RESULTS ON LATTICE VECTOR QUANTIZATION WITH DITHERING[‡]

Ahmet Kirac, Student Member, IEEE, and P. P. Vaidyanathan, Fellow, IEEE
Dept. of Electrical Engineering, 116-81
California Institute of Technology
Pasadena, CA 91125.
Contact author: P. P. Vaidyanathan, ppvnath@sys.caltech.edu

IEEE Transactions on Circuits and Systems, II: Analog and Digital Signal Processing, Vol CAS-43, No 12, December 1996

Paper Category: Digital Signal Processing

Abstract. The statistical properties of the error in uniform scalar quantization have been analyzed by a number of authors in the past, and is a well understood topic today. The analysis has also been extended to the case of dithered quantizers, and the advantages and limitations of dithering have been studied and well documented in the literature. Lattice vector quantization is a natural extension into multiple dimensions of the uniform scalar quantization. Accordingly, there is a natural extension of the analysis of the quantization error. It is the purpose of this paper to present this extension and to elaborate on some of the new aspects that come with multiple dimensions.

We show that, analogous to the one dimensional case, the quantization error vector can be rendered independent of the input in subtractive vector-dithering. In this case, the total mean square error is a function of only the underlying lattice and there are lattices that minimize this error. We give a necessary condition on such lattices. In nonsubtractive vector dithering, we show how to render moments of the error vector independent of the input by using appropriate dither random vectors.

These results can readily be applied for the case of wide sense stationary (WSS) vector random processes, by use of iid dither sequences. We consider the problem of pre- and post filtering around a dithered lattice quantizer, and show how these filters should be designed in order to minimize the overall quantization error in the mean square sense. For the special case where the WSS vector process is obtained by blocking a WSS scalar process, the optimum prefilter matrix reduces to the blocked version of the well-known scalar half-whitening filter.

‡ Work supported in parts by NSF grant MIP 92-15785, Tektronix, Inc., and Rockwell International.

1. INTRODUCTION

Lattice vector quantizers have recently become attractive because they are simple to implement and in most cases, they constitute good alternatives to the computationally more complex vector quantization algorithms like the LBG and ECVQ [1], [2]. The geometric regularity of lattices allow very fast quantization algorithms, and there are already efficient algorithms for several well-known lattice structures [3] -[7].

Dithering was first applied by Roberts [8] to image coding. It was seen that by adding an independent random variable called dither before the quantization and subtracting after it, the perceptual quality of the image improves substantially. After that pioneering idea, there has been considerable work on the theory and applications of dithering. Dithered quantizers were theoretically analyzed by Schuchman [9] using the so called characteristic function method which uses the Fourier transform of the input probability density function (pdf). An analysis of the undithered uniform quantization was provided by Sripad and Snyder [10], using a similar style. More recently, Lipshitz et al. [11] published an excellent survey on quantization and dither. Gray and Stockham [12] gave new insightful proofs for the cases of subtractive and nonsubtractive dithering.

In this paper we use the idea of dithering in lattice quantization. The idea has already been introduced by Ziv [13] as a means of universal quantization. Interesting results on the rate distortion efficiency of dithered lattice quantizers have already been obtained by Zamir and Feder [14]-[17], and by Linder and Zeger [18]. In this paper our major concern is the analysis of the lattice quantization error for dithered and undithered cases. The only overlap between our work and the literature that we are aware of is Theorem 5. This was also reported by Zamir and Feder as a small part of their recent paper [16]. Even in our work, this result, independently found by us, is only a minor ingredient.

In Sec. 2, we review some preliminaries and definitions pertaining to lattice quantization. In Sec. 3, we provide exact analysis of the lattice quantization system. This can be regarded as a multidimensional extension of the work in [10]. The main tool, accordingly, is again Fourier series, but this time multidimensional. Since lattices are uniform structures, there is inherent periodicity in the error statistics, which motivates the use of multidimensional Fourier series. However, unlike in the one dimensional case, the choice of lattice is no longer unique and there exist optimum lattices in the sense that they minimize the familiar dimensionless second moment [19]. After giving the exact relationships between input and error probability densities, we consider dithered, or so called randomized lattice quantization schemes. As in one dimensional case [11], we investigate the possibility of rendering error statistics independent from the input. Sec. 4 covers subtractive dithering where an appropriate random vector is added before the quantizer and subtracted after it. In Sec. 5 nonsubtractive dithering is examined. Sec. 6 is devoted to finding optimum linear time invariant pre- and postfilters to be used in conjunction with dithered lattice quantizers.

Demonstration of the perceptive advantages of vector-dithering in image coding: For motivational purposes, we show in Fig. 1 a demonstration of the improvement of perceptual quality in image compression achieved by the use of vector-dithered lattice quantization. Fig. 1(a) shows the original 8 bit/pixel, 512×512 image of Lenna. Fig. 1(b) shows the output of a lattice quantizer with dimension 24. Vectors are formed by taking 4×6 blocks. The bit rate is about 0.4 bits/pixel. Fig. 1(c) shows the quantization error. Fig. 1(d) shows the output of the same lattice quantizer but with subtractive dithering. The bit rate is about the same. The corresponding quantization error is shown in Fig. 1(e). It is clear that the lattice quantization error in Fig. 1(c) is highly correlated with the input while the dithered quantization error in Fig. 1(e) seems completely uncorrelated with the input and uniform. The output of dithered lattice quantizer is perceptually more pleasant than that of undithered one.

Summary of the main results of the paper

1. In Sec. 3 we provide the necessary and sufficient condition for the quantization error of an undithered lattice quantizer to be uniform in its quantization basic cell. This is the so-called Nyquist-$\mathbf{V}$ condition, where $\mathbf{V}$ is the lattice generator matrix. We provide examples and general classes of random vectors that satisfy this condition (Sec. 3.1). We then examine the error statistics when the input is arbitrary (Sec. 3.2).
2. We next consider subtractive vector dithering, and establish the necessary and sufficient condition for the quantization error to be statistically independent of the input, and be uniform in the quantization basic cell. A comparison of the dimensionless second moment of lattice quantizers [19] is then given. A necessary condition for a lattice quantizer to have minimum dimensionless second moment (among all lattice quantizers of the same dimension) is established (Theorem 5).
3. For nonsubtractive vector dithering, first and second order moments of the quantization error conditioned on the input vector are derived (Sec. 5). Necessary and sufficient conditions for these moments to be independent of the input are provided. Examples of nonsubtractive dither vectors satisfying the moment independence conditions are given, and the dither that produces the minimum error for a given lattice is distinguished (Theorem 7).
4. In Sec. 6 we consider the use of a linear prefilter prior to the lattice quantization of a wide sense stationary (WSS) vector random process $\mathbf{x}(n)$. Under the assumption that the lattice quantizer satisfies certain mild conditions, we will derive an expression for the best choice of prefilter, as a function of the power spectral density matrix of the input process. We will also clarify the similarity and differences between this problem and the problem of designing optimal biorthogonal subband coders.

2. PRELIMINARIES AND DEFINITIONS

Let R^D and Z^D denote the D-dimensional Euclidean space of real numbers and the D-dimensional space of integers respectively. Let $\mathbf{V} = [\,\mathbf{v}_1 \quad \mathbf{v}_2 \quad \ldots \quad \mathbf{v}_D\,]$ be a nondegenerate lattice base in R^D. The lattice is the set of vectors defined as

$$\mathcal{L}(\mathbf{V}) = \{\mathbf{x} : \mathbf{x} = \mathbf{V}\mathbf{n}, \quad \mathbf{n} \in Z^D\} \tag{2.1}$$

Fig. 2 shows an example of a lattice in 2 dimensions. In lattice quantization, the codewords are the lattice points. The partition of the space for decision regions can be done in many ways. This partitioning can be uniform, i.e. each codeword may have the same quantization cell called a *basic cell* (defined below) [20]. From the necessary conditions for distortion-minimal quantizers [1], the quantization cell should be the so called Voronoi region [21] which is defined below. The resulting uniform partition is also known as the nearest-neighbor partition. Note that, from the same necessary conditions, codewords should be the centroids of the quantization cells with respect to the given distortion measure and the input probability density function. However, as in the uniform scalar quantization, one chooses lattice points as reproduction points avoiding the knowledge of probability density function.

If overflow is avoided at all times, then we have a periodic structure for the quantization error and the tools of the following analysis are applicable. In this paper, overflow is always assumed to be avoided. If one uses entropy coding [22] after the quantization, or if the given density has finite support, the resulting bit rate will be finite and by scaling the lattice one can tradeoff bit rate against distortion.

Definition 1. A *basic cell* of a lattice $\mathcal{L}(\mathbf{V})$: Let $\mathcal{P}$ be a region in R^D such that any $\mathbf{x} \in R^D$ can be written as $\mathbf{x} = \mathbf{x}_0 + \mathbf{V}\mathbf{n}$ for a unique $\mathbf{x}_0 \in \mathcal{P}$ and $\mathbf{n} \in Z^D$. Then $\mathcal{P}$ is called a basic cell of the lattice $\mathcal{L}(\mathbf{V})$. It is also said to generate a tiling of R^D with respect to $\mathbf{V}$. ♢

This definition does not imply that a basic cell is convex. In fact, one can partition a convex basic cell into subregions, and then translate each of these subregions by some distinct lattice vectors. The resulting nonconvex region is another basic cell.

Definition 2. The *Voronoi region* of a lattice point $\mathbf{x}_0 \in \mathcal{L}(\mathbf{V})$ is the set of points that are nearer (with respect to Euclidean distance) to that point than to any other lattice point. That is,

$$VOR(\mathbf{x}_0) = \{\mathbf{x} : \|\mathbf{x} - \mathbf{x}_0\| \leq \|\mathbf{x} - \mathbf{V}\mathbf{n}\|, \quad \forall \mathbf{n} \in Z^D\} \tag{2.2}$$

♢

The Voronoi region of the lattice point $\mathbf{0}$, $VOR(\mathbf{0})$, will be denoted by $VOR(\mathbf{V})$ for convenience. Fig. 3(a) shows the $VOR(\mathbf{V})$ of the lattice given in Fig. 2.

The Euclidean distance, used in the definition, leads to the mean square error as a distortion measure. In this paper, our interest will be only in the mean square error.

Definition 3. The *Symmetric Parallelpiped* of a lattice point $\mathbf{x}_0 \in \mathcal{L}(\mathbf{V})$ is defined as [23]:

$$SPD(\mathbf{x}_0) = \{\mathbf{x} : \mathbf{x} = \mathbf{x}_0 + \mathbf{V}\mathbf{u}, \quad \forall \mathbf{u} \in [-\frac{1}{2}, \frac{1}{2})^D\} \tag{2.3}$$

$\diamondsuit$

We will denote the Symmetric Parallelpiped region of the lattice point $\mathbf{0}$, $SPD(\mathbf{0})$, by $SPD(\mathbf{V})$. Fig. 3(b) shows the $SPD(\mathbf{V})$ of the lattice given in Fig. 2.

It can be verified that both $VOR(\mathbf{V})$ and $SPD(\mathbf{V})$ are basic cells of the lattice $\mathcal{L}(\mathbf{V})$ as long as some modifications are done to the boundary points in order to satisfy the uniqueness requirement in the definition of a basic cell. Furthermore, they are symmetric with respect to the origin.

Definition 4. A *lattice quantizer* $Q(\mathcal{P}_0, \mathbf{V})$ with the lattice $\mathcal{L}(\mathbf{V})$ and the quantization basic cell $\mathcal{P}_0$ is a nonlinear mapping from R^D to $\mathcal{L}(\mathbf{V})$ as given by the relation:

$$Q(\mathbf{x}) = \mathbf{V}\mathbf{n} \tag{2.4}$$

where $\mathbf{n}$ is the unique vector satisfying

$$\mathbf{x} = \mathbf{x}_0 + \mathbf{V}\mathbf{n}, \quad \mathbf{x}_0 \in \mathcal{P}_0. \tag{2.5}$$

We will denote the quantizer with $\mathcal{P}_0 = VOR(\mathbf{V})$ by $Q(VOR, \mathbf{V})$. $\diamondsuit$

Definition 5. The *second moment matrix* of a basic cell $\mathcal{P}$ of a lattice $\mathcal{L}(\mathbf{V})$, denoted by $\mathbf{G}_D(\mathcal{P}, \mathbf{V})$, is the second moment of a random vector that is uniformly distributed in $\mathcal{P}$. That is,

$$\mathbf{G}_D(\mathcal{P}, \mathbf{V}) = \frac{1}{|det\mathbf{V}|} \int_{\mathcal{P}} \mathbf{e}\mathbf{e}^T d\mathbf{e}. \tag{2.6}$$

If $\mathcal{P} = VOR(\mathbf{V})$, we will denote the second moment matrix by $\mathbf{G}_D(VOR, \mathbf{V})$. $\diamondsuit$

Definition 6. The characteristic function of a random vector with pdf $f_{\mathbf{X}}(\mathbf{x})$ is defined to be:

$$\Phi_{\mathbf{X}}(\mathbf{\Omega}) = E[e^{j\mathbf{\Omega}^T\mathbf{x}}] = \int f_{\mathbf{X}}(\mathbf{x}) e^{j\mathbf{\Omega}^T\mathbf{x}} d\mathbf{x}. \tag{2.7}$$

$\diamondsuit$

Next we will define a Nyquist-$\mathbf{V}$ vector. We say a function $f(\mathbf{\Omega})$ is Nyquist-$\mathbf{A}$ if $f(\mathbf{A}\mathbf{n}) = c\delta(\mathbf{n}), \forall \mathbf{n} \in Z^D$, where c is a constant and $\delta(\mathbf{n})$ is dirac delta function, which is 1 when $\mathbf{n} = \mathbf{0}$, and 0 otherwise.

