QUANTIZATION FOR NOISY CHANNELS USING STRUCTURED CODES

BY

ANDRÁS MÉHES

Dipl., Technical University of Budapest, 1993

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Electrical Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 1998

Urbana, Illinois

QUANTIZATION FOR NOISY CHANNELS USING STRUCTURED CODES

András Méhes, Ph.D.
Department of Electrical and Computer Engineering
University of Illinois at Urbana-Champaign, 1998
Kenneth Zeger, Advisor

Lossy source coding, or quantization, is a data compression technique used in many practical voice and image coding systems. The increasing demand to transmit large amounts of data with very small delay over bandwidth-limited communication channels calls for efficient and error-resilient solutions for source and channel coding. In this dissertation, structured codes are studied with the aim of reducing complexity but retaining performance. In particular, structured classes of source codes, channel codes, and index assignments are investigated.

A thorough study of the index assignment problem is presented for a class of structured source coders we call binary lattice vector quantizers (BLVQ) and linear error correcting codes on binary memoryless channels. Distortion formulas of affine index assignments are derived and their performances compared. The asymptotic performance of index assignments with increasing blocklengths is studied for BLVQ on a uniform source with no channel coding. A "worst" assignment in terms of mean squared error is derived and shown to be affine. The expected distortion of a randomly chosen index assignment is shown to be asymptotically equivalent to that of the worst index assignment, as the blocklength grows. In this sense, randomly chosen index assignments are asymptotically bad for uniform sources.

Bounds for the optimal rate allocation between source and channel coding are developed for the cascade of "good" quantizers and channel codes that meet the Gilbert-Varshamov or Tsfasman-Vlădut-Zink bounds. Corresponding high-resolution distortion formulas are also given. High-resolution distortion bounds based on a rate allocation for the cascade of BLVQs and classical linear block channel codes are also obtained. The

distortion of these practical systems decays to zero exponentially fast as the transmission rate grows, although the exponent is a sublinear function of the transmission rate. As a result, the performance increase per unit rate increase is diminishing for larger transmission rates.

Necessary conditions for the optimality of binary lattice vector quantizers with respect to the mean squared error distortion criterion are derived for both noiseless and noisy discrete memoryless channels.

ACKNOWLEDGMENTS

I would like to thank my advisor Professor Ken Zeger for his guidance and support. I would also like to thank Professors Thomas S. Huang, P. R. Kumar, Sean Meyn, and Dilip V. Sarwate for reading this dissertation and serving on my committee. I wish to thank Professor Alexander Vardy for serving on the committee for my preliminary examination.

Finally, I thank my family, my parents and my brothers. Even from a distance, they were always there for me. I am very grateful for their encouragement and support.

TABLE OF CONTENTS

LIST OF TABLES

Table **Page**

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

This thesis concerns the use of structured codes in a digital communication system to transmit analog data over a noisy communication channel, subject to a fidelity criterion. Given the statistical description of a data source, the statistics of the noisy channel, and a measure of distortion, the goal is to transmit the source samples with minimal average end-to-end distortion, and to do so under practical constraints on delay and complexity.

This introductory chapter describes the results of the dissertation in a broader context, illuminating their relationship with each other and with other work in the field.

1.1 Communication System Model

A standard model which effectively describes many practical digital communication systems, such as speech and image coders, includes a source encoder and decoder, a channel encoder and decoder, and a discrete memoryless channel, as shown in Figure 1.1. The first detailed mathematical analysis of this model was given by Shannon in 1948 [1, 2]. While a more general model might include fully analog systems, or digital systems interfacing an analog channel with a modulator/demodulator pair, the model in Figure 1.1 is sufficiently general to cover a wide range of digital telecommunication applications and allows certain mathematical analyses.

The task of the source encoder is to efficiently represent the source data as a sequence of bits (or symbols) allowing the source decoder to accurately reconstruct the original data

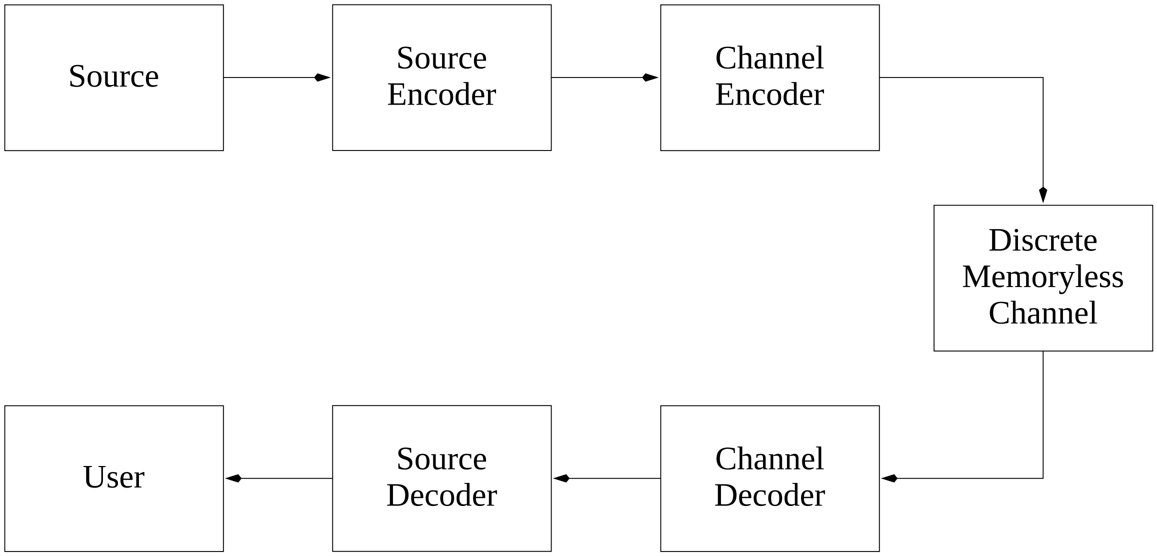


Figure 1.1 Communication system model.

from the encoded stream. Lossless source coding schemes (such as Huffman, Lempel-Ziv, or arithmetic coding) create invertible bit streams, and thus reproduce an exact replica of the original source. Lossy source coding methods, on the other hand, aim to recover only an approximation of the original data, either to produce a more compact representation than lossless coding would make possible, or because the source cannot be represented exactly with a finite number of bits, as is the case for analog sources. The accuracy of a lossy source coder's reproduction is measured by a fidelity criterion, a mathematical function that quantifies the distortion between a source sample and the corresponding approximation. For a fixed transmission rate (i.e., number of bits per source sample), the performance of a lossy source coder is generally measured by the expected value of this distortion. Shannon's *source coding theorem with respect to a fidelity criterion* [3, 4] says that for a given source and source coding rate, the performance of a source coder with a fidelity criterion can be made arbitrarily close to the theoretical optimum (for that source and rate) by coding increasingly longer blocks of source samples. In this dissertation, the source samples are assumed to be random vectors of fixed dimension and known statistics, and the source encoder/decoder is a vector quantizer (VQ). The distortion is given by some positive power of the Euclidean norm of the difference between a source vector and a reconstructed vector. When this positive power is two, one gets the frequently studied mean squared error distortion criterion.

Vector quantization is a powerful lossy block source coding technique [5]. The VQ encoder is characterized by a finite partition of the multidimensional input space of source vectors. For each input vector, the output of the VQ encoder is the index of the encoder region to which the given source vector belongs. To each region of the encoder partition, the VQ decoder associates a reconstruction vector or codevector from a preselected set known as the codebook. Thus, for each received index, the output of the VQ decoder is the codevector corresponding to the encoder partition the given index represents. The blocklength of the source for a VQ is the dimension of the input space and the source coding rate, which is defined as the number of index bits per vector component, is referred to as the quantizer resolution. For norm-based distortion measures, the encoder is commonly implemented as a nearest neighbor search of the codevectors and the decoder as a table lookup operation.

In a traditional setting, the task of the channel encoder is to map the source encoded stream into a stream of channel symbols in such a way that even after possible transmission errors on a noisy channel the channel decoder can with high probability recover the original source encoded stream. In this case, for a fixed channel code rate (the ratio of source encoded symbols to channel symbols), the performance of channel coders is usually measured by the average or expected probability of decoding error. The *channel coding theorem* [2, 6] says that for channel code rates below the capacity of the channel, the probability of decoding error can be made arbitrarily small by coding increasingly larger blocks of channel symbols. The theory and design of channel coders is a thoroughly studied field of research with various branches, and numerous effective coding methods are found in the literature [7–9].

In quantizer systems operating on noisy channels, the role of the channel coder is to reduce the effect of transmission errors on the end-to-end average distortion of the communication system. In this thesis, the channel encoder/decoder is a linear block channel coder with maximum likelihood decoding. Asymptotically good codes satisfying the Gilbert-Varshamov or Tsfasman-Vlădut-Zink bounds, and families of classical error correcting codes such as BCH and Reed-Muller codes are studied in conjunction with vector quantization for a noisy channel.

1.2 Joint Source-Channel Coding

Traditionally, source coders are designed for an ideal noiseless channel, channel coders are designed to minimize the average probability of decoding error irrespective of the actual source, and the resulting encoders and decoders are cascaded as suggested by Figure 1.1. The "separation principle" of Shannon states that in the limit of increasing blocklengths there is no loss of optimality in using the cascade of a block source coder optimized for a given source independent of the channel coder and a channel coder optimized for a given memoryless channel independent of the source coder. However, under practical delay and storage constraints, independent design of source and channel coders is not optimal. This motivates a coupled design of source coders and channel coders. One approach is to use a combined source-channel coder to replace the cascade of a source coder and a channel coder. The complexity of the joint optimization of a combined source-channel coder, however, quickly becomes prohibitive for systems of practical size. Thus, not only does the traditional separation approach require infinite complexity for optimality, but a completely coupled design is also infeasible. This dissertation therefore investigates low complexity techniques which increase the performance of cascaded systems by introducing some amount of coupling between the source coder and the channel coder. In particular, index assignments and source/channel rate allocation are studied for structured codes, from a vector quantization perspective.

Historically, quantization theory was first developed without concern for channel errors. Necessary conditions for the optimality of noiseless quantizers along with an iterative design algorithm based on these conditions were obtained for the scalar case by Lloyd [10] and Max [11], and later generalized for vector quantizers in [12]. The generalized Lloyd algorithm is a useful technique for obtaining good vector quantizers, but the resulting quantizers are often only locally optimal. Sufficient conditions for optimality have been obtained for the scalar case [13, 14], but the optimal design of a vector quantizer in general is unknown. An analytic description of the performance of an optimal system is also lacking. Bennett's integral [15] and Zador's formula [16] provide analytic performance descriptions for asymptotically high resolution quantizers. For large vector dimensions and source coding rates, unstructured quantizers generated with the Lloyd algorithm require extremely large computational complexities (exponential in both dimension and rate) for full-search nearest neighbor encoding. As a result, much recent research has focused on structured (but suboptimal) quantizers which trade off reduced complexity for reduced performance [5]. Examples of structured quantizers include "tree-structured" and "multistage" (or "residual") vector quantizers. The focus of this dissertation is a class of robust structured quantizers we call *binary lattice vector quantizers* (BLVQ). These are a special case of multistage VQ, with two codevectors per stage.

Several studies have considered combined source and channel coding with vector quantization. The three main avenues of research in this field are summarized in the next subsections.

1.2.1 VQ with index assignment

When a vector quantizer operates on an ideal noiseless channel, the index produced by the VQ encoder is received by the VQ decoder unaltered. On a noisy channel, however, the transmitted index may get changed to a different index, resulting in increased distortion. The most direct low-complexity technique for noisy channel quantization is to use a quantizer designed for an ideal noiseless channel and, instead of cascading it with a channel coder for error protection, to find an *index assignment* (i.e., an ordering of quantizer reconstruction vectors) which minimizes the expected distortion given that the codevector indices may get corrupted by channel noise. The idea is to assign indices, which are likely to be mistaken due to transmission errors, to codevectors which are "close" with respect to the given distortion measure. An index assignment can be thought of as a nonredundant channel code which, instead of minimizing the probability of decoding error, minimizes the expected distortion. Finding an optimal index assignment is made difficult by the combinatorial explosion of the number of codevector orderings with increasing transmission rate. Several greedy methods have been proposed for index assignment design including "pseudo-Gray coding" [17], simulated annealing [18,19], Hadamard transform based algorithms [20], and others [21–23]. Other nonredundant methods exploiting a specific quantizer structure are described in [24–26]. Less is known about index assignments analytically. Yamaguchi and Huang [27], Huang [28], and Huang et al. [29] derived mean squared error formulas for the Natural Binary Code and the Gray Code for the case of uniform scalar sources, uniform scalar quantizers, and binary symmetric channels. Huang also asserted the optimality of the Natural Binary Code under these assumptions [28]. The first published proof of Huang's result was given by Crimmins et al. [30], and later extended from uniform scalar quantization to binary lattice vector quantization by McLaughlin, Neuhoff, and Ashley [31]. Optimal index assignments for other sources, quantizers, and channels are not presently known. Analytic bounds on the performance of an optimal index assignment under fairly general conditions are derived in [32].

This thesis advances index assignment research in two directions. Chapter 2 gives a detailed study of structured index assignments for binary lattice vector quantization, including distortion formulas for both symmetric and nonsymmetric channels, a comparison of various useful index assignments, and an extension of the index assignment paradigm to include linear error correcting codes. Chapter 3 revisits the well-studied case of uniform sources, binary lattice quantizers, and binary symmetric channels, and further motivates intelligent index assignment design by showing that the majority of index assignments are asymptotically bad in a certain sense.

1.2.2 Cascade of VQ and channel coder

The traditional "separation approach" of source and channel coding, although not optimal for bounded blocklengths, is still useful and attractive because of its modularity and relatively low complexity. The "independent" design principle of cascaded systems based on the channel coding and source coding theorems of Shannon suggests that the channel coder should operate at a rate close to the capacity of the channel and that the remainder of the available transmission rate should be allocated to the source coder. Experimental results for the cascade of a variety of efficient but suboptimal source coding schemes, such as DPCM and transform coding, with channel codes have been reported in the literature [33–36]. These indicate that given a constraint on the delay and complexity of a cascaded system trading off source coding bits for channel coding bits can result in a significant improvement in performance. Thus, under practical constraints, the optimal channel code rate may be well below capacity. Asymptotic bounds on this optimal rate for the cascade of vector quantizers and binary linear channel coders have been determined for binary symmetric channels by Hochwald and Zeger [37] and for Gaussian channels by Hochwald [38]. However, both of these works exploit the availability of channel codes which achieve the reliability function of the channel. While random coding arguments show the existence of such codes, actual code constructions are presently lacking. Hence, it is of interest to find similar bounds for good but suboptimal channel codes.

In this dissertation, the optimality requirements on the channel coder in [37] are relaxed to include less powerful structured codes. Chapter 4 provides bounds on the asymptotically optimal tradeoff between source and channel coding for classes of channel coders that attain the Gilbert-Varshamov or Tsfasman-Vlădut-Zink bounds. These families of codes are asymptotically good, and their resulting rate allocation bounds, while weaker than those in [37], are very similar in nature. Polynomial constructions for algebraic geometry codes that achieve the Gilbert-Varshamov bound are known [39, 40], but their complexities render them impractical at present. Thus, in Chapter 5 families of error-correcting codes, such as repetition codes, BCH codes, and Reed-Muller codes are considered, and the source coders are taken to be BLVQs. Asymptotic upper bounds on the mean squared distortion of these systems are obtained based on a rate allocation between source and channel coding.

Since channel coding can only provide perfect error-correction in the limit of increasing blocklengths, finding an optimal index assignment is relevant for any practical cascaded system. Also, since channel coders are generally designed to minimize the probability of decoding error, an index assignment is important to map the vector quantizer indices to channel codewords with the aim of minimizing the end-to-end distortion. The channel encoder, the channel, and the channel decoder can be thought of as a new (hopefully less) noisy channel. This effectively reduces the problem to that of a vector quantizer followed by an index assignment on a noisy channel. This is the approach taken in Chapter 2, where results are also given for the performance of cascaded systems using BLVQs, structured index assignments, and linear error-correcting codes.

1.2.3 Channel-optimized VQ

Similar to the necessary conditions for the optimality of unstructured quantizers for an ideal noiseless channel, necessary conditions for the optimality of the encoder and the decoder of noisy channel quantizers (with no explicit index assignment or channel coder) have been derived for the scalar case by Kurtenbach and Wintz [41], and for the vector case by Dunham and Gray [42] and Kumazawa et al. [43]. Source coders satisfying both optimality conditions are often referred to as "channel-optimized vector quantizers" (COVQ), despite the fact that these optimality conditions are not sufficient. In contrast to the noiseless case, sufficient conditions for optimality are not known even for scalar quantizers on a noisy channel. Various studies described modified versions of Lloyd's suboptimal iterative design algorithm for channel-optimized quantizer design and the properties of the resulting quantizers [19, 44–51]. The implementation complexity of channel-optimized vector quantizers is at least that of full-search vector quantizers for noiseless channels, but they do not use a channel coder for protection against channel transmission errors. Even so, for large source coding rates and large vector dimensions, the use of unstructured channel-optimized vector quantizers is severely limited by their complexity. To overcome the complexity of unstructured COVQ, Phamdo et al. [52] considered tree-structured and multistage vector quantizers, and Jafarkhani and Farvardin [53, 54] considered hierarchical table-lookup vector quantizers "matched" to a noisy channel. The resulting structured vector quantizers are called "channel-matched" instead of channel-optimized, because in addition to the suboptimality implied by the structural constraints, the various channel-matched design algorithms are also greedy. Since BLVQs are a subclass of multistage vector quantizers, the above channel-matched approach is also applicable to BLVQs. The special structure of BLVQ, however, also admits a locally optimal design algorithm. Necessary conditions for the optimality of BLVQ for the ideal noiseless case and the noisy channel case (i.e., channel-optimized BLVQ) are derived in Appendix A. A modified Lloyd-algorithm based on these optimality conditions can be used to design locally optimal BLVQs and CO-BLVQs. However, the optimality conditions for noisy channels place increased storage requirements on the CO-BLVQ encoder. Also, CO-BLVQ suffers from the channel-mismatch problem: regular BLVQ outperforms CO-BLVQ for a wide range of channel parameters that do not match the parameters of the channel for which the given CO-BLVQ was designed. Therefore, CO-BLVQ is not studied in this dissertation. There are also open questions in the theory of unstructured COVQ, which are left for future research. Obtaining high-resolution distortion formulas for COVQ, for example, would enable a more thorough comparison with cascaded techniques.

1.3 Thesis Contributions

This section describes in more detail the contributions of this dissertation. The unifying theme of the thesis is the use of structured codes to solve the communication problem outlined in Section 1.1. The perspective chosen is that of noisy channel quantization. Except for Chapter 4, all source coders in this work are binary lattice quantizers. The results of Chapter 4 are also applicable to a large class of BLVQs, but the structure of

BLVQs is not exploited there. The structure and properties of BLVQs are summarized in Section 1.3.1. All channel coders in this dissertation are linear block channel coders. Standard references on error-correcting codes (including other types of channel codes) are [7–9]. The majority of index assignments studied in this thesis are also structured. The rationale for using structured index assignments and the results of the thesis for the index assignment problem are summarized in Section 1.3.2. Finally, Section 1.3.3 contains a summary of rate allocation results for structured codes. Since each chapter is self-contained and includes material previously published or submitted for publication, this thematic presentation of contributions is believed to be illuminating to the reader.

1.3.1 Binary lattice vector quantization

A vector quantizer is commonly decomposed into an encoder given by a partition of the input space and a decoder given by a collection of codevectors. Naturally, BLVQs share this basic VQ structure, but there is an additional constraint on the placement of BLVQ codevectors. The codevectors are required to be a linear combination of their index bits with vector "coefficients" called generator vectors. An additional offset vector allows a nonzero choice for the codevector with index 0. Thus, a b-bit d-dimensional BLVQ codebook is completely specified by b generator vectors and an offset vector, or equivalently by a $d\times(b+1)$ matrix. In contrast, to describe a b-bit d-dimensional unstructured VQ all 2^b of its codevectors must be given, which amounts to a matrix of size $d\times 2^b$. Of course, there is a penalty for the reduction of complexity as illustrated in Figure 1.2, which shows a 4-bit two-dimensional unstructured VQ and BLVQ designed for the same independent Gaussian source. Note how the unrestricted VQ allows a closer match with the circular symmetry of the source. The linear structure of BLVQ, however, turns out to be an advantage on a noisy channel. In [20], Knagenhjelm and Agrell observe a high correlation between the linearity and performance of the index assignments obtained by their algorithms for general sources and quantizers. This motivates the study of binary lattice quantizers on a noisy channel. Also in [20], another view of linear codebooks (that is, BLVQs) is introduced as a linear projection of the b-dimensional hypercube spanned

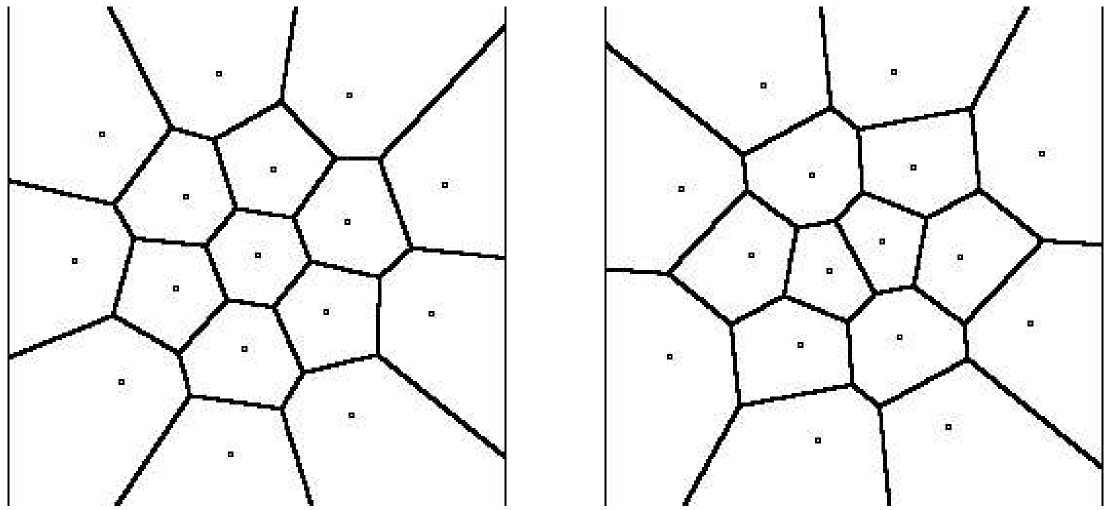


Figure 1.2 Unstructured VQ (left) and BLVQ (right) trained on the same two-dimensional independent Gaussian source.

by the b-bit binary indices to d-dimensional space. Figure 1.3 illustrates the structure of the BLVQ of Figure 1.2. The set of generator vectors is drawn by solid lines, and the dashed-line copies indicate how each point is obtained as a binary linear combination. The two-dimensional projection of a four-dimensional hypercube is easily identifiable.

BLVQs include a large class of useful source coders. A number of equivalent formulations of BLVQ as direct-sum VQ, VQ by a linear mapping of a block code, and truncated lattice VQ are derived in Chapter 2. Figure 1.4 presents two examples: a 3-bit uniform scalar quantizer with a set of generating vectors and an explicit hypercube projection, and a 5-bit BLVQ with generator vectors which span a truncated piece of the well-known hexagonal lattice.

Necessary conditions for the optimality of vector quantizers obtained by a linear mapping of a block code and a corresponding modified Lloyd-algorithm were presented by Hagen and Hedelin in [55]. Since a BLVQ can be viewed as a linear mapping of a nonredundant code, the same algorithm applies for BLVQ design. Appendix A provides a new derivation of the optimality conditions for binary lattice vector quantization not only for the ideal noiseless case but also for noisy channels (i.e., channel-optimized quantization).

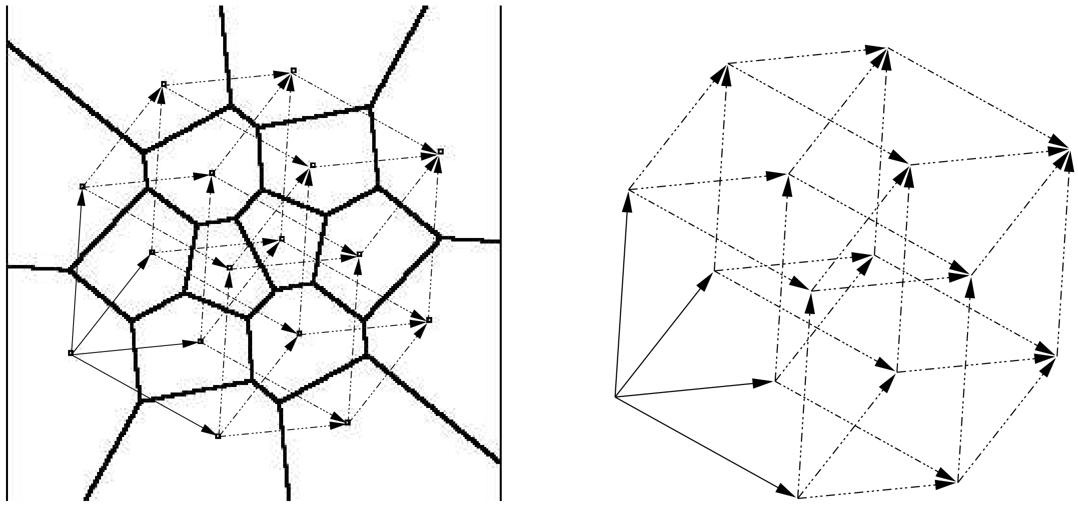


Figure 1.3 The hypercube structure and generator vectors of the BLVQ in Figure 1.2.

The development closely follows the familiar steps of obtaining similar conditions for unstructured quantizers.

Fast nearest neighbor search algorithms developed for unstructured quantization are also applicable to BLVQs. Although the BLVQ structure permits significantly faster greedy encoding algorithms used for general multistage VQ, these algorithms are all suboptimal. Certain subclasses of BLVQ admit a faster nearest neighbor encoding, but no fast uniformly optimal encoding algorithm is presently known. The situation is similar to that of lattice quantizers: very efficient algorithms can be devised for particular lattices, but no known fast algorithm can handle every imaginable lattice equally well. This is not surprising given that the "closest point problem" (the term used for the problem of finding a nearest neighbor in the lattice literature) for lattices is known to be NP-hard [56], and that the class of BLVQs includes appropriately truncated lattice vector quantizers.

Binary lattice vector quantization for noisy channels is extensively studied in Chapters 2, 3, and 5. Chapter 2 treats the index assignment problem for BLVQ with or without error correcting codes. Chapter 3 investigates the asymptotic performance of index assignments for uniform scalar quantization of a uniform source, and shows how

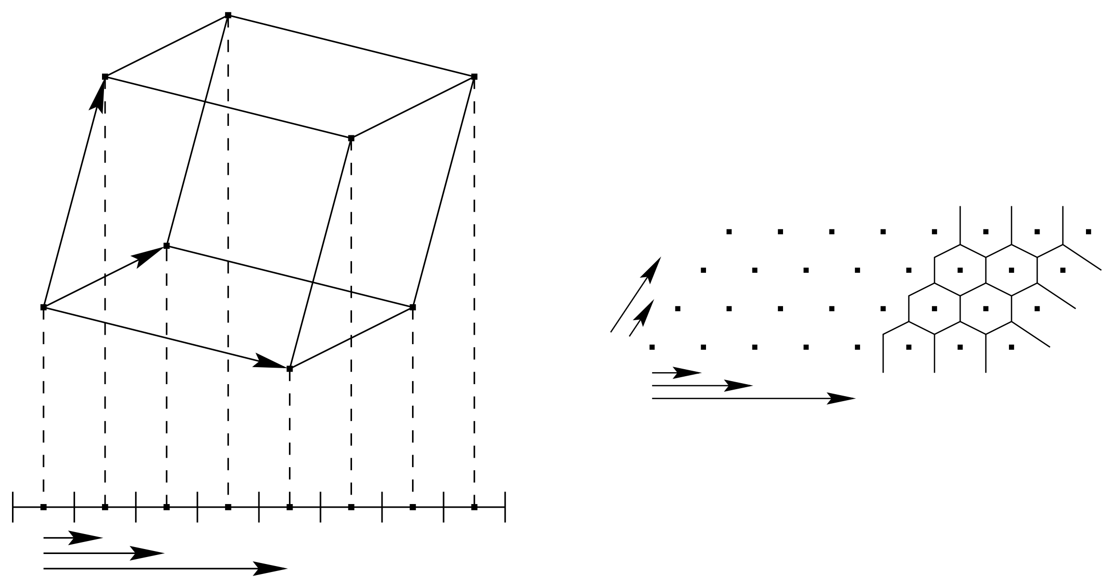


Figure 1.4 The hypercube projection and generator vectors of a 3-bit uniform scalar quantizer (left), and the generator vectors for a 5-bit truncation of the well-known hexagonal lattice (right).

the results extend to general BLVQ. Finally, Chapter 5 obtains high-resolution distortion bounds for the cascade of BLVQ with families of classical algebraic error-correcting codes based on a rate allocation between source and channel coding.

1.3.2 Affine index assignments

When the output of a quantizer encoder designed for an ideal noiseless channel is transmitted to the corresponding quantizer decoder on a noisy channel with no explicit channel coding, the assignment of codebook indices to codevectors becomes important. To minimize distortion, the quantizer codebook should be reordered such that indices that are mistaken with high probability due to channel errors correspond to codevectors whose distance according to the distortion metric is as small as possible. Figure 1.5 shows a best and a worst index assignment for the 3-bit uniform scalar quantizer of Figure 1.4 in terms of mean squared error for a uniform source and a binary symmetric channel. Note, for example, the placement of the 1-bit neighbors of the all-zero index.

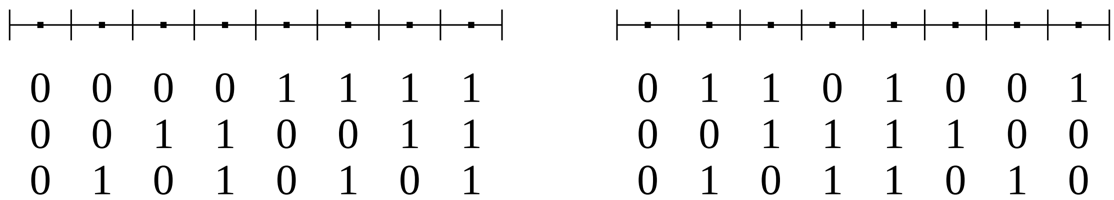


Figure 1.5 A best (left) and a worst (right) index assignment for the 3-bit uniform scalar quantizer of Figure 1.4 in terms of mean squared error for a uniform source and a binary symmetric channel.

A good index assignment can significantly reduce the distortion of a vector quantizer designed for a noiseless channel but used on a noisy channel with no additional transmission rate for error protection. Also, since the entire codebook of an unstructured quantizer must be stored, an index assignment can be implemented with no additional storage in this case by simply reordering the codevectors accordingly. However, when the increased complexity associated with large source coding rates and vector dimensions forces the use of (suboptimal) structured vector quantizers, the codebooks are generally not explicitly stored, and the cost of specifying an arbitrary index assignment separately can be prohibitive. This motivates the study of structured (but possibly suboptimal) classes of index assignments with reduced storage complexities.

Several families of structured index assignments have been studied in the past, including the well-known Natural Binary Code, Folded Binary Code, Two's Complement Code, and Gray Code [57]. Experimental results for the Natural Binary Code, the Folded Binary Code, and the Gray Code were reported in [58–60], for example, for speech sources, but the only analytical results previously available were those of Yamaguchi and Huang [27], Huang [28], and Huang et al. [29] for the mean squared error performance of the Natural Binary Code and the Gray Code for the case of a uniform source, uniform scalar quantizer, and binary symmetric channel. Traditionally, these index assignments were specified recursively, but in fact, all four of them are affine functions in a vector space over the binary field. Affine index assignments can be represented by an invertible binary matrix and a translation vector (also binary). This allows a significant storage reduction compared to unstructured index assignments. For a b-bit quantizer, an affine assignment can be specified by only $(b+1)b$ bits as opposed to the $2^b b$ bits required to describe an arbitrary unstructured index assignment. Generator matrices and translation vectors for the Natural Binary Code, the Folded Binary Code, the Two's Complement Code, and the Gray Code are given in Chapter 2.

Chapter 2 presents a detailed study of affine index assignments for binary lattice vector quantization with or without a binary linear block channel code. An exact analytic expression for the mean squared error performance of such systems with an arbitrary source and a binary symmetric channel is derived based on the Hadamard transform (a version of the Fourier transform for finite groups). The expression is given in terms of the generator vectors of the BLVQ, the generator matrix of the affine index assignment, and the Hadamard transforms of the source and channel statistics. The Hadamard transform of the channel transition probabilities depends on the channel crossover probability and, if a linear block channel code is used, it also depends on the coset weight distribution of the dual code. A similar result is also obtained for the mean squared distortion of a BLVQ for a uniform source followed by an affine index assignment to transmit across a nonsymmetric channel with no explicit channel coding. The general formulas are specialized to the case of the Natural Binary Code, the Folded Binary Code, the Two's Complement Code, and the Gray Code, and the performances of these well-known assignments are compared under various conditions. In particular, it is shown that, although optimal on a binary symmetric channel, the Natural Binary Code is outperformed by the Two's Complement Code on any nonsymmetric channel (and a uniform source). Also, the Folded Binary Code is shown to perform better than the Natural Binary Code for low-variance sources (on a binary symmetric channel). This confirms the experimental findings of Noll [59] for certain speech sources. The material in Chapter 2 appears in part in [61–63] and has been published as [64].

Chapter 3 examines the range of performances achievable using index assignments and BLVQs to transmit a uniform source across a binary symmetric channel. A lower bound on the achievable distortion of these systems is given by the Natural Binary Code, which is known to minimize the mean squared error for the given assumptions [28,30,31]. To obtain an upper bound, a distortion maximizing affine index assignment we call the *Worst Code* is derived. For uniform scalar quantizers and a subclass of BLVQs, the Worst Code is not only the "worst" affine index assignment, but it also maximizes the mean squared error over the set of all possible index assignments. However, a counterexample proves that this property does not extend to the entire class of BLVQs. The main result compares the mean squared error of the Natural Binary Code, the Worst Code, and an "average" index assignment (an index assignment chosen uniformly at random from all possible index assignments). For a given binary symmetric channel in the limit of increasing quantizer resolution, the performance of a randomly chosen index assignment is asymptotically equivalent to that of the Worst Code. This shows that the majority of index assignments perform rather poorly, and thus the search for a good index assignment is indeed very important. The material in Chapter 2 appears in part in [65] and has been accepted for publication as [66].

As shown in [32], the asymptotic performance of any index assignment is bounded away from zero in the limit of increasing resolution. In fact, from the results of Chapter 5 it follows that, depending on how bad the channel is, for a randomly chosen index assignment this limit can be larger than the variance of the underlying source. But a distortion as low as the variance can be achieved with no data transmission by simply reporting the mean of the source to the receiver. Thus, without explicit channel coding, a good index assignment is essential for system performance. Using powerful channel codes, the end-to-end distortion can be forced to decay to zero in the limit of increasing transmission rates independent of the index assignment by properly adjusting the amount of redundancy. How to allocate the available transmission rate between the source coder and the channel coder is the topic of the next subsection. Note, however, that for a fixed transmission rate and source/channel rate allocation, finding a good index assignment is still very important.

1.3.3 Rate allocation for structured codes

Although cascading the best source coder and the best channel coder is not (necessarily) optimal for a fixed transmission rate, the development of source coding and channel coding as virtually independent disciplines over the past few decades shows how widespread the separation approach has become. Solving the two problems independently is conceptually simpler, and the resulting coders have lower complexity than a jointly optimal design would allow. A fundamental question for this modular approach is how to allocate the available transmission rate between source coding and channel coding. Hochwald and Zeger [37] and Hochwald [38] obtained bounds on the asymptotically optimal rate allocation for the cascade of vector quantizers and channel coders operating on binary symmetric channels and Gaussian channels, respectively. Both of these works consider sequences of "good" vector quantizers whose distortion on a noiseless channel decays at the Zador rate and families of channel codes that achieve the reliability function of the channel, which we call "Shannon-optimal" channel codes for easier reference. Various structured classes of quantizers are "good" vector quantizers including certain BLVQs, such as uniform quantizers and lattice-based quantizers. However, no structured classes of Shannon-optimal channel codes are presently known. Therefore, this thesis examines the rate allocation problem for families of structured codes with known constructions.

Chapter 4 extends the existing results of [37] to q-ary symmetric channels, and provides new bounds on the asymptotically optimal tradeoff between source and channel coding for classes of channel coders that attain the Gilbert-Varshamov or Tsfasman-Vlădut-Zink bounds. As a by-product, error exponents including the random coding, sphere packing, and expurgated exponents for q-ary symmetric channels are derived, and appear to be new. Since the channel codes that meet or exceed the Gilbert-Varshamov bound are asymptotically good (i.e., both their channel code rates and their relative minimum distances are bounded away from zero in the limit of increasing blocklength), the rate allocation bounds obtained for these codes reflect only a small penalty when compared to the rate allocation bounds of [37] for Shannon-optimal codes. Similar to the asymptotic behavior of good vector quantizers on a noiseless channel, the overall distortion of the cascade of good vector quantizers and asymptotically good channel coders also decays to zero exponentially fast as the available total transmission rate increases, but the rate of decay is slower due to the noisy channel. Thus, the constant 6 in the famous "6 dB/bit" rule for the distortion decay rate of ideal noiseless quantization is replaced by a smaller constant dependent on the channel, but independent of the transmission rate. The material in Chapter 4 appears in part in [67] and has been submitted for publication as [68].

From a strictly practical perspective, the assumptions of Chapter 4 are still somewhat optimistic. Although infinite families of polynomially constructible codes better than the Gilbert-Varshamov bound are known [39], the complexity of the best known algorithms is presently prohibitive. Chapter 5 investigates the rate allocation problem for known structured codes, namely BLVQs and families of practical binary linear block channel codes, including repetition codes, BCH codes, and Reed-Muller codes. High-resolution distortion bounds based on a rate allocation are derived for these cascaded systems. Since the channel code rates of these families of channel codes approach zero with increasing blocklength, a rate allocation in this case is a decay schedule of how the channel code rate should approach zero as a function of the overall transmission rate. It is shown that by carefully choosing such a decay schedule the mean squared distortion of these systems can be made to decay to zero exponentially fast, although the exponent is a sublinear function of the overall transmission rate. As a result, in contrast to the case of Shannon-optimal or asymptotically good channel codes, there is no fixed dB/bit performance increase. Instead, a case of diminishing returns is observed as the transmission rate grows. The material in Chapter 5 has been submitted for publication as [69].

1.4 References

[1] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423, July 1948.

[2] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 623–656, October 1948.

[3] C. E. Shannon, "Coding theorems for a discrete source with a fidelity criterion," in *IRE National Convention Record, Part 4*, pp. 142–163, 1959.

[4] T. Berger, *Rate Distortion Theory; A Mathematical Basis for Data Compression.* Englewood Cliffs, New Jersey: Prentice-Hall, 1971.

[5] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression.* Boston: Kluwer Academic Publishers, 1992.

[6] R. G. Gallager, *Information Theory and Reliable Communication.* New York: Wiley, 1968.

[7] S. Lin and D. J. Costello, *Error Control Coding: Fundamentals and Applications.* Englewood Cliffs, New Jersey: Prentice-Hall, 1983.

[8] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes.* Amsterdam, The Netherlands: North-Holland, 1977–1993.

[9] M. A. Tsfasman and S. G. Vlăduţ, *Algebraic-Geometric Codes.* Dordrecht: Kluwer Academic Publishers, 1991.

[10] S. P. Lloyd, "Least squares quantization in PCM," *IEEE Trans. Info. Theory*, vol. IT-28, pp. 127–135, March 1982.

[11] J. Max, "Quantizing for minimum distortion," *IEEE Trans. Info. Theory*, pp. 7–12, March 1960.

[12] Y. Linde, A. Buzo, and R. M. Gray, "An algorithm for vector quantizer design," *IEEE Trans. Communication Technology*, vol. COM-28, pp. 84–95, January 1980.

[13] P. E. Fleischer, "Sufficient conditions for achieving minimum distortion in a quantizer," in *IEEE Int. Conv. Rec., Part 1*, pp. 104–111, 1964.

[14] A. V. Trushkin, "Sufficient conditions for uniqueness of a locally optimal quantizer for a class of convex error weighting functions," *IEEE Trans. Info. Theory*, vol. IT-28, pp. 187–198, March 1982.

[15] W. R. Bennett, "Spectra of quantized signals," *Bell Syst. Tech. J.*, vol. 27, pp. 446–472, July 1948.

[16] P. Zador, "Asymptotic quantization error of continuous signals and the quantization dimension," *IEEE Trans. Info. Theory*, vol. IT-28, pp. 139–149, March 1982.

[17] K. Zeger and A. Gersho, "Pseudo-Gray coding," *IEEE Trans. Communications*, vol. COM-38, pp. 2147–2158, December 1990.

[18] D. J. Goodman and T. J. Moulsley, "Using simulated annealing to design transmission codes for analogue sources," *Electronics Letters*, vol. 24, pp. 617–618, May 1988.

[19] N. Farvardin, "A study of vector quantization for noisy channels," *IEEE Trans. Info. Theory*, vol. IT-36, pp. 799–809, July 1990.

[20] P. Knagenhjelm and E. Agrell, "The Hadamard transform – a tool for index assignment," *IEEE Trans. Info. Theory*, vol. IT-42, pp. 1139–1151, July 1996.

[21] N. Rydbeck and C.-E. W. Sundberg, "Analysis of digital errors in nonlinear PCM systems," *IEEE Trans. Communications*, vol. 24, pp. 59–65, January 1976.

[22] J. D. Marca and N. Jayant, "An algorithm for assigning binary indices to the codevectors of a multi-dimensional quantizer," in *Proceedings IEEE Inter. Conf. on Communications*, vol. 1, (Seattle), pp. 1128–1132, June 1987.

[23] H.-S. Wu and J. Barba, "Index allocation in vector quantisation for noisy channels," *Electronics Letters*, vol. 29, pp. 1318–1320, July 1993.

[24] N. Moayeri, "Joint source and channel coding based on tree structured vector quantization," in *CISS-89*, (Baltimore, MD), pp. 729–734, March 1989.

[25] K. C. Chua, W. C. Wong, and K. N. Ngan, "Error detection and correction of vector quantised digital images," *IEE Proceedings*, vol. 137, pp. 417–423, December 1990.

[26] R.-Y. Wang, E. A. Riskin, and R. Ladner, "Codebook organization to enhance maximum a posterori detection of progressive transmission of vector quantized images over noisy channels," in *ICASSP-93*, vol. V, (Minneapolis, MN), pp. 233–236, April 1993.

[27] Y. Yamaguchi and T. S. Huang, "Optimum binary fixed-length block codes," Quarterly Progress Report 78, M.I.T. Research Lab. of Electronics, Cambridge, Mass., July 1965.

[28] T. S. Huang, "Optimum binary code," Quarterly Progress Report 82, M.I.T. Research Lab. of Electronics, Cambridge, Mass., July 1966.

[29] T. S. Huang et al., "Design considerations in PCM transmission of low-resolution monochrome still pictures," *Proceedings of the IEEE*, vol. 55, pp. 331–335, March 1967.

[30] T. R. Crimmins, H. M. Horwitz, C. J. Palermo, and R. V. Palermo, "Minimization of mean-square error for data transmitted via group codes," *IEEE Trans. Info. Theory*, vol. IT-15, pp. 72–78, January 1969.

[31] S. W. McLaughlin, D. L. Neuhoff, and J. J. Ashley, "Optimal binary index assignments for a class of equiprobable scalar and vector quantizers," *IEEE Trans. Info. Theory*, vol. IT-41, pp. 2031–2037, November 1995.

[32] K. Zeger and V. Manzella, "Asymptotic bounds on optimal noisy channel quantization via random coding," *IEEE Trans. Info. Theory*, vol. IT-40, pp. 1926–1938, November 1994.

[33] J. W. Modestino and D. G. Daut, "Combined source-channel coding of images," *IEEE Trans. Communications*, vol. COM-27, pp. 1644–1659, November 1979.

[34] J. W. Modestino, D. G. Daut, and A. L. Vickers, "Combined source-channel coding of images using the block cosine transform," *IEEE Trans. Communications*, vol. COM-29, pp. 1261–1274, September 1981.

[35] D. R. Comstock and J. D. Gibson, "Hamming coding of DCT-compressed images over noisy channels," *IEEE Trans. Communications*, vol. COM-32, pp. 856–861, July 1984.

[36] K. Sayood, F. Liu, and J. D. Gibson, "A constrained joint source/channel coder design," *IEEE J. Sel. Areas Comm.*, vol. JSAC-12, pp. 1584–1593, December 1994.

[37] B. Hochwald and K. Zeger, "Tradeoff between source and channel coding," *IEEE Trans. Info. Theory*, vol. IT-43, pp. 1412–1424, September 1997.

[38] B. Hochwald, "Tradeoff between source and channel coding on a gaussian channel," to appear in *IEEE Trans. Info. Theory.*

[39] G. L. Katsman, M. A. Tsfasman, and S. G. Vlădut, "Modular curves and codes with a polynomial construction," *IEEE Trans. Info. Theory*, vol. IT-30, pp. 353–355, March 1984.

[40] C. Voss and T. Høholdt, "An explicit construction of a sequence of codes attaining the Tsfasman-Vlădut-Zink bound the first steps," *IEEE Trans. Info. Theory*, vol. IT-43, pp. 128–135, January 1997.

[41] A. J. Kurtenbach and P. A. Wintz, "Quantizing for noisy channels," *IEEE Trans. Communication Technology*, vol. COM-17, pp. 291–302, April 1969.

[42] J. Dunham and R. M. Gray, "Joint source and noisy channel trellis encoding," *IEEE Trans. Info. Theory*, vol. IT-27, pp. 516–519, July 1981.

[43] H. Kumazawa, M. Kasahara, and T. Namekawa, "A construction of vector quantizers for noisy channels," *Electronics and Engineering in Japan*, vol. 67-B, no. 4, pp. 39–47, 1984.

[44] E. Ayanoğlu and R. M. Gray, "The design of joint source and channel trellis waveform coders," *IEEE Trans. Info. Theory*, vol. IT-33, pp. 855–865, November 1987.

[45] N. Farvardin and V. Vaishampayan, "Optimal quantizer design for noisy channels: an approach to combined source-channel coding," *IEEE Trans. Info. Theory*, vol. IT-33, pp. 827–838, November 1987.

[46] K. Zeger and A. Gersho, "Vector quantization design for memoryless noisy channels," in *Proc. IEEE Int. Conf. Commun.*, (Philadelphia, PA), June 1988.

[47] N. Farvardin and V. Vaishampayan, "On the performance and complexity of channel-optimized vector quantizers," *IEEE Trans. Info. Theory*, vol. IT-37, pp. 155–160, January 1991.

[48] D. Miller and K. Rose, "Combined source-channel vector quantization using deterministic annealing," *IEEE Trans. Communications*, vol. COM-42, pp. 347–356, February/March/April 1994.

[49] K. Chang and R. W. Donaldson, "Analysis, optimization, and sensitivity study of differential PCM systems operating on noisy communication channels," *IEEE Trans. Communications*, vol. COM-20, pp. 338–350, June 1972.

[50] Q. Chen and T. R. Fischer, "Robust quantization for image coding and noisy digital transmission," in *Proc. Data Compression Conf.*, (Snowbird, UT), pp. 3–12, March 1996.

[51] Q. Chen and T. R. Fischer, "Image coding using robust quantization for noisy digital transmission," *IEEE Trans. Image Processing*, vol. 7, pp. 496–505, April 1998.

[52] N. Phamdo, N. Farvardin, and T. Moriya, "A unified approach to tree-structured and multistage vector quantization for noisy channels," *IEEE Trans. Info. Theory*, vol. IT-39, pp. 835–850, May 1993.

[53] H. Jafarkhani and N. Farvardin, "Channel-matched hierarchical table-lookup vector quantization for transmission of video over wireless channels," in *Proc. IEEE Int. Conf. Image Proc.*, vol. 3, (Lausanne, Switzerland), p. 755, September 1996.

[54] H. Jafarkhani and N. Farvardin, "Channel-matched hierarchical table-lookup vector quantization for finite-state channels without feedback," in *Proc. IEEE Int. Symp. Information Theory*, (Ulm, Germany), p. 512, June 1997.

[55] R. Hagen and P. Hedelin, "Robust vector quantization by a linear mapping of a block-code," to appear in *IEEE Trans. Info. Theory.*

[56] P. van Emde Boas, "Another NP-complete partition problem and the complexity of computing short vectors in a lattice," Report 81-04, Mathematical Institute, University of Amsterdam, Amsterdam, 1981.

[57] N. S. Jayant and P. Noll, *Digital Coding of Waveforms: Principles and Applications to Speech and Video.* Englewood Cliffs, New Jersey: Prentice-Hall, 1984.

[58] I. Dostis, "The effect of digital errors on PCM transmission of compandored speech," *Bell Syst. Tech. J.*, vol. 44, pp. 2227–2243, December 1965.

[59] P. Noll, "Effects of channel errors on the signal-to-noise performance of speech encoding systems," *Bell Syst. Tech. J.*, vol. 54, pp. 1615–1636, November 1975.

[60] R. Zelinski, "Effects of transmission errors on the mean-squared error performance of transform coding systems," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-27, pp. 531–537, October 1979.

[61] A. Méhes and K. Zeger, "Redundancy free codes for binary discrete memoryless channels," in *Proc. Conf. Information Science and Systems*, (Princeton, NJ), pp. 1057–1062, 1994.

[62] A. Méhes and K. Zeger, "On the performance of affine index assignments for redundancy free source-channel coding," in *Proc. Data Compression Conf.*, (Snowbird, UT), p. 433, March 1995.

[63] A. Méhes and K. Zeger, "Affine index assignments for binary lattice quantization with channel noise," in *Proc. IEEE Int. Symp. Information Theory*, (Whistler, BC, Canada), p. 377, September 1995.

[64] A. Méhes and K. Zeger, "Binary lattice vector quantization with linear block codes and affine index assignments," *IEEE Trans. Info. Theory*, vol. IT-44, pp. 79–95, January 1998.

[65] A. Méhes and K. Zeger, "Randomly chosen index assignments are asymptotically bad for uniform sources," in *Proc. IEEE Int. Symp. Information Theory*, (Cambridge, MA), August 1998.

[66] A. Méhes and K. Zeger, "Randomly chosen index assignments are asymptotically bad for uniform sources," to appear in *IEEE Trans. Info. Theory.*

[67] A. Méhes and K. Zeger, "Tradeoff between source and channel coding for codes satisfying the Gilbert-Varshamov bound," in *Proc. Conf. Information Science and Systems*, (Baltimore, MD), March 1997.

[68] A. Méhes and K. Zeger, "Source and channel rate allocation for channel codes satisfying the Gilbert-Varshamov or Tsfasman-Vlăduț-Zink bounds," submitted to *IEEE Trans. Info. Theory.*

[69] A. Méhes and K. Zeger, "Performance of quantizers on noisy channels using structured families of codes," submitted to *IEEE Trans. Info. Theory.*

CHAPTER 2

BINARY LATTICE VECTOR QUANTIZATION WITH LINEAR BLOCK CODES AND AFFINE INDEX ASSIGNMENTS

In this chapter, we determine analytic expressions for the performance of some low-complexity combined source-channel coding systems. The main tool used is the Hadamard transform. In particular, we obtain formulas for the average distortion of binary lattice vector quantization with affine index assignments, linear block channel coding, and a binary symmetric channel. The distortion formulas are specialized to nonredundant channel codes for a binary symmetric channel, and then extended to affine index assignments on a binary asymmetric channel. Various structured index assignments are compared. Our analytic formulas provide a computationally efficient method for determining the performance of various coding schemes. One interesting result shown is that for a uniform source and uniform quantizer, the Natural Binary Code is never optimal for a nonsymmetric channel, even though it is known to be optimal for a symmetric channel.

2.1 Introduction

A useful and frequently studied communication system model includes a source encoder and decoder, a channel encoder and decoder, a noisy channel, and a mapping of source codewords to channel codewords (known as an index assignment). We consider

The material in this chapter has been published as: A. Méhes and K. Zeger, "Binary lattice vector quantization with linear block codes and affine index assignments," *IEEE Transactions on Information Theory*, vol. IT-44, pp. 79–95, January 1998.

the situation where the source encoder/decoder is a vector quantizer (VQ), the channel encoder/decoder is a binary linear block code with maximum likelihood decoding, and the channel is binary and memoryless, as shown in Figure 2.1. The source is assumed to be a random vector of a fixed dimension and whose statistics are known *a priori*. The end-to-end vector mean squared error (MSE) is used to measure the performance.

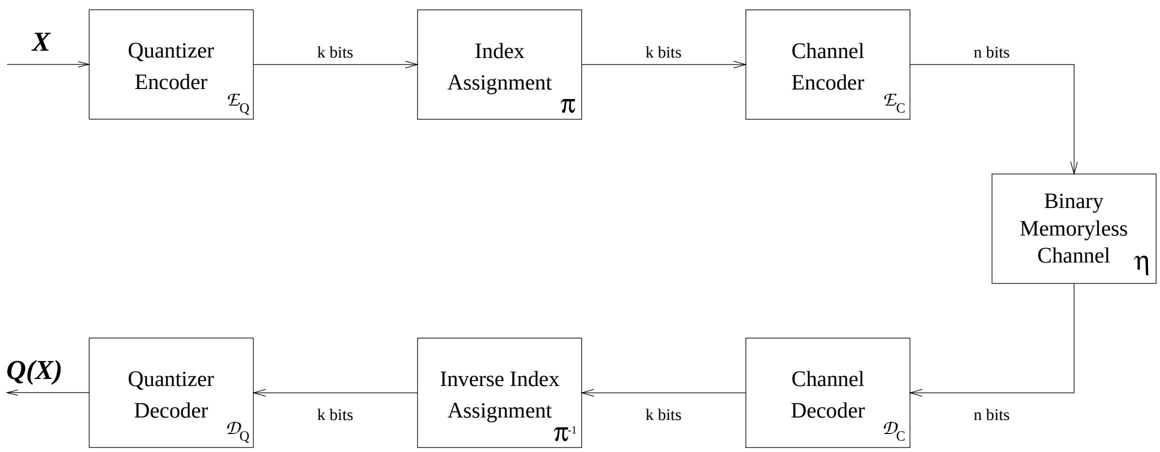


Figure 2.1 Communication system model.

