On optimal wavelet reconstructions from Fourier samples: linearity and universality of the stable sampling rate

B. Adcock[a], A. C. Hansen[b,c], and C. Poon[*b]

[a]Department of Mathematics, Purdue University
[b]Department of Applied Mathematics and Theoretical Physics, University of Cambridge
[c]Institut für Mathematik, Universität Wien,

Abstract

In this paper we study the problem of computing wavelet coefficients of compactly supported functions from their Fourier samples. For this, we use the recently introduced framework of generalized sampling. Our first result demonstrates that using generalized sampling one obtains a stable and accurate reconstruction, provided the number of Fourier samples grows linearly in the number of wavelet coefficients recovered. For the class of Daubechies wavelets we derive the exact constant of proportionality.

Our second result concerns the optimality of generalized sampling for this problem. Under some mild assumptions we show that generalized sampling cannot be outperformed in terms of approximation quality by more than a constant factor. Moreover, for the class of so-called perfect methods, any attempt to lower the sampling ratio below a certain critical threshold necessarily results in exponential ill-conditioning. Thus generalized sampling provides a nearly-optimal solution to this problem.

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

1   Introduction

One of the most fundamental problems in sampling theory is the issue of how to recover an object – an image or signal, for example – from a finite, and typically fixed, collection of its measurements. This problem lies at the heart of countless algorithms, with applications ranging from medical imaging to astronomy.

An important instance of this problem is that of Magnetic Resonance Imaging (MRI). Mathematically, this 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 recover f by computing a discrete Fourier transform (DFT) of the given data. However, this approach suffers from a number of drawbacks, including the sensitivity to motion and the presence of unpleasant Gibbs ringing [27, 46]. Such phenomena present serious problems in MRI.

1.1   Wavelets in MRI

It is known that typical images can be much more efficiently represented using wavelets than by their Fourier series. Images may be sparse in wavelets, or their coefficients may have improved decay properties. Representing an MR image in this way also has several other benefits over the classical Fourier representation. These include better compressibility, improved feature detection (see [42, 43] and references therein), and easier and more effective denoising [33, 35, 45]. For these reasons, the use of wavelets in biomedical imaging applications has been a significant area of research for several decades [33, 42, 43].

---

[*]Corresponding author
Email addresses: adcock@purdue.edu(B.Adcock), a.hansen@damtp.cam.ac.uk(A.C.Hansen), cmhsp2@cam.ac.uk(C.Poon)

Seeking to exploit these beneficial properties, an approach to recover wavelet coefficients directly was introduced in 1992 by Weaver et al [27, 46] (see also [25, 33, 37] and references therein). This is known as *wavelet-encoded* MRI. In this technique, the MR scanner itself is modified to sample wavelet coefficients along one dimension, with Fourier sampling, followed by a one-dimensional DFT, applied in the other. The resulting reconstructed image suffers less from Gibbs ringing, has fewer motion artefacts, and can in principle be acquired more rapidly [36, 37]. For a medical perspective on wavelet encoding, and a discussion on how it can be combined with other imaging techniques such as parallel MRI, see [32].

Unfortunately, there are a number of disadvantages to wavelet encoding, which limit its applicability. These include low signal-to-noise ratio [36, 46], and the extra complications encountered in the acquisition process due to having to modify the MR scanner [33]. Moreover, the state-of-the-art wavelet encoding allows only for reconstructions of wavelet coefficients of a 2D image in one direction, and thus does not permit one to take full advantage of general wavelets.

Nonetheless, the intensity of work on wavelets in MRI, and in particular on wavelet encoding techniques, indicates the importance of the problem of computing wavelet coefficients of biomedical imaging. It also serves to highlight the fact that this problem remains largely unsolved.

With this in mind, the purpose of this paper is to introduce and analyse a different solution to this problem, known as *generalized sampling*. Unlike wavelet encoding, which is primarily an engineering exercise in which the scanner itself is modified to produce different samples, we take the mathematical viewpoint and consider the samples as being fixed Fourier samples, and then seek to reconstruct wavelet coefficients directly via a post-processing algorithm. Our main conclusion is that one can perform wavelet encoding in MRI with generalized sampling without altering the scanner at all. This allows for the use of arbitrary wavelets and removes any hardware restrictions.

Remark 1.1. The reader may wonder at this stage why wavelet encoding is necessary. Why could one not simply recover wavelet coefficients from standard MRI data by applying the DFT and DWT (discrete wavelet transform) in turn? There are two reasons. First, the use of DFT yields a discrete (pixel-based) version of the truncated Fourier series. Hence, by applying the DWT one (at best) obtains the wavelet coefficients of the truncated Fourier series and not the actual wavelet coefficients of the image itself. Second, the recovery algorithm using DFT and DWT would be as follows. The "wavelet coefficients" are obtained by

$$x = \text{DWT} \cdot \text{DFT}^{-1} y,$$

where y is a vector of the Fourier samples. However, when mapping these coefficients back to the pixel domain, one gets

$$\tilde{x} = \text{DWT}^{-1} x = \text{DFT}^{-1} y,$$

which is exactly what we would get in the first place using DFT. In particular, nothing is gained here in terms of the quality of the reconstructed image. By contrast, wavelet encoding techniques seek to reconstruct the true wavelet coefficients directly. This yields a different reconstruction with qualities determined by the wavelet used, and not by the original Fourier series.

1.2 Generalized sampling

In sampling theory, the mathematical problem of recovering the coefficients of a signal or image in a particular basis from samples taken with respect to another basis has been studied for several decades [40]. Motivating this is the fact that many images and signals can be better represented in terms of a different basis (e.g. splines [39] or the aforementioned wavelets) than the basis in which they are sampled (e.g. the Fourier basis). Some of the earliest work on this problem in its abstract form was carried out by Unser & Aldroubi, who introduced a mathematical reconstruction framework known as *consistent reconstructions* for shift-invariant sampling and reconstruction spaces [41] (see also [44]). This was later considered by Eldar et al, who extended this framework to frames in arbitrary Hilbert spaces [15, 16, 17, 21]. Further developments to more general types of signal models were introduced in [34] (see also [7, 18]).

Whilst consistent reconstructions are quite popular in engineering applications, there are a number of issues. As discussed in [2, 3, 19, 29], consistent reconstructions have the significant drawback of being, in general, neither numerically stable or convergent as the number of samples is increased. Hence, when applied to the important problem of recovering wavelet coefficients of MR images, they can result in severe amplification of noise and round-off error.

Nonetheless, it transpires that these issues can be overcome completely by using a different approach, known as *generalized sampling*. Introduced by Adcock & Hansen in [3, 4], based on elements from [26], this framework allows one to recover a signal f modelled as an element of a separable Hilbert space $\mathcal{H}$ in terms of any Riesz basis $\{\varphi_j\}_{j=1}^{\infty}$ from samples $\{\langle f, s_j \rangle\}_{j=1}^{\infty}$ taken with respect to any other Riesz basis $\{s_j\}_{j=1}^{\infty}$ of $\mathcal{H}$. The resulting reconstruction is both convergent and numerically stable, and therefore an obvious candidate for the wavelet recovery problem. The extension of this framework to frames, as opposed to bases, was presented in [5]. See also [6].

Keeping this in mind, the aim of this paper is to show that generalized sampling effectively solves the longstanding problem of recovering wavelet coefficients from Fourier samples. Our main results are explained in more detail in the next section.

1.3 Main results

Generalized sampling obtains a reconstruction by performing a simple least-squares procedure. The fundamental principle which gives this method its stability and accuracy (as opposed to a consistent reconstruction) is that the number of computed coefficients N in the reconstruction basis $\{\varphi_j\}_{j=1}^{\infty}$ (i.e. the wavelet basis) should be allowed to differ from the number M of acquired samples $\{\langle f, s_j \rangle\}_{j=1}^{M}$ (i.e. Fourier samples). In [6], this was posed in terms of the so-called *stable sampling rate* $\Theta(N; \theta)$. Given N coefficients to be recovered, sampling *at a rate* $M \geq \Theta(N; \theta)$ ensures a numerically stable and quasi-optimal reconstruction of f (see Section 2 for definitions), with the stability and quasi-optimality constants depending on the fixed parameter θ.

Understanding the behaviour of $\Theta(N; \theta)$ is critically important from a practical standpoint. In the problem we consider in this paper, for example, it allows one to determine *a priori* how many Fourier samples are required to compute N wavelet coefficients in a manner that is stable and accurate (i.e. the computed wavelet coefficients closely approximate the exact wavelet coefficients). Clearly, it is both wasteful and time-consuming to acquire more samples than necessary. Hence a good estimate of the stable sampling rate is vital.

To this end, the first result we prove in this paper is that the stable sampling rate is linear for any compactly supported wavelet basis. Thus, if N wavelet coefficients are required, one only needs $\mathcal{O}(N)$ Fourier samples of f to apply generalized sampling. In this sense, wavelets give rise to *ideal* bases for the Fourier samples reconstruction problem: up to a constant factor, there is a one-to-one ratio correspondence between Fourier samples and wavelet coefficients. Hence generalized sampling not only solves the long-standing problem of how to recover wavelet coefficients from MR data, but it also does so in a way that is, up to a constant factor, optimal.

This result suggests that little can be gained in terms of reconstruction quality by altering the MR scanner, as is done in wavelet encoding techniques. The problem of recovering wavelet coefficients can be readily solved without altering the scanner by post-processing of the standard Fourier-encoded MR data with generalized sampling. We remark that this conclusion is due completely to the linear scaling of $\Theta(N; \theta)$. Had the scaling been more severe, as can be the case for other reconstruction bases – orthogonal polynomials, for example, have quadratic stable sampling rates, $\Theta(N; \theta) = \mathcal{O}(N^2)$ (a result due originally to Hrycak & Gröchenig [30], see also [4]) – then generalized sampling may well not be as good an approach to the problem as alternatives based on modifying the sampling process.

Given that wavelets have linear stable sampling rates, it is natural to ask how large the ratio $\eta(\theta) = M/N$, which we henceforth refer to as the *stable sampling ratio*, is required to be. Specifically, is it possible to have the optimal ratio $\eta(\theta) = 1$ for some moderate value of θ, and thus get a stable, accurate reconstruction using an equal number of wavelets as Fourier samples? Our second result shows that in general this is not the case. Indeed, every pair of Fourier and wavelet bases is associated with a critical threshold η^* below which the reconstruction becomes exponentially unstable. On the other hand, for certain wavelet bases, such as Daubechies wavelets, a ratio of at least η^* will ensure complete stability.

The third issue we address in this paper is the question of optimality of generalized sampling: that is, whether or not it can be outperformed by a different method. This question is equivalent to asking whether the stable sampling rate is a quantity intrinsic to generalized sampling, or whether it is in fact *universal*. In other words, does the stable sampling rate place a fundamental limit on the number of Fourier samples required to recover N wavelet coefficients in a stable, accurate manner, regardless of the method used?

Optimality of generalized sampling was first discussed in [6]. Using a general result proved therein, we show that the stable sampling rate is indeed universal for all so-called *perfect* methods (i.e. methods which recover finite sums of wavelets in a reasonable way; see Section 2 for a definition). As a result of this, we show that for wavelet reconstructions, any perfect method with ratio less than η^* must be exponentially unstable. Hence, there is always a limit to the amount of improvement over generalized sampling that any perfect method can offer.

Unfortunately, perfect methods represent only a subclass of all possible reconstruction techniques. Hence it cannot be claimed that the stable sampling rate is truly universal. Indeed, perfectness of a method implies that it recovers all functions in a particular class rather well. This leads to the following question: is it possible to devise a different method which outperforms generalized sampling for a single function f? Using our results on the linear scaling of the stable sampling rate, we show under a mild assumption that such a method can at best give a reconstruction whose approximation error is a constant factor smaller than that of generalized sampling. Thus, although it is possible to outperform generalized sampling in terms of approximation error, only the constant can be improved and not the asymptotic rate. In this sense, generalized sampling is, up to a constant factor, an oracle for the problem.

1.4 Outline

The outline for the remainder of this paper is as follows. In Section 2, we recap the generalized sampling framework of [3, 4, 6]. The main results of the paper are presented and discussed in Section 3, and proofs are given in Sections 4–6. In Section 7, we provide numerical results.

2 Generalized sampling

2.1 Generalized sampling

In this section, we recap the main details of generalized sampling from [3, 4], and in particular [6]. Let $\mathcal{H}$ be a separable Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and norm $\|\cdot\|$. Suppose that $\mathcal{S}$ and $\mathcal{T}$ are closed subspaces of $\mathcal{H}$ satisfying the subspace condition

$$\mathcal{T} \cap \mathcal{S}^\perp = \{0\} \text{ and } \mathcal{T} + \mathcal{S}^\perp \text{ is closed in } \mathcal{H}. \tag{1}$$

Let $\{s_j\}_{j=1}^\infty$ be an orthonormal basis for $\mathcal{S}$, and for $f \in \mathcal{H}$, let

$$\hat{f}_j = \langle f, s_j \rangle, \quad j \in \mathbb{N},$$

be the *samples* of f. The reconstruction problem is to recover f with an element $\tilde{f} \in \mathcal{T}$ from its samples $\{\hat{f}_j\}_{j=1}^\infty$.

In practice, one does not have access to the whole set $\{\hat{f}_j\}_{j=1}^\infty$ of samples, nor can one process infinite amounts of information. Hence, in computations we consider the problem of recovering f from its first M samples

$$\hat{f}_1, \ldots, \hat{f}_M.$$

Also, it is usual to assume that there exists a sequence $\{\mathcal{T}_N\}_{N=1}^\infty$ of finite-dimensional subspaces of $\mathcal{T}$ satisfying

$$\mathcal{T}_1 \subseteq \mathcal{T}_2 \subseteq \cdots \subseteq \mathcal{T}, \quad \overline{\bigcup_{N=1}^\infty \mathcal{T}_N} = \mathcal{T}. \tag{2}$$

For example, if $\{\varphi_j\}_{j=1}^\infty$ is a frame or a Riesz basis for $\mathcal{T}$, then one typically has

$$\mathcal{T}_N = \text{span} \{\varphi_1, \ldots, \varphi_N\}.$$

The reconstruction problem is now formulated as follows: given $N \in \mathbb{N}$, compute a reconstruction $\tilde{f}_{N,M} \in \mathcal{T}_N$ of f from the samples $\{\hat{f}_j\}_{j=1}^M$.

In order to formulate what constitutes a 'good' reconstruction, we consider the following two definitions [6]:

Definition 2.1. *Let $F_{N,M} : \mathcal{H} \to \mathcal{T}_N$. The quasi-optimality constant $\mu = \mu(F_{N,M})$ is the least constant such that*

$$\|f - F_{N,M}(f)\| \le \mu \|f - Q_N f\|, \quad \forall f \in \mathcal{H},$$

where $Q_N : \mathcal{H} \to \mathcal{T}_N$ is the orthogonal projection onto $\mathcal{T}_N$. If no such constant exists, we write $\mu = \infty$. We say that $F_{N,M}$ is quasi-optimal if $\mu(F_{N,M})$ is small.

Note that $Q_N f$ is the best approximation in norm to f from $\mathcal{T}_N$. So quasi-optimality means that the difference in norm between f and $F_{N,M}(f)$ is at most a constant factor μ of the difference between f and its best approximation in the subspace $\mathcal{T}_N$.

