Error Analysis of Frame Reconstruction from Noisy Samples

Akram Aldroubi†*, Casey Leonetti †, Qiyu Sun#

EDICS Sampling, extrapolation, and interpolation.

***Abstract*— This paper addresses the problem of reconstructing a continuous function defined on $\mathbb{R}^d$ from a countable collection of samples corrupted by noise. The additive noise is assumed to be i.i.d. with mean zero and variance σ^2. We sample the continuous function f on the uniform lattice $\frac{1}{m}\mathbb{Z}^d$, and show for large enough m that the variance of the error between the frame reconstruction $f_{\varepsilon,m}$ from noisy samples of f and the function f satisfy var$(f_{\varepsilon,m}(x)-f(x)) \approx \frac{\sigma^2}{m^d}\mathcal{C}_x$ where $\mathcal{C}_x$ is the best constant for every $x \in \mathbb{R}^d$. We also prove a similar result in the case that our data are weighted-average samples of f corrupted by additive noise.**

***Index Terms*— Sampling, reconstruction from averages, frames.**

I. Introduction

Sampling and function reconstruction have been widely studied in recent decades, particularly within the setting of shift-invariant spaces (see [1] - [9]). However, the problem of reconstructing a function in shift-invariant spaces from data corrupted by noise has not been given as much attention. For bandlimited functions and L^2 with some regularity properties, Pawlak, Rafajlowowicz and Krzyzak give a reconstruction algorithms and detailed analyses of the error of reconstruction for white noise, colored noise and finite samples [23], see also [21]. In [12], Eldar and Unser provide optimal results for filtering noisy samples of signals from shift-invariant and bandlimited spaces. Smale and Zhou reconstruct signals from noisy data in [25] and give error estimates for the reconstructed signal. In [24] Rohde et al. show that reconstruction from noisy data introduces spatial dependent artifacts that are undesirable for sub-pixel signal processing. The main problem is that reconstruction from noisy samples introduces spatially-dependent (or time-dependent) noise in the reconstructed signal. Thus, an accurate estimate of the noise at each point of the reconstructed signal is desirable. For a general review of reconstruction of functions from noisy samples we refer to [22], [8] and the references therein. In this paper, we provide error estimates for frame reconstruction of a continuous function from a countable collection of sampled data that is corrupted by noise, and give an exact formula for the variance as a function of the position x, of the oversampling factor m, and of the signal and sampling models.

In particular, given data $Y = \{y_j\}_{j\in J}$ of the form $y_j = f(x_j)+\varepsilon_j$, we analyze the frame reconstruction algorithm that produces a continuous function f_ε from the noisy samples $Y = \{y_j\}_{j\in J}$ of a function f in a shift invariant space. We assume the noise sequence $\{\varepsilon_j\}_{j\in J}$ to be a collection of i.i.d. random variables with $\mathrm{E}(\varepsilon_j) = 0$ and var$(\varepsilon_j) = \sigma^2$. We consider uniform sets of sampling of the form $\frac{1}{m}\mathbb{Z}^d$, where m is a positive integer, and find precise estimates of var$(f_{\varepsilon,m}(x) - f(x))$ which is a function of x.

We address this problem not only for exact sampling, but also for weighted average sampling as in [1] and [5]. Specifically, instead of assuming the data $\{y_j\}_{j\in J}$ arise from exact samples of f, we assume the data are of the form $y_j = \langle f, \psi(\cdot - x_j)\rangle + \varepsilon_j$, or even $y_j^i = \langle f, \psi^i(\cdot - x_j)\rangle + \varepsilon_j^i$, $1 \le i \le s$, for some vector function $\Psi = (\psi^1,\dots,\psi^s)^T$. In this case, the uncorrupted data can be interpreted as weighted averages of f at x_j.

We begin this paper by precisely defining and characterizing the underlying shift-invariant space from which our continuous signals originate. As is common in much of the current research (see [1]-[5],[14]), our underlying space will be of the form

$$V^2(\Phi) = \left\{ \sum_{k\in\mathbb{Z}^d} C(k)^T\Phi(\cdot - k) : C \in (l^2)^{(r)} \right\} \tag{1}$$

for some real-valued vector function $\Phi = \left(\phi^1,\dots,\phi^r\right)^T \in (L^2)^{(r)}$, where $C = (c^1,\dots,c^r)^T$ is a real-valued vector sequence such that $c^i := \{c^i(k)\}_{k\in\mathbb{Z}^d} \in l^2$, i.e., $C \in (l^2)^{(r)}$. Thus $\sum_{k\in\mathbb{Z}^d} C(k)^T\Phi(\cdot - k) = \sum_{i=1}^r \sum_{k\in\mathbb{Z}^d} c^i(k)\phi^i(\cdot - k)$. We also introduce definitions and tools crucial to our analysis. Then the case of exact sampling is considered first, and the main theorem is stated. While the complete proof is saved for section V, the main ideas behind the proof are illustrated by looking at the simpler case in section III-A. Then in section IV, we address the weighted-average sampling problem and state the main result there. Once again, the complete proof is saved for section V, while we illustrate the ideas through a simpler setting in IV-A.

The work of the first author was supported in part by grant NSF/DMS-0504788.

†* A. Aldroubi (corresponding author) is with the Department of Mathematics, Vanderbilt University, Nashville, TN 37240. The telephone number is (304)293-2011. The fax number is (304)-293-3982. Email address: aldroubi@math.vanderbilt.edu

† C. Leonetti is with the Department of Mathematics, Vanderbilt University, Nashville, TN 37240. The telephone number is (713)-289-4319. Email address: caseyleonetti@gmail.com

Q. Sun is with the Department of Mathematics, University of Central Florida, Orlando, FL 32816. The telephone number is (407)-8234839. The fax number is (407)-8236253. Email address: qsun@mail.ucf.edu

II. Notation and preliminaries

We begin by defining $V^2(\Phi)$ more precisely. As mentioned before, shift-invariant spaces are commonly used in sampling models. Moreover, it is common to consider continuous shift-invariant spaces that are subspaces of $L^2(\mathbb{R}^d)$ in order to take advantage of reproducing kernel Hilbert space properties.

Let $\Phi = (\phi^1, \dots, \phi^r)^T$, where $\phi^i : \mathbb{R}^d \to \mathbb{R}$ is a function in $L^2(\mathbb{R}^d)$, and assume Φ is such that

$$G_\Phi(\xi) := \sum_{k\in\mathbb{Z}^d} \widehat{\Phi}(\xi+k)\overline{\widehat{\Phi}(\xi+k)}^T = I, \quad \text{a.e. } \xi \in \mathbb{R}^d, \tag{2}$$

where I is the $r \times r$ identity matrix. Define the shift-invariant space

$$V^2(\Phi) := \left\{ \sum_{k\in\mathbb{Z}^d} C(k)^T \Phi(\cdot - k) : C \in (l^2)^{(r)} \right\}.$$

Then $V^2(\Phi)$ is a Hilbert space, $V^2(\Phi)$ is a subspace of $L^2(\mathbb{R}^d)$, and $\{\phi^i(\cdot - k) : 1 \le i \le r, k \in \mathbb{Z}^d\}$ forms an orthonormal basis for $V^2(\Phi)$ [1], [3]. Assumption (2) insures that $\{\phi^i(\cdot - k) : 1 \le i \le r, k \in \mathbb{Z}^d\}$ is an orthonormal basis for the space $V^2(\Phi)$. However, what is important is the space $V^2(\Phi)$ since there are many bases generating the same space. Some of these bases may be more useful for computational purposes or for certain applications as we will see below. Also assume $\phi^i \in W_0^1 := W^1 \cap C^0$, where C^0 is the set of continuous functions, and

$$W^1 = \left\{ f : \sum_{k\in\mathbb{Z}^d} \operatorname*{ess\,sup}_{x\in[0,1]^d} \{|f(x+k)|\} < \infty \right\}.$$

Under this assumption, $V^2(\Phi)$ is a space of continuous functions [3]. Furthermore, with this assumption, for each x in $\mathbb{R}^d$, the point evaluation map $f \mapsto f(x)$, from $V^2(\Phi)$ to $\mathbb{R}$, is bounded. To see this, denote the sequence $a_x^i(k) := \phi^i(x-k)$, and notice that for every $x \in \mathbb{R}^d$, $\left\|a_x^i\right\|_{l^1(\mathbb{Z}^d)} \le \left\|\phi^i\right\|_{W^1}$. Let $f = \sum_{i=1}^r \sum_{k\in\mathbb{Z}^d} c^i(k)\phi^i(\cdot - k) \in V^2(\Phi)$. Then

$$\begin{aligned} |f(x)| &\le \sum_{i=1}^r \sum_{k\in\mathbb{Z}^d} \left|c^i(k)\right| \left|\phi^i(x-k)\right| \\ &= \sum_{i=1}^r \left\langle \left|c^i\right|, \left|a_x^i\right| \right\rangle_{l^2} \\ &\le \sum_{i=1}^r \left\|c^i\right\|_{l^2} \left\|a_x^i\right\|_{l^1} \\ &\le \left(\sum_{i=1}^r \left\|c^i\right\|_{l^2}^2 \right)^{1/2} \left(\sum_{i=1}^r \left\|\phi^i\right\|_{W^1}^2 \right)^{1/2} \\ &\le \left(\sum_{i=1}^r \left\|\phi^i\right\|_{W^1}^2 \right)^{1/2} \|f\|_{L^2}. \end{aligned}$$

We conclude that point evaluation is a continuous linear functional on $V^2(\Phi)$. Therefore, by the Riesz representation Theorem, for every $x \in \mathbb{R}^d$, there exists a reproducing kernel $K_x \in V^2(\Phi)$ satisfying $\langle f, K_x \rangle = f(x)$ for all $f \in V^2(\Phi)$. In fact, it can easily be shown that

$$K_x(y) = \sum_{i=1}^r \sum_{l\in\mathbb{Z}^d} \phi^i(x-l)\phi^i(y-l). \tag{3}$$

Most of signal space models use the assumption that $\phi^i \in W_0^1 := W^1 \cap C^0$ (e.g., signals modelled with multiresolution spaces, B-spline spaces of degree $n \ge 1$). However, there are two spaces that are often used that do not satisfy this assumption. One of these spaces is the space of piecewise constant case where $\phi = \chi_{[0,1)}$ (Obviously this space does not belong to C^0). The other space is the space of bandlimited functions generated by $\phi = \operatorname{sinc}$ which does not belong to W^1 (the bandlimited function space belongs to W^p for any $p > 1$). However minor modifications show that both of these spaces can be treated in similar ways). In particular, in both cases point evaluations are bounded linear functionals, and the reproducing kernel for the bandlimited case has a simple expression given by $K_x(y) = \sum_{l\in\mathbb{Z}^d} \operatorname{sinc}(x-l)\operatorname{sinc}(y-l) = \operatorname{sinc}(x-y)$. Once the underlying space $V^2(\Phi)$ is fixed, the ability to recover a function f in $V^2(\Phi)$ from its samples, $\{f(x_j)\}_{j\in J}$, depends on the sampling set $X := \{x_j : j \in J\}$. Let X be a countable subset of $\mathbb{R}^d$.

Definition 1: We say that X is a set of sampling for $V^2(\Phi)$ if there exist positive constants α and β such that

$$\alpha||f||_{L^2} \le ||\{f(x_j)\}_{x_j\in X}||_{l^2(J)} \le \beta||f||_{L^2}, \ \forall\ f \in V^2(\Phi). \tag{4}$$

Notice that if $X := \{x_j : j \in J\}$ is a set of sampling for $V^2(\Phi)$, then the collection $\{K_{x_j}\}_{j\in J}$ forms a frame for $V^2(\Phi)$, which gives us the following stable reconstruction formula for $f \in V^2(\Phi)$:

$$f = \sum_{j\in J} \left\langle f, K_{x_j} \right\rangle \widetilde{K}_{x_j}, \tag{5}$$

where $\{\widetilde{K}_{x_j}\}_{j\in J}$ is the canonical dual frame associated with $\{K_{x_j}\}_{j\in J}$. Namely, $\widetilde{K}_{x_j} := S^{-1}K_{x_j}$, where S is the frame operator on $V^2(\Phi)$ associated with the frame $\{K_{x_j}\}_{j\in J}$, i.e.

$$Sf = \sum_{j\in J} \left\langle f, K_{x_j} \right\rangle K_{x_j}. \tag{6}$$

The operator S is positive and invertible. Moreover, given any sequence $c \in l^2(J)$, the function $\sum_{j\in J} c(j)\widetilde{K}_{x_j}$ is in $V^2(\Phi)$. See [10] and [15] for more information and background on frames.

