Universal Bound on the Performance of Lattice Codes

Vahid Tarokh, *Member, IEEE*, Alexander Vardy, *Fellow, IEEE*, and Kenneth Zeger, *Senior Member, IEEE*

***Abstract*— We present a lower bound on the probability of symbol error for maximum-likelihood decoding of lattices and lattice codes on a Gaussian channel. The bound is tight for error probabilities and signal-to-noise ratios of practical interest, as opposed to most existing bounds that become tight asymptotically for high signal-to-noise ratios. The bound is also universal; it provides a limit on the highest possible coding gain that may be achieved, at specific symbol error probabilities, using *any* lattice or lattice code in n dimensions. In particular, it is shown that the effective coding gains of the densest known lattices are much lower than their nominal coding gains. The asymptotic (as $n \to \infty$) behavior of the new bound is shown to coincide with the Shannon limit for Gaussian channels.**

***Index Terms*— Coding gain, Gaussian channels, lattice codes, lattices, Shannon limit.**

I. Introduction

DETERMINING the maximum possible coding gain of an n-dimensional lattice code is a fundamental problem in communications. This problem has been extensively studied, for instance in [6], [9], [10], [11], [16], [19], [20], [23], and references therein.

It is well known [9], [11], [16] that, assuming high rates and high signal-to-noise ratio (SNR), the gain of a lattice code over uncoded QAM transmission can be separated into a shaping gain due to the shape of a support region $\mathbb{D}$ and a coding gain due to the structure of the underlying lattice Λ. Asymptotically, as SNR $\to \infty$, the latter approaches the nominal coding gain of Λ which, in turn, depends only on the density of Λ. Thus for very high SNR's, determining the maximum possible coding gain of an n-dimensional lattice code is equivalent to finding the densest possible lattice packing in n dimensions.

Nevertheless, there is usually a sharp discrepancy between the nominal coding gain and the effective coding gain observed at signal-to-noise ratios of practical interest. Hence a more careful analysis of the effective coding gain of lattices and lattice codes is necessary.

Manuscript received June 1, 1996; revised April 1, 1998. This work was supported in part by the National Science Foundation.

V. Tarokh is with AT&T Research Labs, Florham Park, NJ 07932 USA (e-mail: tarokh@research.att.com).

A. Vardy is with Laboratoire Informatique, Signaux, et Systémes a Sophia-Antipolis, Centre Nationale de la Recherche Scientifique, F-06410 Biot, France, on leave from the University of California at San Diego, La Jolla, CA 92093 USA (e-mail: vardy@golay.csl. uiuc.edu).

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

Communicated by R. Laroia, Associate Editor for Source Coding.

Publisher Item Identifier S 0018-9448(99)01404-2.

The union bound [6, p. 70] is a well-known upper bound on the probability P_e of symbol error for maximum-likelihood decoding of lattices and lattice codes on a Gaussian channel. This bound is tight for SNR $\to \infty$. However, as noted in [15], [19], and [20], the union bound is often inadequate. More elaborate *upper* bounds on the probability of error for block-coded and lattice-coded modulation schemes were developed by Berlekamp [3] and, more recently, by De Oliveira–Battail [23] and by Herzberg–Poltyrev [19], [20]. Our objective in this paper is to provide a *lower* bound on the probability of symbol error P_e, which would be reasonably accurate at signal-to-noise ratios of practical interest.

We note that the union bound as well as the bounds of [3], [19], [20], and [23] all depend on the distance spectrum (or theta series [6, p. 45]) of the lattice code at hand. In contrast, we develop a *universal* bound, namely, a bound that provides a limit on the highest possible coding gain that may be achieved, using *any* lattice or lattice code in n dimensions.

Such a bound is presented in the next two sections. Specifically, Section II is concerned with maximum-likelihood decoding of lattices, while Section III deals with lattice codes.

The bound developed in the next section is based on the well-known geometric notion of an equivalent sphere [19], [25], [28] and on the fact that no decoding region of a given volume can be better than a spherical decoding region of the same volume (cf. [28]). Furthermore, although it is difficult, and often impossible, to compute the integral of a Gaussian distribution over the Voronoi region of a nontrivial lattice, the integral over a sphere may be computed in closed form. This computation, detailed in Section II, leads to a lower bound on P_e in terms of the fundamental volume of the lattice at hand. The resulting bound is further converted into a powerful upper bound on the highest possible *coding gain* that may be achieved, for specific symbol error probabilities, using any n-dimensional lattice. Invoking the latter form of our bound, we show that the effective coding gains (at symbol error rates of 10^{-5} to 10^{-7}) of the densest known lattices are much lower than their nominal coding gains. Finally, we investigate the asymptotic (as $n \to \infty$) behavior of the new bounds, and show that it coincides with the Shannon limit for Gaussian channels. This is consistent with the converse to the Shannon theorem, although our proof of this result relies solely on geometric notions.

In practice, only a *finite* set of points of a lattice Λ can be used as a signal constellation in a communication system. This set consists of those points of Λ that are contained in a bounded support region $\mathbb{D}$, and is known as the lattice code $\mathbb{C}(\Lambda, \mathbb{D})$

based on $\mathbb{D}$ and Λ. The performance of a lattice code $\mathbb{C}(\Lambda, \mathbb{D})$ on a Gaussian channel depends not only on the underlying lattice Λ but also on the shape of the support region $\mathbb{D}$.

Lattice codes are considered in Section III of this paper. To extend the bounds of Section II to lattice codes, we rely on the continuous approximation [4], [9], [16], and express our results in terms of the normalized signal-to-noise ratio [9], [16]. For completeness, an overview of some basic facts about lattice codes and a brief primer on the continuous approximation technique are also included in Section III. Herein, we observe that the continuous approximation becomes exact for high-rate codes [1], [18].

Another issue with lattice codes is that the Voronoi region of a point $\boldsymbol{x}$ in the lattice code $\mathbb{C}(\Lambda, \mathbb{D})$ may be larger than the Voronoi region of $\boldsymbol{x}$ in the lattice Λ, if this point lies sufficiently close to the boundary of the support region $\mathbb{D}$. Thus, lower bounds on P_e based on the examination of the Voronoi regions of Λ are not necessarily valid for the lattice code $\mathbb{C}(\Lambda, \mathbb{D})$. This issue is considered in detail in the Appendix, where it is shown that this boundary effect becomes negligible for high-rate codes.

II. Bounds for Lattices

We first establish the relevant notation. An infinite set $\Lambda \subset \mathbb{R}^n$ is called a *sphere packing* if the minimum distance between distinct points of Λ is $d(\Lambda) > 0$. Such a set is called a lattice packing, or simply a *lattice*, if Λ is a group under addition in $\mathbb{R}^n$. The density $\Delta(\Lambda)$ of Λ is the fraction of the space covered by spheres of radius $\frac{1}{2}d(\Lambda)$ about the points of Λ. The center density $\delta(\Lambda)$ of Λ is the density divided by the volume V_n of a unit sphere in $\mathbb{R}^n$. It is known [6, p. 9] that

$$V_n = \frac{\pi^{n/2}}{(n/2)!} = \begin{cases} \frac{\pi^k}{k!}, & n = 2k \\ \frac{2^n \pi^k k!}{n!}, & n = 2k+1 \end{cases} \tag{1}$$

where $(n/2)! \stackrel{\text{def}}{=} \Gamma(\frac{n}{2}+1)$ for both odd and even n, and

$$\Gamma(t) = \int_0^\infty u^{t-1} e^{-u}\, du$$

is Euler's gamma function. For even $n = 2k$, we have $(n/2)! \approx (k/e)^k$ by the Stirling approximation, where e is the natural base of logarithms.

The *Voronoi region* of a point $\boldsymbol{x} \in \Lambda$ is a convex polytope, which consists of all the points in $\mathbb{R}^n$ that are at least as close to $\boldsymbol{x}$ as to any other point in Λ. We let Π denote the Voronoi region of the origin of $\mathbb{R}^n$. (It is easy to see that for lattice packings, Voronoi regions of all the points are congruent to each other.) The *volume* of a lattice Λ is defined as the volume of Π, that is $V(\Lambda) = V(\Pi)$. The asymptotic, or the *nominal*, coding gain (cf. [10]) of Λ can then be expressed as

$$\gamma(\Lambda) \stackrel{\text{def}}{=} 4\delta(\Lambda)^{2/n} = \frac{d(\Lambda)^2}{V(\Lambda)^{2/n}}. \tag{2}$$

The highest possible value γ_n of the coding gain $\gamma(\Lambda)$ of an n-dimensional lattice is called the Hermite parameter. The value of γ_n is presently known [6] only for $n = 1, 2, \cdots, 8$.

A. Lower Bound on the Probability of Error

In this section, we derive a universal lower bound on the probability of symbol error for maximum-likelihood decoding of n-dimensional lattices on an additive white Gaussian noise (AWGN) channel. In contrast to the nominal coding gain, our bound is not asymptotic in SNR; it is reasonably tight at signal-to-noise ratios of practical interest. Moreover, since the bound applies to any lattice of a given volume, we will effectively bound the performance of the *densest* lattices.

