Averaging Bounds for Lattices and Linear Codes

Hans-Andrea Loeliger, *Member, IEEE*

***Abstract*— General random coding theorems for lattices are derived from the Minkowski–Hlawka theorem and their close relation to standard averaging arguments for linear codes over finite fields is pointed out. A new version of the Minkowski–Hlawka theorem itself is obtained as the limit, for $p \to \infty$, of a simple lemma for linear codes over GF (p) used with p-level amplitude modulation. The relation between the combinatorial packing of solid bodies and the information-theoretic "soft packing" with arbitrarily small, but positive, overlap is illuminated. The "soft-packing" results are new. When specialized to the additive white Gaussian noise channel, they reduce to (a version of) the de Buda–Poltyrev result that spherically shaped lattice codes and a decoder that is unaware of the shaping can achieve the rate $1/2 \log_2 (P/N)$.**

***Index Terms*— Coded modulation, lattices, linear codes, Minkowski–Hlawka theorem, random coding, shaping.**

I. Introduction

EUCLIDEAN-space lattices have become a standard tool for the construction of both block codes and (convolutional-type) trellis codes for the additive white Gaussian noise (AWGN) channel at high signal-to-noise ratio [1], [2]. Only block codes will be considered in this paper. Such lattice codes consist of the intersection of a lattice Λ (or a translate of a lattice) with a bounded *shaping region* S, which is typically a ball or a "thick shell" centered at the origin.

For the analysis of such codes, it is common to separate the "coding gain," which is provided by the lattice, from the "shaping gain," which stems from the shaping region [2], [3]. In particular, it is usually assumed that the decoder is unaware of the shaping, i.e., it always decodes to the nearest lattice point, whether or not this point lies in S. Such a decoder will be called a *lattice decoder* and should be distinguished from a *nearest-codeword decoder*, which decodes to the nearest lattice point *inside* S. Note that the attractive symmetry properties commonly associated with lattice codes, such as congruent decoding regions, uniform distance profile, and codeword-independent error probability, apply only to a lattice decoder and not to a nearest-codeword decoder.

The main prior work on the information-theoretic limits of lattice codes are two papers by de Buda [4], [5] and a more recent paper by Poltyrev [6]. In his first paper [4], de Buda considered spherically shaped lattice codes with lattice decoding and showed that arbitrarily small error probability can be achieved at all rates below $1/2 \log_2 (P/N)$ (bits per dimension), where P and N are the signal power and the noise variance, respectively; he also gave an exponential error bound.

In his second paper [5], de Buda seemed to have proved that lattice codes can even achieve the full channel capacity $1/2 \log_2 (1 + P/N)$, with the same exponential error bounds as Shannon's random codes [7]. For technical reasons, he considered "thick-shell" shaping rather than spherical shaping; moreover, he assumed a nearest-codeword decoder rather than a lattice decoder. However, an error in [5] was reported by Linder *et al.* [8]. They were able to fix the problem, at the price of replacing de Buda's "thick" shells with "thin" shells. Consequently, the corresponding codes lose most of their lattice structure and rather resemble random spherical codes.

Even in its corrected form, the upper bound on error probability in de Buda's second paper [5] applies only to the average over all codewords of some fixed code. It is thus possible that his codes contain some weak codewords with error probability close to one. Such weak codewords can, of course, be deleted from the code without noticeable loss in rate, but the resulting code is no longer the intersection of (a translate of) a lattice with a spherical shell. In any case, de Buda's second paper says nothing about lattice codes with *lattice* decoding.

Magalhäes and Battail [21] also considered lattice codes and derived error exponents. In fact, they seemed to have proved that, even with lattice decoding, the full capacity $1/2 \log_2 (1 + P/N)$ is achievable. However, a mistake in the proof (the lattice points inside a $(P + N)$-sphere are tacitly assumed to have average power P) invalidated their argument. Indeed, we conjecture that lattice decoding is, in fact, limited to the rate $1/2 \log_2 (P/N)$.

Poltyrev [6] considered unbounded constellations and lattices, for which he gave a Gallager-type exponential random coding bound. He also proved the achievability of $1/2 \log_2 (P/N)$. (However, neither [4] nor [6] paid attention to a subtle problem that will be discussed at the beginning of Section IV.)

All these authors based their lattice results on the following version (due to Hlawka) of the Minkowski–Hlawka theorem [9]–[11], [12, ch. 3, Theorem 1]: for any Riemann integrable function f: $\mathcal{R}^n \to \mathcal{R}$ of bounded support and any positive ϵ, there exists a lattice Λ in $\mathcal{R}^n$ with fundamental volume 1 such that

$$\sum_{x \in \Lambda} f(x) < \int_{\mathcal{R}^n} f(x)\, dx + \epsilon.$$

(A version of this theorem will be proved in Section II.)

Manuscript received July 22, 1994; revised April 15, 1997. The material in this paper was presented in part at the DIMACS/IEEE Workshop on Coding and Quantization, Piscataway, NJ, October 1992, and the 1993 IEEE International Symposium on Information Theory, San Antonio, TX, January 1993.

The author was with Linköping University, Linköping, Sweden. He is now with Endora Tech AG, CH-4052 Basel, Switzerland.

Publisher Item Identifier S 0018-9448(97)06719-9.

