Binary Lattice Vector Quantization with Linear Block Codes and Affine Index Assignments

András Méhes and Kenneth Zeger, *Senior Member, IEEE*

***Abstract*— 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.**

***Index Terms*— Index assignment, lattices, linear error-correcting codes, source and channel coding, vector quantization.**

I. 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 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 Fig. 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.

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 high 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 paper.

A number of studies have considered the communication system in Fig. 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 and centroid conditions. 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 for vector quantizers followed by channel coding exist. For a given transmission rate and fixed vector dimension, the optimal tradeoff be-

Manuscript received May 15, 1996; revised May 15, 1997. This work was supported in part by the National Science Foundation and the Joint Services Electronics Program. The material in this paper was presented in part at the Data Compression Conference, Snowbird, UT, March 1995, and at the IEEE International Symposium on Information Theory, Whistler, BC, Canada, September 1995.

A. Méhes is with the Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA.

K. Zeger is with the Department of Electrical and Computer Engineering, University of California, San Diego, La Jolla, CA 92093-0407 USA.

Publisher Item Identifier S 0018-9448(98)00076-5.

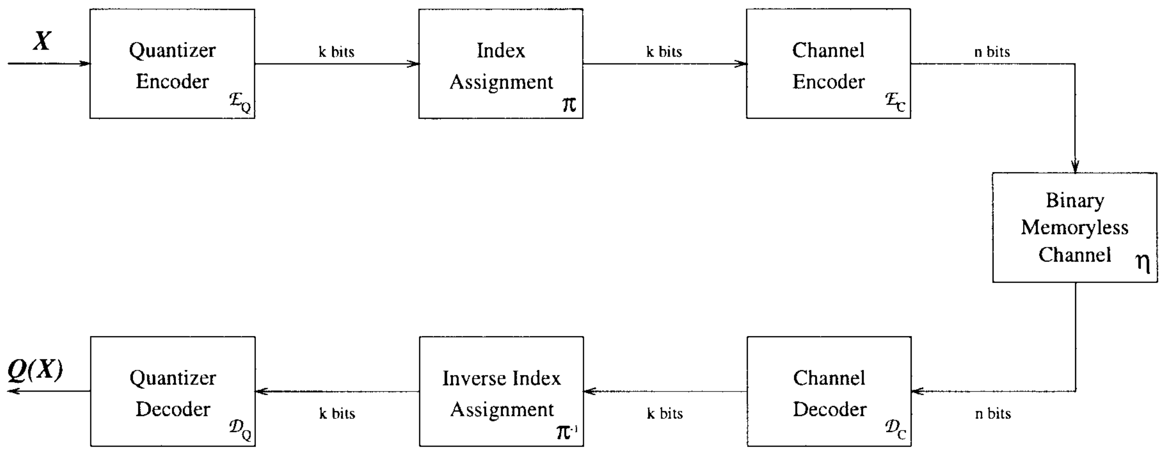

Fig. 1. Communication system model.

tween source and channel coding was 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 nor 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 (NBC), Folded Binary Code (FBC), Two's Complement Code (TCC), and Gray Code (GC) [29]. Huang [30]–[32] computed distortion formulas for the Natural Binary Code and the Gray Code for uniform scalar quantizers and uniform scalar sources. He 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 mean-squared error instead of reducing the probability of channel error. In [33], [38], and [39], Crimmins *et al.* 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 mean-squared error (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 the present paper, 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-

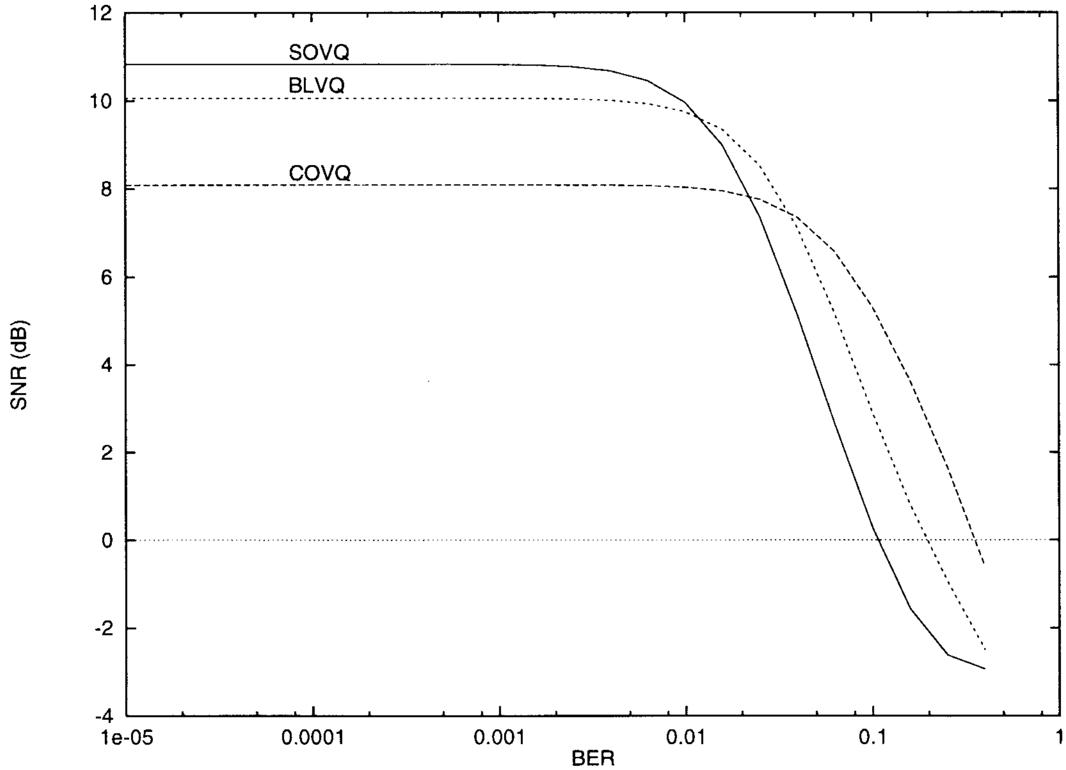

Fig. 2. The channel-mismatch performance of source-optimized VQ, channel-optimized VQ, and binary-lattice VQ. The input vectors are taken from a Gauss–Markov process with correlation 0.9. The COVQ was designed for a BER of 0.1. The 16-dimensional 2048-point quantizers are followed by a $(16, 11, 4)$ extended binary Hamming code.

