

FAST RECONSTRUCTION METHODS FOR BANDLIMITED FUNCTIONS FROM PERIODIC NONUNIFORM SAMPLING*

THOMAS STROHMER[†] AND JARED TANNER[‡]

Abstract. A well-known generalization of Shannon’s sampling theorem states that a bandlimited function can be reconstructed from its periodic nonuniformly spaced samples if the effective sampling rate is at least the Nyquist rate. Analogous to Shannon’s sampling theorem this generalization requires that an infinite number of samples be available, which, however, is never the case in practice. Most existing reconstruction methods for periodic nonuniform sampling yield very low order (often not even first order) accuracy when only a finite number of samples is given. In this paper we propose a fast, numerically robust, root-exponential accurate reconstruction method. The efficiency and accuracy of the algorithm is obtained by fully exploiting the sampling structure and utilizing localized Fourier analysis. We discuss applications in analog-to-digital conversion where nonuniform periodic sampling arises in various situations. Finally, we demonstrate the performance of our algorithm by numerical examples.

Key words. Shannon’s sampling theorem, oversampling, nonuniform periodic sampling, analog-to-digital conversion, Gevrey regularity, uniform interleaved sampling

AMS subject classifications. 41A05, 41A30, 42C15, 65T50, 94A12, 94A20

DOI. 10.1137/040609586

1. Introduction. The classical Shannon sampling theorem plays a crucial role in signal processing and communications, indicating how to transfer between analog signals and discrete sequences [26]. Shannon’s sampling theorem states that if a function¹ belongs to the space of bandlimited functions B_σ , i.e.,

$$(1.1) \quad f(t) := \frac{1}{\sqrt{2\pi}} \int_{-\sigma}^{\sigma} e^{2\pi i w t} F(w) dw, \quad F(w) \in L^2_0[-\sigma, \sigma],$$

then it can be recovered exactly from its equidistant samples

$$(1.2) \quad f(t) \equiv \sum_{k=-\infty}^{\infty} f\left(\frac{k}{2\sigma}\right) \frac{\sin(2\pi\sigma t - \pi k)}{2\pi\sigma t - \pi k} := \sum_{k=-\infty}^{\infty} f\left(\frac{k}{2\sigma}\right) \operatorname{sinc}(2\pi\sigma t - \pi k).$$

Shannon’s sampling theorem assumes that an infinite number of samples is available, which is of course never true in practice. Truncation of the cardinal series (1.2) results in rather poor approximation of the original bandlimited signal, and the truncation error is of the unacceptable low order of $1/\sqrt{L}$, where L is the number of samples; cf. [26]. In the presence of noise or quantization errors convergence may even break down completely [4]. To avoid these problems in practice, one usually

*Received by the editors June 7, 2004; accepted for publication (in revised form) January 26, 2006; published electronically June 2, 2006. This work was partially supported by NSF DMS grant 0208568.

<http://www.siam.org/journals/sinum/44-3/60958.html>

[†]Department of Mathematics, University of California, Davis, CA 95616-8633 (strohmer@math.ucdavis.edu).

[‡]Department of Mathematics, University of California, Davis, CA 95616-8633. Current address: Department of Mathematics, University of Utah, Salt Lake City, UT (tanner@math.utah.edu). The work of this author was supported by NSF DMS VIGRE grant 0135345.

¹Throughout this paper we will use lowercase letters to designate functions in the time domain, and uppercase for the Fourier transform of the same function.

resorts to oversampling of the signal, since this gives rise to vastly better convergence rates and as a result greater robustness to noise.

While oversampling is therefore desirable in practice, it is not always easily done in real world applications. For example, consider advanced wireless communication systems, where demand for data rate is steadily increasing, requiring communication systems that use transmission signals with a (baseband) bandwidth in the range of tens of megahertz up to one gigahertz (as in currently developed ultrawideband systems). Such a huge bandwidth necessitates very high sampling rates in the analog-to-digital conversion which puts enormous demands on the analog sampling devices. While it is possible to construct signal acquisition systems that sample a signal even at nanosecond scales with high precision, such devices become increasingly expensive. More specifically, a linear increase in precision of a sampling device often goes hand in hand with a superlinear increase in the costs of constructing such a device.

One possible way to remedy this problem is to combine several analog-to-digital converters (ADCs) with lower sampling rate to obtain one virtual sensor with high sampling rate. We describe this concept in more detail. A standard ADC uniformly (over)samples an analog signal (a continuous-time function) at rate T^{-1} , say, where T is the time between two successive samples. The so obtained discrete-time signal is then subject to quantization, and the quantized signal is further processed by a digital signal processor (DSP). Instead of using one ADC with sampling rate T^{-1} we could run N ADCs in parallel, each operating at the slower rate $(NT)^{-1}$. The sampling instances of the n th ADC are chosen at $\{(kN + n)T\}_{k \in \mathbb{Z}}$, $n = 0, \dots, N - 1$, so that the combined sampling instances are $\{kT\}_{k \in \mathbb{Z}}$, which is equivalent to the output of one ADC that operates at the N times higher rate T^{-1} .

Many companies such as Maxim (http://www.maxim-ic.com/appnotes.cfm/appnote_number/2094), Agilent Technologies (http://www.agilent.com/labs/news/2003features/fea_adc03.html), and Analog Devices have been developing or are currently developing such time-interleaved ADCs.

To give another concrete current example for the need of time-interleaved ADCs consider the 10 Gigabit Ethernet over copper standard (which is part of the IEEE 802.3 standard; see <http://www.ieee802.org>). There 4 Cat6 copper pairs are used, so 2.5 Gigabits/sec are transmitted. Since a 12-PAM code with error correction is used, this means the baud rate is about 800 MHz. At least 8 bits precision is necessary. The fastest ADC with the desired precision runs about 1/2 of that. Consequently 2 or 4 time-interleaved ADC channels must be used to achieve the required precision.

While a time-interleaved ADC structure obviously has its merits, it does not come without caveats. The coordination of the N ADCs has to be done with high precision, but in practice timing errors between the individual ADCs result in sampling sets of the form $\{kNT + T_n\}_{k \in \mathbb{Z}}$, where the T_n are distinct random timing shifts. In other words, the combined sampling set does not form a uniform sampling set but consists of nonuniformly shifted unions of uniform sampling points, which is often referred to as *periodic nonuniform sampling* or *bunched sampling* [2, 18]. This poses two problems: (i) How can we estimate the unknown shifts T_n ? (ii) How can we reconstruct quickly and stably the original signal from its periodically nonuniformly spaced samples? In this paper we focus on the second question, with shift estimation discussed in [23]. While there are several algorithms in the literature that deal with the reconstruction of bandlimited signals from periodic nonuniform samples (see, e.g., [14, 18, 2, 25, 24, 5, 12, 15, 20]), none of these algorithms provide high order accuracy with respect to truncation error.

Sometimes it can be advantageous to deliberately perform nonuniform periodic sampling in connection with analog-to-digital conversion. For instance, in [20] the use of periodic nonuniform sampling is proposed to avoid noise coupling in a mixed-signal integrated circuit, which contains analog and digital signal processing circuits, as is the case for an ADC. After the analog input signal has been sent through the ADC, the digital output is processed further by a DSP. However, switching of the digital circuits generates noise that can couple into the analog signal path through so-called parasite signal paths. Such noise coupling distorts the analog signal, which degrades the signal-to-noise ratio at the input of the ADC.

To avoid this noise coupling, it is proposed in [20] to have the ADC acquire a group (bunch) of samples at high rate while the digital signal processor is inactive, and allow digital processing of the ADC output during a second phase when the ADC is not sampling. This reduces the noise coupled to the analog signal, since the DSP operates only during the second phase. As the final step, one has to convert the bunched samples to uniformly spaced samples. Practical restrictions in terms of available memory and tolerable time delay imply an upper limit on the number of samples that can be processed during the conversion from bunched to uniform samples. This sampling pattern is obviously a special case of the periodic nonuniform sampling pattern described in the previous paragraph, with the simplification that all $T_n - T_{n-1}$ are (nearly) equal, but with the difficulty that we have a potentially large gap between two clusters of samples. This large gap may cause some instabilities; therefore it is vital to have a numerical reconstruction algorithm that is robust to such large gaps.

In this paper we develop the first method for reconstruction of a bandlimited signal from its periodic nonuniformly spaced samples that achieves root-exponential accuracy from a finite number of samples. The proposed method is numerically robust, and since its computationally most expensive steps consist of fast Fourier transforms (FFTs), it is numerically very efficient.

The paper is organized as follows. In section 2 we review some results on oversampling and localization of functions and their Fourier transforms. The Gevrey class arises as a natural candidate space for compactly supported smooth filter functions in connection with oversampled bandlimited signals. In section 3 we briefly describe how these smooth filters correspond to a localized reconstruction, resulting in a root-exponential accurate, fast algorithm for recovering a bandlimited signal from its uniformly spaced oversampled values. This simple observation is integral for the derivation of the main algorithm for the case of periodic nonuniformly spaced samples; cf. section 4. Numerical simulations that demonstrate the performance of the proposed method are presented in section 5. Finally, section 6 contains our conclusion and an outlook of future research.

2. Oversampling and localization. As mentioned in the introduction, the formulation in (1.2) is unsuitable for practical applications, where only a finite number of samples is available, $\{f(k/2\sigma)\}_{|k|\leq L}$. For truncated samples the error, classically referred to as the truncation error, is controlled by the atom's localization

$$\begin{aligned}
 (2.1) \quad \epsilon(t, L, T) &:= \left| f(t) - \frac{\sqrt{2\pi}}{2\sigma} \sum_{|k|\leq L} f\left(\frac{k}{2\sigma}\right) \psi\left(t - \frac{k}{2\sigma}\right) \right| \\
 &\leq \frac{\sqrt{2\pi}}{2\sigma} \cdot \|f\|_{L^\infty} \sum_{|k|>L} \left| \psi\left(t - \frac{k}{2\sigma}\right) \right|.
 \end{aligned}$$

In the case of the classical Shannon sampling theorem, the atom, $\psi(\tau) := \text{sinc}(\tau)$, suffers from an unacceptably slow decay, $\lim_{\tau \rightarrow \infty} \psi(\tau) \sim 1/\tau$, resulting in a first order convergence rate while moving from the sample boundaries, $\pm L/2\sigma$, to the interior. Moreover, if the samples $f(k/(2\sigma))$ are replaced by noisy samples $f(k/(2\sigma)) + \epsilon_k$, then the corresponding approximation via the cardinal series in (1.2) may differ significantly from $f(t)$; cf. [4].