Definition 7. A *Nyquist-$\mathbf{V}$ random vector* is a vector whose characteristic function is Nyquist-$\mathbf{U}$, that is $\Phi_{\mathbf{X}}(\mathbf{U}\mathbf{n}) = \delta(\mathbf{n})$, where $\mathbf{U}$ is the generating matrix of the reciprocal lattice:

$$\mathbf{U} = 2\pi\mathbf{V}^{-T} = [\, \mathbf{u}_1 \quad \mathbf{u}_2 \quad \ldots \quad \mathbf{u}_D \,] \tag{2.8}$$

$\diamondsuit$

Note that, by definition $\Phi_{\mathbf{X}}(\mathbf{0}) = 1$. The space domain equivalent of the Nyquist condition is:

$$\sum_{\mathbf{n}} f_{\mathbf{X}}(\mathbf{x} + \mathbf{V}\mathbf{n}) = \frac{1}{|det\mathbf{V}|} \tag{2.9}$$

The motivation for Definition 7 comes from the quantization analysis provided in Sec. 3 where the left hand side of (2.9) appears in the expression of probability density function of the error vector. Examples of Nyquist-$\mathbf{V}$ random vectors are given in Sec. 3. Whenever the matrix $\mathbf{V}$ is clear from the context, we will just say Nyquist for both random vectors and their characteristic functions.

3. QUANTIZATION ANALYSIS

Define the error vector of a lattice quantizer, $Q(\mathcal{P}_0, \mathbf{V})$, as $\mathbf{e} = \mathbf{x} - Q(\mathbf{x})$. From the definition of the lattice quantizer, this error necessarily lies in $\mathcal{P}_0$. Each error vector $\mathbf{e} \in \mathcal{P}_0$ is produced by infinitely many input vectors of the form $\mathbf{e} + \mathbf{Vn}, \quad \mathbf{n} \in Z^D$ (see Fig. 4). Note that the error itself is one of these input vectors. Hence, the probability density function of error vector is:

$$f_{\mathbf{E}}(\mathbf{e}) = \begin{cases} \sum_{\mathbf{n}} f_{\mathbf{X}}(\mathbf{e} + \mathbf{Vn}), & \mathbf{e} \in \mathcal{P}_0; \\ 0 & \text{elsewhere} \end{cases} \tag{3.1}$$

One can find the Fourier series expansion, $\tilde{f}_{\mathbf{E}}(\mathbf{e})$, of $f_{\mathbf{E}}(\mathbf{e})$ with respect to the lattice generator matrix $\mathbf{V}$ and express it in the following form:

$$\tilde{f}_{\mathbf{E}}(\mathbf{e}) = \frac{1}{|det\mathbf{V}|} \sum_{\mathbf{n}} \Phi_{\mathbf{X}}(\mathbf{Un}) e^{-j\mathbf{e}^T\mathbf{Un}} \tag{3.2}$$

The restriction of $\tilde{f}_{\mathbf{E}}(\mathbf{e})$ to the basic cell $\mathcal{P}_0$ is $f_{\mathbf{E}}(\mathbf{e})$. For a brief summary of the relation between multidimensional Fourier series and Fourier transform see Appendix A. For further details of multidimensional Fourier series representation, the reader is referred to [24].

Theorem 1. *The quantization error of a lattice quantizer* $Q(\mathcal{P}_0, \mathbf{V})$ *is uniform in* $\mathcal{P}_0$*, that is*

$$f_{\mathbf{E}}(\mathbf{e}) = \begin{cases} \frac{1}{|det\mathbf{V}|}, & \mathbf{e} \in \mathcal{P}_0; \\ 0 & \textit{elsewhere} \end{cases} \tag{3.3}$$

if and only if the input vector $\mathbf{x}$ *is Nyquist-*$\mathbf{V}$*, that is* $\Phi_{\mathbf{X}}(\mathbf{Un}) = \delta(\mathbf{n})$. $\diamondsuit$

Proof. From the properties of Fourier series, we know that $\tilde{f}_{\mathbf{E}}(\mathbf{e})$ in (3.2) is a constant for all $\mathbf{e}$ if and only if the Fourier series coefficients $\Phi_{\mathbf{X}}(\mathbf{Un}) = 0, \forall \mathbf{n} \neq \mathbf{0}$. Thus $f_{\mathbf{E}}(\mathbf{e})$, the restriction of $\tilde{f}_{\mathbf{E}}(\mathbf{e})$ to $\mathcal{P}_0$, is constant in $\mathcal{P}_0$ if and only if $\Phi_{\mathbf{X}}(\mathbf{Un}) = \delta(\mathbf{n})$. $\blacksquare$

If $\mathcal{P}_0 = VOR(\mathbf{V})$ and if the condition of the theorem is satisfied, then $E[\mathbf{e}] = \mathbf{0}$ because $VOR(\mathbf{V})$ is symmetric with respect to the origin, and $E[\mathbf{e}\mathbf{e}^T] = \mathbf{G}_D(VOR, \mathbf{V})$, where $\mathbf{G}_D(VOR, \mathbf{V})$ is defined as in (2.6).

3.1. Nyquist-$\mathbf{V}$ random vectors

The next theorem shows some general classes of Nyquist-$\mathbf{V}$ random vectors:

Theorem 2. *The following random vectors* $\mathbf{x}$ *are Nyquist-*$\mathbf{V}$ *and therefore have uniform quantization errors in the quantization basic cell* $\mathcal{P}_0$ *when quantized by a lattice quantizer* $Q(\mathcal{P}_0, \mathbf{V})$:

1. $\mathbf{x}$ *is uniform in any basic cell* $\mathcal{P}$ *of the lattice* $\mathcal{L}(\mathbf{V})$.

2. $\mathbf{x}$ *is piecewise uniform in an arbitrary union of nonoverlapping basic cells of* $\mathcal{L}(\mathbf{V})$*, that is*

$$f_{\mathbf{X}}(\mathbf{x}) = \begin{cases} c_i, & \mathbf{x} \in \mathcal{P}_i; \\ 0 & \textit{elsewhere} \end{cases} \tag{3.4}$$

where c_i's are positive and $\sum_i c_i = \frac{1}{|det\mathbf{V}|}$.

3. $\mathbf{x}$ is a sum of several independent random vectors, one of which is Nyquist-$\mathbf{V}$. $\diamondsuit$

Proof.

1. Let $Q(\mathcal{P}_0, \mathbf{V})$ be a lattice quantizer with the basic cell $\mathcal{P}_0 = \mathcal{P}$. Then, $Q(\mathbf{x}) = \mathbf{0}$, and therefore $\mathbf{e} = \mathbf{x}$. Hence $\mathbf{e}$ is uniform in $\mathcal{P}_0$. By Theorem 1, $\mathbf{x}$ is Nyquist-$\mathbf{V}$ and therefore it has a uniform quantization error in $\mathcal{P}_0$ even if it is quantized with a lattice quantizer $Q(\mathcal{P}_0, \mathbf{V})$ with $\mathcal{P}_0 \neq \mathcal{P}$.

2. Writing the characteristic function explicitly we have:

$$\begin{aligned}
\Phi_{\mathbf{X}}(\mathbf{Un}) &= \int f_{\mathbf{X}}(\mathbf{x}) e^{j\mathbf{x}^T\mathbf{Un}} d\mathbf{x} \\
&= \sum_i \int_{\mathcal{P}_i} c_i e^{j\mathbf{x}^T\mathbf{Un}} d\mathbf{x} \quad \text{(by nonoverlapping assumption)} \\
&= |det\mathbf{V}| \sum_i c_i \delta(\mathbf{n}) \quad \text{(from part 1)} \\
&= \delta(\mathbf{n}).
\end{aligned} \tag{3.5}$$

3. Let $\mathbf{x} = \mathbf{v} + \mathbf{z}$, where $\mathbf{v}$ and $\mathbf{z}$ are independent. Then, $\Phi_{\mathbf{X}}(\mathbf{\Omega}) = \Phi_{\mathbf{V}}(\mathbf{\Omega})\Phi_{\mathbf{Z}}(\mathbf{\Omega})$. Therefore,

$$\Phi_{\mathbf{X}}(\mathbf{Un}) = \Phi_{\mathbf{V}}(\mathbf{Un})\Phi_{\mathbf{Z}}(\mathbf{Un}) = \delta(\mathbf{n}) \ \text{ if } \ \Phi_{\mathbf{V}}(\mathbf{Un}) = \delta(\mathbf{n}) \ \text{ or } \ \Phi_{\mathbf{Z}}(\mathbf{Un}) = \delta(\mathbf{n}). \tag{3.6}$$

Hence, if one of $\mathbf{v}$ or $\mathbf{z}$ is Nyquist-$\mathbf{V}$ then the sum is Nyquist-$\mathbf{V}$ as well. The extension to arbitrary number of independent random vectors is straightforward. ■

Example 1. If $\mathbf{x}$ is uniform in $SPD(\mathbf{V})$ or $VOR(\mathbf{V})$, then it is Nyquist-$\mathbf{V}$ because both $SPD(\mathbf{V})$ and $VOR(\mathbf{V})$ are basic cells of the lattice $\mathcal{L}(\mathbf{V})$.

The importance of Theorem 1 rests on the fact that we can make any given input vector satisfy the Nyquist condition by applying dither subtractively. (Sec. 4). If the dither is Nyquist-$\mathbf{V}$ and independent of the input (which is quite easy to manage as we will see) then from Theorem 2, part 3 the dithered random vector is Nyquist-$\mathbf{V}$ as well.

3.2. Error statistics when the input is arbitrary

What if the input vector is not Nyquist-$\mathbf{V}$ and we do not want to manipulate it by a dither? In that case, we have the following theorem that states the expected value of any funtion of the error vector $\mathbf{e}$:

Theorem 3. *Let $\mathbf{e}$ be the error vector of a lattice quantizer $Q(\mathcal{P}_0, \mathbf{V})$. Let $g(\mathbf{e})$ be an arbitrary function of $\mathbf{e}$. The expected value of $g(\mathbf{e})$ is $E[g(\mathbf{e})] = \sum_{\mathbf{n}} c_{\mathbf{n}} \Phi_X(\mathbf{Un})$ where $c_{\mathbf{n}} = \frac{1}{|det\mathbf{V}|}\int_{\mathcal{P}_0} g(\mathbf{e}) e^{-j\mathbf{e}^T\mathbf{Un}} d\mathbf{e}$.* $\diamondsuit$

Proof.

$$\begin{aligned}
E[g(\mathbf{e})] &= \int_{\mathcal{P}_0} g(\mathbf{e}) f_{\mathbf{E}}(\mathbf{e}) d\mathbf{e} \\
&= \int_{\mathcal{P}_0} g(\mathbf{e}) \sum_{\mathbf{n}} f_{\mathbf{X}}(\mathbf{e}+\mathbf{V}\mathbf{n}) d\mathbf{e} \quad \text{(from (3.1))} \\
&= \int_{\mathcal{P}_0} g(\mathbf{e}) \frac{1}{|det\mathbf{V}|} \sum_{\mathbf{n}} \Phi_{\mathbf{X}}(\mathbf{U}\mathbf{n}) e^{-j\mathbf{e}^T\mathbf{U}\mathbf{n}} d\mathbf{e} \quad \text{(from (3.2))} \\
&= \sum_{\mathbf{n}} \Phi_{\mathbf{X}}(\mathbf{U}\mathbf{n}) \frac{1}{|det\mathbf{V}|} \int_{\mathcal{P}_0} g(\mathbf{e}) e^{-j\mathbf{e}^T\mathbf{U}\mathbf{n}} d\mathbf{e} \\
&= \sum_{\mathbf{n}} c_{\mathbf{n}} \Phi_X(\mathbf{U}\mathbf{n})
\end{aligned} \tag{3.7}$$

as claimed. Note that $g(\mathbf{e})$ and therefore $c_{\mathbf{n}}$ can be a vector or even a matrix. We assume that interchanging the infinite sum and the integral in the above proof is permissible. ■