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 (BER) is required. In this paper 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 quantization* (BLVQ), 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, Fig. 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 $10\,\log_{10}(\sigma^2/(D/d))$, 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 tradeoff 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 BER's. 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 due to the structured codebook of the BLVQ is relatively small. Fig. 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 paper we generalize much of the previously mentioned work to the case of redundant channel codes and BLVQ's. 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] used the notion of "Hadamard classes" to search for an optimal index assignment in the Hadamard transform domain.

The main contributions of our paper 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; and 3) comparison between the performances of NBC, FBC, GC, and TCC.

In Section II, we give the necessary notation and terminology. In Section III, we prove Theorem 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 IV, we consider quantization systems without the use of redundancy for error control. For binary-symmetric channels, Corollaries 2–4 give explicit formulas for the channel distortions of the NBC, FBC, and GC in this case. Corollary 5 characterizes the class of sources for which the NBC outperforms the FBC. For binary-asymmetric channels, Theorem 2 gives a general formula for the channel distortion of BLVQ using affine index assignments and no channel-coding redundancy. Corollaries 6–8 give explicit formulas for the channel distortions of the NBC, TCC, FBC, and GC in this case. Corollary 9 identifies the best assignment among all affine translates of the NBC for a nonsymmetric channel. Finally, Theorem 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.

II. Definitions

A. 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 paper 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 $i = [i_{k-1}, i_{k-2}, \cdots, i_1, i_0]$, 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 $\boldsymbol{x} \in \mathbb{R}^d$, we write $\boldsymbol{x} = (x_1, x_2, \cdots, x_d)^t$, where x_i is the ith component of $\boldsymbol{x}$.

In this paper 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 $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d$ by

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

The following definition corresponds to Fig. 1.

Definition 1: A d-dimensional, 2^k-point *noisy channel vector quantizer* with *codebook* $\mathcal{Y} = \{\boldsymbol{y}_i \in \mathbb{R}^d : i \in \mathbb{Z}_2^k\}$, and with an (n, k) channel code $\mathcal{C} = \{c_i : i \in \mathbb{Z}_2^k\} \subset \mathbb{Z}_2^n$, 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 $\mathcal{E}_Q$: $\mathbb{R}^d \to \mathbb{Z}_2^k$ 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*, π: $\mathbb{Z}_2^k \to \mathbb{Z}_2^k$ is an *index assignment* (bijection), and η: $\mathbb{Z}_2^n \to \mathbb{Z}_2^n$ 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 $\boldsymbol{X}$ be a random vector in $\mathbb{R}^d$. Let $p_i = \mathrm{P}[\mathcal{E}_Q(\boldsymbol{X})] = i$ denote the probability that the quantizer encoder produces the index i, and define $p_{j|i} = \mathrm{P}[\mathcal{D}_C(\eta(\mathcal{E}_C(i))) = j]$, 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 Fig. 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 $\boldsymbol{X}$ by its mean-squared error $D = \mathrm{E}||\boldsymbol{X} - \mathcal{Q}(\boldsymbol{X})||^2$. Define the *source distortion* $D_S \triangleq \mathrm{E}||\boldsymbol{X} - \tilde{\mathcal{Q}}(\boldsymbol{X})||^2$ (the distortion incurred on a noiseless channel), and the *channel distortion* $D_C \triangleq \mathrm{E}||\tilde{\mathcal{Q}}(\boldsymbol{X}) - \mathcal{Q}(\boldsymbol{X})||^2$ (the component due to channel errors). If the centroid condition is satisfied (i.e., $\boldsymbol{y}_i = \mathrm{E}[\boldsymbol{X}|\mathcal{E}_Q(\boldsymbol{X}) = i]$, $\forall i$), 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 $D_{sc} = \mathrm{E}[\langle \boldsymbol{X} - \tilde{\mathcal{Q}}(\boldsymbol{X}) | \tilde{\mathcal{Q}}(\boldsymbol{X}) - \mathcal{Q}(\boldsymbol{X}) \rangle]$. 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 I lists the three components of D for the BLVQ example of Fig. 2. It

TABLE I
THE THREE COMPONENTS OF THE DISTORTION (PER DIMENSION) FOR THE BLVQ OF FIG. 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				

can be seen in these cases that D_{sc} is negligible compared to D_S and D_C.

Hence, under the assumption $D_{sc} = 0$, for a given quantizer one should minimize D_C to optimize the overall performance. This paper 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 that yields the lowest D_C (i.e., the best index assignment π). Indices that are likely to be mistaken due to channel errors should correspond to code vectors whose Euclidean distance is small.

Fact 1: Let $\boldsymbol{X} \in \mathbb{R}^d$ be a random vector encoded by a noisy-channel vector quantizer. The channel distortion can be written as

$$D_C = \sum_{i \in \mathbb{Z}_2^k} \sum_{j \in \mathbb{Z}_2^k} p_i p_{\pi(j)|\pi(i)} \|\boldsymbol{y}_i - \boldsymbol{y}_j\|^2. \tag{1}$$

B. Linear Codes on a Binary-Symmetric Channel

Definition 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

$$\mathcal{C}^{\perp} = \{i \in \mathbb{Z}_2^n : ij^t = 0, \forall j \in \mathcal{C}\}.$$

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

$$\rho_u \triangleq \mathrm{P}[\eta(v) = v + u] = \epsilon^{w(u)}(1-\epsilon)^{n-w(u)}$$

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

$$q_j \triangleq p_{i+j|i} = \sum_{u \in \mathcal{S}} \rho_{u+j\mathrm{G_C}}, \qquad i, j \in \mathbb{Z}_2^k. \tag{2}$$