If a point $\boldsymbol{y} \in \Lambda$ is transmitted through an AWGN channel, the received point is given by $\boldsymbol{y} + \boldsymbol{X}$, where $\boldsymbol{X}$ is a vector of independent and identically distributed (i.i.d.) Gaussian random variables with zero mean and variance σ^2 per dimension. The channel output $\boldsymbol{y} + \boldsymbol{X}$ is decoded to $\boldsymbol{y}$ under maximum-likelihood decoding if and only if $\boldsymbol{y}+\boldsymbol{X}$ belongs to the Voronoi region of $\boldsymbol{y}$ in the lattice Λ. Thus the probability of correct decoding is given by

$$P_c = \int_\Pi f(\boldsymbol{x})\, d\boldsymbol{x} \tag{3}$$

where

$$f(\boldsymbol{x}) = \frac{1}{(\sqrt{2\pi}\sigma)^n} \exp\left(\frac{-\|\boldsymbol{x}\|^2}{2\sigma^2}\right)$$

is the probability density function of $\boldsymbol{X}$. Now let $\mathcal{S}_\Pi$ denote the n-dimensional sphere of radius r about the origin, having the same volume as Π. This sphere is sometimes called [17], [28] the *equivalent sphere* of Π. Forney [13] defines the *normalized radius* ρ of $\mathcal{S}_\Pi$ by the relation $r^2 = n\rho^2$. The volume of $\mathcal{S}_\Pi$ is $V_n r^n = V(\Pi)$, and, therefore,

$$r = \frac{V(\Pi)^{1/n}}{V_n^{1/n}} = \frac{V(\Lambda)^{1/n} \Gamma\left(\frac{n}{2}+1\right)^{1/n}}{\sqrt{\pi}} \tag{4}$$

in view of (1). The corresponding expression for the normalized radius of $\mathcal{S}_\Pi$ is considerably simpler. For even $n = 2k$, the normalized radius of $\mathcal{S}_\Pi$ is lower-bounded by

$$\rho^2 \geq \frac{V(\Lambda)^{1/k}}{2\pi e}. \tag{5}$$

This follows from (4) and the fact that $k^k/e^k \leq k!$ for all $k \geq 1$. Later in this section, we will use this lower bound as a crude approximation of ρ^2 for $k \to \infty$.

The following simple, but key, observation dates back to the work of Shannon [25], and was mentioned in several recent papers [19], [28].

Lemma 2.1:

$$\int_\Pi f(\boldsymbol{x})\, d\boldsymbol{x} \leq \int_{\mathcal{S}_\Pi} f(\boldsymbol{x})\, d\boldsymbol{x}. \tag{6}$$

Proof: Let $\Phi = \Pi \backslash \mathcal{S}_\Pi$ and $\Psi = \mathcal{S}_\Pi \backslash \Pi$. Notice that $V(\Phi) = V(\Psi)$, by the definition of the equivalent sphere $\mathcal{S}_\Pi$. It is obvious that (6) is equivalent to

$$\int_\Phi f(\boldsymbol{x})\, d\boldsymbol{x} \leq \int_\Psi f(\boldsymbol{x})\, d\boldsymbol{x}.$$

Furthermore,

$$f(\boldsymbol{x}) \le \frac{1}{(\sqrt{2\pi}\sigma)^n}\exp\left(\frac{-r^2}{2\sigma^2}\right), \qquad \text{for all } \boldsymbol{x} \in \Phi$$
$$f(\boldsymbol{x}) \ge \frac{1}{(\sqrt{2\pi}\sigma)^n}\exp\left(\frac{-r^2}{2\sigma^2}\right), \qquad \text{for all } \boldsymbol{x} \in \Psi$$

as $f(\cdot)$ is a decreasing function of the distance from the origin. Since Φ and Ψ have the same volume, this completes the proof of the lemma. □

The probability of error $P_e = 1 - P_c$ under maximum-likelihood decoding of Λ is the probability that a white Gaussian n-tuple $\boldsymbol{X}$ with variance σ^2 per dimension falls outside the Voronoi region Π of Λ. Similarly, we define

$$P_{e,\mathcal{S}} \stackrel{\text{def}}{=} 1 - \int_{\mathcal{S}_\Pi} f(\boldsymbol{x})\,d\boldsymbol{x}.$$

Thus $P_{e,\mathcal{S}}$ is the probability that the n-tuple $\boldsymbol{X}$ falls outside the equivalent sphere $\mathcal{S}_\Pi$. With this notation, Lemma 2.1 reduces to the inequality $P_e \ge P_{e,\mathcal{S}}$. This means that no decoding region can be better that a spherical decoding region of the same volume.

The usefulness of Lemma 2.1 lies in the fact that the integral on the left-hand side of (6) is difficult, often impossible, to compute, whereas the integral on the right-hand side of (6) can be computed in closed form. Indeed, consider a change of variables to the following spherical coordinates:

$$\begin{aligned}
x_1 &= \sigma u \cos\theta_1\\
x_2 &= \sigma u \sin\theta_1 \cos\theta_2\\
&\vdots\\
x_{n-1} &= \sigma u \sin\theta_1 \cdots \sin\theta_{n-2}\cos\theta_{n-1}\\
x_n &= \sigma u \sin\theta_1 \cdots \sin\theta_{n-2}\sin\theta_{n-1}
\end{aligned}$$

where

$$\theta_1, \theta_2, \cdots, \theta_{n-2} \in [0, \pi]$$

while

$$\theta_{n-1} \in [0, 2\pi].$$

The Jacobian of this transformation is given by

$$\sigma^{n-1}u^{n-1}h(\theta_1, \cdots, \theta_{n-1})$$

where the function $h(\cdot)$ does not depend on u. Thus

$$\int_{\mathcal{S}_\Pi} f(\boldsymbol{x})\,d\boldsymbol{x} = \frac{\sigma^{n-1}}{(\sqrt{2\pi}\sigma)^n}\int_0^\pi \cdots \int_0^{2\pi}\int_0^r u^{n-1}e^{-u^2/2} \times h(\theta_1, \cdots, \theta_{n-1})\,d(\sigma u)\,d\theta_{n-1}\cdots d\theta_1.$$

The integral on the right-hand side of this expression separates into a product of two independent integrals, one of which is given by

$$\mathcal{J}(n) \stackrel{\text{def}}{=} \int_0^\pi \cdots \int_0^{2\pi} h(\theta_1, \cdots, \theta_{n-1})\,d\theta_{n-1}\cdots d\theta_1.$$

But

$$\begin{aligned}
\frac{\mathcal{J}(n)}{n} &= \mathcal{J}(n)\int_0^1 u^{n-1}\,du\\
&= \int_0^\pi \cdots \int_0^{2\pi}\int_0^1 u^{n-1}h(\theta_1, \cdots, \theta_{n-1})\,du\,d\theta_{n-1}\cdots d\theta_1
\end{aligned}$$

is precisely the volume of a unit sphere in $\mathbb{R}^n$. In view of (1), this yields

$$\mathcal{J}(n) = nV_n = \frac{n\pi^{n/2}}{(n/2)!}.$$

Thus we have

$$\mathcal{I}(n) \stackrel{\text{def}}{=} \int_{\mathcal{S}_\Pi} f(\boldsymbol{x})\,d\boldsymbol{x} = \frac{n}{2^{n/2}(n/2)!}\int_0^{r/\sigma} u^{n-1}e^{-u^2/2}\,du \tag{7}$$

and integrating by parts gives the following recurrence relation:

$$\mathcal{I}(n) = \mathcal{I}(n-2) - e^{-z}\frac{z^{\frac{1}{2}n-1}}{\left(\frac{n}{2}-1\right)!} \tag{8}$$

where $z = r^2/2\sigma^2$. Further, it can be easily verified that

$$\mathcal{I}(1) = \sqrt{2/\pi}\int_0^{r/\sigma} e^{-u^2/2}\,du = 1 - \mathsf{erfc}\left(z^{\frac{1}{2}}\right)$$
$$\mathcal{I}(2) = \int_0^{r/\sigma} ue^{-u^2/2}\,du = 1 - e^{-z}$$

where $\mathsf{erfc}(\cdot)$ is the complementary error-function given by $\mathsf{erfc}(x) = (2/\sqrt{\pi})\int_x^\infty e^{-t^2}\,dt$. We are now ready to prove one of our main results in this section.

Theorem 2.2: If points of an n-dimensional lattice Λ are transmitted over an AWGN channel with noise variance σ^2 per dimension, the probability of symbol error under maximum-likelihood decoding is lower-bounded as follows:

$$P_e \ge e^{-z}\left(1 + \frac{z}{1!} + \frac{z^2}{2!} + \cdots + \frac{z^{\frac{n}{2}-1}}{\left(\frac{n}{2}-1\right)!}\right) \tag{9}$$

for n even, while for odd n, we have

$$P_e \ge \mathsf{erfc}\left(z^{\frac{1}{2}}\right) + e^{-z}\left(\frac{z^{1/2}}{(1/2)!} + \frac{z^{3/2}}{(3/2)!} + \cdots + \frac{z^{\frac{n}{2}-1}}{\left(\frac{n}{2}-1\right)!}\right) \tag{10}$$

where

$$z = \frac{V(\Lambda)^{2/n}\Gamma\left(\frac{n}{2}+1\right)^{2/n}}{2\pi\sigma^2}. \tag{11}$$

Proof: By Lemma 2.1 and (3), we have

$$P_e \ge P_{e,\mathcal{S}} = 1 - \int_{\mathcal{S}_\Pi} f(\boldsymbol{x})\,d\boldsymbol{x}.$$

The expressions (9) and (10) follow immediately by induction on (8). The expression for $z = r^2/2\sigma^2$ in (11) follows from (1) and (4). □

Remark: For even $n = 2k$, the value of $P_{e,\mathcal{S}}$ in (9) can be also computed as follows. Consider a sequence of k pairs (X_{2i-1}, X_{2i}) of i.i.d. Gaussian random variables with zero mean and variance σ^2, where $i = 1, 2, \cdots, k$. The energy $Y_i = X_{2i-1}^2 + X_{2i}^2$ of each pair is then an exponential random variable with mean $2\sigma^2$. Furthermore, Y_i and Y_j are independent for $i \neq j$. With this notation, we have

$$\begin{aligned} P_{e,\mathcal{S}} &= \Pr\{X_1^2 + X_2^2 + \cdots + X_n^2 \geq r^2\} \\ &= \Pr\{Y_1 + Y_2 + \cdots + Y_k \geq r^2\}. \end{aligned} \tag{12}$$