We also define the condition number of a reconstruction:

Definition 2.2. *Let $F_{N,M} : \mathcal{H} \to \mathcal{T}_N$ be a mapping such that, for each $f \in \mathcal{H}$, $F_{N,M}(f)$ depends only on the samples samples $\{\hat{f}_j\}_{j=1}^M$. The condition number of $\kappa(F_{N,M})$ is given by*

$$\kappa(F_{N,M}) = \sup_{f \in \mathcal{H}} \lim_{\epsilon \to 0^+} \sup_{\substack{g \in \mathcal{H} \\ 0 < \|\hat{g}\|_{l^2} \le \epsilon}} \frac{\|F_{N,M}(f+g) - F_{N,M}(f)\|}{\|\hat{g}\|_{l^2}},$$

where $\hat{g} = \{\hat{g}_j\}_{j=1}^M \in \mathbb{C}^M$. The mapping $F_{N,M}$ is well-conditioned if $\kappa(F_{N,M})$ is small and ill-conditioned otherwise.

We say that the reconstruction $F_{N,M}$ is 'good' if it is stable and quasi-optimal. In other words, if the *reconstruction constant*

$$C(F_{N,M}) = \max\{\kappa(F_{N,M}), \mu(F_{N,M})\},$$

is small.

As we shall explain in a moment, the key to obtaining a good reconstruction is to allow the parameter M, the number of samples, to vary independently from N. To this end, suppose now we write $P_M : \mathcal{H} \to \mathcal{S}_M$ for the orthogonal projection onto the subspace $\mathcal{S}_M = \text{span}\{s_1, \ldots, s_M\}$, i.e.

$$P_M g = \sum_{j=1}^M \langle g, s_j \rangle s_j, \quad g \in \mathcal{H}.$$

The method of tackling the reconstruction problem proposed in [3] is to let $\tilde{f}_{N,M} = F_{N,M}(f) \in \mathcal{T}_N$ be defined by

$$\left\langle P_M \tilde{f}_{N,M}, \varphi_j \right\rangle = \langle P_M f, \varphi_j \rangle, \quad j = 1, \ldots, N. \tag{3}$$

Note that solving (3) is equivalent to finding $\alpha^{[N,M]} = \left\{ \alpha_1^{[N,M]}, \ldots, \alpha_N^{[N,M]} \right\} \in \mathbb{C}^N$ as the least-squares solution to the problem

$$U^{[N,M]} \alpha^{[N,M]} = \hat{f}^{[M]},$$

where $\hat{f}^{[M]} = \{\langle f, s_1 \rangle, \ldots, \langle f, s_M \rangle\}$ and $U^{[N,M]}$ is the M by N matrix whose $(i,j)^{th}$ entry is $\langle \varphi_j, s_i \rangle$. The reconstruction $\tilde{f}_{N,M}$ is then given by $\sum_{j=1}^N \alpha_j^{[N,M]} \varphi_j$.

In [6], it was established that the reconstruction constant $C(F_{N,M})$ of generalized sampling satisfies

$$C(F_{N,M}) = \kappa(F_{N,M}) = \mu(F_{N,M}) = \frac{1}{C_{N,M}},$$

where

$$C_{N,M} = \sqrt{\inf_{\substack{\varphi \in \mathcal{T}_N \\ \|\varphi\| = 1}} \langle P_M \varphi, \varphi \rangle},$$

is the subspace angle between $\mathcal{T}_N$ and $\mathcal{S}_M$. Moreover, since $P_M \to P$ strongly on $\mathcal{H}$ as $M \to \infty$ (where $P : \mathcal{H} \to \mathcal{H}$ is the projection onto $\mathcal{S}$), one has, via (1) and (2), that

$$C_{N,M} \to 1, \quad M \to \infty,$$

for fixed $N \in \mathbb{N}$. Thus, one obtains a good reconstruction by allowing M to be sufficiently large in comparison to N.

To quantify how large M is required to be, the concept of the *stable sampling rate* was introduced in [6]:

Definition 2.3. *For $N \in \mathbb{N}$ and $\theta \in (1, \infty)$, the stable sampling rate is given by*

$$\Theta(N; \theta) = \min \left\{ M \in \mathbb{N} : \frac{1}{C_{N,M}} < \theta \right\}.$$

This notion of the stable sampling rate is important as it determines the number of samples required for guaranteed, quasi-optimal and numerically stable reconstructions. In particular, for all $M \geq \Theta(N; \theta)$, we have that $\tilde{f}_{N,M}$ is quasi-optimal to f from $\mathcal{T}_N$ with constant a most θ, and the condition number $\kappa(F_{N,M})$ is at worst θ.

2.2 Optimality of generalized sampling

In [6] the question of optimality of generalized sampling was also discussed. We now recap the main results proved, since they will be of use later. We first recall the definition of a perfect method:

Definition 2.4. *Let $G_{N,M} : \mathcal{H} \rightarrow \mathcal{T}_N$ be a mapping such that, for each $f \in \mathcal{H}$, $G_{N,M}(f)$ depends only on the samples $\{\hat{f}_j\}_{j=1}^M$. If $G_{N,M}(f) = f$ for all $f \in \mathcal{T}_N$, then $G_{N,M}$ is said to be perfect.*

Observe that the notion of perfectness is strictly weaker than quasi-optimality. Also, we remark that generalized sampling is a perfect method, as can be seen from (3).

The first result of [6] concerns such methods:

Theorem 2.5. *For $M \geq N$ let $G_{N,M} : \mathcal{H} \rightarrow \mathcal{T}_N$ be a perfect reconstruction method such that, for each $f \in \mathcal{H}$, $G_{N,M}(f)$ depends only on the samples $\{\hat{f}_j\}_{j=1}^M$. Then the condition number*

$$\kappa(G_{N,M}) \geq \kappa(F_{N,M}),$$

where $F_{N,M}$ is the generalized sampling reconstruction.

This result implies the following: for any perfect reconstruction method, one must sample at a rate higher than that of generalized sampling – namely, the stable sampling rate – to obtain a stable reconstruction. In other words, generalized sampling cannot be improved upon in terms of its stability (at least for perfect methods).

The case of non-perfect methods was also studied in [6]. The following result was proved:

Theorem 2.6. *Suppose that the stable sampling rate $\Theta(N; \theta)$ is linear in N for a particular sampling and reconstruction problem. Let $f \in \mathcal{H}$ be fixed, and suppose that there exists a sequence of mappings*

$$G_M : \{\hat{f}_j\}_{j=1}^M \mapsto G_M(f) \in \mathcal{T}_{\Psi_f(M)},$$

where $\Psi_f : \mathbb{N} \rightarrow \mathbb{N}$ with $\Psi_f(M) \leq M$. Suppose also that there exist constants $c_1(f), c_2(f), \alpha_f > 0$ such that

$$c_1(f)N^{-\alpha_f} \leq \|f - Q_N f\| \leq c_2(f)N^{-\alpha_f}, \quad \forall N \in \mathbb{N}. \tag{4}$$

Then, given $\theta \in (1, \infty)$, there exist constants $c(\theta) \in (0, 1)$ and $c_f(\theta) > 0$ such that

$$\|f - F_{c(\theta)M,M}(f)\| \leq c_f(\theta)\|f - G_M(f)\|, \quad \forall M \in \mathbb{N}, \tag{5}$$

where $F_{N,M}$ is the generalized sampling reconstruction.

This theorem demonstrates that for problems with linear stable sampling rates, even if one is allowed to design a method that depends on f in a completely non-trivial way, it is still not possible to obtain a faster asymptotic rate of convergence than that of generalized sampling. As we explain Section 3, the stable sampling rate is linear for wavelets, making this theorem directly applicable.

Observe that a consequence of this theorem is that generalized sampling is, up to a constant, an oracle for the wavelet coefficient reconstruction problem. Suppose there was some method that, for a particular f satisfying (4), could recover the first $N = M$ wavelet coefficients of f *exactly* (i.e. with no error) from M Fourier samples. The conclusion of the above corollary is that generalized sampling commits an error that is at worst a constant factor larger than that of this method.

2.3 The wavelet reconstruction and Fourier sampling spaces

In the remainder of this paper we focus on the problem of recovering wavelet coefficients from Fourier samples. To this end, we now specify the corresponding sampling space $\mathcal{S}$, with its corresponding sampling vectors $\{s_j\}_{j\in\mathbb{N}}$, as well as the reconstruction space $\mathcal{T}$ with the reconstruction vectors $\{\varphi_j\}_{j\in\mathbb{N}}$. Throughout we let $\mathcal{H} = L^2(\mathbb{R})$ with its usual inner product and will consider the recovery of functions in $\mathcal{H}$ that are compactly supported on $[0, a]$ for some $a \geq 1$. Moreover, given $f \in \mathcal{H}$, its Fourier Transform is defined as

$$\hat{f}(\omega) = \int_{\mathbb{R}} f(x)e^{-i\omega x}\, dx, \quad \omega \in \mathbb{R}.$$

2.3.1 The wavelet reconstruction space

Suppose that the reconstruction space $\mathcal{T}$ is generated by a mother wavelet ψ and a scaling function ϕ such that $\mathrm{supp}(\psi) = \mathrm{supp}(\phi) = [0, a]$. Then the only wavelets of interest are those whose support intersects $[0, a]$. In particular, for

$$\phi_{j,k} = 2^{j/2}\phi(2^j \cdot -k), \qquad j, k \in \mathbb{Z},$$
$$\psi_{j,k} = 2^{j/2}\psi(2^j \cdot -k), \qquad j, k \in \mathbb{Z},$$

the wavelets of interest are

$$\Omega_a = \left\{\phi_{0,k} : |k| = 0, 1, \ldots, \lceil a \rceil - 1\right\} \cup \left\{\psi_{j,k} : j \in \mathbb{Z}+, k \in \mathbb{Z}, -\lceil a \rceil + 1 \leq k \leq 2^j \lceil a \rceil - 1\right\}.$$

So,

$$\mathcal{T} = \overline{\mathrm{span}\left\{\varphi : \varphi \in \Omega_a\right\}},$$

and for sufficiently large T_1 and T_2, namely, $T_1 \geq \lceil a \rceil - 1$ and $T_2 \geq 2\lceil a \rceil - 1$:

$$L^2[0, a] \subset \mathcal{T} \subset L^2[-T_1, T_2]. \tag{6}$$

Elements of Ω_a are ordered as follows:

$$\{\varphi_j\}_{j\in\mathbb{N}} = \{\phi_{0,-\lceil a \rceil+1}, \phi_{0,-\lceil a \rceil+2}, \ldots, \phi_{0,\lceil a \rceil-1}, \psi_{0,-\lceil a \rceil+1}, \ldots, \psi_{0,\lceil a \rceil-1}, \psi_{1,-\lceil a \rceil+1}, \ldots, \psi_{1,2^j\lceil a \rceil-1}, \ldots\}, \tag{7}$$

and thus

$$\mathcal{T}_N = \mathrm{span}\left\{\varphi_j : j = 1, \ldots, N\right\}. \tag{8}$$

Although one can in principle consider arbitrary values of $N \in \mathbb{N}$, it is natural instead to consider only those N for which $\mathcal{T}_N$ contains all wavelets up to a certain scale. To this end, we now write

$$N_R = 2^R \lceil a \rceil + (R+1)(\lceil a \rceil - 1), \quad R \in \mathbb{N}. \tag{9}$$

We will verify in Lemma 4.1 that the subspace $\mathcal{T}_{N_R}$ consists of all wavelets $\psi_{j,k}$ of scale $0 \leq j \leq R - 1$.

2.3.2 The Fourier sampling space

Given $[-T_1, T_2]$ (the support of $\mathcal{T}$), we let $\epsilon \leq 1/(T_1+T_2)$ be the *sampling density* (or *sampling distance*). 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

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

the sampling space by

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

and the space spanned by the first M sampling vectors by

$$\mathcal{S}_M^\epsilon = \mathrm{span}\left\{s_l^\epsilon : -\left\lfloor \frac{M}{2} \right\rfloor \leq l \leq \left\lceil \frac{M}{2} \right\rceil - 1\right\}. \tag{10}$$

Moreover, P^ϵ and P_M^ϵ will denote the orthogonal projections from $\mathcal{H}$ onto $\mathcal{S}^\epsilon$ and $\mathcal{S}_M^\epsilon$ respectively. Where there is no ambiguity about the value of the sampling density, we will drop the ϵ notation and simply write $\mathcal{S}$, $\mathcal{S}_M$, P and P_M instead.

Remark 2.7. Observe that for all $\epsilon \leq 1/(T_1 + T_2)$, $\mathcal{T} \subset \mathcal{S}^\epsilon$. So, $\mathcal{T} + (\mathcal{S}^\epsilon)^\perp$ is a closed subspace of $\mathcal{H}$ and $\mathcal{T} \cap (\mathcal{S}^\epsilon)^\perp = \{0\}$. Thus, the subspace condition (1) of generalized sampling is satisfied.

3 Main results

We now state the main results of this paper. Proofs are provided in Sections 4 and 5.

3.1 Linearity of the stable sampling rate

The first result of this paper is that the stable sampling rate for wavelet reconstructions from Fourier samples is linear for any compactly supported wavelet basis. In other words, up to a constant factor there is a one-to-one correspondence between Fourier samples and wavelet coefficients. In particular, all information about a function that can be retrieved from its wavelet coefficients can still be retrieved even in the situation where only Fourier samples are available (and not the wavelet coefficients themselves).

More formally, we have the following theorem:

Theorem 3.1. *Let $\mathcal{S}$ and $\mathcal{T}$ be the sampling and reconstruction spaces defined in Section 2.3 and recall N_R from (9). Let $N \leq N_R$ for $R \in \mathbb{N}$. Then for all $\theta \in (1, \infty)$ there exists $S_\theta \in \mathbb{N}$, independent of R, such that for*

$$M = \left\lceil \frac{S_\theta 2^{R+1}}{\epsilon} \right\rceil,$$

we have

$$C_{N,M} \geq \frac{1}{\theta}.$$

In particular,

$$\Theta(N, \theta) \leq \left\lceil \frac{2 S_\theta N}{\epsilon \lceil a \rceil} \right\rceil.$$

Hence, $\Theta(N, \theta) = \mathcal{O}(N)$ for any $\theta \in (1, \infty)$.

Since the stable sampling rate is linear for wavelets, it makes sense to introduce the notion of a *stable sampling ratio*. We define

$$\eta(\theta) = \limsup_{N \to \infty} \frac{\Theta(N; \theta)}{N}, \qquad \theta \in (1, \infty). \tag{11}$$

Note the difference between $\Theta(N; \theta)$, which determines how many samples are required for each N, and $\eta(\theta)$, which stipulates asymptotically how many are required as $N \to \infty$. We will also discuss sampling ratios in the context of other methods. To this end, suppose that an arbitrary method G uses $\Theta_G(N) \in \mathbb{N}$ samples to reconstruct the first N wavelet coefficients. We define the sampling ratio for that method as

$$\eta_G = \limsup_{N \to \infty} \frac{\Theta(N)}{N}.$$

Since the stable sampling rate is linear, we shall only consider methods G for which η_G is defined (all other methods necessarily give worse reconstructions asymptotically as $N \to \infty$).

3.2 Universality of the stable sampling rate

The second collection of results concerns the universality of the stable sampling rate, or equivalently, the optimality of generalized sampling amongst all methods which recover N wavelet coefficients from $M \geq N$ Fourier samples. Our first result is simply a corollary of Theorem 2.5 for the wavelet reconstruction problem from Fourier samples:

Corollary 3.2. *For $N \in \mathbb{N}$, let G_N be a sequence of perfect reconstruction methods with sampling ratio $\eta_G \geq 1$. If η_G is such that $\kappa(G_N) \leq \theta$ for some $\theta \in (1, \infty)$ and all sufficiently large N, then $\eta_G \geq \eta(\theta)$, where $\eta(\theta)$ is the stable sampling ratio for generalized sampling.*

This corollary states that, for any perfect method, the stable sampling ratio $\eta(\theta)$ cannot be lowered. In particular, any perfect method requires at least the same number of Fourier samples to achieve as stable a reconstruction as that of generalized sampling.