C. Affine Index Assignments

Definition 3: An *affine index assignment* $\pi: \mathbb{Z}_2^k \to \mathbb{Z}_2^k$ 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)*

$$\mathrm{G_I^{(FBC)}} = \begin{bmatrix} 1 & 1 \ \cdots \ 1 \\ 0 & \\ \vdots & \mathcal{I}_{k-1} \\ 0 & \end{bmatrix} \qquad t = [01 \ \cdots \ 1]$$

$$(\mathrm{G_I^{(FBC)}})^{-1} = \mathrm{G_I^{(FBC)}}.$$

- *Gray Code (or Reflected Binary Code)*

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

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

- *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

TABLE II
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

source resulting in a uniform distribution on the BLVQ codevectors. Using similar arguments it can be shown that a "worst" affine assignment (i.e., maximizing D_C under the same conditions among affine index assignments) is the following linear code:

$$\mathrm{G}_{\mathrm{I}}^{(\mathrm{WC})} = \begin{bmatrix} a_k & 1 & \cdots & 1 \\ 1 & & & \\ \vdots & & \mathcal{I}_{k-1} & \\ 1 & & & \end{bmatrix} \qquad t = [0 \ \cdots \ 0]$$

$$(\mathrm{G}_{\mathrm{I}}^{(\mathrm{WC})})^{-1} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & & & \\ \vdots & & \tilde{\mathcal{I}}_{k-1} & \\ 1 & & & \end{bmatrix}$$

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

Table II 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]):

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

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

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

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

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

D. Binary Lattice VQ

Definition 4: A d-dimensional, 2^k-point *binary lattice vector quantizer* is a vector quantizer, whose codevectors are of the form

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

for

$$i = [i_{k-1}, i_{k-2}, \cdots, i_1, i_0] \in \mathbb{Z}_2^k, \ \boldsymbol{y}_0 \in \mathbb{R}^d$$

and where $\mathcal{V} = \{\boldsymbol{v}_l\}_{l=0}^{k-1} \subset \mathbb{R}^d$ is the *generating set*, ordered by

$$\|\boldsymbol{v}_0\| \le \|\boldsymbol{v}_1\| \le \cdots \le \|\boldsymbol{v}_{k-1}\|.$$

A binary lattice quantizer can be considered a direct sum quantizer (or multistage, or residual quantizer) with two codevectors at each stage, when the codebook is written as $\bigoplus_{l=0}^{k-1}\{\boldsymbol{y}_0/k, (\boldsymbol{y}_0/k) + \boldsymbol{v}_l\}$. Conversely, any direct sum vector quantizer with two vectors per component codebook, $\bigoplus_{l=0}^{k-1}\{\boldsymbol{a}_l, \boldsymbol{b}_l\}$, can be viewed as a binary lattice VQ by setting

$$\boldsymbol{y}_0 = \sum_{l=0}^{k-1} \boldsymbol{a}_l$$

and $\boldsymbol{v}_l = \boldsymbol{b}_l - \boldsymbol{a}_l \ \forall \, l$, and reordering the generating set if needed.

Given an arbitrary lattice with basis vectors $\{\boldsymbol{u}_j\}_{j=1}^L \subset \mathbb{R}^d$, any set $\{k_j\}_{j=1}^L$ of nonnegative integers satisfying

$$\sum_{j=1}^{L} k_j = k$$

defines a 2^k-point *lattice vector quantizer* with codebook

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

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

The class of binary lattice VQ's includes lattice VQ's (or any of their cosets). In this case, the generating set of the BLVQ is

$$\mathcal{V} = \{2^{l_j}\boldsymbol{u}_j\colon l_j \in \{0, \cdots, k_j - 1\}, \qquad j \in \{1, \cdots, L\}\}$$

and the index i of the vector

$$\boldsymbol{y}_i = \sum_{j=1}^{L} m_j \boldsymbol{u}_j$$

is the concatenation of the binary representations of the lattice coefficients $m_1, m_2, \cdots, m_L$. The codebook of this binary lattice VQ contains the origin ($\boldsymbol{y}_0 = \boldsymbol{0}$). By choosing $\boldsymbol{y}_0 \neq \boldsymbol{0}$ (while keeping the same generating set $\mathcal{V}$) other BLVQ's 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 binary lattice VQ is similar to the nonredundant version of the "VQ by a Linear Mapping of a Block Code" (LMBC-VQ) presented in [44]–[46]. The jth codevector of an LMBC-VQ is defined as

$$\boldsymbol{y}_j = \Theta \boldsymbol{b}_j$$

where Θ is a $d \times (n+1)$ real matrix with columns $\{\boldsymbol{\theta}_i\}_{i=0}^{n+1}$, and the $(n+1)$-dimensional column vector with $\boldsymbol{b}_j \in \{-1, 1\}^{n+1}$ is obtained from the jth codeword of a systematic (n, k) linear code by the mapping $b \rightarrow (-1)^b = 1 - 2b$ for each bit, and a leading 1 is prepended to allow translation of the codebook by $\boldsymbol{\theta}_0$. In the nonredundant case (i.e., $n = k$), $\boldsymbol{b}_j = [1, (-1)^{j_{k-1}}, \cdots, (-1)^{j_0}]^t$. Hence

$$\begin{aligned}\boldsymbol{y}_j &= \boldsymbol{\theta}_0 + \sum_{l=1}^{k} \boldsymbol{\theta}_l (1 - 2j_{k-l}) \\ &= \sum_{l=0}^{k} \boldsymbol{\theta}_l + \sum_{l=0}^{k-1} (-2\boldsymbol{\theta}_{k-l}) j_l.\end{aligned}$$

Thus setting $\boldsymbol{y}_0 = \sum_{l=0}^{k} \boldsymbol{\theta}_l$ and $\boldsymbol{v}_l = -2\boldsymbol{\theta}_{k-l}$ gives the codevector $\boldsymbol{y}_j$ in the form of a binary lattice VQ codevector. Conversely, given a binary lattice VQ we obtain a nonredundant LMBC-VQ by setting

$$\boldsymbol{\theta}_0 = \boldsymbol{y}_0 + \frac{1}{2} \sum_{l=0}^{k-1} \boldsymbol{v}_l$$

and

$$\boldsymbol{\theta}_l = -\frac{1}{2} \boldsymbol{v}_{k-l}$$

for $l = 1, \cdots, k$.

Hagen and Hedelin [44]–[46] adapted the generalized Lloyd algorithm for the design of LMBC-VQ's, and obtained locally optimal "noiseless" LMBC-VQ codebooks. Their scheme does not include error control coding, and there is no explicit mention of index assignments, either. 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].