To remedy these problems in applications, one usually introduces oversampling. Sampling a function in the time domain introduces a periodization in the associated Fourier dual space, where sampling rate $T^{-1} = 2\sigma$ corresponds to a 2σ periodization. In (1.2) the reproducing atom (time domain), $\text{sinc}(2\pi\sigma t - \pi k)$, removes the periodization introduced by sampling, through the action of its associated filter (Fourier dual space), $\chi_{[-\sigma, \sigma]}$. For this critical, Nyquist sampling rate, $\text{sinc}(\cdot)$ is the unique atom that can be used to remove the periodization. However, if the bandlimited signal is sampled at a faster rate, $T^{-1} := 2\sigma/r$, where $r < 1$, then the dual space periodization is increased to $2\sigma/r$, allowing a large family of reproducing filters. Specifically, any function satisfying²

$$(2.2) \quad \Psi(w) = \begin{cases} 1, & |w| \leq \sigma, \\ 0, & |w| > \sigma(2-r)/r =: \Omega \\ \text{anything} & \text{else} \end{cases} \quad \Rightarrow \quad \psi \in B_\Omega,$$

gives rise to a Shannon-type series expansion

$$(2.3) \quad f(t) \equiv \sqrt{2\pi}T \sum_{k=-\infty}^{\infty} f(kT) \psi(t - kT).$$

For $r = 1$ the above filter reduces to $\chi_{[-\sigma, \sigma]}$ and the classical Shannon's sampling theorem, whereas for $r < 1$ a gap³ is introduced between σ and $\sigma(2-r)/r$, allowing for a host of other reproducing filters, including those with a high degree of regularity; see Figure 2.1.

Asymptotically the atom's localization is reflected in the filter's smoothness; consequently, the filter's regularity controls the convergence rate of the truncation error (2.1). By constructing infinitely differentiable filters with precise regularity estimates, we obtain root-exponential accuracy for the approximation of a bandlimited signal, as the point to be approximated moves from the sampling boundary, $\pm LT$, to the interior; see subsection 2.1. Unfortunately, unlike classical finite regularity filters, such as the raised cosine, which have a closed form expression for their corresponding atoms, to the authors' knowledge, there is no known infinitely differentiable compactly supported filter whose atom allows an explicit closed form expression. Alternatively to approximating the atoms as proposed in [16, 22], we introduce and analyze a direct Fourier domain implementation that does not adversely affect the high resolution

²It has been noted in [11] that the reproducing property is somewhat less strict than as stated in (2.2), in that the filter need not be zero for all $|w| \geq (2-r)/r$. Rather, the reproducing property is satisfied if the filter is one for $|w| \leq \sigma$ and zero at the points where the periodic extension of the signal's dual representation is nonzero. However, this added flexibility cannot increase the regularity of the filter or decrease its regularity constants and, as such, cannot improve the asymptotic convergence rate. Although this added flexibility can be used to increase the atom's immediate localization about the origin, it introduces substantial peaks away from the origin [11], decreasing the overall convergence rate.

³Filters which are nonzero for $\sigma < |w| < \Omega$ necessarily decrease noise less than the characteristic filter, $\chi_{[-\sigma, \sigma]}$, but the vastly improved convergence rate more than makes up for this modestly lower denoising which is overcome in other ways.

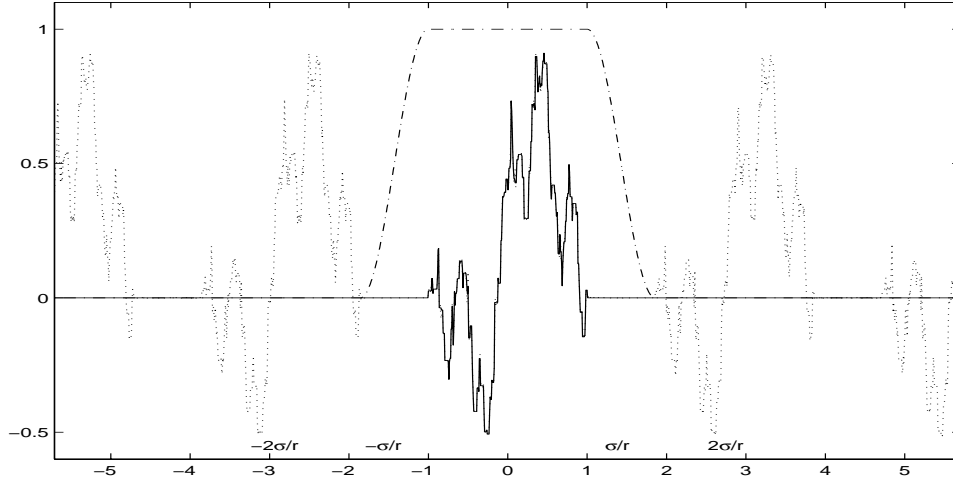


FIG. 2.1. The signal's dual space representation (solid line) and the signal's periodization due to sampling at the rate, $T^{-1} := 2\sigma/r$, for $r < 1$ (dotted line). With the gap between periodization, a smooth filter (dot-dash line) can be used to remove the periodization introduced by sampling.

achieved by smooth filters. The time domain localization of an atom, $\psi(\cdot)$, is reflected in the regularity⁴ of its corresponding filter, $\Psi(\cdot)$,

$$(2.4) \quad |\psi(t)| \leq (2\pi t)^{-s} \|\Psi\|_{C^s} \cdot \frac{2\Omega}{\sqrt{2\pi}} \quad \forall s, \quad \psi \in B_\Omega,$$

where $\|\Psi\|_{C^s} := \|\Psi^{(s)}\|_{L^\infty}$.

Consequently, combining the bounds in (2.1) and (2.4), convergence is gained at the polynomial rate⁵

$$(2.5) \quad \epsilon(t, L, T) \leq \text{Const} \cdot (LT - |t|)^{1-s} (2\pi)^{-s} \|\Psi\|_{C^s} \cdot \left(\frac{\Omega}{\sigma}\right), \quad s \geq 2,$$

as t passes from the boundary, $\pm LT$, to the interior, where $T := r/2\sigma$.

Rather than improving the atom's localization by increasing its corresponding filter's regularity, attempts have been made to construct highly localized atoms by maximizing the atom's local weight, $\int_{-R}^R \psi^2(t) dt / \int_{-\infty}^\infty \psi^2(t) dt$. However, such approaches have resulted in discontinuous filters [17] and atoms which do not decay globally [11]. A much more successful approach for polynomial order filters is to minimize the filter's regularity constant, $\|\Psi\|_{C^s}$. The classical raised cosine is such a filter [21]:

$$(2.6) \quad \Psi_{rc}(w) = \begin{cases} 1, & |w| \leq \sigma, \\ 0, & |w| > \sigma(2-r)/r, \\ \frac{1}{2} \left(1 + \cos\left(\frac{\pi}{2} \left(\frac{r}{1-r}\right) \left(\frac{w}{\sigma} - 1\right)\right)\right), & \sigma < w < \sigma \frac{2-r}{r}, \\ \frac{1}{2} \left(1 + \cos\left(\frac{\pi}{2} \left(\frac{r}{1-r}\right) \left(\frac{w}{\sigma} + 1\right)\right)\right), & -\sigma > w > -\sigma \frac{2-r}{r}, \end{cases}$$

⁴This is achieved by s integrations by parts, where the derivatives are transferred onto the filter.

⁵If the underlying filter possesses $\|\Psi\|_{C^{s+1}} < \infty$, then the bound (2.5) can be tightened by one order of $(LT - |t|)$ to the rate $(LT - |t|)^{-s}$ [9].

where the bounded regularity constants are given explicitly by $\|\Psi\|_{C^1} = \frac{1}{2}(\frac{T}{2(1-r)})$ and $\|\Psi\|_{C^2} = \frac{1}{2}(\frac{T}{2(1-r)})^2$.

When a comparatively small number of sampling points is taken, low regularity constant polynomial order methods give extremely good reconstructions. However, when a larger number of samples is available, atoms with significantly improved asymptotic localization can be achieved by constructing infinitely regular filters, $\Psi \in C_0^\infty$. It should be noted that for infinitely differentiable functions, the optimal bound in (2.4) is not necessarily obtained for large s , as the regularity constant $\|\Psi\|_{C^s}$ grows rapidly in s . Rather, for functions where precise regularity estimates are known, the optimal s can be determined, resulting in an exponential decay without necessarily large constants. These claims will be realized in the numerical experiments presented in section 5, contrary to the assertion in [3], where it is claimed that the increased regularity does not improve numerical convergence. In a direct numerical comparison with the raised cosine filter, our infinitely differentiable filter (2.10) achieves dramatically superior convergence in the interior of the samples, and quantitatively similar errors near the sampling boundaries; see Figure 5.1.

2.1. Localization and Gevrey regularity. To achieve exponential accuracy and satisfy the reproducing condition, (2.2), requires a filter which is infinitely differentiable and compactly supported. The natural space for infinitely differentiable compactly supported functions is the Gevrey class which consists of functions satisfying the smoothness bound

$$(2.7) \quad \|h\|_{C^s} := \|h^{(s)}\|_{L^\infty} \leq \text{Const} \cdot \frac{(s!)^\alpha}{\eta_h^s} \iff h \in G^\alpha,$$

where η_h is a constant independent of s . Incorporating the regularity information in the localization bound, (2.4), and minimizing over all admissible s , we conclude that Gevrey class filters satisfy a root-exponential localization decay,⁶

$$(2.8) \quad |\psi(t)| \leq \text{Const} \cdot \sqrt{|t|} \exp(-\alpha(2\pi\eta|t|)^{1/\alpha}), \quad \Psi \in G^\alpha,$$

and root-exponential truncation error

$$(2.9) \quad \epsilon(t, L, T) \leq \text{Const}_{\alpha,\eta} \exp(-(2\pi\eta(LT - |t|))^{1/\alpha}),$$

where $\text{Const}_{\alpha,\eta} \sim \eta^{-2} \sum_{l=0}^q q! \eta^{l/2} / (q-l)!$, with q the smallest integer greater than or equal to $(3\alpha - 2)/2$.