Despite this result, in some cases it might seemingly be acceptable to forgo complete stability to obtain a better reconstruction. Our next theorem, which is specific to the wavelet reconstruction problem, shows that this cannot be done in practice:

Theorem 3.3. *Let G_N be as in Corollary 3.2 with sampling rate $\eta_G \geq 1$. If $\eta_G < \frac{1}{\epsilon \lceil a \rceil}$, where ϵ is as in Section 2.3, then $\kappa(G_N)$ is unbounded and $\kappa(G_{N_R})$ becomes exponentially large as $N_R \to \infty$, where N_R is as defined in (9).*

This theorem demonstrates that any attempt to improve upon generalized sampling by lowering the sampling rate will result in extremely poor stability, and consequently extreme sensitivity to noise and round-off error. Prior to this result, one may have hoped that sampling below the critical threshold $\eta = \frac{1}{\epsilon \lceil a \rceil}$ might only result in mildly growing condition numbers. This theorem demonstrates that this is not the case: stability rapidly declines dramatically once $\eta < \frac{1}{\epsilon \lceil a \rceil}$.

Corollary 3.2 and Theorem 3.3 establish the *universality* of the stable sampling rate, and the pitfalls of trying to circumvent the stability barrier $\eta \geq \frac{1}{\epsilon \lceil a \rceil}$. However, they are valid only for perfect methods. Recall that the question of non-perfect methods was addressed by Theorem 2.6. In terms of the sampling ratio, this implies that any non-perfect method which has a lower sampling ratio for a particular function f satisfying (4) can only outperform generalized sampling by a constant factor. Note that the problem of recovering wavelet coefficients from Fourier samples certainly satisfies the assumptions of Theorem 2.6: as we prove, the stable sampling rate is linear, and for typical functions f, it is usually the case that the wavelet coefficients decay algebraically (which implies (4)).

3.3 Sharp results for the Daubechies wavelets

Although Theorem 3.1 establishes linearity of the stable sampling rate for any compactly supported wavelet basis, it does not provide the precise constant of proportionality. Nor is it straightforward to determine an upper bound, since the quantity S_θ is not given explicitly. Although one can in theory estimate S_θ by carefully following the steps of the proof, we shall not do this. Instead, in this section we show that for the important case of Daubechies wavelets the constant can be determined exactly.

Remark 3.4. The fact that the constant may not be known in general does not necessary prohibit implementation of generalized sampling. As discussed in [6], the stable sampling rate is explicitly computable, and thus the constant can actually be determined *a priori* for each particular case through numerical means.

Our main result is as follows:

Theorem 3.5. *Let $\mathcal{S}$ and $\mathcal{T}$ be the sampling and reconstruction spaces defined in Section 2.3, where $\mathcal{T}$ is generated by a Daubechies wavelet, and recall N_R from (9). Then, there exists $\theta \in (1, \infty)$ and $R_0 \in \mathbb{N}$ such that for all $R \geq R_0$,*

$$\Theta(N_R, \theta) = \lceil 2^R / \epsilon \rceil .$$

In particular, when $1/\epsilon \in \mathbb{Z}$ it suffices to let

$$\theta > \left(\inf_{\xi \in [-\pi, \pi]} \left| \hat{\phi}(\xi) \right| \right)^{-1} .$$

Moreover, in addition to this, for Haar wavelets, where $a = 1$, we have that $\Theta(N_R, \theta) \leq \lceil 2^R / \epsilon \rceil$ for all $R \in \mathbb{N}$.

Remark 3.6. Note also that for such values of θ and R in Theorem 3.5, if N is such that $N_{R-1} + 1 \leq N \leq N_R$, then

$$\Theta(N, \theta) \leq \left\lceil \frac{2^R}{\epsilon} \right\rceil .$$

Therefore, we have that

$$\frac{1}{\epsilon \lceil a \rceil} \le \eta(\theta) \le \lim_{R \to \infty} \frac{\lceil 2^R / \epsilon \rceil}{N_{R-1} + 1} = \frac{2}{\epsilon \lceil a \rceil}.$$

However, our numerical results in Section 7 suggest that the optimal ratio is $(\epsilon \lceil a \rceil)^{-1}$ and is attained only when $N = N_R$.

4 Proof of Theorem 3.1

The proof of Theorem 3.1 requires a series of lemmas and propositions that will be presented below. The actual proof can be found at the very end of this section.

4.1 Expressing wavelets in terms of the scaling function

It is known that any compactly supported orthonormal wavelet is a Multiresolution Analysis (MRA) wavelet [28]. Hence, as we will demonstrate in this section, given any $N \in \mathbb{N}$, all basis elements of $\mathcal{T}_N$ may be expressed as a linear combination of finitely many basis elements of $\{\phi_{R,k} : k \in \mathbb{Z}\}$ for some $R \in \mathbb{N}$. Let therefore, for $j \in \mathbb{Z}_+$,

$$
\begin{aligned}
V_j &= \overline{\operatorname{span}\{\phi_{j,k} : k \in \mathbb{Z}\}}, \\
W_j &= \overline{\operatorname{span}\{\psi_{j,k} : k \in \mathbb{Z}\}}, \\
V_0^{(a)} &= \operatorname{span}\{\phi_{0,k} : k \in \mathbb{Z}, |k| \le \lceil a \rceil - 1\}, \\
W_j^{(a)} &= \operatorname{span}\{\psi_{j,k} : k \in \mathbb{Z}, -\lceil a \rceil + 1 \le k \le 2^j \lceil a \rceil - 1\}.
\end{aligned}
$$

The following lemma relates $\mathcal{T}_N$ to the two latter types of subspaces.

Lemma 4.1. *For $R \in \mathbb{N}$, let*

$$
\begin{aligned}
A_{R,1} &= -(2^R + 1)\lceil a \rceil + 2^R + 1, \\
A_{R,2} &= 2^{R+1}\lceil a \rceil - 2^R - 1, \\
\mathbb{V}_{R,a} &= \operatorname{span}\{\phi_{R,k} : A_{R,1} \le k \le A_{R,2}\}.
\end{aligned}
\tag{12}
$$

Then, the following holds:

(i)

$$V_0^{(a)} \oplus \left(\bigoplus_{j=0}^{R-1} W_j^{(a)} \right) \subset \mathbb{V}_{R,a}. \tag{13}$$

(ii) Let $N = N_R$, as defined in (9). Then

$$\mathcal{T}_N = V_0^{(a)} \oplus W_0^{(a)} \oplus \cdots \oplus W_{R-1}^{(a)} \subset \mathbb{V}_{R,a},$$

where $\mathcal{T}_N$ is defined in (7) and (8). Moreover, if $\|\phi\|_\infty$ and $\|\psi\|_\infty$ exist, then given any $\varphi \in \mathcal{T}_N$ such that $\|\varphi\| = 1$ and $R \ge \log_2(\lceil a \rceil - 1)$, the following holds:

$$\varphi = \sum_{j=A_{R,1}}^{A_{R,2}} \alpha_j \phi_{R,j}, \qquad \sum_{j=A_{R,1}}^{A_{R,2}} |\alpha_j|^2 = 1$$

and

$$\sum_{j=A_{R,2}-\lceil a \rceil + 1}^{A_{R,2}} |\alpha_j|^2 \le \frac{(\|\phi\|_\infty + \|\psi\|_\infty)^2 \lceil a \rceil (\lceil a \rceil + 1)}{2^{R+1}}.$$

(iii)

$$\phi_{R,k} \in V_0^{(a)} \oplus W_0^{(a)} \oplus \cdots \oplus W_{R-1}^{(a)} \qquad whenever\ 0 \le k \le (2^R - 1)\lceil a \rceil \tag{14}$$

10

Proof. To prove (i) we start by observing that in MRA, we have that, for $R \in \mathbb{N}$

$$V_R = V_0 \oplus \left(\underset{l=0}{\overset{R-1}{\oplus}} W_l, \right)$$

so

$$V_R \supset V_0^{(a)} \oplus \left(\underset{l=0}{\overset{R-1}{\oplus}} W_l^{(a)} \right).$$

Thus, since $\{\phi_{R,k} : k \in \mathbb{Z}\}$ is an orthonormal basis for the closed subspace V_R, it follows that, given $l \in \mathbb{Z}$ such that $|l| \leq \lceil a \rceil - 1$,

$$\phi_{0,l} = \sum_{k \in \mathbb{Z}} \beta_k \phi_{R,k}, \qquad \beta_k = \int_{\mathbb{R}} \phi_{0,l}(x) \overline{\phi_{R,k}(x)} \mathrm{d}x.$$

Note that ϕ has compact support, so finitely many β_k's are non-zero. In particular, $\beta_k = 0$ if k is such that measure $(\mathrm{supp}(\phi_{0,l}) \cap \mathrm{supp}(\phi_{R,k})) = 0$. So, $\beta_k \neq 0$ only if

$$2^R l - \lceil a \rceil + 1 \leq k \leq 2^R(\lceil a \rceil + 1) - 1$$

and this is only if

$$-(2^R + 1)\lceil a \rceil + 2^R + 1 \leq k \leq 2^{R+1} \lceil a \rceil - 2^R - 1. \tag{15}$$

Similarly, given $j, l \in \mathbb{Z}$ such that $0 \leq j \leq R - 1$ and $-\lceil a \rceil + 1 \leq l \leq 2^j \lceil a \rceil - 1$,

$$\psi_{j,l} = \sum_{k \in \mathbb{Z}} \gamma_k \phi_{R,k}, \qquad \gamma_k = \int_{\mathbb{R}} \psi_{j,l}(x) \overline{\phi_{R,k}(x)} \mathrm{d}x.$$

Note that $\gamma_k = 0$ if k is such that measure $(\mathrm{supp}(\psi_{j,l}) \cap \mathrm{supp}(\phi_{R,k})) = 0$. Thus, $\gamma_k \neq 0$ only if

$$2^R \left(\frac{l}{2^j} \right) - \lceil a \rceil + 1 \leq k \leq 2^R \left(\frac{l + \lceil a \rceil}{2^j} \right) - 1.$$

Hence, $\gamma_k \neq 0$ only if k satisfies

$$-(2^{R-j} + 1)\lceil a \rceil + 2^{R-j} + 1 \leq k \leq (2^R + 2^{R-j})\lceil a \rceil - 2^{R-j} - 1. \tag{16}$$

Since we have shown that β_k and γ_k are non-zero only if (15) and (16) are satisfied, we have demonstrated that all elements in $V_0^{(a)} \oplus W_0^{(a)} \oplus \cdots \oplus W_{R-1}^{(a)}$ may be represented as a linear combination of elements in $\mathbb{V}_{R,a}$, and we have proved (13).

To prove (ii), note that $W_j^{(a)}$ has $(2^j + 1)\lceil a \rceil - 1$ basis elements. So, $V_0^{(a)} \oplus W_0^{(a)} \oplus \cdots \oplus W_{R-1}^{(a)}$ has precisely

$$2\lceil a \rceil - 1 + \sum_{j=0}^{R-1} \left((2^j + 1)\lceil a \rceil - 1 \right) = 2^R \lceil a \rceil + (R+1)(\lceil a \rceil - 1)$$

basis elements. Thus whenever $N = 2^R \lceil a \rceil + (R+1)(\lceil a \rceil - 1)$, it follows by the ordering in (7), that

$$\mathcal{T}_N = V_0^{(a)} \oplus W_0^{(a)} \oplus \cdots \oplus W_{R-1}^{(a)}.$$

Hence, $\varphi \in \mathcal{T}_N$ and $\|\varphi\| = 1$ implies that

$$\varphi = \sum_{|l| \leq \lceil a \rceil - 1} b_l \phi_{0,l} + \sum_{j=0}^{R-1} \sum_{l=-\lceil a \rceil + 1}^{2^j \lceil a \rceil - 1} c_{j,l} \psi_{j,l}$$

$$= \sum_{k=A_{R,1}}^{A_{R,2}} \alpha_k \phi_{R,k}$$

11

for some complex numbers $\{\alpha_k\}$, $\{b_l\}$ and $\{c_{j,l}\}$, where

$$\sum_{|l|\leq\lceil a\rceil-1} |b_l|^2 + \sum_{j=0}^{R-1}\sum_{l=-\lceil a\rceil+1}^{2^j\lceil a\rceil-1} |c_{j,l}|^2 = \sum_{j=A_{R,1}}^{A_{R,2}} |\alpha_j|^2 = 1$$

by the orthonormality of the scaling functions and wavelets.

By a similar argument to the proof of (i), it is straightforward to verify that when $2^R \geq \lceil a\rceil - 1$,

$$\text{span}\,\{\phi_{0,l}, \psi_{0,l} : l = -\lceil a\rceil + 1, \ldots, \lceil a\rceil - 2\} \oplus \bigoplus_{j=1}^{R-1} W_j^{(a)}$$

$$\subset \text{span}\,\{\phi_{R,k} : A_{R,1} \leq k \leq A_{R,2} - \lceil a\rceil\}\,.$$

Thus, it follow that for $k = A_{R,2} - \lceil a\rceil + 1, \ldots, A_{R,2}$,

$$\alpha_k = b_{\lceil a\rceil-1}\left\langle \phi_{0,\lceil a\rceil-1}, \phi_{R,k}\right\rangle + c_{0,\lceil a\rceil-1}\left\langle \psi_{0,\lceil a\rceil-1}, \phi_{R,k}\right\rangle.$$

Let $B_k = b_{\lceil a\rceil-1}\left\langle \phi_{0,\lceil a\rceil-1}, \phi_{R,k}\right\rangle$ and $C_k = c_{0,\lceil a\rceil-1}\left\langle \psi_{0,\lceil a\rceil-1}, \phi_{R,k}\right\rangle$ and suppose that both $\|\phi\|_\infty$ and $\|\psi\|_\infty$ exist.

Then, for $j = 0, \ldots, \lceil a\rceil - 1$,

$$\left|B_{A_{R,2}-j}\right| \leq \left\|\phi_{R,A_{R,2}-j}\right\| \left\|\phi_{0,\lceil a\rceil-1}\chi_{I_j}\right\|$$

$$\leq \|\phi\|_\infty \sqrt{\frac{j+1}{2^R}}$$

where $I_j = \text{supp}\phi_{R,A_{R,2}-j} \cap \text{supp}\phi_{0,\lceil a\rceil-1} \subset \left[2\lceil a\rceil - 1 - (j+1)/2^R, 2\lceil a\rceil - 1\right]$. Thus

$$\sum_{j=0}^{\lceil a\rceil-1} \left|B_{A_{R,2}-j}\right|^2 \leq \frac{\|\phi\|_\infty^2\,\lceil a\rceil(\lceil a\rceil+1)}{2^{R+1}}. \tag{17}$$

Similarly, for $j = 0, \ldots, \lceil a\rceil - 1$,

$$\left|C_{A_{R,2}-j}\right| \leq \|\psi\|_\infty \sqrt{\frac{j+1}{2^R}}$$

and

$$\sum_{j=0}^{\lceil a\rceil-1} \left|C_{A_{R,2}-j}\right|^2 \leq \frac{\|\psi\|_\infty^2\,\lceil a\rceil(\lceil a\rceil+1)}{2^{R+1}}. \tag{18}$$

Hence, if $R \geq \log_2(\lceil a\rceil - 1)$ and both $\|\phi\|_\infty$ and $\|\psi\|_\infty$ exist, then by Cauchy Schwarz, and estimates (17), (18),

$$\sum_{k=A_{R,2}-\lceil a\rceil+1}^{A_{R,2}} |\alpha_k|^2 = \sum_{k=A_{R,2}-\lceil a\rceil+1}^{A_{R,2}} |B_k + C_k|^2$$

$$\leq \sum_{k=A_{R,2}-\lceil a\rceil+1}^{A_{R,2}} \left(|B_k|^2 + |C_k|^2\right) + 2\sqrt{\sum_{k=A_{R,2}-\lceil a\rceil+1}^{A_{R,2}} |B_k|^2}\sqrt{\sum_{k=A_{R,2}-\lceil a\rceil+1}^{A_{R,2}} |C_k|^2}$$

$$= \frac{(\|\phi\|_\infty + \|\psi\|_\infty)^2\,\lceil a\rceil(\lceil a\rceil+1)}{2^{R+1}}$$

and this completes the proof of (ii).