E. The Hadamard Transform

Definition 5: For each $i, j \in \mathbb{Z}_2^k$ let $h_{i,j} = (-1)^{ij^t}$ and let $f\colon \mathbb{Z}_2^k \rightarrow \mathbb{R}$. The *Hadamard transform* $\hat{f}\colon \mathbb{Z}_2^k \rightarrow \mathbb{R}$ 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 $H = [h_{i,j}]\,(i, j \in \mathbb{Z}_2^k)$ and viewing the functions as 2^k-dimensional row vectors (i.e., $f = [f_0, f_1, \cdots, f_{2^k-1}]$):

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

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

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

or equivalently

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

where $F = [\boldsymbol{f}_0, \boldsymbol{f}_1, \cdots, \boldsymbol{f}_{2^k-1}]$ 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
$$i_m = \frac{1 - h_{i, e_m}}{2}, \qquad m \in \{0, 1, \cdots, k-1\} \tag{3}$$
where $e_m \in \mathbb{Z}_2^k$ is the binary row vector with its only nonzero component in the mth position.

- $$\sum_{i \in \mathbb{Z}_2^k} h_{i,j} = 2^k \chi_{\{j=0\}}, \qquad j \in \mathbb{Z}_2^k. \tag{4}$$

III. 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 $\hat{\boldsymbol{z}}$'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 1: Let $\boldsymbol{X} \in \mathbb{R}^d$ 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 $r_i = \mathrm{P}[\pi(\mathcal{E}_Q(\boldsymbol{X})) = i]$, $\boldsymbol{z}_i = \mathcal{D}_Q(\pi^{-1}(i))$, and $q_i = p_{j+i|j}$. Then the channel distortion in the Hadamard transform domain is
$$D_C = 4^{-k} \sum_{i \in \mathbb{Z}_2^k} \sum_{j \in \mathbb{Z}_2^k} \langle \hat{\boldsymbol{z}}_i | \hat{\boldsymbol{z}}_j \rangle \hat{r}_{i+j} (\hat{q}_0 - \hat{q}_i - \hat{q}_j + \hat{q}_{i+j}).$$

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

In Lemma 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 1 to Binary Lattice VQ's and affine index assignments.

Theorem 1: The channel distortion of a 2^k-point binary lattice vector quantizer with generating set $\{\boldsymbol{v}_l\}_{l=0}^{k-1}$, 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
$$\begin{aligned}
D_C = \frac{1}{4} \sum_{l=0}^{k-1} \sum_{m=0}^{k-1} \langle \boldsymbol{v}_l | \boldsymbol{v}_m \rangle \hat{p}_{e_l + e_m} \\
\cdot (\hat{q}_0 - \hat{q}_{e_l \mathrm{F_I}} - \hat{q}_{e_m \mathrm{F_I}} + \hat{q}_{(e_l + e_m)\mathrm{F_I}})
\end{aligned} \tag{5}$$
where
$$\hat{q}_i = 2^{k-n} \sum_{a \in (\mathcal{C}^\perp + i\mathrm{F_C})} \hat{J}_a (1 - 2\epsilon)^{w(a)} \tag{6}$$
$\mathrm{F_I} = (\mathrm{G_I^{-1}})^t$, $\mathrm{F_C}$ is a $k \times n$ binary matrix satisfying $\mathrm{G_C F_C^t} = \mathcal{I}_k$, $\mathcal{C}^\perp$ 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 lth 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 lth position.

Theorem 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 (1) requires $O(N^2)$ complexity for a codebook of size $N = 2^k$, whereas using (5) reduces the complexity of computing D_C to $O(N \log^2 N)$. (In (1) each of the nested sums contributes a factor of N, whereas in (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 $\mathrm{G'_C} = \mathrm{G_I G_C}$. Then $\mathrm{F'_C} = \mathrm{F_I F_C}$, the transpose of an inverse of $\mathrm{G'_C}$. To obtain $\hat{q}_{e_l \mathrm{F_I}}$, the sum in (6) is taken over the coset of the dual code containing $e_l \mathrm{F_I F_C} = e_l \mathrm{F'_C}$, the lth row of $\mathrm{F'_C}$.

Proof: We find expressions for the Hadamard transform quantities of Lemma 1. The transform of the BLVQ codevectors is

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

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

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

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

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

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

$$\hat{q}_i = \sum_{j\in\mathbb{Z}_2^k} \left(\sum_{r\in\mathcal{S}} \rho_{r+j\mathrm{G_C}}\right) h_{i,j} \tag{7}$$

$$\begin{aligned}
\hat{q}_i &= \sum_{j\in\mathbb{Z}_2^k} h_{i,j} \sum_{r\in\mathcal{S}} 2^{-n} \sum_{a\in\mathbb{Z}_2^n} \hat{\rho}_a h_{a,r+j\mathrm{G_C}} \\
&= 2^{-n} \sum_{a\in\mathbb{Z}_2^n} \hat{\rho}_a \left(\sum_{r\in\mathcal{S}} h_{r,a}\right) \left(\sum_{j\in\mathbb{Z}_2^k} h_{i,j} h_{a,j\mathrm{G_C}}\right) \\
&= 2^{-n} \sum_{a\in\mathbb{Z}_2^n} \hat{\rho}_a \left(\sum_{r\in\mathbb{Z}_2^n} J_r h_{r,a}\right) \left(\sum_{j\in\mathbb{Z}_2^k} h_{j,i+a\mathrm{G_C}^t}\right) \\
&= 2^{-n} \sum_{a\in\mathbb{Z}_2^n} \hat{\rho}_a \hat{J}_a 2^k \chi_{\{i=a\mathrm{G_C}^t\}} \\
&= 2^{k-n} \sum_{a\in(\mathcal{C}^\perp+i\mathrm{F_C})} \hat{\rho}_a \hat{J}_a
\end{aligned} \tag{8}$$

where

$$\begin{aligned}
\hat{\rho}_a &= \sum_{i\in\mathbb{Z}_2^k} \epsilon^{w(i)} (1-\epsilon)^{n-w(i)} h_{i,a} \\
&= \sum_{i_0\in\{0,1\}} \sum_{i_1\in\{0,1\}} \cdots \sum_{i_{n-1}\in\{0,1\}} \left(\prod_{l=0}^{n-1} \epsilon^{i_l}(1-\epsilon)^{1-i_l}\right) \\
&\quad \cdot \left(\prod_{l=0}^{n-1} (-1)^{i_l a_l}\right) \\
&= \prod_{l=0}^{n-1} \sum_{i_l\in\{0,1\}} \epsilon^{i_l}(1-\epsilon)^{1-i_l}(-1)^{i_l a_l} \\
&= \prod_{l=0}^{n-1} (1-\epsilon+\epsilon(-1)^{a_l}) \\
&= (1-2\epsilon)^{w(a)}
\end{aligned}$$

which completes the proof. □

A. 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

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

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

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

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

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

as observed in [23], [33], and [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 lth row is selected to be the index of a largest $\hat{q}_i$ that it is linearly independent of the first $l-1$ rows. More formally

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

where f_l denotes the lth row of F_I.

