

# A Computable Fourier Condition Generating Alias-Free Sampling Lattices

Yue M. Lu, *Member, IEEE*, Minh N. Do, *Senior Member, IEEE*, and Richard S. Laugesen

**Abstract**—We propose a Fourier analytical condition linking alias-free sampling with the Fourier transform of the indicator function defined on the given frequency support. Our discussions center around how to develop practical computation algorithms based on the proposed analytical condition. We address several issues along this line, including the derivation of simple closed-form expressions for the Fourier transforms of the indicator functions defined on arbitrary polygonal and polyhedral domains; a complete and nonredundant enumeration of all quantized sampling lattices via the Hermite normal forms of integer matrices; and a quantitative analysis of the approximation of the original infinite Fourier condition by using finite computations. Combining these results, we propose a computational testing procedure that can efficiently search for the optimal alias-free sampling lattices for a given polygonal or polyhedral shaped frequency domain. Several examples are presented to show the potential of the proposed algorithm in multidimensional filter bank design, as well as in applications involving the design of efficient sampling patterns for multidimensional bandlimited signals.

**Index Terms**—densest sampling, critical sampling, packing, tiling, maximal decimation, optimal sampling, nonredundant filter banks, Fourier transforms of indicator functions, Poisson summation formula, divergence theorem.

## I. INTRODUCTION

The classical Whittaker-Shannon-Kotelnikov sampling theorem [1], [2] states that a one-dimensional bandlimited signal can be exactly reconstructed from its uniform samples if the sampling rate is beyond the Nyquist rate. The situation is similar in multidimensional cases [3]. In general, the effect of the uniform sampling process in the frequency domain is that the spectrum of the original bandlimited signal gets replicated over a lattice whose density is inversely proportional to the sampling density. If the shifted copies of the spectrum do not overlap with the baseband, then we have an *alias-free sampling*; consequently, the original signal can be reconstructed from its sampled version by applying an ideal interpolation filter whose passband is supported on the baseband.

Copyright (c) 2008 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to [pubs-permissions@ieee.org](mailto:pubs-permissions@ieee.org).

Y. M. Lu was with the Department of Electrical and Computer Engineering and the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana IL 61801. He is now with the Audio-Visual Communications Laboratory, Swiss Federal Institute of Technology Lausanne (EPFL), Switzerland (e-mail: [yue.lu@epfl.ch](mailto:yue.lu@epfl.ch); WWW: <http://lcav.epfl.ch/~lu>).

M. N. Do is with the Department of Electrical and Computer Engineering, the Coordinated Science Laboratory, and the Beckman Institute, University of Illinois at Urbana-Champaign, Urbana IL 61801 (e-mail: [minhdo@illinois.edu](mailto:minhdo@illinois.edu); WWW: <http://www.ifp.uiuc.edu/~minhdo>).

R. S. Laugesen is with the Department of Mathematics, University of Illinois at Urbana-Champaign, Urbana, IL 61801 (e-mail: [Laugesen@illinois.edu](mailto:Laugesen@illinois.edu); WWW: <http://www.math.uiuc.edu/~laugesen>).

This work was supported by the US National Science Foundation under Grant CCR-0237633 (CAREER).

The search for alias-free sampling lattices for a given frequency support, and in particular for those lattices achieving minimum sampling densities, is a fundamental issue in various signal processing applications that involve the design of efficient acquisition schemes for bandlimited signals. Some examples include image and video processing [4], volume sampling in computer graphics [5], Fourier imaging, and the distributed sensing of various physical phenomena [6]–[8]. As a special case of alias-free sampling, the concept of *critical sampling* also plays an important role in the theory and design of critically sampled (a.k.a. maximally decimated) multidimensional filter banks [9]–[11].

The study of optimal sampling lattices is a classical problem [3], [12]. Earlier efforts often focus on cases when the signals are bandlimited to spherical regions (the sphere packing problem [13]), or to some particular regions relevant to certain target applications (e.g. [7], [8]). It remains a challenging open problem as to whether one can find a general and systematic approach determining the optimal alias-free sampling lattices for an arbitrary frequency support in multiple dimensions. On a broader scale, alias-free sampling is mathematically equivalent to the lattice packing of a given domain, for which lots of studies can be found in disciplines such as computational geometry and operational research. So far, most practical algorithms proposed for densest lattice packing (e.g. [14]–[16]) approach the problem from a geometrical perspective. The primary tools employed are the theories from Minkowski’s work [13], as well as various geometrical intuitions and heuristics obtained for particular domains in lower dimensions.

Instead of adopting the usual geometrical viewpoint, we propose in this paper a Fourier analytical approach to the problem of alias-free sampling (*i.e.* packing). Central to this approach is a novel condition linking the alias-free sampling with the Fourier transform of the indicator function defined on the underlying frequency support (see Theorem 1). An important feature of the proposed condition is that it opens the door to purely analytical and computational solutions to our sampling lattice selection problem. Compared with geometry-based methods, the proposed analytical approach can be potentially advantageous in situations when the underlying frequency regions have complicated nonconvex shapes, or in higher dimensions, where it is increasingly difficult to invoke geometrical intuitions.

A conceptually similar idea of using the indicator functions to study lattice tiling was proposed by Kolountzakis and Lagarias [17], [18]. Our work can be viewed as an extension of this early mathematical work to the more general lattice packing case, with specific engineering applications in signal sampling. Our extensions, however, provide fundamental new

results in the following directions.

First, the practical applicability of the proposed Fourier condition depends on whether one can easily calculate the Fourier transforms of the indicator functions. Therefore, we study the computation of these Fourier transforms in Section IV. Using the divergence theorem, we derive simple closed-form expressions for situations when the frequency support regions are arbitrary polygonal and polyhedral domains in 2-D and 3-D (see Propositions 3 and 4), either convex or nonconvex. The proposed results can also be generalized to higher dimensional cases involving  $N$ -D polytopes.

Next, we discuss the quantization of sampling lattices in Section V. To systematically investigate all possible sampling geometries, we describe and employ the Hermite normal forms [19], [20] for integer matrices, which provide a complete characterization of quantized sampling lattices. We present a new estimate for the total number of different quantized lattices of a given density, which can be used to predict the size of the search space.

We demonstrate the proposed techniques in two possible signal processing applications. In Section VI, we present a simple algorithm that can efficiently determine if a given polytope-shaped frequency partitioning allows for critical sampling. This algorithm can be useful in the design of multidimensional multirate systems, when one wants to check whether a certain frequency partitioning can be implemented by a critically sampled (*i.e.* nonredundant) filter bank. In Section VII, we propose an algorithm that, when given a polytope-shaped frequency domain, can search for the optimal alias-free sampling lattices among all quantized lattices at a given quantization scale. Several examples are given to show the potential of the proposed algorithms. We conclude the paper in Section VIII.

## II. PRELIMINARIES

**Notation:** Throughout the paper,  $N$  represents the dimension of the signals. Bold face italic letters represent  $N$ -by- $N$  matrices (upper case) and  $N$ -by-1 vectors (lower case). We denote by

$$|\mathbf{M}| \stackrel{\text{def}}{=} |\det \mathbf{M}|$$

the absolute value of the determinant of a matrix  $\mathbf{M}$ . The Fourier transform of a function  $f(\boldsymbol{\omega})$  defined on  $\mathbb{R}^N$  is defined by

$$\widehat{f}(\mathbf{x}) = \int_{\mathbb{R}^N} f(\boldsymbol{\omega}) e^{-2\pi j \mathbf{x} \cdot \boldsymbol{\omega}} d\boldsymbol{\omega}, \quad (1)$$

where  $\mathbf{x}, \boldsymbol{\omega} \in \mathbb{R}^N$  are the vectors of spatial and frequency variables, respectively.<sup>1</sup> Calligraphic letters, such as  $\mathcal{D}$ , represent bounded and open frequency domains in  $\mathbb{R}^N$ , with  $m(\mathcal{D})$  denoting the Lebesgue measure (*i.e.* volume) of  $\mathcal{D}$ . Given a nonsingular matrix  $\mathbf{M}$  and a vector  $\boldsymbol{\tau}$ , we use  $\mathbf{M}(\mathcal{D} + \boldsymbol{\tau})$  to represent the set of points of the form  $\mathbf{M}(\boldsymbol{\omega} + \boldsymbol{\tau})$  for  $\boldsymbol{\omega} \in \mathcal{D}$ . Finally, we denote by  $\mathbb{1}_{\mathcal{D}}(\boldsymbol{\omega})$  the indicator function

<sup>1</sup>Strictly speaking, to obtain the spatial domain function  $\widehat{f}(\mathbf{x})$  from the frequency domain function  $f(\boldsymbol{\omega})$ , one should apply the *inverse* Fourier transform in (1). We choose to use the *forward* Fourier transform here, mainly for notational conveniences in later parts of this paper. Mathematically, the two choices are equivalent up to a sign change.

of the domain  $\mathcal{D}$ , *i.e.*,  $\mathbb{1}_{\mathcal{D}}(\boldsymbol{\omega}) = 1$  if  $\boldsymbol{\omega} \in \mathcal{D}$  and  $\mathbb{1}_{\mathcal{D}}(\boldsymbol{\omega}) = 0$  otherwise.

### A. Multidimensional Sampling on Lattices

In multidimensional multirate signal processing, the sampling operations are usually defined on lattices, each of which can be generated by an  $N \times N$  nonsingular matrix  $\mathbf{M}$  as

$$\Lambda_{\mathbf{M}} \stackrel{\text{def}}{=} \{\mathbf{M}\mathbf{n} : \mathbf{n} \in \mathbb{Z}^N\}. \quad (2)$$

We denote by  $\Lambda_{\mathbf{M}}^*$  the corresponding reciprocal lattice (a.k.a. polar lattice), defined as

$$\Lambda_{\mathbf{M}}^* \stackrel{\text{def}}{=} \{\mathbf{M}^{-T}\boldsymbol{\ell} : \boldsymbol{\ell} \in \mathbb{Z}^N\} \quad (3)$$

In the rest of the paper, when it is clear from the context what the generating matrix is, we will drop the subscripts in  $\Lambda_{\mathbf{M}}$  and  $\Lambda_{\mathbf{M}}^*$ , and use  $\Lambda$  and  $\Lambda^*$  for simplicity.

For an  $\mathbf{M}$ -fold sampling, the input continuous signal  $f(\mathbf{x})$  and the output discrete signal  $s[\mathbf{n}]$  are related by  $s[\mathbf{n}] = f(\mathbf{M}\mathbf{n})$ . Suppose  $f(\mathbf{x})$  is bandlimited, and its frequency region of support is a bounded open set  $\mathcal{D} \subset \mathbb{R}^N$ . Then the discrete-time Fourier transform of the samples  $s[\mathbf{n}]$ , defined as  $S(\boldsymbol{\omega}) = \sum_{\mathbf{n} \in \mathbb{Z}^N} s[\mathbf{n}] e^{-2\pi j \boldsymbol{\omega} \cdot \mathbf{n}}$ , is supported in [9]–[11]

$$S = \mathbf{M}^T \left( \bigcup_{\mathbf{k} \in \Lambda^*} (\mathcal{D} + \mathbf{k}) \right). \quad (4)$$

In words, the frequency support  $S$  of the discrete samples can be obtained by first taking the union of the baseband  $\mathcal{D}$  and all of its shifted copies (*i.e.* aliasing components), and then applying a linear mapping  $\mathbf{M}^T$ .

### B. Alias-Free Sampling and Critical Sampling

For appropriately chosen sampling lattices, the aliasing components in (4) do not overlap with the baseband frequency support  $\mathcal{D}$ . In this important case, we can fully recover the original continuous signal  $f(\mathbf{x})$  by applying an ideal interpolation filter spectrally supported on  $\mathcal{D}$  to the discrete samples  $s[\mathbf{n}]$ .

*Definition 1:* We say a frequency support  $\mathcal{D}$  allows an *alias-free*  $\mathbf{M}$ -fold sampling, if different shifted copies of  $\mathcal{D}$  in (4) are disjoint, *i.e.*,

$$\mathcal{D} \cap (\mathcal{D} + \mathbf{k}) = \emptyset \quad \text{for all } \mathbf{k} \in \Lambda^* \setminus \{\mathbf{0}\}. \quad (5)$$

Furthermore, we say  $\mathcal{D}$  can be *critically sampled* by  $\mathbf{M}$ , if in addition to the alias-free condition in (5), the union of the shifted copies also covers the entire spectrum, *i.e.*,

$$\bigcup_{\mathbf{k} \in \Lambda^*} (\mathcal{D} + \mathbf{k}) = \mathbb{R}^N, \quad \text{up to a set of measure zero.}^2 \quad (6)$$

The density of a sampling lattice  $\Lambda_{\mathbf{M}}$  is defined as  $\rho_{\mathbf{M}} = 1/|\mathbf{M}|$ , which is the number of samples retained per unit volume. Note that for a given frequency support  $\mathcal{D}$ , there exist infinitely many lattices that can achieve alias-free sampling for  $\mathcal{D}$ ; thus, it is possible to reduce the sampling density, while still remain alias-free, through a judicious choice of the sampling

<sup>2</sup>This technicality is due to the assumption that the frequency region of support  $\mathcal{D}$  is an open set.

geometry. We call those alias-free sampling lattices achieving minimum sampling density the optimal sampling lattices. Our main focus in this paper is to propose a Fourier analytical condition and a corresponding computational procedure that can systematically search for the optimal sampling lattices for a given  $\mathcal{D}$ .

Before proceeding, we make some further remarks on the scope of this paper. First, we restrict our attention to the scenario in which the continuous signals are sampled on a *single* lattice. Consequently, the minimum sampling rate we pursue is the Nyquist rate, which is achieved by those lattices that can pack the frequency support  $\mathcal{D}$  in the tightest way. We note that it is possible to go below the Nyquist rate if we can sample the continuous signals with multiple channels and on multiple lattices (see, *e.g.*, [21], [22]). This more general sampling setup is left for future work. Second, in various Fourier imaging modalities (such as MRI), the measurements we obtain are samples of the Fourier transform of some object. The goal, therefore, is to reconstruct the spatial domain function from its Fourier samples, which is the exact dual of the sampling setup considered in this paper. Therefore, if the spatial support of the object is known, our results on optimal sampling lattices can be directly applied to this problem.