With data $\{y_{x_j}\}$ of the form $y_{x_j} = f(x_j) + \varepsilon_{x_j}$, we can estimate f by f_ε, given by

$$f_\varepsilon := \sum_{j\in J} y_{x_j} \widetilde{K}_{x_j}.$$

Notice that $f_\varepsilon = f$ precisely when there is no noise, i.e. when $\varepsilon_{x_j} = 0$ for all $j \in J$. In this paper, our goal is to give estimates on the error $|f_\varepsilon(x) - f(x)|$ in terms of the noise sequence $\{\varepsilon_{x_j}\}$.

A. Fourier Analysis

Our analysis will heavily rely on the Fourier transform and its properties. We denote the Fourier transform of a function $f \in L^2(\mathbb{R}^d)$ by $\widehat{f}$ and define

$$\widehat{f}(\xi) = \int_{\mathbb{R}^d} f(x) e^{-\mathbf{i}2\pi x\cdot\xi}\, dx \qquad \text{a.e. } \xi \in \mathbb{R}^d, \tag{7}$$

where $\mathbf{i} = \sqrt{-1}$. The function $\widehat{f}$ is also in $L^2(\mathbb{R}^d)$, and $\|f\|_{L^2(\mathbb{R}^d)} = \|\widehat{f}\|_{L^2(\mathbb{R}^d)}$. Similarly, we denote the Fourier series of a sequence $c \in l^2(\mathbb{Z}^d)$ by $\widehat{c}$ and define

$$\widehat{c}(\xi) = \sum_{k\in\mathbb{Z}^d} c(k) e^{-\mathbf{i}2\pi k\cdot\xi} \qquad \text{a.e. } \xi \in [0,1]^d. \tag{8}$$

The function $\widehat{c}$ is in $L^2([0,1]^d)$, and $\|c\|_{l^2(\mathbb{Z}^d)} = \|\widehat{c}\|_{L^2([0,1]^d)}$. The following properties of the Fourier transform will frequently be used.

(i) $\widehat{\tau_y f}(\xi) = e^{-\mathbf{i}2\pi y\cdot\xi}\widehat{f}(\xi)$ where $\tau_y f = f(\cdot - y)$.

(ii) $\hat{\hat{f}} = f^\vee$ where $f^\vee(x) = f(-x)$.

(iii) $\widehat{f^\vee} = \overline{\widehat{f}}$ if f is real-valued.

(iv) $\widehat{f * g} = \widehat{f}\,\widehat{g}$.

For vector functions $F = (f^1, \dots, f^n)^T$, the notation $\widehat{F}$ will represent the vector $(\widehat{f^1}, \dots, \widehat{f^n})^T$.

Another valuable tool from Fourier analysis is the Poisson Summation Formula. If $\sum_{k\in\mathbb{Z}^d} f(x+k) \in L^2([0,1]^d)$, and if $\sum_{k\in\mathbb{Z}^d} \left|\widehat{f}(k)\right|^2 < \infty$, then

$$\sum_{k\in\mathbb{Z}^d} f(x+k) = \sum_{k\in\mathbb{Z}^d} \widehat{f}(k) e^{\mathbf{i}2\pi k\cdot x} \qquad \text{a.e. } x \in \mathbb{R}^d. \tag{9}$$

More often we will use the equivalent version

$$\sum_{k\in\mathbb{Z}^d} \widehat{f}(\xi+k) = \sum_{k\in\mathbb{Z}^d} f(k) e^{-\mathbf{i}2\pi k\cdot\xi} \qquad \text{a.e. } \xi. \tag{10}$$

Notice the right-hand side of the equation is the Fourier series of the sequence whose terms are samples of f on the integer lattice. See [15] for an extensive review of the Fourier transform and its properties.

B. The Zak Transform

The Zak transform of a function f is denoted Zf and defined as

$$Zf(t,\xi) = \sum_{k\in\mathbb{Z}^d} f(t-k) e^{\mathbf{i}2\pi k\cdot\xi}. \tag{11}$$

If $f \in L^2(\mathbb{R}^d)$, then Zf is well-defined almost everywhere in $\mathbb{R}^d \times \mathbb{R}^d$. If $f \in W_0^1$, then Zf is a well-defined, continuous function on $\mathbb{R}^d \times \mathbb{R}^d$ [17]. A simple exercise shows also that (9) implies

$$Zf(x,\xi) = e^{\mathbf{i}2\pi x\cdot\xi}\; Z\widehat{f}(\xi, -x). \tag{12}$$

For a vector function $F = (f^1, \dots, f^n)^T$, we denote by ZF the vector $(Zf^1, \dots, Zf^n)^T$.

III. Exact Sampling

Here we sample on the lattice $\frac{1}{m}\mathbb{Z}^d$, i.e., we assume our data are of the form

$$\{y_{k+j/m} = f(k+j/m) + \varepsilon_{k+j/m} : k \in \mathbb{Z}^d, j \in \Omega_m^d\}$$

for some function $f \in V^2(\Phi)$. For the sake of simplicity, we denote the finite set $\Omega_m^d := \mathbb{Z}^d \cap [0, m-1]^d$, and we use the notation j/m for $\frac{1}{m}j$, where m is a positive integer and j is a vector in Ω_m^d. We also assume that for $m \geq 1$, the lattice $\frac{1}{m}\mathbb{Z}^d$ is a set of sampling for $V^2(\Phi)$, i.e., there exist positive constants α_m and β_m satisfying

$$\alpha_m \|f\|_{L^2}^2 \leq \sum_{j\in\Omega_m^d}\sum_{k\in\mathbb{Z}^d} |f(k+j/m)|^2 \leq \beta_m \|f\|_{L^2}^2 , \tag{13}$$

for all $f \in V^2(\Phi)$. Thus the collection of reproducing kernels $\{K_{k+j/m} : k \in \mathbb{Z}^d, j \in \Omega_m^d\}$ forms a frame for $V^2(\Phi)$, and $f \in V^2(\Phi)$ is uniquely determined by its samples $\{f(k+\frac{1}{m}j) : k \in \mathbb{Z}^d, j \in \Omega_m^d\}$.

Remark 2: It is reasonable to make the assumption that (13) holds. From the results in [5], we know that there exists an $M \in \mathbb{N}$ such that positive α_m and β_m satisfying (13) exist for all $m \geq M$. Moreover, if positive α_1 and β_1 exist (i.e., if $\mathbb{Z}^d$ is a set of sampling for $V^2(\Phi)$), then positive α_m and β_m exist for all $m \in \mathbb{N}$.

Recall from the previous section that f can be recovered from its samples as follows:

$$\begin{aligned} f &= \sum_{j\in\Omega_m^d}\sum_{k\in\mathbb{Z}^d} \langle f, K_{k+j/m}\rangle \widetilde{K}_{k+j/m} \qquad (14)\\ &= \sum_{j\in\Omega_m^d}\sum_{k\in\mathbb{Z}^d} f(k+j/m)\widetilde{K}_{k+j/m}. \end{aligned}$$

Given data $\{y_{k+j/m} = f(k+\frac{1}{m}j) + \varepsilon_{k+j/m}\}$, we define

$$f_{\varepsilon,m} := \sum_{j\in\Omega_m^d}\sum_{k\in\mathbb{Z}^d} y_{k+j/m}\widetilde{K}_{k+j/m}.$$

The expected value and variance of the error between the frame reconstruction $f_{\varepsilon,m}$ and the exact function f is a function of the position x, the oversampling factor m^d, and the noise variance σ^2. The precise estimates and best constants are given by the following theorem.

Theorem 3: Let $\Phi = (\phi^1, \dots, \phi^r)^T$ satisfy $G_\Phi(\xi) = I$ a.e. ξ, and $\phi^i \in W^1 \cap C^0$, $1 \leq i \leq r$. For $m \in \mathbb{N}$, let $\alpha_m, \beta_m > 0$ satisfy (13). Assume, for all $k \in \mathbb{Z}^d$ and $j \in \Omega_m^d$, that $y_{k+j/m} = f(k+j/m) + \varepsilon_{k+j/m}$ for some $f \in V^2(\Phi)$, where $\{\varepsilon_{k+j/m}\}$ is a collection of i.i.d. random variables satisfying $E(\varepsilon_{k+j/m}) = 0$, $\mathrm{var}(\varepsilon_{k+j/m}) = \sigma^2$ and $\varepsilon_{k+j/m} \in [-N, N]$ for some $N < \infty$. Then $E(f_{\varepsilon,m}(x) - f(x)) = 0$, and

$$\mathrm{var}(f_{\varepsilon,m}(x) - f(x)) = \frac{\sigma^2}{m^d} C_x(m),$$

where $C_x(m)$ is given by (35), and we have

$$C_x(m) \;\overset{m\to\infty}{\longrightarrow}\; \sum_{i=1}^r \int_{[0,1]^d} \left|Z\widehat{\phi}^i(-\xi,-x)\right|^2 \, d\xi.$$

Remark 4: In the proofs of section III-A we show that we can also obtain slightly suboptimal estimates that are independent of m or x. In particular, for any $\epsilon > 0$, there exists a number $M \in \mathbb{N}$ such that for all $m \geq M$

$$\operatorname{var}(f_{\varepsilon,m}(x) - f(x)) \leq \frac{(1+\epsilon)\sigma^2}{m^d} \left(\sum_{i=1}^{r} \left\| \phi^i \right\|_{W^1}^2 \right)$$

for all $x \in \mathbb{R}^d$.

A. Exact Sampling in $V^2(\phi)$

Before presenting the proof of the theorem above, we illustrate the simpler case where $r = 1$. In other words, our underlying shift-invariant space has only one generator ϕ. This will also serve to lay the groundwork for the proof of Theorem 3.

Recall that the inequality (13) implies that $\{K_{k+j/m} : k \in \mathbb{Z}^d, j \in \Omega_m^d\}$ is a frame for $V^2(\phi)$, where

$$K_{k+j/m}(x) = \sum_{l \in \mathbb{Z}^d} \phi(k + j/m - l)\phi(x - l), \qquad (15)$$

and f can be reconstructed from its samples on the lattice $\frac{1}{m}\mathbb{Z}^d$ as shown

$$\begin{aligned} f &= \sum_{j \in \Omega_m^d} \sum_{k \in \mathbb{Z}^d} \left\langle f, K_{k+j/m} \right\rangle \widetilde{K}_{k+j/m} \\ &= \sum_{j \in \Omega_m^d} \sum_{k \in \mathbb{Z}^d} f(k + j/m) \widetilde{K}_{k+j/m}. \end{aligned} \qquad (16)$$

Because our sampling set is uniform, we can find the Fourier transform of $\widetilde{K}_{k+j/m} = S_m^{-1} K_{k+j/m}$ explicitly. Recall, for any $f \in V^2(\phi)$, that

$$(S_m f)(x) = \sum_{j \in \Omega_m^d} \sum_{k \in \mathbb{Z}^d} \left\langle f, K_{k+j/m} \right\rangle K_{k+j/m}(x). \qquad (17)$$

Notice that

$$K_{k+j/m} = K_{j/m}(\cdot - k) \qquad \text{for all } k \in \mathbb{Z}^d.$$

We then apply the Fourier transform to (17), and get

$$\widehat{(S_m f)}(\xi) = \sum_{j \in \Omega_m^d} \sum_{k \in \mathbb{Z}^d} \left(f * K_{j/m}^{\vee} \right)(k) \; e^{-\mathrm{i} 2\pi k \cdot \xi} \widehat{K_{j/m}}(\xi),$$

where $K_{j/m}^{\vee}(x) = K_{j/m}(-x)$.