If the random variables $Y_1, Y_2, \cdots, Y_k$ are thought of as the interarrival times of a Poisson random process $\mathcal{Y}(t)$, then $P_{e,\mathcal{S}}$ is precisely the probability that there are fewer than k arrivals in the interval $[0, r^2)$. The probability of m arrivals during the interval $[0, T)$ is given by the Poisson distribution

$$P_{\mathcal{Y}(T)}(m) = e^{-\lambda}\frac{\lambda^m}{m!}, \qquad \text{for } m = 0, 1, \cdots$$

where $\lambda = T/E[Y_i]$. Since $E[Y_i] = 2\sigma^2$ for all i, we have $\lambda = T/2\sigma^2$. It follows that the probability $P_{e,\mathcal{S}}$ of fewer than k arrivals in the interval $[0, r^2)$ is given by

$$\begin{aligned} P_{e,s} &= P_{\mathcal{Y}(r^2)}(0) + \cdots + P_{\mathcal{Y}(r^2)}(k-1) \\ &= e^{-z}\left(1 + \frac{z}{1!} + \frac{z^2}{2!} + \cdots + \frac{z^{k-1}}{(k-1)!}\right) \end{aligned}$$

where $z = r^2/2\sigma^2$. This argument, pointed out by Forney [13], avoids the explicit integration in (7) and (8), but unfortunately does not extend to odd values of n.

B. Upper Bound on Coding Gain

When designing a communication system for a band-limited Gaussian channel, it is more conventional to consider effective coding gains rather than probabilities of symbol error. In this subsection, we will show how the lower bound on P_e obtained in Theorem 2.2 can be converted into an upper bound on the highest possible coding gain that may be achieved, at specific symbol error probabilities, using any n-dimensional lattice. The resulting bounds for $n = 1, 2, \cdots, 32$ are summarized in Table I, and compared with the nominal coding gains of the best known lattices in the corresponding dimensions.

Coding gain is usually defined in terms of the signal-to-noise ratios required by the coded and uncoded systems to achieve a given probability of error. Thus to discuss coding gains, we first need to discuss signal-to-noise ratios. It is not immediately clear what the signal-to-noise ratio *is* in the context of maximum-likelihood decoding of lattices on a Gaussian channel (as opposed to lattice codes, the signal power is, in principle, unlimited in the case of lattices). In this regard, we will follow the suggestion of Forney [13], [14] and use

$$\alpha^2 \stackrel{\text{def}}{=} \frac{\rho^2}{\sigma^2} \tag{13}$$

as a measure of signal-to-noise ratio. Here ρ is the normalized radius of the equivalent sphere $\mathcal{S}_{\Pi}$ and σ^2 is the noise variance per dimension. We will refer to α^2 as the *lattice signal-to-noise ratio*. In the next section, we will show that the lattice SNR for a lattice Λ is closely related to the normalized SNR for a lattice code $\mathbb{C}(\Lambda, \mathbb{D})$ based on Λ, provided the number of points in $\mathbb{C}$ is sufficiently large.

For the sake of brevity, we only consider the case where n is even. The development for n odd is similar [27], and the results are summarized in Table I for all odd $n \leq 31$. For even n, let $k = n/2$ and define the function

$$g_k(x) \stackrel{\text{def}}{=} e^{-x}\left(1 + \frac{x}{1!} + \frac{x^2}{2!} + \cdots + \frac{x^{k-1}}{(k-1)!}\right). \tag{14}$$

Thus the lower bound (9) of Theorem 2.2 can be written as $P_e \geq g_k(z)$. Furthermore, it follows from the definition of α^2 in (13) that $z = r^2/2\sigma^2 = n\rho^2/2\sigma^2 = k\alpha^2$. Thus $P_e \geq g_k(k\alpha^2)$ is a nonasymptotic lower bound on the probability of symbol error in terms of the lattice SNR defined above. Forney [13] suggests that P_e should be normalized per two dimensions, and defines $P_e^* = (2/n)P_e$ as the normalized probability of symbol error. With this notation, the lower bound of Theorem 2.2 can be rewritten as

$$P_e^* \geq \frac{g_k(k\alpha^2)}{k} \tag{15}$$

for all even $n = 2k$. This bound is plotted in Fig. 1 for $n = 4, 16, 64, 256$. We have also included in Fig. 1 simulation results for the 16-dimensional Barnes–Wall lattice BW_{16}, which suggest that the bound of (15) is quite tight at all signal-to-noise ratios. The curve for $n = 1$ plotted in Fig. 1 results by manipulating the lower bound in (10) in a manner analogous to (15). This curve is achieved by the integer lattice $\mathbb{Z}$, since the Voronoi region of $\mathbb{Z}$ is a one-dimensional sphere. It presents an approximate baseline for the measurement of effective coding gains. We will provide a more precise analysis below.

Let P_e denote a *fixed* desired probability of symbol error. We first ask the following question. What is the minimum lattice SNR that is required to achieve a probability of symbol error P_e using an n-dimensional lattice? The answer to this question follows by examining (9) and (14). It is easy to see that the function $g_k(x)$ in (14) is continuous, and is strictly decreasing in the interval $(0, \infty)$. Furthermore, $g_k(0) = 1$ and $\lim_{x\to\infty} g_k(x) = 0$, for all $k \geq 1$. From these properties, it follows that the equation $g_k(x) = P_e$ has a unique solution. We denote this solution by $z(k; P_e)$.

Theorem 2.3: Achieving a probability of symbol error P_e using an n-dimensional lattice Λ requires a lattice signal-to-noise ratio of at least

$$\alpha^2 \geq \frac{z(k; P_e)}{k}. \tag{16}$$

Proof: The probability of symbol error is lower-bounded by $g_k(z)$ for $z = k\alpha^2$. If the lattice signal-to-noise ratio α^2 does not satisfy (16), then $z < z(k; P_e)$. Since $g_k(x)$ is a strictly decreasing function, we then have $g_k(z) > P_e$ and the theorem follows. □

We now consider the uncoded case, namely, the case where a scaled version $c\mathbb{Z}^n$ of the integer lattice $\mathbb{Z}^n$ is used to

TABLE I
UPPER BOUNDS ON THE CODING GAIN OF LOW-DIMENSIONAL LATTICES

n	New upper bound on coding gain			Nominal coding gain ($P_e \to 0$)			
	$P_e^* = 10^{-5}$	$P_e^* = 10^{-6}$	$P_e^* = 10^{-7}$	Best known lattice: Name	Best known lattice: Reference	$\gamma(\Lambda)$	Upper bound on coding gain
1	0	0	0	$\mathbb{Z}$	[6, p.106]	0	0
2	0.62	0.66	0.70	A_2	[6, p.110]	0.62	0.62
3	1.07	1.15	1.22	A_3	[6, p.112]	1.00	1.00
4	1.43	1.54	1.63	D_4	[6, p.118]	1.51	1.51
5	1.72	1.86	1.97	D_5	[6, p.117]	1.81	1.81
6	1.97	2.13	2.25	E_6	[6, p.125]	2.22	2.22
7	2.18	2.36	2.50	E_7	[6, p.124]	2.58	2.58
8	2.37	2.56	2.72	E_8	[6, p.120]	3.01	3.01
9	2.53	2.74	2.91	Λ_9	[6, p.170]	3.01	3.31
10	2.68	2.90	3.08	Λ_{10}	[6, p.170]	3.14	3.57
11	2.80	3.04	3.23	K_{11}	[6, p.165]	3.30	3.82
12	2.92	3.17	3.37	K_{12}	[6, p.127]	3.64	4.05
13	3.04	3.29	3.50	K_{13}	[6, p.165]	3.72	4.27
14	3.14	3.40	3.62	Λ_{14}	[6, p.170]	3.96	4.47
15	3.23	3.51	3.73	Λ_{15}	[6, p.170]	4.21	4.67
16	3.32	3.60	3.83	Λ_{16}	[6, p.129]	4.52	4.86
17	3.40	3.69	3.93	Λ_{17}	[6, p.176]	4.60	5.04
18	3.47	3.77	4.02	Λ_{18}	[6, p.176]	4.75	5.21
19	3.54	3.85	4.10	Λ_{19}	[6, p.176]	4.91	5.37
20	3.61	3.92	4.18	Λ_{20}	[6, p.176]	5.12	5.53
21	3.67	3.99	4.26	Λ_{21}	[6, p.176]	5.30	5.68
22	3.73	4.06	4.33	Λ_{22}	[6, p.176]	5.53	5.83
23	3.78	4.12	4.39	Λ_{23}	[6, p.176]	5.76	5.97
24	3.83	4.18	4.45	Λ_{24}	[6, p.131]	6.02	6.10
25	3.88	4.23	4.51	Λ_{25}	[6, p.177]	5.90	6.24
26	3.93	4.28	4.57	Λ_{26}	[6, p.177]	5.84	6.36
27	3.98	4.34	4.63	Λ_{27}	[6, p.177]	5.80	6.49
28	4.02	4.38	4.68	B_{28}	[2]	5.85	6.61
29	4.06	4.43	4.73	B_{29}	[2]	5.86	6.73
30	4.10	4.47	4.78	Q_{30}	[6, p.220]	5.90	6.84
31	4.14	4.52	4.83	Q_{31}	[6, p.220]	6.07	6.95
32	4.18	4.56	4.87	Q_{32}	[6, p.220]	6.28	7.06

transmit information over a Gaussian channel. The probability of symbol error for the uncoded case can be computed exactly. Indeed, the Voronoi region Π for the lattice $c\mathbb{Z}^n$ is a hypercube of side c, and, therefore

$$1 - P_e = \int_\Pi f(\boldsymbol{x})\, d\boldsymbol{x} = \left(\frac{1}{\sqrt{2\pi}\sigma} \int_{-c/2}^{c/2} e^{-u^2/2\sigma^2}\, du \right)^n$$
$$= \left(1 - \mathrm{erfc}\left(\frac{c}{\sqrt{8}\sigma} \right) \right)^n. \tag{17}$$