Corollary 1: Error moments. For any random vector $\mathbf{x}$, the first and second order moments of the quantization error $\mathbf{e}$ of a lattice quantizer $Q(VOR, \mathbf{V})$ are:

$$E[\mathbf{e}] = \sum_{\mathbf{n}\neq\mathbf{0}} \mathbf{c}_{\mathbf{n}} \Phi_X(\mathbf{U}\mathbf{n}), \qquad E[\mathbf{e}\mathbf{e}^T] = \mathbf{G}_D(VOR, \mathbf{V}) + \sum_{\mathbf{n}\neq\mathbf{0}} \mathbf{C}_{\mathbf{n}} \Phi_{\mathbf{X}}(\mathbf{U}\mathbf{n}) \tag{3.8}$$

where

$$\mathbf{c}_{\mathbf{n}} = \frac{1}{|det\mathbf{V}|} \int_{VOR(\mathbf{V})} \mathbf{e}\ e^{-j\mathbf{e}^T\mathbf{U}\mathbf{n}} d\mathbf{e}, \qquad \mathbf{C}_{\mathbf{n}} = \frac{1}{|det\mathbf{V}|} \int_{VOR(\mathbf{V})} \mathbf{e}\mathbf{e}^T e^{-j\mathbf{e}^T\mathbf{U}\mathbf{n}} d\mathbf{e}. \tag{3.9}$$

◊

Proof. Apply Theorem 3 with $\mathcal{P}_0 = VOR(\mathbf{V})$ to $g(\mathbf{e}) = \mathbf{e}$ and $g(\mathbf{e}) = \mathbf{e}\mathbf{e}^T$ repectively. Because of the symmetry of $VOR(\mathbf{V})$, $\mathbf{c}_{\mathbf{0}} = \mathbf{0}$. Moreover, $\mathbf{C}_{\mathbf{0}}$ is what we defined as $\mathbf{G}_D(VOR, \mathbf{V})$ in (2.6). ■

Note that if $\mathbf{x}$ is Nyquist-$\mathbf{V}$, all the terms of the infinite summations in (3.8) vanish in view of Theorem 1.

4. SUBTRACTIVE DITHERING

Let $\mathbf{v}$ be a random vector, statistically independent of the input vector $\mathbf{x}$. Adding this so called dither vector before the quantization and subtracting after it, we have the subtractive-dithered lattice quantization, as depicted in Fig. 5. The error vector is $\mathbf{e} = \mathbf{x} - (Q(\mathbf{x}+\mathbf{v}) - \mathbf{v}) = \mathbf{x} + \mathbf{v} - Q(\mathbf{x}+\mathbf{v})$. Notice that, this error is the same as the conventional quantization error for an input vector $\mathbf{x}+\mathbf{v}$. The characteristic function of the sum $\mathbf{x}+\mathbf{v}$ is $\Phi_{\mathbf{X}}(\mathbf{\Omega})\Phi_{\mathbf{V}}(\mathbf{\Omega})$ where $\Phi_{\mathbf{X}}(\mathbf{\Omega})$ and $\Phi_{\mathbf{V}}(\mathbf{\Omega})$ are the characteristic functions of $\mathbf{x}$ and $\mathbf{v}$ respectively. Hence $\mathbf{x}+\mathbf{v}$ is Nyquist-$\mathbf{V}$ whenever $\mathbf{v}$ is Nyquist-$\mathbf{V}$ and from Theorem 1, it follows that the quantization error is uniform in the quantization basic cell, $\mathcal{P}_0$. However, more is true as the following theorem shows:

Theorem 4. *In the subtractive quantization scheme of Fig. 5, the error vector $\mathbf{e}$ is statistically independent of the input vector $\mathbf{x}$ and uniformly distributed in $\mathcal{P}_0$ if and only if the dither $\mathbf{v}$ is Nyquist-$\mathbf{V}$, that is $\Phi_{\mathbf{V}}(\mathbf{U}\mathbf{n}) = \delta(\mathbf{n})$, where $\mathbf{U} = 2\pi\mathbf{V}^{-T}$.* ◊

Notice that, in the theorem statement, the error vector is to be independent of the input vector for all possible source statistics.

Proof. Let $\mathbf{u} = \mathbf{x} + \mathbf{v}$. The conditional density of $\mathbf{u}$, conditioned on $\mathbf{x}$, is $f_{\mathbf{U}/\mathbf{X}}(\mathbf{u}/\mathbf{x}) = f_{\mathbf{V}}(\mathbf{u} - \mathbf{x})$ and the corresponding characteristic function is $\Phi_{\mathbf{U}/\mathbf{X}}(\mathbf{\Omega}) = \Phi_{\mathbf{V}}(\mathbf{\Omega})e^{j\mathbf{\Omega}^T\mathbf{x}}$. Hence using (3.2), we can write the conditional density function of the error vector as:

$$f_{\mathbf{E}/\mathbf{X}}(\mathbf{e}/\mathbf{x}) = \frac{1}{|det\mathbf{V}|}\sum_{\mathbf{n}} \Phi_{\mathbf{U}/\mathbf{X}}(\mathbf{Un})e^{-j\mathbf{e}^T\mathbf{Un}} = \frac{1}{|det\mathbf{V}|}\sum_{\mathbf{n}} \Phi_{\mathbf{V}}(\mathbf{Un})e^{j\mathbf{x}^T\mathbf{Un}}e^{-j\mathbf{e}^T\mathbf{Un}} \tag{4.1}$$

for $\mathbf{e} \in \mathcal{P}_0$ and 0 elsewhere. One can think of this as the nonseparable discrete Fourier transform of the sequence $\Phi_{\mathbf{V}}(\mathbf{Un})e^{-j\mathbf{e}^T\mathbf{Un}}$, $\mathbf{x}$ being the transform domain vector. Hence from the uniqueness of Fourier transform, this is independent of $\mathbf{x}$ if and only if $\Phi_{\mathbf{V}}(\mathbf{Un})e^{-j\mathbf{e}^T\mathbf{Un}} = \delta(\mathbf{n})$ which is equivalent to $\Phi_{\mathbf{V}}(\mathbf{Un}) = \delta(\mathbf{n})$. ■

If $\mathcal{P}_0 = VOR(\mathbf{V})$ and the condition of the theorem is satisfied, then $E[\mathbf{e}] = \mathbf{0}$ and $E[\mathbf{ee}^T] = \mathbf{G}_D(VOR, \mathbf{V})$, where $\mathbf{G}_D(VOR, \mathbf{V})$ is defined as in (2.6).

4.1. Nyquist-$\mathbf{V}$ dither vectors: examples and generation

In Theorem 2, we provided some classes of random vectors that are Nyquist-$\mathbf{V}$. Any such vector will serve as a dither vector as long as it is independent of the input vector $\mathbf{x}$. In particular, as given in Example 1, we can use a dither vector that is uniform in $SPD(\mathbf{V})$ or $VOR(\mathbf{V})$. The one that is uniform in $SPD(\mathbf{V})$ is relatively simple to generate and a method for generating such a dither is given next.

Generation of a Nyquist-V vector: We will show how to obtain a random vector that is uniform in $SPD(\mathbf{V})$ and therefore is Nyqusit-$\mathbf{V}$. First generate a set of D independent random variables $z_1, z_2, .., z_D$ each of which is uniform in [-1/2, 1/2). Form the vector $\mathbf{z} = [\, z_1 \quad z_2 \quad \ldots \quad z_D\,]^T$. The vector $\mathbf{v} = \mathbf{Vz}$ is Nyquist-$\mathbf{V}$ because:

$$\Phi_{\mathbf{V}}(\mathbf{Un}) = E_{\mathbf{V}}[e^{j\mathbf{v}^T\mathbf{Un}}] = E_{\mathbf{Z}}[e^{j\mathbf{z}^T\mathbf{V}^T\mathbf{Un}}] = E_{\mathbf{Z}}[e^{j2\pi\mathbf{z}^T\mathbf{n}}] = \delta(\mathbf{n}). \tag{4.2}$$

Since the error vector of a lattice quantizer $Q(\mathcal{P}_0, \mathbf{V})$ can be made uniform in $\mathcal{P}_0$ by applying a Nyquist dither subtractively, we will give our attention to the moments of that error. All of the results stated below can actually be viewed as the properties of the underlying lattice, but the reader should keep in mind that they become the properties of the quantization error if the input is Nyquist, or if a Nyquist-dither is added to the input prior to quantization and subtracted after it.

4.2. Performance comparison of lattice quantizers

Note that $\mathbf{G}_D(\mathcal{P}_0, \mathbf{V})$, the second moment of the error of a lattice quantizer $Q(\mathcal{P}_0, \mathbf{V})$ with Nyquist-$\mathbf{V}$ input, is a positive definite symmetric matrix. The total mean square error of the quantizer is the trace of this matrix: $E[\|\mathbf{e}\|^2] = Tr(\mathbf{G}_D(\mathcal{P}_0, \mathbf{V}))$.

Orthogonal lattices: An orthogonal lattice is a lattice whose generator matrix $\mathbf{V}$ satisfies

$$\mathbf{VV}^T = |det\mathbf{V}|^{2/D}\mathbf{\Lambda}, \tag{4.3}$$

where $\mathbf{\Lambda}$ is a diagonal matrix with diagonal elements $\lambda_i > 0$. To preserve the determinants of both sides of (4.3) we have $\prod_{i=1}^{D} \lambda_i = 1$.

Notice that, an orthogonal lattice quantizer with $VOR(\mathbf{V})$ as its basic cell can be considered as a collection of scalar uniform quantizers for each dimension with possibly different step sizes. We note the following result on the second moment matrix, $\mathbf{G}_D(VOR, \mathbf{V})$, of an orthogonal lattice quantizer $Q(VOR, \mathbf{V})$:

Fact 1. If the lattice $\mathcal{L}(\mathbf{V})$ is orthogonal, that is $\mathbf{V}\mathbf{V}^T = |det\mathbf{V}|^{2/D}\mathbf{\Lambda}$, then

$$\mathbf{G}_D(VOR, \mathbf{V}) = \frac{1}{12}|det\mathbf{V}|^{2/D}\mathbf{\Lambda}. \tag{4.4}$$

$\diamondsuit$

See Appendix B for the proof. As a special case, if $\mathbf{V}\mathbf{V}^T = |det\mathbf{V}|^{2/D}\mathbf{I}$, then $\mathbf{G}_D(VOR, \mathbf{V}) = \frac{1}{12}|det\mathbf{V}|^{2/D}\mathbf{I}$, and therefore $\frac{1}{D}E[\|\mathbf{e}\|^2] = \frac{1}{12}|det\mathbf{V}|^{2/D}$. Taking this as a reference, we can compare the performances of other lattice quantizers. We will normalize the total mean square error per dimension of any lattice quantizer $Q(\mathcal{P}_0, \mathbf{V})$ by $|det\mathbf{V}|^{2/D}$, giving a proper figure of merit for lattices of different volume and dimension D.

Definition 8. The *dimensionless second moment* of a lattice quantizer $Q(\mathcal{P}_0, \mathbf{V})$, denoted by $\sigma_D^2(\mathcal{P}_0, \mathbf{V})$, is defined as

$$\sigma_D^2(\mathcal{P}_0, \mathbf{V}) = \frac{1}{D|detV|^{2/D}} Tr(\mathbf{G}_D(\mathcal{P}_0, \mathbf{V})) = \frac{1}{D|det\mathbf{V}|^{1+2/D}} \int_{\mathcal{P}_0} \|\mathbf{e}\|^2 d\mathbf{e} \tag{4.5}$$

where $\mathbf{G}_D(\mathcal{P}_0, \mathbf{V})$ is as in (2.6). $\diamondsuit$

The quantity $\sigma_D^2(\mathcal{P}_0, \mathbf{V})$ also comes out of high bit rate analysis of lattice quantizers [19], [25], [18]. It is proven in [18] that for an undithered lattice quantizer, as the unit volume, $|detV|$ of a quantizer $Q(\mathcal{P}_0, \mathbf{V})$ goes to 0, the normalized mean square error approaches the limit $\sigma_D^2(\mathcal{P}_0, \mathbf{V})$. The name *dimensionless second moment* is used in [19].

The following fact is on the performance of orthogonal lattice quantizers. The reader is referred to Appendix B for the proof.

Fact 2. Let $Q(\mathcal{P}_0, \mathbf{V})$ be an orthogonal lattice quantizer, that is let the generator matrix $\mathbf{V}$ satisfy (4.3). Then,

$$\sigma_D^2(\mathcal{P}_0, \mathbf{V}) \geq \frac{1}{12}, \tag{4.6}$$

with equality if and only if $\mathbf{V}\mathbf{V}^T = |det\mathbf{V}|^{2/D}\mathbf{I}$ and $\mathcal{P}_0 = VOR(\mathbf{V})$. $\diamondsuit$

The following result is on the performance of lattice quantizers whose quantization basic cells are $SPD(\mathbf{V})$ rather than $VOR(\mathbf{V})$. Note that this result is not a special case of the previous one, where we assumed $\mathbf{V}$ was orthogonal. Here, there is no assumption on $\mathbf{V}$.