It was shown in [33] 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 BLVQ's. 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).

IV. Nonredundant Codes for the BLVQ

A. Binary-Symmetric Channels

Theorem 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 1: The channel distortion of a 2^k-point binary lattice vector quantizer with generating set $\{\boldsymbol{v}_l\}_{l=0}^{k-1}$, 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

$$D_C = \frac{1}{4} \sum_{l=0}^{k-1} \sum_{m=0}^{k-1} \langle \boldsymbol{v}_l | \boldsymbol{v}_m \rangle \hat{p}_{e_l + e_m} (1 - (1-2\epsilon)^{w(e_l F)} - (1-2\epsilon)^{w(e_m F)} + (1-2\epsilon)^{w((e_l+e_m)F)}) \tag{9}$$

where $w(\cdot)$ denotes Hamming weight, $\mathrm{F} = (\mathrm{G}^{-1})^t$, $\hat{p}_l$ is the lth 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 lth position.

1) Formulas for Common Index Assignments: One useful consequence of Theorem 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 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 $\boldsymbol{Y}$ be a random vector distributed according to $\{p_i\}$ over the quantizer codevectors with mean $\overline{\boldsymbol{Y}}$ and $\sigma_{\boldsymbol{Y}}^2 = \mathrm{E}||\boldsymbol{Y} - \overline{\boldsymbol{Y}}||^2$, and let $\boldsymbol{U}$ be a random vector uniformly distributed over the quantizer code points with mean $\overline{\boldsymbol{U}}$ and $\sigma_{\boldsymbol{U}}^2 = \mathrm{E}||\boldsymbol{U} - \overline{\boldsymbol{U}}||^2$. Then

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

$$\begin{aligned} \sigma_{\boldsymbol{U}}^2 &= 2^{-k} \sum_{i \in \mathbb{Z}_2^k} ||\boldsymbol{y}_i - \overline{\boldsymbol{U}}||^2 \\ &= 2^{-k} \sum_{i \in \mathbb{Z}_2^k} \left\| \boldsymbol{y}_0 + \sum_{l=0}^{k-1} \boldsymbol{v}_l i_l - \boldsymbol{y}_0 - \frac{1}{2} \sum_{l=0}^{k-1} \boldsymbol{v}_l \right\|^2 \\ &= 2^{-k} \sum_{i \in \mathbb{Z}_2^k} \left\| \sum_{l=0}^{k-1} \boldsymbol{v}_l (i_l - \frac{1}{2}) \right\|^2 \\ &= \frac{1}{4} \sum_{l=0}^{k-1} \sum_{m=0}^{k-1} \langle \boldsymbol{v}_l | \boldsymbol{v}_m \rangle 2^{-k} \sum_{i \in \mathbb{Z}_2^k} h_{i, e_l} h_{i, e_m} \qquad (10) \\ &= \frac{1}{4} \sum_{l=0}^{k-1} ||\boldsymbol{v}_l||^2 \qquad (11) \end{aligned}$$

where (3) and (4) were used to obtain (10) and (11), respectively.

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

Corollary 2: Given the conditions of Corollary 1, the channel distortion of the Natural Binary Code is

$$D_C^{(\mathrm{NBC})} = 4\epsilon((1-\epsilon)\sigma_{\boldsymbol{U}}^2 + \epsilon(\sigma_{\boldsymbol{Y}}^2 + ||\overline{\boldsymbol{Y}} - \overline{\boldsymbol{U}}||^2)).$$

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

Proof: Using Corollary 1, and the fact that

$$w(e_l F^{(\mathrm{NBC})}) = 1$$

and

$$w((e_l + e_m) F^{(\mathrm{NBC})}) = 2\chi_{\{l \neq m\}}$$

for all l and m, we have

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

□

Corollary 3: Given the conditions of Corollary 1, the channel distortion of the Folded Binary Code is

$$D_C^{(\mathrm{FBC})} = 4\epsilon(1-\epsilon)(\sigma_{\boldsymbol{U}}^2 + \sigma_{\boldsymbol{Y}}^2 + ||\overline{\boldsymbol{Y}} - \overline{\boldsymbol{U}}||^2) - \epsilon(1-2\epsilon)\max_l ||\boldsymbol{v}_l||^2.$$

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

$$w(e_l F^{(\mathrm{FBC})}) = \begin{cases} 1, & l = k-1 \\ 2, & l < k-1 \end{cases}$$

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

and thus

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

Substituting these into (9), and using (11) we obtain

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

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

Corollary 4: Given the conditions of Corollary 1, the channel distortion of the Gray Code is

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

Proof: Substituting

$$w(e_l F^{(\mathrm{GC})}) = k - l$$

and

$$w((e_l + e_m)F^{(\mathrm{GC})}) = |l - m|$$

in (9), the result is immediate. □

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. Corollaries 2 and 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 5 characterizes sources for which the FBC outperforms the NBC, using BLVQ. The variance of the source determines which code is better.

