Performance of Quantizers on Noisy Channels Using Structured Families of Codes

András Méhes and Kenneth Zeger, *Fellow, IEEE*

***Abstract*—Achievable distortion bounds are derived 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 be made to decay to zero exponentially fast as a function of the square root of 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.**

***Index Terms*—Data compression, lattice vector quantization, linear error-correcting codes, source and channel coding.**

I. 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 tool for obtaining an analytical understanding of the properties of optimal quantizers has been the asymptotic theory. Two important types of asymptotic theories exist: 1) fixed transmission rate and growing block length; 2) fixed block length 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 our present paper, 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 paper, 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 fam-

Manuscript received September 1, 1998; revised May 3, 2000. This work was supported in part by the National Science Foundation. The material in this paper was presented in part at the Data Compression Conference, Salt Lake City, UT, March 1999.

A. Mehes is with the Royal Institute of Technology, Stockholm, Sweden (e-mail: andras@code.ucsd.edu).

K. Zeger is with the Department of Electrical and Computer Engineering, University of California at San Diego, La Jolla, CA 92093-0407 (e-mail: zeger@ucsd.edu).

Communicated by P. A. Chou, Associate Editor for Source Coding.

Publisher Item Identifier S 0018-9448(00)09663-2.

ilies of channel codes (e.g., Hamming, Bose–Chaudhuri–Hocquenghem (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 paper, 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 block lengths. 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 [5]–[7], [4]. 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 [5]–[7], [4], 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–Vlăduţ–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 block lengths. In the present paper 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 block length 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 [10]) the structured source coders in this paper 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 paper are collected into Theorem 1 in Section IV, which gives achievable bounds on the asymptotic mean-squared error performance of binary lattice vector quantizers and several useful families of binary linear block channel codes on a binary symmetric channel. The bounds in Theorem 1 show that the minimum distortion with certain structured codes decays to zero as $O(2^{-2Rg(R)})$, 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 paper, the distortion bounds given are also upper bounds on the distortion using optimal unstructured vector quantizer (VQ) with the same structured channel codes. In addition, the derivations of the bounds in Theorem 1 may be useful tools for future research (e.g., see [16]), since they are not specific to the codes used. Section II introduces necessary notations, definitions, and lemmas. Section III gives the framework for the source/channel coding problem and Section IV gives the results of the paper.

II. Preliminaries

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 $|f(n)|\le c|g(n)|$, 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 $\boldsymbol{x}\in\mathbb{R}^d$ will be written as column vectors $\boldsymbol{x}=(x_1,x_2,\ldots,x_d)^t$. The inner product of two Euclidean vectors $\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^d$ is denoted by

$$\langle\boldsymbol{x}\,|\,\boldsymbol{y}\rangle=\sum_{l=1}^{d} x_l y_l\in\mathbb{R}.$$

Also, $\|\boldsymbol{x}\|=\sqrt{\langle\boldsymbol{x}\,|\,\boldsymbol{x}\rangle}$ denotes the usual Euclidean norm of the vector $\boldsymbol{x}\in\mathbb{R}^d$. The symbols $I_{\{\cdot\}}$, $\Pr[\cdot]$, and $E[\cdot]$ are used to denote indicator functions, probabilities, and expectations, respectively.

A. Entropy and Relative Entropy

Definition 1: Let P and $\tilde{P}$ be probability distributions on a finite set.

The *entropy* of P (in bits) is

$$H(P) = -\sum_x P(x)\log_2 P(x). \tag{1}$$

The *relative entropy* between P and $\tilde{P}$ (in bits) is

$$D(P\,\|\,\tilde{P}) = \sum_x P(x)\log_2 \frac{P(x)}{\tilde{P}(x)}. \tag{2}$$

Definition 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

$$h(\epsilon) \triangleq H(\mathcal{P}_\epsilon) = -\epsilon\log_2\epsilon - (1-\epsilon)\log_2(1-\epsilon) \tag{3}$$

and the *binary relative entropy function* (information divergence) is

$$\mathcal{D}_2(\delta\,\|\,\epsilon) \triangleq D(\mathcal{P}_\delta\,\|\,\mathcal{P}_\epsilon) = \delta\log_2\frac{\delta}{\epsilon} + (1-\delta)\log_2\frac{1-\delta}{1-\epsilon}. \tag{4}$$

