A consistent and stable approach to generalized sampling

Clarice Poon*
Department of Applied Mathematics and Theoretical Physics,
University of Cambridge

March 31, 2014

Abstract

We consider the problem of generalized sampling, in which one seeks to obtain reconstructions in arbitrary finite dimensional spaces from a finite number of samples taken with respect to an arbitrary orthonormal basis. Typical approaches to this problem consider solutions obtained via the consistent reconstruction technique or as solutions of an overcomplete linear systems. However, the consistent reconstruction technique is known to be non-convergent and ill-conditioned in important cases, such as the recovery of wavelet coefficients from Fourier samples, and whilst the latter approach presents solutions which are convergent and well-conditioned when the system is sufficiently overcomplete, the solution becomes inconsistent with the original measurements.

In this paper, we consider generalized sampling via a non-linear minimization problem and prove that the minimizers present solutions which are convergent, stable and consistent with the original measurements. We also provide analysis in the case of recovering wavelets coefficients from Fourier samples. We show that for compactly supported wavelets of sufficient smoothness, there is a linear relationship between the number of wavelet coefficients which can be accurately recovered and the number of Fourier samples available.

Keywords: Sampling theory, Generalized sampling, Wavelets, Fourier series, Hilbert space
2010 MSC: 94A20, 42C40, 65T60, 41A65, 46C05

1 Introduction

One of the momentous results of modern sampling theory is the Shannon-Nyquist Sampling Theorem, which enabled bandlimited or compactly supported signals to be fully described via discrete measurements. Formally, by defining the Fourier transform of $f \in L^1(\mathbb{R})$ as

$$\hat{f}(\omega) = \int_{\mathbb{R}} f(t) e^{-i\omega t}\, \mathrm{d}t, \quad \omega \in \mathbb{R}$$

and the sinc-function as $\mathrm{sinc}(x) = \sin(x)/x$, the theorem can be stated as follows:

Let $g \in L^2(\mathbb{R})$ be such that $\mathrm{supp}(g) \subset [-T, T]$ for some $T > 0$ and consider also its Fourier transform $f = \hat{g}$. If $\epsilon \leq \frac{1}{2T}$, then

$$f(t) = \sum_{k\in\mathbb{Z}} f(2\pi k\epsilon)\mathrm{sinc}\left(\frac{t + 2\pi k\epsilon}{2\epsilon}\right)$$

with L^2 and L^∞ convergence and

$$g(t) = \epsilon \sum_{k\in\mathbb{Z}} f(2\pi k\epsilon) e^{2\pi i\epsilon k t}$$

with L^2 convergence. The quantity $\frac{1}{2T}$ is often referred to as the Nyquist criterion.

*C.M.H.S.Poon@maths.cam.ac.uk

A direct computational consequence of this is that the functions f and g as described in the theorem may be approximated by the finite collection of measurements $\{f(2\pi k\epsilon) : |k| \leq N\}$ as follows

$$f_N(t) = \sum_{|k|\leq N} f(2\pi k\epsilon)\operatorname{sinc}\left(\frac{t+2\pi k\epsilon}{2\epsilon}\right), \quad f_N \overset{L^2,L^\infty}{\longrightarrow} f$$
$$g_N(t) = \epsilon \sum_{|k|\leq N} f(2\pi k\epsilon)e^{2\pi i\epsilon kt}, \qquad g_N \overset{L^2}{\longrightarrow} g.$$

However, in many cases, such approximations are not used because the bases generated by the sinc-function or complex exponentials are considered inappropriate representation systems for the underlying signals [30]. In fact, many images and signals can be better represented in terms of a different basis (e.g. splines or wavelets) than the basis in which they are sampled (e.g. the Fourier basis). Consequently, there is much interest in generalising the Shannon-Nyquist Sampling Theorem to recover the coefficients of a signal or image in a particular basis from samples taken with respect to another basis [30].

Thus, in this paper, we will be concerned with the following problem: For some Hilbert space $\mathcal{H}$ with inner product $\langle\cdot,\cdot\rangle$, suppose we are given two orthonormal sets $\{s_j : j \in \mathbb{N}\}$ and $\{w_j : j \in \mathbb{N}\}$ such that the sampling space $\mathcal{S} = \overline{\operatorname{span}\{s_j : j \in \mathbb{N}\}}$ and the reconstruction space $\mathcal{W} = \overline{\operatorname{span}\{w_j : j \in \mathbb{N}\}}$ satisfy $\mathcal{W} \subset \mathcal{S}$. Then for an unknown $f \in \mathcal{H}$, we will seek to obtain an approximation $R(f)$ to f in the reconstruction space $\mathcal{W}_N = \operatorname{span}\{w_j : j = 1,\ldots,N\}$ from the finite set of measurements $\{\langle f, s_j\rangle : j = 1,\ldots,M\}$. This problem of obtaining reconstructions in arbitrary spaces $\mathcal{W}$ from measurements taken with respect to arbitrary sampling vectors $\{s_j : j \in \mathbb{N}\}$ is often referred to as generalized sampling.

1.1 Existing approaches to generalized sampling

The notion of generalized sampling dates back to the work of Aldroubi and Unser [31, 30], in which the framework of consistent sampling was introduced. This framework was later extended by Eldar et al [15, 16, 17]. Given samples $\hat{f}_N = \{\langle f, s_j\rangle : j = 1,\ldots,N\}$, the goal was to find $R_N(f) \in \mathcal{W}_N$ which is consistent with the original measurements, i.e.

$$\langle R_N(f), s_j\rangle = \langle f, s_j\rangle, \quad j = 1,\ldots,N.$$

Equivalently, by letting $U = (\langle w_k, s_j\rangle)_{j,k\in\mathbb{N}}$ and letting $P_{[N]}$ denote the orthogonal projection onto $\operatorname{span}\{e_j : j = 1,\ldots,N\}$ where $\{e_j\}_{j\in\mathbb{N}}$ denotes the standard canonical basis for $\ell^2(\mathbb{N})$, $R(f) = \sum_{j=1}^N \beta_j w_j$ where $\beta = (\beta_j)_{j=1}^N$ is such that $P_{[N]}UP_{[N]}\beta = \hat{f}_N$.

Although this technique has been shown to be successful for certain shift invariant spaces [8, 35, 34], there are a number of problems preventing the use of this technique for arbitrary sampling and reconstruction spaces [6, 22]. Firstly, there is no guarantee of a well defined reconstruction since $P_{[N]}UP_{[N]}$ is note necessarily invertible. Furthermore, even when $P_{[N]}UP_{[N]}$ is invertible, the resultant reconstruction need not converge and may become ill-conditioned, thus this scheme may become computationally intractable for large problem sizes. A notable example of this is in the recovery of wavelet coefficients from Fourier samples, in which the condition number of the matrix $P_{[N]}UP_{[N]}$ becomes exponentially large as N increases [6, 2].

There have been a number of attempts to resolve the non-convergence and ill-conditionedness of consistent sampling. One notable contribution is the idea that the problem of ill-conditioning can be resolved by solving an overdetermined linear problem. Given the samples $\hat{f}_M$ of f, by letting the size of the reconstruction space vary from M, the reconstruction is now taken to be $R(f) = \sum_{j=1}^N \beta_j w_j$, where β is the least squares solution to $P_{[M]}UP_{[N]}\beta = \hat{f}_M$. Equivalently, we find a reconstruction $R_M(f) \in \mathcal{W}_N$ which is consistent with f on the reduced subspace $Q_{\mathcal{S}_M}(\mathcal{W}_N)$, where $Q_{\mathcal{S}_M}$ denotes the orthogonal projection from $\mathcal{H}$ onto $\mathcal{S}_M$.

This idea was initially employed in [27] by Pruessmann et al for the purpose of recovering in voxel coefficients from Fourier samples in the case of magnetic resonance imaging (MRI) and in [23] by Hrycak and Gröchenig where they show that one can stably recover N polynomial coefficients from $\mathcal{O}(N^2)$ Fourier samples. This technique was later generalized to arbitrary spaces through a reduced consistency framework by Adcock and Hansen [6, 5, 7]. Although the property of consistency with the original samples is not preserved, by applying ideas from [20], the work of Adcock and Hansen demonstrates that one is always able to construct a convergent and well-conditioned linear scheme for arbitrary reconstruction and sampling spaces if the number of samples is appropriately chosen with respect to the size of the reconstruction space. This scaling between the number of samples and the size of the reconstruction space is known as the stable sampling rate and can be formally defined with respect to some condition number and some rate of convergence. Furthermore, an understanding of this stable sampling rate is crucial to successful implementation of this reduced consistency framework. See [1] for further details.

1.2 Main results and overview

In this paper, we will propose an alternative framework for the problem of generalized sampling which offers solutions which are both convergent and consistent with the original samples. In particular, for some underlying $f = \sum_{j=1}^{\infty} x_j w_j$, given samples $\hat{f}_M = (\langle f, s_j \rangle)_{j=1}^{M}$, we will let $R_M(f) = \sum_{j\in\mathbb{N}} \beta_j^{[M]} w_j$ where

$$\beta^{[M]} \in \operatorname{argmin} \|\eta\|_{\ell^1} \text{ subject to } P_{[M]} U\eta = \hat{f}_M \tag{1.1}$$

and show that $\left\|\beta^{[M]} - x\right\|_{\ell^1} \to 0$ as $M \to \infty$. We will also show that this reconstruction is stable in that for $f \approx g$, we have that $R_M(f) \approx R_M(g)$. This will be made precise in Section 2. The main contribution of this paper is the mathematical analysis on the error bounds of this non-linear scheme in the context of recovering wavelet coefficients from Fourier samples. We prove that for compactly supported wavelet bases of sufficiently smoothness, the number of wavelet coefficients which are accurately recovered scales linearly with the number of Fourier samples. Furthermore, the scheme perfectly recovers any function represented by its first N wavelet coefficients from $\mathcal{O}(N)$ Fourier samples. This result echoes that of [2], in which it was shown that reduced consistency sampling presents a linear correspondence between the number of wavelet coefficients that one can accurately approximate and the number of Fourier samples available. However, whilst the reduced consistency framework requires an a-priori decision on the size of the reconstruction space with respect to the number of available samples, such a consideration is not required when solving (1.1). Finally, we remark that although sparsity is not considered in this paper, the analysis here offers some insight into the recovery of sparse signals for continuous problems. We will discuss this in Section 3. The proofs of our results are presented in Sections 4 and 5.

1.3 Magnetic resonance imaging and related works

There is great interest in solving generalized sampling in the context of recovering wavelet coefficients from Fourier samples because of the connection to MRI. Mathematically, MRI can be modelled as the recovery of a function (the image) from a collection of pointwise samples of its Fourier transform. The classical approach in MRI is to approximate f by a direct application of the Shannon-Nyquist Sampling Theorem. However, this approach has a number of drawbacks, including slow convergence and artefacts at edges due to the Gibbs phenomenon. On the other hand, the development of powerful wavelet methods for image processing [32, 33] has lead to substantial research in recovering wavelet coefficients directly [18, 26]. Hence, solutions to generalized sampling in this context should be of great relevance, as it recovers the wavelet coefficients via post-processing without modifying the acquisition process.

In fact, the non-linear approach proposed in this paper has already been applied to the problem of wavelet reconstructions from Fourier samples in MRI by Guerquin-Kern et al [19]. Numerical evidence and theoretical analysis of fast algorithms therein demonstrate that such an approach is a practical solution to the MRI problem. Although there is theoretical work on how to compute the reconstructions in [19], we are unaware of any analysis on the convergence to the true image. Hence, our theory can be seen as initial justification for the use of the non-linear scheme proposed in [19] for MRI.

Finally, we remark that via generalized sampling, we analyse the number of wavelet coefficients that can be recovered by finitely many consecutive Fourier samples. A similar question was explored in the development of finite rates of innovation (FRI) [36, 14] and super resolution [9], under it was shown that N Diracs can be perfectly recovered from CN consecutive Fourier samples for some constant C (the theory under FRI gives $C = 2$). Thus, the linear relationship between sufficiently smooth wavelets and Fourier samples in our main result can be seen as a parallel to these types of results.

1.4 Notation

Let $\mathcal{H}$ be a separable Hilbert space with inner product $\langle \cdot, \cdot \rangle$, norm $\|\cdot\|$ and let $\mathcal{B}(\mathcal{H})$ denote the set of bounded linear operators on $\mathcal{H}$. Suppose that

$$\mathcal{S} = \overline{\operatorname{span}} \{s_j : j \in \mathbb{N}\}, \quad \mathcal{W} = \overline{\operatorname{span}} \{w_j : j \in \mathbb{N}\}$$

are closed subspaces of $\mathcal{H}$ where $\{s_j\}_{j\in\mathbb{N}}$ and $\{w_j\}_{j\in\mathbb{N}}$ are orthonormal sets. Furthermore, we will assume that $\mathcal{W} \subseteq \mathcal{S}$. We also define the finite dimensional spaces

$$\mathcal{S}_M = \operatorname{span} \{s_j : j = 1, \dots, M\}$$

and

$$\mathcal{W}_N = \operatorname{span} \{w_j : j = 1, \dots, N\}.$$

Given any subspace $\mathcal{Y} \subseteq \mathcal{H}$, let the operator $Q_{\mathcal{Y}}$ denote the orthogonal projection onto $\mathcal{Y}$.

We recall here some notation relating to $\ell^2(\mathbb{N})$. Let $\{e_j\}_{j\in\mathbb{N}}$ denote the standard canonical basis in $\ell^2(\mathbb{N})$, and let $[M] := \{1, \dots, M\}$. Given any $x = (x_j) \in \ell^2(\mathbb{N})$, $\operatorname{sgn}(x)$ is the vector whose j^{th} element is $x_j/|x_j|$ if $x_j \neq 0$ and zero otherwise. Given any $\Omega \subset \mathbb{N}$, let P_Ω denote the orthogonal projection onto $\overline{\operatorname{span}} \{e_j : j \in \Omega\}$. Finally, we often refer to the measurement matrix generated by $\mathcal{W}$ and $\mathcal{S}$ as the infinite dimensional matrix $U = (u_{ij})_{i,j\in\mathbb{N}}$, with entries $u_{ij} = \langle w_j, s_i \rangle$ for $i, j \in \mathbb{N}$.