Let $\zeta(k;\, P_e)$ denote the unique solution of the equation

$$(1 - \mathrm{erfc}\,(x))^{2k} = 1 - P_e.$$

Then to achieve a symbol error probability of P_e in the uncoded case, we need $c = \sqrt{8}\sigma\zeta(k; P_e)$. The volume of $c\mathbb{Z}^n$ is $c^n V(\mathbb{Z}^n) = c^n$, and hence the corresponding lattice SNR is given by

$$\alpha^2 = \frac{c^2}{\sigma^2} \cdot \frac{(k!)^{1/k}}{2\pi k}$$

assuming that $n = 2k$ is even. It follows that in the uncoded case the desired probability P_e of symbol error is achieved precisely at a lattice SNR of

$$\alpha^2 = \zeta(k; P_e)^2 \cdot \frac{4(k!)^{1/k}}{\pi k}. \tag{18}$$

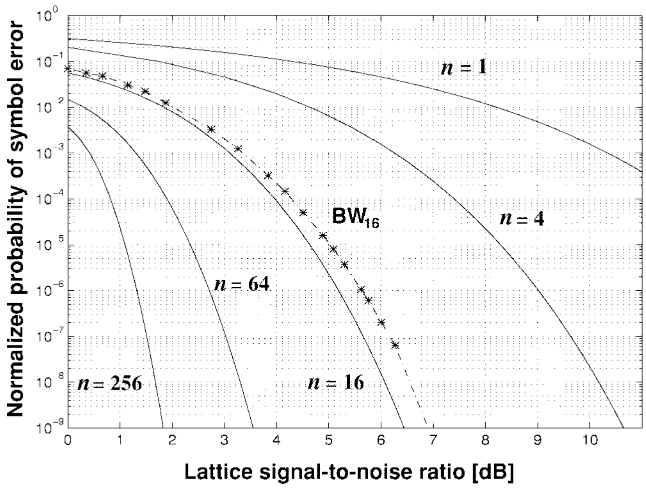

Fig. 1. Lower bounds on the performance of lattices.

The ratio of the expressions on the right-hand side of (16) and (18) is an upper bound on the effective coding gain that can be obtained using any lattice in $n = 2k$ dimensions. Thus we have established the following result.

Theorem 2.4: Let Λ be an n-dimensional lattice, and let $n = 2k$. Then the coding gain of Λ over $\mathbb{Z}^n$ is upper-bounded as follows:

$$\gamma_{\text{eff}}(\Lambda) \le \frac{\zeta(k; P_e)^2}{z(k; P_e)} \cdot \frac{4(k!)^{1/k}}{\pi}. \tag{19}$$

The coding gain $\gamma_{\text{eff}}(\Lambda)$ is defined in terms of lattice SNR, and the foregoing bound is parametrized by both the dimension and the probability of symbol error.

The bound of (19) is tabulated for normalized error probabilities $P_e^* = 10^{-5}, 10^{-6}, 10^{-7}$ and dimensions $n \le 32$ in Table I. All the entries in Table I are given in terms of decibels. The upper bound in the last column is the Rogers bound [24], except for dimensions $n \le 8$ where the densest possible lattice packings are known [6, p. 164]. Observe that the upper bound of Theorem 2.4 is not asymptotic for $P_e \to 0$; it is reasonably tight for symbol error rates of practical interest. As can be seen from Table I, it is generally much tighter than the results obtained by computing the nominal (asymptotic for $P_e \to 0$) coding gains of the densest known n-dimensional lattices, and/or the upper bounds thereupon.

C. Further Results and Asymptotics

We conclude this section with some further remarks and observations. First, we investigate the asymptotic behavior of our bounds as a function of the dimension n, as $n \to \infty$.

All the bounds in this section follow from the assertion $P_e \ge P_{e,\mathcal{S}}$ established in Lemma 2.1. Recall that $P_{e,\mathcal{S}}$ denotes the probability that a white Gaussian n-tuple with variance σ^2 per dimension falls outside the n-dimensional sphere of radius $\alpha\sigma\sqrt{n}$. It follows that $\lim_{n\to\infty} P_{e,\mathcal{S}} = 1$ if $\alpha < 1$, and $\lim_{n\to\infty} P_{e,\mathcal{S}} = 0$ if $\alpha > 1$, by the weak law of large numbers.

Combining this fact and Theorem 2.2 we conclude that asymptotically, reliable communication is impossible unless the lattice SNR satisfies $\alpha^2 \ge 1 = 0$ dB. This result is analogous to the Shannon limit [17], discussed in more detail in the next section. If $\alpha^2 > 1$, then according to our results P_e could, in principle, approach zero as $n \to \infty$. In his 1975 paper, De Buda [7] essentially asserts that there do exist lattices with this property.

Here are some observations regarding the normalized radius of the equivalent sphere $\mathcal{S}_\Pi$ and the lattice SNR. Using the Stirling approximation of $k! \approx (k/e)^k$ in (4), the expression on the right-hand side of (5) becomes an approximation of ρ^2, and we get

$$\alpha^2 \approx \frac{V(\Lambda)^{2/n}}{2\pi e\sigma^2} \tag{20}$$

by the definition of α^2 in (13). This approximation for α^2 was suggested by Forney in [13]. The lattice signal-to-noise ratio has another insightful interpretation as follows. Define the *equivalent noise sphere* $\mathcal{S}_{\boldsymbol{X}}$ as the n-dimensional sphere of squared radius $n\sigma^2$. Then

$$\frac{V(\Lambda)^{2/n}}{V(\mathcal{S}_{\boldsymbol{X}})^{2/n}} = \frac{V(\mathcal{S}_\Pi)^{2/n}}{V(\mathcal{S}_{\boldsymbol{X}})^{2/n}} = \frac{V_n^{2/n}(n\rho^2)}{V_n^{2/n}(n\sigma^2)} = \alpha^2. \tag{21}$$

Thus α^2 is the normalized ratio of the volume of the lattice at hand to the volume of the equivalent noise sphere. Notice that, unlike (20), the derivation in (21) is exact.

Finally, we notice here that the main results of this section hold not only for lattices, but also for geometrically uniform sphere packings [12], since the Voronoi regions in a geometrically uniform packing are always congruent to each other. Forney [13] remarks that these results, in fact, hold for all sphere packings that have a well-defined volume in $\mathbb{R}^n$, for instance, packings that have M points per basic cell of volume V which tiles $\mathbb{R}^n$. As a specific example, consider a coset code $\mathbb{C}$, based on an n-dimensional lattice partition Λ/Λ' and a code over Λ/Λ' whose rate is R bits per two dimensions (cf. [10]). Then the coset code $\mathbb{C}$ has a normalized volume $V(\mathbb{C})^{2/n} = V(\Lambda')^{2/n}/2^R$, regardless of whether it is a lattice or not. Extending this argument, we see that *multilevel* coset codes [10], [15] also have well-defined normalized volumes. The lower bounds on the probability of symbol error and the upper bounds on coding gain derived in this section thus hold for all such codes.

III. Bounds for Lattice Codes

In practice, only a finite set of points of a lattice Λ can be used for transmitting information over a channel. This set of points is usually called a lattice code based on Λ. In this section, we convert the bounds developed in the previous section in the context of *lattices* into powerful bounds on the performance of *lattice codes*. To do so, we rely on the continuous approximation technique [4], [9], [16], which is briefly reviewed in what follows.

A. Preliminaries: Continuous Approximation

Let $\Omega = \Lambda + \boldsymbol{a}$ be the translate of an n-dimensional lattice Λ by a vector $\boldsymbol{a}$, and let $\mathbb{D}$ be a convex, measurable, nonempty bounded region of $\mathbb{R}^n$. Then a *lattice code* $\mathbb{C} = \mathbb{C}(\Lambda, \mathbb{D})$ is defined by $\mathbb{C} = \Omega \cap \mathbb{D}$, and $\mathbb{D}$ is called the *support region* of the code. Because the support region is bounded, a lattice code has finitely many points, say $\mathbb{C} = \{\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_M\}$. The quantity $R = \log_2(M)/n$ is called the *rate* of the code $\mathbb{C}$.