Fact 3. Given a lattice generator matrix $\mathbf{V}$,

$$\sigma_D^2(SPD, \mathbf{V}) \geq \frac{1}{12}, \tag{4.7}$$

with equality if and only if $\mathbf{V}\mathbf{V}^T = |det\mathbf{V}|^{2/D}\mathbf{I}$. $\diamondsuit$

Proof. By making a change of variable as in the proof of Fact 1 in Appendix B, it is easy to see that $\mathbf{G}_D(SPD, \mathbf{V}) = \frac{1}{12}\mathbf{V}\mathbf{V}^T$. Hence,

$$\begin{aligned} \sigma_D^2(SPD) &= \frac{1}{D|detV|^{2/D}} Tr(\frac{1}{12}\mathbf{V}\mathbf{V}^T) \\ &\geq \frac{1}{12|detV|^{2/D}} |det\mathbf{V}\mathbf{V}^T|^{1/D} \quad \text{(see below)} \\ &= \frac{1}{12}. \end{aligned} \tag{4.8}$$

The inequality follows from the AM-GM inequality and the Hadamard inequality [23] as explained next. The diagonal elements of the positive definite matrix $\mathbf{V}\mathbf{V}^T$ are positive. Hence, their arithmetic mean is greater than or equal to their geometric mean. And by the Hadamard inequality, the product of the diagonal elements is greater than or equal to the determinant of $\mathbf{V}\mathbf{V}^T$. The former is an equality if and only if the diagonal elements of $\mathbf{V}\mathbf{V}^T$ are the same and the latter is an equality if and only if $\mathbf{V}\mathbf{V}^T$ is diagonal. Hence, the result follows. $\blacksquare$

As we noted before, for a given lattice $\mathcal{L}(\mathbf{V})$, the minimum dimensionless second moment is achieved by the basic cell $VOR(\mathbf{V})$. One can ask the question: among all the lattices in R^D, what is the optimum lattice that will minimize the dimensionless second moment $\sigma_D^2(VOR, \mathbf{V})$? This question turns out to be theoretically very challenging. The answer is not known for arbitrary D and there is no proof of optimality for dimensions higher than 3 (see for example, [5]).

Examples of optimum lattices. Here are some lattices that have minimum dimensionless second moments:

Case where $D = 1$. The only lattice is the points of the form $\Delta n, \forall n \in Z$, $\Delta \in R$. Any basic cell $\mathcal{P}$ has a total length of Δ. Obviously, the minimum dimensionless second moment is achieved by $VOR(\Delta) = [-\frac{\Delta}{2}, \frac{\Delta}{2})$ and its value is $\sigma_1^2(VOR, \Delta) = \frac{1}{12}$.

Case where $D = 2$. The optimum lattice that minimizes $\sigma_D^2(VOR, \mathbf{V})$ is the one whose $VOR(\mathbf{V})$ is the regular hexagon [26]. A generating matrix for this lattice is

$$\mathbf{V} = \begin{pmatrix} 3 & \frac{3}{2} \\ 0 & \frac{\sqrt{3}}{2} \end{pmatrix} \tag{4.9}$$

The unit volume of the lattice is: $|det\mathbf{V}| = \frac{3\sqrt{3}}{2}$. By explicitly evaluating integrals, we have:

$$E[\mathbf{e}\mathbf{e}^T] = \mathbf{G}_2(VOR, \mathbf{V}) = \frac{5}{24}\mathbf{I}, \tag{4.10}$$

where $\mathbf{I}$ is 2×2 identity matrix. The corresponding dimensionless second moment is: $\sigma_2^2(VOR, \mathbf{V}) = \frac{5}{24} / \frac{3\sqrt{3}}{2} = \frac{5}{36\sqrt{3}} = .08018754$. Compare this to that of optimum one dimensional lattice: $\sigma_1^2(VOR, \Delta) = \frac{1}{12} = 0.0833...$

Case where $D = 3$. The optimum lattice is the body-centered cubic lattice, also called the truncated octahedron as is proven by Barnes and Sloane [27]. This lattice has $\sigma_3^2(VOR, \mathbf{V}) = \frac{19}{192\sqrt[3]{2}} = 0.0785433...$

Case where $D = \infty$. The limiting value of minimum $\sigma_D^2(VOR, \mathbf{V})$ is [19],

$$\liminf_{D\to\infty} \sigma_D^2 = \frac{1}{2\pi e} = 0.058823... \tag{4.11}$$

For a tabulation of lattices that have best known $\sigma_D^2(VOR, \mathbf{V})$ see [19].

After the observation in (4.10) that $\mathbf{G}_2(VOR, \mathbf{V})$ is diagonal with equal elements, these authors suspected that this might be true for any optimum lattice of arbitrary dimension. This turns out to be indeed the case, as elaborated in the next theorem. Assume the dimension D is given and we look at different lattices with the objective of minimizing the dimensionless second moment $\sigma_D^2(\mathcal{P}_0, \mathbf{V})$. Hence the quantization basic cells are chosen to be $VOR(\mathbf{V})$ for each lattice generator matrix $\mathbf{V}$. We have the following result:

Theorem 5. *For a lattice quantizer* $Q(\mathcal{P}_0, \mathbf{V})$ *to be optimum, that is, to have the minimum dimensionless second moment* $\sigma_D^2(\mathcal{P}_0, \mathbf{V})$*, it is necessary that* $\mathcal{P}_0 = VOR(\mathbf{V})$*, and*

$$\mathbf{G}_D(VOR, \mathbf{V}) = c\mathbf{I} \tag{4.12}$$

where $c = \sigma_D^2(VOR, \mathbf{V})|det\mathbf{V}|^{2/D}$ *and* $\mathbf{I}$ *is the* $D \times D$ *identity matrix.* $\diamondsuit$

Comment. Note that $\mathbf{G}_D(\mathcal{P}_0, \mathbf{V})$ is the second moment matrix of a vector $\mathbf{e}$ with uniform pdf in $\mathcal{P}_0$. By Theorem 1, uniformity of the error in $\mathcal{P}_0$ is equivalent to the Nyquist-$\mathbf{V}$ condition on the input vector $\mathbf{x}$. This can be assured by applying an independent Nyquist-$\mathbf{V}$ dither subtractively, as seen from Theorem 4.

During the preparation of this paper, the authors noticed that this result has appeared very recently in [16] and a proof has been provided in [15]. Nevertheless, we provide our proof here for completeness and convenience.

Proof. As we noted before, for any given $\mathbf{V}$, the minimum dimensionless second moment is achieved by the quantization basic cell $VOR(\mathbf{V})$. Hence we take $\mathcal{P}_0 = VOR(\mathbf{V})$. Define a new random vector $\mathbf{z} = \mathbf{Q}^{-1}\mathbf{x}$ for some nonsingular $\mathbf{Q}$, and consider Fig. 6. Since $\hat{\mathbf{x}}$ is on the lattie $\mathcal{L}(\mathbf{V})$, the vector $\hat{\mathbf{z}}$ is on the lattice $\mathcal{L}(\mathbf{Q}^{-1}\mathbf{V})$. We can therefore regard Fig. 6 as a lattice quantizer for the vector $\mathbf{z}$, with the quantized values on $\mathcal{L}(\mathbf{Q}^{-1}\mathbf{V})$. Define the quantization errors $\mathbf{e} = \mathbf{x} - \hat{\mathbf{x}}$ and $\mathbf{f} = \mathbf{z} - \hat{\mathbf{z}}$. Then $\mathbf{f} = \mathbf{Q}^{-1}\mathbf{e}$. Since $\mathbf{e}$ is uniform in $VOR(\mathbf{V})$, the error $\mathbf{f}$ is uniform in a basic cell, $\mathcal{P}$ of $\mathcal{L}(\mathbf{Q}^{-1}\mathbf{V})$. Assuming that $\mathbf{V}$ is optimal for the dimension D , the dimensionless second moments should satisfy

$$\sigma_D^2(\mathcal{P}, \mathbf{Q}^{-1}\mathbf{V}) \geq \sigma_D^2(VOR, \mathbf{V}). \tag{4.13}$$

Observe that $E[\mathbf{ff}^T] = \mathbf{Q}^{-1}E[\mathbf{ee}^T]\mathbf{Q}^{-T}$. Let us choose $\mathbf{Q}$ such that $\mathbf{QQ}^T = E[\mathbf{ee}^T]$, so $E[\mathbf{ff}^T] = \mathbf{I}$. Substituting the expressions

$$\sigma_D^2(\mathcal{P}, \mathbf{Q}^{-1}\mathbf{V}) = \frac{E[\|\mathbf{f}\|^2]}{D|det\mathbf{Q}^{-1}\mathbf{V}|^{2/D}}, \quad \text{and} \quad \sigma_D^2(VOR, \mathbf{V}) = \frac{E[\|\mathbf{e}\|^2]}{D|det\mathbf{V}|^{2/D}} \tag{4.14}$$

into (4.13), we can simplify it to

$$|det\mathbf{Q}\mathbf{Q}^T|^{1/D} \geq \frac{1}{D}Tr(\mathbf{Q}\mathbf{Q}^T). \tag{4.15}$$

Let λ_i be the eigenvalues of the Hermitian matrix $\mathbf{Q}\mathbf{Q}^T$. Hence the determinant and the trace above are, respectively, the product and the sum of these eigenvalues. So the preceding equation is equivalent to $(\prod_{i=1}^{D}\lambda_i)^{1/D} \geq \frac{1}{D}\sum_{i=1}^{D}\lambda_i$. Since by construction $\mathbf{Q}\mathbf{Q}^T$ is positive definite, $\lambda_i > 0$ for all i. We can therefore apply the AM-GM inequality to conclude $\frac{1}{D}\sum_{i=1}^{D}\lambda_i \geq (\prod_{i=1}^{D}\lambda_i)^{1/D}$. The preceding two inequalities on $\{\lambda_i\}$ can be simultaneously true if and only if λ_i is identical for all i. Since $\mathbf{Q}\mathbf{Q}^T$ is Hermitian, this proves that $\mathbf{Q}\mathbf{Q}^T = \lambda\mathbf{I}$. So we have proved that $E[\mathbf{e}\mathbf{e}^T] = \lambda\mathbf{I}$. Combining this with the definition of $\sigma_D^2(VOR, \mathbf{V})$ we obtain (4.12) indeed. ∎

5. NONSUBTRACTIVE DITHERING

In subtractive dithering, one should regenerate the dither vector exactly at the reconstruction end. This is, in most cases, undesirable. The easiest remedy is not to subtract the dither vector, and this results in the nonsubtractive dithering scheme. Referring to Fig. 7, we define the error vector to be $\mathbf{e} = \mathbf{x} - Q(\mathbf{x}+\mathbf{v})$. The error is no longer a periodic function of the input and therefore we do not have a periodical relationship between the error and the input pdf's similar to (3.1) or (3.2). Hence, as can be shown, the error cannot be rendered statistically independent from the input. However, the moments of the error can be rendered independent from the input as will be elaborated next. This result is the generalization of the well-known one dimensional nonsubtractive dithering result [11], [12]. First we will give a lemma that will express the relevant moments in terms of gradients of a function of dither.

Let ∇ and $\nabla\nabla^T$ denote the first and second order gradient operators operating on functions of D variables, $\omega_1, \omega_2, .., \omega_D$:

$$\nabla = [\,\tfrac{\partial}{\partial\omega_1} \quad \tfrac{\partial}{\partial\omega_2} \cdots \quad \tfrac{\partial}{\partial\omega_D}\,]^T, \qquad (\nabla\nabla^T)_{ij} = \frac{\partial^2}{\partial\omega_i\partial\omega_j} \tag{5.1}$$

Let $\mathbf{z}$ be a random vector that is uniform in the quantization basic cell $\mathcal{P}_0$ of the lattice quantizer $Q(\mathcal{P}_0, \mathbf{V})$. Let $f_{\mathbf{V}}(\mathbf{v})$ and $f_{\mathbf{Z}}(\mathbf{z})$ be the pdf's of $\mathbf{v}$ and $\mathbf{z}$ respectively. By definition,

$$f_{\mathbf{Z}}(\mathbf{z}) = \begin{cases} \frac{1}{|det\mathbf{V}|}, & \mathbf{z} \in \mathcal{P}_0; \\ 0 & \text{elsewhere} \end{cases} \tag{5.2}$$

Lemma 1. The first and second moments of the error vector $\mathbf{e}$ of a nonsubtractively dithered lattice quantizer $Q(\mathcal{P}_0, \mathbf{V})$ conditioned on the input vector $\mathbf{x}$ are:

$$E[\mathbf{e}/\mathbf{x}] = \frac{1}{j}\sum_{\mathbf{n}} \nabla H(\mathbf{U}\mathbf{n})e^{j\mathbf{x}^T\mathbf{U}\mathbf{n}} \tag{5.3}$$

$$E[\mathbf{e}\mathbf{e}^T/\mathbf{x}] = \frac{1}{j^2}\sum_{\mathbf{n}} \nabla\nabla^T H(\mathbf{U}\mathbf{n})e^{j\mathbf{x}^T\mathbf{U}\mathbf{n}} \tag{5.4}$$

where,

$$H(\mathbf{\Omega}) = \int h(\mathbf{x})e^{-j\mathbf{\Omega}^T\mathbf{x}}d\mathbf{x} \qquad h(\mathbf{e}) = f_{\mathbf{V}}(\mathbf{e}) * f_{\mathbf{Z}}(-\mathbf{e}) \tag{5.5}$$

$\diamondsuit$

Remark. Note that the extension of the above result to higher moments is straightforward by defining the corresponding operators in an obvious way. However, our interest will only be in the first and second order moments.

