# LATTICE QUANTIZATION ERROR FOR REDUNDANT REPRESENTATIONS

## SERGIY BORODACHOV AND YANG WANG

ABSTRACT. Redundant systems such as frames are often used to represent a signal for error correction, denoising and general robustness. In the digital domain quantization needs to be performed. Given the redundancy, the distribution of quantization errors can be rather complex. In this paper we study quantization error for a signal **X** in  $\mathbb{R}^d$  represented by a frame using a lattice quantizer. We completely characterize the asymptotic distribution of the quantization error as the cell size of the lattice goes to zero. We apply these results to get the necessary and sufficient conditions for the asymptotic form of the White Noise Hypothesis in the case of the pulse-code modulation scheme.

## 1. INTRODUCTION AND MAIN RESULTS

In processing, transmitting, analysing and storing signals analog-to-digital conversion is frequently performed using quantization. Ideas similar to quantization have been present in literature since the end of the nineteenth century. However, the fundamental role of quantization in modulation and analog to digital conversion was first recognized with the early development of the pulse-code modulation schemes in the 1940s. One of the first results on quantization have been obtained in the papers of Oliver, Pierce, and Shannon [17], Bennett [4] and Shannon [19]. Later, a vast amount of engineering and mathematical literature was devoted to this topic. A comprehensive review can be found for example in the paper [14].

The quantized signal is first decomposed using a suitable set of atoms (also called a "dictionary" or a "basis")

$$\mathbf{x} = \sum_{j} c_j \mathbf{u}_j.$$

The elements of the dictionary  $\{\mathbf{u}_j\}$  can be, for example, functions or vectors. In practical applications the dictionary is finite and often has redundancy ("extra" elements) which is

Key words and phrases. White Noise Hypothesis, finite frame, Voronoi cell, vector quantization, linear independence over the rationals, lattice quantization, asymptotically uniformly distributed, PCM.

The second author is partially supported by the National Science Foundation, grant DMS-0813750.

used for error correction, recovery from channel erasures, denoising and general robustness. We may without loss of generality assume that  $\mathbf{u}_j \in \mathbb{R}^d$ . The collection  $\{\mathbf{u}_j\}$  is often chosen to be a *frame*, i.e. the matrix  $F = [\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N]$  has rank d (here and throughout this paper the vectors  $\mathbf{u}_j$  are column vectors). Clearly  $N \ge d$ . When N > d the system is redundant, and as a result the coefficients  $\{c_j\}$  will not be unique. In practical applications, a dual set of vectors  $\{\mathbf{v}_j\}_{j=1}^N$  is chosen so that for any  $\mathbf{x} \in \mathbb{R}^d$  we have

(1.1) 
$$\mathbf{x} := \sum_{j=1}^{N} \langle \mathbf{x}, \mathbf{v}_j \rangle \, \mathbf{u}_j.$$

The vectors  $\{\mathbf{v}_j\}_{j=1}^N$  form the *dual frame* of  $\{\mathbf{u}_j\}_{j=1}^N$ . A standard dual frame of  $\{\mathbf{u}_j\}_{j=1}^N$  is given by the column vectors of the matrix  $G = (FF^T)^{-1}F$ , which is known as the *canonical dual frame*. Thus (1.1) becomes

(1.2) 
$$\mathbf{x} = F\mathbf{y}, \quad \text{where} \quad \mathbf{y} = G^T \mathbf{x}.$$

Next, the coefficient vector  $\mathbf{y} = G^T \mathbf{x}$  is replaced by a vector from some discrete set in  $\mathbb{R}^N$  called the set of *reproduction values* or *points* or *levels*. Quantizing the coefficient vector as a whole has an advantage over quantizing each channel separately, since different channels may be correlated.

In this paper we consider the behavior of the quantization noise (error) when the set of reproduction values is a full rank lattice L in  $\mathbb{R}^N$ . Under this setting, the vector  $\mathbf{y}$  in (1.2) is replaced (quantized) by an element in the lattice L. In general, we replace  $\mathbf{y}$  by the element  $q(\mathbf{y}, L)$  in L, which is the closest element in L to  $\mathbf{y}$  in the Euclidean distance (should there be ties we shall take the first of such elements in the lexicographical order). With quantization we obtain a reconstruction  $\hat{\mathbf{x}}$  of  $\mathbf{x}$  given by

(1.3) 
$$\hat{\mathbf{x}} = Fq(\mathbf{y}, L) = \sum_{j=1}^{N} a_j \mathbf{u}_j$$

where  $q(\mathbf{y}, L) = [a_1, ..., a_N]^T$ .

Now we consider the error from this quantization. Define

$$\tau(\mathbf{y}, L) := \mathbf{y} - q(\mathbf{y}, L).$$

Then the error from the quantization is

(1.4) 
$$\mathbf{x} - \hat{\mathbf{x}} = F(\mathbf{y} - q(\mathbf{y}, L)) = F\tau(\mathbf{y}, L)$$

It is a fact that  $\tau(\mathbf{y}, L)$  lies in a *Voronoi cell* of the lattice *L*. More precisely, for every point  $\mathbf{l} \in L$  let

$$V(\mathbf{l}) := \overline{\{\mathbf{y} \in \mathbb{R}^N : q(\mathbf{y}) = \mathbf{l}\}}.$$

Then  $\{V(\mathbf{l}) : \mathbf{l} \in L\}$  are the Voronoi cells for L, and they form a tiling of  $\mathbb{R}^N$ . The fact that L is a lattice implies that  $V(\mathbf{l}) = V(\mathbf{0}) + \mathbf{l}$ . The vector  $\tau(\mathbf{y}, L)$  is in the Voronoi cell  $V(\mathbf{0})$  of L.