### III. ALIAS-FREE SAMPLING USING FOURIER TECHNIQUES

In this section, we study the problems of alias-free sampling and critical sampling with Fourier techniques. The key observation is a link between the alias-free sampling condition and the Fourier transform of the indicator function  $\mathbb{1}_{\mathcal{D}}(\omega)$  defined on the frequency support  $\mathcal{D}$ .

#### A. A Fourier Analytical Condition for Alias-Free Sampling

As the first step to linking the alias-free sampling problem with Fourier analysis, we consider the autocorrelation function

$$R_{\mathcal{D}}(\omega) = \int \mathbb{1}_{\mathcal{D}}(\tau) \mathbb{1}_{\mathcal{D}}(\tau - \omega) d\tau,$$

and define

$$A_{\mathcal{D},M} \stackrel{\text{def}}{=} \sum_{\mathbf{k} \in \Lambda^* \setminus \{\mathbf{0}\}} R_{\mathcal{D}}(\mathbf{k}). \quad (7)$$

Since  $R_{\mathcal{D}}(\omega)$  measures the volume of intersection between  $\mathcal{D}$  and  $\mathcal{D} + \omega$ , *i.e.*,  $R_{\mathcal{D}}(\omega) = m(\mathcal{D} \cap (\mathcal{D} + \omega))$ , the quantity  $A_{\mathcal{D},M}$  defined above can be interpreted as the total volume of overlapping regions between the original baseband support  $\mathcal{D}$  and all its aliasing components in (4).

*Lemma 1:* A frequency region  $\mathcal{D}$  allows an  $M$ -fold alias-free sampling *if and only if*  $A_{\mathcal{D},M} = 0$ .

*Proof:* By construction,  $R_{\mathcal{D}}(\omega) \geq 0$  for all  $\omega$ ; thus  $A_{\mathcal{D},M} = 0$  if and only if the volume of intersection  $R_{\mathcal{D}}(\mathbf{k}) = 0$  for all  $\mathbf{k} \in \Lambda^* \setminus \{\mathbf{0}\}$ , which is the same as having the alias-free condition given in (5). ■

*Theorem 1:* A frequency region  $\mathcal{D}$  allows an  $M$ -fold alias-free sampling *if and only if*

$$|M| \sum_{\mathbf{n} \in \Lambda} |\widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{n})|^2 = m(\mathcal{D}), \quad (8)$$

where  $\widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{x})$  is the Fourier transform of  $\mathbb{1}_{\mathcal{D}}(\omega)$ .

*Proof:* From the definition of  $R_{\mathcal{D}}(\omega)$ , its Fourier transform is  $\widehat{R}_{\mathcal{D}}(\mathbf{x}) = |\widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{x})|^2$ . Applying the Poisson summation formula (see Appendix A for a justification for the pointwise equality), we have

$$\sum_{\mathbf{k} \in \Lambda^*} R_{\mathcal{D}}(\mathbf{k}) = |M| \sum_{\mathbf{n} \in \Lambda} \widehat{R}_{\mathcal{D}}(\mathbf{n}) = |M| \sum_{\mathbf{n} \in \Lambda} |\widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{n})|^2.$$

The overlapping term  $A_{\mathcal{D},M}$  defined in (7) can then be calculated as

$$\begin{aligned} A_{\mathcal{D},M} &= \sum_{\mathbf{k} \in \Lambda^*} R_{\mathcal{D}}(\mathbf{k}) - R_{\mathcal{D}}(\mathbf{0}) \\ &= |M| \sum_{\mathbf{n} \in \Lambda} |\widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{n})|^2 - m(\mathcal{D}), \end{aligned} \quad (9)$$

where in reaching the second equality we have also used the fact that  $R_{\mathcal{D}}(\mathbf{0}) = m(\mathcal{D})$ . By applying Lemma 1, we are done. ■

*Remark:* The “only if” direction of Theorem 1 can also be *formally* established as the consequence of a standard condition for orthogonal functions (see, for example, [23, p. 132]), which states that a set of functions  $\{f(\omega - \mathbf{k})\}_{\mathbf{k} \in \Lambda^*}$  are mutually orthogonal in  $L^2(\mathbb{R}^N)$  if and only if

$$|M| \sum_{\mathbf{n} \in \Lambda} |\widehat{f}(\mathbf{x} + \mathbf{n})|^2 = \|f\|_{L^2}^2 \quad \text{for almost every } \mathbf{x}. \quad (10)$$

The alias-free sampling condition defined in (5) means that  $\{\mathbb{1}_{\mathcal{D}}(\omega - \mathbf{k})\}_{\mathbf{k} \in \Lambda^*}$  constitutes an orthogonal set of functions in  $L^2(\mathbb{R}^N)$ . Therefore, by applying the characterization (10) to  $\mathbb{1}_{\mathcal{D}}(\omega)$ , we get the formula (8) of Theorem 1. A catch in this derivation though is that (10) holds only for *almost* every  $\mathbf{x}$ , whereas in Theorem 1 we want to be able to evaluate the sum at a specific point  $\mathbf{x} = \mathbf{0}$ . We justify the validity of the pointwise equality in Appendix A.

From Theorem 1, we can obtain an equivalent testing condition for alias-free sampling as described below.

*Proposition 1:* A frequency region  $\mathcal{D}$  allows an  $M$ -fold alias-free sampling *if and only if*

$$|M| \sum_{\mathbf{n} \in \Lambda, \|\mathbf{n}\|_{\infty} \leq r} |\widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{n})|^2 \leq m(\mathcal{D}), \quad \text{for all } r > 0, \quad (11)$$

where  $\|\mathbf{n}\|_{\infty} \stackrel{\text{def}}{=} \max(|n_1|, \dots, |n_N|)$  is the infinity norm.

*Proof:* The necessity of (11) follows immediately from Theorem 1. For the sufficiency, assume (11) holds. Then the left hand side of (8) is less than or equal to the right hand side (just by letting  $r \rightarrow +\infty$ ). On the other hand, because by construction  $A_{\mathcal{D},M}$  is always nonnegative, we can obtain from (9) that  $|M| \sum_{\mathbf{n} \in \Lambda} |\widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{n})|^2 \geq m(\mathcal{D})$ , and thus the equality holds in (8). It then follows from Theorem 1 that  $\mathcal{D}$  allows for an  $M$ -fold alias-free sampling. ■

Compared with the *infinite* sum in Theorem 1, the new condition in Proposition 1 is often more useful for practical computational procedures. This is because, for any given  $r > 0$ , the condition (11) only involves a *finite* sum of  $|\widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{x})|^2$ , which can be easily computed and serves as a necessary condition for alias-free sampling. We will come back to this point in more details in Sections VI and VII.

As an immediate application of Proposition 1, we can get the following well-known lower bound on sampling density,

which is usually proved by geometrical arguments in the literature.

*Corollary 1 (Lower bound on sampling density):* If  $\mathcal{D}$  allows an  $M$ -fold alias-free sampling, then the sampling density  $\rho_M$  satisfies

$$\rho_M = \frac{1}{|M|} \geq m(\mathcal{D}). \quad (12)$$

*Proof:* From (11), we have

$$\frac{m(\mathcal{D})}{|M|} \geq |\widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{0})|^2 = \left( \int_{\mathbb{R}^N} \mathbb{1}_{\mathcal{D}}(\boldsymbol{\omega}) d\boldsymbol{\omega} \right)^2 = m(\mathcal{D})^2,$$

and hence (12).  $\blacksquare$

### B. A Fourier Analytical Condition for Critical Sampling

Here we focus on the special case of critical sampling, and begin by mentioning, without proof, a standard result:

*Lemma 2:* A frequency support  $\mathcal{D}$  can be critically sampled by a sampling matrix  $M$  if and only if  $M$  is an alias-free sampling matrix for  $\mathcal{D}$  with sampling density  $1/|M| = m(\mathcal{D})$ .

Combining Theorem 1 and Lemma 2, we can obtain a simpler set of conditions for the special case of critical sampling as follows.

*Proposition 2:* A frequency support  $\mathcal{D}$  can be critically sampled by a matrix  $M$  if and only if

$$\begin{aligned} \widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{0}) &= m(\mathcal{D}) = \frac{1}{|M|} \quad \text{and} \\ \widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{n}) &= 0 \quad \text{for all } \mathbf{n} \in \Lambda \setminus \{\mathbf{0}\}. \end{aligned} \quad (13)$$

*Proof:* Suppose (13) holds. Then it follows that

$$\sum_{\mathbf{n} \in \Lambda} |\widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{n})|^2 = |\widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{0})|^2 = \frac{m(\mathcal{D})}{|M|},$$

and hence from Theorem 1,  $M$  is an alias-free sampling matrix for  $\mathcal{D}$ . Meanwhile, since  $m(\mathcal{D}) = \frac{1}{|M|}$ , we can apply Lemma 2 to conclude that  $\mathcal{D}$  is critically sampled by  $M$ . By reversing the above line of reasoning, we can also show the necessity of (13).  $\blacksquare$

*Remark:* The result of Proposition 2 is previously known in various disciplines. In approximation theory, the condition (13) is often called the interpolation property, which is satisfied by the Fourier transform of the indicator function defined on any fundamental cell of a lattice (see, for example, [12], [24]). The usefulness of this condition in the context of lattice tiling was first pointed out by Kolountzakis and Lagarias [17] and applied to investigate the tiling of various high dimensional shapes [18].

*Example 1:* Figure 1(a) shows the ideal frequency partitioning of a three-level directional filter bank [25], [26], which divides the 2-D discrete spectrum  $[-\frac{1}{2}, \frac{1}{2}]^2$  into 8 wedge-shaped subbands. How do we verify that each subband can be critically sampled, but without using geometry (*i.e.* drawing figures)? According to Proposition 2, for each subband  $\mathcal{D}$ , we just need to calculate  $\widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{x})$ , and try to find a matrix  $M$  such that (13) holds. For instance, consider the subband marked as the dark region in Figure 1(a). We can work out an expression for  $\widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{x})$  as follows

$$\widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{x}) = \frac{1 - \cos(\pi x_2 - \frac{\pi x_1}{2})}{\pi^2 x_1 (2x_2 - x_1)} - \frac{1 - \cos(\pi x_2 - \pi x_1)}{2\pi^2 x_1 (x_2 - x_1)}. \quad (14)$$

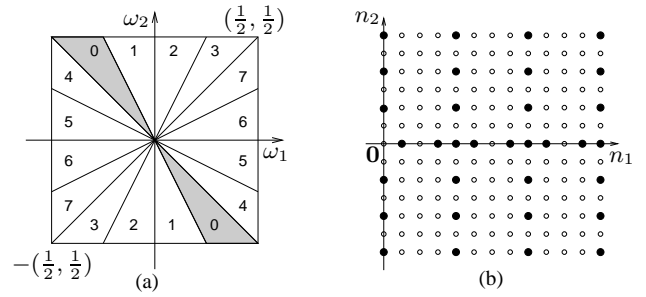


Fig. 1. (a) The ideal frequency partitioning of a three-level directional filter bank. (b) The values of  $\widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{x})$  evaluated at integer points, with  $\mathcal{D}$  being the dark region shown in (a). Black dots represent zero values and white dots represent nonzero values.

Evaluating  $\widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{x})$  at integer points  $\mathbf{n} = (n_1, n_2)$  for  $0 \leq n_1 \leq 12$  and  $|n_2| \leq 6$ , we show the results in Figure 1(b), where each dot represents an integer point, and black dots indicate those locations where  $\widehat{\mathbb{1}}_{\mathcal{D}}(n_1, n_2) = 0$ . We can observe that the zero set (black dots) contains, as a subset,  $\{M\mathbf{n} : 0 \leq n_1 \leq 3, |n_2| \leq 3\} \setminus \{\mathbf{0}\}$ , where  $M = \text{diag}(4, 2)$  is a diagonal matrix whose determinant is equal to  $1/m(\mathcal{D})$ . Substituting  $x_1 = 4n_1, x_2 = 2n_2$  into (14), we can verify that  $\widehat{\mathbb{1}}_{\mathcal{D}}(M\mathbf{n}) = \widehat{\mathbb{1}}_{\mathcal{D}}(4n_1, 2n_2) = 0$  for all integer vectors  $\mathbf{n} \neq \mathbf{0}$ . Therefore, the condition (13) in Proposition 2 is satisfied, and hence  $\mathcal{D}$  can indeed be critically sampled by  $M$ .

We can see from the above example that Proposition 2 provides a purely computational way to determine if a given frequency support allows for critical sampling, and if yes, what the sampling matrices are. However, to develop practical computational algorithms based on this result (as well as on Proposition 1 for alias-free sampling), we must overcome three issues as listed below.

- 1) As a prerequisite to using the Fourier conditions in (11) and (13), we must know the expression for  $\widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{x})$ . This can be a cumbersome task if we need to do the derivation by hand for each given  $\mathcal{D}$ . We address this problem in the next section by presenting simple closed-form expressions for  $\widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{x})$  when  $\mathcal{D}$  are arbitrary polygonal and polyhedral domains.
- 2) To systematically and efficiently search for the optimal sampling lattices, we study in Section V a *complete* and *nonredundant* enumeration of all possible sampling geometries.
- 3) In practical implementations, we can only compute the values of  $\widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{x})$  at a finite number of lattice points, whereas the Fourier conditions in (11) and (13) involve an infinite number of points. We present in Section VII-B a quantitative analysis on the effect of the approximation due to finite computation, for the case of 2-D polygonal domains. The result of the analysis is useful in determining how many points we need to compute to achieve a given precision.