2 A stable and consistent scheme

Suppose we seek to reconstruct $f \in \mathcal{W}$ such that $f = \sum_{j\in\mathbb{N}} x_j w_j$. Letting $x = (x_j)_{j\in\mathbb{N}}$, the measurements can be written as $\hat{f}_M = P_{[M]} U x$. Consider the following non-linear problem

$$\inf_{\eta \in \mathcal{H}} \|\eta\|_{\ell^1} \text{ subject to } P_{[M]} U \eta = P_{[M]} U x \tag{2.1}$$

where U is as defined in Section 1.4. Any solution to this problem will naturally be consistent with the original measurements $\hat{f}_M$. It remains to ascertain whether the solution is convergent to f as M increases.

When considering generalized sampling as a linear reconstruction problem (as considered by consistent sampling and reduced consistency sampling), it was natural to use the $\|\cdot\|_{\ell^2}$ to establish error and stability estimates. However, in solving (2.1), we will instead consider convergence and stability using $\|\cdot\|_{\ell^1}$. We now define stability (see also, [4]).

Definition 2.1. *Let Ω, Δ be finite subsets of $\mathbb{N}$, $U \in \mathcal{B}(\mathcal{H})$. If $\xi \in \mathcal{H}$, $\operatorname{supp}(\xi) = \Delta$ is the unique minimizer of*

$$\inf \{\|\eta\|_{\ell^1} : P_\Omega U \eta = P_\Omega U \xi\},$$

and for any $\delta > 0$ and $\zeta \in \mathcal{H}$ such that $\|\zeta - \xi\|_{\ell^1} \leq \delta$, we have that any solution x to

$$\inf \{\|\eta\|_{\ell^1} : P_\Omega U \eta = P_\Omega U \zeta\},$$

satisfies

$$\|x - \xi\|_{\ell^1} \leq C \cdot \delta$$

for some constant C, then $\{U, \Omega, \Delta\}$ is said to be ℓ^1 stable at ξ. Moreover, if this holds for all $\xi \in \mathcal{H}$, then $\{U, \Omega, \Delta\}$ is said to be globally ℓ^1 stable.

For $p \in [1, \infty)$, it is natural to define (e.g. see [29]) the ℓ^p absolute condition number of a mapping $F : \ell^p(\mathbb{N}) \to \ell^p(\mathbb{N})$ as

$$\kappa_p(F) = \sup_{\hat{f} \in \ell^p} \lim_{\epsilon \to 0+} \sup_{\substack{\hat{g} \in \ell^p \\ 0 < \|\hat{g}\|_{\ell^p} \leq \epsilon}} \left\{ \frac{\left\| F(\hat{f} + \hat{g}) - F(\hat{f}) \right\|_{\ell^p}}{\|\hat{g}\|_{\ell^p}} \right\}. \tag{2.2}$$

So, if we consider Definition 2.1 and define a mapping $G : \ell^1(\mathbb{N}) \to \ell^1(\mathbb{N})$ such that for $\xi \in P_\Delta(\ell^1(\mathbb{N}))$,

$$G(\xi) \in \operatorname{argmin}\left\{\|\xi\|_{\ell^1} : P_\Omega U\eta = P_\Omega U\xi\right\},$$

then ℓ^1 stability at ξ for which $G(\xi) = \xi$ implies that

$$\frac{\|G(\xi) - G(\xi+\eta)\|_{\ell^1}}{\|\eta\|_{\ell^1}} \leq C$$

for all $\eta \in l^1(\mathbb{N})$. Thus, the requirement of ℓ^1 stability is simply the requirement that $\kappa_1(G)$ exists and so Definition 2.1 is related to the standard notion of an absolute condition number.

We remark that well-conditionedness here relates to that of solving an ℓ^1 minimization problem, rather than the well-conditionedness of the reconstruction $R_M(f)$ from M samples $\{\langle f, s_j\rangle : j = 1, \ldots, M\}$. In particular, this stability does not encapsulate robustness to noisy measurements, since consistent reconstructions are only desirable in the absence of noise. We will show that our nonlinear scheme is stable in the sense of Definition 2.1 and consequently, for any $f \in \mathcal{W}$, $\|R_M(f) - f\|_{\mathcal{H}} \to 0$ as $N \to \infty$.

2.1 Computability of (2.1)

The optimization problem (2.1) is infinite-dimensional, and in practice, one would instead solve the following finite dimensional problem:

$$\inf_{\eta \in \mathbb{C}^k} \|\eta\|_{\ell^1} \text{ subject to } P_{[M]}U\eta = P_{[M]}Ux \tag{2.3}$$

for some suitable $k \in \mathbb{N}$ and for sufficiently large k, minimizers of (2.3) will approximate minimizers of (2.1). Indeed, it is proved in [4, Proposition 7.4] that given any $\varepsilon > 0$, there exists $K \in \mathbb{N}$ such that for all $k \geq K$, we have that $\|\xi_k - \zeta_k\|_{\ell^1} < \varepsilon$ where ξ_k and ζ_k satisfy the following:

$$\begin{aligned}
\|\xi_k\|_{\ell^1} &= \inf_{\eta \in \mathbb{C}^k} \|\eta\|_{\ell^1} \text{ subject to } P_{[M]}UP_{[k]}\eta = P_{[M]}Ux, \\
\|\zeta_k\|_{\ell^1} &= \inf_{\eta \in \ell^1(\mathbb{N})} \|\eta\|_{\ell^1} \text{ subject to } P_{[M]}U\eta = P_{[M]}Ux.
\end{aligned}$$

2.2 Main Results

Our first result concerns the stability and convergence of solutions to (2.1). We show that given M samples, there always exists some N such that the solution is an accurate reconstruction of the true signal up to its first N reconstruction coefficients. Thus, although the rate of convergence will depend on the scaling between N and M, the scheme (2.1) will always yield convergent solutions.

Theorem 2.2. *Let U, $\mathcal{W}$, $\mathcal{S}$ be defined as in Section 1.4. Let $x \in \ell^1(\mathbb{N})$. For each $N \in \mathbb{N}$, there exists $m_0 \in \mathbb{N}$, such that for all $M \geq m_0$, if ξ solves (2.1), then*

$$\|\xi - x\|_{\ell^1} \leq C \cdot \left\|P_{[N]}^{\perp} x\right\|_{\ell^1}$$

for some constant C, which is independent of N and x. Hence, given any $f \in \mathcal{W}$ such that $f = \sum_{j=1}^{\infty} x_j w_j$, if ξ solves (2.1) with $x = (x_j)_{j\in\mathbb{N}}$, then

$$\left\| f - \sum_{j=1}^{\infty} \xi_j w_j \right\|_{\mathcal{H}} = \|\xi - x\|_{\ell^2} \leq C \cdot \sum_{j=N+1}^{\infty} |x_j|.$$

Theorem 2.2 shows that for given any $f = \sum_{j\in\mathbb{N}} x_j w_j \in \mathcal{W}$ and any reconstruction resolution N, there always exist M such that (2.1) can obtain an approximation f_M of f with an error of $\mathcal{O}\left(\left\|P_{[N]}^{\perp} x\right\|_{\ell^1}\right)$. So, f_M will converge to f as $M \to \infty$.

2.3 The recovery of wavelet coefficients from Fourier samples

We now let $\mathcal{H} = L^2(\mathbb{R})$ and apply the non-linear scheme (2.1) to the case where the reconstruction space $\mathcal{W}$ is generated by compactly supported orthonormal wavelets and the sampling space is the space of complex exponentials $\mathcal{S} = \overline{\text{span}}\left\{e^{2\pi i\epsilon j\cdot} : j \in \mathbb{Z}\right\}$ for some appropriate $\epsilon > 0$. In particular, for some given $a \geq 1$, we will be concerned with the recovery of wavelet coefficients of elements of $\mathcal{H}$ on the interval $[0, a]$. Our main conclusion is that for sufficiently smooth wavelet bases, the number of reconstruction vectors which can be accurately approximated is linearly proportional to the number of Fourier samples. Consequently, if it is known that the wavelet coefficients x of f has decay $\mathcal{O}\left(\left\|P_N^\perp x\right\|_{\ell^1}\right) = \mathcal{O}\left(N^{-\beta}\right)$ for some $\beta > 0$, then access to N Fourier measurements will yield an approximation with error decay $\mathcal{O}\left(N^{-\beta}\right)$ and acquiring Fourier samples is up to a constant as good as acquiring the wavelet coefficients directly. Before stating our result, we first define the wavelet and Fourier spaces to be considered.

The wavelet reconstruction space

Suppose that we are given an orthonormal mother wavelet ψ and an orthonormal scaling function ϕ such that $\text{supp}(\psi) = \text{supp}(\phi) = [0, a]$ for some $a \geq 1$. We also assume that for some $\alpha \geq 1$ and $C > 0$,

$$\left|\hat{\phi}(\xi)\right| \leq \frac{C}{(1 + |\xi|)^\alpha}, \quad \left|\hat{\psi}(\xi)\right| \leq \frac{C}{(1 + |\xi|)^\alpha}. \tag{2.4}$$

Now, consider the following collection of functions

$$\Omega_a = \{\phi_k, \psi_{j,k} : \text{supp}(\phi_k)^o \cap [0, a] \neq \emptyset,\ \text{supp}(\psi_{j,k})^o \cap [0, a] \neq \emptyset, j \in \mathbb{Z}_+, k \in \mathbb{Z}, \},$$

where

$$\phi_k = \phi(\cdot - k), \qquad \psi_{j,k} = 2^{\frac{j}{2}}\psi(2^j \cdot - k).$$

(the notation K^o denotes the interior of a set $K \subseteq \mathbb{R}$). Setting

$$\mathcal{W} := \overline{\text{span}}\{\varphi : \varphi \in \Omega_a\},$$

this gives

$$\mathrm{L}^2[0, a] \subseteq \mathcal{W} \subseteq \mathrm{L}^2[-T_1, T_2],$$

where $T_1, T_2 > 0$ are such that $[-T_1, T_2]$ contains the support of all functions in Ω_a. Note that the inclusions may be proper (but not always, as is the case with the Haar wavelet.) It is straightforward that

$$\Omega_a = \{\phi_k : |k| = 0, \ldots, \lceil a \rceil - 1\} \cup \{\psi_{j,k} : j \in \mathbb{Z}_+, k \in \mathbb{Z}, -\lceil a \rceil < k < 2^j \lceil a \rceil\}$$

and we may let $T_1 = \lceil a \rceil - 1$, $T_2 = 2\lceil a \rceil - 1$. We will order Ω_a first by translation factors then in increasing order of the scaling factor as follows:

$$\{\phi_{-\lceil a \rceil+1}, \ldots, \phi_{-1}, \phi, \phi_1, \ldots, \phi_{\lceil a \rceil-1}, \psi_{0,-\lceil a \rceil+1}, \ldots, \psi_{0,-1}, \psi_{0,0}, \psi_{0,1}, \ldots, \psi_{0,\lceil a \rceil-1},$$
$$\psi_{1,-\lceil a \rceil+1}, \ldots, \psi_{1,2\cdot\lceil a \rceil-1}, \psi_{2,-\lceil a \rceil+1}, \ldots, \psi_{2,4\cdot\lceil a \rceil-1}, \ldots\}.$$

It is often useful to consider the all elements of Ω_a of wavelet resolution less than $R \in \mathbb{N}$. To this end, we define

$$\Omega_{R,a} = \{\varphi \in \Omega_a : \varphi = \psi_{j,k}, j < R, k \in \mathbb{Z} \text{ or } \varphi = \phi_k, k \in \mathbb{Z}\}$$

and denote the size of $\Omega_{R,a}$ by N_R. It is easy to verify that

$$N_R = 2^R\lceil a \rceil + (R + 1)(\lceil a \rceil - 1) \tag{2.5}$$

and

$$\mathcal{W}_{N_R} \subset \{\phi_{R,k} : A_{R,1} \leq k \leq A_{R,2}\}$$

where

$$A_{R,1} = -(2^R + 1)\lceil a \rceil + 2^R + 1, \qquad A_{R,2} = 2^{R+1}\lceil a \rceil - 2^R - 1. \tag{2.6}$$

The Fourier sampling space

For the Fourier sampling space, we let $\epsilon \leq 1/(T_1+T_2)$ be the *sampling density*. Note that $1/(T_1+T_2)$ is the corresponding Nyquist criterion for functions supported on $[-T_1, T_2]$. We now define the sampling vectors by

$$s_l = \sqrt{\epsilon} e^{2\pi i l \epsilon \cdot} \chi_{[-T_1/(\epsilon(T_1+T_2)), T_2/(\epsilon(T_1+T_2))]},$$

the sampling space by

$$\mathcal{S} = \overline{\mathrm{span}\{s_l : l \in \mathbb{Z}\}} = \left\{ f \in L^2(\mathbb{R}) : \mathrm{supp}(f) \subseteq [-T_1/(\epsilon(T_1+T_2)), T_2/(\epsilon(T_1+T_2))] \right\},$$

and the space spanned by the first M sampling vectors by

$$\mathcal{S}_M = \mathrm{span}\left\{ s_l : -\left\lfloor \frac{M}{2} \right\rfloor \leq l \leq \left\lceil \frac{M}{2} \right\rceil - 1 \right\}.$$

Theorem 2.3. *Let U be the measurement matrix associated with the wavelet reconstruction space $\mathcal{W}$ and the Fourier sampling space $\mathcal{S}$ described above. Suppose that the Fourier sampling density $\epsilon \in \mathbb{Q}$ is such that*

$$0 < \epsilon \leq \delta/(T_1+T_2), \quad \delta \in (0,1)$$

where $T_1, T_2 > 0$ are such that $\mathcal{W} \subset L^2[-T_1, T_2]$.

Then for $x \in \ell^1(\mathbb{N})$ and $N \in \mathbb{N}$, the following holds:

(i) If for some $A > 0$ and $\alpha \geq 1$,

$$\left|\hat{\phi}(\xi)\right| \leq \frac{A}{(1+|\xi|)^\alpha}, \quad \xi \in \mathbb{R}$$

then there exists some constant C independent of N (but dependent on α and ϵ) such that for $M = C \cdot N^{1+1/(2\alpha-1)}$, any solution ξ to (2.1) satisfies

$$\|\xi - x\|_{\ell^1} \leq 6 \cdot \left\| P_{[N]}^{\perp} x \right\|_{\ell^1}.$$

(ii) If for $k = 0, 1, 2$, for some $A > 0$ and $\alpha \geq 1.5$,

$$\left|\hat{\phi}^{(k)}(\xi)\right| \leq \frac{A}{(1+|\xi|)^\alpha}, \quad \left|\hat{\psi}^{(k)}(\xi)\right| \leq \frac{A}{(1+|\xi|)^\alpha}, \quad \xi \in \mathbb{R},$$

then there exists some constant C independent of N (but dependent on ϕ, ψ and ϵ) such that for $M = C \cdot N$, any solution ξ to (2.1) satisfies

$$\|\xi - x\|_{\ell^1} \leq 6 \cdot \left\| P_{[N]}^{\perp} x \right\|_{\ell^1}.$$

Remark 2.1 Intuitively, the restriction on ϵ should be such that $\epsilon \in \mathbb{R}$, and $\epsilon \in (0, 1/(T_1+T_2)]$, since $1/(T_1+T_2)$ is the Nyquist rate. However, the assumptions on ϵ in this theorem are stronger, and this restriction is likely to be an artefact of the proof and does not seem to be necessary in practice.