Given a point $\boldsymbol{y} \in \mathbb{R}^n$, we define the average energy per dimension, or the *power* of $\boldsymbol{y}$ as $\|\boldsymbol{y}\|^2/n$. The average power of the code $\mathbb{C} = \{\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_M\}$ is then given by

$$P_{\rm av} = \frac{1}{M}\sum_{i=1}^{M} \frac{\|\boldsymbol{y}_i\|^2}{n}.$$

If the number of codewords M is large, then it can be approximated as $M \simeq V(\mathbb{D})/V(\Pi)$. Thus we have

$$R = \frac{\log_2(V(\mathbb{D})/V(\Pi))}{n} + o(1) \tag{22}$$

and

$$P_{\rm av} = \frac{\sum_{i=1}^{M} \|\boldsymbol{y}_i\|^2 V(\Pi)}{nV(\mathbb{D})} + o(1) \tag{23}$$

where $o(1)$ is a function of the number of codewords M that tends to zero as $M \to \infty$. The numerator of (23) is a Riemann sum that can be further approximated by $\int_{\mathbb{D}} \|\boldsymbol{x}\|^2 \, d\boldsymbol{x}$. This, along with (22) and (23), is known as the *continuous approximation* [4], [9], [16]. The continuous approximation implies that

$$P_{\rm av} = \mathrm{G}(\mathbb{D})V(\mathbb{D})^{2/n} + o(1) \tag{24}$$

where

$$\mathrm{G}(\mathbb{D}) \stackrel{\rm def}{=} \frac{\int_{\mathbb{D}} \|\boldsymbol{x}\|^2 \, d\boldsymbol{x}}{nV(\mathbb{D})^{\frac{2}{n}+1}}$$

is the normalized second moment of the support region $\mathbb{D}$. Thus under the continuous approximation, the average power of a lattice code $\mathbb{C}(\Lambda, \mathbb{D})$ depends only on $\mathbb{D}$. The quantity $\gamma_s(\mathbb{D}) = 1/12\mathrm{G}(\mathbb{D})$ is known [4], [16] as the *shaping gain* of the support region $\mathbb{D}$. It is well known [4], [9] that the highest possible shaping gain is obtained when $\mathbb{D}$ is a sphere, in which case

$$\gamma_s(\mathbb{D}) = \frac{\pi(n+2)}{12\Gamma\left(\frac{n}{2}+1\right)^{2/n}}. \tag{25}$$

We now introduce the signal-to-noise ratio that will be used in the remainder of this paper. As in [9] and [16], we define the *normalized signal-to-noise ratio* as

$$\mathrm{SNR}_{\rm norm} \stackrel{\rm def}{=} \frac{P_{\rm av}}{(2^{2R}-1)\sigma^2} \tag{26}$$

where σ^2 is the noise variance per dimension. The normalized signal-to-noise ratio allows one to compare lattice codes of different rates on the same scale. Another motivation for the definition in (26) is as follows. Since the capacity of the AWGN channel is given by

$$\frac{1}{2}\log_2\left(1 + \frac{P_{\rm av}}{\sigma^2}\right)$$

Shannon's theorem [25] for Gaussian channels has a concise statement in terms of $\mathrm{SNR}_{\rm norm}$. Namely, arbitrarily small probabilities of symbol error can be achieved arbitrarily close to $\mathrm{SNR}_{\rm norm} = 0$ dB.

For high rates R, we have $2^{2R}-1 \simeq 2^{2R}$ in the denominator of (26), and therefore by (22) and (24), the normalized signal-to-noise ratio can be written as

$$\mathrm{SNR}_{\rm norm} = \frac{\mathrm{G}(\mathbb{D})V(\Lambda)^{2/n}}{\sigma^2} + o(1) \tag{27}$$

where $o(1)$ denotes a function of the rate R that tends to zero as $R \to \infty$. This expression makes it possible to establish the connection between $\mathrm{SNR}_{\rm norm}$ and the lattice signal-to-noise ratio α^2 introduced in the previous section. Specifically, we have

$$\mathrm{SNR}_{\rm norm} = \alpha^2 \cdot \frac{n\pi\mathrm{G}(\mathbb{D})}{\Gamma\left(\frac{n}{2}+1\right)^{2/n}} + o(1). \tag{28}$$

For a spherical support region $\mathbb{D}$, this reduces to

$$\mathrm{SNR}_{\rm norm} = \alpha^2 \cdot \frac{n}{(n+2)} + o(1).$$

Thus the lattice signal-to-noise ratio α^2 is closely related to the more conventional normalized SNR.

Now let d denote the minimum distance between the points of the underlying lattice Λ. Combining (27) with the definition of the coding gain $\gamma(\Lambda)$ in (2) and the definition of the shaping gain $\gamma_s(\mathbb{D})$ gives the following result.

Lemma 3.1:

$$\left(\frac{d}{2\sigma}\right)^2 = 3\gamma_s(\mathbb{D})\gamma(\Lambda)\mathrm{SNR}_{\rm norm} + o(1).$$

Lemma 3.1 is well known; see for instance [9], [16], and [26]. Since the proof of Lemma 3.1 relies, as in [16] and [26], on the continuous approximation, the expression $3\gamma_s(\mathbb{D})\gamma(\Lambda)\mathrm{SNR}_{\rm norm}$ is an accurate estimate of $(d/2\sigma)^2$ only for lattice codes of high rate.[1] For more details on the accuracy of the continuous approximation, we refer the reader to [1] and [18].

B. The Main Results

If a point $\boldsymbol{y} \in \mathbb{C}(\Lambda, \mathbb{D})$ is transmitted through an AWGN channel, then the channel output $\boldsymbol{y}+\boldsymbol{X}$ is decoded to $\boldsymbol{y}$ under maximum-likelihood decoding if and only if $\boldsymbol{y}+\boldsymbol{X}$ belongs to the Voronoi region of $\boldsymbol{y}$ in the code $\mathbb{C}(\Lambda, \mathbb{D})$. Thus if $\boldsymbol{y}$ is sufficiently far from the boundary of the support region $\mathbb{D}$, the probability of correct decoding is still given by

$$P_c = \int_{\Pi} f(\boldsymbol{x}) \, d\boldsymbol{x}, \tag{29}$$

[1] Given a lattice $\mathcal{L}$ and a support region $\mathcal{D}$ of certain shape, there are two ways to make the rate of a lattice code based on $\mathcal{L}$ and $\mathcal{D}$ sufficiently high. The first option is to set $\Lambda = \mathcal{L}$ and $\mathbb{D} = c\mathcal{D}$. By making the scaling constant c arbitrarily large, we can make the rate of $\mathbb{C}(\Lambda, \mathbb{D})$ arbitrarily high. However, in view of (24), this increases the average power of $\mathbb{C}$. Another option is to set $\mathbb{D} = \mathcal{D}$ and $\Lambda = c\mathcal{L}$ for a sufficiently small scaling constant c. This way the rate of $\mathbb{C}(\Lambda, \mathbb{D})$ can be made as high as desired, while keeping the support region (and, hence, the power) constant. This is the option we assume throughout this paper. We note that both the coding gain $\gamma(\mathcal{L})$ and the shaping gain $\gamma_s(\mathcal{D})$ are invariant under scaling.

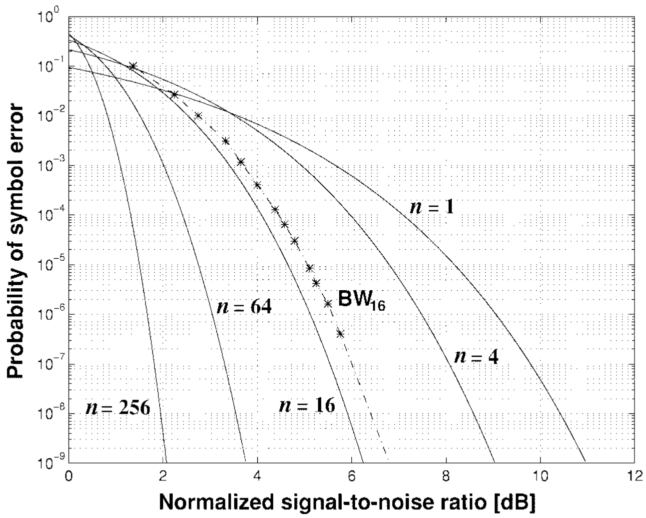

Fig. 2. Lower bounds on the performance of lattice codes.

as in (3), where Π is the Voronoi region of the underlying lattice Λ. If $\boldsymbol{y}$ lies close to the boundary of $\mathbb{D}$, then (29) is not necessarily valid, since then the Voronoi region of $\boldsymbol{y}$ in the lattice code $\mathbb{C}(\Lambda, \mathbb{D})$ is not necessarily equal to the Voronoi region of $\boldsymbol{y}$ in the lattice Λ. In fact, points in $\mathbb{C}(\Lambda, \mathbb{D})$ that lie close to the boundary of $\mathbb{D}$ may have unbounded Voronoi regions. However, we show in the Appendix that for high-rate codes, this boundary effect is negligible. This observation makes it possible to translate the results of the foregoing section into the language of lattice codes. In particular, the following theorem is the counterpart of Theorem 2.2.

Theorem 3.2: If an n-dimensional lattice code $\mathbb{C}(\Lambda, \mathbb{D})$ is used to transmit information over an AWGN channel, then the probability of symbol error under maximum-likelihood decoding is lower-bounded by (9) and (10), with

$$z = \frac{6\Gamma\left(\frac{n}{2}+1\right)^{2/n}}{\pi}\gamma_s(\mathbb{D})\mathrm{SNR}_{\mathrm{norm}} + o(1). \qquad (30)$$

Proof: It is easty to see from (29) that for high-rate latttice codes, the probability of symbol error is still bounded by (9) and (10), as in Theorem 2.2. Further, the expression for $z = r^2/2\sigma^2$ in Theorem 2.2 follows from (2), (4), (11), and Lemma 3.1. The term $o(1)$ in (30) denotes a function of the rate R of $\mathbb{C}(\Lambda, \mathbb{D})$ that tends to zero as $R \to \infty$. □

Theorem 3.2 is a fundamental bound on the performance of any n-dimensional lattice code on an AWGN channel. This bound is plotted in Fig. 2 for $n = 1, 4, 16, 64, 256$, ignoring the term $o(1)$ in (30). We have assumed in Fig. 2 the highest possible shaping gain for a *spherical* support region $\mathbb{D}$ in n dimensions, given by (25). We again include in Fig. 2 the simulation results for a lattice code based on the Barnes–Wall lattice BW_{16} with a spherical support region. It is instructive to compare the corresponding curves in Figs. 1 and 2. These figures confirm that our bounds are tight for both lattices and lattice codes.