We are mainly interested in studying the distribution of the  $\tau(\mathbf{Y}, L)$  where  $\mathbf{Y} = G^T \mathbf{X}$ . Here in our model the signal  $\mathbf{X}$  is assumed to be a random vector in  $\mathbb{R}^d$  with certain (perhaps unknown) absolutely continuous distribution. Once we know the distribution of  $\tau(\mathbf{Y}, L)$  the *Mean Square Error* (MSE) of quantization

(1.5) 
$$\mathbf{MSE}\left(\mathbf{X}, L\right) := \mathbf{E}(|\mathbf{X} - \mathbf{X}|^2)$$

can be estimated, see [15] and the references therein. One natural question is whether  $\tau(\mathbf{Y}, L) = \tau(G^T \mathbf{X}, L)$  is uniformly distributed in the Voronoi cell  $V(\mathbf{0})$  of L. In the important case  $L = \Delta \mathbb{Z}^N$ , where  $\Delta > 0$ , we have the well known Pulse-Code Modulation (PCM) quantization scheme. The corresponding Voronoi cell is simply  $[-\Delta/2, \Delta/2]^N$ . This quantization scheme has been widely studied in mathematical literature (see e.g. [14], [15] for references). For convenience the White Noise Hypothesis (WNH) is often assumed by engineers and mathematicians working in this area (see e.g. [2], [3], [13]). This hypothesis asserts that in the PCM quantization scheme the errors in each channel are independent and uniformly distributed random variables. The WNH is often called the *Bennett's White* Noise Assumption. In the fundamental paper [4] Bennett showed that for d = N = 1 the distribution of the quantization error of the scalar PCM scheme is asymptotically uniform as  $\Delta \rightarrow 0$ . However, despite its wide acceptance the WNH is mathematically false whenever N > d and thus for any redundant system. Even when N = d it holds only in very restrictive conditions, see [15]. For general lattices L of  $\mathbb{R}^N$  there are few results. In the case N = d (no redundancy) Zamir and Feder [21] showed that if **X** has Gaussian distribution in  $\mathbb{R}^d$  and the set of reproduction values is the lattice L in  $\mathbb{R}^d$  optimal for X with respect to the MSE, the error  $\tau(G^T \mathbf{X}, L)$  is uniformly distributed in the Voronoi cell  $V(\mathbf{0})$  of L. We are not aware of any such study for redundant systems (N > d).

For redundant systems (N > d) and PCM quantization a weaker form of the WNH has been studied. For  $L = \Delta \mathbb{Z}^N$ , let  $\mathbf{Z}_{\Delta} = \Delta^{-1} \tau(G^T \mathbf{X}, \Delta \mathbb{Z}^N)$ . This represents the normalized quantization error for the coefficients. This weaker form of the WNH states that as the cell size  $\Delta$  goes to zero, each component (channel) of  $\mathbf{Z}_{\Delta}$  becomes asymptotically uniformly distributed in [-1/2, 1/2] and together they become uncorrelated. This has turned out to be valid much more often. The fact that the components are asymptotically uncorrelated is often found in engineering literature without a rigorous proof (see for example [12] and discussion in [20]). The first rigorous proof of this fact has been given for N = 2 in [20]. Later, in [15] it was proved that if  $\mathbf{X}$  is an absolutely continuous random variable and G has N linearly independent columns, then  $\mathbf{Z}_{\Delta}$  converges in distribution to a random vector uniformly distributed in  $[-1/2, 1/2]^N$  (Note here we do not assume  $N \ge d$ .). It was shown further that if the columns of  $\mathbf{G}$  are pairwise linearly independent vectors over  $\mathbb{Q}$  then asymptotically the components of  $\mathbf{Z}_{\Delta}$  are uniformly distributed and pairwise uncorrelated. For other results related to lattice quantization and quantization using (asymptotically) optimal tesselations see [10], [11], [6], [7], [18], [1], [8], [9], [21], and references in [14].

This paper studies the stronger form of the asymptotic WNH for general lattice by characterizing the asymptotic distribution of the quantization error in the most general setting. More specifically, let L be a full rank lattice in  $\mathbb{R}^N$  and G be a  $d \times N$  matrix. Let **X** be an absolutely continuous random vector in  $\mathbb{R}^d$ . For  $\Delta > 0$ , let

(1.6) 
$$\mathbf{Z}_{\Delta} := \frac{1}{\Delta} \tau(G^T \mathbf{X}, \Delta L)$$

be the normalized quantization error. We are interested in the distribution of  $\mathbf{Z}_{\Delta}$  as  $\Delta$  becomes small. The main difficulty is that in the redundant setting N > d the columns of G are linearly dependent, so even for  $L = \mathbb{Z}^N$  the asymptotic distribution of  $\mathbf{Z}_{\Delta}$  is unknown. The purpose of this paper is to provide a complete characterization of it. We state our main theorems.

**Theorem 1.1.** Let  $L = A\mathbb{Z}^N$  be a full rank lattice in  $\mathbb{R}^N$  and G be a  $d \times N$  matrix. Let  $\mathbf{X}$  be an absolutely continuous random vector in  $\mathbb{R}^d$ . For  $\Delta > 0$ , let

$$\mathbf{Z}_{\Delta} = \frac{1}{\Delta} \tau(G^T \mathbf{X}, \Delta L).$$

Assume that the rows of the matrix  $A^{-1}G^T$  are linearly independent over  $\mathbb{Q}$ . Then  $\mathbb{Z}_{\Delta}$  is asymptotically uniformly distributed in the Voronoi cell  $V(\mathbf{0})$  of the lattice L as  $\Delta \rightarrow 0$ .

Here by asymptotically uniformly distributed we mean  $\mathbf{Z}_{\Delta}$  converges in distribution (as  $\Delta \rightarrow 0$ ) to a random vector that is uniformly distributed in  $V(\mathbf{0})$ .

Note that Theorem 1.1 is a much stronger result than the weaker form of the WNH for PCM quantization. In the PCM case with  $L = \mathbb{Z}^N$ , Theorem 1.1 shows that if the columns of G are linearly independent over  $\mathbb{Q}$  then  $\mathbb{Z}_{\Delta}$  is aymptotically uniformly distributed in  $[-\frac{1}{2}, \frac{1}{2}]^N$ . The weaker form of WNH applies only when channels are asymptotically uncorrelated. Our theorem here also allows us to estimate  $\mathbf{MSE}(\mathbf{X}, \Delta L)$  in the general setting much like the way it was done for the case  $L = \mathbb{Z}^N$  in [15].

