[11] A. S. Holevo, "Towards the mathematical theory of quantum communication channels," *Probl. Pered. Inform.*, vol. 8, no. 1, pp. 63–71, 1972. (English translation: *Probl. Inform. Transm.*, vol. 8, no. 1, pp. 47–56).
[12] ——, "Some estimates of the information transmitted by quantum communication channel," *Probl. Pered. Inform.*, vol. 9, no. 3, pp. 3–11, 1973. (English translation: *Probl. Inform. Transm.*, vol. 9, no. 3, pp. 177–183).
[13] ——, "Problems in the mathematical theory of quantum communication channels," *Rep. Math. Phys.*, vol. 12, no. 2, pp. 273–278, 1977.
[14] ——, "On the capacity of quantum communication channel," *Probl. Pered. Inform.*, vol. 15, no. 4, pp. 3–11, 1979. (English translation: *Probl. Inform. Transm.*, vol. 15, no. 4, pp. 247–253).
[15] ——, *Probabilistic and Statistical Aspects of Quantum Theory*. Amsterdam, The Netherlands: North Holland, 1982.
[16] ——, "On quantum communication channels with constrained inputs," LANL Rep. quant-ph/9705054, May 30, 1997.
[17] R. Jozsa and B. Schumacher, "A new proof of the quantum noiseless coding theorem," *J. Mod. Opt.*, vol. 41, no. 12, pp. 2343–2349, 1994.
[18] M. Ohya, D. Petz, and N. Watanabe, "On capacity of quantum channels," *IEEE Trans. Inform. Theory*. to be published.
[19] A. Peres and W. K. Wootters, "Optimal detection of quantum information," *Phys. Rev. Lett.*, vol. 66, no. 9, pp. 1119–1122, 1991.
[20] B. Schumacher and M. D. Westmoreland, "Sending classical information via noisy quantum channel," *Phys. Rev. A*, 1997, to be published.
[21] A. Wehrl, "General properties of entropy," *Rev. Mod. Phys.*, vol. 50, no. 2, pp. 221–260, 1978.
[22] H. P. Yuen and M. Ozawa, "Ultimate information carrying limit of quantum systems," *Phys. Rev. Lett.*, vol. 70, no. 4, pp. 363–366, Jan. 1993.

Lattice Codes Can Achieve Capacity on the AWGN Channel

Rüdiger Urbanke and Bixio Rimoldi, *Member, IEEE*

***Abstract*— It is shown that lattice codes can achieve capacity on the additive white Gaussian noise channel. More precisely, for any rate R less than capacity and $\epsilon > 0$, there exists a lattice code with rate no less than R and average error probability upper-bounded by ϵ. These lattice codes include *all* points of the (translated) lattice within the spherical bounding region (not just the ones inside a thin spherical shell).**

***Index Terms*— AWGN channel, Blichfeldt's principle, lattice codes, Minkowski–Hlawka theorem.**

I. Introduction

Consider the additive white Gaussian noise (AWGN) channel with peak signal-power constraint S, i.e., each codeword x of an n-dimensional code for this channel must satisfy

$$||x||^2 \leq nS$$

where $||\cdot||$ is the Euclidean norm. It is well known [1] that the capacity of this channel is

$$C = \frac{1}{2}\log_2\left(1+\frac{S}{N}\right) \text{ bits/channel use}$$

where N is the variance of the independent and identically distributed (i.i.d.) Gaussian noise. The proof in [1] is based on a *random coding* argument and, hence, codes that achieve capacity may exhibit little or no structure, making them ill-suited for practical applications. It is, hence, of interest to investigate the maximal reliable transmission rates achievable by structured ensembles of codes.

An important class of structured codes is the class of lattice codes. For the purpose of this correspondence we define a lattice code C_n as the intersection of a (possibly translated) n-dimensional full-rank lattice Λ_n with a region B_n of bounded support. What is the maximal reliable transmission rate achievable by means of lattice codes?

To take full advantage of the underlying lattice structure we would like to neglect the effects of the bounding region B_n and simply decode to the nearest *lattice point* (which may or may not be a code point). This decoding procedure is often referred to as *lattice decoding*. The optimum decoding procedure, on the other hand, is *minimum-distance* decoding which maps the received point into the closest *code point*. Lattice decoding is, in general, significantly less complex and it results in a uniform probability of error among all codewords, i.e., the average and the maximum probability of error are equal. Minimum-distance decoding, on the other hand, minimizes the average probability of error but a "good" code with respect to an average probability of error criterion may contain some "bad" codewords.

The known facts about lattice codes for the AWGN channel can be summarized as follows.

1) For any rate $R < \frac{1}{2}\log_2(S/N)$, there exists a lattice code C_n with arbitrarily small (maximum) probability of error when used with lattice decoding [2]. Moreover, the bounding region B_n can be chosen to be the n-dimensional ball of radius $\sqrt{nS}$.
2) By choosing B_n to be a "thin" spherical shell centered at the origin, rates up to capacity can be achieved with arbitrarily low average probability of error using a minimum distance decoder [3], [4].

To prove the first result, de Buda used the Minkowski–Hlawka Theorem. Loeliger [5], [6] derived the same result using a standard averaging argument for linear codes applied to lattices. He also conjectured that $\frac{1}{2}\log_2(S/N)$ is indeed the highest achievable rate under lattice decoding. Regarding the second result, in [4] it was pointed out that because of the "thin" spherical bounding region these codes loose most of their structure and resemble more "random" codes.

In this correspondence we will use [3] and [4] to close one of the remaining gaps by showing that lattice codes with a bounding region equal to an n-dimensional ball as opposed to a spherical shell also achieve capacity under minimum-distance decoding. Compared to the codes used in [3] and [4], these codes exhibit more structure which should facilitate their encoding and decoding process. Given the already known results this result is not surprising since in high dimensions most of the volume within a ball lies in a thin spherical shell and, hence, one might expect that most code points also lie in this spherical shell. This intuitive idea is made precise in Lemmas 1 and 3.