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

Lemma 1 ([17, p. 531]): For $0 \le \epsilon < \delta \le 1$

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

B. The Hadamard Transform

Definition 3: For each $i, j \in \mathbb{Z}_2^k$ let $h_{i,j} = (-1)^{ij^t}$ and let $\boldsymbol{f} : \mathbb{Z}_2^k \to \mathbb{R}^d$. The *Hadamard transform* $\hat{\boldsymbol{f}} : \mathbb{Z}_2^k \to \mathbb{R}^d$ of the mapping $\boldsymbol{f}$ is defined by

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

and the inverse transform is given by

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

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. The following useful identities also hold for $i, j, j' \in \mathbb{Z}_2^k$:

$$\begin{aligned} h_{i,j} &= h_{j,i},\\ h_{i,j}h_{i,j'} &= h_{i,j+j'},\\ \sum_{i\in\mathbb{Z}_2^k} h_{i,j} &= 2^k I_{\{j=0\}}. \end{aligned} \tag{5}$$

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

$$1 - h_{i,e_m} = 2i_m, \qquad m \in \{0, 1, \ldots, k-1\}. \tag{6}$$

C. Source Coding—Vector Quantization

Definition 4: A d-dimensional, 2^k-point *vector quantizer* (VQ) with *index set* $\mathcal{I} = \{0, \ldots, 2^k - 1\}$, and *codebook* $\mathcal{Y} = \{\boldsymbol{y}_i \in \mathbb{R}^d : i \in \mathcal{I}\}$, is a functional composition $\mathcal{Q}_0 = \mathcal{D}_Q \circ \mathcal{E}_Q$, where $\mathcal{E}_Q : \mathbb{R}^d \to \mathcal{I}$ is a *quantizer encoder* and $\mathcal{D}_Q : \mathcal{I} \to \mathcal{Y}$ is a *quantizer decoder*. (The subscript 0 denotes association with a noiseless channel.) The elements of the codebook $\boldsymbol{y}_i \in \mathcal{Y}$ are called *codevectors*. Associated with each codevector $\boldsymbol{y}_i$ is its *encoder region* $\mathcal{R}_i = \{\boldsymbol{x} \in \mathbb{R}^d \,|\, \mathcal{E}_Q(\boldsymbol{x}) = i\}$. 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 VQ $\mathcal{Q}_0$ for a source random variable $\boldsymbol{X} \in \mathbb{R}^d$ is

$$\Delta_S = E\|\boldsymbol{X} - \mathcal{Q}_0(\boldsymbol{X})\|^2 = \sum_{i\in\mathcal{I}}\int_{\mathcal{R}_i}\|\boldsymbol{x} - \boldsymbol{y}_i\|^2\,d\mu(\boldsymbol{x}) \tag{7}$$

where μ is the probability distribution of the input $\boldsymbol{X}$.

Necessary conditions for the optimality of a vector quantizer using the mean-squared distortion are (see [18], for example) the *Centroid Condition*

$$\boldsymbol{y}_i = E[\boldsymbol{X}\,|\,\boldsymbol{X} \in \mathcal{R}_i] \qquad \forall i \in \mathcal{I} \tag{8}$$

and the *Nearest Neighbor Condition*:

$$\mathcal{R}_i = \{\boldsymbol{x} \in \mathbb{R}^d : \|\boldsymbol{x} - \boldsymbol{y}_i\| < \|\boldsymbol{x} - \boldsymbol{y}_j\|\ \forall j \in \mathcal{I}\backslash\{i\}\} \quad \forall i \in \mathcal{I}. \tag{9}$$

Locally optimal vector quantizers satisfying both necessary conditions (8) and (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, which is stated below in a convenient form.