We now convert the bound of Theorem 3.2 into an upper bound on the coding gain of lattice codes. The derivation of this bound is similar to the development in Section II-B. First observe that for a spherical support region $\mathbb{D}$, the expression for z in (30) reduces to

$$\begin{aligned} z &= \frac{6\Gamma\left(\frac{n}{2}+1\right)^{2/n}}{\pi} \cdot \frac{\pi(n+2)}{12\Gamma\left(\frac{n}{2}+1\right)^{2/n}} \cdot \mathrm{SNR}_{\mathrm{norm}} + o(1) \\ &= (k+1)\mathrm{SNR}_{\mathrm{norm}} + o(1) \end{aligned} \qquad (31)$$

where we have assumed that $n = 2k$ is even. Now recall that $z(k; P_e)$ was defined in Section II-B as the unique solution of the equation $g_k(x) = P_e$, where $g_k(x)$ is the function defined in (14). Together with (31), this establishes the following result, which is the counterpart of Theorem 2.3 for lattice codes.

Theorem 3.3: To achieve a probability of symbol error P_e using a lattice code $\mathbb{C}$ of rate R in $n = 2k$ dimensions, a normalized signal-to-noise ratio of at least

$$\mathrm{SNR}_{\mathrm{norm}} \geq \frac{z(k; P_e)}{k+1} + o(1)$$

is required, where $o(1)$ is a function of the rate R that tends to zero as $R \to \infty$.

Now recall that $\zeta(k; P_e)$ was defined in Section II-B as the unique solution of the equation $(1 - \mathrm{erfc}\,(x))^{2k} = 1 - P_e$. Further observe that $\gamma(c\mathbb{Z}^n) = 1$ from (2), and hence under the continuous approximation we can write

$$\frac{c}{\sqrt{8}\sigma} = \sqrt{\frac{3}{2}\gamma_s(\mathbb{D})\mathrm{SNR}_{\mathrm{norm}} + o(1)} \qquad (32)$$

in view of Lemma 3.1. This is so because the minimum distance between the points of $c\mathbb{Z}^n$ is precisely c. Together with (17), this establishes the following result, which is the counterpart of Theorem 2.4 for lattice codes.

Theorem 3.4: Let $\mathbb{C}(\Lambda, \mathbb{D})$ be a high-rate n-dimensional lattice code with a spherical support region $\mathbb{D}$, and let $n = 2k$. Then the coding gain of $\mathbb{C}(\Lambda, \mathbb{D})$ over $\mathbb{C}(c\mathbb{Z}^n, \mathbb{D})$ is upper-bounded by

$$\gamma_{\mathrm{eff}}(\mathbb{C}) \leq \frac{\zeta(k; P_e)^2}{z(k; P_e)} \cdot \frac{4\Gamma(k+1)^{1/k}}{\pi}.$$

The coding gain $\gamma_{\mathrm{eff}}(\Lambda)$ is defined in terms of the normalized SNR, and the scaling constant c is chosen in such a way that $\mathbb{C}(\Lambda, \mathbb{D})$ and $\mathbb{C}(c\mathbb{Z}^n, \mathbb{D})$ have the same rate.

The asymptotic results for lattice codes are also analogous to the asymptotics for lattices, discussed in Section II-C. Specifically, it is possible to show, using the weak law of large numbers, that

$$\lim_{k \to \infty} \frac{z(k; P_e)}{k+1} = 1$$

regardless of the desired symbol error rate P_e. Thus the lower bound of Theorem 3.3 coincides with the Shannon limit $\mathrm{SNR}_{\mathrm{norm}} = 0$ dB as $k \to \infty$. This is consistent with the converse to the Shannon theorem for lattice codes. Notably, our proof of this result relies solely on the geometric notion of equivalent sphere, and does not involve information-theoretic arguments. A well-known conjecture in coding theory [9], [16] says that lattice codes achieve capacity on the Gaussian channel. (De Buda [8] attempted to prove this for lattice codes whose support region is a "thick" shell. However, his proof was shown to be incorrect in [22], although [22] includes an alternative proof that only applies to "thin" shells. See also the recent work of Loeliger [23] and Urbanke [29] on this problem.) If this conjecture is true, this would imply that the bounds derived in this paper are asymptotically (for $n \to \infty$) exact at all signal-to-noise ratios.

Remark: Although our focus in this paper is on maximum-likelihood decoding, it is possible to extend the lower bounds of Theorem 2.2 and Theorem 3.2 to suboptimal decoding algorithms, such as bounded-distance decoding. Indeed, consider a decoding algorithm with decision region $\Upsilon \neq \Pi$, namely, an algorithm that decodes a channel output $\boldsymbol{y} + \boldsymbol{X}$ to $\boldsymbol{y}$ if and only if $\boldsymbol{y} + \boldsymbol{X} \in \boldsymbol{y} + \Upsilon$ for some $\boldsymbol{y} \in \mathbb{C}(\Lambda, \mathbb{D})$, and declares a decoding failure otherwise. Ignoring for simplicity the distinction between decoding errors and decoding failures, we can proceed as in Lemma 2.1 and Theorem 2.2, with the equivalent sphere of Π replaced by the equivalent sphere of Υ. The radius of the equivalent sphere of Υ is given by

$$\tilde{r} = \frac{V(\Upsilon)^{1/n}}{V_n^{1/n}} = \frac{V(\Upsilon)^{1/n}\Gamma\left(\frac{n}{2}+1\right)^{1/n}}{\sqrt{\pi}}$$

as in (4). This implies that the probability of error or failure is lower-bounded by (9) and (10), with

$$z = \frac{\tilde{r}^2}{2\sigma^2} = \frac{V(\Upsilon)^{2/n}\Gamma\left(\frac{n}{2}+1\right)^{2/n}}{2\pi\sigma^2} \tag{33}$$

for a lattice Λ, and

$$z = \frac{\tilde{r}^2}{2\sigma^2} = \frac{6\Gamma\left(\frac{n}{2}+1\right)^{2/n}}{\pi}\gamma_s(\mathbb{D})\frac{V(\Upsilon)^{2/n}}{V(\Lambda)^{2/n}}\mathrm{SNR}_{\mathrm{norm}} + o(1) \tag{34}$$

for a lattice code $\mathbb{C}(\Lambda, \mathbb{D})$. Thus the volume of the decision region is an important indicator of the performance of a decoding algorithm. We observe, however, that the bounds based on (33) and (34) are likely to be less tight than the corresponding bounds for maximum-likelihood decoding in Theorem 2.2 and Theorem 3.2. This is so because the proximity of our lower bounds to the actual performance (as evidenced for BW_{16} in Figs. 1 and 2) is largely due to the fact that the Voronoi regions of dense lattices are "nearly" spherical [4], [5], [6], [11]. There is no reason why this should be so for the decision regions of a suboptimal decoding algorithm.

APPENDIX

In this Appendix we prove that for a lattice code $\mathbb{C}(\Lambda, \mathbb{D})$ with a spherical support region $\mathbb{D}$, the probability of correct decoding is given by

$$P_c = \int_{\Pi} f(\boldsymbol{x})\, d\boldsymbol{x} + o(1) \tag{35}$$

where Π is the Voronoi region of Λ, and $o(1)$ is a function of the rate R of $\mathbb{C}(\Lambda, \mathbb{D})$ that tends to zero as $R \to \infty$. Thus (29) used in Section III is a valid approximation of the probability

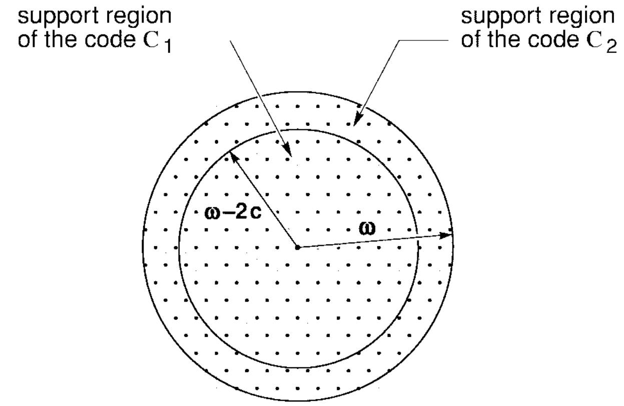

Fig. 3. Partition of $\mathbb{C}(\Lambda, \mathbb{D})$ into boundary points and points with $\Pi(\boldsymbol{y}) = \Pi$.

of correct decoding for high-rate lattice codes. We observe that although only spherical support regions are considered in this Appendix, a similar proof applies for other support regions, such as a cube (QAM constellation [10], [28]) or a Voronoi region of a sublattice (Voronoi constellation [5], [11]).

If a codeword $\boldsymbol{y}$ of a lattice code $\mathbb{C}(\Lambda, \mathbb{D})$ is transmitted through an AWGN channel, and a maximum-likelihood decoder is applied to the channel output, then the probability of correct decoding is

$$P_c = \int_{\Pi(\boldsymbol{y})} f(\boldsymbol{x})\, d\boldsymbol{x} \tag{36}$$

where $\boldsymbol{y}+\Pi(\boldsymbol{y})$ is the Voronoi region of $\boldsymbol{y}$ in the code $\mathbb{C}(\Lambda, \mathbb{D})$. Clearly, $\Pi \subseteq \Pi(\boldsymbol{y})$, and, therefore,

$$P_c \geq \int_{\Pi} f(\boldsymbol{x})\, d\boldsymbol{x}. \tag{37}$$

Furthermore, as illustrated in Fig. 3, $\Pi(\boldsymbol{y})$ is congruent to Π, unless the point $\boldsymbol{y}$ lies sufficiently close to the boundary of $\mathbb{D}$. The basic idea is that for a "well-behaved" support region $\mathbb{D}$, the percentage of such points must approach zero as the rate increases. Although this is intuitively clear, we will provide a proof of this statement in this Appendix.