2.4 Discussion

Recovery of Daubechies wavelet coefficients

We first remark that the assumption in (i) of Theorem 2.3 is natural for Daubechies wavelets, since we know from [12, Proposition 4.7] that there exists $\alpha_N > 0$ such that for all $N \in \mathbb{N}$ with $N \geq 2$, there exists a Daubechies-N scaling function, ϕ_N, such that

$$\left|\hat{\phi}_N(\xi)\right| \leq \frac{1}{(1+|\xi|)^{\alpha_N+1}} \tag{2.7}$$

and the same decay estimate holds for the corresponding wavelet, $\hat{\psi}_N$. Table 1 presents estimates of α_N which were derived in [13] and by direct application of Theorem 2.3, estimates on β_N for which M Fourier samples is guaranteed to recover at least $\mathcal{O}\left(M^{\beta_N}\right)$ wavelet coefficients.

N	α_N	β_N
1	0	0.5
2	0.339	0.627
3	0.636	0.694
4	0.913	0.739
5	1.177	0.770
6	1.432	0.794

Table 1: For each Daubechies-N wavelet, this table shows estimates of α_N such that (2.7) holds and β_N such that M Fourier samples is guaranteed to recover $\mathcal{O}\left(M^{\beta_N}\right)$ wavelet coefficients in the sense of Theorem 2.3.

To understand which Daubechies wavelets satisfy the assumption in (ii) of Theorem 2.3, we first note that there is a natural correspondence between the smoothness and the decay of the Fourier transform. If ϕ has α derivatives, then

$$\left|\hat{\phi}(\xi)\right| \leq \frac{A}{(1+|\xi|)^\alpha}, \quad \xi \in \mathbb{R}.$$

Furthermore, smoothness of ϕ also implies decay in the derivatives of its Fourier transform: For $k = 1, 2$, $\phi_k(x) := x^k \phi(x)$ will also possess α derivatives and $\phi_k \in L^1(\mathbb{R})$ whenever ϕ is of compact support. Thus, for $k = 0, 1, 2$

$$\left|\frac{\mathrm{d}^k}{\mathrm{d}\xi^k}\hat{\phi}(\xi)\right| = \left|\hat{\phi}_k(\xi)\right| \leq \frac{A}{(1+|\xi|)^\alpha}.$$

From [12], it is known that the first Daubechies wavelet which is twice continuously differentiable is the wavelet of 7 vanishing moment, thus our theorem implies that this linear correspondence between Fourier samples and wavelet coefficients is true for Daubechies wavelets of at least 7 vanishing moments. However, numerical results suggest that this linear relationship actually holds for all Daubechies wavelets, thus suggesting that this results is perhaps not sharp. See Section 6 for further details.

Necessity of a linear relationship

In [2], it was shown that if $N \geq N_R$ and $M = c2^R$ where $c < \epsilon^{-1}$, then

$$\inf_{g \in \mathcal{W}_N, \|g\|_{\mathcal{H}}} \|Q_{\mathcal{S}_M} g\|_{\mathcal{H}} = Ae^{-BN_R + \log N_R} \tag{2.8}$$

for some positive constants A and B. Thus, for all N, there exists some $g \in \mathcal{W}_N$ such that $\|g\|_{\mathcal{H}} = 1$, but the samples $\hat{g} = (\langle g, s_j \rangle)_{j=-\lfloor M/2 \rfloor}^{\lceil M/2 \rceil - 1}$ are of exponentially small norm as N increases. Although we do not consider robustness of noise in the measurements, from a computational viewpoint, such a relationship is undesirable. Thus, in this respect, it is perhaps necessary to have at least a linear scaling between the number of Fourier samples, and the number of wavelet coefficients which an algorithm accurately reconstructs.

3 Links with compressed sensing

The recent field of compressed sensing considers the recovery of signals with a reduced number of samples via non-linear schemes by exploiting the fact that typical signals have sparse structures in particular bases, such as wavelet bases.

Generalized sampling makes no assumption on the structure of the underlying signal, however, it reveals to us the number of consecutive sampling coefficients required to accurately recover N consecutive reconstruction coefficients in a stable manner. Such knowledge is important to the understanding of compressed sensing in the context of a separable Hilbert space as demonstrated in [4]. In formulating generalized sampling as an ℓ^1 minimization problem, the main result of this paper shows that N wavelet coefficients can be exactly recovered from CN consecutive Fourier samples for some constant C (and provided that the wavelets are twice continuously differentiable). In this section, we will consider the implications of this linear relationship for compressed sensing in the infinite dimensional setting.

The infinite dimensional framework of compressed sensing introduced in [4] aims to recover signals of the form $x = x + h$, where $\text{supp}(x) = \Delta \subset [N]$, $|\Delta| = s$ and $h \in \ell^1(\mathbb{N})$ by solving the following non-linear problem

$$\inf_{\eta \in \ell^1(\mathbb{N})} \|\eta\|_{\ell^1} \text{ subject to } P_\Omega U \eta = P_\Omega U(x+h) \tag{3.1}$$

where for some $M \in \mathbb{N}$, $\Omega \subset [M]$ is chosen in a uniformly random manner and is of cardinality $\lceil qM \rceil$ for some $q \in [0,1]$. They show that any solution ξ of (3.1) is such that $\|\xi - x\|_{\ell^2} \leq C \cdot \|h\|_{\ell^1}$ with high probability if the following holds:

(i) M and $1/q$ satisfy some *balancing property* with respect to N and s.

(ii) $m \geq C \cdot \mu \cdot s \log(N/q)$.

where $\mu = \max_{i,j \in \mathbb{N}} |u_{ij}|^2$ is known as the incoherence of U. Note that (ii) is a standard requirement of compressed sensing algorithms, and the novelty is the balancing property which we recall in Definition 3.1.

There are two decisions to be made when implementing (3.1):

(1) What should be the range of our samples? i.e. What should M be?

(2) How many samples do we require? i.e. What should q be?

We will consider these two question in the context of wavelet reconstructions from Fourier samples, where the scaling function and wavelet generating the reconstruction space are assumed to have α continuous derivatives. If the choice of M had to be such that $M >> N$, then (3.1) will be less of an attractive scheme, however, as will be explained, the results of this paper show that up to some log factors, it suffices to let M grow linearly with N.

3.1 The range of samples

In order to consider (1), we recall the definition of the balancing property:

Definition 3.1. *[4] Let $U \in \mathcal{B}(\ell^2(\mathbb{N}))$ be an isometry. Then $M \in \mathbb{N}$ and $K \geq 1$ satisfy the balancing property with respect to U, $N \in \mathbb{N}$ and $s \in \mathbb{N}$ if*

$$\|P_N U^* P_M U P_N - P_N\|_{\ell^\infty \to \ell^\infty} \leq \frac{1}{8} \left(\log_2^{1/2} \left(4\sqrt{s} K N \right) \right)^{-1}, \tag{3.2}$$

where $\|\cdot\|_{\ell^\infty \to \ell^\infty}$ is the norm on $\mathcal{B}(\ell^\infty(\mathbb{N}))$ and

$$\|P_N^\perp U^* P_M U P_N\|_{\ell^\infty \to \ell^\infty} \leq \frac{1}{8}. \tag{3.3}$$

The theory from [4] demonstrate that when implementing (3.1), the range of samples M should be chosen such that M and m/M satisfy the balancing property with respect to the measurement matrix U, N and s. Observe that m/M and s will only change the relationship between M and N by a log factor. So, this balancing property is essentially a relationship between M and N and is such that $P_M U P_N$ is close to an isometry. Moreover, it essentially encompasses conditions under which (2.1) guarantees exact recovery of N reconstruction coefficients from M samples.

The constraints of the balancing property are precisely the quantities covered by the analysis of (2.1) in the context of wavelet reconstructions from Fourier samples. Assuming that the wavelet and scaling functions are α continuously differentiable for $\alpha \geq 2$, Proposition 5.2 gives that (3.3) holds whenever $M = C \cdot N$ and Corollary 5.4 gives that (3.2) holds whenever $M = C \cdot N \cdot \left(\log_2(4N\sqrt{s}/q)\right)^{1/(4\alpha-2)}$.

3.2 The amount of subsampling

From the discussion on (1), in order to obtain reconstructions up to wavelet resolution N, up to some log factors, we are required to subsample from $M = \mathcal{O}(N)$ Fourier samples. However, from [4], the number of samples m which we need to take from $[M]$ in order to perfectly reconstruct a signal supported on $\Delta \subset [N]$ with probability exceeding $1-\epsilon$ is

$$m \geq C \cdot (\log(\epsilon^{-1}) + 1) \cdot \log\left(\frac{NM\sqrt{s}}{m}\right) \cdot \mu(U) \cdot |\Delta| \cdot M.$$

So, up to log factors, the amount of subsampling depends on the sparsity of the signal $|\Delta|$ and the coherence of the measurement matrix matrix $\mu(U) = \max_{i,j\in\mathbb{N}} |u_{ij}|^2$. This is problematic as it can be shown that $\mu = o(1)$ in this case. To mitigate this, variable density sampling or half-half schemes are implemented in practice[28, 24], in which one considers problems of the form

$$\inf_{\eta\in\mathcal{H}} \|\eta\|_{\ell^1} \text{ subject to } (P_{[M_1]} \oplus P_\Omega)U\eta = (P_{[M_1]} \oplus P_\Omega)Ux \tag{3.4}$$

where $\Omega \subseteq \{M_1+1,\ldots,M\}$ is chosen in an uniformly random manner. The success of such schemes was mathematically analysed in [3]. We recall some of the key aspects here and link it to the work presented in this paper, although the interested reader should refer to [3] for details. Despite $\mu(U)$ not being small and the underlying signal x may not be sparse, we have that $\mu(P_{[N]}^{\perp}U) \to 0$ as $N \to \infty$ and generally $\operatorname{supp}(x) \subset [N_1] \cup \Delta$, where $\Delta \subset \{N_1+1,\ldots,N\}$ is such that $|\Delta|/(N-N_1)$ is small. These properties are respectively referred to as asymptotic incoherence and asymptotic sparsity. In addition, if M_1, M are correctly chosen with respect to N_1, N, then $|\Omega|$ depends on the smaller values of $\mu(P_{[M_1]}^{\perp}U)$ and $|\Delta|$. So, the amount of subsampling can be based on the sparse part of the signal and the incoherent part of the sampling matrix. In particular, a suitable choice of M_1 with respect to N_1 is such that $\left\|P_{[M_1]}^{\perp}UP_{[N_1]}\right\|_{\ell^2} \leq \frac{\gamma}{\sqrt{M_1}}$ for sufficiently small γ. This is precisely the quantity studied in Lemma 5.1 and holds whenever $M_1 = \mathcal{O}\left(N_1^{1+1/(2\alpha-1)}\right)$. Furthermore, for more complex signal structures, [3] also generalizes (3.4) for the purpose of recovering signals such that

$$\operatorname{supp}(x) \subset \Delta_1 \cup \ldots \cup \Delta_r, \quad \Delta_k \subset \{M_{k-1}+1,\ldots,M_k\}, \quad 0 = M_0 < M_1 < \ldots < M_r$$

to multiple levels of the form

$$\inf_{\eta\in\mathcal{H}} \|\eta\|_{\ell^1} \text{ subject to } (P_{\Omega_1} \oplus \cdots \oplus P_{\Omega_r})U\eta = (P_{\Omega_1} \oplus \cdots \oplus P_{\Omega_r})Ux$$

where $0 = N_0 < N_1 < \ldots < N_r$ and for each $1 \leq k \leq r$, $\Omega_k \subset \{N_{k-1}+1,\ldots,N_k\}$ is taken uniformly at random. Similarly to the two level scheme, effective subsampling can be achieved when for $k = 1,\ldots,r-1$, the values $\left\|P_{M_k}^{\perp}UP_{N_k}\right\|_{\ell^2}$ are sufficiently small, so there is little 'interference' between different levels. Thus, Lemma 5.1 serves as a starting point in understanding the effects of the choice of $\{N_k\}_{k=1}^{r}$ on the recovery of signals.

4 Proofs

4.1 Existence of unique minimizers and stability

In order to prove Theorem 2.2, we will first present a result, now of common usage in the compressed sensing literature. A similar result was proved in [4] (see also [10]) to establish conditions for the existence of unique minimizers. However, we repeat the proof here in order to derive a statement which includes sufficient conditions for ℓ^1 stability.

Proposition 4.1. *Let $U \in \mathcal{B}(\ell^2(\mathbb{N}))$ with $\|U\|_{\ell^2} \leq 1$ and let $\Omega, \Delta \subset \mathbb{N}$ be such that $|\Omega|, |\Delta| < \infty$. Suppose that $x, h \in \ell^1(\mathbb{N})$ and $\operatorname{supp}(x) = \Delta$ and $\operatorname{supp}(h) \cap \Delta = \emptyset$. Consider the optimisation problem*

$$\inf_{\eta\in\ell^1(\mathbb{N})} \|\eta\|_{\ell^1} \text{ subject to } P_\Omega U\eta = P_\Omega U(x+h). \tag{4.1}$$

If there exists ρ such that

(i) $\rho = U^* P_\Omega \eta$ for some η.