Lemma 2 (Zador [3]): The minimum mean-squared error of a rate R_S vector quantizer is asymptotically (as $R_S \to \infty$) given by

$$\Delta_S = 2^{-2R_S + O(1)}. \tag{10}$$

This is often referred to as the "6 dB per bit" rule, since

$$10\log_{10}\left(2^{-2R_S+O(1)}\big/2^{-2(R_S+1)+O(1)}\right) \to 20\log_{10}2 \approx 6\text{ dB}.$$

We say that a sequence of quantizers is *asymptotically good* if

$$\limsup_{R_S\to\infty}\Delta_S 2^{2R_S} < \infty. \tag{11}$$

We say that a sequence of quantizers is *bounded*, if the codepoints of the quantizers are bounded, that is,

$$\sup_k\left(\max_{i\in\mathcal{I}_k}\left\|\boldsymbol{y}_i^{(k)}\right\|\right) < \infty \tag{12}$$

where $\mathcal{I}_k$ denotes the index set and the $\boldsymbol{y}_i^{(k)}$ denote the codepoints of the k-bit quantizer in the sequence. Lemma 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 (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.

1) Binary Lattice VQ:

Definition 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

$$\boldsymbol{y}_i = \boldsymbol{y}_0 + \sum_{l=0}^{k-1} \boldsymbol{v}_l i_l \qquad \forall i \in \mathbb{Z}_2^k \tag{13}$$

where $\boldsymbol{y}_0 \in \mathbb{R}^d$ is an *offset vector* and $\{\boldsymbol{v}_l\}_{l=0}^{k-1} \subset \mathbb{R}^d$ is the set of *generator vectors,* ordered by $\|\boldsymbol{v}_0\| \le \|\boldsymbol{v}_1\| \le \cdots \le \|\boldsymbol{v}_{k-1}\|$.

In this paper, 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 binary lattice vector quantizers.

D. Channel Coding—Linear Codes on a Binary Symmetric Channel

Definition 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 of the $2^k - 1$ nonzero codewords has 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 indexes) 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

$$p_{l\,|\,m} = \sum_{u \in \mathcal{D}_C^{-1}(l)} \epsilon^{w(u+\mathcal{E}_C(m))}\,(1-\epsilon)^{n-w(u+\mathcal{E}_C(m))}.$$

If the code is linear then $\mathcal{E}_C$ and $\mathcal{D}_C$ can be chosen to ensure $p_{l\,|\,m} = p_{l+m\,|\,0}$ (e.g., any linear encoder and coset decoder will do). 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 $P_l^{(\mathrm{bit})}$ 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

$$P_l^{(\mathrm{bit})} = \sum_{i \in \mathbb{Z}_2^k} q_i I_{\{i_l = 1\}}$$

and $P_e = 1 - q_0$. Let $P_{\max}^{(\mathrm{bit})}$ denote the maximum of the error probabilities for decoded bits. Then

$$P_{\max}^{(\mathrm{bit})} = \max_l P_l^{(\mathrm{bit})} \le P_e.$$

Since a code with minimum distance $d_{\min}$ can correct all possible $\lfloor (d_{\min}-1)/2 \rfloor$-bit errors and since $\lfloor (d_{\min}-1)/2 \rfloor + 1 \ge d_{\min}/2$, Lemma 1 can be used to bound P_e as follows.

Lemma 3: For any $[n, k, d_{\min}]$ linear block channel code and for any $\epsilon \le d_{\min}/(2n)$, the probability of a block error with a linear encoder and a symmetric maximum-likelihood (ML) decoder on a binary symmetric channel with bit-error probability ϵ satisfies

$$P_e \le 2^{-n\mathcal{D}_2(\frac{d_{\min}}{2n}\|\epsilon)}.$$

To obtain asymptotic results we consider families of $[n, k, d_{\min}]$ linear channel codes indexed by the block length 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 3 becomes trivial as the block length 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 block length. From a source–channel tradeoff perspective, the best codes in these families are those with small block lengths. 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]), or the Tsfasman–Vlăduţ–Zink bound [20]. Bounds on the asymptotically optimal source/channel rate allocation were derived in [7] 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 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 block length increases, but more reliably.

