braten and T. Tjelta, "Semi-Markov multistate modeling of the land mobile propagation channel for geostationary satellites," *IEEE Trans. Antennas Propag.*, vol. 50, no. 12, pp. 1795–1802, Dec. 2002. doi: 10.1109/TAP.2002.807441 136

[311] J. Wang, J. Cai, and A. S. Alfa, "New channel model for wireless communications: Finite-state phase-type semi-Markov channel model," in *Proc. IEEE Int. Conf. Commun. (ICC 2008)*, May 2008, pp. 4461–4465. doi: 10.1109/ICC.2008.837 136

[312] S. Wang and J.-T. Park, "Modeling and analysis of multi-type failures in wireless body area networks with semi-Markov model," *IEEE Commun. Lett.*, vol. 14, no. 1, pp. 6–8, Jan. 2010. doi: 10.1109/LCOMM.2010.01.091719 136

[313] P. Kaski and P. R. J. Östergård, *Classification Algorithms for Codes and Designs*. Berlin: Springer, 2006. 141

[314] A. Barg and A. McGregor, "Distance distribution of binary codes and the error probability of decoding," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4237–4246, Dec. 2005. doi: 10.1109/TIT.2005.858977 141

[315] L. Weiss, "On the strong converse of the coding theorem for symmetric channels without memory," *Q. Appl. Math.*, vol. 18, no. 3, pp. 209–214, Oct. 1960. 141

[316] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, 3rd ed. Belmont, MA: Athena Scientific, 2005, vol. 1. 149, 152

[317] I. Dietrich and F. Dressler, "On the lifetime of wireless sensor networks," *ACM Trans. Sensor Netw.*, vol. 5, no. 1, p. 5, Feb. 2009. doi: 10.1145/1464420.1464425 155, 157

[318] B. Canton, A. Labno, and D. Endy, "Refinement and standardization of synthetic biological parts and devices," *Nat. Biotechnol.*, vol. 26, no. 7, pp. 787–793, Jul. 2008. doi: 10.1038/nbt1413 155, 181

[319] J. D. Bekenstein, "The limits of information," *Stud. Hist. Philos. Mod. Phys.*, vol. 32, no. 4, pp. 511–524, Dec. 2001. doi: 10.1016/S1355-2198(01)00020-X 155

[320] D. R. Headrick, *The Invisible Weapon: Telecommunications and International Politics, 1851–1945*. New York: Oxford University Press, 1991. 155

[321] J. Jarzebowski, L. Srinivasan, and T. P. Coleman, "Using stochastic control with data compression perspectives to enhance P300 neural communication prostheses," in *Proc. IEEE Inf. Theory Workshop (ITW'08)*, May 2008, pp. 109–113. doi: 10.1109/ITW.2008.4578632 155

[322] L. Benini and G. De Micheli, *Networks on Chips: Technology and Tools*. San Francisco, CA: Morgan Kaufmann Publishers, 2006. 157

[323] R. P. Feynman, *Feynman Lectures on Computation*. Reading, MA: Addison-Wesley Publishing Company, 1996. 157

[324] C. H. Bennett, "Notes on Landauer's principle, reversible computation, and Maxwell's Demon," *Stud. Hist. Philos. Mod. Phys.*, vol. 34, no. 3, pp. 501–510, Sep. 2003. doi: 10.1016/S1355-2198(03)00039-X 157, 158, 159

[325] ——, "The thermodynamics of computation—a review," *Int. J. Theor. Phys.*, vol. 21, no. 12, pp. 905–940, Dec. 1982. doi: 10.1007/BF02084158 157, 159

[326] E. Fredkin and T. Toffoli, "Conservative logic," *Int. J. Theor. Phys.*, vol. 21, no. 3/4, pp. 219–253, Apr. 1982. doi: 10.1007/BF01857727 157, 158, 181

[327] L. R. Varshney and S. D. Servetto, "A distributed transmitter for the sensor reachback problem based on radar signals," in *Advances in Pervasive Computing and Networking*, B. K. Szymanski and B. Yener, Eds. Boston: Kluwer Academic Publishers, 2005, pp. 225–245. doi: 10.1007/0-387-23466-7_12 157

[328] B. Ananthasubramaniam and U. Madhow, "On localization performance in imaging sensor nets," *IEEE Trans. Signal Process.*, vol. 55, no. 10, pp. 5044–5057, Oct. 2007. doi: 10.1109/TSP.2007.896260 157

[329] J. A. Paradiso and T. Starner, "Energy scavenging for mobile and wireless electronics," *IEEE Pervasive Comput.*, vol. 4, no. 1, pp. 18–27, Jan.-Mar. 2005. doi: 10.1109/MPRV.2005.9 157

