Sampling and aliasing without translation-invariance

J. A. Hogan
Department of Mathematical Sciences
University of Arkansas
Fayetteville, AR 72701

J. Lakey
Department of Mathematical Sciences
New Mexico State University
Las Cruces, NM 88003–8001

Abstract— We investigate the sampling of functions in principal shift-invariant spaces with the aid of the Zak transform. The main question we ask is: when can a signal in such a space be reconstructed from its samples in a manner analogous to the classical sampling theorem of Shannon for bandlimited functions? The answer is shown to depend on general properties of the generator of these spaces, primarily orthogonality, smoothness, self-similarity and the nature of the sampling. We address the question of how to control aliasing error and the problems associated with the non-translation-invariance of these spaces.

I. INTRODUCTION

In this paper we investigate several issues surrounding the sampling of signals in principal shift-invariant spaces. The archetype of such spaces is the space V of signals bandlimited to $[-\frac{1}{2}, \frac{1}{2}]$, i.e., $f \in V \iff f \in L^2(\mathbb{R})$ and $\hat{f}(\xi) = 0$ for $|\xi| > \frac{1}{2}$. The Shannon sampling theorem governs the sampling of such functions and has become the mathematical technology that allows for the digitisation of analog signals and the reconstruction of such signals from their samples.

In recent years, the theory of wavelets has made an enormous impact on digital signal processing. In this context, information is thought to be composed of differing levels of resolution. This heirarchy of details is captured by a *multiresolution analysis.* It is appropriate then to consider the sampling of signals in these spaces and to try to provide for multiresolution analysis that which the Shannon sampling theorem provides for Fourier analysis – a means of moving seamlessly between the analog and digital realms.

It should be stressed that while several new techniques for sampling in shift-invariant and multiresolution spaces are introduced in this paper, the principal focus and main results are in the validation of these schemes and those of several other authors, i.e., this paper gives general sufficient conditions on shift-invariant spaces for these sampling procedures to be effective. Questions related to aliasing and the problems associated with the non-translation-invariance of general multiresolution spaces are also considered.

The results here concern sampling rather than the "generalised sampling" of Unser and Zerubia [1]. All of the sampling operators considered in this paper are (non-orthogonal) projections, and are associated with ideal samplers, i.e., given a continuous signal f, we assume we are able to obtain the value of $f(x_0)$ for any x_0, rather than just an averaged or smoothed approximation $a_{t_0} = \int f(t)\Phi(t - t_0)\, dt$ for some distribution Φ supported on a set of positive measure around 0.

II. SAMPLING IN SHIFT-INVARIANT SPACES

A. Definitions and Examples

By a principal shift-invariant space, we mean a closed subspace V of $L^2(\mathbb{R})$ and a function $\varphi \in V$ such that the collection $\{\varphi(x - k)\}_{k\in\mathbb{Z}}$ is an orthonormal basis for V. The orthonormality condition on the integer shifts of φ may be weakened to a Riesz basis condition without difficulty.

As a first example, consider the space of functions

$$V_H = \{f \in L^2;\ f\big|_{[k,k+1)} = \text{const}\}$$

with $\varphi_H(t) = \chi_{[0,1)}(t)$. The pair (V_H, φ_H) is the Haar PSI space. Since each $f \in V_H$ is continuous except at the integers, it is clear that samples taken at $x_0 + \mathbb{Z}$ are well-defined when $0 < x_0 < 1$ and

$$f(x) = \sum_k f(x_0 + k)\varphi_H(x - k),$$

a sampling result for signals in V_H.

In the context of sampling, a more celebrated example, and one that is more illuminating to the path we will follow, is the so-called Shannon PSI space (V_S, φ_S) where

$$V_S = \{f \in L^2(\mathbb{R});\ \hat{f}(\xi) = 0 \text{ if } |\xi| > \frac{1}{2}\}$$

and $\varphi_S(t) = \frac{\sin \pi t}{\pi t}$. The Shannon sampling theorem states that each $f \in V_S$ may be reconstructed from its sampled values $\{f(k)\}_{k\in\mathbb{Z}}$ via

$$f(t) = \sum_k f(k)\varphi_S(t - k).$$

---

[*] Email for JL: `jlakey@nmsu.edu`; Email for JAH: `jeffh@mail.uark.edu`;

Keywords: sampling, Zak transform, wavelets
AMS Subject Classification: 42C40, 94A20.
Thanks Roy. Thanks H.G.