## IV. THE FOURIER TRANSFORMS OF INDICATOR FUNCTIONS

### A. The Divergence Theorem

We consider the computation of  $\widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{x})$  in this section. Suppose the domain  $\mathcal{D}$  has a piecewise smooth boundary, and

let  $\mathbf{F}(\boldsymbol{\omega}) = (F_1(\boldsymbol{\omega}), F_2(\boldsymbol{\omega}), \dots, F_N(\boldsymbol{\omega}))$  be a continuously differentiable vector field defined on  $\mathcal{D}$ . The divergence theorem states that

$$\int_{\mathcal{D}} \operatorname{div} \mathbf{F}(\boldsymbol{\omega}) d\boldsymbol{\omega} = \int_{\partial\mathcal{D}} \mathbf{F}(\boldsymbol{\omega}) \cdot \mathbf{n}_{\boldsymbol{\omega}} dS(\boldsymbol{\omega}), \quad (15)$$

where  $\operatorname{div} \mathbf{F}(\boldsymbol{\omega}) \stackrel{\text{def}}{=} \sum_{i=1}^N \frac{\partial F_i(\boldsymbol{\omega})}{\partial \omega_i}$  is the divergence of  $\mathbf{F}(\boldsymbol{\omega})$ , and  $\partial\mathcal{D}$  is the boundary of  $\mathcal{D}$  oriented by outward unit normal vectors  $\mathbf{n}_{\boldsymbol{\omega}}$ .

In this work, we choose the vector-valued function to be  $\mathbf{F}_{\mathbf{x}}(\boldsymbol{\omega}) = \frac{j\mathbf{x}}{2\pi\|\mathbf{x}\|^2} e^{-2\pi j\mathbf{x}\cdot\boldsymbol{\omega}}$ , where  $\mathbf{x} \neq \mathbf{0}$  is a fixed vector parameter. We compute

$$\begin{aligned} \operatorname{div} \mathbf{F}_{\mathbf{x}}(\boldsymbol{\omega}) &= \sum_{i=1}^N \frac{\partial}{\partial \omega_i} \left( \frac{j x_i}{2\pi\|\mathbf{x}\|^2} e^{-2\pi j \sum_k x_k \omega_k} \right) \\ &= e^{-2\pi j \mathbf{x}\cdot\boldsymbol{\omega}} \sum_{i=1}^N \frac{x_i^2}{\|\mathbf{x}\|^2} = e^{-2\pi j \mathbf{x}\cdot\boldsymbol{\omega}}, \end{aligned}$$

which is exactly the exponential term used in the Fourier transform. The divergence theorem (15) then leads to a formula of Randol [27], namely that

$$\begin{aligned} \widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{x}) &\stackrel{\text{def}}{=} \int_{\mathcal{D}} e^{-2\pi j \mathbf{x}\cdot\boldsymbol{\omega}} d\boldsymbol{\omega} = \int_{\mathcal{D}} \operatorname{div} \mathbf{F}_{\mathbf{x}}(\boldsymbol{\omega}) d\boldsymbol{\omega} \\ &= \frac{j}{2\pi\|\mathbf{x}\|^2} \int_{\partial\mathcal{D}} e^{-2\pi j \mathbf{x}\cdot\boldsymbol{\omega}} (\mathbf{x} \cdot \mathbf{n}_{\boldsymbol{\omega}}) dS(\boldsymbol{\omega}), \quad (16) \end{aligned}$$

for all  $\mathbf{x} \neq \mathbf{0}$ . In other words, we can simplify the computation of  $\widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{x})$  from an integration on the entire volume  $\mathcal{D}$  to one on the boundary surface  $\partial\mathcal{D}$ .

### B. Fourier Transforms of Indicator Functions on Polygonal Domains

In various practical applications, the domains of frequency supports of the underlying signals can often be modeled as, or at least approximated by, polygons in 2-D, polyhedra in 3-D, and in general, polytopes in  $N$ -D. For these types of domain, the boundary  $\partial\mathcal{D}$  assumes a particularly simple shape (*e.g.* the boundary of a polygon is just a finite union of line segments); consequently, we can work out the integral in (16) and obtain closed-form expressions for  $\widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{x})$ . To see this, we start with the following proposition for the 2-D polygonal case. A similar result was first given in [28].

*Proposition 3:* Suppose  $\mathcal{D}$  is a polygon with  $K$  sides, whose vertices, when traced clockwise, are  $\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_K$ . Also, let  $\mathbf{p}_{K+1} \stackrel{\text{def}}{=} \mathbf{p}_1$ . Then  $\widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{0}) = m(\mathcal{D})$ ; and for  $\mathbf{x} \neq \mathbf{0}$ ,

$$\begin{aligned} \widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{x}) &= \frac{j}{2\pi\|\mathbf{x}\|^2} \sum_{k=1}^K d_k (\mathbf{x} \cdot \mathbf{n}_k) \\ &\quad \operatorname{sinc}(\mathbf{x} \cdot (\mathbf{p}_{k+1} - \mathbf{p}_k)) e^{-2\pi j \mathbf{x}\cdot\mathbf{c}_k}, \end{aligned} \quad (17)$$

where  $\operatorname{sinc}(x) \stackrel{\text{def}}{=} \sin(\pi x)/(\pi x)$  is the normalized sinc function;  $d_k = \|\mathbf{p}_{k+1} - \mathbf{p}_k\|$  is the length,  $\mathbf{n}_k$  is the outward unit normal vector, and  $\mathbf{c}_k = (\mathbf{p}_k + \mathbf{p}_{k+1})/2$  is the center, of the  $k$ th side.

*Proof:* When  $\mathbf{x} = \mathbf{0}$ , we have  $\widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{0}) = \int \mathbb{1}_{\mathcal{D}}(\boldsymbol{\omega}) e^{-2\pi j \mathbf{0}\cdot\boldsymbol{\omega}} d\boldsymbol{\omega} = m(\mathcal{D})$ . In what follows, we consider the case when  $\mathbf{x} \neq \mathbf{0}$ . The boundary  $\partial\mathcal{D}$  consists

of  $K$  line segments, denoted by  $B_1, B_2, \dots, B_K$ . We let  $\mathbf{t}_k = (\mathbf{p}_{k+1} - \mathbf{p}_k)/d_k$  be the unit tangent vector of  $B_k$ . We apply (16) to get

$$\begin{aligned} \widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{x}) &= \frac{j}{2\pi\|\mathbf{x}\|^2} \sum_{k=1}^K (\mathbf{x} \cdot \mathbf{n}_k) \int_{B_k} e^{-2\pi j \mathbf{x}\cdot\boldsymbol{\omega}} dS(\boldsymbol{\omega}) \\ &= \frac{j}{2\pi\|\mathbf{x}\|^2} \sum_{k=1}^K (\mathbf{x} \cdot \mathbf{n}_k) \int_{-d_k/2}^{d_k/2} e^{-2\pi j \mathbf{x}\cdot(\mathbf{c}_k + \mathbf{t}_k v)} dv \\ &= \frac{j}{2\pi\|\mathbf{x}\|^2} \sum_{k=1}^K (\mathbf{x} \cdot \mathbf{n}_k) \frac{\sin(\pi d_k \mathbf{x} \cdot \mathbf{t}_k)}{\pi \mathbf{x} \cdot \mathbf{t}_k} e^{-2\pi j \mathbf{x}\cdot\mathbf{c}_k}, \end{aligned}$$

which is equivalent to (17) after replacing  $\mathbf{t}_k$  with its corresponding expression.  $\blacksquare$

Many applications deal with real-valued signals, for which the domains of frequency supports are always symmetric with respect to the origin. We can easily verify the following corollary, which employs the symmetry of the domain and calculates  $\widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{x})$  using only half of the vertices.

*Corollary 2:* Suppose the domain  $\mathcal{D}$  is a disjoint union of a polygon  $\mathcal{D}_1$  and its symmetric copy  $(-\mathcal{D}_1)$ , *i.e.*,  $\mathcal{D} = \mathcal{D}_1 \cup -\mathcal{D}_1$  and  $\mathcal{D}_1 \cap -\mathcal{D}_1 = \emptyset$ . Meanwhile, suppose  $\mathcal{D}_1$  has  $K$  sides whose vertices are as in Proposition 3. Then for all  $\mathbf{x} \neq \mathbf{0}$ ,

$$\begin{aligned} \widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{x}) &= \frac{1}{\pi\|\mathbf{x}\|^2} \sum_{k=1}^K d_k (\mathbf{x} \cdot \mathbf{n}_k) \\ &\quad \operatorname{sinc}(\mathbf{x} \cdot (\mathbf{p}_{k+1} - \mathbf{p}_k)) \operatorname{sinc}(\pi \mathbf{x} \cdot (\mathbf{p}_{k+1} + \mathbf{p}_k)), \end{aligned} \quad (18)$$

where  $d_k$  and  $\mathbf{n}_k$  are the length and outward unit normal vector of the  $k$ th side, respectively.

*Example 2:* Let  $\mathcal{D}$  be the wedge-shaped frequency support shown in Figure 2. It is a centrally symmetric domain, whose upper half is a triangle with vertices  $\mathbf{p}_1 = (-\frac{1}{2}, \frac{1}{2})^T$ ,  $\mathbf{p}_2 = (-\frac{1}{4}, \frac{1}{2})^T$ , and  $\mathbf{p}_3$  at the origin. In Example 1, we have shown an expression for  $\widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{x})$  derived by hand. Here, we achieve the same goal by applying Corollary 2. From the coordinates of the vertices, we can calculate the length and unit normal vector on each of the three sides as  $d_1 = \frac{1}{4}$ ,  $d_2 = \frac{\sqrt{5}}{4}$ ,  $d_3 = \frac{\sqrt{2}}{2}$ ; and  $\mathbf{n}_1 = (0, 1)^T$ ,  $\mathbf{n}_2 = (\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}})^T$ ,  $\mathbf{n}_3 = (-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2})^T$ . Inserting these parameters into (18) leads to

$$\begin{aligned} \pi\|\mathbf{x}\|^2 \widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{x}) &= \frac{x_2}{4} \operatorname{sinc} \frac{\pi(4x_2 - 3x_1)}{4} \operatorname{sinc} \frac{x_1}{4} \\ &\quad + \frac{2x_1 + x_2}{4} \operatorname{sinc} \frac{\pi(2x_2 - x_1)}{4} \operatorname{sinc} \frac{2x_2 - x_1}{4} \\ &\quad + \frac{x_1 + x_2}{2} \operatorname{sinc} \frac{\pi(x_1 - x_2)}{2} \operatorname{sinc} \frac{x_1 - x_2}{2}. \end{aligned} \quad (19)$$

After some manipulation, we can verify that (19) is equivalent to the expression given in (14).

### C. Higher Dimensional Cases: Polyhedron and Polytope Domains

We can generalize the previous results to higher dimensional cases when the domain  $\mathcal{D}$  is a polyhedron in 3-D or a polytope in  $N$ -D. For example, when  $\mathcal{D}$  is a polyhedron, its boundary is a finite union of 2-D polygons, the integration on which has been solved by Proposition 3. Thus, by combining (16) and

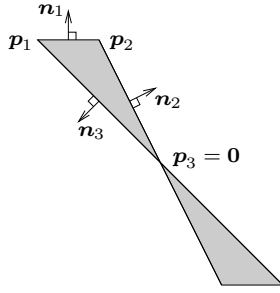


Fig. 2. A centrally symmetric frequency domain  $\mathcal{D}$ , whose vertices are  $\mathbf{p}_1 = (-\frac{1}{2}, \frac{1}{2})^T$ ,  $\mathbf{p}_2 = (-\frac{1}{4}, \frac{1}{2})^T$ , and  $\mathbf{p}_3 = \mathbf{0}$ .

(17) in Proposition 3, we can get the following result, whose proof is omitted.

**Proposition 4:** Suppose  $\mathcal{D}$  is a polyhedron with  $K$  facets. The  $k$ th facet is an  $L_k$ -sided polygon, whose vertices, when traced clockwise, are denoted by  $\mathbf{p}_1^k, \mathbf{p}_2^k, \dots, \mathbf{p}_{L_k}^k$ . Then  $\widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{0}) = m(\mathcal{D})$ ; and for  $\mathbf{x} \neq \mathbf{0}$ ,

$$\widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{x}) = \frac{-1}{\|2\pi\mathbf{x}\|^2} \sum_{k=1}^K \frac{\mathbf{x} \cdot \mathbf{n}_k}{\|\mathbf{x}\|^2 - |\mathbf{x} \cdot \mathbf{n}_k|^2} \sum_{\ell=1}^{L_k} (\mathbf{x} \cdot \mathbf{v}_{\ell}^k) \operatorname{sinc}(\mathbf{x} \cdot (\mathbf{p}_{\ell+1}^k - \mathbf{p}_{\ell}^k)) e^{-2\pi j \mathbf{x} \cdot \mathbf{c}_{\ell}^k}, \quad (20)$$

where  $\mathbf{n}_k$  is the outward unit normal vector of the  $k$ th facet,  $\mathbf{v}_{\ell}^k = \mathbf{n}_k \times (\mathbf{p}_{\ell+1}^k - \mathbf{p}_{\ell}^k)$ , and  $\mathbf{c}_{\ell}^k = (\mathbf{p}_{\ell}^k + \mathbf{p}_{\ell+1}^k)/2$  is the center of the  $\ell$ th side on the  $k$ th facet.

**Remark:** In general, the boundary facets of an  $N$ -D polytope are just a finite union of  $(N-1)$ -D polytopes; hence, one can calculate  $\widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{x})$  on arbitrary  $N$ -D polytope domains by *recursively* applying (16) and the previous results for the  $(N-1)$ -D case. Due to space limitations, we omit here further details of this generalization.

## V. ENUMERATION OF QUANTIZED SAMPLING LATTICES

Another important ingredient of the proposed Fourier conditions in Section III is the sampling lattice. In this section, we discuss the characterization and enumeration of all (quantized) sampling lattices, and the results will be used to define the search space of the computational procedures proposed in the following sections.

### A. Quantization of Lattices

In principle, the sampling lattices can be generated by any nonsingular matrix  $\mathbf{M}$  with *real-valued* entries. In this paper, however, we focus our attention to matrices of the following form:

$$\mathbf{M} = \frac{\mathbf{P}}{q}, \quad (21)$$

where  $\mathbf{P}$  is an *integer-valued* matrix, and  $q \geq 1$  is a pre-determined quantization scale. In other words, we have quantized the entries of  $\mathbf{M}$  to be  $p_{i,j}/q$ , with  $p_{i,j} \in \mathbb{Z}$ .