Ideally, one would optimize the end-to-end MSE over all possible choices of source encoders and decoders, channel encoders and decoders, and index assignments. But because of the large computational complexity of this task, it is presently unknown how to perform the joint optimization. The most common approach to finding good, but suboptimal, systems is to assume that all but one component of the system is fixed and then to optimize the choice of that component. Even this suboptimal approach is often algorithmically very complex, and it is generally difficult to quantify the performance analytically. Finding good algorithms and acquiring theoretical understanding of their performance are two of the most important research goals in this field.

Even when the channel is noiseless, the optimal design of a source coder is in general unknown, as is an analytic description of the performance of an optimal system. The well-known generalized Lloyd algorithm is a useful technique for obtaining good, but possibly suboptimal, vector quantizers, and the Bennett-Zador formulas give analytic performance descriptions for asymptotically high resolution quantizers [1]. For large vector dimensions and source coding rates, quantizers generated with the Lloyd algorithm can require extremely large computational complexities (linear in the codebook size) for full-search nearest neighbor encoding. As a result, much recent research has focused on structured (but suboptimal) quantizers which trade off reduced complexity for reduced performance [1]. One example of a structured quantizer is a "multistage vector quantizer" (sometimes called a "residual quantizer"). A special case, with two codevectors per stage, is referred to here as a "binary lattice vector quantizer" and is studied in this chapter.

A number of studies have considered the communication system in Figure 2.1 when the channel is noisy. Optimality conditions and a suboptimal design algorithm for the quantizer encoder and decoder (with a nonredundant channel coder) have been derived for the scalar case by Kurtenbach and Wintz [2], and for the vector case by Dunham and Gray [3] and Kumazawa et al. [4], and were further studied in [5–13]. The resulting source coder is often referred to as "channel optimized vector quantization" (COVQ) and obeys generalized versions of the well-known Nearest Neighbor Condition and Centroid Condition. Very little is known analytically about the performance of these quantizers, and their implementation complexity is at least that of a full-search vector quantizer for a noiseless channel. Thus, their usefulness diminishes as the source vector dimension increases.

A useful technique to combat channel noise and avoid the large complexity of COVQ is to design a source coder for the noiseless channel and cascade it with an error control code. Results for the cascade of a variety of efficient but suboptimal source coding schemes, such as DPCM and transform coding, with channel codes have been reported in the literature [14–17], but few similar results exist for vector quantizers followed by channel coding. For a given transmission rate and fixed vector dimension, the optimal tradeoff between source and channel coding is examined in [18] for high-resolution quantization. However, little else is known theoretically about this problem, other than Shannon's rate-distortion theorem, which assumes unboundedly large source vector dimensions [19].

Also, little is known about good index assignments when error control codes are used, i.e., assignments of quantizer codevectors to channel codewords.

Another approach to source coding in the presence of channel noise has been to use, on a noisy channel, a source coder designed for a noiseless channel, but with an optimized index assignment and with no explicit channel coder [8, 20–24]. Other nonredundant methods exploiting specific quantizer structures can be found in [25–27]. In [28], a random coding argument is used to give analytic bounds on the performance of an optimal index assignment. One appealing feature of index assignments is that they require no extra channel rate or any extra storage to implement; index assignments are implicitly contained in the ordering of the codevectors in the vector quantizer codebook. However, for large source coding rates and high vector dimensions, the increased complexity of full-search vector quantization often forces system designers to implement structured (and thus suboptimal) source coders. In this case, quantizer codebooks are generally not stored explicitly, and the cost of specifying an index assignment can be equally prohibitive.

This motivates the study of structured (but possibly suboptimal) index assignments with low implementation complexities. Various families of recursively defined index assignments have been extensively studied in the past, including the well-known Natural Binary Code, Folded Binary Code, Two's Complement Code, and Gray Code [29]. Yamaguchi and Huang [30], Huang [31], and Huang et al. [32] computed distortion formulas for the Natural Binary Code and the Gray Code for uniform scalar quantizers and uniform scalar sources. Huang asserted that the Natural Binary Code was optimal among all possible index assignments for the uniform source [31]. This was proven by Crimmins et al. [33] and later in the more general setting of binary lattice vector quantization by McLaughlin, Neuhoff, and Ashley [34]. The exact performance of structured classes of index assignments has not been generally known except for the Natural Binary Code and the Gray Code, and with a uniform source. Experimental results for the NBC, FBC, and GC can be found in [35–37], for example, for speech sources. One of the interesting features of the four index assignments above is that they are all "affine" functions in a vector space over the binary field. In fact, affine index assignments are relatively easy to implement with low storage and computational complexity. Specifying an affine index assignment requires only $O(k^2)$ bits for a 2^k-point quantizer, as opposed to $O(k2^k)$ bits for an unconstrained index assignment.

Affine index assignments have been studied for several decades as an effective zero-redundancy technique for source coders that transmit across noisy channels. Linear index assignments are special cases of affine index assignments. They are also special cases of nonsystematic linear block channel codes whose minimum distance is 1, and their purpose is to reduce end-to-end MSE instead of reducing the probability of channel error. Crimmins et al. [33], Crimmins and Horwitz [38], and Crimmins [39] showed that for uniform scalar quantization of a uniform source, using a linear block code and standard array decoding for transmission over a binary symmetric channel, there exists a linear index assignment that is optimal in the MSE sense. They use a binary alphabet and assume that both the encoding and the decoding index assignments are one-to-one mappings. Redinbo and Wolf extended these results in two directions. In [40] they generalized to q-ary (prime-power) alphabets, and in [41] they allowed the decoder mapping to produce outputs outside the codebook (e.g., linear combinations of codevectors). Ashley considered channel redundancy for uniform scalar quantizers [42], and obtained a formula for the MSE in terms of the weight distribution of the cosets of the dual code. Khayrallah examined the problem of finding the best linear index assignment when an error control code is used with a uniform scalar quantizer on a uniform source [43].

In this chapter, we derive exact formulas for the performance of general affine index assignments when explicit block channel coding is used on a binary symmetric channel. We also derive related formulas for the performance of index assignments on binary asymmetric channels, with no explicit channel coding. These are specialized to several known classes of index assignments. As an interesting special case, we show that while the Natural Binary Code is optimal on the binary symmetric channel for uniform sources, it is inferior in general to the Two's Complement Code on the binary asymmetric channel.

In order for a channel optimized quantizer to perform optimally, a good estimate of the channel's bit error rate is required. In this chapter we study a reduced complexity structured vector quantizer combined with an affine index assignment, which together give enhanced channel robustness over a wide range of error rates. We consider *binary lattice vector quantizers* (BLVQs), the class of source coders studied in [34], and a variant of the VQ by a linear mapping of a block code introduced by Hagen and Hedelin [44–46].

Another motivation for studying BLVQ is its inherent robustness to channel noise, in particular under "channel mismatch" conditions, i.e., when the exact level of channel noise is not perfectly known. While channel-optimized vector quantization is an optimal encoding technique if the statistics of a noisy channel are known, a quantizer designed for a noise-free channel using the generalized Lloyd algorithm, referred to here as source-optimized VQ (SOVQ), delivers nearly optimal performance for small effective (i.e., after channel coding) bit error rates. As an example, Figure 2.2 compares the performance of SOVQ, COVQ, and BLVQ for a Gauss-Markov source with correlation coefficient 0.9 using a $(16, 11, 4)$ extended Hamming code. The plot displays signal-to-noise ratio versus the bit error rate of the binary symmetric channel. The signal-to-noise ratio is defined as $10 \log_{10} \frac{\sigma^2}{D/d}$, where σ^2 is the variance of the source components, D is the average vector distortion, and d is the vector dimension of the source. The source vector dimension of the Hamming coded system is 16. All three quantizers were obtained using appropriate variants of the generalized Lloyd algorithm. The COVQ was designed for the coded channel at the uncoded (BSC) bit error rate of 0.1 (a bit error rate that can occur in certain low-power radio channels and near cell boundaries in cellular telephony).

A trade-off between structured (e.g., BLVQ) and unstructured (e.g., SOVQ) quantizers can be observed over the range of error probabilities where the channel code is effective (i.e., the coded channel can be considered practically noise-free). But as the coding advantage disappears the BLVQ outperforms the SOVQ. The COVQ is inferior to the SOVQ and BLVQ under channel mismatch for small BERs and outperforms the SOVQ and BLVQ for large BERs. Thus BLVQ can offer a reasonable compromise. The BLVQ is uniformly robust and close to optimum over a large range of error rates. The price paid for the memory savings resulting from the structured codebook of the BLVQ

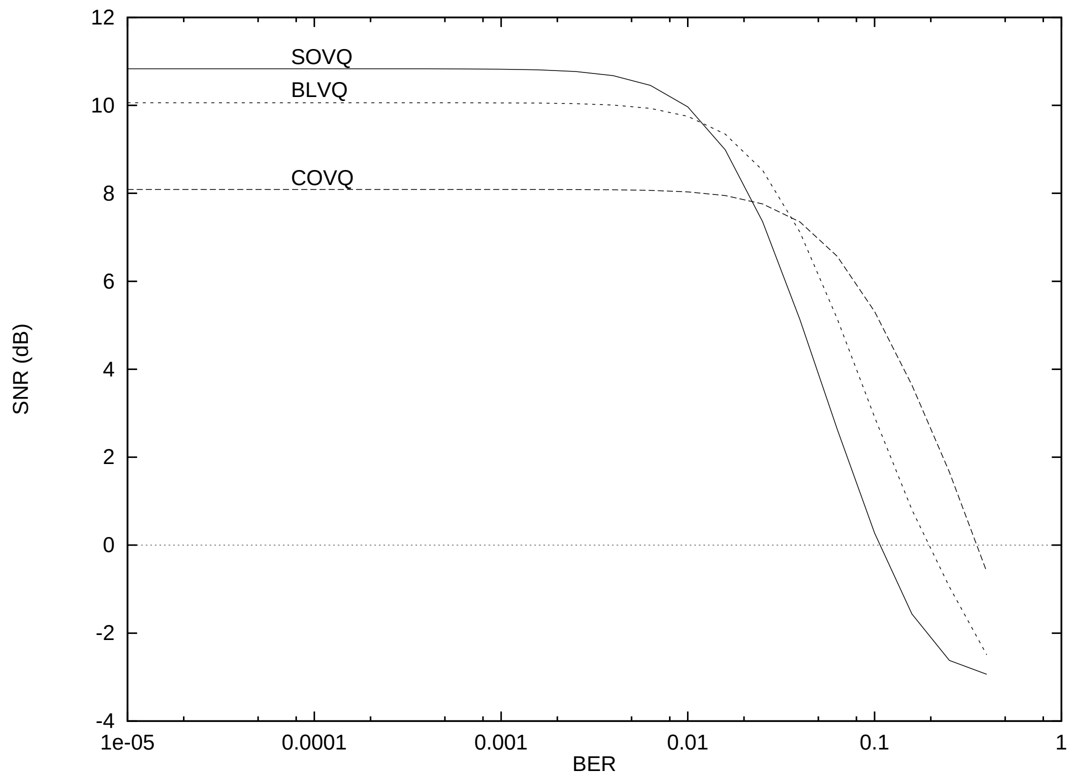


Figure 2.2 The channel-mismatch performance of source optimized VQ, channel optimized VQ, and binary lattice VQ. The input vectors are taken from a Gauss-Markov process with correlation 0.9. The COVQ was designed for BER = 0.1. The 16-dimensional 2048-point quantizers are followed by a $(16, 11, 4)$ extended Hamming code.

is relatively small. Figure 2.2 is not meant to be a comprehensive comparison between SOVQ, BLVQ, and COVQ, but rather is to partially motivate the study of BLVQ.

In this chapter we generalize much of the previously mentioned work to the case of redundant channel codes and BLVQs. We make extensive use of the Hadamard transform, which has been used either implicitly or explicitly in many previous works. The Hadamard transform is also the main tool used by Hagen and Hedelin to construct implicit index assignments without error control coding [44–46]. Also, Knagenhjelm [47], and Knagenhjelm and Agrell [23] use the notion of "Hadamard classes" to search for an optimal index assignment in the Hadamard transform domain. The main contributions of this chapter include: (1) a generalization of the analytic performance calculations of

Hagen and Hedelin for BLVQ, to include error control coding (equivalently, a generalization of the Crimmins et al. formulas to nonuniform sources and to vector quantizers); (2) analytic performance calculations for nonredundant channel coding, which extend the formulas obtained by Huang, by Crimmins et al., and by McLaughlin, Neuhoff, and Ashley from the NBC and GC to any affine index assignment and nonsymmetric channels; (3) comparison between the performances of NBC, FBC, GC, and TCC.

In Section 2.2, we give the necessary notation and terminology. In Section 2.3, we prove Theorem 2.1, which gives a general formula for the channel distortion of a BLVQ using an affine index assignment, a linear error correcting code, and transmission across a BSC. The formula is given in terms of the Hadamard transforms of the source and channel statistics. Our formula reduces the complexity of computing the distortion from $O(N^2)$ to $O(N \log^2 N)$, where N is the vector quantizer codebook size. In Section 2.4, we consider quantization systems without the use of redundancy for error control. For binary symmetric channels, Corollaries 2.2-2.4 give explicit formulas for the channel distortions of the NBC, FBC, and GC in this case. Corollary 2.5 characterizes the class of sources for which the NBC outperforms the FBC. For binary asymmetric channels, Theorem 2.2 gives a general formula for the channel distortion of BLVQ using affine index assignments and no channel coding redundancy. Corollaries 2.6-2.8 give explicit formulas for the channel distortions of the NBC, TCC, FBC, and GC in this case. Corollary 2.9 identifies the best assignment among all affine translates of the NBC for a nonsymmetric channel. Finally, Theorem 2.3 gives explicit comparisons between the performances of the NBC, TCC, FBC, and GC for all possible binary asymmetric channels. In particular, it is shown that the TCC outperforms the other three codes for most useful bit error probabilities, when the channel is nonsymmetric.

2.2 Definitions

2.2.1 Noisy channel VQ with index assignment

For any positive integer k, let $\mathbb{Z}_2^k$ denote the field of k-bit binary words, where arithmetic is performed modulo 2. The results in this chapter are given for binary channels (although generalization to more general channels can be made).

Notation For any binary k-tuple $i \in \mathbb{Z}_2^k$, we write $i = [i_{k-1}, i_{k-2}, \dots, i_1, i_0]$, where $i_l \in \{0, 1\}$ denotes the coefficient of 2^l in the binary representation of i, i.e., $i = \sum_{l=0}^{k-1} i_l 2^l$. For any Euclidean vector $\mathbf{x} \in \mathbb{R}^d$, we write $\mathbf{x} = (x_1, x_2, \dots, x_d)^t$, where x_i is the i^{th} component of $\mathbf{x}$.

In this chapter we assume for convenience that elements of any Euclidean space $\mathbb{R}^d$ are column vectors, whereas we assume that elements of any Hamming space $\mathbb{Z}_2^k$ are binary row vectors. We denote the inner product of two binary vectors $i, j \in \mathbb{Z}_2^k$ by $ij^t = \sum_{l=0}^{k-1} i_l j_l \in \{0, 1\}$, and the inner product of two Euclidean vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$ by $\langle \mathbf{x} \,|\mathbf{y} \rangle = \sum_{l=1}^{d} x_l y_l \in \mathbb{R}$. The following definition corresponds to Figure 2.1.

Definition 2.1 A d-dimensional, 2^k-point *noisy channel vector quantizer* with *codebook* $\mathcal{Y} = \{\mathbf{y}_i \in \mathbb{R}^d \colon i \in \mathbb{Z}_2^k\}$, and with an (n, k) channel code $\mathcal{C} = \{c_i \colon i \in \mathbb{Z}_2^k\} \subset \mathbb{Z}_2^n$, is a functional composition $\mathcal{Q} = \mathcal{D}_Q \circ \pi^{-1} \circ \mathcal{D}_C \circ \eta \circ \mathcal{E}_C \circ \pi \circ \mathcal{E}_Q$, where $\mathcal{E}_Q \colon \mathbb{R}^d \to \mathbb{Z}_2^k$ is a *quantizer encoder*, $\mathcal{D}_Q \colon \mathbb{Z}_2^k \to \mathcal{Y}$ is a *quantizer decoder*, $\mathcal{E}_C \colon \mathbb{Z}_2^k \to \mathcal{C}$ is a *channel encoder*, $\mathcal{D}_C \colon \mathbb{Z}_2^n \to \mathbb{Z}_2^k$ is a *channel decoder*, $\pi \colon \mathbb{Z}_2^k \to \mathbb{Z}_2^k$ is an *index assignment* (bijection), and $\eta \colon \mathbb{Z}_2^n \to \mathbb{Z}_2^n$ is a random mapping representing a noisy channel.

When $n = k$, we say that the noisy channel vector quantizer has a *nonredundant channel code*. Let $\mathbf{X}$ be a random vector in $\mathbb{R}^d$. Let $p_i = \mathrm{P}\left[\mathcal{E}_Q(\mathbf{X}) = i\right]$ denote the probability that the quantizer encoder produces the index i, and define $p_{j|i} = \mathrm{P}\left[\mathcal{D}_C(\eta(\mathcal{E}_C(i))) = j\right]$, the transition probabilities of the coded channel, i.e., the probability that the channel decoder emits the symbol j given that the input to the channel encoder was i. Let χ_S denote the indicator function of a set S.

In Figure 2.1, the quantizer $\tilde{\mathcal{Q}} = \mathcal{D}_Q \circ \mathcal{E}_Q$ is assumed to be designed for a noiseless channel with an optimal (i.e., nearest neighbor) encoder $\mathcal{E}_Q$. (To allow for low-complexity structured codebooks, the decoder is not required to be optimal.) The index assignment π is a permutation of the set $\mathbb{Z}_2^k$. The channel encoder $\mathcal{E}_C$ maps a k-bit binary source index to an n-bit binary channel codeword. This codeword is then transmitted across a binary memoryless channel η, where it may get corrupted by noise. The channel decoder $\mathcal{D}_C$ maps the received n-bit word back to a k-bit source index, which then goes through the inverse index assignment π^{-1}. The quantizer decoder $\mathcal{D}_Q$ then generates the associated output vector in $\mathcal{Y} \subset \mathbb{R}^d$ from the resulting index.

We measure the performance of the noisy channel vector quantizer for a vector source $\mathbf{X}$ by its mean squared error $D = \mathrm{E}\,\|\mathbf{X} - \mathcal{Q}(\mathbf{X})\|^2$. We define the *source distortion* (the distortion incurred on a noiseless channel) as $D_S \triangleq \mathrm{E}\left\|\mathbf{X} - \tilde{\mathcal{Q}}(\mathbf{X})\right\|^2$, and the *channel distortion* (the component due to channel errors) as $D_C \triangleq \mathrm{E}\left\|\tilde{\mathcal{Q}}(\mathbf{X}) - \mathcal{Q}(\mathbf{X})\right\|^2$. If the Centroid Condition is satisfied (i.e., $\mathbf{y}_i = \mathrm{E}\left[\mathbf{X}|\mathcal{E}_Q(\mathbf{X}) = i\right]$, $\forall i$), then $D = D_S + D_C$. If the codevectors are not the centroids of their respective encoding regions then $D = D_S + D_C + 2D_{sc}$, where $D_{sc} = \mathrm{E}\left[\left\langle \mathbf{X} - \tilde{\mathcal{Q}}(\mathbf{X}) \,\middle|\, \tilde{\mathcal{Q}}(\mathbf{X}) - \mathcal{Q}(\mathbf{X}) \right\rangle\right]$. The magnitude of the cross-term D_{sc} is usually very small in practice, and in [28] is shown to asymptotically vanish for regular quantizers (see also [48]). As an example, Table 2.1 lists the three components of D for the BLVQ example of Figure 2.2. It can be seen in these cases that D_{sc} is negligible compared to D_S and D_C.

Table 2.1 The three components of the distortion (normalized by the dimension) for the BLVQ of Figure 2.2.

ϵ	10^{-5}	10^{-4}	10^{-3}	10^{-2}	10^{-1}
D_C	4.08×10^{-8}	4.08×10^{-6}	4.06×10^{-4}	3.83×10^{-2}	2.17
D_{sc}	2.31×10^{-13}	2.31×10^{-11}	2.30×10^{-9}	2.20×10^{-7}	1.34×10^{-5}
D_S	0.52				

Hence, under the assumption $D_{sc} = 0$, for a given quantizer one should minimize D_C to optimize the overall performance. This chapter determines the value of D_C for various noisy channel vector quantizer systems. With no redundancy (i.e., $k = n$), an approach to this problem is to find a reordering of the codevectors (i.e., the best index assignment π) that yields the lowest D_C. Indices that are likely to be mistaken due to channel errors should correspond to code vectors whose Euclidean distance is small.

Fact 2.1 *Let* $\mathbf{X} \in \mathbb{R}^d$ *be a random vector encoded by a noisy channel vector quantizer. The channel distortion can be written as*

$$D_C = \sum_{i \in \mathbb{Z}_2^k} \sum_{j \in \mathbb{Z}_2^k} p_i p_{\pi(j)|\pi(i)} \left\| \mathbf{y}_i - \mathbf{y}_j \right\|^2 . \tag{2.1}$$

2.2.2 Linear codes on a binary symmetric channel

Definition 2.2 A binary (n, k) *linear code* $\mathcal{C}$ with $k \times n$ binary *generator matrix* $\mathrm{G_C}$ is the set of all 2^k n-bit binary words of the form $i\mathrm{G_C}$, for $i \in \mathbb{Z}_2^k$. The *dual code* of $\mathcal{C}$ is defined as $\mathcal{C}^\perp = \{i \in \mathbb{Z}_2^n : ij^t = 0, \quad \forall j \in \mathcal{C}\}$.

We assume a linear code is used with standard array decoding. The channel encoder is given by $\mathcal{E}_C(i) = i\mathrm{G_C}$, and we denote the set of coset leaders by $\mathcal{S}$. Note that $\mathcal{S} = {\mathcal{D}_C}^{-1}(\{0\})$, the set of n-bit binary words decoded into the all-zero codeword, and that by linearity the set of all n-bit words decoded into an arbitrary channel codeword u is $\mathcal{S}_u = \mathcal{S} + u$, a translate of $\mathcal{S}$.

Notation The probability that the error pattern $u \in \mathbb{Z}_2^n$ occurs on a binary symmetric channel with crossover probability ϵ is denoted by

$$\rho_u \stackrel{\triangle}{=} \mathrm{P}\left[\eta(v) = v + u\right] = \epsilon^{w(u)} (1 - \epsilon)^{n - w(u)},$$

where v is an arbitrary element of $\mathbb{Z}_2^n$, and $w(\cdot)$ denotes Hamming weight.

Notation The probability that the information error pattern $j \in \mathbb{Z}_2^k$ occurs when an (n, k) linear block code is used to transmit over a binary symmetric channel is denoted by

$$q_j \stackrel{\triangle}{=} p_{i+j|i} = \sum_{u \in \mathcal{S}} \rho_{u + j\mathrm{G_C}} \qquad i, j \in \mathbb{Z}_2^k. \tag{2.2}$$

2.2.3 Affine index assignments

Definition 2.3 An *affine index assignment* $\pi\colon \mathbb{Z}_2^k \to \mathbb{Z}_2^k$ is a permutation of the form

$$\pi(i) = i\mathrm{G_I} + t, \qquad \pi^{-1}(i) = (i+t){\mathrm{G_I}}^{-1},$$

where $\mathrm{G_I}$ is a binary nonsingular $k \times k$ *generator matrix*, t is a k-dimensional binary *translation vector*, and the operations are performed in $\mathbb{Z}_2^k$. If $t = 0$, then π is called *linear*.

The family of affine index assignments is attractive due to its low implementation complexity. An unconstrained index assignment requires a table of size $O(k2^k)$ bits to implement for a 2^k-point quantizer, whereas affine assignments can be described by $O(k^2)$ bits. The number of unstructured index assignments is $(2^k)!$, whereas the number of affine index assignments is $(2^k)\prod_{i=0}^{k-1}(2^k - 2^i)$. Many well-known useful redundancy free codes are linear or affine, including the Natural Binary Code (NBC), the Folded Binary Code (FBC), the Gray Code (GC), and the Two's Complement Code (TCC):

- *Natural Binary Code*

$${\mathrm{G_I}}^{(NBC)} = \mathcal{I}_k, \qquad t = [0 \cdots 0]\,.$$

- *Folded Binary Code (or Sign-Magnitude Code)*

$${\mathrm{G_I}}^{(FBC)} = \left[\begin{array}{c|ccc} 1 & 1 & \cdots & 1 \\ \hline 0 & & & \\ \vdots & & \mathcal{I}_{k-1} & \\ 0 & & & \end{array}\right], \qquad t = [01 \cdots 1]\,, \qquad \left({\mathrm{G_I}}^{(FBC)}\right)^{-1} = {\mathrm{G_I}}^{(FBC)}.$$

- *Gray Code (or Reflected Binary Code)*

$${\mathrm{G_I}}^{(GC)} = \begin{bmatrix} 1 & 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \cdots & \cdots & 0 & 1 \end{bmatrix}, \qquad t = [0 \cdots 0]\,,$$

$$\left(\mathrm{G_I}^{(GC)}\right)^{-1} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & 1 \end{bmatrix}.$$

- *Two's Complement Code*

$$\mathrm{G_I}^{(TCC)} = \mathcal{I}_k, \qquad t = [10\cdots 0]\,.$$

- *Worst Code*

  It has been shown [33, 34] that the Natural Binary Code is a "best" index assignment on a BSC for any source resulting in a uniform distribution on the BLVQ codevectors. In Chapter 3, we show that a "worst" affine assignment (i.e., maximizing D_C under the same conditions among affine index assignments) is the following linear code:

$$\mathrm{G_I}^{(WC)} = \left[\begin{array}{c|ccc} a_k & 1 & \cdots & 1 \\ \hline 1 & & & \\ \vdots & & \mathcal{I}_{k-1} & \\ 1 & & & \end{array}\right], \qquad t = [0\cdots 0]\,,$$

$$\left(\mathrm{G_I}^{(WC)}\right)^{-1} = \left[\begin{array}{c|ccc} 1 & 1 & \cdots & 1 \\ \hline 1 & & & \\ \vdots & & \bar{\mathcal{I}}_{k-1} & \\ 1 & & & \end{array}\right],$$

  where $a_k = 0$ if k is even, and $a_k = 1$ if k is odd; and $\bar{\mathcal{I}}_{k-1}$ is the one's complement of the identity matrix $\mathcal{I}_{k-1}$.

Table 2.2 gives an explicit listing of these affine index assignments in both decimal and binary. The following recursive relationships between these index assignments can be used to obtain formulas for D_C (e.g., see [49]).

Table 2.2 Examples of 4-bit index assignments.

i	$\pi_4^{(NBC)}(i)$		$\pi_4^{(FBC)}(i)$		$\pi_4^{(GC)}(i)$		$\pi_4^{(TCC)}(i)$		$\pi_4^{(WC)}(i)$	
0	0	0000	7	0111	0	0000	8	1000	0	0000
1	1	0001	6	0110	1	0001	9	1001	9	1001
2	2	0010	5	0101	3	0011	10	1010	10	1010
3	3	0011	4	0100	2	0010	11	1011	3	0011
4	4	0100	3	0011	6	0110	12	1100	12	1100
5	5	0101	2	0010	7	0111	13	1101	5	0101
6	6	0110	1	0001	5	0101	14	1110	6	0110
7	7	0111	0	0000	4	0100	15	1111	15	1111
8	8	1000	8	1000	12	1100	0	0000	7	0111
9	9	1001	9	1001	13	1101	1	0001	14	1110
10	10	1010	10	1010	15	1111	2	0010	13	1101
11	11	1011	11	1011	14	1110	3	0011	4	0100
12	12	1100	12	1100	10	1010	4	0100	11	1011
13	13	1101	13	1101	11	1011	5	0101	2	0010
14	14	1110	14	1110	9	1001	6	0110	1	0001
15	15	1111	15	1111	8	1000	7	0111	8	1000

$$\pi_k^{(NBC)}(i) = \begin{cases} \pi_{k-1}^{(NBC)}(i) & 0 \le i \le 2^{k-1}-1 \\ 2^{k-1} + \pi_{k-1}^{(NBC)}(i-2^{k-1}) & 2^{k-1} \le i \le 2^k - 1 \end{cases}, \qquad \pi_0^{(NBC)}(0) = 0.$$

$$\pi_k^{(FBC)}(i) = \begin{cases} 2^{k-1} - 1 - \pi_{k-1}^{(NBC)}(i) & 0 \le i \le 2^{k-1}-1 \\ 2^{k-1} + \pi_{k-1}^{(NBC)}(i-2^{k-1}) & 2^{k-1} \le i \le 2^k - 1 \end{cases}.$$

$$\pi_k^{(GC)}(i) = \begin{cases} \pi_{k-1}^{(GC)}(i) & 0 \le i \le 2^{k-1}-1 \\ 2^{k-1} + \pi_{k-1}^{(GC)}(2^k - 1 - i) & 2^{k-1} \le i \le 2^k - 1 \end{cases}, \qquad \pi_0^{(GC)}(0) = 0.$$

$$\pi_k^{(TCC)}(i) = \begin{cases} 2^{k-1} + \pi_{k-1}^{(NBC)}(i) & 0 \le i \le 2^{k-1}-1 \\ \pi_{k-1}^{(NBC)}(i-2^{k-1}) & 2^{k-1} \le i \le 2^k - 1 \end{cases}.$$

$$\pi_k^{(WC)}(i) = \begin{cases} \chi_{\{w(i)\ \text{odd}\}} 2^{k-1} + \pi_{k-1}^{(NBC)}(i) & 0 \leq i \leq 2^{k-1}-1 \\ \chi_{\{w(2^k-1-i)\ \text{even}\}} 2^{k-1} + \pi_{k-1}^{(NBC)}(2^k-1-i) & 2^{k-1} \leq i \leq 2^k-1 \end{cases}.$$

2.2.4 Binary lattice VQ

Definition 2.4 A d-dimensional, 2^k-point *binary lattice vector quantizer* is a vector quantizer, whose codevectors are of the form

$$\mathbf{y}_i = \mathbf{y}_0 + \sum_{l=0}^{k-1} \mathbf{v}_l i_l$$

for $i = [i_{k-1}, i_{k-2}, \ldots, i_1, i_0] \in \mathbb{Z}_2^k$, $\mathbf{y}_0 \in \mathbb{R}^d$, and where $\mathcal{V} = \{\mathbf{v}_l\}_{l=0}^{k-1} \subset \mathbb{R}^d$ is the *generating set*, ordered by $\|\mathbf{v}_0\| \leq \|\mathbf{v}_1\| \leq \ldots \leq \|\mathbf{v}_{k-1}\|$.

A BLVQ can be considered a direct sum quantizer (or multistage, or residual quantizer) with two codevectors at each stage, when the codebook is written as $\bigoplus_{l=0}^{k-1}\{\mathbf{y}_0/k, (\mathbf{y}_0/k) + \mathbf{v}_l\}$. Conversely, any direct sum vector quantizer with two vectors per component codebook, $\bigoplus_{l=0}^{k-1}\{\mathbf{a}_l, \mathbf{b}_l\}$, can be viewed as a BLVQ by setting $\mathbf{y}_0 = \sum_{l=0}^{k-1} \mathbf{a}_l$ and $\mathbf{v}_l = \mathbf{b}_l - \mathbf{a}_l\ \forall l$, and reordering the generating set if needed.

Given an arbitrary lattice with basis vectors $\{\mathbf{u}_j\}_{j=1}^{L} \subset \mathbb{R}^d$, any set $\{k_j\}_{j=1}^{L}$ of non-negative integers satisfying $\sum_{j=1}^{L} k_j = k$ defines a 2^k-point *lattice vector quantizer* with codebook

$$\Lambda = \left\{ \sum_{j=1}^{L} m_j \mathbf{u}_j \colon m_j \in \{0, 1, \ldots, 2^{k_j} - 1\} \right\}.$$

For each j, the vectors in the direction of $\mathbf{u}_j$ are addressed with k_j bits.

The class of BLVQs includes lattice VQs (or any of their cosets). In this case, the BLVQ's generating set is $\mathcal{V} = \left\{2^{l_j}\mathbf{u}_j \colon l_j \in \{0, \ldots, k_j - 1\}, j \in \{1, \ldots, L\}\right\}$ and the index i of the vector $\mathbf{y}_i = \sum_{j=1}^{L} m_j \mathbf{u}_j$ is the concatenation of the binary representations of the lattice coefficients $m_1, m_2, \cdots, m_L$. The codebook of this BLVQ contains the origin ($\mathbf{y}_0 = \mathbf{0}$). By choosing $\mathbf{y}_0 \neq \mathbf{0}$ (while keeping the same generating set $\mathcal{V}$), other BLVQs can be obtained corresponding to truncations of cosets of the original lattice. A 2^k-level uniform scalar quantizer with stepsize Δ and granular region (a, b) is a special case of a binary lattice quantizer, obtained by setting $y_0 = a + \Delta/2$, and $v_l = \Delta 2^l$.

A BLVQ is similar to the nonredundant version of the "VQ by a linear mapping of a block code" (LMBC-VQ) presented in [44–46]. The j^{th} codevector of an LMBC-VQ is defined as

$$\mathbf{y}_j = \Theta \mathbf{b}_j,$$

where Θ is a $d \times (n+1)$ real matrix with columns $\{\theta_i\}_{i=0}^{n+1}$, and the $(n+1)$-dimensional column vector $\mathbf{b}_j \in \{-1,1\}^{n+1}$ is obtained from the j^{th} codeword of a systematic (n,k) linear code by the mapping $b \to (-1)^b = 1-2b$ for each bit, and a leading 1 is prepended to allow translation of the codebook by θ_0. In the nonredundant case (i.e., $n=k$), $\mathbf{b}_j = [1, (-1)^{j_{k-1}}, \ldots, (-1)^{j_0}]^t$. Hence

$$\begin{aligned} \mathbf{y}_j &= \theta_0 + \sum_{l=1}^{k} \theta_l (1 - 2j_{k-l}) \\ &= \sum_{l=0}^{k} \theta_l + \sum_{l=0}^{k-1} (-2\theta_{k-l})\, j_l. \end{aligned}$$

Thus, setting $\mathbf{y}_0 = \sum_{l=0}^{k} \theta_l$ and $\mathbf{v}_l = -2\theta_{k-l}$ gives the codevector $\mathbf{y}_j$ in the form of a BLVQ codevector. Conversely, given a BLVQ we obtain a nonredundant LMBC-VQ by setting $\theta_0 = \mathbf{y}_0 + \frac{1}{2}\sum_{l=0}^{k-1} \mathbf{v}_l$, and $\theta_l = -\frac{1}{2}\mathbf{v}_{k-l}$ for $l = 1, \ldots, k$.

Hagen and Hedelin [44–46] adapted the generalized Lloyd algorithm for the design of LMBC-VQs, and obtained locally optimal "noiseless" LMBC-VQ codebooks. Their scheme does not include error control coding, nor do they explicitly mention index assignments. They use a linear block code exclusively as a tool for quantizer design. When this "design code" is nonredundant, their scheme can only implement index assignments corresponding to bit-permutations of the indices (since the 2^n codevectors uniquely determine the $n+1$ columns of Θ up to sign and order). On a memoryless channel these index assignments all have the same value of D_C as the NBC. However, by increasing the redundancy of the "design code," more general index assignments can be obtained. Indeed, in the maximum redundancy case (i.e., $n = 2^k - 1$), the matrix of codevectors is related to Θ by the Hadamard transform as described in [47], and thus any index assignment (reordering of the codevectors) can be modeled by choosing Θ accordingly.

An optimal and a fast suboptimal algorithm for finding a good assignment in that case are presented in [23].

2.2.5 The Hadamard transform

Definition 2.5 For each $i, j \in \mathbb{Z}_2^k$ let $h_{i,j} = (-1)^{ij^t}$ and let $f\colon \mathbb{Z}_2^k \to \mathbb{R}$. The *Hadamard transform* $\hat{f}\colon \mathbb{Z}_2^k \to \mathbb{R}$ of f is defined by

$$\hat{f}_j = \sum_{i \in \mathbb{Z}_2^k} f_i h_{i,j},$$

and the inverse transform is given by

$$f_i = 2^{-k} \sum_{j \in \mathbb{Z}_2^k} \hat{f}_j h_{j,i}.$$

We refer to the numbers $h_{i,j}$ as Hadamard coefficients. The transform equations can be expressed in vector form using the $2^k \times 2^k$ Sylvester-type Hadamard matrix $H = [h_{i,j}] \quad (i, j \in \mathbb{Z}_2^k)$ and viewing the functions as 2^k-dimensional row vectors (i.e., $f = [f_0, f_1, \ldots, f_{2^k-1}]$):

$$\hat{f} = fH \qquad f = 2^{-k} \hat{f} H.$$

The Hadamard transform extends to vector valued functions $\mathbf{f}\colon \mathbb{Z}_2^k \to \mathbb{R}^d$ in a straightforward manner:

$$\hat{\mathbf{f}}_j = \sum_{i \in \mathbb{Z}_2^k} \mathbf{f}_i h_{i,j} \qquad \mathbf{f}_i = 2^{-k} \sum_{j \in \mathbb{Z}_2^k} \hat{\mathbf{f}}_j h_{j,i}$$

or equivalently

$$\hat{F} = FH \qquad F = 2^{-k} \hat{F} H,$$

where $F = [\mathbf{f}_0, \mathbf{f}_1, \ldots, \mathbf{f}_{2^k-1}]$ is a $d \times 2^k$ real matrix.

The Hadamard transform is an orthogonal transform, and the convolution and inner product properties (e.g., Parseval's identity) of Fourier transforms also hold for Hadamard transforms. The following useful identities also hold:

- $$h_{i,j} = h_{j,i} \qquad i, j \in \mathbb{Z}_2^k.$$

- $$h_{i,l} h_{i,m} = h_{i,l+m} \qquad i, l, m \in \mathbb{Z}_2^k.$$

- The bits of any binary word $i \in \mathbb{Z}_2^k$ are related to the Hadamard matrix entries by

$$i_m = \frac{1 - h_{i,e_m}}{2} \qquad m \in \{0, 1, \ldots, k-1\}, \tag{2.3}$$

  where $e_m \in \mathbb{Z}_2^k$ is the binary row vector with its only nonzero component in the m^{th} position.

- $$\sum_{i \in \mathbb{Z}_2^k} h_{i,j} = 2^k \chi_{\{j=0\}} \qquad j \in \mathbb{Z}_2^k. \tag{2.4}$$

2.3 Results

The following lemma gives an expression for the channel distortion of a noisy channel vector quantizer in terms of the Hadamard transforms of the source distribution (the $\hat{r}$'s), the quantizer codebook (the $\hat{\mathbf{z}}$'s), and the channel statistics (the $\hat{q}$'s). A similar expression is found in [50], and a concise proof is provided here for completeness.

Lemma 2.1 *Let $\mathbf{X} \in \mathbb{R}^d$ be a random vector that is quantized by a 2^k-point vector quantizer with encoder $\mathcal{E}_Q$ and decoder $\mathcal{D}_Q$, index assignment π, and using a linear block channel code on a binary symmetric channel. Let $r_i = \mathrm{P}\left[\pi(\mathcal{E}_Q(\mathbf{X})) = i\right]$, $\mathbf{z}_i = \mathcal{D}_Q(\pi^{-1}(i))$, and $q_i = p_{j+i|j}$. Then the channel distortion in the Hadamard transform domain is*

$$D_C = 4^{-k} \sum_{i \in \mathbb{Z}_2^k} \sum_{j \in \mathbb{Z}_2^k} \langle \hat{\mathbf{z}}_i \, | \hat{\mathbf{z}}_j \rangle \, \hat{r}_{i+j} \left(\hat{q}_0 - \hat{q}_i - \hat{q}_j + \hat{q}_{i+j} \right).$$

Proof

Using a linear block channel code on a binary symmetric channel the transition probabilities $p_{j|i}$ only depend on the (modulo 2) sum $i + j$. With the notation $q_{i+j} = p_{j|i}$, Equation (2.1) can be written as

$$
\begin{aligned}
D_C &= \sum_{i\in\mathbb{Z}_2^k}\sum_{j\in\mathbb{Z}_2^k} p_i \left\|\mathbf{y}_i-\mathbf{y}_j\right\|^2 p_{\pi(j)|\pi(i)} \\
&= \sum_{i\in\mathbb{Z}_2^k}\sum_{j\in\mathbb{Z}_2^k} r_i \left\|\mathbf{z}_i-\mathbf{z}_j\right\|^2 q_{i+j} \\
&= \sum_{i\in\mathbb{Z}_2^k} r_i \sum_{j\in\mathbb{Z}_2^k} \left\| 2^{-k}\sum_{l\in\mathbb{Z}_2^k} \hat{\mathbf{z}}_l (h_{l,i}-h_{l,j}) \right\|^2 q_{i+j} \\
&= 4^{-k}\sum_{i\in\mathbb{Z}_2^k} r_i \sum_{j\in\mathbb{Z}_2^k} q_{i+j} \sum_{l\in\mathbb{Z}_2^k}\sum_{m\in\mathbb{Z}_2^k} \langle \hat{\mathbf{z}}_l \,|\hat{\mathbf{z}}_m\rangle \left(h_{l+m,i}-h_{l,j}h_{m,i}-h_{l,i}h_{m,j}+h_{l+m,j}\right) \\
&= 4^{-k}\sum_{l\in\mathbb{Z}_2^k}\sum_{m\in\mathbb{Z}_2^k} \langle \hat{\mathbf{z}}_l \,|\hat{\mathbf{z}}_m\rangle \sum_{i\in\mathbb{Z}_2^k} r_i h_{l+m,i} \sum_{j\in\mathbb{Z}_2^k} q_{i+j}(1-h_{l,i+j}-h_{m,i+j}+h_{l+m,i+j}) \\
&= 4^{-k}\sum_{l\in\mathbb{Z}_2^k}\sum_{m\in\mathbb{Z}_2^k} \langle \hat{\mathbf{z}}_l \,|\hat{\mathbf{z}}_m\rangle \left(\sum_{i\in\mathbb{Z}_2^k} r_i h_{l+m,i}\right) \sum_{c\in\mathbb{Z}_2^k} q_c(h_{0,c}-h_{l,c}-h_{m,c}+h_{l+m,c}) \\
&= 4^{-k}\sum_{l\in\mathbb{Z}_2^k}\sum_{m\in\mathbb{Z}_2^k} \langle \hat{\mathbf{z}}_l \,|\hat{\mathbf{z}}_m\rangle \hat{r}_{l+m} \left(\hat{q}_0-\hat{q}_l-\hat{q}_m+\hat{q}_{l+m}\right).
\end{aligned}
$$

■

In Lemma 2.1 “complete” channel decoding is assumed. That is, *every* received word from the channel is decoded to a nearby channel codeword (to the one in the same coset as the received word), as opposed to incomplete decoding (or bounded distance decoding), where a received word is decoded only if it is within a prescribed Hamming distance (usually, the code's minimum distance) to a codeword — otherwise it is deemed uncorrectable. The form of the expression for D_C for incomplete decoding of a linear block code is similar: D_C has an additional term $\sigma_{\mathcal{Y}}^2(1-\hat{q}_0)$, where $\sigma_{\mathcal{Y}}^2$ is the codebook energy, and $(1-\hat{q}_0)$ is the probability of an uncorrectable error. Since this additional term is independent of π, it is not significant in determining the optimal index assignment.

The following theorem specializes Lemma 2.1 to BLVQs and affine index assignments.

Theorem 2.1 *The channel distortion of a 2^k-point binary lattice vector quantizer with generating set $\{\mathbf{v}_l\}_{l=0}^{k-1}$, affine index assignment with generator matrix $\mathrm{G_I}$, (n,k) linear code $\mathcal{C}$ with generator matrix $\mathrm{G_C}$, and a binary symmetric channel with crossover probability ϵ, is given by*

$$D_C = \frac{1}{4}\sum_{l=0}^{k-1}\sum_{m=0}^{k-1}\langle \mathbf{v}_l \,|\mathbf{v}_m\rangle\, \hat{p}_{e_l+e_m}\left(\hat{q}_0 - \hat{q}_{e_l\mathrm{F_I}} - \hat{q}_{e_m\mathrm{F_I}} + \hat{q}_{(e_l+e_m)\mathrm{F_I}}\right), \tag{2.5}$$

where

$$\hat{q}_i = 2^{k-n}\sum_{a\in(\mathcal{C}^{\perp}+i\mathrm{F_C})} \hat{J}_a(1-2\epsilon)^{w(a)}, \tag{2.6}$$

$\mathrm{F_I} = \left({\mathrm{G_I}}^{-1}\right)^t$, $\mathrm{F_C}$ *is a $k\times n$ binary matrix satisfying $\mathrm{G_C}{\mathrm{F_C}}^t = \mathcal{I}_k$, $\mathcal{C}^{\perp}$ is the dual code of $\mathcal{C}$, $J_r = \chi_{\{r\in\mathcal{S}\}}$ is the characteristic function of the set $\mathcal{S}$ of coset leaders of $\mathcal{C}$, $\hat{p}_l$ is the l^{th} component of the Hadamard transform of the distribution on the quantizer codevectors, $w(\cdot)$ denotes Hamming weight, and e_l is the binary row vector with its only nonzero entry in the l^{th} position.*

Theorem 2.1 makes explicit the dependence of the channel distortion on the BLVQ structure, the affine index assignment, and the channel code. Also, computing D_C based on (2.1) requires $O(N^2)$ complexity for a codebook of size $N = 2^k$, whereas using (2.5) reduces the complexity of computing D_C to $O(N\log^2 N)$. (In (2.1) each of the nested sums contributes a factor of N, whereas in (2.5) the corresponding sums only require $\log N$ steps, but each of the Hadamard transforms inside the sums takes $O(N)$ steps.)

Note that on a binary symmetric channel the translation vector t of the affine assignment is irrelevant. Thus without loss of generality we may assume that the index assignment is linear. A linear index assignment can be incorporated in the channel encoder by setting $\mathrm{G'}_C = \mathrm{G_I}\mathrm{G_C}$. Then $\mathrm{F'}_C = \mathrm{F_I}\mathrm{F_C}$, the transpose of an inverse of $\mathrm{G'}_C$. To obtain $\hat{q}_{e_l\mathrm{F_I}}$, the sum in (2.6) is taken over the coset of the dual code containing $e_l\mathrm{F_I}\mathrm{F_C} = e_l\mathrm{F'}_C$, the l^{th} row of $\mathrm{F'}_C$.

Proof

We find expressions for the Hadamard transform quantities of Lemma 2.1. The transform of the BLVQ codevectors is

$$
\begin{aligned}
\hat{\mathbf{z}}_a &= \sum_{i \in \mathbb{Z}_2^k} \mathbf{z}_i h_{i,a} \\
&= \sum_{i \in \mathbb{Z}_2^k} \mathbf{z}_{\pi(i)} h_{\pi(i),a} \\
&= \sum_{i \in \mathbb{Z}_2^k} \mathbf{y}_i h_{\pi(i),a} \\
&= \sum_{i \in \mathbb{Z}_2^k} \left(\mathbf{y}_0 + \sum_{l=0}^{k-1} \mathbf{v}_l i_l \right) h_{\pi(i),a} \\
&= \chi_{\{a=0\}} 2^k \left(\mathbf{y}_0 + \frac{1}{2} \sum_{l=0}^{k-1} \mathbf{v}_l \right) + \chi_{\{a \neq 0\}} \sum_{l=0}^{k-1} \mathbf{v}_l \sum_{i \in \mathbb{Z}_2^k} \frac{1 - h_{i,e_l}}{2} h_{\pi(i),a} \\
&= \chi_{\{a=0\}} 2^k \left(\mathbf{y}_0 + \frac{1}{2} \sum_{l=0}^{k-1} \mathbf{v}_l \right) - \frac{1}{2} \chi_{\{a \neq 0\}} \sum_{l=0}^{k-1} \mathbf{v}_l \sum_{i \in \mathbb{Z}_2^k} h_{i,e_l} h_{\pi(i),a}.
\end{aligned}
$$

Since π is an affine index assignment, for $a \neq 0$ we have

$$
\begin{aligned}
\sum_{i \in \mathbb{Z}_2^k} h_{i,e_l} h_{\pi(i),a} &= \sum_{i \in \mathbb{Z}_2^k} h_{i,e_l} h_{i\mathrm{G_I}+d,a} \\
&= h_{d,a} \sum_{i \in \mathbb{Z}_2^k} h_{i,e_l + a\mathrm{G_I}^t} \\
&= \chi_{\{a\mathrm{G_I}^t = e_l\}} 2^k h_{d,a}.
\end{aligned}
$$

Thus

$$
\hat{\mathbf{z}}_a = -2^{k-1} h_{d,a} \sum_{l=0}^{k-1} \mathbf{v}_l \chi_{\{a = e_l \mathrm{F_I}\}}
$$

for $a \neq 0$. Exactly one term in this summation is nonzero. The transform of the discrete distribution on the codevectors is

$$
\hat{r}_a = \sum_{i \in \mathbb{Z}_2^k} p_i h_{\pi(i),a} = \sum_{i \in \mathbb{Z}_2^k} p_i h_{i\mathrm{G_I}+d,a} = h_{d,a} \hat{p}_{a\mathrm{G_I}^t}.
$$

If either a or b equals 0, then $\hat{q}_0 - \hat{q}_a - \hat{q}_b + \hat{q}_{a+b} = 0$, so without loss of generality we can write

$$\begin{aligned} D_C = 4^{-k} \sum_{a \in \mathbb{Z}_2^k} \sum_{b \in \mathbb{Z}_2^k} & \left\langle -2^{k-1} h_{d,a} \sum_{l=0}^{k-1} \mathbf{v}_l \chi_{\{a=e_l \mathrm{F_I}\}} \right| \left. -2^{k-1} h_{d,b} \sum_{m=0}^{k-1} \mathbf{v}_m \chi_{\{a=e_m \mathrm{F_I}\}} \right\rangle \\ & \cdot h_{d,a+b} \hat{p}_{(a+b)\mathrm{G_I}^t} \left(\hat{q}_0 - \hat{q}_a - \hat{q}_b + \hat{q}_{a+b} \right) \\ = \frac{1}{4} \sum_{l=0}^{k-1} \sum_{m=0}^{k-1} & \left\langle \mathbf{v}_l \,|\mathbf{v}_m \right\rangle \hat{p}_{e_l+e_m} \left(\hat{q}_0 - \hat{q}_{e_l \mathrm{F_I}} - \hat{q}_{e_m \mathrm{F_I}} + \hat{q}_{(e_l+e_m)\mathrm{F_I}} \right). \end{aligned}$$

Since $q_j = \sum_{r \in \mathcal{S}} \rho_{r+j\mathrm{G_C}}$, the $\hat{q}_i$'s for an (n,k) linear code can be expressed in terms of the $\hat{\rho}_i$'s and the Hadamard transform of $J = \chi_{\mathcal{S}}$ as

$$\begin{aligned} \hat{q}_i &= \sum_{j \in \mathbb{Z}_2^k} \left(\sum_{r \in \mathcal{S}} \rho_{r+j\mathrm{G_C}} \right) h_{i,j} && (2.7) \\ &= \sum_{j \in \mathbb{Z}_2^k} h_{i,j} \sum_{r \in \mathcal{S}} 2^{-n} \sum_{a \in \mathbb{Z}_2^n} \hat{\rho}_a h_{a,r+j\mathrm{G_C}} \\ &= 2^{-n} \sum_{a \in \mathbb{Z}_2^n} \hat{\rho}_a \left(\sum_{r \in \mathcal{S}} h_{r,a} \right) \left(\sum_{j \in \mathbb{Z}_2^k} h_{i,j} h_{a,j\mathrm{G_C}} \right) \\ &= 2^{-n} \sum_{a \in \mathbb{Z}_2^n} \hat{\rho}_a \left(\sum_{r \in \mathbb{Z}_2^n} J_r h_{r,a} \right) \left(\sum_{j \in \mathbb{Z}_2^k} h_{j,i+a\mathrm{G_C}^t} \right) \\ &= 2^{-n} \sum_{a \in \mathbb{Z}_2^n} \hat{\rho}_a \hat{J}_a 2^k \chi_{\{i=a\mathrm{G_C}^t\}} \\ &= 2^{k-n} \sum_{a \in (\mathcal{C}^\perp + i\mathrm{F_C})} \hat{\rho}_a \hat{J}_a, && (2.8) \end{aligned}$$

where

$$\begin{aligned} \hat{\rho}_a &= \sum_{i \in \mathbb{Z}_2^k} \epsilon^{w(i)} (1-\epsilon)^{n-w(i)} h_{i,a} \\ &= \sum_{i_0 \in \{0,1\}} \sum_{i_1 \in \{0,1\}} \cdots \sum_{i_{n-1} \in \{0,1\}} \left(\prod_{l=0}^{n-1} \epsilon^{i_l} (1-\epsilon)^{1-i_l} \right) \left(\prod_{l=0}^{n-1} (-1)^{i_l a_l} \right) \\ &= \prod_{l=0}^{n-1} \sum_{i_l \in \{0,1\}} \epsilon^{i_l} (1-\epsilon)^{1-i_l} (-1)^{i_l a_l} \end{aligned}$$

$$= \prod_{l=0}^{n-1} \left(1 - \epsilon + \epsilon(-1)^{a_l}\right)$$
$$= (1 - 2\epsilon)^{w(a)},$$

which completes the proof. ∎

2.3.1 Uniform output distribution

If the quantizer codevectors are equiprobable, then $p_i = 2^{-k}$ for all $i \in \mathbb{Z}_2^k$, and $\hat{p}_i = \chi_{\{i=0\}}$. In this case the channel distortion of BLVQ with an affine assignment simplifies to

$$D_C = \frac{1}{2} \sum_{l=0}^{k-1} \|\mathbf{v}_l\|^2 \left(\hat{q}_0 - \hat{q}_{e_l \mathrm{F_I}}\right).$$

Since $\hat{q}_0$ is independent of the index assignment, minimizing D_C is equivalent to maximizing

$$\sum_{l=0}^{k-1} \|\mathbf{v}_l\|^2 \, \hat{q}_{e_l \mathrm{F_I}}.$$

By assumption, the $\mathbf{v}_l$'s are ordered by their norms. Thus an affine assignment which minimizes D_C must satisfy

$$\hat{q}_{e_{k-1}\mathrm{F_I}} \geq \hat{q}_{e_{k-2}\mathrm{F_I}} \geq \ldots \geq \hat{q}_{e_0 \mathrm{F_I}},$$

as observed in [23, 33, 43]. This is achieved by making the k-bit index of a maximal $\hat{q}_i$ $(i \neq 0)$ the first row (corresponding to e_{k-1}) of $\mathrm{F_I}$. Then the l^{th} row is selected to be the index of a largest $\hat{q}_i$ that is linearly independent of the first $l-1$ rows. More formally,

$$f_l = \operatorname*{argmax}_{i \notin \mathrm{span}[f_j]_{j=1}^{l-1}} \hat{q}_i,$$

where f_l denotes the l^{th} row of F_I.