Finally, to prove (iii) and (14), note that

$$\phi_{R,k} = \sum_{l\in\mathbb{Z}} \alpha_l \phi_{0,l} + \sum_{j=0}^{R-1}\sum_{l\in\mathbb{Z}} \beta_{j,l} \psi_{j,l}$$

where $\alpha_l = \langle \phi_{R,k}, \phi_{0,l} \rangle$ and $\beta_{j,l} = \langle \phi_{R,k}, \psi_{j,l} \rangle$. If $0 \leq k \leq (2^R - 1) \lceil a \rceil$, then

$$\operatorname{supp}(\phi_{R,k}) \subset [0, \lceil a \rceil].$$

So, if $\phi_{0,l} \notin V_0^{(a)}$, then measure $(\operatorname{supp}(\phi_{0,l}) \cap [0, \lceil a \rceil]) = 0$ and $\alpha_l = 0$. Similarly, if $\psi_{j,l} \notin W_j^{(a)}$, then measure $(\operatorname{supp}(\psi_{j,l}) \cap [0, \lceil a \rceil]) = 0$ and $\beta_{j,l} = 0$. Hence,

$$\phi_{R,k} \in V_0^{(a)} \oplus W_0^{(a)} \oplus \cdots \oplus W_{R-1}^{(a)},$$

as required. $\qquad \square$

4.2 Useful results with trigonometric polynomials

Our proof hinges on some precise estimates on the behaviour of trigonometric polynomials. These estimates are presented below.

Lemma 4.2. *Let $A_1, A_2 \in \mathbb{Z}$ be such that $A_1 \leq A_2$ and consider the trigonometric polynomial $\Phi(z) = \sum_{j=A_1}^{A_2} \alpha_j e^{2\pi i j z}$. If $L \in \mathbb{N}$ is such that $2L \geq A_2 - A_1 + 1$, then*

$$\sum_{j=0}^{2L-1} \frac{1}{2L} \left| \Phi\left(\frac{j}{2L}\right) \right|^2 = \|\Phi\|_{L^2([0,1])}^2 = \sum_{j=A_1}^{A_2} |\alpha_j|^2.$$

Proof. Given $N \in \mathbb{N}$, $x = (x_0, \ldots, x_{N-1}) \in \mathbb{Z}_N$, the Discrete Fourier Transform of x is defined by $\hat{x} = (\hat{x}_0, \ldots, \hat{x}_{N-1})$, where

$$\hat{x}_k = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} x_j e^{-\frac{2\pi i k j}{N}}.$$

Recall that $x \mapsto \hat{x}$ is a unitary operator on $l^2(\mathbb{Z}_N)$ with $\|x\|_{l^2(\mathbb{Z}_N)} = \|\hat{x}\|_{l^2(\mathbb{Z}_N)}$, where

$$\|x\|_{l^2(\mathbb{Z}_N)} = \sqrt{\sum_{j=0}^{N-1} |x_j|^2}.$$

The proof of this lemma is a direct application of the Discrete Fourier Transform, with $N = 2L$. Define $x = (x_0, \ldots, x_{N-1}) \in \mathbb{Z}_{2L}$ as follows:

$$x_{j+L} = \begin{cases} \alpha_{j+A_1+L} & -L \leq j \leq -L + A_2 - A_1 \\ 0 & \text{otherwise} \end{cases}$$

Then:

$$\begin{aligned}
\hat{x}_k &= \frac{1}{\sqrt{2L}} \sum_{j=0}^{2L-1} x_j e^{-\frac{2\pi i k j}{2L}} \\
&= \frac{1}{\sqrt{2L}} \sum_{j=-L}^{L-1} x_{j+L} e^{-\frac{2\pi i k (j+L)}{2L}} \\
&= \frac{e^{\frac{\pi i k A_1}{L}}}{\sqrt{2L}} \sum_{j=A_1}^{A_2} \alpha_j e^{\frac{2\pi i k j}{2L}} \\
&= \frac{e^{-\frac{\pi i k A_1}{L}}}{\sqrt{2L}} \Phi\left(\frac{k}{2L}\right).
\end{aligned}$$

So,

$$\sum_{k=0}^{2L-1} \frac{1}{2L} \left| \Phi\left(\frac{k}{2L}\right) \right|^2 = \sum_{k=0}^{2L-1} |\hat{x}_k|^2 = \sum_{k=0}^{2L-1} |x_k|^2 = \sum_{|k| \leq A} |\alpha_k|^2.$$

$\qquad \square$

The following theorem is a reworking of a result from [23, Proposition 1]:

Theorem 4.3. *Let $D \in \mathbb{N}$, $A \in \mathbb{R}$, $A \leq x_1 < \ldots < x_r < A + 1$ and suppose that*

$$\delta = \max_{j=1,\ldots,r} x_{j+1} - x_j < \frac{1}{2D}$$

where $x_{r+1} = x_1 + 1$. If $\Phi(x) = \sum_{j=D_1}^{D_2} \alpha_j e^{2\pi i j x}$ and $D_2 - D_1 \leq 2D$, then

$$(1 - 2\delta D) \|\Phi\|_{L^2[A,A+1)} \leq \left(\sum_{j=1}^{r} \nu_j |\Phi(x_j)|^2 \right)^{\frac{1}{2}} \leq (1 + 2\delta D) \|\Phi\|_{L^2[A,A+1)}$$

where $\nu_j = \frac{1}{2}(x_{j+1} - x_{j-1})$ and $x_0 = x_r - 1$.

4.3 Bounding the stable sampling rate

We are now ready to prove the linearity of the stable sampling rate. However, before we can present the final proof we need a couple of technical lemmas and propositions. The following lemma is an adaptation of [13, Theorem 6.3.1], the proof has simply been included for clarity:

Lemma 4.4. *Let ϕ be a compactly supported scaling function of some MRA. Let I be any closed interval of length 2π. Then, for each $\gamma \in (0,1)$, there exists N such that for all $\xi \in I$,*

$$\sum_{|l| \leq N} \left| \hat{\phi}(\xi + 2\pi l) \right|^2 \geq \gamma.$$

Proof. First note that the orthonormality of $\{\phi(\cdot - l) : l \in \mathbb{Z}\}$ is equivalent to

$$\sum_{l \in \mathbb{Z}} \left| \hat{\phi}(\xi + 2\pi l) \right|^2 = 1$$

for all $\xi \in \mathbb{R}$. In particular, for all $\xi \in I$, there exists N_ξ such that for some $\tilde{\gamma} \in (\gamma, 1)$,

$$\sum_{|l| \leq N_\xi} \left| \hat{\phi}(\xi + 2\pi l) \right|^2 \geq \tilde{\gamma}.$$

Since $\hat{\phi}$ is continuous, it follows that

$$x \mapsto \sum_{|l| \leq N_\xi} \left| \hat{\phi}(x + 2\pi l) \right|^2$$

is also continuous. Hence, there exists some δ_ξ such that

$$\sum_{|l| \leq N_\xi} \left| \hat{\phi}(\eta + 2\pi l) \right|^2 \geq \gamma, \qquad \forall \eta \in (\xi - \delta_\xi, \xi + \delta_\xi) =: U_\xi.$$

Note that $I \subset \cup_{\xi \in I} U_\xi$ and I is compact, hence, $I = \cup_{\xi \in J} U_\xi$ for some finite subset $J \subset I$. Let $N = \max \{N_\xi : \xi \in J\}$. Then for all $\xi \in I$,

$$\sum_{|l| \leq N} \left| \hat{\phi}(\xi + 2\pi l) \right|^2 \geq \gamma.$$

$\square$

Proposition 4.5. *For $R, l \in \mathbb{Z}$, let the vectors $\phi_{R,l}$, s_l and the finite rank operator P_M be defined as in 2.3. Suppose that $N_1, N_2 \in \mathbb{Z}$ and $\varphi = \sum_{l=N_1}^{N_2} \alpha_l \phi_{R,l}$, with $\alpha_l \in \mathbb{C}$, such that φ is compactly supported in $[-T_1, T_2]$. Then, for all $j \in \mathbb{Z}$,*

$$\langle \varphi, s_j \rangle = \frac{\sqrt{\epsilon}}{\sqrt{2^R}} \Phi\left(\frac{\epsilon j}{2^R}\right) \hat{\phi}\left(-\frac{2\pi \epsilon j}{2^R}\right)$$

where

$$\Phi(z) = \sum_{l=N_1}^{N_2} \alpha_l e^{2\pi i l z}.$$

In particular,

$$\|P_M \varphi\|^2 = \sum_{j=-\lfloor \frac{M}{2} \rfloor}^{\lceil \frac{M}{2} \rceil - 1} |\langle \varphi, s_j \rangle|^2$$

$$= \sum_{j=-\lfloor \frac{M}{2} \rfloor}^{\lceil \frac{M}{2} \rceil - 1} \frac{\epsilon}{2^R} \left| \Phi\left(\frac{\epsilon j}{2^R}\right) \hat{\phi}\left(-\frac{2\pi \epsilon j}{2^R}\right) \right|^2.$$

Proof. Note that $\epsilon \leq \dfrac{1}{T_1 + T_2}$ and $s_j = \sqrt{\epsilon} e^{2\pi i j \epsilon \cdot} \chi_{\left[-\frac{T_1}{\epsilon(T_1+T_2)}, \frac{T_2}{\epsilon(T_1+T_2)}\right]}$, so, by the assumption on the support of φ,

$$\langle \varphi, s_j \rangle = \sqrt{\epsilon} \int_{-\frac{T_1}{\epsilon(T_1+T_2)}}^{\frac{T_2}{\epsilon(T_1+T_2)}} \varphi(x) e^{2\pi i \epsilon j x}$$

$$= \sqrt{\epsilon} \hat{\varphi}(-2\pi \epsilon j)$$

$$= \sqrt{\epsilon} \sum_{l=N_1}^{N_2} \alpha_l \hat{\phi}_{R,l}(-2\pi \epsilon j)$$

$$= \frac{\sqrt{\epsilon}}{\sqrt{2^R}} \sum_{l=N_1}^{N_2} \alpha_l e^{\frac{2\pi i \epsilon j l}{2^R}} \hat{\phi}\left(-\frac{2\pi \epsilon j}{2^R}\right)$$

$$= \frac{\sqrt{\epsilon}}{\sqrt{2^R}} \Phi\left(\frac{\epsilon j}{2^R}\right) \hat{\phi}\left(-\frac{2\pi \epsilon j}{2^R}\right).$$

$\square$

Proposition 4.6. *Given $\gamma \in (0,1)$ and $\epsilon_1, \epsilon_2 \in (0, 1/(T_1 + T_2)]$, choose $\delta(\gamma) \in (0,1)$ and $C(\gamma) > 1$ such that*

$$\sqrt{\delta(\gamma)^2 - \frac{4}{\pi^2(C(\gamma)-1)}} - \sqrt{1 - \delta(\gamma)^2} > \gamma. \tag{19}$$

Suppose that there exists M_1 such that

$$\inf_{\varphi \in \mathcal{T}_N, \|\varphi\|=1} \left\| P_{M_1}^{\epsilon_1} \varphi \right\| \geq \delta(\gamma), \qquad N \in \mathbb{N}. \tag{20}$$

Then, the following holds:

$$\inf_{\varphi \in \mathcal{T}_N, \|\varphi\|=1} \left\| P_{M_2}^{\epsilon_2} \varphi \right\| \geq \gamma, \qquad N \in \mathbb{N}, \tag{21}$$

whenever

$$M_2 = \left\lceil \frac{C(\gamma) M_1 \epsilon_1}{\epsilon_2} \right\rceil. \tag{22}$$

Proof. Without loss of generality, in this proof, M_1 and M_2 will be even. Also, It is easy to see that $\delta(\gamma)$ and $C(\gamma)$ always exist. Observe now that for any $M_2 \in \mathbb{N}$,

$$\inf_{\varphi \in \mathcal{T}_N, \|\varphi\|=1} \left\| P_{M_2}^{\epsilon_2} \varphi \right\| \geq \inf_{\varphi \in \mathcal{T}_N, \|\varphi\|=1} \left(\left\| P_{M_2}^{\epsilon_2} P_{M_1}^{\epsilon_1} \varphi \right\| - \left\| P_{M_2}^{\epsilon_2} (P_{M_1}^{\epsilon_1})^\perp \varphi \right\| \right)$$

$$\geq \inf_{\varphi \in \mathcal{T}_N, \|\varphi\|=1} \left\| P_{M_2}^{\epsilon_2} P_{M_1}^{\epsilon_1} \varphi \right\| - \sqrt{1 - \delta(\gamma)^2},$$

where the last inequality follows from (20). Hence, to prove the proposition, it suffices to determine M_2 such that

$$\inf_{\varphi \in \mathcal{T}_N, \|\varphi\|=1} \left\| P_{M_2}^{\epsilon_2} P_{M_1}^{\epsilon_1} \varphi \right\| - \sqrt{1 - \delta(\gamma)^2} \geq \gamma.$$

In order to understand why M_2 exists, first note that $P_n^{\epsilon_2} \to P^{\epsilon_2}$ strongly as $n \to \infty$ and since $\mathcal{B} = P_{M_1}^{\epsilon_1} (\{\varphi \in \mathcal{T}_N : \|\varphi\| = 1\})$ is finite dimensional, $P_n^{\epsilon_2} \to P^{\epsilon_2}$ uniformly on $\mathcal{B}$ as $n \to \infty$. Also, $\mathcal{T} \subset \mathcal{S}^{\epsilon_1} \cap \mathcal{S}^{\epsilon_2}$, $P^{\epsilon_2} P^{\epsilon_1} \varphi = \varphi$ for all $\varphi \in \mathcal{T}_N$. So, for all $\xi > 0$, there exists M_2 such that

$$\inf_{\varphi \in \mathcal{T}_N, \|\varphi\|=1} \left\| P_{M_2}^{\epsilon_2} P_{M_1}^{\epsilon_1} \varphi \right\| \geq 1 - \sup_{\varphi \in \mathcal{T}_N, \|\varphi\|=1} \left(\left\| P_{M_2}^{\epsilon_2} P_{M_1}^{\epsilon_1} \varphi - P^{\epsilon_2} P_{M_1}^{\epsilon_1} \varphi \right\| + \left\| P^{\epsilon_2} P_{M_1}^{\epsilon_1} \varphi - \varphi \right\| \right)$$

$$\geq 1 - \xi - \sqrt{1 - \delta(\gamma)^2} \geq \gamma$$

Thus, by the choice of $\delta(\gamma)$, for sufficiently small ξ and so sufficiently large M_2,

$$\inf_{\varphi \in \mathcal{T}_N, \|\varphi\|=1} \left\| P_{M_2}^{\epsilon_2} P_{M_1}^{\epsilon_1} \varphi \right\| \geq \gamma.$$

