Perfect Reconstruction Formulas and Bounds on Aliasing Error in Sub-Nyquist Nonuniform Sampling of Multiband Signals

Raman Venkataramani, *Student Member, IEEE,* and Yoram Bresler, *Fellow, IEEE*

***Abstract*—We examine the problem of periodic nonuniform sampling of a multiband signal and its reconstruction from the samples. This sampling scheme, which has been studied previously, has an interesting optimality property that uniform sampling lacks: one can sample and reconstruct the class $\mathcal{B}(\mathcal{F})$ of multiband signals with spectral support $\mathcal{F}$, at rates arbitrarily close to the Landau minimum rate equal to the Lebesgue measure of $\mathcal{F}$, even when $\mathcal{F}$ does not tile $\mathbb{R}$ under translation. Using the conditions for exact reconstruction, we derive an explicit reconstruction formula. We compute bounds on the peak value and the energy of the aliasing error in the event that the input signal is band-limited to the "span of $\mathcal{F}$" (the smallest interval containing $\mathcal{F}$) which is a bigger class than the valid signals $\mathcal{B}(\mathcal{F})$, band-limited to $\mathcal{F}$. We also examine the performance of the reconstruction system when the input contains additive sample noise.**

***Index Terms*—Error bounds, Landau–Nyquist rate, multiband, nonuniform periodic sampling, signal representation.**

I. INTRODUCTION

THE classical sampling theorem states that a signal occupying a finite range in the frequency domain can be represented by its samples taken at a finite rate. Often attributed to Whittaker, Kotelńikov, and Shannon, a more precise statement of this so-called WKS sampling theorem is that a real low-pass signal, whose Fourier transform is limited to the range $(-f_0, f_0)$, can be recovered from its samples taken uniformly at the rate $f_s = 2f_0$ (the Nyquist rate) or higher [1].

Sampling a signal $x(t)$ uniformly at f_s causes the resulting spectrum to contain multiple copies of the original spectrum $X(f)$ located with uniform spacing of f_s between adjacent copies. Hence the choice $f_s \geq 2f_0$ guarantees no overlaps in the sampled spectrum, and thus allows recovery of the original signal by a low-pass filtering operation. This is the key idea behind the classical sampling theorem. For efficient sampling, it is desirable to attain the lowest sampling rate possible, and this is characterized by the absence of gaps or overlaps [2] in the spectrum of the sampled signal. Unfortunately, in the case of bandpass signals, it is not always possible to eliminate gaps in the sampled spectrum, however, it is possible to minimize them. Some results on bandpass sampling can be found in [3], [4]. Thus while uniform sampling theorems work well for low-pass signals, they are quite inefficient for representing certain bandpass signals and, more generally, for multiband signals, i.e., signals containing several bands in the frequency domain. We refer the reader to Papoulis [5] and Jerri's tutorial [6] for some generalizations of the WKS sampling theorem.

To quantify the sampling efficiency for signals with a given spectral support $\mathcal{F}$, we define its spectral span, $[\mathcal{F}]$, as the smallest interval containing $\mathcal{F}$, and its spectral occupancy as $\Omega = \lambda(\mathcal{F})/\lambda([\mathcal{F}])$, where $\lambda(\cdot)$ denotes the Lebesgue measure. The Nyquist rate $f_{\rm nyq}$ for signals with spectral support $\mathcal{F}$ is defined as the smallest uniform sampling rate that guarantees no aliasing

$$f_{\rm nyq} = \inf\{\theta > 0\colon \mathcal{F} \cap (n\theta \oplus \mathcal{F}) = \emptyset,\ \forall n \in \mathbb{Z}\backslash\{0\}\}$$

where

$$\theta \oplus \mathcal{F} \stackrel{\rm def}{=} \{\theta + f\colon f \in \mathcal{F}\}$$

is the translation of the set $\mathcal{F}$ by θ. Then, the Nyquist sampling rate satisfies

$$\lambda(\mathcal{F}) \leq f_{\rm nyq} \leq \lambda([\mathcal{F}]).$$

We say that $\mathcal{F}$ tessellates $\mathbb{R}$ or $\mathcal{F}$ is *packable* if $f_{\rm nyq} = \lambda(\mathcal{F})$, and *nonpackable* otherwise ($f_{\rm nyq} > \lambda(\mathcal{F})$). In other words, the Nyquist rate for nonpackable signals exceeds the total length of its spectral support. At the other extreme is the case $f_{\rm nyq} = \lambda([\mathcal{F}])$ (*totally nonpackable*), where uniform sampling cannot exploit the presence of gaps in $\mathcal{F}$.

The general case of interest in this paper is that of $\mathcal{F}$ being nonpackable such that the Nyquist rate for sampling $x(t)$ with spectral support $\mathcal{F}$ is $f_{\rm nyq} > \lambda(\mathcal{F})$. On the other hand, Landau [7] showed that the sampling rate of an arbitrary sampling scheme for the class of multiband signals with spectral support $\mathcal{F}$ is lower-bounded by the quantity $\lambda(\mathcal{F})$, which may be significantly smaller than the Nyquist rate. Thus the spectral occupancy Ω is a measure of the efficiency of Landau's lower bound over the Nyquist rate. Because Ω can be low for certain nonpackable signals (in fact, it is easy to construct examples of nonpackable $\mathcal{F}$ with arbitrarily small Ω), uniform sampling is highly inefficient for such signals. Fig. 1 illustrates a typical case of such a nonpackable multiband signal. The Nyquist

Manuscript received July 12, 1998; revised March 8, 2000. This work was supported in part by the Joint Services Electronic Program under Grant N00014-96-1-0129, the National Science Foundation under Grant MIP 97-07633, and DARPA under Contract F49620-98-1-0498.

The authors are with the Coordinated Science Laboratory, Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA (e-mail: ybressler@uiuc.edu; raman@ifp.uiuc.edu).

Communicated by C. Herley, Associate Editor for Estimation.
Publisher Item Identifier S 0018-9448(00)06996-0.

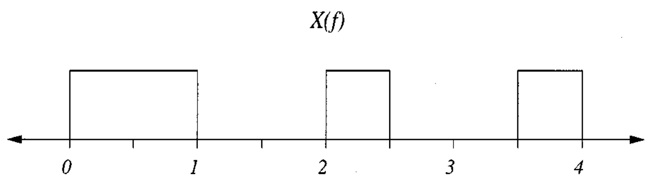

Fig. 1. The spectrum of a nonpackable multiband signal.

rate for this signal is $f_{\rm nyq} = \lambda([\mathcal{F}]) = 4$ (hence $\mathcal{F}$ is totally nonpackable), whereas the Landau lower bound is $\lambda(\mathcal{F}) = 2$. The spectral occupancy for this signal, $\Omega = 0.5$, suggests that it might be possible to sample the signal twice as efficiently as the Nyquist rate. We examine the problem of efficient sampling of nonpackable signals in this paper.

Our results apply to the class of continuous complex-valued band-limited signals of finite energy with spectral support $\mathcal{F}$, namely,

$$\mathcal{B}(\mathcal{F}) = \{x(t) \in L^2(\mathbb{R}) \cap C(\mathbb{R}): X(f) = 0, f \notin \mathcal{F}\}$$

where $X(f)$ is the Fourier transform of $x(t)$. All the results except those about the peak aliasing error also apply to the bigger class

$$\mathcal{B}_2(\mathcal{F}) = \{x(t) \in L^2(\mathbb{R}): X(f) = 0, f \in \mathcal{F}\}$$

if interpreted in the sense of L^2 convergence.

A. Nonuniform Sampling

Uniform sampling is not well suited for nonpackable signals. However, it turns out that there is a clever way of sampling the signal $x(t)$ called "multicoset sampling" or "periodic nonuniform sampling" at a rate lower than the Nyquist rate, that captures enough information to recover $x(t)$ exactly. Multicoset sampling and reconstruction from the samples will be described more fully in the following sections. First, we survey some known work on nonuniform sampling. Kahn and Liu [8] showed how to represent and reconstruct signals from a multiple-channel sampling scheme. They provide conditions for exact reconstruction from the sampling trains and relate it to the maximum number of overlaps. Their sampling scheme, which is essentially a filter bank, is more general than nonuniform sampling since their "analysis filters" are not required to be simple delays. They express the reconstruction as the solution to a matrix equation, but do not provide an explicit interpolation formula. Cheung and Marks [9], [10] showed that multicoset sampling allows sampling of two–dimensional (2-D) signals below their Nyquist density. A similar treatment for one–dimensional (1-D) and 2-D signals was done by Feng and Bresler [11], and Bresler and Feng [12], respectively, in the broader context of spectrum blind sampling. Filter bank theory and periodic nonuniform sampling was also used to obtain sampling rate reductions in [13]–[16]. Shenoy [17] and Higgins [18] apply multicoset sampling to multiband signals that do not tessellate under translation. Their results indicate that signals with certain spectral supports require a *single* interpolation filter as opposed to *more than one* in the other analyses of the problem. In other words, the sampling expansion is composed of time translates of a single function. Simplicity of its implementation is an obvious advantage of their scheme. However, because it only works for a restricted class of signals, we do not consider their scheme in this paper, focusing instead on multicoset sampling.

Herley and Wong [16], following [8], used filter bank theory instead, to suggest a sampling scheme for minimum rate sampling. They choose the analysis filters of the filter bank to be simple delays, $H_i(z) = z^{-i}$, and then show that some of the analysis channel outputs can be discarded, and yet, the input signal can be reconstructed from the other channels. It is clear that the reconstruction is performed by processing subsamples (obtained nonuniformly) of the original sample train at the Nyquist rates. As the number of channels goes to infinity, the average sampling rate converges to Landau's minimum sampling rate, as expected. In fact, all of the schemes proposed in [8]–[11], [16] achieve the Landau minimum rate asymptotically. Although we do not adopt a filter bank approach and use a different notation, the work in [16] will be the basis for all the analysis in this paper.