In [33] it was shown that among all possible index assignments the best affine index assignment achieves the minimum MSE possible for a uniform scalar quantizer and a uniform distribution. We conjecture that the same result is valid for BLVQs. It is known to be true for nonredundant channel codes [34], and we have verified that it is true for some simple codes such as the $(7,4,3)$ Hamming code and the $(8,4,4)$ first order Reed-Muller code, and it trivially holds for all $(n,1,n)$ repetition codes. One can also use a result in [38] to determine the best choice of a coset leader set $\mathcal{S}$ and an affine index assignment (even if the best affine assignment does not coincide with the global optimum).

2.4 Nonredundant Codes for the BLVQ

2.4.1 Binary symmetric channels

Theorem 2.1 can be specialized to nonredundant codes (i.e., $n = k$, $\mathcal{C} = \mathbb{Z}_2^k$, $\mathcal{C}^\perp = \mathcal{S} = \{0\}$, $\mathrm{G_C} = \mathrm{F_C} = \mathcal{I}_k$), giving the following result (similar to a result obtained in [46]).

Corollary 2.1 *The channel distortion of a 2^k-point binary lattice vector quantizer with generating set $\{\mathbf{v}_l\}_{l=0}^{k-1}$, which uses an affine index assignment with generator matrix* G, *and nonredundant channel coding, to transmit across a binary symmetric channel with crossover probability ϵ, is given by*

$$\begin{aligned} D_C = \frac{1}{4} \sum_{l=0}^{k-1} \sum_{m=0}^{k-1} \langle \mathbf{v}_l \,|\mathbf{v}_m \rangle \, \hat{p}_{e_l+e_m} \\ \cdot \left(1 - (1-2\epsilon)^{w(e_l F)} - (1-2\epsilon)^{w(e_m F)} + (1-2\epsilon)^{w((e_l+e_m)F)}\right), \end{aligned} \tag{2.9}$$

where $w(\cdot)$ denotes Hamming weight, $\mathrm{F} = (\mathrm{G}^{-1})^t$, *$\hat{p}_l$ is the l^{th} component of the Hadamard transform of the distribution on the quantizer code points, and e_l is the binary row vector with its only nonzero entry in the l^{th} position.*

2.4.1.1 Formulas for common index assignments

One useful consequence of Theorem 2.1 is that exact expressions for the channel distortion D_C can be obtained for certain well-known structured classes of index assignments, such as the NBC, the FBC, and the GC. Since on a binary symmetric channel the TCC and the NBC have the same channel distortion, the NBC formula also holds for the TCC. In the formula for the GC the double sum of Theorem 2.1 cannot be further simplified, since $(1-2\epsilon)^{|l-m|}$ is not a separable function of l and m. For the NBC and the FBC we can express D_C in terms of the means and the component variances of two discrete random variables as follows. Let $\mathbf{Y}$ be a random vector distributed according to $\{p_i\}$ over the quantizer codevectors with mean $\bar{\mathbf{Y}}$ and $\sigma_{\mathbf{Y}}^2 = \mathrm{E}\left\|\mathbf{Y}-\bar{\mathbf{Y}}\right\|^2$, and let $\mathbf{U}$ be a random vector uniformly distributed over the quantizer code points with mean $\bar{\mathbf{U}}$ and $\sigma_{\mathbf{U}}^2 = \mathrm{E}\left\|\mathbf{U}-\bar{\mathbf{U}}\right\|^2$. Then

$$\begin{aligned}
\bar{\mathbf{U}} &= 2^{-k}\sum_{i\in\mathbb{Z}_2^k}\mathbf{y}_i \\
&= 2^{-k}\sum_{i\in\mathbb{Z}_2^k}\left(\mathbf{y}_0+\sum_{l=0}^{k-1}\mathbf{v}_l i_l\right) \\
&= \mathbf{y}_0+\frac{1}{2}\sum_{l=0}^{k-1}\mathbf{v}_l .
\end{aligned}$$

$$\begin{aligned}
\sigma_{\mathbf{U}}^2 &= 2^{-k}\sum_{i\in\mathbb{Z}_2^k}\left\|\mathbf{y}_i-\bar{\mathbf{U}}\right\|^2 \\
&= 2^{-k}\sum_{i\in\mathbb{Z}_2^k}\left\|\mathbf{y}_0+\sum_{l=0}^{k-1}\mathbf{v}_l i_l-\mathbf{y}_0-\frac{1}{2}\sum_{l=0}^{k-1}\mathbf{v}_l\right\|^2 \\
&= 2^{-k}\sum_{i\in\mathbb{Z}_2^k}\left\|\sum_{l=0}^{k-1}\mathbf{v}_l\left(i_l-\frac{1}{2}\right)\right\|^2 \\
&= \frac{1}{4}\sum_{l=0}^{k-1}\sum_{m=0}^{k-1}\langle\mathbf{v}_l\,|\mathbf{v}_m\rangle\, 2^{-k}\sum_{i\in\mathbb{Z}_2^k}h_{i,e_l}h_{i,e_m} && (2.10)\\
&= \frac{1}{4}\sum_{l=0}^{k-1}\left\|\mathbf{v}_l\right\|^2, && (2.11)
\end{aligned}$$

where (2.3) and (2.4) were used to obtain (2.10) and (2.11), respectively.

Note that $\bar{\mathbf{U}}$ and $\sigma_{\mathbf{U}}^2$ do not depend on the index assignment or the input distribution.

Corollary 2.2 *Given the conditions of Corollary 2.1, the channel distortion of the Natural Binary Code is*

$$D_C^{(NBC)} = 4\epsilon\left((1-\epsilon)\sigma_{\mathbf{U}}^2 + \epsilon(\sigma_{\mathbf{Y}}^2 + \left\|\bar{\mathbf{Y}} - \bar{\mathbf{U}}\right\|^2)\right).$$

A related formula also appears in [46], and we provide a short proof for completeness.

Proof Using Corollary 2.1, and the fact that $w(e_l F^{(NBC)}) = 1$ and $w((e_l + e_m)F^{(NBC)}) = 2\chi_{\{l \neq m\}}$, for all l and m, we have

$$\begin{aligned}
D_C^{(NBC)} &= \sum_{l=0}^{k-1} \|\mathbf{v}_l\|^2 \epsilon + \sum_{\substack{l=0 \\ l\neq m}}^{k-1} \sum_{m=0}^{k-1} \langle \mathbf{v}_l \,|\mathbf{v}_m\rangle\, \epsilon^2 \hat{p}_{e_l+e_m} \\
&= 4\epsilon\sigma_{\mathbf{U}}^2 + \epsilon^2 \left(\sum_{l=0}^{k-1}\sum_{m=0}^{k-1} \langle \mathbf{v}_k \,|\mathbf{v}_l\rangle \sum_{i\in\mathbb{Z}_2^k} p_i h_{i,e_l+e_m} - \sum_{l=0}^{k-1} \|\mathbf{v}_l\|^2 \right) \\
&= 4\epsilon(1-\epsilon)\sigma_{\mathbf{U}}^2 + \epsilon^2 \sum_{i\in\mathbb{Z}_2^k} p_i \left\| \sum_{l=0}^{k-1} \mathbf{v}_l\,(2i_l - 1) \right\|^2 \\
&= 4\epsilon(1-\epsilon)\sigma_{\mathbf{U}}^2 + 4\epsilon^2 \mathrm{E}\left\|\mathbf{Y} - \bar{\mathbf{U}}\right\|^2 \\
&= 4\epsilon\left((1-\epsilon)\sigma_{\mathbf{U}}^2 + \epsilon\mathrm{E}\left\|\mathbf{Y} - \bar{\mathbf{Y}} + \bar{\mathbf{Y}} - \bar{\mathbf{U}}\right\|^2\right) \\
&= 4\epsilon\left((1-\epsilon)\sigma_{\mathbf{U}}^2 + \epsilon(\sigma_{\mathbf{Y}}^2 + \left\|\bar{\mathbf{Y}} - \bar{\mathbf{U}}\right\|^2)\right).
\end{aligned}$$

■

Corollary 2.3 *Given the conditions of Corollary 2.1, the channel distortion of the Folded Binary Code is*

$$D_C^{(FBC)} = 4\epsilon(1-\epsilon)\left(\sigma_{\mathbf{U}}^2 + \sigma_{\mathbf{Y}}^2 + \left\|\bar{\mathbf{Y}} - \bar{\mathbf{U}}\right\|^2\right) - \epsilon(1-2\epsilon)\max_l \|\mathbf{v}_l\|^2.$$

Proof The Hamming weights of the rows of $\mathrm{F}^{(FBC)}$ are

$$\begin{aligned}
w(e_l F^{(FBC)}) &= \begin{cases} 1 & l = k-1 \\ 2 & l < k-1 \end{cases} \\
w((e_l + e_m)F^{(FBC)}) &= \begin{cases} 0 & l = m \\ 1 & m < l = k-1 \quad \text{or} \quad l < m = k-1 \\ 2 & l < k-1, \;\; m < k-1, \;\; l \neq m \end{cases},
\end{aligned}$$

and thus

$$\frac{1}{4}\left(1-(1-2\epsilon)^{w(e_lF)}-(1-2\epsilon)^{w(e_mF)}+(1-2\epsilon)^{w((e_l+e_m)F)}\right) = \begin{cases} \epsilon & l=m=k-1 \\ 2\epsilon(1-\epsilon) & l=m<k-1 \\ \epsilon(1-\epsilon) & l\neq m \end{cases}.$$

Substituting these into (2.9), and using (2.11) we obtain

$$\begin{aligned}
D_C^{(FBC)} &= \sum_{l=0}^{k-1} \|\mathbf{v}_l\|^2 2\epsilon(1-\epsilon) - \|\mathbf{v}_{k-1}\|^2 \left(2\epsilon(1-\epsilon)-\epsilon\right) \\
&\quad + \sum_{\substack{l=0 \\ l\neq m}}^{k-1}\sum_{m=0}^{k-1} \langle \mathbf{v}_l \,|\mathbf{v}_m \rangle \,\hat{p}_{e_l+e_m}\epsilon(1-\epsilon) \\
&= 8\epsilon(1-\epsilon)\sigma_{\mathbf{U}}^2 - \epsilon(1-2\epsilon)\|\mathbf{v}_{k-1}\|^2 \\
&\quad + \epsilon(1-\epsilon)\left(\sum_{l=0}^{k-1}\sum_{m=0}^{k-1}\langle \mathbf{v}_l \,|\mathbf{v}_m \rangle \sum_{i\in\mathbb{Z}_2^k} p_i h_{i,e_l+e_m} - \sum_{l=0}^{k-1}\|\mathbf{v}_l\|^2\right) \\
&= 4\epsilon(1-\epsilon)\sigma_{\mathbf{U}}^2 + \epsilon(1-\epsilon)\sum_{i\in\mathbb{Z}_2^k}\left\|\sum_{l=0}^{k-1}\mathbf{v}_l\,(2i_l-1)\right\|^2 - \epsilon(1-2\epsilon)\|\mathbf{v}_{k-1}\|^2 \\
&= 4\epsilon(1-\epsilon)\left(\sigma_{\mathbf{U}}^2 + \mathrm{E}\left\|\mathbf{Y}-\bar{\mathbf{U}}\right\|^2\right) - \epsilon(1-2\epsilon)\|\mathbf{v}_{k-1}\|^2 \\
&= 4\epsilon(1-\epsilon)\left(\sigma_{\mathbf{U}}^2 + \mathrm{E}\left\|\mathbf{Y}-\bar{\mathbf{Y}}+\bar{\mathbf{Y}}-\bar{\mathbf{U}}\right\|^2\right) - \epsilon(1-2\epsilon)\|\mathbf{v}_{k-1}\|^2 \\
&= 4\epsilon(1-\epsilon)\left(\sigma_{\mathbf{U}}^2 + \sigma_{\mathbf{Y}}^2 + \left\|\bar{\mathbf{Y}}-\bar{\mathbf{U}}\right\|^2\right) - \epsilon(1-2\epsilon)\max_l\|\mathbf{v}_l\|^2,
\end{aligned}$$

where $\|\mathbf{v}_{k-1}\| = \max_l \|\mathbf{v}_l\|$ follows, since the basis vectors are ordered by their norms. ■

Corollary 2.4 *Given the conditions of Corollary 2.1, the channel distortion of the Gray Code is*

$$\begin{aligned}
D_C^{(GC)} &= \frac{1}{2}\sum_{l=0}^{k-1}\|\mathbf{v}_l\|^2\left(1-(1-2\epsilon)^{k-l}\right) \\
&\quad + \sum_{\substack{l=0 \\ l\neq m}}^{k-1}\sum_{m=0}^{k-1}\langle \mathbf{v}_l \,|\mathbf{v}_m \rangle \,\hat{p}_{e_l+e_m}\left(1-(1-2\epsilon)^{k-l}-(1-2\epsilon)^{k-m}+(1-2\epsilon)^{|l-m|}\right).
\end{aligned}$$

Proof

Substituting $w(e_l F^{(GC)}) = k - l$ and $w((e_l + e_m)F^{(GC)}) = |l - m|$ in (2.9), the result is immediate. ∎

2.4.1.2 Comparison of the NBC and the FBC

While the NBC is known to be optimal for the binary symmetric channel with a uniform source, little is known about optimal codes for nonuniform sources. Corollary 2.2 and Corollary 2.3 can be used to compare the MSE performance of the NBC and the FBC for nonredundant source-channel coding. Noll found that for certain speech data the FBC achieves better performance than the NBC when used in conjunction with the optimal noiseless quantizer [36]. Corollary 2.5 characterizes sources for which the FBC outperforms the NBC, using BLVQ. The variance of the source determines which code is better.

Corollary 2.5 *Given the conditions of Corollary 2.1, and for all* $\epsilon < 1/2$,

$$D_C^{(FBC)} < D_C^{(NBC)} \quad \Leftrightarrow \quad \sigma_{\mathbf{Y}}^2 + \left\|\bar{\mathbf{Y}} - \bar{\mathbf{U}}\right\|^2 < \frac{1}{4}\max_l \left\|\mathbf{v}_l\right\|^2 .$$

2.4.2 Codes for binary asymmetric channels

Definition 2.6 For $i \in \mathbb{Z}_2^k$, let $B_i = \{l \in \{0, 1, \ldots, k-1\} : \langle i \,| e_l \rangle = 1\}$, i.e., the set of positions where the binary row vector i has nonzero coordinates. Then

$$i \prec j \Leftrightarrow B_i \subset B_j$$

defines a partial ordering "$\prec$" of the elements of $\mathbb{Z}_2^k$. Equivalently, $i \prec j \Leftrightarrow w(i + j) = w(j) - w(i) \qquad \forall i, j \in \mathbb{Z}_2^k$.

Theorem 2.2 *If a* 2^k*-point binary lattice vector quantizer with generating set* $\{\mathbf{v}_l\}_{l=0}^{k-1}$ *induces equiprobable quantizer codevectors, and an affine index assignment with generator matrix* G *and translation vector* t *is used to transmit across a binary asymmetric channel* *with transition probabilities* $p_{1|0} = \epsilon$ *and* $p_{0|1} = \delta$ *and with a nonredundant channel code, then the channel distortion is given by*

$$\begin{aligned} D_C &= \frac{1}{2}\sum_{l=0}^{k-1} \|\mathbf{v}_l\|^2 \left(1-\gamma^{w(e_l F)}\right) + \frac{1}{4}\sum_{\substack{l=0\\ l\neq m}}^{k-1}\sum_{m=0}^{k-1} \langle \mathbf{v}_l \,|\mathbf{v}_m\rangle \, h_{t,(e_l+e_m)F}\,\beta^{w((e_l+e_m)F)} \\ &\quad \cdot \left(1-\chi_{\{(e_l+e_m)F\prec e_l F\}}\gamma^{w(e_m F)} - \chi_{\{(e_l+e_m)F\prec e_m F\}}\gamma^{w(e_l F)}\right), \end{aligned} \tag{2.12}$$

where $\gamma = 1-\epsilon-\delta$, $\beta = \delta-\epsilon$, $w(\cdot)$ *denotes Hamming weight,* $\mathrm{F} = \mathrm{G}^{-t}$, e_l *is the binary row vector with its only nonzero entry in the* l^{th} *position, and* $h_{i,j} = (-1)^{ij^t}$ *denotes a Hadamard transform coefficient.*

Note that at most one of the two indicator functions in Theorem 2.2 can be nonzero for any pair l and m ($l \neq m$).

Proof

For nonsymmetric channels, $p_{j|i}$ does not depend only on the (modulo 2) sum $i+j$, so an approach different from the one used in the proof of Theorem 2.1 is necessary. Let the random variable $I = \mathcal{E}_Q(\mathbf{X})$ denote the k-bit source coded index, and let W denote the k-bit binary channel error vector. The decoded k-bit index J is then

$$J = \pi^{-1}(\eta(\pi(I))) = ([(I\mathrm{G}+t)+W]+t)\mathrm{G}^{-1} = I + W\mathrm{G}^{-1},$$

as depicted in Figure 2.3.

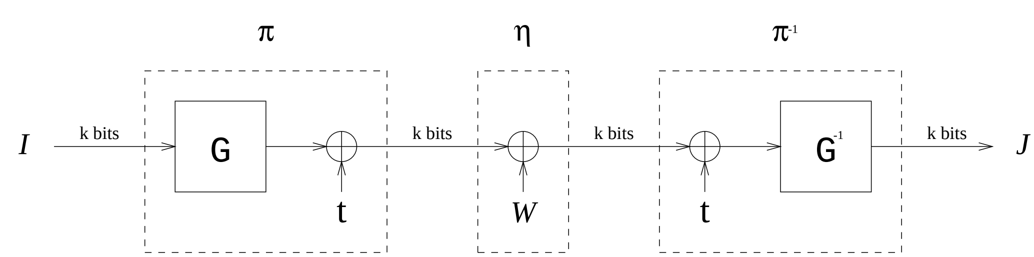


Figure 2.3 Channel subsystem with affine index assignment.

Thus, the channel distortion of a BLVQ (with codevectors $\mathbf{y}_i = \mathbf{y}_0 + \sum_{l=0}^{k-1} \mathbf{v}_l i_l$ for $i \in \mathbb{Z}_2^k$) can be written as

$$
\begin{aligned}
D_C &= \mathrm{E}\left\|\mathbf{y}_J - \mathbf{y}_I\right\|^2 \\
&= \mathrm{E}\left\|\sum_{l=0}^{k-1} \mathbf{v}_l (J_l - I_l)\right\|^2 \\
&= \mathrm{E}\left\|\sum_{l=0}^{k-1} \mathbf{v}_l \left(\frac{1 - h_{I+W\mathrm{G}^{-1},e_l}}{2} - \frac{1 - h_{I,e_l}}{2}\right)\right\|^2 \\
&= \frac{1}{4}\sum_{l=0}^{k-1}\sum_{m=0}^{k-1} \langle \mathbf{v}_l \,|\mathbf{v}_m\rangle \,\mathrm{E}\left[h_{I,e_l}\left(1 - h_{W,e_l F}\right) h_{I,e_m}\left(1 - h_{W,e_m F}\right)\right] \\
&= \frac{1}{4}\sum_{l=0}^{k-1}\sum_{m=0}^{k-1} \langle \mathbf{v}_l \,|\mathbf{v}_m\rangle \,\mathrm{E}\left[h_{I,e_l+e_m}\mathrm{E}\left[\left(1 - h_{W,e_l F} - h_{W,e_m F} + h_{W,(e_l+e_m)F}\right)|I\right]\right]. \quad (2.13)
\end{aligned}
$$

The k bits of W are conditionally independent given I, and satisfy

$$
\begin{aligned}
\mathrm{P}\left[W_l = 1 | (IG + t)_l = 0\right] &= \epsilon \\
\mathrm{P}\left[W_l = 1 | (IG + t)_l = 1\right] &= \delta
\end{aligned}
$$

on a binary asymmetric channel. Thus,

$$
\begin{aligned}
\mathrm{E}\left[h_{W,e_r}|I\right] &= \frac{1 + h_{IG+t,e_r}}{2}(1 - 2\epsilon) + \frac{1 - h_{IG+t,e_r}}{2}(1 - 2\delta) \\
&= (1 - \epsilon - \delta) + (\delta - \epsilon) h_{IG+t,e_r} \\
&= \gamma + \beta h_{IG+t,e_r}.
\end{aligned}
$$

Hence, for any k-bit binary row vector f,

$$
\begin{aligned}
\mathrm{E}\left[h_{W,f}|I\right] &= \prod_{e_r \prec f} \mathrm{E}\left[h_{W,e_r}|I\right] \\
&= \prod_{e_r \prec f} \left(\gamma + \beta h_{IG+t,e_r}\right) \\
&= \sum_{a \prec f} \gamma^{w(f)-w(a)} \beta^{w(a)} h_{IG+t,a}.
\end{aligned}
$$

With equiprobable codepoints, the k bits of I are independent and are equally likely to be 0 or 1. Hence, $\mathrm{E}\left[h_{I,a}\right] = 0$ for any nonzero k-bit binary row vector a, and we have

$$\begin{aligned}
\mathrm{E}\left[h_{I,e_l+e_m}\mathrm{E}\left[h_{W,f}|I\right]\right] &= \sum_{a\prec f}\gamma^{w(f)-w(a)}\beta^{w(a)}h_{t,a}\mathrm{E}\left[h_{I,e_l+e_m+aG^t}\right] \\
&= \sum_{a\prec f}\gamma^{w(f+a)}\beta^{w(a)}h_{t,a}\chi_{\{a=(e_l+e_m)F\}} \\
&= \chi_{\{(e_l+e_m)F\prec f\}}\gamma^{w(f+(e_l+e_m)F)}\beta^{w((e_l+e_m)F)}h_{t,(e_l+e_m)F}.
\end{aligned}$$

Substituting e_lF, e_mF, and $(e_l+e_m)F$ for f, the last three terms within the expectations in (2.13) are obtained. Noting that $\mathrm{E}\left[h_{I,e_l+e_m}\right] = \chi_{\{l=m\}}$ and factoring out common terms gives

$$\begin{aligned}
D_C &= \frac{1}{4}\sum_{l=0}^{k-1}\sum_{m=0}^{k-1}\langle \mathbf{v}_l\,|\mathbf{v}_m\rangle\,\chi_{\{l=m\}} + \frac{1}{4}\sum_{l=0}^{k-1}\sum_{m=0}^{k-1}\langle \mathbf{v}_l\,|\mathbf{v}_m\rangle\,\beta^{w((e_l+e_m)F)}h_{t,(e_l+e_m)F} \\
&\quad\cdot\left(-\chi_{\{(e_l+e_m)F\prec e_lF\}}\gamma^{w(e_mF)} - \chi_{\{(e_l+e_m)F\prec e_mF\}}\gamma^{w(e_lF)} + 1\right) \\
&= \frac{1}{2}\sum_{l=0}^{k-1}\|\mathbf{v}_l\|\left(1-\gamma^{w(e_lF)}\right) + \frac{1}{4}\sum_{\substack{l=0\\ l\neq m}}^{k-1}\sum_{m=0}^{k-1}\langle \mathbf{v}_l\,|\mathbf{v}_m\rangle\,\beta^{w((e_l+e_m)F)}h_{t,(e_l+e_m)F} \\
&\quad\cdot\left(1-\chi_{\{(e_l+e_m)F\prec e_lF\}}\gamma^{w(e_mF)} - \chi_{\{(e_l+e_m)F\prec e_mF\}}\gamma^{w(e_lF)}\right),
\end{aligned}$$

and the proof is complete. ■

2.4.2.1 Formulas for structured index assignments

Here we specialize Theorem 2.2 to the FBC, the GC, the NBC, and the affine translates of the NBC. The formulas presented generalize those given in [49] for the uniform scalar quantizer case (the α, β, γ notation is consistent with [49]), and generalize those given in [30, 32] to nonsymmetric channels. Also, by letting $\epsilon = \delta$, the special cases of Corollaries 2.2-2.4 for the uniform output distribution case are recovered.

Corollary 2.6 *Given the conditions of Theorem 2.2, the channel distortion of the affine translate of the Natural Binary Code corresponding to translation vector t is*

$$\begin{aligned}
D_C^{(NBC+t)} &= \frac{1}{2}\left(\alpha\sum_{l=0}^{k-1}\|\mathbf{v}_l\|^2 + \beta^2\sum_{l=0}^{k-1}\sum_{m=0}^{l-1}\langle \mathbf{v}_l\,|\mathbf{v}_m\rangle\,(-1)^{t_l+t_m}\right) \\
&= (2\alpha-\beta^2)\sigma_{\mathbf{U}}^2 + \beta^2\left\|\mathbf{y}_t - \bar{\mathbf{U}}\right\|^2,
\end{aligned}$$

where $\alpha = \epsilon + \delta = 1 - \gamma$, *and* $\mathbf{y}_t = \mathbf{y}_0 + \sum_{l=0}^{k-1} \mathbf{v}_l t_l$. *In particular, the channel distortion of the NBC* $(t = [00\ldots0])$ *is*

$$D_C^{(NBC)} = \frac{1}{4}\left((2\alpha - \beta^2)\sum_{l=0}^{k-1}\|\mathbf{v}_l\|^2 + \beta^2\left\|\sum_{l=0}^{k-1}\mathbf{v}_l\right\|^2\right)$$

and the channel distortion of the TCC $(t = [10\ldots0])$ *is*

$$D_C^{(TCC)} = \frac{1}{4}\left((2\alpha - \beta^2)\sum_{l=0}^{k-1}\|\mathbf{v}_l\|^2 + \beta^2\left\|\mathbf{v}_{k-1} - \sum_{l=0}^{k-2}\mathbf{v}_l\right\|^2\right).$$

Proof Since $w(e_l F^{(NBC)}) = 1$ and $w((e_l + e_m)F^{(NBC)}) = 2\chi_{\{l \neq m\}}$, for all l and m, no row of $F^{(NBC)}$ can precede the sum of two rows in the partial ordering. Using this and Theorem 2.2, the statement follows. ■

For a uniform scalar quantizer with step size Δ, we have $\mathbf{v}_l = 2^l\Delta$ and the above expressions simplify to

$$D_C^{(NBC)} = \frac{\Delta^2}{6}\left(\alpha\left(4^k - 1\right) + \beta^2\left(4^k - 3\cdot 2^k + 2\right)\right),$$

and

$$D_C^{(TCC)} = \frac{\Delta^2}{6}\left(\alpha\left(4^k - 1\right) - 2\beta^2\left(4^{k-1} - 1\right)\right).$$

The formula for $D_C^{(NBC)}$ generalizes results in [31,33] to the asymmetric channel case.

Corollary 2.7 *Given the conditions of Theorem 2.2, the channel distortion of the Folded Binary Code is*

$$D_C^{(FBC)} = \frac{1}{2}\left(\alpha\sum_{l=0}^{k-1}\|\mathbf{v}_l\|^2 + \alpha(1-\alpha)\sum_{l=0}^{k-2}\|\mathbf{v}_l\|^2 \right.$$
$$\left. + \beta^2\sum_{l=0}^{k-2}\sum_{m=0}^{l-1}\langle \mathbf{v}_l \,|\mathbf{v}_m\rangle - \alpha\beta\sum_{l=0}^{k-2}\langle \mathbf{v}_l\,|\mathbf{v}_{k-1}\rangle\right),$$

where $\alpha = \epsilon + \delta = 1 - \gamma$.

For a uniform scalar quantizer with step size Δ, we have $\mathbf{v}_l = 2^l\Delta$. Thus, $\langle \mathbf{v}_l \,|\mathbf{v}_m \rangle = 4^{l+m}\Delta^2$ and the above formula becomes

$$
\begin{aligned}
D_C^{(FBC)} &= \frac{\Delta^2}{6}\left(\alpha\left(4^k-1\right)+\alpha(1-\alpha)\left(4^{k-1}-1\right)\right. \\
&\quad \left. +\beta^2\left(4^{k-1}-3\cdot 2^{k-1}+2\right)-3\alpha\beta\left(4^{k-1}-2^{k-1}\right)\right).
\end{aligned}
$$

Proof For $l \neq m$,

$$
(e_l+e_m)F^{(FBC)} = \begin{cases} e_l & m=k-1 \\ e_m & l=k-1 \\ e_l+e_m & \text{otherwise} \end{cases},
$$

and thus the indicator functions in Theorem 2.2 will only be nonzero if either $l = k-1$ or $m = k-1$. Using this and $t = [01\dots1]$, and substituting the Hamming weights of the rows of $\mathrm{F}^{(FBC)}$ in (2.12), one gets

$$
\begin{aligned}
D_C^{(FBC)} &= \frac{1}{2}\left\|\mathbf{v}_{k-1}\right\|^2(1-\gamma)+\frac{1}{2}\sum_{l=0}^{k-2}\left\|\mathbf{v}_l\right\|^2\left(1-\gamma^2\right)+\frac{1}{4}\sum_{\substack{l=0\\ l\neq m}}^{k-2}\sum_{m=0}^{k-2}\langle \mathbf{v}_l \,|\mathbf{v}_m \rangle\, h_{t,e_l+e_m}\beta^2 \\
&+ \frac{1}{4}\sum_{l=0}^{k-2}\langle \mathbf{v}_l \,|\mathbf{v}_{k-1} \rangle\, h_{t,e_l}\beta^2(1-\gamma)+\frac{1}{4}\sum_{m=0}^{k-2}\langle \mathbf{v}_{k-1} \,|\mathbf{v}_m \rangle\, h_{t,e_m}\beta^2(1-\gamma) \\
&= \frac{1}{2}\left(\alpha\sum_{l=0}^{k-1}\left\|\mathbf{v}_l\right\|^2+\left(2\alpha-\alpha^2-\alpha\right)\sum_{l=0}^{k-2}\left\|\mathbf{v}_l\right\|^2\right. \\
&\quad \left. +\beta^2\sum_{l=0}^{k-2}\sum_{m=0}^{l-1}\langle \mathbf{v}_l \,|\mathbf{v}_m \rangle-\alpha\beta^2\sum_{l=0}^{k-2}\langle \mathbf{v}_l \,|\mathbf{v}_{k-1} \rangle\right).
\end{aligned}
$$

■

Corollary 2.8 *Given the conditions of Theorem 2.2, the channel distortion of the Gray Code is*

$$
D_C^{(GC)} = \frac{1}{2}\left(\sum_{l=0}^{k-1}\left\|\mathbf{v}_l\right\|^2\left(1-\gamma^{k-l}\right)+\sum_{l=0}^{k-1}\sum_{m=0}^{l-1}\langle \mathbf{v}_l \,|\mathbf{v}_m \rangle\, \beta^{l-m}\left(1-\gamma^{k-l}\right)\right).
$$

For a uniform scalar quantizer with step size Δ, we have $\mathbf{v}_l = 2^l\Delta$. Thus, $\langle \mathbf{v}_l \,|\mathbf{v}_m \rangle = 4^{l+m}\Delta^2$ and the above formula becomes

$$D_C^{(GC)} = \frac{\Delta^2}{2}\left(\frac{4^k-1}{3} - \gamma\frac{4^k-\gamma^k}{4-\gamma}\right.$$
$$\left. + \frac{\beta}{2-\beta}\left(\frac{4^k-1}{3} - \gamma\frac{4^k-\gamma^k}{4-\gamma} - \frac{(2\beta)^k-1}{2\beta-1} + \gamma\frac{(2\beta)^k-\gamma^k}{2\beta-\gamma}\right)\right).$$

Proof The statement follows by observing that the precedence $(e_l + e_m)F \prec e_m F$ is satisfied if and only if $l > m$, and substituting $w(e_l F) = k - l$, $w((e_l + e_m)F) = l - m$ for $l > m$, and $t = 0$ in (2.12). ■

2.4.2.2 Affine translates of the NBC

The family of affine translates of the NBC is known to perform optimally for BLVQs with a uniform output distribution on a BSC. If, however, the channel is asymmetric, different translates result in different distortions. The best one is identified next.

Corollary 2.9 *If a 2^k-point binary lattice vector quantizer induces equiprobable quantizer codevectors for a given source, and if it transmits an affine translation of the Natural Binary Code across a binary asymmetric channel with crossover probabilities $p_{1|0} = \epsilon$ and $p_{0|1} = \delta$ and with a nonredundant channel code, then the channel distortion is minimized if and only if the translation vector t satisfies*

$$t = \operatorname*{argmin}_{\{i\in\mathbb{Z}_2^k\}} \left\|\mathbf{y}_i - \bar{\mathbf{U}}\right\|,$$

where $\bar{\mathbf{U}} = 2^{-k}\sum_{i\in\mathbb{Z}_2^k}\mathbf{y}_i$ is the arithmetic mean of the codebook. In particular, the Two's Complement Code is optimal among the NBC translates for uniform scalar quantization.

Proof Immediate from Corollary 2.6. ■

For a uniform scalar quantizer with step size Δ, $v_l = 2^l\Delta$, and $\bar{U} = y_0 + (2^{k-1} - \frac{1}{2})\Delta$. Thus both $t = [01\ldots1]$ $(y_t = y_0 + (2^{k-1} - 1)\Delta)$, and $t = [10\ldots0]$ $(y_t = y_0 + 2^{k-1}\Delta)$ have the same performance (optimal among the translates of the NBC). The latter translate is the Two's Complement Code (a rotation of the Odd-Even Code of [49]).

2.4.2.3 Comparisons for uniform scalar quantization

Based on the formulas presented in Corollaries 2.6-2.9 for uniform scalar quantization, the structured index assignments we have considered can be compared. First, we define the one's complement of an index assignment. This corresponds to changing 0's to 1's and 1's to 0's in the binary representation of the indices. Unless the performance of an assignment is symmetric in ϵ and δ, it is advantageous to use the one's complement of the assignment instead of the assignment itself, either when $\epsilon < \delta$ or $\epsilon > \delta$.

Definition 2.7 The *one's complement index assignment* $\overline{X}$, of an index assignment X, is defined by

$$\pi^{\overline{X}}(i) = \pi^{(X)}(i) + \mathbf{1} \qquad \forall i \in \mathbb{Z}_2^k,$$

where $\mathbf{1} = [11 \cdots 1]$ (the vector of weight k).

The one's complement of an affine index assignment can be obtained by replacing its translation vector t by $\bar{t}$, the one's complement of t (the generator matrix remains unchanged). The distortion formulas are also easily updated, as only the roles of ϵ and δ have to be exchanged (or equivalently, β is to be replaced by $-\beta$). Hence, the one's complement of an index assignment whose distortion formula includes only even powers of β (e.g., NBC, TCC) has the same performance as the original assignment. Furthermore, since odd powers of β change sign when $\epsilon = \delta$, the one's complement outperforms the original assignment either when $\epsilon < \delta$ or $\epsilon > \delta$.

Theorem 2.3 *Given a uniform 2^k-level scalar quantizer for a uniform source, the channel distortions of the Natural Binary Code (NBC), the Folded Binary Code (FBC), the Gray Code (GC), and the Two's Complement Code (TCC) on a binary memoryless channel with $p_{1|0} = \epsilon$ and $p_{0|1} = \delta$ and with a nonredundant channel code satisfy (assuming $0 \neq \delta \geq \epsilon$ and $\epsilon + \delta < 1$):*

$$(i) \qquad D_C^{(FBC)} < D_C^{(\overline{FBC})} \qquad \forall k > 1, \quad \forall \epsilon \neq \delta$$

$$(ii) \qquad D_C^{(\overline{GC})} < D_C^{(GC)} \qquad \forall k > 1, \quad \forall \epsilon \neq \delta$$

$$(iii) \qquad D_C^{(TCC)} < D_C^{(NBC)} \qquad \forall k > 1, \quad \forall \epsilon \neq \delta$$

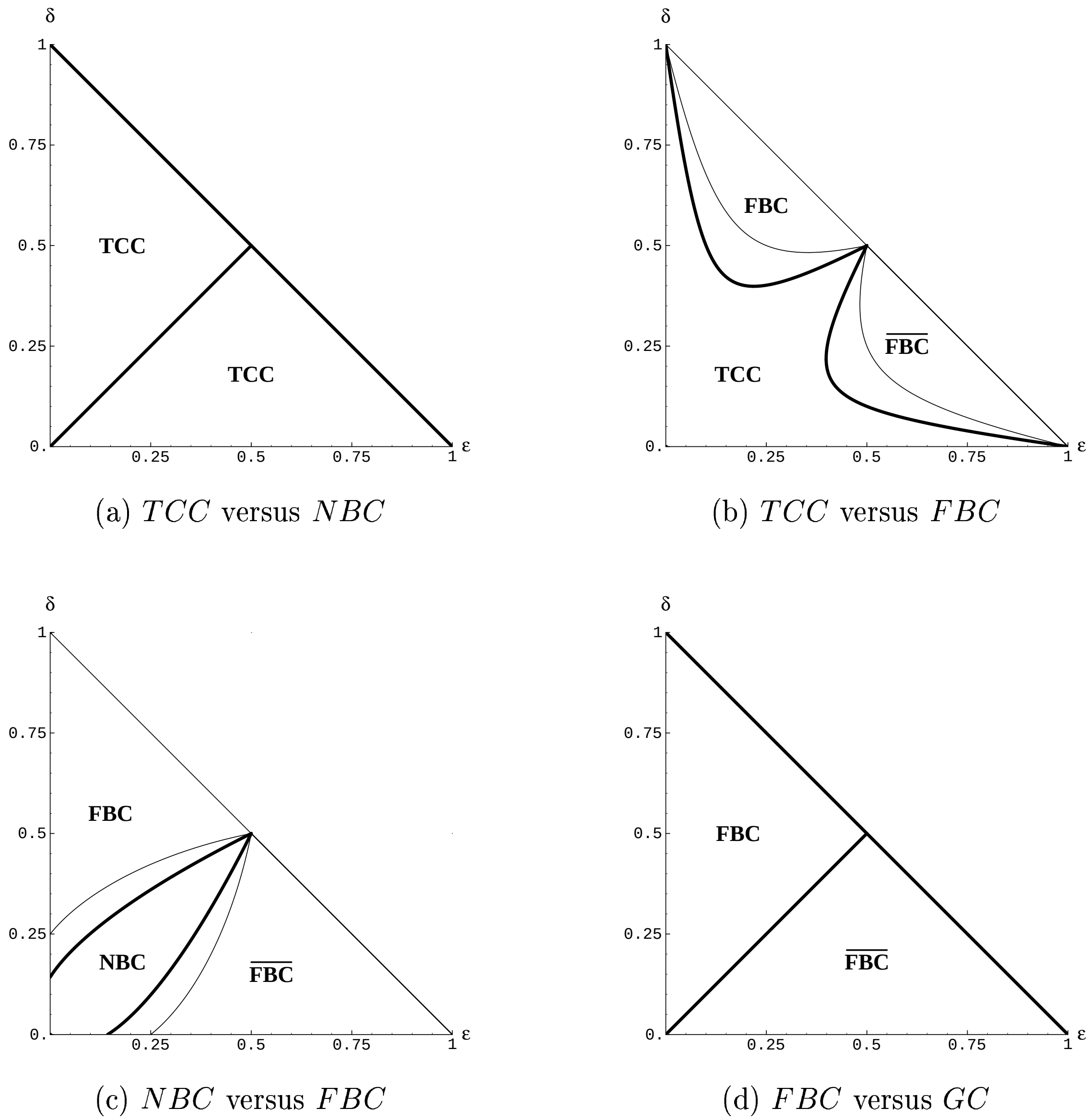
$$(iv) \qquad D_C^{(FBC)} < D_C^{(\overline{GC})} \qquad \forall k > 2, \quad \forall \epsilon, \delta$$

$$(v) \qquad D_C^{(TCC)} < D_C^{(FBC)} \qquad \begin{array}{l} \forall k > 1 \qquad \textit{if} \;\; \epsilon + \delta + 5\epsilon^2 - 8\epsilon\delta - \delta^2 \geq 0 \\ \forall k \leq 1 + \log_2 \dfrac{(\epsilon+\delta)(\epsilon+\delta-1)}{\epsilon+\delta+5\epsilon^2-8\epsilon\delta-\delta^2} \\ \qquad \textit{if} \;\; \epsilon + \delta + 5\epsilon^2 - 8\epsilon\delta - \delta^2 < 0 \\ \qquad \textit{and} \;\; \epsilon + \delta + 2\epsilon^2 - 5\epsilon\delta - \delta^2 \geq 0, \end{array}$$

$$(vi) \qquad D_C^{(NBC)} < D_C^{(FBC)} \qquad \begin{array}{l} \forall k > 1 \qquad \textit{if} \;\; \epsilon + \delta - \epsilon^2 + 4\epsilon\delta - 7\delta^2 \geq 0 \\ \forall k \leq 1 + \log_2 \dfrac{(\epsilon+\delta)(\epsilon+\delta-1)}{\epsilon+\delta-\epsilon^2+4\epsilon\delta-7\delta^2} \\ \qquad \textit{if} \;\; \epsilon + \delta - \epsilon^2 + 4\epsilon\delta - 7\delta^2 < 0 \\ \qquad \textit{and} \;\; \epsilon + \delta - \epsilon^2 + \epsilon\delta - 4\delta^2 \geq 0, \end{array}$$

The inequalities (i),(ii),(iii) hold with equality if $\epsilon = \delta$.

The above inequalities follow from Corollaries 2.6-2.8 by straightforward algebraic manipulations; thus, their proofs are omitted. The code comparisons of the above theorem are shown in Figure 2.4. In each graph two index assignments (and/or their one's complements) are compared for binary asymmetric channels (each point (ϵ, δ) corresponds to a different channel). The region where one code is uniformly better (i.e., $\forall k > 2$) than the other is marked by the name of the superior one. In the unmarked area (between the thick and the thin curves, where applicable) the winner depends on the value of n.



(a) TCC versus NBC

(b) TCC versus FBC

(c) NBC versus FBC

(d) FBC versus GC

Figure 2.4 Performance comparisons of various nonredundant codes for a uniform scalar source on a binary asymmetric channel with $p_{1|0} = \epsilon$ and $p_{0|1} = \delta$, and $\epsilon + \delta < 1$. The region where one code is uniformly better (i.e., $\forall k > 2$) than the other is marked by the name of the superior one. In the unmarked area (between the thick and the thin curves, where applicable) the winner depends on the value of n. The thick curves correspond to ties between the codes.

2.5 References

[1] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression.* Boston: Kluwer Academic Publishers, 1992.

[2] A. J. Kurtenbach and P. A. Wintz, "Quantizing for noisy channels," *IEEE Trans. Communication Technology*, vol. COM-17, pp. 291–302, April 1969.

[3] J. Dunham and R. M. Gray, "Joint source and noisy channel trellis encoding," *IEEE Trans. Info. Theory*, vol. IT-27, pp. 516–519, July 1981.

[4] H. Kumazawa, M. Kasahara, and T. Namekawa, "A construction of vector quantizers for noisy channels," *Electronics and Engineering in Japan*, vol. 67-B, no. 4, pp. 39–47, 1984.

[5] E. Ayanoğlu and R. M. Gray, "The design of joint source and channel trellis waveform coders," *IEEE Trans. Info. Theory*, vol. IT-33, pp. 855–865, November 1987.

[6] N. Farvardin and V. Vaishampayan, "Optimal quantizer design for noisy channels: an approach to combined source-channel coding," *IEEE Trans. Info. Theory*, vol. IT-33, pp. 827–838, November 1987.

[7] K. Zeger and A. Gersho, "Vector quantization design for memoryless noisy channels," in *Proc. IEEE Int. Conf. Commun.*, (Philadelphia, PA), June 1988.

[8] N. Farvardin, "A study of vector quantization for noisy channels," *IEEE Trans. Info. Theory*, vol. IT-36, pp. 799–809, July 1990.

[9] N. Farvardin and V. Vaishampayan, "On the performance and complexity of channel-optimized vector quantizers," *IEEE Trans. Info. Theory*, vol. IT-37, pp. 155–160, January 1991.

[10] D. Miller and K. Rose, "Combined source-channel vector quantization using deterministic annealing," *IEEE Trans. Communications*, vol. COM-42, pp. 347–356, February/March/April 1994.

[11] K. Chang and R. W. Donaldson, "Analysis, optimization, and sensitivity study of differential PCM systems operating on noisy communication channels," *IEEE Trans. Communications*, vol. COM-20, pp. 338–350, June 1972.

[12] Q. Chen and T. R. Fischer, "Robust quantization for image coding and noisy digital transmission," in *Proc. Data Compression Conf.*, (Snowbird, UT), pp. 3–12, March 1996.

[13] Q. Chen and T. R. Fischer, "Image coding using robust quantization for noisy digital transmission," *IEEE Trans. Image Processing*, vol. 7, pp. 496–505, April 1998.

[14] J. W. Modestino and D. G. Daut, "Combined source-channel coding of images," *IEEE Trans. Communications*, vol. COM-27, pp. 1644–1659, November 1979.

[15] J. W. Modestino, D. G. Daut, and A. L. Vickers, "Combined source-channel coding of images using the block cosine transform," *IEEE Trans. Communications*, vol. COM-29, pp. 1261–1274, September 1981.

[16] D. R. Comstock and J. D. Gibson, "Hamming coding of DCT-compressed images over noisy channels," *IEEE Trans. Communications*, vol. COM-32, pp. 856–861, July 1984.

[17] K. Sayood, F. Liu, and J. D. Gibson, "A constrained joint source/channel coder design," *IEEE J. Sel. Areas Comm.*, vol. JSAC-12, pp. 1584–1593, December 1994.

[18] B. Hochwald and K. Zeger, "Tradeoff between source and channel coding," *IEEE Trans. Info. Theory*, vol. IT-43, pp. 1412–1424, September 1997.

[19] R. G. Gallager, *Information Theory and Reliable Communication.* New York: Wiley, 1968.

[20] J. D. Marca and N. Jayant, "An algorithm for assigning binary indices to the codevectors of a multi-dimensional quantizer," in *Proceedings IEEE Inter. Conf. on Communications*, vol. 1, (Seattle), pp. 1128–1132, June 1987.

[21] H.-S. Wu and J. Barba, "Index allocation in vector quantisation for noisy channels," *Electronics Letters*, vol. 29, pp. 1318–1320, July 1993.

[22] K. Zeger and A. Gersho, "Pseudo-Gray coding," *IEEE Trans. Communications*, vol. COM-38, pp. 2147–2158, December 1990.

[23] P. Knagenhjelm and E. Agrell, "The Hadamard transform – a tool for index assignment," *IEEE Trans. Info. Theory*, vol. IT-42, pp. 1139–1151, July 1996.

[24] D. J. Goodman and T. J. Moulsley, "Using simulated annealing to design transmission codes for analogue sources," *Electronics Letters*, vol. 24, pp. 617–618, May 1988.

[25] N. Moayeri, "Joint source and channel coding based on tree structured vector quantization," in *CISS-89*, (Baltimore, MD), pp. 729–734, March 1989.

[26] K. C. Chua, W. C. Wong, and K. N. Ngan, "Error detection and correction of vector quantised digital images," *IEE Proceedings*, vol. 137, pp. 417–423, December 1990.

[27] R.-Y. Wang, E. A. Riskin, and R. Ladner, "Codebook organization to enhance maximum a posterori detection of progressive transmission of vector quantized images over noisy channels," in *ICASSP-93*, vol. V, (Minneapolis, MN), pp. 233–236, April 1993.

[28] K. Zeger and V. Manzella, "Asymptotic bounds on optimal noisy channel quantization via random coding," *IEEE Trans. Info. Theory*, vol. IT-40, pp. 1926–1938, November 1994.

[29] N. S. Jayant and P. Noll, *Digital Coding of Waveforms: Principles and Applications to Speech and Video.* Englewood Cliffs, New Jersey: Prentice-Hall, 1984.

[30] Y. Yamaguchi and T. S. Huang, "Optimum binary fixed-length block codes," Quarterly Progress Report 78, M.I.T. Research Lab. of Electronics, Cambridge, Mass., July 1965.

[31] T. S. Huang, "Optimum binary code," Quarterly Progress Report 82, M.I.T. Research Lab. of Electronics, Cambridge, Mass., July 1966.

[32] T. S. Huang et al., "Design considerations in PCM transmission of low-resolution monochrome still pictures," *Proceedings of the IEEE*, vol. 55, pp. 331–335, March 1967.

[33] T. R. Crimmins, H. M. Horwitz, C. J. Palermo, and R. V. Palermo, "Minimization of mean-square error for data transmitted via group codes," *IEEE Trans. Info. Theory*, vol. IT-15, pp. 72–78, January 1969.

[34] S. W. McLaughlin, D. L. Neuhoff, and J. J. Ashley, "Optimal binary index assignments for a class of equiprobable scalar and vector quantizers," *IEEE Trans. Info. Theory*, vol. IT-41, pp. 2031–2037, November 1995.

[35] I. Dostis, "The effect of digital errors on PCM transmission of compandored speech," *Bell Syst. Tech. J.*, vol. 44, pp. 2227–2243, December 1965.

[36] P. Noll, "Effects of channel errors on the signal-to-noise performance of speech encoding systems," *Bell Syst. Tech. J.*, vol. 54, pp. 1615–1636, November 1975.

[37] R. Zelinski, "Effects of transmission errors on the mean-squared error performance of transform coding systems," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-27, pp. 531–537, October 1979.

[38] T. R. Crimmins and H. M. Horwitz, "Mean-square-error optimum coset leaders for group codes," *IEEE Trans. Info. Theory*, vol. IT-16, pp. 429–432, July 1970.

[39] T. R. Crimmins, "On encoding and decoding maps for group codes," *IEEE Trans. Info. Theory*, vol. IT-22, pp. 763–764, November 1976.

[40] G. R. Redinbo and G. A. Wolf, "On minimum mean-square error linear block codes when the data have q-adic weighting," *Information and Control*, vol. 26, pp. 154–177, 1974.

[41] G. A. Wolf and G. R. Redinbo, "The optimum mean-square estimate for decoding binary block codes," *IEEE Trans. Info. Theory*, vol. IT-20, pp. 344–351, May 1974.

[42] J. Ashley, "Error correcting codes and quantization," preprint.

[43] A. S. Khayrallah, "On the mean squared error of linear codes," in *Proc. 32nd Ann. Allerton Conf.*, (Illinois), pp. 503–512, September 1994.

[44] R. Hagen and P. Hedelin, "Robust vector quantization by linear mappings of block-codes," in *Proc. IEEE Int. Symp. Information Theory*, (San Antonio, TX), p. 171, January 1993.

[45] R. Hagen and P. Hedelin, "Design methods for VQ by linear mappings of block codes," in *Proc. IEEE Int. Symp. Information Theory*, (Trondheim, Norway), p. 241, June 1994.

[46] R. Hagen and P. Hedelin, "Robust vector quantization by a linear mapping of a block-code," to appear in *IEEE Trans. Info. Theory.*

[47] P. Knagenhjelm, "Robust vector quantization (tutorial)," Technical Report 12, Dept. of Information Theory, Chalmers Univ. of Technology, May 1993.

[48] W. M. Lam and A. R. Reibman, "Asymptotic quantization for signal estimation with noisy channels," in *ICASSP-91*, (Toronto, Canada), pp. 1901–1904, May 1991.

[49] A. Méhes and K. Zeger, "Redundancy free codes for binary discrete memoryless channels," in *Proc. Conf. Information Science and Systems*, (Princeton, NJ), pp. 1057–1062, 1994.

[50] P. Hedelin, P. Knagenhjelm, and M. Skoglund, *Theory for Transmission of Vector Quantization Data*, ch. 10 in *Speech Coding and Synthesis.* Amsterdam, The Netherlands: Elsevier Science, 1995. Editors W. B. Kleijn and K. K. Paliwal.

CHAPTER 3

RANDOMLY CHOSEN INDEX ASSIGNMENTS ARE ASYMPTOTICALLY BAD FOR UNIFORM SOURCES

It is known that among all redundancy-free codes (or index assignments) the Natural Binary Code minimizes the mean squared error of the uniform source and uniform quantizer on a binary symmetric channel. In this chapter, we derive a code which maximizes the mean squared error, and we demonstrate that the code is linear and that its distortion is asymptotically equivalent, as the blocklength grows, to the expected distortion of an index assignment chosen uniformly at random.

3.1 Introduction

An index assignment is a mapping of source code symbols to channel code symbols. The usual goal of index assignment design for noisy channel vector quantizers is to minimize the end-to-end mean squared error (MSE) over all possible index assignments. The MSE is computed with respect to the statistics of both the source and the channel. Previous work has examined the theoretical and practical aspects of index assignment in noisy channel vector quantizer systems. In particular, it is known that the performance of such a system can be significantly affected by the choice of index assignment.