The converse of Theorem 1.1 is also valid. It in fact follows from a stronger theorem concerning the distribution of  $\mathbf{Z}_{\Delta}$  in case the rows of  $A^{-1}G^T$  are not linearly independent over  $\mathbb{Q}$ . Assume that m is the maximal number of linearly independent rows of  $A^{-1}G^T$  over  $\mathbb{Q}$ . Now let

$$W_0 = \left\{ \mathbf{x} \in \mathbb{Q}^N : \ \mathbf{x}^T A^{-1} G^T = \mathbf{0} \right\}.$$

 $W_0$  is a nontrivial subspace in  $\mathbb{Q}^N$  of dimension N - m, whose closure  $\overline{W}_0$  in  $\mathbb{R}^N$  is a rational subspace of the same dimension. It is easy to see that  $\overline{W}_0^{\perp}$  is a rational subspace in  $\mathbb{R}^N$  of dimension m. Note that a rational subspace projected onto the torus  $\mathbb{T}^N$  is a compact manifold of the same dimension, and here the projection of  $\overline{W}_0^{\perp}$  onto  $\mathbb{T}^N$  is precisely the closure of the projection of the subspace  $A^{-1}G^T(\mathbb{R}^N)$  onto  $\mathbb{T}^N$ , see [16]. Set  $V(L,G) = A\overline{W}_0^{\perp}$  and

(1.7) 
$$\Lambda(L,G) = \overline{\{\mathbf{x} - q(\mathbf{x},L) : \mathbf{x} \in V(L,G)\}} = \overline{\{\tau(\mathbf{x},L) : \mathbf{x} \in V(L,G)\}}$$

Thus  $\Lambda(L,G)$  is the projection of V(L,G) onto the Voronoi cell  $V(\mathbf{0})$  of L.

**Theorem 1.2.** Let  $L = A\mathbb{Z}^N$  be a full rank lattice in  $\mathbb{R}^N$  and G be a  $d \times N$  matrix. Let  $\mathbf{X}$  be an absolutely continuous random vector in  $\mathbb{R}^d$ . For  $\Delta > 0$ , let

$$\mathbf{Z}_{\Delta} = \frac{1}{\Delta} \tau(G^T \mathbf{X}, \Delta L).$$

Assume that the maximal number of linearly independent rows of the matrix  $A^{-1}G^T$  over  $\mathbb{Q}$  is m. Then  $\mathbb{Z}_{\Delta}$  is asymptotically uniformly distributed in  $\Lambda(L,G)$  with respect to the m-dimensional Hausdorff measure  $\mathcal{H}_m$  as  $\Delta \rightarrow 0$ .

**Remark 1.** For every matrix G of size  $d \times N$ , one can always find a non-degenerate  $N \times N$  matrix A such that the collection of all rows of  $A^{-1}G^T$  is independent over the rationals. Indeed, let  $r = \operatorname{rank} G$ . Without loss of generality we can assume that first r columns of  $G^T$  are independent over the reals. There exists an  $r \times N$  matrix  $G_1$  of rank r, whose columns are independent over the rationals. One can take A to be a non-degenerate  $N \times N$  matrix, which transforms columns of  $G_1^T$  into first r columns of  $G^T$ . Then  $A^{-1}G^T$  will have rows independent over the rationals.

Thus, whatever matrix G is given, we can always find a full-rank lattice L in  $\mathbb{R}^N$  such that for every absolutely continuous random variable X in  $\mathbb{R}^d$ , the quantization error  $\overline{Z}_{\Delta}$ defined as in (1.6), will be asymptotically uniformly distributed as  $\Delta \to 0$  in the Voronoi cell  $V(\mathbf{0})$  of lattice L.

**Remark 2.** Lattice L used as a set of quantizers, is an infinite set. However, when the distribution of random variable X is compactly supported in  $\mathbf{R}^d$ , there are only finitely many points in L, whose Voronoi cells intersect with the support of the distribution of X. In this case, all other points of L will be used with probability zero, which will make the quantizer essentially finite.

We would like to thank David Jimenez for very helpful discussions.

## 2. Proof of Main Theorems

In this section we prove the main theorems by establishing a series of lemmas. Our first step is to prove Theorem 1.1 for the case of  $L = \mathbb{Z}^N$ . Not only this result will be used to prove the more general results, but it also serves to show the main ideas behind the proof of the main theorems.

Key to the proof of our main theorems is a theorem on uniform distribution. Recall that for any  $\mathbf{y} \in \mathbb{R}^n$ , we use  $\tau(\mathbf{y}, \mathbb{Z}^n)$  to denote the vector  $\mathbf{y} - q(\mathbf{y}, \mathbb{Z}^n)$ , where  $q(\mathbf{y}, \mathbb{Z}^n)$  denotes the element in  $\mathbb{Z}^n$  that is the closest to  $\mathbf{y}$ . In other words,  $\tau(\mathbf{y}, \mathbb{Z}^n)$  is the error when  $\mathbf{y}$  is rounded off to its nearest integer point.

We will need the following definition (see Cassels [5], page 61). Let  $\mathbf{z}_{\alpha}$  be a sequence of points in  $\left[-\frac{1}{2}, \frac{1}{2}\right]^k$  labeled with vectors  $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_d) \in \mathbb{Z}^d$ . For  $\mathbf{a}, \mathbf{b} \in \left[-\frac{1}{2}, \frac{1}{2}\right]^k$ , such that  $-\frac{1}{2} \leq a_i < b_i \leq \frac{1}{2}, i = 1, \ldots, k$ , let  $F_{n_1,\ldots,n_d}(\mathbf{a}, \mathbf{b})$  be the number of points  $\mathbf{z}_{\alpha}$ , such that  $1 \leq \alpha_1 \leq n_1, \ldots, 1 \leq \alpha_d \leq n_d$ , which lie in the parallelepiped  $[a_1, b_1] \times \ldots \times [a_k, b_k]$ . Denote

$$\mathcal{D}_{n_1,\dots,n_d} = \sup_{\mathbf{a},\mathbf{b}} \left| \frac{1}{n_1 \cdot \dots \cdot n_d} F_{n_1,\dots,n_d}(\mathbf{a},\mathbf{b}) - \prod_{j=1}^k (b_j - a_j) \right|.$$

We say that a sequence  $\{\mathbf{z}_{\alpha}\}_{\alpha \in \mathbb{Z}^d}$  is uniformly distributed in  $\left[-\frac{1}{2}, \frac{1}{2}\right]^k$  if

$$\lim_{n_1,\dots,n_d\to\infty}\mathcal{D}_{n_1,\dots,n_d}=0.$$

**Proposition 2.1.** Let C be an  $n \times d$  matrix. Assume that the only  $\mathbf{u} \in \mathbb{Z}^n$  such that  $\mathbf{u}^T C$  has integer entries is  $\mathbf{u} = \mathbf{0}$ . Then  $\{\tau(C\boldsymbol{\alpha},\mathbb{Z}^n) : \boldsymbol{\alpha} \in \mathbb{Z}^d\}$  is uniformly distributed in  $[-\frac{1}{2},\frac{1}{2}]^n$ .