0018–9448/97$10.00 © 1997 IEEE

De Buda stressed the point that, for lattices, the Minkowski–Hlawka theorem can replace the usual random coding arguments. However, all known proofs of the Minkowski–Hlawka theorem are obtained from averaging over a large, usually infinite, class of lattices; in this sense, the Minkowski–Hlawka theorem *is* random coding and may be regarded as a pre-Shannon result in information theory. The proof by Cassels [10] (as cited in [4]), is actually based on averaging over the set of $(n, n-1)$ linear codes over GF(p), although the connection to coding is not made explicit.

From the stated version of the Minkowski–Hlawka theorem, one can derive various existence results for "packing lattices." In fact, it was the problem of packing n-dimensional spheres and other bodies (posed by Hilbert [13]) that motivated Minkowski and Hlawka, and the name "Minkowski–Hlawka theorem" was originally (and sometimes still is) used only for one such result. Results of this type will also be considered in Section II.

In this line of development, Rush and Sloane [14], [15] observed that applying a version of the Varshamov–Gilbert bound to certain linear codes over GF(p) used with "Construction A" [16] proves the existence of lattices with the same asymptotic sphere-packing density as those known earlier from the Minkowski–Hlawka theorem. For large n, this is still the best existence result known for such packings. Rush and Sloane did not recognize, however, that (as in Cassels' proof) the Minkowski–Hlawka theorem itself can be obtained from averaging over Construction A lattices. This approach is summarized in [17], which emphasizes the role of (versions of) the Minkowski–Hlawka theorem as the lattice analog to the Varshamov–Gilbert bound.

The present paper aims at making explicit the mentioned connections between the familiar averaging bounds ("random coding") for linear codes and the Minkowski–Hlawka theorem and its applications, both for the combinatorial packing of solid bodies and for the information-theoretic "soft packing" with arbitrarily small, but positive, overlap. First, it is shown as a generalization of Cassels' proof, how the Minkowski–Hlawka theorem may be obtained from a simple averaging lemma for (n, k) linear codes over GF(p), $1 < k < n$, used with p-level amplitude modulation. After a brief discussion of the application to the packing of solid bodies, we will turn to the information-theoretic packing problem and give a simple proof of a new general upper bound on error probability for additive errors. We will then consider shaping and obtain a general existence result for lattice codes, which, when specialized to the AWGN channel, reduces to the de Buda–Poltyrev result that lattice codes can achieve the rate $1/2 \log_2(P/N)$.

The analogy between lattices and linear codes is further elaborated in the Appendix: for each lattice theorem in the main text, the Appendix contains an analogous theorem for linear codes.

Throughout the paper, we will use the following notation. The symbols $\mathcal{Z}$, $\mathcal{R}$, $\mathcal{Z}_p$, and F_q denote the integers, the real numbers, the integers modulo p, and the finite field with q elements, respectively. For a set E (e.g., $E \subset \mathcal{R}^n$ or $E \subset F_q^n$), we write $E^* \triangleq E\backslash\{0\}$ and $\Delta E \triangleq \{e - e'\colon e, e' \in E\}$. The space $\mathcal{R}^n$ will always be assumed to be equipped with the Euclidean metric. For a subset $E \subset \mathcal{R}^n$, $V(E)$ denotes the volume of E. (It will usually be required that $V(E)$ is the Riemann integral of the indicator function of E, which means that E must be Jordan measurable.) Random variables will be denoted by bold-faced letters (e.g., $\boldsymbol{e}$).

II. The Minkowski–Hlawka Theorem

We will derive a new version of the Minkowski–Hlawka theorem as the limit case (for $p \to \infty$) of a simple lemma about linear codes over $\mathcal{Z}_p$ used with p-level amplitude modulation. We begin with the mentioned simple lemma for linear codes. As usual, an (n,k) *linear code over* F_q is a k-dimensional subspace of F_q^n. We say that a set $\mathcal{C}$ of linear (n, k) codes over F_q is *balanced* if every nonzero element of F_q^n is contained in the same number, denoted by $N_{\mathcal{C}}$, of codes from $\mathcal{C}$. It is easily seen that, for fixed n, k, and q, the set of *all* linear (n, k) codes over F_q is balanced.

The term "balanced" is borrowed from [18] where, however, it was used as a property of a set of affine encoders rather than of linear codes. Being balanced is the key property for most averaging arguments for linear codes such as Varshamov–Gilbert-type bounds or random coding bounds; in fact, these bounds can be derived from the following lemma (cf., the Appendix and [19]), versions of which are routinely (and usually implicitly) used in coding texts.

Lemma 1 (Basic Averaging Lemma): Let $f(\cdot)$ be an arbitrary mapping $F_q^n \to \mathcal{R}$; let $\mathcal{C}$ be a balanced set of linear (n, k) codes over F_q. Then the average, over all linear codes C in $\mathcal{C}$, of the sum $\sum_{c\in C^*} f(c)$ (over all nonzero codewords of C) is given by

$$\frac{1}{|\mathcal{C}|}\sum_{C\in\mathcal{C}}\sum_{c\in C^*} f(c) = \frac{q^k-1}{q^n-1}\sum_{v\in(F_q^n)^*} f(v). \tag{1}$$

Proof: It follows from the definition of a balanced set of codes that

$$\sum_{C\in\mathcal{C}}\sum_{c\in C^*} f(c) = N_{\mathcal{C}}\sum_{v\in(F_q^n)^*} f(v). \tag{2}$$

Equation (1) then follows from the relation

$$|\mathcal{C}|\,(q^k-1) = N_{\mathcal{C}}\,(q^n-1) \tag{3}$$

which follows from counting the total number of nonzero codewords of all codes in $\mathcal{C}$, each with its multiplicity. □

We will see that the Minkowski–Hlawka theorem is, in a sense, the lattice version of Lemma 1.