B. Zak transform

The principal tool in the analysis of sampling operators is the Zak transform whose definition and basic properties we briefly outline. An excellent reference for properties and applications of the Zak transform is [2]. Given $f \in \mathcal{S}(\mathbb{R})$ (the Schwarz space of rapidly decreasing functions) and $x, \xi \in \mathbb{R}$, let the Zak transform of f at the point (x, ξ) in phase space be given by

$$Zf(x,\xi) = \sum_k f(x+k)e^{2\pi ik\xi}.$$

The Zak transform satisfies a quasi-periodicity property: $Zf(x+k, \xi+l) = e^{-2\pi ik\xi}Zf(x,\xi) \quad (k,l \in \mathbb{Z})$. Hence, the values of Zf on the square $Q = [0,1]\times[0,1]$ imply the values of f on all of phase space and we think of Q as the domain of Zf. The Zak transform is unitary on $\mathcal{S}(\mathbb{R})$: $\int_0^1\int_0^1 |Zf(x,\xi)|^2\,dx\,d\xi = \int_{-\infty}^{\infty}|f(x)|^2\,dx$ and hence Z extends to a unitary operator $Z : L^2(\mathbb{R}) \to L^2(Q)$. We will also want to extend the Zak transform to $\mathbb{R}\times\mathbb{C}$ – for $x \in \mathbb{R}$, $z \in \mathbb{C}$ let $Z_{\mathbb{C}}f(x,z)$ denote the Laurent polynomial $Z_{\mathbb{C}}f(x,z) = \sum_k f(x+k)z^k$. Note that $Z_{\mathbb{C}}(x, e^{2\pi i\xi}) = Zf(x,\xi)$ and when f is supported on $[0,M]$ (M a positive integer) and $0 \le x < 1$ then $Zf(x,z)$ is a polynomial of degree at most $M-1$.

C. Multiresolution analysis and wavelets

Let φ be a scaling function for an MRA of $L^2(\mathbb{R})$ and H be the scaling filter associated with φ, i.e., φ satisfies a dilation equation of the form

$$\varphi(\frac{x}{2}) = \sum_k h_k\varphi(x-k) \tag{1}$$

and $H(z) = \sum_k h_k z^k$ $(z \in \mathbb{C}\setminus\{0\})$. We assume φ is supported on $[0,L]$. Then $h_k = 0$ for $k < 0$ or $k > L$, i.e., H is a polynomial of degree L. Further, we assume the three conditions: (a) $H(1) = 1$, (b) $H(z)\overline{H(\frac{1}{\bar{z}})} + H(-z)\overline{H(\frac{-1}{\bar{z}})} \equiv 1$ $(z \neq 0)$, and (c) the τ-cycle condition: there is no non-trivial cycle $\{\zeta_1,\dots,\zeta_n\}$ for the operation $\tau\zeta = \zeta^2$ acting on the circle $\mathbb{T}$ (i.e., $\tau\zeta_i = \zeta_{i+1}$ and $\zeta_1 = \zeta_n$) such that $|H(\zeta_i)| = 1$ for all i. Together, these conditions are necessary and sufficient to ensure that the integer translates of φ are orthonormal. The wavelet ψ associated with the scaling function φ is given by

$$\psi(\frac{x}{2}) = \sum_k g_k\varphi(x-k) \tag{2}$$

with $g_k = (-1)^k\overline{h_{1-k}}$. Further details on wavelets and multiresolution analysis may be found in [3].

It was observed by Janssen [4] that equation (1) forces a self-similarity property on $Z_{\mathbb{C}}\varphi$. In fact, if we define an operator T on Laurent series by $Tf(z^2) = H(z)f(z) + H(-z)f(-z)$, then by taking the Zak transform of (1) we have

$$Z_{\mathbb{C}}\varphi(x,z) = T(Z_{\mathbb{C}}\varphi(2x,\cdot))(z). \tag{3}$$

As an operator on $L^2(\mathbb{T})$, the adjoint T^* acts via $T^*f(z) = 2\overline{H(\frac{1}{\bar{z}})}f(z^2)$, is a multiple of an isometry, and is a right inverse for T.

For applications of these ideas to the construction of compactly supported wavelets and wavelets amenable to extrapolation and sampling, the reader is referred to [5].

D. Critical Sampling

The basic sampling problem for a PSI space is that which generalises the Shannon sampling formula: can any signal $f \in V$ be recovered from its sampled values $f(k)$ $(k \in \mathbb{Z})$? Equivalently, one asks whether there exists $S \in V$ such that $f(x) = \sum_k c_k\varphi(x-j) = \sum_k f(k)S(x-k)$.

Let $f(x) = \sum_k c_k\varphi(x-k) \in V$. Then $Zf(x,\xi) = C(\xi)Z\varphi(x,\xi)$. If there exists $0 \le y < 1$ such that $\inf_\xi|Z\varphi(y,\xi)| > 0$, then we have $C(\xi) = \frac{Zf(y,\xi)}{Z\varphi(y,\xi)}$ and a simple calculation gives $f(x) = \sum_k f(y+k)S^y(x-k)$ where

$$S^y(x) = \int_0^1 \frac{Z\varphi(x,\xi)}{Z\varphi(y,\xi)}\,d\xi = \sum_k b_k\varphi(x-k) \tag{4}$$

and $b_k = \int_0^1 \frac{e^{-2\pi ik\xi}}{Z\varphi(y,\xi)}\,d\xi$. Janssen [4] (and Walter [6] when $y = 0$) gives explicit computations of these coefficients and associated sampling functions S^y for several examples.