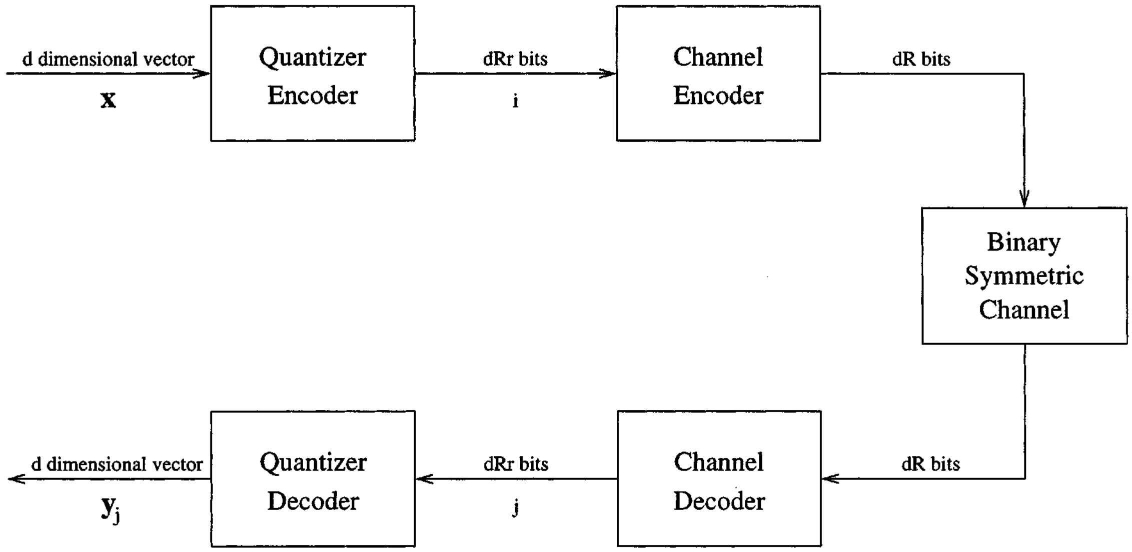

Fig. 1. Cascaded vector quantizer and channel coder system.

In this paper, we focus attention on the third category. One seeks an optimal "schedule" of the rate k/n converging to 0 as a function of the block length n.

E. The Cascaded System

The following definition corresponds to Fig. 1.

Definition 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 $\mathcal{E}_Q : \mathbb{R}^d \to \mathbb{Z}_2^k$ is a quantizer encoder, $\mathcal{D}_Q : \mathbb{Z}_2^k \to \mathcal{Y}$ is a quantizer decoder, $\mathcal{E}_C : \mathbb{Z}_2^k \to \mathcal{C}$ is a channel encoder, $\mathcal{D}_C : \mathbb{Z}_2^n \to \mathbb{Z}_2^k$ is a channel decoder, and $\eta : \mathbb{Z}_2^n \to \mathbb{Z}_2^n$ is a random mapping representing a noisy channel.

The mean-squared distortion of a noisy channel vector quantizer for a source random variable $\boldsymbol{X} \in \mathbb{R}^d$ is

$$\Delta = E\|\boldsymbol{X} - \mathcal{Q}(\boldsymbol{X})\|^2 = \sum_{i \in \mathbb{Z}_2^k} \sum_{j \in \mathbb{Z}_2^k} q_{i+j} \int_{\mathcal{R}_i} \|\boldsymbol{x} - \boldsymbol{y}_j\|^2 \, d\mu(\boldsymbol{x}) \tag{14}$$

where μ is the probability distribution of the input $\boldsymbol{X}$, and for $i, j \in \mathbb{Z}_2^k$ the $q_{i+j} = \Pr[\mathcal{D}_C(\eta(\mathcal{E}_C(i))) = j]$ are the transition probabilities of the coded channel.

III. Rate Allocation Tradeoff

Analogous to the source distortion in (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

$$\Delta_C \triangleq E\|\mathcal{Q}_0(\boldsymbol{X}) - \mathcal{Q}(\boldsymbol{X})\|^2 \tag{15}$$

(the component of the distortion influenced by channel errors). If the quantizer $\mathcal{Q}_0$ satisfies the Centroid Condition then

$$\Delta = \Delta_S + \Delta_C. \tag{16}$$

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 (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

$$\Delta \le (\sqrt{\Delta_S} + \sqrt{\Delta_C})^2. \tag{17}$$

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 binary lattice vector quantizers 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 (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.

A. Rate Allocation for BLVQ

Consider a d-dimensional 2^k-point binary-lattice vector quantizer 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 Fig. 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 (17), we seek an exponentially decaying (in R) expression for the channel distortion Δ_C of the cascade of a binary lattice vector quantizer 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 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 paper, we do not use an explicit (i.e., nonidentity) 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 [10] to the case of the Natural Binary Code (identity index assignment).