Assuming that $\mathbb{D}$ is a sphere centered at the origin, let ω denote the radius of $\mathbb{D}$. We further assume for simplicity that $\mathbb{C}(\Lambda, \mathbb{D}) = \Lambda \cap \mathbb{D}$, and let c denote the covering radius of Λ. It is clear that $\omega \geq 3c$, provided the rate of $\mathbb{C}$ is high enough, and we assume that this relation holds in what follows. Later in this Appendix we will prove that $\omega \geq 3c$ for all lattice codes of sufficiently high rate, based on nontrivial lattices.

Lemma A.1: If $\boldsymbol{y} \in \mathbb{C}(\Lambda, \mathbb{D})$ is at distance $\geq 2c$ from the boundary of $\mathbb{D}$, namely if

$$||\boldsymbol{y}|| \leq \omega - 2c$$

then $\Pi(\boldsymbol{y})$ is congruent to Π.

Proof: The faces of the Voronoi region of $\boldsymbol{y}$ in Λ are determined by the neighbors of $\boldsymbol{y}$ in Λ, and all these neighbors are within distance $2c$ from $\boldsymbol{y}$. Thus if $||\boldsymbol{y}|| \leq \omega - 2c$, then all the neighbors of $\boldsymbol{y}$ in Λ are contained in $\mathbb{C}(\Lambda, \mathbb{D}) = \Lambda \cap \mathbb{D}$, and the Voronoi region of $\boldsymbol{y}$ in $\mathbb{C}(\Lambda, \mathbb{D})$ coincides with the Voronoi region of $\boldsymbol{y}$ in Λ. □

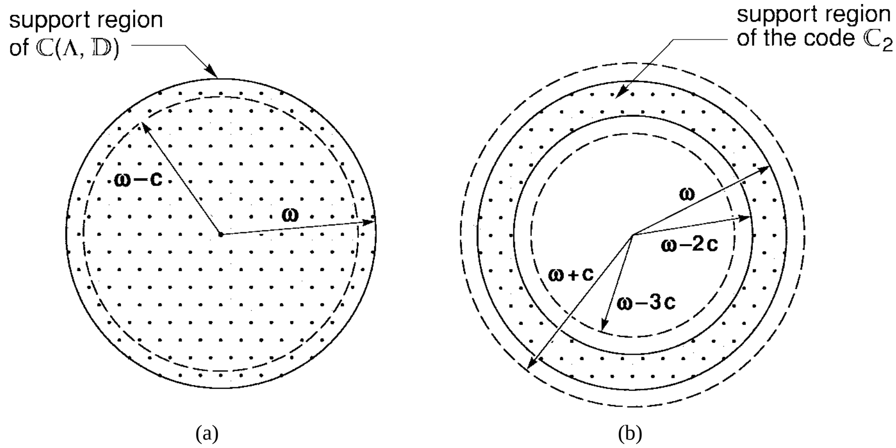

Fig. 4. Estimation of the number of boundary points in a spherical constellation. (a) Lower bound on the total number of points in $\mathbb{C}(\Lambda, \mathbb{D})$. (b) Upper bound on the number of points in the code $\mathbb{C}_2$.

In view of Lemma A.1, we now partition $\mathbb{C}(\Lambda, \mathbb{D})$ as illustrated in Fig. 3. Namely, we partition $\mathbb{C}(\Lambda, \mathbb{D})$ into two disjoint subcodes $\mathbb{C}_1$ and $\mathbb{C}_2$ defined by

$$\mathbb{C}_1 \stackrel{\text{deff}}{=} \{\boldsymbol{y} \in \mathbb{C}(\Lambda, \mathbb{D}) : \|\boldsymbol{y}\| \le \omega - 2c\}$$
$$\mathbb{C}_2 \stackrel{\text{deff}}{=} \{\boldsymbol{y} \in \mathbb{C}(\Lambda, \mathbb{D}) : \|\boldsymbol{y}\| > \omega - 2c\}.$$

Let M_1 and M_2 denote the number of codewords in $\mathbb{C}_1$ and $\mathbb{C}_2$, respectively. Then, assuming equal a priori probability of transmission we have

$$\begin{aligned} P_c &\le \Pr\{\boldsymbol{y} \in \mathbb{C}_1\} \int_\Pi f(\boldsymbol{x})\, d\boldsymbol{x} + \Pr\{\boldsymbol{y} \in \mathbb{C}_2\} \\ &\le \int_\Pi f(\boldsymbol{x})\, d\boldsymbol{x} + \frac{M_2}{M_1 + M_2}. \end{aligned} \tag{38}$$

It follows from (37) and (38) that in order to establish the claim of (35), it remains to show that the ratio $M_2/(M_1+M_2)$ tends to zero as the rate $R \to \infty$.

As illustrated (for the hexagonal lattice) in Fig. 4, the number of codewords in $\mathbb{C}(\Lambda, \mathbb{D})$ is lower-bounded by

$$M_1 + M_2 \ge \frac{V_n(\omega - c)^n}{V(\Pi)} \tag{39}$$

since every point in a sphere of radius $\omega - c$ about the origin belongs to $\boldsymbol{y} + \Pi$ for a unique point $\boldsymbol{y} \in \mathbb{C}(\Lambda, \mathbb{D})$. On the other hand,

$$M_2 \le \frac{V_n(\omega + c)^n - V_n(\omega - 3c)^n}{V(\Pi)} \tag{40}$$

since the set $\cup_{\boldsymbol{y} \in \mathbb{C}_2}(\boldsymbol{y} + \Pi)$ is properly contained within a shell of inner radius $\omega - 3c$ and outer radius $\omega + c$ (see, again, Fig. 4). Combining (39) and (40) yields

$$\frac{M_2}{M_1 + M_2} \le \frac{(\omega + c)^n - (\omega - 3c)^n}{(\omega - c)^n} \le 2^n \left[1 - \left(\frac{\omega - 3c}{\omega + c}\right)^n\right] \tag{41}$$

where the second inequality follows from the relation $\omega \ge 3c$. Furthermore,

$$\begin{aligned} 1 - \left(\frac{\omega - 3c}{\omega + c}\right)^n &= \frac{4c}{\omega + c}\left(1 + \frac{\omega - 3c}{\omega + c} + \frac{(\omega - 3c)^2}{(\omega + c)^2}\right. \\ &\quad \left. + \cdots + \frac{(\omega - 3c)^{n-1}}{(\omega + c)^{n-1}}\right) \le 4n\left(\frac{c}{\omega + c}\right) \end{aligned}$$

for all n, which together with (41) implies

$$\frac{M_2}{M_1 + M_2} \le n2^{n+2}\left(\frac{c}{\omega + c}\right). \tag{42}$$

We can now relate the upper bound in (42) to the rate R of $\mathbb{C}(\Lambda, \mathbb{D})$ as follows. Clearly,

$$V_n(\omega + c)^n \ge MV(\Pi) = 2^{nR}V(\Pi) \tag{43}$$

and, therefore,

$$\omega + c \ge 2^R \frac{V(\Lambda)^{1/n}}{V_n^{1/n}} = \frac{2^R \rho}{\Delta(\Lambda)^{1/n}} \tag{44}$$

where ρ is the packing radius of Λ and $\Delta(\Lambda)$ is the packing density of Λ.

It follows from (42) and (44) that

$$\frac{M_2}{M_1 + M_2} \le \frac{1}{2^R}\left(\frac{c}{\rho}\right) n2^{n+2}\Delta(\Lambda)^{1/n}. \tag{45}$$

Since for a given n-dimensional lattice Λ, the parameters c/ρ and $\Delta(\Lambda)$ are fixed, the bound in (45) implies that for all $\varepsilon > 0$ and for *any given* lattice, there exists a rate R_0, such that $M_2/(M_1 + M_2) \le \varepsilon$ for all $R \ge R_0$. In view of (37) and (38), this suffices to prove the claim in (35), which is our main objective in this Appendix.

However, it is worthwhile to establish the following slightly stronger result. We say that an n-dimensional lattice Λ is *non-trivial* if it is at least as dense as the integer lattice $\mathbb{Z}^n$. In the remainder of this Appendix, we prove that for all $\varepsilon > 0$ and for *all* nontrivial n-dimensional lattices, there exists a rate R_0 such that $M_2/(M_1+M_2) \leq \varepsilon$ for all $R \geq R_0$. To this end, we first recast the bound of (45) as

$$\frac{M_2}{M_1+M_2} \leq \frac{1}{2^R}\left(\frac{c}{\rho}\right) n 2^{n+1} \frac{\sqrt{\pi}}{\Gamma\left(\frac{n}{2}+1\right)^{1/n}} \sqrt{\gamma(\Lambda)} \tag{46}$$

where $\gamma(\Lambda)$ is the coding gain defined in (2). Next, we make use of the following lemma, due to Hermite [21].