Proof. Since we do not subtract the dither after the quantizer, the reproduction points are the lattice points of the form $\mathbf{Vn}$. That is, $Q(\mathbf{x}+\mathbf{v}) = \mathbf{Vn}$ for some $\mathbf{n} \in Z^D$. Hence, the corresponding error vector is $\mathbf{e} = \mathbf{x} - \mathbf{Vn}$. Note that, given $\mathbf{x}$, this is a discrete random vector. It has the probality mass function:

$$\begin{aligned} P_{\mathbf{E}/\mathbf{X}}(\mathbf{x}-\mathbf{Vn}) &= Prob\{Q(\mathbf{x}+\mathbf{v}) = \mathbf{Vn}\} \\ &= Prob\{\mathbf{x}+\mathbf{v} = \mathbf{x}_0 + \mathbf{Vn}, \mathbf{x}_0 \in \mathcal{P}_0\} \\ &= \int_{\mathcal{P}_0(\mathbf{Vn}-\mathbf{x})} f_{\mathbf{V}}(\mathbf{v})d\mathbf{v} \end{aligned} \tag{5.6}$$

where $\mathcal{P}_0(\mathbf{Vn}-\mathbf{x})$ denotes the translated region of $\mathcal{P}_0$ by the vector $\mathbf{Vn}-\mathbf{x}$. Using the artificial random vector $\mathbf{z}$ defined by the pdf in (5.2), one can express the preceding as a convolution:

$$P_{\mathbf{E}/\mathbf{X}}(\mathbf{x}-\mathbf{Vn}) = |det\mathbf{V}| \int f_{\mathbf{V}}(\mathbf{v}) f_{\mathbf{Z}}(\mathbf{v}-\mathbf{Vn}+\mathbf{x})d\mathbf{v} \tag{5.7}$$

Hence,

$$\begin{aligned} P_{\mathbf{E}/\mathbf{X}}(\mathbf{e}) &= |det\mathbf{V}| \int f_{\mathbf{V}}(\mathbf{v}) f_{\mathbf{Z}}(\mathbf{v}+\mathbf{e})d\mathbf{v} \\ &= |det\mathbf{V}| \; h(-\mathbf{e}), \end{aligned} \tag{5.8}$$

where,

$$h(\mathbf{e}) = f_{\mathbf{V}}(\mathbf{e}) * f_{\mathbf{Z}}(-\mathbf{e}). \tag{5.9}$$

Now, the first order moment of the error vector is:

$$\begin{aligned} E[\mathbf{e}/\mathbf{x}] &= \sum \mathbf{e} P_{\mathbf{E}/\mathbf{X}}(\mathbf{e}) \\ &= \sum_{\mathbf{n}} (\mathbf{x}-\mathbf{Vn})|det\mathbf{V}|h(-\mathbf{x}+\mathbf{Vn}) \\ &= \sum_{\mathbf{n}} g(\mathbf{x}+\mathbf{Vn}) \end{aligned} \tag{5.10}$$

where $g(\mathbf{x})$ is defined as $|det\mathbf{V}|\mathbf{x}h(-\mathbf{x})$. The Fourier transform of $g(\mathbf{x})$ is, $G(\mathbf{\Omega}) = \frac{1}{j}|det\mathbf{V}|\nabla H(-\mathbf{\Omega})$, where $H(\mathbf{\Omega})$ is the Fourier transform of $h(\mathbf{e})$, that is

$$H(\mathbf{\Omega}) = \Phi_{\mathbf{V}}(-\mathbf{\Omega})\Phi_{\mathbf{Z}}(\mathbf{\Omega}). \tag{5.11}$$

By using the Fourier series representation, (see Appendix A), one can write (5.10) as:

$$E[\mathbf{e}/\mathbf{x}] = \frac{1}{|det\mathbf{V}|} \sum_{\mathbf{n}} G(\mathbf{Un})e^{-j\mathbf{x}^T\mathbf{Un}} \tag{5.12}$$

which reduces to (5.3). The derivation of (5.4) is through the same steps and is omitted. ■

Using these results and noting the uniqueness property of Fourier series, the next theorem follows:

Theorem 6. *Consider the nonsubtractive quantization scheme of Fig. 7. Let $H(\mathbf{\Omega})$ be as in (5.11).*

1. The first order moment of the error vector is independent of the input if and only if $\nabla H(\mathbf{\Omega})$ is Nyquist-$\mathbf{U}$, that is $\nabla H(\mathbf{Un}) = \mathbf{c}\ \delta(\mathbf{n})$,

2. The second order moment matrix of the error vector is independent of the input if and only if $\nabla\nabla^T H(\mathbf{\Omega})$ is Nyquist-$\mathbf{U}$, that is $\nabla\nabla^T H(\mathbf{Un}) = \mathbf{C}\ \delta(\mathbf{n})$.

If the corresponding conditions are satisfied then,

$$E[\mathbf{e}/\mathbf{x}] = E[\mathbf{e}] = E[\mathbf{z}] - E[\mathbf{v}], \quad E[\mathbf{e}\mathbf{e}^T/\mathbf{x}] = E[\mathbf{e}\mathbf{e}^T] = E[(\mathbf{z}-\mathbf{v})(\mathbf{z}-\mathbf{v})^T] \tag{5.13}$$

respectively, where $\mathbf{z}$ is uniform in $\mathcal{P}_0$ and indepedent of $\mathbf{v}$. $\diamondsuit$

Remark. If the conditions are satisfied with a symmetric basic cell $\mathcal{P}_0$, then $E[\mathbf{e}] = -E[\mathbf{v}]$, and $E[\mathbf{e}\mathbf{e}^T] = \mathbf{G}_D(\mathcal{P}_0, \mathbf{V}) + E[\mathbf{v}\mathbf{v}^T]$, where $\mathbf{G}_D(\mathcal{P}_0, \mathbf{V})$ is defined as in (2.6). In particular, if $\mathcal{P}_0 = VOR(\mathbf{V})$, then $E[\mathbf{e}\mathbf{e}^T] = \mathbf{G}_D(VOR, \mathbf{V}) + E[\mathbf{v}\mathbf{v}^T]$.

Proof. The necessary and sufficient conditions follow from Lemma 1. If the corresponding conditions are satisfied, then

$$E[\mathbf{e}] = \frac{1}{j}\nabla H(\mathbf{0}), \quad \text{and} \quad E[\mathbf{e}\mathbf{e}^T] = \frac{1}{j^2}\nabla\nabla^T H(\mathbf{0}), \tag{5.14}$$

respectively. Now, by (5.9), h($\mathbf{e}$) can be considered as the pdf of a random vector $\mathbf{v} - \mathbf{z}$, where $\mathbf{z}$ is independent from $\mathbf{v}$ and uniform in $\mathcal{P}_0$. Hence, from the moment generating property of characteristic functions, we have

$$\frac{1}{j}\nabla H(\mathbf{0}) = E[\mathbf{z} - \mathbf{v}], \quad \text{and} \quad \frac{1}{j^2}\nabla\nabla^T H(\mathbf{0}) = E[(\mathbf{z}-\mathbf{v})(\mathbf{z}-\mathbf{v})^T]. \tag{5.15}$$

■

Example 2. Let $\mathbf{v}$ be any Nyquist-$\mathbf{V}$ random vector, that is, $\Phi_{\mathbf{V}}(\mathbf{Un}) = \delta(\mathbf{n})$. Then the condition for the first part of the theorem is satisfied. To see this:

$$\nabla H(\mathbf{\Omega}) = \Phi_{\mathbf{V}}(-\mathbf{\Omega})\nabla\Phi_{\mathbf{Z}}(\mathbf{\Omega}) - \Phi_{\mathbf{Z}}(\mathbf{\Omega})\nabla\Phi_{\mathbf{V}}(-\mathbf{\Omega}) \tag{5.16}$$

Since $\Phi_{\mathbf{Z}}$ itself is Nyquist and $\Phi_{\mathbf{V}}$ is chosen to be so, ∇H is Nyquist as well. Hence $E[\mathbf{e}/\mathbf{x}] = E[\mathbf{e}] = E[\mathbf{z}] - E\mathbf{v}]$. This is zero if (i) the dither is uniform in the quantization basic cell or (ii) the dither is uniform in any symmetric basic cell and the quantization basic cell is symmetric. The dither vector that is uniform in $SPD(\mathbf{V})$ satisfies the condition of the theorem and it produces zero-mean error if the quantization basic cell is symmetric.

Example 3. Let $\mathbf{v} = \mathbf{z}_1 + \mathbf{z}_2$ where, $\mathbf{z}_1$ and $\mathbf{z}_2$ are independent random vectors each of which is Nyquist-$\mathbf{V}$. Then the condition for the second part of the theorem is satisfied, because:

$$\begin{aligned}\nabla\nabla^T H(\mathbf{\Omega}) &= \nabla(\Phi_{\mathbf{V}}(-\mathbf{\Omega})\nabla^T\Phi_{\mathbf{Z}}(\mathbf{\Omega}) - \Phi_{\mathbf{Z}}(\mathbf{\Omega})\nabla^T\Phi_{\mathbf{V}}(-\mathbf{\Omega}))\\ &= \Phi_{\mathbf{V}}(-\mathbf{\Omega})\nabla\nabla^T\Phi_{\mathbf{Z}}(\mathbf{\Omega}) - \nabla\Phi_{\mathbf{V}}(-\mathbf{\Omega})\nabla^T\Phi_{\mathbf{Z}}(\mathbf{\Omega}) - \nabla\Phi_{\mathbf{Z}}(\mathbf{\Omega})\nabla^T\Phi_{\mathbf{V}}(-\mathbf{\Omega}) + \Phi_{\mathbf{Z}}(\mathbf{\Omega})\nabla\nabla^T\Phi_{\mathbf{V}}(-\mathbf{\Omega})\end{aligned} \tag{5.17}$$

The first term is Nyquist because, $\Phi_{\mathbf{V}}$, being the product of two Nyquist functions, is Nyquist. From the previous example, $\nabla\Phi_{\mathbf{V}}$ is also Nyquist and therefore second and third terms are Nyquist. Since $\Phi_{\mathbf{Z}}$ is given to be Nyquist, the last term is Nyquist too, making $\nabla\nabla^T H$ Nyquist as desired. Hence,

$$E[\mathbf{e}\mathbf{e}^T/\mathbf{x}] = E[\mathbf{e}\mathbf{e}^T] = E[(\mathbf{z}-\mathbf{v})(\mathbf{z}-\mathbf{v})^T]. \tag{5.18}$$

If the quantization basic cell and the regions of supports of the random vectors $\mathbf{z}_1$ and $\mathbf{z}_2$ are symmetric with respect to the origin, then $E[\mathbf{e}\mathbf{e}^T] = E[\mathbf{z}\mathbf{z}^T + \mathbf{v}\mathbf{v}^T] = E[\mathbf{z}\mathbf{z}^T] + E[\mathbf{z}_1\mathbf{z}_1^T] + E[\mathbf{z}_2\mathbf{z}_2^T]$. Note that, the dither in this example satisfies the condition for the first part of the theorem as well, hence the first order moment is also independent of the input. In particular, notice the following special cases:
Assume the quantization basic cell, $\mathcal{P}_0$ is symmetric with respect to the origin.
(i) if both $\mathbf{z}_1$ and $\mathbf{z}_2$ are uniform in $SPD(\mathbf{V})$, then

$$E[\mathbf{e}\mathbf{e}^T/\mathbf{x}] = E[\mathbf{e}\mathbf{e}^T] = \mathbf{G}_D(\mathcal{P}_0, \mathbf{V}) + \frac{1}{6}\mathbf{V}\mathbf{V}^T \tag{5.19}$$