A *lattice* is a discrete additive subgroup of $\mathcal{R}^n$ (Euclidean n-space). The *fundamental volume* $V_f(\Lambda)$ of a lattice Λ is the reciprocal of the number of lattice points per unit volume. We will use $V_f(\Lambda)$ instead of the more usual *determinant* of Λ, which is (usually) defined as $V_f(\Lambda)^2$.

For fixed positive integers n and p, let $\mathcal{Z}^n \to \mathcal{Z}_p^n\colon v \mapsto \overline{v}$ denote the componentwise reduction modulo p. The lattices of this paper are *mod-p lattices*, i.e., of the form $\Lambda_C \triangleq \{v \in \mathcal{Z}^n\colon \overline{v} \in C\}$, where p is a prime and C is a linear code over $\mathcal{Z}_p$ (i.e., [16, "Construction A"]).

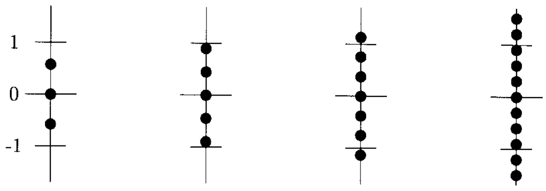

Fig. 1. Scaled p-level signal constellation (amplitude modulation) for $p = 3, 5, 7,$ and 11. The scaling factor γ is adjusted such that the fundamental volume $V_f \triangleq \gamma^n p^{n-k}$ is kept constant (for $k/n = 1/2$).

In fact, we will actually consider *scaled* mod-p lattices, i.e., lattices of the form $\gamma\Lambda_C \triangleq \{\gamma v\colon v \in \Lambda_C\}$ for some $\gamma \in \mathcal{R}$. The fundamental volume of such a lattice is

$$V_f(\gamma\Lambda_C) = \gamma^n p^{n-k} \tag{4}$$

where $n,\ k,$ and p are the blocklength, the dimension, and the alphabet size, respectively, of C. It is usual to consider scaled versions of some same lattice as essentially identical; for the purpose of this paper, however, it is necessary to clearly distinguish between differently scaled versions of a lattice.

Theorem 1 (Minkowski–Hlawka Theorem): Let f be a Riemann integrable function $\mathcal{R}^n \to \mathcal{R}$ of bounded support (i.e., $f(v) = 0$ if $||v||$ exceeds some bound). Then, for any integer k, $0 < k < n$, and any fixed V_f, the approximation

$$\frac{1}{|\mathcal{C}|}\sum_{C\in\mathcal{C}}\ \sum_{v\in\gamma\Lambda_C^*} f(v) \approx V_f^{-1}\int_{\mathcal{R}^n} f(v)\,dv \tag{5}$$

where $\mathcal{C}$ is any balanced set of linear (n, k) codes over $\mathcal{Z}_p$, becomes exact in the limit $p \to \infty,\ \gamma \to 0,\ \gamma^n p^{n-k} = V_f$ fixed.

Before we prove the theorem, we will examine more closely this simultaneous limit $p \to \infty,\ \gamma \to 0$. Since $\gamma^n p^{n-k} = V_f$ is kept fixed, we have $(\gamma p)^n = p^k V_f \to \infty$, which implies $\gamma p \to \infty$. An engineering interpretation of this limit is shown in Fig. 1. We consider p-level amplitude modulation, p an odd prime, with levels

$$\{-\gamma(p-1)/2,\ -\gamma(p-3)/2, \cdots, 0, \cdots, \gamma(p-1)/2\}$$

together with a linear (n, k) code C over $\mathcal{Z}_p$. For $\gamma \to 0,\ p \to \infty$, and $\gamma p \to \infty$, the signal set develops as shown in Fig. 1 and the signal space image $\tilde{C}$ of C becomes essentially equivalent with the lattice $\gamma\Lambda_C$—more precisely, the intersection of $\gamma\Lambda_C$ with any bounded subset S of $\mathcal{R}^n$ equals $S \cap \tilde{C}$. This shows that, in a practical sense, Theorem 1 is about linear codes used with p-level amplitude modulation rather than about lattices.

