OPTIMAL SPARSE-QIM CODES FOR ZERO-RATE BLIND WATERMARKING

Pierre Moulin, Anil K. Goteti and Ralf Koetter

University of Illinois
Beckman Inst., Coord. Sci. Lab & ECE Dept.
405 N. Mathews Ave., Urbana, IL 61801
Email: {*moulin,goteti,koetter*}*@ifp.uiuc.edu*

ABSTRACT

The problem of blind watermarking of an arbitrary host signal in $\mathbb{R}^n$ under squared-error distortion constraints and Gaussian attacks is considered in this paper. While distortion-compensated lattice quantization index modulation (QIM) using nearly spherical Voronoi cells is known to be asymptotically capacity-achieving in this setup, our results suggest that such schemes are suboptimal in terms of error probability when the number of possible messages is subexponential in n. Our conjecture is substantiated by examples involving low-dimensional lattices and is related to the simplex conjecture in coding theory.

1. INTRODUCTION

Much of the current research in blind watermarking has focused on the development of QIM schemes [1], which are connected to fundamental information-theoretic binning ideas and outperform spread-spectrum modulation (SSM) techniques in various scenarios. Lattice QIM schemes in particular can be implemented relatively simply. Erez and Zamir [2, 3, 4] recently proved that the family of lattice QIM schemes contains capacity-achieving codes for the Gaussian channel (aka watermarking with quadratic distortion constraints for the embedder and Gaussian noise attacks). These codes are random linear codes, and loosely speaking, the associated Voronoi cells are high-dimensional and nearly spherical.

In some watermarking problems however, only a small number of bits need to be embedded, i.e., the transmission rate is well below capacity. For such problems, we conjecture in this paper that low-dimensional QIM lattice schemes applied to a few signal components are ideal in an error probability sense. This is analogous to the classical problem of communicating a few bits over a Gaussian channel. Under a transmit power constraint, the optimal codes are not random-like but have a simple geometrical property: they form the vertices of a low-dimensional simplex.

The sparse lattice QIM codes we study are related to Chen and Wornell's Spread Transform Dither Modulation (STDM) codes [1]: we quantify their advantage over other lattice QIM schemes and show they come very close to the performance of optimal *private* schemes in which the host signal is known to the receiver.

WORK SUPPORTED BY NSF GRANTS CCR 00-81268, CCR 02-08809, AND CDA 96-24396.

2. MATHEMATICAL MODEL

A message $m \in \{1, 2, \cdots, M\}$ is to be embedded into a host signal $s \in \mathbb{R}^n$, using a secret key $k \in \mathcal{K}$. The marked signal is of the form $x = f(s, m, k)$, where f is the embedding function. The embedding is subject to the distortion constraint $\mathbb{E}\|x - s\|^2 \leq nD_e$ for all $s \in \mathbb{R}^n$, where D_e is the squared-error distortion per sample, and the expectation is with respect to k and m. The variables k and m are assumed to be random, independent, and uniformly distributed. No statistical model for s is needed.

The marked signal x is subject to Gaussian attacks: $y = x+w$, where y is the degraded signal, and w is white Gaussian noise with mean zero and variance D_w.

A decoder produces an estimate $\hat{m} \in \{1, 2, \cdots, M\}$ of the original message based on the degraded signal and the secret key: $\hat{m} = g(y, k)$, where g is the decoding function. The host signal s is not available to the decoder.

2.1. QIM

The embedding methods considered in this paper are in the class of lattice-based QIM methods, which are capacity-achieving under the above mathematical model [2, 3, 4]. Recall the ingredients of a lattice-based QIM scheme:

1. A coarse lattice $\Lambda = \{x \ : \ x = \sum_{i=1}^{L} \zeta_i g_i, \ \zeta \in \mathbb{Z}^L\}$, which is the set of all integral combinations of basis vectors $g_1, g_2, \cdots, g_L$ in $\mathbb{R}^L$;
2. a quantizer function $Q \ : \ \mathbb{R}^L \rightarrow \Lambda$ mapping each vector in $\mathbb{R}^L$ to the nearest (in the Euclidean metric) lattice point;
3. a set $\mathcal{C}$ of M minimum-length vectors $c_1, \cdots, c_M$ and associated cosets $\Lambda_m \stackrel{\triangle}{=} c_m + \Lambda$, $1 \leq m \leq M$. The vectors $\{c_m\}$ are termed coset vectors, or dither vectors. The union of cosets $\mathcal{C} + \Lambda = \cup_{m=1}^{M} \Lambda_m$ may (but need not) form a lattice;
4. a lattice inflation factor ("Costa parameter") $\alpha \in (0, 1)$.