Notice $\sum_{k \in \mathbb{Z}^d} \left(f * K_{j/m}^{\vee} \right)(k) \; e^{-\mathrm{i}2\pi k \cdot \xi}$ is the Fourier series of the sequence whose terms are samples of the function $f * K_{j/m}^{\vee}$ on the integer lattice. Thus, by (10) and properties (iii) and (iv) of the Fourier transform, we have

$$\widehat{(S_m f)}(\xi) = \sum_{j \in \Omega_m^d} \left(\sum_{k \in \mathbb{Z}^d} \widehat{f}(\xi + k) \overline{\widehat{K_{j/m}}(\xi + k)} \right) \widehat{K_{j/m}}(\xi).$$

For any $f = \sum_{l \in \mathbb{Z}^d} c(l)\phi(\cdot - l)$ in $V^2(\phi)$, we can use the fact that convolution becomes multiplication in the Fourier domain to express $\widehat{f}(\xi) = \widehat{c}(\xi)\widehat{\phi}(\xi)$. Thus we can use (8) and (15) to show $\widehat{K_{j/m}}(\xi) = Z\phi(j/m, -\xi)\widehat{\phi}(\xi)$. Then we can write

$$\begin{aligned} \widehat{(S_m f)}(\xi) &= \sum_{j \in \Omega_m^d} \widehat{c}(\xi) \left(\sum_{k \in \mathbb{Z}^d} \left| \widehat{\phi}(\xi + k) \right|^2 \right) \\ &\qquad \times \left| Z\phi(j/m, -\xi) \right|^2 \widehat{\phi}(\xi) \\ &= \left(\sum_{j \in \Omega_m^d} \left| Z\phi(j/m, -\xi) \right|^2 \right) \widehat{f}(\xi) \qquad \text{a.e. } \xi. \end{aligned}$$

Thus, for any $f \in V^2(\phi)$, we have

$$\widehat{(S_m^{-1} f)}(\xi) = \left(\sum_{j \in \Omega_m^d} \left| Z\phi(j/m, -\xi) \right|^2 \right)^{-1} \widehat{f}(\xi) \qquad (18)$$

Specifically, for fixed $j \in \Omega_m^d$,

$$\begin{aligned} (\widehat{S_m^{-1} K_{j/m}})(\xi) \;&=\; \Big(\sum_{j' \in \Omega_m^d} \left| Z\phi(j'/m, -\xi) \right|^2 \Big)^{-1} \\ &\qquad \times Z\phi(j/m, -\xi)\widehat{\phi}(\xi). \end{aligned} \qquad (19)$$

Using (18) and the fact that translation corresponds to modulation in the Fourier domain, it can easily be verified that $\widetilde{K}_{k+j/m} = S_m^{-1} K_{k+j/m} = (S_m^{-1} K_{j/m})(\cdot - k) = \widetilde{K}_{j/m}(\cdot - k)$.

Remark 5: Using equation (13), one can verify that $0 < \alpha_m \leq \sum_{j \in \Omega_m^d} \left| Z\phi(j/m, -\xi) \right|^2$ for all ξ, and hence that the formulas (18) and (19) are well defined. In the proof of Theorem 3, we will prove the stronger result that when m is large, there is a positive lower bound for $\frac{1}{m^d} \sum_{j \in \Omega_m^d} \left| Z\phi(j/m, -\xi) \right|^2$ that does not depend on m.

Given data $\{y_{k+j/m} = f(k + \frac{1}{m}j) + \varepsilon_{k+j/m}\}$, we define

$$\begin{aligned} f_{\varepsilon,m} &:= \sum_{j \in \Omega_m^d} \sum_{k \in \mathbb{Z}^d} y_{k+j/m} (S_m^{-1} K_{k+j/m}) \\ &= \sum_{j \in \Omega_m^d} \sum_{k \in \mathbb{Z}^d} y_{k+j/m} (\widetilde{K}_{j/m})(\cdot - k). \end{aligned}$$

We assume that the error $\{\varepsilon_{k+j/m}\}$ is a collection of i.i.d. random variables with mean zero and variance σ^2. A simple calculation shows that

$$E(f_{\varepsilon,m}(x) - f(x)) = \sum_{j \in \Omega_m^d} \sum_{k \in \mathbb{Z}^d} E(\varepsilon_{k+j/m})(\widetilde{K}_{j/m})(x - k) = 0.$$

We can compute $\operatorname{var}(f_{\varepsilon,m}(x) - f(x))$.

$$
\begin{aligned}
&\operatorname{var}(f_{\varepsilon,m}(x)-f(x)) \\
&= \operatorname{var}\left(\sum_{j\in\Omega_m^d}\sum_{k\in\mathbb{Z}^d}\varepsilon_{k+j/m}\widetilde{K}_{j/m}(x-k)\right) \\
&= \sigma^2\sum_{j\in\Omega_m^d}\sum_{k\in\mathbb{Z}^d}\left|S_m^{-1}K_{j/m}(x-k)\right|^2 \\
&= \sigma^2\sum_{j\in\Omega_m^d}\int_{[0,1]^d}\left|e^{-\mathbf{i}2\pi x\cdot\xi}\sum_{k\in\mathbb{Z}^d}e^{\mathbf{i}2\pi x\cdot k}\widehat{S_m^{-1}K_{j/m}}(k-\xi)\right|^2 d\xi \\
&= \sigma^2\sum_{j\in\Omega_m^d}\int_{[0,1]^d}\left|\sum_{k\in\mathbb{Z}^d}e^{\mathbf{i}2\pi x\cdot k}\widehat{S_m^{-1}K_{j/m}}(k-\xi)\right|^2 d\xi \\
&= \sigma^2\sum_{j\in\Omega_m^d}\int_{[0,1]^d}\left|\sum_{k\in\mathbb{Z}^d}e^{\mathbf{i}2\pi x\cdot k}\frac{Z\phi(j/m,\xi)\widehat{\phi}(k-\xi)}{\sum_{j'\in\Omega_m^d}|Z\phi(j'/m,\xi)|^2}\right|^2 d\xi \\
&= \sigma^2\int_{[0,1]^d}\frac{\sum_{j\in\Omega_m^d}|Z\phi(j/m,\xi)|^2\left|Z\widehat{\phi}(-\xi,-x)\right|^2}{\left|\sum_{j'\in\Omega_m^d}|Z\phi(j'/m,\xi)|^2\right|^2} d\xi \\
&= \frac{\sigma^2}{m^d}\int_{[0,1]^d}\frac{\left|Z\widehat{\phi}(-\xi,-x)\right|^2}{\frac{1}{m^d}\sum_{j\in\Omega_m^d}|Z\phi(j/m,\xi)|^2} d\xi \\
&= \frac{\sigma^2}{m^d}C_x(m).
\end{aligned}
$$

Focusing on the denominator in the above calculation, we notice

$$\frac{1}{m^d}\sum_{j\in\Omega_m^d}|Z\phi(j/m,\xi)|^2 \quad\stackrel{m\to\infty}{\longrightarrow}\quad \int_{[0,1]^d}|Z\phi(t,\xi)|^2\ dt$$

for all $\xi\in[0,1]^d$. By further investigating this denominator, we find that

Lemma 6: For every $\xi\in[0,1]^d$, $\int_{[0,1]^d}|Z\phi(t,\xi)|^2\ dt=1$.

Proof:

$$
\begin{aligned}
\int_{[0,1]^d}|Z\phi(t,\xi)|^2\ dt &\stackrel{(12)}{=}\int_{[0,1]^d}\left|Z\widehat{\phi}(\xi,-t)\right|^2\ dt \\
&= \int_{[0,1]^d}\left|\sum_{k\in\mathbb{Z}^d}\widehat{\phi}(\xi+k)e^{\mathbf{i}2\pi t\cdot k}\right|^2\ dt \\
&= \sum_{k\in\mathbb{Z}^d}\left|\widehat{\phi}(\xi+k)\right|^2 \stackrel{(2)}{=} 1 \quad \text{a.e. } \xi
\end{aligned}
$$

Because $Z\phi$ is continuous, $\int_{[0,1]^d}|Z\phi(t,\xi)|^2\ dt$ is a continuous function of ξ. Therefore $\int_{[0,1]^d}|Z\phi(t,\xi)|^2\ dt=1$ for every $\xi\in[0,1]^d$.

□

Now, for each positive integer m, define the function

$$g_m(\xi):=\frac{1}{m^d}\sum_{j\in\Omega_m^d}|Z\phi(j/m,\xi)|^2 \qquad \xi\in[0,1]^d.$$

Lemma 6 tells us that $g_m(\xi)\to 1$ pointwise. In fact, it will be shown in the proof of Theorem 3 that g_m converges uniformly to the constant function 1 on the unit cube $[0,1]^d$.

Therefore, for any $\epsilon>0$, there exists a number $M\in\mathbb{N}$ such that for all $m\geq M$, sampling on the lattice $\frac{1}{m}\mathbb{Z}^d$ gives the estimate

$$\operatorname{var}(f_{\varepsilon,m}(x)-f(x))\leq\frac{(1+\epsilon)\sigma^2}{m^d}\int_{[0,1]^d}\left|Z\widehat{\phi}(-\xi,-x)\right|^2\ d\xi. \tag{20}$$

Using the argument from the above proof of Lemma 6, we can see that, equivalently, for large enough m, we have

$$\operatorname{var}(f_{\varepsilon,m}(x)-f(x))\leq\frac{(1+\epsilon)\sigma^2}{m^d}\sum_{k\in\mathbb{Z}^d}|\phi(x+k)|^2$$

for every $x\in\mathbb{R}^d$. In other words, we obtain the slightly suboptimal estimate that depends on x but does not depend on m for large m. Also notice that $\operatorname{var}(f_{\varepsilon,m}(x)-f(x))$ is periodic with period 1. Then for any $x\in[0,1]^d$,

$$\sum_{k\in\mathbb{Z}^d}|\phi(x+k)|^2\leq\left(\sum_{k\in\mathbb{Z}^d}\sup_{x\in[0,1]^d}|\phi(x+k)|\right)^2=\|\phi\|_{W^1}^2, \tag{21}$$

giving the coarser estimate from Remark 4 that does not depend on x or m.

IV. AVERAGE SAMPLING

Here we assume our data are of the form

$$\Big\{\left\langle f,\psi^i(\cdot-(k+j/m))\right\rangle+\varepsilon^i_{k+j/m}: \\ k\in\mathbb{Z}^d,\ j\in\Omega_m^d, 1\leq i\leq s\Big\} \tag{22}$$

for some $f\in V^2(\Phi)$ and some real-valued vector function $\Psi=(\psi^1,\ldots,\psi^s)^T$, where $\Psi\in\left[L^1(\mathbb{R}^d)\cap L^2(\mathbb{R}^d)\right]^{(s)}$. We use the notation $\psi^i_{k+j/m}$ to denote $\psi^i(\cdot-(k+j/m))$. We continue to assume $\Phi\in\left(L^2(\mathbb{R}^d)\right)^{(r)}$ satisfies (2) and that $\Phi\in\left(W_0^1\right)^{(r)}$.

In order to recover a function f in $V^2(\Phi)$ from its weighted averages using shifts of the functions ψ^i, Ψ must satisfy certain conditions. We require that the Grammian

$$G_\Psi(\xi):=\sum_{k\in\mathbb{Z}^d}\widehat{\Psi}(\xi+k)\overline{\widehat{\Psi}(\xi+k)}^T$$

be bounded, i.e. there exists a number η such that $G_\Psi(\xi)\leq\eta I$, a.e. $\xi\in\mathbb{R}^d$ [5]. Furthermore, we assume Ψ is such that, for each $m\in\mathbb{N}$, there exist positive constants α_m and β_m satisfying

$$\alpha_m\|f\|_2^2\leq\sum_{i=1}^s\sum_{j\in\Omega_m^d}\sum_{k\in\mathbb{Z}^d}\left|\left\langle f,\psi^i_{k+j/m}\right\rangle\right|^2\leq\beta_m\|f\|_2^2 \tag{23}$$

for all f in $V^2(\Phi)$. Finally, we also assume

$$\lim_{N\to\infty}\sup_{\xi\in[0,1]^d}\sum_{i=1}^s\sum_{|k|\geq N}\left|\widehat{\psi^i}(\xi+k)\right|^2=0. \tag{24}$$

Condition (24) comes from [5] and serves to exclude pathological examples. Because condition (23) is satisfied, $f\in V^2(\Phi)$

is uniquely determined by, and can be stably reconstructed from, the collection

$$\{\langle f, \psi^i(\cdot - (k + j/m))\rangle : k \in \mathbb{Z}^d,\ j \in \Omega_m^d, 1 \le i \le s\}.$$

Recall that ψ^i is not necessarily in $V^2(\Phi)$, so although (23) is satisfied, the collection $\{\psi^i_{k+j/m}\}$ does not constitute a frame for $V^2(\Phi)$. As in [1], consider the orthogonal projection P from $L^2(\mathbb{R}^d)$ onto $V^2(\Phi)$, and define

$$\theta^i_{k+j/m} := P\psi^i_{k+j/m}.$$

Then for all $f \in V^2(\Phi)$,

$$\left\langle f, \theta^i_{k+j/m}\right\rangle = \left\langle f, P\psi^i_{k+j/m}\right\rangle = \left\langle Pf, \psi^i_{k+j/m}\right\rangle = \left\langle f, \psi^i_{k+j/m}\right\rangle.$$