Lemma 4 ([10]): Let $\boldsymbol{X} \in \mathbb{R}^d$ be a source random variable quantized by a 2^k-point binary-lattice vector quantizer with generating set $\{\boldsymbol{v}_l\}_{l=0}^{k-1}$ and transmitted on a binary symmetric channel using the Natural Binary index assignment and an $[n,k]$ binary linear channel code. Let $p_i = \Pr[\boldsymbol{X} \in \mathcal{R}_i]$ 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

$$\Delta_C = \frac{1}{4}\sum_{l=0}^{k-1}\sum_{m=0}^{k-1}\langle \boldsymbol{v}_l \,|\, \boldsymbol{v}_m\rangle \hat{p}_{e_l+e_m}(\hat{q}_0 - \hat{q}_{e_l} - \hat{q}_{e_m} + \hat{q}_{e_l+e_m}) \tag{18}$$

where the hats denote Hadamard transforms, and e_l is the binary row vector with its only nonzero entry in the lth position.

Equation (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 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 (18). Using the Hadamard transform definition and its identities gives

$$\begin{aligned}
&\hat{q}_0 - \hat{q}_{e_l} - \hat{q}_{e_m} + \hat{q}_{e_l+e_m}\\
&\quad= \sum_{i\in\mathbb{Z}_2^k} q_i\,(1-h_{i,e_l})\,(1-h_{i,e_m})\\
&\quad= 4\sum_{i\in\mathbb{Z}_2^k} q_i i_l i_m\\
&\quad= 4\Pr[l\text{th and } m\text{th bits both in error}]\\
&\quad\le 4\min\Big(P_l^{(\mathrm{bit})}, P_m^{(\mathrm{bit})}\Big)\\
&\quad\le 4P_{\max}^{(\mathrm{bit})}.
\end{aligned} \tag{19}$$

Next, we examine the remaining portion of the sum in (18), the source-dependent component. Again using the Hadamard transform definition and its identities, we obtain

$$\begin{aligned}
&\frac{1}{4}\sum_{l=0}^{k-1}\sum_{m=0}^{k-1}\langle \boldsymbol{v}_l \,|\, \boldsymbol{v}_m\rangle \hat{p}_{e_l+e_m}\\
&\quad= \frac{1}{4}\sum_{l=0}^{k-1}\sum_{m=0}^{k-1}\langle \boldsymbol{v}_l \,|\, \boldsymbol{v}_m\rangle \sum_{i\in\mathbb{Z}_2^k} p_i h_{i,e_l+e_m}\\
&\quad= \frac{1}{4}\sum_{i\in\mathbb{Z}_2^k} p_i \left\|\sum_{l=0}^{k-1}\boldsymbol{v}_l(1-2i_l)\right\|^2\\
&\quad= \frac{1}{4}\sum_{i\in\mathbb{Z}_2^k} p_i \|\boldsymbol{y}_{\bar{i}} - \boldsymbol{y}_i\|^2 &(20)\\
&\quad\le \frac{1}{4}\sum_{i\in\mathbb{Z}_2^k} p_i (\|\boldsymbol{y}_{\bar{i}}\| + \|\boldsymbol{y}_i\|)^2 &(21)\\
&\quad\le \rho^2 &(22)
\end{aligned}$$

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 (12).

Combining (19) and (22), the channel distortion in (18) can be upper-bounded as

$$\Delta_C \le 4P_{\max}^{(\mathrm{bit})}\rho^2. \tag{23}$$

It remains to show that $P_{\max}^{(\mathrm{bit})}$, 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 nondecreasing function of n, which implies a one-to-one relationship between the channel code rate r and the block length n (e.g., this holds for repetition and Hamming codes). We divide the Rd bits per source vector into blocks of shorter channel codes from the same family of $[n,k,d_{\min}]$ codes, and assume that each has the same block length n (a divisor of Rd). Thus the length Rd channel code is the (Rd/n)-ary Cartesian product of identical length n codes. This maintains the overall transmission rate 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 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,

$$P_{\max}^{(\mathrm{bit})} \le 2^{-n\mathcal{D}_2\left(\frac{d_{\min}}{2n}\|\epsilon\right)}. \tag{24}$$

Thus $P_{\max}^{(\mathrm{bit})}$ can be made to decay to zero exponentially fast in R by choosing the constituent block length n to satisfy $n \to \infty$ as $R \to \infty$.