(ii) $\langle \rho, e_j \rangle = \langle \mathrm{sgn}(x), e_j \rangle, \quad j \in \Delta.$

(iii) $|\langle \rho, e_j \rangle| \leq 1/2, \quad j \notin \Delta$

and

(iv) $(P_\Delta U^* P_\Omega U P_\Delta)^{-1}$ *exists on* $P_\Delta(\ell^1(\mathbb{N}))$ *and* $\left\| P_\Delta^\perp U^* P_\Omega U P_\Delta (P_\Delta U^* P_\Omega U P_\Delta)^{-1} \right\|_{\ell^\infty} \leq 2$*,*

then any minimizer ξ *of (4.1) satisfies*

$$\|\xi - x\|_{\ell^1} \leq 11 \|h\|_{\ell^1}.$$

Proof. First observe that since ξ satisfies the constraint of (4.1), we have that $P_\Omega U(x - P_\Delta \xi) = P_\Omega U P_\Delta^\perp(\xi - h)$. Thus,

$$\begin{aligned}
\|x - P_\Delta \xi\|_{\ell^1} &= \left\| (P_\Delta U^* P_\Omega U P_\Delta)^{-1} P_\Delta U^* P_\Omega U P_\Delta (x - \xi) \right\|_{\ell^1} \\
&= \left\| (P_\Delta U^* P_\Omega U P_\Delta)^{-1} P_\Delta U^* P_\Omega U P_\Delta^\perp (\xi - h) \right\|_{\ell^1} \\
&\leq \left\| (P_\Delta U^* P_\Omega U P_\Delta)^{-1} P_\Delta U^* P_\Omega U P_\Delta^\perp \right\|_{\ell^1} \left\| P_\Delta^\perp (\xi - h) \right\|_{\ell^1} \leq 2 \left(\left\| P_\Delta^\perp \xi \right\|_{\ell^1} + \|h\|_{\ell^1} \right)
\end{aligned}$$

where the last line follows from (iv). It then follows that

$$\|x - \xi\|_{\ell^1} \leq 3 \left\| P_\Delta^\perp \xi \right\|_{\ell^1} + 2 \|h\|_{\ell^1}. \tag{4.2}$$

We will now proceed to bound $\left\| P_\Delta^\perp \xi \right\|_{\ell^1}$ in terms of $\|h\|_{\ell^1}$:

$$\begin{aligned}
\|\xi\|_{\ell^1} &= \|P_\Delta \xi - x + x\|_{\ell^1} + \left\| P_\Delta^\perp \xi \right\|_{\ell^1} \geq \|x\|_{\ell^1} + \mathrm{Re} \langle P_\Delta \xi - x, \mathrm{sgn}(x) \rangle + \left\| P_\Delta^\perp \xi \right\|_{\ell^1} \\
&\geq \|x + h\|_{\ell^1} - \|h\|_{\ell^1} + \mathrm{Re} \langle P_\Delta \xi - x, \mathrm{sgn}(x) \rangle + \left\| P_\Delta^\perp \xi \right\|_{\ell^1}.
\end{aligned}$$

Since (ii) holds and $\|\xi\|_{\ell^1} \geq \|x + h\|_{\ell^1}$, it follows that

$$\begin{aligned}
\left\| P_\Delta^\perp \xi \right\|_{\ell^1} &\leq |\langle P_\Delta \xi - x, \mathrm{sgn}(x) \rangle| + \|h\|_{\ell^1} \\
&= |\langle P_\Delta \xi - x, P_\Delta \rho \rangle| + \|h\|_{\ell^1} = \left| \langle \xi - x, \rho \rangle - \langle P_\Delta^\perp \xi, \rho \rangle \right| + \|h\|_{\ell^1}.
\end{aligned}$$

Now, since $\rho = U^* P_\Omega U \eta$ and $P_\Omega U(x - \xi) = -P_\Omega U h$, we have that

$$\left\| P_\Delta^\perp \xi \right\|_{\ell^1} \leq |\langle h, \rho \rangle| + \left| \langle P_\Delta^\perp \xi, \rho \rangle \right| + \|h\|_{\ell^1} \leq \frac{3}{2} \|h\|_{\ell^1} + \frac{1}{2} \left\| P_\Delta^\perp \xi \right\|_{\ell^1}$$

where the last line follows from (iii). Finally, plugging $\left\| P_\Delta^\perp \xi \right\|_{\ell^1} \leq 3 \|h\|_{\ell^1}$ into (4.2) yields

$$\|x - \xi\|_{\ell^1} \leq 11 \|h\|_{\ell^1}.$$

□

4.2 Proof of Theorem 2.2

Proof of Theorem 2.2. Fix $N \in \mathbb{N}$. By Proposition 4.1, it is sufficient to show that for M sufficiently large,

(a) $P_{[N]} U^* P_{[M]} U P_{[N]}$ is invertible on $P_{[N]}(\ell^1(\mathbb{N}))$ such that

$$\left\| P_{[N]}^\perp U^* P_{[M]} U P_{[N]} (P_{[N]} U^* P_{[M]} U P_{[N]})^{-1} \right\|_{\ell^\infty} \leq 2,$$

(b) there exists $\rho \in \mathrm{range}(U^* P_{[M]})$ such that $P_{[N]} \rho = \mathrm{sgn}(P_{[N]} x)$ and $\left\| P_{[N]}^\perp \rho \right\|_{\ell^\infty} < 1/2$.

We know from the analysis of [6] that since U^*U is self adjoint and positive, there exist $C > 0$ and $m_0 \in \mathbb{N}$ such that for all $M \geq m_0$, $P_{[N]}U^*P_{[M]}UP_{[N]}$ is invertible on $P_{[N]}(\ell^1(\mathbb{N}))$ and $\|(P_{[N]}U^*P_{[M]}UP_{[N]})^{-1}\| \leq C$. Furthermore, for fixed $N \in \mathbb{N}$, $\left\|P_{[M]}^{\perp}UP_{[N]}\right\|_{\ell^2} \to 0$ as $M \to \infty$. Thus, using the fact that U is an isometry,

$$\begin{aligned}\left\|P_{[N]}^{\perp}U^*P_{[M]}UP_{[N]}(P_{[N]}U^*P_{[M]}UP_{[N]})^{-1}\right\|_{\ell^\infty} &= \left\|P_{[N]}^{\perp}U^*P_{[M]}^{\perp}UP_{[N]}(P_{[N]}U^*P_{[M]}UP_{[N]})^{-1}\right\|_{\ell^\infty} \\ &\leq \sqrt{N}\left\|P_{[M]}^{\perp}UP_{[N]}\right\|_{\ell^2}\left\|(P_{[N]}U^*P_{[M]}UP_{[N]})^{-1}\right\|_{\ell^2} \leq 2\end{aligned}$$

for M sufficiently large. So, (a) is satisfied.

Now,

$$\rho = U^*P_{[M]}UP_{[N]}(P_{[N]}U^*P_{[M]}UP_{[N]})^{-1}\mathrm{sgn}(P_{[N]}x)$$

is well defined. Moreover, $P_{[N]}\rho = P_{[N]}\mathrm{sgn}(x)$. For $j > N$

$$\begin{aligned}&|\langle \rho, e_j\rangle| = \left|\langle U^*P_{[M]}^{\perp}UP_{[N]}(P_{[N]}U^*P_{[M]}UP_{[N]})^{-1}\mathrm{sgn}(P_{[N]}x), e_j\rangle\right| \\ &\leq \left\|P_{[M]}^{\perp}UP_{[N]}\right\|_{\ell^2}\left\|(P_{[N]}U^*P_{[M]}UP_{[N]})^{-1}\right\|_{\ell^2}\sqrt{N} \leq C\sqrt{N}\left\|P_{[M]}^{\perp}UP_{[N]}\right\|_{\ell^2} \to 0 \quad \text{as } M \to \infty.\end{aligned}$$

So, (b) is satisfied. Therefore, by Proposition 4.1, for sufficiently large M, the optimisation problem (2.1) has a unique solution if $\mathrm{supp}(x) \subset \{1, \ldots, N\}$ and for any $x \in \mathcal{H}$,

$$\|\xi - x\|_{\ell^1} \leq 12 \cdot \left\|P_{[N]}^{\perp}x\right\|_{\ell^1}$$

□

5 Proof of Theorem 2.3

Throughout this section, we will be concerned with the reconstruction of functions compactly supported on $[0, a]$ for some $a \geq 1$ and the reconstruction space $\mathcal{W}$ and sampling space $\mathcal{S}$ will be the wavelet and Fourier spaces as defined in Section 2.3.

Before presenting the proof of Theorem 2.3, we first require some results on the relationship between the Fourier sampling space $\mathcal{S}$ and the wavelet reconstruction space $\mathcal{W}$. This will be the purpose of the next section.

5.1 Preliminary results

The following is a result from [2], we however derive more precise bounds due the additional assumption of polynomial decay on the Fourier transform of the scaling function ϕ.

Lemma 5.1. *Assume that the scaling function and wavelet, ϕ and ψ, are supported on $[0, a]$ and for some $A > 0$ and $\alpha \geq 1$,*

$$\left|\hat{\phi}(\xi)\right| \leq \frac{A}{(1 + |\xi|)^{\alpha}}, \quad \xi \in \mathbb{R}.$$

Suppose also that $\epsilon \in \mathbb{Q}$ is such that $0 < \epsilon \leq 1/(T_1 + T_2)$ where $T_1, T_2 > 0$ are such that $\mathcal{W} \subset L^2[-T_1, T_2]$.

Then, given any $\gamma \in (0, 1)$ and $N \in \mathbb{N}$, there exists some $C_{\alpha,A,\epsilon} \in \mathbb{N}$ which depends only on ϵ, α and A such that

$$\sup_{\substack{\varphi \in \mathcal{W}_N, \\ \|\varphi\|_{\mathcal{H}} \leq 1}} \left\|Q_{\mathcal{S}_M}^{\perp}\varphi\right\|_{\mathcal{H}} \leq \gamma. \tag{5.1}$$

whenever

$$M \geq C_{\alpha,A,\epsilon} \cdot N \cdot \gamma^{-2/(2\alpha - 1)}. \tag{5.2}$$

Remark 5.1 In the case where $\epsilon^{-1} \in \mathbb{N}$ and $N \leq N_R$, it suffices to replace (5.2) by

$$M \geq \epsilon^{-1} \left(\frac{4A^2}{(2\pi)^{2\alpha}(2\alpha - 1)} \right)^{\frac{1}{2\alpha-1}} 2^{R+1} \gamma^{-\frac{2}{2\alpha-1}}.$$

This dependence on ϵ is necessary because, as mentioned in (2.8), it can be shown that $\inf_{\varphi \in \mathcal{T}_N} \|Q_{\mathcal{S}_M} \varphi\|_{\mathcal{H}}$ becomes exponentially small as M increases if $M = c2^R$ with $c < \epsilon^{-1}$. However, when $\epsilon^{-1} \in \mathbb{Q}$, the dependence of C_ϵ on ϵ which arises out of the proof is more complex and this is likely to be an artefact of the proof technique.

Proof of Lemma 5.1. Recalling the definition of N_R from (2.5), we first choose $R \in \mathbb{N}$ such that $N_{R-1} < N \leq N_R$, so $N = \mathcal{O}\left(2^R\right)$. Also, since $\epsilon \in \mathbb{Q}$, we can choose $C_\epsilon \in \mathbb{N}$ such that $\epsilon^{-1} C_\epsilon \in \mathbb{N}$. In particular, if $\epsilon^{-1} \in \mathbb{N}$, then we can let $C_\epsilon = 1$. Let $M = \epsilon^{-1} C_\epsilon \cdot S \cdot 2^{R+1}$, for some $S \in \mathbb{N}$. The goal is to determine S such that

$$\sup_{\substack{\varphi \in \mathcal{W}_N, \\ \|\varphi\|_{\mathcal{H}} \leq 1}} \left\| Q_{\mathcal{S}_M}^{\perp} \varphi \right\|_{\mathcal{H}} \leq \gamma.$$

Let $\varphi \in \mathcal{W}_N$ such that $\|\varphi\|_{\mathcal{H}} = 1$. From (2.6), $\varphi = \sum_{l=A_{R,1}}^{A_{R,2}} \beta_l \phi_{R,l}$ such that $\sum_{l=A_{R,1}}^{A_{R,2}} |\beta_l|^2 = 1$. Observe that

$$\left\| Q_{\mathcal{S}_M}^{\perp} \varphi \right\|_{\mathcal{H}}^2 = \sum_{k \geq \lceil \frac{M}{2} \rceil, k < \lfloor \frac{M}{2} \rfloor} \left| \sum_{l=A_{R,1}}^{A_{R,2}} \beta_l \langle \phi_{R,l}, s_k \rangle \right|^2 = \sum_{k \geq \lceil \frac{M}{2} \rceil, k < \lfloor \frac{M}{2} \rfloor} \frac{\epsilon}{2^R} \left| \Phi\left(\frac{\epsilon k}{2^R}\right) \hat{\phi}\left(\frac{-2\pi\epsilon k}{2^R}\right) \right|^2$$

where

$$\Phi(z) = \sum_{l=A_{R,1}}^{A_{R,2}} \beta_l e^{2\pi i l z}.$$

Let $L = C_\epsilon 2^R/\epsilon \in \mathbb{N}$, then, by our choice of M, we can write $k = mL + j$ for $m < -S, m \geq S$ and $j = 0, \ldots, L-1$, giving

$$\left\| Q_{\mathcal{S}_M}^{\perp} \varphi \right\|_{\mathcal{H}}^2 = \sum_{j=0}^{L-1} \frac{C_\epsilon}{L} \left| \Phi\left(\frac{C_\epsilon j}{L}\right) \right|^2 \sum_{k < -S, k \geq S} \left| \hat{\phi}\left(-2\pi C_\epsilon \left(\frac{j}{L} + k\right)\right) \right|^2.$$

Now, it is a consequence of the Parseval property of the Discrete Fourier transform that given any even $B \in \mathbb{N}$, and $A_1, A_2 \in \mathbb{N}$ such that $B \geq A_2 - A_1 + 1$,

$$\sum_{j=1}^{B-1} \frac{1}{B} \left| \Psi\left(\frac{j}{B}\right) \right|^2 = \sum_{j=A_1}^{A_2} |\xi_l|^2 \tag{5.3}$$

where $\Psi(z) = \sum_{l=A_1}^{A_2} \xi_l e^{2\pi i l z}$. In our case, recalling the definition of $A_{R,2}$ and $A_{R,1}$ from (2.6), we have that

$$L \geq 2^R/\epsilon \geq (3\lceil a \rceil - 2)2^R, \quad A_{R,2} - A_{R,1} + 1 = 2^R(3\lceil a \rceil - 2) + \lceil a \rceil - 1.$$

So, to apply (5.3), we let $A_{R,3} = \lceil (A_{R,2} - A_{R,1} + 1)/2 \rceil - 1$ and

$$\Phi_1(z) = \sum_{l=A_{R,1}}^{A_{R,3}} \beta_l e^{2\pi i l z}, \quad \Phi_2(z) = \sum_{l=A_{R,3}+1}^{A_{R,2}} \beta_l e^{2\pi i l z}.$$