There are, however, several drawbacks to this scheme. If $Z\varphi$ is continuous it will have a zero at some point $(x_0,\xi_0) \in Q$. (See [2].) For such a value of x_0, the sampling function S^{x_0} is not well-defined and the scheme will fail, i.e., the sampling operator $f \in V \to \{f(x_0+k)\}$ does not have a bounded inverse. Further, it is not known, even if φ is a (compactly supported) scaling function for an MRA of $L^2(\mathbb{R})$, whether there exsits $0 \le y < 1$ for which $\inf_\xi|Z\varphi(y,\xi)| > 0$. In this situation, shifted sampling would fail for all shifts y.

Another perhaps more serious problem is the following. Suppose (V,φ) is a PSI space, φ is compactly supported and $0 \le y < 1$ is such that $|Z\varphi(y,\xi)| \ge c > 0$ for all ξ. Then $Z\varphi(y,\xi)$ is a trigonometric polynomial in ξ and $(Z\varphi(y,\xi))^{-1}$ is not a trigonometric polynomial unless $\varphi(y+k) = \delta_{k-l}$ for some integer l. If φ is a scaling function, it has been shown by Xia and Zhang in [7] that this is not possible with $y = 0$. So in general there are infinitely many non-zero coefficients b_k in (4) and S^y will not be compactly supported. Note that if φ is supported on $[0,L]$ and $f \in V$ then for any given x the sum $f(x) = \sum_k c_k\varphi(x-k)$ has at most $L-1$ non-zero terms while in the sampling expansion

$f(x) = \sum_k f(y+k)S^y(x-k)$ there are infinitely many non-zero terms.

E. Staggered Sampling

In [8], Djokovic and Vaidyanathan propose several schemes to overcome the difficulties inherent in critical sampling, the first of which is known as staggered sampling. With staggered sampling, the sampling rate can be critical *on average*, but uniform samples are staggered by a collection of points in the support of φ. There are several of these schemes, but we discuss just one of them and leave the discussion of its general effectiveness – an issue not dealt with in [8] except through particular examples – until section III A.

Suppose (V, φ) is a PSI space, φ is supported on $[0, L]$ and $0 \le x_0 < x_1 < \cdots < x_{L-1} < 1$. Let $M^\varphi = M^\varphi(x_0, \dots, x_{L-1})$ be the $L \times L$ matrix with $(i,j)^{\text{th}}$ entry $[M^\varphi]_{ij} = \varphi(x_i + j)$ $(0 \le i, j \le L-1)$. If $f(x) = \sum_k c_k \varphi(x-k) \in V$ then $f(x_j + kL) = \sum_l c_{kL-l}[M^\varphi]_{jl}$. If M^φ is invertible, we have $c_{kL-m} = \sum_{j=0}^{L-1}[(M^\varphi)^{-1}]_{mj} f(x_j + kL)$ so that $f(x) = \sum_k \sum_{j=0}^{L-1} f(x_j + kL)S_j(x - kL)$ where

$$S_j(x) = \sum_{m=0}^{L-1} [(M^\varphi)^{-1}]_{mj} \varphi(x+m) \tag{5}$$

is supported on $[1-L, L]$.

The invertibility of M^φ for a general function φ is not addressed in [8] although several examples are examined. In particular, when φ is the quadratic B-spline $(L = 3)$, critical sampling at the integers is not possible since $Z\varphi_2(0, \frac{1}{2}) = 0$. However, $\det(M^\varphi(x_0, x_1, x_2))$ is a Vandermonde determinant and hence does not vanish as long as x_0, x_1, x_2 are distinct. To make general statements about the invertibility of M^φ, some restriction of the class of admissible φ is required.

F. Oversampling

Another way to obtain compactly supported sampling functions is to sample at a rate higher than the critical rate, i.e, by *oversampling*. The higher rate is compensated for by the improved flexibility such schemes allow and improved aliasing results, as we shall see in section III B.

On top of the orthonormality condition on φ, there are other conditions one might impose, including compact support, bandlimitedness, or the satisfaction of a dilation equation. However, with no extra conditions, we have the following result.

Theorem 1 *Suppose (V, φ) is a PSI space, $0 \le y < 1$ and $M \ge 2$ is a positive integer for which*

$$\sum_{l=0}^{M-1} |Z\varphi(y + \frac{l}{M}, \xi)|^2 \ge c > 0 \tag{6}$$

for all ξ. Then any $f \in V$ may be written as $f(x) = \sum_p \sum_{l=0}^{M-1} f(\frac{l}{M} + p + y) S_l^y(x-p)$ with

$$S_l^y(x) = \int_0^1 \frac{Z\varphi(x,\xi)\overline{Z\varphi(\frac{l}{M} + y, \xi)}}{\sum_{n=0}^{M-1} |Z\varphi(\frac{n}{M} + y, \xi)|^2}\, d\xi \tag{7}$$

for $0 \le l \le M-1$.