[330] R. Dinyari, J. D. Loudin, P. Huie, D. Palanker, and P. Peumans, "A curvable silicon retinal implant," in *Proc. IEEE Int. Electron Devices Meeting (IEDM 2009)*, Dec. 2009. 157, 182

[331] R. R. Harrison, P. T. Watkins, R. J. Kier, R. O. Lovejoy, D. J. Black, B. Greger, and F. Solzbacher, "A low-power integrated circuit for a wireless 100-electrode neural recording system," *IEEE J. Solid-State Circuits*, vol. 42, no. 1, pp. 123–133, Jan. 2007. doi: 10.1109/JSSC.2006.886567 157, 182

[332] R. Want, "Enabling ubiquitous sensing with RFID," *IEEE Computer*, vol. 37, no. 4, pp. 84–86, Apr. 2004. doi: 10.1109/MC.2004.1297315 157, 182

[333] I. Wasserman, D. Hahn, D. H. Nguyen, H. Reckmann, and J. Macpherson, "Mud-pulse telemetry sees step-change improvement with oscillating shear valves," *Oil Gas J.*, vol. 106, no. 24, pp. 39–47, Jun. 2008. 157

[334] M. Gastpar, "On capacity under receive and spatial spectrum-sharing constraints," *IEEE Trans. Inf. Theory*, vol. 53, no. 2, pp. 471–487, Feb. 2007. doi: 10.1109/TIT.2006.889016 158, 162

[335] R. Landauer, "Irreversibility and heat generation in the computing process," *IBM J. Res. Develop.*, vol. 5, no. 3, pp. 183–191, Jul. 1961. 158

[336] A. Ephremides, "Energy concerns in wireless networks," *IEEE Wireless Commun. Mag.*, vol. 9, no. 4, pp. 48–59, Aug. 2002. doi: 10.1109/MWC.2002.1028877 159

[337] D. Attwell and S. B. Laughlin, "An energy budget for signaling in the grey matter of the brain," *J. Cereb. Blood Flow Metab.*, vol. 21, no. 10, pp. 1133–1145, Oct. 2001. doi: 10.1097/00004647-200110000-00001 159

[338] A. F. MacAskill, J. E. Rinholm, A. E. Twelvetrees, I. L. Arancibia-Carcamo, J. Muir, A. Fransson, P. Aspenstrom, D. Attwell, and J. T. Kittler, "Miro1 is a calcium sensor for glutamate receptor-dependent localization of mitochondria at synapses," *Neuron*, vol. 61, no. 4, pp. 541–555, Feb. 2009. doi: 10.1016/j.neuron.2009.01.030 159

[339] S. L. Howard, C. Schlegel, and K. Iniewski, "Error control coding in low-power wireless sensor networks: When is ECC energy-efficient?" *EURASIP J. Wireless Commun. Netw.*, vol. 2006, p. 74812, 2006. doi: 10.1155/WCN/2006/74812 159

[340] S. Kar, S. Aldosari, and J. M. F. Moura, "Topology for distributed inference on graphs," Jun. 2006, arXiv:cs/0606052. 159, 182

[341] M. Bhardwaj, "Communications in the observation limited regime," Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA, Jun. 2009. 161

[342] D. Siegmund, *Sequential Analysis: Tests and Confidence Intervals*. New York: Springer-Verlag, 1985. 161

[343] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, 3rd ed. Budapest: Akadémiai Kiadó, 1997. 162

[344] S. W. Golomb, "The limiting behavior of the Z-channel," *IEEE Trans. Inf. Theory*, vol. IT-26, no. 3, p. 372, May 1980. 165

[345] W. B. Levy and R. A. Baxter, "Using energy efficiency to make sense out of neural information processing," in *Proc. 2002 IEEE Int. Symp. Inf. Theory*, Jul. 2002, p. 18. doi: 10.1109/ISIT.2002.1023290 166

[346] K. Dostert, *Powerline Communications*. Prentice Hall, 2001. 166, 174, 182

[347] J. G. Smith, "The information capacity of amplitude- and variance-constrained scalar Gaussian channels," *Inf. Control*, vol. 18, no. 3, pp. 203–219, Apr. 1971. doi: 10.1016/S0019-9958(71)90346-9 167, 171

[348] L. Itti and P. Baldi, "Bayesian surprise attracts human attention," in *Advances in Neural Information Processing Systems 18*, Yair, B. Schölkopf, and J. Platt, Eds. Cambridge, MA: MIT Press, 2006, pp. 547–554. 167