The problem of algorithmically finding good index assignments has been previously studied in [1–6], and analytic formulas have been found for binary symmetric channels

The material in this chapter will appear in the *IEEE Transactions on Information Theory* as: A. Méhes and K. Zeger, "Randomly chosen index assigments are asymptotically bad for uniform sources."

and certain sources [7–12]. The optimality of the Natural Binary Code was conjectured in [8] and proven in [10] for uniform scalar quantization of a uniform source, and later extended to binary lattice vector quantizers (BLVQs) with equiprobable quantization points in [11].

In this chapter we derive an index assignment which *maximizes* the MSE for a uniform scalar source, and we show that the worst case performance thus obtained is asymptotically equivalent to the expected performance of an index assignment chosen uniformly at random. This indicates that the majority of index assignments are asymptotically bad. Also, this result analytically reveals the entire range of possible performances achievable by different index assignments.

The overall mean squared error of a quantizer optimized for a noiseless channel can be decomposed into a "source distortion" due to quantization and a "channel distortion" due to channel noise [13]. The source component is a result of representing the source with a finite number of quantization points, and thus is independent of the index assignment. The channel component, on the other hand, results from confusing the indices of quantization points because of channel errors. Hence, we focus on the channel distortion when evaluating index assignments. With this in mind, the index assignment problem can be reformulated as a discrete problem with no direct reference to quantization. The usual index assignment problem is to assign indices to quantization points to minimize the mean squared error within the finite set of quantization points. For n-bit uniform scalar quantization of a uniform source, this finite set of points is a scaled and translated version of the set $\{0, \dots, 2^n - 1\}$. In this chapter, however, we maximize the MSE.

This chapter is organized as follows. Section 3.2 gives notation and definitions. In Section 3.3, we derive a distortion-maximizing index assignment (the Worst Code) for uniform scalar quantization of a uniform source (Theorem 3.1), and we compare the performances of the best, worst, and randomly chosen index assignments (Corollary 3.1). A counterexample in Section 3.4 shows that the MSE-maximizing property of the Worst Code does not extend to arbitrary BLVQs (Corollary 3.3), even though it is known that the MSE-minimizing property of the Natural Binary Code does extend to BLVQs.

We establish, however, that among all affine index assignments the Worst Code does maximize the MSE of arbitrary BLVQs (Corollary 3.2).

3.2 Preliminaries

For any positive integer n, let $\mathbb{Z}_2^n$ denote the field of n-bit binary words, where arithmetic is performed modulo 2. Every integer $i \in \mathcal{S} = \{0, \ldots, 2^n - 1\}$ has a unique binary representation $i = \sum_{l=0}^{n-1} 2^l i_l$, where $i_l \in \{0, 1\}$. We denote by $\mathbf{i} \in \mathbb{Z}_2^n$ the binary n-tuple (row vector) corresponding to i, i.e.,

$$\mathbf{i} = [i_{n-1}, i_{n-2}, \ldots, i_1, i_0] .$$

The transpose of $\mathbf{i} \in \mathbb{Z}_2^n$ is denoted by $\mathbf{i}^T$. For $\mathbf{i}, \mathbf{j} \in \mathbb{Z}_2^n$, $\mathbf{i}^T\mathbf{j}$ is a binary matrix, while $\mathbf{i}\mathbf{j}^T = \sum_{l=0}^{n-1} i_l j_l \in \{0, 1\}$ is the binary inner product of the two vectors. We denote by $\mathbf{e}^{(m)}$ the binary vector corresponding to 2^m, i.e., $e_l^{(m)} = I_{\{m=l\}}$, where I is the indicator function. The all-zero vector is denoted by $\mathbf{0}$, and the all-one vector by $\mathbf{1}$.

Definition 3.1 An *index assignment* is a mapping $\pi\colon \mathbb{Z}_2^n \to \mathbb{Z}_2^n$ which is a bijection. An index assignment is a permutation of $\mathbb{Z}_2^n$, and thus there are $(2^n)!$ different index assignments.
An *affine index assignment* $\pi\colon \mathbb{Z}_2^n \to \mathbb{Z}_2^n$ is an index assignment of the form

$$\pi(\mathbf{i}) = \mathbf{i}\mathsf{G} + \mathbf{t}, \qquad \pi^{-1}(\mathbf{i}) = (\mathbf{i} + \mathbf{t})\mathsf{G}^{-1},$$

where G is a binary nonsingular $n \times n$ *generator matrix*, $\mathbf{t}$ is an n-dimensional binary *translation vector*, and the arithmetic is performed in $\mathbb{Z}_2^n$. If $\mathbf{t} = \mathbf{0}$, then π is called *linear.*

The family of affine index assignments is attractive due to its low implementation complexity, and was first systematically studied in [12, 14–16]. An unstructured index assignment requires a table of size $O(n2^n)$ bits to implement, whereas affine assignments can be described by $O(n^2)$ bits. Many useful index assignments are known to be affine, including the Natural Binary Code, Folded Binary Code, Gray Code, and Two's Complement Code [12].

Definition 3.2 The *Natural Binary Code* π_N is the identity index assignment $\pi_N(\mathbf{i}) = \mathbf{i}$. It is a linear index assignment with generator matrix $\mathsf{G}_N = \mathsf{I}$, the identity matrix.

We define the *Worst Code* π_W to be the linear index assignment with generator matrix

$$\mathsf{G}_W = \left[\begin{array}{c|cccc} n & 1 & 1 & \cdots & 1 \\ \hline 1 & 1 & 0 & \cdots & 0 \\ 1 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 1 & 0 & \cdots & 0 & 1 \end{array}\right],$$

where the "n" in the top-left component of G_W is taken modulo 2. The inverse of the generator matrix is

$$\mathsf{G}_W^{-1} = \left[\begin{array}{c|cccc} 1 & 1 & 1 & \cdots & 1 \\ \hline 1 & 0 & 1 & \cdots & 1 \\ 1 & 1 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{array}\right].$$

Table 3.1 gives an explicit listing (in both decimal and binary) of these two index assignments for $n = 4$.

Let the channel transition probabilities of a binary symmetric channel be denoted by (for $\epsilon < 1/2$)

$$p(a|b) = \begin{cases} (1-\epsilon) & \text{if } a = b \\ \epsilon & \text{if } a \neq b \end{cases} \qquad a, b \in \{0, 1\}.$$

Definition 3.3 The *Hamming weight* of a binary n-tuple $\mathbf{a} \in \mathbb{Z}_2^n$ is the number of its nonzero components,

$$w(\mathbf{a}) \triangleq \sum_{l=0}^{n-1} I_{\{a_l \neq 0\}}.$$

Table 3.1 A 4-bit example of the Natural Binary Code and the Worst Code.

i	$\pi_N(\mathbf{i})$		$\pi_W(\mathbf{i})$	
0	0	0000	0	0000
1	1	0001	9	1001
2	2	0010	10	1010
3	3	0011	3	0011
4	4	0100	12	1100
5	5	0101	5	0101
6	6	0110	6	0110
7	7	0111	15	1111
8	8	1000	7	0111
9	9	1001	14	1110
10	10	1010	13	1101
11	11	1011	4	0100
12	12	1100	11	1011
13	13	1101	2	0010
14	14	1110	1	0001
15	15	1111	8	1000

The *transition probabilities* for binary n-tuples on a binary symmetric channel are

$$P(\mathbf{a}|\mathbf{b}) = \prod_{l=0}^{n-1} p(a_l|b_l) = \epsilon^{w(\mathbf{a}+\mathbf{b})}(1-\epsilon)^{n-w(\mathbf{a}+\mathbf{b})}, \qquad \mathbf{a}, \mathbf{b} \in \mathbb{Z}_2^n.$$

We denote the probability that an error pattern $\mathbf{a} \in \mathbb{Z}_2^n$ occurs on a binary symmetric channel by

$$\rho_{\mathbf{a}} \triangleq P(\mathbf{b}+\mathbf{a}|\mathbf{b}) = \epsilon^{w(\mathbf{a})}(1-\epsilon)^{n-w(\mathbf{a})}, \qquad \mathbf{b} \in \mathbb{Z}_2^n. \tag{3.1}$$

Definition 3.4 Let π be an index assignment and suppose an element i is chosen uniformly at random from the set $\mathcal{S} = \{0, \ldots, 2^n - 1\}$, where the binary n-tuple $\pi(\mathbf{i})$ is transmitted over a binary symmetric channel with error probability ϵ. The end-to-end *mean squared error* is defined as

$$D \triangleq 2^{-n} \sum_{i \in \mathcal{S}} \sum_{j \in \mathcal{S}} (i-j)^2 \rho_{\pi(\mathbf{i})+\pi(\mathbf{j})}. \tag{3.2}$$

It may be assumed without loss of generality that $\pi(\mathbf{0}) = \mathbf{0}$, which for an affine index assignment $\pi(\mathbf{i}) = \mathbf{i}\mathsf{G} + \mathbf{t}$, is equivalent to setting $\mathbf{t} = \mathbf{0}$. Thus, we omit the translation vector $\mathbf{t}$ in what follows.

Definition 3.5 For each $\mathbf{i}, \mathbf{j} \in \mathbb{Z}_2^n$, let $h_{\mathbf{i},\mathbf{j}} = (-1)^{\mathbf{i}\mathbf{j}^T}$. The *Hadamard transform* $\hat{f} \colon \mathbb{Z}_2^n \to \mathbb{R}$ of a mapping $f \colon \mathbb{Z}_2^n \to \mathbb{R}$ is defined by

$$\hat{f}(\mathbf{j}) = \sum_{\mathbf{i} \in \mathbb{Z}_2^n} f(\mathbf{i}) h_{\mathbf{i},\mathbf{j}},$$

and the inverse transform is given by

$$f(\mathbf{i}) = 2^{-n} \sum_{\mathbf{j} \in \mathbb{Z}_2^n} \hat{f}(\mathbf{j}) h_{\mathbf{j},\mathbf{i}}.$$

The Hadamard transform provides a tool for analyzing the mean squared distortion [5, 12, 16–19]. The following properties of Hadamard transforms were given in Chapter 2, and are repeated here for convenience. For any $\mathbf{i}, \mathbf{j}, \mathbf{a}, \mathbf{b} \in \mathbb{Z}_2^n$,

(i) $$h_{\mathbf{i},\mathbf{j}} = h_{\mathbf{j},\mathbf{i}}$$

(ii) $$h_{\mathbf{i},\mathbf{a}+\mathbf{b}} = h_{\mathbf{i},\mathbf{a}} h_{\mathbf{i},\mathbf{b}}$$

(iii) $$\sum_{\mathbf{i} \in \mathbb{Z}_2^n} h_{\mathbf{i},\mathbf{j}} = \begin{cases} 2^n & \text{if } \mathbf{j} = \mathbf{0} \\ 0 & \text{otherwise} \end{cases}$$

(iv) $$\sum_{\mathbf{i} \in \mathbb{Z}_2^n} i_m h_{\mathbf{i},\mathbf{j}} = \begin{cases} 2^{n-1} & \text{if } \mathbf{j} = \mathbf{0} \\ -2^{n-1} & \text{if } \mathbf{j} = \mathbf{e}^{(m)} \\ 0 & \text{otherwise} \end{cases} \qquad m \in \{0, 1, \dots, n-1\}$$

The first two properties are straightforward. Property (iii) follows from the fact that exactly half of the binary vectors in $\mathbb{Z}_2^n$ are orthogonal to any fixed nonzero vector $\mathbf{j} \in \mathbb{Z}_2^n$. To see Property (iv), let $\mathbf{i}', \mathbf{j}' \in \mathbb{Z}_2^{n-1}$ respectively denote the binary vectors $\mathbf{i}, \mathbf{j} \in \mathbb{Z}_2^n$ but with the m^{th} component removed. Then we can rewrite Property (iv) as

$$\sum_{\mathbf{i}' \in \mathbb{Z}_2^{n-1}} \sum_{i_m \in \{0,1\}} i_m (-1)^{i_m j_m} h_{\mathbf{i}',\mathbf{j}'} = (-1)^{j_m} \sum_{\mathbf{i}' \in \mathbb{Z}_2^{n-1}} h_{\mathbf{i}',\mathbf{j}'} = (-1)^{j_m} 2^{n-1} I_{\{\mathbf{j}'=\mathbf{0}\}},$$

where the last equality follows from Property (iii) for $\mathbb{Z}_2^{n-1}$.

Lemma 3.1 *The Hadamard transform $\hat{\rho}$ of the error pattern distribution ρ is*

$$\hat{\rho}_{\mathbf{j}} = (1-2\epsilon)^{w(\mathbf{j})}.$$

Proof

$$\begin{aligned}
\hat{\rho}_{\mathbf{j}} &= \sum_{\mathbf{i}\in\mathbb{Z}_2^n} \epsilon^{w(\mathbf{i})}(1-\epsilon)^{n-w(\mathbf{i})} h_{\mathbf{i},\mathbf{j}} \\
&= \sum_{i_0\in\{0,1\}} \sum_{i_1\in\{0,1\}} \cdots \sum_{i_{n-1}\in\{0,1\}} \left(\prod_{l=0}^{n-1} \epsilon^{i_l}(1-\epsilon)^{1-i_l}(-1)^{i_l j_l} \right) \\
&= \prod_{l=0}^{n-1} \sum_{i_l\in\{0,1\}} \epsilon^{i_l}(1-\epsilon)^{1-i_l}(-1)^{i_l j_l} \\
&= \prod_{l=0}^{n-1} \left(1-\epsilon+\epsilon(-1)^{j_l}\right) \\
&= (1-2\epsilon)^{w(\mathbf{j})}
\end{aligned}$$

■

3.3 Construction of the Worst Code

The following lemma gives an expression for the distortion D in the Hadamard transform domain. Variants of this result were used to show the optimality of the Natural Binary Code in [10, 11]. The lemma is useful for identifying a "worst" code.

Lemma 3.2 *Let $\eta(\mathbf{i}) = \pi^{-1}(\mathbf{i})$ for all $\mathbf{i} \in \mathbb{Z}_2^n$. Then the distortion in the Hadamard transform domain is*

$$D = 2 \sum_{\mathbf{a}\in\mathbb{Z}_2^n\setminus\{\mathbf{0}\}} \left[2^{-n}\hat{\eta}(\mathbf{a})\right]^2 \left(1-(1-2\epsilon)^{w(\mathbf{a})}\right).$$

Proof

Rewriting (3.2) using η yields

$$
\begin{aligned}
D &= 2^{-n} \sum_{\mathbf{i}\in\mathbb{Z}_2^n} \sum_{\mathbf{j}\in\mathbb{Z}_2^n} \left(\eta(\mathbf{i}) - \eta(\mathbf{j})\right)^2 \rho_{\mathbf{i}+\mathbf{j}} \\
&= 2^{-n} \sum_{\mathbf{i}\in\mathbb{Z}_2^n} \sum_{\mathbf{j}\in\mathbb{Z}_2^n} \left[2^{-n} \sum_{\mathbf{a}\in\mathbb{Z}_2^n} \hat{\eta}(\mathbf{a}) \left(h_{\mathbf{a},\mathbf{i}} - h_{\mathbf{a},\mathbf{j}}\right) \right]^2 \rho_{\mathbf{i}+\mathbf{j}} \\
&= 8^{-n} \sum_{\mathbf{i}\in\mathbb{Z}_2^n} \sum_{\mathbf{j}\in\mathbb{Z}_2^n} \rho_{\mathbf{i}+\mathbf{j}} \sum_{\mathbf{a}\in\mathbb{Z}_2^n} \sum_{\mathbf{b}\in\mathbb{Z}_2^n} \hat{\eta}(\mathbf{a})\hat{\eta}(\mathbf{b}) \left(h_{\mathbf{a}+\mathbf{b},\mathbf{i}} - h_{\mathbf{a},\mathbf{i}}h_{\mathbf{b},\mathbf{j}} - h_{\mathbf{a},\mathbf{j}}h_{\mathbf{b},\mathbf{i}} + h_{\mathbf{a}+\mathbf{b},\mathbf{j}}\right) \\
&= 8^{-n} \sum_{\mathbf{a}\in\mathbb{Z}_2^n} \sum_{\mathbf{b}\in\mathbb{Z}_2^n} \hat{\eta}(\mathbf{a})\hat{\eta}(\mathbf{b}) \sum_{\mathbf{i}\in\mathbb{Z}_2^n} h_{\mathbf{a}+\mathbf{b},\mathbf{i}} \sum_{\mathbf{j}\in\mathbb{Z}_2^n} \rho_{\mathbf{i}+\mathbf{j}} \left(h_{\mathbf{0},\mathbf{i}+\mathbf{j}} - h_{\mathbf{a},\mathbf{i}+\mathbf{j}} - h_{\mathbf{b},\mathbf{i}+\mathbf{j}} + h_{\mathbf{a}+\mathbf{b},\mathbf{i}+\mathbf{j}}\right) \\
&= 4^{-n} \sum_{\mathbf{a}\in\mathbb{Z}_2^n} \sum_{\mathbf{b}\in\mathbb{Z}_2^n} \hat{\eta}(\mathbf{a})\hat{\eta}(\mathbf{b}) \left(2^{-n} \sum_{\mathbf{i}\in\mathbb{Z}_2^n} h_{\mathbf{a}+\mathbf{b},\mathbf{i}} \right) \left(\sum_{\mathbf{c}\in\mathbb{Z}_2^n} \rho_{\mathbf{c}} \left(h_{\mathbf{0},\mathbf{c}} - h_{\mathbf{a},\mathbf{c}} - h_{\mathbf{b},\mathbf{c}} + h_{\mathbf{a}+\mathbf{b},\mathbf{c}}\right) \right) \\
&= 4^{-n} \sum_{\mathbf{a}\in\mathbb{Z}_2^n} \sum_{\mathbf{b}\in\mathbb{Z}_2^n} \hat{\eta}(\mathbf{a})\hat{\eta}(\mathbf{b}) I_{\{\mathbf{a}=\mathbf{b}\}} \left(\hat{\rho}_{\mathbf{0}} - \hat{\rho}_{\mathbf{a}} - \hat{\rho}_{\mathbf{b}} + \hat{\rho}_{\mathbf{a}+\mathbf{b}}\right) \\
&= 2 \sum_{\mathbf{a}\in\mathbb{Z}_2^n} \left[2^{-n}\hat{\eta}(\mathbf{a})\right]^2 \left(\hat{\rho}_{\mathbf{0}} - \hat{\rho}_{\mathbf{a}}\right) \\
&= 2 \sum_{\mathbf{a}\in\mathbb{Z}_2^n\setminus\{\mathbf{0}\}} \left[2^{-n}\hat{\eta}(\mathbf{a})\right]^2 \left(1 - (1-2\epsilon)^{w(\mathbf{a})}\right).
\end{aligned}
$$

■

The following bounds on D follow from Lemma 3.2 using $1 \le w(\mathbf{a}) \le n$ for $\mathbf{a} \in \mathbb{Z}_2^n \setminus \{\mathbf{0}\}$:

$$
4\epsilon \sum_{\mathbf{a}\in\mathbb{Z}_2^n\setminus\{\mathbf{0}\}} \left[2^{-n}\hat{\eta}(\mathbf{a})\right]^2 \le D \le 2\left(1 - (1-2\epsilon)^n\right) \sum_{\mathbf{a}\in\mathbb{Z}_2^n\setminus\{\mathbf{0}\}} \left[2^{-n}\hat{\eta}(\mathbf{a})\right]^2. \tag{3.3}
$$

The lower bound was established in [10] and can be achieved with equality if $\hat{\eta}(\mathbf{a}) = 0$ for every $\mathbf{a} \in \mathbb{Z}_2^n$ with Hamming weight $w(\mathbf{a}) > 1$. For example, the Natural Binary Code satisfies this requirement [10, 11]. To achieve the upper bound with equality, we must have $\hat{\eta}(\mathbf{a}) = 0$ for every $\mathbf{a} \in \mathbb{Z}_2^n$ such that $w(\mathbf{a}) < n$, i.e., $\hat{\eta}(\mathbf{1})$ must be the only nonzero

Hadamard transform component. Note, however, that for any index assignment π,

$$
\begin{aligned}
\sum_{\mathbf{a}\in\mathbb{Z}_2^n\setminus\{\mathbf{0}\}} \left[2^{-n}\hat{\eta}(\mathbf{a})\right]^2 &= \left(\sum_{\mathbf{a}\in\mathbb{Z}_2^n}\left[2^{-n}\sum_{\mathbf{i}\in\mathbb{Z}_2^n}\eta(\mathbf{i})h_{\mathbf{i},\mathbf{a}}\right]^2\right) - \left[2^{-n}\sum_{\mathbf{i}\in\mathbb{Z}_2^n}\eta(\mathbf{i})h_{\mathbf{i},\mathbf{0}}\right]^2 \\
&= 2^{-n}\sum_{i\in\mathcal{S}}\sum_{j\in\mathcal{S}} ij\left(2^{-n}\sum_{\mathbf{a}\in\mathbb{Z}_2^n} h_{\pi(\mathbf{i})+\pi(\mathbf{j}),\mathbf{a}}\right) - \left[2^{-n}\sum_{i\in\mathcal{S}} i\right]^2 \\
&= 2^{-n}\sum_{i\in\mathcal{S}} i^2 - \left[2^{-n}\sum_{i\in\mathcal{S}} i\right]^2 \qquad (3.4)\\
&= \frac{4^n-1}{12} \qquad (3.5)\\
&= \sigma_{\mathcal{S}}^2,
\end{aligned}
$$

the variance of a random variable chosen uniformly at random from $\mathcal{S}$. On the other hand, for any π,

$$
\begin{aligned}
\left[2^{-n}\hat{\eta}(\mathbf{1})\right]^2 &= \left[2^{-n}\sum_{\mathbf{i}\in\mathbb{Z}_2^n}\eta(\mathbf{i})h_{\mathbf{i},\mathbf{1}}\right]^2 \\
&= \left[2^{-n}\sum_{i\in\mathcal{S}} i(-1)^{\pi(\mathbf{i})\mathbf{1}^T}\right]^2 \\
&\le \left[2^{-n}\left(\sum_{i=2^{n-1}}^{2^n-1} i - \sum_{i=0}^{2^{n-1}-1} i\right)\right]^2 \\
&= \left[2^{-n}4^{n-1}\right]^2 \\
&= 4^{n-2} \qquad (3.6)\\
&< \frac{4^n-1}{12},
\end{aligned}
$$

for every $n > 1$. Thus, the upper bound given in (3.3) is not achievable when $n > 1$. The next tightest upper bound from Lemma 3.2 is obtained using $w(\mathbf{a}) \le n-1$ for all $\mathbf{a} \ne \mathbf{1}$, i.e.,

$$
\begin{aligned}
D \le\ & 2\left(1-(1-2\epsilon)^n\right)\max_{\eta_0}\left[2^{-n}\hat{\eta}_0(\mathbf{1})\right]^2 \\
&+ 2\left(1-(1-2\epsilon)^{n-1}\right)\left(\sum_{\mathbf{a}\in\mathbb{Z}_2^n\setminus\{\mathbf{0}\}}\left[2^{-n}\hat{\eta}(\mathbf{a})\right]^2 - \max_{\eta_0}\left[2^{-n}\hat{\eta}_0(\mathbf{1})\right]^2\right). \qquad (3.7)
\end{aligned}
$$

Indeed, the Worst Code achieves (3.7). To prove this, consider the Hadamard transform components $\hat{\eta}(\mathbf{a})$ of an arbitrary linear index assignment $\pi(\mathbf{i}) = \mathbf{i}\mathsf{G}$, for any $\mathbf{a} \in \mathbb{Z}_2^n \setminus \{\mathbf{0}\}$:

$$\begin{aligned}
\hat{\eta}(\mathbf{a}) &= \sum_{\mathbf{i}\in\mathbb{Z}_2^n} \eta(\mathbf{i}) h_{\mathbf{i},\mathbf{a}} \\
&= \sum_{i\in\mathcal{S}} i h_{\pi(\mathbf{i}),\mathbf{a}} \\
&= \sum_{i\in\mathcal{S}} \left(\sum_{l=0}^{n-1} 2^l i_l \right) h_{\mathbf{i}\mathsf{G},\mathbf{a}} \\
&= \sum_{l=0}^{n-1} 2^l \left(\sum_{\mathbf{i}\in\mathbb{Z}_2^n} i_l h_{\mathbf{i},\mathbf{a}\mathsf{G}^T} \right) \\
&= -2^{n-1} \sum_{l=0}^{n-1} 2^l I_{\{\mathbf{a}\mathsf{G}^T = \mathbf{e}^{(l)}\}},
\end{aligned} \tag{3.8}$$

where (3.8) follows from Property (iv) since $\mathbf{a}\mathsf{G}^T \neq \mathbf{0}$ for $\mathbf{a} \neq \mathbf{0}$ by the nonsingularity of G. Therefore, the only nonzero Hadamard transform components are those corresponding to $\mathbf{a} = \mathbf{e}^{(l)} \left(\mathsf{G}^T\right)^{-1}$, for $l = 0, \ldots, n-1$. Thus, to achieve the lower bound given in (3.3) every row of $\left(\mathsf{G}^T\right)^{-1}$ must have Hamming weight 1, as with the Natural Binary Code.

Similarly, the upper bound given in (3.7) can be achieved by setting $\mathbf{e}^{(n-1)} \left(\mathsf{G}^T\right)^{-1} = \mathbf{1}$ (i.e., the first row of the inverse of the transposed generator matrix must be all ones), and choosing the remaining $n-1$ rows of $\left(\mathsf{G}^T\right)^{-1}$ to have Hamming weight $n-1$. An example is G_W^{-1}, given in the definition of the Worst Code. Note that the all-one vector has to be the first row of $\left(\mathsf{G}^T\right)^{-1}$ to ensure the maximization of

$$\left[2^{-n}\hat{\eta}(\mathbf{1})\right]^2 = \left[2^{-n}\left(-2^{n-1}2^{n-1}\right)\right]^2 = 4^{n-2}.$$

Thus, combining the lower bound from (3.3) and the upper bound from (3.7), and using (3.5) and (3.6) to eliminate the remaining Hadamard transforms from the expressions, we obtain the following theorem.

Theorem 3.1 *Suppose an integer* $\mathbf{i}$ *is chosen uniformly at random from* $\mathcal{S} = \{0, \ldots, 2^n - 1\}$ *and the* n*-bit word* $\pi(\mathbf{i})$ *is transmitted over a binary symmetric channel with bit error probability* $\epsilon \in [0, 1/2]$*, using an index assignment* π*. Then the resulting mean squared* *error D satisfies*

$$\epsilon \, \frac{4^n - 1}{3} \leq D \leq \epsilon \left(1 - 2\epsilon\right)^{n-1} 4^{n-1} + \left(1 - \left(1 - 2\epsilon\right)^{n-1}\right) \frac{4^n - 1}{6},$$

where the lower bound is achieved by the Natural Binary Code and the upper bound by the Worst Code.

Let us denote the distortion of the Natural Binary Code and the Worst Code respectively by $D_{\text{min}} = \epsilon \, \frac{4^n-1}{3}$ and $D_{\text{max}} = \epsilon \left(1 - 2\epsilon\right)^{n-1} 4^{n-1} + \left(1 - \left(1 - 2\epsilon\right)^{n-1}\right) \frac{4^n-1}{6}$. If an index assignment is chosen uniformly at random, then the average distortion is

$$D_{\text{ave}} \triangleq \frac{1}{(2^n)!} \sum_{\pi} 2^{-n} \sum_{i \in \mathcal{S}} \sum_{j \in \mathcal{S}} (i - j)^2 \rho_{\pi(\mathbf{i}) + \pi(\mathbf{j})}. \tag{3.9}$$

Since $\rho_{\mathbf{0}} = (1 - \epsilon)^n$, and

$$\begin{aligned}
\frac{1}{(2^n)!} \sum_{\pi} \rho_{\pi(\mathbf{i}) + \pi(\mathbf{j})} &= \sum_{\mathbf{a} \in \mathbb{Z}_2^n} \sum_{\mathbf{b} \in \mathbb{Z}_2^n} \rho_{\mathbf{a}+\mathbf{b}} \left(\frac{1}{(2^n)!} \sum_{\pi} I_{\{\pi(\mathbf{i}) = \mathbf{a}, \pi(\mathbf{j}) = \mathbf{b}\}} \right) \\
&= \sum_{\mathbf{a} \in \mathbb{Z}_2^n} \sum_{\mathbf{b} \in \mathbb{Z}_2^n} \rho_{\mathbf{a}+\mathbf{b}} \left(I_{\{\mathbf{i}=\mathbf{j}, \mathbf{a}=\mathbf{b}\}} \frac{(2^n - 1)!}{(2^n)!} + I_{\{\mathbf{i} \neq \mathbf{j}, \mathbf{a} \neq \mathbf{b}\}} \frac{(2^n - 2)!}{(2^n)!} \right) \\
&= I_{\{\mathbf{i}=\mathbf{j}\}} \rho_{\mathbf{0}} + I_{\{\mathbf{i} \neq \mathbf{j}\}} \frac{1 - \rho_{\mathbf{0}}}{2^n - 1},
\end{aligned}$$

this gives

$$D_{\text{ave}} = \frac{1 - (1 - \epsilon)^n}{2^n (2^n - 1)} \sum_{i \in \mathcal{S}} \sum_{j \in \mathcal{S}} (i - j)^2 = \left(1 - (1 - \epsilon)^n\right) \frac{4^n + 2^n}{6}.$$

The values of D_{min} and D_{ave} were apparently first reported in [7].

Clearly, the inequalities $D_{\text{min}} \leq D_{\text{ave}} \leq D_{\text{max}}$ hold for every $\epsilon \in [0, 1/2]$ and $n \geq 1$. It is interesting to examine the asymptotic behavior of the minimum, maximum, and average distortions, both as the blocklength n grows and as the channel error probability ϵ decreases. The partial derivatives of $\frac{D_{\text{max}}}{D_{\text{min}}}$, $\frac{D_{\text{ave}}}{D_{\text{min}}}$, and $\frac{D_{\text{max}}}{D_{\text{ave}}}$, with respect to ϵ are all strictly negative for all $\epsilon \in (0, 1/2)$ and for all $n > 1$. Hence, the largest performance gain of a best index assignment over a worst index assignment or over an average index assignment occurs in the limit as $\epsilon \to 0$. Asymptotically as $\epsilon \to 0$, for a fixed blocklength n, these gains are given by

$$\begin{aligned}
\lim_{\epsilon\to 0}\frac{D_{\text{max}}}{D_{\text{min}}} &= n-1+\left(\frac{3}{4}\right)\frac{1}{1-4^{-n}},\\
\lim_{\epsilon\to 0}\frac{D_{\text{ave}}}{D_{\text{min}}} &= \frac{n}{2(1-2^{-n})},\\
\lim_{\epsilon\to 0}\frac{D_{\text{max}}}{D_{\text{ave}}} &= 2(1-n^{-1})(1-2^{-n})+\left(\frac{3}{2}\right)\frac{1}{n(1+2^{-n})}.
\end{aligned}$$

On the other hand, for a fixed bit error probability ϵ, letting $n\to\infty$ yields

$$\begin{aligned}
\lim_{n\to\infty}\frac{D_{\text{max}}}{D_{\text{min}}} &= \frac{1}{2\epsilon},\\
\lim_{n\to\infty}\frac{D_{\text{ave}}}{D_{\text{min}}} &= \frac{1}{2\epsilon},\\
\lim_{n\to\infty}\frac{D_{\text{max}}}{D_{\text{ave}}} &= 1.
\end{aligned}$$

Thus for asymptotically large block lengths, the performance gain of a best index assignment over a worst index assignment or an average index assignment is $1/2\epsilon$, which can be very large. In this sense, a large fraction of index assignments can be considered "bad."

Corollary 3.1 *For any fixed large n, as $\epsilon\to 0$ the relative mean squared errors of worst, average, and best index assignments for the uniform source obey the following ratios:*

$$D_{\text{max}} : D_{\text{ave}} : D_{\text{min}} \approx 1 : 1/2 : 1/n,$$

and for any fixed ϵ, as $n\to\infty$ the relative mean squared errors obey the ratios:

$$D_{\text{max}} : D_{\text{ave}} : D_{\text{min}} = 1 : 1 : 2\epsilon.$$

That is, for any $\epsilon > 0$, the expected distortion of a randomly chosen index assignment asymptotically equals (as the blocklength grows) that of the worst index assignment. If an integer chosen uniformly at random from $\mathcal{S}$ is normalized to have zero mean and unit variance then the resulting distortions corresponding to the best, worst, and random index assignments are given by

$$\begin{aligned}
\tilde{D}_{\text{min}} &\triangleq \frac{D_{\text{min}}}{\sigma_{\mathcal{S}}^2} = 4\epsilon\\
\tilde{D}_{\text{max}} &\triangleq \frac{D_{\text{max}}}{\sigma_{\mathcal{S}}^2} = 2\left(1-(1-2\epsilon)^{n-1}\right)+3\epsilon\frac{(1-2\epsilon)^{n-1}}{1-4^{-n}}\\
\tilde{D}_{\text{ave}} &\triangleq \frac{D_{\text{ave}}}{\sigma_{\mathcal{S}}^2} = 2\frac{1-(1-\epsilon)^n}{1-2^{-n}}.
\end{aligned}$$

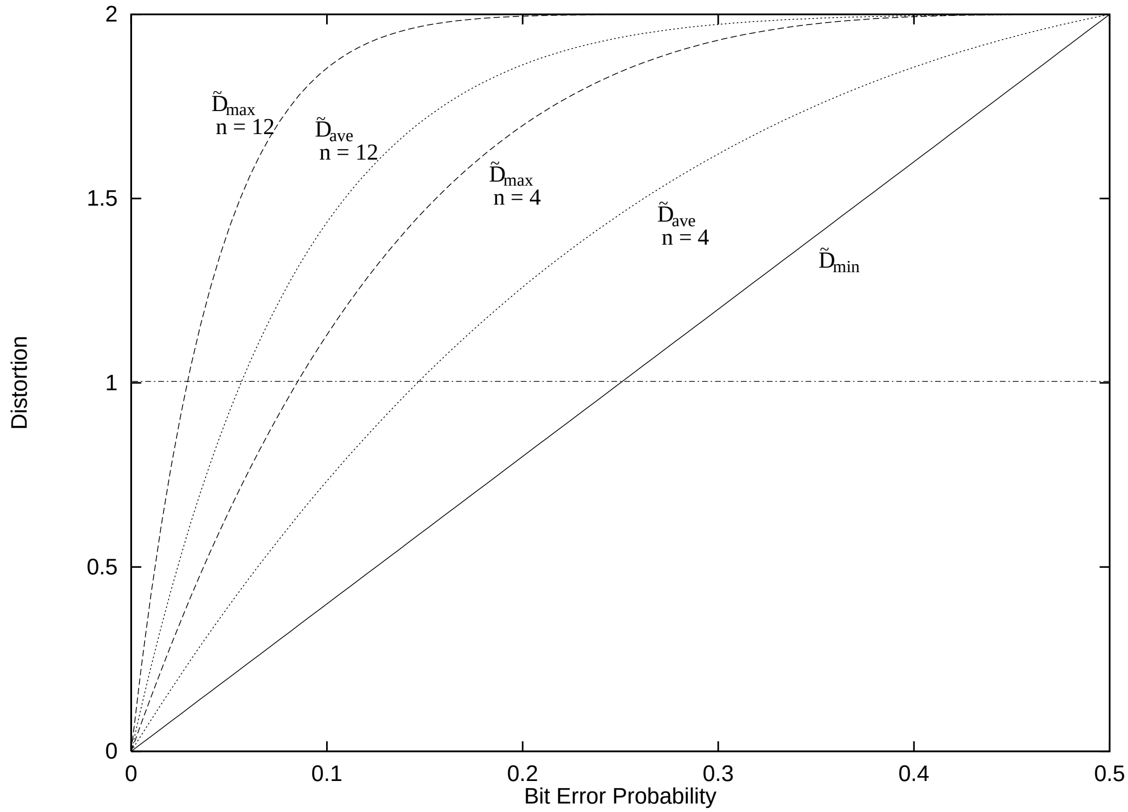


Figure 3.1 The best, worst, and average performance achievable by index assignments for a uniform source. The solid line corresponds to $\tilde{D}_{\text{min}}$. The dashed and dotted curves show $\tilde{D}_{\text{max}}$ and $\tilde{D}_{\text{ave}}$, respectively, for $n = 4, 12$. The horizontal line represents the variance of the source, an achievable distortion at zero transmission rate.

Figure 3.1 compares $\tilde{D}_{\text{min}}$, $\tilde{D}_{\text{max}}$, and $\tilde{D}_{\text{ave}}$. The horizontal line at normalized distortion 1.0 represents the distortion achievable with no information transmission (by simply reproducing the mean of the source at the receiver). Thus, the usefulness of any index assignment is limited to values of ϵ smaller than the bit error probability determined by the intersection of this horizontal line and the distortion curve corresponding to the index assignment. Since $\lim_{n\to\infty} \tilde{D}_{\text{max}} = \lim_{n\to\infty} \tilde{D}_{\text{ave}} = 2$ for any $\epsilon \in (0, 1/2]$, the useful region of bit error probabilities for the worst and average index assignments shrinks steadily as the blocklength increases. If $n\epsilon << 1$, then we obtain the approximations (linear in ϵ)

$$\tilde{D}_{\text{max}} \approx \left(4(n-1) + \frac{3}{1-4^{-n}}\right)\epsilon \qquad \text{and} \qquad \tilde{D}_{\text{ave}} \approx \left(\frac{2n}{1-2^{-n}}\right)\epsilon.$$

These hold for small ϵ on the curves in Figure 3.1 for which n is not too large. Suppose these linearized approximations hold, and suppose that n is large but not too large (i.e., if $2^{-n} << 1$ while maintaining $n\epsilon << 1$). Then the useful regions of the worst and average index assignments can be approximated as $\epsilon \in (0, 1/(4n))$ and $\epsilon \in (0, 1/(2n))$, respectively. These intervals are obtained by examing which values of ϵ yield distortions less than 1. Note that $\tilde{D}_{\min} = 4\epsilon$ is independent of n and linear on the full range $\epsilon \in [0, 1/2]$. Thus, the useful region of the best index assignment is $(0, 1/4)$ irrespective of the blocklength n.

3.4 Generalization to Vector Quantizers

The Natural Binary Code was shown to minimize the distortion D for a uniform scalar quantizer and a uniform source in [10] and was generalized to a class of vector quantizers in [11]. The class of vector quantizers in [11] is the same class studied in [12, 17–19] and was referred to in Chapter 2 as "binary lattice vector quantization." In contrast, we demonstrate by means of a counterexample that the distortion maximization property of the Worst Code for a uniform scalar quantizer cannot be generalized to arbitrary BLVQs. We do, however, show that the Worst Code maximizes the distortion among all *affine* index assignments for arbitrary BLVQs.

For any positive integer d, let $\mathbb{R}^d$ denote d-dimensional Euclidean space. We use a horizontal bar to distinguish between real vectors $\bar{\mathbf{x}} \in \mathbb{R}^d$ and binary vectors $\mathbf{i} \in \mathbb{Z}_2^n$. The Euclidean norm of a vector $\bar{\mathbf{x}} \in \mathbb{R}^d$ is denoted by $\|\bar{\mathbf{x}}\|$. As in Chapter 2, we define a BLVQ as follows.

Definition 3.6 A d-dimensional, 2^n-point *binary lattice vector quantizer* is a vector quantizer with codevectors of the form $\bar{\mathbf{y}}_i = \bar{\mathbf{y}}_0 + \sum_{l=0}^{n-1} \bar{\mathbf{v}}_l i_l$ for $i \in \mathcal{S}$, where $\bar{\mathbf{y}}_0 \in \mathbb{R}^d$, and $\mathcal{V} = \{\bar{\mathbf{v}}_l\}_{l=0}^{n-1} \subset \mathbb{R}^d$ is a *generating set*, ordered by $\|\bar{\mathbf{v}}_0\| \leq \|\bar{\mathbf{v}}_1\| \leq \ldots \leq \|\bar{\mathbf{v}}_{n-1}\|$.

Analogous to (3.2), the *channel distortion* of a BLVQ with equiprobable codevectors is defined as

$$D \stackrel{\triangle}{=} 2^{-n} \sum_{i \in \mathcal{S}} \sum_{j \in \mathcal{S}} \|\bar{\mathbf{y}}_i - \bar{\mathbf{y}}_j\|^2 \rho_{\pi(\mathbf{i})+\pi(\mathbf{j})}. \tag{3.10}$$

A uniform scalar quantizer with step size Δ is a special case of a BLVQ with $d = 1$ and $\bar{\mathbf{v}}_l = 2^l \Delta$ for $l \in \{0, 1, \ldots, n-1\}$.

The results of Section 3.3 also apply for binary lattice quantizers if we replace $\eta(\mathbf{i})$ by $\bar{\mathbf{z}}(\mathbf{i}) = \bar{\mathbf{y}}_{\pi^{-1}(i)}$. In particular, Lemma 3.2 becomes

$$D = 2 \sum_{\mathbf{a} \in \mathbb{Z}_2^n \setminus \{\mathbf{0}\}} \left\|2^{-n} \hat{\bar{\mathbf{z}}}(\mathbf{a})\right\|^2 \left(1 - (1-2\epsilon)^{w(\mathbf{a})}\right), \tag{3.11}$$

and thus (3.3) becomes

$$4\epsilon \sum_{\mathbf{a} \in \mathbb{Z}_2^n \setminus \{\mathbf{0}\}} \left\|2^{-n} \hat{\bar{\mathbf{z}}}(\mathbf{a})\right\|^2 \le D \le 2\left(1 - (1-2\epsilon)^n\right) \sum_{\mathbf{a} \in \mathbb{Z}_2^n \setminus \{\mathbf{0}\}} \left\|2^{-n} \hat{\bar{\mathbf{z}}}(\mathbf{a})\right\|^2. \tag{3.12}$$

We also have by (3.4) that

$$\sum_{\mathbf{a} \in \mathbb{Z}_2^n \setminus \{\mathbf{0}\}} \left\|2^{-n} \hat{\bar{\mathbf{z}}}(\mathbf{a})\right\|^2 = 2^{-n} \sum_{i \in \mathcal{S}} \|\bar{\mathbf{y}}_i\|^2 - \left\|2^{-n} \sum_{i \in \mathcal{S}} \bar{\mathbf{y}}_i\right\|^2 = \frac{1}{4} \sum_{l=0}^{n-1} \|\bar{\mathbf{v}}_l\|^2, \tag{3.13}$$

for any choice of index assignment π. It is difficult, however, to find $\max_\pi \left\|2^{-n} \hat{\bar{\mathbf{z}}}(\mathbf{1})\right\|^2$ for an arbitrary index assignment. For affine index assignments, we have by (3.8),

$$\hat{\bar{\mathbf{z}}}(\mathbf{a}) = -2^{n-1} \sum_{l=0}^{n-1} \bar{\mathbf{v}}_l I_{\{\mathbf{a}\mathbf{G}^T = \mathbf{e}^{(l)}\}},$$

and thus

$$\max_{\pi \text{ affine}} \left\|2^{-n} \hat{\bar{\mathbf{z}}}(\mathbf{1})\right\|^2 = \frac{1}{4} \|\bar{\mathbf{v}}_{n-1}\|^2. \tag{3.14}$$

Using (3.14) and (3.13), the same argument that led to the upper bound in Theorem 3.1 yields the following corollary.

Corollary 3.2 *The channel distortion of a binary lattice vector quantizer, with generators* $\bar{\mathbf{v}}_0, \ldots, \bar{\mathbf{v}}_{n-1}$*, followed by an affine index assignment and a binary symmetric channel with bit error probability* $\epsilon \in [0, 1/2]$ *satisfies*

$$D \leq \frac{1}{2}\left(1 - (1-2\epsilon)^n\right) \|\bar{\mathbf{v}}_{n-1}\|^2 + \frac{1}{2}\left(1 - (1-2\epsilon)^{n-1}\right) \sum_{l=0}^{n-2} \|\bar{\mathbf{v}}_l\|^2,$$

and the Worst Code achieves the upper bound with equality.

Note that from (3.14) Corollary 3.2 can be generalized to all (i.e., affine and non-affine) index assignments if $\max_\pi \left\|2^{-n}\hat{\hat{\mathbf{z}}}(\mathbf{1})\right\|^2 = \frac{1}{4}\|\bar{\mathbf{v}}_{n-1}\|^2$. However, in general Corollary 3.2 cannot be generalized in this manner, as demonstrated in the following corollary.

Corollary 3.3 *The Worst Code does not maximize the mean squared error of an arbitrary binary lattice vector quantizer over all index assignments for a binary symmetric channel.*

Proof We show by means of a counterexample that in general there exists an index assignment π_X yielding a higher MSE than that of the Worst Code π_W. Specifically, define the non-affine 3-bit index assignment π_X by

$$\pi_X(\mathbf{i}) = \begin{cases} \pi_W(100) & \text{if } \mathbf{i} = 011, \\ \pi_W(011) & \text{if } \mathbf{i} = 100, \\ \pi_W(\mathbf{i}) & \text{otherwise.} \end{cases}$$

Table 3.2 explicitly lists the index assignments π_W and π_X, along with the Hadamard transform vectors $\hat{\hat{\mathbf{z}}}_W$ and $\hat{\hat{\mathbf{z}}}_X$. The identity

$$\|\bar{\mathbf{u}} + \bar{\mathbf{w}}\|^2 + \|\bar{\mathbf{u}} - \bar{\mathbf{w}}\|^2 = 2\|\bar{\mathbf{u}}\|^2 + 2\|\bar{\mathbf{w}}\|^2 \qquad \bar{\mathbf{u}}, \bar{\mathbf{w}} \in \mathbb{R}^d$$

will be used frequently in what follows. Also, to simplify notation we set $\gamma = 1 - 2\epsilon$. By (3.11) the channel distortion of the Worst Code is

$$\begin{aligned} D_W &= 2(1-\gamma^3)\|(-4/8)\bar{\mathbf{v}}_2\|^2 + 2(1-\gamma^2)\left(\|(-4/8)\bar{\mathbf{v}}_0\|^2 + \|(-4/8)\bar{\mathbf{v}}_1\|^2\right) \\ &= \frac{1}{2}\left[\left(\|\bar{\mathbf{v}}_2\|^2 + \|\bar{\mathbf{v}}_1\|^2 + \|\bar{\mathbf{v}}_0\|^2\right) - \gamma^2\left(\|\bar{\mathbf{v}}_1\|^2 + \|\bar{\mathbf{v}}_0\|^2\right) - \gamma^3\|\bar{\mathbf{v}}_2\|^2\right], \end{aligned}$$

Table 3.2 Hadamard transform vectors of the Worst Code and a non-affine index assignment which performs even worse.

$\mathbf{i}$	$\pi_W(\mathbf{i})$	$\hat{\hat{\mathbf{z}}}_W(\mathbf{i})$	$\pi_X(\mathbf{i})$	$\hat{\hat{\mathbf{z}}}_X(\mathbf{i})$
000	000	$8\bar{\mathbf{y}}_0 + 4(\bar{\mathbf{v}}_0 + \bar{\mathbf{v}}_1 + \bar{\mathbf{v}}_2)$	000	$8\bar{\mathbf{y}}_0 + 4(\bar{\mathbf{v}}_0 + \bar{\mathbf{v}}_1 + \bar{\mathbf{v}}_2)$
001	101	$\bar{\mathbf{0}}$	101	$\bar{\mathbf{0}}$
010	110	$\bar{\mathbf{0}}$	110	$\bar{\mathbf{0}}$
011	011	$\bar{\mathbf{0}}$	111	$\bar{\mathbf{0}}$
100	111	$\bar{\mathbf{0}}$	011	$+2(\bar{\mathbf{v}}_2 - \bar{\mathbf{v}}_1 - \bar{\mathbf{v}}_0)$
101	010	$-4\bar{\mathbf{v}}_1$	010	$-2(\bar{\mathbf{v}}_2 + \bar{\mathbf{v}}_1 - \bar{\mathbf{v}}_0)$
110	001	$-4\bar{\mathbf{v}}_0$	001	$-2(\bar{\mathbf{v}}_2 - \bar{\mathbf{v}}_1 + \bar{\mathbf{v}}_0)$
111	100	$-4\bar{\mathbf{v}}_2$	100	$-2(\bar{\mathbf{v}}_2 + \bar{\mathbf{v}}_1 + \bar{\mathbf{v}}_0)$

and the channel distortion of the index assignment π_X is

$$\begin{aligned}
D_X &= 2(1-\gamma^3)\left\|(-2/8)(\bar{\mathbf{v}}_2 + \bar{\mathbf{v}}_1 + \bar{\mathbf{v}}_0)\right\|^2 \\
&\quad +2(1-\gamma^2)\left(\left\|(-2/8)(\bar{\mathbf{v}}_2 - \bar{\mathbf{v}}_1 + \bar{\mathbf{v}}_0)\right\|^2 + \left\|(-2/8)(\bar{\mathbf{v}}_2 + \bar{\mathbf{v}}_1 - \bar{\mathbf{v}}_0)\right\|^2\right) \\
&\quad +2(1-\gamma)\left\|(2/8)(\bar{\mathbf{v}}_2 - \bar{\mathbf{v}}_1 - \bar{\mathbf{v}}_0)\right\|^2 \\
&= \frac{1}{8}\Big[4\left(\|\bar{\mathbf{v}}_2\|^2 + \|\bar{\mathbf{v}}_1\|^2 + \|\bar{\mathbf{v}}_0\|^2\right) - \gamma\left\|\bar{\mathbf{v}}_2 - (\bar{\mathbf{v}}_1 + \bar{\mathbf{v}}_0)\right\|^2 \\
&\quad -2\gamma^2\left(\|\bar{\mathbf{v}}_2\|^2 + \|\bar{\mathbf{v}}_1 - \bar{\mathbf{v}}_0\|^2\right) - \gamma^3\left\|\bar{\mathbf{v}}_2 + (\bar{\mathbf{v}}_1 + \bar{\mathbf{v}}_0)\right\|^2\Big].
\end{aligned}$$

Thus, $D_X > D_W$ whenever

$$\begin{aligned}
0 &> \gamma^2\left(\left\|\bar{\mathbf{v}}_2 + (\bar{\mathbf{v}}_1 + \bar{\mathbf{v}}_0)\right\|^2 - 4\|\bar{\mathbf{v}}_2\|^2\right) \\
&\quad +2\gamma\left(\|\bar{\mathbf{v}}_2\|^2 + \|\bar{\mathbf{v}}_1 - \bar{\mathbf{v}}_0\|^2 - 2\left(\|\bar{\mathbf{v}}_1\|^2 + \|\bar{\mathbf{v}}_0\|^2\right)\right) + \left\|\bar{\mathbf{v}}_2 - (\bar{\mathbf{v}}_1 + \bar{\mathbf{v}}_0)\right\|^2 \\
&= \gamma^2\left[2\left(\|\bar{\mathbf{v}}_1 + \bar{\mathbf{v}}_0\|^2 - \|\bar{\mathbf{v}}_2\|^2\right) - \|\bar{\mathbf{v}}_1 + \bar{\mathbf{v}}_0 - \bar{\mathbf{v}}_2\|^2\right] \\
&\quad -2\gamma\left(\|\bar{\mathbf{v}}_1 + \bar{\mathbf{v}}_0\|^2 - \|\bar{\mathbf{v}}_2\|^2\right) + \|\bar{\mathbf{v}}_1 + \bar{\mathbf{v}}_0 - \bar{\mathbf{v}}_2\|^2 \\
&= \left(\left[2\left(\|\bar{\mathbf{v}}_1 + \bar{\mathbf{v}}_0\|^2 - \|\bar{\mathbf{v}}_2\|^2\right) - \|\bar{\mathbf{v}}_1 + \bar{\mathbf{v}}_0 - \bar{\mathbf{v}}_2\|^2\right]\gamma - \|\bar{\mathbf{v}}_1 + \bar{\mathbf{v}}_0 - \bar{\mathbf{v}}_2\|^2\right)(\gamma - 1).
\end{aligned}$$

Hence, for any 8-point BLVQ satisfying

$$\|\bar{\mathbf{v}}_1 + \bar{\mathbf{v}}_0\|^2 > \|\bar{\mathbf{v}}_2\|^2 + \|\bar{\mathbf{v}}_1 + \bar{\mathbf{v}}_0 - \bar{\mathbf{v}}_2\|^2, \tag{3.15}$$

the index assignment π_X is worse than the Worst Code if

$$\frac{\|\bar{\mathbf{v}}_1 + \bar{\mathbf{v}}_0 - \bar{\mathbf{v}}_2\|^2}{2\left(\|\bar{\mathbf{v}}_1 + \bar{\mathbf{v}}_0\|^2 - \|\bar{\mathbf{v}}_2\|^2\right) - \|\bar{\mathbf{v}}_1 + \bar{\mathbf{v}}_0 - \bar{\mathbf{v}}_2\|^2} < \gamma < 1,$$

or equivalently, whenever

$$0 < \epsilon < \frac{\left(\|\bar{\mathbf{v}}_1 + \bar{\mathbf{v}}_0\|^2 - \|\bar{\mathbf{v}}_2\|^2\right) - \|\bar{\mathbf{v}}_1 + \bar{\mathbf{v}}_0 - \bar{\mathbf{v}}_2\|^2}{2\left(\|\bar{\mathbf{v}}_1 + \bar{\mathbf{v}}_0\|^2 - \|\bar{\mathbf{v}}_2\|^2\right) - \|\bar{\mathbf{v}}_1 + \bar{\mathbf{v}}_0 - \bar{\mathbf{v}}_2\|^2}. \tag{3.16}$$

In particular, if $\bar{\mathbf{v}}_1 + \bar{\mathbf{v}}_0 = \alpha\bar{\mathbf{v}}_2$ for $\alpha > 1$, then (3.15) is satisfied and (3.16) reduces to

$$0 < \epsilon < \frac{2\|\bar{\mathbf{v}}_2\|}{\|\bar{\mathbf{v}}_1 + \bar{\mathbf{v}}_0\| + 3\|\bar{\mathbf{v}}_2\|} = \frac{2}{\alpha + 3}. \tag{3.17}$$

The right-hand side of (3.17) can be arbitrarily close to 1/2 as $\alpha \to 1$. Thus, for any $\epsilon \in (0, 1/2)$ a BLVQ can be found for which the index assignment π_X is worse than the Worst Code. ∎

3.5 References

[1] J. D. Marca and N. Jayant, "An algorithm for assigning binary indices to the codevectors of a multi-dimensional quantizer," in *Proceedings IEEE Inter. Conf. on Communications*, vol. 1, (Seattle), pp. 1128–1132, June 1987.