A similarly localized atom was constructed in [8, 16] by multiplying the sinc function with the inverse Fourier transform of an appropriately dilated G^2 function. Alternatively, such G^α filters can be expressed explicitly in the dual space, such as

$$(2.10) \quad \Psi_{G^2}(w) = \begin{cases} 1, & |w| \leq \sigma, \\ 0, & |w| > \sigma(2-r)/r, \\ \rho(\frac{w-\sigma}{\sigma(2-r)/r}), & \sigma < w < \sigma\frac{2-r}{r}, \\ \rho(\frac{-w-\sigma}{\sigma(2-r)/r}), & -\sigma > w > -\sigma\frac{2-r}{r}, \end{cases}$$

where $\rho(w) := \exp[\beta(w-1)^{-1} e^{-1/w}] \in G^2$ [13].

⁶Compact support is inconsistent with analyticity, G^1 , so reproducing atoms can at most be in the space G^α , for $\alpha > 1$, excluding true exponential decay, i.e., $\alpha = 1$, as was shown in the classical paper [1]. The Gevrey class of functions is essentially similar to ultradifferentiable functions [19].

Although the filter $\Psi_{G2}(\cdot)$ and the one in [16] result in rapid convergence while approaching the interior, $|t| \leq LT$, their associated atoms lack an explicit construction. As a result, to reconstruct a bandlimited signal at an arbitrary point has required the costly implementation of a quadrature evaluation, or a global approximation of the atom, such as the Padé and Gabor approximations proposed in [16] and [22], respectively. Alternatively, in the next section we introduce and analyze a direct Fourier domain implementation that does not adversely effect the high resolution achieved by smooth filters.

3. Dual space implementation for uniform oversampling. In this section we introduce an implementation which removes the sampling induced periodization through the direct action of the filter in the Fourier dual space. More specifically, if the bandlimited signal is sampled on the mesh $R := \{kT\}_{|k| \leq L}$, with $T := r/2\sigma$ and $r < 1$, we seek to compute an approximation to the signal on the refined mesh $P := \{kT/p\}_{|k| \leq pL}$, where $p \in \mathbb{N}/\{1\}$. This implementation is extremely efficient, as it only requires the FFT of the zero inserted signal, (3.2), and the pointwise multiplication in the dual space.

Define the approximation on the fine mesh as

$$(3.1) \quad \text{Approx}_\psi f\left(\frac{qT}{p}\right) := \sum_{|k| \leq L} f(kT)\psi\left[\left(\frac{q}{p} - k\right)T\right].$$

We zero insert the samples from the coarse mesh to the fine mesh

$$(3.2) \quad f_o(x) := \begin{cases} f(x), & x \in R, \\ 0, & x \in P/R, \end{cases}$$

and note that the approximation in (3.1) is a discrete convolution,

$$(3.3) \quad \text{Approx}_\psi f(hq) = \sum_{|k| \leq pL} f_o(hk)\psi([q - k]h),$$

where $h := T/p$. To transfer the discrete convolution to pointwise multiplication in the dual space, we define the discrete, pseudo-Fourier transforms of a function as

$$(3.4) \quad \tilde{G}(w_j) := \frac{h}{\sqrt{2\pi}} \sum_{|k| \leq pL} g(hk) \exp(-2\pi i h k w_j),$$

$$(3.5) \quad g(hq) := \frac{\sqrt{2\pi}}{h(2pL + 1)} \sum_{|j| \leq pL} \tilde{G}(w_j) \exp(2\pi i h q w_j),$$

where $w_j := j/h(2pL + 1)$.

In this notation the time domain implementation can be expressed in the Fourier domain,

$$(3.6) \quad \begin{aligned} \text{Approx}_\psi f(hq) &= \sum_{|k| \leq pL} f_o(kh)\psi([q - k]h) \\ &= \frac{2\pi}{h^2(2pL + 1)} \sum_{|j| \leq pL} \tilde{F}_o(w_j)\tilde{\Psi}(w_j) \exp(2\pi i h w_j q). \end{aligned}$$

Although this illustrates an efficient method for computing the approximation on a refined grid without increasing the overall error, it remains necessary to compute the atom in order to determine $\tilde{\Psi}(w_j)$. However, the pseudo-Fourier transform is the spectral projection (truncated Fourier series) of the true Fourier representation,

$$\begin{aligned}
 \tilde{\Psi}\left(\frac{w_j}{2\pi h}\right) &= \frac{h}{\sqrt{2\pi}} \sum_{|k|\leq pL} \psi(hk) \exp(-ikw_j) \\
 &= h \int_{-\sigma(2-r)/r}^{\sigma(2-r)/r} \Psi\left(\frac{w}{2\pi h}\right) D_{pL}(2\pi h(w-w_j)) dw \\
 (3.7) \qquad &= S_{pL}\Psi\left(\frac{w_j}{2\pi h}\right),
 \end{aligned}$$

where $D_R(x) := \sin((R + 1/2)x)/2\pi \sin(x/2)$ is the Dirichlet kernel of order R , and $S_R f(\cdot)$ is the R term truncated Fourier series projection of $f(\cdot)$, i.e., $S_R f := D_R * f$. As such, for highly smooth filters, the two Fourier representations are root-exponentially close⁷ for all w_j ,

$$\begin{aligned}
 (3.8) \qquad \left| \Psi(w_j) - \tilde{\Psi}(w_j) \right| &= \left| \Psi(w_j) - S_{pL}\Psi(w_j) \right| \\
 &\leq \text{Const} \cdot e^{-\alpha(\eta_\Psi LT)^{1/\alpha}}, \quad \Psi(\cdot) \in G^\alpha.
 \end{aligned}$$

The composite error is then composed of the traditional truncation error, (2.1), and the error in replacing the pseudo-Fourier transform with the exact dual space representation of the atom, i.e., the filter. We summarize the above results in the following theorem.

THEOREM 3.1. *Let $f(t)$ be a signal bandlimited to $[-\sigma, \sigma]$, and $\Psi(\cdot)$ a filter in G^α satisfying (2.2). From the function's oversampled values on the mesh, $R := \{kT\}_{|k|\leq L}$, where $T := r/2\sigma$; its approximation on the fine mesh, $P := \{kh\}_{|k|\leq pL/R}$, where $h := r/2\sigma p$ for $p \in \mathbb{N}/\{1\}$, can be computed by pointwise multiplying the pseudo-Fourier transform, (3.4), of the signal's zero insertion onto P and the filter, $\Psi(\cdot)$. The resulting error is bounded by*

$$\begin{aligned}
 &\left| f(hk) - \frac{2\pi}{h^2(2pL+1)} \sum_{|j|\leq pL} \tilde{F}_o(w_j)\Psi(w_j) \exp(2\pi i h w_j k) \right| \\
 &\leq \text{Const} \cdot \frac{p^2 L}{\sigma r} \|f\|_{L^\infty} \left(e^{-\alpha(\eta_\Psi LT)^{1/\alpha}} + e^{-\alpha(\eta_\Psi(LT-|t|))^{1/\alpha}} \right) \\
 &\leq \text{Const} \cdot \frac{p^2 L}{\sigma r} \|f\|_{L^\infty} e^{-\alpha(\eta_\Psi(LT-|t|))^{1/\alpha}}.
 \end{aligned}$$

Remark. Although the approximation of an arbitrary point is not possible by the method put forth in Theorem 3.1, the filter can be modulated to give an approximation on a shifted submesh, $P + s_0 := \{kh + s_0\}_{|k|\leq pL}$. This can be seen directly by replacing $\tilde{\Psi}(w_j)$ in (3.6) with its modulation $\tilde{\Psi}(w_j) \exp(2\pi i s_0 w_j)$ and absorbing the modulation into the exponential, resulting in $Approx_\psi f(hq + s_0)$. An arbitrary point in $|t| \leq LT$ can be approximated for a small modulation, $|s_0| \leq h/2$, and

⁷This is true for any $p \in \mathbb{N}/\{1\}$ as the trapezoidal quadrature is taken over the support of the filter, $[-\sigma(2-r)/r, \sigma(2-r)/r] \subset [w_{-pL}, w_{pL}]$.

as a result, the difference between the modulated pseudo-filter and the true modulated filter remain exponentially small, only modestly increasing the constant in (3.8). Consequently any point in the interior of the samples, $|t| \leq LT$, can be approximated with the exponential accuracy stated in Theorem 3.1 through the application of the appropriately modulated filter and a zero insertion of $p = 2$.

4. An algorithm for periodic nonuniform sampling. Although uniform oversampling achieves the optimal convergence rate for a given sampling rate, applications exist where for various reasons one is confronted with more general sampling geometries. In this section we devote our attention to the case of periodic nonuniform sampling (bunched sampling), i.e., when multiple uniform undersampled sets are combined to achieve an effective sampling density similar to the uniform sampling case. First order methods exist for the reconstruction of the bandlimited signal from periodic nonuniform sampling [18, 2, 25, 5]. Here we investigate the dual space structure induced by periodic nonuniform sampling, and derive direct high resolution reconstruction methods.

It is well known that for periodic nonuniform sampling a generalization of the cardinal series (1.2) holds by using N atoms in the series expansion (in combination with appropriate coefficients) instead of just one atom. Various essentially equivalent versions of this generalized sampling theorem have been derived, which all revolve around exploiting in the Fourier domain certain periodicities induced by the sampling geometry. Initially we follow a similar path, but unlike other approaches we pay special attention in our derivations to our goal of using highly localized atoms for the reconstruction of the sampled function.