The embedding function for a subvector (block) $s \in \mathbb{R}^L$ is then defined as

$$x = f(s, m) = Q(\alpha s - c_m) + c_m + (1 - \alpha)s \quad \in \mathbb{R}^L. \quad (1)$$

The decoder is a lattice decoder:

$$\hat{m} = g(y) = \mathrm{argmin}_{1 \leq m \leq M} \mathrm{dist}(\alpha y, \Lambda_m) \quad (2)$$

where $\mathrm{dist}(x, \Lambda_m) \stackrel{\triangle}{=} \min_{y \in \Lambda_m} \|x - y\|$.

2.2. Private Spread-Spectrum Watermarking

The performance of a QIM scheme cannot surpass that of an ideally designed SSM scheme in which *the decoder knows the host* s. The SSM embedding function is $x = s + \sqrt{nD_e}u_m$, where $\{u_m,\ 1 \le m \le M\}$ are unit-norm vectors. The decoder chooses among the M hypotheses:

$$H_m\ :\ y - s = \sqrt{nD_e}u_m + w, \quad w \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, D_w),$$

for $1 \le m \le M$. The maximum-likelihood decoder outputs

$$\hat{m} = \text{argmin}_{1\le m\le M}\|y - s - \sqrt{nD_e}\,u_m\|.$$

For large $SNR = nD_e/D_w$, the error probability P_e of the decoder is dominated by the two nearest codewords: $P_e \approx \Phi\left(\frac{d_{\min}}{2\sqrt{D_w}}\right)$, where $\Phi(t) = \int_t^\infty (2\pi)^{-1/2}e^{-u^2/2}\,du \le \frac{1}{2}e^{-t^2/2}$ is the tail probability of the Gaussian pdf. Here $d_{\min} = \sqrt{nD_e}\,\min_{1\le i,j\le M}\|u_i - u_j\|$. For $M = 2$ we can choose u_1 arbitrarily and let $u_2 = -u_1$; then $d_{\min} = \sqrt{4nD_e}$. For increasing M the value of $d_{\min}$ decreases to $\sqrt{2nD_e}$ for $M = 2n$ and decreases very slowly for $M > 2n$ [5]. Note the simple structure of optimal codes when $M \le n+1$: their codewords form the vertices of a M-dimensional simplex. Random codebook constructions would be clearly suboptimal for such problems.

3. SPARSE QIM

Choose $L << n$ and $\alpha = 1$ (the latter restriction will be justified in Sec. 4). Throughout, we assume the following construction: a length-L host vector is obtained by applying an orthonormal transformation to the host data n-vector, and selecting the first L transform coefficients. The ratio L/n is viewed as a sparsity, or time-sharing parameter: communication takes place during a fraction L/n of the time, but the transmission power is boosted to $D_e n/L$ during that time. When Λ is the cubic lattice $\Delta\mathbb{Z}^L$, this scheme is the STDM method [1]. The transformation is used for security and perceptual-transparency reasons. Our grand goal is to solve the following problem: **Given** M, D_e, D_w**, select** L**, a lattice** Λ **and dither vectors** $c_1, \cdots, c_M \in \mathbb{R}^L$ **to minimize the probability of error of the decoder.** This problem may be too hard to solve, but we present some conjectures and evaluate the performance of several lattice codes.

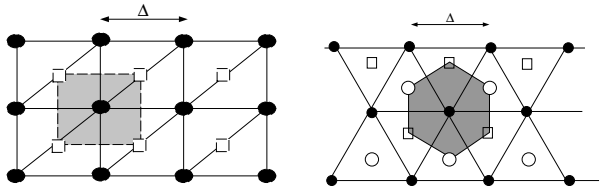
3.1. Lattice Code Properties

The Voronoi cell $\mathcal{V}$ associated with lattice Λ is the set of points that lie closer to the origin than to any other element of Λ. Voronoi cells are shown in Fig. 1 for two lattices used in this paper.