[2] N. Farvardin, "A study of vector quantization for noisy channels," *IEEE Trans. Info. Theory*, vol. IT-36, pp. 799–809, July 1990.

[3] H.-S. Wu and J. Barba, "Index allocation in vector quantisation for noisy channels," *Electronics Letters*, vol. 29, pp. 1318–1320, July 1993.

[4] K. Zeger and A. Gersho, "Pseudo-Gray coding," *IEEE Trans. Communications*, vol. COM-38, pp. 2147–2158, December 1990.

[5] P. Knagenhjelm and E. Agrell, "The Hadamard transform – a tool for index assignment," *IEEE Trans. Info. Theory*, vol. IT-42, pp. 1139–1151, July 1996.

[6] D. J. Goodman and T. J. Moulsley, "Using simulated annealing to design transmission codes for analogue sources," *Electronics Letters*, vol. 24, pp. 617–618, May 1988.

[7] Y. Yamaguchi and T. S. Huang, "Optimum binary fixed-length block codes," Quarterly Progress Report 78, M.I.T. Research Lab. of Electronics, Cambridge, Mass., July 1965.

[8] T. S. Huang, "Optimum binary code," Quarterly Progress Report 82, M.I.T. Research Lab. of Electronics, Cambridge, Mass., July 1966.

[9] T. S. Huang et al., "Design considerations in PCM transmission of low-resolution monochrome still pictures," *Proceedings of the IEEE*, vol. 55, pp. 331–335, March 1967.

[10] T. R. Crimmins, H. M. Horwitz, C. J. Palermo, and R. V. Palermo, "Minimization of mean-square error for data transmitted via group codes," *IEEE Trans. Info. Theory*, vol. IT-15, pp. 72–78, January 1969.

[11] S. W. McLaughlin, D. L. Neuhoff, and J. J. Ashley, "Optimal binary index assignments for a class of equiprobable scalar and vector quantizers," *IEEE Trans. Info. Theory*, vol. IT-41, pp. 2031–2037, November 1995.

[12] A. Méhes and K. Zeger, "Binary lattice vector quantization with linear block codes and affine index assignments," *IEEE Trans. Info. Theory*, vol. IT-44, pp. 79–95, January 1998.

[13] R. E. Totty and G. C. Clark, "Reconstruction error in waveform transmission," *IEEE Trans. Info. Theory*, vol. IT-13, pp. 336–338, April 1967.

[14] A. Méhes and K. Zeger, "Redundancy free codes for binary discrete memoryless channels," in *Proc. Conf. Information Science and Systems*, (Princeton, NJ), pp. 1057–1062, 1994.

[15] A. Méhes and K. Zeger, "On the performance of affine index assignments for redundancy free source-channel coding," in *Proc. Data Compression Conf.*, (Snowbird, UT), p. 433, March 1995.

[16] A. Méhes and K. Zeger, "Affine index assignments for binary lattice quantization with channel noise," in *Proc. IEEE Int. Symp. Information Theory*, (Whistler, BC, Canada), p. 377, September 1995.

[17] R. Hagen and P. Hedelin, "Robust vector quantization by a linear mapping of a block-code," to appear in *IEEE Trans. Info. Theory.*

[18] R. Hagen and P. Hedelin, "Design methods for VQ by linear mappings of block codes," in *Proc. IEEE Int. Symp. Information Theory*, (Trondheim, Norway), p. 241, June 1994.

[19] R. Hagen and P. Hedelin, "Robust vector quantization by linear mappings of block-codes," in *Proc. IEEE Int. Symp. Information Theory*, (San Antonio, TX), p. 171, January 1993.

CHAPTER 4

SOURCE AND CHANNEL RATE ALLOCATION FOR CHANNEL CODES SATISFYING THE GILBERT-VARSHAMOV OR TSFASMAN-VLĂDUT-ZINK BOUNDS

In this chapter, we derive bounds for optimal rate allocation between source and channel coding for linear channel codes that meet the Gilbert-Varshamov or Tsfasman-Vlădut-Zink bounds. Formulas giving the high-resolution vector quantizer distortion of these systems are also derived. In addition, we give bounds on how far below channel capacity the transmission rate should be for a given delay constraint.

4.1 Introduction

One commonly used approach to transmit source information across a noisy channel is to cascade a vector quantizer designed for a noiseless channel, and a block channel coder designed independently of the source coder. A fundamental question for this traditional "separation" technique is to determine the optimal allocation of available transmission rate between source coding and channel coding. Upper [1] and lower [2] distortion bounds on the optimal tradeoff between source and channel coding were previously derived for a binary symmetric channel. They exploit the fact that optimal source coding and optimal channel coding each contribute an exponentially decaying amount to the total distortion

The material in this chapter has been submitted to the *IEEE Transactions on Information Theory* as: A. Méhes and K. Zeger, "Source and channel rate allocation for channel codes satisfying the Gilbert-Varshamov or Tsfasman-Vlădut-Zink bounds."

(averaged over all index assignments), as a function of the overall transmission rate of the system.

In practice, there is usually a constraint on the overall delay and complexity of such a system. This constraint limits the lengths of source blocks and of channel codewords. As a result, the classical approach of Shannon, to transmit channel information at a rate close to the channel's capacity and to encode the source with the corresponding amount of available information, cannot be used in practice. Instead, one must often transmit data at a rate substantially below capacity. The amount below capacity was determined in [2] for binary symmetric channels and in [3] for Gaussian channels. However, the results in both [2] and [3] exploit the existence of codes which have exponentially decaying error probabilities achieving the expurgated error exponent. Although such codes are known to exist, no efficiently decodable ones have yet been discovered. Various suboptimal algorithms do exist for vector quantizer design for noisy channels, but their implementation and design complexities generally grow exponentially fast as a function of the transmission rate of the system.

In this chapter we determine bounds on the optimal tradeoff between source and channel coding for classes of channel codes that attain the Gilbert-Varshamov bound. It is known that, asymptotically, a random linear code achieves the Gilbert-Varshamov bound with probability one [4, 5], although most known structured classes of codes fall short of the bound. The existence of certain Goppa codes, alternant codes, self-dual codes, and double circulant or quasi-cyclic codes, all of which meet the Gilbert-Varshamov bound, has been discussed in [6, p. 557]. A significant breakthrough was achieved by Tsfasman, Vlădut, and Zink [7], where sequences of algebraic-geometry codes over $GF(q)$ (with $q = p^{2m}$ and p prime) were constructed from reductions of modular curves. These codes exceed the Gilbert-Varshamov bound (in an interval of rates) if $q \geq 49$. Katsman, Tsfasman, and Vlădut [8] showed that there is an infinite family of polynomially constructible codes better than the Gilbert-Varshamov bound, although the best presently known (polynomial) algorithms are not yet practical. Another explicit construction of codes above the Gilbert-Varshamov curve was given recently in [9], but a detailed analysis of the algorithmic complexity of the construction is presently lacking. No binary constructions of codes with parameters exceeding the Gilbert-Varshamov bound are known. In fact, it is widely believed that the Gilbert-Varshamov bound is the tightest possible for $q = 2$. The best known binary codes are obtained from good q-ary codes by concatenation. Corresponding bounds are also available, but are generally weaker than the binary version of the Gilbert-Varshamov bound. There are several other bounds for the parameters of both linear and nonlinear, and both binary and nonbinary codes based on algebraic-geometry codes. A summary of these bounds is found in [10] and a standard reference on algebraic-geometry codes is [11]. In [12] variable inner codes and an algebraic geometry outer code are concatenated to obtain exponentially decaying probability of error.

Due to the current lack of practical constructions for channel codes attaining the Gilbert-Varshamov or Tsfasman-Vlădut-Zink bounds, the results of this chapter are presently of theoretical nature. However, with advances in the field of algebraic geometry codes, we anticipate the discovery of more efficient construction algorithms in the future. In addition, the results obtained for codes of this type broaden the class of known channel codes for which quantizer distortions decay to zero exponentially fast with increasing transmission rate. We demonstrate that this class includes certain suboptimal coding schemes. Note that families of channel codes which are not asymptotically good need not have a constant asymptotic decay rate as a function of the overall transmission rate. Indeed, repetition codes and other classes of codes with asymptotically vanishing channel code rates do not possess this property.

To obtain results for families of channel codes attaining the Gilbert-Varshamov or Tsfasman-Vlădut-Zink bounds, we only use the property that a positive monotone decreasing function g (in Proposition 2) exists describing the relationship between the channel code rate and the relative minimum distance of these codes. Thus, the same method of derivation could potentially be used to obtain similar bounds for other classes of asymptotically good channel codes, some of which (e.g., Justesen codes, Blokh-Zyablov codes) are practical. However, it is often difficult to exhibit the function g in an analytically tractable form.

The main results of this chapter are as follows. In Theorem 4.1, upper and lower bounds are given for the optimal tradeoff between source and channel coding for channel codes satisfying the Gilbert-Varshamov or Tsfasman-Vlădut-Zink inequalities. Theorem 4.2 extends a result of [2] for the optimal source-channel coding tradeoff over an unrestricted class of channel codes. Theorems 4.1 and 4.2 enable a comparison of channel codes that achieve the reliability function of the channel (and in this sense are optimal for the given channel) and certain asymptotically good channel codes that are independent of the underlying channel. Figure 4.4 (on page 112) presents an example of the penalty in channel code rate for suboptimality. Note that the bounds compared need not be the tightest possible in all cases. Theorem 4.3 gives the large dimension performance of the optimal tradeoff determined in Theorem 4.1. In [2], the upper and lower bounds on the optimal rate allocation for "optimal" channel codes were shown to coincide for large enough dimensions (dependent on the bit error probability). Thus, we do not derive the large dimension performance corresponding to Theorem 4.2, but in the example shown in Figure 4.7 (on page 117) we include bounds for both optimal and suboptimal channel codes for comparison.

Throughout this chapter we assume a randomized index assignment (i.e., a uniformly random mapping of vector quantizer codevectors to channel codewords). While this assumption is certainly suboptimal from an implementation standpoint, it provides a powerful mathematical tool for obtaining tight performance bounds, analogous in spirit to the classical randomization techniques used to prove Shannon's channel coding theorem. The same index assignment randomization method was used in [1–3] as well. Furthermore, it is not presently known if randomization of index assignments is in general asymptotically suboptimal.

Section 4.2 gives necessary notations, definitions, and lemmas and Section 4.3 presents the source/channel coding tradeoff problem. Section 4.4 gives basic results on bounds and error exponents. The main results of this chapter are given in Section 4.5, and one technically complicated proof is left to Section 4.7.

4.2 Preliminaries

The following notations will be useful in our asymptotic analysis.

Notation Let $f(n)$ and $g(n)$ be real-valued sequences. Then, we write

- $f = O(g)$, if there is a positive real number c, and a positive integer n_0 such that $|f(n)| \leq c|g(n)|$, whenever $n > n_0$;
- $f = o(g)$, if g has only a finite number of zeros, and $f(n)/g(n) \to 0$ as $n \to \infty$;
- $f = \Theta(g)$, if there are positive real numbers c_1 and c_2, and a positive integer n_0, such that $c_1|g(n)| \leq |f(n)| \leq c_2|g(n)|$, for all $n > n_0$.

We obtain bounds on the optimal rate allocation for the cascaded system depicted in Figure 4.1. Similar to Chapter 2, the source coder is a vector quantizer.

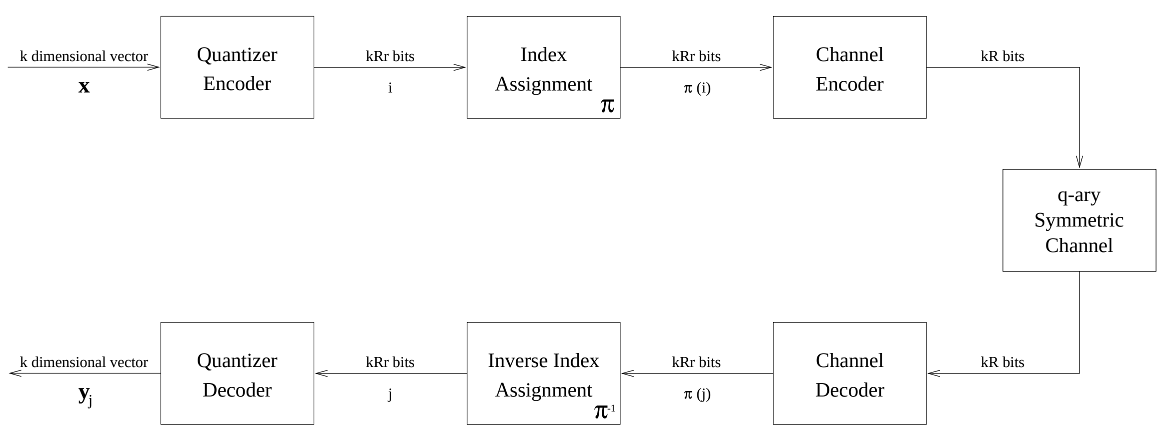


Figure 4.1 Cascaded vector quantizer and channel coder system.

Definition 4.1 A k-dimensional, M-point *vector quantizer* is a mapping from k-dimensional Euclidean space $\mathbb{R}^k$ to a set of *codevectors* $\{\mathbf{y}_1 \ldots, \mathbf{y}_M\} \subset \mathbb{R}^k$. Associated with each codevector $\mathbf{y}_i$ is an *encoder region* $\mathcal{R}_i \subset \mathbb{R}^k$, the set of all points in $\mathbb{R}^k$ that are mapped by the quantizer to $\mathbf{y}_i$. The set of encoder regions forms a partition of $\mathbb{R}^k$. The *rate* (or resolution) of a vector quantizer is defined as $R_s = (\log_2 M)/k$.

A vector quantizer is commonly decomposed into a quantizer *encoder* and a quantizer *decoder*. For each input vector, the encoder produces the index $i \in \{1, \ldots, M\}$ of the encoder region $\mathcal{R}_i$ containing the input vector. For each index i, the decoder outputs the codevector $\mathbf{y}_i$.

The p^{th}*-power distortion* of a vector quantizer is

$$D_0 = \sum_{i=1}^{M} \int_{\mathcal{R}_i} \|\mathbf{x} - \mathbf{y}_i\|^p d\mu(\mathbf{x}), \tag{4.1}$$

where $\|\cdot\|$ is the usual Euclidean norm, and μ is the probability distribution of a k-dimensional source vector. The subscript 0 is used to distinguish the distortion on an error-free channel from the distortion due to a noisy channel (to be discussed later). The high-resolution (i.e., large R_s) behavior of D_0 can be described by Zador's formula.

Lemma 4.1 (Zador [13]) *The minimum p^{th}-power distortion of a rate R_s vector quantizer is asymptotically given by*

$$D_0 = 2^{-pR_s + O(1)}.$$

This is often referred to as the "6 dB/bit/component rule" for $p = 2$, since

$$10 \log_{10} \left(2^{-pR_s}/2^{-p(R_s+1)}\right) \approx 3p.$$

In addition to the minimum distortion achieved by optimal quantizers, the asymptotic distortion of several other classes of vector quantizers, including uniform quantizers and other lattice-based quantizers, has the same high-resolution decay rate.

Definition 4.2 We call a vector quantizer that achieves the asymptotic distortion of Lemma 4.1 a *good vector quantizer.*

Motivated by the (nonbinary) alphabet size requirements of algebraic geometry codes, we consider channel codes over GF(q) and use a q-ary symmetric channel in our system model shown in Figure 4.1. The following two definitions formally introduce q-ary symmetric channels and q-ary linear block channel codes.

Definition 4.3 A *discrete memoryless channel* is a probabilistic mapping from an input alphabet $\mathcal{A}$ to an output alphabet $\mathcal{B}$ characterized by *channel transition probabilities* $P(b|a)$, i.e., the probability that the channel maps an input symbol $a \in \mathcal{A}$ to the output symbol $b \in \mathcal{B}$. A *q-ary symmetric channel* with symbol error probability $\epsilon \in [0, 1-q^{-1}]$ is a discrete memoryless channel having $\mathcal{A} = \mathcal{B} = \{0, \ldots, q-1\}$ and channel transition probabilities

$$P(a|b) = I_{\{a=b\}}(1-\epsilon) + I_{\{a \neq b\}}\frac{\epsilon}{q-1} \qquad a, b \in \{0, \ldots, q-1\}, \tag{4.2}$$

where I denotes the indicator function.

Definition 4.4 An (n, k) *block channel code* is a set of length n strings of q-ary symbols, called *codewords*. A *linear q-ary $[n, k, d]_q$ block channel code* is a linear subspace of $[\mathrm{GF}(q)]^n$, containing $M = q^k$ codewords, each with at least d nonzero components. The number $r = \frac{k}{n} \in (0, 1]$ is the *channel code rate*.

Associated with a channel code is a *channel encoder* and a *channel decoder*. The channel encoder is a one-to-one mapping of messages (e.g., quantizer indices) to channel codewords for transmission. The channel decoder, on the other hand, is a many-to-one mapping. It maps received sequences of channel symbols (not necessarily codewords) to messages. Denoting the channel codeword corresponding to m by $c^{(m)}$, and the set of length n sequences decoded into l by S_l, the transition probabilities of the coded channel are

$$\beta_{l|m} = \sum_{u \in S_l} \prod_{i=1}^{n} P(u_i | c_i^{(m)}),$$

where $u_i, c_i^{(m)} \in \{0, \ldots, q-1\}$ are the i^{th} symbols of u and $c^{(m)}$, respectively. The *average probability of decoding error* (for a uniform source) is

$$P_e = \frac{1}{M} \sum_{l=1}^{M} \left(1 - \beta_{l|l}\right). \tag{4.3}$$

Although we never assume a uniform source, this definition of P_e is notationally convenient in what follows. The following two lemmas state classical asymptotic upper and lower bounds on P_e.

Lemma 4.2 **[14, pp. 140, 153]** *For every $r < C$, there exist sequences of (n, rn) channel codes such that*

$$P_e < e^{-nE_{\max}(r)+o(n)},$$

where C denotes the capacity of the channel, and $E_{\max}(r) = \max\left(E_{\text{rc}}(r), E_{\text{ex}}(r)\right)$ is the maximum of the "random coding" and the "expurgated" error exponents.[1]

Lemma 4.2 characterizes the class of channel codes considered in [2]. For easier reference, we introduce the following terminology.

Definition 4.5 We call a block channel code that achieves the asymptotic error exponent in Lemma 4.2 an *efficient channel code.*

Lemma 4.3 **[14, p. 157]** *Any sequence of (n, rn) channel codes on a discrete memoryless channel must satisfy*

$$P_e > e^{-nE_{\text{sp}}(r)+o(n)},$$

where $E_{\text{sp}}(r)$ is the "sphere packing" error exponent.

While Lemma 4.2 is an existence result, Lemma 4.3 holds for all channel codes. The error exponent functions depend on the channel statistics. Definitions of E_{rc}, E_{ex}, and E_{sp} in terms of the transition probabilities of a discrete memoryless channel, and a derivation

[1] The notation E_{rc} is used instead of the usual E_r to avoid confusing the subscript and the rate r.

of closed-form expressions for q-ary symmetric channels are given in Section 4.7. All three of these error exponent functions are known to be positive and convex in the range $0 < r < C$.

Another element of our system model shown in Figure 4.1 is an index assignment.

Definition 4.6 An *index assignment* π is a permutation of the index set $\{1, \ldots, M\}$.

The purpose of an index assignment is to match a vector quantizer and a channel coder in a cascaded system in order to minimize the end-to-end distortion. Distance properties of channel codewords and quantizer codevectors should be aligned, so that on average a likely channel error (small Hamming distance) results in a tolerable quantization error (small Euclidean distance).

4.3 Problem Formulation

Consider a k-dimensional vector quantizer cascaded with a channel coder operating over a q-ary symmetric channel with a fixed overall transmission rate R measured in bits per vector component, as shown in Figure 4.1. For each k-dimensional input vector a channel codeword consisting of n q-ary symbols is transmitted across the channel to the receiver. The transmission rate is $R = (n \log_2 q)/k$. Let $r \in [0, 1]$ denote the rate of a q-ary $[n, rn, d]_q$ linear block channel code, where d is the minimum distance of the code (in q-ary symbols). The source coding rate and the overall transmission rate are related by $R_s = Rr$. Let M denote the number of quantizer codevectors (equivalently, the number of channel codewords). Then, $M = 2^{kR_s} = 2^{kRr} = q^{rn}$. For each input vector $\mathbf{x} \in \mathbb{R}^k$, the quantizer encoder produces an integer index $i \in \{1, \ldots, M\}$, which in turn is mapped to another index $\pi(i)$ by an index assignment. The channel encoder transmits the $\pi(i)^{\text{th}}$ channel codeword through a q-ary symmetric channel (n q-ary symbols corresponding to kR bits). At the receiver, the channel decoder reconstructs an index $\pi(j)$ from the (possibly corrupted) n q-ary symbols received from the channel. Then the inverse index assignment is performed and the quantizer codevector $\mathbf{y}_j \in \mathbb{R}^k$ corresponding to the resulting index j is presented at the output.

For a given index assignment π, the average p^{th}-power distortion can be expressed as

$$D(\pi) = \sum_{i=1}^{M} \sum_{j=1}^{M} \beta_{\pi(j)|\pi(i)} \int_{\mathcal{R}_i} \|\mathbf{x} - \mathbf{y}_j\|^p d\mu(\mathbf{x}). \tag{4.4}$$

There are no known general techniques for analytically determining $\min_\pi D(\pi)$. As an alternative, we randomize the choice of index assignment. This technique serves as a tool in obtaining an existence theorem, and also models the choice of index assignment in systems where index design is ignored. Hence, we examine the following distortion:

$$D = \frac{1}{M!} \sum_{\pi} D(\pi) = \sum_{i=1}^{M} \sum_{j=1}^{M} \left[\frac{1}{M!} \sum_{\pi} \beta_{\pi(j)|\pi(i)} \right] \int_{\mathcal{R}_i} \|\mathbf{x} - \mathbf{y}_j\|^p d\mu(\mathbf{x}), \tag{4.5}$$

where the sums over π are taken over all $M!$ permutations of the integers $\{1, \ldots, M\}$. The averaging effectively replaces the original q-ary symmetric channel by a "new" M-ary symmetric channel whose symbol error probability equals the average probability of channel decoding error P_e of the underlying channel. We have

$$\begin{aligned}
\frac{1}{M!} \sum_{\pi} \beta_{\pi(j)|\pi(i)} &= \frac{1}{M!} \sum_{\pi} \sum_{l=1}^{M} \sum_{m=1}^{M} \beta_{m|l} I_{\{\pi(i)=l, \pi(j)=m\}} \\
&= I_{\{i=j\}} \sum_{l=0}^{M} \beta_{l|l} \frac{1}{M!} \sum_{\pi} I_{\{\pi(i)=l\}} + I_{\{i \neq j\}} \sum_{l=1}^{M} \sum_{\substack{m=1 \\ m \neq l}}^{M} \beta_{m|l} \frac{1}{M!} \sum_{\pi} I_{\{\pi(i)=l, \pi(j)=m\}} \\
&= I_{\{i=j\}} \sum_{l=1}^{M} \beta_{l|l} \frac{(M-1)!}{M!} + I_{\{i \neq j\}} \sum_{l=1}^{M} \left(1 - \beta_{l|l}\right) \frac{(M-2)!}{M!} \\
&= I_{\{i=j\}} \left(1 - P_e\right) + I_{\{i \neq j\}} \frac{P_e}{M-1}.
\end{aligned} \tag{4.6}$$

Substituting (4.6) into (4.5) yields

$$D = (1 - P_e) \sum_{i=1}^{M} \int_{\mathcal{R}_i} \|\mathbf{x} - \mathbf{y}_i\|^p d\mu(\mathbf{x}) + \frac{P_e}{M-1} \sum_{i=1}^{M} \sum_{\substack{j=1 \\ j \neq i}}^{M} \int_{\mathcal{R}_i} \|\mathbf{x} - \mathbf{y}_j\|^p d\mu(\mathbf{x}). \tag{4.7}$$

The sum in the first term of (4.7) is the distortion for a noiseless channel. We assume that the source has compact support, in which case

$$\begin{aligned}
\frac{1}{M-1} \sum_{i=1}^{M} \sum_{\substack{j=1 \\ j \neq i}}^{M} \int_{\mathcal{R}_i} \|\mathbf{x} - \mathbf{y}_j\|^p d\mu(\mathbf{x}) &\leq \frac{1}{M-1} \sum_{i=1}^{M} (M-1) \max_{\mathbf{x} \in \mathcal{R}_i, j \neq i} \|\mathbf{x} - \mathbf{y}_j\|^p \int_{\mathcal{R}_i} d\mu(\mathbf{x}) \\
&\leq \operatorname{diam}(\mu),
\end{aligned}$$

where $\mathrm{diam}(\mu)$ is the diameter of the support region. Unless the source is deterministic, a nonzero lower bound on the same double sum can be obtained using the p^{th}-moment type quantity $\nu_p(\mu) = \min_{\mathbf{y}} \int_{\mathbb{R}^k} \|\mathbf{x}-\mathbf{y}\|^p d\mu(\mathbf{x})$. Namely,

$$
\begin{aligned}
\frac{1}{M-1}\sum_{i=1}^{M}\sum_{\substack{j=1\\ j\neq i}}^{M}\int_{\mathcal{R}_i}\|\mathbf{x}-\mathbf{y}_j\|^p d\mu(\mathbf{x}) &= \frac{1}{M-1}\sum_{j=1}^{M}\sum_{i=1}^{M}\int_{\mathcal{R}_i}\|\mathbf{x}-\mathbf{y}_j\|^p d\mu(\mathbf{x}) \\
&\quad - \frac{1}{M-1}\sum_{i=1}^{M}\int_{\mathcal{R}_i}\|\mathbf{x}-\mathbf{y}_i\|^p d\mu(\mathbf{x}) \\
&\geq \frac{M\,\nu_p(\mu) - D_0}{M-1} \\
&\geq \nu_p(\mu).
\end{aligned}
$$

Note that both the upper and lower bounds above depend solely on the source and not on the channel. Thus, returning to (4.7) we have

$$D = (1-P_e)\,D_0 + P_e\Theta(1). \tag{4.8}$$

We assume a vector quantizer and an efficient channel code. Then, using Lemma 4.1 to bound D_0, and Lemma 4.3 and Lemma 4.2 to bound P_e, the average p^{th}-power distortion D of a cascaded source coder and rate r channel coder, with transmission rate R can asymptotically (as $R \to \infty$) be bounded as

$$2^{-pRr+O(1)} + 2^{-kRE_{\mathrm{sp}}(r)+o(R)} \leq D \leq 2^{-pRr+O(1)} + 2^{-kRE_{\max}(r)+o(R)} \tag{4.9}$$

where the error exponents have been scaled by a factor of $\ln q$ as compared to Lemmas 4.2 and 4.3, in order to change the unit of block length from symbols to bits. The minimum value of the right side of (4.9) over all $r \in [0,1]$ is an asymptotically achievable (as $R \to \infty$) distortion D, and the minimum value of the left side of (4.9) is a lower bound on D for any choice of r. Let $r_{\max}$ and r_{sp} respectively denote the values of r which minimize (asymptotically) the right and left sides of (4.9). Then $r_{\max} \leq r^* \leq r_{\mathrm{sp}}$, where r^* is the optimal rate allocation. It can be seen that to minimize the bounds in (4.9), the exponents of the two decaying exponentials in each bound have to be balanced, so that

$$E_X(r_X) = \frac{p}{k} r_X + o(1), \tag{4.10}$$

where formally $X \in \{\mathrm{sp}, \max\}$ and $o(1) \to 0$ as $R \to \infty$. The distortion achieved with a channel code rate r^* in this case is

$$D = 2^{-pRr^* + O(1)}.$$

The values of $r_{\max}$ and r_{sp} were determined in [2] for efficient binary channel codes. We investigate the problem of optimal rate allocation for channel codes that attain the Gilbert-Varshamov bound and/or the Tsfasman-Vlăduţ-Zink bound (or "basic algebraic-geometry bound"). Such codes are in general weaker than those in [2], but are potentially less algorithmically complex. Our results also generalize those in [2] to q-ary channels.

4.4 Error Exponents

In this section, we present the classical channel coding error exponents E_{rc}, E_{ex}, and E_{sp} specialized to a q-ary symmetric channel, and we derive two new q-ary error exponents E_{GV} and E_{TVZ} for channel codes that satisfy the Gilbert-Varshamov and Tsfasman-Vlăduţ-Zink inequalities, respectively. All five of these error exponents can be concisely written using q-ary versions of the entropy, the relative entropy, and Rényi's entropy of order $1/2$. We start with the general definitions of these information measures.

Definition 4.7 Let P and $\hat{P}$ be probability distributions on a finite set.

The *entropy* of P is

$$H(P) = -\sum_x P(x) \log_2 P(x). \tag{4.11}$$

The *relative entropy* between P and $\hat{P}$ is

$$D(P \| \hat{P}) = \sum_x P(x) \log_2 \left(P(x) / \hat{P}(x) \right). \tag{4.12}$$

The *Rényi entropy of order* α of P is

$$H_\alpha(P) = \frac{1}{1-\alpha}\log_2 \sum_x [P(x)]^\alpha, \tag{4.13}$$

for $\alpha > 0$, $\alpha \neq 1$. Jensen's inequality implies $H_\alpha(P) \geq H(P)$ for $\alpha \in (0,1)$, and $H_\alpha(P) \leq H(P)$ for $\alpha > 1$. Details of Rényi's information measures are given in [15].

Next, we introduce the various q-ary entropy functions defined for one-parameter distributions related to the transition probabilities of a q-ary symmetric channel.

Definition 4.8 Let $\epsilon, \delta \in [0, 1-q^{-1}]$, and let $\mathcal{P}_\epsilon$ and $\mathcal{P}_\delta$ be probability distributions on $\{0, \ldots, q-1\}$ with respective probabilities $(1-\epsilon, \frac{\epsilon}{q-1}, \ldots, \frac{\epsilon}{q-1})$, and $(1-\delta, \frac{\delta}{q-1}, \ldots, \frac{\delta}{q-1})$.

The *q-ary entropy function* is defined as

$$\mathcal{H}_q(\epsilon) \triangleq H(\mathcal{P}_\epsilon)/\log_2 q = \epsilon \log_q(q-1) - \epsilon \log_q \epsilon - (1-\epsilon)\log_q(1-\epsilon). \tag{4.14}$$

For $q = 2$ this gives the *binary entropy function* $h(\epsilon) = -\epsilon \log_2 \epsilon - (1-\epsilon)\log_2(1-\epsilon)$. The derivative of $\mathcal{H}_q$ with respect to ϵ is

$$\mathcal{H}'_q(\epsilon) = \log_q(q-1) - \log_q \epsilon + \log_q(1-\epsilon), \tag{4.15}$$

and the second derivative is

$$\mathcal{H}''_q(\epsilon) = -\frac{\log_q e}{\epsilon(1-\epsilon)}. \tag{4.16}$$

Thus, $\mathcal{H}_q(\epsilon)$ is concave, strictly increasing on $[0, 1-q^{-1}]$, and achieves its maximum $\mathcal{H}_q(1-q^{-1}) = 1$ and its minimum $\mathcal{H}_q(0) = 0$. The notation $\mathcal{H}_q^{-1}$ denotes the inverse of $\mathcal{H}_q\colon [0, 1-q^{-1}] \to [0,1]$. Clearly, $\mathcal{H}_q^{-1}$ is convex, from (4.16). The *capacity* of a q-ary symmetric channel with symbol error probability $\epsilon \in [0, 1-q^{-1}]$ expressed in q-ary symbols is

$$C_q = 1 - \mathcal{H}_q(\epsilon). \tag{4.17}$$

The *(q-ary) relative entropy (information divergence) function* is defined as

$$\mathcal{D}_q(\delta \| \epsilon) = D(\mathcal{P}_\delta \| \mathcal{P}_\epsilon)/\log_2 q = \delta \log_q \frac{\delta}{\epsilon} + (1-\delta)\log_q \frac{1-\delta}{1-\epsilon}, \tag{4.18}$$

which can also be expressed in terms of the q-ary entropy function as

$$\mathcal{D}_q\left(\delta \| \epsilon\right) = \mathcal{H}_q(\epsilon) + (\delta - \epsilon)\mathcal{H}_q'(\epsilon) - \mathcal{H}_q(\delta). \tag{4.19}$$

For $|\delta - \epsilon|$ small, a Taylor series approximation of $\mathcal{H}_q(\delta)$ around ϵ gives

$$\mathcal{D}_q\left(\delta \| \epsilon\right) = -\frac{1}{2}(\delta - \epsilon)^2\mathcal{H}_q''(\epsilon) + O\left(|\delta - \epsilon|^3\right). \tag{4.20}$$

We restrict attention to Rényi's entropy of order 1/2, and the corresponding channel capacity of order 1/2 for a q-ary symmetric channel. The *q-ary entropy function of order 1/2* is defined as

$$\mathcal{H}_q^{(1/2)}(\epsilon) \stackrel{\triangle}{=} H_{\frac{1}{2}}(\mathcal{P}_\epsilon)/\log_2 q = 2\log_q\left(\sqrt{1-\epsilon} + \sqrt{\epsilon(q-1)}\right). \tag{4.21}$$

The capacity of order 1/2 of a q-ary symmetric channel with symbol error probability $\epsilon \in [0, 1 - q^{-1}]$ expressed in q-ary symbols is

$$C_q^{(1/2)} = 1 - \mathcal{H}_q^{(1/2)}(\epsilon), \tag{4.22}$$

which Csiszár [15] showed to equal the "cutoff rate" of the channel.

The error exponents of Lemmas 4.2 and 4.3 can be specialized to a q-ary symmetric channel as follows (the proof of Proposition 1 is given in Section 4.7).

Proposition 1

$$E_{\mathrm{sp}}(r) = \mathcal{D}_q\left(\mathcal{H}_q^{-1}(1-r)\big\| \epsilon\right) \qquad r \in (0, C_q) \tag{4.23}$$

$$E_{\mathrm{rc}}(r) = \begin{cases} C_q^{(1/2)} - r & r \in (0, r_2] \\ \mathcal{D}_q\left(\mathcal{H}_q^{-1}(1-r)\big\| \epsilon\right) & r \in [r_2, C_q) \end{cases} \tag{4.24}$$

$$E_{\mathrm{ex}}(r) = \begin{cases} \mathcal{H}_q^{-1}(1-r)\log_q \frac{q-1}{q^{\mathcal{H}_q^{(1/2)}(\epsilon)}-1} & r \in (0, r_1] \\ C_q^{(1/2)} - r & r \in [r_1, C_q^{(1/2)}) \end{cases} \tag{4.25}$$

where $r_1 = 1 - \mathcal{H}_q\left(1 - q^{-\mathcal{H}_q^{(1/2)}(\epsilon)}\right)$ *and* $r_2 = 1 - \mathcal{H}_q\left(\frac{\sqrt{(q-1)\epsilon}}{\sqrt{1-\epsilon}+\sqrt{(q-1)\epsilon}}\right)$.

Also, since $r_2 \leq C_q^{(1/2)} \leq C_q$, and $r_1 \leq r_2$ for $\epsilon < 1 - q^{-1}$, we have

$$E_{\max}(r) = \begin{cases} \mathcal{H}_q^{-1}(1-r)\log_q \frac{q-1}{q^{\mathcal{H}_q^{(1/2)}(\epsilon)}-1} & r \in (0, r_1] \\ C_q^{(1/2)} - r & r \in [r_1, r_2] \\ \mathcal{D}_q\left(\mathcal{H}_q^{-1}(1-r) \middle\| \, \epsilon\right) & r \in [r_2, C_q) \end{cases}. \tag{4.26}$$

The lower bound on P_e given in Lemma 4.3 holds for an arbitrary code. The upper bound of Lemma 4.2, however, is an existence result. Analogous upper bounds and corresponding error exponents can be obtained for "asymptotically good" families of codes.

For a sequence of (n, nr, d) codes to be asymptotically good, both the rate r and the relative minimum distance d/n must be bounded away from zero as the block length n increases. Usually, bounds are given in the form $d \geq ng(r)$ or $r \geq g^{-1}(d/n)$, for some monotonic decreasing function g. In this chapter we consider two of the best known such bounds, the Gilbert-Varshamov bound and the Tsfasman-Vlădut-Zink bound (see [11, p. 609] for a summary of these and several related bounds).

Definition 4.9 *An $[n, nr, d]_q$ code is said to satisfy the*

- *Gilbert-Varshamov bound, if*

$$r \geq 1 - \mathcal{H}_q(d/n),$$

- *Tsfasman-Vlădut-Zink bound, if*

$$r \geq 1 - d/n - (\sqrt{q} - 1)^{-1}.$$

The following lemma provides a bound on the tail of a binomial distribution.

Lemma 4.4 ([14, p. 531]) *For $\delta > \epsilon \geq 0$,*

$$\sum_{i=n\delta}^{n} \binom{n}{i} \epsilon^i (1-\epsilon)^{n-i} \leq 2^{-n\mathcal{D}_2(\delta \| \epsilon)}.$$

Proposition 2 *If an* $[n, rn, d]_q$ *linear block channel code has minimum distance* $d \geq ng(r)$ *for some positive monotone decreasing function* g*, then the average probability of decoding error on a q-ary symmetric channel with symbol error probability* ϵ *satisfies*

$$P_e \leq 2^{-n\mathcal{D}_2\left(\frac{1}{2}g(r)\|\epsilon\right)} \qquad r \in (0, g^{-1}(2\epsilon)).$$

Proof

Since a code with minimum distance d can correct at least $\lfloor \frac{d-1}{2} \rfloor$ errors,

$$\begin{aligned} P_e &\leq \sum_{i=\lfloor \frac{d-1}{2} \rfloor + 1}^{n} \binom{n}{i} (q-1)^i (1-\epsilon)^{n-i} \left(\frac{\epsilon}{q-1}\right)^i && (4.27) \\ &\leq \sum_{i=n\frac{1}{2}g(r)}^{n} \binom{n}{i} \epsilon^i (1-\epsilon)^{n-i} && (4.28) \\ &\leq 2^{-n\mathcal{D}_2\left(\frac{1}{2}g(r)\|\epsilon\right)}, && (4.29) \end{aligned}$$

where inequality (4.28) follows from $\lfloor \frac{d-1}{2} \rfloor + 1 \geq d/2 \geq ng(r)/2$, and inequality (4.29) from Lemma 4.4. ■

The bound on P_e in (4.27) used to obtain Proposition 2 may suggest the use of bounded distance decoding in the channel decoder. Since this is generally suboptimal, using maximum likelihood decoding is preferable and yields a lower probability of decoding error in most cases. While using tighter bounds on P_e may also improve the rate allocation bounds derived later in this chapter, we opted for the "standard" bound (inequality (4.27)) because it depends only on the minimum distance. This enables us to directly apply the function g relating the rate and the relative minimum distance, without any further assumptions on the structure of the channel codes. The upper bound on P_e given in Proposition 2 only depends on the code parameters n and r, the symbol error probability ϵ, and the function g. The following two corollaries follow immediately from Proposition 2 and will also be useful in what follows.

Corollary 4.1 *Consider the cascade of a good k-dimensional vector quantizer, a q-ary linear block channel coder that achieves the Gilbert-Varshamov bound, and a q-ary symmetric channel with symbol error probability* ϵ *and overall transmission rate* R*. For every*

$r < 1 - \mathcal{H}_q(2\epsilon) = C_q(2\epsilon)$, *the average probability of channel decoding error satisfies*

$$P_e \leq 2^{-kRE_{\mathrm{GV}}(r)},$$

where

$$E_{\mathrm{GV}}(r) = \mathcal{D}_q\left(\frac{1}{2}\mathcal{H}_q^{-1}(1-r)\middle\|\,\epsilon\right) \tag{4.30}$$

is the Gilbert-Varshamov error exponent.

Corollary 4.2 *Consider the cascade of a good k-dimensional vector quantizer, a q-ary linear block channel coder that achieves the Tsfasman-Vlădut-Zink bound, and a q-ary symmetric channel with symbol error probability ϵ and overall transmission rate R. For every $r < 1 - (\sqrt{q}-1)^{-1} - 2\epsilon$, the average probability of channel decoding error satisfies*

$$P_e \leq 2^{-kRE_{\mathrm{TVZ}}(r)},$$

where

$$E_{\mathrm{TVZ}}(r) = \mathcal{D}_q\left(\frac{1}{2}\left(1 - r - (\sqrt{q}-1)^{-1}\right)\middle\|\,\epsilon\right) \tag{4.31}$$

is the Tsfasman-Vlădut-Zink error exponent.

Analogous to $E_{\max}(r) = \max\left(E_{\mathrm{rc}}(r), E_{\mathrm{ex}}(r)\right)$, we define

$$E'_{\max}(r) = \max\left(E_{\mathrm{GV}}(r), E_{\mathrm{TVZ}}(r)\right).$$

For $q \leq 49$, $E'_{\max}(r) = E_{\mathrm{GV}}(r)$. For $q > 49$,

$$E'_{\max}(r) = \begin{cases} \mathcal{D}_q\left(\frac{1}{2}\left(1 - r - (\sqrt{q}-1)^{-1}\right)\middle\|\,\epsilon\right) & r \in [r'_1, r'_2] \\ \mathcal{D}_q\left(\frac{1}{2}\mathcal{H}_q^{-1}(1-r)\middle\|\,\epsilon\right) & r \in (0, r'_1]\bigcup[r'_2, C_q(2\epsilon)) \end{cases}$$

where $r'_1 < r'_2$ are roots of $\mathcal{H}_q^{-1}(1-r) = 1 - r + (\sqrt{q}-1)^{-1}$. (In [7] this equation is shown to have two distinct roots for $q > 49$.)

4.5 Optimal Rate Allocation

The bounds we obtain on the optimal rate allocation in a cascaded vector quantizer and channel coder system are functions of the vector dimension k, the channel symbol error probability ϵ, and the parameter p in the distortion criterion. These bounds do not depend, however, on the source statistics. We obtain analytic bounds on the optimal rate allocation for two important cases of interest: a large vector dimension k, and a small symbol error probability ϵ. In each case the remaining parameters are assumed fixed but arbitrary. To obtain these bounds, we analyze the error exponents E_{sp}, $E_{\max}$, and $E'_{\max}$.

First, we note that on the interval $[r_1, r_2]$, the function $E_{\max}(r) = C_q^{(1/2)} - r$ is linear. Let r_{lin} be a solution of (4.10) (for $X = \max$) such that $r_{\text{lin}} \in [r_1, r_2]$, whenever such a solution exists. Then,

$$C_q^{(1/2)} - r_{\text{lin}} = \frac{p}{k} r_{\text{lin}},$$

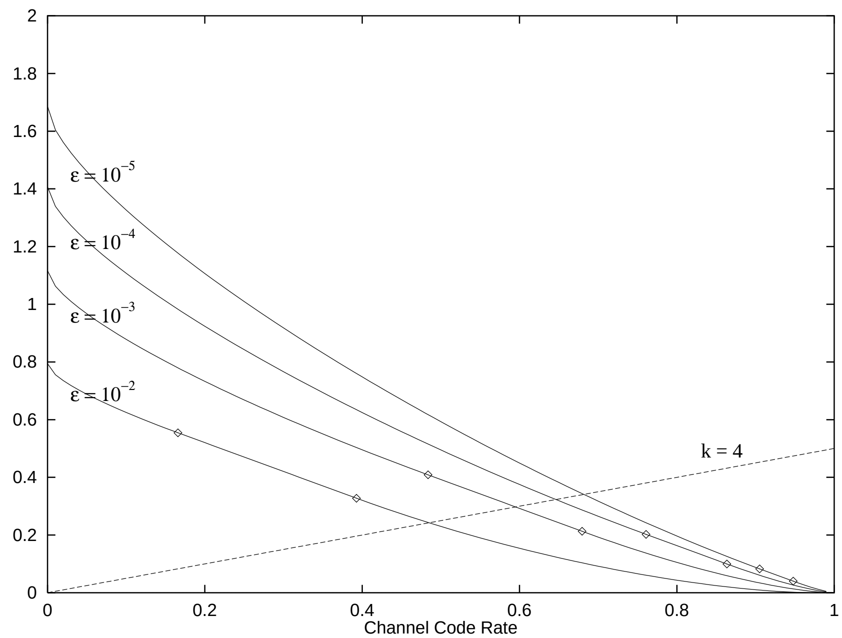
or equivalently, $r_{\text{lin}} = C_q^{(1/2)} / \left(1 + (p/k)\right)$. If k is fixed, and ϵ approaches zero, then $r_{\text{lin}} \to \left(1 + (p/k)\right)^{-1}$ and $r_1 \to 1$. Hence, $r_{\text{lin}} < r_1$ for ϵ sufficiently small. Thus, for ϵ sufficiently small, $r_{\max} < r_1$ and it therefore suffices to restrict attention to E_{ex} instead of $E_{\max}$ (see Figure 4.2(a)).

If ϵ is fixed and k increases, then $r_{\text{lin}} \to C_q^{(1/2)}$. Hence, $r_{\text{lin}} > r_2$ for k sufficiently large. Thus, for k sufficiently large, $r_{\max} > r_2$ and thus it suffices to restrict attention to E_{rc} instead of $E_{\max}$ (see Figure 4.2(b)). Also note that $E_{\text{rc}}(r) = E_{\text{sp}}(r)$ for all $r \in [r_2, C_q)$. Thus, the upper and lower bounds coincide as k increases, and hence, it suffices to consider E_{sp}.

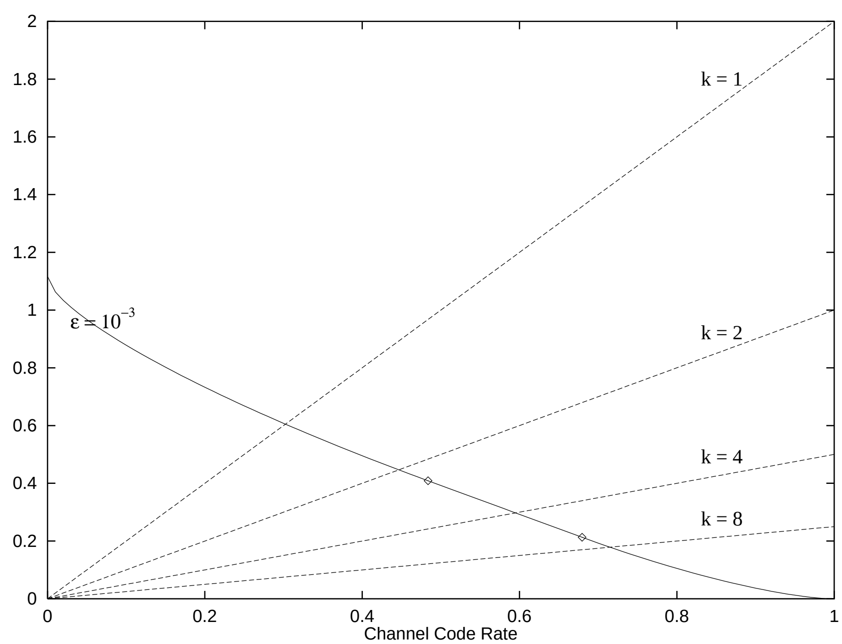
Next, we examine $E'_{\max}$. For $q < 49$, $E'_{\max} = E_{\text{GV}}$ for all r. For $q \geq 49$, note that r'_2 is independent of both k and ϵ and depends only on q. Thus, for k fixed, and ϵ decreasing, $\frac{p}{k} r'_2$ (the right-hand side of (4.10)) is constant, whereas $\mathcal{D}_q\left(\mathcal{H}_q^{-1}(1 - r'_2)/2 \,\|\, \epsilon\right)$ (the left-hand side of (4.10)) increases without bound. Hence, for ϵ small enough,

$$\mathcal{D}_q\left(\mathcal{H}_q^{-1}(1 - r'_2)/2 \,\|\, \epsilon\right) > \frac{p}{k} r'_2. \tag{4.32}$$

Since $\frac{p}{k} r$ is a monotone increasing function of r, and $E'_{\max}(r)$ is monotone decreasing in r, (4.32) implies that if $E'_{\max}(r'_{\max}) = \frac{p}{k} r'_{\max}$ then $r'_{\max} > r'_2$ (see Figure 4.3(a)).

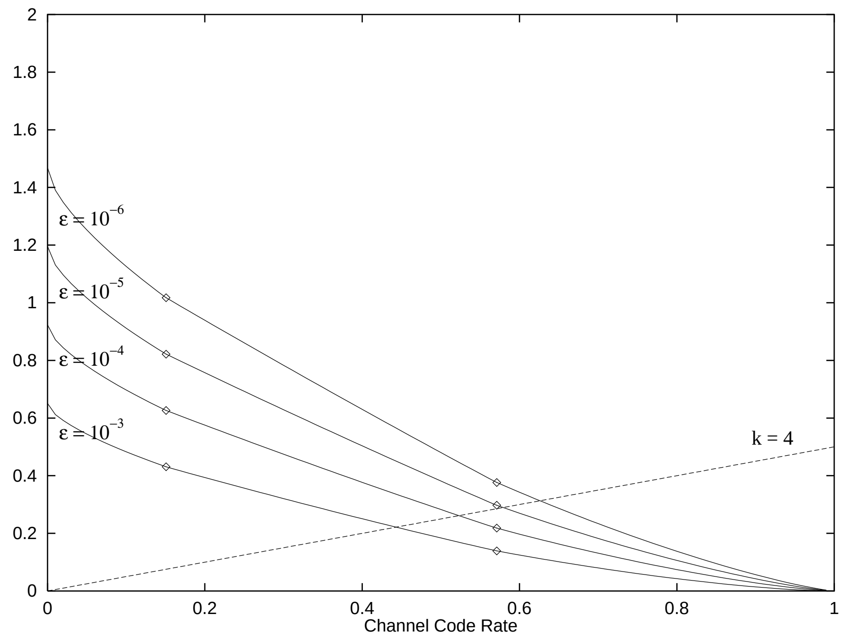


(a) Small bit error probability

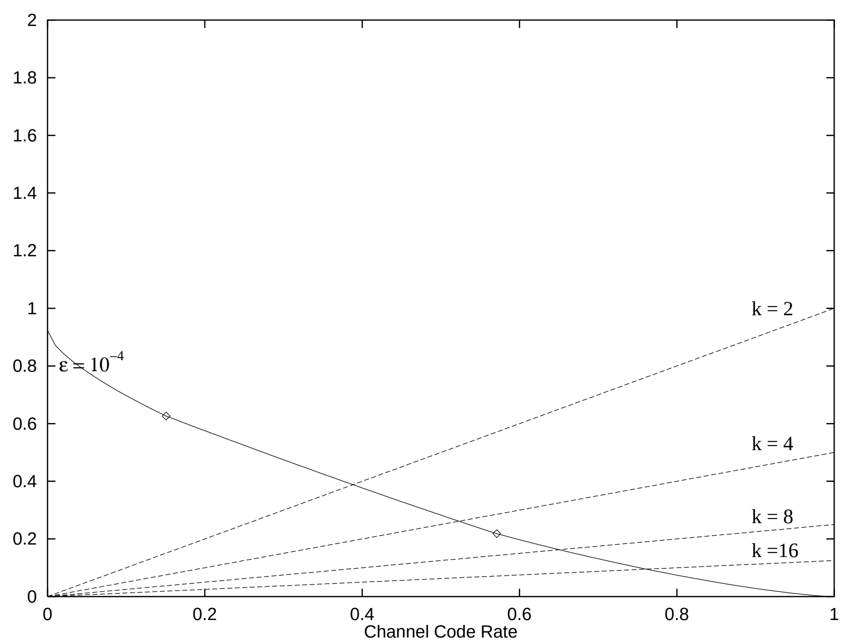


(b) Large vector dimension

Figure 4.2 A graphical solution to (4.10) for $E_{\max}$ ($p = 2$, $q = 64$). The solid curves show $E_{\max}(r)$ for different values of ϵ, and the dashed lines have slope p/k. The two dots on each error exponent curve correspond to r_1 and r_2.



(a) Small bit error probability



(b) Large vector dimension

Figure 4.3 A graphical solution to (4.10) for $E'_{\max}$ ($p = 2$, $q = 64$). The solid curves show $E'_{\max}(r)$ for different values of ϵ, and the dashed lines have slope p/k. The two dots on each error exponent curve correspond to r'_1 and r'_2.

For ϵ fixed and k increasing, $\frac{p}{k}r_2'$ (the right-hand side of (4.10)) is decreasing, while $\mathcal{D}_q\left(\mathcal{H}_q^{-1}(1-r_2')/2\|\,\epsilon\right)$ (the left-hand side of (4.10)) is constant. Hence, for k large enough, (4.32) holds, and by the same monotonicity argument used above, $r'_{\max} > r_2'$ (see Figure 4.3(b)). Consequently, it suffices to work with E_{GV} instead of $E'_{\max}$. Thus, we henceforth omit E_{TVZ} from our analysis.

We note that a slightly more complicated differentiable bound relating d/n and r is also known. This bound, called "Vlăduts bound" [sic] in [16], effectively "smoothes the edges" of the maximum of the Gilbert-Varshamov and Tsfasman-Vlădut-Zink bounds. Applying Proposition 2, a "Vlădut error exponent" could also be obtained, but there exists a rate, analogous to r_2' (independent from ϵ and k), beyond which the Vlădut and Gilbert-Varshamov error exponents coincide. Hence, by the same argument given above, it suffices to restrict attention to E_{GV} instead of the Vlădut error exponent.