Thus, by applying (5.3), we have that

$$\begin{aligned}
\sum_{j=0}^{L-1} \frac{C_\epsilon}{L} \left| \Phi\left(\frac{C_\epsilon j}{L}\right) \right|^2 &\leq 2 \sum_{j=0}^{L-1} \frac{C_\epsilon}{L} \left(\left| \Phi_1\left(\frac{C_\epsilon j}{L}\right) \right|^2 + \left| \Phi_2\left(\frac{C_\epsilon j}{L}\right) \right|^2 \right) \\
&\leq 2C_\epsilon^2 \sum_{l=A_{R,1}}^{A_{R,2}} |\beta_l|^2 = 2C_\epsilon^2.
\end{aligned}$$

From the estimate on the decay of $\hat{\phi}$, for $j = 0, \dots, L-1$,

$$\sum_{k<-S, k\geq S} \left| \hat{\phi}\left(-\frac{2\pi C_\epsilon j}{L} - 2\pi k C_\epsilon \right) \right|^2 \leq \sum_{|k|\geq S} \frac{A^2}{(2\pi C_\epsilon)^{2\alpha} k^{2\alpha}} \leq \frac{2A^2}{(2\pi C_\epsilon)^{2\alpha}(2\alpha-1)S^{2\alpha-1}}.$$

Thus, $\left\| Q_{\mathcal{S}_M}^{\perp} \varphi \right\|_{\mathcal{H}} \leq \gamma$ whenever

$$S \geq \left(\frac{4A^2}{\gamma^2 (2\pi)^{2\alpha}(2\alpha-1)} \right)^{\frac{1}{2\alpha-1}},$$

i.e. whenever $M = \epsilon^{-1} C_\epsilon C_{\alpha,A} 2^{R+1} \gamma^{-\frac{2}{2\alpha-1}}$ for some constant

$$C_{\alpha,A} \geq \left(\frac{4A^2}{(2\pi)^{2\alpha}(2\alpha-1)} \right)^{\frac{1}{2\alpha-1}}.$$

□

Proposition 5.2. *Assume that the scaling and wavelet functions, ϕ and ψ, are supported on $[0,a]$ and for $k = 0, 1, 2$, for some $A > 0$ and $\alpha \geq 1.5$,*

$$\left| \hat{\phi}^{(k)}(\xi) \right| \leq \frac{A}{(1+|\xi|)^\alpha}, \quad \left| \hat{\psi}^{(k)}(\xi) \right| \leq \frac{A}{(1+|\xi|)^\alpha}, \quad \xi \in \mathbb{R}.$$

Suppose also that the Fourier sampling density $\epsilon \in \mathbb{Q}$ is such that

$$0 < \epsilon \leq \delta/(T_1+T_2), \quad \delta \in (0,1)$$

where $T_1, T_2 > 0$ are such that $\mathcal{W} \subset L^2[-T_1, T_2]$. Then, for $N \in \mathbb{N}$ and any $\gamma \in (0,1)$, there exists S independent of R (but dependent on ϕ, ψ and ϵ) such that

(i) for $M \geq S \cdot \gamma^{-1/(\alpha-1)} \cdot N$,

$$\sup_{\left\{\beta \in \mathbb{C}^N : \|\beta\|_{\ell^\infty}=1\right\}} \sup_{l\in\mathbb{N}} \left| \langle Q_{\mathcal{S}_M} \sum_{\substack{j=1,\\ j\neq l}}^{N} \beta_j \varphi_j, \varphi_l \rangle \right| \leq \gamma. \tag{5.4}$$

(ii) for $M \geq S \cdot \gamma^{-1/(2\alpha-1)} \cdot N$,

$$\sup_{\left\{\beta \in \mathbb{C}^N : \|\beta\|_{\ell^\infty}=1\right\}} \sup_{l\in[N]} \left| \langle Q_{\mathcal{S}_M} \sum_{\substack{j=1,\\ j\neq l}}^{N} \beta_j \varphi_j, \varphi_l \rangle \right| \leq \gamma. \tag{5.5}$$

Proof. Without loss of generality, we will assume throughout that $N = N_R$ for some $R \in \mathbb{N}$. First note that φ_l takes the form of either $\phi_{J,j}$ or $\psi_{J,j}$ for some $J \in \mathbb{N}$ and $j \in \mathbb{Z}$. In this proof, when considering φ_l, we will only be making use of the decay of $\hat{\psi}$ and $\hat{\phi}$ which are assumed to have the same decay. Consequently, we will assume that l is sufficiently large such that the l^{th} element of Ω_a is denoted by $\varphi_l = \psi_{R_l, j_l}$, in accordance with the ordering defined in (2.3). Let $\tilde{\beta} = (\tilde{\beta}_j)$ be such that $\left\|\tilde{\beta}\right\|_{\ell^\infty} \leq 1$. Then, from (2.6), for such $\tilde{\beta}$, there exists (β_j) such that

$$\sum_{j=1}^{N_R} \tilde{\beta}_j \varphi_j = \sum_{j=A_{R,1}}^{A_{R,2}} \beta_j \phi_{R,j}. \tag{5.6}$$

Observe also that for any $\xi \in \mathbb{R}$, there are at most $\lceil a \rceil + 1$ elements of $\{\psi_{j,l} : l \in \mathbb{Z}\}$ whose support contains ξ. Hence,

$$\begin{aligned} \left| \sum_{j=1}^{N_R} \tilde{\beta}_j \varphi_j(\xi) \right| &\leq (\lceil a \rceil + 1) \max\left\{ \|\phi\|_{L^\infty}, \|\psi\|_{L^\infty} \right\} \left(1 + 1 + \sqrt{2} + \cdots + \sqrt{2^{R-1}} \right) \\ &= (\lceil a \rceil + 1) \max\left\{ \|\phi\|_{L^\infty}, \|\psi\|_{L^\infty} \right\} \left(1 + \frac{2^{R/2}-1}{\sqrt{2}-1} \right), \end{aligned}$$

so

$$
\begin{aligned}
|\beta_j| &= \left|\langle \sum_{j=1}^{N_R} \tilde{\beta}_j \varphi_j, \phi_{R,j}\rangle\right| = \left|\int \sum_{j=1}^{N_R} \tilde{\beta}_j \varphi_j(t) \overline{\phi_{R,j}(t)} \mathrm{d}t\right| \leq \left\|\sum_{j=1}^{N_R} \tilde{\beta}_j \varphi_j\right\|_{L^\infty} \int |\phi_{R,j}(t)| \,\mathrm{d}t \\
&\leq \left\|\sum_{j=1}^{N_R} \tilde{\beta}_j \varphi_j\right\|_{L^\infty} \sqrt{\int |\phi_{R,j}(t)|^2 \,\mathrm{d}t} \sqrt{|\mathrm{supp}(\phi_{R,j})|} \qquad (5.7) \\
&\leq \sqrt{\lceil a \rceil}(\lceil a \rceil + 1) \max\left\{\|\phi\|_{L^\infty}, \|\psi\|_{L^\infty}\right\} \left(1 + \frac{1}{\sqrt{2}-1}\right) =: M_{\phi,\psi}
\end{aligned}
$$

and $\|(\beta_j)\|_{\ell^\infty}$ is bounded by a constant $M_{\phi,\psi}$, independent of R.

Using the observations (5.6) and (5.7),

$$
\begin{aligned}
&\sup_{l\in\mathbb{N}} \left|\langle Q_{\mathcal{S}_M} \sum_{\substack{j=1\\ j\neq l}}^{N_R} \tilde{\beta}_j \varphi_j, \varphi_l\rangle\right| = \sup_{l\in\mathbb{N}} \left|\langle Q_{\mathcal{S}_M}^\perp \sum_{\substack{j=1\\ j\neq l}}^{N_R} \tilde{\beta}_j \varphi_j, \varphi_l\rangle\right| \\
&\leq \sup_{l\in\mathbb{N}} \left| \sum_{j=A_{R,1}}^{A_{R,2}} \beta_j \sum_{\substack{k<-\lfloor \frac{M}{2}\rfloor,\\ k\geq \lceil \frac{M}{2}\rceil}} \frac{\epsilon}{\sqrt{2^{R+R_l}}} \hat{\phi}\left(-\frac{2\pi\epsilon k}{2^R}\right) e^{2\pi i\epsilon k j/2^R} \overline{\hat{\psi}\left(-\frac{2\pi\epsilon k}{2^{R_l}}\right)} e^{-2\pi i\epsilon k j_l/2^{R_l}} \right|. \qquad (5.8)
\end{aligned}
$$

Since $\epsilon \in \mathbb{Q}$, we can choose $C_\epsilon \in \mathbb{N}$ such that $C_\epsilon \epsilon^{-1} \in \mathbb{N}$. Now, for some $S \in \mathbb{N}$ which will be determined, let $L = 2^R C_\epsilon \epsilon^{-1}$ and $M = 2SL$. The goal is to show that there exists some constant $S \in \mathbb{N}$ independent of R such that (5.8) is bounded by γ.

Now, by writing $k = mL + b$ with m such that $m < -S$ and $m \geq S$ and $b = 0, \dots, L-1$,

$$
\begin{aligned}
&\sum_{k<-\lfloor \frac{M}{2}\rfloor, k\geq \lceil \frac{M}{2}\rceil} \hat{\phi}\left(-\frac{2\pi\epsilon k}{2^R}\right) e^{2\pi i\epsilon k j/2^R} \overline{\hat{\psi}\left(-\frac{2\pi\epsilon k}{2^{R_l}}\right)} e^{-2\pi i\epsilon k j_l/2^{R_l}} \\
&= \sum_{b=0}^{L-1} H_l\left(\frac{b}{L}\right) \exp\left(\frac{2\pi i\epsilon b j}{2^R} - \frac{2\pi i\epsilon b j_l}{2^{R_l}}\right)
\end{aligned}
$$

where

$$
H_l(\xi) = \sum_{m<-S, m\geq S} \hat{\phi}\left(-2\pi C_\epsilon(m+\xi)\right) \overline{\hat{\psi}\left(-2\pi \frac{2^R}{2^{R_l}} C_\epsilon (m+\xi)\right)} \exp\left(-\frac{i2^{R+1} j_l \pi m C_\epsilon}{2^{R_l}}\right).
$$

So, plugging this back into (5.8) and using (5.7),

$$
\begin{aligned}
&\sup_{l\in\mathbb{N}} \left|\langle Q_{\mathcal{S}_M} \sum_{\substack{j=1\\ j\neq l}}^{N_R} \tilde{\beta}_j \varphi_j, \varphi_l\rangle\right| = \sup_{l\in\mathbb{N}} \left| \sum_{j=A_{R,1}}^{A_{R,2}} \beta_j \frac{\epsilon}{\sqrt{2^{R+R_l}}} \sum_{b=0}^{L-1} H_l\left(\frac{b}{L}\right) \exp\left(\frac{2\pi i\epsilon b j}{2^R} - \frac{2\pi i\epsilon b j_l}{2^{R_l}}\right)\right| \\
&\leq M_{\phi,\psi} \sup_{l\in\mathbb{N}} \sum_{j=A_{R,1}}^{A_{R,2}} \left| \frac{\epsilon}{\sqrt{2^{R+R_l}}} \sum_{b=0}^{L-1} H_l\left(\frac{b}{L}\right) \exp\left(\frac{2\pi i\epsilon b j}{2^R} - \frac{2\pi i\epsilon b j_l}{2^{R_l}}\right)\right| \qquad (5.9) \\
&\leq M_{\phi,\psi} \sup_{l\in\mathbb{N}} \frac{C_\epsilon}{\sqrt{2^{R_l - R}}} \sum_{j=A_{R,1}}^{A_{R,2}} |g_l(j)|
\end{aligned}
$$

where

$$
g_l(j) = \frac{1}{L} \sum_{b=0}^{L-1} H_l\left(\frac{b}{L}\right) \exp\left(\frac{2\pi i\epsilon b j}{2^R} - \frac{2\pi i\epsilon b j_l}{2^{R_l}}\right).
$$

Note that $H_l(0) = H_l(1) = 0$ since it is known that $\hat{\phi}(2\pi m) = 0$ whenever $m \in \mathbb{Z}\backslash\{0\}$ [21, Proposition 2.17, pg 64].

To gain some intuition as to why (5.9) should be bounded by a small constant, first observe that

$$g_l(j) \approx \int_0^1 H_l(t) e^{2\pi i C_\epsilon (j - 2^{R-R_l} j_l) t} \mathrm{d}t = \hat{F}_l(-2\pi i C_\epsilon (j - 2^{R-R_l} j_l))$$

where $F_l = H_l \cdot \chi_{[0,1]}$. Now, H_l is zero at 0 and at 1, and if it is twice differentiable on [0,1), then (by integration by parts) $\hat{F}_l(-2\pi i C_\epsilon (j - 2^{R-R_l} j_l)) \leq C(2\pi C_\epsilon (j - 2^{R-R_l} j_l))^{-2}$ where C depends on $H_l^{(k)}$ for $k = 1, 2$. Moreover, we will aim to make C small by an appropriate choice of S in the definition of H_l. We then use that the fact that $\sum_{n \in \mathbb{N}} n^{-2} < \infty$ to estimate $\sum_{j=A_{R,1}}^{A_{R,2}} |g_l(j)|$. With this in mind, we will now proceed to approximate g_l.

Given sequences (a_k) and (b_k), the following summation by parts formula holds:

$$\sum_{k=0}^{N} a_k b_k = a_N \sum_{k=0}^{N} b_k - \sum_{k=0}^{N-1} (a_{k+1} - a_k) \sum_{j=0}^{k} b_j. \tag{5.10}$$

Observe also that

$$\frac{-\lceil a \rceil + 1}{2^{R_l}} \leq \frac{j_l}{2^{R_l}} \leq \lceil a \rceil - \frac{1}{2^{R_l}},$$

$$-(1 + 2^{-R})\lceil a \rceil - 1 - 2^{-R} = \frac{A_{R,1}}{2^R} \leq \frac{j}{2^R} \leq \frac{A_{R,2}}{2^R} = 2\lceil a \rceil - 1 - 2^{-R}.$$

Thus, $\epsilon \leq \delta/(3\lceil a \rceil - 2)$ implies that

$$-\delta < \frac{\epsilon j}{2^R} - \frac{\epsilon j_l}{2^{R_l}} < \delta \tag{5.11}$$

and

$$\exp\left(2\pi i \left(\frac{\epsilon j}{2^R} - \frac{\epsilon j_l}{2^{R_l}}\right)\right) \neq 1$$

whenever $\left|j - \frac{j_l}{2^{R_l - R}}\right| \neq 0$.