(ii) if both both $\mathbf{z}_1$ and $\mathbf{z}_2$ are uniform in $\mathcal{P}_0$, then

$$E[\mathbf{e}\mathbf{e}^T/\mathbf{x}] = E[\mathbf{e}\mathbf{e}^T] = 3\mathbf{G}_D(\mathcal{P}_0, \mathbf{V}) \tag{5.20}$$

Assume, we use a dither as in Example 3, which satisfies the first and second order moment independence conditions. Among all such schemes, the minimum total mean square error is achieved by using the lattice quantizer with $\mathcal{P}_0 = VOR(\mathbf{V})$, and a dither vector that is sum of two independent vectors that are uniform in $VOR(\mathbf{V})$ as in the second special case given above. The resulting total mean square error is 3 times that of the subtractive dithered quantization and that is true for any dimension D. Making use of Theorem 5 on optimum lattices, we have the following result:

Theorem 7. *Let $\mathbf{V}$ be the generating matrix of the optimum lattice (i.e. the lattice with minimum $\sigma_D^2(VOR, \mathbf{V})$). In subtractive dithering, the minimum total mean square error is achieved by any dither that is Nyquist-$\mathbf{V}$. In nonsubtractive dithering, among all dithers as in Example 3, the minimum total mean square error is achieved by the optimal lattice $\mathbf{V}$, and by the dither that is the sum of two independent Nyquist-$\mathbf{V}$ vectors each of which is uniform in $VOR(\mathbf{V})$. The resulting second moment matrices are:*

$$E[\mathbf{e}\mathbf{e}^T] = \mathbf{G}_D(VOR, \mathbf{V}) = \sigma_D^2(VOR, \mathbf{V})|det\mathbf{V}|^{2/D}\mathbf{I} \qquad \text{(subtractive dithering)} \tag{5.21}$$

$$E[\mathbf{e}\mathbf{e}^T] = 3\mathbf{G}_D(VOR, \mathbf{V}) = 3\sigma_D^2(VOR, \mathbf{V})|det\mathbf{V}|^{2/D}\mathbf{I} \qquad \text{(nonsubtractive dithering)} \tag{5.22}$$

$\diamondsuit$

Necessary and sufficient condition for total mean square error independence. In Theorem 6, we gave the necessary and sufficient conditions for the first order moment vector and the second order moment matrix of the error to be independent from the input. One can desire to make the total mean square error, $E[\|\mathbf{e}\|^2]$ instead of the second order matrix, $E[\mathbf{e}\mathbf{e}^T]$ independent from the input. The following corollary to Theorem 6 states the necessary and sufficient condition for this weaker requirement:

Corollary 1. In the nonsubtractive quantization scheme of Fig. 7, the total mean square error is independent of the input vector, i.e. $E[\|\mathbf{e}\|^2/\mathbf{x}] = E[\|\mathbf{e}\|^2]$, if and only if $Tr(\nabla\nabla^T H(\mathbf{\Omega})$ is Nyquist-$\mathbf{U}$, that is $Tr(\nabla\nabla^T H(\mathbf{Un})) = d\ \delta(\mathbf{n})$. $\diamondsuit$

If the quantization basic cell is symmetric with respect to the origin and if the above condition holds, then $E[\|\mathbf{e}\|^2] = E[\|\mathbf{z}\|^2] + E[\|\mathbf{v}\|^2] = Tr(\mathbf{G}_D(\mathcal{P}_0, \mathbf{V})) + E[\|\mathbf{v}\|^2]$, where $\mathbf{z}$ is defined as in (5.2) and is independent of $\mathbf{v}$.

Proof. From (5.4) in Lemma 1,

$$\begin{aligned} E[\|\mathbf{e}\|^2/\mathbf{x}] &= Tr(\frac{1}{j^2}\sum_{\mathbf{n}} \nabla\nabla^T H(\mathbf{Un}) e^{j\mathbf{x}^T\mathbf{Un}}) \\ &= \frac{1}{j^2}\sum_{\mathbf{n}} Tr(\nabla\nabla^T H(\mathbf{Un})) e^{j\mathbf{x}^T\mathbf{Un}}. \end{aligned} \tag{5.23}$$

Hence, by the uniquenes of Fourier series, the necessary and sufficient condition follows. If the condition is satisfied, then $E[\|\mathbf{e}\|^2] = \frac{1}{j^2} Tr(\nabla\nabla^T H(\mathbf{0})) = Tr(E[(\mathbf{z}-\mathbf{v})(\mathbf{z}-\mathbf{v})^T])$, which leads to the result, since $\mathbf{v}$ and $\mathbf{z}$ are independent. $\blacksquare$

Generation of the dither vector for nonsubtractive case: We need a random vector that is uniform in $VOR(\mathbf{V})$ in the scheme of Example 3 to achieve minimum mean square error. Here is a simple method to generate such a vector: Obtain a dither vector $\mathbf{z}$ that is uniform in $SPD(\mathbf{V})$ using the method given in Sec. 3.1. Quantize $\mathbf{z}$ using the lattice quantizer $Q(VOR, \mathbf{V})$. Take the dither vector $\mathbf{v}$ to be the quantization error: $\mathbf{v} = \mathbf{z} - Q(\mathbf{z})$. Then $\mathbf{v}$ is uniform in $VOR(\mathbf{V})$ because of Theorem 1. More generally, one can generate a uniform random vector in any basic cell $\mathcal{P}$ of the lattice $\mathcal{L}(\mathbf{V})$ by replacing the quantizer with $Q(\mathcal{P}, \mathbf{V})$.

Remark. In subtractive dithering, any Nyquist-$\mathbf{V}$ dither produces an error that is independent of the input and uniform in the quantization basic cell. Hence the resulting mean square error is independent of the particular dither used. In nonsubtractive dithering, on the other hand, the total mean square error depends on the dither as well. In particular, the dither should be confined in as small volume as possible in order to obtain the lowest total mean square error.

6. OPTIMUM PRE AND POST FILTERING FOR LATTICE QUANTIZERS

In traditional scalar quantization schemes where a random process $x(n)$ is uniformly quantized, one assumes that the quantizer noise process $e(n)$ is WSS, white and has a power proportional to the input power. That is, $e(n)$ has a power spectral density $S_{ee}(e^{j\omega}) = c\sigma_x^2$. With these assumptions, one considers the possibility of improvement of the noise level by pre filtering the input process before quantization and post filtering it after the quantization with the inverse of the original filter (see Fig. 8). It is known [28] that the best prefilter $F(e^{j\omega})$ is given by

$$|F(e^{j\omega})|^2 = \frac{1}{\sqrt{S_{xx}(e^{j\omega})}} \tag{6.1}$$

and that the phase of $F(e^{j\omega})$ is arbitrary. This is commonly referred as *half-whitening* since the power spectral density of the output of $F(e^{j\omega})$ is $\sqrt{S_{xx}(e^{j\omega})}$, which is flatter than $S_{xx}(e^{j\omega})$ but not completely flat.

The assumptions that lead to the half-whitening solution are valid if the number of levels of the uniform quantizer is very large. However if one uses a dithered quantizer with proper choice of dither, then the assumptions are not only valid but are precisely true regardless of the bit rate. Hence the half-whitening filter is the optimum filter for a dithered quantizer. After making this elementary observation, we now ask the same question in the lattice vector quantization context: what is the optimum prefilter matrix $\mathbf{F}(e^{j\omega})$ that produces minimum total mean square error? In this section we proceed to answer this question.

Dithering of WSS vector random processes. Let $\mathbf{x}(n)$ be a WSS vector process with power spectral density matrix $\mathbf{S_{xx}}(e^{j\omega})$. Let $\mathbf{v}(n)$ be a vector process independent of $\mathbf{x}(n)$. Assume we add the two processes together and then quantize the sum at each time instant n with a lattice quantizer $Q(VOR, \mathbf{V})$. After the quantization, we can either subtract the original dither process resulting in subtractive dithering or we can leave it as it is, resulting in nonsubtractive dithering. This is a generalization of Fig. 5 and Fig. 7, with all the vectors replaced by vector random processes. First consider the subtractive case. It is not difficult to see that, if the dither process is chosen to be iid and Nyquist-$\mathbf{V}$, then the error process $\mathbf{e}(n) = \mathbf{x}(n) + \mathbf{v}(n) - Q(\mathbf{x}(n) + \mathbf{v}(\mathbf{n}))$ will be independent of $\mathbf{x}(n)$ and iid, with uniform distribution in $VOR(\mathbf{V})$. Next, for the nonsubtractive case, if the dither process is chosen to be the sum of two independent random process each of which is iid and uniform in $VOR(\mathbf{V})$, then the second moment of the error vector $\mathbf{e}$ will be independent of $\mathbf{x}(n)$. Assume that we are using the optimum lattice $\mathcal{L}(\mathbf{V})$ for the given dimension. Then from Theorem 7, we have,

$$E[\mathbf{e}(n)\mathbf{e}^T(n+k)] = \sigma_D^2(VOR, \mathbf{V})|det\mathbf{V}|^{2/D}\delta(k)\mathbf{I} \qquad \text{(subtractive dithering)} \tag{6.2}$$

$$E[\mathbf{e}(n)\mathbf{e}^T(n+k)] = 3\sigma_D^2(VOR, \mathbf{V})|det\mathbf{V}|^{2/D}\delta(k)\mathbf{I} \qquad \text{(nonsubtractive dithering)} \tag{6.3}$$

Prefiltering of dithered lattice quantization. Assume we filter $\mathbf{x}(n)$ by $\mathbf{F}(z)$ before quantization and by $\mathbf{F}^{-1}(z)$ after the quantization as shown in Fig. 9. Let $\mathbf{S_{qq}}(e^{j\omega})$ be the power spectral density of the dithered-quantizer noise process $\mathbf{q}(n) = \mathbf{u}(n) - \hat{\mathbf{u}}(n)$. Then, by (6.2) and (6.3), it follows that

$$\mathbf{S_{qq}}(e^{j\omega}) = c\mathbf{I} \tag{6.4}$$

where c depends only on lattice. To be precise, $c = \sigma_D^2(VOR, \mathbf{V})|det\mathbf{V}|^{2/D}$ in subtractive case and $c = 3\sigma_D^2(VOR, \mathbf{V})|det\mathbf{V}|^{2/D}$ in nonsubtractive case.

Assumption about the dependence of c on the input variance. Dithering analysis is valid only if the overflow is avoided. If the total bit rate is constrained to be fixed, then obviously there should be a relation between the unit volume $|det\mathbf{V}|$ of the lattice $\mathcal{L}(\mathbf{V})$ and the statistics of the input. If the bit rate is defined by the logarithm of the total number of codewords, then the support of D-dimensional pdf of the process can not be infinite. If, on the other hand, the bit rate is defined to be the entropy of the quantized process then D-dimensional pdf can have infinite support as in the cases of well-known distributions like Gaussian, Laplacian, etc. Without going into the detailed discussion of the rate-distortion analysis of dithered quantizers, we are going to assume that the constant c in (6.4) is proportional to the total variance of the quantizer input, that is $c = d\sigma_{\mathbf{u}}^2$. Hence (6.4) becomes:

$$\mathbf{S_{qq}}(e^{j\omega}) = d\sigma_{\mathbf{u}}^2\mathbf{I} \tag{6.5}$$

where $\sigma_{\mathbf{u}}^2$ is the total variance of $\mathbf{u}(n)$ in Fig. 9.