In this paper we continue along the lines of [16] to examine the problem of nonuniform sampling. First, we present some new results about the sampling and reconstruction scheme itself. Herley and Wong [16] suggest using an iterative projection onto convex sets (POCS) algorithm to design the reconstruction filters, rather than derive an explicit reconstruction formula. Unlike in their analysis, we provide exact expressions for the interpolation filters, or equivalently the explicit reconstruction formulas for the sampling scheme. Beyond their obvious practical advantage, these expressions are useful for analytical purposes. For instance, they are useful in a) analyzing the reconstruction error of the system; b) quantifying the effects of signal mismodeling, i.e., computing the aliasing error and output noise; and c) optimizing the system for the given class of signals.

B. Error Bounds

Bounds on sensitivity to mismodeling of the signal are important to any sampling scheme. They are particularly important for the sampling schemes considered in this paper because these schemes achieve what is impossible with other schemes: approach the Landau lower bound arbitrarily closely. This raises the question of a possibly increased sensitivity to signal mismodeling and sample noise, leading to an increased reconstruction error. Using our new explicit reconstruction formulas, we derive bounds on the peak amplitude or the energy of the error signal. We compute bounds on the aliasing error that results from input signals in the class of functions $\mathcal{B}([\mathcal{F}])$, which is larger than the class $\mathcal{B}(\mathcal{F})$. We find that the upper bound on the peak aliasing error takes the form

$$\sup_t |x(t) - \hat{x}(t)| \le \psi_\infty \int_{[\mathcal{F}]\setminus\mathcal{F}} |X(f)|\, df$$

as it usually does for various other schemes. The bounding constant ψ_∞ can be used as a performance measure of the system. Different systems can be compared based on their corresponding bounding constants. In particular, Beaty and Higgins [19] derive a similar bound on the aliasing error for packable signals. The bounding constant for their case is $\psi_\infty = 2$. We also derive a bound on the energy of the aliasing error which takes the form

$$\|x - \hat{x}\|_2 \le \psi_2 \sqrt{\int_{[\mathcal{F}]\setminus\mathcal{F}} |X(f)|^2 \, df}.$$

Finally, we derive an expression for the output noise power when the input is contaminated by additive white sample noise with variance σ^2

$$\langle E|x(t) - \tilde{x}(t)|^2\rangle_t = \psi_{\rm n}\sigma^2.$$

It turns out that the constants ψ_∞, ψ_2, and $\psi_{\rm n}$ depend on some parameters that are free to be chosen. An optimal choice of these free parameters would minimize ψ_∞, ψ_2, or $\psi_{\rm n}$. These results can then applied to the design of sampling patterns, but this problem will be addressed elsewhere.

II. Multicoset Sampling

We begin with a few definitions essential to the development and analysis of the sampling scheme. The class of continuous complex-valued of finite energy, band-limited to a real set $\mathcal{F}$ (consisting of a finite union of bounded intervals) is defined by $\mathcal{B}(\mathcal{F})$

$$\mathcal{B}(\mathcal{F}) = \{x(t) \in L^2(\mathbb{R}) \cap C(\mathbb{R}): X(f) = 0,\ f \notin \mathcal{F}\} \quad \text{where } \mathcal{F} = \bigcup_{i=1}^{n} [a_i, b_i) \tag{1}$$

and

$$X(f) = \int_{-\infty}^{\infty} x(t) \exp(-j2\pi f t)\, dt$$

is the Fourier transform of $x(t)$. The *span of* $\mathcal{F}$, denoted by $[\mathcal{F}]$, represents the convex hull of $\mathcal{F}$, i.e., the smallest interval containing $\mathcal{F}$.

Let $x(t) \in \mathcal{B}(\mathcal{F})$. We shall assume, with no loss of generality, that $\inf \mathcal{F} = 0$. In multicoset sampling, we first pick a suitable sampling period T (such that uniform sampling at rate $1/T$ causes no aliasing), and a suitable integer $L > 0$, and then sample the input signal $x(t)$ nonuniformly at the instants $t = (nL + c_i)T$ for $1 \le i \le p$ and $n \in \mathbb{Z}$. The set $\{c_i\}$ contains p distinct integers chosen from the set $\mathcal{L} \stackrel{\rm def}{=} \{0, 1, \ldots, L-1\}$. The sampling process just described can be viewed as first sampling the signal at the "base sampling rate" of $1/T$ and then discarding all but p samples in every block of L samples periodically. The samples that are retained in each block are specified by $\{c_i\}$. The base sampling rate could be chosen equal to the Nyquist rate, i.e., $1/T = f_{\rm nyq}$, but never lower. However, we choose $1/T = \lambda([\mathcal{F}])$, because, sampling at this rate always guarantees no aliasing for any $\mathcal{F}$.

For a given c_i, it is clear that the coset of sampling instants $t = (nL + c_i)T,\ n \in \mathbb{Z}$ is uniform with intersample spacing equal to LT. We call this the ith *active coset*. The set $\mathcal{C} = \{c_i: 1 \le i \le p\}$ is referred to as an (L, p) *sampling pattern* and the integer L as the period of the pattern. Fig. 2 shows two multicoset sampling patterns corresponding to parameters $(L, p) = (5, 3)$. All other patterns for these parameters can be obtained by cyclic shifts of the patterns shown, namely, $\mathcal{C} = \{0, 1, 2\}$ and $\mathcal{C} = \{0, 1, 3\}$. Patterns related to each other by cyclic shifts with or without reflections are essentially equivalent in terms of the associated reconstruction problems and their error sensitivities.

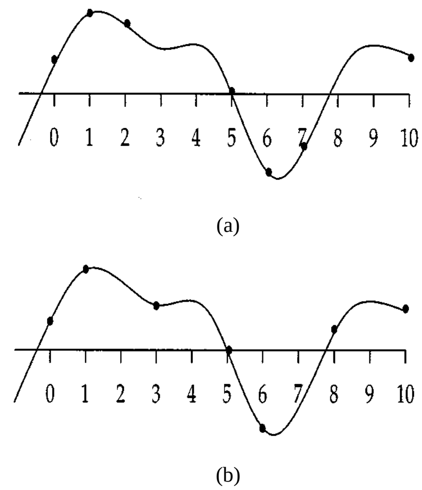

Fig. 2. Two distinct sampling patterns for $(L, p) = (5, 3)$. (a) $\mathcal{C} = \{0, 1, 2\}$. (b) $\mathcal{C} = \{0, 1, 3\}$.

Now, consider the following L discrete-time sequences obtained by zeroing out all samples $\{x(nT)\}$ except those at $t = (mL + l)T,\ m \in \mathbb{Z}$

$$\underline{x}_l(n) \stackrel{\rm def}{=} x(nT) \sum_{m \in \mathbb{Z}} \delta(n - (mL + l)), \qquad 0 \le l < L$$

where $\delta(n)$ is the Kronecker delta function. It is clear that the sequence $\underline{x}_{c_i}(n)$ contains the samples of the ith active coset with samples separated by $L - 1$ interleaving zeros. It is straightforward to verify that the discrete-time Fourier transform of the lth sequence is

$$\begin{aligned}\underline{X}_l(e^{j2\pi fT}) &= \sum_{n=-\infty}^{\infty} \underline{x}_l(n) \exp(-j2\pi n f T) \\ &= \frac{1}{LT} \sum_{r \in \mathbb{Z}} X\left(f + \frac{r}{LT}\right) \exp\left(\frac{j2\pi r l}{L}\right)\end{aligned} \tag{2}$$

which, using the fact that $X(f) = 0$ for $f \notin [0, 1/T)$, gives us

$$\underline{X}_l(e^{j2\pi fT}) = \frac{1}{LT} \sum_{r=0}^{L-1} X_r(f) \exp\left(\frac{j2\pi r l}{L}\right), \quad f \in \mathcal{F}_0 \tag{3}$$

where $X_r(f)$ is defined as

$$X_r(f) \stackrel{\rm def}{=} X\left(f + \frac{r}{LT}\right) \chi(f \in \mathcal{F}_0) \tag{4}$$

$$\mathcal{F}_0 \stackrel{\rm def}{=} \left[0, \frac{1}{LT}\right) \tag{5}$$

and $\chi(f \in \mathcal{H})$ denotes the indicator function of a real set $\mathcal{H}$ consisting of a finite union of bounded intervals

$$\chi(f \in \mathcal{H}) \stackrel{\rm def}{=} \begin{cases} 1, & \text{if } f \in \mathcal{H} \\ 0, & \text{if } f \notin \mathcal{H}. \end{cases}$$

In other words, the "spectral component" $X_r(f)$ is obtained by first using an ideal bandpass filter to extract the signal in the frequency range $r/LT \le f < r+1/LT$ and then performing a frequency shift to the left by r/LT units. Denoting the inverse Fourier transform of $X_r(f)$ by $x_r(t)$, it is evident from the above definition that

$$x(t) = \sum_{r=0}^{L-1} x_r(t) \exp\left(\frac{j2\pi rt}{LT}\right). \tag{6}$$

Another result that can be deduced from (2) or directly from the definition of $\underline{x}_l(nT)$ is

$$\underline{X}_l(e^{j2\pi(f+\frac{r}{LT})T}) = e^{(-j2\pi\frac{lr}{L})}\underline{X}_l(e^{j2\pi fT}), \qquad r \in \mathbb{Z}. \tag{7}$$

We will use (6) and (7) later in deriving the reconstruction equations. We now let $l = c_i, i = 1, 2, \dots, p$ in (3) to get

$$\underline{X}_{c_i}(e^{j2\pi fT}) = \frac{1}{LT}\sum_{r=0}^{L-1} \exp\left(\frac{j2\pi c_i r}{L}\right) X_r(f), \qquad f \in \mathcal{F}_0 \tag{8}$$

This is the main equation relating the spectral components $X_r(f)$ to the information contained in the observed samples. Note that (8) on the interval $\mathcal{F}_0$, contains all the relevant information present in the samples since, from (7), $X_{c_i}(e^{j2\pi ft})$ is essentially "periodic" with period $1/LT$. Reconstruction of the original signal $x(t)$ is achieved if we recover its spectral components $\{x_r(t)\}$.