Conditions which ensure the existence of a constant c for which (6) holds are easy to find. In fact, the sum in (6) may be viewed as a Riemann sum approximation to the integral $\int_0^1 |Z\varphi(x,\xi)|^2\, dx$. If $\{\varphi(x-k)\}$ is an orthonormal set, we have $\int_0^1 |Z\varphi(x,\xi)|^2\, dx = \sum_k |\hat{\varphi}(\xi + k)|^2 = 1$ for a.e. ξ. This is the key observation in the following result.

Theorem 2 *Let φ be a compactly supported Hölder continuous function of order $\alpha > 0$ such that $\{\varphi(x-k)\}_k$ is an orthonormal set. Then there exists a positive integer N such that*

$$\sum_{k=0}^{N-1} |Z_{\mathbb{C}}\varphi(\frac{k}{N}, z)|^2 \ge c > 0 \tag{8}$$

for all $z \in \mathbb{C}$.

This verifies condition (6) of Theorem 1 but still does not overcome the the non-compact support of the associated sampling functions. If φ is bandlimited, then for each fixed ξ, $|Z\varphi(x,\xi)|^2$ is a trigonometric polynomial in x, and as a consequence we have the following result.

Theorem 3 *Suppose φ is bandlimited to $[0, M]$ and $\{\varphi(x-k)\}$ is an othornormal set. Then for all $0 \le y < 1$, $\sum_{k=0}^{M-1} |Z\varphi(y + \frac{k}{M}, \xi)|^2 = M$ for all ξ.*

Theorem 4 *If (V, φ) is a PSI space and φ is bandlimited to $[0, M]$, then each $f \in V$ may be represented by $f(x) = \sum_p \sum_{l=0}^{M-1} f(\frac{l}{M} + p + y) S_l^y(x-p)$ with $S_l^y(x) = \sum_k \overline{\varphi(\frac{k}{M} + y + k)}\varphi(y+k)$.*

Remark. Note that since φ is bandlimited, so too is S_l^y and for each y, l, $\operatorname{supp}(\hat{S}_l^y) \subset \operatorname{supp}(\hat{\varphi}) = [0, M]$.

If we take a closer look at Theorem 2 we see that we can in fact construct compactly supported sampling functions. To see this, note that the condition (8) may be interpreted to mean that the polynomials $\{Z_{\mathbb{C}}(\frac{n}{M}, z)\}_{n=0}^{M-1}$ have no common zeroes. With an application of the (extended) euclidean algorithm, we can construct polynomials $P^0, \dots P^{M-1}$ of degree $\le L-2$ such that

$$\sum_{n=0}^{M-1} P^n(z) Z_{\mathbb{C}}\varphi(\frac{n}{M}, z) \equiv 1. \tag{9}$$

Restricting z to lie in the unit circle does the trick — if $f \in V$ then $Zf(\frac{n}{M}, \xi) = C(\xi) Z\varphi(\frac{n}{M}, \xi)$ and hence $\sum_{n=0}^{M-1} Zf(\frac{n}{M}, \xi) P^n(\xi) = C(\xi)$. With $P^n(z) = \sum_{j=0}^{L-2} p_j^n z^j$, f may be expressed as $f(x) = \sum_r \sum_{n=0}^{M-1} f(\frac{n}{M} + r) S_n(x-r)$ where $S_n(x) = \sum_{j=0}^{L-2} p_j^n \varphi(x+j)$ is supported on $[2-L, L]$.

III. VALIDATION OF SAMPLING SCHEMES

A. Staggered sampling

The staggered sampling scheme relies on the invertibility of the matrix M^φ defined in section II E. Suppose φ is an orthogonal scaling function supported on $[0, L]$ and the staggered samples are taken at dyadic rationals $x_j = l_j/2^J$; $0 \le j \le L-1$, $0 \le l_0 < l_1 < \cdots < l_{L-1} \le 2^J - 1$. . If M^φ were singular, there would exist $\mathbf{d} = (d_0, \dots d_{L-1}) \in \mathbb{C}^L \setminus \{0\}$ such that $\sum_{j=0}^{L-1} \varphi(\frac{l_j}{2^J} + k) d_j = 0$ for all k. Taking the Fourier series of both sides gives, by (3), $T^J(D\Phi)(\xi) = 0$ where $D(\xi) = \sum_{j=0}^{M-1} d_j e^{2\pi i l_j \xi}$, i.e., the singularity of M^φ is equivalent to the existence of the trigonometric polynomial D of degree $2^J - 1$ as above for which $T^J(D\Phi) = 0$.

We can use this criterion to check that this scheme is valid in the case where $L = 3$ and sampling is done at three of the four points $0, 1/4, 1/2, 3/4$. For the proof of this claim, the reader is referred to [9].

B. Oversampling

To validate this sampling scheme, we need to find an upper bound on integers $N \ge 2$ for which the polynomials $\{Z_{\mathbb{C}}\varphi(\frac{n}{N}, \cdot)\}_{n=0}^{N-1}$ have no common zeroes. If Φ has the property that $\Phi(z_0) = 0 \implies \Phi(z_0^2) \ne 0$ then $N = 2$ will do the trick, i.e., 1/2-integer sampling suffices. More generally, we have the following result.