Notice that, by setting the quantization scale  $q = 1$ , the above model includes discrete signal processing as a special case, where we only use integer sampling lattices. For continuous signal processing, the quantization in (21) is still a reasonable simplification, since any real-valued sampling

matrix can be well-approximated by the quantized matrices in (21) when  $q$  is large enough. Meanwhile, due to constraints in hardware precision, the sampling locations in practical systems are often quantized anyway. As we will see in the discussions below, the main advantage of considering this quantized model is to reduce the search space of sampling lattices to a finite set.

### B. The Hermite Normal Form and Characterization of Lattices

Given a frequency support  $\mathcal{D}$ , we know from Corollary 1 that all suitable matrices  $\mathbf{M}$  for alias-free sampling of  $\mathcal{D}$  must have  $|\mathbf{M}| \leq 1/m(\mathcal{D})$ . Correspondingly, the integer matrix  $\mathbf{P}$  in (21) must satisfy

$$|\mathbf{P}| \in \left\{ 1, 2, 3, \dots, \left\lfloor \frac{q^N}{m(\mathcal{D})} \right\rfloor \right\}, \quad (22)$$

where  $\lfloor \alpha \rfloor$  is the largest integer less than or equal to a real number  $\alpha$ . Similarly, for critical sampling, we know from Proposition 2 that

$$|\mathbf{P}| = \frac{q^N}{m(\mathcal{D})}. \quad (23)$$

Although infinitely many integer matrices satisfy (22) or (23), we only need to check a finite number of them, as there exist only a finite number of *geometrically distinct* lattices generated by these matrices [20]. To see this, the first step of our discussions is the following well-known result (see, *e.g.*, [4]):

**Proposition 5:** A frequency support  $\mathcal{D}$  allows an  $M_1$ -fold alias-free sampling, if and only if it allows an  $M_2$ -fold alias-free sampling, where  $M_2 = M_1 \mathbf{U}$  and  $\mathbf{U}$  is an arbitrary unimodular matrix (*i.e.* an integer matrix with  $|\mathbf{U}| = 1$ ). The same is true for critical sampling.

The above statements can be easily verified by noting that the matrices  $\mathbf{M}_1$  and  $\mathbf{M}_1 \mathbf{U}$  always generate the same sampling lattices. To apply this result, we define the following binary relation between two matrices: we say  $\mathbf{M}_1 \sim \mathbf{M}_2$  if  $\mathbf{M}_2 = \mathbf{M}_1 \mathbf{U}$  for some unimodular matrix  $\mathbf{U}$ . We can verify that  $\mathbf{M}_1 \sim \mathbf{M}_2$  is an equivalence relation, and therefore, the corresponding equivalence classes form a partition of the set of all matrices whose determinants have the same absolute value. Proposition 5 implies that, for each equivalence class, we just need to pick a representative matrix from that class, and check if it forms an alias-free sampling (or critical sampling). The result then carries over to all members in that equivalence class. The following theorem [19], [20] provides a convenient class of representative matrices.

**Theorem 2 (Hermite normal form):** For every matrix of the form  $\mathbf{M}_1 = \mathbf{P}/q$  with  $\mathbf{P}$  being an integer matrix, there is a *unique* matrix  $\mathbf{M}_2 = \mathbf{H}/q$ , such that  $\mathbf{M}_1 \sim \mathbf{M}_2$  and that  $\mathbf{H} = [h_{i,j}]_{1 \leq i, j \leq N}$  is an integer matrix having the Hermite normal form, which means

- 1)  $\mathbf{H}$  is upper triangular;
- 2)  $h_{N,N} > 0$  and  $0 \leq h_{i,j} < h_{i,i}$  for  $1 \leq i < j \leq N$ .

It follows from Theorem 2 that we only need to consider sampling matrices of the form  $\mathbf{M} = \mathbf{H}/q$ , where  $\mathbf{H}$  is a Hermite normal matrix. Such matrices provide a complete and nonredundant enumeration of all possible sampling lattices at a

quantization level  $q$ . The following procedure, first introduced to the sampling literature by Cortelazzo and Manduchi [20], can be used to list the generating matrices of all quantized sampling lattices (*i.e.* sublattices of  $\frac{1}{q}\mathbb{Z}^N$ ) at a given density.

*Procedure 1 (Enumeration of Hermite normal matrices):*

Let the determinant  $\delta$  be a positive integer.

- 1) Form a set  $V = \{(d_1, d_2, \dots, d_N) : \prod_{n=1}^N d_n = \delta\}$ , where  $d_n (1 \leq n \leq N)$  are positive integers. Note that  $V$  is always a finite set.
- 2) For each vector  $(d_1, d_2, \dots, d_N)$  from  $V$ , put the corresponding elements  $d_1, d_2, \dots, d_N$  at the main diagonal of an  $N$ -by- $N$  matrix. Select the off-diagonal elements of the matrix according to the rule that  $h_{i,j} = 0$  for  $1 \leq j < i \leq N$  and  $0 \leq h_{i,j} < h_{i,i}$  for  $1 \leq i < j \leq N$ . There are  $\prod_{i=1}^N d_i^{N-i}$  different choices.

For example, when  $N = 2$  and the determinant is equal to 3, there are a total of four such Hermite normal matrices:

$$\begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}, \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 3 & 1 \\ 0 & 1 \end{pmatrix}, \text{ and } \begin{pmatrix} 3 & 2 \\ 0 & 1 \end{pmatrix}.$$

### C. Size of the Search Space

In the following sections, we will discuss algorithms that examine all possible Hermite normal matrices satisfying (22) and test each of them for alias-free sampling or critical sampling. It is therefore important to estimate the size of the search space, which depends on the number of distinct Hermite normal matrices. In what follows, we use  $H_N(\delta)$  to denote the total number of  $N$ -by- $N$  Hermite normal matrices with determinant equal to a positive integer  $\delta$ .

In 2-D,  $H_2(\delta)$  equals the ‘‘divisor function’’ of  $\delta$  (often written as  $\sigma(\delta)$  [29]), which is defined as the sum of all positive divisors of  $\delta$ . To see this, we observe that all 2-by-2 Hermite matrices have the form  $\begin{pmatrix} a & c \\ 0 & b \end{pmatrix}$ , where  $a, b$  are positive integers with  $ab = \delta$  and  $0 \leq c < a$ . That is, both  $a$  and  $b$  are divisors of  $\delta$ . Given the divisor  $a$ , we determine  $b$  from  $b = \delta/a$ , and observe that the number of choices for  $c$  equals exactly  $a$ . Summing over all possible divisors  $a$  gives the total number of Hermite matrices as  $\sum_{a|\delta} a = \sigma(\delta)$ . For large  $\delta$ , it can be shown that the growth rate of  $\sigma(\delta)$  (*i.e.* the total number of 2-by-2 Hermite matrices of determinant  $\delta$ ) is asymptotic to  $e^\gamma \delta \ln \ln \delta$  [29], where  $\gamma \approx 0.577$  is the Euler constant.

For the general  $N$ -D case, we obtain the following estimates on the total number of Hermite normal matrices, whose proof is given in Appendix B.

*Proposition 6:* For  $N \geq 2$ , we have

$$1 \leq \frac{H_N(\delta)}{\delta^{N-1}} < 22 \delta^{0.001}. \quad (24)$$

*Remark:* The term  $\delta^{0.001}$  grows very slowly — it is less than  $10^{0.1} \approx 1.259$  for  $\delta < 10^{100}$  — and can thus be treated as a small constant from a practical point of view. Consequently, the estimates in (24) imply that  $H_N(\delta)$  is essentially equal to  $C_\delta \delta^{N-1}$  for some bounded quantity  $C_\delta$ . Meanwhile, the constant factor 22 and the exponent 0.001 in (24) are not the only choices. The exponent can be chosen arbitrarily close

to 0; however the corresponding constant must necessarily be increased towards infinity.

Proposition 6 can be used to predict the size of the search space of sampling matrices. Let  $F_N(q)$  denote the total number of Hermite normal matrices satisfying (22); then it follows from (24) that

$$\begin{aligned} F_N(q) &\approx \mathcal{O}\left(1 + 2^{N-1} + 3^{N-1} + \dots + \left(\frac{q^N}{m(\mathcal{D})}\right)^{N-1}\right) \\ &= \mathcal{O}\left(q^{N^2}\right). \end{aligned} \quad (25)$$

We can see that  $F_N(q)$  grows rapidly as the quantization scale  $q$  increases. However, since the search for optimal sampling lattices is usually a design optimization problem that needs to be carried out only once for each frequency support shape,  $F_N(q)$  can still be a feasible number for off-line computation, especially when we work in lower dimensions ( $N = 2, 3$ ) and choose reasonably sized quantization scales. Moreover, as we will see later, not all candidate sampling matrices in the search space require the same amount of effort to check for their suitability for alias-free sampling. In fact, the majority of the candidate sampling matrices can be eliminated fairly quickly with only a small amount of computation. We leave the details to Section VII.

## VI. APPLICATIONS IN NONREDUNDANT FILTER BANKS

We dedicate the remaining part of this paper to two applications of the proposed theoretical results. In this section we focus on the special case of critical sampling, and describe its application in the design of multidimensional critically sampled filter banks. We then deal with the more general problem of identifying optimal sampling lattices in Section VII.

### A. Frequency Partitioning of Critically Sampled Filter Banks

Consider a general multidimensional filter bank, where each channel contains a subband filter and a sampling operator. As an important step in filter bank design, we need to specify the ideal passband support of each subband filter, all of which form a partitioning of the frequency spectrum. For example, as shown in Figure 1(a), a 2-D directional filter bank partitions the spectrum into wedge-shaped subbands.

Not every possible frequency partitioning can be used for filter bank implementation though. In particular, if we want to have a nonredundant filter bank, then the ideal passband support of each subband filter must be critically sampled by the sampling matrix in that channel. Consequently, whenever given a possible frequency partitioning, we must first perform a ‘‘reality check’’ of seeing whether the above condition is met, before proceeding to actual filter design.<sup>3</sup>

The critical sampling condition is commonly verified geometrically (*i.e.* by drawing figures). Although intuitive and straightforward, this geometrical approach becomes cumbersome when the shape of the passband support is complicated,

<sup>3</sup>In practice, suitable passband supports must also satisfy some additional ‘‘permissibility’’ conditions [30], which take into account the nonideal frequency responses of realistic filters. These further conditions are beyond the scope of this paper.

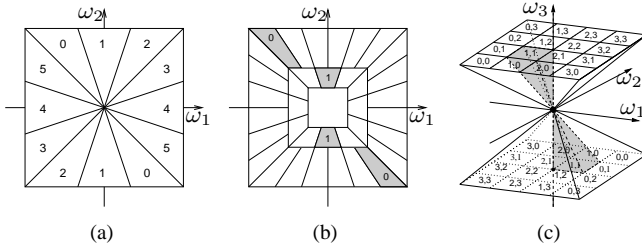


Fig. 3. The ideal frequency partitioning of several filter banks. (a) An alternative directional filter bank which decomposes the frequency cell  $(-\frac{1}{2}, \frac{1}{2}]^2$  into 6 subbands. (b) A directional multiresolution frequency partitioning. (c) A 3-D directional frequency decomposition with pyramid-shaped passband supports.

or when we work in 3-D and higher dimensional cases where it is difficult to invoke geometrical intuition. Applying the result of Proposition 2 and generalizing the ideas presented in Example 1, we propose in the following a computational procedure, which can systematically check and determine the critical sampling matrices of a given polytope region. Notice that the algorithm only searches among integer matrices, since the filter banks considered here operate on discrete-time signals.

*Procedure 2 (Critical Sampling):* Let  $\mathcal{D}$  be a given polytope-shaped frequency support region.

- 1) Calculate  $\delta = 1/m(\mathcal{D})$ . From (13), any matrix  $M$  that can critically-sample  $\mathcal{D}$  must satisfy  $|M| = \delta$ . If  $\delta$  is not an integer, then stop the procedure, since in this case it is impossible for  $\mathcal{D}$  to be critically sampled by any integer matrix.
- 2) According to the specification of  $\mathcal{D}$ , construct the explicit formula for  $\hat{\mathbb{1}}_{\mathcal{D}}(\mathbf{x})$  by using the results in Section IV.
- 3) Apply Procedure 1 to construct a list of all Hermite normal matrices with determinant equal to  $\delta$ .
- 4) For every matrix  $M$  in the above list, test the following condition

$$\hat{\mathbb{1}}_{\mathcal{D}}(M\mathbf{n}) = 0 \text{ for all } \mathbf{n} \in \mathbb{Z}^N \setminus \{\mathbf{0}\} \text{ with } \|\mathbf{n}\|_{\infty} \leq r, \quad (26)$$

where  $r$  is a large positive integer.

- 5) Present all the matrices in the list that satisfy (26). If there is no such matrix, then  $\mathcal{D}$  cannot be critically sampled by any integer matrix.

To be clear, the expression (26) is a necessary condition for  $\mathcal{D}$  to be critically sampled by  $M$ . It is not sufficient since we only check for integer points within a finite radius  $r$ , and so in principle, even if  $M$  satisfies (26) for all  $\|\mathbf{n}\|_{\infty} \leq r$ , it might happen that  $\hat{\mathbb{1}}_{\mathcal{D}}(M\mathbf{n}) \neq 0$  for some  $\mathbf{n}$  with  $\|\mathbf{n}\|_{\infty} > r$ . However, by choosing  $r$  sufficiently large, we can gain confidence in the validity of the original infinite condition (13) as required in Proposition 2. We leave the quantitative analysis of this approximation due to finite computation to Section VII. In the following examples, we choose  $r = 10000$ , and assume that the issue of numerical precision in evaluating the equality (26) is negligible.