III. Reconstruction

We now focus on the problem of reconstructing $x(t) \in \mathcal{B}(\mathcal{F})$ from its multicoset samples. Herley and Wong [16] considered the analogous problem for real signals, but did not, however, provide an explicit reconstruction formula or system. We shall derive the reconstruction equations formally and devise a multirate system to perform the reconstruction.

Our objective is to invert the set of linear equations (8) to obtain $X_r(f)$. The recovery of $x(t)$ is then merely an application of (6). Notice that, if $\mathcal{C}$ and $\mathcal{F}$ satisfy certain conditions, the inversion of (8) can be accomplished even though there are fewer equations (p) than unknowns (L) for each $f \in \mathcal{F}_0$. This is possible because our signals belong to $\mathcal{B}(\mathcal{F})$, which is smaller than $\mathcal{B}([\mathcal{F}])$.

Let $\mathcal{F}$ be the union of n bounded intervals as in (1) with the additional assumption that

$$0 = a_1 < b_1 < a_2 < b_2 < \cdots < a_n < b_n = \frac{1}{T}$$

made with no loss of generality. Consider the finite set Γ defined below

$$\Gamma \stackrel{\text{def}}{=} \left\{a_i - \frac{\lfloor LTa_i \rfloor}{LT} : 1 \le i \le n\right\} \cup \left\{b_i - \frac{\lfloor LTb_i \rfloor}{LT} : 1 \le i \le n\right\} \tag{9}$$

where $\lfloor\cdot\rfloor$ is the floor function. Let $\Gamma = \{\gamma_1, \gamma_2, \dots, \gamma_M\}$ be the $M \le 2n$ elements of Γ arranged in increasing order. We have $\gamma_1 = 0$ as a consequence of $a_1 = 0$. We define $\gamma_{M+1} \stackrel{\text{def}}{=} 1/LT$ to get

$$0 = \gamma_1 < \gamma_2 < \cdots < \gamma_{M+1} = \frac{1}{LT}$$

and a collection of intervals $\{\mathcal{G}_m\}$ that partitions the set $\mathcal{F}_0$

$$\mathcal{G}_m = [\gamma_m, \gamma_{m+1}), \qquad 1 \le m \le M.$$

We prove (in the Appendix) the following fact involving the indicator function of $\mathcal{F}$ and these sets $\{\mathcal{G}_m\}$.

Lemma 1: For each $r \in \mathcal{L}$ and $m \in \{1, 2, \dots, M\}$ the function $\chi(f + r/LT \in \mathcal{F})$ is constant over the interval $\mathcal{G}_m$.

Equivalently, the theorem states that each of the "subcells" $(r/LT) \oplus \mathcal{G}_m$ for $l \in \mathcal{L}$ is either fully contained in $\mathcal{F}$ or disjoint from it. This interpretation of the theorem motivates the following definition of the "spectral index sets" $\mathcal{K}_m$ and their complements $\mathcal{K}_m^c$ for $m \in \{1, 2, \dots, M\}$:

$$\mathcal{K}_m \stackrel{\text{def}}{=} \left\{r \in \mathcal{L} : \frac{r}{LT} \oplus \mathcal{G}_m \subset \mathcal{F}\right\} \quad \text{and} \quad \mathcal{K}_m^c = \mathcal{L} \backslash \mathcal{K}_m.$$

The index set $\mathcal{K}_m$ tells us which subcells in the collection $\{(r/LT) \oplus \mathcal{G}_m : r \in \mathcal{L}\}$ are "active," while $\mathcal{K}_m^c$ indicates which of them are not. The following theorem, which is a restatement of the main theorem in [16] using our notation, provides a necessary condition for reconstruction:

Theorem 1: Equation (8) admits a unique solution for $X_r(f)$ only if the indicator function of $\mathcal{F}$ satisfies

$$q(f) \stackrel{\text{def}}{=} \sum_{r=0}^{L-1} \chi\left(f + \frac{r}{LT} \in \mathcal{F}\right) \le p, \qquad f \in \mathcal{F}_0. \tag{10}$$

Furthermore, an equality in (10) is necessary for attaining Landau's lower bound on the sampling rate.

To make the paper self-contained, we provide a Proof of Theorem 1 in the Appendix. Since, by Lemma 1, $q(f)$ is constant on each $\mathcal{G}_m$, the inequality (10) reduces to

$$p \ge \max_m \, q_m, \qquad \text{where } q_m = q(f), \; f \in \mathcal{G}_m \tag{11}$$

Evidently, q_m is the cardinality of the set $\mathcal{K}_m$. We denote the elements of $\mathcal{K}_m$ and $\mathcal{K}_m^c$ by

$$\{k_m(l) : 1 \le l \le q_m\}$$

and

$$\{k_m^c(l) : 1 \le l \le L - q_m\}$$

respectively. Later, we shall see that, for a suitable choice of $\mathcal{C}$, (11) is also sufficient for unique reconstruction. In the following example, we show how to construct the relevant $\mathcal{K}$-sets for the spectrum illustrated in Fig. 3.

Example 1: Let the spectral support of our class of signals be $\mathcal{F} = [0, 1.3) \cup [2.7, 3.7) \cup [4.5, 5)$. Comparing this with (1), we find that

$$a_1 = 0, \; a_2 = 2.7, \; a_3 = 4.5, \; b_1 = 1.3, \; b_2 = 3.7, \; \text{and } b_3 = 5.$$

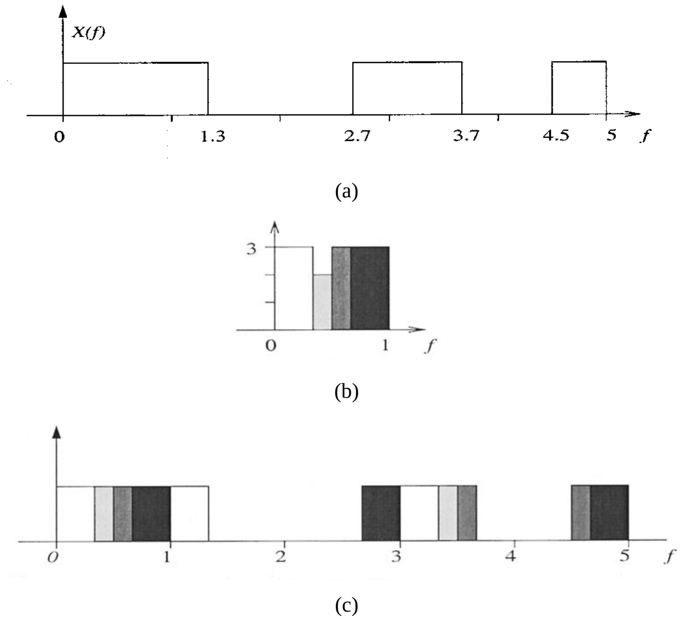

Fig. 3. (a) Indicator function of the spectral support $\mathcal{F}$ of span $1/T = 5$. (b) Left-hand side of (10) gives the number of overlapping pieces of the sampled spectrum (for $L = 5$) which is constant on each of the sets $\mathcal{G}_m, m = 1, \ldots, 4$. (c) The active spectral subcells. For each $m = 1, \ldots, 4$, the translates $(r/LT) \oplus \mathcal{G}_m, r \in \mathcal{K}_m$ are shown in the same color.

The indicator function $\chi(f \in \mathcal{F})$ of the set $\mathcal{F}$ is shown in Fig. 3(a). The length of $\mathcal{F}$ is $\lambda(\mathcal{F}) = 2.8$ and its span is $[\mathcal{F}] = [0, 5)$. It can be checked easily that $\mathcal{F}$ is nonpackable and hence the Nyquist rate for signals in $\mathcal{B}(\mathcal{F})$ is $1/T = \lambda[\mathcal{F}] = 5$. Yet Landau's lower bound gives a rate of 2.8 and the corresponding occupancy is $\Omega = 2.8/5 = 0.56$. Suppose we pick $L = 5$. We see that (9) yields

$$\Gamma \equiv \{\gamma_0, \gamma_1, \gamma_2, \gamma_3\} = \{0, 0.3, 0.5, 0.7\}$$

containing $M = 4$ elements, from which we construct sets $\mathcal{G}_m$ that partition $\mathcal{F}_0 = [0, 1/LT) = [0, 1)$

$$\mathcal{G}_1 = [0, 0.3) \qquad \mathcal{G}_2 = [0.3, 0.5)$$
$$\mathcal{G}_3 = [0.5, 0.7) \qquad \mathcal{G}_4 = [0.7, 1).$$

The following are immediately apparent:

$$\begin{aligned} q_1 &= 3, \quad \mathcal{K}_1 = \{0, 1, 3\}, \quad \mathcal{K}_1^c = \{2, 4\}. \\ q_2 &= 2, \quad \mathcal{K}_2 = \{0, 3\}, \qquad \mathcal{K}_2^c = \{1, 2, 4\}. \\ q_3 &= 3, \quad \mathcal{K}_3 = \{0, 3, 4\}, \quad \mathcal{K}_3^c = \{1, 2\}. \\ q_4 &= 3, \quad \mathcal{K}_4 = \{0, 2, 4\}, \quad \mathcal{K}_4^c = \{1, 3\}. \end{aligned}$$

Fig. 3(b) shows the left-hand side of (10) plotted on the interval $[0, 1)$. Note that this function is piecewise-constant, equal to q_m on each of the intervals $\{\mathcal{G}_m\}$ which are color-coded for convenience. Fig. 3(c) shows the indicator function of $\mathcal{F}$ color-coded to show the active subcells derived by translating $\mathcal{G}_m$.