**Proof.** See Cassels [5], Theorem I, page 64. It should be pointed out that in Theorem I it states that  $\{C\boldsymbol{\alpha} : \boldsymbol{\alpha} \in \mathbb{Z}^d \pmod{1}\}$  is uniformly distributed. But it clearly applies to  $\{\tau(C\boldsymbol{\alpha},\mathbb{Z}^n) : \boldsymbol{\alpha} \in \mathbb{Z}^d\}.$ 

Going back to  $\mathbf{Z}_{\Delta} = \frac{1}{\Delta} \tau(G^T \mathbf{X}, \Delta \mathbb{Z}^N)$ , where G is a  $d \times N$  matrix whose columns represent the dual frame, we further simplify the setting by assuming G has the form

$$(2.1) G^T = \begin{bmatrix} I_d \\ B \end{bmatrix}$$

where B is a  $(N-d) \times d$  matrix. We prove that if the rows of  $G^T$  are linearly independent over  $\mathbb{Q}$  then  $\mathbb{Z}_{\Delta}$  converges in distribution to the uniform distribution in  $[-\frac{1}{2}, \frac{1}{2}]^N$ . To do so it suffices to prove that for any cube  $\Omega = \Omega_1 \times \Omega_2$  in  $[-\frac{1}{2}, \frac{1}{2}]^N$  where  $\Omega_1 \subset [-\frac{1}{2}, \frac{1}{2}]^d$  and  $\Omega_2 \subset [-\frac{1}{2}, \frac{1}{2}]^{N-d}$  we have

(2.2) 
$$\lim_{\Delta \to 0} \operatorname{Prob}\left(\mathbf{Z}_{\Delta} \in \Omega\right) = \mu(\Omega).$$

Note that  $\mathbf{Z}_{\Delta} \in \Omega$  is equivalent to  $\frac{\mathbf{X}}{\Delta} \in E(\Omega)$ , or  $\mathbf{X} \in \Delta E(\Omega)$ , where

$$E(\Omega) := \left\{ \mathbf{x} \in \mathbb{R}^d : \mathbf{x} \in \Omega_1 + \mathbb{Z}^d \text{ and } B\mathbf{x} \in \Omega_2 + \mathbb{Z}^{N-d} \right\}.$$

Thus (2.2) is equivalent to

(2.3) 
$$\lim_{\Delta \to 0} \int_{\Delta E(\Omega)} f(\mathbf{x}) d\mathbf{x} = \mu(\Omega) = \mu(\Omega_1) \cdot \mu(\Omega_2),$$

where  $f \in L^1(\mathbb{R}^d)$  denotes the probability density function of **X** and  $\mu$  denotes the Lebesgue measure.

**Lemma 2.2.** Suppose that G has the form (2.1) and the rows of  $G^T$  are independent over  $\mathbb{Q}$ . Then  $\{\tau(B\boldsymbol{\alpha},\mathbb{Z}^{N-d}): \boldsymbol{\alpha}\in\mathbb{Z}^d\}$  are uniformly distributed in  $[-\frac{1}{2},\frac{1}{2}]^{N-d}$ .

**Proof.** By Proposition 2.1 we only need to show that if  $\mathbf{u} \in \mathbb{Z}^{N-d}$  is such that  $\mathbf{u}^T B$  has integer entries, then  $\mathbf{u} = \mathbf{0}$ . Let  $\mathbf{v}^T := \mathbf{u}^T B \in \mathbb{Z}^d$ . Then  $[-\mathbf{v}^T, \mathbf{u}^T]G^T = \mathbf{0}$ . But the rows of

 $G^T$  are linearly independent over  $\mathbb{Q}$ , and both  $\mathbf{u}$  and  $\mathbf{v}$  are integer vectors. It follows that  $[-\mathbf{v}^T, \mathbf{u}^T] = \mathbf{0}$ . Hence  $\mathbf{u} = \mathbf{0}$ . The lemma follows.

**Lemma 2.3.** Suppose that G has the form (2.1) and the rows of  $G^T$  are independent over  $\mathbb{Q}$ . Let  $g(\mathbf{x})$  be any indicator function of a cube in  $\mathbb{R}^d$ . Then for any cube  $\Omega = \Omega_1 \times \Omega_2$  in  $\mathbb{R}^d \times \mathbb{R}^{N-d}$  we have

$$\lim_{\Delta \to 0} \int_{\Delta E(\Omega)} g(\mathbf{x}) d\, \mathbf{x} = \mu(\Omega) \int g.$$

**Proof.** Note that  $\mathbf{x} \in \Delta E(\Omega)$  is equivalent to  $\frac{\mathbf{x}}{\Delta} \in E(\Omega)$ . As one will see from the proof, we may without loss of generality assume that  $g = \chi_{[-a,a]^d}$ , where a > 0. If the cube is not centered at the origin then we can make a simple shift without affecting the proof. Set  $g_{\Delta}(\mathbf{x}) = g(\Delta \mathbf{x}) = \chi_{J_{\Delta}}(\mathbf{x})$  where  $J_{\Delta} := [-\frac{a}{\Delta}, \frac{a}{\Delta}]^d$ . Then

(2.4) 
$$\int_{\Delta E(\Omega)} g(\mathbf{x}) d\mathbf{x} = \Delta^d \int_{E(\Omega)} g_{\Delta}(\mathbf{y}) d\mathbf{y} = \Delta^d \mu \big( J_{\Delta} \cap E(\Omega) \big).$$

Let  $\mathbf{z}_0$  be a point in  $\Omega_1$  and  $\varepsilon > 0$  be sufficiently small. We first assume that  $\operatorname{diam}(B(\Omega_1)) < \varepsilon$ . Define