Theorem 5 *Let φ be a scaling function for an MRA of $L^2(\mathbb{R})$, supported on $[0, L]$. Then the polynomials $\{Z_{\mathbb{C}}\varphi(\frac{l}{2^{L-2}}, z)\}_{l=0}^{2^{L-2}-1}$ have no common zeroes.*

Proof: Suppose z_0 is a common zero for the polynomials $\{Z_{\mathbb{C}}\varphi(\frac{l}{2^{L-2}}, z)\}_{l=0}^{2^{L-2}-1}$ and $z_1^2 = z_0$. Then by (3) for each $0 \le l \le 2^{L-2} - 1$, $0 = Z_{\mathbb{C}}\varphi(\frac{l}{2^{L-2}}, z_0) = H(z_1) Z_{\mathbb{C}}\varphi(\frac{l}{2^{L-3}}, z_1) + H(-z_1) Z_{\mathbb{C}}\varphi(\frac{l}{2^{L-3}}, -z_1)$. Replacing l by $l + 2^{L-3}$ and using the quasiperiodicity property gives $0 = Z_{\mathbb{C}}\varphi(\frac{l+2^{L-3}}{2^{L-2}}, z_0) = \frac{1}{z_1}[H(z_1) Z_{\mathbb{C}}\varphi(\frac{l}{2^{L-3}}, z_1) - H(-z_1) Z_{\mathbb{C}}\varphi(\frac{l}{2^{L-3}}, -z_1)]$. But by the QMF condition, $H(z_1)$ and $H(-z_1)$ cannot both be zero. Hence, by swapping z_1 and $-z_1$ if necessary, we have $Z_{\mathbb{C}}\varphi(\frac{l}{2^{L-3}}, z_1) = 0$ for all $0 \le l \le 2^{L-3} - 1$. In particular, we have produced another zero of $\Phi(z) = Z_{\mathbb{C}}\varphi(0, z)$: $\Phi(z_0) = \Phi(z_1) = 0$. We continue this process. If $|z_0| \ne 1$, then we produce L distinct zeroes $0, z_0, z_1, \dots, z_{L-2}$ of Φ, a polynomial of degree at most $L - 1$. If $|z_0| = 1$, there may be a repetition: $z_j = z_0$ for some $0 < j < L - 2$. Suppose that for each $0 < m \le j$ we have $Z_{\mathbb{C}}\varphi(\frac{m}{2^{L-3}}, -z_1) \ne 0$. Then $H(-z_m) = 0$ and again by the QMF condition we have $|H(z_m)| = 1$. Consequently we have a non-trivial cycle $z_j, z_{j-1} = z_j^2, z_{j-2} = z_j^4, \dots z_0 = z_j^{2^j} = z_j$ for the action $\tau : z \to z^2$ on the circle with $|H(z_{j-1})| = \cdots = |H(z_0)| = 1$, thus contradicting the τ-cycle condition. We conclude that $Z_{\mathbb{C}}\varphi(\frac{l}{2^{L-m-2}}, -z_m) = 0$ for at least one $0 \le m < j$ and all l. We now begin a new string of zeroes $w_m = -z_m, w_{m+1}, \dots$. It is easy to show that the w's are distinct from the z's. Either the zeroes $z_0, z_1, \dots, z_{m-1}, w_m = -z_m, w_{m+1}, \dots, w_{L-2}$ are distinct, or this string will branch as before. Continuing this way we are assured of finding a branch of the tree containing $L - 1$ distinct zeroes. Hence Φ has L zeroes ($L - 1$ from the tree and 0), yet is of degree no greater than $L - 1$. We conclude that the polynomials $\{Z_{\mathbb{C}}\varphi(\frac{l}{2^{L-2}}, z)\}_{l=0}^{2^{L-2}-1}$ are co-prime. $\square$

As a corollary to the theorem, we have that sampling at a rate of 2^{L-2} always produces a valid oversampling scheme, e.g., for sampling functions such as D_4 which are supported on $[0, 3]$, $\frac{1}{2}$-integer sampling is effective. In general, sampling rates much lower than that given in the theorem are achievable, but the theorem suggests that by calculating the Zak transfoms $Z_{\mathbb{C}}\varphi(0, z)$, $Z_{\mathbb{C}}\varphi(1/2, z)$, $Z_{\mathbb{C}}\varphi(1/4, z)$, $Z_{\mathbb{C}}\varphi(3/4, z), \dots$ eventually we will produce polynomials without common zeroes.

IV. ALIASING

In [4], a very simple way of thinking about aliasing in general multiresolution spaces is introduced. Janssen's sampling operator $\mathcal{S}^y f(x) = \sum_k f(k) S^y(x - k)$ is a projection onto V_0 but, thought of as an operator from $C_0(\mathbb{R}) \cap L^2(\mathbb{R})$ to V_0, it will not be an orthogonal projection. So one has to be concerned about aliasing effects.

As a measure of aliasing error, Janssen proposed looking at the size of $\mathcal{S}^y f$ when $f \in W_0$ (the orthogonal complement of V_0 in V_1).