In direct analogy to (4), we define

$$X_{k_m(l)}(f) \stackrel{\text{def}}{=} X\left(f + \frac{k_m(l)}{LT}\right)\chi(f \in \mathcal{G}_m). \tag{12}$$

However, beware of this definition since it is not equal to (4) evaluated at $r = k_m(l)$. The extra factor of $\chi(f \in \mathcal{G}_m)$ in (12) makes it different. The "spectral component" $X_{k_m(l)}(f)$ is obtained by shifting the values of $X(f)$ on the subcell $(k_m(l)/LT) \oplus \mathcal{G}_m$ to the origin. For each m, we define a $p \times q_m$ matrix $\boldsymbol{A}_m$, and vectors $\boldsymbol{y}(\boldsymbol{f}) \in \mathbb{C}^p$, $\boldsymbol{x}_m^+(f) \in \mathbb{C}^{q_m}$ and $\boldsymbol{x}_m^-(f) \in \mathbb{C}^{L-q_m}$ as follows:

$$\begin{aligned} [\boldsymbol{y}(f)]_i &= (T\sqrt{L})\underline{X}_{c_i}(e^{j2\pi fT})\chi(f \in \mathcal{F}_0) \\ [\boldsymbol{x}_m^+(f)]_l &= X_{k_m(l)}(f) \\ [\boldsymbol{x}_m^-(f)]_l &= X_{k_m^c(l)}(f) \\ [\boldsymbol{A}_m]_{il} &= \frac{1}{\sqrt{L}} \exp\left(\frac{j2\pi c_i k_m(l)}{L}\right). \end{aligned} \tag{13}$$

Note that $\boldsymbol{A}_m$ is the submatrix of the $L \times L$ DFT matrix W_L obtained by extracting its rows indexed by $\mathcal{C}$, and columns indexed by $\mathcal{K}_m$. We denote this by $\boldsymbol{A}_m = \boldsymbol{W}_L(\mathcal{C}, \mathcal{K})$. Next, using the fact that $\boldsymbol{x}_m^-(f) = 0$ whenever $x(t) \in \mathcal{B}(\mathcal{F})$, we can rewrite (8) in matrix form over each subcell $\mathcal{G}_m$ as follows:

$$\boldsymbol{y}(f) = \boldsymbol{A}_m \boldsymbol{x}_m^+(f) \qquad \forall f \in \mathcal{G}_m,\ 1 \le m \le M. \tag{14}$$

At this point, we introduce the following two definitions that characterize the sampling pattern $\mathcal{C}$, of size p, in terms of the $L \times L$ DFT matrix $\boldsymbol{W}_L$.

Definition 1: Given an index set $\mathcal{K}$ with $|\mathcal{K}| = q \le p$, we call $\mathcal{C}$ a $\mathcal{K}$-reconstructive sampling pattern if the matrix $\boldsymbol{W}_L(\mathcal{C}, \mathcal{K})$ has full row rank.

Definition 2: A pattern $\mathcal{C}$ with $|\mathcal{C}| = p \ge q$ is (p, q) universal if the matrix $\boldsymbol{W}_L(\mathcal{C}, \mathcal{K})$ has full row rank for every index set of q elements, i.e., whenever $|\mathcal{K}| = q$. A (p, p)-universal pattern is simply called universal.

For fixed values of p and q, the second definition is stronger than the first. For every L and $p \le L$, there always exists a (p, p) universal pattern. The "bunched" sampling pattern $\mathcal{C} = \{0, 1, \ldots, p-1\}$ is an example since the resulting matrix $\boldsymbol{A} = \boldsymbol{W}_L(\mathcal{C}, \mathcal{K})$ is a Vandermonde matrix for any choice of $\mathcal{K}$. This guarantees $\operatorname{rank}(\boldsymbol{A}) = q$ for *any* spectral support with $q \le p$ active cells.

Equation (11) together with the assumption that $\boldsymbol{A}_m$ has full rank for each m (i.e., $\mathcal{C}$ being $\boldsymbol{K}_m$-reconstructive for each m) is necessary and sufficient for reconstruction. A simpler sufficient condition is universality of $\mathcal{C}$. For convenience, we assume throughout that $\mathcal{C}$ is universal. This guarantees the existence of left-inverses $\boldsymbol{A}_m^{-1}$ of $\boldsymbol{A}_m$. Therefore, inverting (14) gives

$$\begin{aligned} \boldsymbol{x}_m^+(f) &= \boldsymbol{A}_m^{-1}\boldsymbol{y}(f) \\ \boldsymbol{x}_m^-(f) &= \boldsymbol{C}_m \boldsymbol{y}(f) \end{aligned} \qquad f \in \mathcal{G}_m, m \in \{1, \ldots, M\} \tag{15}$$

where, in order that $\boldsymbol{x}_m^-(f) = 0$ hold, $\boldsymbol{C}_m$ is any $(L - q_m) \times p$ matrix satisfying

$$\boldsymbol{C}_m \boldsymbol{A}_m = \boldsymbol{0} \tag{16}$$

for each m. The matrices $\boldsymbol{A}_m^{-1}$ and $\boldsymbol{C}_m$ are nonunique unless $p = q_m$. In other words, there is some freedom that can be used in designing a reconstruction system. The actual choice of matrices does not affect the reconstruction, but does influence the bounds on aliasing error, as described later. This suggests that finding the optimal matrices is of some interest. These equations

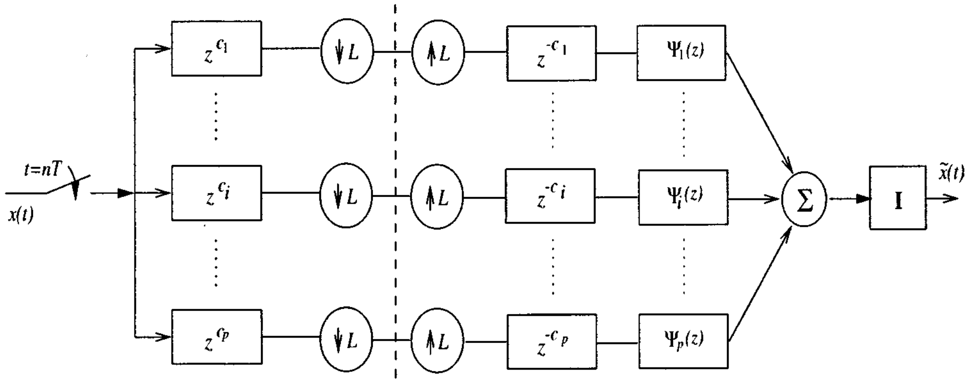

Fig. 4. Multicoset sampling and reconstruction. The block "I" is an ideal sinc interpolator.

specify all the information required to reconstruct the spectrum $X(f)$ on all its spectral subcells, and hence $x(t)$ itself.

The interpolation equation may be calculated from (15) in a rather messy but straightforward manner. The result is summarized in the theorem below, and we prove it in the Appendix.

Theorem 2: Let $x(t) \in \mathcal{B}(\mathcal{F})$ be sampled on an (L, p) multicoset pattern $\mathcal{C}$, and $\mathcal{K}_m, m = 1, 2, \ldots, M$ be the spectral index sets of $\mathcal{F}$. Then, if $\mathcal{C}$ is $\mathcal{K}_m$-reconstructive for each m, $x(t)$ can be uniquely interpolated from its multicoset samples according to the following formula:

$$\begin{aligned} x(t) &= \sum_{i=1}^{p} \sum_{n=-\infty}^{\infty} \underline{x}_{c_i}(nT)\phi_i(t - nT) \\ &\equiv \sum_{i=1}^{p} \sum_{j=-\infty}^{\infty} x((c_i + Lj)T)\phi_i(t - (c_i + Lj)T) \end{aligned} \quad (17)$$

where the functions $\phi_i(t), i = 1, \ldots, p$ have Fourier transforms $\Phi_i(f)$ that are piecewise-constant on $[\mathcal{F}]$

$$\Phi_i(f) = \begin{cases} T\sqrt{L}[\boldsymbol{A}_m^{-1}]_{li} e^{j2\pi \frac{c_i k_m(l)}{L}}, & \text{if } f \in \dfrac{k_m(l)}{LT} \oplus \mathcal{G}_m \\ T\sqrt{L}[\boldsymbol{C}_m]_{li} e^{j2\pi \frac{c_i k_m^c(l)}{L}}, & \text{if } f \in \dfrac{k_m^c(l)}{LT} \oplus \mathcal{G}_m. \end{cases} \quad (18)$$

Corollary 1: The result in Theorem 2 holds if $\max_m q_m \le p$, and $\mathcal{C}$ is universal.

The reconstruction scheme is illustrated in Fig. 4. In the figure $\Psi_i(z)$ is a digital filter whose impulse response is $\phi_i(nT)$. The filters used are ideal. In practice, causal, possibly finite impulse response (FIR) approximations are used, introducing some delay and distortion. The analysis of the resulting error is analogous to the truncation error in classical cardinal series expansion and is beyond the scope of this paper. Instead, we assume that the filters are ideal and concentrate on the aliasing errors due to signal mismodeling.

IV. Error Bounds

For a system designed to sample and reconstruct signals in $\mathcal{B}(\mathcal{F})$, it is necessary that (11) hold. In this case, the reconstruction of $x(t)$ would be exact. However, if $x(t) \notin \mathcal{B}(\mathcal{F})$ then the signal $\tilde{x}(t)$ reconstructed using the right-hand side of (17) would be in error. For example, this would happen if, in the system design, we underestimated the spectral support of signals we expected to encounter, i.e., if we would choose to ignore certain frequencies that contained negligible signal energy. The purpose of this section is to obtain bounds on the aliasing error $e(t) = \tilde{x}(t) - x(t)$ resulting from an underestimation of the spectral support.