Poisson’s summation theorem relates the physical space sampling to the resulting dual space periodization. Specifically, a bandlimited signal, $f \in B_\sigma$, sampled at the rate $T^{-1} := 2\sigma/r$ causes a dual space periodization of $2\sigma/r$. For oversampling, $r < 1$, the signal’s dual space representation is separated (Figure 2.1), and the introduced periodization can be removed in one step by applying a smooth filter which satisfies the reproducing condition (2.2). Alternatively, when a signal is undersampled ($r > 1$), the dual space periodizations overlap, and a general signal cannot be reconstructed from those samples alone. More precisely, if a signal is sampled at the points $\{lT + T_n\}_{l \in \mathbb{Z}}$, the dual space representation is given by

$$(4.1) \quad S_{T_n}(w) := e^{2\pi iT_n w} \sum_{l=-\infty}^{\infty} e^{-2\pi i l T_n T^{-1}} F(w - lT^{-1}).$$

If $N \geq \lceil r \rceil$ such undersampled sets are available, then an *effective sampling rate* of $2\sigma N/r$ is obtained and the overlapping can be removed for $|w| \leq \sigma$, allowing the recovery of the bandlimited signal. We now present general conditions, reminiscent of (2.2), for the recovery of a bandlimited signal from its bunched sampling. We then conclude this section with an algorithm for the construction of a family of filters which remove the sampling induced periodization.

For a given value of k , $F(w)$ can be recovered from the N bunched sampling sets of the signal in the interval $I_k := [(-N + k - 1)T^{-1} + \sigma, kT^{-1} - \sigma]$ by multiplying each undersampled set’s dual space representation, S_{T_n} by an undetermined coefficient, $c_{k,n}$, selected to remove the periodizations for $l = -N + k, \dots, k - 1$, i.e.,

$$(4.2) \quad F_k(w) := \sum_{n=1}^N c_{k,n} e^{-2\pi i T_n w} S_{T_n}(w) \quad \text{with } F_k(w) = F(w) \quad \text{for } w \in I_k.$$

The coefficients can then be determined by solving the resulting system⁸

$$(4.3) \quad AR(k)c(k) = e_{N-k+1} \quad \text{with} \quad A_{m,n} := \exp(2\pi iT_n T^{-1}m), \quad m, n = 1, \dots, N,$$

$$c(k) = (c_{k,1} \ c_{k,2} \ \dots \ c_{k,N})^T, \quad e_{N-k+1} = \delta_{j,N-k+1}, \quad \text{and}$$

$$R(k) := \text{diag}(\gamma_1(k) \ \gamma_2(k) \ \dots \ \gamma_N(k)), \quad \text{where} \quad \gamma_n(k) := \exp(2\pi iT_n T^{-1}(k - N - 1)).$$

Repeating this process for a sufficient set of intervals to cover the bandwidth of the signal, $[-\sigma, \sigma] \subseteq \cup_{j=1}^{\kappa} I_{k_j}$, removes the sampling induced periodization for the bandwidth of interest.⁹ However, the overall domain is segmented into κ overlapping intervals which must be spliced together with partitioning functions $\Phi_{k_j}(w)$ constructed appropriately to recover $F(w)$,

$$(4.4) \quad F(w) = \sum_{j=1}^{\kappa} F_{k_j}(w)\Phi_{k_j}(w), \quad |w| \leq \sigma.$$

For partitions which are not dependent on the translates, $\{T_n\}_{n=1}^N$ requires $\Phi_k(w) = 0$ for $w \notin I_k$ and consequently $\sum_{j=1}^{\kappa} \Phi_{k_j}(w) = 1$ for $|w| \leq \sigma$. In summary, the signal's dual space representation can be recovered from its bunched sampling if the following conditions on the intervals and partitioning functions are satisfied:

$$(4.5) \quad [-\sigma, \sigma] \subseteq \cup_{j=1}^{\kappa} I_{k_j}, \quad \Phi_{k_j}(w) = 0 \quad \text{for} \quad w \notin I_{k_j}, \quad \sum_{j=1}^{\kappa} \Phi_{k_j}(w) = 1 \quad \text{for} \quad |w| \leq \sigma.$$

Accordingly, a bandlimited signal's ($f(\cdot) \in B_\sigma$) Fourier transform can be recovered from its overlapping induced periodization by the set of filters $\{\Psi_n(\cdot)\}_{n=1}^N$,

$$(4.6) \quad \begin{aligned} F(w) &= \sum_{j=1}^{\kappa} \Phi_{k_j}(w) \sum_{n=1}^N S_{T_n}(w) c_{k_j,n} e^{-2\pi iT_n w} \\ &= \sum_{n=1}^N \left(\sum_{j=1}^{\kappa} c_{n,k_j} \Phi_{k_j}(w) \right) \sum_{l=-\infty}^{\infty} e^{-2\pi i l T_n T^{-1}} F(w - lT^{-1}) \\ &= \sum_{n=1}^N \Psi_n(w) \sum_{l=-\infty}^{\infty} e^{-2\pi i l T_n T^{-1}} F(w - lT^{-1}) \\ &= T \sum_{n=1}^N \Psi_n(w) \sum_{l=-\infty}^{\infty} f(lT + T_n) e^{-2\pi i w(lT + T_n)}, \end{aligned}$$

where the last equality is due to the Poisson summation formula, and the filters are defined as

$$(4.7) \quad \Psi_n(w) := \sum_{j=1}^{\kappa} c_{k_j,n} \Phi_{k_j}(w),$$

⁸The matrix A is of Vandermonde type and is therefore invertible for distinct translates, $\{T_n\}_{n=1}^N$.

⁹The effective oversampling rate $N/r > 1$ guarantees the full set of intervals cover the signal's bandwidth, $[-\sigma, \sigma] \subseteq \cup_{k=1}^N I_k$.

with I_{k_j} and Φ_{k_j} selected to satisfy conditions (4.5), and $c_{k_j,n}$ being the solutions of the system (4.3). Note that this N stage filtering is in contrast to the one-step filtering, $F(w) = \Psi(w) \sum_{l=-\infty}^{\infty} F(w-lT^{-1})$, used to remove the periodization induced by uniform oversampling.

Consequently, by taking the inverse Fourier transform of (4.6), the signal can be represented by its bunched samples and translated atoms:

$$(4.8) \quad f(t) = T \sum_{n=1}^N \sum_{l=-\infty}^{\infty} f(lT - T_n) \psi_n(t - (lT + T_n)),$$

where ψ_n , the inverse Fourier transform of Ψ_n , is the atom associated with the under-sampled set $\{lT - T_n\}_{l \in \mathbb{Z}}$. This is the corresponding generalization of the oversampling representation (2.3) to the case of bunched sampling.

Similar to the case of uniform oversampling, the truncation error for bunched sampling, (4.8), is governed by the atom's localization:

$$(4.9) \quad \begin{aligned} \epsilon_b(t, L, T) &:= \left| f(t) - T \sum_{n=1}^N \sum_{|l| \leq L} f(lT - T_n) \psi_n(t - (lT + T_n)) \right| \\ &\leq T \|f\|_{L^\infty} \sum_{n=1}^N \sum_{|l| > L} |\psi_n(t - (lT + T_n))|. \end{aligned}$$

In [5] the Nyquist sampling rate $N = r$ was considered, where the interval $|w| \leq \sigma$ was necessarily partitioned with characteristic functions, $\Phi_k(w) = \chi_{I_k}$. However, similar to uniform Nyquist sampling the abrupt filtering results in atoms with first order decay, and consequently the convergence rate for truncated sets of samples is first order, making it an impractical method for real world applications.

For smooth G^α filters, the atoms, $\psi_n(t)$, possess root-exponential localization, (2.8), and consequently the truncation error satisfies

$$(4.10) \quad \epsilon_b(t, L, T) \leq \text{Const}_{\alpha,\eta} N T \|A^{-1}\| \|f\|_{L^\infty} \cdot \exp(-\alpha(2\pi\eta((L-1)T - |t|))^{1/\alpha}),$$

where $\text{Const}_{\alpha,\eta}$ is as before and $\|A^{-1}\|$ is determined solely by the set of translates, $\{T_n\}_{n=1}^N$.

Remark. (i) The truncation error bound in (4.10) is overly pessimistic in the factor $\|A^{-1}\|$ due to inherent structure in the system of equations $AR(k)c(k) = e_{N-k+1}$. The $(N - k + 1)$ th row of this system simplifies to $\sum_{n=1}^N c_{k,n} = 1$ for each k , imposing the additional structure on the atoms that the sum of the filters is the sum of the partitions,

$$(4.11) \quad \sum_{n=1}^N \Psi_n(w) = \sum_{n=1}^N \sum_{j=1}^{\kappa} c_{n,k_j} \Phi_{k_j}(w) = \sum_{j=1}^{\kappa} \Phi_{k_j}(w),$$

which is by construction a smooth function satisfying condition (2.2) with a modified bandwidth Ω . Applying the Fourier transform to (4.11), this structure implies that the atoms sum to a fixed function, independent of the set of translates. As a result, even when the translates are such that the matrix A is ill conditioned, a substantial amount of cancellation between the

atoms significantly reduces its effect on the truncation error. To see this more quantitatively, first group those translates that are near one another, say $\{T_n\}_{n \in \lambda}$. Then let x_λ be the overall amount of e_{N-k+1} contained in the space spanned by the associated columns of $AR(k)$. We observe in (4.3) that due to the linear dependence of the columns in $AR(k)$ associated with $n \in \lambda$ the coefficients $\{c_{k,n}\}_{n \in \lambda}$ are often large in magnitude. Nonetheless, the sum of these coefficients is dictated by x_λ , which by construction is order one, not by the individual translates. For this reason, although the individual *atoms* associated with $n \in \lambda$ may have large magnitude, determined by $\{c_{k,n}\}_{n \in \lambda}$ rather than x_λ , their sum will not. Combined with the values of $f(lT - T_n)$ for $n \in \lambda$ being nearly equal results in a substantial amount of cancellation between the associated atoms. To capture the effect of this cancellation quantitatively, rather than pass the absolute value onto each element as in (4.9), the bound can be left as

$$(4.12) \quad \epsilon_b(t, L, T) \leq T \sum_{|l| > L} \left| \sum_{n=1}^N f(lT - T_n) \psi_n(t - (lT + T_n)) \right|,$$

where for each l , the atoms corresponding to near translates possess substantial cancellation. The numerical example in Figure 5.3 illustrates this effect, where the error near the sampling boundaries does not increase substantially for highly ill conditioned matrices A , but rather roundoff error in the cancellations pollutes the high resolution near the origin.