Definition 1 *The aliasing norm $\mathcal{N}(\mathcal{S}^y)$ of the sampling operator $\mathcal{S}^y : V_1 \to V_0$ is defined as*

$$\mathcal{N}(\mathcal{S}^y) = \sup\{\|\mathcal{S}^y f\|_{V_0} / \|f\|_{W_0} : f \in W_0\}.$$

If the aliasing norm is large then aliasing effects are very strong. Usually this is an undesirable feature in signal analysis. Consequently one seeks ways in which to minimise or at least control the effects of aliasing. Janssen [4] showed that $\mathcal{N}(\mathcal{S}^y) = \max_\xi |\frac{Z\psi(y,\xi)}{Z\varphi(y,\xi)}|$. Minimising the effect of aliasing then amounts to finding y such that $\mathcal{N}(\mathcal{S}^y)$ is minimised and Janssen showed that in the case of the Daubechies D_4 scaling function the sampling norm could in fact be made quite small for $y \approx 0.37$.

A. Aliasing with staggered samples

As discussed above, in the case of regular critical sampling, the sampling function cannot have compact support but it is possible to get good control of aliasing. In Djokovic and Vaidyanathan's staggered sampling scheme one has average critical sampling in which the sampling functions can have compact support because the relationship between sample values and multiresolution coefficients is completely local. In fact, one should suspect that this localisation might lead to instability from the point of view of aliasing errors and, indeed, this appears to be the case. To make matters more precise we need to extend Definition 1 to more general sampling operators.

Definition 2 *Any continuous operator* $\mathcal{S} : V_1 \to V_0$ *of the form* $f \mapsto \sum_{m=0}^{M-1} \sum_{k\in\mathbb{Z}} f(x_{mk}) S_m(x - y_{mk})$ *such that* $\mathcal{S}|_{V_0} = I|_{V_0}$ *is called a sampling operator. For any such operator* $\mathcal{S}$ *we define the aliasing norm* $\mathcal{N}(\mathcal{S})$ *of* $\mathcal{S}$ *to be* $\mathcal{N}(\mathcal{S}) = \sup\{\|\mathcal{S}f\|_{V_0}/\|f\|_{W_0} : f \in W_0\}$.

Lemma 1 *Let* $\mathcal{S} : V_1 \to V_0$ *be the staggered sampling operator* $f \mapsto \sum_k \sum_{j=0}^{L-1} f(x_j + kL) S_j(x - kL)$ *where* $0 \le x_j < 1$ *and* $\{S_j\}_{j=0}^{L-1}$ *are the sampling functions given in (5). Then the aliasing norm of* $\mathcal{S}$ *is given by* $\mathcal{N}(\mathcal{S}) = \|M^{\psi}(M^{\varphi})^{-1}\|$ *where* $M^{\psi}_{jl} = \psi(x_j + l)$.

B. Controlling aliasing by oversampling

Classical sampling of bandlimited functions provides ample reason for expecting that oversampling can lead to control of aliasing. In the present context we are trying to define a projection operator onto V_0 in which the projection is defined in terms of samples taken, at least on average, at a rate higher than one per unit interval. We will specifically focus on sampling projections of the form $\mathcal{S}f(x) = \sum_{k\in\mathbb{Z}} \sum_{j=0}^{M-1} f(x_j + k) S_j(x-k)$ where the x_j are distinct numbers in $[0, 1)$.

The sampling functions are finite linear combinations of translates of the scaling function and the number of translates needed will typically depend on the sampling rate: the higher the sampling rate the fewer translates needed. We will show how the added degrees of freedom provided by oversampling allow one to control aliasing.

Suppose that φ is a scaling function supported in $[0, L]$ and let $0 \le x_0 < x_1 < \ldots < x_{M-1} < L$. In the notation of section II F, when $M(D+1) > L + D$ the coefficients of the polynomials P^j are typically underdetermined. This leaves open the possibility that, in addition to reproducing V_0, the sampling operator might also annihilate part or all of W_0, thereby eliminating aliasing. To this end, in addition to satisfying condition (9) we also seek to satisfy

$$\sum_{j=0}^{M-1} Z_{\mathbb{C}}\psi(x_j, z) P^j(z) \equiv 0. \tag{10}$$

The effect of this condition on aliasing is given by the following lemma.

Lemma 2 *Let* φ *be a compactly supported orthgonal scaling function with wavelet* ψ*, and polynomials* $P^j(z)$ *satisfying (9) and (10). If* $\mathcal{S}$ *is the sampling operator* $\mathcal{S} : f \mapsto \sum_k \sum_{j=0}^{M-1} f(x_j + k) S_j(x - k)$ *where* $S_j(x) = \sum_l p^j_l \varphi(x - l)$*, then the aliasing norm* $\mathcal{N}(\mathcal{S}) = 0$*, i.e.,* W_0 *lies in the kernel of* $\mathcal{S}$.