High-rate lattice quantization theory is often invoked to model the quantization noise $Q(s) - s$ as random and uniformly distributed over $\mathcal{V}$. This model is *exact* if instead of $Q(\cdot)$, one uses a dithered quantizer $Q(\cdot - d) + d$, where the dither vector d is uniformly distributed over $\mathcal{V}$ [2]. The use of a dither vector that is a function of the key k approximately satisfies this model, provided that the key is long enough.

Two geometrical properties of $\mathcal{V}$ are fundamental in our analysis. The first is the covering radius of $\mathcal{V}$:

$$\rho_{\text{cov}} \overset{\triangle}{=} \max_{x\in\mathcal{V}}\|x\| = \max_{x\in\mathbb{R}^L}\text{dist}(x, \Lambda). \tag{3}$$



(a) square lattice $\mathbb{Z}^2$ (b) hexagonal lattice A_2

Fig. 1. Voronoi cells (shaded areas) for $L = 2$: (a) square lattice; (b) hexagonal lattice. Deep holes are marked with white squares and circles.

The second is the embedding distortion per sample:

$$\overline{D}_e = \frac{1}{L}\frac{1}{\text{Vol}(\mathcal{V})}\int_{\mathcal{V}}\|x\|^2\,dx. \tag{4}$$

The total squared-error distortion is

$$nD_e = L\overline{D}_e. \tag{5}$$

Other properties of the lattice, such as the packing radius or the volume, have no bearing in the analysis.

Deep holes. The deep holes of a lattice Λ are the points in $\mathbb{R}^L$ that are furthest away from Λ, i.e., v is a deep hole $\Leftrightarrow \text{dist}(v, \Lambda) = \rho_{\text{cov}}$. In general for a given lattice Λ there is a minimal set of vectors $v_1, \cdots, v_{n_{DH}}$ such that the set of deep holes is equal to $\cup_{i=1}^{n_{DH}} v_i + \Lambda$. In the case $L = 2$, we have $n_{DH} = 1$ and $n_{DH} = 2$ for the square lattice and the hexagonal lattice, respectively; the corresponding deep holes are shown in Fig. 1. Deep holes are ideal candidates for the dither vectors, as elaborated below.

Lattice Code Minimum Distance. Define

$$\overline{d}_{\min} \overset{\triangle}{=} \min_{1\le i,j\le M}\frac{1}{\sqrt{L}}\text{dist}(\Lambda_i, \Lambda_j) \le \frac{1}{\sqrt{L}}\rho_{\text{cov}} \tag{6}$$

which is a function of $\mathcal{V}$ and the dither vectors $\{c_i, 1 \le i \le M\}$. We also define the normalized quantity

$$\gamma \overset{\triangle}{=} \overline{d}_{\min}/\sqrt{\overline{D}_e} \tag{7}$$

which is invariant to scalings of the lattice code. Without loss of generality, we often take $c_1 = 0$. To have equality in (6), c_i must be a deep hole of $c_j + \Lambda$ for each $1 \le i, j \le M$. Then the code is ideal in the following geometric sense: $\text{dist}(\Lambda_i, \Lambda_j) = \rho_{\text{cov}}$ for all i, j. This condition also requires $M \le n_{DH} + 1$.

The parameter γ defined in (7) is important in our analysis. A different normalization (using $[\text{Vol}(\mathcal{V})]^{1/L}$ instead of $\sqrt{\overline{D}_e}$) is appropriate in the study of capacity-achieving coset codes [6].