Proof of Theorem 1: We have

$$\begin{aligned}
&\frac{1}{|\mathcal{C}|}\sum_{C\in\mathcal{C}}\ \sum_{v\in\gamma\Lambda_C^*} f(v)\\
&\quad= \frac{1}{|\mathcal{C}|}\sum_{C\in\mathcal{C}}\left[\sum_{v\in(\mathcal{Z}^n)^*\colon\, \overline{v}=0} f(\gamma v) + \sum_{v\in\mathcal{Z}^n\colon\, \overline{v}\in C^*} f(\gamma v)\right] &(6)\\
&\quad= \sum_{v\in(\mathcal{Z}^n)^*\colon\, \overline{v}=0} f(\gamma v)\\
&\qquad + \frac{1}{|\mathcal{C}|}\sum_{C\in\mathcal{C}}\ \sum_{c\in C^*}\left[\sum_{v\in\mathcal{Z}^n\colon\, \overline{v}=c} f(\gamma v)\right] &(7)\\
&\quad= \sum_{v\in(\mathcal{Z}^n)^*\colon\, \overline{v}=0} f(\gamma v)\\
&\qquad + \frac{p^k-1}{p^n-1}\sum_{c\in(\mathcal{Z}_p^n)^*}\left[\sum_{v\in\mathcal{Z}^n\colon\, \overline{v}=c} f(\gamma v)\right] &(8)\\
&\quad= \sum_{v\in(\mathcal{Z}^n)^*\colon\, \overline{v}=0} f(\gamma v) + \frac{p^k-1}{p^n-1}\sum_{v\in\mathcal{Z}^n\colon\, \overline{v}\neq 0} f(\gamma v) &(9)
\end{aligned}$$

where the step from (7) to (8) follows from Lemma 1. Since f has bounded support, the left term of (9) vanishes for sufficiently large γp and the right term of (9) becomes

$$\frac{p^k-1}{p^n-1}\sum_{v\in(\mathcal{Z}^n)^*} f(\gamma v) \approx p^{k-n}\gamma^{-n}\int_{\mathcal{R}^n} f(v)\,dv \tag{10}$$

which becomes exact in the limit $\gamma \to 0,\ \gamma p \to \infty.$ □

In the next section, we will use the Minkowski–Hlawka theorem in the following form.

Theorem 2: Let E be a bounded subset of $\mathcal{R}^n$ that is Jordan measurable (i.e., $V(E)$ is the Riemann integral of the indicator function of E); let k be an integer such that $0 < k < n$ and let V_f be a positive real number. Then the approximation

$$\frac{1}{|\mathcal{C}|}\sum_{C\in\mathcal{C}} |\gamma\Lambda_C^* \cap E| \approx V(E)\,/\,V_f \tag{11}$$

where $\mathcal{C}$ is any balanced set of linear (n, k) codes over $\mathcal{Z}_p$, becomes exact in the limit $p \to \infty,\ \gamma \to 0,\ \gamma^n p^{n-k} = V_f$ fixed.

Proof: Let f be the indicator function for E (i.e., $f(v)=1$ if $v\in E$ and $f(v)=0$ otherwise) and apply Theorem 1. □

We conclude the discussion of the Minkowski–Hlawka theorem with its application to the classical packing problem. A *packing lattice* for a subset $E \subset \mathcal{R}^n$ is a lattice $\Lambda \subset \mathcal{R}^n$ such that, for any two points $x,\ y \in \Lambda,\ x \neq y$, the sets $x + E$ and $y + E$ are disjoint. Recall our notation $\Delta E \triangleq \{e - e'\colon e - e' \in E\}$. It is easily seen that a lattice Λ is a packing lattice for E if and only if $\Lambda \cap \Delta E = \{0\}$. The following theorem is the lattice analog of the Varshamov–Gilbert bound (cf., Theorem 8 of the Appendix); for $n \to \infty$, it is still the best known general existence result for packing lattices.

Theorem 3: Let E be a bounded subset of $\mathcal{R}^n,\ n \geq 2$, such that the volume $V(\Delta E)$ of ΔE is well defined (i.e., ΔE is Jordan measurable). Then, for any $V_f > V(\Delta E)/2$, there exists a packing lattice Λ for E with fundamental volume $V_f(\Lambda) = V_f$. Moreover, Λ may be chosen to be a scaled version of a mod-p lattice Λ_C for some linear (n, k) code C over $\mathcal{Z}_p$, where k may be chosen freely between 1 and $n-1$ and p is a sufficiently large prime.

Proof: Since $V(\Delta E)\,/\,V_f < 2$, Theorem 2 implies that, for sufficiently large p and sufficiently small $\gamma,\ \gamma^n p^{n-k} = V_f$, there exists a lattice $\gamma\Lambda_C$ such

$$|\gamma\Lambda_C^* \cap \Delta E| < 2. \tag{12}$$

But, since both ΔE and $\gamma\Lambda_C$ are closed under multiplication by -1, the left side of (12) must be even and thus equals zero. □

In the important special case where E is a sphere, we have $V(\Delta E) = 2^n V(E)$, and Theorem 3 guarantees the existence of a packing lattice with density $V(E)/V_f$ arbitrarily close to 2^{-n+1}. We also note that, with a more complicated argument [12, pp. 202–203], the slightly better density $2^{-n+1}(1+2^{-n}+3^{-n}+\cdots)$ can be obtained from Theorem 1.

III. A Ramdom Coding Theorem for Lattices

Assume that a transmitter selects a codeword $\boldsymbol{x}$ from a lattice $\Lambda \in \mathcal{R}^n$. That codeword is then transmitted over a channel that adds a random "noise" vector $\boldsymbol{e} \in \mathcal{R}^n$. We assume that $\boldsymbol{e}$ is independent of $\boldsymbol{x}$ but may have an arbitrary probability density. The receiver obtains $\boldsymbol{y} \overset{\Delta}{=} \boldsymbol{x}+\boldsymbol{e}$ and tries to recover $\boldsymbol{x}$.