The main result of this correspondence is summarized by

Theorem 1: Let $S, N,$ and $\epsilon > 0$ be given. If

$$R < \frac{1}{2}\log_2\left(1+\frac{S}{N}\right)$$

then there exists a lattice code C_n for the additive white Gaussian noise channel with peak power constraint S and noise variance N, where B_n is the n-dimensional ball of radius $\sqrt{nS}$, such that C_n

Manuscript received October 23, 1994; revised May 28, 1997. This work was supported by the National Science Foundation under Grants NCR-9357689 and NCR-9304763. The material in this correspondence was presented in part at the IEEE International Symposium on Information Theory, Whistler, BC, Canada, Sept. 17–22, 1995.

R. Urbanke is with Bell Laboratories, Murray Hill, NJ 07974 USA.
B. Rimoldi is with the Department of Electrical Engineering, ESSRL, Washington University, St. Louis, MO 63130 USA.

Publisher Item Identifier S 0018-9448(98)00005-4.

has rate lower-bounded by R and average probability of error of a minimum-distance decoder upper-bounded by ϵ.

Although the techniques applied in our proof are drawn from the ones used in [3] and [4], we introduce a simpler decoding rule which significantly simplifies the proof.

There are still several open questions concerning lattice codes. First, a proof of the conjecture that $\frac{1}{2}\log_2\frac{S}{N}$ is the capacity of lattice codes under lattice decoding is still missing. Secondly, as stated above, minimum-distance decoding leads, in general, to a nonuniform probability of error for individual codewords. Loeliger pointed out that the usual approach of deleting the worst half of all codewords is not applicable to lattice codes as this procedure destroys the structure of these codes. Hence, it still remains to be seen if lattice codes under a *maximum* probability of error condition do exist for rates close to $\frac{1}{2}\log_2\left(1+\frac{S}{N}\right)$. Both questions will be settled if one can prove that there exist a lattice code that achieves $\frac{1}{2}\log_2\left(1+\frac{S}{N}\right)$ with a lattice decoder. Forney [7] has conjectured that this is the case if one interprets S as the *average* rather than the peak power constraint.

To make this correspondence self-contained, the Minkowski–Hlawka Theorem and Blichfeldt's principle are stated in Appendix I.

II. The Proof

Let S be the peak signal energy per dimension and N be the noise variance. Let R be given such that

$$R < \frac{1}{2}\log_2\left(1+\frac{S}{N}\right).$$

There exist numbers R' and S' such that

$$R < R' < \frac{1}{2}\log_2\left(1+\frac{S'}{N}\right) < \frac{1}{2}\log_2\left(1+\frac{S}{N}\right).$$

Let T_n be the n-dimensional closed ball of radius $\sqrt{nS}$ and volume

$$V_n = \frac{(\pi n S)^{\frac{n}{2}}}{\Gamma(n/2+1)}$$

[8, p. 135], where $\Gamma(x)$ is the well-known *Gamma* function [9, p. 350]. Let T'_n be the n-dimensional closed ball of radius $\sqrt{nS'}$ and volume

$$V'_n = \frac{(\pi n S')^{\frac{n}{2}}}{\Gamma(n/2+1)}$$

see Fig. 1, and let $\partial T'_n$ be the sphere of radius $\sqrt{nS'}$ and area

$$A'_n = \frac{n\pi^{\frac{n}{2}}(nS')^{\frac{n-1}{2}}}{\Gamma(n/2+1)}.$$

Further, define $T_n^\Delta = T_n \backslash T'_n$ with volume $V_n^\Delta = V_n - V'_n$. Given a lattice Λ_n with fundamental region P_n and $s \in P_n$, define the lattice code

$$\mathcal{C}_n = \mathcal{C}_n(\Delta_n, s) = (\Lambda_n + s) \cap T_n.$$

Similarly, define the subcodes

$$\mathcal{C}'_n = \mathcal{C}'_n(\Delta_n, s) = (\Lambda_n + s) \cap T'_n$$

and

$$\mathcal{C}_n^\Delta = \mathcal{C}_n^\Delta(\Delta_n, s) = (\Lambda_n + s) \cap T_n^\Delta = \mathcal{C}_n(\Delta_n, s)\backslash \mathcal{C}'_n(\Lambda_n, s).$$

Let $M_n = M_n(\Delta_n, s), M'_n = M'_n(\Lambda_n, s)$, and $M_n^\Delta = M_n^\Delta(\Lambda_n, s)$ be the cardinalities of these codes, respectively. In the sequel we will also sometimes require a particular point x_0 with coordinates $x_0 = (\sqrt{nS'}, 0, \cdots, 0)$.

For an arbitrary code $\mathcal{C}$, $P^{\mathcal{C}}$ will denote the average probability of error under minimum-distance decoding.

Lemma 1:

$$P^{\mathcal{C}_n} \leq \frac{M'_n}{M_n} + P^{\mathcal{C}_n^\Delta}.$$

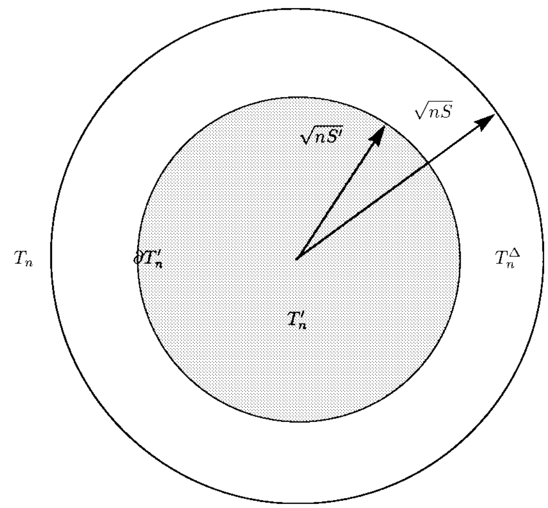

Fig. 1. Balls T_n, T'_n, spherical shell T_n^Δ, and sphere $\partial T'_n$.