*Example 3:* Figure 3(a) presents an alternative way to decompose the frequency spectrum into directional subbands:

instead of having 8 wedge-shaped subbands as in Figure 1(a), this new directional filter bank (DFB) generates 6 subbands. Applying the algorithm in Procedure 2, we can easily verify that this new frequency decomposition can also be critically sampled. The corresponding sampling matrices, denoted by  $M_k$  for the  $k$ th subband, are

$$M_0 = M_1 = M_2 = \begin{pmatrix} 6 & 3 \\ 0 & 1 \end{pmatrix}$$

and

$$M_3 = M_4 = M_5 = \begin{pmatrix} 2 & 1 \\ 0 & 3 \end{pmatrix}.$$

*Example 4:* We show in Figure 3(b) a directional and multiresolution decomposition of the 2-D frequency spectrum [31], [32]. Applying Procedure 2 confirms that such a frequency partitioning can be critically sampled as well. The sampling matrices for two representative subbands (marked as dark regions in the figure) are

$$M_0 = \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix} \text{ and } M_1 = \begin{pmatrix} 8 & 4 \\ 0 & 4 \end{pmatrix}.$$

*Example 5:* Figure 3(c) shows an extension of the original 2-D DFB to the 3-D case [33], where the ideal passbands of the component filters are rectangular-based pyramids radiating out from the origin at different orientations and tiling the entire frequency space. Unlike the original 2-D DFB, which is a nonredundant filter bank, the 3-D DFB constructed in [33] is an oversampled filter bank. A natural question then becomes: can the 3-D frequency partitioning shown in Figure 3(c) be critically sampled, *i.e.*, allow for a nonredundant filter bank implementation? Applying Procedure 2, we find that the answer is negative; in other words, redundancy is unavoidable for a 3-D DFB.

### B. Critical Sampling of General Cone-Shaped Frequency Regions in Higher Dimensions

The result in Example 5 can be generalized to higher dimensions, and to cases where the subbands take different directional shapes. As an application of the Fourier condition in Proposition 2, we show here a much more general statement: it is impossible to implement any cone-shaped frequency partitioning by a nonredundant filter bank, except for the 2-D case. Though this result seems to be expected [34] in the filter bank community, to our knowledge, the following is the first rigorous proof in the literature.

Since we will be discussing general  $N$ -D cases, it is necessary to introduce some algebraic notations. We consider ideal subband supports of the following “truncated-cone” shape:

$$\mathcal{D} = \{(\omega_1, \dots, \omega_N) : a \leq |\omega_N| \leq b, (\omega_1, \dots, \omega_{N-1}) \in \omega_N \mathcal{B}\}. \quad (27)$$

Geometrically,  $\mathcal{D}$  takes the form of a two-sided cone in  $\mathbb{R}^N$ , truncated by hyperplanes  $|\omega_N| = a$  and  $|\omega_N| = b$ , where  $0 \leq a < b$ . The “base” region  $\mathcal{B}$  in (27) is the intersection between the cone and the hyperplane  $\omega_N = 1$ .

The formulation in (27) is flexible enough to characterize, up to a rotation, any directional subband shown in Figure 3. For example, to describe subband-0 in Figure 3(a), we can



set  $a = 0, b = \frac{1}{2}$ , and let the base  $\mathcal{B} = [-1, -\frac{1}{3}]$ ; by choosing  $a = \frac{1}{8}, b = \frac{1}{4}$ , and  $\mathcal{B} = [-\frac{1}{3}, \frac{1}{3}]$ , we obtain subband-1 in Figure 3(b); similarly, the 3-D pyramid-shaped subband (1,1) in Figure 3(c) can be presented by  $a = 0, b = \frac{1}{2}$ , and  $\mathcal{B} = [-\frac{1}{2}, 0]^2$ . However, the class of frequency shapes that can be described by (27) is far beyond those mentioned above, since the formulation (27) allows for arbitrary configuration of the cross sections heights  $a$  and  $b$  (not necessarily the dyadic decomposition as in Figure 3(b)) and arbitrary shape for the base  $\mathcal{B}$  (not necessarily lines or squares).

*Lemma 3:* If a frequency support  $\mathcal{D}$  can be critically sampled by an integer matrix  $\mathbf{M}$ , then

$$\widehat{\mathbb{1}}_{\mathcal{D}}(|\mathbf{M}| \mathbf{n}) = 0, \quad \text{for all } \mathbf{n} \in \mathbb{Z}^N \setminus \{0\}. \quad (28)$$

*Proof:* We just need to show that, for any integer matrix  $\mathbf{M}$ , the vector  $|\mathbf{M}| \mathbf{n}$  belongs to the lattice  $\Lambda$  generated by  $\mathbf{M}$ . Once we have established this claim, then (28) becomes a natural consequence of (13) in Proposition 2. By definition, verifying  $|\mathbf{M}| \mathbf{n} \in \Lambda$  is equivalent to checking that  $\mathbf{m} = \mathbf{M}^{-1}(|\mathbf{M}| \mathbf{n})$  is an integer vector. By writing  $\mathbf{M}^{-1} = \text{adj}(\mathbf{M})/|\mathbf{M}|$ , where  $\text{adj}(\mathbf{M})$  is the adjugate matrix of  $\mathbf{M}$ , we get  $\mathbf{m} = \text{adj}(\mathbf{M}) \mathbf{n}$ . Since both  $\text{adj}(\mathbf{M})$  and  $\mathbf{n}$  are integer-valued,  $\mathbf{m}$  is indeed an integer vector. ■

*Theorem 3:* For arbitrary choice of  $0 \leq a < b$  and the base shape  $\mathcal{B}$ , the frequency domain support  $\mathcal{D}$  given in (27) cannot be critically sampled by any integer matrix, except for the 2-D case.

*Proof:* We have already established the positive result for 2-D in Examples 1, 3, and 4, where several 2-D cone-shaped regions have been shown to allow for critical sampling. Next, we show the negative result in higher dimensional cases by using contradiction. Suppose for  $N \geq 3$ , and for some particular choices of  $0 \leq a < b$  and  $\mathcal{B}$ , the corresponding frequency region  $\mathcal{D}$  in (27) can be critically sampled by an integer matrix  $\mathbf{M}$ . It follows from (28) in Lemma 3 that

$$\widehat{\mathbb{1}}_{\mathcal{D}}(0, \dots, 0, |\mathbf{M}| \mathbf{n}) = 0, \quad \text{for all } \mathbf{n} \in \mathbb{Z} \setminus \{0\}. \quad (29)$$

From the definition of  $\mathcal{D}$ , we have

$$\begin{aligned} & \widehat{\mathbb{1}}_{\mathcal{D}}(0, \dots, 0, x) \\ &= \int_{a \leq |\omega_N| \leq b} d\omega_N \left( e^{-2\pi j x \omega_N} \int_{\omega_N \mathcal{B}} 1 d\omega_1 \dots d\omega_{N-1} \right) \\ &= \int_{a \leq |\omega| \leq b} e^{-2\pi j x \omega} m(\omega \mathcal{B}) d\omega \\ &= \int_{a \leq |\omega| \leq b} e^{-2\pi j x \omega} |\omega|^{N-1} m(\mathcal{B}) \\ &= 2 m(\mathcal{B}) \int_a^b \omega^{N-1} \cos(2\pi x \omega) d\omega. \end{aligned}$$

Using the above formula and after a change of variable, we can rewrite (29) as  $\int_{2\pi|\mathbf{M}|a}^{2\pi|\mathbf{M}|b} \omega^{N-1} \cos(n\omega) d\omega = 0$ , for all  $n \in \mathbb{Z} \setminus \{0\}$ , which is impossible when  $N \geq 3$  by Appendix C. ■

## VII. FINDING OPTIMAL ALIAS-FREE SAMPLING LATTICES

### A. Algorithm

In this section, we extend Procedure 2 in the previous section for critical sampling to the more general case of

identifying optimal alias-free sampling lattices for a given frequency support. Based on the Fourier condition presented in Proposition 1, we propose the following algorithm.

*Procedure 3 (Optimal alias-free sampling lattices):* Let  $\mathcal{D}$  be a polytope-shaped frequency support region, and  $q$  a fixed quantization scale.

- 1) Since all suitable sampling lattices must satisfy (22), we start with the largest possible determinant, *i.e.*, set  $\delta = \lfloor q^N / m(\mathcal{D}) \rfloor$ .
- 2) Use Procedure 1 to construct a list of all Hermite normal matrices with determinant equal to  $\delta$ .
- 3) For every Hermite normal matrix  $\mathbf{H}$  in the above list, let  $\mathbf{M} = \mathbf{H}/q$  and do the following:

- a) Run a sequence of tests

$$|\mathbf{M}| \sum_{\mathbf{n} \in \mathbb{Z}^N, \|\mathbf{n}\|_{\infty} \leq r} |\widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{M}\mathbf{n})|^2 \leq m(\mathcal{D}), \quad (30)$$

where the integer-valued radius  $r$  increases from 1 to  $r_{\max}$ . Note that  $r_{\max}$  represents the maximum search radius, whose value will be determined in Section VII-B.

- b) If condition (30) holds for all  $r$  up to  $r_{\max}$ , then record the current matrix  $\mathbf{M}$  as a suitable sampling matrix; otherwise, proceed to the next Hermite normal matrix as soon as we reach a radius  $r$  at which (30) fails.

- 4) If there exist previously recorded suitable sampling matrices, present all these matrices and stop the procedure. Otherwise, set  $\delta \leftarrow \delta - 1$ .
- 5) If  $\delta \geq 1$ , then return to Step 2); otherwise, stop the procedure, in which case the frequency region  $\mathcal{D}$  does not allow alias-free sampling by any quantized matrices at the given quantization scale  $q$ .

In practical implementations, the computational efficiency of the above algorithm can be improved in the following two ways. First, since  $\mathbb{1}_{\mathcal{D}}(\omega)$  is a real-valued function, we have  $|\widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{x})| = |\widehat{\mathbb{1}}_{\mathcal{D}}(-\mathbf{x})|$ . Using this symmetry, we can save about half of the computations in checking the conditions in (30). Second, we can see that most of the running time of the above algorithm is spent on evaluating  $|\widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{M}\mathbf{n})|^2$  for various choices of  $\mathbf{M}$  and  $\mathbf{n}$ . Therefore, to reduce repetitive computations, we can pre-calculate and save the values of  $|\widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{n}/q)|^2$  for all integer vectors  $\mathbf{n}$  within a given radius  $r_0$  at the beginning of the algorithm. The radius  $r_0$  is determined by the amount of available computer memory. Later in the algorithm, only when  $\mathbf{M}\mathbf{n}$  goes outside of the pre-calculated range, do we need to compute  $|\widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{M}\mathbf{n})|^2$ ; otherwise we can just directly pull out the corresponding values from the saved array.

### B. The Precision of the Approximation due to Finite Computation

In the above procedure, we can only check condition (30) for a finite number of radii  $r$ , ranging from 1 to  $r_{\max}$ . In contrast, the original condition (11) in Proposition 1 requires  $r$  to go to infinity. Therefore, our algorithm is only evaluating a necessary condition for alias-free sampling. Intuitively, though,

by using a larger  $r_{\max}$ , we should gain more confidence in the validity of the original infinite condition (11).

To quantify this intuition, we recall that the term  $A_{\mathcal{D},M}$  defined in (7) represents the total volume of overlapping regions between  $\mathcal{D}$  and all its aliasing components. From (9), we can write

$$\begin{aligned} A_{\mathcal{D},M} &= \left( |\mathbf{M}| \sum_{\mathbf{n} \in \mathbb{Z}^N, \|\mathbf{n}\|_\infty \leq r} |\widehat{\mathbf{1}}_{\mathcal{D}}(\mathbf{M}\mathbf{n})|^2 - m(\mathcal{D}) \right) \\ &\quad + |\mathbf{M}| \sum_{\mathbf{n} \in \mathbb{Z}^N, \|\mathbf{n}\|_\infty > r} |\widehat{\mathbf{1}}_{\mathcal{D}}(\mathbf{M}\mathbf{n})|^2 \quad (31) \\ &= A_1(r) + A_2(r), \quad (32) \end{aligned}$$

where  $A_1(r)$  and  $A_2(r)$  represent the first and second terms in the right side of (31), respectively.

When the frequency domain  $\mathcal{D}$  is a 2-D polygon, the following proposition provides an upper bound for the second term  $A_2(r)$ . Note that the techniques used in the proof can be generalized and lead to similar results for cases when  $\mathcal{D}$  is an  $N$ -D polytope (e.g. a polyhedron). We omit this generalization due to space limitations.

*Proposition 7:* Let  $\mathcal{D}$  be a polygon with  $K$  sides, whose vertices, when traced clockwise, are  $\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_K$ . Also, let  $\mathbf{p}_{K+1} \stackrel{\text{def}}{=} \mathbf{p}_1$ . For any nonsingular matrix  $\mathbf{M}$  and radius  $r \geq 1$ , we have

$$\begin{aligned} A_2(r) &= |\mathbf{M}| \sum_{\mathbf{n} \in \mathbb{Z}^2, \|\mathbf{n}\|_\infty > r} |\widehat{\mathbf{1}}_{\mathcal{D}}(\mathbf{M}\mathbf{n})|^2 \\ &< \frac{3K(1 + \frac{1}{\sqrt{2}r})^2}{4\pi^2(r - \frac{1}{2})} \sum_{k=1}^K \frac{d_k(\mathbf{M})}{|\mathbf{M}|} \left(1 + \pi d_k(\mathbf{M})/\sqrt{2}\right)^2, \quad (33) \end{aligned}$$

where  $d_k(\mathbf{M}) \stackrel{\text{def}}{=} \|\mathbf{M}^T(\mathbf{p}_{k+1} - \mathbf{p}_k)\|$  is the length of the  $k$ th side of the warped polygon  $\mathbf{M}^T\mathcal{D}$ .

*Proof:* See Appendix D.  $\blacksquare$

When  $r$  is sufficiently large, the inequality in (33) can be well approximated by  $A_2(r) < c(\mathcal{D}, \mathbf{M})/r$ , where the constant  $c(\mathcal{D}, \mathbf{M}) \stackrel{\text{def}}{=} \frac{3K}{4\pi^2} \sum_{k=1}^K d_k(\mathbf{M}) \left(1 + \pi d_k(\mathbf{M})/\sqrt{2}\right)^2 / |\mathbf{M}|$ . Now for any sampling matrix  $\mathbf{M}$  that satisfies (30) for radii up to  $r_{\max}$ , we have  $A_1(r_{\max}) \leq 0$ , and hence it follows from (32) that

$$A_{\mathcal{D},M} \leq \frac{c(\mathcal{D}, \mathbf{M})}{r_{\max}}. \quad (34)$$

By increasing the maximum search radius  $r_{\max}$ , we can reduce the area of overlapping so that

$$A_{\mathcal{D},M} \leq \varepsilon m(\mathcal{D}) \quad \text{for an arbitrary } \varepsilon > 0. \quad (35)$$

The number  $\varepsilon$  can be used to control the desired precision of our algorithm. After fixing  $\varepsilon$ , we should choose<sup>4</sup>  $r_{\max} = c(\mathcal{D}, \mathbf{M})/(\varepsilon m(\mathcal{D}))$ . In this case, although we can only test a necessary condition for alias-free sampling in Procedure 3, the amount of aliasing  $A_{\mathcal{D},M}$  of the obtained matrices can always be kept within the desired precision range.