In the following analysis $X(f)$ need not vanish on $[\mathcal{F}]\backslash\mathcal{F}$ although we assume, for simplicity, that $X(f) = 0$ for $f \notin [\mathcal{F}]$. In other words, we assume that the spectral span $[\mathcal{F}]$ is correctly specified, but the multiband structure to which $x(t)$ is bandlimited within $[\mathcal{F}]$ may be misspecified. We shall first derive bounds on the sup- and 2-norms of the aliasing error $e(t)$ for the nonuniform sampling process described in the last section.

Recall that $\mathcal{K}_m^c = \{k_m^c(l): 1 \le l \le L - q_m\}$ tells us exactly which spectral subcells in the collection $\{(r/LT) \oplus \mathcal{G}_m : r \in \mathcal{L}\}$ are inactive. Now, for each $m \in \{1, 2, \ldots, M\}$, let $\boldsymbol{B}_m = \boldsymbol{W}_L(\mathcal{C}, \mathcal{K}_m)$ denote the $p \times (L - q_m)$ submatrix of the $L \times L$ DFT matrix $\boldsymbol{W}_L$, with rows and columns indexed by $\mathcal{C}$ and $\mathcal{K}_m^c$ respectively, i.e.,

$$[\boldsymbol{B}_m]_{il} = \frac{1}{\sqrt{L}} \exp\left(\frac{j2\pi c_i k_m^c(l)}{L}\right). \quad (19)$$

We can then rewrite (8) in matrix form as

$$\boldsymbol{y}(f) = \boldsymbol{A}_m \boldsymbol{x}_m^+(f) + \boldsymbol{B}_m \boldsymbol{x}_m^-(f), \qquad f \in \mathcal{G}_m \quad (20)$$

where $\boldsymbol{x}_m^-(f)$ is the $(L - q_m) \times 1$ vector defined in (13). If $x(t) \in \mathcal{B}(\mathcal{F})$ then $\boldsymbol{x}_m^-(f)$ would vanish. Denoting the reconstructed signal by $\tilde{x}(t)$, it immediately follows from (15), (16), and (20) that

$$\begin{aligned} \tilde{\boldsymbol{x}}_m^+(f) &= \boldsymbol{x}_m^+(f) + \boldsymbol{A}_m^{-1}\boldsymbol{B}_m \boldsymbol{x}_m^-(f) \\ \tilde{\boldsymbol{x}}_m^-(f) &= \boldsymbol{C}_m \boldsymbol{B}_m \boldsymbol{x}_m^-(f) \end{aligned} \qquad f \in \mathcal{G}_m \quad (21)$$

where $\tilde{\boldsymbol{x}}_m^+(f)$ and $\tilde{\boldsymbol{x}}_m^-(f)$ have definitions analogous to $\boldsymbol{x}_m^+(f)$ and $\boldsymbol{x}_m^-(f)$, respectively. We define matrices $\boldsymbol{D}_m$ and $\boldsymbol{F}_m$ (of sizes $q_m \times (L - q_m)$ and $(L - q_m) \times (L - q_m)$, respectively) for each m

$$\begin{aligned} \boldsymbol{D}_m &= \boldsymbol{A}_m^{-1}\boldsymbol{B}_m \\ \boldsymbol{F}_m &= \boldsymbol{C}_m \boldsymbol{B}_m - \boldsymbol{I}. \end{aligned} \quad (22)$$

In the following subsections, we present bounds on the peak aliasing error, the aliasing error energy, and evaluate the performance of the system in the presence of input noise.

A. The Sup-Norm of the Error

The following theorem provides the time-domain expression for the aliasing error.

Theorem 3: The aliasing error $e(t)$ takes the form

$$e(t) = \sum_{m=1}^{M} \sum_{l=1}^{L-q_m} \mu_{m,l}(t) x_{k_m^c(l)}(t)$$

where $\mu_{m,l}(t)$ for each $1 \le m \le M$ and $1 \le l \le L - q_m$ are continuous, LT-periodic functions defined as

$$\mu_{m,l}(t) = \sum_{r=1}^{L-q_m} [\boldsymbol{F}_m]_{rl} e^{\frac{j2\pi k_m^c(r)t}{LT}} + \sum_{r=1}^{q_m} [\boldsymbol{D}_m]_{rl} e^{\frac{j2\pi k_m(r)t}{LT}}.$$

Furthermore, the peak value of $e(t)$ satisfies the tight bound

$$\sup_t |e(t)| \le \psi \int_{[\mathcal{F}]\setminus\mathcal{F}} |X(f)|\, df$$
$$\psi \stackrel{\text{def}}{=} \max_m \Big(\max_{1\le l \le L-q_m, t\in[0,LT]} |\mu_{m,l}(t)| \Big).$$

We prove this theorem in the Appendix. The tightness of this bound is proved by demonstrating an input signal that satisfies the equality in the bound. Note that the constant ψ can be bounded from above as follows:

$$\begin{aligned}\psi &= \max_{m,l,t} \left| \sum_{r=1}^{L-q_m} [\boldsymbol{F}_m]_{rl} \exp\left(\frac{j2\pi k_m^c(r)t}{LT}\right) + \sum_{r=1}^{q_m} [\boldsymbol{D}_m]_{rl} \exp\left(\frac{j2\pi k_m(r)t}{LT}\right) \right| \\ &\le \max_m \left(\max_l \left(\sum_{r=1}^{L-q_m} |[\boldsymbol{F}_m]_{rl}| + \sum_{r=1}^{q_m} |[\boldsymbol{D}_m]_{rl}| \right) \right) \\ &= \max_m \left\| \begin{pmatrix} \boldsymbol{D}_m \\ \boldsymbol{F}_m \end{pmatrix} \right\|_1\end{aligned}$$

where $\|\cdot\|_1$ is the maximum-column-sum norm for matrices (or the ℓ_1 norm for column vectors.) Hence we obtain the weaker, but more tractable bound

$$\sup_t |e(t)| \le \psi_\infty \int_{[\mathcal{F}]\setminus\mathcal{F}} |X(f)|\, df, \quad \text{where } \psi_\infty = \max_m \left\| \begin{pmatrix} \boldsymbol{D}_m \\ \boldsymbol{F}_m \end{pmatrix} \right\|_1. \tag{23}$$

B. 2-Norm of the Error

Theorem 4: The energy of the aliasing error is bounded by

$$\int_{-\infty}^{\infty} e^2(t)\, dt \le \max_m [\lambda_{\max}(\boldsymbol{F}_m^* \boldsymbol{F}_m + \boldsymbol{D}_m^* \boldsymbol{D}_m)] \mathcal{E}_{\text{out}}$$
$$\int_{-\infty}^{\infty} e^2(t)\, dt \ge \min_m [\lambda_{\min}(\boldsymbol{F}_m^* \boldsymbol{F}_m + \boldsymbol{D}_m^* \boldsymbol{D}_m)] \mathcal{E}_{\text{out}}$$

where $\mathcal{E}_{\text{out}}$, the out-of-band energy, equals

$$\mathcal{E}_{\text{out}} = \int_{[\mathcal{F}]\setminus\mathcal{F}} |X(f)|^2\, df$$

with both bounds being tight.

Once again, the proof can be found in the Appendix. We are particularly interested in the upper bound on the 2-norm of the error

$$\|e(t)\|_2 \stackrel{\text{def}}{=} \|e\|_2 = \sqrt{\textstyle\int |e(t)|^2\, dt}.$$

It is clearly related to the spectral norm of the matrix composed of $\boldsymbol{F}_m$ and $\boldsymbol{D}_m$ as

$$\|e\|_2 \le \psi_2 \sqrt{\mathcal{E}_{\text{out}}}, \qquad \text{where } \psi_2 = \max_m \left\| \begin{pmatrix} \boldsymbol{D}_m \\ \boldsymbol{F}_m \end{pmatrix} \right\|_2. \tag{24}$$

C. Performance in the Presence of Noise

Finally, we consider the effect of additive white sample noise, representing, e.g., quantization noise. The sampled signal can be modeled as

$$\overline{x}(nT) = x(nT) + w(nT)$$

where $w(n)$ is a noise process with

$$E[w(mT)w(nT)] = \sigma^2 \delta(m-n)$$

and $x(t)$ is the actual signal we would like to be sampling. Owing to its linearity, (17) directly gives us the following expression for the output noise $\tilde{w}(t)$ which is independent of $x(t)$:

$$\tilde{w}(t) = \sum_{i=1}^{p} \sum_{j=-\infty}^{\infty} w((c_i + Lj)T) \phi_i(t - (c_i + Lj)T).$$

Theorem 5: The output noise $\tilde{w}(t)$ is possibly nonstationary, with average power given by

$$\langle E|\tilde{w}(t)|^2 \rangle_t = \sigma^2 \psi_{\mathbf{n}}, \quad \text{where } \psi_{\mathbf{n}} = T \sum_{m=1}^{M} \lambda(\mathcal{G}_m)(\|\boldsymbol{A}_m^{-1}\|_F^2 + \|\boldsymbol{C}_m\|_F^2).$$

The norm $\|\cdot\|_F$ represents the Frobenius norm. Theorem 5 is proved in the Appendix.

V. Conclusion

We have presented the analysis of a scheme for sampling multiband signals below the Nyquist rate. The sampling scheme uses multicoset sampling and achieves the Landau minimum sampling rate in the limit $L \to \infty$, where L is the period of the sampling pattern. However, for many spectra, the minimum rate can be achieved for a finite L. Typically, this scheme is useful for sampling signals with sparse and nonpackable spectra.

We determined necessary and sufficient conditions for the reconstruction of a multiband signal from its multicoset samples and derived an explicit reconstruction equation. There are free parameters in the reconstruction equation when the Landau minimum rate is not achieved for the particular L chosen. We computed bounds on the aliasing error occurring in the event that the signal lies outside the valid class of multiband signals and determined the sensitivity of the system to input sample noise. The constants in the bounds and the noise-sensitivity factor reveal that some sampling patterns are better than others. In other words, these bounds, which quantify the goodness of sampling patterns, can be minimized to produce the optimal sampling pattern and parameter choice in the reconstruction formula.