We now assume that some set $E \subset \mathcal{R}^n$ of typical noise vectors is specified. (For the moment, the term "typical" is used informally; the theorem below holds for any Jordan measurable bounded subset of $\mathcal{R}^n$.) We say that an *ambiguity* occurs if the received vector $\boldsymbol{y}$ can be written in more than one way as $\boldsymbol{y} = x + e$ with $x \in \Lambda$ and $e \in E$.

Note that an ambiguity occurs if and only if $\boldsymbol{e}$ can be written as $x+e$ with $x \in \Lambda^*$ and $e \in E$, i.e., if and only if $\boldsymbol{e} \in \Lambda^* + E$. In particular, the probability of an ambiguity does not depend on the transmitted codeword $\boldsymbol{x}$.

Let $P_{\mathrm{amb}\,|E}$ be the probability of an ambiguity, conditioned on the event that $\boldsymbol{e}$ is in E. If we assume that the receiver is able to recover $\boldsymbol{x}$ whenever $\boldsymbol{e} \in E$ and no ambiguity occurs, then the probability P_e of a transmission error (or failure) is upper-bounded by

$$P_e \leq P_{\mathrm{amb}\,|E} + P(\boldsymbol{e} \notin E). \tag{13}$$

The term $P(\boldsymbol{e} \notin E)$ is independent of the lattice Λ; for the term $P_{\mathrm{amb}\,|E}$, we have the following "random coding" theorem for scaled mod-p lattices.

Theorem 4: Let E be a Jordan measurable bounded subset of $\mathcal{R}^n$; let k be an integer such that $0 < k < n$. Then, for any $\delta > 0$, for all sufficiently small γ, and all sufficiently large primes p, the arithmetic average of $P_{\mathrm{amb}\,|E}$ over all lattices $\gamma\Lambda_C,\ C \in \mathcal{C}$, is bounded by

$$\overline{P_{\mathrm{amb}\,|E}} < (1+\delta)\, V(E)\,/\,V_f \tag{14}$$

where $\mathcal{C}$ is any balanced set of linear (n, k) codes over $\mathcal{Z}_p$ and where $V_f \overset{\Delta}{=} \gamma^n p^{n-k}$ is the fundamental volume of the lattices $\gamma\Lambda_C,\ C \in \mathcal{C}$.

Proof: (See also the analogous proof of Theorem 9 in the Appendix.) Let $f_{\mathbf{e}|E}$ be the probability density function of $\boldsymbol{e}$ conditioned on the event $\boldsymbol{e} \in E$. We first consider a fixed lattice Λ. For any $e \in E$, the event $\boldsymbol{e} = e$ is an ambiguity if and only if $\Lambda^* \cap (e - E) \neq \emptyset$. We thus have the bound

$$P_{\mathrm{amb}\,|E} \leq \int_E f_{\mathbf{e}|E}(v) \cdot |\Lambda^* \cap (v - E)|\, dv. \tag{15}$$

Averaging over all lattices $\gamma\Lambda_C$ and applying Theorem 2 yields

$$\begin{aligned} \overline{P_{amb|E}} &= \frac{1}{|\mathcal{C}|} \sum_{C \in \mathcal{C}} P_{\mathrm{amb}\,|E} && (16)\\ &\leq \int_E f_{\mathbf{e}|E}(v)\, \frac{1}{|\mathcal{C}|} \sum_{C \in \mathcal{C}} |\gamma\Lambda_C^* \cap (v - E)|\, dv && (17)\\ &\approx \int_E f_{\mathbf{e}|E}(v)\, V(E)\,/\,V_f\, dv && (18)\\ &= V(E)\,/\,V_f && (19) \end{aligned}$$

and the approximation (18) becomes exact in the limit $\gamma \to 0$, $p \to \infty$. (The sum in (17) is a Riemann sum by (10) and thus uniformly well approximated by $V(E)\,/\,V_f$.) □

Note that the lattices $\gamma\Lambda_C,\ C \in \mathcal{C}$, in Theorem 4 can be replaced by any other set of lattices for which the Minkowski–Hlawka theorem can be proved.

It is instructive to rewrite the bound (14) as

$$\overline{P_{\mathrm{amb}\,|E}} < (1+\delta)\, 2^{n[r+h(E)]} \tag{20}$$

where $h(E) \overset{\Delta}{=} 1/n \log_2 V(E)$ is the *(geometric) entropy rate* of E and where $r \overset{\Delta}{=} 1/n \log_2 (1/V_f)$ is the *information density rate* of a lattice with fundamental volume V_f.

Note that, if E is chosen as a set of typical noise vectors (where, this time, "typical" is meant in the formal sense of information theory, e.g., as defined in [20]) then $h(E)$ will, for $n \to \infty$, converge to the (information-theoretic) differential entropy rate $h(\boldsymbol{e})$ of the noise. (The density of the noise $\boldsymbol{e}$ must be sufficiently "nice" so that E is Jordan measurable; if E is not bounded, it can be made so by intersecting it by a large enough sphere.) Expression (20) then shows that arbitrarily reliable transmission is possible with lattices of information density rate r provided only that $r + h(\boldsymbol{e}) < 0$.