Having established the existence of M_2 we now demonstrate that (21) follows when M_2 takes the value in (22). We begin by letting

$$\mathcal{B}_{M_2} = \left\{ l \in \mathbb{Z} : l \geq \frac{M_2}{2} \text{ or } l \leq -\frac{M_2}{2} - 1 \right\}.$$

Then

$$\left\| (P_{M_2}^{\epsilon_2})^\perp P_{M_1}^{\epsilon_1} \varphi \right\|^2 = \left\| \sum_{l \in \mathcal{B}_{M_2}} \left\langle P_{M_1}^{\epsilon_1} \varphi, s_l^{\epsilon_2} \right\rangle s_l^{\epsilon_2} \right\|^2$$

$$= \left\| \sum_{l \in \mathcal{B}_{M_2}} \left\langle \sum_{j=-\frac{M_1}{2}}^{\frac{M_1}{2}-1} \left\langle \varphi, s_j^{\epsilon_1} \right\rangle s_j^{\epsilon_1}, s_l^{\epsilon_2} \right\rangle s_l^{\epsilon_2} \right\|^2$$

$$= \sum_{l \in \mathcal{B}_{M_2}} \left| \sum_{j=-\frac{M_1}{2}}^{\frac{M_1}{2}-1} \left\langle \varphi, s_j^{\epsilon_1} \right\rangle \left\langle s_j^{\epsilon_1}, s_l^{\epsilon_2} \right\rangle \right|^2 \tag{23}$$

$$\leq \sum_{l \in \mathcal{B}_{M_2}} \left(\sum_{j=-\frac{M_1}{2}}^{\frac{M_1}{2}-1} \left| \left\langle \varphi, s_j^{\epsilon_1} \right\rangle \right|^2 \sum_{j=-\frac{M_1}{2}}^{\frac{M_1}{2}-1} \left| \left\langle s_l^{\epsilon_1}, s_l^{\epsilon_2} \right\rangle \right|^2 \right)$$

$$\leq \sum_{l \in \mathcal{B}_{M_2}} \sum_{j=-\frac{M_1}{2}}^{\frac{M_1}{2}-1} \left| \left\langle s_j^{\epsilon_1}, s_l^{\epsilon_2} \right\rangle \right|^2.$$

Let $\epsilon_+ = \max\{\epsilon_1, \epsilon_2\}$, and note that

$$
\begin{aligned}
\left|\langle s_j^{\epsilon_1}, s_l^{\epsilon_2}\rangle\right| &= \left|\sqrt{\epsilon_1 \epsilon_2} \int_{-\frac{1}{2\epsilon_+}}^{\frac{1}{2\epsilon_+}} e^{2\pi i \epsilon_1 j x} e^{-2\pi i \epsilon_2 l x} \mathrm{d}x\right| \\
&= \left|\sqrt{\epsilon_1 \epsilon_2} \frac{\exp\left(\dfrac{\pi i(\epsilon_1 j - \epsilon_2 l)}{\epsilon_+}\right) - \exp\left(\dfrac{-\pi i(\epsilon_1 j - \epsilon_2 l)}{\epsilon_+}\right)}{2\pi i(\epsilon_1 j - \epsilon_2 l)}\right| \\
&= \sqrt{\epsilon_1 \epsilon_2} \left|\frac{\sin\left(\dfrac{\pi(\epsilon_1 j - \epsilon_2 l)}{\epsilon_+}\right)}{\pi(\epsilon_1 j - \epsilon_2 l)}\right|.
\end{aligned}
\tag{24}
$$

So, by substituting (24) into (23), we have that

$$
\begin{aligned}
\left\|(P_{M_2}^{\epsilon_2})^\perp P_{M_1}^{\epsilon_1} \varphi\right\|^2 &\leq \sum_{l \in \mathcal{B}_{M_2}} \sum_{j=-\frac{M_1}{2}}^{\frac{M_1}{2}-1} \epsilon_1 \epsilon_2 \left|\frac{\sin\left(\dfrac{\pi(\epsilon_1 j - \epsilon_2 l)}{\epsilon_+}\right)}{\pi(\epsilon_1 j - \epsilon_2 l)}\right|^2 \\
&\leq \frac{\epsilon_1 \epsilon_2}{\pi^2} \sum_{l \in \mathcal{B}_{M_2}} \sum_{j=-\frac{M_1}{2}}^{\frac{M_1}{2}-1} \frac{1}{|\epsilon_1 j - \epsilon_2 l|^2}.
\end{aligned}
$$

Suppose that $M_2 = \lceil C(\gamma) M_1 \epsilon_1 / \epsilon_2 \rceil$ where $C(\gamma)$ stems from (19), then

$$
\begin{aligned}
\left\|(P_{M_2}^{\epsilon_2})^\perp P_{M_1}^{\epsilon_1} \varphi\right\|^2 &\leq \frac{\epsilon_1 \epsilon_2}{\pi^2} M_1 \sum_{l > \frac{M_2}{2}} \frac{2}{\left|\epsilon_1 \frac{M_1}{2} - \epsilon_2 l\right|^2} \\
&\leq \frac{\epsilon_1 \epsilon_2}{\pi^2} M_1 \frac{4}{\epsilon_2(-\epsilon_1 M_1 + \epsilon_2 M_2)} \\
&\leq \frac{\epsilon_1}{\pi^2} M_1 \frac{4}{(\epsilon_1 M_1(C(\gamma) - 1))} \\
&\leq \frac{4}{\pi^2(C(\gamma) - 1)}.
\end{aligned}
$$

Therefore,

$$
\left\|P_{M_2}^{\epsilon_2} P_{M_1}^{\epsilon_1} \varphi\right\|^2 = \left\|P_{M_1}^{\epsilon_1} \varphi\right\|^2 - \left\|(P_{M_2}^{\epsilon_2})^\perp P_{M_1}^{\epsilon_1} \varphi\right\|^2 \geq \delta(\gamma)^2 - \frac{4}{\pi^2(C(\gamma) - 1)}.
$$

whenever

$$
M_2 = \left\lceil \frac{C(\gamma) M_1 \epsilon_1}{\epsilon_2} \right\rceil.
$$

Hence,

$$
\inf_{\varphi \in \mathcal{T}_N, \|\varphi\|=1} \left\|P_{M_2}^{\epsilon_2} \varphi\right\| \geq \sqrt{\delta(\gamma)^2 - \frac{4}{\pi^2(C(\gamma) - 1)}} - \sqrt{1 - \delta(\gamma)^2} > \gamma
$$

by the choice of $\delta(\gamma)$ and $C(\gamma)$ in (19). $\qquad\square$

4.4   The proof

We are now ready to present the proof of Theorem 3.1.

Proof of Theorem 3.1. Recall that the reconstruction space $\mathcal{S}$ is defined for sampling distances $\epsilon \in (0, 1/(T_1 + T_2)]$. Fix $\epsilon = 1/(T_1 + T_2 + \lceil a \rceil)$ and let $N \leq N_R$, with $R \in \mathbb{N}$.

Suppose it is known that for all $\theta \in (1, \infty)$, there exists $S_\theta \in \mathbb{N}$, independent of R, such that for

$$M_\theta = \frac{S_\theta 2^{R+1}}{\epsilon},$$

we have that

$$\inf_{\varphi \in \mathcal{T}_N, \|\varphi\|=1} \left\| P_{M_\theta}^\epsilon \varphi \right\| \geq \frac{1}{\theta}.$$

Then, by Proposition 4.6, given $\theta \in (1, \infty)$ and any sampling distance $\epsilon_1 \in (0, 1/(T_1 + T_2)]$, if $C(\theta)$ and $\delta(\theta)$ satisfy

$$\sqrt{\delta(\theta)^2 - \frac{4}{\pi^2 (C(\theta) - 1)}} - \sqrt{1 - \delta(\theta)^2} > \frac{1}{\theta},$$

then we have that

$$\inf_{\varphi \in \mathcal{T}_N, \|\varphi\|=1} \left\| P_M^{\epsilon_1} \varphi \right\| \geq \frac{1}{\theta} \qquad \text{whenever} \qquad M = \left\lceil \frac{C(\theta) \epsilon M_{\delta(\theta)}}{\epsilon_1} \right\rceil = \left\lceil \frac{C(\theta) S_{\delta(\theta)} 2^{R+1}}{\epsilon_1} \right\rceil.$$

Hence, it is sufficient to prove this theorem for $\epsilon = 1/(T_1 + T_2 + \lceil a \rceil)$.

Recall $A_{R,1}$ and $A_{R,2}$ from (12), then by the choice of N and Lemma 4.1,

$$\mathcal{T}_N \subset \text{span}\left\{ \phi_{R,k} : A_{R,1} \leq k \leq A_{R,2} \right\}. \tag{25}$$

Let $\varphi \in \mathcal{T}_N$ such that $\|\varphi\| = 1$. Then, by (25), we have that

$$\varphi = \sum_{l=A_{R,1}}^{A_{R,2}} \alpha_l \phi_{R,l}$$

and

$$\sum_{l=A_{R,1}}^{A_{R,2}} |\alpha_l|^2 = 1. \tag{26}$$

Moreover, φ is compactly supported in $[-T_1, T_2]$ since it is a linear combination of elements in Ω_a. Thus, by Proposition 4.5,

$$\|P_M \varphi\|^2 = \sum_{j=-\lfloor \frac{M}{2} \rfloor}^{\lceil \frac{M}{2} \rceil - 1} |\langle \varphi, s_j \rangle|^2$$

$$= \sum_{j=-\lfloor \frac{M}{2} \rfloor}^{\lceil \frac{M}{2} \rceil - 1} \frac{\epsilon}{2^R} \left| \Phi\left(\frac{\epsilon j}{2^R}\right) \hat{\phi}\left(-\frac{2\pi \epsilon j}{2^R}\right) \right|^2$$

where

$$\Phi(z) = \sum_{l=A_{R,1}}^{A_{R,2}} \alpha_l e^{2\pi i l z}. \tag{27}$$

Let $L = 2^R / \epsilon$, then L is some even integer since $1/\epsilon = T_1 + T_2 + \lceil a \rceil = 4\lceil a \rceil - 2 \in \mathbb{N}$. Furthermore, suppose that $M/2 = SL$ for some $S \in \mathbb{N}$. Then:

$$\|P_M \varphi\|^2 = \sum_{j=0}^{L-1} \sum_{k=-S}^{S-1} \frac{\epsilon}{2^R} \left| \Phi\left(\frac{\epsilon}{2^R}(j + kL)\right) \right|^2 \left| \hat{\phi}\left(-\frac{2\pi \epsilon}{2^R}(j + kL)\right) \right|^2 \tag{28}$$

$$= \sum_{j=0}^{L-1} \frac{1}{L} \left| \Phi\left(\frac{j}{L}\right) \right|^2 \sum_{k=-S}^{S-1} \left| \hat{\phi}\left(-\frac{2\pi j}{L} - 2\pi k\right) \right|^2. \tag{29}$$

By applying Lemma 4.4 to the interval $[-2\pi, 0]$, given any $\theta \in (1, \infty)$, there exists $S_\theta \in \mathbb{N}$ such that for all $j = 0, \ldots L-1$,

$$\sum_{k=-S_\theta}^{S_\theta-1} \left| \hat{\phi}\left(-\frac{2\pi j}{L} - 2\pi k\right) \right|^2 \geq \frac{1}{\theta^2}.$$

Since

$$L = \frac{2^R}{\epsilon} = 2^R(4\lceil a \rceil - 2) > 2^R(3\lceil a \rceil - 2) + \lceil a \rceil - 1 = A_{R,2} - A_{R,1} + 1,$$

Lemma 4.2 (via (26) and (27)) implies that

$$\sum_{j=0}^{L-1} \frac{1}{L} \left| \Phi\left(\frac{j}{L}\right) \right|^2 = 1.$$

Thus,

$$\|P_M \varphi\|^2 \geq \frac{1}{\theta^2} \sum_{j=0}^{L-1} \frac{1}{L} \left| \Phi\left(\frac{j}{L}\right) \right|^2 = \frac{1}{\theta^2}.$$

Hence, for $N \leq N_R$ and $M = S_\theta 2^{R+1}/\epsilon$, where S_θ depends only on the scaling function ϕ and θ,

$$C_{N,M} = \inf_{\varphi \in \mathcal{T}_N, \|\varphi\|=1} \|P_M \varphi\| \geq \frac{1}{\theta},$$

and the theorem is proven. $\qquad\square$

5 Proof of Theorem 3.3

The proof of Theorem 3.3 hinges on the following proposition.

Proposition 5.1. *Let $N \geq N_R$, and suppose $M = c2^R$ for $c < \frac{1}{\epsilon}$, then $C_{N,M} \to 0$ exponentially as $N \to \infty$.*

With this result at hand the proof of Theorem 3.3 is straightforward.

Proof of Theorem 3.3. Suppose that $\eta_G < \frac{1}{\epsilon \lceil a \rceil}$. Then, by Corollary 3.2 and Proposition 5.1, $\kappa(G_N)$ cannot be bounded. Moreover, from [6], for $M = \Theta_G(N_R)$, we have that

$$\kappa(G_{N_R}) \geq \kappa(F_{N_R,M}) \geq \frac{1}{C_{N_R,M}}.$$

Hence, by Proposition 5.1, $\kappa(G_{N_R})$ becomes exponentially large as N_R grows. $\qquad\square$

The rest of this section is devoted to the proof of Proposition 5.1, however, before we can state the proof, we need the following results on trigonometric polynomials and Chebyshev polynomials from [22].

Proposition 5.2. *Let $\omega \in [0, \pi]$ and consider the following function, defined over $[-\pi, \pi]$:*

$$Q_{n,\omega}(z) = Q_{2n}\left(\frac{\sin(z/2)}{\sin(\omega/2)}\right)$$

where $Q_{2n}(x) = \cos(2n \arccos x)$ for $x \in [-1, 1]$ is the Chebyshev polynomial of degree $2n$. Then the following holds:

(i) $Q_{n,\omega}$ is a trigonometric polynomial in z of degree n, i.e. $Q_{n,\omega}(z) = \sum_{|j| \leq n} \alpha_j e^{izj}$.

(ii) $\|Q_{n,\omega}\|_{L^\infty[-\omega,\omega]} = 1$.