(ii) To improve the robustness of the proposed method even further we could multiply each uniform sampling set by some weight, similar to the general nonuniform sampling case discussed in [7]. In fact, by introducing properly chosen weights we can obtain estimates for the condition number of A , since the Toeplitz matrix A^*A is of the same form as the Toeplitz matrix appearing in [10]. We leave the details to the reader.

We now turn our attention to constructing intervals and smooth G^α partitioning functions satisfying condition (4.5) where the number of undersampled sets is sufficient to achieve a density similar to oversampling, $N > r$, with an *effective oversampling rate* of $N/r > 1$.

For minimal oversampling, $N = \lceil r \rceil > r$, the recovered regions only overlap with their immediate neighbors, i.e., $I_k \cap I_j = \emptyset$ for $|k - j| > 1$, and the full set of recovered zones $\{I_k\}_{k=1}^N$ is required to cover the interval $[-\sigma, \sigma]$. Moreover, for each I_k there is a subset that is not contained in the other intervals; consequently, the condition $\sum_k \Phi_k(w) = 1$ for $|w| \leq \sigma$ implies that the partitioning functions must satisfy

$$(4.13) \quad \Phi_k(w) = \begin{cases} 1, & w \in [\max(-\sigma, (k-1)T^{-1} - \sigma), \min(\sigma, (-N+k)T^{-1} + \sigma)], \\ 0, & w \notin I_k. \end{cases}$$

An example of G^α partitioning functions satisfying conditions (4.13) and $\sum_k \Phi_k(w) = 1$ for $|w| \leq \sigma$ is

$$(4.14) \quad \Phi_{k_1}(w) := \begin{cases} 0, & w \leq \sigma + (k_1 - 1)T^{-1}, \\ \rho\left(\frac{-w - \sigma}{-(k_1 - 1)T^{-1} - 2\sigma}\right), & \sigma + (k_1 - 1)T^{-1} < w < -\sigma, \\ 1, & -\sigma \leq w \leq \sigma + (k_2 - 1)T^{-1}, \\ \rho\left(\frac{w - (\sigma + (k_2 - 1)T^{-1})}{(k_1 - k_2 + N + 1)T^{-1} - 2\sigma}\right), & \sigma + (k_2 - 1)T^{-1} < w < (k_1 + N - 1)T^{-1} - \sigma, \\ 0, & (k_1 + N - 1)T^{-1} - \sigma \leq w, \end{cases}$$

for the leftmost partition,

$$(4.15) \quad \Phi_{k_j}(w) := \begin{cases} 0, & w \leq \sigma + (k_j - 1)T^{-1}, \\ 1 - \rho\left(\frac{w - (\sigma + (k_j - 1)T^{-1})}{(k_{j-1} - k_j + N + 1)T^{-1} - 2\sigma}\right), & \sigma + (k_j - 1)T^{-1} < w < (k_{j-1} + N - 1)T^{-1} - \sigma, \\ 1, & (k_{j-1} + N - 1)T^{-1} - \sigma \leq w \leq \sigma + (k_{j+1} - 1)T^{-1}, \\ \rho\left(\frac{w - (\sigma + (k_{j+1} - 1)T^{-1})}{(k_j - k_{j+1} + N + 1)T^{-1} - 2\sigma}\right), & \sigma + (k_{j+1} - 1)T^{-1} < w < (k_{j+1} + 1)T^{-1} - \sigma, \\ 0, & (k_{j+1} + 1)T^{-1} - \sigma \leq w, \end{cases}$$

for interior regions $j = 2, 3, \dots, \kappa - 1$, and

$$(4.16) \quad \Phi_{k_\kappa}(w) := \begin{cases} 0, & w \leq \sigma + (k_\kappa - 1)T^{-1}, \\ 1 - \rho\left(\frac{w - (\sigma + (k_\kappa - 1)T^{-1})}{(k_{\kappa-1} - k_\kappa + N + 1)T^{-1} - 2\sigma}\right), & \sigma + (k_\kappa - 1)T^{-1} < w < (k_{\kappa-1} + N)T^{-1} - \sigma, \\ 1, & (k_{\kappa-1} + N)T^{-1} - \sigma \leq w \leq \sigma, \\ \rho\left(\frac{w - \sigma}{(k_\kappa + N)T^{-1} - 2\sigma}\right), & \sigma < w < (k_\kappa + N)T^{-1} - \sigma, \\ 0, & (k_\kappa + N)T^{-1} - \sigma \leq w, \end{cases}$$

for the rightmost partition, where $k_j := j - N$ for $j = 1, 2, \dots, N$ and $\rho(\cdot)$ is defined as in (2.10).

For more general effective oversampling rates, $N > r$, the recovered zones I_k often overlap many of their neighbors, and as such constructing the full set of N partitioning functions satisfying $\sum_k \Phi_k(w) = 1$ for $|w| \leq \sigma$ becomes substantially more complicated. However, for such higher effective oversampling, a smaller number of partitions, $\kappa \leq N$, is required to cover the support of $F(w)$. The ideal subset of intervals and partitioning functions selected through $\{k_j\}_{j=1}^\kappa$ possess minimal slope, requiring that the intervals have equal size of internal and boundary overlaps, i.e., $length(I_{k_j} \cap I_{k_{j+1}}) = length(I_{k_1} / [-\sigma, \sigma]) = length(I_{k_\kappa} / [-\sigma, \sigma])$. Combined with the fixed length of I_k the optimal subset for a given κ is selected as $k_j^* := j \frac{N+1}{\kappa+1} - N$ for $j = 1, 2, \dots, \kappa$. The overlap length for κ sets with k_j as defined before is $T^{-1}(1 - r - \frac{\kappa N - 1}{\kappa + 1})$. The minimum number of partitions κ necessary to cover the bandwidth

$[-\sigma, \sigma]$ is then determined by requiring the overlap interval to be nonnegative, yielding $\kappa_{min} := \lceil \frac{r}{N+1-r} \rceil$.

For computational purposes using the minimum number of partitions, κ_{min} results in unnecessarily steep partitions. Alternatively, a reasonable balance between simplicity of construction and minimizing the partitions slope is achieved by using the maximum number of partitions subject to the constraint that the intervals only interact with their immediate neighbors, yielding $\kappa^* := \min(N, \lfloor \frac{N+1+r}{N+1-r} \rfloor)$. The above results for bunched sampling are summarized in the following theorem.

THEOREM 4.1. *A bandlimited signal $f \in B_\sigma$ can be expressed in terms of its samples on the $N > \lceil r \rceil$ uniform meshes $\{lT + T_n\}_{l \in \mathbb{Z}}$, where $\{T_n\}_{n=1}^N$ are distinct and $T := r/2\sigma$. Reminiscent of the classical Shannon sampling theorem, the signal is decomposed into the translates of N atoms, $\psi_n(\cdot)$, each of which are associated with a particular uniform sampling mesh,*

$$f(t) = T \sum_{n=1}^N \sum_{l=-\infty}^{\infty} f(lT - T_n) \psi_n(t - (lT + T_n)).$$

A particularly simple construction of filters is achieved by solving the system of equations (4.3) for $k_j^ = \text{round}(j \frac{N+1}{\kappa^*+1} - N)$, where $\kappa^* = \min(N, \lfloor \frac{N+1+r}{N+1-r} \rfloor)$. The N filters are then given by*

$$\Psi_n(w) := \sum_{j=1}^{\kappa^*} c_{k_j^*, n} \Phi_{k_j^*}(w),$$

where the coefficient $c_{k_j^, n}$ are determined by solving (4.3) for $\{k_j^*\}_{j=1}^{\kappa^*}$, and the partitioning functions are given by (4.14), (4.15), (4.16).*

Before developing a dual space implementation for truncated bunched sampling, we illustrate partitions and representative atoms for bunched samples in the case of minimal oversampling, $N = \lceil r \rceil > r$.

EXAMPLE 1. *The partitions for $r = 2.4$ as expressed in (4.14), (4.15), (4.16) for the case $N = \lceil r \rceil > r$ are shown in Figure 4.1, where $\rho(w) := \exp[\beta(w-1)^{-1} e^{-1/w}] \in G^2$, with $\beta = e^2/3$. It should be noted that the partitioning does not depend on the translates $\{T_n\}_{n=1}^N$, rather solely on the number of undersampled sets, N .*

The atoms, $\{\psi_n\}_{n=1}^N$, associated with $r = 2.4$, $N = 3$, and random translates $T_n/T = \{-0.4484, 0.3419, -0.0984\}$ are given in Figure 4.2. This distribution of shifts is near the distribution that would correspond to uniform oversampling, $T_n/T = \{-1/3, 0, 1/3\}$ at the rate $N/r = 1.25$, and as such the atoms for this bunched sampling are qualitatively similar to the atom associated with uniform oversampling at the rate $N/r = 1.25$. Figure 4.3 illustrates the atoms for the set of translates $T_n/T = \{-1/3, 0, 10^{-6}\}$, where $\text{cond}(A) = 5.5 \times 10^5$. Having constructed the partitions Φ_k from $\rho(\cdot) \in G^2$, the atoms possess root-exponential localization.

4.1. Direct dual space implementation for bunched sampling. Just as in the case of the single oversampled set, the approximation of $f(\cdot)$ on the zero inserted grid $\{hq\}_{|q| \leq pL}$, where $h = T/p$, can be implemented directly in the dual space. Define

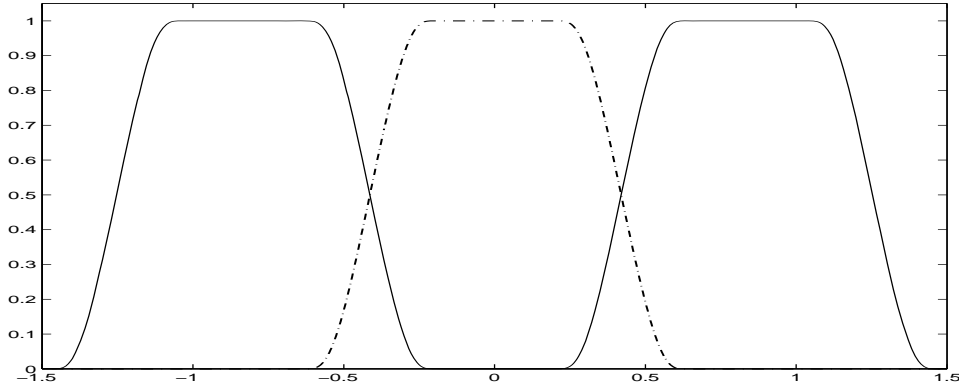


FIG. 4.1. The partitions for $r = 2.4$ and $N = 3$ as described in (4.14), (4.15), (4.16); the end partitions Φ_1, Φ_3 (solid line) and the center partition Φ_2 (dot-dashed line).