Proof: For $x \in \mathcal{C}_n$, let $P^{\mathcal{C}_n}(x)$ be the probability of error of a minimum-distance decoder for $\mathcal{C}_n$ given that x was transmitted. Denote by $\tilde{P}^{\mathcal{C}_n}$ the probability of error of the suboptimum decoder which maps the received point into the nearest point within $\mathcal{C}_n^\Delta$. Then

$$\begin{aligned} P^{\mathcal{C}_n} &\leq \tilde{P}^{\mathcal{C}_n} \\ &= \frac{1}{M_n}\sum_{x\in\mathcal{C}'_n}\tilde{P}^{\mathcal{C}_n}(x) + \frac{1}{M_n}\sum_{x\in\mathcal{C}_n^\Delta}\tilde{P}^{\mathcal{C}_n}(x) \\ &= \frac{M'_n}{M_n} + \frac{1}{M_n}\sum_{x\in\mathcal{C}_n^\Delta}P^{\mathcal{C}_n^\Delta}(x) \\ &\leq \frac{M'_n}{M_n} + \frac{1}{M'_n}\sum_{x\in\mathcal{C}_n^\Delta}P^{\mathcal{C}_n^\Delta}(x) \\ &= \frac{M'_n}{M_n} + P^{\mathcal{C}_n^\Delta}. \end{aligned}$$

□

Let $\pi: \mathbb{R}^n\backslash\{0\} \to \partial T'_n$ be the mapping defined by $\pi(x) = (\sqrt{nS'}/||x||)x$. This mapping radially projects a nonzero point onto the sphere of radius $\sqrt{nS'}$.

Lemma 2 [4, p. 737]:

$$P^{\mathcal{C}_n^\Delta} \leq P^{\pi(\mathcal{C}_n^\Delta)}.$$

Proof: Let $\tilde{P}^{\mathcal{C}}$ be the average probability of error of a code $\mathcal{C}$ subject to a "minimum-angle" decision scheme defined as follows. If y is the received point then the decoder decides for[1]

$$\lambda_{\mathcal{C}}(y) = \underset{x\in\mathcal{C}}{\operatorname{argmin}}\, \angle(x, y)$$

where $\angle(x, y)$ is defined as the unique solution of

$$\cos\angle(x, y) = \frac{x\cdot y}{||x||\,||y||}$$

in the range of $[0, \pi)$. Clearly, this decision scheme is suboptimal unless all codewords have the same modulus in which case this decoder coincides with a minimum-distance decoder. Hence

$$P^{\mathcal{C}_n^\Delta} \leq \tilde{P}^{\mathcal{C}_n^\Delta} \tag{1}$$

and

$$P^{\pi(\mathcal{C}_n^\Delta)} = \tilde{P}^{\pi(\mathcal{C}_n^\Delta)}. \tag{2}$$

For a minimum-angle decoder the decoding region associated to codeword x does not change if x is multiplied by an arbitrary positive constant. Let $x' = \pi(x), x \in \mathcal{C}_n^\Delta$. The minimum-angle

[1] We do not worry about ties since they are elements of a set of zero probability.

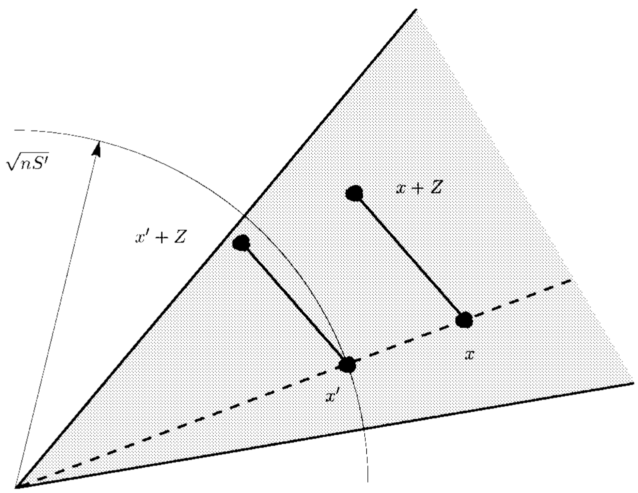

Fig. 2. Minimum-angle decision regions for x and x'.

decision regions for x and x' coincide (see Fig. 2). Assume that $\lambda_{\pi(\mathcal{C}_n^{\Delta})}(x'+Z) = x'$ where Z represents the additive noise with i.i.d. components distributed according to $\mathcal{N}(0, N)$. It is clear from Fig. 2 (and straightforward to prove analytically) that

$$\angle(x'+Z, x') \geq \angle(x+Z, x)$$

and, hence, $\lambda_{\mathcal{C}_n^{\Delta}}(x+Z) = x$. This shows that

$$\tilde{P}^{\mathcal{C}_n^{\Delta}} \leq \tilde{P}^{\pi(\mathcal{C}_n^{\Delta})}$$

which together with (1) and (2) proves the claim. □

Together with Lemma 1 this shows that

$$P^{\mathcal{C}_n} \leq \frac{M_n'}{M_n} + P^{\pi(\mathcal{C}_n^{\Delta})}. \tag{3}$$

Next, we will bound $P^{\pi(\mathcal{C}_n^{\Delta})}$ by defining a suitable suboptimum decoder.