Thus condition (23) implies that $\{\theta^i_{k+j/m} : k \in \mathbb{Z}^d, j \in \Omega_m^d, 1 \le i \le s\}$ forms a frame for $V^2(\Phi)$. Furthermore, using the orthonormality of $\{\phi^l(\cdot - k)\}$, we can write

$$\begin{aligned}\theta^i_{k+j/m}(x) &= \sum_{l=1}^{r}\sum_{n\in\mathbb{Z}^d}\left\langle \theta^i_{k+j/m}, \phi^l(\cdot - n)\right\rangle \phi^l(x-n)\\ &= \sum_{l=1}^{r}\sum_{n\in\mathbb{Z}^d}\left\langle \psi^i_{k+j/m}, \phi^l(\cdot - n)\right\rangle \phi^l(x-n),\end{aligned}$$

and we see that $\theta^i_{k+j/m} = \theta^i_{j/m}(\cdot - k)$. There exists a dual frame $\{\widetilde{\theta}^i_{k+j/m} : k \in \mathbb{Z}^d, j \in \Omega_m^d 1 \le i \le s\}$, defined by

$$\widetilde{\theta}^i_{k+j/m} := S_m^{-1}\theta^i_{k+j/m},$$

where S_m is the frame operator on $V^2(\Phi)$ corresponding to the frame $\{\theta^i_{k+j/m}\}$, i.e.

$$S_m f = \sum_{i=1}^{s}\sum_{j\in\Omega_m^d}\sum_{k\in\mathbb{Z}^d}\left\langle f, \theta^i_{k+j/m}\right\rangle \theta^i_{k+j/m}. \tag{25}$$

Then for any scalar-valued sequence $\{a^i_{k+j/m} : k \in \mathbb{Z}^d, j \in \Omega_m^d, 1 \le i \le s\}$ satisfying

$$\sum_{i=1}^{s}\sum_{j\in\Omega_m^d}\sum_{k\in\mathbb{Z}^d}\left|a^i_{k+j/m}\right|^2 < \infty,$$

the function defined by

$$\sum_{i=1}^{s}\sum_{j\in\Omega_m^d}\sum_{k\in\mathbb{Z}^d} a^i_{k+j/m}\ \widetilde{\theta}^i_{k+j/m}$$

is in $V^2(\Phi)$ [15]. Furthermore, we have the following reconstruction formula for any function $f \in V^2(\Phi)$:

$$f = \sum_{i=1}^{s}\sum_{j\in\Omega_m^d}\sum_{k\in\mathbb{Z}^d}\left\langle f, \psi^i_{k+j/m}\right\rangle \widetilde{\theta}^i_{k+j/m}. \tag{26}$$

Given data

$$\left\{y^i_{k+j/m} = \left\langle f, \psi^i_{k+j/m}\right\rangle + \varepsilon^i_{k+j/m}\right\}, \tag{27}$$

we define

$$f_{\varepsilon,m} := \sum_{i=1}^{s}\sum_{j\in\Omega_m^d}\sum_{k\in\mathbb{Z}^d} y^i_{k+j/m}\ \widetilde{\theta}^i_{k+j/m}. \tag{28}$$

In the case of average sampling, we arrive at results for $\mathrm{var}(f_{\varepsilon,m}(x) - f(x))$ similar to those of Theorem 3. For $\xi \in [0,1]^d$, define the self-adjoint matrix

$$G^{\Psi}_{\Phi}(\xi) := \sum_{i=1}^{s}\sum_{k\in\mathbb{Z}^d}\widehat{\Phi}(\xi+k)\overline{\widehat{\Phi}(\xi+k)}^T\left|\widehat{\psi^i}(\xi+k)\right|^2.$$

As in Theorem 3, the expected value and variance of the error between the frame reconstruction $f_{\varepsilon,m}$ and the exact function f is a function of the position x, the oversampling factor m^d, and the noise variance σ^2. The precise estimates and best constants are given by the following theorem.

Theorem 7: Let $\Phi = (\phi^1, \ldots, \phi^r)^T$ satisfy $G_\Phi(\xi) = I$ a.e. ξ, and $\phi^i \in W^1 \cap C^0$, $1 \le i \le r$. Assume $G_\Psi(\xi) \le \eta I$, a.e. $\xi \in \mathbb{R}^d$ and also that equations (23) and (24) are satisfied. Assume, for all $k \in \mathbb{Z}^d$, $j \in \Omega_m^d$, and $1 \le i \le s$, the data $\{y^i_{k+j/m}\}$ are of the form (27) for some $f \in V^2(\Phi)$, where $\{\varepsilon^i_{k+j/m}\}$ is a collection of i.i.d. random variables satisfying $E(\varepsilon^i_{k+j/m}) = 0$, $\mathrm{var}(\varepsilon^i_{k+j/m}) = \sigma^2$, and $\varepsilon^i_{k+j/m} \in [-N, N]$ for some $N < \infty$. Then $E(f_{\varepsilon,m}(x) - f(x)) = 0$, and

$$\mathrm{var}(f_{\varepsilon,m}(x) - f(x)) = \frac{\sigma^2}{m^d}D_x(m),$$

where $D_x(m)$ is given by (39), and

$$D_x(m) \overset{m\to\infty}{\longrightarrow} \int_{[0,1]^d}\overline{Z\widehat{\Phi}(-\xi,-x)}^T\left(G^{\Psi}_{\Phi}(\xi)\right)^{-1}Z\widehat{\Phi}(-\xi,-x)\,d\xi.$$

Remark 8: In [5] it is shown that (23) and (24) imply that there exists a positive number δ_0 such that $\delta_0 I \le G^{\Psi}_{\Phi}(\xi)$ for all ξ. It follows from (21) that there exists a number $\delta > 0$ and a number $M \in \mathbb{N}$ such that for every $m \ge M$, we obtain the suboptimal but uniform estimate:

$$\mathrm{var}(f_{\varepsilon,m}(x) - f(x)) \le \frac{\sigma^2}{m^d}\left(\frac{1}{\delta}\right)\left(\sum_{n=1}^{r}\|\phi^n\|^2_{W^1}\right)$$

for all $x \in \mathbb{R}^d$.

A. Average Sampling in $V^2(\phi)$

Once again, before presenting the proof of the theorem above, we will lay the groundwork for that proof by illustrating the simpler case where $r = 1$. In other words, our underlying shift-invariant space has only one generator, ϕ.

As we did in the example in the previous section, in this uniform case, we can find the Fourier transform of $\widetilde{\theta}^i_{k+j/m} = S_m^{-1}\theta^i_{k+j/m}$ explicitly.

Let S_m be the frame operator on $V^2(\phi)$ associated to the frame $\left\{\theta^i_{k+j/m} : k \in \mathbb{Z}^d, j \in \Omega_m^d, 1 \le i \le s\right\}$. Recall that

$$(S_m f)(x) = \sum_{i=1}^{s}\sum_{j\in\Omega_m^d}\sum_{k\in\mathbb{Z}^d}\left\langle f, \psi^i_{k+j/m}\right\rangle \theta^i_{k+j/m}(x), \tag{29}$$

$$\theta^i_{k+j/m}(x) = \sum_{l\in\mathbb{Z}^d}\left\langle \psi^i_{k+j/m}, \phi(\cdot - l)\right\rangle \phi(x-l), \tag{30}$$

and also that $\theta^i_{k+j/m} = \theta^i_{j/m}(\cdot - k)$. For any $f \in V^2(\phi)$, we apply the Fourier transform to (29) and rewrite the inner product as convolution to get

$$\widehat{(S_m f)}(\xi) = \sum_{i=1}^{s}\sum_{j\in\Omega_m^d}\sum_{k\in\mathbb{Z}^d}\left(f * (\psi^i_{j/m})^{\vee}\right)(k)\ e^{-\mathbf{i}2\pi k\cdot\xi}\widehat{\theta^i_{j/m}}(\xi),$$

where $(\psi^i_{j/m})^\vee(x) = \psi^i_{j/m}(-x)$.

Notice $\sum_{k\in\mathbb{Z}^d}\left(f*(\psi^i_{j/m})^\vee\right)(k)\ e^{-\mathbf{i}2\pi k\cdot\xi}$ is the Fourier series of the sequence whose terms are samples of the function $f*(\psi^i_{j/m})^\vee$ on the integer lattice. Thus, by (10) and properties (iii) and (iv) of the Fourier transform, we have

$$\widehat{(S_m f)}(\xi) = \sum_{i=1}^{s}\sum_{j\in\Omega^d_m}\left(\sum_{k\in\mathbb{Z}^d}\widehat{f}(\xi+k)\overline{\widehat{\psi^i_{j/m}}(\xi+k)}\right)\widehat{\theta^i_{j/m}}(\xi).$$

Similarly, we can use (30) to show that

$$\widehat{\theta^i_{j/m}}(\xi) = \left(\sum_{l\in\mathbb{Z}^d}\widehat{\psi^i_{j/m}}(\xi+l)\overline{\widehat{\phi}(\xi+l)}\right)\widehat{\phi}(\xi).$$

Notice that

$$\begin{aligned}
&\sum_{l\in\mathbb{Z}^d}\overline{\widehat{\phi}(\xi+l)}\widehat{\psi^i_{j/m}}(\xi+l)\\
&\overset{(10)}{=}\sum_{l\in\mathbb{Z}^d}(\phi^\vee*\psi^i_{j/m})(l)e^{-\mathbf{i}2\pi l\cdot\xi}\\
&=\sum_{l\in\mathbb{Z}^d}(\phi^\vee*\psi^i)(l-j/m)e^{-\mathbf{i}2\pi l\cdot\xi}\\
&=\sum_{l\in\mathbb{Z}^d}(\phi*\psi^{i^\vee})(j/m-l)e^{-\mathbf{i}2\pi l\cdot\xi}\\
&=Z(\phi*\psi^{i^\vee})(j/m,-\xi).
\end{aligned}$$

Thus for any $f=\sum_{l\in\mathbb{Z}^d}c(l)\phi(\cdot-l)$ in $V^2(\phi)$, we have

$$\begin{aligned}
\widehat{(S_m f)}(\xi) &= \sum_{i=1}^{s}\sum_{j\in\Omega^d_m}\widehat{c}(\xi)\left|Z(\phi*\psi^{i^\vee})(j/m,-\xi)\right|^2\widehat{\phi}(\xi)\\
&=\left(\sum_{i=1}^{s}\sum_{j\in\Omega^d_m}\left|Z(\phi*\psi^{i^\vee})(j/m,-\xi)\right|^2\right)\widehat{f}(\xi),
\end{aligned}$$

and therefore

$$\widehat{(S_m^{-1}f)}(\xi) = \frac{\widehat{f}(\xi)}{\sum_{i=1}^{s}\sum_{j\in\Omega^d_m}\left|Z(\phi*\psi^{i^\vee})(j/m,-\xi)\right|^2} \quad (31)$$

provided that the denominator is nonzero. Then for fixed i and j,

$$\widehat{(S_m^{-1}\theta^i_{j/m})}(\xi) = \frac{Z(\phi*\psi^{i^\vee})(j/m,-\xi)\widehat{\phi}(\xi)}{\sum_{i'=1}^{s}\sum_{j'\in\Omega^d_m}\left|Z(\phi*\psi^{i'^\vee})(j'/m,-\xi)\right|^2}.$$

Using (31) and property (i) of the Fourier transform, it can be verified that $S_m^{-1}\theta^i_{k+j/m} = \left(S_m^{-1}\theta^i_{j/m}\right)(\cdot-k)$. Now we can use (26) and (28) to begin computing $\mathrm{var}(f_{\varepsilon,m}(x)-f(x))$.

$$\begin{aligned}
&\mathrm{var}(f_{\varepsilon,m}(x)-f(x))\\
&=\mathrm{var}\left(\sum_{i=1}^{s}\sum_{j\in\Omega^d_m}\sum_{k\in\mathbb{Z}^d}\varepsilon^i_{k+j/m}\widetilde{\theta}^i_{j/m}(x-k)\right)\\
&=\sigma^2\sum_{i=1}^{s}\sum_{j\in\Omega^d_m}\sum_{k\in\mathbb{Z}^d}\left|S_m^{-1}\theta^i_{j/m}(x-k)\right|^2\\
&=\sigma^2\sum_{i=1}^{s}\sum_{j\in\Omega^d_m}\int_{[0,1]^d}\left|e^{\mathbf{i}2\pi x\cdot\xi}\sum_{k\in\mathbb{Z}^d}e^{\mathbf{i}2\pi x\cdot k}\widehat{S_m^{-1}\theta^i_{j/m}}(\xi+k)\right|^2\ d\xi\\
&=\sigma^2\sum_{i=1}^{s}\sum_{j\in\Omega^d_m}\int_{[0,1]^d}\left|\sum_{k\in\mathbb{Z}^d}e^{\mathbf{i}2\pi x\cdot k}\widehat{S_m^{-1}\theta^i_{j/m}}(\xi+k)\right|^2\ d\xi\\
&=\frac{\sigma^2}{m^d}\int_{[0,1]^d}\frac{\left|Z\widehat{\phi}(\xi,-x)\right|^2}{\sum_{i=1}^{s}\frac{1}{m^d}\sum_{j\in\Omega^d_m}\left|Z(\phi*\psi^{i^\vee})(j/m,-\xi)\right|^2}\ d\xi\\
&=\frac{\sigma^2}{m^d}D_x(m).
\end{aligned}$$