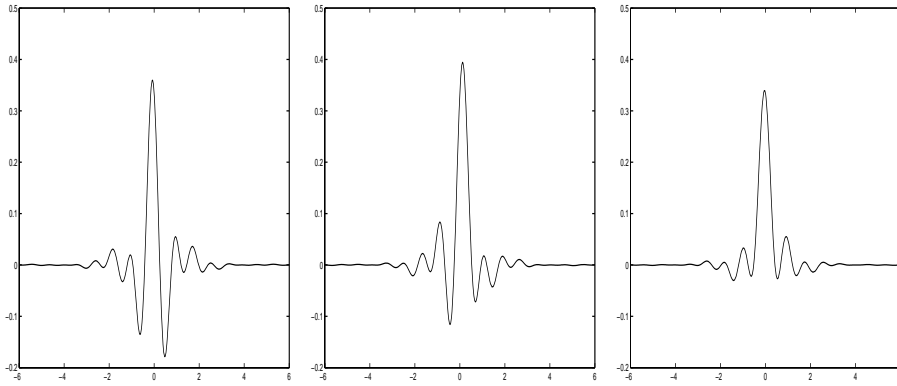


FIG. 4.2. The atoms $\{\psi_n\}_{n=1}^N$ associated with $r = 2.4$, $N = 3$, and random translates $T_n/T = \{-0.4484, 0.3419, -0.0984\}$ shown from left to right, respectively. The atoms are qualitatively similar due to the low condition number $\text{cond}(A) = 1.8939$.

the approximation on this mesh from the bunched sampling as

$$\begin{aligned}
 \text{Approx}_{\psi, B} f(hq) &:= T \sum_{n=1}^N \sum_{|k| \leq L} f(kT - T_n) \psi_n(hq - (kT + T_n)) \\
 (4.17) \qquad &= T \sum_{n=1}^N \sum_{|j| \leq pL} f_0(jh - T_n) \psi_n(h(q - j)T_n),
 \end{aligned}$$

where $f_0(x)$ is zero unless $x = kT - T_n$ for $k = -L, \dots, L$. Replacing the point values in (4.17) with their pseudo-Fourier transform, we can express the evaluation in terms

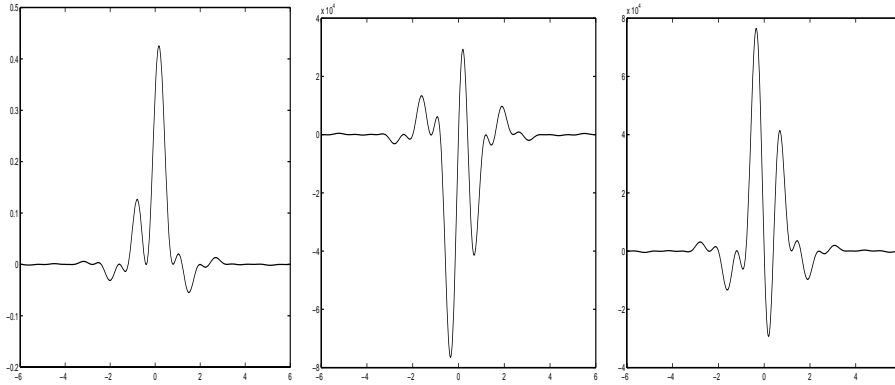


FIG. 4.3. The atoms $\{\psi_n\}_{n=1}^N$ associated with $r = 2.4$, $N = 3$, and translates $T_n/T = \{-1/3, 0, 10^{-6}\}$ shown from left to right, respectively. The first atom is similar to those in Figure 4.2, and although the remaining atoms have significant amplitude, due to the relatively large condition number $\text{cond}(A) = 5.5 \times 10^5$, they are nearly the negative of each other, allowing for significant cancellation.

of the sampling sets pseudo-Fourier transforms,

$$\begin{aligned}
 \text{Approx}_{\psi, B} f(hq) &= T \sum_{n=1}^N \sum_{|j| \leq pL} f_0(jh - T_n) \psi_n(h(q - j)T_n) \\
 &= \frac{2\pi T}{h^2(2pL + 1)^2} \sum_{n=1}^N \sum_{|j| \leq pL} \left(\sum_{|l| \leq pL} \tilde{F}_0(w_l) e^{2\pi i w_l(jh - T_n)} \right) \\
 &\quad \times \left(\sum_{|k| \leq pL} \tilde{\Psi}_n(w_k) e^{2\pi i w_k(h(q - j) - T_n)} \right) \\
 (4.18) \quad &= T\sqrt{2\pi} \sum_{n=1}^N \sum_{|l| \leq pL} \left(\frac{\sqrt{2\pi}}{h(2pL + 1)} \tilde{F}_0(w_l) e^{2\pi i w_l(hq - T_n)} \right) \tilde{\Psi}_n(w_l) e^{-2\pi i T_n w_l}.
 \end{aligned}$$

The last line can be viewed as an algorithm, where first each undersampled set is zero inserted of order p , its pseudo-Fourier transform is computed, and it is multiplied by the appropriately modulated filter, $\tilde{\Psi}_n(\cdot)$, to remove the overlapping periodization. These filtered dual space representations are then summed and their inverse pseudo-Fourier transform computed to achieve an approximation of the bandlimited signal at the set of point $\{hq\}_{|q| \leq pL}$. Just as in the case of the uniform oversampled dual space representation it is computationally advantageous to avoid the construction of $\psi_n(\cdot)$ in order to compute $\tilde{\Psi}_n$; rather for a fast algorithm the true filters $\{\Psi_n(\cdot)\}_{n=1}^N$ should be applied directly. Again we note that the atom's pseudo-Fourier transform is the pL order spectral projection of the atom's associated filter, (3.7). As such their difference is exponentially small,

$$(4.19) \quad \left| \Psi_n(w_l) - \tilde{\Psi}_n(w_l) \right| \leq \text{Const} \cdot \|A^{-1}\| e^{-\alpha(\eta_n LT)^{1/\alpha}}, \quad \Psi_n(\cdot) \in G^\alpha,$$

if the zero padding is sufficient to extend the dual axis beyond the support of the filter, $p \geq 2N - r$.

In addition to the usual truncation error, this additional error gives a threshold below which the error does not fall, determined by the condition number of A as dictated by the set of translates. Unlike the truncation error, the structure in A does not result in cancellation to reduce the effects of $\|A^{-1}\|$; however, the bound does not depend on t and is well below the truncation error for all but the most ill conditioned sets of translates. We summarize the above results in the following theorem.

THEOREM 4.2. *Let $f(t) \in B_\sigma$ be sampled at the points $\{kT + T_n\}_{|k| \leq L}$ with sampling rate $T^{-1} := 2\sigma/r$ and translate $|T_n| \leq T/2$. From $N > r$ such distinct sampling sets, and filters $\{\Psi_n\}_{n=1}^N \in G^\alpha$ constructed as in Theorem 4.1, the signal can be approximated on the set $\{kT/N\}_{|k| \leq pL}$, with $p \geq 2N - r$, within the bound*

$$|f(t) - \text{Approx}_B f(hk)| \leq \text{Const} \frac{N^2}{\sigma} \|f\|_{L^\infty} \|A^{-1}\| \cdot e^{-\alpha(\eta((L-1)T-|t|))^{1/\alpha}},$$

where $\text{Approx}_B f(hk)$ is computed by the following algorithm:

1. Zero insert each of the uniform sampling sets $\{kT + T_n\}_{|k| \leq L}$ to the fine mesh $\{kh + T_n\}_{|k| \leq pL}$, where $p \geq 2N - r$ and $h = T/p$, yielding $\{f_{o,n}\}_{n=1}^N$.
2. Compute the pseudo-Fourier transform, as defined in (3.4), of each set of zero inserted samples from step 1, labeled $\{\tilde{F}_{o,n}(w_l)\}_{n=1}^N$, and pointwise multiply by $\exp(-2\pi iT_n w_l)$, respectively.
3. Pointwise multiply each of the pseudo-Fourier transforms from step 2, $\tilde{F}_{o,n}(w_l)$, by their corresponding filters, $\Psi_n(w_l)$, and sum over N , yielding an approximation of $F(w)$, $\tilde{F}_A(w_l) := \sum_{n=1}^N e^{-2\pi iT_n w_l} \tilde{F}_{o,n}(w_l) \Psi_n(w_l)$.
4. Compute the inverse pseudo-Fourier transform of \tilde{F}_A formed in step 4, and multiply by p .

Similar to the remark following Theorem 4.1, the signal can be recovered on a shifted mesh, $\{kT/N + s_0\}_{|k| \leq pL}$, by multiplying $\tilde{F}_A(w_l)$ with $\exp(2\pi i s_0 w_l)$ between steps 3 and 4 of the algorithm in Theorem 4.2.

The direct implementation for uniform oversampling has a computational cost limited by the FFT, proportional to $L \log(L)$, where L is the number of samples used in the reconstruction. For the algorithm described in Theorem 4.2 for bunched sampling, the computational cost is again limited by the overall FFT evaluations. To allow for a direct comparison to the uniform oversampling let each of the N sampling sets contain L/N samples for a similar total number of samples being available for the algorithm. Each of the N FFTs then requires $\frac{L}{N} \log(pL/N)$, where p is the level of zero insertion required to be proportional to the number of sampling sets, $p \approx N$. The overall computational cost for the algorithm in Theorem 4.2 is then $NL \log(L)$, where L is the total number of samples used in the approximation. In principle we also have to take into account the costs for inverting the matrix A in (4.3). Since A is a Vandermonde matrix and A^*A is a Toeplitz matrix, there are plenty of fast standard algorithms for the solution of the system in (4.3) at our disposal. Moreover, in practice N is small compared to L , and thus the computational costs of this step have little impact on the overall complexity of the proposed method. In summary, the bunched sampling algorithm requires an additional factor of N in the total computational cost, when compared to uniform sampling at the same effective sampling rate and number of available samples.