Substituting (24) in (23) yields

$$\Delta_C \le 2^{-n\mathcal{D}_2(\frac{d_{\min}}{2n} \| \epsilon) + O(1)}. \tag{25}$$

Combining this with the formula for the source distortion of asymptotically good quantizers in (11) and using (17), the total distortion is bounded as

$$\Delta \le \left(2^{-R\frac{k}{n} + O(1)} + 2^{-\frac{n}{2}\mathcal{D}_2(\frac{d_{\min}}{2n}\|\epsilon)+O(1)}\right)^2. \tag{26}$$

The value of the right side of (26) for any n that divides Rd represents an achievable distortion, since there exist binary lattice vector quantizers and families of channel codes that satisfy such a bound (it can easily be shown that $n_R \le Rd$). We seek n as a function of R to minimize the right side of (26) asymptotically in R. As argued in [5], since one of the exponents is increasing in n while the other is decreasing, in the asymptotic sense the minimum is achieved when the two exponents are asymptotically equal, for then both terms on the right side of (26) decay at the same rate. (A recent application of the same idea is found in [16].) Let n_R denote a value of n obtained by equating the exponents. Asymptotically (in R), $n_R \to \infty$ must hold, for otherwise the second term in (26) would be bounded away from zero. 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 (26) is a finite nonzero constant which we denote by

$$\beta \triangleq \mathcal{D}_2(1/2 \lim_{n\to\infty} (d_{\min}/n) \,\|\, \epsilon).$$

Thus the asymptotically minimizing n_R satisfies

$$\lim_{R\to\infty} \frac{2Rk}{n_R^2 \beta} = 1. \tag{27}$$

Let r_R denote the channel code rate corresponding to the n_R which solves (27). Then by (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 (27) for various code families.

IV. Asymptotic Distortion Decay Rates

First, two lemmas are given that solve (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: 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 (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 follows by direct substitution, since

$$\frac{2Rk}{n_R^2\beta} = \left(\frac{2Rc}{\beta}\right) n_R^{\alpha-2} = 1;$$

and

$$r_R = cn_R^{\alpha-1} = c\left(\frac{2c}{\beta}R\right)^{-\frac{1-\alpha}{2-\alpha}}. \qquad \square$$

Note that $\alpha = 1$ corresponds to asymptotically good codes (where the optimal rate is asymptotically constant as shown in [7]) and $\alpha = 0$ corresponds to repetition codes. For $\alpha \approx 0$, the channel code rate decays as $r_R = O(1/\sqrt{R})$, 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 6: If $k/(\log_2 n)^l \to c$ as $n \to \infty$ for some finite $c > 0$ and some $l \ge 0$, then

$$n_R = \sqrt{\frac{2c}{\beta}R\left(\frac{1}{2}\log_2 R\right)^l}$$

satisfies (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 6 follows by direct substitution

$$\begin{aligned}\lim_{R\to\infty}\frac{2Rk}{n_R^2\beta} &= \lim_{R\to\infty}\left(\frac{2R}{\beta}\right)\frac{k}{(\log_2 n_R)^l}\frac{(\log_2 n_R)^l}{n_R^2}\\ &= \lim_{R\to\infty}\frac{2Rc}{\beta}\frac{\left(\log_2\left[\sqrt{R\frac{2c}{\beta}\left(\frac{1}{2}\log_2 R\right)^l}\right]\right)^l}{\frac{2c}{\beta}R\left(\frac{1}{2}\log_2 R\right)^l}\\ &= \lim_{R\to\infty}\left(\frac{\frac{1}{2}\log_2 R + \log_2\left[\sqrt{\frac{2c}{\beta}\left(\frac{1}{2}\log_2 R\right)^l}\right]}{\frac{1}{2}\log_2 R}\right)^l\\ &= 1\end{aligned}$$