Conversely, it is obvious that, for reliable transmission, the fundamental volume V_f of a lattice cannot be smaller that the volume of a high-probability error set; for $n \to \infty$, this implies that the information density rate is upper-bounded by $-h(\boldsymbol{e})$. We summarize these observations as

Theorem 5: Assume that, for $n \to \infty$, the random additive error $\boldsymbol{e}$ has a sufficiently nice density (such that the set of typical errors is Jordan measurable). Then arbitrarily reliable transmission is possible with lattices of information density rate r if $r + h(\boldsymbol{e}) < 0$. Conversely, reliable transmission is not possible for $r + h(\boldsymbol{e}) > 0$.

To conclude this section, it seems worth pointing out that the probabilistic "Shannon packings" of this section are, in a certain sense, much tighter than the packings of the previous section: the asymptotic information density rate guaranteed by Theorem 3 is

$$\lim_{n \to \infty} (1/n)\, \log_2 (2/V(\Delta E)) = -h(\Delta E)$$

which is the best asymptotic existence bound known for nonoverlapping packings (i.e., for $P_{\mathrm{amb}\,|E} = 0$).

IV. Shaping

If we form the intersection $\Lambda \cap S$ of a lattice $\Lambda \subset \mathcal{R}^n$ with a shaping region $S \subset \mathcal{R}^n$, we would expect to obtain a code with about $V(S)\,/\,V_f(\Lambda)$ codewords. In fact, we know from Theorem 2 that the value $V(S)\,/\,V_f(\Lambda)$ is correct in the average over a suitable set of lattices. Combining this observation with Theorem 4, we would thus expect the existence of lattice codes $\Lambda \cap S$ with at least M codewords satisfying

$$P_{\mathrm{amb}\,|E} < (1+\delta) M\, V(E)\,/\,V(S). \tag{21}$$

Unfortunately, we were not able to prove the existence of such lattice codes. The problem here is that all those lattices with sufficiently low $P_{\mathrm{amb}\,|E}$ might have too few points in S while those lattices with sufficiently many points in S might have too large $P_{\mathrm{amb}\,|E}$. (This problem is skipped in both [4] and [6].)

In order to overcome this problem, we resort to translates of lattices, i.e., we consider codes of the form $(v+\Lambda)\cap S$ for some $v \in \mathcal{R}^n$. The existence of such codes satisfying (21) is easily established by means of the following lemma (cf., [12, ch. 2, Theorem 2]).

Lemma 2: Let $\Lambda \subset \mathcal{R}^n$ be a lattice; let S be a Jordan measurable bounded subset of $\mathcal{R}^n$. Then, there exists a $v \in \mathcal{R}^n$ such that the translate $v+\Lambda$ satisfies

$$|(v+\Lambda)\cap S| \geq V(S)\,/\,V_f(\Lambda). \tag{22}$$

Proof: Let R_f be a fundamental region of Λ (e.g., the Voronoi region of the origin); let μ: $\mathcal{R}^n \to \{0, 1\}$ be the indicator function of S (i.e., $\mu(v)=1$ if $v \in S$ and $\mu(v)=0$ otherwise). Averaging $|(v+\Lambda)\cap S|$ over all $v \in R_f$ yields

$$\begin{aligned} 1/V_f(\Lambda) \int_{R_f} & |(v+\Lambda)\cap S|\,dv \\ &= 1/V_f(\Lambda) \int_{R_f} \sum_{x\in\Lambda} \mu(v+x)\,dv && (23)\\ &= 1/V_f(\Lambda) \int_{\mathcal{R}^n} \mu(v)\,dv && (24)\\ &= V(S)\,/\,V_f(\Lambda) && (25)\end{aligned}$$

which implies that there must be at least one $v \in \mathcal{R}_f$ such that $|(v+\Lambda)\cap S| \geq V(S)\,/\,V_f(\Lambda)$. □

For any desired number of codewords M, we can thus choose $V_f = V(S)/M$, take any "good" lattice Λ with $V_f(\Lambda) = V_f$ from Theorem 4, and find a translate $v+\Lambda$ such that the code $(v+\Lambda)\cap S$ has at least M codewords. We have proved

Theorem 6: Assume that both the error set E and the shaping region S are Jordan measurable bounded subsets of $\mathcal{R}^n$, $n \geq 2$; let M be an arbitrary positive integer. Then, for any $\delta > 0$, there exists a lattice code $(v+\Lambda)\cap S$ with at least M codewords such that, for lattice decoding,

$$P_{\mathrm{amb}\,|E} < (1+\delta) M\, V(E)\,/\,V(S). \tag{26}$$

Moreover, Λ may be chosen to be a scaled version of a mod-p lattice Λ_C for some linear (n, k) code C over $\mathcal{Z}_p$, where k may be chosen freely between 1 and $n-1$ and p is a sufficiently large prime.

With the notation $h(E)$ and $h(S) \triangleq 1/n \log_2 V(S)$ as in Section III, the bound (26) can be rewritten as

$$P_{\mathrm{amb}\,|E} < (1+\delta)\, 2^{-n[h(S)-h(E)-R]} \tag{27}$$

where $R \triangleq 1/n \log_2 M$ is the information rate of the code in bits per dimension. As in the previous section, we can choose E to be a set of typical (in the formal sense) noise vectors and obtain

Theorem 7: Assume that, for $n \to \infty$, the random additive error $\boldsymbol{e}$ has a sufficiently nice density (such that the set of typical errors is Jordan measurable). Then arbitrarily reliable transmission is possible with lattice codes (of the form $(v+\Lambda)\cap S$) of rate at least R provided $R < h(S) - h(\boldsymbol{e})$.