[349] M. Gastpar, B. Rimoldi, and M. Vetterli, "To code, or not to code: Lossy source-channel communication revisited," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1147–1158, May 2003. doi: 10.1109/TIT.2003.810631 167, 170

[350] J. G. Smith, "On the information capacity of peak and average power constrained Gaussian channels," Ph.D. dissertation, University of California, Berkeley, Berkeley, CA, 1969. 167, 168, 171, 172

[351] M. Loéve, *Probability Theory*, 3rd ed. New York: Van Nostrand Reinhold Company, 1963. 167, 168

[352] R. L. Dobrushin, "Mathematical problems in the Shannon theory of optimal coding of information," in *Proc. 4th Berkeley Symp. Math. Stat. Probab.*, J. Neyman, Ed., vol. 1. Berkeley: University of California Press, 1961, pp. 211–252. 167

[353] R. G. Bartle, *The Elements of Real Analysis*. New York: John Wiley & Sons, 1964. 170, 176

[354] K. Knopp, *Theory of Functions: Parts I and II*. Mineola, NY: Dover Publications, 1996. 170

[355] J. Huang and S. P. Meyn, "Characterization and computation of optimal distributions for channel coding," *IEEE Trans. Inf. Theory*, vol. 5, no. 7, pp. 2336–2351, Jul. 2005. doi: 10.1109/TIT.2005.850108 171

[356] J. Dauwels, "Numerical computation of the capacity of continuous memoryless channels," in *Proc. 26th Symp. Inf. Theory Benelux*, May 2005, pp. 221–228. 171, 172

[357] J. Aczél and Z. Daróczy, *On Measures of Information and Their Characterization*. New York: Academic Press, 1975. 172

[358] N. Wiener, *Cybernetics: or Control and Communication in the Animal and the Machine*, 2nd ed. Cambridge, MA: MIT Press, 1961. 172

[359] W. B. Carlson, "Nikola Tesla and the idea of broadcasting electric power, 1890–1905," in *2007 IEEE Conf. History Electric Power*, Aug. 2007. 174

[360] J. Giles and B. Hajek, "An information-theoretic and game-theoretic study of timing channels," *IEEE Trans. Inf. Theory*, vol. 48, no. 9, pp. 2455–2477, Sep. 2002. doi: 10.1109/TIT.2002.801405 174

[361] M. Schwartz, "The early history of carrier-wave telephony over power lines," in *2007 IEEE Conf. History Electric Power*, Aug. 2007. 174

[362] E. Biglieri, S. Galli, Y.-H. Lee, H. V. Poor, and A. J. H. Vinck, "Power line communications," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 7, pp. 1261–1266, Jul. 2006. doi: 10.1109/JSAC.2006.874398 174

[363] D. G. Luenberger, *Optimization by Vector Space Methods*. New York: John Wiley & Sons, 1969. 174, 176

[364] A. A. Al-Yamani, S. Ramsundar, and D. K. Pradhan, "A defect tolerance scheme for nanotechnology circuits," *IEEE Trans. Circuits Syst. I*, vol. 54, no. 11, pp. 2402–2409, Nov. 2007. doi: 10.1109/TCSI.2007.907875 181

[365] S.-L. Jeng, J.-C. Lu, and K. Wang, "A review of reliability research on nanotechnology," *IEEE Trans. Rel.*, vol. 56, no. 3, pp. 401–410, Sep. 2007. doi: 10.1109/TR.2007.903188 181

[366] M. Feder and N. Merhav, "Relations between entropy and error probability," *IEEE Trans. Inf. Theory*, vol. 40, no. 1, pp. 259–266, Jan. 1994. doi: 10.1109/18.272494 181

[367] T. Koch, A. Lapidoth, and P. P. Sotiriadis, "Channels that heat up," *IEEE Trans. Inf. Theory*, vol. 55, no. 8, pp. 3594–3612, Aug. 2009. doi: 10.1109/TIT.2009.2023753 181

[368] K. Eswaran, "Communication and third parties: Costs, cues, and confidentiality," Ph.D. dissertation, University of California, Berkeley, Berkeley, CA, 2009. 182

[369] A. Mohan, G. Woo, S. Hiura, Q. Smithwick, and R. Raskar, "Bokode: Imperceptible visual tags for camera based interaction from a distance," *ACM Trans. Graphics*, vol. 28, no. 3, p. 98, Aug. 2009. doi: 10.1145/1531326.1531404 182

[370] C. H. Bennett, "Dissipation-error tradeoff in proofreading," *Bio Syst.*, vol. 11, no. 2-3, pp. 85–91, Aug. 1979. 183

[371] J. Campbell, *Grammatical Man: Information, Entropy, Language, and Life*. New York: Simon and Schuster, 1982. 183