APPENDIX

Lemma 2: If u, v are integers and $\alpha, \beta \in [0, 1/LT)$ satisfy

$$\frac{u}{LT} + \alpha < \frac{v}{LT} + \beta \tag{A.1}$$

then $u \leq v$. Further if $u = v$ then $\alpha < \beta$ follows trivially.

Proof: Equation (A.1) implies that

$$(u - v)/LT < \beta - \alpha < 1/LT$$

where the strictness of the second inequality comes from the fact that $[0, 1/LT)$ is open on the right. Hence $u - v < 1$, or equivalently, $u \leq v$. □

Proof of Lemma 1: Let r and m be fixed. Then for each $i \in \{1, 2, \ldots, n\}$ we can express a_i and b_i uniquely as

$$a_i = \frac{u_i}{LT} + \alpha_i \quad \text{and} \quad b_i = \frac{v_i}{LT} + \beta_i \tag{A.2}$$

where u_i, v_i are integers and α_i, β_i are elements of Γ with $\alpha_i, \beta_i \in [0, 1/LT)$. Now we shall prove by contradiction that exactly one of the two conditions

$$(r/LT) \oplus \mathcal{G}_m \subset [a_i, b_i)$$

and

$$(r/LT) \oplus \mathcal{G}_m \cap [a_i, b_i) = \emptyset$$

holds. It is clear that both statements cannot be true simultaneously. So if neither of them holds, then

$$\gamma_m + r/LT < a_i < \gamma_{m+1} + r/LT$$

or

$$\gamma_m + r/LT < b_i < \gamma_{m+1} + r/LT$$

or both must hold. If the first condition holds then (A.2) along with Lemma 2 implies that $r \leq u_i \leq r$. Therefore, $r = u_i$ and $\gamma_m < \alpha_i < \gamma_{m+1}$. The last observation contradicts $\alpha_i \in \Gamma$ because the γ's are arranged in increasing order. Similarly, the second statement above would also lead to a contradiction. This proves our claim that the "subcell" $(r/LT) \oplus \mathcal{G}_m$ is either contained in or disjoint from $[a_i, b_i)$. This fact, being true for each i, now implies that either $(f + r/LT) \in \mathcal{F}$ for *every* $f \in \mathcal{G}_m$ or it is so for *no* such f. Therefore, $\chi(f + r/LT \in \mathcal{F})$ is constant over the interval $\mathcal{G}_m$. □

Proof of Theorem 1: Observe that the quantity $q(f)$ in (10) equals the number of nonzero entries in $\{X_r(f): r = 0, 1, \ldots, L-1\}$. Hence the summation in (8) (repeated below) contains $q(f)$ nonzero terms, for each $f \in \mathcal{F}_0$

$$\underline{X}_{c_i}(e^{j2\pi fT}) = \frac{1}{LT} \sum_{r=0}^{L-1} \exp\left(\frac{j2\pi c_i r}{L}\right) X_r(f), \qquad i = 1, \ldots, p. \tag{A.3}$$

These equations form a set of p linear equations with $q(f)$ unknown variables on the right-hand side, and solving them requires $p \geq q(f)$. Hence, $p \geq q(f)$, $\forall\ f \in \mathcal{F}_0$ is necessary for reconstruction of the spectral components $X_r(f)$. Next as a consequence of $p \geq q(f)$ and the definition of $q(f)$ in (10), we can bound the average sampling density p/LT from below by the Landau minimum rate

$$\frac{p}{LT} \geq \int_0^{\frac{1}{LT}} q(f)\, df = \lambda(\mathcal{F})$$

with equality holding if and only if $p = q(f)$ for all $f \in \mathcal{F}_0$. □

Proof of Theorem 2: To derive the interpolation equations, we begin by expressing (15) in scalar form

$$\left.\begin{aligned} [\boldsymbol{x}_m^+(f)]_l &= \sum_{i=1}^{p} [\boldsymbol{A}_m^{-1}]_{li} [\boldsymbol{y}(f)]_i \\ [\boldsymbol{x}_m^-(f)]_l &= \sum_{i=1}^{p} [\boldsymbol{C}_m]_{li} [\boldsymbol{y}(f)]_i \end{aligned}\right\} \quad f \in \mathcal{G}_m.$$

We use the expressions for $\boldsymbol{y}(f)$, $\boldsymbol{x}_m^+(f)$, and $\boldsymbol{x}_m^-(f)$ from (13) to obtain

$$X_{k_m(l)}(f) = T\sqrt{L} \sum_{i=1}^{p} [\boldsymbol{A}_m^{-1}]_{li} \underline{X}_{c_i}(e^{j2\pi fT}) \chi(f \in \mathcal{G}_m), \qquad 1 \leq l \leq q_m$$

$$X_{k_m^c(l)}(f) = T\sqrt{L} \sum_{i=1}^{p} [\boldsymbol{C}_m]_{li} \underline{X}_{c_i}(e^{j2\pi fT}) \chi(f \in \mathcal{G}_m), \qquad 1 \leq l \leq L - q_m$$

for $f \in \mathcal{G}_m$ and each m. Or equivalently, using (12) we have

$$X(f) = \begin{cases} T\sqrt{L} \displaystyle\sum_{i=1}^{p} [\boldsymbol{A}_m^{-1}]_{li} \underline{X}_{c_i}(e^{j2\pi (f - \frac{k_m(l)}{LT})T}), & \text{if } f \in \dfrac{k_m(l)}{LT} \oplus \mathcal{G}_m \\ T\sqrt{L} \displaystyle\sum_{i=1}^{p} [\boldsymbol{C}_m]_{li} \underline{X}_{c_i}(e^{j2\pi (f - \frac{k_m^c(l)}{LT})T}), & \text{if } f \in \dfrac{k_m^c(l)}{LT} \oplus \mathcal{G}_m. \end{cases}$$

for each m, which in view of (7), leads to

$$X(f) = \begin{cases} T\sqrt{L} \displaystyle\sum_{i=1}^{p} [\boldsymbol{A}_m^{-1}]_{li} e^{j2\pi \frac{c_i k_m(l)}{L}} \underline{X}_{c_i}(e^{j2\pi fT}), & \text{if } f \in \dfrac{k_m(l)}{LT} \oplus \mathcal{G}_m \\ T\sqrt{L} \displaystyle\sum_{i=1}^{p} [\boldsymbol{C}_m]_{li} e^{j2\pi \frac{c_i k_m^c(l)}{L}} \underline{X}_{c_i}(e^{j2\pi fT}), & \text{if } f \in \dfrac{k_m^c(l)}{LT} \oplus \mathcal{G}_m. \end{cases} \tag{A.4}$$

for $1 \leq m \leq M$. Equations (A.4) specify the spectrum $X(f)$ over the active subcells $(k_m(l)/LT) \oplus \mathcal{G}_m$ that partition $\mathcal{F}$. For the range $f \notin \mathcal{F}$, $X(f) = 0$ holds by our choice of $\boldsymbol{C}_m$'s. We multiply the right-hand sides of (A.4) by the indicator functions corresponding to their regions of validity and add them together. This gives us a single equation for $X(f)$

$$\begin{aligned}X(f) = & T\sqrt{L}\sum_{m=1}^{M}\sum_{l=1}^{q_m}\sum_{i=1}^{p} \\ & \cdot [\boldsymbol{A}_m^{-1}]_{li} e^{j2\pi\frac{c_i k_m(l)}{L}} \underline{X}_{c_i}(e^{j2\pi fT})\chi\left(f \in \frac{k_m(l)}{LT} \oplus \mathcal{G}_m\right) \\ & + T\sqrt{L}\sum_{m=1}^{M}\sum_{l=1}^{L-q_m}\sum_{i=1}^{p} \\ & \cdot [\boldsymbol{C}_m]_{li} e^{j2\pi\frac{c_i k_m^c(l)}{L}} \underline{X}_{c_i}(e^{j2\pi fT})\chi\left(f \in \frac{k_m^c(l)}{LT} \oplus \mathcal{G}_m\right) \\ \equiv & \sum_{i=1}^{p} \Phi_i(f)\underline{X}_{c_i}(e^{j2\pi fT})\end{aligned}$$

where

$$\begin{aligned}\Phi_i(f) = & T\sqrt{L}\sum_{m=1}^{M}\sum_{l=1}^{q_m} \\ & \cdot [\boldsymbol{A}_m^{-1}]_{li} e^{j2\pi\frac{c_i k_m(l)}{L}} \chi\left(f \in \frac{k_m(l)}{LT} \oplus \mathcal{G}_m\right) \\ & + T\sqrt{L}\sum_{m=1}^{M}\sum_{l=1}^{L-q_m} \\ & \cdot [\boldsymbol{C}_m]_{li} e^{j2\pi\frac{c_i k_m^c(l)}{L}} \chi\left(f \in \frac{k_m^c(l)}{LT} \oplus \mathcal{G}_m\right) \\ = & \begin{cases} T\sqrt{L}[\boldsymbol{A}_m^{-1}]_{li} e^{j2\pi\frac{c_i k_m(l)}{L}}, & \text{if } f \in \frac{k_m(l)}{LT} \oplus \mathcal{G}_m \\ T\sqrt{L}[\boldsymbol{C}_m]_{li} e^{j2\pi\frac{c_i k_m^c(l)}{L}}, & \text{if } f \in \frac{k_m^c(l)}{LT} \oplus \mathcal{G}_m. \end{cases}\end{aligned}$$