Verifying (10) is just another nondegeneracy condition. In general it requires more oversampling. If the M polynomials P^j each have degree D then we now have $M(D+1)$ unknown coefficients for which to solve but now we have $L + L' + 2D$ equations, when we assume that ψ is supported on the interval $[0, L']$. So $M(D+1) \ge L + L' + 2D$ is required for any hope of a solution, together with the rank condition. In the case of the D_4 scaling function and wavelet this says that we need $M \ge 6$ samples per interval to eliminate aliasing when $D = 0$ and $M \ge 4$ samples per unit interval when $D = 1$. If one uses polynomials of degree four it is possible to use only three samples per interval. This scheme cannot eliminate aliasing if one samples twice per unit interval and uses sampling functions with compact support. It is not yet known whether aliasing can be eliminated if one samples at twice the critical rate with non-compactly supported sampling functions.

For proofs of the results of this section and a numerical investigation of aliasing norms of critical sampling, staggered sampling and oversampling operators in the case of scaling functions supported on $[0, 3]$, the reader is referred to [9].

V. Non-translation-invariance

A. Discrepancy

As is shown in [10], the only MRA (V_j, φ) which is translation-invariant (in the sense that $f \in V_0 \Rightarrow \tau_\alpha f = f(\cdot - \alpha) \in V_0$ for all α) is the Shannon MRA. In this section we use the Zak transform to investigate the degree of non-translation-invariance of MRA spaces and the difficulties it causes for sampling. We provide feasible schemes for overcoming these difficulties through oversampling.

For each $\alpha \in \mathbb{R}$ and $f \in V_0$, we measure the degree to which the energy of f leaks out of V_0 when translated by α by the quantity $\|\tau_\alpha f - P_{V_0}\tau_\alpha f\|_2^2/\|f\|_2^2$. (Here P_{V_0} is the orthogonal projection onto V_0.) Hence, we define $d_\varphi(\alpha) = \sup_{f\in V_0,\ \|f\|_2=1} \|\tau_\alpha f - P_{V_0}\tau_\alpha f\|_2^2$ as the *translation discrepancy* of φ with respect to α. It is natural then the consider quantities such as

$$d_\varphi = \sup_{0\le\alpha<1} d_\varphi(\alpha) \text{ and } \overline{d}_\varphi = \int_0^1 d_\varphi(\alpha)\, d\alpha$$

as measures of the non-translation-invariance of (V_0, φ). For the purposes of this paper, we'll consider the maximum discrepancy d_φ only.

The discrepancy is significant in the analysis of signals via the wavelet spectrogram. The energy of a signal f in the wavelet space W_j is $E_j(f) = \|P_{W_j} f\|_2^2$. Then the (discrete) wavelet spectrogram of f is the sequence $E(f) = \{E_j(f)\}_{j=-\infty}^{\infty}$. We define the cumulant $F_J(f) = \sum_{j=-\infty}^{J-1} E_j(f) = \|P_{V_J}\|_2^2$ $(j \in \mathbb{Z})$.

A high value of $d_\varphi(\alpha)$ means that there is a signal $f \in V_0 = \oplus_{j=-\infty}^{-1} W_j$ (thought of as a space of lower order details) such that $\tau_\alpha f$ has most of its energy in

$V_0^\perp$ (thought of as the space of higher order details). In fact, we have the following result.

Theorem 6 *Let* φ *be a scaling function for an MRA of* $L^2(\mathbb{R})$ *with continuous QM filter* H*. Then* $d_\varphi(\frac{1}{2}) = 1$*, i.e., given* $\varepsilon > 0$ *there exists* $f \in V_0$ *such that* $\|P_{V_0}\tau_{\frac{1}{2}}f\|_2^2 < \varepsilon$*.*

B. Determining the shift

The basic problem of critical sampling in an MRA space is the following: given incoming data $a_k = f(\alpha + k)$ $(k \in \mathbb{Z})$ consisting of shifted integer samples of $f \in V_0$, how do we reconstruct f? In the case of the Shannon MRA which is translation-invariant, we have $\sum_k a_k\varphi(x-k) = \sum_k f(k+\alpha)\varphi(x-k) = f(x+\alpha) \in V_0$. Even without prior knowledge of the shift α, the left-hand-side of this equation gives a signal in V_0 which interpolates the data. As we have seen, this can only be done in the Shannon case. For other MRA's, the schemes we have developed and validated require knowledge of the shift parameter α. Oversampling can be used to determine the shift and hence initialise our sampling procedures.

We now give an algorithm for determining the shift from oversampled data. The assumptions we make on φ is that it be compactly supported and for all $0 \le \alpha < \frac{1}{2}$, the polynomials $Z_{\mathbb{C}}\varphi(\alpha, z)$ and $Z_{\mathbb{C}}\varphi(\alpha + \frac{1}{2}, z)$ are co-prime.