 $U^{\varepsilon} = \{ \mathbf{x} \in \Omega_2 : \text{ dist}(\mathbf{x}, \partial \Omega_2) \ge \varepsilon \} \quad \text{ and } \quad T^{\varepsilon} = \left\{ \mathbf{x} \in \mathbb{R}^{N-d} : \text{ dist}(\mathbf{x}, \Omega_2) < \varepsilon \right\}.$ Let  $m_1 = m_1(\Delta, \varepsilon)$  and  $m_2 = m_2(\Delta, \varepsilon)$  be defined respectively by

$$m_1 = \# \left\{ \boldsymbol{\alpha} \in \mathbb{Z}^d : \boldsymbol{\alpha} \in J_{\Delta}^- \text{ and } \tau(B(\boldsymbol{\alpha} + \mathbf{z}_0), \mathbb{Z}^{N-d}) \in U^{\varepsilon} \right\},$$
  
$$m_2 = \# \left\{ \boldsymbol{\alpha} \in \mathbb{Z}^d : \boldsymbol{\alpha} \in J_{\Delta}^+ \text{ and } \tau(B(\boldsymbol{\alpha} + \mathbf{z}_0), \mathbb{Z}^{N-d}) \in T^{\varepsilon} \right\},$$

where  $J_{\Delta}^- := [-\frac{a}{\Delta} + 1, \frac{a}{\Delta} - 1]^d$  and  $J_{\Delta}^+ := [-\frac{a}{\Delta} - 1, \frac{a}{\Delta} + 1]^d$ . It is easy to see that

$$m_1 \mu(\Omega_1) \le \mu(J_\Delta \cap E(\Omega)) \le m_2 \mu(\Omega_1).$$

By Lemma 2.2 we have  $\lim_{\Delta \to 0} \frac{m_1}{(2\Delta^{-1}a)^d} = \mu(U^{\varepsilon})$  and  $\lim_{\Delta \to 0} \frac{m_2}{(2\Delta^{-1}a)^d} = \mu(T^{\varepsilon})$ . But there exists a C > 0 such that  $\mu(T^{\varepsilon}) < \mu(\Omega_2) + C\varepsilon$  and  $\mu(\Omega_2) < \mu(U^{\varepsilon}) + C\varepsilon$ . Thus for sufficiently small  $\Delta > 0$  we have

(2.5) 
$$(2a)^d (\mu(\Omega_2) - 2C\varepsilon) \mu(\Omega_1) \le \Delta^d \mu (J_\Delta \cap E(\Omega)) \le (2a)^d (\mu(\Omega_2) + 2C\varepsilon) \mu(\Omega_1).$$

In general we may partition  $\Omega_1$  as a disjoint union of cubes  $D_1, \ldots, D_k$  with diam $(D_j) < \varepsilon$ for each j. Then for each j and sufficiently small  $\Delta$  we have

$$(2a)^d (\mu(\Omega_2) - 2C\varepsilon) \mu(D_j) \le \Delta^d \mu (J_\Delta \cap E(D_j \times \Omega_2)) \le (2a)^d (\mu(\Omega_2) + 2C\varepsilon) \mu(D_j).$$

Summing up the above inequalities we obtain (2.5) for arbitrary  $\Omega_1$  whenever  $\Delta$  is sufficiently small. Thus

$$\lim_{\Delta \to 0} \Delta^d \mu \big( J_\Delta \cap E(\Omega) \big) = (2a)^d \mu(\Omega_1) \mu(\Omega_2) = \mu(\Omega) \int g.$$

The lemma follows from (2.4).

**Lemma 2.4.** Suppose that G has the form (2.1) and the rows of  $G^T$  are independent over  $\mathbb{Q}$ . Let **X** be an absolutely continuous random vector in  $\mathbb{R}^d$ . Then  $\mathbf{Z}_{\Delta} = \frac{1}{\Delta} \tau(G^T \mathbf{X}, \Delta \mathbb{Z}^N)$  is aymptotically uniformly distributed in  $[-\frac{1}{2}, \frac{1}{2}]^N$  as  $\Delta \rightarrow 0$ .

**Proof.** Let  $f(\mathbf{x})$  be the probability density function of  $\mathbf{X}$ . We prove (2.3) for any cube  $\Omega \subset [-\frac{1}{2}, \frac{1}{2}]^N$ . For any  $\varepsilon > 0$  we may approximate f by  $g_{\varepsilon} = \sum_{j=1}^k c_j g_j$  such that each  $g_j$  is an indicator function of a cube in  $\mathbb{R}^d$  and  $||f - g_{\varepsilon}||_{L^1} < \varepsilon$ . Now

$$\int_{\Delta E(\Omega)} g_{\varepsilon} - \mu(\Omega) \int g_{\varepsilon} = \sum_{j=1}^{k} c_j \Big( \int_{\Delta E(\Omega)} g_j - \mu(\Omega) \int g_j \Big).$$

It follows from Lemma 2.3 that for  $\Delta$  sufficiently small we have  $|\int_{\Delta E(\Omega)} g_{\varepsilon} - \mu(\Omega) \int g_{\varepsilon}| < \varepsilon$ . Since  $\int f = 1$  and  $||f - g_{\varepsilon}||_{L^1} < \varepsilon$  we have  $|\int g_{\varepsilon} - 1| < \varepsilon$ , which yields

$$\left|\int_{\Delta E(\Omega)} g_{\varepsilon} - \mu(\Omega)\right| < (1 + \mu(\Omega))\varepsilon \le 2\varepsilon.$$

Finally

$$\left|\int_{\Delta E(\Omega)} f - \mu(\Omega)\right| \le \left|\int_{\Delta E(\Omega)} g_{\varepsilon} - \mu(\Omega)\right| + \|f - g_{\varepsilon}\|_{L^{1}} < 3\varepsilon.$$

This establishes (2.3), which proves the lemma.