4.5.1 Small bit error probability

In this section we determine the behavior of the solution to (4.10) for small ϵ, and fixed k and p. First, we set $\delta = \mathcal{H}_q^{-1}(1-r)$ and rewrite the error exponents as

$$E_{\mathrm{sp}}(\delta) = \mathcal{D}_q\left(\delta\|\,\epsilon\right) \qquad \delta \in (\epsilon, 1-q^{-1}) \tag{4.33}$$

$$E_{\mathrm{ex}}(\delta) = \delta \log_q \frac{q-1}{q^{\mathcal{H}_q^{(1/2)}(\epsilon)}-1} \qquad \delta \in \left[1-q^{-\mathcal{H}_q^{(1/2)}(\epsilon)}, 1-q^{-1}\right) \tag{4.34}$$

$$E_{\mathrm{GV}}(\delta) = \mathcal{D}_q\left(\delta/2\|\,\epsilon\right) \qquad \delta \in \left(2\epsilon, 1-q^{-1}\right). \tag{4.35}$$

Next we find a real number δ_X that satisfies

$$E_X(\delta_X) = cr(\delta_X), \tag{4.36}$$

where formally $X \in \{\mathrm{sp}, \mathrm{ex}, \mathrm{GV}\}$,

$$r(\delta) = 1 - \mathcal{H}_q(\delta), \tag{4.37}$$

and $c = p/k$. Then, we obtain the solution to (4.10) by setting

$$r_X^* = r(\delta_X) + o(1), \tag{4.38}$$

where the $o(1)$ term vanishes as $R \to \infty$.

Observe that the sphere packing and Gilbert-Varshamov exponents can both be written as

$$\begin{aligned} E_X(\delta) &= \mathcal{D}_q\left(\delta/i \| \, \epsilon\right) \\ &= \frac{\delta}{i}\log_q \frac{\delta}{i} + \frac{\delta}{i}\log_q 1/\epsilon + (1-\frac{\delta}{i})\log_q(1-\frac{\delta}{i}) - (1-\frac{\delta}{i})\log_q(1-\epsilon), \end{aligned} \tag{4.39}$$

where $i = 1$ when $X = \text{sp}$, and $i = 2$ when $X = \text{GV}$. Using (4.72) and (4.70), the expurgated exponent can be rewritten as

$$\begin{aligned} E_{\text{ex}}(\delta) &= -\delta \log_q\left(2\sqrt{(1-\epsilon)\frac{\epsilon}{q-1}} + (q-2)\frac{\epsilon}{q-1}\right) \\ &= \delta\left[\frac{1}{2}\log_q 1/\epsilon - \log_q\left(2\sqrt{\frac{1-\epsilon}{q-1}} + (q-2)\frac{\sqrt{\epsilon}}{q-1}\right)\right]. \end{aligned} \tag{4.40}$$

Since δ is bounded, the dominant term on the left-hand side of Equation (4.36) (as given in (4.39) and (4.40)) equals $(\delta/i)\log_q 1/\epsilon$ in all three cases, while the right-hand side is bounded between 0 and c, independent of ϵ. Hence, as ϵ approaches zero, for equality to hold in (4.36), δ has to approach zero at least as fast as $\left(\log_q 1/\epsilon\right)^{-1}$. On the other hand, the right-hand side of (4.36) approaches the finite constant c if $\delta \to 0$. Thus, δ cannot converge to zero faster than $\left(\log_q 1/\epsilon\right)^{-1}$ for the left-hand side to stay bounded away from zero. We therefore conclude that the solution to (4.36) must be of the form

$$\delta_X = \frac{ic + \alpha_X}{\log_q 1/\epsilon}, \tag{4.41}$$

where $\alpha_X \to 0$ as $\epsilon \to 0$, and $i = 1$ when $X = \text{sp}$, and $i = 2$ when $X \in \{\text{GV}, \text{ex}\}$. To characterize δ_X more precisely, α_X has to be determined. In what follows all $O(\cdot)$ terms go to zero as $\epsilon \to 0$.

Substituting (4.41) in (4.39), and applying power series expansions yields

$$\begin{aligned} E_X(\alpha_X) = &\frac{ic+\alpha_X}{i\log_q 1/\epsilon}\left[\log_q \frac{ic+\alpha_X}{i\log_q 1/\epsilon} + \log_q 1/\epsilon\right] \\ &+ \left(1 - \frac{ic+\alpha_X}{i\log_q 1/\epsilon}\right)\left[\log_q\left(1 - \frac{ic+\alpha_X}{i\log_q 1/\epsilon}\right) + \log_q(1-\epsilon)\right] \end{aligned}$$

$$
\begin{aligned}
&= c + (\alpha_X/i) - \frac{ic+\alpha_X}{i\log_q 1/\epsilon}\left[\log_q \log_q 1/\epsilon - \log_q\left(c + (\alpha_X/i)\right)\right] \\
&\quad - \left(1 - \frac{ic+\alpha_X}{i\log_q 1/\epsilon}\right)\left[\frac{ic+\alpha_X}{i\log_q 1/\epsilon}\log_q e + O\left(\frac{1}{\log_q^2 1/\epsilon}\right) + O(\epsilon)\right] \\
&= c + (\alpha_X/i) - \frac{ic+\alpha_X}{i\log_q 1/\epsilon}\left[\log_q \log_q 1/\epsilon - \log_q\left(c + (\alpha_X/i)\right) + \log_q e\right] \\
&\quad + O\left(\frac{1}{\log_q^2 1/\epsilon}\right),
\end{aligned}
\tag{4.42}
$$

where $i = 1$ for $X = \text{sp}$, and $i = 2$ for $X = \text{GV}$.

The same steps applied to (4.40) result in

$$
\begin{aligned}
E_{\text{ex}}(\alpha_{\text{ex}}) &= \frac{2c+\alpha_{\text{ex}}}{\log_q 1/\epsilon}\left[\frac{1}{2}\log_q 1/\epsilon - \log_q\left(2\sqrt{\frac{1-\epsilon}{q-1}} + (q-2)\frac{\sqrt{\epsilon}}{q-1}\right)\right] \\
&= c + (\alpha_{\text{ex}}/2) - \frac{2c+\alpha_{\text{ex}}}{\log_q 1/\epsilon}\left[\frac{1}{2}\log_q(q-1) - \log_q 2 + O(\sqrt{\epsilon})\right] \\
&= c + (\alpha_{\text{ex}}/2) - \frac{2c+\alpha_{\text{ex}}}{2\log_q 1/\epsilon}\left[\log_q(q-1) - \log_q 4\right] + O\left(\frac{1}{\log_q^2 1/\epsilon}\right).
\end{aligned}
\tag{4.43}
$$

To obtain the right-hand side of (4.36) as a function of α_X, we write

$$
\begin{aligned}
r_X(\alpha_X) &= 1 - \frac{ic+\alpha_X}{\log_q 1/\epsilon}\left(\log_q(q-1) - \log_q \frac{ic+\alpha_X}{\log_q 1/\epsilon}\right) \\
&\quad + \left(1 - \frac{ic+\alpha_X}{\log_q 1/\epsilon}\right)\log_q\left(1 - \frac{ic+\alpha_X}{\log_q 1/\epsilon}\right) \\
&= 1 - \frac{ic+\alpha_X}{\log_q 1/\epsilon}\left(\log_q(q-1) - \log_q(ic+\alpha_X) + \log_q \log_q 1/\epsilon\right) \\
&\quad - \left(1 - \frac{ic+\alpha_X}{\log_q 1/\epsilon}\right)\frac{ic+\alpha_X}{\log_q 1/\epsilon}\log_q e + O\left(\frac{1}{\log_q^2 1/\epsilon}\right) \\
&= 1 - \frac{ic+\alpha_X}{\log_q 1/\epsilon}\left(\log_q \log_q 1/\epsilon + \log_q(q-1) - \log_q(ic+\alpha_X) + \log_q e\right) \\
&\quad + O\left(\frac{1}{\log_q^2 1/\epsilon}\right).
\end{aligned}
\tag{4.44}
$$

Next, we proceed to solve (4.36) for α_X. Comparing Equations (4.36), (4.42), (4.43), and (4.44), we conclude that $\alpha_X = O\left(\frac{\log_q \log_q 1/\epsilon}{\log_q 1/\epsilon}\right)$. Based on this observation, the

$\log_q(c+(\alpha_X/i))$ terms in (4.42) and (4.44) can be further expanded to obtain

$$
\begin{aligned}
E_X(\alpha_X) &= c+(\alpha_X/i) - \frac{ic+\alpha_X}{i\log_q 1/\epsilon}\Big[\log_q\log_q 1/\epsilon \\
&\qquad -\log_q c - \frac{\alpha_X \log_q e}{ic} + O\left(\alpha_X^2\right) + \log_q e\Big] + O\left(\frac{1}{\log_q^2 1/\epsilon}\right) \\
&= c+(\alpha_X/i) - \frac{c}{\log_q 1/\epsilon}\left[\log_q\log_q 1/\epsilon - \log_q c + \log_q e\right] \\
&\qquad - \frac{\alpha_X}{i\log_q 1/\epsilon}\left[\log_q\log_q 1/\epsilon - \log_q c\right] + O\left(\frac{1}{\log_q^2 1/\epsilon}\right) \\
&= c - \frac{c}{\log_q 1/\epsilon}\left[\log_q\log_q 1/\epsilon - \log_q c + \log_q e\right] \\
&\qquad + (\alpha_X/i)\left(1 - \frac{\log_q\log_q 1/\epsilon - \log_q c}{\log_q 1/\epsilon}\right) + O\left(\frac{1}{\log_q^2 1/\epsilon}\right), \qquad (4.45)
\end{aligned}
$$

and

$$
\begin{aligned}
r_X(\alpha_X) &= 1 - \frac{ic+\alpha_X}{\log_q 1/\epsilon}\left(\log_q\log_q 1/\epsilon - \log_q(ic) - \frac{\alpha_X\log_q e}{ic} + O\left(\alpha_X^2\right) + \log_q e(q-1)\right) \\
&\qquad + O\left(\frac{1}{\log_q^2 1/\epsilon}\right) \\
&= 1 - \frac{ic}{\log_q 1/\epsilon}\left(\log_q\log_q 1/\epsilon - \log_q(ic) + \log_q e(q-1)\right) \\
&\qquad - \frac{\alpha_X}{\log_q 1/\epsilon}\left(\log_q\log_q 1/\epsilon + \log_q(q-1) - \log_q(ic)\right) + O\left(\frac{1}{\log_q^2 1/\epsilon}\right). \qquad (4.46)
\end{aligned}
$$

4.5.1.1 Sphere packing and Gilbert-Varshamov exponents

Substituting (4.45) and (4.46) into (4.36) and rearranging terms yields

$$
\begin{aligned}
0 &= i\left(E_X(\alpha_X) - c r_X(aX)\right) \\
&= \alpha_X\left(1 - \frac{(1-ic)\left(\log_q\log_q 1/\epsilon - \log_q c\right) - ic\left(\log_q(q-1) - \log_q i\right)}{\log_q 1/\epsilon}\right) \\
&\quad - \frac{ic}{\log_q 1/\epsilon}\left[(1-ic)\left(\log_q\log_q 1/\epsilon - \log_q c + \log_q e\right) - ic\left(\log_q(q-1) - \log_q i\right)\right] \\
&\quad + O\left(\frac{1}{\log_q^2 1/\epsilon}\right).
\end{aligned}
$$

Thus,

$$\alpha_X = \frac{ic(1-ic)\left(\log_q \log_q 1/\epsilon - \log_q c + \log_q e\right) - (ic)^2\left(\log_q(q-1) - \log_q i\right)}{\log_q 1/\epsilon - (1-ic)\left(\log_q \log_q 1/\epsilon - \log_q c\right) + ic\left(\log_q(q-1) - \log_q i\right)} + O\left(\frac{1}{\log_q^2 1/\epsilon}\right).$$

Substituting this in (4.46), gives

$$r_X = 1 - \frac{ic\left(\log_q \log_q 1/\epsilon - \log_q ic + \log_q e(q-1)\right)}{\log_q 1/\epsilon - (1-ic)(\log_q \log_q 1/\epsilon - \log_q c) + ic(\log_q(q-1) - \log_q i)} + O\left(\frac{1}{\log_q^2 1/\epsilon}\right).$$

Now, using (4.38) the bounds on the optimal rate are summarized in the following two lemmas.

Lemma 4.5 *For any p and k, and sufficiently small $\epsilon > 0$,*

$$r^*_{\mathrm{sp}} = 1 - \frac{\frac{p}{k}\left(\log_q \log_q 1/\epsilon - \log_q \frac{p}{k} + \log_q e(q-1)\right)}{\log_q 1/\epsilon - (1-\frac{p}{k})(\log_q \log_q 1/\epsilon - \log_q \frac{p}{k}) + \frac{p}{k}\log_q(q-1)} + O\left(\frac{1}{\log_q^2 1/\epsilon}\right) + o(1),$$

where the $O\left(\frac{1}{\log_q^2 1/\epsilon}\right)$ term goes to zero as $\epsilon \to 0$ for any R, and the $o(1)$ term goes to zero as $R \to \infty$.

Lemma 4.6 *For any p and k, and sufficiently small $\epsilon > 0$,*

$$r^*_{\mathrm{GV}} = 1 - \frac{2\frac{p}{k}\left(\log_q \log_q 1/\epsilon - \log_q\left(2\frac{p}{k}\right) + \log_q e(q-1)\right)}{\log_q 1/\epsilon - (1-2\frac{p}{k})(\log_q \log_q 1/\epsilon - \log_q \frac{p}{k}) + 2\frac{p}{k}(\log_q(q-1) - \log_q 2)} + O\left(\frac{1}{\log_q^2 1/\epsilon}\right) + o(1),$$

where the $O\left(\frac{1}{\log_q^2 1/\epsilon}\right)$ term goes to zero as $\epsilon \to 0$ for any R, and the $o(1)$ term goes to zero as $R \to \infty$.

Combining Lemmas 4.5 and 4.6 gives the desired bounds for optimal rate allocation for codes attaining the Gilbert-Varshamov bound, as summarized in the following theorem.

Theorem 4.1 *Consider the cascade of a good k-dimensional vector quantizer, a q-ary linear block channel coder that meets the Gilbert-Varshamov bound or the Tsfasman-Vlădut-Zink bound, and a q-ary symmetric channel with symbol error probability ϵ. The channel code rate r^* that minimizes the p^{th}-power distortion (averaged over all index assignments) satisfies*

$$
\begin{aligned}
&1 - \frac{\frac{2p}{k}\left(\log_q \log_q 1/\epsilon - \log_q \frac{2p}{k} + \log_q e(q-1)\right)}{\log_q 1/\epsilon - (1 - 2\frac{p}{k})(\log_q \log_q 1/\epsilon - \log_q \frac{p}{k}) + 2\frac{p}{k}(\log_q(q-1) - \log_q 2)} \\
&\qquad\qquad + O\left(\frac{1}{\log_q^2 1/\epsilon}\right) + o(1) \\
\leq r^* \leq \\
&1 - \frac{\frac{p}{k}\left(\log_q \log_q 1/\epsilon - \log_q \frac{p}{k} + \log_q e(q-1)\right)}{\log_q 1/\epsilon - (1 - \frac{p}{k})(\log_q \log_q 1/\epsilon - \log_q \frac{p}{k}) + \frac{p}{k}\log_q(q-1)} \\
&\qquad\qquad + O\left(\frac{1}{\log_q^2 1/\epsilon}\right) + o(1),
\end{aligned}
$$

where the $O\left(\frac{1}{\log_q^2 1/\epsilon}\right)$ term goes to zero as $\epsilon \to 0$ for any transmission rate R, and the $o(1)$ term goes to zero as $R \to \infty$.

A crude comparison of the upper and lower bounds in Theorem 4.1 shows a factor of 2 difference in the asymptotically dominant ϵ-dependent term for channel codes that attain the Gilbert-Varshamov bound. The same phenomenon was observed in [2] for efficient binary channel codes. Next, we derive more precise bounds for efficient q-ary channel codes based on the expurgated error exponent, and we present an example comparing the optimal rate allocation bounds for channel codes that achieve the reliability function of the channel and channel codes that attain the Gilbert-Varshamov bound.

4.5.1.2 Expurgated exponent

Substituting (4.43) and (4.46) into (4.36) and rearranging terms yields

$$
\begin{aligned}
0 &= 2\left(E_{\text{ex}}(\alpha_{\text{ex}}) - c r_{\text{ex}}(\alpha_{\text{ex}})\right) \\
&= \alpha_{\text{ex}} \Bigg(1 - \frac{1}{\log_q 1/\epsilon} \Big[\log_q(q-1) - \log_q 4 \\
&\qquad -2c\left(\log_q \log_q 1/\epsilon + \log_q(q-1) - \log_q 2 - \log_q c\right)\Big]\Bigg) \\
&\qquad - \frac{2c}{\log_q 1/\epsilon}\Bigg[\log_q(q-1) - \log_q 4 \\
&\qquad - 2c\left(\log_q \log_q 1/\epsilon + \log_q(q-1) - \log_q 2 - \log_q c + \log_q e\right)\Bigg] \\
&\qquad + O\left(\frac{1}{\log_q^2 1/\epsilon}\right).
\end{aligned}
$$

Thus,

$$
\begin{aligned}
\alpha_{\text{ex}} = 2c&\frac{\log_q(q-1) - \log_q 4 - 2c\left(\log_q \log_q 1/\epsilon + \log_q(q-1) - \log_q 2 - \log_q c + \log_q e\right)}{\log_q 1/\epsilon - \log_q(q-1) + \log_q 4 + 2c\left(\log_q \log_q 1/\epsilon + \log_q(q-1) - \log_q 2 - \log_q c\right)} \\
&+ O\left(\frac{1}{\log_q^2 1/\epsilon}\right),
\end{aligned}
$$

which when substituted into (4.46), gives

$$
\begin{aligned}
r_{\text{ex}} = 1 - &\frac{2c\left(\log_q \log_q 1/\epsilon - \log_q 2c + \log_q e(q-1)\right)}{\log_q 1/\epsilon + 2c(\log_q \log_q 1/\epsilon - \log_q 2c + \log_q(q-1)) - \log_q(q-1) + \log_q 4} \\
&+ O\left(\frac{1}{\log_q^2 1/\epsilon}\right).
\end{aligned}
$$

Now, using (4.38) the bound on the optimal rate is summarized in the following lemma.

Lemma 4.7 *For any p and k, and sufficiently small $\epsilon > 0$,*

$$
\begin{aligned}
r_{\text{ex}}^* = 1 - &\frac{\frac{2p}{k}\left(\log_q \log_q 1/\epsilon - \log_q 2\frac{p}{k} + \log_q e(q-1)\right)}{\log_q 1/\epsilon + 2\frac{p}{k}(\log_q \log_q 1/\epsilon - \log_q 2\frac{p}{k} + \log_q(q-1)) - \log_q(q-1) + \log_q 4} \\
&+ O\left(\frac{1}{\log_q^2 1/\epsilon}\right) + o(1),
\end{aligned}
$$

where the $O\left(\frac{1}{\log_q^2 1/\epsilon}\right)$ *term goes to zero as* $\epsilon \to 0$ *for any* R*, and the* $o(1)$ *term goes to zero as* $R \to \infty$*.*

Combining Lemmas 4.5 and 4.7 establishes bounds for the optimal rate allocation based on "random coding." These extend the results of [2] to q-ary channels.

Theorem 4.2 *Consider the cascade of a good* k*-dimensional vector quantizer, an efficient* q*-ary linear block channel coder, and a* q*-ary symmetric channel with symbol error probability* ϵ*. The channel code rate* r^* *that minimizes the* p^{th}*-power distortion (averaged over all index assignments) satisfies*

$$
\begin{aligned}
& 1 - \frac{\frac{2p}{k}\left(\log_q \log_q 1/\epsilon - \log_q \frac{2p}{k} + \log_q e(q-1)\right)}{\log_q 1/\epsilon + 2\frac{p}{k}(\log_q \log_q 1/\epsilon - \log_q 2\frac{p}{k} + \log_q(q-1)) - \log_q(q-1) + \log_q 4} \\
& \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad + O\left(\frac{1}{\log_q^2 1/\epsilon}\right) + o(1) \\
\leq r^* \leq & \\
& 1 - \frac{\frac{p}{k}\left(\log_q \log_q 1/\epsilon - \log_q \frac{p}{k} + \log_q e(q-1)\right)}{\log_q 1/\epsilon - (1-\frac{p}{k})(\log_q \log_q 1/\epsilon - \log_q \frac{p}{k}) + \frac{p}{k}\log_q(q-1)} \\
& \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad + O\left(\frac{1}{\log_q^2 1/\epsilon}\right) + o(1),
\end{aligned}
$$

where the $O\left(\frac{1}{\log_q^2 1/\epsilon}\right)$ *term goes to zero as* $\epsilon \to 0$ *for any transmission rate* R*, and the* $o(1)$ *term goes to zero as* $R \to \infty$*.*

The dominant ϵ-dependent terms in the upper and lower bounds of Theorem 4.2 differ only by a factor of 2. This was derived in [2] for efficient binary channel codes and observed in Theorem 4.1 for general q-ary channel codes that meet the Gilbert-Varshamov bound. The upper bounds of Theorems 4.1 and 4.2 are identical, since they both derive from the sphere-packing error exponent. The lower bounds, however, are identical only to the precision afforded in [2]. The main rationale for our more complex formulas is to pinpoint the difference between the two categories of channel codes (i.e., those that achieve the reliability function of the channel and those that attain the Gilbert-Varshamov bound). Figure 4.4 compares the rate allocation bounds of Theorems 4.1 and 4.2 for $q = 64$. The choice of alphabet size is motivated by the requirement on algebraic geometry codes needed in order to lie above the Gilbert-Varshamov bound, but the bounds for other values of q display similar behavior. The solid curves correspond to r_{sp}, the upper bound in both Theorem 4.1 and Theorem 4.2. The dotted curves represent $r_{\max}$, the lower bound in Theorem 4.2. The dashed curves show $r'_{\max}$, the lower bound in Theorem 4.1. Thus, the dark shaded regions represent the uncertainty of the bounds of Theorem 4.2, and the corresponding light shaded regions show the discrepancy between the bounds of Theorems 4.1 and 4.2. The curves plotted do not omit any $O(\cdot)$ terms.

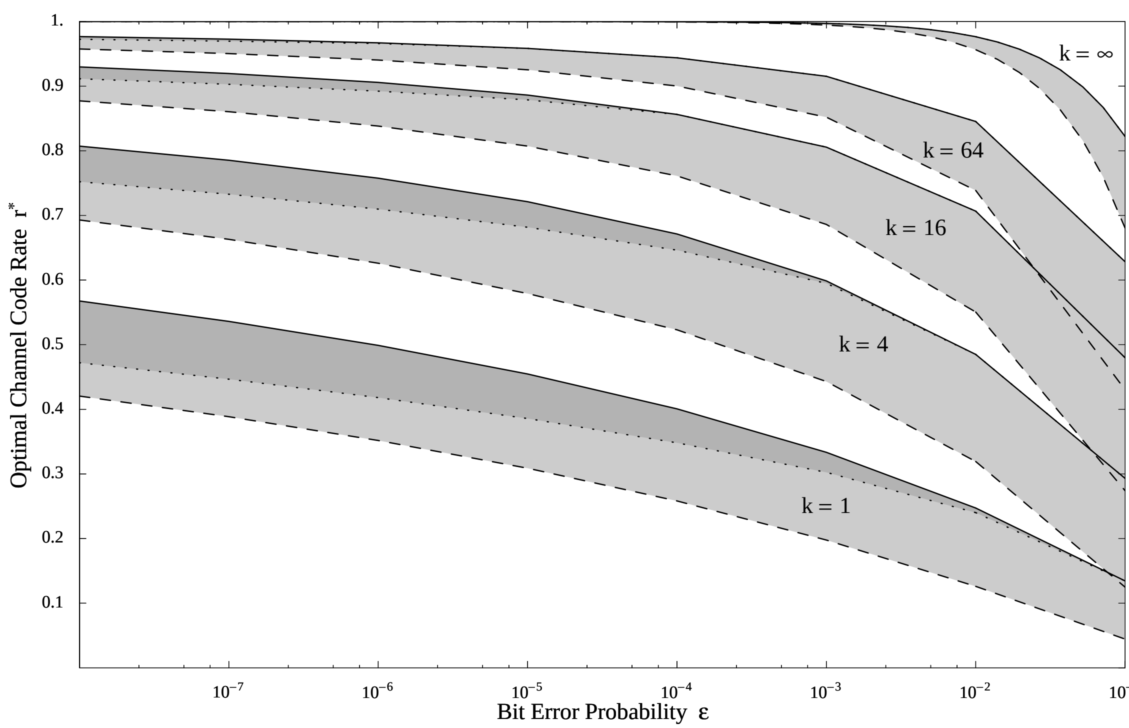


Figure 4.4 A comparison of the upper and lower bounds on the optimal channel code rate as given in Theorems 4.1 and 4.2 for $p = 2$, $q = 64$. For $k = 1, 4, 16, 64$, the solid curves show r_{sp}, the dotted curves $r_{\max}$, and the dashed curves $r'_{\max}$, respectively. For $k = \infty$, $r_{\text{sp}} = r_{\max} = C_q(\epsilon)$ and $r'_{\max} = C_q(2\epsilon)$ are displayed. The corresponding triplets of bounds are indicated by shading. The lightly shaded regions illustrate the penalty for suboptimality.

Note that by substituting $\alpha_X = O\left(\frac{\log_q \log_q 1/\epsilon}{\log_q 1/\epsilon}\right)$ directly into (4.41), simpler expressions for r_X can be obtained at the expense of precision. Figures 4.5 and 4.6 provide additional motivation for the more intricate analysis. The curves obtained numerically without omitting any $O(\cdot)$ terms are denoted by r_X. The approximations (omitting the $O(\cdot)$ terms) given in [2] are denoted by $r_X^{(HZ)}$, and those given in Lemmas 4.5 and 4.7 by $r_X^{(MZ)}$. The expressions we used for $r_X^{(HZ)}$ (see Theorem 1 in [2]) are

$$r_{\text{sp}}^{(HZ)} = 1 - \frac{p}{k}\frac{\log_2 \log_2 1/\epsilon}{\log_2 1/\epsilon} \qquad \text{and} \qquad r_{\text{ex}}^{(HZ)} = 1 - \frac{2p}{k}\frac{\log_2 \log_2 1/\epsilon}{\log_2 1/\epsilon}.$$

The illustrations in Figures 4.5 and 4.6 are for $q = 2$, $p = 2$, and $k = 8$. We note that the curve for r_{sp} is closely approximated by Lemma 4.5. The situation is similar for r_{GV} (not plotted here).

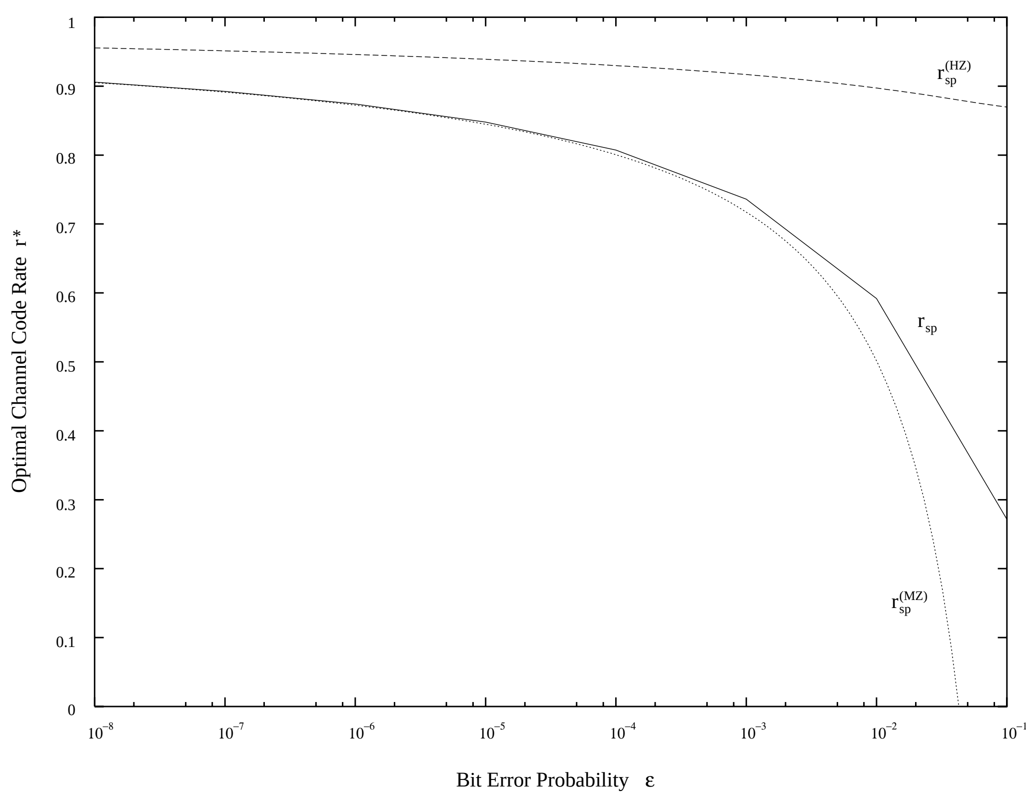


Figure 4.5 Approximations to r_{sp} given by Theorem 1 in [2] (dashed curve $r_{\text{sp}}^{(HZ)}$), and by our Lemma 4.5 (dotted curve $r_{\text{sp}}^{(MZ)}$). The solid curve r_{sp} was obtained by numerical solution of (4.10) for $q = 2$, $p = 2$, and $k = 8$.

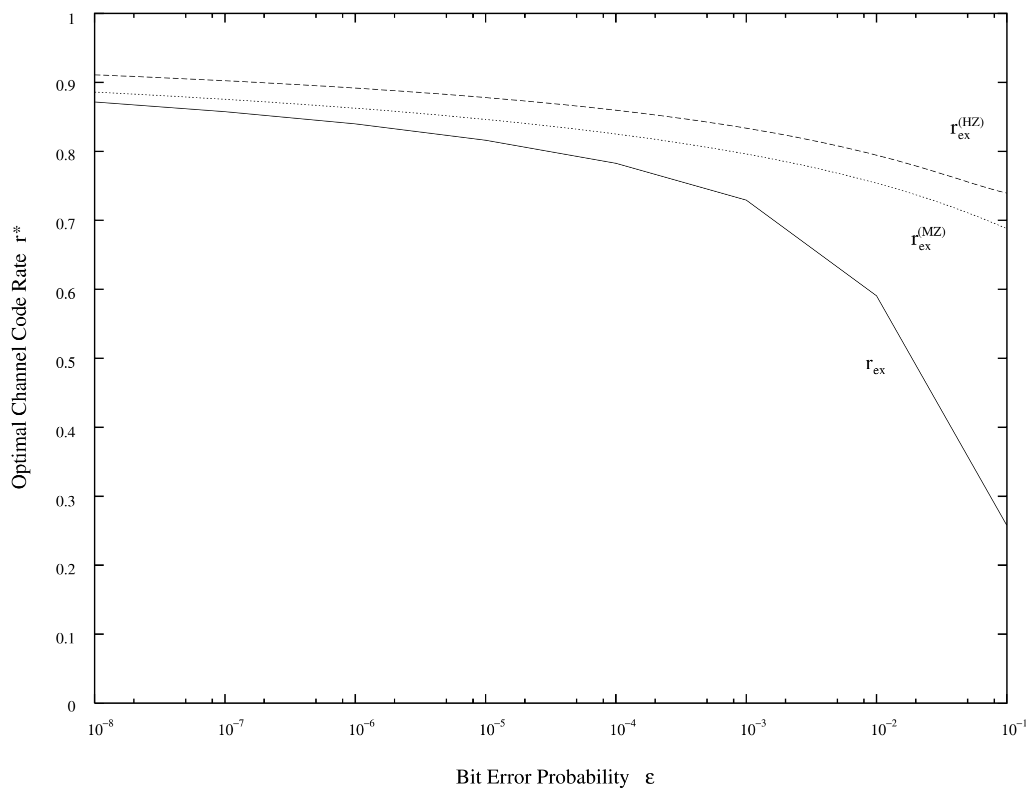


Figure 4.6 Approximations to r_{ex} given by Theorem 1 in [2] (dashed curve $r_{\text{ex}}^{(HZ)}$), and by our Lemma 4.7 (dotted curve $r_{\text{ex}}^{(MZ)}$). The solid curve r_{ex} was obtained by numerical solution of (4.10) for $q = 2$, $p = 2$, and $k = 8$.

4.5.2 Large source vector dimension

In this section, we analyze the optimal rate allocation for large source vector dimensions. As noted earlier, it suffices to examine the sphere packing and Gilbert-Varshamov exponents for k sufficiently large. Solutions of (4.10) based on these two exponents provide upper and lower bounds for the optimal rate allocation for systems using asymptotically good codes that meet the Gilbert-Varshamov bound. The upper and lower bounds on the optimal rate coincide for efficient channel codes, since E_{sp} and E_{rc} are identical for k large. Hence, an exact asymptotic solution of the rate allocation problem is possible, if the channel codes used obey Lemma 4.2. Thus, in what follows we concentrate on the Gilbert-Varshamov case.

We combine (4.23) and (4.30) to obtain

$$E_X(r) = \mathcal{D}_q\left(\frac{1}{i}\mathcal{H}_q^{-1}(1-r)\middle\|\epsilon\right) \qquad r \in (0, C_q(i\epsilon)), \tag{4.47}$$

where $C_q(i\epsilon) = 1 - \mathcal{H}_q(i\epsilon)$, and $i = 1$ when $X = \text{sp}$, and $i = 2$ when $X = \text{GV}$. First, the value of r_X satisfying

$$\mathcal{D}_q\left(\frac{1}{i}\mathcal{H}_q^{-1}(1-r_X)\middle\|\epsilon\right) = \frac{p}{k}r_X \tag{4.48}$$

is found, and then a solution to (4.10) is obtained by setting

$$r_X^* = r_X + o(1), \tag{4.49}$$

where $o(1) \to 0$ as $R \to \infty$.

The error exponents are decreasing functions of the rate and vanish for rates above $C_q(i\epsilon)$. As k increases, the right-hand side of (4.48) decreases, so $r_X \to C_q(i\epsilon)$ as $k \to \infty$. Thus, for large k, $\mathcal{H}_q^{-1}(1-r_X)$ can be approximated by its Taylor series around $1-C_q(i\epsilon)$ as

$$\mathcal{H}_q^{-1}(1-r_X) = \mathcal{H}_q^{-1}(1-C_q(i\epsilon)) + \frac{C_q(i\epsilon) - r_X}{\mathcal{H}_q'(\mathcal{H}_q^{-1}(1-C_q(i\epsilon)))} + O((C_q(i\epsilon)-r_X)^2), \tag{4.50}$$

where $\mathcal{H}_q'$ is the first derivative of the q-ary entropy function, given in (4.15).

Let $(x_i)_k \triangleq C_q(i\epsilon) - r_X$ (note that r_X depends on k via (4.48)). Then $(x_i)_k \geq 0$, and $(x_i)_k \to 0$ as $k \to \infty$. Thus, using $\mathcal{H}_q^{-1}(1-C_q(i\epsilon)) = \mathcal{H}_q^{-1}(\mathcal{H}_q(i\epsilon)) = i\epsilon$ we can rewrite (4.50) as

$$\frac{1}{i}\mathcal{H}_q^{-1}(1-r_X) = \epsilon + \frac{(x_i)_k}{i\mathcal{H}_q'(i\epsilon)} + O((x_i)_k^2)$$

which approaches ϵ, as $k \to \infty$. Applying (4.20), gives

$$\begin{aligned}\mathcal{D}_q\left(\frac{1}{i}\mathcal{H}_q^{-1}(1-r_X)\middle\|\epsilon\right) &= -\frac{\mathcal{H}_q''(\epsilon)}{2}\left(\frac{(x_i)_k}{i\mathcal{H}_q'(i\epsilon)} + O((x_i)_k^2)\right)^2 + O((x_i)_k^3) \\ &= -\frac{\mathcal{H}_q''(\epsilon)(x_i)_k^2}{2i^2\left[\mathcal{H}_q'(i\epsilon)\right]^2} + O((x_i)_k^3),\end{aligned} \tag{4.51}$$

where $\mathcal{H}_q''$ is the second derivative of the q-ary entropy function, given in (4.16). Let $\gamma_i \triangleq \frac{2p\left[i\mathcal{H}_q'(i\epsilon)\right]^2}{-\mathcal{H}_q''(\epsilon)}$. Substituting (4.51) and $r_X = C_q(i\epsilon) - (x_i)_k$ in (4.48), yields

$$(x_i)_k^2 = \frac{\gamma_i}{k}(C_q(i\epsilon) - (x_i)_k) + O((x_i)_k^3). \tag{4.52}$$

The nonnegative root of the quadratic in (4.52) is

$$\begin{aligned}(x_i)_k &= \sqrt{\frac{\gamma_i^2}{4k^2} + \frac{\gamma_i}{k} C_q(i\epsilon) + O((x_i)_k^3)} - \frac{\gamma_i}{2k} \\ &= \sqrt{\frac{\gamma_i C_q(i\epsilon)}{k}} + O\left(\frac{1}{k}\right).\end{aligned}$$

Then, by (4.49) we have

$$r_X^* = C_q(i\epsilon) - \frac{i\mathcal{H}_q(i\epsilon)'}{\sqrt{k}} \left(\frac{2pC_q(i\epsilon)}{-\mathcal{H}_q''(\epsilon)}\right)^{\frac{1}{2}} + O\left(\frac{1}{k}\right) + o(1),$$

and we can state the following theorem.

Theorem 4.3 *Consider the cascade of a good k-dimensional vector quantizer, a q-ary linear block channel code that attains the Gilbert-Varshamov bound or the Tsfasman-Vlădut-Zink bound, and a q-ary symmetric channel with symbol error probability ϵ. The channel code rate r^* that minimizes the p^{th}-power distortion (averaged over all index assignments) satisfies*

$$\begin{aligned} & C_q(2\epsilon) - \frac{2\mathcal{H}_q'(2\epsilon)}{\sqrt{k}} \left(\frac{2pC_q(2\epsilon)}{-\mathcal{H}_q''(\epsilon)}\right)^{\frac{1}{2}} + O\left(\frac{1}{k}\right) + o(1) \\ \leq r^* \leq & \\ & C_q(\epsilon) - \frac{\mathcal{H}_q'(\epsilon)}{\sqrt{k}} \left(\frac{2pC_q(\epsilon)}{-\mathcal{H}_q''(\epsilon)}\right)^{\frac{1}{2}} + O\left(\frac{1}{k}\right) + o(1),\end{aligned}$$

where the $O\left(\frac{1}{k}\right)$ terms approach zero as $k \to \infty$ for any transmission rate R, the $o(1)$ terms approach zero as $R \to \infty$, $\mathcal{H}_q$ is the q-ary entropy function, and $C_q(x)$ is the capacity of a q-ary symmetric channel with symbol error probability x.

Figure 4.7 illustrates the upper and lower bounds of Theorem 4.3. As before, the solid curves represent the upper bound r_{sp}, and the dashed curves show the lower bound $r'_{\max}$. For the sake of comparison we also plotted the corresponding lower bounds for efficient channel codes using dotted curves. As shown in [2] for the binary case, the dotted and solid curves converge for k sufficiently large. The light shading illustrates how the bounds for codes attaining the Gilbert-Varshamov bound compare to the bounds obtained for efficient codes. The curves plotted do not omit any $O(\cdot)$ terms.

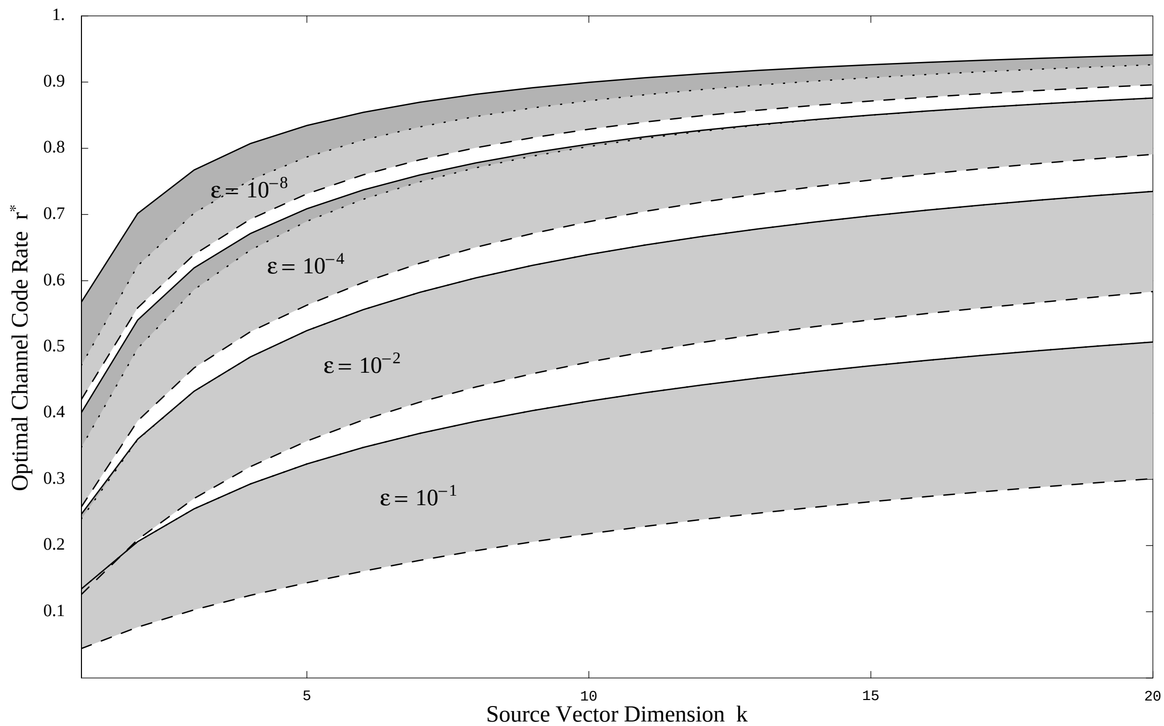


Figure 4.7 Upper and lower bounds on the optimal channel code rate as given in Theorem 4.3 for $p = 2$, $q = 64$ are shown by solid (upper bound, r_{sp}) and dashed (lower bound, $r'_{\max}$) curves for $\epsilon = 10^{-1}, 10^{-2}, 10^{-4}, 10^{-8}$. For comparison, lower bounds ($r_{\max}$) corresponding to efficient channel codes are plotted by dotted curves. The matching triplets of bounds for the same value of ϵ are indicated by shading. The lightly shaded regions illustrate the penalty for suboptimality.

4.6 Conclusion

To determine the optimal tradeoff between source and channel coding for certain structured linear block channel codes, we have derived upper and lower bounds on the channel code rate that minimizes the p^{th}-power distortion of a k-dimensional vector quantizer cascaded with a linear block channel coder on a q-ary symmetric channel. We have presented bounds based on the Gilbert-Varshamov and Tsfasman-Vlădut-Zink bounds as well as random coding arguments for q-ary alphabets. Comparisons of the two types of results were also given.

4.7 Proof of Proposition 1

We use definitions and notation from [14], but we scale the error exponents by a factor of $\ln q$. We denote by $\mathbf{Q} = (Q(0), Q(1), \ldots, Q(q-1))$ the input distribution, and by $P(i|j)$, $i, j \in \{0, 1, \ldots, q-1\}$ the transition probabilities of a q-ary channel.

Random Coding and Sphere Packing Exponents:

Let us slightly reformulate some definitions from [14, pp. 144, 157]:

$$F_0(\rho, \mathbf{Q}) \triangleq \sum_{i=0}^{q-1} \left(\sum_{j=0}^{q-1} Q(j) P(i|j)^{1/(1+\rho)} \right)^{1+\rho} \tag{4.53}$$

$$E_0(\rho, \mathbf{Q}) \triangleq -\log_q F_0(\rho, \mathbf{Q}) \tag{4.54}$$

$$\hat{E}_0(\rho) \triangleq -\rho r + \max_{\mathbf{Q}} E_0(\rho, \mathbf{Q}) \tag{4.55}$$

$$E_{\mathrm{rc}}(r) \triangleq \max_{0 \le \rho \le 1} \hat{E}_0(\rho) \tag{4.56}$$

$$E_{\mathrm{sp}}(r) \triangleq \sup_{\rho > 0} \hat{E}_0(\rho). \tag{4.57}$$

Since the expressions for E_{sp} and E_{rc} differ only in the range of ρ, much of our forthcoming derivation is common to both. Clearly, $\mathrm{argmax}_{\mathbf{Q}}\, E_0(\rho, \mathbf{Q}) = \mathrm{argmin}_{\mathbf{Q}}\, F_0(\rho, \mathbf{Q})$. For a q-ary symmetric channel (see (4.2)),

$$F_0(\rho, \mathbf{Q}) = \sum_{i=0}^{q-1} \left[Q(i)(1-\epsilon)^{1/(1+\rho)} + (1 - Q(i)) \left(\frac{\epsilon}{q-1} \right)^{1/(1+\rho)} \right]^{1+\rho}.$$

Thus, the Jacobian is a diagonal matrix

$$\begin{aligned} &\left[\frac{\partial^2}{\partial Q(i) \partial Q(j)} F_0(\rho, \mathbf{Q}) \right] \\ &\quad = \mathrm{diag} \left\{ \rho(1+\rho) \left[Q(i)(1-\epsilon)^{1/(1+\rho)} + (1 - Q(i)) \left(\frac{\epsilon}{q-1} \right)^{1/(1+\rho)} \right]^{\rho - 1} \right. \\ &\qquad\qquad \left. \times \left((1-\epsilon)^{1/(1+\rho)} - \left(\frac{\epsilon}{q-1} \right)^{1/(1+\rho)} \right)^2 \right\}, \end{aligned}$$

which is positive definite for $\epsilon \in (0, 1-q^{-1})$. Hence, setting

$$\frac{\partial}{\partial Q(j)}\left(F_0(\rho, \mathbf{Q}) - \lambda \sum_{i=0}^{q-1} Q(i)\right) = 0$$

yields a minimum. By symmetry, the minimizing distribution is uniform (this is also easily verified using Theorem 5.6.5. of [14]). Then,

$$\begin{aligned}\hat{E}_0(\rho) &= -\rho r - \log_q q\left[\frac{(1-\epsilon)^{1/(1+\rho)} + (q-1)\left(\epsilon/(q-1)\right)^{1/(1+\rho)}}{q}\right]^{1+\rho} \\ &= \rho(1-r) - (1+\rho)\log_q\left[(1-\epsilon)^{1/(1+\rho)} + (q-1)\left(\epsilon/(q-1)\right)^{1/(1+\rho)}\right], \end{aligned} \tag{4.58}$$

and thus

$$\begin{aligned}\frac{d}{d\rho}\hat{E}_0(\rho) &= 1 - r - \log_q\left[(1-\epsilon)^{1/(1+\rho)} + (q-1)\left(\epsilon/(q-1)\right)^{1/(1+\rho)}\right] \\ &\quad + \frac{(1-\epsilon)^{1/(1+\rho)}\log_q(1-\epsilon)^{1/(1+\rho)} + (q-1)\left(\frac{\epsilon}{q-1}\right)^{1/(1+\rho)}\log_q\left(\frac{\epsilon}{q-1}\right)^{1/(1+\rho)}}{(1-\epsilon)^{1/(1+\rho)} + (q-1)\left(\epsilon/(q-1)\right)^{1/(1+\rho)}} \\ &= 1 - r - \mathcal{H}_q(\delta), \end{aligned} \tag{4.59}$$

where

$$\delta = \frac{(q-1)\left(\frac{\epsilon}{q-1}\right)^{1/(1+\rho)}}{(1-\epsilon)^{1/(1+\rho)} + (q-1)\left(\epsilon/(q-1)\right)^{1/(1+\rho)}}. \tag{4.60}$$

Then, since

$$\begin{aligned}\frac{\partial\delta}{\partial\rho} &= -\frac{1}{(1+\rho)^2}\left\{(q-1)\left(\frac{\epsilon}{q-1}\right)^{1/(1+\rho)}\ln\left(\frac{\epsilon}{q-1}\right)\right. \\ &\quad \times\left[(1-\epsilon)^{1/(1+\rho)} + (q-1)\left(\epsilon/(q-1)\right)^{1/(1+\rho)}\right] - (q-1)\left(\frac{\epsilon}{q-1}\right)^{1/(1+\rho)} \\ &\quad \times\left.\left[(1-\epsilon)^{1/(1+\rho)}\ln(1-\epsilon) + (q-1)\left(\frac{\epsilon}{q-1}\right)^{1/(1+\rho)}\ln\left(\frac{\epsilon}{q-1}\right)\right]\right\} \\ &\quad \times\left[(1-\epsilon)^{1/(1+\rho)} + (q-1)\left(\epsilon/(q-1)\right)^{1/(1+\rho)}\right]^{-2} \\ &= \frac{\delta(1-\delta)}{1+\rho}\left(\log_q(q-1) - \log_q\delta + \log_q(1-\delta)\right)\ln q \\ &= \frac{\delta(1-\delta)}{1+\rho}\mathcal{H}_q'(\delta)\ln q,\end{aligned}$$

we obtain

$$\frac{d^2}{d\rho^2}\hat{E}_0(\rho) = -\mathcal{H}_q'(\delta)\frac{\partial\delta}{\partial\rho}$$
$$= -\frac{\delta(1-\delta)}{1+\rho}\left(\mathcal{H}_q'(\delta)\right)^2 \ln q$$
$$< 0$$

for all values of δ corresponding to $\rho \geq 0$. Thus, the stationary point ρ^*, found by setting $d\hat{E}_0/d\rho = 0$, is a maximum. Instead of solving explicitly for ρ^*, we obtain a parametric expression for the error exponents in terms of $\delta^* = \delta|_{\rho=\rho^*}$. From (4.59), we have

$$r = 1 - \mathcal{H}_q(\delta^*), \tag{4.61}$$

which when substituted into (4.58) gives

$$\begin{aligned}
\hat{E}_0(\rho^*) &= \rho^*\mathcal{H}_q(\delta^*) - (1+\rho^*)\log_q\left[(1-\epsilon)^{1/(1+\rho^*)} + (q-1)\left(\epsilon/(q-1)\right)^{1/(1+\rho^*)}\right] \\
&= -\mathcal{H}_q(\delta^*) + (1+\rho^*)\left\{\delta^*\log_q(q-1) - \delta^*\log_q\delta^* - (1-\delta^*)\log_q(1-\delta^*)\right. \\
&\qquad \left. - \log_q\left[(1-\epsilon)^{1/(1+\rho^*)} + (q-1)\left(\epsilon/(q-1)\right)^{1/(1+\rho^*)}\right]\right\} \\
&= -\mathcal{H}_q(\delta^*) - \delta^*\log_q\left(\frac{\delta^*}{q-1}\left[(1-\epsilon)^{1/(1+\rho^*)} + (q-1)\left(\epsilon/(q-1)\right)^{1/(1+\rho^*)}\right]\right)^{1+\rho^*} \\
&\qquad - (1-\delta^*)\log_q\left((1-\delta^*)\left[(1-\epsilon)^{1/(1+\rho^*)} + (q-1)\left(\epsilon/(q-1)\right)^{1/(1+\rho^*)}\right]\right)^{1+\rho^*} \\
&= \delta^*\log_q\frac{\delta^*}{q-1} + (1-\delta^*)\log_q(1-\delta^*) - \delta^*\log_q\frac{\epsilon}{q-1} - (1-\delta^*)\log_q(1-\epsilon)
\end{aligned} \tag{4.62}$$

$$\begin{aligned}
&= \mathcal{D}_q\left(\delta^* \| \epsilon\right) \\
&= \mathcal{D}_q\left(\mathcal{H}_q^{-1}(1-r) \middle\| \epsilon\right),
\end{aligned} \tag{4.63}$$

where (4.62) follows by (4.60), and (4.61) was used in the last equality. Since $\delta|_{\rho=0} = \epsilon$, $\delta|_{\rho=1} = \frac{\sqrt{\epsilon(q-1)}}{\sqrt{1-\epsilon}+\sqrt{\epsilon(q-1)}}$, and $\delta|_{\rho\to\infty} = 1 - q^{-1}$, (4.63) gives the sphere packing exponent for $r \in (0, C_q)$ and the random coding exponent for $r \in [r_2, C_q)$, where

$$r_2 = 1 - \mathcal{H}_q\left(\frac{\sqrt{\epsilon(q-1)}}{\sqrt{1-\epsilon}+\sqrt{\epsilon(q-1)}}\right). \tag{4.64}$$

As shown in [14], $\rho = 1$ maximizes $\hat{E}_0(\rho)$ for rates less than r_2. Hence, for $r \in (0, r_2]$,

$$\begin{aligned}
E_{\mathrm{rc}}(r) &= \hat{E}_0(1) \\
&= -r + 1 - 2\log_q \left[\sqrt{1-\epsilon} + \sqrt{\epsilon(q-1)}\right] \\
&= C_q^{(1/2)} - r,
\end{aligned}$$

which completes the proof of (4.23) and (4.24). ∎

Expurgated Exponent:

Let us define (see [14, p. 153]):

$$F_x(\rho, \mathbf{Q}) \triangleq \sum_{i=0}^{q-1}\sum_{j=0}^{q-1} Q(i)Q(j)\left[\sum_{l=0}^{q-1}\sqrt{P(l|i)P(l|j)}\right]^{1/\rho} \tag{4.65}$$

$$E_x(\rho, \mathbf{Q}) \triangleq -\rho \log_q F_x(\rho, \mathbf{Q}) \tag{4.66}$$

$$\hat{E}_x(\rho) \triangleq -\rho r + \max_{\mathbf{Q}} E_x(\rho, \mathbf{Q}) \tag{4.67}$$

$$E_{\mathrm{ex}}(r) \triangleq \sup_{\rho \geq 1} \hat{E}_x(\rho). \tag{4.68}$$

Again, $\operatorname{argmax}_{\mathbf{Q}} E_0(\rho, \mathbf{Q}) = \operatorname{argmin}_{\mathbf{Q}} F_0(\rho, \mathbf{Q})$. For a q-ary symmetric channel (see (4.2)),

$$\sum_{l=0}^{q-1}\sqrt{P(l|i)P(l|j)} = \begin{cases} 1 & i = j \\ 2\sqrt{(1-\epsilon)\frac{\epsilon}{q-1}} + (q-2)\frac{\epsilon}{q-1} & i \neq j \end{cases}. \tag{4.69}$$