Assuming that j is such that $\left|j - \frac{j_l}{2^{R_l - R}}\right| \geq \max\left\{(2^{R-R_l})^{\alpha - 1/2}, 1\right\}$, we may apply (5.10) to obtain the following.

$$\begin{aligned}
g_l(j) &= \frac{1}{L} \sum_{b=0}^{L} H_l\left(\frac{b}{L}\right) \exp\left(\frac{2\pi i \epsilon b j}{2^R} - \frac{2\pi i \epsilon b j_l}{2^{R_l}}\right) \\
&= -\frac{1}{L} \sum_{b=0}^{L-1} \left(H_l\left(\frac{b+1}{L}\right) - H_l\left(\frac{b}{L}\right)\right) \sum_{k=0}^{b} \exp\left(2\pi i k \left(\frac{\epsilon j}{2^R} - \frac{\epsilon j_l}{2^{R_l}}\right)\right) \\
&= -\frac{1}{L} \sum_{b=0}^{L-1} \left(H_l\left(\frac{b+1}{L}\right) - H_l\left(\frac{b}{L}\right)\right) \left(\frac{\exp\left(2\pi i (b+1) \left(\frac{\epsilon j}{2^R} - \frac{\epsilon j_l}{2^{R_l}}\right)\right) - 1}{\exp\left(2\pi i \left(\frac{\epsilon j}{2^R} - \frac{\epsilon j_l}{2^{R_l}}\right)\right) - 1}\right) \\
&= -\frac{1}{L} \sum_{b=0}^{L-1} \left(H_l\left(\frac{b+1}{L}\right) - H_l\left(\frac{b}{L}\right)\right) \left(\frac{e^{2\pi i b \left(\frac{\epsilon j}{2^R} - \frac{\epsilon j_l}{2^{R_l}}\right)}}{1 - e^{-2\pi i \left(\frac{\epsilon j}{2^R} - \frac{\epsilon j_l}{2^{R_l}}\right)}}\right) \\
&+ \frac{1}{L} \sum_{b=0}^{L-1} \left(H_l\left(\frac{b+1}{L}\right) - H_l\left(\frac{b}{L}\right)\right) \left(\frac{e^{-2\pi i \left(\frac{\epsilon j}{2^R} - \frac{\epsilon j_l}{2^{R_l}}\right)}}{1 - e^{-2\pi i \left(\frac{\epsilon j}{2^R} - \frac{\epsilon j_l}{2^{R_l}}\right)}}\right) \\
&= -\frac{1}{L} \sum_{b=0}^{L-1} \left(H_l\left(\frac{b+1}{L}\right) - H_l\left(\frac{b}{L}\right)\right) \left(\frac{e^{2\pi i b \left(\frac{\epsilon j}{2^R} - \frac{\epsilon j_l}{2^{R_l}}\right)}}{1 - e^{-2\pi i \left(\frac{\epsilon j}{2^R} - \frac{\epsilon j_l}{2^{R_l}}\right)}}\right)
\end{aligned} \tag{5.12}$$

where the last line follows because

$$\sum_{b=0}^{L-1} \left(H_l\left(\frac{b+1}{L}\right) - H_l\left(\frac{b}{L}\right)\right) = H_l(1) - H_l(0) = 0.$$

We now let

$$J(L-1)=H_l\left(\frac{L-1}{L}\right),\qquad J(b)=H_l\left(\frac{b+2}{L}\right)-2H_l\left(\frac{b+1}{L}\right)+H_l\left(\frac{b}{L}\right),\quad b=0,\dots,L-2,$$

then, by applying (5.10) to (5.12), we obtain

$$\begin{aligned} g_l(j) &= \frac{1}{L\left(1-e^{-2\pi i\epsilon\left(\frac{j}{2^R}-\frac{j_l}{2^{R_l}}\right)}\right)}\left\{J(L-1)\cdot\left(\frac{e^{-\frac{2\pi i j_l 2^R}{2^{R_l}}}-1}{1-e^{2\pi i\epsilon\left(\frac{j}{2^R}-\frac{j_l}{2^{R_l}}\right)}}\right)-\sum_{b=0}^{L-2}J(b)\cdot\left(\frac{e^{2\pi i\epsilon(b+1)\left(\frac{j}{2^R}-\frac{j_l}{2^{R_l}}\right)}-1}{1-e^{2\pi i\epsilon\left(\frac{j}{2^R}-\frac{j_l}{2^{R_l}}\right)}}\right)\right\} \\ &= \frac{1}{L\left|2\sin\left(\pi\epsilon\left(\frac{j}{2^R}-\frac{j_l}{2^{R_l}}\right)\right)\right|^2}\left\{J(L-1)\cdot\left(e^{-\frac{2\pi i j_l 2^R}{2^{R_l}}}-1\right)-\sum_{b=0}^{L-2}J(b)\cdot\left(e^{2\pi i\epsilon(b+1)\left(\frac{j}{2^R}-\frac{j_l}{2^{R_l}}\right)}-1\right)\right\}. \end{aligned} \tag{5.13}$$

Note that H_l is twice continuously differentiable on $[0,1]$ by Lemma 5.3. Hence, by the mean value theorem, for $b=0,1,\dots,L-2$,

$$\begin{aligned} J(b)=H_l\left(\frac{b+2}{L}\right)-2H_l\left(\frac{b+1}{L}\right)+H_l\left(\frac{b}{L}\right) &= \frac{1}{L}\left(H_l'\left(\frac{\xi_{b+1}}{L}\right)-H_l'\left(\frac{\xi_b}{L}\right)\right) \\ &= \frac{(\xi_{b+1}-\xi_b)}{L^2}H_l''\left(\frac{\xi_{b+1/2}}{L}\right) \end{aligned}$$

where $\xi_b\in[b,b+1]$ and $\xi_{b+1/2}\in[\xi_b,\xi_{b+1}]\subset[b,b+2]$ and also, since $H_l(1)=0$,

$$J(L-1)=-H_l\left(\frac{L-1}{L}\right)=H_l(1)-H_l\left(\frac{L-1}{L}\right)=\frac{h'(\eta)}{L}$$

for some $\eta\in[1-1/L,1]$. Since $\delta\in(0,1)$, there exists $c_\delta>0$ such that for all $x\in[-\delta,\delta]$,

$$1\geq|\sin(x)/x|\geq c_\delta.$$

So from (5.11), for $\left|j-\frac{j_l}{2^{R_l-R}}\right|\geq\max\left\{(2^{R-R_l})^{\alpha-1/2},1\right\}$

$$\begin{aligned} |g_l(j)| &\leq \frac{1}{4c_\delta^2L\pi^2\left|\left(\frac{\epsilon j}{2^R}-\frac{\epsilon j_l}{2^{R_l}}\right)\right|^2}\left|J(L-1)\left(e^{-\frac{2\pi i\epsilon j_l L}{2^{R_l}}}-1\right)-\sum_{b=0}^{L-2}J(b)\left(e^{2\pi i\epsilon(b+1)\left(\frac{j}{2^R}-\frac{j_l}{2^{R_l}}\right)}-1\right)\right| \\ &\leq \frac{1}{4c_\delta^2\pi^2C_\epsilon^2\left|\left(j-\frac{j_l}{2^{R_l-R}}\right)\right|^2}\left\{2\left|H_l'(\eta)\right|+4\left|\sum_{b=0}^{L-2}\frac{H_l''\left(\frac{\xi_{b+1/2}}{L}\right)}{L}\right|\right\} \\ &\leq \frac{1}{4c_\delta^2\pi^2C_\epsilon^2\left|\left(j-\frac{j_l}{2^{R_l-R}}\right)\right|^2}\left(2\left\|H_l'\right\|_{L^\infty[0,1]}+4\left\|H_l''\right\|_{L^\infty[0,1]}\right) \end{aligned} \tag{5.14}$$

Let $K=\max\left\{(2^{R-R_l})^{\alpha-1/2},1\right\}$ and note that

$$\left|\left\{j:\left|j-\frac{j_l}{2^{R_l-R}}\right|<K\right\}\right|\leq 2K,$$

and

$$|g_l(j)|\leq\|H_l\|_{L^\infty[0,1]}. \tag{5.15}$$

By (5.9),(5.14) and (5.15),

$$
\begin{aligned}
&\sup_{l\in\mathbb{N}}\left|\langle Q_{\mathcal{S}_M}\sum_{j=1}^{N_R}\tilde{\beta}_j\varphi_j,\varphi_l\rangle\right| \le \sup_{l\in\mathbb{N}}\frac{M_{\phi,\psi}C_\epsilon}{\sqrt{2^{R_l-R}}}\sum_{j=A_{R,1}}^{A_{R,2}}|g_l(j)| \\
&\le \sup_{l\in\mathbb{N}}\frac{M_{\phi,\psi}}{\sqrt{2^{R_l-R}}}\left(2C_\epsilon K\left\|H_l\right\|_{L^\infty[0,1]}+\frac{\|H_l'\|_{L^\infty[0,1]}+2\left\|H_l''\right\|_{L^\infty[0,1]}}{2c_\delta^2\pi^2C_\epsilon}\sum_{\left\{j:\left|j-\frac{j_l}{2^{R_l-R}}\right|\ge K\right\}}\frac{1}{\left|(j-\frac{j_l}{2^{R_l-R}})\right|^2}\right) \\
&\le \sup_{l\in\mathbb{N}}\frac{M_{\phi,\psi}}{\sqrt{2^{R_l-R}}}\left(2KC_\epsilon\left\|H_l\right\|_{L^\infty[0,1]}+\min\left\{\frac{\pi^2}{6},\frac{2}{K}\right\}\frac{\|H_l'\|_{L^\infty[0,1]}+2\left\|H_l''\right\|_{L^\infty[0,1]}}{c_\delta^2\pi^2C_\epsilon}\right) \\
&\le \sup_{l\in\mathbb{N}}2M_{\phi,\psi}\left(2^{(R-R_l)\alpha}C_\epsilon\left\|H_l\right\|_{L^\infty[0,1]}+\min\left\{\frac{\pi^2\sqrt{2^{R-R_l}}}{3},2^{(R-R_l)(-\alpha+1)}\right\}\frac{\|H_l'\|_{L^\infty[0,1]}+2\left\|H_l''\right\|_{L^\infty[0,1]}}{c_\delta^2\pi^2C_\epsilon}\right)
\end{aligned} \tag{5.16}
$$

where the penultimate line follows from

$$
\sum_{\left\{j:\left|j-j_l2^{R-R_l}\right|\ge K\right\}}\frac{1}{\left|(j-j_l2^{R-R_l})\right|^2}\le 2\sum_{j\ge K}\frac{1}{j^2}\le\min\left\{\frac{\pi^2}{3},\frac{4}{K}\right\}.
$$

Now, by Lemma 5.3 , we know that for $k=0,1,2$ and some constant $C>0$,

$$
\left\|H_l^{(k)}\right\|_{L^\infty[0,1]}\le C\cdot\begin{cases}\frac{2^k}{(2\pi C_\epsilon)^{\alpha-k}(\alpha-1)S^{\alpha-1}} & R\le R_l\\ \frac{1}{(2\pi C_\epsilon)^{2\alpha-k}(\alpha-1)S^{2\alpha-1}}\left(\frac{2^{R_l}}{2^R}\right)^{\alpha-k} & R>R_l.\end{cases} \tag{5.17}
$$

Thus, by plugging the estimates in (5.17) into (5.16) and recalling that $\alpha\ge 1.5$ and $C_\epsilon\ge 1$, we have that

$$
\begin{aligned}
&\sup_{\{l\in\mathbb{N}:R\le R_l\}}\left|\langle Q_{\mathcal{S}_M}\sum_{j=1}^{N_R}\tilde{\beta}_j\varphi_j,\varphi_l\rangle\right| \\
&\le\frac{2M_{\phi,\psi}C}{S^{\alpha-1}(\alpha-1)}\left(\frac{1}{(2\pi)^\alpha C_\epsilon^{\alpha-1}}+\frac{4}{c_\delta^2 2^{\alpha-1}\pi^{\alpha+1}C_\epsilon^\alpha}+\frac{8}{c_\delta^2 2^{\alpha-2}\pi^\alpha C_\epsilon^{\alpha-1}}\right)\le\frac{6M_{\phi,\psi}C}{S^{\alpha-1}(\alpha-1)c_\delta^2}
\end{aligned}
$$

and

$$
\begin{aligned}
&\sup_{\{l\in\mathbb{N}:R>R_l\}}\left|\langle Q_{\mathcal{S}_M}\sum_{j=1}^{N_R}\tilde{\beta}_j\varphi_j,\varphi_l\rangle\right| \\
&\le\frac{2M_{\phi,\psi}C}{S^{2\alpha-1}(\alpha-1)}\left(\frac{1}{(2\pi)^{2\alpha}C_\epsilon^{2\alpha-1}}+\frac{2^{(R_l-R)(2\alpha-2)}}{c_\delta^2 2^{2\alpha-1}\pi^{2\alpha+1}C_\epsilon^{2\alpha}}+\frac{2^{(R_l-R)(2\alpha-3)}}{c_\delta^2 2^{2\alpha-3}\pi^{2\alpha}C_\epsilon^{2\alpha-1}}\right)\le\frac{2M_{\phi,\psi}C}{S^{2\alpha-1}(\alpha-1)c_\delta^2}.
\end{aligned}
$$

Recalling that $M=2^{R+1}\epsilon^{-1}C_\epsilon S$, (i) follows by choosing

$$
S\ge\left(\frac{6M_{\phi,\psi}C}{(\alpha-1)c_\delta^2\gamma}\right)^{1/(\alpha-1)}
$$

and (ii) follows by choosing

$$
S\ge\left(\frac{6M_{\phi,\psi}C}{(\alpha-1)c_\delta^2\gamma}\right)^{1/(2\alpha-1)}.
$$

□

Lemma 5.3. *Assume that ϕ, ψ are compactly supported and for $k = 0, 1, 2$ and $\varphi = \hat{\phi}, \hat{\psi}$, $C > 0$, $Y \in \mathbb{N}$ and $\alpha \geq 1.5$,*

$$\left| \frac{\mathrm{d}^k}{\mathrm{d}\xi^k} \varphi(\xi) \right| \leq \frac{C}{(1 + |\xi|)^\alpha}.$$

Let

$$H(\xi) = \sum_{m < -S, m \geq S} h_m(\xi).$$

where

$$h_m(\xi) = \hat{\phi}\left(-2\pi Y(m + \xi)\right) \overline{\hat{\psi}\left(-\frac{2^{R+1}\pi Y}{2^{R_l}}(m + \xi)\right)} \exp\left(-\frac{i2^{R+1}j_l \pi Y m}{2^{R_l}}\right).$$

Then there exists some constant C such that for all $\xi \in [0, 1]$, the following holds:

(i)

$$\sum_{m < -S, m \geq S} |h_m(\xi)| \leq \frac{C}{(2\pi Y)^\alpha (\alpha - 1) S^{\alpha - 1} (1 + 2^{R - R_l + 1}\pi Y S)^\alpha},$$

(ii)

$$\sum_{m < -S, m \geq S} |h'_m(\xi)| \leq \frac{C\left(1 + 2^{R - R_l}\right)}{(2\pi Y)^{\alpha - 1} (\alpha - 1) S^{\alpha - 1} (1 + 2^{R - R_l + 1}\pi Y S)^\alpha},$$

(iii)

$$\sum_{m < -S, m \geq S} |h''_m(\xi)| \leq \frac{C(1 + 2^{R - R_l})^2}{(2\pi Y)^{\alpha - 2} (\alpha - 1) S^{\alpha - 1} (1 + 2^{R - R_l + 1}\pi Y S)^\alpha}.$$

Then, $H \in C^2[0, 1]$ and there exists some constant C such that for $k = 0, 1, 2$,

$$\left\| H^{(k)} \right\|_{L^\infty[0,1]} \leq \sup_{\xi \in [0,1]} \left(\sum_{m < -S, m \geq S} \left| h_m^{(k)}(\xi) \right| \right) \leq \frac{C(1 + 2^{R - R_l})^k}{(2\pi Y)^{\alpha - k} (\alpha - 1) S^{\alpha - 1} (1 + 2^{R - R_l + 1}\pi Y S)^\alpha}. \tag{5.18}$$

Proof. First note that $\hat{\phi} \in C^\infty$ and $\hat{\psi} \in C^\infty$, so $h_m \in C^\infty$. Moreover, the absolute convergence in (i), (ii) and (iii) implies that $h \in C^2[0, 1]$ and (5.18). So, it remains to show (i), (ii) and (iii).