**Proof of Theorem 1.1.** Again we first consider the case  $L = \mathbb{Z}^N$ . Hence  $A = I_d$ . Assume that the matrix G has rank  $r \leq d$ . It follows that we can find a nonsingular matrix  $Q \in M_d(\mathbb{R})$  and a permutation matrix  $P \in M_N(\mathbb{R})$  such that

$$PG^TQ = \begin{bmatrix} I_r & 0\\ B & 0 \end{bmatrix}$$

where B is a  $(N - r) \times r$  matrix. Let  $G_1 = [I_r, B^T]$ , which is a  $r \times N$  matrix such that the rows of  $G_1^T$  are linearly independent over  $\mathbb{Q}$ . Let  $\tilde{\mathbf{X}}_r \in \mathbb{R}^r$  be the vector whose entries are the first r entries of the vector  $Q^{-1}\mathbf{X}$ . Note that  $Q^{-1}\mathbf{X}$  is absolutely continuous, and thus  $\tilde{\mathbf{X}}_r$  is also absolutely continuous. Let

$$\tilde{\mathbf{Z}}_{\Delta} := \frac{1}{\Delta} \tau(G_1^T \tilde{\mathbf{X}}_r, \Delta \mathbb{Z}^N).$$

By Lemma 2.4  $\tilde{\mathbf{Z}}_{\Delta}$  is asymptotically uniformly distributed in  $[-\frac{1}{2}, \frac{1}{2}]^N$ . Now P is a permutation matrix, and hence so is  $P^{-1}$ . Thus  $P^{-1}\tilde{\mathbf{Z}}_{\Delta}$  is asymptotically uniformly distributed in  $[-\frac{1}{2}, \frac{1}{2}]^N$ . Finally we can easily check that for  $L = \mathbb{Z}^N$  we in fact have  $\mathbf{Z}_{\Delta} = P^{-1}\tilde{\mathbf{Z}}_{\Delta}$ . This proves the theorem for  $L = \mathbb{Z}^N$ .

In the general case  $L = A\mathbb{Z}^N$  for some nonsingular  $A \in M_N(\mathbb{R})$ , to prove that  $\mathbf{Z}_{\Delta}$  is asymptotically uniformly distributed we only need to show that for any  $\Omega \subset V(\mathbf{0})$  with  $\operatorname{diam}(A^{-1}(\Omega)) < 1$  we have

$$\lim_{\Delta \to 0} \operatorname{Prob}\left(\mathbf{Z}_{\Delta} \in \Omega\right) = \frac{\mu(\Omega)}{\mu(V(\mathbf{0}))}$$

But  $\mathbf{Z}_{\Delta} \in \Omega$  is precisely  $G^T \mathbf{X} \in \Delta \Omega + \Delta A \mathbb{Z}^N$ , which is equivalent to

$$A^{-1}G^T \mathbf{X} \in \Delta A^{-1}(\Omega) + \Delta \mathbb{Z}^N.$$

This is the PCM case for the matrix  $A^{-1}G^T$ , and hence we have

$$\lim_{\Delta \to 0} \operatorname{Prob}\left(\mathbf{Z}_{\Delta} \in \Omega\right) = \lim_{\Delta \to 0} \operatorname{Prob}\left(A^{-1}G^{T}\mathbf{X} \in \Delta A^{-1}(\Omega) + \Delta \mathbb{Z}^{N}\right) = \mu\left(A^{-1}(\Omega)\right).$$
  
But  $\mu\left(A^{-1}(\Omega)\right) = \mu(\Omega)/|\det(A)| = \mu(\Omega)/\mu(V(\mathbf{0})).$  This proves the theorem.

**Lemma 2.5.** Let  $\mathcal{G}$  and  $\mathcal{F}$  be compact Abelian topological groups with probability Haar measures  $\nu_{\mathcal{G}}$  and  $\nu_{\mathcal{F}}$ , respectively. Let  $\varphi : \mathcal{G} \longrightarrow \mathcal{F}$  be a continuous homomorphism that is surjective and has finite kernel. Assume that  $\mathbf{Z}$  is a random variable on  $\mathcal{G}$  that is uniformly distributed with respect to  $\nu_{\mathcal{G}}$ . Then  $\varphi(\mathbf{Z})$  is uniformly distributed on  $\mathcal{F}$ .

**Proof.** For any  $E \subset \mathcal{F}$  define  $\nu(E) = \nu_{\mathcal{G}}(\varphi^{-1}(E))$ . It is standard that  $\nu(\cdot)$  defines a probability (Radon) measure on  $\mathcal{F}$ . For any  $a \in \mathcal{F}$  we have  $\varphi^{-1}(a+E) = \varphi^{-1}(E) + c$ , where  $c \in \mathcal{G}$  is any element with  $\varphi(c) = a$ . Thus

$$\nu(a+E) = \nu_{\mathcal{G}} \left( c + \varphi^{-1}(E) \right) = \nu_{\mathcal{G}} \left( \varphi^{-1}(E) \right) = \nu(E).$$

It follows that  $\nu$  is a probability Haar measure and hence  $\nu = \nu_{\mathcal{F}}$  by the uniqueness. Now,

$$\operatorname{Prob}\left(\varphi(\mathbf{Z})\in E\right) = \operatorname{Prob}\left(\mathbf{Z}\in\varphi^{-1}(E)\right) = \nu_{\mathcal{G}}\left(\varphi^{-1}(E)\right) = \nu_{\mathcal{F}}(E).$$

Thus  $\varphi(\mathbf{Z})$  is uniformly distributed on  $\mathcal{F}$ .

**Proof of Theorem 1.2.** Again, we first consider the PCM case  $L = \mathbb{Z}^N$ . Assume without loss of generality that the first m rows of  $G^T$  are linearly independent over  $\mathbb{Q}$  while the

remaining rows are rational combinations of the first m rows. Write  $G = [G_1, G_2]$ , where  $G_1$  consists of the first m columns of G. Let  $\mathbf{Y} = G_1^T \mathbf{X}$ . Then

(2.6) 
$$G^T \mathbf{X} = \begin{bmatrix} \mathbf{Y} \\ B \mathbf{Y} \end{bmatrix},$$

where B is an  $(N-m) \times m$  rational matrix. Let  $B = \frac{1}{K}B_0$  where  $K \in \mathbb{N}$  and  $B_0$  is an integer matrix. Now set  $\mathbf{W}_{\Delta} := \frac{1}{\Delta}\tau(\mathbf{Y}, \Delta K\mathbb{Z}^m)$ . By Theorem 1.1,  $\mathbf{W}_{\Delta}$  is asymptotically uniformly distributed in  $[-\frac{K}{2}, \frac{K}{2}]^m$ . We now identify  $[-\frac{K}{2}, \frac{K}{2})^m$  with the torus  $\mathbb{T}_K^m := \mathbb{R}^m/K\mathbb{Z}^m$ . Thus  $\mathbf{W}_{\Delta}$  is asymptotically uniformly distributed on  $\mathbb{T}_K^m$  with respect to the probability Haar measure on  $\mathbb{T}_K^m$ .