and

$$\begin{aligned}\lim_{R\to\infty}\frac{k}{n_R r_R} &= \lim_{R\to\infty}\frac{k}{(\log_2 n_R)^l}\frac{(\log_2 n_R)^l}{n_R r_R}\\ &= \lim_{R\to\infty} c\frac{\left(\log_2\left[\sqrt{R\frac{2c}{\beta}\left(\frac{1}{2}\log_2 R\right)^l}\right]\right)^l}{\sqrt{\frac{2c}{\beta}R\left(\frac{1}{2}\log_2 R\right)^l\frac{c\beta\left(\frac{1}{2}\log_2 R\right)^l}{2R}}}\\ &= \lim_{R\to\infty}\frac{\left(\log_2\left[\sqrt{R\frac{2c}{\beta}\left(\frac{1}{2}\log_2 R\right)^l}\right]\right)^l}{\left(\frac{1}{2}\log_2 R\right)^l}\\ &= 1. \qquad \square\end{aligned}$$

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 $O(\sqrt{\log R/R})$.

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 shortened 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 1: Let $\boldsymbol{X} \in \mathbb{R}^d$ 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

$$\Delta \leq 2^{-2Rr_R + O(1)} \tag{28}$$

which is achieved by a channel code rate r_R, for various channel code families as follows:

i) for a family of $[n, 1, n]$ repetition codes $(n \geq 1)$

$$r_R = \sqrt{\frac{-\log_2 2\sqrt{\epsilon(1-\epsilon)}}{2R}}, \qquad \epsilon \in (0, 1/2); \tag{29}$$

ii) for a family of lth-order $[2^m, \sum_{i=0}^{l}\binom{m}{i}, 2^{m-l}]$ Reed–Muller codes $(m \geq 1)$

$$r_R = \sqrt{\frac{-\left(\log_2 2^{l+1}\left(\epsilon\left(\frac{1-\epsilon}{2^{l+1}-1}\right)^{2^{l+1}-1}\right)^{\frac{1}{2^{l+1}}}\right)(\log_2 R)^l}{l!2^{l+1}R}}, \qquad \epsilon \in (0, 1/2^{l+1}); \tag{30}$$

iii) and for a family of duals of extremal t-error-correcting $[2^m - 1, mt, 2^{m-1} - \lfloor\log_2(2t-1)\rfloor]$ BCH codes $(m \geq 1)$

$$r_R = \sqrt{\frac{-t\left(\log_2 4\left(\epsilon\left(\frac{1-\epsilon}{3}\right)^3\right)^{\frac{1}{4}}\right)\log_2 R}{4R}}, \qquad \epsilon \in (0, 1/4). \tag{31}$$

Proof: The inequality in (28) is a direct consequence of (26) and the ensuing discussion. The various expressions for r_R are obtained from the solutions n_R of (27) as given by Lemma 6 (alternatively, Lemma 5 for repetition codes) with β substituted using the actual code parameters.

i) Since $d_{\min}/n = 1$ for repetition codes,

$$\beta = \mathcal{D}_2(1/2 \,\|\, \epsilon) = -\log_2(2\sqrt{\epsilon(1-\epsilon)}).$$

Substituting this in Lemma 5 with $\alpha = 0$ and $c = 1$ (or in Lemma 6 with $l = 0$ and $c = 1$), the result follows.

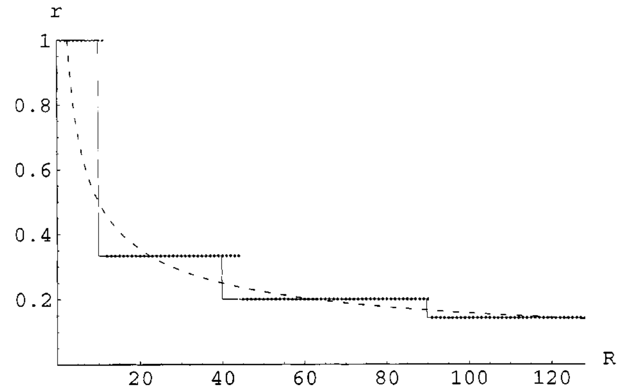

Fig. 2. An illustration of Theorem 1 for uniform scalar quantization of a uniform source on $(0, 1)$ using repetition codes to transmit on a binary symmetric channel with $\epsilon = 10^{-3}$. The distortion minimizing channel code rate r is plotted against the overall transmission rate R. The dashed curve is obtained directly from (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.