Theorem 8. *In the scheme of Fig. 9, assuming the relation (6.5), the optimum prefilter matrix that minimizes the total mean square error is given by*

$$\mathbf{F}(e^{j\omega}) = [\mathbf{\Lambda}(e^{j\omega})]^{-1/4}\mathbf{U}(e^{j\omega}) \tag{6.6}$$

where $\mathbf{\Lambda}(e^{j\omega})$ is a diagonal matrix with positive elements, and $\mathbf{U}(e^{j\omega})$ is a paraunitary matrix, i.e.
$\mathbf{U}^\dagger(e^{j\omega})\mathbf{U}(e^{j\omega}) = \mathbf{I}, \forall\omega$ [23]. The matrices $\mathbf{U}(e^{j\omega})$ and $\mathbf{\Lambda}(e^{j\omega})$ are related to the power spectral density $\mathbf{S_{xx}}(e^{j\omega})$ of $\mathbf{x}(n)$ as

$$\mathbf{S_{xx}}(e^{j\omega}) = \mathbf{U}^\dagger(e^{j\omega})\mathbf{\Lambda}(e^{j\omega})\mathbf{U}(e^{j\omega}). \tag{6.7}$$

The resulting total mean square error is

$$\sigma_{\mathbf{e}}^2 = d\Big[\int_{-\pi}^{\pi} Tr([\mathbf{\Lambda}(e^{j\omega})]^{1/2})\frac{d\omega}{2\pi}\Big]^2 \tag{6.8}$$

$\diamondsuit$

Proof. Let $\mathbf{S_1}(e^{j\omega}) = \mathbf{U}^\dagger(e^{j\omega})[\mathbf{\Lambda}(e^{j\omega})]^{1/2}\mathbf{U}(e^{j\omega})$, where $\mathbf{U}(e^{j\omega})$ and $\mathbf{\Lambda}(e^{j\omega})$ are as defined in the theorem statement. Then, $\mathbf{S_{xx}}(e^{j\omega}) = \mathbf{S_1}(e^{j\omega})\mathbf{S_1}^\dagger(e^{j\omega})$. Now,

$$\begin{aligned}\mathbf{R_{ee}}(0) = E[\mathbf{e}(n)\mathbf{e}^T(n)] &= \int_{-\pi}^{\pi} \mathbf{F}^{-1}(e^{j\omega})\mathbf{S_{qq}}(e^{j\omega})[\mathbf{F}^{-1}(e^{j\omega})]^\dagger\frac{d\omega}{2\pi} \\ &= d\int_{-\pi}^{\pi} \sigma_{\mathbf{u}}^2\mathbf{F}^{-1}(e^{j\omega})[\mathbf{F}^{-1}(e^{j\omega})]^\dagger\frac{d\omega}{2\pi}\end{aligned} \tag{6.9}$$

$$
\begin{aligned}
E[\|\mathbf{e}\|^2] &= d\sigma_{\mathbf{u}}^2 Tr\int_{-\pi}^{\pi} \mathbf{F}^{-1}(e^{j\omega})[\mathbf{F}^{-1}(e^{j\omega})]^\dagger \frac{d\omega}{2\pi} \\
&= d\ Tr\int_{-\pi}^{\pi} \mathbf{F}(e^{j\omega})\mathbf{S}_{\mathbf{xx}}(e^{j\omega})\mathbf{F}^\dagger(e^{j\omega})\frac{d\omega}{2\pi} Tr\int_{-\pi}^{\pi} \mathbf{F}^{-1}(e^{j\omega})[\mathbf{F}^{-1}(e^{j\omega})]^\dagger \frac{d\omega}{2\pi} \\
&= d\int_{-\pi}^{\pi} Tr(\mathbf{F}(e^{j\omega})\mathbf{S}_{\mathbf{xx}}(e^{j\omega})\mathbf{F}^\dagger(e^{j\omega}))\frac{d\omega}{2\pi} \int_{-\pi}^{\pi} Tr(\mathbf{F}^{-1}(e^{j\omega})[\mathbf{F}^{-1}(e^{j\omega})]^\dagger)\frac{d\omega}{2\pi} \\
&\geq d\Big[\int_{-\pi}^{\pi} \sqrt{Tr(\mathbf{F}(e^{j\omega})\mathbf{S}_{\mathbf{xx}}(e^{j\omega})\mathbf{F}^\dagger(e^{j\omega}))Tr(\mathbf{F}^{-1}(e^{j\omega})[\mathbf{F}^{-1}(e^{j\omega})]^\dagger)}\frac{d\omega}{2\pi}\Big]^2 \\
&\geq d\Big[\int_{-\pi}^{\pi} Tr(\mathbf{F}(e^{j\omega})\mathbf{S_1}(e^{j\omega})\mathbf{F}^{-1}(e^{j\omega}))\frac{d\omega}{2\pi}\Big]^2 \\
&= d\Big[\int_{-\pi}^{\pi} Tr(\mathbf{S_1}(e^{j\omega}))\frac{d\omega}{2\pi}\Big]^2 \\
&= d\Big[\int_{-\pi}^{\pi} Tr([\mathbf{\Lambda}(e^{j\omega})]^{1/2})\frac{d\omega}{2\pi}\Big]^2.
\end{aligned}
\tag{6.10}
$$

The first inequality is Cauchy-Schwarz inequality for integrals and the equality holds if and only if

$$Tr(\mathbf{F}(e^{j\omega})\mathbf{S}_{\mathbf{xx}}(e^{j\omega})\mathbf{F}^\dagger(e^{j\omega})) = k'\ Tr(\mathbf{F}^{-1}(e^{j\omega})[\mathbf{F}^{-1}(e^{j\omega})]^\dagger) \text{ for all } \omega \tag{6.11}$$

The second ineqality is another Cauchy-Schwarz inequality, applied to the following inner product space: $(\mathbf{A},\mathbf{B}) = Tr(\ \mathbf{B}^\dagger\mathbf{A})$, (see for example, [29], p. 360)

$$|Tr(\mathbf{B}^\dagger\mathbf{A})|^2 \leq Tr(\mathbf{A}^\dagger\mathbf{A})Tr(\mathbf{B}^\dagger\mathbf{B}) \tag{6.12}$$

with equality if and only if $\mathbf{A} = k\mathbf{B}$. Letting $\mathbf{A} = \mathbf{F}(e^{j\omega})\mathbf{S_1}(e^{j\omega})$ and $\mathbf{B} = [\mathbf{F}^{-1}(e^{j\omega})]^\dagger$ we have the second inequality and therefore the equality holds if and only if

$$\mathbf{F}(e^{j\omega})\mathbf{S_1}(e^{j\omega}) = k\ [\mathbf{F}^{-1}(e^{j\omega})]^\dagger \tag{6.13}$$

or equivalently,

$$[\mathbf{S_1}(e^{j\omega})]^{-1} = \mathbf{F}^\dagger(e^{j\omega})\mathbf{F}(e^{j\omega}) \tag{6.14}$$

where $\mathbf{S_1}(e^{j\omega})$ is the spectral factor of $\mathbf{S}_{\mathbf{xx}}(e^{j\omega})$, i.e. $\mathbf{S}_{\mathbf{xx}}(e^{j\omega}) = \mathbf{S_1}(e^{j\omega})\mathbf{S_1}^\dagger(e^{j\omega})$. We can choose $k = 1$ as it will not affect the final result. So, $\mathbf{F}(e^{j\omega})$ should be a spectral factor of the inverse of the spectral factor of the positive definite matrix $\mathbf{S}_{\mathbf{xx}}(e^{j\omega})$. Note that, (6.11) is satisfied automatically if $\mathbf{F}(e^{j\omega})$ is chosen as in (6.14). The filter defined given by (6.6) satisfies (6.14) as can be verified by direct substitution. Hence it is an optimal filter matrix with the resulting total mean-square error as in (6.8). When the dimension is 1, the solution reduces to the well known half-whitening filter as in (6.1). ■

Comment. The solution (6.6) can be understood in the following way: The optimum $\mathbf{F}(e^{j\omega})$ is the cascade of two systems. The first system, $\mathbf{U}(e^{j\omega})$, which is a paraunitary filter bank, decorrelates the components of the vector process $\mathbf{x}(n)$ (assuming zero-mean for simplicity). The second system, $[\mathbf{\Lambda}(e^{j\omega})]^{-1/4}$, is nothing but half-whitening of each of the decorrelated components! See Fig. 10.

Prefiltered lattice quantization of scalar WSS processes. Assume now, that the vector process $\mathbf{x}(n)$ is formed by blocking a WSS random process $x(n)$ [23] (see Fig. 11). Then one way to diagonalize the power spectral density is to use a set of ideal filters. Let $\{H_i(e^{j\omega})\}$ be a set of ideal filters that have nonoverlapping frequency supports as shown in Fig. 12. Using these filters as in Fig. 13, it can be verified that the components after the decimation in Fig. 13 are uncorrelated. It is not difficult to see that the set of half-whitening prefilters after the ideal filter bank is equivalent to one half-whitening prefilter preceding the ideal filter bank. Similarly, the set of corresponding postfilters followed by the ideal filter bank is equivalent to the ideal filter bank followed by one postfilter corresponding to the unblocked output (Fig. 14). This system can be redrawn as in Fig. 15 using the polyphase representation [23]. By construction, the polyphase matrix $\mathbf{E}(e^{j\omega})$ is paraunitary. Let $\mathbf{r}(n)$ and $\mathbf{u}(n)$ be the input and output of the system $\mathbf{E}(e^{j\omega})$. It can be shown (Appendix C of [23]) that $E[\mathbf{u}^T(n)\mathbf{u}(n)] = E[\mathbf{r}^T(n)\mathbf{r}(n)]$. The quantity $\sigma^2_{\mathbf{u}}$ in (6.5) is $E[\mathbf{u}^T(n)\mathbf{u}(n)]$ with the assumption that the processes have zero mean. Hence $\sigma^2_{\mathbf{u}}$ is unaffected by the choice of $\mathbf{E}(e^{j\omega})$. So we can eliminate $\mathbf{E}(e^{j\omega})$ and $\mathbf{E}^\dagger(e^{j\omega})$ and obtain the simplified form of Fig. 16. We have proved:

Theorem 9. *In the lattice quantization scheme of Fig. 9, if the input vector process $\mathbf{x}(n)$ is obtained by blocking a WSS scalar process $x(n)$, then the optimum prefilter $\mathbf{F}(e^{j\omega})$ is equivalent to the scalar half-whitening filter applied to the input $x(n)$, as depicted in Fig. 16.* $\diamondsuit$

Relation to the optimum subband coding problem. In subband coding systems, the channels are often quantized with one dimensional uniform quantizers. Let $u_i(n)$ be the ith subband signal and $q_i(n)$ the corresponding quantization noise. Since each of the channels is quantized separately, the total bit rate is the sum of bit rates of each channel. Let b_i be the rate assigned to the channel i. In subband coding problems, the following is assumed:

$$\sigma^2_{q_i} = c\,\sigma^2_{u_i} 2^{-2b_i} \tag{6.15}$$

This assumption is justified when the bit rate is high and the overload effect is negligible [1]. The same constant c is assumed for all channels although, in [1] it is shown that c depends on the source statistics. For the prefiltered lattice quantization scheme we assumed (6.5). Since $\sigma^2_{\mathbf{u}} = \sum_{k=1}^{D} \sigma^2_{u_k}$, this asumption implies

$$\sum_{i=1}^{D} \sigma^2_{q_i} = d \sum_{i=1}^{D} \sigma^2_{u_i} \tag{6.16}$$

Compare this with (6.15) which is traditionally used in subband coding with separate subband quantizers. (6.15) yields

$$\sum_{i=1}^{D} \sigma_{q_i}^2 = c \sum_{i=1}^{D} 2^{-2b_i} \sigma_{u_i}^2 \tag{6.17}$$

Thus the set of quantizer noise variances $\{\sigma_{q_i}^2\}$ is assumed to be related to the set of quantizer input variances $\{\sigma_{u_i}^2\}$ by (6.16) in the prefiltered lattice quantizer, and by (6.17) in the case of traditional subband coding. These two assumptions create significant difference in the formulation and solution of these two problems, which should not, therefore, be compared. In particular, the line of reasoning which allowed us to reduce Fig. 15 into the simpler form of Fig. 16 will not hold in the traditional subband coding case. As mentioned earlier, the problem of optimizing the prefilter under the subband coding constraint (6.15) is equivalent to finding the best biorthogonal subband coder for a given input and a fixed number of channels D. This is an open problem, and is outside the scope of this paper.

7. SUMMARY

In this paper we provided the error analysis of dithered and undithered lattice quantizers. In Sec. 3, we analyzed the lattice quantization system. In Sec. 4, we saw that, for any input, we can make the quantization error independent from the input and uniform in the quantization basic cell. We provided some results on the moments of the error and gave a necessary condition for a lattice to have minimum dimensionless second moment. Sec. 5 covered nonsubtractive dithering of lattice quantization and we saw that we can make the moments of the error vector independent from the input. We gave one set of dither vectors that can be used in nonsubtractive dithering to achieve the first and second order moment independence conditions. Among them, we outlined how to choose a dither vector that results in minimum total mean square error. We saw that this dither should be a sum of two independent random vectors, each uniform in $VOR(\mathbf{V})$, where $\mathbf{V}$ is the generator matrix of the optimum lattice for its dimension. We emphasized that the requirement to make the total mean square error independent from the input is weaker than the requirement to make the second moment matrix independent from the input. We provided two methods of generating Nyquist-$\mathbf{V}$ vectors, one for the dither that is uniform in $SPD(\mathbf{V})$, the other for the dither that is uniform in $VOR(\mathbf{V})$. The former was sufficient for all purposes in subtractive dithering and the latter was necessary to have minimum mean square error in nonsubtractive dithering. Finally, using the results on optimum lattices from Sec. 4, in Sec. 6, we addressed the problem of optimum linear prefiltering of dithered lattice quantizers. With the assumption that the sum of the variances of the noise vector components is proportional to the sum of the variances of the input components, we came up with a general solution. In the special case of blocking one dimensional WSS processes, we saw that our solution reduces to the scalar half-whitening filter.

APPENDIX A

The definitions of multidimensional Fourier transform, Fourier series and their interrelations are summarized here in a way most suited to our notations. Details can be found in many standard references, for example [24].