Next we consider the map  $\varphi : \mathbb{T}_K^m \longrightarrow \mathbb{T}_K^m \times \mathbb{T}^{N-m}$  given by

$$\varphi(\mathbf{x}) = \begin{bmatrix} \mathbf{x} \\ B\mathbf{x} \end{bmatrix},$$

where  $\mathbb{T} = \mathbb{T}_1 = \mathbb{R}/\mathbb{Z}$  is the standard torus. Since KB is an integer matrix,  $\varphi$  is well defined, and satisfies assumptions of Lemma 2.5. Thus  $\varphi(\mathbf{W}_{\Delta})$  is asymptotically uniformly distributed on  $\varphi(\mathbb{T}_K^m)$ . Next let  $\pi_K : \mathbb{T}_K \longrightarrow \mathbb{T}$  be the standard projection  $\pi_K(x) = x \pmod{1}$ . We construct another map  $\phi : \mathbb{T}_K^m \times \mathbb{T}^{N-m} \longrightarrow \mathbb{T}^N$  by the standard projection of the first mentries into  $\mathbb{T}$  and leaving the remaining N - m entries the same. Then  $\phi \circ \varphi$  also satisfies the assumptions of Lemma 2.5. Thus,  $\phi \circ \varphi(\mathbf{W}_{\Delta})$  is asymptotically uniformly distributed on  $\mathcal{G} := \phi \circ \varphi(\mathbb{T}_K^m)$  with respect to the probability Haar measure on  $\mathcal{G}$ .

Now observe that by identifying  $[-\frac{1}{2}, \frac{1}{2})^N$  with  $\mathbb{T}^N$ , the random vector  $\mathbf{Z}_\Delta$  projected onto  $\mathbb{T}^N$  is precisely  $\phi \circ \varphi(\mathbf{W}_\Delta)$ . Thus  $\mathbf{Z}_\Delta$  projected onto  $\mathbb{T}^N$  is asymptotically uniformly distributed on  $\mathcal{G}$  with respect to the probability Haar measure on  $\mathcal{G}$ . It is easy to see that  $\mathcal{G}$  is the projection of the subspace  $V(\mathbb{Z}^N, G) = \left\{ \begin{bmatrix} \mathbf{y} \\ B\mathbf{y} \end{bmatrix} : \mathbf{y} \in \mathbb{R}^m \right\}$  onto the torus  $\mathbb{T}^N$ . The standard projection from  $\mathbb{R}^N$  to  $\mathbb{T}^N$  maps bijectively the set  $\Lambda(\mathbb{Z}^N, G) \cap [-\frac{1}{2}, \frac{1}{2})^N$  onto  $\mathcal{G}$ . Furthermore  $\Lambda(\mathbb{Z}^N, G)$  is locally an *m*-dimensional hyperplane, and all these hyperplanes have the same normal vectors. Therefore the probability Haar measure on  $\mathcal{G}$  lifted onto  $\Lambda(\mathbb{Z}^N, G)$  is precisely the normalized Hausdorff measure  $\mathcal{H}_m$ . This proves Theorem 1.2 for the PCM case  $L = \mathbb{Z}^N$ .

The proof for the general lattice  $L = A\mathbb{Z}^N$  case from the PCM case is exactly the same as that in the proof of Theorem 1.1. We omit it here.

**Example.** It is not always easy to check whether a set of vectors are linearly dependent over  $\mathbb{Q}$ . Here we examine an important class of tight frames called the *harmonic frames*. Let  $N \geq d$ . The harmonic frame  $H_{d,N}$  in  $\mathbb{R}^d$  is defined as follows: For d = 2d' we have  $H_{d,N} = \{\mathbf{w}_j\}_{j=0}^{N-1}$  where

$$\mathbf{w}_j = \sqrt{\frac{2}{d}} \left[ \cos \frac{2\pi j}{N}, \sin \frac{2\pi j}{N}, \cos \frac{4\pi j}{N}, \sin \frac{4\pi j}{N}, \dots, \cos \frac{2d'\pi j}{N}, \sin \frac{2d'\pi j}{N} \right]^T.$$

For d = 2d' + 1 we have  $H_{d,N} = \{\mathbf{w}_j\}_{j=0}^{N-1}$  where

$$\mathbf{w}_j = \sqrt{\frac{2}{d}} \Big[ \frac{1}{\sqrt{2}}, \cos\frac{2\pi j}{N}, \sin\frac{2\pi j}{N}, \dots, \cos\frac{2d'\pi j}{N}, \sin\frac{2d'\pi j}{N} \Big]^T$$

It is well known that  $\mathcal{H}_{d,N}$  is a unit norm tight frame with frame bound  $\lambda = \frac{N}{d}$ . Let  $F = [\mathbf{w}_0, \mathbf{w}_1, \dots, \mathbf{w}_{N-1}]$  be the corresponding frame matrix. The canonical dual frame matrix is  $G = \frac{d}{N}F$ . Using the notation of Theorem 1.2 we now consider the case  $L = \mathbb{Z}^N$ . Thus A = I. We claim: For even d the columns of G are never linearly independent over  $\mathbb{Q}$ . For odd d the columns of G are linearly independent over  $\mathbb{Q}$  if and only if N has no proper divisor greater than d/2.

To see this, for even d it is easy to check that  $\sum_{j=0}^{N-1} \mathbf{w}_j = \mathbf{0}$ . Therefore they are linearly dependent over  $\mathbb{Q}$ . For odd d = 2d' + 1, assume that the columns of G are linearly dependent over  $\mathbb{Q}$ . Then there exist  $a_0, a_1, \ldots, a_{N-1} \in \mathbb{Q}$  such that  $\sum_{j=0}^{N-1} a_j \mathbf{w}_j = \mathbf{0}$ . Let  $f(z) = \sum_{j=0}^{N-1} a_j z^j$ . This is equivalent to  $f(\omega_N^k) = 0$  for  $0 \le k \le d'$ , where  $\omega_N$  is a primitive N-th root of unity. Of the roots  $\omega_N^k$  for  $0 \le k \le d'$  the algebraic conjugating classes are represented by 1 and  $\{\omega_N^r : r | N, r \le d'\}$ . Let  $\Phi_n(z)$  denote the cyclotomic polynomial of order n. Then z - 1|f(z) and  $\Phi_{N/r}(z)|f(z)$ , where  $r|N, r \le d'$ . It follows that