For $y \in \mathbb{R}^n \setminus \{0\}$ and $0 < \theta < \pi/2$, let $B_\theta(y)$ be the n-dimensional closed circular cone with apex at 0, axis passing through y, and half-angle θ and let $B_\theta^c(y)$ denote its complement in $\mathbb{R}^n$. We now define a (suboptimum) decoding function by defining for each $x \in \pi(\mathcal{C}_n^{\Delta})$ the associated decoding region $A_\theta(x)$

$$A_\theta(x) := B_\theta(x) \Big\backslash \bigcup_{x' \in \pi(\mathcal{C}_n^{\Delta}) \setminus \{x\}} B_\theta(x').$$

In words, the decoding region for a codeword consists of those parts of its associated cone which do not overlap with any cone associated to another codeword. To completely specify the decoding function, we declare a decoding error whenever the received word is not an element of $\cup_{x \in \pi(\mathcal{C}_n^{\Delta})} A_\theta(x)$. Note that

$$A_\theta^c(x) = B_\theta^c(x) \bigcup_{x' \in \pi(\mathcal{C}_n^{\Delta}) \setminus \{x\}} B_\theta(x'). \tag{4}$$

Let $P_\theta^{\pi(\mathcal{C}_n^{\Delta})}$ denote the probability of error for the proposed suboptimum decoder and recall that $x_0 = (\sqrt{nS'}, 0, \cdots, 0)$. For any $x \in \mathcal{C}_n^{\Delta}$ we have

$$\begin{aligned}
P_\theta^{\pi(\mathcal{C}_n^{\Delta})}&(\pi(x)) \\
&= \Pr(\pi(x) + Z \in A_\theta^c(\pi(x))) \\
&\overset{a)}{\leq} \Pr(\pi(x) + Z \in B_\theta^c(\pi(x))) \\
&\quad + \sum_{x' \in \mathcal{C}_n^{\Delta} \setminus \{x\}} \Pr(\pi(x) + Z \in B_\theta(\pi(x'))) \\
&= \Pr(x_0 + Z \notin B_\theta(x_0)) \\
&\quad + \sum_{x' \in \mathcal{C}_n^{\Delta} \setminus \{x\}} \Pr(\pi(x) + Z \in B_\theta(\pi(x'))) \\
&\overset{b)}{=} \Pr(x_0 + Z \notin B_\theta(x_0)) \\
&\quad + \sum_{x' \in \mathcal{C}_n^{\Delta} \setminus \{x\}} \Pr(\pi(x') + Z \in B_\theta(\pi(x))) \\
&\overset{c)}{=} \Pr(x_0 + Z \notin B_\theta(x_0)) \\
&\quad + \sum_{x' \in \mathcal{C}_n^{\Delta} \setminus \{x\}} p_\theta(x', x) \\
&= \Pr(x_0 + Z \notin B_\theta(x_0)) \\
&\quad + \sum_{g \in \Lambda_n \setminus \{0\}} p_\theta(g + x, x) \chi_{T_n^{\Delta}}(g + x)
\end{aligned} \tag{5}$$

where a) follows from (4) and an application of the union bound, b) follows from the spherical symmetry of Gaussian noise and the symmetry of the cone, and c) follows from the definition

$$p_\theta(x', x) := \Pr(\pi(x') + Z \in B_\theta(\pi(x)))$$

where

$$\chi_{T_n^{\Delta}}(x) = \begin{cases} 1, & \text{if } x \in T_n^{\Delta} \\ 0, & \text{if } x \notin T_n^{\Delta}. \end{cases}$$

Combining (3) and (5) we get

$$\begin{aligned}
P^{\mathcal{C}_n} &\leq \frac{M_n'}{M_n} + P^{\pi(\mathcal{C}_n^{\Delta})} \\
&= \frac{M_n'}{M_n} + \frac{1}{M_n^{\Delta}} \sum_{x \in \mathcal{C}_n^{\Delta}} P^{\pi(\mathcal{C}_n^{\Delta})}(\pi(x)) \\
&\leq \frac{M_n'}{M_n} + \frac{1}{M_n^{\Delta}} \sum_{x \in \mathcal{C}_n^{\Delta}} P_\theta^{\pi(\mathcal{C}_n^{\Delta})}(\pi(x)) \\
&\leq \frac{M_n'}{M_n} + \Pr(x_0 + Z \notin B_\theta(x_0)) \\
&\quad + \frac{1}{M_n^{\Delta}} \sum_{x \in \mathcal{C}_n^{\Delta}} \sum_{g \in \Lambda_n \setminus \{0\}} p_\theta(g + x, x) \chi_{T_n^{\Delta}}(g + x).
\end{aligned} \tag{6}$$

In Lemma 4 of Appendix III it is shown that for any given $d_n \in \mathbb{R}^+$ there exists a lattice Λ_n^* with determinant[2] $\det(\Lambda_N^*) = d_n$ and a shift s^* such that

$$\begin{aligned}
\frac{1}{M_n^{\Delta}(\Lambda_n^*, s^*)} &\sum_{x \in \mathcal{C}_n^{\Delta}(\Lambda_n^*, s^*)} \sum_{g \in \Lambda_n^* \setminus \{0\}} p_\theta(g + x, x) \chi_{T_n^{\Delta}}(g + x) \\
&\leq \frac{2\sqrt{nS}(n-1)\pi^{n-1/2}(nS')^{n/2}}{d_n \Gamma\left(\frac{n+1}{2}\right)} \int_0^\theta (\sin x)^{n-2}\, dx.
\end{aligned} \tag{7}$$

Moreover, s^* can be chosen in such a way that

$$\frac{M_n'(\Lambda_n^*, s^*)}{M_n(\Lambda_n^*, s^*)} \leq 4 \frac{V_n'}{V_n} \tag{8}$$

and

$$M_n^{\Delta}(\Lambda_n^*, s^*) \geq \frac{V_n^{\Delta}}{4 d_n}. \tag{9}$$

Inserting (8) and (7) into (6) we get

$$\begin{aligned}
P^{\mathcal{C}_n(\Lambda_n^*, s^*)} &\leq 4 \frac{V_n'}{V_n} + \Pr(x_0 + Z \notin B_\theta(x_0)) \\
&\quad + \frac{2\sqrt{nS}(n-1)\pi^{n-1/2}(nS')^{n/2}}{d_n \Gamma\left(\frac{n+1}{2}\right)} \int_0^\theta (\sin x)^{n-2}\, dx.
\end{aligned} \tag{10}$$

[2] The determinant of a lattice is equal to its fundamental volume [10].