(i) For $\xi \in [0, 1]$,

$$\begin{aligned}
\sum_{m < -S, m \geq S} |h_m(\xi)| &\leq \sup_{|\eta| \geq 2^{R - R_l + 1}\pi Y S} \left| \hat{\psi}(\eta) \right| \sum_{m < -S, m \geq S} \left| \hat{\phi}(-2\pi Y(m + \xi)) \right| \\
&\leq \frac{C}{(2\pi Y)^\alpha (1 + 2^{R - R_l + 1}\pi Y S)^\alpha} \sum_{m < -S, m \geq S} \frac{1}{|m + \xi|^\alpha} \leq \frac{C}{(2\pi Y)^\alpha (1 + 2^{R - R_l + 1}\pi Y S)^\alpha} \sum_{m > S} \frac{1}{m^\alpha} \\
&\leq \frac{C}{(2\pi Y)^\alpha (\alpha - 1) S^{\alpha - 1} (1 + 2^{R - R_l + 1}\pi Y S)^\alpha}.
\end{aligned}$$

(ii)

$$\begin{aligned}
h'_m(\xi) = &- 2\pi Y \hat{\phi}'\left(-2\pi Y(m + \xi)\right) \overline{\hat{\psi}\left(-\frac{2^{R+1}\pi Y}{2^{R_l}}(m + \xi)\right)} \exp\left(-\frac{i2^{R+1}j_l \pi Y m}{2^{R_l}}\right) \\
&- \frac{2^{R+1} Y \pi}{2^{R_l}} \hat{\phi}\left(-2\pi Y(m + \xi)\right) \overline{\hat{\psi}'\left(-\frac{2^{R+1} Y \pi}{2^{R_l}}(m + \xi)\right)} \exp\left(-\frac{i2^{R+1}j_l \pi Y m}{2^{R_l}}\right).
\end{aligned}$$

So, for $\xi \in [0, 1]$, by arguing as in (i),

$$\begin{aligned}
\sum_{m < -S, m \geq S} |h'_m(\xi)| &\leq \frac{C}{(2\pi Y)^{\alpha - 1} (1 + 2^{R - R_l + 1}\pi Y S)^\alpha} \left((1 + 2^{R - R_l}) \sum_{m < -S, m \geq S} \frac{1}{|m + \xi|^\alpha} \right) \\
&\leq \frac{C\left(1 + 2^{R - R_l}\right)}{(2\pi Y)^{\alpha - 1} (\alpha - 1) S^{\alpha - 1} (1 + 2^{R - R_l + 1}\pi Y S)^\alpha}.
\end{aligned}$$

(iii)

$$\begin{aligned} h_m''(\xi) =& 4\pi^2 Y^2 \hat{\phi}''\left(-2\pi Y(m+\xi)\right) \overline{\hat{\psi}\left(-\frac{2^{R+1}\pi Y}{2^{R_l}}(m+\xi)\right)} \exp\left(-\frac{i2^{R+1}j_l \pi Y m}{2^{R_l}}\right) \\ &+ \frac{2^{R+3}\pi^2 Y^2}{2^{R_l}} \hat{\phi}'\left(-2\pi Y(m+\xi)\right) \overline{\hat{\psi}'\left(-\frac{2^{R+1}\pi Y}{2^{R_l}}(m+\xi)\right)} \exp\left(-\frac{i2^{R+1}j_l \pi Y m}{2^{R_l}}\right) \\ &+ \frac{2^{2R+2}\pi^2 Y^2}{2^{2R_l}} \hat{\phi}\left(-2\pi Y(m+\xi)\right) \overline{\hat{\psi}''\left(-\frac{2^{R+1}\pi Y}{2^{R_l}}(m+\xi)\right)} \exp\left(-\frac{i2^{R+1}j_l \pi Y m}{2^{R_l}}\right) \end{aligned}$$

So, for $\xi \in [0,1]$, by arguing as in (i),

$$\begin{aligned} \sum_{m<-S, m\geq S} |h_m''(\xi)| &\leq \frac{C(1+2^{R-R_l+1}+2^{2R-2R_l})}{(2\pi Y)^{\alpha-2}(1+2^{R-R_l+1}\pi Y S)^\alpha} \sum_{m<-S, m\geq S} \frac{1}{|m+\xi|^\alpha} \\ &\leq \frac{C(1+2^{R-R_l})^2}{(2\pi Y)^{\alpha-2}(\alpha-1)S^{\alpha-1}(1+2^{R-R_l+1}\pi Y S)^\alpha}. \end{aligned}$$

□

5.2 Proof of Theorem 2.3

We begin with an corollary.

Corollary 5.4. *Consider the setting of Proposition 5.2 and let C be some constant independent of N but dependent on ϕ, ψ and ϵ. Then*

$$\left\| P_{[N]} U^* P_{[M]} U P_{[N]} - P_{[N]} \right\|_{\ell^\infty} \leq \gamma$$

wherever $M \geq C\gamma^{-1/(2\alpha-1)} N$ and

$$\left\| P_{[N]}^\perp U^* P_{[M]} U P_{[N]} \right\|_{\ell^\infty} \leq \gamma$$

wherever $M \geq C\gamma^{-1/(\alpha-1)} N$.

Let $x = (x_j)_{j=1}^N$ such that $\|x\|_{\ell^\infty} = 1$. First observe that

Proof.

$$\left\| (P_{[N]} U^* P_{[M]} U P_{[N]} - P_{[N]}) x \right\|_{\ell^\infty} = \sup_{l=1,\dots,N} \left| \langle Q_{\mathcal{S}_M} \sum_{j=1}^N x_j \varphi_j, \varphi_l \rangle - x_l \right|$$

$$\leq \sup_{l=1,\dots,N} \left| \langle Q_{\mathcal{S}_M} \sum_{\substack{j=1 \\ j\neq l}}^N x_j \varphi_j, \varphi_l \rangle \right| + \|x\|_{\ell^\infty} \left| \|Q_{\mathcal{S}_M} \varphi_l\|_{\mathcal{H}}^2 - 1 \right| = \sup_{l=1,\dots,N} \left| \langle Q_{\mathcal{S}_M} \sum_{\substack{j=1 \\ j\neq l}}^N x_j \varphi_j, \varphi_l \rangle \right| + \left\| Q_{\mathcal{S}_M}^\perp \varphi_l \right\|_{\mathcal{H}}^2. \tag{5.19}$$

It now follows from (ii) of Proposition 5.2 and Lemma 5.1 that for $M \geq C\gamma^{-1/(2\alpha-1)} N$

$$\left\| (P_{[N]} U^* P_{[M]} U P_{[N]} - P_{[N]}) x \right\|_{\ell^\infty} \leq \gamma.$$

Finally,

$$\left\| P_{[N]}^\perp U^* P_{[M]} U P_{[N]} x \right\|_{\ell^\infty} = \sup_{l\geq N} \left| \langle Q_{\mathcal{S}_M} \sum_{j=1}^N x_j \varphi_j, \varphi_l \rangle \right| \leq \gamma$$

whenever $M \geq C\gamma^{-1/(\alpha-1)} N$ by (i) of Proposition 5.2. □

Proof of Theorem 2.3. As in the proof of Theorem 2.2, we will show that

(a) $P_{[N]}U^*P_{[M]}UP_{[N]}$ is invertible on $P_{[N]}(\ell^1(\mathbb{N}))$ such that

$$\left\|P_{[N]}^{\perp}U^*P_{[M]}UP_{[N]}(P_{[N]}U^*P_{[M]}UP_{[N]})^{-1}\right\|_{\ell^\infty} \leq 2,$$

(b) there exists $\rho \in \text{range}(U^*P_{[M]})$ such that $P_{[N]}\rho = \text{sgn}(P_{[N]}x)$ and $\left\|P_{[N]}^{\perp}\rho\right\|_{\ell^\infty} < 1/2$.

and a direct application of Proposition 4.1 concludes this proof.

Suppose first that

$$\left\|(P_{[N]}U^*P_{[M]}UP_{[N]} - P_{[N]})\right\|_{\ell^\infty} \leq \frac{1}{2} \tag{5.20}$$

and

$$\left\|(P_{[N]}^{\perp}U^*P_{[M]}UP_{[N]}\right\|_{\ell^\infty} \leq \frac{1}{4}. \tag{5.21}$$

Then $(P_{[N]}U^*P_{[M]}UP_{[N]})^{-1}$ exists,

$$\left\|(P_{[N]}U^*P_{[M]}UP_{[N]})^{-1}\right\|_{\ell^\infty} \leq \sum_{j=0}^{\infty} \left\|P_{[N]}U^*P_{[M]}UP_{[N]} - P_{[N]}\right\|_{\ell^\infty}^j \leq 2.$$

and

$$\left\|P_{[N]}^{\perp}U^*P_{[M]}UP_{[N]}(P_{[N]}U^*P_{[M]}UP_{[N]})^{-1}\right\|_{\ell^\infty} \leq \frac{1}{2}. \tag{5.22}$$

So (a) is satisfied.

We now let

$$\rho = U^*P_{[M]}UP_{[N]}(P_{[N]}U^*P_{[M]}UP_{[N]})^{-1}P_{[N]}\text{sgn}(x) \in \text{range}(U^*P_{[M]}).$$

Clearly, $P_{[N]}\rho = P_{[N]}\text{sgn}(x)$ and since (5.22) holds, we have that $\left\|P_{[N]}^{\perp}\rho\right\|_{\ell^\infty} \leq \frac{1}{2}$ and (b) holds. Thus, it remains to show that (5.20) and (5.20) hold.

Under the assumptions of (ii), by Corollary 5.4, (5.20) and (5.20) hold whenever $M \geq C \cdot N$ for some constant dependent only on ϕ, ψ and ϵ.

Finally, under the assumptions of (i), by Lemma 5.1, $\left\|P_{[M]}^{\perp}UP_{[N]}\right\|_{\ell^2} \leq \frac{1}{4\sqrt{N}}$ whenever $M \geq C \cdot N^{1+1/(2\alpha-1)}$ for some constant C dependent only on ϕ, ψ and ϵ. Thus, using the fact that U is an isometry,

$$\left\|P_{[N]}U^*P_{[M]}UP_{[N]} - P_{[N]}\right\|_{\ell^\infty} \left\|P_{[N]}U^*P_{[M]}^{\perp}UP_{[N]}\right\|_{\ell^\infty} \leq \sqrt{N}\left\|P_{[M]}^{\perp}UP_{[N]}\right\|_{\ell^2} \leq \frac{1}{4}$$

and

$$\left\|P_{[N]}^{\perp}U^*P_{[M]}UP_{[N]}\right\|_{\ell^\infty} \leq \sqrt{N}\left\|P_{[M]}^{\perp}UP_{[N]}\right\|_{\ell^2} \leq \frac{1}{4}.$$

□

6 A numerical example

In this section, we will be considering the use of generalized sampling in the recovery of wavelet coefficients from Fourier samples. The wavelet bases used will be the orthogonal Daubechies wavelets adapted to the unit interval, by modifying the boundary wavelets [11]. Note that although the previous sections only considered orthogonal wavelets with zero boundary conditions, all the results proved hold also for these interval adapted wavelet bases. This is because the proofs are only dependent on the decay properties of the Fourier transform of the scaling function and wavelet, which are preserved in the construction of boundary adapted wavelets.

We will consider the recovery of the function $f(x) = x^3 + 1$ on the interval $[0,1]$ from its Fourier samples, with Fourier sampling density $\epsilon = 1/2$. Namely, for $M \in \mathbb{N}$, let $\Lambda_M = \{-\lfloor M/2 \rfloor, \ldots, \lceil M/2 \rceil - 1\}$, then the samples observed are $\mathcal{F}_M = \left(2^{-1/2}\hat{f}(\pi k)\right)_{k \in \Lambda_M}$.

For fixed $v \in \mathbb{N}$, let φ_j^v denote the j^{th} Daubechies-v wavelet, where the ordering is in increasing order of wavelet scale. We will compute two types of reconstructions.