(2.7) 
$$\deg(f) \ge 1 + \sum_{r \mid N, r \le d'} \deg(\Phi_{N/r}) = 1 + \sum_{r \mid N, r \le d'} \phi\left(\frac{N}{r}\right),$$

where  $\phi$  is the Euler function. However, it is well known that  $\sum_{r|N} \phi\left(\frac{N}{r}\right) = N$ . If N has no proper divisor greater than d/2 then

$$\deg(f) \ge 1 + \sum_{r \mid N, r \le d'} \phi\left(\frac{N}{r}\right) = \sum_{r \mid N} \phi\left(\frac{N}{r}\right) = N.$$

But  $\deg(f) \leq N - 1$ . This is a contradiction. So in this case the columns of G are independent over  $\mathbb{Q}$ . Conversely, if N does have a proper divisor greater than d/2 then

$$1 + \sum_{r \mid N, r \le d'} \deg(\Phi_{N/r}) \le N - 1.$$

Thus by taking f(z) as the product of z-1 and  $\Phi_{N/r}(z)$ , r|N,  $r \leq d'$ , we have  $\deg(f) \leq N-1$ . For every  $1 \leq k \leq d'$ , if c is the greatest common divisor of k and N, we have  $\omega_N^k = (\omega_N^c)^{k_1}$ , where  $k_1$  is coprime with N/c. Hence,  $\Phi_{N/c}(\omega_N^k) = 0$ , c|N, and  $c \leq d'$ . Since f(1) = 0, we have  $f(\omega_N^k) = 0$ ,  $0 \leq k \leq d'$ , which yields the linear dependence of the columns of G over  $\mathbb{Q}$ .

The above argument can in fact be used to obtain m, the maximal number of linearly independent columns of G over  $\mathbb{Q}$ . It is given by  $m = \sum_{r|N,r \leq d'} \phi\left(\frac{N}{r}\right) - 1$  for d = 2d', and  $m = \sum_{r|N,r \leq d'} \phi\left(\frac{N}{r}\right)$  for d = 2d' + 1. We'll omit the proof here.

#### References

- E.S. Barnes, N.J.A. Sloane, The optimal lattice quantizer in three dimensions, SIAM J. Alg. Discr. Methods 4 (Mar. 1983), 30–41.
- [2] J. Benedetto, A. M. Powell, Ö. Yilmaz, Sigma-Delta ΣΔ quantization of finite frames, *IEEE Trans. Inform. Theory* 52 (2006), no. 5, 1990–2005.
- [3] J. Benedetto, A. M. Powell, Ö. Yilmaz, Second order sigma-delta ΣΔ quantization of finite frame expansions, Appl. Comput. Harmon. Anal. 20 (2006), no. 1, 126–148.
- [4] W.R. Bennett, Spectra of quantized signals, Bell Syst. Tech. J. 27 (July 1948), 446–472.
- [5] J.W.S. Cassels, An Introduction to diophantine approximation, Cambridge Tracts in Mathematics and Mathematical Physics, Hafner Publishing Company, New York, 1972.
- J.H. Conway, N.J.A. Sloane, Voronoi regions of lattices, second moments of polytopes, and quantization, IEEE Trans. Inform. Theory IT-28 (Mar. 1982), 211–226.
- [7] J.H. Conway, N.J.A. Sloane, Fast quantizing and decoding algorithms for lattice quantizers and codes, *IEEE Trans. Inform. Theory* IT-28 (Mar. 1982), 227–232.
- [8] J.H. Conway, N.J.A. Sloane, A fast encoding method for lattice codes and quantizers, *IEEE Trans. Inform. Theory* IT-29 (Nov. 1983), 820–824.
- [9] J.H. Conway, N.J.A. Sloane, Sphere packings, lattices and groups, New York: Springer-Verlag, 1988.
- [10] L. Fejes-Tóth, Sur la representation d'une population infinie par un nombre fini d'elements, Acta Math. Acad. Sci. Hung. 10 (1959), 76–81.
- [11] A. Gersho, Asymptotically optimal block quantization, *IEEE Trans. Inform. Theory* IT-25 (July 1979), 373–380.
- [12] A. Gersho, R.M. Gray, Vector quantization and signal compression, Kluwer, Boston, 1992.
- [13] V.K. Goyal, Vetterli, N.T. Thao, Quantized overcomplete expansions in ℝ<sup>N</sup>: analysis, synthesis, and algorithms, *IEEE Trans. Inform. Theory* 44 (1998), 16–31.
- [14] R.M. Gray, D.L. Neuhoff, Quantization, IEEE Trans. Inform. Theory, 44 (1998), 2325–2383.
- [15] D. Jimenez, L. Wang, Y. Wang, White noise hypothesis for uniform quantization errors, SIAM J. Math. Anal. 38 (2007), no. 6, 2042–2056.
- [16] J. Lagarias and Y. Wang, Orthogonality criteria for compactly supported refinable functions and refinable function vectors, J. Fourier Anal. and Appl. 6, Issue 2 (2000), 173-190.
- [17] B.M. Oliver, J. Pierce, C.E. Shannon, The philosophy of PCM, Proc. IRE 36 (Nov. 1948), 1324–1331.
- [18] D.J. Newmann, The hexagon theorem, Bell Lab. Tech. Memo, 1964, published in the special issue on quantization of the IEEE Trans. Inform. Theory IT-82 (Mar. 1982), 137–139.
- [19] C.E. Shannon, A mathematical theory of communication, Bell Syst. Tech. J. 27 (1948), 379–423, 623– 656.
- [20] H. Viswanathan, R. Zamir, On the whiteness of the high-resolution quantization errors, *IEEE Trans. Inform. Theory* 47 (2001), 2029–2038.

[21] R. Zamir, M. Feder, On lattice quantization noise, IEEE Trans. Inform. Theory 42 (July 1996), 1152– 1159.

DEPARTMENT OF MATHEMATICS, TOWSON UNIVERSITY, BALTIMORE, MD 21252, USA *E-mail address*: borodachov@towson.edu

DEPARTMENT OF MATHEMATICS, MICHIGAN STATE UNIVERSITY, EAST LANSING, MI 48824-1027, USA. *E-mail address:* ywang@math.msu.edu