This shows that for every $n \in \mathbb{N}, d_n \in \mathbb{R}^+$, and $\theta, 0 < \theta < \pi/2$, there exists a lattice code $\mathcal{C}_n(\Lambda_n^*, s^*)$ with probability of error upper-bounded by (10) and rate $\frac{1}{n}\log M_n(\Lambda_n^*, s^*)$ lower-bounded by

$$\frac{1}{n}\log M_n^{\Delta}(\Lambda_n^*, s^*) \geq (1/n)\log \frac{V_n^{\Delta}}{4d_n}$$

where for the last inequality we used (9).

Choose

$$d_n = 2^{-nR'}V_n^{\Delta} \quad \text{and} \quad \sin\theta = 2^{-R'}. \tag{11}$$

Then

$$\frac{1}{n}\log M_n(\Lambda_n^*, s^*) \geq \frac{1}{n}\log M_n^{\Delta}(\Lambda_n^*, s^*) \geq \frac{1}{n}\log \frac{V_n^{\Delta}}{4d_n} = \frac{1}{n}\log 2^{nR'-2} = R' - \frac{2}{n}$$

which together with $R < R'$ shows that the achieved rate is at least R for sufficiently large n. Further, that the right side of (10) goes to zero as $n \to \infty$ can be seen as follows. Clearly, $(V_n'/V_n^{\Delta}) \overset{n\to\infty}{\longrightarrow} 0$. Note that

$$\sin\angle(x_0 + Z, x_0) = \sqrt{\frac{||Z^{\perp}||^2}{(\sqrt{nS'} + Z^{||})^2 + ||Z^{\perp}||^2}} \overset{n\to\infty}{\longrightarrow} \sqrt{\frac{N}{S'+N}}$$

with probability 1 [11, p. 131]. Recall that by assumption

$$R' < \frac{1}{2}\log_2\left(1 + \frac{S'}{N}\right).$$

Hence

$$\sin\theta = 2^{-R'} > \sqrt{\frac{N}{S'+N}}.$$

It follows that

$$\Pr(x_0 + Z \notin B_{\theta}(x_0)) \overset{n\to\infty}{\longrightarrow} 0.$$

Finally, the last term on the right of (10) can be bounded as follows:

$$\begin{aligned}
&\frac{2\sqrt{nS}(n-1)\pi^{n-1/2}(nS')^{n/2}}{d_n\Gamma\left(\frac{n+1}{2}\right)}\int_0^{\theta}(\sin x)^{n-2}\,dx \\
&\overset{a)}{\leq} \frac{2\sqrt{nS}(n-1)\pi^{n+1/2}(nS')^{n/2}2^{-R'(n-2)}}{2^{-nR'}V_n^{\Delta}\Gamma\left(\frac{n+1}{2}\right)} \\
&\overset{b)}{\leq} \frac{2\sqrt{\pi nS}(n-1)2^{2R'}\Gamma(n/2)}{[(S/S')^{n/2}-1]\frac{2}{n}(\Gamma(n/2)-1)} \overset{n\to\infty}{\longrightarrow} 0
\end{aligned}$$

where a) follows from (11) and the fact that $\theta < \pi/2$, and b) is true since the Gamma function satisfies

$$\begin{aligned}
\Gamma\left(\frac{n}{2}+1\right) &= \frac{n}{2}\Gamma\left(\frac{n}{2}\right) = \frac{n}{2}\int_0^{\infty} x^{(n/2)-1}e^{-x}\,dx \\
&= \frac{n}{2}\left[\int_0^1 x^{(n/2)-1}e^{-x}\,dx + \int_1^{\infty} x^{(n/2)-1}e^{-x}\,dx\right] \\
&< \frac{n}{2}\left[1 + \int_1^{\infty} x^{(n+1/2)-1}e^{-x}\,dx\right] \\
&< \frac{n}{2}\left[1 + \Gamma\left(\frac{n+1}{2}\right)\right].
\end{aligned}$$

This proves Theorem 1. □

Appendix I
Blichfeldt's Principle and Minkowski–Hlawka Theorem

Theorem 2 (Blichfeldt's Principle) [12, p. 45]: Let f be an integrable function with bounded support. If Λ_n is a lattice with fundamental region P_n then

$$\int f(s)\,dV(s) = \int_{P_n}\left(\sum_{h\in\Lambda_n} f(h+s)\right)dV(s).$$

With T a bounded measurable subset of $\mathbb{R}^n$ and $f = \chi_T$ we get

Corollary 1:

$$\int \chi_T\,dV(s) = \int_{P_n} |(\Lambda_n + s)\cap T|\,dV(s)$$

where the left-hand side is the volume of T and the right-hand side is the average cardinality of the points of the lattice $\Lambda_n + s$ within T.

Theorem 3 (Minkowski–Hlawka) [12, p. 65]: Let f be a nonnegative integrable function with bounded support. Then for every $d \in \mathbb{R}^+$ there exists a lattice Λ_n with determinant $\det(\Lambda_n) = d$ such that

$$d\sum_{g\in\Lambda_n\setminus\{0\}} f(g) \leq \int f\,dx.$$

We can think of the left-hand side as a Rieman sum with samples taken at the lattice points (excluding the origin).