Focusing on the denominator, we can see that

$$\begin{aligned}
&\frac{1}{m^d}\sum_{j\in\Omega^d_m}\left|Z(\phi*\psi^{i^\vee})(j/m,-\xi)\right|^2\\
&\overset{m\to\infty}{\longrightarrow}\int_{[0,1]^d}\left|Z(\phi*\psi^{i^\vee})(t,-\xi)\right|^2\ dt
\end{aligned}$$

for each ξ. In the proof of Theorem 7, we will see that this convergence is uniform on $[0,1]^d$. Further analysis of this denominator gives us:

Lemma 9: For every $\xi\in[0,1]^d$,

$$\sum_{i=1}^{s}\int_{[0,1]^d}\left|Z(\phi*\psi^{i^\vee})(t,-\xi)\right|^2\ dt\geq\delta>0.$$

Proof:

$$\begin{aligned}
&\sum_{i=1}^{s}\int_{[0,1]^d}\left|Z(\phi*\psi^{i^\vee})(t,-\xi)\right|^2\ dt\\
&\overset{(12)}{=}\sum_{i=1}^{s}\int_{[0,1]^d}\left|Z(\widehat{\phi*\psi^{i^\vee}})(-\xi,-t)\right|^2\ dt\\
&=\sum_{i=1}^{s}\int_{[0,1]^d}\left|\sum_{l\in\mathbb{Z}^d}\widehat{\phi}(l-\xi)\overline{\widehat{\psi^i}(l-\xi)}e^{\mathbf{i}2\pi t\cdot l}\right|^2\ dt\\
&=\sum_{i=1}^{s}\sum_{l\in\mathbb{Z}^d}\left|\widehat{\phi}(l-\xi)\overline{\widehat{\psi^i}(l-\xi)}\right|^2\\
&=\sum_{i=1}^{s}\sum_{l\in\mathbb{Z}^d}\left|\widehat{\phi}(l-\xi)\right|^2\left|\widehat{\psi^i}(l-\xi)\right|^2.
\end{aligned}$$

In [5] it is shown that (23) and (24) imply that there exists a positive number δ such that $\delta\leq\sum_{i=1}^{s}\sum_{l\in\mathbb{Z}^d}\left|\widehat{\phi}(l-\xi)\right|^2\left|\widehat{\psi^i}(l-\xi)\right|^2$ for all ξ.

□

Therefore, using this lemma and (21), we see that there exists a number $M\in\mathbb{N}$ such that for all $m\geq M$, average sampling on $\frac{1}{m}\mathbb{Z}^d$ gives

$$\mathrm{var}(f_{\varepsilon,m}(x)-f(x)) \le \frac{\sigma^2}{m^d}\left(\frac{2}{\delta}\right)\|\phi\|^2_{W^1} \qquad \text{for all } x\in\mathbb{R}^d.$$

V. Proofs

A. Proof of Theorem 3

We wish to compute the variance of the error as in Section III-A. First we must find $S_m^{-1}K_{j/m}$ explicitly. In section III-A we showed that

$$\widehat{(S_m f)}(\xi) = \sum_{j\in\Omega_m^d}\left(\sum_{k\in\mathbb{Z}^d}\widehat{f}(\xi+k)\overline{\widehat{K_{j/m}}(\xi+k)}\right)\widehat{K_{j/m}}(\xi).$$

For any $f=\sum_{l\in\mathbb{Z}^d}C(l)^T\Phi(\cdot-l)$ in $V^2(\Phi)$, we then get

$$\widehat{(S_m f)}(\xi) = \widehat{C}(\xi)^T\left(\sum_{j\in\Omega_m^d}\overline{Z\Phi(j/m,-\xi)}Z\Phi(j/m,-\xi)^T\right)\widehat{\Phi}(\xi)$$

a.e. ξ. Notice in the equation above that $\left(\sum_{j\in\Omega_m^d}\overline{Z\Phi(j/m,-\xi)}Z\Phi(j/m,-\xi)^T\right)$ is a self-adjoint $r\times r$ matrix. Define the matrix

$$A_m(\xi) := \sum_{j\in\Omega_m^d}\overline{Z\Phi(j/m,-\xi)}Z\Phi(j/m,-\xi)^T.$$

Remark 10: It can be shown that $\alpha_m I\le A_m(\xi)$ for all ξ, and hence the matrix $A_m(\xi)$ is invertible. Instead, for large m we provide a stronger result in Lemma 11 below. Still, it should be noted that the following formulas (32) and (33) make sense as long as (13) holds.

Therefore, we have

$$\widehat{(S_m^{-1} f)}(\xi) = \widehat{C}(\xi)^T\left(A_m(\xi)\right)^{-1}\widehat{\Phi}(\xi). \tag{32}$$

Finally, using (15) and (32), for any fixed $j\in\Omega_m^d$ we have

$$(\widehat{S_m^{-1}K_{j/m}})(\xi) = Z\Phi(j/m,-\xi)^T\left(A_m(\xi)\right)^{-1}\widehat{\Phi}(\xi). \tag{33}$$

Using (32) and the fact that translation corresponds to modulation in the Fourier domain, it can easily be verified that $S_m^{-1}K_{k+j/m}=(S_m^{-1}K_{j/m})(\cdot-k)=\widetilde{K}_{j/m}(\cdot-k)$.

We are now ready to compute the expected value and the variance of the error $(f_{\varepsilon,m}(x)-f(x))$.

$$E(f_{\varepsilon,m}(x)-f(x)) = \sum_{j\in\Omega_m^d}\sum_{k\in\mathbb{Z}^d}E(\varepsilon_{k+j/m})\widetilde{K}_{j/m}(x-k)=0$$

and

$$\begin{aligned}
&\mathrm{var}(f_{\varepsilon,m}(x)-f(x))\\
&=\mathrm{var}\left(\sum_{j\in\Omega_m^d}\sum_{k\in\mathbb{Z}^d}\varepsilon_{k+j/m}\widetilde{K}_{j/m}(x-k)\right)\\
&=\sigma^2\sum_{j\in\Omega_m^d}\sum_{k\in\mathbb{Z}^d}\left|S_m^{-1}K_{j/m}(x-k)\right|^2\\
&=\sigma^2\sum_{j\in\Omega_m^d}\int_{[0,1]^d}\left|e^{-\mathbf{i}2\pi x\cdot\xi}\sum_{k\in\mathbb{Z}^d}\widehat{S_m^{-1}K_{j/m}}(k-\xi)e^{\mathbf{i}2\pi k\cdot x}\right|^2 d\xi\\
&=\sigma^2\sum_{j\in\Omega_m^d}\int_{[0,1]^d}\left|\sum_{k\in\mathbb{Z}^d}\widehat{S_m^{-1}K_{j/m}}(k-\xi)e^{\mathbf{i}2\pi k\cdot x}\right|^2 d\xi\\
&=\sigma^2\int_{[0,1]^d}\sum_{j\in\Omega_m^d}\left|Z\Phi(j/m,\xi)^T\left(A_m(\xi)\right)^{-1}\left(Z\widehat{\Phi}(-\xi,-x)\right)\right|^2 d\xi.
\end{aligned}$$

The matrix $(A_m(\xi))^{-1}$ is self-adjoint because it is the inverse of a self-adjoint matrix. Next we use the fact that $a^TAb=b^T\overline{A}a$ for any vectors a and b and any self-adjoint matrix A, and hence

$$\left|a^TAb\right|^2=\overline{a^TAb}a^TAb=\overline{b}^TA\overline{a}a^TAb. \tag{34}$$

If $\left(\overline{a}a^T\right)^{-1}=A$, then we have $\overline{b}^TAb$. Now we have

$$\begin{aligned}
&\mathrm{var}(f_{\varepsilon,m}(x)-f(x))\\
&=\sigma^2\int_{[0,1]^d}\left(\overline{Z\widehat{\Phi}(-\xi,-x)}\right)^T\left(A_m(\xi)\right)^{-1}\left(Z\widehat{\Phi}(-\xi,-x)\right)\,d\xi\\
&=\frac{\sigma^2}{m^d}\int_{[0,1]^d}\left(\overline{Z\widehat{\Phi}(-\xi,-x)}\right)^T\left(\frac{1}{m^d}A_m(\xi)\right)^{-1}\left(Z\widehat{\Phi}(-\xi,-x)\right)\,d\xi\\
&=\frac{\sigma^2}{m^d}C_x(m)
\end{aligned} \tag{35}$$

Lemma 11: For every $\epsilon>0$ there is a number $M\in\mathbb{N}$ such that for every $m\ge M$

$$(1-\epsilon)I\le\frac{1}{m^d}A_m(\xi) \qquad \text{for all } \ \xi\in[0,1]^d.$$

We prove this lemma in Section V-B.

Using Lemma 11, we conclude that there is a number $M\in\mathbb{N}$ such that for all $m\ge M$, sampling on the set $\frac{1}{m}\mathbb{Z}^d$ gives

$$\begin{aligned}
&\mathrm{var}(f_\epsilon(x)-f(x))\\
&\le\frac{(1+\epsilon)\sigma^2}{m^d}\int_{[0,1]^d}\left(\overline{Z\widehat{\Phi}(-\xi,-x)}\right)^T\left(Z\widehat{\Phi}(-\xi,-x)\right)\,d\xi.\\
&=\frac{(1+\epsilon)\sigma^2}{m^d}\sum_{i=1}^r\int_{[0,1]^d}\left|Z\widehat{\phi^i}(-\xi,-x)\right|^2\,d\xi.
\end{aligned}$$

In section III-A, we saw that

$$\begin{aligned}
&\int_{[0,1]^d}\left|Z\widehat{\phi^i}(-\xi,-x)\right|^2\,d\xi\\
&\qquad=\sum_{k\in\mathbb{Z}^d}\left|\phi^i(x+k)\right|^2\le\left\|\phi^i\right\|^2_{W^1}
\end{aligned}$$

for all $x\in\mathbb{R}^d$. Thus when m is large enough,

$$\begin{aligned}&\mathrm{var}(f_{\varepsilon,m}(x)-f(x))\\&\le \frac{(1+\epsilon)\sigma^2}{m^d}\left(\sum_{i=1}^r\left(\sum_{k\in\mathbb{Z}^d}\left|\phi^i(x+k)\right|^2\right)\right)\\&\le \frac{(1+\epsilon)\sigma^2}{m^d}\left(\sum_{i=1}^r\left\|\phi^i\right\|_{W^1}^2\right)\qquad \text{for all } x\in\mathbb{R}^d.\end{aligned}$$

□

B. Proof of Lemma 11

Notice that, for $1 \le n,n' \le r$, the (n,n')-entry of $\frac{1}{m^d}A_m(\xi)$ is

$$\begin{aligned}&\left[\frac{1}{m^d}A_m(\xi)\right]_{(n,n')}\\&=\frac{1}{m^d}\sum_{j\in\Omega_m^d}(Z\phi^n(j/m,\xi))\left(\overline{Z\phi^{n'}(j/m,\xi)}\right)\\&=\frac{1}{m^d}\sum_{j\in\Omega_m^d}\left(\sum_{l\in\mathbb{Z}^d}\sum_{k\in\mathbb{Z}^d}\phi^n(j/m-l)\phi^{n'}(j/m-l-k)e^{-\mathbf{i}2\pi k\cdot\xi}\right)\\&=\sum_{k\in\mathbb{Z}^d}e^{-\mathbf{i}2\pi k\cdot\xi}\frac{1}{m^d}\sum_{j\in\Omega_m^d}\left(\sum_{l\in\mathbb{Z}^d}\phi^n(j/m-l)\phi^{n'}(j/m-l-k)\right).\end{aligned}$$

Taking the limit as m goes to infinity, we have

$$\begin{aligned}&\lim_{m\to\infty}\left[\frac{1}{m^d}A_m(\xi)\right]_{(n,n')}\\&=\sum_{k\in\mathbb{Z}^d}e^{-\mathbf{i}2\pi k\cdot\xi}\int_{\mathbb{R}^d}\phi^n(x)\phi^{n'}(x-k)\,dx=\delta_{n,n'}.\end{aligned}$$

Thus the diagonal entries of the matrix converge to 1 and the off-diagonal entries of the matrix converge to 0 for each ξ.