3.2. Error Probability

When n is large, $\frac{1}{n}\log_2 M << 1$ and $\alpha = 1$, the calculation of error probabilities is relatively simple. Indeed the detector must choose between M composite hypotheses:

$$H_m\ :\ y = x + w, \quad x \in \Lambda_m,\ w \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, D_w),$$

for $1 \le m \le M$. Also define

$$d_{\min} = \min_{1 \le i,j \le M} \mathrm{dist}(\Lambda_i, \Lambda_j) = \sqrt{L} d_{\min} = \gamma \sqrt{n D_e}.$$

For large n, the error probability P_e of the decoder (2) is dominated by the two closest lattices, and $P_e \approx \Phi(\frac{d_{\min}}{2\sqrt{D_w}})$. More precisely, the approximation is tight in the exponent: $\lim_{n\to\infty} -\frac{1}{n} \ln P_e = \frac{\gamma^2 D_e}{8 D_w}$. Hence **the problem of minimizing P_e is asymptotically equivalent to the problem of maximizing γ.**

3.3. Are nearly spherical Voronoi cells desirable?

While lattice codes with nearly spherical Voronoi cells are asymptotically (as $n \to \infty$) capacity-achieving (i.e., $\lim_{n\to\infty} \frac{1}{n} \log_2 M = C > 0$) [3, 4], the same need not be true for small, fixed M. To gain some insight into this problem, let $M = L = 2$ and compare the performance of square and hexagonal lattices. The corresponding Voronoi cells are squares and hexagons, respectively. For the square lattice, we have $\rho_{\mathrm{cov}} = \sqrt{24 \overline{D}_e}$; for the hexagonal lattice, $\rho_{\mathrm{cov}} = \sqrt{19.2 \overline{D}_e}$. Note that if $\mathcal{V}$ were a disk, we would have $\rho_{\mathrm{cov}} = \sqrt{8 \overline{D}_e}$. Consider now two watermarking codes based on Fig. 1.

Square lattice: Here $\Delta = \sqrt{6 n D_e}$ to satisfy the distortion constraint, and $\mathcal{V} = [-\frac{\Delta}{2}, \frac{\Delta}{2}]^2$. Choose $c_1 = 0$ and $c_2 = (\frac{\Delta}{2}, \frac{\Delta}{2})^T$, which is a deep hole of $\mathcal{V}$. The distance between the lattice Λ and its translate $c_2 + \Lambda$ is $\|c_2\| = \frac{\Delta}{\sqrt{2}} = \sqrt{3 n D_e} = d_{\min}$ (thus $\gamma = \sqrt{3}$).

Hexagonal lattice: $\gamma = \sqrt{12/5}$, as calculated in Sec. 3.5.

The "more spherical" of the two lattices is therefore the worse one. In fact, "nearly spherical" lattices can be constructed using random generator matrices when $L \to \infty$. It follows from [4] that $\gamma \to 1$ for such lattices, which is 42% lower than the best γ.

3.4. Construction A

While generally not optimal, Conway and Sloane's Construction A [5, Ch. 5] can be used to construct good QIM codes. Let Λ be the cubic lattice $\Delta \mathbb{Z}^L$; its Voronoi cell $\mathcal{V}$ is the cube $[-\frac{\Delta}{2}, \frac{\Delta}{2}]^L$. Due to the embedding distortion constraint, we have $\frac{n D_e}{L} = \overline{D}_e = \frac{\Delta^2}{12}$. Choose a (L, k_c, d_H) binary code $\mathcal{C}$ (where L = dimension, $M = 2^{k_c}$ = number of codewords, and d_H = minimum Hamming distance between codewords), and let c_m be the m-th codeword in $\mathcal{C}$, scaled by $\frac{\Delta}{2}$. We have $d^2_{\min} = d_H (\frac{\Delta}{2})^2$ and hence $\gamma^2 = 3 \frac{d_H}{L}$.

The Plotkin bound [5] yields $d_H \le \frac{LM}{2(M-1)}$, and a variety of good codes attain that bound. For $M = 2$, equality is achieved only for the $(L, 1, L)$ repetition code. Therefore, under Construction A, simple scalar QIM with arbitrary L is optimal when $M = 2$, and achieves $\gamma^2 = 3$. For $M = 4$, we choose $L = 3$ and the (3,2,2) code $\mathcal{C} = \{(0,0,0), (0,1,1), (1,0,1), (1,1,0)\}$; then $\gamma^2 = 2$. Note that $\mathcal{C} + 2\mathbb{Z}^3$ is D_3, the fcc lattice. For any M, we can achieve $\gamma^2 = \frac{3M}{2(M-1)}$, but no better.