ii) An lth-order length $n = 2^m$ Reed–Muller code has

$$k = \sum_{i=0}^{l}\binom{m}{i} = m^l/l!(1 + o(1))$$

information symbols as $n \to \infty$. Hence, Lemma 6 can be applied with $c = 1/l!$. 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 6 with $l = 1$ and $c = t$. By the Carlitz–Uchiyama bound ([19, p. 280]), $\lim_n d_{\min}/n = 1/2$. The result then follows by substitution. □

Fig. 2 provides an illustration of Theorem 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 (29) (dashed curve), we also plotted the channel code rate corresponding to the closest odd block length (step function).

As with Zador's lemma, Theorem 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 "decibels per bit" of performance gain in the bound (28) 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

$$\begin{aligned} &\text{SNR}(R+1) - \text{SNR}(R) \\ &= 10\log_{10}\left(2^{-2\sqrt{R}+O(1)} \Big/ 2^{-2\sqrt{R+1}+O(1)}\right) \\ &\approx \frac{3}{\sqrt{R}} \text{ [dB]}. \end{aligned}$$

However, the bounds presented might be improved in the future.

V. Conclusion

The paper 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 $O\left(\sqrt{\frac{\log_2^l R}{R}}\right)$ 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).

References

[1] C. E. Shannon, "Coding theorems for a discrete source with a fidelity criterion," in *IRE Nat. Conv. Rec., Pt. 4*, 1959, pp. 142–163.

[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. Inform. Theory*, vol. IT-28, pp. 139–149, Mar. 1982.

[4] K. Zeger and V. Manzella, "Asymptotic bounds on optimal noisy channel quantization via random coding," *IEEE Trans. Inform. Theory*, vol. 40, pp. 1926–1938, Nov. 1994.

[5] B. Hochwald and K. Zeger, "Tradeoff between source and channel coding," *IEEE Trans. Inform. Theory*, vol. 43, pp. 1412–1424, Sept. 1997.

[6] B. Hochwald, "Tradeoff between source and channel coding on a Gaussian channel," *IEEE Trans. Inform. Theory*, vol. 44, pp. 3044–3055, Nov. 1998.

[7] A. Méhes and K. Zeger, "Source and channel rate allocation for channel codes satisfying the Gilbert-Varshamov or Tsfasman–Vlăduţ–Zink bounds," *IEEE Trans. Inform. Theory*, vol. 46, pp. 2133–2151, Sept. 2000.

[8] G. L. Katsman, M. A. Tsfasman, and S. G. Vlăduţ, "Modular curves and codes with a polynomial construction," *IEEE Trans. Inform. Theory*, vol. IT-30, pp. 353–355, Mar. 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. Inform. Theory*, vol. 43, pp. 128–135, Jan. 1997.

[10] A. Méhes and K. Zeger, "Binary lattice vector quantization with linear block codes and affine index assignments," *IEEE Trans. Inform. Theory*, vol. 44, pp. 79–95, Jan. 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, Jan. 1993, p. 171.

[12] ——, "Design methods for VQ by linear mappings of block codes," in *Proc. IEEE Int. Symp. Information Theory*, Trondheim, Norway, June 1994, p. 241.

[13] ——, "Robust vector quantization by a linear mapping of a block code," *IEEE Trans. Inform. Theory*, vol. 45, pp. 200–218, Jan. 1999.

[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. Inform. Theory*, vol. 41, pp. 2031–2037, Nov. 1995.

[15] A. Méhes and K. Zeger, "Randomly chosen index assignments are asymptotically bad for uniform sources," *IEEE Trans. Inform. Theory*, vol. 45, pp. 788–794, Mar. 1999.

[16] P. W. Moo and D. L. Neuhoff, "Polar quantization revisited," *IEEE Trans. Inform. Theory*, submitted for publication.

[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, MA: Kluwer, 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. Vlăduţ, and T. Zink, "Modular curves, Shimura curves, and Goppa codes, better than Varshamov–Gilbert bound," *Math. Nachrichtentech.*, vol. 109, pp. 21–28, 1982.