Now we will show the collection $\left\{\left[\frac{1}{m^d}A_m(\cdot)\right]_{(n,n')}:m\in\mathbb{N}\right\}$ is equicontinuous and conclude that convergence is uniform on the unit cube $[0,1]^d$. Recall that a collection $\mathcal{G}$ of continuous functions on $[0,1]^d$ is equicontinuous if for every $\epsilon>0$ there is a $\delta>0$ such that for all $g\in\mathcal{G}$, $|g(\xi_1)-g(\xi_2)|<\epsilon$ for all $\xi_1,\xi_2\in[0,1]^d$ satisfying $|\xi_1-\xi_2|<\delta$.

Let $1\le n,n'\le r$. Let $\epsilon>0$. There exists a number $N\in\mathbb{N}$ such that

$$\sum_{|l|>N}\sup_{x\in[0,1]^d}|\phi^n(x-l)|<\frac{\epsilon}{6\left\|\phi^{n'}\right\|_{W^1}}.$$

Then there exists a number $N'\in\mathbb{N}$ such that

$$\sum_{|k|>N'}\sup_{x\in[0,1]^d}\left|\phi^{n'}(x-l-k)\right|<\frac{\epsilon}{6\left\|\phi^n\right\|_{W^1}}$$

for all l such that $|l|\le N$. Then there exists a number $\delta>0$ such that whenever $|\xi_1-\xi_2|<\delta$,

$$\left|e^{-\mathbf{i}2\pi k\cdot\xi_1}-e^{-\mathbf{i}2\pi k\cdot\xi_2}\right|<\frac{\epsilon}{3\left\|\phi^n\right\|_{W^1}\left\|\phi^{n'}\right\|_{W^1}}$$

for every k such that $|k|\le N'$. Notice

$$\begin{aligned}&\left|\left[\frac{1}{m^d}A_m(\xi_1)\right]_{(n,n')}-\left[\frac{1}{m^d}A_m(\xi_2)\right]_{(n,n')}\right|\\=\ &\Big|\frac{1}{m^d}\sum_{j\in\Omega_m^d}\sum_{l\in\mathbb{Z}^d}\phi^n(j/m-l)\\&\times\sum_{k\in\mathbb{Z}^d}\phi^{n'}(j/m-l-k)\left(e^{-\mathbf{i}2\pi k\cdot\xi_1}-e^{-\mathbf{i}2\pi k\cdot\xi_2}\right)\Big|\\\le\ &\frac{1}{m^d}\sum_{j\in\Omega_m^d}\Big[\sum_{|l|\le N}|\phi^n(j/m-l)|\\&\times\Big(\sum_{|k|\le N'}\left|\phi^{n'}(j/m-l-k)\right|\left|e^{-\mathbf{i}2\pi k\cdot\xi_1}-e^{-\mathbf{i}2\pi k\cdot\xi_2}\right|\\&+\sum_{|k|>N'}\left|\phi^{n'}(j/m-l-k)\right|\left|e^{-\mathbf{i}2\pi k\cdot\xi_1}-e^{-\mathbf{i}2\pi k\cdot\xi_2}\right|\Big)\\&+\sum_{|l|>N}|\phi^n(j/m-l)|\sum_{k\in\mathbb{Z}^d}\left|\phi^{n'}(j/m-l-k)\right|\\&\times\left|e^{-\mathbf{i}2\pi k\cdot\xi_1}-e^{-\mathbf{i}2\pi k\cdot\xi_2}\right|\Big]\\<\ &\frac{1}{m^d}\sum_{j\in\Omega_m^d}\left(\frac{\epsilon}{3}+\frac{\epsilon}{3}+\frac{\epsilon}{3}\right)=\epsilon.\end{aligned}$$

Thus the collection $\left\{\left[\frac{1}{m^d}A_m(\cdot)\right]_{(n,n')}:m\in\mathbb{N}\right\}$ is equicontinuous, and hence for each pair (n,n'), $\left[\frac{1}{m^d}A_m(\cdot)\right]_{(n,n')}\to\delta_{n,n'}$ uniformly on $[0,1]^d$.

Therefore, for any $\epsilon>0$, there is a number $M\in\mathbb{N}$ such that for all $m\ge M$

$$\left\|\frac{1}{m^d}A_m(\xi)-I\right\|<\epsilon\qquad\text{for all }\ \xi\in[0,1]^d.$$

Hence our lemma is proved.

□

C. Proof of Theorem 7

Once again, our objective is to compute the expected value and the variance of $\big(f_{\varepsilon,m}(x)-f(x)\big)$, where in this case

$$f=\sum_{i=1}^s\sum_{j\in\Omega_m^d}\sum_{k\in\mathbb{Z}^d}\left\langle f,\psi^i_{k+j/m}\right\rangle\widetilde{\theta}^i_{k+j/m}$$

and

$$f_{\varepsilon,m}:=\sum_{i=1}^s\sum_{j\in\Omega_m^d}\sum_{k\in\mathbb{Z}^d}y^i_{k+j/m}\ \widetilde{\theta}^i_{k+j/m}.$$

A simple calculation shows

$$\begin{aligned}&E\big(f_{\varepsilon,m}(x)-f(x)\big)\\=\ &\sum_{i=1}^s\sum_{j\in\Omega_m^d}\sum_{k\in\mathbb{Z}^d}E\left(\varepsilon^i_{k+j/m}\right)\ \widetilde{\theta}^i_{k+j/m}(x)=0.\end{aligned}$$

To compute the variance, we first need to compute $\widetilde{\theta}^i_{k+j/m}=S_m^{-1}\theta^i_{k+j/m}$ explicitly. In section IV-A, we showed that

$$\widehat{(S_mf)}(\xi)=\sum_{i=1}^s\sum_{j\in\Omega_m^d}\left(\sum_{k\in\mathbb{Z}^d}\hat{f}(\xi+k)\overline{\widehat{\psi^i_{j/m}}(\xi+k)}\right)\widehat{\theta^i_{j/m}}(\xi),\tag{36}$$

and

$$\widehat{\theta^i_{j/m}}(\xi) = \left(\sum_{l\in\mathbb{Z}^d} \widehat{\psi^i_{j/m}}(\xi+l)\overline{\widehat{\Phi}(\xi+l)} \right)^T \widehat{\Phi}(\xi). \qquad (37)$$

Define the self-adjoint matrix

$$[A_m]^\Psi_\Phi(\xi) := \sum_{i=1}^s \sum_{j\in\Omega_m^d} \left(\sum_{l\in\mathbb{Z}^d} \widehat{\Phi}(\xi+l)\overline{\widehat{\psi^i_{j/m}}(\xi+l)} \right) \times \left(\sum_{l'\in\mathbb{Z}^d} \overline{\widehat{\Phi}(\xi+l')}\widehat{\psi^i_{j/m}}(\xi+l') \right)^T.$$

For any $f = \sum_{l\in\mathbb{Z}^d} C(l)^T\Phi(\cdot - l)$ in $V^2(\Phi)$, we see from (36) and (37) that

$$\widehat{(S_m f)}(\xi) = \widehat{C}(\xi)^T \left([A_m]^\Psi_\Phi(\xi)\right)\widehat{\Phi}(\xi).$$

Define B^i_j to be the coefficient vector sequence for the function $\theta^i_{j/m}$, i.e., $B^i_j = \left((b^i_j)^1, \ldots, (b^i_j)^r\right)^T$, where $(b^i_j)^n(l) = \left\langle \theta^i_{j/m}, \phi^n(\cdot - l)\right\rangle = \left\langle \psi^i_{j/m}, \phi^n(\cdot - l)\right\rangle$. Then

$$\widehat{(S_m\theta^i_{j/m})}(\xi) = \widehat{B^i_j}(\xi)^T \left([A_m]^\Psi_\Phi(\xi)\right)\widehat{\Phi}(\xi). \qquad (38)$$

If $[A_m]^\Psi_\Phi(\xi)$ is invertible, then

$$\widehat{S_m^{-1}\theta^i_{j/m}}(\xi) = \widehat{B^i_j}(\xi)^T \left([A_m]^\Psi_\Phi(\xi)\right)^{-1}\widehat{\Phi}(\xi).$$

Using property (i) of the Fourier transform, it can easily be verified that

$$S_m^{-1}\theta^i_{k+j/m} = (S_m^{-1}\theta^i_{j/m})(\cdot - k) = \widetilde{\theta}^i_{j/m}(\cdot - k).$$

Remark 12: It can be shown that $\alpha_m I \le [A_m]^\Psi_\Phi(\xi)$ for almost every ξ, where α_m is the positive lower bound in (23). Thus, for almost every ξ, $[A_m]^\Psi_\Phi(\xi)$ is invertible for every $m \ge 1$. However, for large enough m, we will show a stronger result below, namely that there is a positive number δ (that does not depend on m) such that $\delta I \le \frac{1}{m^d}[A_m]^\Psi_\Phi(\xi)$ for every ξ.

We are now ready to compute $\mathrm{var}(f_{\varepsilon,m}(x) - f(x))$.

$$\begin{aligned}
&\mathrm{var}(f_{\varepsilon,m}(x) - f(x)) \\
&= \mathrm{var}\left(\sum_{i=1}^s \sum_{j\in\Omega_m^d} \sum_{k\in\mathbb{Z}^d} \varepsilon^i_{k+j/m}\widetilde{\theta}^i_{j/m}(x-k) \right) \\
&= \sigma^2 \sum_{i=1}^s \sum_{j\in\Omega_m^d} \sum_{k\in\mathbb{Z}^d} \left| S_m^{-1}\theta^i_{j/m}(x-k)\right|^2 \\
&= \sigma^2 \sum_{i=1}^s \sum_{j\in\Omega_m^d} \int_{[0,1]^d} \left| \sum_{k\in\mathbb{Z}^d} \widehat{S_m^{-1}\theta^i_{j/m}}(k-\xi)e^{\mathrm{i}2\pi x\cdot(k-\xi)} \right|^2 d\xi \\
&= \sigma^2 \sum_{i=1}^s \sum_{j\in\Omega_m^d} \int_{[0,1]^d} \left| \sum_{k\in\mathbb{Z}^d} \widehat{S_m^{-1}\theta^i_{j/m}}(k-\xi)e^{\mathrm{i}2\pi x\cdot k} \right|^2 d\xi \\
&= \sigma^2 \int_{[0,1]^d} \sum_{i=1}^s \sum_{j\in\Omega_m^d} \left| \widehat{B^i_j}(-\xi)^T \left([A_m]^\Psi_\Phi(-\xi)\right)^{-1} \left(Z\widehat{\Phi}(-\xi,-x)\right)\right|^2 d\xi.
\end{aligned}$$

We notice that the matrix $\left([A_m]^\Psi_\Phi(-\xi)\right)^{-1}$ is self-adjoint, and use the argument (34) from the proof of Theorem 3, along with the fact that

$$\sum_{i=1}^s \sum_{j\in\Omega_m^d} \overline{\widehat{B^i_j}(-\xi)}\widehat{B^i_j}(-\xi)^T = [A_m]^\Psi_\Phi(-\xi),$$

to get

$$\begin{aligned}
&\mathrm{var}\big(f_{\varepsilon,m}(x) - f(x)\big) \\
&= \sigma^2 \int_{[0,1]^d} \left(\overline{Z\widehat{\Phi}(-\xi,-x)}\right)^T \left([A_m]^\Psi_\Phi(-\xi)\right)^{-1} \times \left(Z\widehat{\Phi}(-\xi,-x)\right) d\xi \\
&= \frac{\sigma^2}{m^d} D_x(m).
\end{aligned}$$

where

$$D_x(m) = \int_{[0,1]^d} \left(\overline{Z\widehat{\Phi}(-\xi,-x)}\right)^T \left(\frac{1}{m^d}[A_m]^\Psi_\Phi(-\xi)\right)^{-1} \times \left(Z\widehat{\Phi}(-\xi,-x)\right) d\xi. \qquad (39)$$

Lemma 13: There exists a number $\delta > 0$ and a number $M \in \mathbb{N}$ such that for every $m \ge M$,

$$\delta I \le \frac{1}{m^d}[A_m]^\Psi_\Phi(\xi) \qquad \text{for all } \xi \in [0,1]^d.$$

This lemma is proved in Section V-D. Thus for large enough m, we have

$$\begin{aligned}
&\mathrm{var}\big(f_{\varepsilon,m}(x) - f(x)\big) \\
&\le \frac{\sigma^2}{m^d}\left(\frac{1}{\delta}\right) \int_{[0,1]^d} \sum_{n=1}^r \left| Z\widehat{\phi^n}(-\xi,-x)\right|^2 d\xi \\
&= \frac{\sigma^2}{m^d}\left(\frac{1}{\delta}\right) \left(\sum_{n=1}^r \left(\sum_{k\in\mathbb{Z}^d} |\phi^n(x+k)|^2 \right)\right) \\
&\le \frac{\sigma^2}{m^d}\left(\frac{1}{\delta}\right) \left(\sum_{n=1}^r \|\phi^n\|^2_{W^1} \right) \qquad \text{for all } x \in \mathbb{R}^d.
\end{aligned}$$