Let $\{a_k = f(\alpha + \frac{k}{2})\}_k$ be twice oversampled data arising from samples of $f \in V_0$. Split the data into even and odd components: $a_k^e = a_{2k}$; $a_k^o = a_{2k+1}$, and form the Fourier series of these sequences:

$$A^e(\xi) = C(\xi)Z\varphi(\alpha, \xi); \quad A^o(\xi) = C(\xi)Z\varphi(\alpha + \frac{1}{2}, \xi).$$

With S^β the critical sampling operator described in section II D, we define an operator $\mathcal{T}^\beta : l^2(\mathbb{Z}) \to V_0$ by $\mathcal{T}^\beta d(x) = \sum_k d_k S^\beta(x-k)$. Then

$$\begin{aligned}
&\sum_m |\mathcal{T}^\beta a^e(m + \frac{1}{2} + \beta) - a^o(m)|^2 \\
&\quad = \int_0^1 |A^e(\xi) Z S^\beta(\beta + \frac{1}{2}, \xi) - A^o(\xi)|^2 \, d\xi \\
&\quad = \int_0^1 \frac{|C(\xi)|^2}{|Z\varphi(\beta, \xi)|^2} |Z\varphi(\alpha, \xi) Z\varphi(\beta + \frac{1}{2}, \xi) \\
&\qquad\qquad - Z\varphi(\alpha + \frac{1}{2}, \xi) Z\varphi(\beta, \xi)|^2 \, d\xi
\end{aligned}$$

and we conclude that $\mathcal{T}^\beta a^e(m + \frac{1}{2} + \beta) = a^0(m)$ for all m if and only if

$$Z\varphi(\alpha, \xi) Z\varphi(\beta + \frac{1}{2}, \xi) = Z\varphi(\alpha + \frac{1}{2}, \xi) Z\varphi(\beta, \xi)$$

for all ξ. Since the polynomials $Z_{\mathbb{C}}\varphi(\alpha, z)$ and $Z_{\mathbb{C}}\varphi(\alpha + \frac{1}{2}, z)$ are co-prime, we see that $Z\varphi(\alpha, \xi) = Z\varphi(\beta, \xi)$ for all ξ, i.e, $\varphi(\alpha + k) = \varphi(\beta + k)$ for all integers k. Since φ satisfies a dilation equation, we also have $\varphi(\frac{\alpha+k}{2^J}) = \varphi(\frac{\beta+k}{2^J})$ for all $J, k \in \mathbb{Z}$. Now if φ is continuously differentiable, this implies that for each J, k there exists $x_{J,k} \in [\frac{k}{2^J}, \frac{k+1}{2^J}]$ such that $\varphi'(x_{Jk}) = 0$ and hence that $\varphi \equiv 0$. Hence, the shift α is the unique β for which

$$\begin{aligned}
D(\beta) &= \int_0^1 |A^e(\xi) Z\varphi(\beta + \frac{1}{2}, \xi) \\
&\qquad - A^o(\xi) Z\varphi(\beta, \xi)|^2 \, d\xi \\
&= \sum_m |\sum_k [a_{2k}\varphi(\beta + \frac{1}{2} + m - k) \\
&\qquad - a_{2k+1}\varphi(\beta + m - k)]|^2 = 0.
\end{aligned}$$

In practice, we choose the β which minimises this quantity.

REFERENCES

[1] M. Unser and J. Zerubia, "A generalized sampling theory without band-limiting constraints," *IEEE Trans. Circuits and Systems II: Analog and Digital Sig. Proc.*, **45** pp. 959–969, 1998.

[2] A. J. E. M. Janssen, "The Zak transform: a signal transform for sampled time-continuous signals," *Philips J. Res.* **39**, pp. 23–69, 1988.

[3] I. Daubechies, "Ten Lectures on Wavelets", CBMS–NSF Reg. Conf. Ser. Applied Math. **61**, SIAM, Philadelphia, 1992.

[4] A. J. E. M. Janssen, "The Zak transform and sampling theorems for wavelet subspaces," *IEEE Trans. Sig. Proc.*, **41**, pp. 3360–3364, 1993.

[5] J. E. Gilbert, J. A. Hogan and J. D. Lakey, "Wavelet subspaces for sampling and extrapolation," in *Proc. SampTA*, pp. 273–278, 1997.

[6] G. Walter, "A sampling theorem for wavelet subspaces," *IEEE Trans. Info. Th.*, **38**, pp. 881–884, 1992.

[7] X.-G. Xia and Z. Zhang, "On sampling theorem, wavelets, and wavelet transforms," *IEEE Trans. Sig. Proc.* **41**, pp. 3524–3535, 1993.

[8] I. Djokovic and P.P. Vaidyanathan, "Generalized sampling theorems in multiresolution subspaces," *IEEE Trans. Sig. Proc.* **45**, pp. 583–599, 1997.

[9] J. A. Hogan and J. Lakey, "Sampling for shift-invariant and wavelet subspaces," in *Wavelet Applications in Signal and Image Processing VIII*, Akram Aldroubi, A. F. Laine and M. Unser, Editors, Proc. SPIE **4119**, pp. 36–47, 2000.

[10] A. J. Bastys, "Orthogonal and bi-orthogonal scaling functions with good translation invariance characteristic," in *Proc. SampTA*, pp. 239–244, 1997.