(iii) For $\omega \in [\pi/2, \pi)$, there exists constants $c_1, c_2 > 0$ such that

$$\exp(c_1 n(\pi - \omega)) \leq \|Q_{n,\omega}\|_{L^\infty[-\pi,\pi]} = Q_{n,\omega}(\pi) \leq \exp(c_2 n(\pi - \omega)).$$

Proof of Proposition 5.1. The goal is to use Proposition 5.2, and the first part of the proof is a setup for that. In particular, let $M = c2^R$ for some $c < 1/\epsilon$ and $R \geq 1$. By Lemma 4.1, if $0 \leq l \leq (2^R - 1)\lceil a \rceil$, then

$$\phi_{R,l} \in V_0^{(a)} \oplus W_0^{(a)} \oplus \cdots \oplus W_{R-1}^{(a)}.$$

Hence, for $N \geq N_R$ and $p = (2^{R-1} - 1)\lceil a \rceil$, it follows that

$$\mathcal{T}_N \supset V_0^{(a)} \oplus W_0^{(a)} \oplus \cdots \oplus W_{R-1}^{(a)} \supset \{\phi_{R,l} : 0 \leq l \leq 2p\}.$$

Thus, we get that

$$(C_{N,M})^2 = \inf_{\substack{\|\varphi\|=1 \\ \varphi \in \mathcal{T}_N}} \|P_M \varphi\|^2 \leq \inf\{\|P_M \varphi\|^2 : \sum_{l=0}^{2p} |\beta_l|^2 = 1, \varphi = \sum_{l=0}^{2p} \beta_l \phi_{R,l}\}.$$

Hence, by Proposition 4.5 and the choice of $M = c2^R$, it follows that

$$(C_{N,M})^2 \leq \inf\left\{ \sum_{j=-\lfloor \frac{M}{2} \rfloor}^{\lceil \frac{M}{2} \rceil - 1} \frac{\epsilon}{2^R} \left| \Phi\left(\frac{\epsilon j}{2^R}\right) \hat{\phi}\left(-\frac{2\pi\epsilon j}{2^R}\right) \right|^2 : \Phi(z) = \sum_{l=0}^{2p} \beta_l e^{2\pi izl}, \sum_{l=0}^{2p} |\beta_l|^2 = 1 \right\}$$

$$\leq \left\|\hat{\phi}\right\|_{L^\infty[-\pi c\epsilon, \pi c\epsilon]}^2 \inf\left\{ \sum_{j=-\lfloor \frac{M}{2} \rfloor}^{\lceil \frac{M}{2} \rceil - 1} \frac{\epsilon}{2^R} \left| \Phi\left(\frac{\epsilon j}{2^R}\right) \right|^2 : \Phi(z) = \sum_{l=0}^{2p} \beta_l e^{2\pi izl}, \|\Phi\|_{L^2[0,1]}^2 = 1 \right\} \quad (30)$$

$$= \left\|\hat{\phi}\right\|_{L^\infty[-\pi c\epsilon, \pi c\epsilon]}^2 \inf\left\{ \sum_{j=-\lfloor \frac{M}{2} \rfloor}^{\lceil \frac{M}{2} \rceil - 1} \frac{\epsilon}{2^R} \left| \Phi\left(\frac{\epsilon j}{2^R}\right) \right|^2 : \Phi(z) = \sum_{|l| \leq p} \beta_l e^{2\pi izl}, \|\Phi\|_{L^2[0,1]}^2 = 1 \right\}$$

The last equality above is a consequence of the following: For $\Phi(z) = \sum_{l=0}^{2p} \beta_l e^{2\pi izl}$,

$$\sum_{j=-\lfloor \frac{M}{2} \rfloor}^{\lceil \frac{M}{2} \rceil - 1} \frac{\epsilon}{2^R} \left| \Phi\left(\frac{\epsilon j}{2^R}\right) \right|^2 = \sum_{j=-\lfloor \frac{M}{2} \rfloor}^{\lceil \frac{M}{2} \rceil - 1} \frac{\epsilon}{2^R} \left| \sum_{|l| \leq p} \beta_{l+p} e^{2\pi i\epsilon jl/2^R} e^{2\pi i\epsilon jp/2^R} \right|^2$$

$$= \sum_{j=-\lfloor \frac{M}{2} \rfloor}^{\lceil \frac{M}{2} \rceil - 1} \frac{\epsilon}{2^R} \left| \sum_{|l| \leq p} \beta_{l+p} e^{2\pi i\epsilon jl/2^R} \right|^2.$$

Note that we have carried out this shift in indices in order to later show that the infimum is taken over a set of functions which include those of the form $Q_{n,\omega}$ defined in Proposition 5.2. From (30) it follows easily that

$$(C_{N,M})^2 \leq c\epsilon \left\|\hat{\phi}\right\|_{L^\infty[-\pi c\epsilon, \pi c\epsilon]}^2 \inf\left\{ \|\Phi\|_{L^\infty[-\pi c\epsilon, \pi c\epsilon]}^2 : \Phi(z) = \sum_{|l| \leq p} \beta_l e^{izl}, \|\Phi\|_{L^2[-\pi,\pi]}^2 = 1 \right\},$$

where we have again used that $M = c2^R$. Also, by Cauchy Schwarz, for $\Phi(z) = \sum_{|l| \leq p} \beta_l e^{izl}$,

$$|\Phi(z)|^2 \leq (2p+1) \sum_{|l| \leq p} |\beta_l|^2 = (2p+1) \|\Phi\|_{L^2[-\pi,\pi]}^2.$$

So,

$$\|\Phi\|_{L^\infty[-\pi,\pi]}^2 = 2p+1 \implies \|\Phi\|_{L^2[-\pi,\pi]}^2 \geq 1.$$

Thus,

$$(C_{N,M})^2 \le c\epsilon \left\|\hat{\phi}\right\|^2_{L^\infty[-\pi c\epsilon, \pi c\epsilon]} \inf\left\{ \|\Phi\|^2_{L^\infty[-\pi c\epsilon, \pi c\epsilon]} : \Phi(z) = \sum_{|l|\le p} \beta_l e^{izl}, \|\Phi\|^2_{L^2[-\pi,\pi]} \ge 1 \right\}$$

$$\le D_R \inf\left\{ \|\Phi\|^2_{L^\infty[-\pi c\epsilon, \pi c\epsilon]} : \Phi(z) = \sum_{|l|\le p} \beta_l e^{izl}, \|\Phi\|^2_{L^\infty[-\pi,\pi]} = 1 \right\},$$

(31)

where

$$D_R = (2p+1)c\epsilon \left\|\hat{\phi}\right\|^2_{L^\infty[-\pi c\epsilon, \pi c\epsilon]}.$$

Having established (31) we can now make use of Proposition 5.2. Indeed, for $\omega \in [\pi/2, \pi)$, let

$$q_\omega = \frac{Q_{p,\omega}}{\|Q_{p,\omega}\|_{L^\infty[-\pi,\pi]}},$$

where $Q_{p,\omega}$ is defined in Proposition 5.2. Then, by Proposition 5.2,

$$q_\omega \in \left\{ \Phi : \Phi(z) = \sum_{|l|\le p} \beta_l e^{izl}, \|\Phi\|^2_{L^\infty[-\pi,\pi]} = 1 \right\},$$

(32)

and there exists some constant $\eta > 0$, independent of p, such that

$$\|q_\omega\|_{L^\infty[-\omega,\omega]} \le \frac{1}{\|Q_{p,\omega}\|_{L^\infty[-\pi,\pi]}} \le \exp(-\eta p(\pi - \omega)).$$

(33)

We now split the proof into two cases, and we will show that $C_{N,M} \to 0$ exponentially as $R \to \infty$ when

$$\text{Case 1:} \quad c \in \left[\frac{1}{2\epsilon}, \frac{1}{\epsilon}\right), \qquad \text{Case 2:} \quad c \in \left(0, \frac{1}{2\epsilon}\right).$$

Case 1: By (31), (32) and (33) (and recalling the value of $p = (2^{R-1} - 1)\lceil a \rceil$),

$$C_{N,M} \le D_R \|q_{\pi c\epsilon}\|^2_{L^\infty[-\pi c\epsilon, \pi c\epsilon]}$$

$$\le (2^R \lceil a \rceil - 2\lceil a \rceil + 1)c\epsilon \left\|\hat{\phi}\right\|^2_{L^\infty[-\pi c\epsilon, \pi c\epsilon]} \exp((-\eta\pi(1 - c\epsilon)(2^R \lceil a \rceil - 2\lceil a \rceil)))$$

Thus, we have shown that $C_{N,M}$ decays exponentially as $N \to \infty$ in the first case scenario.

Case 2: Clearly, we still have exponential decay in $C_{N,M}$, since, again by (31), (32) and (33),

$$C_{N,M} \le D_R \|q_{\pi/2}\|^2_{L^\infty[-\pi c\epsilon, \pi c\epsilon]}$$

$$\le D_R \|q_{\pi/2}\|^2_{L^\infty[-\pi/2, \pi/2]}$$

$$\le (2^R \lceil a \rceil - 2\lceil a \rceil + 1)c\epsilon \left\|\hat{\phi}\right\|^2_{L^\infty[-\pi c\epsilon, \pi c\epsilon]} \exp(-\eta\pi(2^{R-1} \lceil a \rceil - \lceil a \rceil)).$$

$\square$

6 Proof of Theorem 3.5

We are now ready to present the proof of Theorem 3.5.

Remark 6.1. In the construction of Daubechies wavelets [13], the scaling function ϕ is defined such that

$$\hat{\phi}(\xi) := \prod_{s=1}^\infty m_0\left(\frac{\xi}{2^s}\right)$$

21

where

$$m_0(\xi) = \left(\frac{1 + e^{-i\xi}}{2}\right)^N \mathcal{L}(\xi)$$

for some $N \in \mathbb{N}$ and $\mathcal{L}$ is such that

$$|\mathcal{L}(\xi)|^2 = \sum_{k=0}^{N-1} \binom{N-1+k}{k} \sin^{2k}\left(\frac{\xi}{2}\right).$$

Note that in this case, $|m_0(\xi)| > 0$ for all $\xi \in (-\pi, \pi)$ and since $\hat{\phi}(0) = 1$, there exists $K \in \mathbb{N}$ such that $\left|\hat{\phi}\left(\xi/2^K\right)\right| > 0$ for all $\xi \in (-2\pi, 2\pi)$. Hence,

$$\hat{\phi}(\xi) = \hat{\phi}\left(\frac{\xi}{2^K}\right) \prod_{s=1}^{K} m_0\left(\frac{\xi}{2^s}\right) \neq 0 \qquad \text{for all } \xi \in (-2\pi, 2\pi). \tag{34}$$

Proof of Theorem 3.5. Recall that $\epsilon \in (0, 1/(T_1 + T_2)]$ and from Proposition 5.1, for all $c < 1/\epsilon$, $C_{N_R,M}$ will tend to 0 exponentially if $M < 2^R/\epsilon$. So, for each $\theta \in (1, \infty)$, there exists $R_0 \in \mathbb{N}$ such that for all $R \geq R_0$,

$$\Theta(N_R, \theta) \geq \left\lceil \frac{2^R}{\epsilon} \right\rceil.$$

Hence, if it is known that there exists R_1 and $\theta \in (1, \infty)$ such that for all $R \geq R_1$

$$\Theta(N_R, \theta) \leq \left\lceil \frac{2^R}{\epsilon} \right\rceil. \tag{35}$$

then for such θ and all $R \geq \max\{R_0, R_1\}$

$$\Theta(N_R, \theta) = \left\lceil \frac{2^R}{\epsilon} \right\rceil.$$

So, it remains to show the existence of $\theta \in (1, \infty)$ such that (35) holds. Let $\varphi \in \mathcal{T}_{N_R}$ such that $\|\varphi\| = 1$. Then, by Lemma 4.1, we have that

$$\varphi = \sum_{l=A_{R,1}}^{A_{R,2}} \alpha_l \phi_{R,l}, \qquad \sum_{l=A_{R,1}}^{A_{R,2}} |\alpha_l|^2 = 1, \tag{36}$$

where $A_{R,1}$ and $A_{R,2}$ are as defined in (12). Now, let $M = \lceil 2^R/\epsilon \rceil$. Then, by Proposition 4.5,

$$\|P_M \varphi\|^2 = \sum_{j=-\lfloor \frac{M}{2} \rfloor}^{\lceil \frac{M}{2} \rceil - 1} \frac{\epsilon}{2^R} \left| \Phi\left(\frac{\epsilon j}{2^R}\right) \hat{\phi}\left(-\frac{2\pi \epsilon j}{2^R}\right) \right|^2 \geq \gamma_1^2 \sum_{j=-\lfloor \frac{M}{2} \rfloor}^{\lceil \frac{M}{2} \rceil - 1} \frac{\epsilon}{2^R} \left| \Phi\left(\frac{\epsilon j}{2^R}\right) \right|^2 \tag{37}$$

where

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

and

$$\gamma_1 = \inf_{\xi \in [-\pi \epsilon M 2^{-R}, \pi \epsilon M 2^{-R}]} \left| \hat{\phi}(\xi) \right| \geq \inf_{\xi \in [-(1+\epsilon 2^{-R})\pi, (1+\epsilon 2^{-R})\pi]} \left| \hat{\phi}(\xi) \right| > 0.$$

Note that $\gamma_1 > 0$ due to (34) and $\gamma_1 = \inf_{\xi \in [-\pi, \pi]} \left| \hat{\phi}(\xi) \right|$ whenever $2^R/\epsilon \in \mathbb{Z}$. Hence, it remains to obtain a positive lower bound for

$$\sum_{j=-\lfloor \frac{M}{2} \rfloor}^{\lceil \frac{M}{2} \rceil - 1} \frac{\epsilon}{2^R} \left| \Phi\left(\frac{\epsilon j}{2^R}\right) \right|^2. \tag{38}$$

22

We will split the proof into several cases. The case of $a = 1$ is treated separately mainly for pedagogical reasons as the proof is simpler in this case.

Case 1: $a = 1$ and $1/\epsilon \in \mathbb{N}$.

Since $a = 1$, we have that $2^R/\epsilon \geq 2^R(3\lceil a \rceil - 2) = A_{R,2} - A_{R,1} + 1$ and for $1/\epsilon \in \mathbb{N}$ (in which case, $M = 2^R/\epsilon$ is even), Lemma 4.2 gives that

$$\sum_{j=-\frac{M}{2}}^{\frac{M}{2}-1} \frac{\epsilon}{2^R} \left| \Phi\left(\frac{\epsilon j}{2^R}\right) \right|^2 = 1.$$

So, given any $R \in \mathbb{N}$ and

$$\theta \geq \left(\inf_{\xi \in [-\pi, \pi]} \left| \hat{\phi}(\xi) \right| \right)^{-1},$$

we have that

$$\Theta(N_R, \theta) \leq 2^R/\epsilon.$$

Case 2: $a = 1$ and $1/\epsilon \notin \mathbb{N}$.

In this case we must have $\epsilon < 1/(3\lceil a \rceil - 2)$, and an application of Theorem 4.3 to Φ with $r = \lceil 2^R/\epsilon \rceil$, $2D = 2^R(3\lceil a \rceil - 2) \geq A_{R,2} - A_{R,1}$, $\delta = \epsilon/2^R$ and

$$x_j = \frac{\epsilon}{2^R}\left(-\left\lfloor \frac{M}{2} \right\rfloor + j - 1 \right), \qquad j = 1, \ldots, r$$

gives that

$$\sum_{j=-\left\lfloor \frac{M}{2} \right\rfloor}^{\left\lceil \frac{M}{2} \right\rceil - 1} \frac{\epsilon}{2^R} \left| \Phi\left(\frac{\epsilon j}{2^R}\right) \right|^2 \geq (1 - \epsilon(3\lceil a \rceil - 2))^2 > 0.$$

So, given any $R \in \mathbb{N}$ and

$$\theta \geq \left((1 - \epsilon(3\lceil a \rceil - 2)) \inf_{\xi \in [-(1+\epsilon)\pi, (1+\epsilon)\pi]} \left| \hat{\phi}(\xi) \right| \right)^{-1},$$

we have that

$$\Theta(N_R, \theta) \leq 2^R/\epsilon.$$

Case 3: $a > 1$ and $2^R/\epsilon \in \mathbb{N}$ for some R.

When $a > 1$, Lemma 4.2 and Theorem 4.3 cannot be applied directly because $2^R/\epsilon$ may be less than $A_{R,2} - A_{R,1} + 1 = 2^R(3\lceil a \rceil - 2) + \lceil a \rceil - 1$ and so, we will first decompose Φ into two other trigonometric polynomials for which we can obtain bounds.