Corollary 5: Given the conditions of Corollary 1, and for all $\epsilon < 1/2$

$$D_C^{(\mathrm{FBC})} < D_C^{(\mathrm{NBC})} \Leftrightarrow \sigma_{\boldsymbol{Y}}^2 + ||\overline{\boldsymbol{Y}} - \overline{\boldsymbol{U}}||^2 < \frac{1}{4}\max_l ||\boldsymbol{v}_l||^2.$$

B. Codes for Binary-Asymmetric Channels

Definition 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

$$i \prec j \Leftrightarrow w(i+j) = w(j) - w(i)\ \forall i, j \in \mathbb{Z}_2^k.$$

Theorem 2: If a 2^k-point binary lattice vector quantizer with generating set $\{\boldsymbol{v}_l\}_{l=0}^{k-1}$ 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

$$\begin{aligned} D_C &= \frac{1}{2} \sum_{l=0}^{k-1} ||\boldsymbol{v}_l||^2 \Big(1 - \gamma^{w(e_l F)}\Big) \\ &\quad + \frac{1}{4} \sum_{\substack{l=0 \\ l\neq m}}^{k-1} \sum_{m=0}^{k-1} \langle \boldsymbol{v}_l | \boldsymbol{v}_m \rangle h_{t,\,(e_l+e_m)F} \beta^{w((e_l+e_m)F)} \\ &\quad \cdot (1 - \chi_{\{(e_l+e_m)F \prec e_l F\}} \gamma^{w(e_m F)} \\ &\quad - \chi_{\{(e_l+e_m)F \prec e_m F\}} \gamma^{w(e_l F)}) \end{aligned} \tag{12}$$

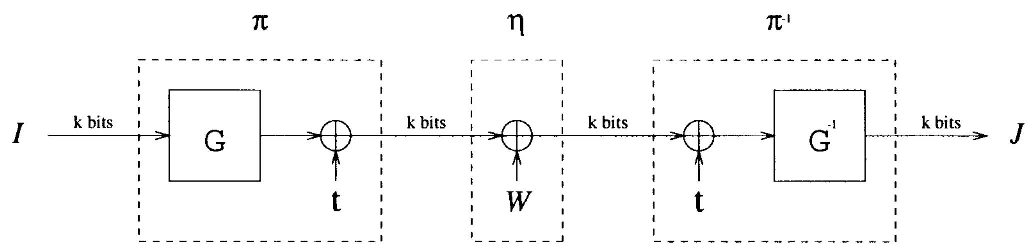

Fig. 3. Channel subsystem with affine index assignment.

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 lth 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 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 1 is necessary. Let the random variable $I=\mathcal{E}_Q(\boldsymbol{X})$ 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 Fig. 3.

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

$$\begin{aligned}D_C &= \mathrm{E}\|\boldsymbol{y}_J-\boldsymbol{y}_I\|^2\\ &=\mathrm{E}\left\|\sum_{l=0}^{k-1}\boldsymbol{v}_l(J_l-I_l)\right\|^2\\ &=\mathrm{E}\left\|\sum_{l=0}^{k-1}\boldsymbol{v}_l\left(\frac{1-h_{I+W\mathrm{G}^{-1},e_l}}{2}-\frac{1-h_{I,e_l}}{2}\right)\right\|^2\\ &=\frac{1}{4}\sum_{l=0}^{k-1}\sum_{m=0}^{k-1}\langle\boldsymbol{v}_l|\boldsymbol{v}_m\rangle\mathrm{E}[h_{I,e_l}(1-h_{W,e_lF})\\ &\quad\cdot h_{I,e_m}(1-h_{W,e_mF})]\\ &=\frac{1}{4}\sum_{l=0}^{k-1}\sum_{m=0}^{k-1}\langle\boldsymbol{v}_l|\boldsymbol{v}_m\rangle\mathrm{E}[h_{I,e_l+e_m}\mathrm{E}[(1-h_{W,e_lF}\\ &\quad -h_{W,e_mF}+h_{W,(e_l+e_m)F})|I]].\end{aligned}\tag{13}$$

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

$$\mathrm{P}[W_l=1|(IG+t)_l=0]=\epsilon$$
$$\mathrm{P}[W_l=1|(IG+t)_l=1]=\delta$$

on a binary-asymmetric channel. Thus

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

Hence, for any k-bit binary row vector f

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

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

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

Substituting e_lF, e_mF, and $(e_l+e_m)F$ for f, the last three terms within the expectations in (13) are obtained. Noting that $\mathrm{E}[h_{I,e_l+e_m}]=\chi_{\{l=m\}}$ and factoring out common terms gives

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

and the proof is complete. □

1) Formulas for Structured Index Assignments: Here we specialize Theorem 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] and [32] to nonsymmetric channels. Also, by letting $\epsilon=\delta$, the special cases of

Corollaries 2–4 for the uniform output distribution case are recovered.

Corollary 6: Given the conditions of Theorem 2, the channel distortion of the affine translate of the Natural Binary Code corresponding to translation vector t is

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

where $\alpha = \epsilon + \delta = 1 - \gamma$, and

$$\boldsymbol{y}_t = \boldsymbol{y}_0 + \sum_{l=0}^{k-1} \boldsymbol{v}_l t_l.$$

In particular, the channel distortion of the NBC ($t = [00 \cdots 0]$) is

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

and the channel distortion of the TCC ($t = [10 \cdots 0]$) is

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

Proof: Since

$$w(e_l F^{(\mathrm{NBC})}) = 1$$

and

$$w((e_l + e_m) F^{(\mathrm{NBC})}) = 2\chi_{\{l \neq m\}}$$

for all l and m, no row of $F^{(\mathrm{NBC})}$ can precede the sum of two rows in the partial ordering. Using this and Theorem 2, the statement follows. □

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

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

and

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

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

Corollary 7: Given the conditions of Theorem 2, the channel distortion of the Folded Binary Code is

$$\begin{aligned} D_C^{(\mathrm{FBC})} = \frac{1}{2}\Bigg(&\alpha \sum_{l=0}^{k-1} \|\boldsymbol{v}_l\|^2 + \alpha(1-\alpha) \sum_{l=0}^{k-2} \|\boldsymbol{v}_l\|^2 \\ &+ \beta^2 \sum_{l=0}^{k-2}\sum_{m=0}^{l-1} \langle \boldsymbol{v}_l | \boldsymbol{v}_m\rangle - \alpha\beta \sum_{l=0}^{k-2} \langle \boldsymbol{v}_l | \boldsymbol{v}_{k-1}\rangle \Bigg) \end{aligned}$$

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

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

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

Proof: For $l \neq m$

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

and thus the indicator functions in Theorem 2 will only be nonzero if either $l = k-1$ or $m = k-1$. Using this and $t = [01 \cdots 1]$, and substituting the Hamming weights of the rows of $\mathrm{F}^{(\mathrm{FBC})}$ in (12), one gets