For a convex shaping region S, we conjecture that the converse to Theorem 7 is also true, viz., that reliable transmission is not possible with such codes if $R > h(S) - h(\boldsymbol{e})$.

For the AWGN channel and spherical shaping, we have

$$h(\boldsymbol{e}) = 1/2 \log_2 (2\pi e N)$$

where N is the noise variance (per dimension), and

$$\lim_{n\to\infty} h(S) = 1/2 \log_2 (2\pi e P)$$

where P is the signal power per dimension. (The latter follows either from the formula for the volume of an n-dimensional sphere and a Sterling approximation or from noting that $h(S)$ asymptotically equals the differential entropy of a Gaussian random variable with variance P.) Theorem 7 thus guarantees that arbitrarily small (but positive) error probability is achievable with lattice codes and lattice decoding at any rate below $1/2 \log_2 (P/N)$.

V. Concluding Remarks

The two main themes of this paper were 1) an information-theoretic investigation of lattice codes used with lattice decoding and 2) the interpretation of various versions of the Minkowski–Hlawka theorem as (a limit of) familiar "random coding" (i.e., averaging) theorems for linear codes (cf., Appendix). It should be emphasized that our existence theorems for lattice codes hold for any class of lattices for which the Minkowski–Hlawka theorem can be proved. We have seen that scaled mod-p lattices for $p \to \infty$ (i.e., for "sufficiently large" p) are such a class; (shaped) codes from such lattices are actually rather (shaped) linear codes over $\mathcal{Z}_p$ used with p-level amplitude modulation. It is interesting that the dimension k of these linear codes could be chosen freely between 1 and $n-1$.

Appendix
Analogous Theorems for Linear Codes

For every theorem on lattices that was considered in this paper, there exists a corresponding theorem on linear codes which we give in this Appendix. The proofs are usually somewhat easier since all sets are finite. We give most of these theorems below without much comment; for an in-depth discussion we refer to [19].

The theorems of this Appendix are based on Lemma 1, precisely as the lattice theorems of this paper are based on the Minkowski–Hlawka theorem (Theorem 1). Just as Theorem 1 was obtained as a limit ($p \to \infty$, $\gamma \to 0$, $\gamma p \to \infty$) of Lemma 1, the other lattice theorems of this paper could have been obtained by such a limit from the theorems of this Appendix.

The analog of Theorem 2 is

Lemma 3 (Average Intersection Cardinality Lemma): Let $\mathcal{C}$ be a balanced set of linear (n, k) codes over F_q; let E be an arbitrary subset of F_q^n. Then the average cardinality of $C^* \cap E$ over all codes C in $\mathcal{C}$ is given by

$$\frac{1}{|\mathcal{C}|} \sum_{C \in \mathcal{C}} |C^* \cap E| = \frac{q^k - 1}{q^n - 1} |E^*|. \tag{28}$$

Proof: Define f: $F_q^n \rightarrow \{0, 1\}$ as $f(v) = 1$ if $v \in E$ and $f(v) = 0$ otherwise and apply Lemma 1. □

It is then easy to prove the following version of the Varshamov–Gilbert bound, which is the analog to Theorem 3.

Theorem 8 (Varshamov–Gilbert Bound): Let $\mathcal{C}$ be a balanced set of linear (n, k) codes over F_q and let E be an arbitrary nonempty subset of F_q^n. If

$$|\Delta E| \leq q^{n-k} \tag{29}$$

then there exists a code in $\mathcal{C}$ that corrects all errors in E.

If q is a power of an *odd* prime, then (29) can actually be replaced by the weaker condition $|\Delta E| \leq 2q^{n-k}$, which corresponds more closely to the condition $V(\Delta E) < 2V_f$ of Theorem 3.

As in Section III, we now let the additive error pattern $\boldsymbol{e}$ be a random variable. We consider a subset E of F_q^n of typical (meant informally) error patterns and define an ambiguity as the event that two (or more) error patterns in E are consistent with the received vector. We then have the following analog to Theorem 4.

Theorem 9 (Random Coding Bound): The arithmetic average, over all codes of a balanced set of q-ary linear (n, k) codes, of $P_{amb|E}$ is bounded by

$$\overline{P_{\text{amb}\,|E}} \leq q^{k-n}\, |E|. \tag{30}$$

Proof: The proof is analogous to (but simpler than) that of Theorem 4. For a fixed code C and any fixed $e \in E$, the event $\boldsymbol{e} = e$ is an ambiguity if and only if $C^* \cap (e - E) \neq \emptyset$. We thus have the bound

$$P_{\text{amb}\,|E} \leq \sum_{e \in E} P(\boldsymbol{e} = e | \boldsymbol{e} \in E) \cdot |C^* \cap (e - E)|. \tag{31}$$

Averaging over all codes $C \in \mathcal{C}$ and applying Lemma 3 yields

$$\begin{aligned} \overline{P_{\text{amb}\,|E}} &= \frac{1}{|\mathcal{C}|} \sum_{C \in \mathcal{C}} P_{\text{amb}\,|E} & (32)\\ &\leq \sum_{e \in E} P(\boldsymbol{e} = e | \boldsymbol{e} \in E)\, \frac{1}{|\mathcal{C}|} \sum_{C \in \mathcal{C}} |C^* \cap (e - E)| & (33)\\ &= \sum_{e \in E} P(\boldsymbol{e} = e | \boldsymbol{e} \in E)\, \frac{q^k - 1}{q^n - 1}\, |(e - E)^*| & (34)\\ &\leq q^{k-n}\, |E|. & (35) \end{aligned}$$

□

The analog to Lemma 2 is the following simple lemma, which, as the exception in this Appendix, does not depend on Lemma 1.