Appendix II
A Lemma Concerning P_n^*

Lemma 3 Let Λ_n be a lattice with fundamental region P_n and determinant $\det(\Lambda_n)$ and define

$$P_n^* = \left\{s \in P_n\colon M_n^{\Delta} \geq \frac{V_n^{\Delta}}{4\det(\Lambda_n)}; \frac{M_n'(\Lambda_n, s)}{M_n^{\Delta}(\Lambda_n, s)} \leq 4\frac{V_n'}{V_n^{\Delta}}\right\}.$$

Then

$$V_n^{\Delta} \leq 2\int_{P_n^*} M_n^{\Delta}(\Lambda_n, s)\,dV(s).$$

Proof: Define

$$R_n = \{s \in P_n\colon M_n^{\Delta}(\Lambda_n, s) \geq (V_n^{\Delta}/4\det(\Lambda_n))\}$$

and $R_n^c = P_n\setminus R_n$. Further let

$$O_n = \{s \in P_n\colon (M_n'(\Lambda_n, s)/M_n^{\Delta}(\Lambda_n, s)) \leq 4(V_n'/V_n^{\Delta})\}$$

and $O_n^c = P_n\setminus O_n$. Note that $P_n^* = R_n \cap O_n$. We then have the following sequence of inequalities:

$$\begin{aligned}
V_n^{\Delta} &= \int_{P_n} M_n^{\Delta}(\Lambda_n, s)\,dV(s) \\
&= \int_{P_n^*} M_n^{\Delta}(\Lambda_n, s)\,dV(s) + \int_{R_n\cap O_n^c} M_n^{\Delta}(\Lambda_n, s)\,dV(s) \\
&\quad + \int_{R_n^c} M_n^{\Delta}(\Lambda_n, s)\,dV(s) \\
&\leq \int_{P_n^*} M_n^{\Delta}(\Lambda_n, s)\,dV(s) + \int_{O_n^c} M_n^{\Delta}(\Lambda_n, s)\,dV(s) \\
&\quad + \frac{V_n^{\Delta}}{4\det(\Lambda_n)}\int_{R_n^c} dV(s) \\
&\leq \int_{P_n^*} M_n^{\Delta}(\Lambda_n, s)\,dV(s) + \frac{V_n^{\Delta}}{4V_n'}\int_{O_n^c} M_n'(\Lambda_n, s)\,dV(s) \\
&\quad + \frac{V_n^{\Delta}}{4\det(\Lambda_n)}\int_{P_n} dV(s) \\
&\leq \int_{P_n^*} M_n^{\Delta}(\Lambda_n, s)\,dV(s) + \frac{V_n^{\Delta}}{4V_n'}\int_{P_n} M_n'(\Lambda_n, s)\,dV(s) + \frac{V_n^{\Delta}}{4} \\
&\leq \int_{P_n^*} M_n^{\Delta}(\Lambda_n, s)\,dV(s) + \frac{V_n^{\Delta}}{4} + \frac{V_n^{\Delta}}{4}
\end{aligned}$$

where we used Corollary 1 repeatedly (see Appendix I). The claim follows by subtracting $V_n^{\Delta}/2$ from both sides and multiplying by 2. □

APPENDIX III
THE EXISTENCE OF SUITABLE LATTICES

Lemma 4: Given $d_n \in \mathbb{R}^+$ there exists a lattice Λ_n^* with determinant $\det(\Lambda_n^*) = d_n$ and an $s^* \in P_n^*$ such that

$$(M_n'(\Lambda_n^*, s^*)/M_n(\Lambda_n^*, s^*)) \le 4(V_n'/V_n), M_n^\Delta(\Lambda_n^*, s^*) \ge (V_n^\Delta/4d_n)$$

and

$$\frac{1}{M_n^\Delta(\Lambda_n^*, s^*)} \sum_{x \in \mathcal{C}_n^\Delta(\Lambda_n^*, s^*)} \sum_{g \in \Lambda_n^* \setminus \{0\}} p_\theta(g+x, x)\chi_{T_n^\Delta}(g+x) \le \frac{2\sqrt{nS}(n-1)\pi^{n-1/2}(nS')^{n/2}}{d_n \Gamma\left(\frac{n+1}{2}\right)} \int_0^\theta (\sin x)^{n-2}\, dx.$$

Proof: Let

$$f(z) = \int p_\theta(z+s, s)\chi_{T_n^\Delta}(z+s)\chi_{T_n^\Delta}(s)\, dV(s)$$

and let $d_n \in \mathbb{R}^+$. Since f is nonnegative, integrable, and has support only in $T_n^\Delta + T_n^\Delta$ (Minkowski sum) we can apply the Minkowski–Hlawka Theorem as stated in Theorem 3 to conclude that there exists a lattice Λ_n^* with determinant $\det(\Lambda_n^*) = d_n$ such that

$$\begin{aligned}\sum_{g \in \Lambda_n^* \setminus \{0\}} \int p_\theta(g+s, s)\chi_{T_n^\Delta}(g+s)\chi_{T_n^\Delta}(s)\, dV(s) &= \sum_{g \in \Lambda_n^* \setminus \{0\}} f(g) \\ &\le \frac{1}{d_n}\int f(z)\, dV(z) \\ &= \frac{1}{d_n}\int\left[\int p_\theta(z,s)\chi_{T_n^\Delta}(z)\chi_{T_n^\Delta}(s)\, dV(s)\right] dV(z). \qquad (12)\end{aligned}$$

Let $P_n^*(\Lambda_n^*)$ be as defined in Lemma 3, i.e.,

$$P_n^*(\Lambda_n^*) := \left\{ s \in P_n(\Lambda_n^*)\colon\ M_n^\Delta(\Lambda_n^*, s) \ge \frac{V_n^\Delta}{4d_n};\ \frac{M_n'(\Lambda_n^*, s)}{M_n^\Delta(\Lambda_n^*, s)} \le 4\frac{V_n'}{V_n^\Delta}\right\}. \qquad (13)$$