Each of the filters $\Phi_i(f)$, $1 \le i \le p$ has a piecewise-constant frequency response. Therefore, the reconstruction equation is

$$\begin{aligned}x(t) &= \sum_{i=1}^{p}\sum_{n=-\infty}^{\infty} \underline{x}_{c_i}(nT)\phi_i(t - nT) \\ &= \sum_{i=1}^{p}\sum_{j=-\infty}^{\infty} x((c_i + Lj)T)\phi_i(t - (c_i + Lj)T)\end{aligned}$$

where $\phi_i(t)$ is the inverse Fourier transform of $\Phi_i(f)$. □

Proof of Theorem 3: The following equations (for each m and $f \in \mathcal{G}_m$) are (21) rewritten in scalar form:

$$\begin{aligned}\tilde{X}_{k_m(r)}(f) &= X_{k_m(r)}(f) + \sum_{l=1}^{L-q_m} [\boldsymbol{D}_m]_{rl} X_{k_m^c(l)}(f), \quad 1 \le r \le q_m \\ \tilde{X}_{k_m^c(r)}(f) &= X_{k_m^c(r)}(f) + \sum_{l=1}^{L-q_m} [\boldsymbol{F}_m]_{rl} X_{k_m^c(l)}(f), \quad 1 \le r \le L - q_m.\end{aligned}$$

Consider the spectral components $E_{k_m(l)}(f)$ and $E_{k_m^c(l)}(f)$ of the aliasing error ($E(f) \stackrel{\text{def}}{=} \tilde{X}(f) - X(f)$) defined exactly as in (12) with $E(f)$ in place of $X(f)$

$$\begin{aligned}E_{k_m(r)}(f) &= \sum_{l=1}^{L-q_m} [\boldsymbol{D}_m]_{rl} X_{k_m^c(l)}(f) \\ E_{k_m^c(s)}(f) &= \sum_{l=1}^{L-q_m} [\boldsymbol{F}_m]_{sl} X_{k_m^c(l)}(f)\end{aligned} \tag{A.5}$$

for $f \in \mathcal{G}_m$, $1 \le r \le q_m$, $1 \le s \le L - q_m$, and each m. When expressed in the time domain, (A.5) becomes

$$\begin{aligned}e_{k_m(r)}(t) &= \sum_{l=1}^{L-q} [\boldsymbol{D}_m]_{rl}[x_{k_m^c(l)}(t)] \\ e_{k_m^c(s)}(t) &= \sum_{l=1}^{L-q} [\boldsymbol{F}_m]_{sl}[x_{k_m^c(l)}(t)].\end{aligned} \tag{A.6}$$

The total aliasing error can be obtained by modulating the errors on each subcell appropriately and adding. The result is that

$$e(t) = \sum_{m=1}^{M} e^{(m)}(t)$$

where

$$e^{(m)}(t) = \sum_{l=1}^{q_m} e_{k_m(l)}(t) e^{\frac{j2\pi k_m(l)t}{LT}} + \sum_{l=1}^{L-q_m} e_{k_m^c(l)}(t) e^{\frac{j2\pi k_m^c(l)t}{LT}}.$$

Employing (A.6) in the above equation gives

$$\begin{aligned}e(t) &= \sum_{m=1}^{M}\sum_{r=1}^{q_m} e^{\frac{j2\pi k_m(r)t}{LT}} \left(\sum_{l=1}^{L-q_m} [\boldsymbol{D}_m]_{rl} x_{k_m^c(l)}(t)\right) \\ &\quad + \sum_{m=1}^{M}\sum_{r=1}^{L-q_m} e^{\frac{j2\pi k_m^c(r)t}{LT}} \left(\sum_{l=1}^{L-q_m} [\boldsymbol{F}_m]_{rl} x_{k_m^c(l)}(t)\right) \\ &\equiv \sum_{m=1}^{M}\sum_{l=1}^{L-q_m} \mu_{m,l}(t) x_{k_m^c(l)}(t)\end{aligned} \tag{A.7}$$

where $\mu_{m,l}(t)$ for each $1 \le m \le M$ and $1 \le l \le L - q_m$ are continuous, LT-periodic functions defined as

$$\mu_{m,l}(t) = \sum_{r=1}^{L-q_m} [\boldsymbol{F}_m]_{rl} e^{\frac{j2\pi k_m^c(r)t}{LT}} + \sum_{r=1}^{q_m} [\boldsymbol{D}_m]_{rl} e^{\frac{j2\pi k_m(r)t}{LT}}.$$

The sup-norm of $e(t)$ can be computed directly from (A.7) as follows:

$$\begin{aligned}\sup_t |e(t)| &\le (\max_{m,l,t} |\mu_{m,l}(t)|) \sum_{m=1}^{M}\sum_{l=1}^{L-q_m} \sup_t |x_{k_m^c(l)}(t)| \\ &\le (\max_{m,l,t} |\mu_{m,l}(t)|) \sum_{m=1}^{M}\sum_{l=1}^{L-q_m} \int |X_{\overline{k}_m(l)}(f)|\, df \\ &= \max_{m,l,t} |\mu_{m,l}(t)| \sum_{m=1}^{M}\sum_{l=1}^{L-q_m} \\ &\quad \cdot \int_{[\mathcal{F}]} |X(f)| \chi\left(f \in \frac{k_m^c(l)}{LT} \oplus \mathcal{G}_m\right) df.\end{aligned}$$

The double sum of the integrals reduces to the integral of $|X(f)|$ over the union of the subcells $\{(k_m^c(l)/LT) \oplus \mathcal{G}_m\}$, or equivalently, over the out-of-band region $[\mathcal{F}]\backslash\mathcal{F}$. Hence

$$\sup_t |e(t)| \leq \psi \int_{[\mathcal{F}]\backslash\mathcal{F}} |X(f)|\, df \tag{A.8}$$

where ψ is the constant

$$\psi \stackrel{\text{def}}{=} \max_m \Big(\max_{1\leq l\leq L-q_m, t\in[0,LT]} |\mu_{m,l}(t)| \Big).$$

We have used the fact that $\mu_{m,l}(t)$ is a periodic function of period LT to restrict the range for t in the above maximization. We demonstrate that ψ is the smallest possible coefficient in the bound (A.8) for the sup-norm of $e(t)$. First note that $\psi = |\mu_{m_0,l_0}(t_0)|$ for some $m_0 \in \{1,\ldots,M\}$ and $l_0 \in \{1,2,\ldots,L-q_{m_0}\}$ and $t_0 \in [0,LT]$ since $\mu_{m,l}(t)$ is continuous on $[0,LT]$. Now define

$$X(f) = \begin{cases} \exp(-j2\pi f t_0), & \text{if } f \in \dfrac{k_{m_0}^c(l_0)}{LT} \oplus \mathcal{G}_{m_0} \\ 0, & \text{otherwise.} \end{cases} \tag{A.9}$$

In the time domain this is equivalent to

$$x(t) = \lambda(\mathcal{G}_{m_0})\,\mathrm{sinc}\,(\lambda(\mathcal{G}_{m_0})(t-t_0)) \times \exp\left(j2\pi\left(\frac{k_{m_0}^c(l_0)}{LT} + \frac{1}{2}(\gamma_{m_0}+\gamma_{m_0+1})\right)(t-t_0)\right) \tag{A.10}$$

For the choice of $X(f)$ in (A.9) it is clear that there is only one nonzero term in the right-hand side of (A.7), namely, the term corresponding to $m=m_0$ and $l=l_0$. Hence we obtain

$$\begin{aligned} \sup_t |e(t)| &= \sup_t |\mu_{m_0,l_0}(t) x_{k_{m_0}^c(l_0)}(t)| \\ &\geq |\mu_{m_0,l_0}(t_0)| \times |x_{k_{m_0}^c(l_0)}(t_0)| \\ &= \psi \times \lambda(\mathcal{G}_{m_0}) = \psi \int_{[\mathcal{F}]\backslash\mathcal{F}} |X(f)|\, df. \end{aligned}$$

In fact, both sides of the above inequality are equal since (A.8) holds. This proves that the bound in (A.8) is sharp with the extremal $x(t)$ in (A.10) achieving the bound. □

Proof of Theorem 4: We shall now derive a bound on the energy of the error $e(t)$. First observe that by Parseval's theorem

$$\begin{aligned} \int_{-\infty}^{\infty} e^2(t)\, dt &= \int_{[\mathcal{F}]} |E(f)|^2\, df \\ &= \int_{[\mathcal{F}]\backslash\mathcal{F}} |E(f)|^2\, df + \int_{\mathcal{F}} |E(f)|^2\, df \\ &= \sum_{m=1}^{M}\sum_{r=1}^{L-q_m} \int_{\frac{k_m^c(r)}{LT}\ominus\mathcal{G}_m} |E(f)|^2\, df \\ &\quad + \sum_{m=1}^{M}\sum_{r=1}^{q_m} \int_{\frac{k_m(r)}{LT}\ominus\mathcal{G}_m} |E(f)|^2\, df \\ &= \sum_{m=1}^{M}\sum_{r=1}^{L-q_m} \int_{\mathcal{G}_m} |E_{k_m^c(r)}(f)|^2\, df \\ &\quad + \sum_{m=1}^{M}\sum_{r=1}^{q_m} \int_{\mathcal{G}_m} |E_{k_m(r)}(f)|^2\, df. \end{aligned}$$

Now it readily follows from (A.5), (13), and the above equation that

$$\begin{aligned} \int_{-\infty}^{\infty} e^2(t)\, dt &= \sum_{m=1}^{M} \int_{\mathcal{G}_m} (\boldsymbol{x}_m^-(f))^*(\boldsymbol{F}_m^*\boldsymbol{F}_m + \boldsymbol{D}_m^*\boldsymbol{D}_m)\boldsymbol{x}_m^-(f)\, df \\ &= \sum_{m=1}^{M}\sum_{l=1}^{L-q_m} \int_{\mathcal{G}_m} |X_{k_m^c(l)}(f)|^2\, df \\ &= \int_{\mathcal{H}_m} |X(f)|^2\, df \end{aligned} \tag{A.11}$$

where, the last step follows from the definition of $\boldsymbol{x}_m^-(f)$, and $\mathcal{H}_m$ is defined as

$$\mathcal{H}_m = \bigcup_{l=1}^{L-q_m} \frac{k_m^c(l)}{LT} \oplus \mathcal{G}_m. \tag{A.12}$$