<sup>4</sup>By doing this, we will have different maximum search radius  $r_{\max}$  for different  $\mathbf{M}$ , as  $c(\mathcal{D}, \mathbf{M})$  is a function of  $\mathbf{M}$ .

A potential problem in choosing a small  $\varepsilon$  to achieve high precision is that, since  $r_{\max}$  is inversely proportional to  $\varepsilon$ , the amount of computations might be dramatically increased. Fortunately, this scenario will not happen, as most candidate sampling matrices can be quickly eliminated well before the radius  $r$  in (30) reaches  $r_{\max}$ . To see this, we rewrite (32) as

$$A_1(r) = A_{\mathcal{D},M} - A_2(r) \geq A_{\mathcal{D},M} - \frac{c(\mathcal{D}, \mathbf{M})}{r}.$$

For a matrix that does cause aliasing, we have  $A_{\mathcal{D},M} > 0$ ; in this case,  $A_1(r) > 0$  (i.e. the condition (30) fails) as soon as

$$r > r_c \stackrel{\text{def}}{=} c(\mathcal{D}, \mathbf{M})/A_{\mathcal{D},M}.$$

As shown in the following numerical experiments, for the majority of the sampling matrices, the corresponding cut-off radius  $r_c$  is typically much smaller than the required maximum search radius  $r_{\max}$ .

### C. Examples

We test the proposed algorithm in Procedure 3 on a set of four different frequency support regions shown in Figures 4(a)–(d), including two convex polygons  $\mathcal{D}_1$  (a 28-sided regular polygon, approximating a circle) and  $\mathcal{D}_2$ , and two nonconvex polygons  $\mathcal{D}_3$  and  $\mathcal{D}_4$ . For each region, we apply the algorithm to search for the corresponding optimal sampling lattices. In our experiment, we choose the quantization scale to be  $q = 50$ , and set the precision threshold  $\varepsilon$  in (35) to be 0.005.

Figures 4(e)–(h) demonstrate the densest frequency packings achieved by our algorithm. We measure the quality of the results in terms of the sampling efficiency, defined as

$$\text{efficiency} \stackrel{\text{def}}{=} |\mathbf{M}| m(\mathcal{D}) \times 100\%,$$

where  $\mathbf{M}$  is the obtained sampling matrix. Note that a sampling efficiency of 100% corresponds to the case of critical sampling. The highest sampling efficiencies achieved for  $\mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3$ , and  $\mathcal{D}_4$  are 90.66%, 92.24%, 89.94%, and 86.81%, respectively. Note that the obtained sampling density for  $\mathcal{D}_1$  (90.66%) is fairly close to  $\pi/\sqrt{12} \approx 90.69\%$ , which is known to be the highest packing density of circles, achieved by a hexagonal (“honeycomb”) arrangement.

The algorithm is implemented in C++, and the running time ranges from 69 seconds (for  $\mathcal{D}_2$ ) to 355 seconds (for  $\mathcal{D}_3$ ) on a computer with a 2.2 GHz CPU. We observe that the algorithm spends only a small fraction of the total running time to eliminate all unsuitable matrices and reach the optimal sampling lattices. Afterwards, however, the remaining majority of the running time is actually spent on verifying that the obtained matrices indeed provide alias-free sampling within the given precision range. To explain this interesting fact, we recall that the proposed algorithm checks condition (30) against a large number of sampling matrices (with decreasing densities), until it finds the first sampling matrix for which (30) holds for all radii up to  $r_{\max}$ . Let  $N_m$  denote the total number of sampling matrices checked by the algorithm. For every matrix eliminated in the process, there is a corresponding cut-off radius  $r$  at which (30) fails. We denote by  $P(r)$  the

TABLE I  
SOME KEY STATISTICS GATHERED IN THE EXPERIMENT.

Frequency Region	$N_m$	$r_{\max}$	$P(1)$	$P(2)$	$P(3)$	$P(4)$	$P(5)$	$\sum_{r=1}^{10} P(r)$
$\mathcal{D}_1$	11652660	2807	94.79%	3.28%	0.91%	0.38%	0.20%	99.88%
$\mathcal{D}_2$	7642267	2512	96.30%	2.68%	0.59%	0.20%	0.09%	99.97%
$\mathcal{D}_3$	35911314	5666	96.34%	2.99%	0.36%	0.14%	0.06%	99.98%
$\mathcal{D}_4$	25302982	4231	91.75%	6.41%	1.02%	0.33%	0.19%	99.93%

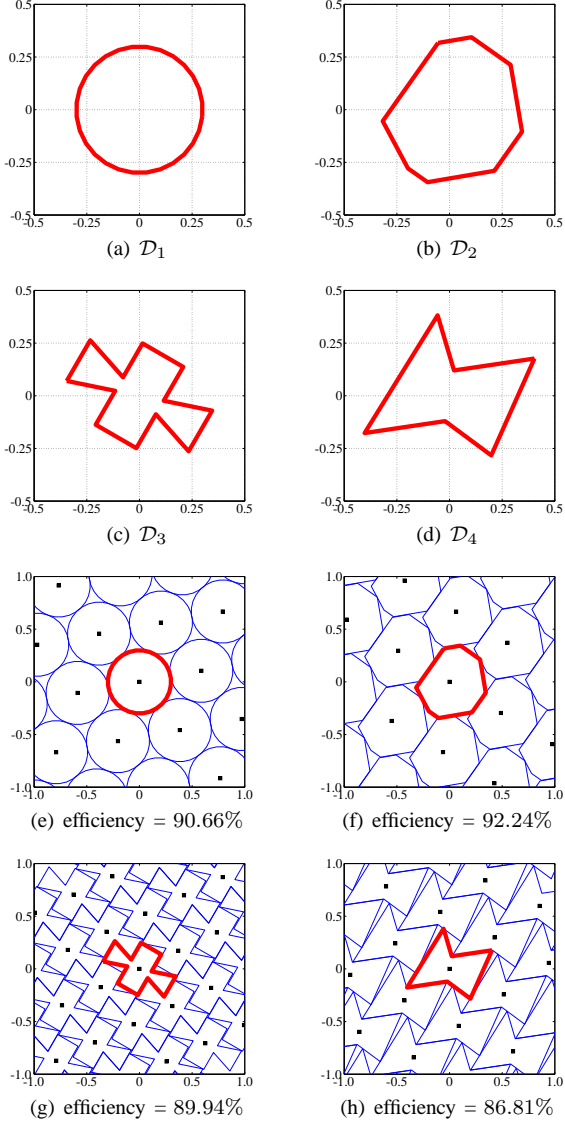


Fig. 4. Top two rows: Four frequency support regions used in our experiment. Bottom two rows: The tightest frequency packing obtained by the proposed algorithm. The basebands (drawn in thick lines) are shown together with their aliasing copies (drawn in thin lines).

relative percentage (with respect to  $N_m$ ) of those matrices whose cut-off radius is equal to  $r$  ( $r = 1, 2, 3, \dots$ ).

Table I summarizes the values of  $N_m$ ,  $r_{\max}$ , and  $P(r)$  ( $r = 1 \dots 5$ ) for the four frequency shapes tested in our experiment. We can see that, although both  $N_m$  and  $r_{\max}$  can be quite large, the majority of the tested matrices can be eliminated by using a fairly small radius. In the case of  $\mathcal{D}_2$ , 96.30% of the tested matrices can be eliminated by using only  $r = 1$ , and 99.97% of the matrices can be eliminated by a radius up to

10. Note that the smaller the cut-off radius is, the less time it takes the algorithm to verify condition (30). For  $r = 1$  (the most likely case as seen from Table I), the algorithm simply needs to compute the values of  $|\hat{\mathbb{1}}_{\mathcal{D}}(M\mathbf{n})|^2$  at four points<sup>5</sup> to eliminate a matrix.

#### D. Fourier Analytical versus Geometrical Approach: A Brief Comment

The proposed algorithm described in Procedure 3 differs from the standard geometrical approach to finding optimal alias-free sampling lattices. The latter often builds upon the following argument [13]: one can verify that the alias-free sampling condition in (5) is equivalent to requiring

$$\Lambda^* \cap (\mathcal{D} - \mathcal{D}) = \{\mathbf{0}\}, \quad (36)$$

where  $\mathcal{D} - \mathcal{D} \stackrel{\text{def}}{=} \{\omega - \tau : \text{for all } \omega, \tau \in \mathcal{D}\}$  is the Minkowski sum of  $\mathcal{D}$  and  $-\mathcal{D}$ . When  $\mathcal{D}$  is a polytope, the corresponding Minkowski sum  $\mathcal{D} - \mathcal{D}$  is also a polytope, whose bounding hyperplanes can be calculated from those of  $\mathcal{D}$  [35]. If we further assume  $\mathcal{D} - \mathcal{D}$  is convex (and hence can be specified by a set of linear inequalities), then for each candidate sampling matrix  $M$ , the condition (36) can be readily checked by evaluating a system of linear inequalities for a finite number of points in  $\Lambda^*$ . The task becomes more difficult, however, when  $\mathcal{D} - \mathcal{D}$  is nonconvex, in which case one has to first decompose it into a union of convex subregions. As a potential advantage, the proposed Fourier analytical approach is purely computational, and in particular, does not depend on the convexity of the domain  $\mathcal{D}$ .

In terms of computational efficiency, the experiments in Section VII-C suggest that the proposed algorithm can be very efficient in eliminating most of the candidate sampling lattices (by only evaluating the Fourier transforms at four points), but needs to spend more time testing those lattices that are close to being alias-free. Consequently, the Fourier analytical approach is likely to be faster than the geometrical approach for most of the candidate lattices, but will be slower than the latter in the remaining cases. It is therefore promising, from a computational point of view, to develop a “hybrid” algorithm—utilizing both the proposed Fourier condition and the geometrical condition in (36)—for finding optimal sampling lattices.

## VIII. CONCLUSIONS AND FUTURE WORK

The main contribution of this paper is the Fourier analytical condition presented in Theorem 1. By linking the alias-free

<sup>5</sup>These points are  $\mathbf{n} = (1, 0), (1, 1), (0, 1), (-1, 1)$ . The remaining points can be inferred by symmetry since  $|\hat{\mathbb{1}}_{\mathcal{D}}(M\mathbf{n})| = |\hat{\mathbb{1}}_{\mathcal{D}}(-M\mathbf{n})|$ .

sampling of a given frequency support region with the Fourier transform of the indicator function, this simple but powerful result provides a versatile computational tool in the search for optimal sampling lattices.

The indicator function plays a pivotal role in this work; however, it is not the only choice. In fact, we can verify that the Fourier condition in Theorem 1 still holds if we replace  $\mathbb{1}_{\mathcal{D}}(\omega)$  with any square-integrable function  $f(\omega)$  that is positive and supported on  $\mathcal{D}$ , and correspondingly, replace  $m(\mathcal{D})$  with  $\|f(\omega)\|_{L^2}^2$ . This generalization leads to many possible lines of research. For example, it is possible to improve computational efficiency by using a function  $f(\omega)$  that has smooth transition at the boundary of  $\mathcal{D}$ , since the smoothness will translate into faster decay (compared with  $|\widehat{\mathbb{1}_{\mathcal{D}}(\mathbf{x})}|$ ) in the spatial domain. Second, the spectral contents of practical signals are unlikely to be uniformly distributed on the frequency support  $\mathcal{D}$ . Therefore, an interesting class of research would choose a function  $f(\omega)$  that is adapted to the power spectral density of the signals. This would potentially allow us to have a more sensible criterion in choosing the optimal sampling lattices.

## APPENDIX

### A. The Poisson Summation Formula

The Poisson summation formula (PSF) or ‘‘lattice sampling formula’’ relates the infinite summation of a function over a lattice with a summation of the Fourier transform of that function over the dual lattice. In this paper, we need the following incarnation of the PSF.

*Theorem 4:* Suppose functions  $f(\omega), g(\omega) \in L^2(\mathbb{R}^N)$  are supported on a bounded domain  $\mathcal{D}$ , i.e.,  $\text{supp } f(\omega) \subset \mathcal{D}$  and  $\text{supp } g(\omega) \subset \mathcal{D}$ , and let  $R(\omega) \stackrel{\text{def}}{=} \int_{\mathbb{R}^N} f(\tau) g(\tau - \omega)^* d\tau$ . Then for all nonsingular matrices  $M$ ,

$$\sum_{\mathbf{k} \in \Lambda^*} R(\mathbf{k}) = |M| \sum_{\mathbf{n} \in \Lambda} \widehat{f}(\mathbf{n}) \widehat{g}(\mathbf{n})^*, \quad (37)$$

where  $\Lambda = \{M\mathbf{m} : \mathbf{m} \in \mathbb{Z}^N\}$  and  $\Lambda^* = \{M^{-T}\boldsymbol{\ell} : \boldsymbol{\ell} \in \mathbb{Z}^N\}$ .

*Remark:* A difficulty in directly applying the PSF to show (37) is that the standard versions of the PSF (such as given in [36, pp. 250-257]) require the Fourier transform to decay sufficiently fast. In particular, for our case we would need

$$|\widehat{f}(\mathbf{x}) \widehat{g}(\mathbf{x})| \leq A(1 + \|\mathbf{x}\|)^{-N-\varepsilon} \quad (38)$$

for some  $A > 0$  and  $\varepsilon > 0$ . The condition (38) need not be satisfied in multidimensional cases ( $N \geq 2$ ) in our work, because we take  $f(\omega) = g(\omega) = \mathbb{1}_{\mathcal{D}}(\omega)$  and the discontinuity of this indicator function can lead to slow decay of its Fourier transform.

We provide a direct proof of Theorem 4 that circumvents the assumption (38).