We have

$$\begin{aligned}
&\int_{P_n^*} M_n^\Delta(\Lambda_n, s)\left[\frac{1}{M_n^\Delta(\Lambda_n, s)} \sum_{x \in \mathcal{C}_n^\Delta(\Lambda_n^*, s)} \cdot \sum_{g \in \Lambda_n^* \setminus \{0\}} p_\theta(g+x, x)\chi_{T_n^\Delta}(g+x)\right] dV(s) \\
&\overset{a)}{\le} \int_{P_n} \sum_{x \in \mathcal{C}_n^\Delta(\Lambda_n^*, s)} \sum_{g \in \Lambda_n^* \setminus \{0\}} p_\theta(g+x, x) \cdot \chi_{T_n^\Delta}(g+x)\, dV(s) \\
&\overset{b)}{=} \int_{P_n} \sum_{h \in \Lambda_n^*} \sum_{g \in \Lambda_n^* \setminus \{0\}} p_\theta(g+h+s, h+s) \cdot \chi_{T_n^\Delta}(g+h+s)\chi_{T_n^\Delta}(h+s)\, dV(s) \\
&\overset{c)}{=} \int \sum_{g \in \Lambda_n^* \setminus \{0\}} p_\theta(g+s, s)\chi_{T_n^\Delta}(g+s)\chi_{T_n^\Delta}(s)\, dV(s) \\
&\overset{d)}{=} \sum_{g \in \Lambda_n^* \setminus \{0\}} \int p_\theta(g+s, s)\chi_{T_n^\Delta}(g+s)\chi_{T_n^\Delta}(s)\, dV(s) \\
&\overset{e)}{\le} \frac{1}{d_n}\int\left[\int p_\theta(z,s)\chi_{T_n^\Delta}(z)\chi_{T_n^\Delta}(s)\, dV(s)\right] dV(z) \\
&\overset{f)}{=} \frac{1}{d_n}\int\left[\int_{z \in T_n^\Delta} p_\theta(z,s)\, dV(z)\right]\chi_{T_n^\Delta}(s)\, dV(s) \\
&\overset{g)}{=} \frac{V_n^\Delta}{d_n}\int_{z \in T_n^\Delta} p_\theta(z, x_0)\, dV(z) \\
&= \frac{V_n^\Delta}{d_n}\int_{\sqrt{nS'}}^{\sqrt{nS}}\left[\int_{z:||z||=r} p_\theta(z, x_0)\, dA(z)\right] dr \\
&\overset{h)}{=} \frac{V_n^\Delta}{d_n}\int_{\sqrt{nS'}}^{\sqrt{nS}}\left[\int_{z:||z||=r} \Pr(\pi(z) + Z \in B_\theta(x_0)) \cdot dA(z)\right] dr \\
&\overset{i)}{=} \frac{V_n^\Delta}{d_n}\int_{\sqrt{nS'}}^{\sqrt{nS}}\left[\frac{(n-1)\pi^{n-1/2}(nS')^{n/2}}{\Gamma\left(\frac{n+1}{2}\right)} \cdot \int_0^\theta (\sin x)^{n-2}\, dx\right] dr \\
&= \frac{V_n^\Delta(\sqrt{nS} - \sqrt{nS'})}{d_n}\left[\frac{(n-1)\pi^{n-1/2}(nS')^{n/2}}{\Gamma\left(\frac{n+1}{2}\right)}\int_0^\theta \cdot (\sin x)^{n-2}\, dx\right] \\
&\le \frac{\sqrt{nS}V_n^\Delta(n-1)\pi^{n-1/2}(nS')^{n/2}}{d_n\Gamma\left(\frac{n+1}{2}\right)}\int_0^\theta (\sin x)^{n-2}\, dx \\
&\overset{j)}{\le} \int_{P_n^*} M_n^\Delta(\Lambda_n^*, s)\left[\frac{2\sqrt{nS}(n-1)\pi^{n-1/2}(nS')^{n/2}}{d_n\Gamma\left(\frac{n+1}{2}\right)} \cdot \int_0^\theta (\sin x)^{n-2}\, dx\right] dV(s)
\end{aligned}$$

where a) follows from the nonnegativity of the integrand, b) follows from the definition of the code $\mathcal{C}_n^\Delta(\Delta_n^*, s)$, c) is a consequence of Blichfeldt's principle as stated in Theorem 2 in Appendix I, d) is valid since only a finite number of elements in the summation are nonzero, e) is (12), f) follows from the nonnegativity of the integrand, g) is valid since the inner integral is independent of the choice of s, h) follows from the definition of $p_\theta(\cdot,\cdot)$, i) is true since the integral is equal to the area which the cone $B_\theta(x_0)$ cuts out from the sphere $\partial T_n'$ and this area is given by

$$((n-1)\pi^{n-1/2}(nS')^{n/2}/\Gamma(n+1/2)) \int_0^\theta (\sin x)^{n-2}\, dx$$

[11, p. 131], and finally, j) is a consequence of Lemma 3 proved in Appendix II which states that $V_n^\Delta \le 2\int_{P_n^*} M_n^\Delta(\Lambda_n^*, s)\, dV(s)$. The claim now follows since P_n^* has nonzero measure (see Lemma 3). □

ACKNOWLEDGMENT

The authors wish to thank V. Tarokh for his helpful and detailed comments on a first draft of this correspondence and A. Wyner for

pointing out the proof methods in Wolfowitz [11] which lead to a simplification of the proof.