Let us define

$$\omega \triangleq 2\sqrt{(1-\epsilon)\frac{\epsilon}{q-1}} + (q-2)\frac{\epsilon}{q-1}. \tag{4.70}$$

Alternatively, ω can be expressed in terms of $\mathcal{H}_q^{(1/2)}(\epsilon)$. Using

$$\begin{aligned}
1 + (q-1)\omega &= 1 + 2\sqrt{(1-\epsilon)\epsilon(q-1)} + (q-2)\epsilon \\
&= \left(\sqrt{1-\epsilon} + \sqrt{(q-1)\epsilon}\right)^2 \\
&= q^{\mathcal{H}_q^{(1/2)}(\epsilon)},
\end{aligned} \tag{4.71}$$

we obtain

$$\omega = \frac{q^{\mathcal{H}_q^{(1/2)}(\epsilon)} - 1}{q-1}. \tag{4.72}$$

Using ω, (4.65) can be rewritten as

$$F_x(\rho, \mathbf{Q}) = \sum_{i=0}^{q-1} Q(i)^2 + \omega^{1/\rho}\left(1 - \sum_{i=0}^{q-1} Q(i)^2\right).$$

Thus, the Jacobian is a diagonal matrix (in fact, it is the identity matrix scaled)

$$\left[\frac{\partial^2}{\partial Q(i)\partial Q(j)} F_x(\rho, \mathbf{Q})\right] = 2(1-\omega^{1/\rho})\mathbf{I},$$

which is positive definite for $\epsilon \in (0, 1-q^{-1})$, since (4.72) implies that $\omega < 1$, unless $\epsilon = 1 - q^{-1}$. Hence, setting $\frac{\partial}{\partial Q(j)}\left(F_x(\rho, \mathbf{Q}) - \lambda \sum_{i=0}^{q-1} Q(i)\right) = 0$ yields a minimum. By symmetry, the minimizing distribution is uniform. Thus,

$$\begin{aligned}
\hat{E}_x(\rho) &= -\rho r - \rho \log_q\left[q^{-1} + \omega^{1/\rho}(1-q^{-1})\right] \\
&= \rho\left(1 - r - \log_q\left[1 + (q-1)\omega^{1/\rho}\right]\right),
\end{aligned} \tag{4.73}$$

which implies

$$\begin{aligned}
\frac{d}{d\rho}\hat{E}_x(\rho) &= 1 - r - \log_q\left[1 + (q-1)\omega^{1/\rho}\right] \\
&\quad + \rho \frac{1}{1+(q-1)\omega^{1/\rho}}(q-1)\omega^{1/\rho}\frac{1}{\rho^2}\log_q \omega \\
&= 1 - r + \frac{1}{1+(q-1)\omega^{1/\rho}} \log_q\left(\frac{1}{1+(q-1)\omega^{1/\rho}}\right) \\
&\quad + \frac{(q-1)\omega^{1/\rho}}{1+(q-1)\omega^{1/\rho}} \log_q\left(\frac{\omega^{1/\rho}}{1+(q-1)\omega^{1/\rho}}\right) \\
&= 1 - r - \mathcal{H}_q(\delta),
\end{aligned} \tag{4.74}$$

where

$$\delta = \frac{(q-1)\omega^{1/\rho}}{1+(q-1)\omega^{1/\rho}}. \tag{4.75}$$

Now, since

$$\begin{aligned}
\frac{\partial \delta}{\partial \rho} &= -\left(\frac{1}{\rho^2}\right)\frac{\left((q-1)\omega^{1/\rho}\ln\omega\right)\left[1+(q-1)\omega^{1/\rho}\right] - (q-1)\omega^{1/\rho}\left[(q-1)\omega^{1/\rho}\ln\omega\right]}{\left[1+(q-1)\omega^{1/\rho}\right]^2} \\
&= \frac{\delta(1-\delta)}{\rho}\mathcal{H}_q'(\delta)\ln q,
\end{aligned}$$

we obtain

$$\begin{aligned}\frac{d^2}{d\rho^2}\hat{E}_x(\rho) &= -\mathcal{H}_q'(\delta)\frac{\partial \delta}{\partial \rho} \\ &= -\frac{\delta(1-\delta)}{\rho}\left(\mathcal{H}_q'(\delta)\right)^2 \ln q \\ &< 0\end{aligned}$$

for all values of δ corresponding to $\rho \geq 1$. Thus, the stationary point ρ^*, found by setting $d\hat{E}_x/d\rho = 0$, is a maximum. Instead of solving explicitly for ρ^*, we obtain a parametric expression for the error exponent in terms of $\delta^* = \delta|_{\rho=\rho^*}$. From (4.74), we have

$$r = 1 - \mathcal{H}_q(\delta^*), \tag{4.76}$$

which when substituted into (4.73) gives

$$\begin{aligned}\hat{E}_x(\rho^*) &= \rho^*\left(\mathcal{H}_q(\delta^*) - \log_q\left[1+(q-1)\omega^{1/\rho^*}\right]\right) \\ &= \rho^*\left(\delta^*\log_q(q-1) - \delta^*\log_q\delta^* - (1-\delta^*)\log_q(1-\delta^*) - \log_q\left[1+(q-1)\omega^{1/\rho^*}\right]\right) \\ &= -\delta^*\log_q\left(\frac{\delta^*}{q-1}\left[1+(q-1)\omega^{1/\rho^*}\right]\right)^{\rho^*} \\ &\quad - \rho^*(1-\delta^*)\log_q\left((1-\delta^*)\left[1+(q-1)\omega^{1/\rho^*}\right]\right) \qquad (4.77) \\ &= -\delta^*\log_q\omega \\ &= -\mathcal{H}_q^{-1}(1-r)\log_q\omega, \qquad (4.78)\end{aligned}$$

where (4.77) follows by (4.75), and (4.78) follows by (4.76) and (4.71). Since

$$\delta|_{\rho=1} = 1 - \frac{1}{1+(q-1)\omega} = 1 - q^{-\mathcal{H}_q^{(1/2)}(\epsilon)},$$

and $\delta|_{\rho\to\infty} = 1 - q^{-1}$, (4.78) gives the expurgated exponent for $r \in (0, r_1]$, where

$$r_1 = 1 - \mathcal{H}_q\left(1 - q^{-\mathcal{H}_q^{(1/2)}(\epsilon)}\right). \tag{4.79}$$

For $r > r_1$, $\hat{E}_x(\rho)$ is a decreasing function of ρ, since $d\hat{E}_x/d\rho < 0$ in this case (see (4.74)). Hence, for $r > r_1$,

$$E_{\text{ex}}(r) = \hat{E}_x(1) = 1 - r - \log_q[1+(q-1)\omega] = 1 - r - \mathcal{H}_q^{(1/2)}(\epsilon) = C_q^{(1/2)} - r. \tag{4.80}$$

Combining (4.78) and (4.80), we obtain (4.25). ■

Maximum of the Random Coding and Expurgated Exponents:

It is easy to argue that both

$$E_{\mathrm{rc}}(r) \geq C_q^{(1/2)} - r \qquad \text{and} \qquad E_{\mathrm{ex}}(r) \geq C_q^{(1/2)} - r \tag{4.81}$$

for all rates $r \in (0,1)$, since for every r, the value $\rho = 1$ corresponding to $C_q^{(1/2)} - r$ is in the range of maximization in both (4.56) and (4.68). Hence, provided that $0 \leq r_1 \leq r_2 \leq C_q^{(1/2)} \leq C_q$ holds, (4.26) is obtained. Consulting (4.79), the properties of the entropy function imply that $0 \leq r_1$. To see that $r_1 \leq r_2$, we rewrite equation (4.79) as

$$r_1 = 1 - \mathcal{H}_q\left(1 - \frac{1}{\left(\sqrt{1-\epsilon} + \sqrt{\epsilon(q-1)}\right)^2}\right),$$

and (4.64) as

$$r_2 = 1 - \mathcal{H}_q\left(1 - \frac{\sqrt{1-\epsilon}}{\sqrt{1-\epsilon} + \sqrt{\epsilon(q-1)}}\right).$$

Then $r_1 \leq r_2$ is equivalent to

$$\begin{aligned}
1 &\leq \sqrt{1-\epsilon}\left(\sqrt{1-\epsilon} + \sqrt{\epsilon(q-1)}\right) = 1 - \epsilon + \sqrt{\epsilon(1-\epsilon)(q-1)} \\
\epsilon &\leq (1-\epsilon)(q-1) \\
\epsilon &\leq 1 - q^{-1}.
\end{aligned}$$

Next, we show $r_2 \leq C_q^{(1/2)}$. By the concavity of log, for any $\delta \in [0,1]$,

$$\mathcal{H}_q(\delta) = -\delta \log_q \frac{\delta}{q-1} - (1-\delta)\log_q(1-\delta) \geq -\log_q\left(\frac{1}{q-1}\delta^2 + (1-\delta)^2\right).$$

Substituting this with $\delta = \frac{\sqrt{\epsilon(q-1)}}{\sqrt{1-\epsilon}+\sqrt{\epsilon(q-1)}}$ in (4.64), we can upper bound r_2 as

$$\begin{aligned}
r_2 &\leq 1 + \log_q\left(\frac{1}{q-1}\frac{\epsilon(q-1)}{\left(\sqrt{1-\epsilon} + \sqrt{\epsilon(q-1)}\right)^2} + \frac{1-\epsilon}{\left(\sqrt{1-\epsilon} + \sqrt{\epsilon(q-1)}\right)^2}\right) \\
&= 1 - 2\log_q\left(\sqrt{1-\epsilon} + \sqrt{\epsilon(q-1)}\right) \\
&= C_q^{(1/2)},
\end{aligned}$$

which is what we wanted to prove. The remaining inequality $C_q^{(1/2)} \leq C_q$ can be equivalently stated as $\mathcal{H}_q^{(1/2)}(\epsilon) \geq \mathcal{H}_q(\epsilon)$, which follows by Jensen's inequality. ■

4.8 References

[1] K. Zeger and V. Manzella, "Asymptotic bounds on optimal noisy channel quantization via random coding," *IEEE Trans. Info. Theory*, vol. IT-40, pp. 1926–1938, November 1994.

[2] B. Hochwald and K. Zeger, "Tradeoff between source and channel coding," *IEEE Trans. Info. Theory*, vol. IT-43, pp. 1412–1424, September 1997.

[3] B. Hochwald, "Tradeoff between source and channel coding on a Gaussian channel," to appear in *IEEE Trans. Info. Theory.*

[4] J. N. Pierce, "Limit distribution of the minimum distance of random linear codes," *IEEE Trans. Info. Theory*, vol. IT-13, pp. 595–599, October 1967.

[5] J. T. Coffey and R. M. Goodman, "Any code of which we cannot think is good," *IEEE Trans. Info. Theory*, vol. IT-36, pp. 1453–1461, November 1990.

[6] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes.* Amsterdam, The Netherlands: North-Holland, 1977–1993.

[7] M. A. Tsfasman, S. G. Vlăduţ, and T. Zink, "Modular curves, Shimura curves, and Goppa codes, better than Varshamov-Gilbert bound," *Math. Nachr.*, vol. 109, pp. 21–28, 1982.

[8] G. L. Katsman, M. A. Tsfasman, and S. G. Vlăduţ, "Modular curves and codes with a polynomial construction," *IEEE Trans. Info. Theory*, vol. IT-30, pp. 353–355, March 1984.

[9] C. Voss and T. Høholdt, "An explicit construction of a sequence of codes attaining the Tsfasman-Vlăduţ-Zink bound the first steps," *IEEE Trans. Info. Theory*, vol. IT-43, pp. 128–135, January 1997.

[10] M. A. Tsfasman, "Algebraic-geometric codes and asymptotic problems," *Discrete Applied Mathematics*, vol. 33, pp. 241–256, 1991.

[11] M. A. Tsfasman and S. G. Vlăduţ, *Algebraic-Geometric Codes.* Dordrecht: Kluwer Academic Publishers, 1991.

[12] T. Uyematsu and E. Okamoto, "A construction of codes with exponential error bounds on arbitrary discrete memoryless channels," *IEEE Trans. Info. Theory*, vol. IT-43, pp. 992–996, May 1997.

[13] P. Zador, "Asymptotic quantization error of continuous signals and the quantization dimension," *IEEE Trans. Info. Theory*, vol. IT-28, pp. 139–149, March 1982.

[14] R. G. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.

[15] I. Csiszár, "Generalized cutoff rates and Rényi's information measures," *IEEE Trans. Info. Theory*, vol. IT-41, pp. 26–34, January 1995.

[16] S. N. Litsyn and M. A. Tsfasman, "A note on lower bounds," *IEEE Trans. Info. Theory*, vol. IT-32, pp. 705–706, September 1986.

CHAPTER 5

PERFORMANCE OF QUANTIZERS ON NOISY CHANNELS USING STRUCTURED FAMILIES OF CODES

In this chapter, we derive achievable distortion bounds for the cascade of structured families of binary linear channel codes and binary lattice vector quantizers. It is known that for the cascade of asymptotically good channel codes and asymptotically good vector quantizers the end-to-end distortion decays to zero exponentially fast as a function of the overall transmission rate, and is achieved by choosing a channel code rate that is independent of the overall transmission rate. We show that for certain families of practical channel codes and binary lattice vector quantizers, the overall distortion can still be made to decay to zero exponentially fast as the transmission rate grows, although the exponent is a sublinear function of the transmission rate. This is achieved by carefully choosing a channel code rate that decays to zero as the transmission rate grows. Explicit channel code rate schedules are obtained for several well-known families of channel codes.

5.1 Introduction

Lossy source coding, or quantization, plays an important role in many practical data compression systems such as voice and image transmission devices. The primary mathematical apparatus for obtaining an analytical understanding of the properties of optimal

The material in this chapter has been submitted to the *IEEE Transactions on Information Theory* as: A. Méhes and K. Zeger, "Performance of quantizers on noisy channels using structured families of codes."

quantizers has been the asymptotic theory. Two important types of asymptotic theories exist: (1) fixed transmission rate and growing blocklength; (2) fixed blocklength and growing transmission rate. The first type of asymptotic theory was studied by Shannon [1] and is known as rate-distortion theory. The second type is the study of high-resolution quantization theory [2, 3]. The high-resolution theory indirectly assumes delay and complexity constraints and thus is typically more closely related to practical considerations. The high-resolution results in [2, 3] specifically assume a noiseless channel. In this chapter, we will exploit results from the high-resolution theory to obtain new quantization results for noisy channels.

High-resolution quantization theory for noisy channels gives analytic descriptions of the minimum achievable average distortion, as a function of the transmission rate, the source density, and the vector dimension. For distortion functions which are powers of Euclidean distances and with no channel noise, the minimum average distortion is known to decay to zero exponentially fast as the transmission rate increases [3]. It was shown in [4, 5] that when the source information is transmitted across a noisy channel, the minimum average distortion again decays to zero exponentially fast as the transmission rate increases, although the exponential decay constant is reduced by an amount dependent on how poor the channel is. In fact, the rate of decay of distortion in the noisy channel case is closely related to the optimal allocation of transmission rate between source coding and channel coding (via the channel code rate).

The results in [5] provide mathematical guarantees for a potentially achievable minimum quantizer distortion in the presence of channel noise. However, those results assume the existence of optimal channel codes, namely those described in Shannon's channel coding theorem using random coding arguments. Similar techniques were used to generalize the results of [5] to Gaussian channels [6] and to certain algebraic geometry codes [7]. Hence, the results in [5–7] are existence constructions and do not necessarily correspond to achievable performance based on the best presently known implementable channel codes. There is thus motivation to find a high-resolution theory for quantization with a noisy channel, using families of structured algebraic channel codes.

However, finding such a high-resolution theory appears to be a difficult task for general unstructured source coders, even if the channel coders are structured. In this chapter, we approach the problem by examining systems with structure in both the source coder and channel coder. Such systems are practical to implement and also give insight (via distortion bounds) into the unstructured source coder case.

To illustrate the problem at hand by way of an example, suppose a random variable uniformly distributed on $[0,1]$ is uniform scalar quantized, and transmitted across a binary symmetric channel using a repetition code. For a fixed number of available bits R per transmission, how many times should each information bit be repeated in the repetition code to minimize the end-to-end mean squared error? In other words, what is the optimal rate allocation between source and channel coding? If the channel code rate is decreased, fewer uncorrected bit errors occur but at the expense of coarser quantization, and vice versa if the channel code rate is increased.

A key assumption in [5, 7] is that by keeping the channel code rate fixed (below capacity) while increasing the overall transmission rate R, the probability of decoding error P_e can decay to zero exponentially fast as a function of R. This assumption is valid for "Shannon-optimal" codes and more generally for asymptotically good codes, but most known structured families of channel codes (e.g., Hamming, BCH, Reed-Muller) do not have this property. In the repetition code example, keeping the channel code rate fixed is equivalent to keeping the number of repetitions constant. This in turn implies that the probability of incorrectly decoding an information bit does not change. Therefore, P_e is bounded away from zero, since the probability of decoding error (i.e., an incorrect block) is at least as large as the probability of a single bit error. In this chapter, we investigate the rate allocation problem for structured families of source coders which are asymptotically good and for structured families of channel coders which are not asymptotically good, but which can be used in practice.

A common method for lossy transmission of source data across a noisy channel uses independently designed source coders and channel coders. This follows Shannon's basic "separation principle" in source and channel coding, which is known to be optimal for asymptotically large blocklengths. An important design parameter is the allocation of the available transmission rate between source and channel coding. Tight upper and lower bounds on the optimal tradeoff between source and channel coding are known for certain codes and channels and pth-power distortion measures [4–7]. These results exploit the fact that the distortion contributions of optimal source coding and optimal channel coding decay exponentially fast as functions of the overall transmission rate. The source coder is taken to be a 'good' vector quantizer (one that obeys Zador's decay rate) in [4–7], and index assignment randomization is used. In both [5] and [6], the channel codes are assumed to have exponentially decaying error probabilities achieving the expurgated error exponent for the given channel (a binary symmetric channel in [5] and an additive white Gaussian noise channel in [6]). Although such codes are known to exist, no efficiently decodable ones have yet been discovered. In [7], the results of [5] are extended to q-ary symmetric channels, and a class of asymptotically good channel codes (namely those attaining the Gilbert-Varshamov and Tsfasman-Vlădut-Zink bounds) is examined. Constructions of channel codes better than the Gilbert-Varshamov bound are known [8, 9], but the best known algorithms are not currently practical.

The channel codes considered in [5–7] all have the property that their channel code rates are bounded away from zero for increasing blocklengths. In this chapter, we investigate the tradeoff between source and channel coding for structured classes of codes whose channel code rates approach zero in the limit as the blocklength grows. Hence, we seek a decay schedule of the channel code rate as a function of the overall transmission rate which minimizes the overall distortion. The channel codes we examine are classical binary linear block codes including repetition codes, Reed-Muller codes, and BCH codes. We call (as in Chapter 2) the structured source coders in this chapter binary lattice vector quantizers (BLVQs). Vector quantizers with essentially identical structure have been extensively studied under various different names in [10–15].

The main results of this chapter are collected into Theorem 5.1 in Section 5.4, which gives achievable bounds on the asymptotic mean squared error performance of BLVQs and several useful families of binary linear block channel codes on a binary symmetric channel.

The bounds in Theorem 5.1 show that the minimum distortion with certain structured codes decays to zero as $O\left(2^{-2Rg(R)}\right)$, where $g(R) \to 0$ as $R \to \infty$. The distortion bounds are obtained by choosing $g(R) = O\left(\frac{1}{\sqrt{R}}\right)$ for repetition codes and $g(R) = O\left(\sqrt{\frac{(\log R)^l}{R}}\right)$ for Reed-Muller codes and duals of BCH codes. The constants inside the $O(\cdot)$ depend on the channel noise level. In contrast, for optimal unstructured vector quantizers and no channel noise, $g(R) = 1$ for all R, and for optimal unstructured vector quantizers and optimal channel codes on a noisy channel, $g(R) < 1$ (depending on the channel noise level) and g is bounded away from zero. Since structured source coders are assumed in this chapter, the distortion bounds given are also upper bounds on the distortion using optimal unstructured VQ with the same structured channel codes. In addition, the derivations of the bounds in Theorem 5.1 may be useful tools for future research (e.g., see [16]), since they are not specific to the codes used. Section 5.2 introduces necessary notations, definitions, and lemmas. Section 5.3 gives the framework for the source/channel coding problem and Section 5.4 gives the results of this chapter.

5.2 Preliminaries

As in Chapter 4, for real-valued sequences $f(n)$ and $g(n)$, we write

- $f = O(g)$, if there is a positive real number c, and a positive integer n_0 such that $|f(n)| \leq c|g(n)|$, whenever $n > n_0$;
- $f = o(g)$, if g has only a finite number of zeros, and $f(n)/g(n) \to 0$ as $n \to \infty$.

For any positive integer k, let $\mathbb{Z}_2^k$ denote the field of k-bit binary words. Arithmetic in $\mathbb{Z}_2^k$ is performed modulo 2. Binary k-tuples $i \in \mathbb{Z}_2^k$ will be written as row vectors $i = [i_{k-1}, i_{k-2}, \ldots, i_1, i_0]$, where $i_l \in \{0, 1\}$ denotes the coefficient of 2^l in the binary representation of the corresponding integer i, i.e., $i = \sum_{l=0}^{k-1} i_l 2^l$. We denote by e_l the binary row vector with its only nonzero entry in the lth position, thus $i_l = ie_l^t$. The inner product of two binary vectors $i, j \in \mathbb{Z}_2^k$ is denoted by $ij^t = \sum_{l=0}^{k-1} i_l j_l \in \{0, 1\}$. The

Hamming weight (the number of nonzero bits) of a binary vector $i \in \mathbb{Z}_2^k$ is denoted by $w(i)$.

Euclidean vectors $\mathbf{x} \in \mathbb{R}^d$ will be written as column vectors $\mathbf{x} = (x_1, x_2, \ldots, x_d)^t$. The inner product of two Euclidean vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$ is denoted by $\langle \mathbf{x} \mid \mathbf{y} \rangle = \sum_{l=1}^{d} x_l y_l \in \mathbb{R}$. Also, $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x} \mid \mathbf{x} \rangle}$ denotes the usual Euclidean norm of the vector $\mathbf{x} \in \mathbb{R}^d$. The symbols $I_{\{\cdot\}}$, $\Pr[\cdot]$, and $\mathrm{E}[\cdot]$ are used to denote indicator functions, probabilities, and expectations, respectively.

5.2.1 Entropy and relative entropy

For convenience, we restate the definitions of entropy and relative entropy from Chapter 4 specialized to binary alphabets.

Definition 5.1 Let P and $\tilde{P}$ be probability distributions on a finite set.

The *entropy* of P (in bits) is

$$H(P) = -\sum_x P(x) \log_2 P(x). \tag{5.1}$$

The *relative entropy* between P and $\tilde{P}$ (in bits) is

$$D(P\|\tilde{P}) = \sum_x P(x) \log_2 \frac{P(x)}{\tilde{P}(x)}. \tag{5.2}$$

Definition 5.2 Let $\epsilon, \delta \in [0, 1/2]$, and let $\mathcal{P}_\epsilon$ and $\mathcal{P}_\delta$ be probability distributions on $\{0,1\}$ with $\mathcal{P}_\epsilon(1) = 1 - \mathcal{P}_\epsilon(0) = \epsilon$ and $\mathcal{P}_\delta(1) = 1 - \mathcal{P}_\delta(0) = \delta$.

The *binary entropy function* is

$$h(\epsilon) \triangleq H(\mathcal{P}_\epsilon) = -\epsilon \log_2 \epsilon - (1-\epsilon)\log_2(1-\epsilon), \tag{5.3}$$

and the *binary relative entropy function* (information divergence) is

$$\mathcal{D}_2\left(\delta \| \epsilon\right) \triangleq D(\mathcal{P}_\delta \| \mathcal{P}_\epsilon) = \delta \log_2 \frac{\delta}{\epsilon} + (1-\delta)\log_2 \frac{1-\delta}{1-\epsilon}. \tag{5.4}$$

The following lemma provides a bound on the tail of a binomial distribution.

Lemma 5.1 ([17, p. 531]) For $0 \le \epsilon < \delta \le 1$,

$$\sum_{i=n\delta}^{n} \binom{n}{i} \epsilon^i (1-\epsilon)^{n-i} \le 2^{-n\mathcal{D}_2(\delta\|\epsilon)}.$$

5.2.2 The Hadamard transform

As in Chapters 2 and 3, the Hadamard transform will also be a useful tool in this chapter.

Definition 5.3 For each $i, j \in \mathbb{Z}_2^k$ let $h_{i,j} = (-1)^{ij^t}$ and let $\mathbf{f}\colon \mathbb{Z}_2^k \to \mathbb{R}^d$. The *Hadamard transform* $\hat{\mathbf{f}}\colon \mathbb{Z}_2^k \to \mathbb{R}^d$ of the mapping $\mathbf{f}$ is defined by

$$\hat{\mathbf{f}}(j) = \sum_{i \in \mathbb{Z}_2^k} \mathbf{f}(i) h_{i,j},$$

and the inverse transform is given by

$$\mathbf{f}(i) = 2^{-k} \sum_{j \in \mathbb{Z}_2^k} \hat{\mathbf{f}}(j) h_{j,i}.$$

We refer to the numbers $h_{i,j}$ as Hadamard coefficients. The Hadamard transform is an orthogonal transform equipped with the same convolution and inner product properties (e.g., Parseval's identity) as other Fourier transforms. For easier reference, we restate the Hadamard transform identities seen in Chapters 2 and 3. For $i, j, j' \in \mathbb{Z}_2^k$:

$$\begin{aligned} h_{i,j} &= h_{j,i}, \\ h_{i,j} h_{i,j'} &= h_{i,j+j'}, \\ \sum_{i \in \mathbb{Z}_2^k} h_{i,j} &= 2^k I_{\{j=0\}}. \end{aligned} \tag{5.5}$$

The bits of any binary word $i \in \mathbb{Z}_2^k$ are related to the Hadamard coefficients by

$$1 - h_{i,e_m} = 2 i_m \qquad m \in \{0, 1, \ldots, k-1\}. \tag{5.6}$$

5.2.3 Source coding – vector quantization

Definition 5.4 A d-dimensional, 2^k-point *vector quantizer* (VQ) with *index set* $\mathcal{I} = \{0, \ldots, 2^k - 1\}$, and *codebook* $\mathcal{Y} = \{\mathbf{y}_i \in \mathbb{R}^d \colon i \in \mathcal{I}\}$, is a functional composition $\mathcal{Q}_0 = \mathcal{D}_Q \circ \mathcal{E}_Q$, where $\mathcal{E}_Q\colon \mathbb{R}^d \to \mathcal{I}$ is a *quantizer encoder* and $\mathcal{D}_Q\colon \mathcal{I} \to \mathcal{Y}$ is a *quantizer decoder*. (The subscript 0 denotes association with a noiseless channel.) The elements of the codebook $\mathbf{y}_i \in \mathcal{Y}$ are called *codevectors.* Associated with each codevector $\mathbf{y}_i$ is its *encoder region* $\mathcal{R}_i = \{\mathbf{x} \in \mathbb{R}^d | \mathcal{E}_Q(\mathbf{x}) = i\}$. The set of encoder regions forms a partition of $\mathbb{R}^d$. The source coding *rate* (or resolution) of a vector quantizer is defined as $R_S = k/d$.

The *mean squared error* (or *source distortion*) of a vector quantizer $\mathcal{Q}_0$ for a source random variable $\mathbf{X} \in \mathbb{R}^d$ is

$$\Delta_S = \mathrm{E}\,\|\mathbf{X} - \mathcal{Q}_0(\mathbf{X})\|^2 = \sum_{i \in \mathcal{I}} \int_{\mathcal{R}_i} \|\mathbf{x} - \mathbf{y}_i\|^2 d\mu(\mathbf{x}), \tag{5.7}$$

where μ is the probability distribution of the input $\mathbf{X}$.

Necessary conditions for the optimality of a vector quantizer using the mean squared distortion are (see [18] for example) the *Centroid Condition:*

$$\mathbf{y}_i = \mathrm{E}\,[\mathbf{X}|\mathbf{X} \in \mathcal{R}_i] \qquad \forall i \in \mathcal{I} \tag{5.8}$$

and the *Nearest Neighbor Condition:*

$$\mathcal{R}_i = \{\mathbf{x} \in \mathbb{R}^d: \ \|\mathbf{x} - \mathbf{y}_i\| < \|\mathbf{x} - \mathbf{y}_j\| \ \ \forall j \in \mathcal{I} \setminus \{i\}\} \qquad \forall i \in \mathcal{I}. \tag{5.9}$$

Locally optimal vector quantizers satisfying both necessary conditions (5.8) and (5.9) can be obtained using the Generalized Lloyd Algorithm [18].

The high-resolution (i.e., large R_S) behavior of Δ_S for optimal quantization of a bounded source is described by Zador's formula (also given in Chapter 4), which is stated below in a convenient form for the mean squared error case.

Lemma 5.2 (Zador [3]) *The minimum mean squared error of a rate R_S vector quantizer is asymptotically (as $R_S \to \infty$) given by*

$$\Delta_S = 2^{-2R_S + O(1)}. \tag{5.10}$$

This is often referred to as the "6 dB per bit" rule, since

$$10 \log_{10} \left(2^{-2R_S+O(1)}/2^{-2(R_S+1)+O(1)}\right) \to 20 \log_{10} 2 \approx 6 \text{ dB}.$$

We say that a sequence of quantizers is *asymptotically good* if

$$\limsup_{R_S \to \infty} \Delta_S 2^{2R_S} < \infty. \tag{5.11}$$

We say that a sequence of quantizers is *bounded* if the codepoints of the quantizers are bounded, that is,

$$\sup_k \left(\max_{i \in \mathcal{I}_k} \left\| \mathbf{y}_i^{(k)} \right\| \right) < \infty, \tag{5.12}$$

where $\mathcal{I}_k$ denotes the index set and the $\mathbf{y}_i^{(k)}$ denote the codepoints of the k-bit quantizer in the sequence. Lemma 5.2 shows that optimal quantizers are asymptotically good. In fact, a large class of quantizers including uniform quantizers and other lattice-based vector quantizers are also asymptotically good, although the limit in (5.11) may be larger than for optimal quantizers. Unrestricted optimal quantizers for a bounded source are also bounded, as are large classes of other useful quantizers including truncated lattice VQs, for example.

5.2.3.1 Binary lattice VQ

Definition 5.5 For positive integers d and k, a d-dimensional, 2^k-point *binary lattice vector quantizer* is a vector quantizer with index set $\mathcal{I} = \mathbb{Z}_2^k$, whose codevectors are of the form

$$\mathbf{y}_i = \mathbf{y}_0 + \sum_{l=0}^{k-1} \mathbf{v}_l i_l \qquad \forall i \in \mathbb{Z}_2^k, \tag{5.13}$$

where $\mathbf{y}_0 \in \mathbb{R}^d$ is an *offset vector* and $\{\mathbf{v}_l\}_{l=0}^{k-1} \subset \mathbb{R}^d$ is the set of *generator vectors*, ordered by $\|\mathbf{v}_0\| \leq \|\mathbf{v}_1\| \leq \ldots \leq \|\mathbf{v}_{k-1}\|$.

In this chapter, we focus on BLVQs. There are several equivalent formulations of BLVQ as, for example, truncated lattice VQ, direct sum (or residual) VQ, and VQ by a linear mapping of a (nonredundant) block code. BLVQs can save in memory requirements and encoding complexity. They can also be used for progressive transmission and possess a certain natural robustness to channel noise (see [10] for details).

BLVQs encompass a broad class of useful structured quantizers. For example, a 2^k-level uniform scalar quantizer on the interval (a, b) is a special case of a binary lattice quantizer, obtained by setting $y_0 = a + s/2$ and $v_l = 2^l s$, where $s = (b - a)2^{-k}$ denotes the quantizer stepsize. As a consequence, sequences of asymptotically good BLVQs exist. In fact, for any bounded source, a sequence of increasingly finer (properly truncated and rotated) cubic lattices containing the support of the source is both bounded and asymptotically good. Thus, in what follows, we restrict attention to asymptotically good bounded sequences of BLVQs.

5.2.4 Channel coding – linear codes on a binary symmetric channel

Definition 5.6 A *linear binary* $[n, k, d_{\min}]$ *block channel code* is a linear subspace of $\mathbb{Z}_2^n$ containing 2^k binary n-tuples called *codewords*, each with at least $d_{\min}$ nonzero components. The *channel code rate* is given by $r = k/n$, and the *relative minimum distance* by $\delta = d_{\min}/n$.

Associated with a channel code is a *channel encoder* $\mathcal{E}_C$ and a *channel decoder* $\mathcal{D}_C$. The channel encoder is a one-to-one mapping of messages (e.g., quantizer indices) to channel codewords for transmission. The channel decoder, on the other hand, is a many-to-one mapping. It maps received n-bit blocks (not necessarily codewords) to messages. Let $\mathcal{E}_C(m)$ denote the channel codeword corresponding to message m and $\mathcal{D}_C^{-1}(l)$ the set of n-bit blocks decoded into message l. Then on a binary symmetric channel with crossover probability ϵ, the transition probabilities of the coded channel are

$$p_{l|m} = \sum_{u \in \mathcal{D}_C^{-1}(l)} \epsilon^{w(u+\mathcal{E}_C(m))} \left(1 - \epsilon\right)^{n-w(u+\mathcal{E}_C(m))},$$

and if the code is linear then $p_{l|m} = p_{l+m|0}$. In what follows, let $q_i \triangleq p_{i|0}$ denote the probability that the information error pattern $i \in \mathbb{Z}_2^k$ occurs when an $[n, k]$ linear block code is used to transmit over a binary symmetric channel.

Let $P_l^{(\text{bit})}$ denote the probability that the lth bit of the decoded block is in error and let P_e denote the probability that the decoded block is in error (i.e., at least one of its bits is incorrect). Then, $P_l^{(\text{bit})} = \sum_{i \in \mathbb{Z}_2^k} q_i I_{\{i_l=1\}}$, and $P_e = 1 - q_0$. Let $P_{\max}^{(\text{bit})}$ denote the maximum of the error probabilities for decoded bits. Then $P_{\max}^{(\text{bit})} = \max_l P_l^{(\text{bit})} \leq P_e$.

Since a code with minimum distance $d_{\min}$ can correct all possible $\lfloor \frac{d_{\min}-1}{2} \rfloor$-bit errors and since $\lfloor \frac{d_{\min}-1}{2} \rfloor + 1 \geq d_{\min}/2$, Lemma 5.1 can be used to bound P_e as follows.

Lemma 5.3 *For any $[n, k, d_{\min}]$ linear block channel code and for any $\epsilon \leq d_{\min}/(2n)$, the probability of a block error with maximum likelihood decoding on a binary symmetric channel with bit error probability ϵ satisfies*

$$P_e \leq 2^{-n\mathcal{D}_2\left(\frac{d_{\min}}{2n}\middle\|\epsilon\right)}.$$

To obtain asymptotic results we consider families of $[n, k, d_{\min}]$ linear channel codes indexed by the blocklength n. All families of channel codes fall into exactly one of the following three categories (assuming the limits of $d_{\min}/n$ and k/n exist as $n \to \infty$):

- $\lim_{n\to\infty} \frac{d_{\min}}{n} = 0$

  For codes of this type, the upper bound on the probability of decoding error in Lemma 5.3 becomes trivial as the blocklength increases. The best known families of block channel codes in this category have $k/n \to 1$ as $n \to \infty$. Examples include Hamming codes, families of t-error-correcting binary BCH codes for any fixed t, and lth-order Reed-Muller codes if l is an increasing function of the blocklength. From a source-channel tradeoff perspective, the best codes in these families are those with small blocklengths. Hence, these codes are not relevant to our asymptotic investigations, although their duals are.

- $\lim_{n\to\infty} \frac{d_{\min}}{n} > 0$ and $\lim_{n\to\infty} \frac{k}{n} > 0$

  Families of codes with both their rate and relative minimum distance bounded away from 0 are called *asymptotically good* [19]. Examples include Justesen codes [19, p. 306 ff] and codes satisfying the Zyablov bound [19, p. 315], the Gilbert-Varshamov bound [19, p. 557], and the Tsfasman-Vlădut-Zink bound [20]. Bounds on the asymptotically optimal source/channel rate allocation were derived in Chapter 4 for some of these codes.

- $\lim_{n\to\infty} \frac{d_{\min}}{n} > 0$ and $\lim_{n\to\infty} \frac{k}{n} = 0$
  Codes that fall into this category include repetition codes, lth-order Reed-Muller codes for any fixed order l, t-error-correcting binary BCH codes with $t = O(n)$, and duals of t-error-correcting binary BCH families for any fixed t. Lemma 5.3 guarantees that the probability of decoding error decays to zero exponentially fast for families of this type. Since $k/n \to 0$, relatively less information is transmitted as the blocklength increases, but more reliably.

In this chapter, we focus attention on the third category above. One seeks an optimal "schedule" of the rate k/n converging to 0 as a function of the blocklength n.

5.2.5 The cascaded system

The following definition corresponds to Figure 5.1.

Definition 5.7 A d-dimensional, 2^k-point *noisy channel vector quantizer* with *index set* $\mathbb{Z}_2^k$, *codebook* $\mathcal{Y}$, and with an $[n, k]$ linear channel code $\mathcal{C}$ operating on a binary channel, is a functional composition $\mathcal{Q} = \mathcal{D}_Q \circ \mathcal{D}_C \circ \eta \circ \mathcal{E}_C \circ \mathcal{E}_Q$, where $\mathcal{E}_Q \colon \mathbb{R}^d \to \mathbb{Z}_2^k$ is a quantizer encoder, $\mathcal{D}_Q \colon \mathbb{Z}_2^k \to \mathcal{Y}$ is a quantizer decoder, $\mathcal{E}_C \colon \mathbb{Z}_2^k \to \mathcal{C}$ is a channel encoder, $\mathcal{D}_C \colon \mathbb{Z}_2^n \to \mathbb{Z}_2^k$ is a channel decoder, and $\eta \colon \mathbb{Z}_2^n \to \mathbb{Z}_2^n$ is a random mapping representing a noisy channel.

The mean squared distortion of a noisy channel vector quantizer for a source random variable $\mathbf{X} \in \mathbb{R}^d$ is

$$\Delta = \mathrm{E}\,\|\mathbf{X} - \mathcal{Q}(\mathbf{X})\|^2 = \sum_{i\in\mathbb{Z}_2^k} \sum_{j\in\mathbb{Z}_2^k} q_{i+j} \int_{\mathcal{R}_i} \|\mathbf{x} - \mathbf{y}_j\|^2 d\mu(\mathbf{x}), \tag{5.14}$$

where μ is the probability distribution of the input $\mathbf{X}$, and for $i, j \in \mathbb{Z}_2^k$ the $q_{i+j} = \Pr\left[\mathcal{D}_C(\eta(\mathcal{E}_C(i))) = j\right]$ are the transition probabilities of the coded channel.

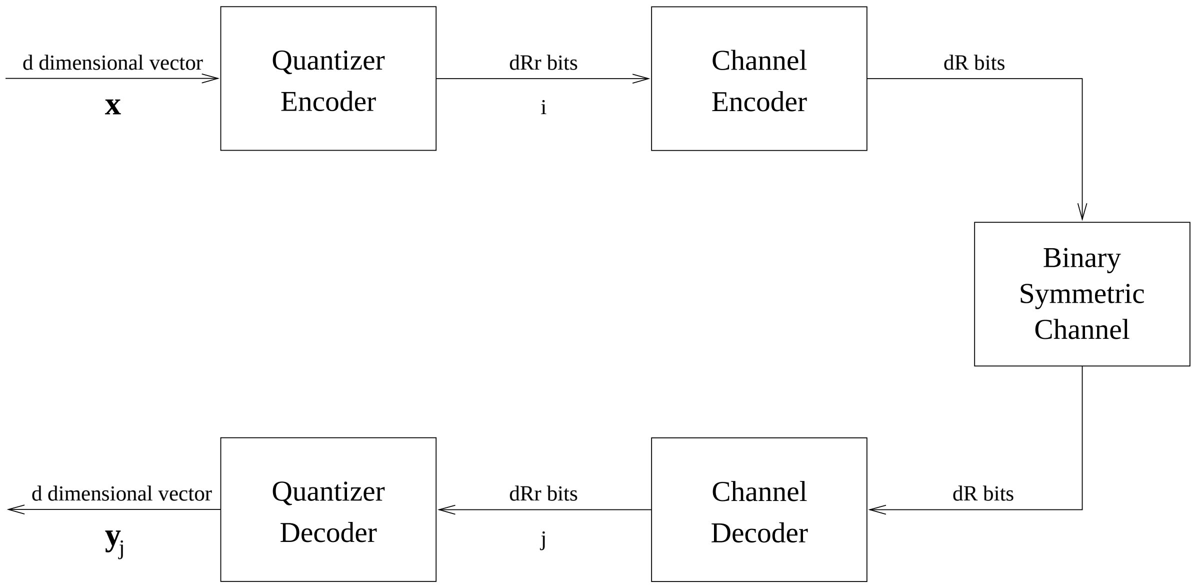


Figure 5.1 Cascaded vector quantizer and channel coder system.

5.3 Rate Allocation Tradeoff

Analogous to the source distortion in (5.7) (i.e., the distortion incurred on a noiseless channel, due to quantization only), we define the *channel distortion* of a noisy channel vector quantizer as

$$\Delta_C \triangleq \mathrm{E} \left\| \mathcal{Q}_0(\mathbf{X}) - \mathcal{Q}(\mathbf{X}) \right\|^2 \tag{5.15}$$

(the component of the distortion influenced by channel errors). If the quantizer $\mathcal{Q}_0$ satisfies the Centroid Condition, then

$$\Delta = \Delta_S + \Delta_C. \tag{5.16}$$

As a function of the overall transmission rate R, both Δ_S and Δ_C decay to zero exponentially fast for optimal quantization of a bounded source and with optimal channel coding. The exact rate of decay is determined by the channel code rate r. An asymptotically optimal channel code rate implies that both terms in (5.16) must decay at the same exponential rate [5].

Structured vector quantizers, however, are often suboptimal. In most cases, the structure dictates the placement of codevectors and the encoding regions are chosen to satisfy the Nearest Neighbor Condition (i.e., the Centroid Condition need not hold). When the codevectors are not the centroids of their respective encoding regions, the Minkowski inequality can be used to bound the distortion as

$$\Delta \leq \left(\sqrt{\Delta_S} + \sqrt{\Delta_C}\right)^2. \tag{5.17}$$

For asymptotically good sequences of BLVQs, the source distortion decays to zero exponentially fast as the source coding rate $R_S \to \infty$. In what follows, we find the asymptotic behavior of the channel distortion for the cascade of BLVQs and practical families of channel codes (which are not asymptotically good), and obtain the channel code rate which asymptotically (in R) minimizes the bound in (5.17) for this system. This is done by equating the exponents of Δ_S and Δ_C. In contrast to [5–7], however, for this system the minimizing channel code rate is a (decreasing) function of the overall transmission rate R.

5.3.1 Rate allocation for BLVQ

Consider a d-dimensional 2^k-point BLVQ cascaded with an $[n, k, d_{\min}]$ binary linear channel code on a binary symmetric channel with an overall transmission rate R. The source coding rate R_S is related to the overall transmission rate R and the channel code rate r by $R_S = Rr$. Each d-dimensional input vector is quantized to $k = dRr$ bits and channel coded with $n = dR$ bits, as shown in Figure 5.1.

For a fixed transmission rate R, increasing the channel code rate results in higher quantizer resolution and a decrease in the BLVQ source distortion Δ_S, but leaves less redundancy to protect against channel errors, which results in an increase in the channel distortion Δ_C. There is thus a tradeoff between source and channel coding governed by the choice of the channel code rate. In order to minimize the right-hand side of (5.17), we seek an exponentially decaying (in R) expression for the channel distortion Δ_C of the cascade of a BLVQ with certain practical channel codes (i.e., with $k/n \to 0$ as $n \to \infty$), and we wish to find the dependence of Δ_C on the channel code rate r.

Lemma 5.4 gives a formula for the channel distortion of a binary lattice quantizer cascaded with the identity index assignment and a linear channel code on a binary symmetric channel. In this chapter, we do not use an explicit index assignment. Instead, the original ordering of the BLVQ codevectors is preserved (the BLVQ basis vectors are ordered by their Euclidean norms). Not using an explicit index assignment is equivalent to specializing the result from Chapter 2 to the case of the Natural Binary Code (identity index assignment).

Lemma 5.4 *Let $\mathbf{X} \in \mathbb{R}^d$ be a source random variable quantized by a 2^k-point binary lattice vector quantizer with generating set $\{\mathbf{v}_l\}_{l=0}^{k-1}$ and transmitted on a binary symmetric channel using the Natural Binary index assignment and an $[n,k]$ binary linear channel code. Let $p_i = \Pr[\mathbf{X} \in \mathcal{R}_i]$ denote the source distribution on the codevectors, and let $q_i = \Pr[\mathcal{D}_C(\eta(\mathcal{E}_C(u))) = u + i]$ denote the transition probabilities of the coded channel. Then, the channel distortion is given by*

$$\Delta_C = \frac{1}{4} \sum_{l=0}^{k-1} \sum_{m=0}^{k-1} \langle \mathbf{v}_l \mid \mathbf{v}_m \rangle \, \hat{p}_{e_l+e_m} \left(\hat{q}_0 - \hat{q}_{e_l} - \hat{q}_{e_m} + \hat{q}_{e_l+e_m} \right), \tag{5.18}$$

where the hats denote Hadamard transforms, and e_l is the binary row vector with its only nonzero entry in the lth position.

Equation (5.18) can be viewed as containing a source-dependent component and a channel-dependent component. We show that the source component is positive and bounded for all transmission rates R and that the channel component can be made to approach zero exponentially fast as $R \to \infty$, and thus the desired bound on Δ_C is obtained.

We first examine the channel-dependent component of (5.18). Using the Hadamard transform definition and its identities gives

$$\hat{q}_0 - \hat{q}_{e_l} - \hat{q}_{e_m} + \hat{q}_{e_l+e_m} = \sum_{i \in \mathbb{Z}_2^k} q_i (1 - h_{i,e_l})(1 - h_{i,e_m})$$

$$
\begin{aligned}
&= 4 \sum_{i \in \mathbb{Z}_2^k} q_i i_l i_m \\
&= 4 \Pr\left[l\text{th and } m\text{th bits both in error}\right] \\
&\leq 4 \min\left(P_l^{(\text{bit})}, P_m^{(\text{bit})}\right) \\
&\leq 4 P_{\max}^{(\text{bit})}. \qquad (5.19)
\end{aligned}
$$

Next, we examine the remaining portion of the sum in (5.18), the source-dependent component. Again using the Hadamard transform definition and its identities, we obtain

$$
\begin{aligned}
\frac{1}{4} \sum_{l=0}^{k-1} \sum_{m=0}^{k-1} \langle \mathbf{v}_l \mid \mathbf{v}_m \rangle \hat{p}_{e_l + e_m} &= \frac{1}{4} \sum_{l=0}^{k-1} \sum_{m=0}^{k-1} \langle \mathbf{v}_l \mid \mathbf{v}_m \rangle \sum_{i \in \mathbb{Z}_2^k} p_i h_{i, e_l + e_m} \\
&= \frac{1}{4} \sum_{i \in \mathbb{Z}_2^k} p_i \left\| \sum_{l=0}^{k-1} \mathbf{v}_l \left(1 - 2 i_l\right) \right\|^2 \\
&= \frac{1}{4} \sum_{i \in \mathbb{Z}_2^k} p_i \left\| \mathbf{y}_{\bar{i}} - \mathbf{y}_i \right\|^2 \qquad (5.20) \\
&\leq \frac{1}{4} \sum_{i \in \mathbb{Z}_2^k} p_i \left(\|\mathbf{y}_{\bar{i}}\| + \|\mathbf{y}_i\| \right)^2 \qquad (5.21) \\
&\leq \rho^2, \qquad (5.22)
\end{aligned}
$$

where $\bar{i}$ is the one's complement of the binary index i (i.e., $\bar{i}_l = 1 - i_l$), and ρ is the radius of some sphere containing every codevector of every quantizer in a sequence of bounded quantizers (independent of the source coding rate) as guaranteed by (5.12).

Combining (5.19) and (5.22), the channel distortion in (5.18) can be upper bounded as

$$
\Delta_C \leq 4 P_{\max}^{(\text{bit})} \rho^2. \qquad (5.23)
$$

It remains to show that $P_{\max}^{(\text{bit})}$, the largest of the error probabilities for a decoded bit, can be made to go to zero exponentially fast as a function of the overall transmission rate R.

We consider a family of $[n, k, d_{\min}]$ channel codes satisfying $\lim_{n \to \infty} k/n = 0$ and $\lim_{n \to \infty} d_{\min}/n > 2\epsilon > 0$, where ϵ is the crossover probability of the underlying binary symmetric channel. We further assume that k is a monotone increasing function of

n, which implies a one-to-one relationship between the channel code rate r and the blocklength n (e.g., this holds for repetition codes and Hamming codes). We divide the Rd bits per sample into blocks of shorter channel codes from the same family of $[n, k, d_{\min}]$ codes, and assume that each has the same blocklength n (a divisor of Rd). Thus, the length Rd channel code is the (Rd/n)-ary Cartesian product of identical codes of length n. This maintains the overall transmission rate of R bits per vector component, and allows a variety of channel code rates r.

This channel coding scheme is not in general optimal, but it provides a conceptually simple means of obtaining *achievable* bounds. Within each n-bit block, the decoding error probability of any given bit is upper-bounded by the decoding error probability of that block. Since the n-bit blocks have identical code parameters, the same bound applies to the decoding error probability of any bit in the overall length Rd code. Then for each n (and consequently, for each corresponding channel code rate r), Lemma 5.3 can be used to upper-bound the largest bit error probability of decoding in the length Rd code using the block error probabilities of the length n constituent codes, namely,

$$P_{\max}^{(\mathrm{bit})} \leq 2^{-n\mathcal{D}_2\left(\frac{d_{\min}}{2n}\middle\|\epsilon\right)}. \tag{5.24}$$

Thus, $P_{\max}^{(\mathrm{bit})}$ can be made to decay to zero exponentially fast in R by choosing the constituent blocklength n to satisfy $n \to \infty$ as $R \to \infty$.

Substituting (5.24) in (5.23) yields

$$\Delta_C \leq 2^{-n\mathcal{D}_2\left(\frac{d_{\min}}{2n}\middle\|\epsilon\right)+O(1)}. \tag{5.25}$$

Combining this with the formula for the source distortion of asymptotically good quantizers in (5.11) and using (5.17), the total distortion is bounded as

$$\Delta \leq \left(2^{-R\frac{k}{n}+O(1)} + 2^{-\frac{n}{2}\mathcal{D}_2\left(\frac{d_{\min}}{2n}\middle\|\epsilon\right)+O(1)}\right)^2. \tag{5.26}$$

The value of the right side of (5.26) for any n that divides Rd represents an achievable distortion, since there exist BLVQs and families of channel codes that satisfy such a bound. In particular, we minimize the right side of (5.26) over n. Let n_R denote a value of n which achieves the minimum. Asymptotically (in R), $n_R \to \infty$ must hold, for otherwise the second term in (5.26) would be bounded away from zero. In fact, to minimize the bound in (5.26) the exponents of the two decaying exponentials have to be asymptotically equal. Since $n_R \to \infty$ as $R \to \infty$ and the families of codes considered satisfy $\lim_{n\to\infty} d_{\min}/n > 2\epsilon$ by assumption, the limit of the information divergence in the exponent of the second term in (5.26) is a finite nonzero constant which we denote by $\beta \triangleq \mathcal{D}_2\left(\frac{1}{2}\lim_{n\to\infty}\frac{d_{\min}}{n}\middle\|\,\epsilon\right)$. Thus, the asymptotically minimizing n_R satisfies

$$\lim_{R\to\infty}\frac{2Rk}{n_R^2\beta} = 1. \tag{5.27}$$

Let r_R denote the channel code rate corresponding to the n_R which solves (5.27). Then by (5.26), the overall distortion vanishes at least as fast as $2^{-2Rr_R+O(1)}$. The next section presents the rate allocations r_R obtained from solutions to (5.27) for various code families.

5.4 Asymptotic Distortion Decay Rates

First, two lemmas are given that solve (5.27) for different dependencies of k on n. Then, the main theorem describing the behavior of several families of codes cascaded with BLVQ follows.

Lemma 5.5 *If $k = cn^\alpha$ for some $c > 0$ and some $\alpha \in [0, 1)$, then*

$$n_R = \left(\frac{2c}{\beta}R\right)^{\frac{1}{2-\alpha}}$$

solves (5.27) *(asymptotically in R), and the corresponding channel code rate is*

$$r_R = c\left(\frac{\beta}{2cR}\right)^{\frac{1-\alpha}{2-\alpha}}.$$

Proof

Lemma 5.5 follows by direct substitution, since

$$\frac{2Rk}{n_R^2\beta} = \left(\frac{2Rc}{\beta}\right)n_R^{\alpha-2} = 1;$$

and

$$r_R = cn_R^{\alpha-1} = c\left(\frac{2c}{\beta}R\right)^{-\frac{1-\alpha}{2-\alpha}}.$$

■

Note that $\alpha = 1$ corresponds to asymptotically good codes (where the optimal rate is asymptotically constant as shown in Chapter 4) and $\alpha = 0$ corresponds to repetition codes. For $\alpha \approx 0$, the channel code rate decays as $r_R = O(\frac{1}{\sqrt{R}})$, in contrast to the case in [5] for Shannon optimal codes where r_R is a positive constant. The distortion decays at least as fast as $O(2^{-2\sqrt{R}})$, in contrast to the $O(2^{-2R})$ Zador rate. Many structured families of codes that satisfy $k/n \to 0$ as $n \to \infty$, however, have a logarithmic dependence between k and n.