□

D. Proof of Lemma 13

First, for $\xi \in [0,1]^d$, define the self-adjoint matrix

$$G^\Psi_\Phi(\xi) := \sum_{i=1}^s \sum_{k\in\mathbb{Z}^d} \widehat{\Phi}(\xi+k)\overline{\widehat{\Phi}(\xi+k)}^T \left|\widehat{\psi^i}(\xi+k)\right|^2.$$

We will now show that

$$\frac{1}{m^d}[A_m]^\Psi_\Phi(\xi) \quad \overset{m\to\infty}{\longrightarrow} \quad G^\Psi_\Phi(\xi) \qquad \text{for every } \xi \in [0,1]^d,$$

i.e., for each $\xi \in [0,1]^d$, each entry of the matrix $\frac{1}{m^d}[A_m]^\Psi_\Phi(\xi)$ converges to the corresponding entry of the matrix $G^\Psi_\Phi(\xi)$. For $1 \le n, n' \le r$, we look at the (n,n')-entry of the matrix $\frac{1}{m^d}[A_m]^\Psi_\Phi(\xi)$.

$$\left(\frac{1}{m^d}\left[A_m\right]_\Phi^\Psi(\xi)\right)_{(n,n')}$$
$$= \sum_{i=1}^s \frac{1}{m^d}\sum_{j\in\Omega_m^d}\left(\sum_{k\in\mathbb{Z}^d}\widehat{\phi^n}(\xi+k)\overline{\widehat{\psi^i_{j/m}}(\xi+k)}\right) \times\left(\sum_{k'\in\mathbb{Z}^d}\overline{\widehat{\phi^{n'}}(\xi+k')}\widehat{\psi^i_{j/m}}(\xi+k')\right)$$
$$= \sum_{i=1}^s \frac{1}{m^d}\sum_{j\in\Omega_m^d}\left(Z(\widehat{\phi^n*\psi^i}^\vee)(\xi,-j/m)\right) \times\left(\overline{Z(\widehat{\phi^{n'}*\psi^i}^\vee)(\xi,-j/m)}\right)$$
$$\stackrel{m\to\infty}{\longrightarrow} \sum_{i=1}^s\int_{[0,1]^d}\left(Z(\widehat{\phi^n*\psi^i}^\vee)(\xi,-x)\right) \times\left(\overline{Z(\widehat{\phi^{n'}*\psi^i}^\vee)(\xi,-x)}\right)\,dx$$
$$= \sum_{i=1}^s\left\langle q_\xi^{n,i},q_\xi^{n',i}\right\rangle_{L^2([0,1]^d)}$$

where, for $1\le l\le r$, q_ξ^l is the function on $[0,1]^d$ whose Fourier coefficients $\widehat{q_\xi^{l,i}}(k)$ are given by

$$\widehat{q_\xi^{l,i}}(k) = \widehat{\phi^l}(\xi+k)\overline{\widehat{\psi^i}(\xi+k)}.$$

Invoking Plancherel identity, we have

$$\sum_{i=1}^s\left\langle q_\xi^{n,i},q_\xi^{n',i}\right\rangle_{L^2([0,1]^d)}$$
$$= \sum_{i=1}^s\left\langle \widehat{q_\xi^{n,i}},\widehat{q_\xi^{n',i}}\right\rangle_{l^2(\mathbb{Z}^d)}$$
$$= \sum_{i=1}^s\sum_{k\in\mathbb{Z}^d}\widehat{\phi^n}(\xi+k)\overline{\widehat{\phi^{n'}}(\xi+k)}\left|\widehat{\psi^i}(\xi+k)\right|^2$$
$$= \left[G_\Phi^\Psi(\xi)\right]_{(n,n')}$$

Thus $\frac{1}{m^d}\left[A_m\right]_\Phi^\Psi(\xi) \stackrel{m\to\infty}{\longrightarrow} G_\Phi^\Psi(\xi)$ for each $\xi\in[0,1]^d$. Now we claim, for fixed (n,n'), the collection $\left\{\left(\frac{1}{m^d}\left[A_m\right]_\Phi^\Psi(\cdot)\right)_{(n,n')} : m\in\mathbb{N}\right\}$ is equicontinuous, which implies that $\left(\frac{1}{m^d}\left[A_m\right]_\Phi^\Psi(\cdot)\right)_{(n,n')}$ converges uniformly to $\left[G_\Phi^\Psi(\cdot)\right]_{(n,n')}$ on $[0,1]^d$.

In a manner similar to that in the proof of Lemma 11, it can be verified that

$$\left(\frac{1}{m^d}\left[A_m\right]_\Phi^\Psi(\xi)\right)_{(n,n')}$$
$$= \sum_{i=1}^s\frac{1}{m^d}\times\sum_{j\in\Omega_m^d}\sum_{l\in\mathbb{Z}^d}\sum_{k\in\mathbb{Z}^d}(\phi^n*\psi^{i\vee})(j/m+l) \times(\phi^{n'}*\psi^{i\vee})(j/m+l+k)e^{-\mathbf{i}2\pi\xi\cdot k}.$$

Because $W^1*L^1\subset W^1$, we know that $\phi^n*\psi^{i\vee}\in W^1$, and therefore, the argument from Lemma 11 can be used to show the collection is equicontinuous.

In [5] it is shown that (23) and (24) imply that there exists a positive number δ_0 such that $\delta_0 I\le G_\Phi^\Psi(\xi)$ for all ξ. Let $\delta=\frac{\delta_0}{2}$. Because $\left(\frac{1}{m^d}\left[A_m\right]_\Phi^\Psi(\cdot)\right)_{(n,n')}$ converges uniformly to $\left[G_\Phi^\Psi(\cdot)\right]_{(n,n')}$ on $[0,1]^d$, there exists a number $M\in\mathbb{N}$ such that for all $m\ge M$

$$\delta I\le\frac{1}{m^d}\left[A_m\right]_\Phi^\Psi(\xi)\qquad\text{for all }\xi.$$

□

VI. Discussion and examples

A. Computational aspects

Theorems 3 and 7 give the exact values of the variance of the reconstruction error in term of the position x. These values depend on the terms $C_x(m)$ and $D_x(m)$ that involve integrals of expressions in the Zak transforms of the generator Φ and the sampler Ψ. Although the computations can be obtained by numerical calculations and integrations, several observations can simplify these calculations and render them more precise. In particular, $C_x(m)$ in Theorem 3 whose expression is given by (35) can be computed using any generator $\Phi^{\approx}$ generating the same space as Φ, i.e., $V^2(\Phi^{\approx})=V^2(\Phi)$. Moreover, using (12), we can write $C_x(m)$ as

$$C_x(m) = \int_{[0,1]^d}\left(\overline{Z\Phi^{\approx}(x,-\xi)}\right)^T\left(\frac{1}{m^d}A_m^{\approx}(\xi)\right)^{-1} \times\left(Z\Phi^{\approx}(x,-\xi)\right)\,d\xi,$$

where

$$A_m^{\approx}(\xi):=\sum_{j\in\Omega_m^d}\overline{Z\Phi^{\approx}(j/m,-\xi)}Z\Phi^{\approx}(j/m,-\xi)^T.$$

We can also write an equivalent expression for $C_x(m)$ (thanks to (12)) by replacing $\Phi^{\approx}$ by $\widehat{\Phi^{\approx}}$ in the two equations above. If we know that $V^2(\Phi)$ can be generated by $\Phi^{\approx}=(\phi^1,\dots,\phi^r)^T$, where all the ϕ^is have compact support, then the infinite sums in the Zak transforms become finite sums, and no truncation is necessary. For this case, the terms appearing in the integrand are trigonometric polynomials, and the computation of $C_x(m)$ become easier, more accurate and precise, as in the example in Section VI-B below.

Similarly, the computation of $D_x(m)$ can be obtained by a similar formula

$$D_x(m) = \int_{[0,1]^d}\left(\overline{Z\Phi^{\approx}}(x,-\xi)\right)^T\left(\frac{1}{m^d}\left[A_m\right]_{\Phi^{\approx}}^\Psi(\xi)\right)^{-1} \times\left(Z\Phi^{\approx}(x,-\xi)\right)\,d\xi,$$

where

$$\left[A_m\right]_{\Phi^{\approx}}^\Psi(\xi) := \sum_{i=1}^s\sum_{j\in\Omega_m^d}\overline{\left(Z\Phi^{\approx}*\psi^i(j/m,-\xi)\right)} \times\left(Z\Phi^{\approx}*\psi^i(j/m,-\xi)\right)^T.$$

We can also write an equivalent expression for $D_x(m)$ (thanks to (12)) by replacing $\Phi^{\approx} * \psi^i$ by $\widehat{\Phi^{\approx}}\widehat{\psi^i}$ in the two equations above. However if we wish to compute the the limit of $C_x(m)$ or $D_x(m)$ from their given expression in Theorems 3 and 7, we must use a generator Φ that generates an orthonormal basis and not any generator $\Phi^{\approx}$.

B. Examples

For a single generator ϕ, i.e., $r = 1$ which is often the case in applications, the formula for $C_x(m)$ is given by

$$C_x(m) = \int_{[0,1]^d} \frac{|Z\phi(x, -\xi)|^2}{\frac{1}{m^d}\sum_{j\in\Omega_m^d} |Z\phi(j/m, \xi)|^2}\, d\xi.$$

Therefore $C_x(m)$ is a periodic function of x, i.e., $C_{x+1}(m) = C_x(m)$. For the spline space models, $C_x(m)$ can be easily computed using $\phi_{\approx} = \beta^n, n \geq 1$ which are compactly supported generators. In fact for the spline space models, explicit formulae can be obtained. In particular, for the linear spline space (i.e., $n = 1$) we have that

$$C_x(m) = \int_0^1 \frac{1 - 4x\cos^2\pi\xi + 4x^2\cos^2(\pi\xi)}{1 - 2(1-\frac{1}{m})\cos^2\pi\xi + \frac{2}{3}(1-\frac{1}{m})(2-\frac{3}{m})\cos^2(\pi\xi)} d\xi$$

and

$$\begin{aligned} C_x(\infty) &= \lim_{m\to\infty} C_x(m) \\ &= \int_0^1 \frac{1 - 4x\cos^2\pi\xi + 4x^2\cos^2(\pi\xi)}{1 - 2\cos^2\pi\xi + \frac{4}{3}\cos^2(\pi\xi)} d\xi \end{aligned}$$

where $x \in [0, 1]$, while for the quadratic spline model space (i.e., $n = 2$),

$$C_x(m) = \int_0^1 \frac{P_1(x,\xi)}{m^{-1}\sum_{j=0}^{m-1} P_1(j/m,\xi)} d\xi$$

and

$$C_x(\infty) = \lim_{m\to\infty} C_x(m) = \int_0^1 \frac{P_1(x,\xi)}{\frac{11}{10} + \frac{13}{15}\cos 2\pi\xi + \frac{19}{30}\cos 4\pi\xi} d\xi$$

where $x \in [0, 1]$ and $P_1(x, \xi) = (1+3x^2-6x^3+3x^4)+(1-4x^2+8x^3-4x^4)\cos 2\pi\xi+(x^2-2x^3+x^4)\cos 4\pi\xi$. Similarly for the cubic spline model (i.e. $n = 3$), we obtain that

$$C_x(m) = \int_0^1 \frac{P_2(x,\xi)}{m^{-1}\sum_{j=0}^{m-1} P_2(j/m,\xi)} d\xi$$

and

$$\begin{aligned} C_x(\infty) &= \lim_{m\to\infty} C_x(m) \\ &= \int_0^1 \frac{P_2(x,\xi)}{\frac{604}{35} + \frac{211}{70}\cos 2\pi\xi + \frac{12}{7}\cos 4\pi\xi + \frac{1}{70}\cos 6\pi\xi} d\xi \end{aligned}$$

where $P_2(x, \xi) = (18-18x^2+16x^3+42x^4-60x^5+20x^6)+(16+24x^2-18x^3-66x^4+90x^5-30x^6)\cos 2\pi\xi+(2-6x^2+30x^4-36x^5+12x^6)\cos 4\pi\xi+(2x^3-6x^4+6x^5-2x^6)\cos 6\pi\xi$.