Lemma A.2: For any n-dimensional lattice Λ, there exists a basis $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ with

$$\prod_{i=1}^{n} \|\boldsymbol{v}_i\| \leq V(\Lambda)\left(\frac{2}{\sqrt{3}}\right)^{\frac{n(n-1)}{2}}. \tag{47}$$

Dividing both sides of (47) by $d(\Lambda)^n$ yields

$$\prod_{i=1}^{n} \frac{\|\boldsymbol{v}_i\|}{d(\Lambda)} \leq \frac{1}{\gamma(\Lambda)^{n/2}}\left(\frac{2}{\sqrt{3}}\right)^{\frac{n(n-1)}{2}}.$$

Since $\|\boldsymbol{v}\| \geq d(\Lambda)$ for all nonzero $\boldsymbol{v} \in \Lambda$, we may further conclude that

$$\|\boldsymbol{v}_i\| \leq \frac{d(\Lambda)}{\gamma(\Lambda)^{n/2}}\left(\frac{2}{\sqrt{3}}\right)^{\frac{n(n-1)}{2}}, \qquad \text{for } i = 1, 2, \cdots, n. \tag{48}$$

Now, consider the fundamental parallelotope (or the fundamental region) of the lattice Λ spanned by the basis $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$, that is,

$$\Xi \stackrel{\text{deff}}{=} \{\alpha_1\boldsymbol{v}_1 + \alpha_2\boldsymbol{v}_2 + \cdots + \alpha_n\boldsymbol{v}_n : 0 \leq \alpha_1, \alpha_2, \cdots, \alpha_n < 1\}.$$

By the definition of a fundamental region, every point of $\mathbb{R}^n$ lies in the translate of Ξ by some point of Λ. Thus the distance from any point of $\mathbb{R}^n$ to the lattice Λ cannot be greater than the distance from the farthest point of Ξ to the origin. This leads to the following upper bound on the covering radius of Λ:

$$c \leq \max_{0 \leq \alpha_1, \cdots, \alpha_n < 1} \|\alpha_1\boldsymbol{v}_1 + \alpha_2\boldsymbol{v}_2 + \cdots + \alpha_n\boldsymbol{v}_n\|.$$

Since $\alpha_i < 1$ for all $i = 1, 2, \cdots, n$, the triangle inequality now yields

$$c \leq \|\boldsymbol{v}_1\| + \|\boldsymbol{v}_2\| + \cdots + \|\boldsymbol{v}_n\|.$$

Combining this with (48), we conclude that the ratio of the covering radius to the packing radius of Λ is upper-bounded as follows:

$$\frac{c}{\rho} = \frac{2c}{d(\Lambda)} \leq \frac{2n}{\gamma(\Lambda)^{n/2}}\left(\frac{2}{\sqrt{3}}\right)^{\frac{n(n-1)}{2}}. \tag{49}$$

Note that this is a very crude bound. In fact (cf. [6]), in many of the densest known lattice packings the ratio c/ρ is less than 2. Nevertheless, the bound in (49) will suffice for our purposes.

First, we use (49) to prove that $\omega \geq 3c$ for all lattice codes $\mathbb{C}(\Lambda, \mathbb{D})$ of sufficiently high rate based on a nontrivial lattice Λ. Assume to the contrary that $\omega < 3c$. Then, as in (43), the number of codewords in $\mathbb{C}(\Lambda, \mathbb{D})$ is upper-bounded by

$$M \leq \frac{V_n(\omega+c)^n}{V(\Pi)} \leq \frac{V_n(4c)^n}{V_n\rho^n} = 2^{2n}\left(\frac{c}{\rho}\right)^n$$

since a sphere of packing radius ρ is properly contained in the Voronoi region Π. Hence, if Λ is a nontrivial lattice, then the rate R of $\mathbb{C}(\Lambda, \mathbb{D})$ is at most

$$\begin{aligned} R = \frac{\log_2 M}{n} &\leq 2 + \log_2\left(\frac{c}{\rho}\right) \\ &\leq \frac{n(n-1)}{2}\log_2\left(\frac{2}{\sqrt{3}}\right) + \log_2 n + 3 \end{aligned} \tag{50}$$

where the second inequality follows from (49), along with the observation that if Λ is nontrivial then $\gamma(\Lambda) \geq \gamma(\mathbb{Z}^n) = 1$. It follows that for all nontrivial n-dimensional lattice codes whose rate exceeds (50), we must have $\omega \geq 3c$.

Finally, combining (49) with (46), we see that for nontrivial lattice packings

$$\frac{M_2}{M_1+M_2} \leq \frac{1}{2^R}\frac{\sqrt{\pi}n^2 2^{n+2}}{\Gamma\left(\frac{n}{2}+1\right)^{1/n}}\left(\frac{2}{\sqrt{3}}\right)^{\frac{n(n-1)}{2}}. \tag{51}$$

As the right-hand side of (51) depends only on the rate R and the dimension n of $\mathbb{C}(\Lambda, \mathbb{D})$, we may conclude that the percentage of codewords that lie close to the boundary of the support region $\mathbb{D}$ converges to zero uniformly (and exponentially) as $R \to \infty$, for all nontrivial n-dimensional lattice codes.

Acknowledgment

We are grateful to G. David Forney, Jr. who was a referee of the earlier version of this work. His comments significantly improved both the content and the presentation of this paper. We thank Dakshi Agrawal for providing the simulation results for BW_{16} that are included in Figs. 1 and 2. We are also grateful to Bert Hochwald and Dakshi Agrawal for valuable comments and discussions.

References

[1] G. E. Andrews, "A lower bound for the volume of strictly convex bodies with many boundary lattice points," *Trans. Amer. Math. Soc.*, vol. 106, pp. 270–279, 1963.
[2] R. Bacher, "Dense lattices in dimensions 28 and 29," *Inventiones Math.*, vol. 130, pp. 153–158, 1997.
[3] E. R. Berlekamp, "The technology of error-correction codes," in *Proc. IEEE*, vol. 68, pp. 564–593, 1980,
[4] A. R. Calderbank, *Bandwidth Efficient Communication*, in preparation.
[5] J. H. Conway and N. J. A. Sloane, "A fast encoding method for lattice codes and quantizers," *IEEE Trans. Inform. Theory*, vol. IT-29, pp. 820–824, 1983.
[6] J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Groups*. New York: Springer-Verlag, 1988.
[7] R. DeBuda, "The upper error bound of a new near-optimal code," *IEEE Trans. Inform. Theory*, vol. IT-21, pp. 441–445, 1975.
[8] ——, "Some optimal codes have structure," *IEEE J. Select. Areas Commun.*, vol. 7, pp. 893–899, 1989.
[9] M. V. Eyuboğlu and G. D. Forney Jr., "Lattice and trellis quantization with lattice and trellis-bounded codebooks—High-rate theory for memoryless sources," *IEEE Trans. Inform. Theory*, vol. 39, pp. 46–59, 1993.
[10] G. D. Forney Jr., "Coset codes—Part II: Binary lattices and related codes," *IEEE Trans. Inform. Theory*, vol. 34, pp. 1152–1187, 1988.
[11] ——, "Multidimensional constellations—Part II: Voronoi constellations," *IEEE J. Select. Areas Commun.*, vol. 7, pp. 941–958, 1989.
[12] ——, "Geometrically uniform codes," *IEEE Trans. Inform. Theory*, vol. 37, pp. 1241–1260, 1991.
[13] ——, personal communication, Aug. 1996.
[14] ——, "Approaching AWGN channel capacity with coset codes and multilevel coset codes," *IEEE Trans. Inform. Theory*, submitted for publication, 1997.

[15] G. D. Forney, Jr. and A. Vardy, "Generalized minimum distance decoding of Euclidean-space codes and lattices," *IEEE Trans. Inform. Theory*, vol. 42, pp. 1992–2026, 1996.

[16] G. D. Forney, Jr. and L.-F. Wei, "Multidimensional constellations: Introduction, figures of merit, and generalized cross-constellations," *IEEE J. Select. Areas Commun.*, vol. 7, pp. 877–892, 1989.

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

[18] P. M. Gruber and C. G. Lekkerkerker, *Geometry of Numbers*, 2nd ed. New York: Elsevier, 1987.

[19] H. Herzberg and G. Poltyrev, "Techniques of bounding the probability of decoding error for block coded modulation structures," *IEEE Trans. Inform. Theory*, vol. 40, pp. 903–911, 1994.

[20] ——, "The error-probability of M-ary PSK block coded modulation schemes," *IEEE Trans. Commun.*, vol. 44, pp. 427–433, 1996.

[21] S. Lang, "Conjectured Diophantine estimates on elliptic curves," in *Arithmetic and Geometry*, vol. I, M. Artin and J. Tate, Eds. Boston, MA: Birkhäuser, 1983, pp. 155–171.

[22] T. Linder, C. Schlegel, and K. Zeger, "Corrected proof of De Buda's theorem," *IEEE Trans. Inform. Theory*, vol. 39, pp. 1735–1737, 1993.

[23] H.-A. Loeliger, "Averaging bounds for lattices and linear codes," *IEEE Trans. Inform. Theory*, vol. 43, pp. 1767–1773, Nov. 1997.

[24] H. M. de Oliveira and G. Battail, "Performance of lattice codes over the Gaussian channel," *Ann. Télécommun.*, vol. 47, pp. 293–305, 1992.

[25] C. A. Rogers, "The packing of equal spheres," *Proc. London Math. Soc.*, vol. 8, pp. 609–620, 1958.

[26] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423 and pp. 623–656, 1948.

[27] V. Tarokh, "Trellis complexity versus the coding gain of lattice-based communication systems," Ph.D. dissertation, Univ. Waterloo, Ont., Canada, 1995.

[28] V. Tarokh, A. Vardy, and K. Zeger, "On the performance of lattice codes," Res. Rep. UILU-ENG-96-2538, Univ. Illinois, Urbana-Champaign, Jan. 1996.

[29] R. Urbanke and B. Rimoldi, "Lattice codes can achieve capacity on the AWGN channel," *IEEE Trans. Inform. Theory*, vol. 44, pp. 273–278, Jan. 1998.

[30] J. M. Wozencraft and I. M. Jacobs, *Principles of Communication Engineering*. New York: Wiley Interscience, 1965.