References

[1] C. Shannon, "Probability of error for optimal codes in a Gaussian channel," *Bell Syst. Tech. J.*, vol. 38, pp. 611–656, May 1959.
[2] R. De Buda, "The upper error bound of a new near-optimal code," *IEEE Trans. Inform. Theory*, vol. IT-21, pp. 441–445, July 1975.
[3] ——, "Some optimal codes have structure," *IEEE J. Select. Areas Commun.*, vol. 6, pp. 893–899, Aug. 1989.
[4] T. Linder, C. Schlegel, and K. Zeger, "Corrected proof of de Buda's theorem," *IEEE Trans. Inform. Theory*, vol. 39, pp. 1735–1737, Sept. 1993.
[5] H. A. Loeliger, "Averaging bounds for lattices and linear codes," *IEEE Trans. Inform. Theory*, vol. 43, pp. 1767–1773, Nov. 1997.
[6] ——, "On the basic averaging arguments for linear codes," in *Communications and Cryptography: Two Sides of One Tapestry*, R. Blahut, D. Costello, U. Maurer, and T. Mittelholzer, Eds. Norwell, MA: Kluwer, 1994, pp. 251–262.
[7] J. G. D. Forney, "Approaching the capacity of the AWGN channel with coset codes and multilevel coset codes," preprint, 1997.
[8] D. M. Y. Sommerville, *An Introduction to the Geometry of N Dimensions*. New York: Dutton, 1929.
[9] W. H. Beyer, Ed., *CRC, Standard Mathematical Tables and Formulae*. Boca Raton, FL: CRC Press, 1991.
[10] J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Groups*. New York: Springer-Verlag, 1992.
[11] J. Wolfowitz, *Coding Theorems of Information Theory*, 2nd ed. New York: Springer-Verlag, 1964.
[12] E. Hlawka, J. Schoißengeier, and R. Taschner, *Geometric and Analytic Number Theory*. New York: Springer-Verlag, 1991.

Asymptotic Analysis of Multiple Description Quantizers

V. A. Vaishampayan, *Member, IEEE*,
and J.-C. Batllo, *Student Member, IEEE*

***Abstract*— A high-rate analysis of multiple description quantizers is presented for rth-power distortions and general source densities. Both, fixed-length and variable-length encoding of the quantizer indices are considered. Optimal companding functions are shown to be the same as for single-channel quantizers. As compared to the bound of Ozarow, a gap of 8.69 dB and 3.07 dB exists between the entropy-constrained and level-constrained cases, respectively, for a memoryless Gaussian source and $r = 2$.**

***Index Terms*— Data compression, diversity systems, multiple descriptions, quantization, source coding.**

I. Introduction

We present an asymptotic analysis of multiple description scalar quantizers (MDSQ's), introduced in [1] for rth-power difference distortion measures and general source densities. A multiple description quantizer is designed for a communication system that connects the source to the destination via two (or more) channels. It is assumed that either of the two channels may be broken and that this information is known to the decoder but not the encoder. The objective is to design the encoder and decoder so as to minimize the average distortion when both channels work (referred to as the average central distortion) subject to constraints on the average distortion when only one of the channels works (referred to as the average side distortion).

An achievable rate region for the multiple description source coding problem was given by El Gamal and Cover [3],[1] for a memoryless source and a single-letter fidelity criterion. Ozarow [2] constructed the rate-distortion region for the special case of a memoryless Gaussian source and the squared-error distortion criterion. The binary-symmetric memoryless source with an error frequency distortion criterion has been studied by Berger and Zhang [4], [5], Ahlswede [6], Witsenhausen and Wyner [7], Wolf, Wyner, and Ziv [8]. It was conjectured that the achievable rate region given in [3] coincided with the rate-distortion region in cases other than the Gaussian memoryless source and the squared-error distortion criterion. However, this conjecture was disproved in [5]. There have been no results to date exactly characterizing the entire rate-distortion region for non-Gaussian sources and for sources with memory. An important special case of the multiple description problem is the problem of successive refinement of information [9]. In [9], a necessary and sufficient condition for a rate distortion problem to be successively refinable is derived.

The design of MDSQ's was addressed in [1] and [10], respectively. Vector quantizer design for the multiple description problem was addressed in [11]. Applications of multiple description source codes that have been explored so far are to speech transmission over packet-switched networks and to communication over correlated Rayleigh fading channels. Speech transmission over packet-switched networks was considered in [12]–[14]. In [15], it was shown that for transmitting information from a memoryless Gaussian source over a Rayleigh fading channel, the multiple description approach results in good performance at low interleaving delays as compared to standard channel coding approaches. This conclusion was extended to sources with memory in [16] where, on an equal interleaving delay basis, significant performance improvements are obtained over channel codes for speech transmission over Rayleigh fading channels. Applications of multiple description codes to image and video transmission over lossy packet networks are considered in [17].

The asymptotic analysis of single-channel quantizers is very relevant to the work presented here. Bennett [18] derived an asymptotic formula for the mean-squared error using a companding approach. This was extended to rth-power difference distortions by Algazi [19]. Panter and Dite [20] derived the optimum MSE performance of a level-constrained quantizer and Gish and Pierce [21] derived the performance of optimum entropy-constrained quantizers and showed that uniform threshold quantizers were optimum for this purpose in the limit of high level density. Bucklew and Wise [22] presented a sufficient condition under which Bennett's approximation is accurate.

The contribution of this correspondence is a derivation of the asymptotic performance of multiple description quantizers. Specifically, we derive expressions for the average side and central distortions and for the entropy when the number of quantization levels is large. We show that the optimum companding functions for the

Manuscript received October 11, 1996; revised May 15, 1997. This work was supported by the NSF under Grants NCR-9104566 and NSF NCR-9314221. The material in this correspondence was presented in part at the IEEE International Symposium on Information Theory, Trondheim, Norway, June 27–July 1, 1994. This work was performed while V. A. Vaishampayan was at Texas A&M University.

V. A. Vaishampayan is with AT&T Labs–Research, Florham Park, NJ 07932 USA.

J.-C. Batllo is with the Department of Electrical Engineering, Texas A&M University, College Station, TX 77843 USA.

Publisher Item Identifier S 0018-9448(98)00019-4.

[1] Formulation of the multiple description problem is attributed in [3] to Gersho, Witsenhausen, Wolf, Wyner, Ziv, and Ozarow.