Figures 1 show $C_x(m)$ for various values of the oversampling parameter m, while Figure 2 shows the value of $D_x(m)$ for several values of m for the case of average sampling. In all cases the function $C_x(m)$ and $D_x(m)$ approach a limit predicted by the theory.

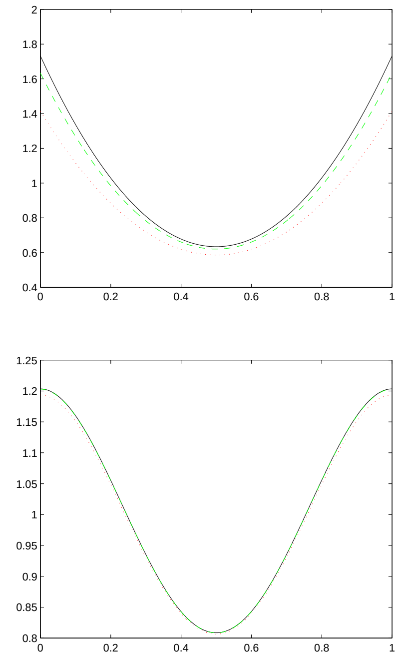

Fig. 1. Polynomial spline models for ideal sampling: (a) The function $C_x(m)$ for polynomial splines of degree one: $m = 255$ continuous line; $m = 4$ dashed line; $m = 2$ dotted line. (b) The function $C_x(m)$ for polynomial splines of degree three: $C_x(\infty)$ continuous line; $m = 4$ dashed line; $m = 2$ dotted line. Notice that at this scale, the two curves are almost indistinguishable.

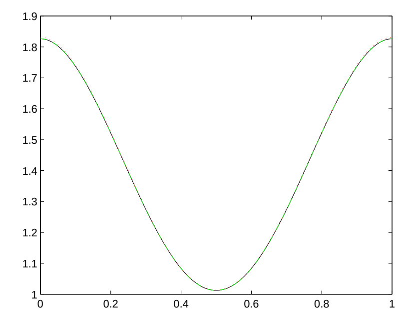

Fig. 2. Polynomial spline of degree three: The function $D_x(m)$ for polynomial splines of degree three with average sampler $\psi = \chi_{[0,1)}$: $D_x(\infty)$ continuous line; $m = 4$ dashed line; $m = 2$ dotted line. Notice that at this scale, the three curves are almost indistinguishable.

For spaces that do not have any generator with compact support, we can approximate the the infinite sums in the formula by finite sums since the generator ϕ has decay (recall that we assume that $\phi \in W^1$), and the value of $C_x(m)$ can be found by numerical integration.

For the bandlimited case, the integrand is equal to 1 for all values of x and $\xi \in [0, 1)$. Thus for this case $C_x(m)$ is independent of x and the variance of the error is $\frac{\sigma^2}{m^d}$.

Finally for the case of piecewise constants we have that the variance of the error is independent of x and it is given by $\frac{\sigma^2}{m^d}$.

C. L^2-estimates

The variance of the error between the reconstructed function and the original function is given pointwise for each x. The knowledge of this pointwise variance is important in some applications as in [24]. However, in other applications, L^2 estimates are more appropriate, see e.g., [21]. Since our assumption is that the noise $\{\varepsilon_{k+j/m}\}$ is a collection of i.i.d. random variables satisfying $E(\varepsilon_{k+j/m}) = 0$, $\text{var}(\varepsilon_{k+j/m}) = \sigma^2$, the variance of the $L^2(\mathbb{R}^d)$-norm of the error is infinite for each m. This problem may be circumvented by assuming, the more practical situation in which we have only a finite number of samples n. This type of analysis has been done for the case of bandlimited functions by Pawlak, and Rafajlowicz in [21] yielding an error estimate of order $O(n^{-1/3})$. We do not analyze this situation in this paper. However, because of the variance of the error is $\mathbb{Z}^d$-periodic, it is natural to compute the variance of the L^2 error over a cube of side length 1. Using the assumption that the added noise is an i.i.d random variable, it can be checked that this error can be computed directly from the pointwise estimate and we get that

$$\text{var}(\|f_{\varepsilon,m}(x) - f(x)\|^2_{L^2(I)}) = \frac{\sigma^2}{m^d} \int\limits_{[0,1]^d} C_x(m)dx,$$

where I is any unit cube in $\mathbb{R}^d$. Similarly, for the case of reconstructions from averages we get

$$\text{var}(\|f_{\varepsilon,m}(x) - f(x)\|^2_{L^2(I)}) = \frac{\sigma^2}{m^d} \int\limits_{[0,1]^d} D_x(m)dx.$$

VII. Concluding remarks

We have analyzed the frame reconstruction of a function from its noisy samples. In our analysis, it is assumed that the sampled function f belongs to a shift invariant space $V^2(\Phi)$ yielding an estimated reconstructed function $f_{\varepsilon,m}(x)$ which is unbiased. It would be interesting to assume that f does not belong to the reconstructed space $V^2(\Phi)$ (e.g., f belongs to some Sobolev space), and only finitely many samples n are available. Methods similar to the ones used by Pawlak, Rafajlowicz, and Krzyzak in [23] may be possible for such analysis. We also assume that the noise $\{\varepsilon_{k+j/m}\}$ is bounded. This technical assumption insures the convergence of the infinite series in the frame reconstruction. However formally, unbounded Gaussian noise yields the same exact result, but a more technical mathematical justification would be needed for the argument. Another interesting and important situation is when the noise is colored. Such analysis for the case of bandlimited signals as well as for L^2 signals with some regularity properties has been done in [23]. This situation is not covered in this paper.

References

[1] A. Aldroubi, *Non-uniform weighted average sampling and reconstruction in shift-invariant and wavelet spaces*, Appl. Comput. Harmon. Anal. 13 (2002) 151-161.

[2] A. Aldroubi, K. Gröchenig, *Beurling-Landau-type theorems for non-uniform sampling in shift invariant spline spaces*, J. Fourier Anal. Appl. 6 (2000) 93-103.

[3] A. Aldroubi, K. Gröchenig, *Nonuniform sampling and reconstruction in shift-invariant spaces*, SIAM Rev. 43 (4) (2001) 585-620.

[4] A. Aldroubi, I. Krishtal, *Robustness of sampling and reconstruction and Beurling-Landau-type theorems for shift invariant spaces*, Appl. Comput. Harmon. Anal., 20 (2006) 256–260.

[5] A. Aldroubi, Q. Sun, W. Tang, *Convolution, average sampling, and a Calderon resolution of the identity for shift-invariant spaces*, J. Fourier Anal. Appl. 22 (2005) 215-244.

[6] J.J. Benedetto, *Irregular sampling and frames*, in Wavelets: A Tutorial in Theory and Applications, C.K. Chui, ed., Academic Press, Boston, 1992.

[7] C. Blanco, C. Cabrelli, S. Heineken, *Functions in sampling spaces*, STPIP, 5 (3) (2006) 275–296.

[8] N. Bissantz, H. Holzmann, and A. Munk, *Testing parametric assumptions on band- or time-limited signals under noise*, IEEE Trans. Informations Theory, (2005), 3797–3805.

[9] P.L. Butzer, G. Schmeisser, R.L. Stens, *An introduction to sampling analysis*, in Nonuniform Sampling Theory and Practice, F. Marvasti, ed., Kluwer/Plenum, New York, 2001.

[10] O. Christensen, *An Introduction to Frames and Riesz Bases*, Birkhäuser, Boston, 2003.

[11] Z. Cvetkovic, M. Vetterli, *Oversampled filter banks*, IEEE Trans. Signal Process 46 (5) (1998) 1245-1255.

[12] Y. Eldar, M. Unser, *Nonideal sampling and interpolation from noisy observations in shift-invariant spaces*, IEEE Trans. Signal Process 54 (7) (2006) 2636-2651.

[13] H.G. Feichtinger, K. Gröchenig, *Error analysis in regular and irregular sampling theory*, Appl. Anal. 50 (1993) 167-189.

[14] K. Gröchenig, H. Schwab, *Fast local reconstruction methods for nonuniform sampling in shift-invariant spaces*, SIAM J. Matrix Anal. Appl. 24 (4) (2003) 899-913.

[15] K. Gröchenig, *Foundations of Time Frequency Analysis*, Birkhäuser, Boston, 2001.

[16] C. Heil, D. Walnut, *Continuous and discrete wavelet transforms*, SIAM Review, 31 (1989) 628-666.

[17] A.J.E.M. Janssen, *The Zak transform: a signal transform for sampled time-continuous signals*, Philips J. Res. 39 (1998), 23-69.

[18] A. Jerri, *The Shannon sampling theorem - its various extensions and applications: a tutorial review*, Proc. IEEE 65 (1977) 1565-1596.

[19] B. Lacaze, *Reconstruction of stationary processes sampled at random times*, in Nonuniform Sampling Theory and Practice, F. Marvasti, ed., Kluwer/Plenum, New York, 2001.

[20] F. Marvasti, *Random topics in nonuniform sampling*, in Nonuniform Sampling Theory and Practice, F. Marvasti, ed., Kluwer/Plenum, New York, 2001.

[21] M. Pawlak, and E. Rafajlowicz, *On restoring bandlimited signals*, IEEE Trans. Information Theory, (5) (1994), 1490–1503.

[22] M. Pawlak and U. Stadtmuller, *Ststistical aspects of sampling for noisy and grouped data*, In advances in Sannon Sampling Theory: Mathematics and Applications, eds. J.J. Benedetto and P. Ferreira, 2001.

[23] M. Pawlak, E. Rafajlowicz, and A. Krzyzak, *Postfiltering versus prefiltering for signal recovery from noisy samples*, IEEE Trans. Information Theory, 49 (12) (2003, 3195–3212.

[24] G.K. Rohde, C.A. Berenstein, D.M. Healy, Jr., *Measuring image similarity in the presence of noise*, Proceedings of the SPIE, vol. 5747, Medical Imaging 2005: Image Processing; pp. 132-143, 2005.

[25] S. Smale, D.X. Zhou, *Shannon sampling and function reconstruction from point values*, Bull. Amer. Math. Soc. 41 (3) (2004) 279-305.

[26] T. Strohmer, *Numerical analysis of the non-uniform sampling problem*, J. Comp. Appl. Math. 122 (2000) 297-316.

[27] G.G. Walter, *A sampling theorem for wavelet subspaces*, IEEE Trans. Inform. Theory 38 (2) (1992) 881-884.

[28] J.M. Whittaker, *Interpolatory Function Theory*, Cambridge University Press, London, 1935.

[29] X. Zhou, W. Sun, *On the sampling theorem for wavelet subspaces*, J. Fourier Anal. Appl. 5 (4) (1999) 347-354.

Akram Aldroubi received the M.S. degree in electrical engineering from the Swiss Federal Institute of Technology, Lausanne, Switzerland in 1982. He received the MS and Ph.D. degrees in mathematics from Carnegie-Mellon University, PA, in 1984,1987 respectively.

He is currently Professor of Mathematics at Vanderbilt University. His research interest includes sampling theory, frame theory, wavelets, signal and image processing, functional analysis, harmonic analysis, and mathematical biology.

Prof. Aldroubi is a member of the American Mathematical Society. He received the Rene Cousin Prize for excellence in mathematics from the Swiss Federal Institute of Technology in 1982, and the IEEE Signal Processing Society's 1995 Best Senior Paper Award (IMDSP technical area) for a Transaction paper (with M. Unser and M. Eden) on B-spline signal processing. He is the co-editor-in-chief of the Journal of Sampling Theory in Signal and Image Processing. He is on the editorial board The Journal of Mathematical Imaging and Vision, The Journal of Fourier Analysis and Applications. He is co-editor (with M. Unser) of the book "Wavelet in Medicine and Biology" CRC Press 1996, and co-editor (with E.B. Lin) of the book, "Wavelets, Multiwavelets, and Applications," in Contemporary Mathematics, published by the American Mathematical Society, 1998.

Qiyu Sun received the BSc and PhD degree in mathematics from Hangzhou University, China in 1985,1990 respectively. He is currently with Department of Mathematics, University of Central Florida. His prior position was with the Zhejiang University (China), the National University of Singapore (Singapore), Vanderbilt University, and University of Houston. His research interest includes sampling theory, multi-band wavelets, wavelet and frame theory, and Fourier analysis. He has published more than 80 papers and written a book "An Introduction to Multiband Wavelets" with Ning Bi and Daren Huang. He is the editor of the journal of Advance in Computational Mathematics.

Casey Leonetti earned her Ph.D. in mathematics from Vanderbilt University in 2007 and B.A. from Rice University in 2000. She currently works as a consultant for The Boston Consulting Group in Houston, TX.