*Proof:* First assume  $M$  is the identity matrix. Denote by  $\mathcal{C}_0$  the unit cube  $[-\frac{1}{2}, \frac{1}{2}]^N$ . We may suppose  $f$  is supported in a cube of the form  $\boldsymbol{\ell} + \mathcal{C}_0$  for some integer vector  $\boldsymbol{\ell} \in \mathbb{Z}^N$ , because the original  $f$  can be written as a finite linear combination of functions supported in such cubes (noting  $f$  has bounded support and that  $f$  appears linearly in the definition

of  $R$ ). Similarly we may assume  $g$  is supported in a cube  $\mathbf{m} + \mathcal{C}_0$  for some  $\mathbf{m} \in \mathbb{Z}^N$ . Hence

$$\sum_{\mathbf{k} \in \mathbb{Z}^N} R(\mathbf{k}) = \int_{\boldsymbol{\ell} + \mathcal{C}_0} f(\boldsymbol{\tau}) g(\boldsymbol{\tau} - \boldsymbol{\ell} + \mathbf{m})^* d\boldsymbol{\tau},$$

since the support cubes of  $f(\boldsymbol{\tau})$  and  $g(\boldsymbol{\tau} - \mathbf{k})$  are disjoint except when  $\mathbf{k} = \boldsymbol{\ell} - \mathbf{m}$ . Now changing variable with  $\boldsymbol{\tau} \mapsto \boldsymbol{\tau} + \boldsymbol{\ell}$  yields:

$$\sum_{\mathbf{k} \in \mathbb{Z}^N} R(\mathbf{k}) = \int_{\mathcal{C}_0} f(\boldsymbol{\tau} + \boldsymbol{\ell}) g(\boldsymbol{\tau} + \mathbf{m})^* d\boldsymbol{\tau}. \quad (39)$$

Applying the Plancherel theorem for Fourier series, we have

$$\int_{\mathcal{C}_0} f(\boldsymbol{\tau} + \boldsymbol{\ell}) g(\boldsymbol{\tau} + \mathbf{m})^* d\boldsymbol{\tau} = \sum_{\mathbf{n} \in \mathbb{Z}^N} c_{\mathbf{n}}(f) c_{\mathbf{n}}^*(g), \quad (40)$$

where the Fourier coefficients on the right-hand side are  $c_{\mathbf{n}}(f) \stackrel{\text{def}}{=} \int_{\mathcal{C}_0} f(\boldsymbol{\tau} + \boldsymbol{\ell}) e^{-2\pi j \mathbf{n} \cdot \boldsymbol{\tau}} d\boldsymbol{\tau} = \widehat{f}(\mathbf{n})$ , and similarly  $c_{\mathbf{n}}(g) = \widehat{g}(\mathbf{n})$ . Substituting these identities into (40) gives that

$$\sum_{\mathbf{k} \in \mathbb{Z}^N} R(\mathbf{k}) = \sum_{\mathbf{n} \in \mathbb{Z}^N} \widehat{f}(\mathbf{n}) \widehat{g}(\mathbf{n})^* \quad (41)$$

as desired.

So far we have proved (37) for the case when  $M$  is an identity matrix. For general  $M$ , we define  $\widetilde{f}(\omega) = |M|^{-1/2} f(M^{-T}\omega)$  and  $\widetilde{g}(\omega) = |M|^{-1/2} g(M^{-T}\omega)$ . Applying (41) to  $\widetilde{f}$  and  $\widetilde{g}$  then proves the general case of the theorem, after some straightforward calculations. ■

### B. The Number of Hermite Normal Matrices with a Given Determinant

We first state the following result about the total number of Hermite normal matrices, whose proof can be found in [19, pp.19-21].

*Lemma 4:* Let  $H_N(\delta)$  be the number of  $N$ -by- $N$  Hermite normal matrices with determinant  $\delta$ . If  $\delta = \mu\nu$ , for coprime  $\mu$  and  $\nu$ , then  $H_N(\mu\nu) = H_N(\mu) H_N(\nu)$ . Furthermore, if  $\delta = \pi^k$ ,  $\pi$  a prime, then

$$H_N(\pi^k) = \prod_{i=1}^{N-1} \frac{\pi^{k+i} - 1}{\pi^i - 1}. \quad (42)$$

Note that we can factorize any integer  $\delta \geq 2$  into  $\delta = \prod_{j=1}^J \pi_j^{k_j}$ , where  $\pi_1, \pi_2, \dots, \pi_J$  are distinct prime factors and the exponents  $k_1, k_2, \dots, k_J$  are positive integers. Applying the above result, we have  $H_N(\delta) = \prod_{j=1}^J H_N(\pi_j^{k_j})$ , where each term can be calculated by (42). We can see that the formula for the exact value of  $H_N(\delta)$  depends on the prime factorization of  $\delta$ . It is therefore often simpler and more convenient to use the estimates given in Proposition 6. To prove that proposition, we establish its lower bound and upper bounds in the following two propositions.

*Proposition 8:*  $H_N(1) = 1$ , and  $H_N(\delta) \geq \frac{\delta^N - 1}{\delta - 1}$  for  $\delta \geq 2$ , where the equality is achieved when  $\delta$  is a prime number.

*Proof:* For  $\delta \geq 2$  and a given  $k$ ,  $1 \leq k \leq N$ , we consider the following class of integer matrices  $\mathbf{H} = [h_{i,j}]$ , whose entries are all zeros except for those on the diagonal line, where  $h_{k,k} = \delta$  and  $h_{i,i} = 1$  for  $i \neq k$ , and those on the  $k$ th

row, where  $0 \leq h_{k,j} < \delta$  for  $k < j \leq N$ . By construction, we can easily verify that these matrices are all in the Hermite normal form (as specified in Theorem 2) and with determinant  $\delta$ . Meanwhile, there are a total of  $\delta^{N-k}$  of such matrices. By varying  $k$  from 1 to  $N$ , we get  $H_N(\delta) \geq \delta^{N-1} + \delta^{N-2} + \dots + \delta + 1 = \frac{\delta^N - 1}{\delta - 1}$ . Furthermore, when  $\delta$  is prime, putting  $\delta$  in one of the diagonal elements and 1 in all other diagonal elements is the only way for the determinant of a Hermite normal matrix to equal  $\delta$ . Therefore, the equality is achieved in this case.  $\blacksquare$

*Proposition 9:* As an upper bound, we have  $H_N(\delta) < 22 \delta^{N-1+0.001}$ .

*Proof:* For a prime number  $\pi$ , we have from (42) that

$$H_N(\pi^k) < \prod_{i=1}^{N-1} \frac{\pi^{k+i}}{\pi^i - 1} = (\pi^k)^{N-1} \prod_{i=1}^{N-1} \frac{1}{1 - \pi^{-i}} \quad (43)$$

$$< (\pi^k)^{N-1} C(\pi),$$

where  $C(\pi) \stackrel{\text{def}}{=} \prod_{i=1}^{\infty} (1 - \pi^{-i})^{-1}$ . We can easily verify that  $C(\pi) > 1$  for all  $\pi$ , and that it is a decreasing function of  $\pi$ , i.e.,  $C(\pi_1) > C(\pi_2)$  for  $\pi_1 < \pi_2$ .

For any integer  $\delta \geq 2$ , we factorize it into  $\delta = \prod_{j=1}^J \pi_j^{k_j}$ , where  $\pi_1, \pi_2, \dots, \pi_J$  are distinct prime factors. It follows from Lemma 4 and inequality (43) that

$$H_N(\delta) = \prod_{j=1}^J H_N(\pi_j^{k_j}) < \prod_{j=1}^J (\pi_j^{k_j})^{N-1} \prod_{j=1}^J C(\pi_j)$$

$$= \delta^{N-1} \prod_{j=1}^J C(\pi_j). \quad (44)$$

Next, we just need to show  $\prod_{j=1}^J C(\pi_j) < 22 \delta^{0.001}$ .

For a given positive integer  $n$ , we define a set  $A$  consisting of all prime factors  $\{\pi_j\}$  of  $\delta$  that are less than  $n$ , i.e.,  $A \stackrel{\text{def}}{=} \{\pi_j : 1 \leq j \leq J \text{ and } \pi_j < n\}$ . Similarly, let  $B = \{\pi_j : 1 \leq j \leq J \text{ and } \pi_j \geq n\}$ . We can then write

$$\prod_{j=1}^J C(\pi_j) = \prod_{\pi \in A} C(\pi) \prod_{\pi \in B} C(\pi). \quad (45)$$

In the above expression, whenever  $A$  or  $B$  is empty, the corresponding product on that set is understood to equal 1. Let  $P$  represent the set of all prime numbers. It then follows from the definition of  $A$  that

$$\prod_{\pi \in A} C(\pi) \leq \prod_{\pi \in P, \pi < n} C(\pi). \quad (46)$$

Denote by  $|B|$  the cardinality of  $B$ . If  $|B| > 0$ , then for all  $\pi \in B$ , we have  $\pi \geq n$  and hence  $C(\pi) \leq C(n)$ . Meanwhile, since  $\delta \geq \prod_{\pi \in B} \pi \geq n^{|B|}$ , we have  $|B| \leq \log_n \delta$ . It follows that

$$\prod_{\pi \in B} C(\pi) \leq C(n)^{|B|} \leq C(n)^{\log_n \delta} = \delta^{\log_n C(n)}. \quad (47)$$

Substituting (46) and (47) into (45), we obtain  $\prod_{j=1}^J C(\pi_j) \leq \delta^{\log_n C(n)} \prod_{\pi \in P, \pi < n} C(\pi)$ . We have the freedom in choosing  $n$ . For example, when  $n = 192$ , the exponent becomes  $\log_{192} C(192) \approx 9.98 \times 10^{-4} < 0.001$ ; the constant factor becomes  $\prod_{\pi \in P, \pi < n} C(\pi) \approx 21.7 < 22$ .

Substituting these numbers into (44) yields the desired result.  $\blacksquare$

### C. A Lemma Used in the Proof of Theorem 3

*Lemma 5:* For arbitrarily chosen  $0 \leq a < b$  and  $c > 1$ , there must exist some  $n \in \mathbb{Z} \setminus \{0\}$  such that  $\int_a^b t^c \cos(nt) dt \neq 0$ .

*Proof:* We show this by contradiction. Suppose there is a particular set of parameters  $0 \leq a < b$  and  $c > 1$  for which

$$\int_a^b t^c \cos(nt) dt = 0, \quad \text{for all } n \in \mathbb{Z} \setminus \{0\}. \quad (48)$$

Defining the function  $f(t) = t^c \mathbb{1}_{[a,b]}(t)$ , we can rewrite the above equality as

$$0 = \int_{-\infty}^{\infty} f(t) \cos(nt) dt$$

$$= \sum_{k \in \mathbb{Z}} \int_{-\pi+2\pi k}^{\pi+2\pi k} f(t) \cos(nt) dt$$

$$= \int_{-\pi}^{\pi} \sum_{k \in \mathbb{Z}} f(t+2\pi k) \cos(n(t+2\pi k)) dt$$

$$= \int_{-\pi}^{\pi} g(t) \cos(nt) dt, \quad n \in \mathbb{Z} \setminus \{0\}, \quad (49)$$

where  $g(t) \stackrel{\text{def}}{=} \sum_{k \in \mathbb{Z}} f(t+2\pi k)$ . Actually, since  $f(t)$  is compactly supported, the sum generating  $g(t)$  contains only a finite number of nonzero terms. The equality (49) means that all the even Fourier coefficients (except for the DC term) of  $g(t)$  are zero, which implies that  $g(t)$  is an odd function plus some constant.

For  $c > 1$ , the function  $f(t)$  is strictly convex on  $[a, b]$ ; periodizing the endpoints of  $[a, b]$  gives an ‘‘exceptional’’ set  $E \stackrel{\text{def}}{=} (a+2\pi\mathbb{Z}) \cup (b+2\pi\mathbb{Z})$ , and so  $g(t)$  is *strictly convex* on each subinterval of  $[-\pi, \pi] \setminus E$ . However, from the symmetry of  $g(t)$  (i.e. an odd function plus some constant),  $g(t)$  will be *strictly concave* on the reflected (w.r.t. the origin) versions of these intervals. This is a contradiction, and hence the initial assumption (48) does not hold.  $\blacksquare$

### D. Proof of Proposition 7

We start by proving (33) for the special case when  $M$  is the identity matrix. Using the formula (17) of  $\widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{x})$ , and writing  $\boldsymbol{\theta}_k = \pi(\mathbf{p}_{k+1} - \mathbf{p}_k)$ , we get

$$|\widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{x})| \leq \frac{1}{2\pi \|\mathbf{x}\|} \sum_{k=1}^K d_k |\text{sinc}(\mathbf{x} \cdot \boldsymbol{\theta}_k / \pi)|$$

$$< \frac{1}{2\pi \|\mathbf{x}\|} \sum_{k=1}^K \frac{d_k \sqrt{1.5}}{\sqrt{1 + (\mathbf{x} \cdot \boldsymbol{\theta}_k)^2}},$$

where the second inequality is due to the fact that  $|\text{sinc}(x)| = |\sin(\pi x)/(\pi x)| < \sqrt{1.5}/(1 + \pi^2 x^2)$ . Applying the Cauchy-Schwarz inequality  $(\sum_{k=1}^K 1 \cdot x_k)^2 \leq K \sum_{k=1}^K x_k^2$  to the above

expression yields

$$\begin{aligned} |\widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{x})|^2 &< \frac{3K}{8\pi^2 \|\mathbf{x}\|^2} \sum_{k=1}^K \frac{d_k^2}{1 + (\mathbf{x} \cdot \boldsymbol{\theta}_k)^2} \\ &= \frac{3K}{8\pi^2} \sum_{k=1}^K d_k^2 G_{\boldsymbol{\theta}_k}(\mathbf{x}), \end{aligned} \quad (50)$$

where  $G_{\boldsymbol{\theta}}(\mathbf{x}) \stackrel{\text{def}}{=} \|\mathbf{x}\|^{-2} (1 + (\mathbf{x} \cdot \boldsymbol{\theta})^2)^{-1}$ .