3.5. General Approach

Construction A is simple but generally nonoptimal for $M > 2$. Consider for instance $M = 3$, and let Λ be the scaled hexagonal lattice A_2 of Fig. 1(b). Due to the embedding distortion constraint, we have $\frac{n D_e}{L} = \overline{D}_e = \frac{5\Delta^2}{72}$, where Δ is the distance between adjacent lattice points. We choose $c_1 = (0,0)$, $c_2 = (\frac{\Delta}{2}, \frac{\Delta}{2\sqrt{3}})$, and $c_3 = (0, \frac{\Delta}{\sqrt{3}})$; the latter two are deep holes of Λ. Here $d^2_{\min} = \frac{\Delta^2}{3} = \frac{12}{5} n D_e$, which is 20% better than the performance using cubic lattices. Note that $\mathcal{C} + A_2$ is A_2^*, the dual hexagonal lattice; and the code is ideal in a geometric sense (see Sec. 3.1).

embedding scheme $(\mathcal{C}+\Lambda)/\Lambda$		$M=2$	$M=3$	$M=4$	$M \to \infty$
$\mathbb{Z}/M\mathbb{Z}$	$L=1$	$\boxed{3}$*	$\frac{4}{3}$	$\frac{3}{4}$	$\frac{12}{M^2}$
$\mathbb{Z}^L/L\mathbb{Z}^L$	$L=\log_2 M$	3		$\frac{3}{2}$	$\frac{3}{L}$
$\mathcal{C}+2\mathbb{Z}^L/2\mathbb{Z}^L$	$L \ge \log_2 M$	3	2	$\boxed{2}$	$\boxed{\sim \frac{3}{2}}$
A_2^*/A_2	$L=2$	$\frac{12}{5}$	$\boxed{\frac{12}{5}}$*	$\frac{9}{5}$	$\sim \frac{6}{5}$
$D_3/2\mathbb{Z}^3$	$L=3$	2	2	$\boxed{2}$	1
$\mathbb{Z}^3/D_3$	$L=3$	$\frac{8}{3}$			$\sim \frac{4}{3}$
$\mathbb{Z}^4 \cup D_4^+/D_4$	$L=4$	$\frac{30}{13}$	$\frac{30}{13}$		$\sim \frac{15}{13}$
Private SSM		4	3	$\frac{8}{3}$	2

Table 1. Values of $\gamma^2 = \frac{d^2_{\min}}{n D_e}$ for different watermarking schemes and different numbers of messages. Boxes indicate the best performance we have obtained for each value of M. Asterisks indicate the performance of *ideal codes*.

The performance of various QIM schemes is summarized in Table 1. Given an arbitrary Λ and $\mathcal{C}$, we have $\frac{n D_e}{L} = \overline{D}_e = G(\Lambda)[\mathrm{Vol}(\mathcal{V})]^{2/L}$, where $G(\Lambda) \ge \frac{1}{2\pi e}$ is the normalized second moment of the lattice. [1] Then we calculate

$$\gamma^2 = \frac{d^2_{\min}}{n D_e} = \overline{d}^2_{\min}(\mathcal{C}, \Lambda) [\overline{D}_e(\Lambda)]^{-2/L}$$

using an arbitrary lattice scale parameter (e.g., $\Delta = 1$).

Consider the case $M = 2$ again. Let Λ be the checkerboard lattice $D_L = \{x : \sum_{i=1}^L x_i \text{ is an even integer}\}$. The lattice D_3 is the fcc lattice and its deep holes are the elements of the shifted lattice $(1,0,0) + D_3$. Here $\gamma^2 = \frac{8}{3}$. The deep holes of D_4 are the elements of the two shifted lattices $(1,0,0,0) + D_4$ and $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}) + D_4$. Here $\gamma^2 = \frac{30}{13} \approx 2.30$.

4. CHOICE OF α

With the choice $\alpha = 1$ used in the previous section, the noise in the decoder's statistical test (8) is Gaussian; the error exponent is given in Sec. 3.2. The choice $\alpha = 1$ finds a theoretical justification when $L << n$, as discussed below.