5. Numerical examples. We illustrate the convergence rates and algorithms for the results presented in Theorems 3.1 and 4.2 for characteristic bandlimited signals. For the approximation of an arbitrary bandlimited signal, we form a test signal

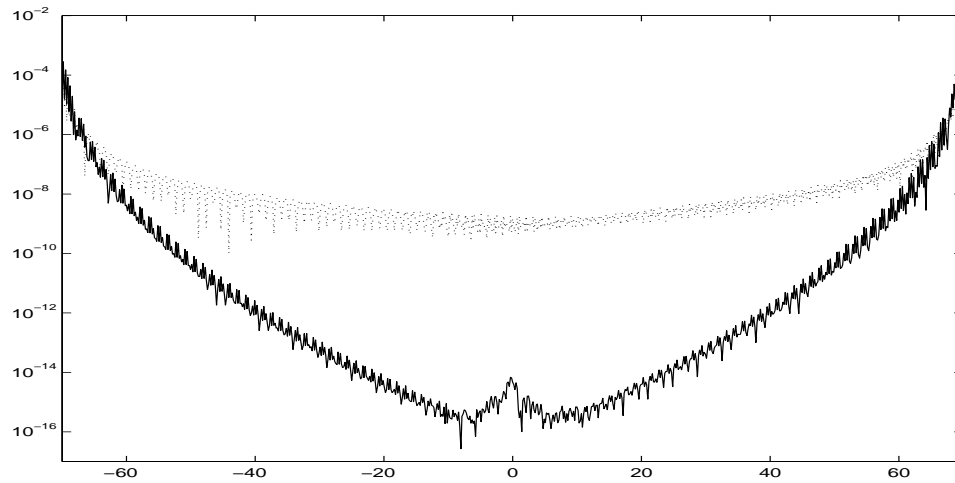


FIG. 5.1. The error with oversample rate $r^{-1} = 1.43$ in recovering the signal in B_1 whose real portion is shown in Figure 2.1. With the log axis, note the log convergence exhibited by the reconstruction using the raised cosine (dotted line), as compared to the root convergence obtained by the dual space implementation using Ψ_{G2} (solid line). The approximations were computed with zero insertion rate $p = 2$ on the grid translated by $s_0 = T/\sqrt{5}$.

whose Fourier transform is composed of one hundred characteristic functions with random complex valued amplitudes normalized to unit l_2 norm, and with random widths and centers, normalized so that the largest magnitude bandwidth is σ . The resulting numerics shown are characteristic of arbitrary complex valued bandlimited signals. The dual space representation, real portion, of such a function can be seen in Figure 2.1. Before illustrating the main results of Theorem 4.2 we briefly contrast the exponential convergence of Gevrey class filters with the polynomial order convergence of classical finite regular filters. To compare representative filters with finite and infinite regularity, we use the canonical raised cosine filter (2.6) and the Gevrey order two filter given in (2.10), respectively; see Figure 5.1.

Much of the success of the raised cosine filter is due to the optimally small first two regularity constants, $\|\Psi\|_{C^s}$ for $s = 1, 2$, which result in rapid initial localization. Infinitely regular filters possess bounded regularity constants for all s , but at the cost of necessarily larger regularity constants for small s . However, a great deal of freedom exists in the selection of G^α regular filters, for example the constant β used in the filter of (2.10). A good approximation of the β which minimizes the first regularity constant in Ψ_{G2} can be obtained by selecting β such that the filter's points of inflection are at the middle of the region connecting zero and one, i.e., $\Psi^{(2)}(\pm\sigma/r) = 0$. As such, for the numerical experiments involving the filter Ψ_{G2} , we use $\beta := \frac{1}{3}e^2$ which satisfies $\rho^{(2)}(\frac{1}{2}) = 0$.

Various properties of Theorem 4.2 are demonstrated in the following numerical examples. First we begin with the simplest case of bunched sampling where only just sufficiently many sampling sets are available for effective oversampling, $N := \lceil r \rceil > r$; in particular for $r = 2.4$ and a random set of well separated translates with corresponding atoms presented in Figure 4.2, and the exponentially small error shown in Figure 5.2. To illustrate the computational robustness, Figure 5.3 shows the error as two sampling sets approach one another, resulting in a poorly conditioned matrix A in system (4.3). However, as noted in the Remark following (4.10), the ill conditioning

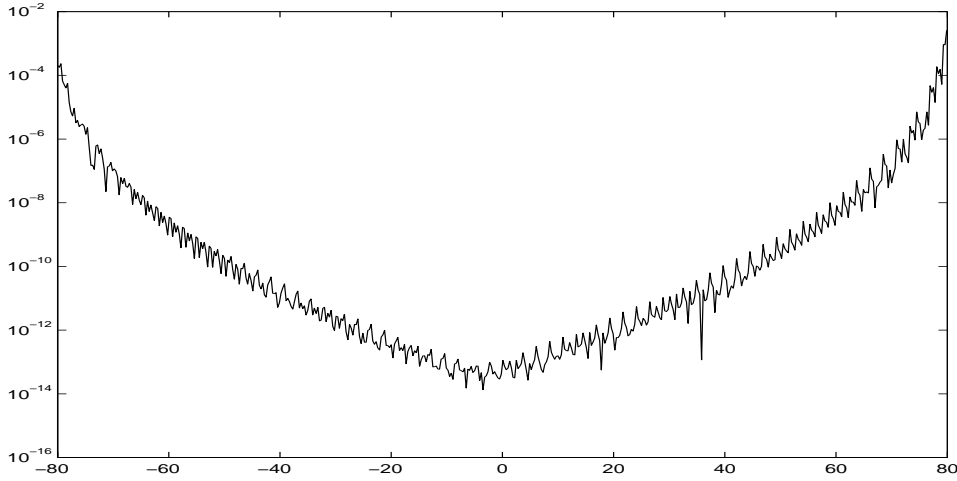


FIG. 5.2. The error with $N = 3$ undersampled sets with effective sampling rate $N/r = 1.25$ in recovering the signal in B_1 whose real portion is shown in Figure 2.1. The translates were $T_n/T = \{-0.4484, -0.0984, 0.3419\}$ with $\text{cond}(A) = 1.8939$ and atoms illustrated in Figure 4.2.

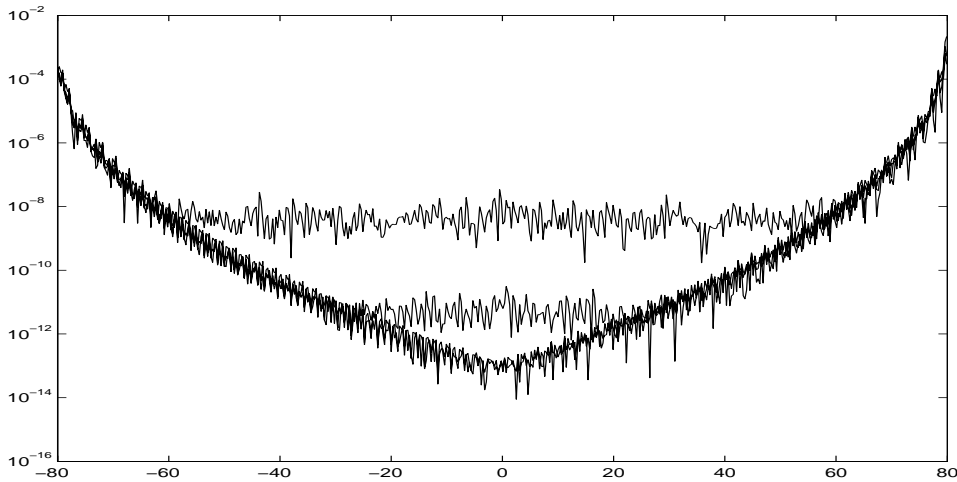


FIG. 5.3. The error with $N = 3$ undersampled sets and effective sampling rate $N/r = 1.25$ in recovering the signal in B_1 whose real portion is shown in Figure 2.1. The translates used, $T_n/T = \{-1/3, 0, 10^{-3j}\}$ for $j = 1, 2, 3$, result in systems (4.3) with respective condition numbers $\text{cond}(A) = 5.5 \times 10^{3j-1}$. Note that although the condition number becomes very large, the error near the boundaries does not suffer. Rather, the minimal error is increased due to roundoff errors. The atoms associated with the set of translations, $j = 2$, are shown in Figure 4.3.

of the matrix does not increase the entire error by the factor $\|A^{-1}\|$ as stated in the pessimistic bound of Theorem 4.2; rather, the ill conditioning results in a rounding error that limits the achievable error for a given precision arithmetic.