Note that the union of these sets $\cup\mathcal{H}_m$ equals the total out-of-band region $[\mathcal{F}]\backslash\mathcal{F}$. Therefore, we deduce from (A.11) and (A.12) that

$$\int_{-\infty}^{\infty} e^2(t)\, dt \leq \sum_{m=1}^{M} \lambda_{\max}(\boldsymbol{F}_m^*\boldsymbol{F}_m + \boldsymbol{D}_m^*\boldsymbol{D}_m) \int_{\mathcal{H}_m} |X(f)|^2\, df.$$

This bound is not very useful in this form. We can weaken it a little to express it in terms of the out-of-band signal energy $\mathcal{E}_{\text{out}}$

$$\int_{-\infty}^{\infty} e^2(t)\, dt \leq \max_m [\lambda_{\max}(\boldsymbol{F}_m^*\boldsymbol{F}_m + \boldsymbol{D}_m^*\boldsymbol{D}_m)]\mathcal{E}_{\text{out}} \tag{A.13}$$

where $\mathcal{E}_{\text{out}}$ is defined as

$$\mathcal{E}_{\text{out}} = \sum_{m=1}^{M} \int_{\mathcal{H}_m} |X(f)|^2\, df \equiv \int_{[\mathcal{F}]\backslash\mathcal{F}} |X(f)|^2\, df \tag{A.14}$$

By a very similar argument, we obtain the following lower bound on the energy of $e(t)$:

$$\int_{-\infty}^{\infty} e^2(t)\, dt \geq \min_m [\lambda_{\min}(\boldsymbol{F}_m^*\boldsymbol{F}_m + \boldsymbol{D}_m^*\boldsymbol{D}_m)]\mathcal{E}_{\text{out}}. \tag{A.15}$$

These bounds are indeed sharp. The constants multiplying $\mathcal{E}_{\text{out}}$ are the best. To demonstrate this we construct extremal functions satisfying each of the above bounds. It is sufficient to specify the active and inactive spectral components, $\boldsymbol{x}_m^+(f)$ and $\boldsymbol{x}_m^-(f)$ rather than $X(f)$. After all, one can be determined in terms of the other. Consider the bound (A.13) first. Let

$$m_0 = \arg\max_m (\lambda_{\max}(\boldsymbol{F}_m^*\boldsymbol{F}_m + \boldsymbol{D}_m^*\boldsymbol{D}_m))$$

and define for each m

$$\begin{aligned} \boldsymbol{x}_m^+(f) &= \boldsymbol{0}, & f \in \mathcal{G}_m \\ \boldsymbol{x}_m^-(f) &= \begin{cases} \boldsymbol{0}, & \text{if } m \neq m_0 \\ \boldsymbol{p}_{m_0}, & \text{if } m = m_0 \end{cases} & f \in \mathcal{G}_{m_0} \end{aligned}$$

where $\boldsymbol{p}_{m_0}$ is the eigenvector of $[\boldsymbol{F}_{m_0}^*\boldsymbol{F}_{m_0} + \boldsymbol{D}_{m_0}^*\boldsymbol{D}_{m_0}]$ corresponding to its largest eigenvalue. Starting from (A.11), one can readily verify that the above function is an extremal for the bound (A.13). An extremal for (A.15) is constructed analogously. □

Proof of Theorem 5: Recall that

$$\tilde{w}(t) = \sum_{i=1}^{p} \sum_{j=-\infty}^{\infty} w((c_i + Lj)T)\phi_i(t - (c_i + Lj)T)$$

It is clear that

$$E[w^*((Lj + c_i)T)w((Lj' + c_{i'})T)] = \sigma^2 \delta_{jj'}\delta_{ii'}.$$

Hence, distinct terms in the above summation are uncorrelated and we obtain

$$E|\tilde{w}(t)|^2 = \sigma^2 \sum_{i=1}^{p} \sum_{j=-\infty}^{\infty} |\phi_i(t - LjT - c_iT)|^2$$

for the output noise power at time t. The above expression, although not necessarily independent of time, is certainly periodic with period LT. Hence the *average* noise power can be computed as follows:

$$\begin{aligned}\langle E|\tilde{w}(t)|^2\rangle_t &= \frac{1}{LT}\int_0^{LT} E|\tilde{w}(t)|^2\,dt \\ &= \frac{\sigma^2}{LT}\int_0^{LT} \sum_{i=1}^{p}\sum_{j=-\infty}^{\infty} |\phi_i(t - LTj - c_iT)|^2\,dt \\ &\equiv \frac{\sigma^2}{LT}\sum_{i=1}^{p}\int_{-\infty}^{\infty} |\phi_i(t - c_iT)|^2\,dt = \frac{\sigma^2}{LT}\sum_{i=1}^{p}\mathcal{E}_{\phi_i}\end{aligned} \tag{A.16}$$

where $\mathcal{E}_{\phi_i}$ is the energy contained in $\phi_i(t)$. Using Parseval's theorem and (18) we compute $\mathcal{E}_{\phi_i}$

$$\begin{aligned}\mathcal{E}_{\phi_i} &= \int |\Phi_i(f)|^2\,df \\ &= T^2 L \sum_{m=1}^{M} \lambda(\mathcal{G}_m)\left(\sum_{l=1}^{q_m} |[\boldsymbol{A}_m^{-1}]_{li}|^2 + \sum_{l=1}^{L-q_m} |[\boldsymbol{C}_m]_{li}|^2\right).\end{aligned} \tag{A.17}$$

This computation was quite simple owing to the fact that $\Phi_i(f)$ is piecewise-constant. Combining (A.16) and (A.17) gives

$$\begin{aligned}\langle E|\tilde{w}(t)|^2\rangle_t &= \sigma^2 T \sum_{i=1}^{p}\sum_{m=1}^{M} \lambda(\mathcal{G}_m) \\ &\quad \cdot \left(\sum_{l=1}^{q_m} |[\boldsymbol{A}_m^{-1}]_{li}|^2 + \sum_{l=1}^{L-q_m} |[\boldsymbol{C}_m]_{li}|^2\right) \\ &= \sigma^2 \psi_{\mathbf{n}},\end{aligned}$$

where $\psi_{\mathbf{n}} = T\sum_{m=1}^{M} \lambda(\mathcal{G}_m)(\|\boldsymbol{A}_m^{-1}\|_F^2 + \|\boldsymbol{C}_m\|_F^2)$

where $\|\cdot\|_F$ is the Frobenius norm. □

REFERENCES

[1] C. E. Shannon, "Communication in the presence of noise," *Proc. IRE* , vol. 37, pp. 10–21, Jan. 1949.

[2] J. L. Brown, Jr., "Sampling expansions for multiband signals," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 33, pp. 312–315, Feb. 1985.

[3] J. R. Higgins, *Sampling Theory in Fourier and Signals Analysis Foundations*. Oxford, U.K.: Oxford Science, 1996.

[4] R. J. Marks II, *Advanced Topics in Shannon Sampling and Interpolation Theory*, R. J. Marks II, Ed. New York: Springer-Verlag, 1993.

[5] R. J. Papoulis, "Generalized sampling expansions," *IEEE Trans. Circuits Syst.*, vol. CAS-24, pp. 652–654, Nov. 1977.

[6] A. J. Jerri, "The Shannon sampling theorem—Its various extensions and applications: A tutorial review," *Proc. IEEE*, vol. 65, pp. 1565–1596, Nov. 1977.

[7] H. Landau, "Necessary density conditions for sampling and interpolation of certain entire functions," *Acta Math.*, 1967.

[8] R. E. Kahn and B. Liu, "Sampling representations and the optimum reconstruction of signals," *IEEE Trans. Inform. Theory*, vol. IT-11, pp. 339–347, July 1965.

[9] K. Cheung and R. Marks, "Image sampling below the Nyquist density without aliasing," *J. Opt. Soc. Amer. A*, 1990.

[10] K. Cheung, *Advanced Topics in Shannon Sampling and Interpolation Theory*. New York: Springer-Verlay, 1993, ch. 3.

[11] P. Feng and Y. Bresler, "Spectrum-blind minimum-rate sampling and reconstruction of multiband signals," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*. Atlanta, GA, May 1996.

[12] Y. Bresler and P. Feng, "Spectrum-blind minimum-rate sampling and reconstruction of 2-D multiband signals," in *Proc. 3rd IEEE Int. Conf. on Image Processing, ICIP'96*, vol. 1, Lausanne, Switzerland, Sept. 1996, pp. 701–704.

[13] P. Vaidyanathan and V. Liu, "Efficient recosntruction of band-limited sequences from nonuniformly decimated versions by use of polyphase filter bands," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, pp. 1927–1936, Nov. 1990.

[14] B. Foster and C. Herley, "Exact reconstruction from periodic nonuniform sampling of signals with arbitrary frequency support," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*. Detroit, MI, May 1998.

[15] S. C. Scoluar and W. J. Fitzgerald, "Periodic nonuniform sampling of multiband signals," *Sig. Processing*, vol. 28, no. 2, pp. 195–200, Aug. 1992.

[16] C. Herley and P. W. Wong, "Minimum rate sampling of signals with arbitrary frequency support," in *IEEE Int. Conf. Image Processing*. Lausanne, Switzerland, Sept. 1996.

[17] R. G. Shenoy, "Nonuniform sampling of signals and applications," in *Int. Symp. Circuits and Systems*, vol. 2. London, U.K., May 1994, pp. 181–184.

[18] J. R. Higgins, "Some gap sampling series for multiband signal," *Signal Processing*, vol. 12, no. 3, pp. 313–319, Apr. 1986.

[19] M. G. Beaty and J. R. Higgins, "Aliasing and poisson summation in the sampling theory of Paley-Wiener spaces," *J. Fourier Anal. Applic.*, vol. 1, pp. 67–85, 1994.