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

□

Corollary 8: Given the conditions of Theorem 2, the channel distortion of the Gray Code is

$$\begin{aligned} D_C^{(\mathrm{GC})} = \frac{1}{2}\Bigg(&\sum_{l=0}^{k-1} \|\boldsymbol{v}_l\|^2(1 - \gamma^{k-l}) \\ &+ \sum_{l=0}^{k-1}\sum_{m=0}^{l-1} \langle \boldsymbol{v}_l | \boldsymbol{v}_m\rangle \beta^{l-m}(1 - \gamma^{k-l})\Bigg). \end{aligned}$$

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

$$\begin{aligned} D_C^{(\mathrm{GC})} &= \\ &\frac{\Delta^2}{2}\Bigg(\frac{4^k - 1}{3} - \gamma\frac{4^k - \gamma^k}{4 - \gamma} + \frac{\beta}{2 - \beta} \\ &\cdot \left(\frac{4^k - 1}{3} - \gamma\frac{4^k - \gamma^k}{4 - \gamma} - \frac{(2\beta)^k - 1}{2\beta - 1} + \gamma\frac{(2\beta)^k - \gamma^k}{2\beta - \gamma}\right)\Bigg). \end{aligned}$$

$$\begin{array}{lll}
\text{i)} & D_C^{(\mathrm{FBC})} < D_C^{\overline{(\mathrm{FBC})}} & \forall k > 1, \forall \epsilon \neq \delta \\
\text{ii)} & D_C^{\overline{(\mathrm{GC})}} < D_C^{(\mathrm{GC})} & \forall k > 1, \forall \epsilon \neq \delta \\
\text{iii)} & D_C^{(\mathrm{TCC})} < D_C^{(\mathrm{NBC})} & \forall k > 1, \forall \epsilon \neq \delta \\
\text{iv)} & D_C^{(\mathrm{FBC})} < D_C^{\overline{(\mathrm{GC})}} & \forall k > 2, \forall \epsilon, \delta \\
\text{v)} & D_C^{(\mathrm{TCC})} < D_C^{(\mathrm{FBC})} & \forall k > 1 \quad \text{if } \epsilon + \delta + 5\epsilon^2 - 8\epsilon\delta - \delta^2 \geq 0 \\
 & & \forall k \leq 1 + \log_2 \dfrac{(\epsilon+\delta)(\epsilon+\delta-1)}{\epsilon+\delta+5\epsilon^2-8\epsilon\delta-\delta^2} \\
 & & \text{if } \epsilon + \delta + 5\epsilon^2 - 8\epsilon\delta - \delta^2 < 0 \\
 & & \text{and } \epsilon + \delta + 2\epsilon^2 - 5\epsilon\delta - \delta^2 \geq 0, \\
\text{vi)} & D_C^{(\mathrm{NBC})} < D_C^{(\mathrm{FBC})} & \forall k > 1 \quad \text{if } \epsilon + \delta - \epsilon^2 + 4\epsilon\delta - 7\delta^2 \geq 0 \\
 & & \forall k \leq 1 + \log_2 \dfrac{(\epsilon+\delta)(\epsilon+\delta-1)}{\epsilon+\delta-\epsilon^2+4\epsilon\delta-7\delta^2} \\
 & & \text{if } \epsilon + \delta - \epsilon^2 + 4\epsilon\delta - 7\delta^2 < 0 \\
 & & \text{and } \epsilon + \delta - \epsilon^2 + \epsilon\delta - 4\delta^2 \geq 0.
\end{array}$$

Proof: The statement follows by observing that the precedence $(e_l+e_m)F \prec e_mF$ is satisfied if and only if $l > m$, and substituting $w(e_lF) = k - l$, $w((e_l + e_m)F) = l - m$ for $l > m$, and $t = 0$ in (12). □

2) Affine Translates of the NBC: The family of affine translates of the NBC is known to perform optimally for BLVQ's 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 9: If a 2^k-point binary lattice vector quantizer induces equiprobable quantizer codevectors for a given source, and 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

$$t = \operatorname*{argmin}_{\{i \in \mathbb{Z}_2^k\}} ||\boldsymbol{y}_i - \overline{\boldsymbol{U}}||$$

where

$$\overline{\boldsymbol{U}} = 2^{-k} \sum_{i \in \mathbb{Z}_2^k} \boldsymbol{y}_i$$

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 6. □

For a uniform scalar quantizer with step size Δ, $v_l = 2^l\Delta$, and $\overline{U} = y_0 + (2^{k-1} - \frac{1}{2})\Delta$. Thus both

$$t = [01 \ \cdots \ 1] \quad (y_t = y_0 + (2^{k-1} - 1)\Delta)$$

and

$$t = [10 \ \cdots \ 0] \quad (y_t = y_0 + 2^{k-1}\Delta)$$

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]).

3) Comparisons for Uniform Scalar Quantization: Based on the formulas presented in Corollaries 6–9 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 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 $\overline{t}$, 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 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 $\epsilon + \delta < 1$) the inequalities shown at the top of this page. The inequalities i), ii), and iii) hold with equality if $\epsilon = \delta$.