A more general example of Theorem 4.2 is shown in Figure 5.4 where $N = 17$ uniform sampling sets with random translates undersampled at the rate $r = 12.4$ are given for an effective sampling rate of $N/r \approx 1.37$. The resulting error is typical for the algorithm of Theorem 4.2 when the system (4.3) has a relatively modest condition

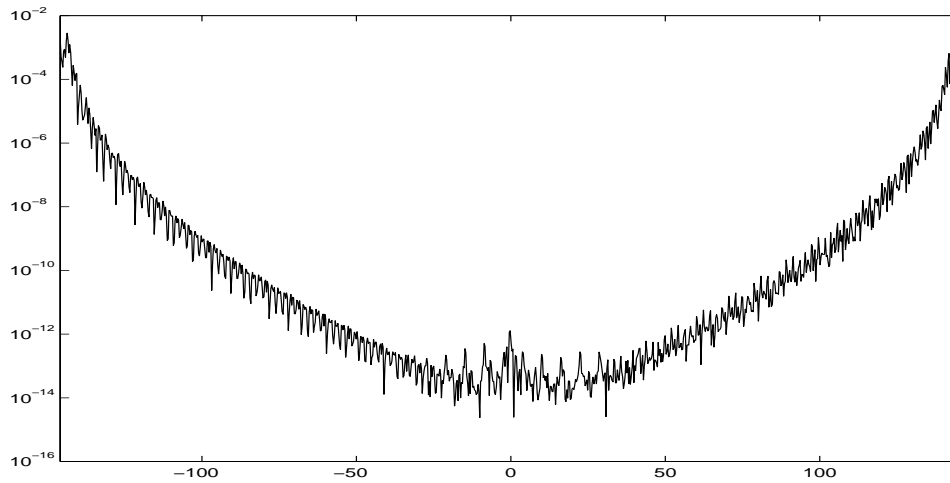


FIG. 5.4. *The error in recovering a random signal in B_1 from $N = 17$ undersampled sets of random translation with effective sampling rate $N/r = 1.37$. The filters were composed of five partitions, $\kappa^* = 5$, and the system of the random translates had condition number $\text{cond}(A) = 1.4 \times 10^4$.*

number.

Before concluding the numerical examples we illustrate the algorithm's performance for a particular application discussed in the introduction. Due to feedback interactions between sampling and processing chips in close proximity, it is advantageous to use a sampling structure that includes relatively large sampling gaps, when the signal processing can be applied and not interact with the sampling. However, the introduced sampling gap could potentially introduce stability problems. The extraordinary robustness of the algorithm in Theorem 4.2 overcomes any stability issues, even for relatively large sampling gaps. For example, Figure 5.5 shows the approximation error with eight sets interleaved over a third of the effective sampling rate, i.e., $T_n := T(n/24)$ for $n = 1, 2, \dots, 8$.

6. Final remarks. We have derived a fast algorithm for reconstructing a bandlimited signal from its periodic nonuniform samples that achieves root-exponential accuracy with respect to the given number of samples. Due to its high accuracy the method can be easily realized in practice via finite impulse response (FIR) filters. Furthermore, since the numerically most expensive steps are FFTs the proposed method lends itself to a simple implementation on standard DSP processors. Furthermore, the high accuracy provided by the algorithms derived in this paper will not be lost in the subsequent reconstruction of the signal from its quantized samples due to the recently developed highly accurate algorithms for recovering a quantized bandlimited signal; cf. [4].

Another application where periodic nonuniform sampling arises is image processing. For instance, in astronomical imaging one is confronted with images that are blurred and notoriously undersampled. The goal is to combine these blurred low-resolution images to one high-resolution image. This problem is also referred to as superresolution; see, e.g., [6]. The low-resolution images contain (blurred, noisy) samples of the high-resolution image where the sampling sets can be thought of as a union of arbitrarily shifted (and/or rotated) uniform sampling sets. One step in reconstruct-

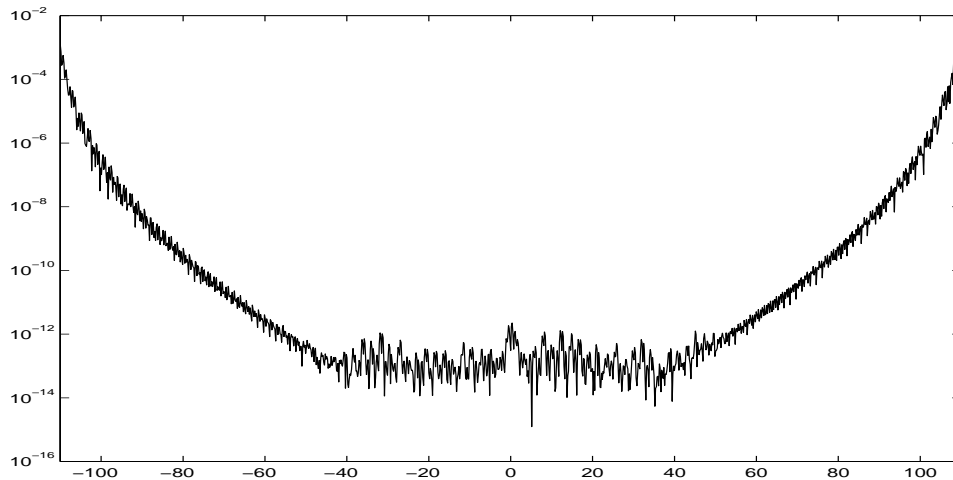


FIG. 5.5. The error in recovering a random signal in B_1 from $N = 8$ undersampled sets with effective sampling rate $N/r = 1.82$. The filters were composed of two partitions, $\kappa^* = 2$, and the translates, $T_n := T(n/24)$ for $n = 1, 2, \dots, 8$, had condition number $\text{cond}(A) = 3.1 \times 10^4$.

ing a high-resolution image is thus the conversion of the periodic nonuniform image samples to a uniform sampling set at a fine sampling grid. In our future research we will address two-dimensional reconstruction algorithms and their effect on deblurring and noise.

Acknowledgment. We are indebted to Professor Bernard Levy for insightful discussions on practical aspects of the research presented in this paper.

REFERENCES

- [1] A. BEURLING AND P. MALLIAVIN, *On the closure of characters and the zeros of entire functions*, Acta Math., 118 (1967), pp. 79–93.
- [2] P. L. BUTZER AND G. HINSEN, *Reconstruction of bounded signals from pseudo-periodic irregularly spaced samples*, Signal Process., 17 (1989), pp. 1–17.
- [3] I. DAUBECHIES, *Ten Lectures on Wavelets*, CBMS-NSF Regional Conf. Ser. in Appl. Math. 61, SIAM, Philadelphia, 1992.
- [4] I. DAUBECHIES AND R. DEVORE, *Approximating a bandlimited function using very coarsely quantized data: A family of stable sigma-delta modulators of arbitrary order*, Ann. of Math. (2), 158 (2003), pp. 679–710.
- [5] A. FARIDANI, *A generalized sampling theorem for locally compact abelian groups*, Math. Comp., 63 (1994), pp. 307–327.
- [6] S. FARSIU, D. ROBINSON, M. ELAD, AND P. MILANFAR, *Advances and challenges in super-resolution*, Int. J. Imaging Systems and Technology, 14 (2004), pp. 47–57.
- [7] H. G. FEICHTINGER AND K. H. GRÖCHENIG, *Theory and practice of irregular sampling*, in Wavelets: Mathematics and Applications, J. Benedetto and M. Frazier, eds., CRC Press, Boca Raton, FL, 1994, pp. 305–363.
- [8] R. GERVAIS, Q. I. RAHMAN, AND G. SCHMEISSER, *A bandlimited function simulating a duration-limited one*, in Anniversary Volume on Approximation Theory and Functional Analysis (Oberwolfach, 1983), Internat. Schriftenreihe Numer. Math. 65, Birkhäuser-Verlag, Basel, Switzerland, 1984, pp. 355–362.
- [9] D. GOTTLIEB AND S. A. ORSZAG, *Numerical Analysis of Spectral Methods: Theory and Applications*, CBMS-NSF Regional Conf. Ser. in Appl. Math. 26, SIAM, Philadelphia, 1977.
- [10] K. GRÖCHENIG, *Irregular sampling of wavelet and short-time Fourier transforms*, Constr. Approx. Theory, 9 (1993), pp. 283–297.
- [11] S. HUESTIS, *Optimum kernels for oversampled signals*, J. Acoust. Soc. Amer., 92 (1992), pp.

- 1172–1173.
- [12] H. JOHANSSON AND P. LÖWENBERG, *Reconstruction of nonuniformly sampled bandlimited signals by means of digital fractional delay filters*, IEEE Trans. Acoust. Speech Sig. Proc., 50 (2002), pp. 2757–2767.
 - [13] F. JOHN, *Partial Differential Equations*, 4th ed., Appl. Math. Sci. 1, Springer-Verlag, New York, 1991.
 - [14] A. KOHLEBERG, *Exact interpolation of bandlimited functions*, J. Appl. Phys., 24 (1953), pp. 1432–1436.
 - [15] F. MARVASTI, ED., *Theory and Practice of Nonuniform Sampling*, Kluwer Academic Publishers, Dordrecht, The Netherlands, Plenum, New York, 2001.
 - [16] F. NATTERER, *Efficient evaluation of oversampled functions*, J. Comput. Appl. Math., 14 (1986), pp. 303–309.
 - [17] R. A. NILAND, *Optimum oversampling*, J. Acoust. Soc. Amer., 86 (1989), pp. 1805–1812.
 - [18] A. PAPOULIS, *Signal Analysis*, McGraw–Hill, New York, 1977.
 - [19] S. PILIPOVIĆ AND N. TEOFANOV, *Wilson bases and ultramodulation spaces*, Math. Nachr., 242 (2002), pp. 179–196.
 - [20] R. S. PRENDERGAST, B. C. LEVY, AND P. J. HURST, *Reconstruction of band-limited periodic nonuniformly sampled signals from multirate filter banks*, IEEE Trans. Circuits Syst. I Regul. Pap., 51 (2004), pp. 1612–1622.
 - [21] T. S. RAPPAPORT, *Wireless Communications: Principles & Practice*, Prentice–Hall, Upper Saddle River, NJ, 1996.
 - [22] T. STROHMER AND J. TANNER, *Implementations of Shannon’s sampling theorem, A time-frequency approach*, Sampl. Theory Signal Image Process., 4 (2005), pp. 1–17.
 - [23] T. STROHMER AND J. XU, *Blind calibration of multiple time-interleaved analog-to-digital converters*, IEEE Trans. Signal Process., submitted, 2006.
 - [24] P. P. VAIDYANATHAN, *Multirate Systems and Filter Banks*, Prentice–Hall, Englewood Cliffs, NJ, 1993.
 - [25] P. P. VAIDYANATHAN AND V. C. LIU, *Efficient reconstruction of band-limited sequences from nonuniformly decimated versions by use of polyphase filter banks*, IEEE Trans. Acoust. Speech Signal Process, 38 (1990), pp. 1927–1936.
 - [26] A. I. ZAYED, *Advances in Shannon’s Sampling Theory*, CRC Press, Boca Raton, FL, 1993.