We now let $R \geq \log_2(\lceil a \rceil - 1)$. Since ϕ and ψ are continuous and compactly supported, $\|\phi\|_\infty$ and $\|\psi\|_\infty$ exist. So, by Lemma 4.1 (ii) and Proposition 4.5,

$$\sum_{j=-\left\lfloor \frac{M}{2} \right\rfloor}^{\left\lceil \frac{M}{2} \right\rceil - 1} \frac{\epsilon}{2^R} \left| \Phi\left(\frac{\epsilon j}{2^R}\right) \right|^2 = \sum_{j=-\left\lfloor \frac{M}{2} \right\rfloor}^{\left\lceil \frac{M}{2} \right\rceil - 1} \frac{\epsilon}{2^R} \left| \Phi_1\left(\frac{\epsilon j}{2^R}\right) + \Phi_2\left(\frac{\epsilon j}{2^R}\right) \right|^2$$

where

$$\Phi_1(z) = \sum_{j=A_{R,1}}^{A_{R,2}-\lceil a \rceil} \alpha_j e^{2\pi i z j}, \qquad \Phi_2(z) = \sum_{j=A_{R,2}-\lceil a \rceil+1}^{A_{R,2}} \alpha_j e^{2\pi i z j},$$

and

$$\sum_{j=A_{R,2}-\lceil a \rceil+1}^{A_{R,2}} |\alpha_j|^2 \leq \frac{(\|\phi\|_\infty + \|\psi\|_\infty)^2 \lceil a \rceil (\lceil a \rceil + 1)}{2^{R+1}}. \tag{39}$$

So, as argued in (37),
$$\|P_M\varphi\|^2 \geq \gamma_1^2\left(C_{\Phi_1}^2 + C_{\Phi_2}^2 - 2C_{\Phi_1}C_{\Phi_2}\right)$$
where
$$C_{\Phi_s} = \sqrt{\sum_{j=-\lfloor\frac{M}{2}\rfloor}^{\lceil\frac{M}{2}\rceil-1}\frac{\epsilon}{2^R}\left|\Phi_s\left(\frac{\epsilon j}{2^R}\right)\right|^2}, \qquad s = 1, 2. \tag{40}$$

If $2^R/\epsilon \in \mathbb{N}$ for some R, then we may apply Lemma 4.2 to Φ_1 since
$$(A_{R,2} - A_{R,1}) - \lceil a\rceil + 1 = 2^R(3\lceil a\rceil - 2) \leq 2^R/\epsilon,$$
and to Φ_2 since
$$A_{R,2} - (A_{R,2} - \lceil a\rceil + 1) + 1 = \lceil a\rceil \leq 2^R/\epsilon.$$
We thus obtain,
$$C_{\Phi_1}^2 = \sum_{j=A_{R,1}}^{A_{R,2}-\lceil a\rceil}|\alpha_j|^2$$
and
$$C_{\Phi_2}^2 = \sum_{j=A_{R,2}-\lceil a\rceil+1}^{A_{R,2}}|\alpha_j|^2.$$
Note that

$$C_{\Phi_1}^2 + C_{\Phi_2}^2 - 2C_{\Phi_1}C_{\Phi_2} = \sum_{j=A_{R,1}}^{A_{R,2}}|\alpha_j|^2 - 2\left(\sum_{j=A_{R,1}}^{A_{R,2}-\lceil a\rceil}|\alpha_j|^2\right)^{1/2}\left(\sum_{j=A_{R,2}-\lceil a\rceil+1}^{A_{R,2}}|\alpha_j|^2\right)^{1/2}$$
$$\geq 1 - 2\left(\frac{(\|\phi\|_\infty + \|\psi\|_\infty)(\lceil a\rceil + 1)}{2^{(R+1)/2}}\right)$$

by (36) and (39). Hence, for all $\mu \in (0,1)$, there exists R_0 such that for all $R \geq R_0$,

$$\|P_M\varphi\|^2 \geq \inf_{\xi\in[-\pi,\pi]}\left|\hat{\phi}(\xi)\right|^2\left(1 - \frac{(\|\phi\|_\infty + \|\psi\|_\infty)(\lceil a\rceil + 1)}{2^{(R-1)/2}}\right) > \mu\inf_{\xi\in[-\pi,\pi]}\left|\hat{\phi}(\xi)\right|^2,$$

and so given any
$$\theta > \left(\inf_{\xi\in[-\pi,\pi]}\left|\hat{\phi}(\xi)\right|\right)^{-1},$$
there exists R_0 such that for all $R \geq R_0$,
$$\Theta(N_R, \theta) \leq 2^R/\epsilon.$$

Case 4: $a > 1$ and $2^R/\epsilon \notin \mathbb{N}$ for all $R \in \mathbb{N}$.

In this case, $\epsilon < 1/(3\lceil a\rceil - 2)$ and as in Case 3, obtaining appropriate estimates for C_{Φ_1} and C_{Φ_2} defined in (40) will provide the required lower bound for (38).

In the case of C_{Φ_1}, applying Theorem 4.3 to Φ_1 with $r = \lceil 2^R/\epsilon\rceil$, $2D = 2^R(3\lceil a\rceil - 2) - 2 = (A_{R,2} - \lceil a\rceil) - A_{R,1}$, $\delta = \epsilon/2^R$ and
$$x_j = \frac{\epsilon}{2^R}\left(-\left\lfloor\frac{M}{2}\right\rfloor + j - 1\right), \qquad j = 1, \ldots, r$$
gives that

$$(1 - \delta_1)\sum_{j=A_{R,1}}^{A_{R,2}-\lceil a\rceil}|\alpha_j|^2 \leq \left(\sum_{j=1}^{r}\nu_j|\Phi_1(x_j)|^2\right)^{\frac{1}{2}} \leq (1 + \delta_1)\sum_{j=A_{R,1}}^{A_{R,2}-\lceil a\rceil}|\alpha_j|^2$$

24

where $\delta_1 = \epsilon(3\lceil a \rceil - 2 - 1/2^{R-1}) < \epsilon(3\lceil a \rceil - 2) < 1$, $\nu_j = (x_{j+1} - x_{j-1})/2$ and $x_0 = x_r - 1$. Note that $\epsilon/2^{R+1} \leq \nu_j \leq 2^R/\epsilon$. Hence, by (40),

$$(1 - \delta_1)^2 \sum_{j=A_{R,1}}^{A_{R,2}-\lceil a \rceil} |\alpha_j|^2 \leq C_{\Phi_1}^2 \leq 2(1 + \delta_1)^2 \sum_{j=A_{R,1}}^{A_{R,2}-\lceil a \rceil} |\alpha_j|^2.$$

In the case of C_{Φ_2}, applying Theorem 4.3 to Φ_2 with $r = \lceil 2^R/\epsilon \rceil$, $2D = 2\lceil(\lceil a \rceil - 1)/2\rceil \geq A_{R,2} - (A_{R,2} - \lceil a \rceil + 1)$, $\delta = \epsilon/2^R$ and

$$x_j = \frac{\epsilon}{2^R}\left(-\left\lfloor \frac{M}{2} \right\rfloor + j - 1\right), \qquad j = 1, \ldots, r$$

gives that

$$(1 - \delta_2) \sum_{j=A_{R,2}-\lceil a \rceil+1}^{A_{R,2}} |\alpha_j|^2 \leq \left(\sum_{j=1}^r \nu_j |\Phi_1(x_j)|^2\right)^{\frac{1}{2}} \leq (1 + \delta_2) \sum_{j=A_{R,2}-\lceil a \rceil+1}^{A_{R,2}} |\alpha_j|^2$$

where $\delta_2 \leq \epsilon(\lceil a \rceil + 1)/2^R < 1$, $\nu_j = (x_{j+1} - x_{j-1})/2$ and $x_0 = x_r - 1$. Again, $\epsilon/2^{R+1} \leq \nu_j \leq 2^R/\epsilon$. So,

$$(1 - \delta_2)^2 \sum_{j=A_{R,2}-\lceil a \rceil+1}^{A_{R,2}} |\alpha_j|^2 \leq C_{\Phi_2}^2 \leq 2(1 + \delta_2)^2 \sum_{j=A_{R,2}-\lceil a \rceil+1}^{A_{R,2}} |\alpha_j|^2.$$

Hence,

$$\|P_M\varphi\|^2 \geq \inf_{\xi \in [-(1+\epsilon2^{-R})\pi, (1+\epsilon2^{-R})\pi]} \left|\hat{\phi}(\xi)\right|^2 \left((1 - \delta_1)^2 - (1 + \delta_2)(1 + \delta_1)\frac{(\|\phi\|_\infty + \|\psi\|_\infty)(\lceil a \rceil + 1)}{2^{(R-3)/2}}\right)$$

$$\to (1 - \epsilon(3\lceil a \rceil - 2))^2 \inf_{\xi \in [-\pi, \pi]} \left|\hat{\phi}(\xi)\right|^2 > 0 \quad \text{as} \quad R \to \infty.$$

So, for all $\mu \in (0, 1)$, there exists R_0 such that for all $R \geq R_0$,

$$\|P_M\varphi\|^2 \geq \mu(1 - \epsilon(3\lceil a \rceil - 2))^2 \inf_{\xi \in [-\pi, \pi]} \left|\hat{\phi}(\xi)\right|^2 > 0.$$

and for all

$$\theta > \left((1 - \epsilon(3\lceil a \rceil - 2)) \inf_{\xi \in [-\pi, \pi]} \left|\hat{\phi}(\xi)\right|\right)^{-1},$$

there exists R_0 such that for all $R \geq R_0$,

$$\Theta(N_R, \theta) \leq 2^R/\epsilon.$$

$\square$

7 Numerical Examples

In this section we provide numerical examples to illustrate the behaviour of the stable sampling rate as well as demonstrating sharpness of our estimates. We also show that, because of the linearity of the stable sampling rate, any convergence properties of a series expansion of a function in a particular wavelet basis will be inherited (up to a constant) by the generalized sampling reconstruction based on Fourier samples. In other words, as discussed in Section 2.2, generalized sampling is, up to a constant, an oracle for the wavelet reconstruction problem.
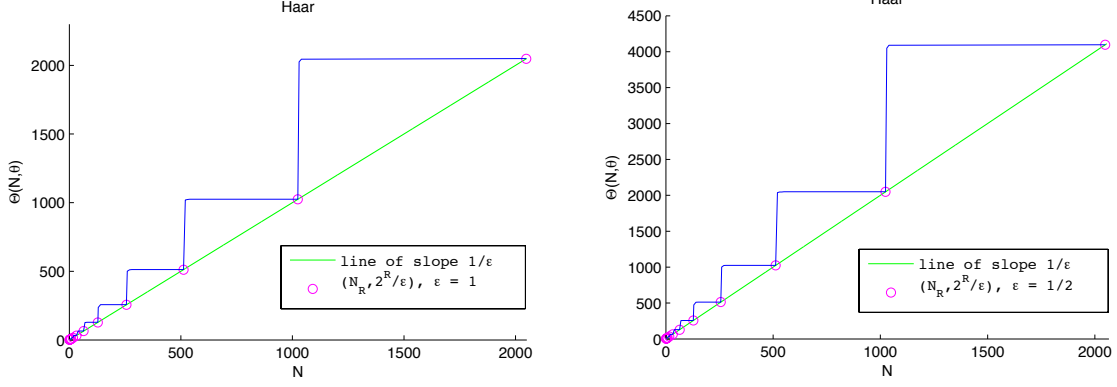
Figure 1: The figure displays the stable sampling rate $\Theta(N, \theta)$ in blue for the Haar wavelet with Fourier samples for $\theta = \pi/2$ at a sampling distance $\epsilon = 1$ (left) and $\epsilon = 1/2$ (right).

7.1 Sharpness of the stable sampling rate estimates

Before we demonstrate the sharpness of our estimates numerically, let us recall the result from Theorem 3.5. In particular, for $N_R = 2^R \lceil a \rceil + (R+1)(\lceil a \rceil - 1)$ and when $2^R/\epsilon \in \mathbb{Z}$, then for all sufficiently large R

$$\Theta(N_R, \theta) = \frac{2^R}{\epsilon}, \tag{41}$$

where

$$\theta > \left(\inf_{\xi \in [-\pi, \pi]} \left| \hat{\phi}(\xi) \right| \right)^{-1}, \tag{42}$$

and ϕ is the scaling function of the wavelet. Recall also the asymptotic result

$$\lim_{R \to \infty} \frac{\Theta(N_R, \theta)}{N_R} = \frac{1}{\epsilon \lceil a \rceil}. \tag{43}$$

In this section we demonstrate these sharp results numerically. We consider the Haar wavelet (supported on $[0, 1]$), the Daubechies-4 wavelet (supported in $[0, 3]$), and the Daubechies-6 wavelet (supported in $[0, 5]$).

For the Haar wavelet, the Fourier sampling distance must be $\epsilon \leq 1$. Since

$$\left(\inf_{\xi \in [-\pi, \pi]} \left| \hat{\phi}(\xi) \right| \right)^{-1} = \frac{\pi}{2},$$

in this case, from the proof of Theorem 3.5, we see that (41) applies whenever $\theta \geq \frac{\pi}{2}$.

Figure 1 shows the growth of $\Theta(N, \pi/2)$ for sampling distances $\epsilon = 1$ and $\epsilon = 1/2$ respectively. We observe from the figure that

$$\Theta(N_R, \pi/2) = 2^R, \qquad \Theta(N_R, \pi/2) = 2^{R+1}, \qquad R \in \mathbb{N}$$

respectively, exactly as suggested in (41). Moreover, by (43), we have that

$$\Theta(N_R, \pi/2) \sim N_R \frac{1}{\epsilon \lceil a \rceil},$$

which is verified in Figure 1 via the green line.

In the case of the DB4 and DB6 wavelets, the Fourier sampling space must be of sampling distance $\epsilon \leq 1/7$ and $\epsilon \leq 1/13$ respectively . Computationally we may observe that

$$\theta_1^{-1} = 0.684 < \inf_{x \in [-\pi, \pi]} \left| \hat{\phi}_{DB4}(x) \right|, \qquad \theta_2^{-1} = 0.698 < \inf_{x \in [-\pi, \pi]} \left| \hat{\phi}_{DB6}(x) \right|,$$
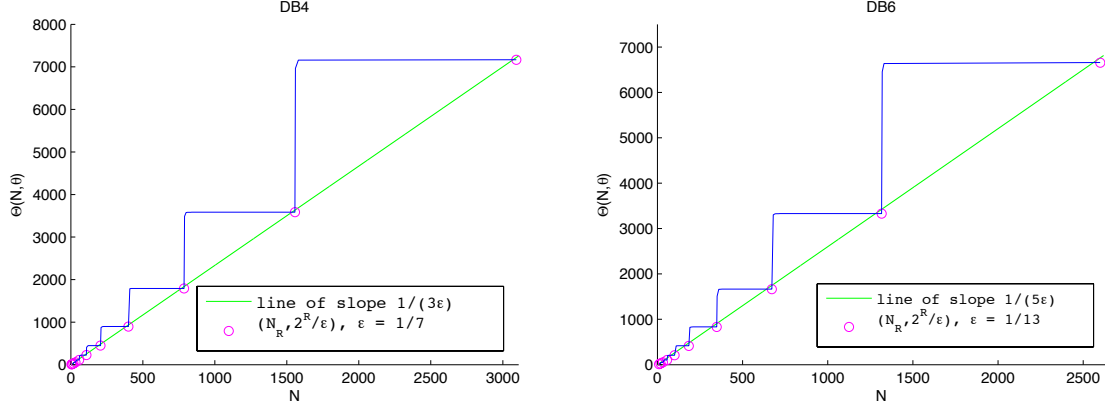
Figure 2: The figure displays the stable sampling rate $\Theta(N, \theta_1)$ and $\Theta(N, \theta_2)$ in blue for the Daubechies-4 wavelet (left) and the Daubechies-6 wavelet (right) with Fourier samples at a sampling distance $\epsilon = 1/7$ and $\epsilon = 1/13$ respectively.

where ϕ_{DB4} and ϕ_{DB6} are the scaling function of the DB4 and DB6 wavelets respectively. So, again, as displayed in Figure 2, we have

$$\Theta(N_R, \theta_1) = 7 \cdot 2^R, \qquad \Theta(N_R, \theta_2) = 13 \cdot 2^R, \qquad R \in \mathbb{N},$$

which confirms (41). Moreover, by (43), we have that

$$\Theta(N_R, \theta_1) \sim N_R \frac{1}{3\epsilon}, \qquad \Theta(N_R, \theta_2) \sim N_R \frac{1}{5\epsilon}$$

which is verified in Figure 2 via the green line.