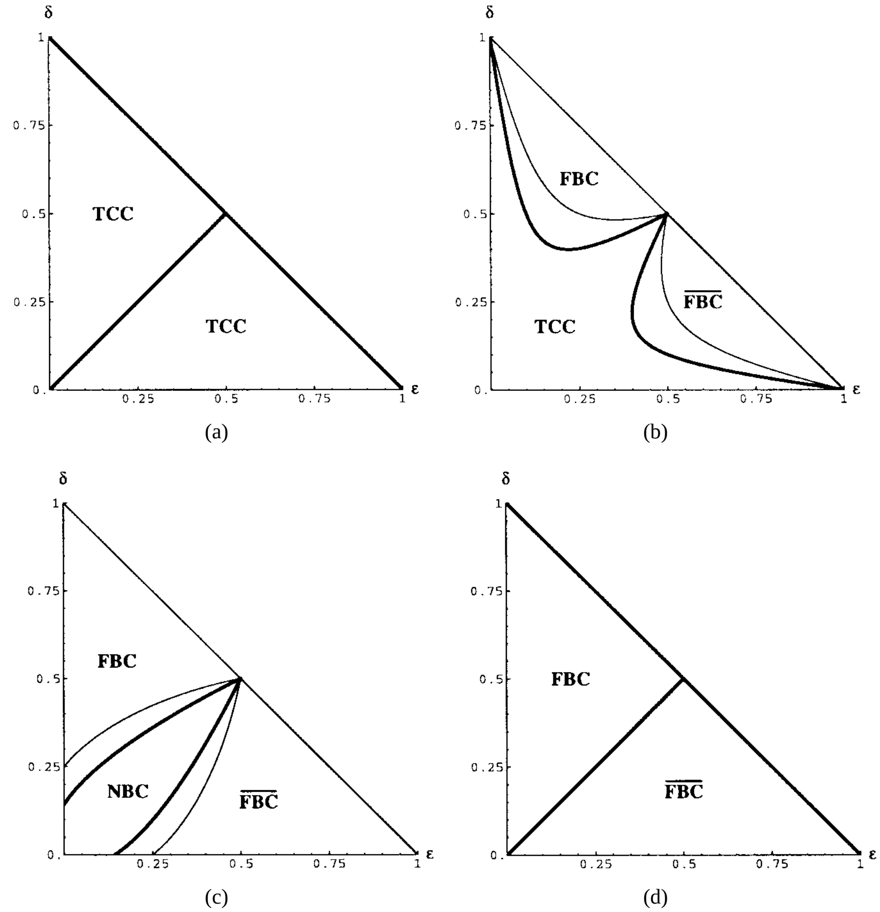

Fig. 4. Performance comparisons of various nonredundant codes for a uniform scalar source on a binary-asymmetric channel with $p_{1|0} = \epsilon$ and $p_{0|1} = \delta$, and $\epsilon + \delta < 1$. The region where one code is uniformly better (i.e., $\forall\, k > 2$) than the other is marked by the 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 k. The thick curves correspond to ties between the codes. (a) TCC versus NBC. (b) TCC versus FBC. (c) NBC versus FBC. (d) FBC versus GC.

The above inequalities follow from Corollaries 6–8 by straightforward algebraic manipulations; thus their proofs are omitted. The code comparisons of the above theorem are shown in Fig. 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 k.

REFERENCES

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

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

[3] J. Dunham and R. M. Gray, "Joint source and noisy channel trellis encoding," *IEEE Trans. Inform. 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. Inform. Theory*, vol. IT-33, pp. 855–865, Nov. 1987.

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

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

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

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

[10] D. Miller and K. Rose, "Combined source-channel vector quantization using deterministic annealing," *IEEE Trans. Commun.*, vol. 42, pp. 347–356, Feb./Mar./Apr. 1994.

[11] K. Chang and R. W. Donaldson, "Analysis, optimization, and sensitivity study of differential PCM systems operating on noisy communication channels," *IEEE Trans. Commun.*, 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, Mar. 1996), pp. 3–12.

[13] ——, "Image coding using robust quantization for noisy digital transmission," submitted to *IEEE Trans. Image Processing*.

[14] J. W. Modestino and D. G. Daut, "Combined source-channel coding of images," *IEEE Trans. Commun.*, vol. COM-27, pp. 1644–1659, Nov. 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. Commun.*, vol. COM-29, pp. 1261–1274, Sept. 1981.

[16] D. R. Comstock and J. D. Gibson, "Hamming coding of DCT-compressed images over noisy channels," *IEEE Trans. Commun.*, 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. Select. Areas Commun.*,

vol. 12, pp. 1584–1593, Dec. 1994.

[18] B. Hochwald and K. Zeger, "Tradeoff between source and channel coding," *IEEE Trans. Inform. Theory*, vol. 43, pp. 1412–1424, Sept. 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 *Proc. IEEE Int. Conf. on Communications* (Seattle, WA, June 1987), vol. 1, pp. 1128–1132.

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

[22] K. Zeger and A. Gersho, "Pseudo-gray coding," *IEEE Trans. Commun.*, vol. 38, pp. 2147–2158, Dec. 1990.

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

[24] D. J. Goodman and T. J. Moulsley, "Using simulated annealing to design transmission codes for analogue sources," *Electron. Lett.*, 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, Mar. 1989), pp. 729–734.

[26] K. C. Chua, W. C. Wong, and K. N. Ngan, "Error detection and correction of vector quantised digital images," *Proc. Inst. Elec. Eng.*, vol. 137, pp. 417–423, Dec. 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* (Minneapolis, MN, Apr. 1993), vol. V, pp. 233–236.

[28] 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.

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

[30] Y. Yamaguchi and T. S. Huang, "Optimum binary fixed-length block codes," MIT Res. Lab. of Electronics, Cambridge, MA, Quart. Progr. Rep. 78, July 1965.

[31] T. S. Huang, "Optimum binary code," MIT Res. Lab. of Electronics, Cambridge, MA, Quart. Progr. Rep. 82, July 1966.

[32] T. S. Huang *et al.*, "Design considerations in PCM transmission of low-resolution monochrome still pictures," *Proc. IEEE*, vol. 55, pp. 331–335, Mar. 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. Inform. Theory*, vol. IT-15, pp. 72–78, Jan. 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. Inform. Theory*, vol. 41, pp. 2031–2037, Nov. 1995.

[35] I. Dostis, "The effect of digital errors on PCM transmission of compandored speech," *Bell Syst. Tech. J.*, vol. 44, pp. 2227–2243, Dec. 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, Nov. 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, Oct. 1979.

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

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

[40] G. R. Redinbo and G. A. Wolf, "On minimum mean-square error linear block codes when the data have q-adic weighting," *Inform. Contr.*, 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. Inform. 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, Sept. 1994), pp. 503–512.

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

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

[46] ——, "Robust vector quantization by a linear mapping of a block-code," preprint.

[47] P. Knagenhjelm, "Robust vector quantization (Tutorial)," Dept. of Information Theory, Chalmers Univ. of Technology, Tech. Rep. 12, May 1993.

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

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

[50] P. Hedelin, P. Knagenhjelm, and M. Skoglund, "Theory for transmission of vector quantization data," in *Speech Coding and Synthesis*, W. B. Kleijn and K. K. Paliwal, Eds. Amsterdam, The Netherlands: Elsevier, 1995, ch. 10.

[51] 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, Mar. 1995), p. 433.

[52] ——, "Affine index assignments for binary lattice quantization with channel noise," in *Proc. IEEE Int. Symp. on Information Theory* (Whistler, BC, Canada, Sept. 1995), p. 377.