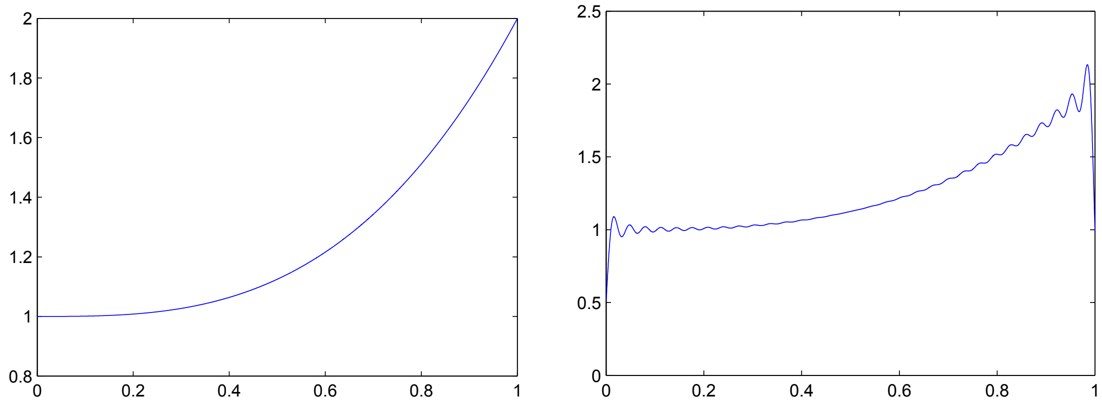

Figure 1: The two figures are reconstructions of $f(x) = x^3+1$ on the unit interval from its first 128 Fourier samples. The left figure shows the generalized sampling reconstruction using Daubechies-3 wavelets, and the right figure shows the truncated Fourier representation.

(i) The truncated Fourier representation $T_M(f) = \epsilon \sum_{j\in\Lambda_M} \hat{f}(2\pi\epsilon j)e^{2\pi i\epsilon j\cdot}$.

(ii) The generalized sampling reconstruction $R_M(f,v) = \sum_{j=1}^{2^{14}} \beta_j \varphi_j^v$ where $\beta = (\beta_j)_{j=1}^{2^{14}}$ is such that

$$\beta \in \operatorname{argmin}_{\eta\in\mathbb{C}^{2^{14}}} \|\eta\|_{\ell^1} \text{ subject to } P_{\Lambda_M} U^v \eta = \mathcal{F}_M. \tag{6.1}$$

where $P_{\Lambda_M} U^v \eta = \left(\langle \sum_{j=1}^{2^{14}} \eta_j \varphi_j^v, \sqrt{\epsilon} e^{2\pi i\epsilon k\cdot}\rangle\right)_{k\in\Lambda_M}$. Recall from Section 2.1 that although β is finite dimensional, it can then be understood as an approximation to an infinite dimensional generalized sampling reconstruction.

Figure 1 shows plots of the $R_M(f,3)$ and $T_M(f)$ for $M = 128$ and Table 2 shows the errors of the reconstructions $T_M(f)$ and $R_M(f,v)$ for $v = 1,2,3$.

Since f is smooth but not periodic, it is expected that the truncated Fourier representation of f has slow decay in its coefficients: $\|\mathcal{T}_M(f) - f\|_{\mathcal{H}} = \mathcal{O}\left(M^{-1/2}\right)$. This is reflected in Table 2. On the other hand, if one has direct access to the wavelet coefficients of f, such that we can compute

$$Q_{\mathcal{W}_M}(f) = \sum_{j=1}^{M} \langle f, \varphi_j^v\rangle \varphi_j^v,$$

then it is known [25] (since the boundary corrected Daubechies wavelets preserve v vanishing moments on the unit interval), that

$$\|Q_{\mathcal{W}_M}(f) - f\|_{\mathcal{H}} = \mathcal{O}\left(M^{-v}\right).$$

In the results of Table 2, although $R_M(f,v)$ is constructed only from M Fourier samples, it achieves an error close to $\mathcal{O}(M^{-v})$. This suggests that there is a linear correspondence between the number of Fourier samples and the number of wavelet coefficients recovered. In particular, one does not need to directly access the wavelets coefficients to attain their benefits.

Our results in Theorem 2.3 suggests such a linear correspondence for sufficiently smooth wavelets, however the results here suggest that this relationship might hold even for nonsmooth wavelets. Thus, the question of whether the result in Theorem 2.3 can be improved remains an open problem. Finally, we remark that this example demonstrates that the nonlinear approach to generalized sampling not only achieves consistent reconstructions but also the same error bounds as the generalized sampling reconstructions derived from least squares approaches.

7 Concluding remarks

We have provided analysis of a non-linear scheme for generalized sampling in arbitrary spaces and have proved that the scheme is consistent, convergent and stable in an ℓ^1 sense. Furthermore, we have

M	Reconstruction	Error_M	$-\frac{\log_2(\text{Error}_M)}{\log_2(M)}$
64	$R_M(f,1)$	0.0113	1.0783
128	$R_M(f,1)$	0.0058	1.0597
256	$R_M(f,1)$	0.0030	1.0490
512	$R_M(f,1)$	0.0015	1.0421
64	$R_M(f,2)$	3.58×10^{-4}	1.9079
128	$R_M(f,2)$	9.21×10^{-5}	1.9152
256	$R_M(f,2)$	2.33×10^{-5}	1.9236
512	$R_M(f,2)$	5.88×10^{-6}	1.9307
64	$R_M(f,3)$	8.43×10^{-6}	2.8092
128	$R_M(f,3)$	1.11×10^{-6}	2.8258
256	$R_M(f,3)$	2.24×10^{-7}	2.7616
512	$R_M(f,3)$	1.75×10^{-7}	2.4937
64	$T_M(f)$	0.0911	0.5760
128	$T_M(f)$	0.0642	0.5658
256	$T_M(f)$	0.0457	0.5563
512	$T_M(f)$	0.0330	0.5466

Table 2: Comparison of the truncated Fourier representation of f, $T_M(f)$ and the generalized sampling reconstruction of f with Daubechies-v wavelets, $R_M(f,v)$. Note that Daubechies-1 refers to the Haar wavelet. In this table, for each reconstruction, say $F_M(f)$, we let $\text{Error}_M = \|F_M(f) - f\|_{\mathcal{H}}$. Note that all reconstructions have as input the first M Fourier samples.

derived error bounds for this scheme in the context of wavelet reconstructions from Fourier samples. In particular, for wavelets of sufficient smoothness, there is a linear correspondence between the number of Fourier samples and the number of wavelets that can be accurately recovered.

Although the work presented in this paper make no assumption on sparsity of the underlying signal, we have remarked upon the relevance of the analysis of this paper for the implementation of compressed sensing. In particular, the choice of an effective sampling strategy in the presence of sparsity requires an a-priori understanding of how many wavelet coefficients can be accurately recovered from M Fourier coefficients.

As previously mentioned, the scheme presented in this paper is already used in practice for the reconstruction of MR images. Although the analysis in this paper has been conducted in one-dimension and for orthonormal systems of wavelets with zero boundary conditions only, the actual properties required for the proofs are very general. The key properties of the space $\mathcal{W}$ that our proofs exploit are the smoothness and hence Fourier decay of the scaling function ϕ and the wavelet ψ and that the existence of an increasing sequence

$$0 < N_1 < \cdots < N_R < N_{R+1} < \cdots$$

such that $N_R = \mathcal{O}\left(2^R\right)$, $\overline{\bigcup_{R\in\mathbb{N}} \mathcal{W}_{N_R}} = \mathcal{W}$ and

$$\mathcal{W}_{N_R} \subseteq \operatorname{span}\left\{\phi_{R,j} : A_{R,1} \leq j \leq A_{R,2}\right\}, \quad A_{R,2} - A_{R,1} = \mathcal{O}\left(2^R\right). \tag{7.1}$$

Consequently, the results of this paper can be readily extended to other types of boundary conditions, in particular, the Daubechies wavelets with special boundary wavelet and scaling functions as described in [11], since it is known from the construction that the boundary scaling function can be written as a linear combination of finitely many elements in $\{\phi(\cdot - k) : k \in \mathbb{Z}\}$. Furthermore, since the properties of the wavelet bases described here are easily preserved when extended to separable two-dimensional wavelet bases, the results here are applicable also to two-dimensions. Thus, the work here may be seen as theoretical foundations for the use of this non-linear framework of generalized sampling in MRI.

8 Acknowledgements

This work was supported by the UK Engineering and Physical Sciences Research Council (EPSRC) grant EP/H023348/1 for the University of Cambridge Centre for Doctoral Training, the Cambridge Centre for Analysis. The author would also like to thank Anders C. Hansen for useful discussions and comments.

References

[1] B. Adcock, A. Hansen, and C. Poon. Beyond consistent reconstructions: Optimality and sharp bounds for generalized sampling, and application to the uniform resampling problem. *SIAM Journal on Mathematical Analysis*, 45(5):3132–3167, 2013.

[2] B. Adcock, A. Hansen, and C. Poon. On optimal wavelet reconstructions from fourier samples: Linearity and universality of the stable sampling rate. *Applied and Computational Harmonic Analysis*, (0):–, 2013.

[3] B. Adcock, A. Hansen, C. Poon, and B. Roman. Breaking the coherence barrier: A new theory for compressed sensing. *Preprint*, 2014.

[4] B. Adcock and A. C. Hansen. Generalized sampling and infinite-dimensional compressed sensing. *Technical report NA2011/02, DAMTP, University of Cambridge*, 2011.

[5] B. Adcock and A. C. Hansen. Reduced consistency sampling in Hilbert spaces. In *Proceedings of the 9th International Conference on Sampling Theory and Applications*, 2011.

[6] B. Adcock and A. C. Hansen. A generalized sampling theorem for stable reconstructions in arbitrary bases. *J. Fourier Anal. Appl.*, 18(4):685–716, 2012.

[7] B. Adcock and A. C. Hansen. Stable reconstructions in Hilbert spaces and the resolution of the Gibbs phenomenon. *Appl. Comput. Harmon. Anal.*, 32(3):357–388, 2012.

[8] A. Aldroubi. Oblique projections in atomic spaces. *Proceedings of the American Mathematical Society*, 124(7):2051–2060, 1996.

[9] E. Candes and C. Fernandez-Granda. Towards a mathematical theory of super-resolution. *arXiv preprint arXiv:1203.5871*, 2012.

[10] E. Candes, J. Romberg, and T. Tao. Stable signal recovery from incomplete and inaccurate measurements. *Communications on pure and applied mathematics*, 59(8):1207–1223, 2006.

[11] A. Cohen, I. Daubechies, and P. Vial. Wavelets on the interval and fast wavelet transforms. *Applied and Computational Harmonic Analysis*, 1(1):54–81, 1993.

[12] I. Daubechies. Orthonormal bases of compactly supported wavelets. *Communications on pure and applied mathematics*, 41(7):909–996, 1988.

[13] I. Daubechies. *Ten Lectures on Wavelets*. Cbms-Nsf Regional Conference Series in Applied Mathematics. Society for Industrial and Applied Mathematics, 1992.

[14] P. L. Dragotti, M. Vetterli, and T. Blu. Sampling moments and reconstructing signals of finite rate of innovation: Shannon meets Strang–Fix. *IEEE Trans. Signal Process.*, 55(5):1741–1757, 2007.

[15] Y. C. Eldar. Sampling with arbitrary sampling and reconstruction spaces and oblique dual frame vectors. *J. Fourier Anal. Appl.*, 9(1):77–96, 2003.

[16] Y. C. Eldar. Sampling without input constraints: Consistent reconstruction in arbitrary spaces. In A. I. Zayed and J. J. Benedetto, editors, *Sampling, Wavelets and Tomography*, pages 33–60. Boston, MA: Birkhäuser, 2004.

[17] Y. C. Eldar and T. Werther. General framework for consistent sampling in Hilbert spaces. *Int. J. Wavelets Multiresolut. Inf. Process.*, 3(3):347, 2005.

[18] N. Gelman and M. L. Wood. Wavelet encoding for 3-d gradient echo MR-imaging. *Magn. Reson. Med.*, 36:613–19, 1996.

[19] M. Guerquin-Kern, M. Haberlin, K. Pruessmann, and M. Unser. A fast wavelet-based reconstruction method for magnetic resonance imaging. *Medical Imaging, IEEE Transactions on*, 30(9):1649–1660, 2011.

[20] A. C. Hansen. On the solvability complexity index, the n-pseudospectrum and approximations of spectra of operators. *J. Amer. Math. Soc.*, 24(1):81–124, 2011.

[21] E. Hernández and G. Weiss. *A First Course on Wavelets.* Studies in Advanced Mathematics. CRC Press, 1996.

[22] A. Hirabayashi and M. Unser. Consistent sampling and signal recovery. *IEEE Trans. Signal Process.*, 55(8):4104–4115, 2007.

[23] T. Hrycak and K. Gröchenig. Pseudospectral Fourier reconstruction with the modified inverse polynomial reconstruction method. *J. Comput. Phys.*, 229(3):933–946, 2010.

[24] F. Krahmer and R. Ward. Compressive imaging: stable and robust recovery from variable density frequency samples. *arXiv preprint arXiv:1210.2380*, 2012.

[25] S. Mallat. *A Wavelet Tour of Signal Processing: The Sparse Way.* Elsevier Science, 2008.

[26] L. P. Panych, P. D. Jakab, and F. A. Jolesz. Implementation of wavelet-encoded MR imaging. *J. Magn. Reson. Imaging*, 3:649–55, 1993.

[27] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, P. Boesiger, et al. Sense: sensitivity encoding for fast mri. *Magnetic Resonance in Medicine*, 42(5):952–962, 1999.

[28] V. Studer, J. Bobin, M. Chahid, H. Mousavi, E. Candes, and M. Dahan. Compressive fluorescence microscopy for biological and hyperspectral imaging. *Proceedings of the National Academy of Sciences*, 109(26):E1679–E1687, 2012.

[29] L. Trefethen and D. Bau. *Numerical Linear Algebra.* Society for Industrial and Applied Mathematics, 1997.

[30] M. Unser. Sampling–50 years after Shannon. *Proc. IEEE*, 88(4):569–587, 2000.

[31] M. Unser and A. Aldroubi. A general sampling theory for nonideal acquisition devices. *IEEE Trans. Signal Process.*, 42(11):2915–2925, 1994.

[32] M. Unser and A. Aldroubi. A review of wavelets in biomedical applications. *Proc. IEEE*, 84(4):626–638, 1996.

[33] M. Unser, A. Aldroubi, and A. F. Laine. Guest editorial: wavelets in medical imaging. *IEEE Trans. Med. Imaging*, 22(3):285–288, 2003.

[34] M. Unser and J. Zerubia. Generalized sampling: Stability and performance analysis. *Signal Processing, IEEE Transactions on*, 45(12):2941–2950, 1997.

[35] M. Unser and J. Zerubia. A generalized sampling theory without band-limiting constraints. *IEEE Trans. Circuits Syst. II.*, 45(8):959–969, 1998.

[36] M. Vetterli, P. Marziliano, and T. Blu. Sampling signals with finite rate of innovation. *IEEE Trans. Signal Process.*, 50(6):1417–1428, 2002.