One may suspect that performance could be improved by using $\alpha < 1$. The motivation is Zamir, Shamai and Erez's analysis [2, 3, 4] which has proved that QIM schemes using $\alpha = \frac{D_e}{D_e + D_w}$, "nearly spherical" Voronoi cells, and $L = n$ are capacity-achieving. That is, the number of messsages is exponential: $M = 2^{nR}$, where R is just below the capacity $C = \frac{1}{2} \log_2(1 + D_e/D_w)$. The above value of α is the one that minimizes the variance $\sigma^2 = \frac{D_e D_w}{D_e + D_w}$ of the noise at the decoder. When M is fixed or subexponential in n, one could ask whether the results of the previous section could be improved by selecting $\alpha = \frac{D_e}{D_e + D_w}$ and $L = n$. The answer is not obvious – the variance of the noise is reduced

[1] For the cubic, hexagonal, D_3 (fcc) and D_4 lattices, we have $G(\Lambda) = \frac{1}{12}, \frac{5}{36\sqrt{3}}, \frac{19}{192\, 2^{1/3}}$ and $\frac{13}{120\sqrt{2}}$, respectively.

when $\alpha < 1$, but the noise becomes non-Gaussian, and *might* thus have heavier tails. In an asymptotic probability of error analysis, the latter effect could dominate.

To gain some insight into this tradeoff, we explored the choice of scalar uniform quantizers for binary hypothesis testing ($M = 2$), with L ranging from 1 to n. Here $\Delta = \sqrt{12\frac{n}{L}D_e}$. The two codebooks are: $\Delta\mathbb{Z}^L$ and $(\frac{\Delta}{2}, \cdots, \frac{\Delta}{2}) + \Delta\mathbb{Z}^L$, as illustrated in Fig. 1(a). The detector is presented with two hypotheses:

$$\begin{cases} H_0 \ : \ \alpha y \bmod \Delta = v \bmod \Delta \\ H_1 \ : \ \alpha y \bmod \Delta = (\frac{\Delta}{2} + v) \bmod \Delta \end{cases}$$

where $v = (1-\alpha)z + \alpha w$ is a weighted combination of the quantization noise z (uniformly distributed over $[-\frac{\Delta}{2}, \frac{\Delta}{2}]$) and the attack noise $w \sim \mathcal{N}(0, D_w)$. The value of α that minimizes the variance of the noise v at the detector is

$$\alpha = \frac{(n/L)D_e}{(n/L)D_e + D_w}. \tag{8}$$

Clearly $\alpha \to 1$ as $\frac{n}{L} \to \infty$, which motivated our choice of $\alpha = 1$ in the previous section. Also note that the value of α that minimizes P_e is slightly different from (8). Let $\tilde{v} = v \bmod \Delta$. The probability of error exponent is given by

$$\lim_{n\to\infty} -\frac{1}{n} \ln P_e = \beta$$

where $\beta = -\ln \int_0^{\Delta} \sqrt{p_{\tilde{V}}(\tilde{v}) p_{\tilde{V}}(\tilde{v} + \frac{\Delta}{2})}\, d\tilde{v}$ is the Bhattacharyya coefficient [7] between the rival distributions at the detector. Fig. 2 compares $\log_{10} P_e$ and the Bhattacharyya bound as a function of L for $D_e = D_w$ and $n = 15$. The Bhattacharyya bound is a useful predictor of $\log_{10} P_e$ in the sense that the gap, normalized by $\frac{1}{n}$, is indeed small. Fig. 3 shows the Bhattacharyya bound on $\log_{10} P_e$ *vs* D_e/D_w for several values of L and α selected as in (8). The standard QIM scheme ($L = n$) performs worse than the STDM scheme ($L = 1$, $\alpha \approx 1$). The suboptimality of the former is attributed to the non-Gaussian nature of the quantization noise, as discussed above. In these experiments, performance improved monotonically as a function of L.

5. DISCUSSION

For $M = 2$ and $M = 3$, we conjecture that the scalar and hexagonal QIM schemes with $L = 1$ and $L = 2$, respectively, are optimal. The codes are ideal in the geometric sense mentioned in Sec. 3.1. Remarkably, distortion compensation is not needed in this setup. The error exponents are $\frac{3}{4}$ and $\frac{4}{5}$ times the optimal error exponent in the known host case, respectively.