Next, we derive an upper bound for the infinite sum of  $G_{\boldsymbol{\theta}}(\mathbf{x})$ . For arbitrary  $\|\mathbf{u}\| \leq \frac{\sqrt{2}}{2} < r < \|\mathbf{n}\|_{\infty}$ ,

$$\begin{aligned} \frac{G_{\boldsymbol{\theta}}(\mathbf{n})}{G_{\boldsymbol{\theta}}(\mathbf{n} + \mathbf{u})} &= \frac{\|\mathbf{n} + \mathbf{u}\|^2}{\|\mathbf{n}\|^2} \frac{1 + (\mathbf{n} \cdot \boldsymbol{\theta} + \mathbf{u} \cdot \boldsymbol{\theta})^2}{1 + (\mathbf{n} \cdot \boldsymbol{\theta})^2} \\ &\leq \left(1 + \frac{\|\mathbf{u}\|}{\|\mathbf{n}\|}\right)^2 (1 + |\mathbf{u} \cdot \boldsymbol{\theta}|)^2 \\ &\leq \left(1 + \frac{1}{\sqrt{2}r}\right)^2 (1 + \|\boldsymbol{\theta}\|/\sqrt{2})^2 \stackrel{\text{def}}{=} C_{\boldsymbol{\theta}}. \end{aligned} \quad (51)$$

It then follows that  $G_{\boldsymbol{\theta}}(\mathbf{n}) \leq C_{\boldsymbol{\theta}} \int_{\mathbf{u} \in [-\frac{1}{2}, \frac{1}{2}]^2} G_{\boldsymbol{\theta}}(\mathbf{n} + \mathbf{u}) d\mathbf{u} = C_{\boldsymbol{\theta}} \int_{\mathbf{n} + [-\frac{1}{2}, \frac{1}{2}]^2} G_{\boldsymbol{\theta}}(\mathbf{u}) d\mathbf{u}$ , and therefore

$$\begin{aligned} &\sum_{\mathbf{n} \in \mathbb{Z}^2, \|\mathbf{n}\|_{\infty} > r} G_{\boldsymbol{\theta}}(\mathbf{n}) \\ &\leq C_{\boldsymbol{\theta}} \int_{\|\mathbf{x}\| > r - \frac{1}{2}} G_{\boldsymbol{\theta}}(\mathbf{x}) d\mathbf{x} \\ &= C_{\boldsymbol{\theta}} \int_{r - \frac{1}{2}}^{\infty} \rho d\rho \int_0^{\pi/2} \frac{4 d\alpha}{\rho^2 (1 + \|\boldsymbol{\theta}\|^2 \rho^2 \cos^2 \alpha)} \\ &= C_{\boldsymbol{\theta}} \int_{r - \frac{1}{2}}^{\infty} \left( \frac{2\pi}{\rho \sqrt{1 + \|\boldsymbol{\theta}\|^2 \rho^2}} \right) d\rho < \frac{2\pi C_{\boldsymbol{\theta}}}{\|\boldsymbol{\theta}\| (r - \frac{1}{2})}, \end{aligned} \quad (52)$$

where (53) is obtained from (52) based on the following identity: for any constant  $b \stackrel{\text{def}}{=} \|\boldsymbol{\theta}\|^2 \rho^2$ ,

$$\frac{d}{d\alpha} \arctan\left(\frac{\tan \alpha}{\sqrt{1+b}}\right) = \frac{\sqrt{1+b}}{1+b \cos^2 \alpha}, \quad \text{for } 0 \leq \alpha < \frac{\pi}{2}.$$

Substituting (51) and (53) into (50), and using  $\|\boldsymbol{\theta}_k\| = \pi d_k$ , we reach the inequality (33) for the case when  $\mathbf{M}$  is an identity matrix. For general  $\mathbf{M}$ , we use the following change of variables:  $\mathbb{1}_{\mathcal{D}}(\boldsymbol{\omega}) = \mathbb{1}_{\mathbf{M}^T \mathcal{D}}(\mathbf{M}^T \boldsymbol{\omega})$ , and hence  $\widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{x}) = |\mathbf{M}|^{-1} \widehat{\mathbb{1}}_{\mathbf{M}^T \mathcal{D}}(\mathbf{M}^{-1} \mathbf{x})$ . It follows that

$$\begin{aligned} |\mathbf{M}| \sum_{\mathbf{n} \in \mathbb{Z}^2, \|\mathbf{n}\|_{\infty} > r} |\widehat{\mathbb{1}}_{\mathcal{D}}(\mathbf{M}\mathbf{n})|^2 \\ = \frac{1}{|\mathbf{M}|} \sum_{\mathbf{n} \in \mathbb{Z}^2, \|\mathbf{n}\|_{\infty} > r} |\widehat{\mathbb{1}}_{\mathbf{M}^T \mathcal{D}}(\mathbf{n})|^2. \end{aligned}$$

Applying the previous result to  $\widehat{\mathbb{1}}_{\mathbf{M}^T \mathcal{D}}(\mathbf{n})$  leads to (33).

## REFERENCES

- [1] A. Jerri, "The Shannon sampling theorem — its various extensions and applications: A tutorial review," *Proc. IEEE*, vol. 65, no. 11, pp. 1565–1596, Nov. 1977.
- [2] M. Unser, "Sampling - 50 years after Shannon," *Proc. IEEE*, vol. 88, no. 4, pp. 569–587, Apr. 2000.
- [3] D. P. Peterson and D. Middleton, "Sampling and reconstruction of wavenumber-limited functions in  $N$ -dimensional Euclidean spaces," *Inform. Contr.*, vol. 5, pp. 279–323, 1962.
- [4] E. Dubois, "The sampling and reconstruction of time-varying imagery with application in video systems," *Proc. IEEE*, vol. 73, no. 4, pp. 502–522, April 1985.
- [5] A. Entezari, D. Van De Ville, and T. Möller, "Practical box splines for reconstruction on the body centered cubic lattice," *IEEE Trans. Vis. Comput. Graphics*, vol. 14, no. 2, pp. 1–16, Mar./Apr. 2008.
- [6] J.-X. Chai, X. Tong, S.-C. Chan, and H.-Y. Shum, "Plenoptic sampling," in *Proc. of SIGGRAPH*, New York, NY, USA, 2000, pp. 307–318.
- [7] J. O. Coleman, "Three-phase sample timing on a wideband triangular array of 4/3 the usual density reduces the Nyquist rate for far-field signals by two thirds," in *Proc. 38th Asilomar Conference on Signals, Systems and Computers*, vol. 1, 2004, pp. 284–288.
- [8] T. Ajdler, L. Sbaiz, and M. Vetterli, "The plenacoustic function and its sampling," *IEEE Trans. Signal Process.*, vol. 54, no. 10, pp. 3790–3804, 2006.
- [9] E. Viscito and J. P. Allebach, "The analysis and design of multidimensional FIR perfect reconstruction filter banks for arbitrary sampling lattices," *IEEE Trans. Circuits Syst.*, vol. 38, no. 1, pp. 29–41, January 1991.
- [10] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
- [11] M. Vetterli and J. Kovačević, *Wavelets and Subband Coding*. Englewood Cliffs: Prentice Hall, 1995.
- [12] H. R. Künsch, E. Agrell, and F. A. Hamprecht, "Optimal lattices for sampling," *IEEE Trans. Inf. Theory*, vol. 51, no. 2, pp. 634–47, Feb. 2005.
- [13] P. M. Gruber and C. G. Lekkerkerker, *Geometry of Numbers*, 2nd ed. Amsterdam: Elsevier Science Publishers, 1987.
- [14] P. D. Watson and A. M. Tobias, "An efficient algorithm for the regular W1 packing of polygons in the infinite plane," *Journal of the Operational Research Society*, vol. 50, no. 10, pp. 1054–1062, Oct. 1999.
- [15] U. Betke and M. Henk, "Densest lattice packings of 3-polytopes," *Computational Geometry*, vol. 16, no. 3, pp. 157–186, 2000.
- [16] V. J. Milenkovic, "Densest translational lattice packing of non-convex polygons," in *Proc. 16th Annual Symposium on Computational Geometry*, New York, NY, 2000, pp. 280–289.
- [17] M. N. Kolountzakis and J. C. Lagarias, "Tilings of the line by translates of a function," *Duke Math. J.*, vol. 82, no. 3, pp. 653–678, 1996.
- [18] M. N. Kolountzakis, "The study of translational tiling with Fourier analysis," in *Fourier Analysis and Convexity*, L. Brandolini, L. Colzani, A. Iosevich, and G. Travaglino, Eds. Boston: Birkhäuser, 2004, pp. 131–187.
- [19] M. Newman, *Integral Matrices*. San Diego: Academic Press, 1972.
- [20] G. Cortelazzo and R. Manduchi, "On the determination of all the sublattices of preassigned index and its application to multidimensional subsampling," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 3, no. 4, pp. 318–320, August 1993.
- [21] A. Papoulis, "Generalized sampling expansion," *IEEE Trans. Circuits Syst.*, vol. CAS-24, no. 11, pp. 652–654, Nov. 1977.
- [22] K. F. Cheung and R. Marks, "Imaging sampling below the Nyquist density without aliasing," *Journal of the Optical Society of America A: Optics, Image Science, Vision*, vol. 7, no. 1, pp. 92–105, January 1990.
- [23] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia, PA: SIAM, 1992.
- [24] D. Van De Ville, T. Blu, M. Unser, W. Philips, I. Lemahieu, and R. V. de Walle, "Hex-Splines: A novel spline family for hexagonal lattices," *IEEE Trans. Image Process.*, vol. 13, no. 6, pp. 758–772, Jun. 2004.
- [25] R. H. Bamberger and M. J. T. Smith, "A filter bank for the directional decomposition of images: theory and design," *IEEE Trans. Signal Process.*, vol. 40, no. 4, pp. 882–893, April 1992.
- [26] M. N. Do, "Directional multiresolution image representations," Ph.D. dissertation, Swiss Federal Institute of Technology, Lausanne, Switzerland, December 2001.
- [27] B. Randol, "On the asymptotic behavior of the Fourier transform of the indicator function of a convex set," *Trans. Amer. Math. Soc.*, vol. 139, pp. 279–285, May 1969.
- [28] L. Brandolini, L. Colzani, and G. Travaglino, "Average decay of Fourier transforms and integer points in polyhedra," *Ark. Mat.*, vol. 35, pp. 253–275, 1997.
- [29] G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*, 5th ed. Oxford, England: Oxford University Press, 1979.
- [30] T. Chen and P. P. Vaidyanathan, "Considerations in multidimensional filter bank design," in *Proc. IEEE Int. Symp. Circ. and Syst.*, Chicago, Illinois, USA, May 1993, pp. 643–646.
- [31] Y. Lu and M. N. Do, "CRISP-contourlets: A critically-sampled directional multiresolution image representation," in *Proc. of SPIE Conference*

on *Wavelet Applications in Signal and Image Processing X*, San Diego, USA, August 2003, pp. 655–665.

- [32] T. T. Nguyen and S. Orantara, “Multiresolution direction filter banks: Theory, design and applications,” *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 3895–3905, October 2005.
- [33] Y. M. Lu and M. N. Do, “Multidimensional directional filter banks and surfacelets,” *IEEE Trans. Image Process.*, vol. 16, no. 4, pp. 918–931, April 2007.
- [34] R. H. Bamberger, “New results on two and three dimensional directional filter banks,” in *Twenty-Seventh Asilomar Conf. on Signals, Systems and Computers*, vol. 2, 1993, pp. 1286–1290.
- [35] P. K. Ghosh, “A unified computational framework for Minkowski operations,” *Comput. Graphics*, vol. 17, no. 4, pp. 357–378, 1993.
- [36] E. M. Stein and G. Weiss, *Introduction to Fourier Analysis on Euclidean Spaces*. Princeton: Princeton University Press, 1971.



**Richard S. Laugesen** received his Ph.D. from Washington University in St. Louis in 1993. After postdoctoral appointments at the University of Michigan, Johns Hopkins University and the Institute for Advanced Study, he joined the faculty at the University of Illinois in 1997. His mathematical interests span partial differential equations and harmonic analysis.



**Yue M. Lu** received the B.Eng and M.Eng degrees in electrical engineering from Shanghai Jiao Tong University, China, in 1999 and 2002, respectively. He received the M.Sc degree in mathematics and the Ph.D. degree in electrical engineering from the University of Illinois at Urbana-Champaign in 2007.

He was a Research Assistant at the University of Illinois at Urbana-Champaign, and has worked for Microsoft Research Asia, Beijing, China and Siemens Corporate Research, Princeton, NJ. He is now with the Audio-Visual Communications Laboratory at the Swiss Federal Institute of Technology Lausanne (EPFL), Switzerland. His research interests include signal processing for sensor networks; the theory, constructions, and applications of multiscale geometric representations for multidimensional signals; image and video processing; and sampling theory.

He received the Most Innovative Paper Award of IEEE International Conference on Image Processing (ICIP) in 2006 for his paper (with Minh N. Do) on the construction of directional multiresolution image representations, and the Student Paper Award of IEEE ICIP in 2007.



**Minh N. Do** was born in Thanh Hoa, Vietnam, in 1974. He received the B.Eng. degree in computer engineering from the University of Canberra, Australia, in 1997, and the Dr.Sci. degree in communication systems from the Swiss Federal Institute of Technology Lausanne (EPFL), Switzerland, in 2001.

He joined the University of Illinois at Urbana-Champaign in 2002, where he is currently an Associate Professor with the Department of Electrical and Computer Engineering and a Research Associate Professor with the Coordinated Science Laboratory

and the Beckman Institute. His research interests include image and multi-dimensional signal processing, wavelets and multiscale geometric analysis, computational imaging, and visual information representation.

He received a Silver Medal from the 32nd International Mathematical Olympiad in 1991, a University Medal from the University of Canberra in 1997, the best doctoral thesis award from EPFL in 2001, and a CAREER award from the National Science Foundation in 2003. He was named a Beckman Fellow at the Center for Advanced Study, UIUC, in 2006, and received of a Xerox Award for Faculty Research from the College of Engineering, UIUC, in 2007. He is a member of the IEEE Signal Processing Society Signal Processing Theory and Methods and Image and MultiDimensional Signal Processing Technical Committees, and an Associate Editor of the IEEE Transactions on Image Processing.