Lemma 5.6 *If $k/\left(\log_2 n\right)^l \to c$ as $n \to \infty$ for some finite $c > 0$ and some $l \geq 0$, then*

$$n_R = \sqrt{\frac{2c}{\beta}R\left(\frac{1}{2}\log_2 R\right)^l}$$

satisfies (5.27)*, and the corresponding asymptotic channel code rate is*

$$r_R = \sqrt{\frac{c\beta\left(\frac{1}{2}\log_2 R\right)^l}{2R}}.$$

Proof

Lemma 5.6 follows by direct substitution:

$$\begin{aligned}
\lim_{R\to\infty}\frac{2Rk}{n_R^2\beta} &= \lim_{R\to\infty}\left(\frac{2R}{\beta}\right)\frac{k}{\left(\log_2 n_R\right)^l}\frac{\left(\log_2 n_R\right)^l}{n_R^2} \\
&= \lim_{R\to\infty}\frac{2Rc}{\beta}\frac{\left(\log_2\left[\sqrt{R\frac{2c}{\beta}\left(\frac{1}{2}\log_2 R\right)^l}\right]\right)^l}{\frac{2c}{\beta}R\left(\frac{1}{2}\log_2 R\right)^l} \\
&= \lim_{R\to\infty}\left(\frac{\frac{1}{2}\log_2 R + \log_2\left[\sqrt{\frac{2c}{\beta}\left(\frac{1}{2}\log_2 R\right)^l}\right]}{\frac{1}{2}\log_2 R}\right)^l \\
&= 1;
\end{aligned}$$

and

$$
\begin{aligned}
\lim_{R\to\infty} \frac{k}{n_R r_R} &= \lim_{R\to\infty} \frac{k}{(\log_2 n_R)^l} \frac{(\log_2 n_R)^l}{n_R r_R} \\
&= \lim_{R\to\infty} c \frac{\left(\log_2 \left[\sqrt{R\frac{2c}{\beta}\left(\frac{1}{2}\log_2 R\right)^l}\right]\right)^l}{\sqrt{\frac{2c}{\beta} R \left(\frac{1}{2}\log_2 R\right)^l \frac{c\beta\left(\frac{1}{2}\log_2 R\right)^l}{2R}}} \\
&= \lim_{R\to\infty} \frac{\left(\log_2 \left[\sqrt{R\frac{2c}{\beta}\left(\frac{1}{2}\log_2 R\right)^l}\right]\right)^l}{\left(\frac{1}{2}\log_2 R\right)^l} \\
&= 1.
\end{aligned}
$$

■

The case $l = 0$ corresponds to repetition codes, while larger values of l correspond to more powerful codes (lth-order Reed-Muller codes, for example). For simplex codes ($l = 1$), the channel code rate r_R decays as $O(\sqrt{\frac{\log R}{R}})$. Reed-Muller codes (often punctured) cover a large range of code families. Zeroth-order Reed-Muller codes are themselves repetition codes. Simplex codes (the duals of Hamming codes) are punctured first-order Reed-Muller codes. Punctured Reed-Muller codes are cyclic and as such are related to BCH codes. See [19, p. 384] for the nesting properties of BCH and Reed-Muller codes.

Theorem 5.1 *Let $\mathbf{X} \in \mathbb{R}^d$ be a bounded random variable which is transmitted at a rate R bits per component across a binary symmetric channel with crossover probability ϵ. Suppose the source coder is chosen from a sequence of asymptotically good bounded binary lattice vector quantizers, and the channel coder is chosen from a family of $[n, k, d_{\min}]$ linear block channel codes satisfying $\lim_{n\to\infty} k/n = 0$ and $\lim_{n\to\infty} d_{\min}/n > 2\epsilon$. Then, the overall minimum mean squared error decays (asymptotically in R) at least as fast as*

$$\Delta \leq 2^{-2Rr_R + O(1)}, \tag{5.28}$$

which is achieved by a rate allocation r_R between source and channel coding, for various channel code families as follows:

(*i*) *for a family of* $[n,1,n]$ *repetition codes* $(n \geq 1)$

$$r_R = \sqrt{\frac{-\log_2 2\sqrt{\epsilon(1-\epsilon)}}{2R}}, \qquad \epsilon \in (0,1/2); \tag{5.29}$$

(*ii*) *for a family of* l*th-order* $\left[2^m, \sum_{i=0}^{l}\binom{m}{i}, 2^{m-l}\right]$ *Reed-Muller codes* $(m \geq 1)$

$$r_R = \sqrt{\frac{-\left(\log_2 2^{l+1}\left(\epsilon\left(\frac{1-\epsilon}{2^{l+1}-1}\right)^{2^{l+1}-1}\right)^{\frac{1}{2^{l+1}}}\right)(\log_2 R)^l}{l!\; 2^{l+1}R}}, \qquad \epsilon \in (0,1/2^{l+1}); \tag{5.30}$$

(*iii*) *and for a family of duals of extremal* t*-error-correcting* $[2^m-1, mt, 2^{m-1}-[\log_2(2t-1)]]$ *BCH codes* $(m \geq 1)$

$$r_R = \sqrt{\frac{-t\left(\log_2 4\left(\epsilon\left(\frac{1-\epsilon}{3}\right)^3\right)^{\frac{1}{4}}\right)\log_2 R}{4R}}, \qquad \epsilon \in (0,1/4). \tag{5.31}$$

Proof

The inequality in (5.28) is a direct consequence of (5.26) and the ensuing discussion. The various expressions for r_R are obtained from the solutions n_R of (5.27) as given by Lemma 5.6 (alternatively, Lemma 5.5 for repetition codes) with β substituted using the actual code parameters.

(i) Since $d_{\min}/n = 1$ for repetition codes, $\beta = \mathcal{D}_2\left(\frac{1}{2}\middle\|\,\epsilon\right) = -\log_2(2\sqrt{\epsilon(1-\epsilon)})$. Substituting this in Lemma 5.5 with $\alpha = 0$ and $c = 1$ (or in Lemma 5.6 with $l = 0$ and $c = 1$), the result follows.

(ii) An lth-order length $n = 2^m$ Reed-Muller code has $k = \sum_{i=0}^{l}\binom{m}{i} = \frac{m^l}{l!}(1+o(1))$ information symbols as $n \to \infty$. Hence, Lemma 5.6 can be applied with $c = \frac{1}{l!}$. Substituting $d_{\min}/n = 2^{-l}$ in β yields the desired expression for r_R.

(iii) Often, only bounds are available on the parameters of BCH codes. For simplicity, we assume a family of "extremal" BCH codes at our disposal, which meet these bounds with equality. A t-error-correcting binary BCH code of length $n = 2^m - 1$ has at least $n - mt$ information bits. Thus, its dual has $k \leq mt = t \log_2 n(1 + o(1))$ (which we treat as an equality). This corresponds to Lemma 5.6 with $l = 1$ and $c = t$. By the Carlitz-Uchiyama bound [19, p.280], $\lim_n d_{\min}/n = \frac{1}{2}$. The result then follows by substitution. ■

Figure 5.2 provides an illustration of Theorem 5.1 for the special case of using a uniform scalar quantizer for a uniform source on $(0, 1)$ and a family of repetition codes on a binary symmetric channel with $\epsilon = 10^{-3}$. For each $R = 1, 2, 3, \ldots, 128$, the repetition code with the smallest distortion was found by exhaustive search and the resulting rate was plotted (discrete dots). Since deleting a bit of an even length repetition code results in an odd-length repetition code with the same bit error probability, using the extra bit for source coding always results in a smaller overall distortion. Hence, in addition to the analytic expression for r_R from (5.29) (dashed curve), we also plotted the channel code rate corresponding to the closest odd blocklength (step function).

As with Zador's lemma, Theorem 5.1 also gives a rule of thumb for the expected gain in system performance per bit increase in the overall transmission rate. Unlike on an error-free channel or on a noisy channel using asymptotically good codes (as in [5–7]), however, there is no fixed increase in the signal-to-noise ratio per "bit investment." Instead, the number of "dB's per bit" of performance gain diminishes as the rate R grows. For example, increasing the total transmission rate R by 1 bit per component for a cascaded system using repetition codes yields a signal-to-noise ratio increase of

$$\begin{aligned} \mathrm{SNR}(R+1) - \mathrm{SNR}(R) &= 10 \log_{10} \left(2^{-2\sqrt{R} + O(1)} / 2^{-2\sqrt{R+1} + O(1)} \right) \\ &= 2(\sqrt{R+1} - \sqrt{R}) 10 \log_{10} 2 \\ &\approx \frac{3}{\sqrt{R}} \text{ [dB]}. \end{aligned}$$

However, the bounds presented might be improved in the future.

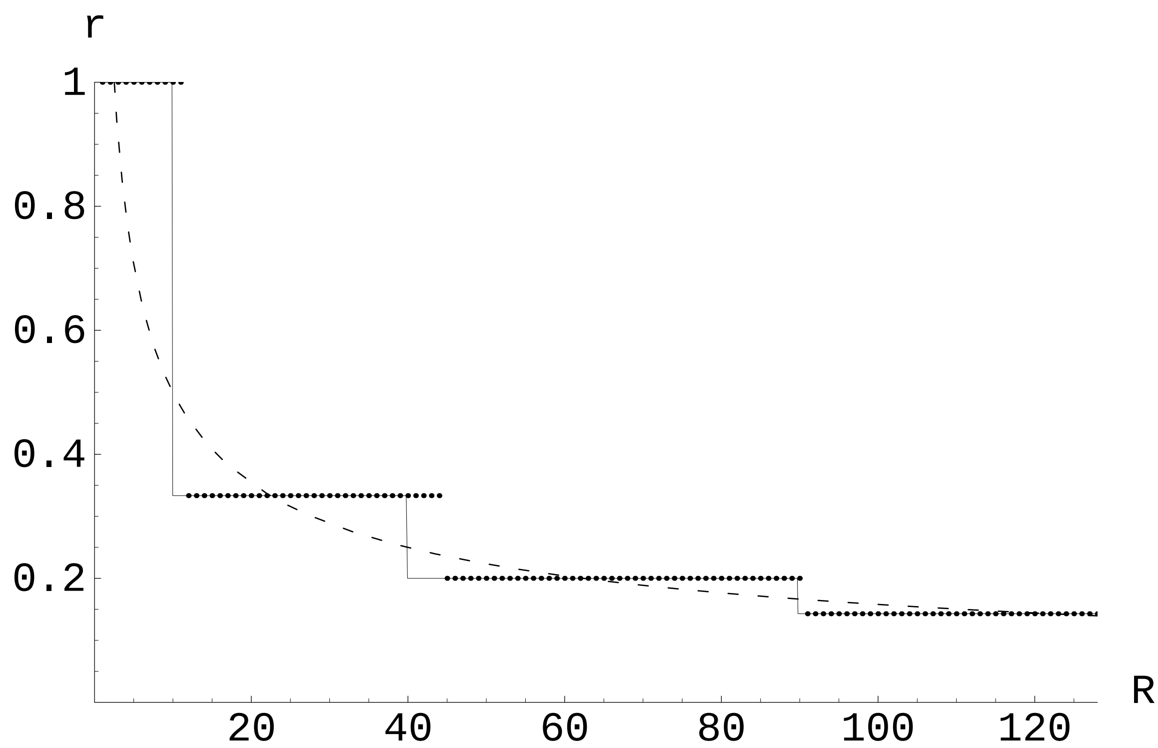


Figure 5.2 An illustration of Theorem 5.1 for uniform scalar quantization of a uniform source on $(0,1)$ using repetition codes to transmit on a binary symmetric channel with $\epsilon = 10^{-3}$. The distortion minimizing channel code rate r is plotted against the overall transmission rate R. The dashed curve is obtained directly from (5.29), the solid-line step function is the closest channel code rate for an odd-length repetition code, and the individual dots represent the rates of the best repetition codes found by exhaustive search.

5.5 Conclusion

This chapter presented bounds on the performance of implementable communication systems as a function of the overall transmission rate R. The systems employ a binary lattice vector quantizer for source coding a bounded random input, and a binary linear channel code for transmission over a binary symmetric channel. The channel code is obtained as a Cartesian product of short codes from channel code families with vanishing rate. Many well studied $[n,k]$ linear channel codes have k proportional to some power of $\log_2 n$. We showed that for such codes, using a rate allocation between source and channel coding of $O(\sqrt{\frac{\log_2^l R}{R}})$ as $R \to \infty$, one gets an asymptotic distortion decay of

$2^{-2\sqrt{R \log_2^l R}}$. Since the exponent is sublinear in R, we see diminishing returns in the per-bit performance increase instead of the usual 6 dB/bit for error-free transmission (or some other constant return for optimal or asymptotically good codes).

5.6 References

[1] C. E. Shannon, "Coding theorems for a discrete source with a fidelity criterion," in *IRE National Convention Record, Part 4*, pp. 142–163, 1959.

[2] W. R. Bennett, "Spectra of quantized signals," *Bell Syst. Tech. J.*, vol. 27, pp. 446–472, July 1948.

[3] P. Zador, "Asymptotic quantization error of continuous signals and the quantization dimension," *IEEE Trans. Info. Theory*, vol. IT-28, pp. 139–149, March 1982.

[4] K. Zeger and V. Manzella, "Asymptotic bounds on optimal noisy channel quantization via random coding," *IEEE Trans. Info. Theory*, vol. IT-40, pp. 1926–1938, November 1994.

[5] B. Hochwald and K. Zeger, "Tradeoff between source and channel coding," *IEEE Trans. Info. Theory*, vol. IT-43, pp. 1412–1424, September 1997.

[6] B. Hochwald, "Tradeoff between source and channel coding on a Gaussian channel," to appear in *IEEE Trans. Info. Theory.*

[7] A. Méhes and K. Zeger, "Source and channel rate allocation for channel codes satisfying the Gilbert-Varshamov or Tsfasman-Vlădut-Zink bounds," submitted to *IEEE Trans. Info. Theory.*

[8] G. L. Katsman, M. A. Tsfasman, and S. G. Vlădut, "Modular curves and codes with a polynomial construction," *IEEE Trans. Info. Theory*, vol. IT-30, pp. 353–355, March 1984.

[9] C. Voss and T. Høholdt, "An explicit construction of a sequence of codes attaining the Tsfasman-Vlădut-Zink bound the first steps," *IEEE Trans. Info. Theory*, vol. IT-43, pp. 128–135, January 1997.

[10] A. Méhes and K. Zeger, "Binary lattice vector quantization with linear block codes and affine index assignments," *IEEE Trans. Info. Theory*, vol. IT-44, pp. 79–95, January 1998.

[11] R. Hagen and P. Hedelin, "Robust vector quantization by linear mappings of block-codes," in *Proc. IEEE Int. Symp. Information Theory*, (San Antonio, TX), p. 171, January 1993.

[12] R. Hagen and P. Hedelin, "Design methods for VQ by linear mappings of block codes," in *Proc. IEEE Int. Symp. Information Theory*, (Trondheim, Norway), p. 241, June 1994.

[13] R. Hagen and P. Hedelin, "Robust vector quantization by a linear mapping of a block-code," to appear in *IEEE Trans. Info. Theory.*

[14] S. W. McLaughlin, D. L. Neuhoff, and J. J. Ashley, "Optimal binary index assignments for a class of equiprobable scalar and vector quantizers," *IEEE Trans. Info. Theory*, vol. IT-41, pp. 2031–2037, November 1995.

[15] A. Méhes and K. Zeger, "Randomly chosen index assignments are asymptotically bad for uniform sources," to appear in *IEEE Trans. Info. Theory.*

[16] P. W. Moo and D. L. Neuhoff, "Polar quantization revisited," submitted to *IEEE Trans. Info. Theory.*

[17] R. G. Gallager, *Information Theory and Reliable Communication.* New York: Wiley, 1968.

[18] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression.* Boston: Kluwer Academic Publishers, 1992.

[19] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes.* Amsterdam, The Netherlands: North-Holland, 1977–1993.

[20] M. A. Tsfasman, S. G. Vlădut, and T. Zink, "Modular curves, Shimura curves, and Goppa codes, better than Varshamov-Gilbert bound," *Math. Nachr.*, vol. 109, pp. 21–28, 1982.

CHAPTER 6

CONCLUSION AND FUTURE DIRECTIONS

This thesis has studied low-complexity techniques for the transmission of source data over a noisy communication channel, subject to a fidelity criterion. The use of structured codes in most components of a digital communication system, including the source coder, the channel coder, and the index assignment, was emphasized throughout the dissertation. This chapter summarizes the work and discusses directions for future research in this area.

The initial part of the thesis treated the index assignment problem. Exact expressions for the mean squared error performance of affine index assignments were derived when used with binary lattice vector quantizers (BLVQs) on binary memoryless channels with or without channel codes for explicit error-control. These results enabled a thorough comparison of the well-known Natural Binary Code, Folded Binary Code, Gray Code, and Two's Complement Code for both symmetric and nonsymmetric channels. No single index assignment was found to be optimal for all channel conditions, but the Two's Complement Code was shown to be superior to the other three families of index assignments for a wide range of channel parameters. In addition, a Worst Code maximizing the mean squared distortion of a uniform quantizer with a uniform source and a binary symmetric channel was derived. Combined with the optimality of the Natural Binary Code under these conditions, the Worst Code enabled a complete description of the range of performances achievable by index assignments. The distortion of a randomly chosen index assignment was also evaluated and found to approach that of the Worst Code for increasing blocklengths. The Worst Code was demonstrated to be affine, and it was shown to maximize the mean squared error among all affine index assignments not only for uniform quantizers, but also for BLVQs. The Hadamard transform was an important tool in obtaining results for the index assignment problem.

The second part of the dissertation was concerned with the problem of rate allocation between source and channel coding for cascaded source/channel coder systems. High-resolution quantization theory was used to obtain rate allocation results for suboptimal structured codes. Upper and lower bounds on a distortion-minimizing channel code rate were derived for the cascade of good vector quantizers, linear block channel coders satisfying the Gilbert-Varshamov or Tsfasman-Vlădut-Zink bounds, and q-ary symmetric channels. The bounds were obtained by balancing the source coding and channel coding error exponents. Analytic expressions were derived for small channel error probabilities and arbitrary vector dimensions, and arbitrary channel error probabilities and large source dimensions. The resulting high-resolution distortion was shown to decay to zero exponentially fast for increasing transmission rates, with the rate of decay bounded above and below by constants dependent on the channel noise level. Similar techniques were used to derive high-resolution distortion bounds for cascaded systems of BLVQs, algebraic codes with asymptotically vanishing channel code rates, and binary symmetric channels. Explicit rate allocations were given for families of repetition codes, BCH codes, and Reed-Muller codes, which showed that the minimum mean squared distortion of these systems can be made to decay to zero asymptotically as the transmission rate grows at least exponentially fast with the square root of the transmission rate.

Optimality conditions for the encoder and for the decoder of BLVQs were also given for both noisy and noiseless discrete memoryless channels and the mean squared error distortion criterion. These provide the update equations for a modified Generalized Lloyd Algorithm to design locally optimal noiseless and channel-optimized BLVQs.

A variety of research ideas can be suggested for future work in the areas investigated in this dissertation. BLVQs, affine index assignments, and binary linear block channel codes share the same basic structure. All three types of codes are affine functions over the binary field. The same approach followed in this thesis can be extended to obtain not only channel coders but also new structured source coders and index assignments over other finite fields.

There are several important open questions regarding channel-optimized vector quantizers. Analytic performance formulas are lacking. Even for the simple case of a uniform source and a binary symmetric channel the minimum achievable distortion using channel-optimized quantizers (as a function of the channel error probability and the quantizer resolution) is presently unknown. There is also a need for a high-resolution theory. In general, analytic results would enable a more thorough evaluation of the tradeoff between performance and complexity for unstructured vs. structured, or channel-optimized vs. cascaded systems. Another long-standing question is to determine under what circumstances channel-optimized vector quantizers are regular. The optimal encoder regions satisfying the weighted nearest neighbor condition are known to be convex, but it is currently unclear what conditions (on the source and/or channel statistics) characterize when the codevectors are contained in their respective encoder cells.

A possible direction of future research is to extend the results of this thesis to include other classes of channel codes. Another well-known question in channel coding is to develop practical constructions and decoding algorithms for codes that meet or exceed the Gilbert-Varshamov bound.

In this thesis, noisy channels are modeled as discrete and memoryless. While the theorems obtained for discrete memoryless channels provide valuable insight, it may be of interest for real-world applications (including mobile/satellite communication, magnetic/laser recording) to generalize the results to different channel models. Several of the tools and ideas presented in the dissertation provide useful guidelines for future research in this area.

APPENDIX A

OPTIMALITY CONDITIONS FOR BINARY LATTICE VECTOR QUANTIZATION

In this appendix, we follow a unified approach to obtain necessary conditions for the mean squared optimality of binary lattice vector quantizers (BLVQs) for both noisy and noiseless discrete memoryless channels. The derivation first treats the familiar case of unstructured vector quantization, and then applies the same steps to obtain optimality conditions for binary lattice vector quantization. The relationships between the first two moments of optimal quantizers and those of the quantized source are also described.

A.1 Preliminaries

The symbols $I_{\{\cdot\}}$, $\Pr[\cdot]$, and $\mathrm{E}[\cdot]$ will be used to denote indicator functions, probabilities, and expectations, respectively. Euclidean vectors $\mathbf{x} \in \mathbb{R}^d$ will be written as column vectors $\mathbf{x} = (x_1, x_2, \ldots, x_d)^t$, where the superscript t denotes transpose. Thus, $\mathbf{x}^t\mathbf{y} = \sum_{l=1}^{d} x_l y_l \in \mathbb{R}$ is the inner product, and $\mathbf{x}\mathbf{y}^t \in \mathbb{R}^{d\times d}$ is the outer product of the vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$. Also, $\|\mathbf{x}\| = \sqrt{\mathbf{x}^t\mathbf{x}}$ denotes the Euclidean norm of $\mathbf{x} \in \mathbb{R}^d$. The trace of a square matrix $\mathrm{A} = [a_{l,m}] \in \mathbb{R}^{d\times d}$ is defined as $\mathrm{trace}\{\mathrm{A}\} = \sum_{l=1}^{d} a_{l,l}$. Thus, $\mathbf{x}^t\mathbf{y} = \mathrm{trace}\{\mathbf{x}\mathbf{y}^t\}$ follows. The following standard result will also be needed.

Lemma A.1 *For positive integers d and n, let $\mathrm{A} \in \mathbb{R}^{d\times n}$. If $\mathrm{trace}\{\mathrm{A}\mathrm{T}^t\} = 0$ for every $\mathrm{T} \in \mathbb{R}^{d\times n}$, then $\mathrm{A} = \mathbf{0}_{d\times n}$, the all-zero matrix in $\mathbb{R}^{d\times n}$.*

Proof

Let $\mathrm{A} = [a_{l,m}]$ and $\mathrm{T} = [t_{l,m}]$. Then by assumption,

$$\mathrm{trace}\left\{\mathrm{AT}^t\right\} = \sum_{l=1}^{d}\sum_{m=1}^{n} a_{l,m}t_{l,m} = 0$$

holds for all choices of the $d \cdot n$ parameters $t_{l,m}$. Setting

$$t_{l,m} = \begin{cases} 1 & \text{if } \ l = i \ \text{ and } \ m = j \\ 0 & \text{otherwise} \end{cases}$$

yields $a_{i,j} = 0$. Since i and j are arbitrary, the statement follows. ■

A.2 Vector Quantization

For convenience, we review the definition and properties of vector quantizers.

Definition A.1 A d-dimensional, 2^k-point *vector quantizer* (VQ) with *index set* $\mathcal{I} = \{0, \dots, 2^k - 1\}$, and *codebook* $\mathcal{Y} = \{\mathbf{y}_i \in \mathbb{R}^d : i \in \mathcal{I}\}$, is a functional composition $\mathcal{Q}_0 = \beta \circ \alpha$, where $\alpha \colon \mathbb{R}^d \to \mathcal{I}$ is a *quantizer encoder* and $\beta \colon \mathcal{I} \to \mathcal{Y}$ is a *quantizer decoder.* (The subscript 0 denotes association with an ideal noiseless channel.)

The elements of the codebook $\mathbf{y}_i \in \mathcal{Y}$ are called *codevectors.* The set of codevectors completely specifies the decoder, since $\beta(i) = \mathbf{y}_i$ for all $i \in \mathcal{I}$. Associated with each index i is the *encoder region* $\mathcal{R}_i = \alpha^{-1}(i) = \{\mathbf{x} \in \mathbb{R}^d | \alpha(\mathbf{x}) = i\}$. The set of encoder regions forms a partition of $\mathbb{R}^d$. By definition, the encoder partition completely specifies the encoder.

The *mean squared error* of a vector quantizer $\mathcal{Q}_0$ for a source random variable $\mathbf{X} \in \mathbb{R}^d$ is given by

$$\Delta = \mathrm{E}\left\|\mathbf{X} - \mathcal{Q}_0(\mathbf{X})\right\|^2 = \sum_{i \in \mathcal{I}} \int_{\mathcal{R}_i} \|\mathbf{x} - \mathbf{y}_i\|^2 d\mu(\mathbf{x}), \tag{A.1}$$

where μ is the probability distribution of the input $\mathbf{X}$.

Necessary conditions for the optimality of a vector quantizer using the mean squared distortion are (see [1] for example) the *Centroid Condition:*

$$\mathbf{y}_i = \mathrm{E}\left[\mathbf{X}|\mathbf{X} \in \mathcal{R}_i\right] \qquad \forall i \in \mathcal{I}, \tag{A.2}$$

which gives the optimal decoder for a fixed encoder, and the *Nearest Neighbor Condition:*

$$\mathcal{R}_i = \{\mathbf{x} \in \mathbb{R}^d \colon \; \|\mathbf{x} - \mathbf{y}_i\|^2 < \|\mathbf{x} - \mathbf{y}_j\|^2 \;\; \forall j \in \mathcal{I} \setminus \{i\}\} \qquad \forall i \in \mathcal{I}, \tag{A.3}$$

which gives the optimal encoder for a fixed decoder. Locally optimal vector quantizers satisfying both necessary conditions (A.2) and (A.3) can be obtained using the Generalized Lloyd Algorithm [1].

A.2.1 Binary lattice VQ

Definition A.2 For positive integers d and k, a d-dimensional, 2^k-point *binary lattice vector quantizer* is a vector quantizer whose codevectors are of the form

$$\mathbf{y}_i = \mathbf{y}_0 + \sum_{l=0}^{k-1} \mathbf{v}_l i_l \qquad \forall i \in \mathcal{I}, \tag{A.4}$$

where $\mathbf{y}_0 \in \mathbb{R}^d$ is an *offset vector*, $\{\mathbf{v}_l\}_{l=0}^{k-1} \subset \mathbb{R}^d$ is a set of *generator vectors*, i_l is the lth bit of the index i (i.e., the coefficient of 2^l in the binary expansion of the integer i), and $\mathcal{I} = \{0, \ldots, 2^k - 1\}$ is the index set.

Since only the decoder of a BLVQ is constrained, the Nearest Neighbor Condition remains unchanged for BLVQ.

A.2.2 Noisy channel VQ

When a d-dimensional, 2^k-point *vector quantizer* as given in Definition A.1 is used on a noisy channel, the defining functional composition becomes $\mathcal{Q}_\eta = \beta \circ \eta \circ \alpha$, where $\eta \colon \mathcal{I} \to \mathcal{I}$ is a probabilistic mapping representing the channel. The mapping η is characterized by the transition probabilities $q_{j|i} \triangleq \Pr\left[\eta(i) = j\right] \;\; \forall i, j \in \mathcal{I}$. An ideal noiseless channel is a special case with $q_{j|i} = I_{\{i=j\}}$ for all $i, j \in \mathcal{I}$.

The *overall mean squared error* of a noisy channel vector quantizer $\mathcal{Q}_\eta$ for a source random variable $\mathbf{X} \in \mathbb{R}^d$ is given by

$$\Delta = \mathrm{E}\left\|\mathbf{X} - \mathcal{Q}_\eta(\mathbf{X})\right\|^2 = \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{I}} q_{j|i} \int_{\mathcal{R}_i} \left\|\mathbf{x} - \mathbf{y}_j\right\|^2 d\mu(\mathbf{x}), \tag{A.5}$$

where μ is the probability distribution of the input $\mathbf{X}$.

Necessary conditions for the optimality of a noisy channel vector quantizer using the mean squared error distortion criterion have been derived by Kumazawa et al. [2]. To simplify notation, for each i let $p_i \triangleq \Pr[\mathbf{X} \in \mathcal{R}_i]$ and $\mathbf{c}_i \triangleq \mathrm{E}[\mathbf{X}|\mathbf{X} \in \mathcal{R}_i]$ respectively denote the probability and the centroid of the ith encoder region. The centroid of a zero probability region is not defined, but the product $\mathbf{c}_i p_i = \int_{\mathcal{R}_i} \mathbf{x} d\mu(\mathbf{x})$ is always well-defined. Also, for each $j \in \mathcal{I}$ let $P_j = \sum_{i\in\mathcal{I}} q_{j|i} p_i$ denote the probability that the index j is received.

If the channel is noisy, the optimal decoder for a fixed encoder is given by the *Weighted Centroid Condition:*

$$\mathbf{y}_j = \frac{1}{P_j}\sum_{i\in\mathcal{I}} q_{j|i} p_i \mathbf{c}_i \qquad \forall j \in \mathcal{I}\colon P_j \neq 0, \tag{A.6}$$

and the optimal encoder for a fixed decoder is given by the *Weighted Nearest Neighbor Condition:*

$$\mathcal{R}_i = \{\mathbf{x} \in \mathbb{R}^d\colon \sum_{j\in\mathcal{I}} q_{j|i} \left\|\mathbf{x} - \mathbf{y}_j\right\|^2 < \sum_{j\in\mathcal{I}} q_{j|i'} \left\|\mathbf{x} - \mathbf{y}_j\right\|^2 \;\; \forall i' \in \mathcal{I}\setminus\{i\}\} \qquad \forall i \in \mathcal{I}. \tag{A.7}$$

The Generalized Lloyd Algorithm can be modified to design locally optimal vector quantizers satisfying both (A.6) and (A.7). The resulting quantizers are called *channel-optimized* vector quantizers.

A.3 Optimality Conditions

Analogous to (A.2) and (A.3), necessary conditions for the mean squared optimality of a class of vector quantizers including BLVQs were given by Hagen and Hedelin in [3]. Their result implicitly assumes an ideal noiseless channel. We derive optimality conditions for BLVQ for both noisy and noiseless discrete memoryless channels. Since only the decoder (i.e., codebook) of a BLVQ is constrained, the nearest neighbor conditions (A.3) and (A.7) remain unaffected. We obtain modified versions of the centroid conditions (A.2) and (A.6) using minimum-variance linear estimation. The development parallels the familiar method of obtaining the centroid condition for unstructured vector quantizers [1] and leads to new insight regarding the properties of locally optimal structured quantizers. The method easily extends to include "VQ by a linear mapping of a block code" (LMBC-VQ), the class of quantizers treated in [3].

A.3.1 Encoder optimality

Since the BLVQ constraint only affects the decoder, the optimal encoder partition for a fixed BLVQ codebook is given by the Nearest Neighbor Condition (A.3) for ideal noiseless channels and by the Weighted Nearest Neighbor Condition (A.7) for noisy channels. Nevertheless, the BLVQ structure allows alternative expressions for these conditions, which provide savings in storage and computation. However, these savings are less significant for channel-optimized BLVQs.

By (A.4), for a BLVQ we have

$$\|\mathbf{x}-\mathbf{y}_j\|^2 = \|\mathbf{x}-\mathbf{y}_0\|^2 - 2\sum_{l=0}^{k-1}(\mathbf{x}-\mathbf{y}_0)^t\mathbf{v}_l j_l + \sum_{l=0}^{k-1}\sum_{m=0}^{k-1}\mathbf{v}_l^t\mathbf{v}_m j_l j_m. \tag{A.8}$$

Thus for each $i \in \mathcal{I}$, the Nearest Neighbor Condition can be rewritten as

$$\mathcal{R}_i = \{\mathbf{x}\in\mathbb{R}^d :\, 2\sum_{l=0}^{k-1}(\mathbf{x}-\mathbf{y}_0)^t\mathbf{v}_l\,(j_l - i_l) < \sum_{l=0}^{k-1}\sum_{m=0}^{k-1}\mathbf{v}_l^t\mathbf{v}_m\,(j_l j_m - i_l i_m) \quad \forall j \in \mathcal{I}\setminus\{i\}\}. \tag{A.9}$$

Since the index bits are readily available, (A.9) suggests that only the offset vector $\mathbf{y}_0$, the set of generator vectors $\{\mathbf{v}_l\}_{l=0}^{k-1}$, and the $\binom{k}{2}$ inner products $\mathbf{v}_l^t\mathbf{v}_m$ have to be stored for a total storage complexity of $O\left(k(d+k)\right)$. This approach may also reduce computational complexity, since the k inner products $(\mathbf{x}-\mathbf{y}_0)^t\mathbf{v}_l$ need not be recomputed for every different j.

Using (A.8), the Weighted Nearest Neighbor Condition can be simplified similarly as

$$\mathcal{R}_i = \{\mathbf{x} \in \mathbb{R}^d :$$
$$2\sum_{l=0}^{k-1}(\mathbf{x}-\mathbf{y}_0)^t\mathbf{v}_l \sum_{j\in\mathcal{I}}\left(q_{j|i'}-q_{j|i}\right)j_l < \sum_{l=0}^{k-1}\sum_{m=0}^{k-1}\mathbf{v}_l^t\mathbf{v}_m \sum_{j\in\mathcal{I}}\left(q_{j|i'}-q_{j|i}\right)j_l j_m \;\; \forall i' \in \mathcal{I}\setminus\{i\}\}, \tag{A.10}$$

for each $i \in \mathcal{I}$. Although (A.10) may also enable savings in both storage and computational complexity, these savings are less significant, because $k2^k$ "weighted bits" $\sum_{j\in\mathcal{I}} q_{j|i} i_l$ and $\binom{k}{2}2^k$ "weighted bit pairs" $\sum_{j\in\mathcal{I}} q_{j|i} i_l i_m$ also have to be stored, in addition to the $O\left(k(d+k)\right)$ storage required for the offset and generator vectors and the inner products of the generator vectors. Thus, the total storage complexity of a channel-optimized BLVQ is $O\left(k(d+k2^k)\right)$, which is already comparable to the $O\left(d2^k\right)$ storage requirements of unstructured quantizers (channel-optimized or not). In fact, whenever $k^2 > d$ it is more memory-efficient to disregard the BLVQ structure, although doing so also eliminates the possibility of savings in computational complexity.

A.3.2 Decoder optimality

We follow a unified approach for both noisy and noiseless channels. To find the optimal decoder for given source random variable $\mathbf{X}$ and a fixed encoder, we first define the random variable $J \triangleq \eta(\alpha(\mathbf{X}))$ representing the received index. (For an ideal noiseless channel, η is the identity mapping and thus $J = \alpha(\mathbf{X})$.) By writing a vector quantizer as a linear estimator of $\mathbf{X}$ based on the observables $\{I_{\{J=j\}}\}$, we obtain the optimal codebook as a minimum-variance linear estimator [4]. We first derive the Centroid Condition and Weighted Centroid Condition for unstructured quantizers, and then follow the same steps to obtain the corresponding optimality conditions for BLVQs.

We define the *selector vector* $\mathbf{S}$ as the 2^k-dimensional random vector whose jth component is the indicator random variable of the received index j:

$$\mathbf{S} \triangleq \left(I_{\{J=0\}}, I_{\{J=1\}}, \ldots, I_{\{J=2^k-1\}}\right)^t,$$

and the *codebook matrix* Y as the $d \times 2^k$ real matrix whose ith column is the ith codevector:

$$\mathrm{Y} \triangleq \begin{bmatrix} | & | & & | \\ \mathbf{y}_0 & \mathbf{y}_1 & \cdots & \mathbf{y}_{2^k-1} \\ | & | & & | \end{bmatrix}.$$

Using this notation, any VQ can be written as

$$\mathcal{Q}(\mathbf{X}) = \mathrm{Y}\mathbf{S}, \tag{A.11}$$

where we omitted the subscripts to emphasize that the same approach is valid for both the noisy and the noiseless channel case. Finding the mean squared optimal codebook for a given partition is equivalent to finding the codebook matrix Y which minimizes the mean squared distortion

$$\Delta = \mathrm{E}\,\|\mathbf{X} - \mathrm{Y}\mathbf{S}\|^2 \tag{A.12}$$

over all *allowable* $d \times 2^n$ matrices. For unconstrained quantizers, Y is a free variable, and any $d \times 2^n$ matrix is allowed. Structured quantizers, however, impose constraints on Y.

The codebook of a BLVQ, for example, is constrained by (A.4). To write the constraint in matrix form, we define the *generator matrix* V of a BLVQ as the $d \times (k+1)$ real matrix whose first column is the offset vector $\mathbf{y}_0$ and the subsequent k columns are the generator vectors $\mathbf{v}_l$ for $l = k-1, \ldots, 0$:

$$\mathrm{V} \triangleq \begin{bmatrix} | & | & & | \\ \mathbf{y}_0 & \mathbf{v}_{k-1} & \cdots & \mathbf{v}_0 \\ | & | & & | \end{bmatrix};$$

and we define the *bit-mapping matrix* B as the $(k+1) \times 2^k$ binary matrix whose ith column is $(1, i_{k-1}, i_{k-2}, \ldots, i_1, i_0)^t$ which gives the bits of i with a leading 1 prepended:

$$\mathrm{B} \triangleq \left[\begin{array}{ccccccc} 1 & 1 & 1 & \cdots & 1 & 1 & 1 \\ \hline 0 & 0 & 0 & & 1 & 1 & 1 \\ & \vdots & & \cdots & & \vdots & \\ 0 & 0 & 1 & & 0 & 1 & 1 \\ 0 & 1 & 0 & & 1 & 0 & 1 \end{array}\right].$$

The codebook constraint in (A.4) can be rewritten as

$$\mathrm{Y} = \mathrm{VB}, \tag{A.13}$$

and (A.11) becomes

$$\mathcal{Q}(\mathbf{X}) = \mathrm{VB}\mathbf{S}. \tag{A.14}$$

Since B is fixed, a BLVQ is completely specified by its generator matrix V. Thus, compared to unstructured VQ the BLVQ constraint reduces the free parameters (i.e., d-dimensional vectors) of allowable codebooks from 2^k to $(k+1)$.[1]

First, we re-derive the centroid conditions for unstructured codebooks, and then we show how the same steps lead to an optimal BLVQ codebook. By the Projection Theorem [4], the codebook matrix Y^* which minimizes the mean squared error in (A.12) over all $d \times 2^k$ matrices Y must satisfy the orthogonality condition

$$\mathrm{E}\left[(\mathbf{X} - \mathrm{Y}^*\mathbf{S})^t\, \mathrm{Y}\mathbf{S}\right] = 0 \qquad \forall \mathrm{Y} \in \mathbb{R}^{d\times 2^k}. \tag{A.15}$$

Equivalently,

$$\mathrm{trace}\left\{\mathrm{E}\left[(\mathbf{X} - \mathrm{Y}^*\mathbf{S})\,\mathbf{S}^t\right]\mathrm{Y}^t\right\} = 0 \qquad \forall \mathrm{Y} \in \mathbb{R}^{d\times 2^k}. \tag{A.16}$$

Hence, by Lemma A.1 we have

$$\mathrm{E}\left[(\mathbf{X} - \mathrm{Y}^*\mathbf{S})\,\mathbf{S}^t\right] = \mathbf{0}_{d\times 2^k}, \tag{A.17}$$

which implies

$$\mathrm{Y}^* = \mathrm{E}\left[\mathbf{X}\mathbf{S}^t\right]\left(\mathrm{E}\left[\mathbf{S}\mathbf{S}^t\right]\right)^{-1}, \tag{A.18}$$

where we assumed that the inverse exists, since otherwise the observables would be linearly dependent, and thus a smaller set of linearly independent observables could be chosen. To evaluate the expectations, we define the following $2^k \times 2^k$ real matrices:

[1] The LMBC-VQs of [3] can be included in the above framework by adding "parity rows" to the bit-mapping matrix B and corresponding generator vectors to the generator matrix V.

- the diagonal matrix of *a priori* probabilities p_i for $i \in \mathcal{I}$

$$\Pi \triangleq \begin{bmatrix} p_0 & & & 0 \\ & p_1 & & \\ & & \ddots & \\ 0 & & & p_{2^k-1} \end{bmatrix},$$

- the matrix of transition probabilities $q_{j|i}$ for $i, j \in \mathcal{I}$

$$\mathrm{Q} \triangleq \begin{bmatrix} q_{0|0} & q_{1|0} & \cdots & q_{2^k-1|0} \\ q_{0|1} & q_{1|1} & \cdots & q_{2^k-1|1} \\ \vdots & \vdots & \ddots & \vdots \\ q_{0|2^k-1} & q_{1|2^k-1} & \cdots & q_{2^k-1|2^k-1} \end{bmatrix},$$

- and the diagonal matrix of *a posteriori* probabilities P_i for $i \in \mathcal{I}$

$$\tilde{\Pi} \triangleq \begin{bmatrix} P_0 & & & 0 \\ & P_1 & & \\ & & \ddots & \\ 0 & & & P_{2^k-1} \end{bmatrix}.$$

In addition, we define the $d \times 2^k$ matrix of centroids

$$\mathrm{C} \triangleq \begin{bmatrix} | & | & & | \\ \mathbf{c}_0 & \mathbf{c}_1 & \cdots & \mathbf{c}_{2^k-1} \\ | & | & & | \end{bmatrix}.$$

The expected values in (A.18) can then be written as

$$\mathrm{E}\left[\mathbf{X}\mathbf{S}^t\right] = \mathrm{C}\Pi\mathrm{Q} \qquad \text{and} \qquad \mathrm{E}\left[\mathbf{S}\mathbf{S}^t\right] = \tilde{\Pi}. \tag{A.19}$$

Substituting these in (A.18) yields

$$\mathrm{Y}^* = \mathrm{C}\Pi\mathrm{Q}\tilde{\Pi}^{-1}, \tag{A.20}$$

which is the Weighted Centroid Condition for unstructured VQ in matrix form. For an ideal noiseless channel, we have

$$\mathrm{Q} = \mathrm{I} \qquad \text{and} \qquad \tilde{\Pi} = \Pi, \tag{A.21}$$

where I is the identity matrix. Hence, (A.20) simplifies to

$$\mathrm{Y}^* = \mathrm{C}\Pi\mathrm{I}\Pi^{-1} = \mathrm{C}, \tag{A.22}$$

the usual Centroid Condition for unstructured VQ.

We now derive the corresponding conditions for BLVQ. Equations (A.24), (A.25), (A.26), and (A.27) below follow the steps taken in (A.15), (A.16), (A.17), and (A.18), respectively. Finding the mean squared optimal BLVQ codebook for a given encoder is equivalent to finding the generator matrix V which minimizes the mean squared distortion

$$\Delta = \mathrm{E}\left\|\mathbf{X} - \mathrm{VB}\mathbf{S}\right\|^2, \tag{A.23}$$

over all $d \times (n+1)$ matrices. Although the selector vector $\mathbf{S}$ still describes the complete event space, imposing the BLVQ constraint effectively restricts the set of observables to $\mathrm{B}\mathbf{S}$. By the Projection Theorem, the generator matrix V^* which minimizes the mean squared error in (A.23) over all $d \times (k+1)$ matrices V must satisfy the orthogonality condition

$$\mathrm{E}\left[(\mathbf{X} - \mathrm{V}^*\mathrm{B}\mathbf{S})^t\, \mathrm{VB}\mathbf{S}\right] = 0 \qquad \forall \mathrm{V} \in \mathbb{R}^{d\times(k+1)}. \tag{A.24}$$

Equivalently,

$$\operatorname{trace}\left\{\mathrm{E}\left[(\mathbf{X} - \mathrm{V}^*\mathrm{B}\mathbf{S})\,\mathbf{S}^t\right]\mathrm{B}^t\mathrm{V}^t\right\} = 0 \qquad \forall \mathrm{V} \in \mathbb{R}^{d\times(k+1)}. \tag{A.25}$$

Hence, by Lemma A.1 we have

$$\mathrm{E}\left[(\mathbf{X} - \mathrm{V}^*\mathrm{B}\mathbf{S})\,\mathbf{S}^t\right]\mathrm{B}^t = \mathbf{0}_{d\times(k+1)}, \tag{A.26}$$

which implies

$$\mathrm{V}^* = \mathrm{C}\Pi\mathrm{Q}\mathrm{B}^t \left(\mathrm{B}\tilde{\Pi}\mathrm{B}^t\right)^{-1}, \tag{A.27}$$

where (A.19) was used to replace the expectations, and we assumed that the autocorrelation matrix of the observables is invertible.

Substituting (A.27) in (A.13) gives the optimal codebook matrix as

$$\mathrm{Y}^* = \mathrm{C}\Pi\mathrm{Q}\mathrm{B}^t \left(\mathrm{B}\tilde{\Pi}\mathrm{B}^t\right)^{-1} \mathrm{B}. \tag{A.28}$$

This is the Weighted Centroid Condition for BLVQ on a noisy channel. The optimal codebook for the noiseless case is obtained by substituting (A.21) in (A.28) to yield

$$\mathrm{Y}^* = \mathrm{C}\Pi\mathrm{B}^t \left(\mathrm{B}\Pi\mathrm{B}^t\right)^{-1} \mathrm{B}. \tag{A.29}$$

This is the Centroid Condition for BLVQ on an ideal noiseless channel. For both noisy and noiseless discrete memoryless channels, the optimal BLVQ codevectors are linear combinations of the centroids. The actual coefficients are determined by the bit-mapping matrix and, in the noisy channel case, the transition probabilities of the channel.

A.3.3 Implications of optimality

Proposition 3 *If an unstructured vector quantizer or a binary lattice vector quantizer designed for a noisy or an ideal noiseless discrete memoryless channel satisfies the corresponding version of the Centroid Condition, then*

$$\mathrm{E}\left[\mathcal{Q}(\mathbf{X})\right] = \mathrm{E}\left[\mathbf{X}\right], \tag{A.30}$$

$$\Delta = \mathrm{E}\left\|\mathbf{X}\right\|^2 - \mathrm{E}\left\|\mathcal{Q}(\mathbf{X})\right\|^2, \tag{A.31}$$

and

$$\mathrm{E}\left\|\mathcal{Q}(\mathbf{X}) - \mathrm{E}\left[\mathcal{Q}(\mathbf{X})\right]\right\|^2 = \mathrm{E}\left\|\mathbf{X} - \mathrm{E}\left[\mathbf{X}\right]\right\|^2 - \Delta. \tag{A.32}$$

Proof

Equation (A.32) is a direct consequence of (A.30) and (A.31), since for any random vector $\mathbf{Z}$ we have

$$\mathrm{E}\left\|\mathbf{Z}-\mathrm{E}\left[\mathbf{Z}\right]\right\|^2 = \mathrm{E}\left\|\mathbf{Z}\right\|^2 - 2\mathrm{E}\left[\mathbf{Z}^t\right]\mathrm{E}\left[\mathbf{Z}\right] + \left\|\mathrm{E}\left[\mathbf{Z}\right]\right\|^2.$$

Thus, it suffices to show that (A.30) and (A.31) hold.

Equation (A.31) follows from the orthogonality conditions (A.15) for unstructured VQ and (A.24) for BLVQ (which are valid for both noisy and noiseless channels). To see this, note that by

$$\Delta = \mathrm{E}\left\|\mathbf{X}-\mathcal{Q}(\mathbf{X})\right\|^2 = \mathrm{E}\left\|\mathbf{X}\right\|^2 - 2\mathrm{E}\left[\mathbf{X}^t\mathcal{Q}(\mathbf{X})\right] + \mathrm{E}\left\|\mathcal{Q}(\mathbf{X})\right\|^2,$$

Equation (A.31) is equivalent to $\mathrm{E}\left[\mathbf{X}^t\mathcal{Q}(\mathbf{X})\right] = \mathrm{E}\left\|\mathcal{Q}(\mathbf{X})\right\|^2$, which reduces to

$$\mathrm{E}\left[\left(\mathbf{X}-\mathcal{Q}(\mathbf{X})\right)^t\mathcal{Q}(\mathbf{X})\right] = 0. \tag{A.33}$$

By (A.11) an optimal unstructured VQ can be written as

$$\mathcal{Q}(\mathbf{X}) = \mathrm{Y}^*\mathbf{S}. \tag{A.34}$$

Hence, choosing $\mathrm{Y} = \mathrm{Y}^*$ in (A.15) yields (A.33) for unstructured VQ. Similarly, by (A.14) an optimal BLVQ can be written as

$$\mathcal{Q}(\mathbf{X}) = \mathrm{V}^*\mathrm{B}\mathbf{S}. \tag{A.35}$$

Thus, choosing $\mathrm{V} = \mathrm{V}^*$ in (A.24) gives (A.33) for BLVQ.

It remains to show (A.30). For unstructured quantizers, substituting (A.34) in (A.17) and multiplying both sides by $\mathbf{1}_{2^k}$, a 2^k-dimensional vector of ones, we obtain

$$\mathbf{0}_d = \mathrm{E}\left[\left(\mathbf{X}-\mathcal{Q}(\mathbf{X})\right)\mathbf{S}^t\right]\mathbf{1}_{2^k} \tag{A.36}$$

$$= \mathrm{E}\left[\left(\mathbf{X}-\mathcal{Q}(\mathbf{X})\right)\mathbf{S}^t\mathbf{1}_{2^k}\right] \tag{A.37}$$

$$= \mathrm{E}\left[\mathbf{X}-\mathcal{Q}(\mathbf{X})\right], \tag{A.38}$$

where (A.37) follows by the linearity of expectation, and (A.38) follows from

$$\mathbf{S}^t \mathbf{1}_{2^k} = \sum_{j \in \mathcal{I}} I_{\{J=j\}} = 1.$$

The proof of (A.30) for BLVQ is essentially identical. Substituting (A.35) in (A.26) and multiplying both sides by $\mathbf{e} = (1, 0, \ldots, 0)^t$, a $(k+1)$-dimensional vector whose only nonzero entry is a 1 in the first position, we obtain

$$\mathbf{0}_d = \mathrm{E}\left[(\mathbf{X} - \mathcal{Q}(\mathbf{X}))\, \mathbf{S}^t\right] \mathrm{B}^t \mathbf{e} \tag{A.39}$$

$$= \mathrm{E}\left[(\mathbf{X} - \mathcal{Q}(\mathbf{X}))\, \mathbf{S}^t\right] \mathbf{1}_{2^k}, \tag{A.40}$$

which is identical to (A.36). ■

We note that while in the noiseless case we have

$$\mathrm{E}\left[\mathcal{Q}(\mathbf{X})\right] = \sum_{i \in \mathcal{I}} p_i \mathbf{y}_i \qquad \text{and} \qquad \mathrm{E}\left\|\mathcal{Q}(\mathbf{X})\right\|^2 = \sum_{i \in \mathcal{I}} p_i \left\|\mathbf{y}_i\right\|^2,$$

for noisy channels the corresponding quantities are

$$\mathrm{E}\left[\mathcal{Q}(\mathbf{X})\right] = \sum_{i \in \mathcal{I}} P_i \mathbf{y}_i \qquad \text{and} \qquad \mathrm{E}\left\|\mathcal{Q}(\mathbf{X})\right\|^2 = \sum_{i \in \mathcal{I}} P_i \left\|\mathbf{y}_i\right\|^2.$$

A.4 References

[1] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression.* Boston: Kluwer Academic Publishers, 1992.

[2] H. Kumazawa, M. Kasahara, and T. Namekawa, "A construction of vector quantizers for noisy channels," *Electronics and Engineering in Japan*, vol. 67-B, no. 4, pp. 39–47, 1984.

[3] R. Hagen and P. Hedelin, "Robust vector quantization by a linear mapping of a block-code," to appear in *IEEE Trans. Info. Theory.*

[4] D. G. Luenberger, *Optimization by Vector Space Methods.* New York: Wiley, 1969.

VITA

András Méhes was born on October 30, 1968, in Paks, Hungary. He received the diploma with honours in electrical and computer engineering from the Technical University of Budapest (Hungary) in 1993. He was awarded a "Köztársasági Ösztöndíj" (Fellowship of the Republic of Hungary) in both 1991 and 1992, and was a recipient of the "A Magyar Műszaki Haladásért Emlékérem" (Award for the Advancement of Hungarian Technology) in 1992. From 1993 to 1998 he was a research assistant at the University of Illinois at Urbana-Champaign, and from 1996 to 1998 a visiting scholar at the University of California, San Diego. He has co-authored the following papers:

- A. Méhes and K. Zeger, "Binary lattice vector quantization with linear block codes and affine index assignments," *IEEE Trans. on Information Theory,* vol. IT-44, pp. 79-95, January 1998.

- A. Méhes and K. Zeger, "Randomly chosen index assignments are asymptotically bad for uniform sources," *IEEE Trans. on Information Theory* (to appear).

- A. Méhes and K. Zeger, "Source and channel rate allocation for channel codes satisfying the Gilbert-Varshamov or Tsfasman-Vladut-Zink bounds," submitted to *IEEE Trans. on Information Theory,* (October 1997).

- A. Méhes and K. Zeger, "Performance of quantizers on noisy channels using structured families of codes," submitted to *IEEE Trans. on Information Theory,* (August 1998).

- A. Méhes and K. Zeger, "Randomly chosen index assignments are asymptotically bad for uniform sources," *Proc. IEEE Int. Symp. Information Theory,* Cambridge, MA, August 1998.

- A. Méhes and K. Zeger, "Tradeoff between source and channel coding for codes satisfying the Gilbert-Varshamov bound," *Proc. Conf. Information Science and Systems,* Baltimore, MD, March 1997.

- A. Méhes and K. Zeger, "Affine index assignments for binary lattice quantization with channel noise," *Proc. IEEE Int. Symp. Information Theory,* Whistler, B.C., Canada, p. 377, September 1995.

- A. Méhes and K. Zeger, "On the performance of affine index assignments for redundancy free source-channel coding," *Proc. IEEE Data Compression Conf.,* Snowbird, UT, p. 433, March 1995.

- A. Méhes and K. Zeger, "Redundancy free codes for binary discrete memoryless channels," *Proc. Conf. Information Science and Systems,* Princeton, NJ, pp. 1057-1062, March 1994.

- A. Méhes, "Multiple access channels," *Diploma Work,* Technical University of Budapest, Hungary, June 1993.