6. REFERENCES

[1] B. Chen and G. W. Wornell, "Quantization Index Modulation: A Class of Provably Good Methods for Digital Watermarking and Information Embedding," *IEEE Trans. Info. Thy*, Vol. 47, No. 4, pp. 1423—1443, May 2001.

[2] R. Zamir, S. Shamai (Shitz), and U. Erez, "Nested Linear/Lattice Codes for Structured Multiterminal Binning," *IEEE Trans. IT*, Vol. 48, No. 6, pp. 1250—1276, June 2002.

[3] U. Erez and R. Zamir, "Achieving $\frac{1}{2}\log(1 + SNR)$ on the AWGN Channel with Lattice Encoding and Decoding," *preprint*, May 2001; revised, Sep. 2003.

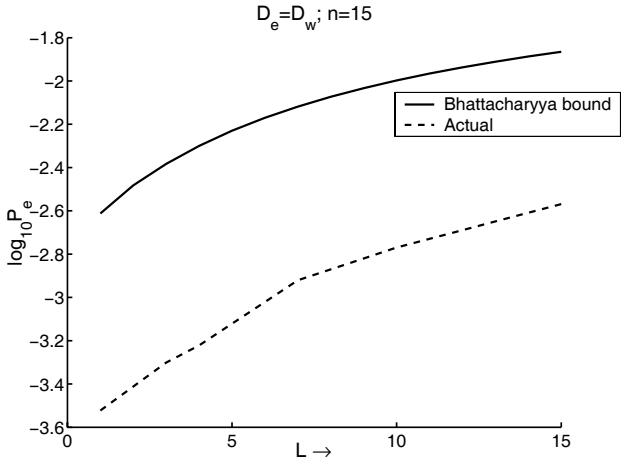


Fig. 2. Binary hypothesis testing ($M = 2$): Comparison of $\log_{10} P_e$ and Bhattacharyya bound for scalar QIM using $D_e = D_w$ and $n = 15$. The variable on the horizontal axis is L. The lattice inflation factor α is optimized for each value of L.

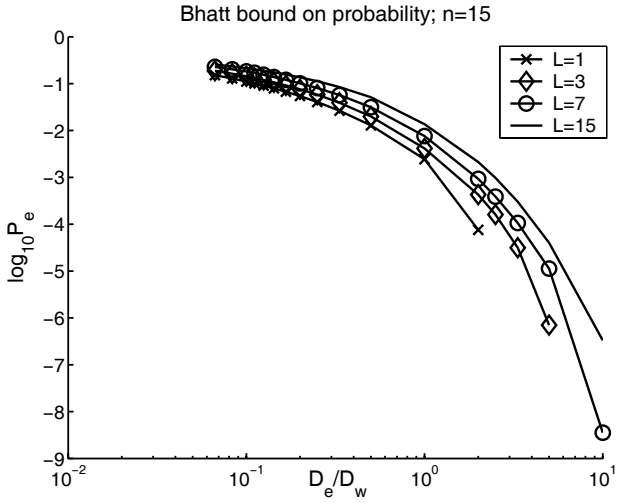


Fig. 3. Binary hypothesis testing ($M = 2$): Performance of scalar QIM (with optimized α, dashed line) applied to $L \leq n$ samples of a host vector. The traditional QIM scheme is obtained when $L = n$; STDM is obtained when $L = 1$. The latter case with $\alpha \approx 1$ is the best choice.

[4] U. Erez and R. Zamir, "Lattices which are Good for (Almost) Everything," *preprint*, June 2003.

[5] J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Groups*, 3rd ed., Springer-Verlag, New York, 1999.

[6] G. D. Forney Jr., M. D. Trott and S.-Y. Chung, "Sphere-Bound-Achieving Coset Codes and Multilevel Coset Codes," *IEEE Trans. IT*, Vol. 46, No. 3, pp. 820—850, May 2000.

[7] H. L. Van Trees, *Detection, Estimation and Modulation Theory I*, Wiley, New York, 1968.