1. The MD Fourier transform of $f(\mathbf{x})$ is defined as

$$F(\mathbf{\Omega}) = \int f(\mathbf{x}) e^{-j\mathbf{\Omega}^T \mathbf{x}} d\mathbf{x} \tag{A.1}$$

We see that the characteristic function (2.7) is therefore $\Phi_{\mathbf{X}}(\mathbf{\Omega}) = F(-\mathbf{\Omega})$.

2. $f(\mathbf{x})$ is said to be periodic-$\mathbf{V}$, if $f(\mathbf{x} + \mathbf{V}\mathbf{n}) = f(\mathbf{x})$ for every $\mathbf{x} \in R^D$ and $\mathbf{n} \in Z^D$. Let $\mathcal{P}$ be a basic cell with respect to $\mathbf{V}$, and let $\mathbf{U}$ be the matrix generating the reciprocal lattice, that is, $\mathbf{U} = 2\pi\mathbf{V}^{-T}$. Then the Fourier series coefficients of $\mathbf{f}(\mathbf{x})$ are given by

$$c_{\mathbf{k}} = \frac{1}{|det\mathbf{V}|} \int_{\mathcal{P}} f(\mathbf{x}) e^{-j\mathbf{x}^T \mathbf{U}\mathbf{k}} d\mathbf{x} \tag{A.2}$$

and the Fourier series representation of $f(\mathbf{x})$ is given by

$$f(\mathbf{x}) = \sum_{\mathbf{k}} c_{\mathbf{k}} e^{j\mathbf{x}^T \mathbf{U}\mathbf{k}} \tag{A.3}$$

3. *Relation between Fourier series and Fourier transform.* Let $F(\mathbf{\Omega})$ be the FT of $f(\mathbf{x})$. Define the periodic-$\mathbf{V}$ function $g(\mathbf{x}) = \sum_{\mathbf{k}} f_{\mathbf{X}}(\mathbf{x} + \mathbf{V}\mathbf{k})$. and let $\{c_{\mathbf{k}}\}$ be its Fourier series as defined above. Then the Fourier series coefficients $\{c_{\mathbf{k}}\}$ are related to the samples of the Fourier transform, taken on the lattice generated by $\mathbf{U}$. More precisely,

$$c_{\mathbf{k}} = \frac{1}{|det\mathbf{V}|} F(\mathbf{U}\mathbf{k}) \tag{A.4}$$

Thus, the periodic function $g(\mathbf{x})$ can be expanded as

$$g(\mathbf{x}) = \frac{1}{|det\mathbf{V}|} \sum_{\mathbf{k}} F(\mathbf{U}\mathbf{k}) e^{j\mathbf{x}^T \mathbf{U}\mathbf{k}} \tag{A.5}$$

APPENDIX B

Proof of Fact 1.

$$\begin{aligned}
\mathbf{G}_D(VOR, \mathbf{V}) &= \frac{1}{|det\mathbf{V}|} \int_{VOR(\mathbf{V})} \mathbf{e}\mathbf{e}^T d\mathbf{e} \\
&= \frac{1}{|det\mathbf{V}|} \int_{SPD(\mathbf{V})} \mathbf{e}\mathbf{e}^T d\mathbf{e} \\
&= \mathbf{V} \int_{[-\frac{1}{2}, \frac{1}{2})^D} \hat{\mathbf{e}}\hat{\mathbf{e}}^T d\hat{\mathbf{e}} \mathbf{V}^T \\
&= \frac{1}{12} \mathbf{V}\mathbf{V}^T \\
&= \frac{1}{12} |det\mathbf{V}|^{2/D} \mathbf{\Lambda}
\end{aligned} \tag{7.1}$$

The reason for the second equality is that $VOR(\mathbf{V}) = SPD(\mathbf{V})$ for an orthogonal lattice. The third equality follows by a change of variable $\hat{\mathbf{e}} = \mathbf{V}^{-1}\mathbf{e}$. ■

Proof of Fact 2.

$$\begin{aligned}
\sigma_D^2(\mathcal{P}_0, \mathbf{V}) &\geq \sigma_D^2(VOR, \mathbf{V}) \quad \text{(by definition of } VOR(\mathbf{V})) \\
&= \frac{1}{D|detV|^{2/D}} Tr(\mathbf{G}_D(VOR, \mathbf{V})) \\
&= \frac{1}{12D} Tr(\mathbf{\Lambda}) \quad \text{(by equation (4.3))} \\
&= \frac{1}{12D} \sum_{i=1}^{D} \lambda_i \\
&\geq \frac{1}{12} (\prod_{i=1}^{D} \lambda_i)^{1/D} \quad \text{(arithmetic-geometric mean inequality)} \\
&= \frac{1}{12}.
\end{aligned} \tag{7.2}$$

The first inequality can be viewed as an application of the necessary condition for an optimal quantizer: the partition of the space for a given codeword should be the Voronoi partition. It is not difficult to see that no other partitioning can give a better error. Hence, equality holds if and only if $\mathcal{P}_0 = VOR(\mathbf{V})$. The other inequality is an application of arithmetic-geometric mean inequality (abbreviated as AM-GM) [23] to the positive diagonal elements λ_i. Hence, the equality holds if and only if $\lambda_i = c, \forall i$. Finally, beacuse of the definition of $\mathbf{\Lambda}$ in (4.3), $\prod_{i=1}^{D} \lambda_i = 1$, implying $c = 1$. Hence the equality holds if and only if $\mathbf{\Lambda} = \mathbf{I}$. ■

REFERENCES

[1] A. Gersho and R.M. Gray, *Vector quantization and signal compression*, Boston: Kluwer, 1992.

[2] P.A. Chou, T. Lookabaugh, and R.M. Gray, "Entropy-constrained vector quantization," IEEE Trans. Acoust. Speech Signal Processing vol. ASSP-37, pp31-42, Jan. 1989.

[3] J.H. Conway and N.J.A. Sloane, "Fast quantizing and decoding algorithms for lattice quantizers and codes," IEEE Trans. Information Theory, vol. IT-28, pp227-232, Mar. 1982.

[4] K. Sayood, J.D. Gibson, and M.C. Rost, "An algorithm for uniform vector quantizer design," IEEE Trans. Information Theory, vol. IT-30, pp805-814, Nov. 1984.

[5] J.D. Gibson and K. Sayood, "Lattice quantization," in *Advances in Electronics and Electron Physics* (P. Hawkes, Ed.) New York: Academic Press, 1988, vol. 72, ch. 3.

[6] D.G. Jeong and J.D. Gibson, "Uniform and piecewise uniform lattice vector quantization for memoryless Gaussian and Laplacian sources," IEEE Trans. Information Theory vol. IT-39, pp786-804, May 1993.

[7] ———, "Image coding with uniform and piece-wise uniform vector quantizers", IEEE Trans. Image Proc. vol. IP-4, pp140-146, Feb. 1995.

[8] L.G. Roberts, "Picture coding using pseudo-random noise," IRE Trans. Information Theory, vol. IT-8, pp. 145-154, Feb. 1962.

[9] L. Schuchman,"Dither signals and their effect on quantization noise", IEEE Trans. Comm. Techn. Dec. 1964.

[10] A.B. Sripad and D.L. Snyder,"A necessary and sufficient condition for quantization errors to be uniform and white", IEEE Trans. Acoust. Speech Signal Processing vol. ASSP-25, pp442-448, Oct. 1977.

[11] S.P. Lipshitz, R.A. Wannamaker, and J. Vanderkooy, "Quantization and dither - a theoretical survey," J. Audio Eng. Soc., vol. 40, pp355-375, May 1992.

[12] R.M. Gray and T.G. Stockham, "Dithered Quantizers," IEEE Trans. Information Theory, vol. IT-39, pp805-812, May 1993.

[13] J. Ziv, "On universal quantization," IEEE Trans. Information Theory, vol. IT-31, pp344-347, May 1985.

[14] R. Zamir and M. Feder, "On universal qunatization by randomized uniform/lattice quantizers", IEEE Trans. Information Theory, vol. IT-38, pp428-436, Mar. 1992.

[15] ———, "On lattice quantization noise," in Proc. Data Compression Conf., DCC-94, pp380-389, Snowbird, UT, Mar. 1994.

[16] ———, "Rate-distortion performance in coding bandlimited sources by sampling and dithered quantization," IEEE Trans. Information Theory, vol. IT-41, Jan. 1995.

[17] ———, "Information rates of pre/post filtered dithered quantizers," IEEE Trans. Information Theory, to appear.

[18] T. Linder and K. Zeger, "Asymptotic entropy-constrained performance of tessellating and universal randomized lattice quantization", IEEE Trans. Information Theory, vol. IT-40, pp575-579, Mar. 1994.

[19] J.H. Conway and N.J.A. Sloane, "Voronoi regions of lattices, second moments of polytopes, and quantization," IEEE Trans. Information Theory vol. IT-28, pp211-226, Mar. 1982.

[20] H.S.M. Coxeter, *Introduction to geometry*, New York: Wiley, 1969.

[21] J.H. Conway and N.J. Sloane, *Sphere packings, lattices and groups*, New York: Springer-Verlag, 1988.

[22] T.M. Cover and J.A. Thomas, *Elements of information theory*, New York: John Wiley & Sons, 1991.

[23] P.P. Vaidyanathan, *Multirate systems and filter banks*, Englewood Cliffs, NJ: Prentice Hall, 1993.

[24] D.E. Dudgeon and R.M. Mersereau, *Multidimensional digital signal processing*, New Jersey: Prentice Hall, 1984.

[25] A. Gersho, "Asymptotically optimal block quantization," IEEE Trans. Information Theory vol. IT-25, pp373-380, Jul. 1979.

[26] D.J. Newman, "The hexagon theorem," IEEE Trans. Information Theory, vol. IT-28, pp137-139, Mar. 1982.

[27] E.S. Barnes and N.J.A. Sloane, "The optimal lattice quantizer in three dimensions," SIAM J. Algebraic Discrete Methods vol. 4, pp30-41, Mar. 1983.

[28] N.S. Jayant and P. Noll, *Digital coding of waveforms : principles and applications to speech and video*, Englewood Cliffs, NJ: Prentice-Hall, 1984.

[29] S.H. Friedberg, A.J. Insel, and L. E. Spence, *Linear Algebra*, Englewood Cliffs, NJ: Prentice-Hall, 1979.

FOOTNOTES

1. Manuscript received
2. This work was supported by NSF grant MIP 92-15785, funds from Tektronix, Inc., and Rockwell, Intl.
3. The authors are with the Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125.

LIST OF FIGURES

Fig. 1. Demonstration of the perceptual advantages of dithered lattice quantizers.

(a) Original image of Lenna, 512x512, 8 bits/pixel.

(b) Output of lattice quantization with dimension 24, bit rate = .4 bits/pixel.

(c) Error of the lattice quantization.

(d) Output of the same lattice quantization with subtractive dithering, bit rate is about the same.

(e) Error of the dithered lattice quantization.

Fig. 2. Lattice example in $2-D$. The heavy dots are the points on the lattice.

Fig. 3. (a) Voronoi regions for the lattice in Fig. 2. The shaded region is $VOR(\mathbf{V})$.

(b) SPD regions for the lattice in Fig. 2. The shaded region is $SPD(\mathbf{V})$.

Fig. 4. Error in lattice quantization.

(a) An error vector $\mathbf{e}$, and an input vector $\mathbf{x}$ that produces it.

(b) A different input vector producing the same error vector.

Fig. 5. Subtractively dithered lattice quantizer.

Fig. 6. Transformation of a lattice quantizer. Here, $\mathbf{e} = \mathbf{x} - \hat{\mathbf{x}}$ is uniform in $VOR(\mathbf{V})$. This can be assured by subtractive dithering. Hence, $\mathbf{f} = \mathbf{z} - \hat{\mathbf{z}}$ is uniform in a basic cell of the transformed lattice.

Fig. 7. Nonsubtractively dithered lattice quantizer.

Fig. 8. Pre and postfiltering of a scalar process. Q denotes a uniform scalar quantizer. The optimum choice of the filter is the half-whitening solution.

Fig. 9. Pre and postfiltering of a vector process in conjunction with a dithered lattice quantizer. The lattice $\mathcal{L}(\mathbf{V})$ is the optimum lattice for its dimension.

Fig. 10. Optimum pre and postfiltering in lattice quantization. $\mathbf{U}(e^{j\omega})$ is the decorrelator filter matrix and the filters $F_1, F_2, ..., F_D$ are the half-whitening filters for their inputs.

Fig. 11. The vector process $\mathbf{x}(n)$, obtained by blocking a scalar WSS process.

Fig. 12. A set of ideal filters to be used as the decorrelating paraunitary system.

Fig. 13. The ideal filter bank of Fig. 12 is used as the decorrelating system.

Fig. 14. Half-whitening filters reduce to a single half-whitening filter.

Fig. 15. Redrawing the system in Fig. 14 using the polyphase decomposition of the ideal filter bank.

Fig. 16. The final simplified form of the system when the input is the blocked version of a scalar WSS process.