Lemma 4: Let $C \subseteq F_q^n$ be a code (not necessarily linear) with q^k codewords; let S be any subset of F_q^n. Then there exists $v \in F_q^n$ such that the translate $v + C$ satisfies

$$|(v + C) \cap S| \geq q^{k-n}|S|. \tag{36}$$

Proof: Let μ: $F_q^n \rightarrow \{0, 1\}$ be the indicator function for S. Averaging over all $v \in F_q^n$ gives

$$\begin{aligned} \frac{1}{q^n} \sum_{v \in F_q^n} |(v + C) \cap S| &= \frac{1}{q^n} \sum_{v \in F_q^n} \sum_{c \in C} \mu(v + c) & (37)\\ &= \frac{1}{q^n} \sum_{c \in C} \sum_{v \in F_q^n} \mu(v) & (38)\\ &= q^{k-n}|S| & (39) \end{aligned}$$

and the lemma follows. □

The analog to Theorem 6 is as follows.

Theorem 10: For an arbitrary error set $E \subseteq F_q^n$, an arbitrary shaping set $S \subseteq F_q^n$, and any integer k, $0 \leq k \leq n$, there exists an (n, k) linear code C over F_q and some $v \in F_q^n$ such that the shaped translate $(v + C) \cap S$ is a code with at least $M \triangleq q^{k-n}|S|$ codewords and satisfies

$$P_{\text{amb}\,|E} \leq M\, |E| \,/\, |S|. \tag{40}$$

The proof is immediate from Theorem 9 and Lemma 4.

Acknowledgment

The helpful comments by the reviewer and the discussions with G. D. Forney, Jr., R. Urbanke, and V. Tarokh are gratefully acknowledged.

References

[1] A. R. Calderbank and N. J. A. Sloane, "New trellis codes based on lattices and cosets," *IEEE Trans. Inform. Theory*, vol. IT-33, pp. 177–195, Mar. 1987.

[2] G. D. Forney, Jr., "Coset codes—Part I: Introduction and geometrical classification," *IEEE Trans. Inform. Theory*, vol. 34, pp. 1123–1151, Sept. 1988.

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

[4] 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.

[5] ——, "Some optimal codes have structure," *IEEE J. Select. Areas Commun.*, vol. 7, pp. 893–899, Aug. 1989.

[6] G. Poltyrev, "On coding without restrictions for the AWGN channel," *IEEE Trans. Inform. Theory*, vol. 40, pp. 409–417, Mar. 1994.

[7] C. E. Shannon, "Probability of error for optimal codes in a Gaussian channel," *Bell Syst. Tech. J.*, vol. 38, pp. 611–656, 1959.

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

[9] E. Hlawka, "Zur Geometrie der Zahlen," *Math. Z.*, vol. 49, pp. 285–312, 1944.

[10] J. W. S. Cassels, *An Introduction to the Geometry of Numbers.* New York: Springer-Verlag, 1959 (Grundlehren der mathematischen Wissenschaften, vol. 99).

[11] C. A. Rogers, *Packing and Covering.* Cambridge, U.K.: Cambridge Univ. Press, 1964.

[12] P. M. Gruber and C. G. Lekkerkerker, *Geometry of Numbers.* Amsterdam, The Netherlands: Elsevier, 1987.

[13] D. Hilbert, "Mathematische Probleme," *Arch. Math. Phys.*, vol. 1, pp. 44–63 and 213–237, 1901.

[14] J. A. Rush and N. J. A. Sloane, "An improvement to the Minkowski–Hlawka bound for packing superballs," *Mathematika*, vol. 34, pp. 8–18, 1987.

[15] J. A. Rush, “A lower bound on packing density,” *Invent. Math.*, vol. 98, pp. 499–509, 1989.

[16] J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Groups.* New York: Springer-Verlag, 1988 (Grundlehren der mathematischen Wissenschaften, vol. 290).

[17] M. A. Tsfasman and S. G. Vlăduţ, *Algebraic–Geometric Codes.* Norwell, MA: Kluwer, 1991.

[18] P. Delsarte and P. Piret, “Algebraic constructions of Shannon codes for regular channels,” *IEEE Trans. Inform. Theory,* vol. IT-28, pp. 593–599, July 1982.

[19] H.-A. Loeliger, “On the basic averaging arguments for linear codes,” in *Communications and Cryptography: Two Sides of One Tapestry,* (festschrift in honor of James L. Massey on the occasion of his 60th birthday), R. E. Blahut *et al.*, Eds. Norwell, MA: Kluwer, 1994, pp. 251–261.

[20] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* New York: Wiley, 1991.

[21] M. Magalhäes de Oliveira and G. Battail, “A capacity theorem for lattice codes on Gaussian channels,” in *Proc. SBT/IEEE Int. Telecommunications Symp.* (Rio de Janeiro, Brazil, Sept. 3–6, 1990).