Remark 7.1. Note that

$$\Theta(N_R, \theta) < \Theta(N, \theta) \leq \Theta(N_{R+1}, \theta), \quad N_R < N \leq N_{R+1}.$$

The staircase effect witnessed in the figures suggests that the upper bound is in fact an equality. Hence, although the stable sampling rate is linear for all N, from the point of view of the stable sampling rate at least, there is nothing to be gained from allowing $N \neq N_R$.

7.2   Generalized sampling and function reconstruction

In this section we demonstrate the power of generalized sampling in practice. Given the result on the stable sampling rate above we have now full control over how to balance the number of Fourier samples versus the number of wavelet coefficients. In these experiments with Daubechies wavelets we will, in order to get a stable and convergent reconstruction, use the predicted value from (43), namely, the number of samples M should asymptotically satisfy

$$M = \frac{N}{\epsilon \lceil a \rceil},$$

where N is the number of coefficients to be computed, ϵ is the sampling distance and a is the maximum value of the support of the mother wavelet.

We will also demonstrate, as predicted by Theorem 3.3, that failure of satisfying the stable sampling rate gives a completely unstable and even non-convergent reconstruction. In this case we will chose the disastrous value

$$M = cN, \qquad c < \frac{1}{\epsilon \lceil a \rceil},$$

which causes the condition number of the algorithm to blow up exponentially. It also makes the constant in the error bound blow up at the same rate and thus one gets a non-convergent method.

(M, N, α)	$\|f - f_M\|_{L^2}$	$\|f - \tilde{f}_{N,M}\|_{L^2}$	$-\frac{\log \|f - \tilde{f}_{N,M}\|}{\log N}$	Wavelet
$(260, 100, 2)$	7.6×10^{-3}	2.4×10^{-4}	1.81	DB 6
$(479, 184, 2)$	3.1×10^{-3}	9.3×10^{-5}	1.78	DB 6
$(905, 348, 2)$	9.6×10^{-4}	9.7×10^{-6}	1.97	DB 6
$(481, 206, 2.5)$	7.9×10^{-3}	6.1×10^{-6}	2.25	DB 4
$(934, 400, 2.5)$	3.1×10^{-3}	9.1×10^{-7}	2.32	DB 4
$(1834, 786, 2.5)$	1.5×10^{-3}	8.6×10^{-8}	2.44	DB 4
$(256, 256, 3)$	2.1×10^{-2}	2.8×10^{-7}	2.72	Haar
$(512, 512, 3)$	1.2×10^{-2}	5.0×10^{-8}	2.69	Haar
$(1024, 1024, 3)$	9.6×10^{-3}	2.6×10^{-9}	2.85	Haar

Table 1: The table shows the error of the reconstructions based on classical Fourier series, f_M, as well as generalized sampling $\tilde{f}_{N,M}$ with different types of wavelets. Note that both f_M and $f_{N,M}$ use exactly the same samples.

(M, N, α)	$\|f - f_M\|_{L^2}$	$\|f - \tilde{f}_{N,M}\|_{L^2}$	Noise Level ε	Wavelet
$(481, 206, 2.5)$	9.5×10^{-2}	1.4×10^{-1}	1.0×10^{-1}	DB 4
$(934, 400, 2.5)$	7.4×10^{-2}	9.8×10^{-2}	7.2×10^{-2}	DB 4
$(1834, 786, 2.5)$	5.0×10^{-3}	7.0×10^{-3}	5.0×10^{-3}	DB 4
$(256, 256, 3)$	2.1×10^{-2}	6.4×10^{-4}	4.2×10^{-4}	Haar
$(512, 512, 3)$	1.2×10^{-2}	4.8×10^{-5}	3.1×10^{-5}	Haar
$(1024, 1024, 3)$	9.6×10^{-3}	3.8×10^{-6}	2.5×10^{-6}	Haar

Table 2: The table shows the error of the reconstructions based on classical Fourier series, f_M, as well as generalized sampling $\tilde{f}_{N,M}$ with different types of wavelets, where the samples are contaminated with noise. Note that both f_M and $f_{N,M}$ use exactly the same samples.

The test functions will be of the form

$$f = \sum_{j=1}^{\infty} \beta_j \varphi_j, \qquad |\beta_j| = \mathcal{O}(j^{-\alpha}), \quad \alpha > 1, \tag{44}$$

where the φ_js are different types of Daubechies wavelets. We let $\tilde{f}_{N,M}$ denote the function that is constructed with generalized sampling using M Fourier coefficients as samples and then reconstructing by computing N approximate wavelet coefficients. In other words, $\tilde{f}_{N,M}$ is the solution to

$$\left\langle P_M \tilde{f}_{N,M}, \varphi_j \right\rangle = \left\langle P_M f, \varphi_j \right\rangle, \quad j = 1, \ldots, N, \tag{45}$$

where P_M is the projection onto the sampling space $\mathcal{S}_M$, where $\mathcal{S}_M$ is defined in (10). As a comparison we will use the truncated Fourier series

$$f_M = P_M f = \sum_{j=1}^{M} \langle f, s_j \rangle s_j.$$

We will sometimes assume that the samples $\langle f, s_j \rangle$ are contaminated with noise and thus we observe

$$\xi = \{\langle f, s_1 \rangle, \ldots, \langle f, s_M \rangle\} + v, \qquad \|v\| = \varepsilon,$$

for some noise level $\varepsilon \geq 0$. Note that f_M and $\tilde{f}_{N,M}$ use exactly the same information sampled.

The fact that

$$\|f - f_M\| = \|P_M^{\perp} f\|, \qquad \|f - \tilde{f}_{N,M}\| \leq \theta \|Q_N^{\perp} f\|,$$

together with (44) show that the reconstruction created by generalized sampling will asymptotically outperform the reconstruction based on the truncated Fourier series on the types of functions described

(M,α)	$\|f - f_M\|_{L^2}$	$\|f - \tilde{f}_{M/c,M}\|_{L^2}$	$\|f - \tilde{f}_{M/c_1,M}\|_{L^2}$	Noise Level ε	Wavelet
$(481,3)$	9.5×10^{-2}	2.3×10^{-7}	2.2×10^{-2}	0	DB 4
$(934,3)$	1.6×10^{-3}	4.6×10^{-8}	2.4×10^{2}	0	DB 4
$(7220,3)$	5.1×10^{-4}	3.0×10^{-10}	3.5×10^{6}	0	DB 4
$(481,3)$	9.5×10^{-2}	9.3×10^{-6}	6.6×10^{1}	1.0×10^{-5}	DB 4
$(934,3)$	1.6×10^{-3}	9.3×10^{-6}	5.3×10^{5}	1.0×10^{-5}	DB 4
$(7220,3)$	5.1×10^{-4}	1.1×10^{-5}	1.0×10^{11}	1.0×10^{-5}	DB 4

Table 3: The table shows the error of the reconstructions based on classical Fourier series, f_M, as well as generalized sampling $\tilde{f}_{N,M}$ with $N = M/c$ and $N = M/c_1$, with noiseless and noisy data. Note that f_M and $f_{N,M}$ use exactly the same samples.

in (44). In particular, since $\tilde{f}_{N,M}$ is quasi-optimal, we have that

$$-\frac{\log \|f - \tilde{f}_{N,M}\|}{\log N} \approx \alpha$$

for large N. This is verified in Table 1.

Also, observe in Table 2 the predicted stability of generalized sampling. In particular, the condition number of generalized sampling is equal to θ which in the case of this experiment is $\pi/2$ for the Haar case and 1.46 for the DB4. In Table 3 we demonstrate that if the number of samples M does not satisfies the stable sampling rate we get an unstable and non-convergent method. In particular, we compare the choices

$$M = cN, \quad c = \frac{1}{\epsilon \lceil a \rceil}, \qquad M = c_1 N, \quad c_1 = 0.95c.$$

As verified in Table 3 the latter choice gives disastrous results.

8 Conclusions and future work

The aim of this paper has been to show that generalized sampling solves the problem of computing wavelet coefficients in a stable and accurate manner from Fourier samples. In particular, we have proved that the stable sampling rate is linear for all wavelets, and thus generalized sampling is, up to a constant factor, an optimal method for this problem. Furthermore, we have shown that, for the class of perfect reconstruction methods, any attempt to lower the stable sampling ratio necessarily results in exponential ill-conditioning.

There are a number of directions for future work. First, we have not considered the efficient implementation of generalized sampling for this problem. The main issue herein is the complexity of computing the reconstruction. As discussed in [4], this is $\mathcal{O}(NM)$ in general (i.e. $\mathcal{O}(N^2)$ whenever the stable sampling rate is linear, such as in the wavelet case), since one is required to solve a dense $M \times N$ well-conditioned least-squares problem. However, for wavelets at least, the corresponding matrix is extremely structured. It is therefore likely that this value can be reduced to $\mathcal{O}(N(\log N)^2)$ or even $\mathcal{O}(N \log N)$.

Another topic we have not addressed is that of sparsity. The generalized sampling framework studied in this paper guarantees recovery of all signals in a wavelet basis from their Fourier samples. However, suppose now that the signal to be recovered is in fact sparse in the wavelet domain, or compressible (i.e. well approximated by a sparse signal). Can this property be exploited to reduce the number of Fourier samples used in recovering the signal?

An abstract framework for sparsity-exploiting generalized sampling was recently developed in [1]. Note that this is intimately related to the field of compressed sensing [8, 20, 24]. However, unlike the standard compressed sensing framework, which models signals as finite length vectors in vector spaces, the framework developed in [1] models signals as elements of separable, infinite-dimensional Hilbert spaces. As discussed in [1], the infinite-dimensional model can often be more faithful to the original problem, leading to significant potential benefits. For example, in the MRI problem – which is best modelled by the continuous, as opposed to the discrete, Fourier transform – it allows one to avoid the issues raised in Remark 1.1.

The aim of future work in this direction is to combine the results of this paper with the framework of [1] so as to obtain a full theory for wavelet reconstructions of compressible signals from Fourier samples. In particular, the analogue of the stable sampling rate in [1], known as the *balancing property*, must be first analysed. Moreover, compressed sensing relies on so-called incoherence between sampling and reconstruction bases. This must also be estimated.

Alongside this, there are several other extensions to be pursued. These include the generalization to higher dimensions, as well as the extension to more exotic objects such as contourlets [14, 38], curvelets [9, 10] and shearlets [11, 12, 31]. Another open problem involves the question of Fourier samples taken non-uniformly. In this paper we have considered only Fourier samples taken on a regular lattice. However, non-uniform sampling patterns are more common in applications. The question of generalized sampling for non-uniform Fourier samples was considered previously in [6] within the setting of Fourier frames. We believe that the key results proved herein regarding the behaviour of the stable sampling rate can be extended to this case.

References

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

[2] 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.

[3] B. Adcock and A. C. Hansen. A generalized sampling theorem for stable reconstructions in arbitrary bases. *J. Fourier Anal. Appl. (to appear)*, 2012.

[4] 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.

[5] B. Adcock, A. C. Hansen, E. Herrholz, and G. Teschke. Generalized sampling: extension to frames and ill-posed problems. *Technical report NA2011/17, DAMTP, University of Cambridge*, 2011.

[6] B. Adcock, A. C. Hansen, and C. Poon. Beyond consistent reconstructions: optimality and sharp bounds for generalized sampling, and application to the uniform resampling problem. *Preprint*, 2012.

[7] T. Blumensath. Sampling theorems for signals from the union of finite-dimensional linear subspaces. *IEEE Trans. Inform. Theory*, 55(4):1872–1882, 2009.

[8] E. J. Candès. An introduction to compressive sensing. *IEEE Signal Process. Mag.*, 25(2):21–30, 2008.

[9] E. J. Candès and D. L. Donoho. Recovering edges in ill-posed inverse problems: optimality of curvelet frames. *Ann. Statist.*, 30(3):784–842, 2002.

[10] E. J. Candès and D. L. Donoho. New tight frames of curvelets and optimal representations of objects with piecewise C^2 singularities. *Comm. Pure Appl. Math.*, 57(2):219–266, 2004.

[11] S. Dahlke, G. Kutyniok, P. Maass, C. Sagiv, H.-G. Stark, and G. Teschke. The uncertainty principle associated with the continuous shearlet transform. *Int. J. Wavelets Multiresolut. Inf. Process.*, 6(2):157–181, 2008.

[12] S. Dahlke, G. Kutyniok, G. Steidl, and G. Teschke. Shearlet coorbit spaces and associated banach frames. *Applied and Computational Harmonic Analysis*, 27(2):195–214, 2009.

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

[14] M. N. Do and M. Vetterli. The Contourlet Transform: An Efficient Directional Multiresolution Image Representation. *IEEE Transactions on Image Processing*, 14(12):2091–2106, 2005.

[15] T. Dvorkind and Y. C. Eldar. Robust and consistent sampling. *IEEE Signal Process. Letters*, 16(9):739–742, 2009.

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

[17] 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.

[18] Y. C. Eldar. Robust recovery of signals from a structured union of subspaces. *IEEE Trans. Inform. Theory*, 55(11):5302–5316, 2009.

[19] Y. C. Eldar and T. Dvorkind. A minimum squared-error framework for generalized sampling. *IEEE Trans. Signal Process.*, 54(6):2155–2167, 2006.

[20] Y. C. Eldar and G. Kutyniok, editors. *Compressed Sensing: Theory and Applications*. Cambridge University Press, 2012.

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

[22] T. Erdélyi. Remez-type inequalities on the size of generalized polynomials. *J. London Math. Soc*, 45:255–264, 1992.

[23] H. G. Feichtinger, K. Gröchenig, and T. Strohmer. Efficient numerical methods in non-uniform sampling theory. *Numerische Mathematik*, 69:423–440, 1995. 10.1007/s002110050101.

[24] M. Fornasier and H. Rauhut. Compressive sensing. In *Handbook of Mathematical Methods in Imaging*, pages 187–228. Springer, 2011.

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

[26] 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.

[27] D. M. Healy and J. B. Weaver. Two applications of wavelet transforms in Magnetic Resonance Imaging. *IEEE Trans. Inform. Theory*, 38(2):840–862, 1992.

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

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

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

[31] G. Kutyniok, J. Lemvig, and W.-Q. Lim. Compactly supported shearlets. In M. Neamtu and L. Schumaker, editors, *Approximation Theory XIII: San Antonio 2010*, volume 13 of *Springer Proceedings in Mathematics*, pages 163–186. Springer New York, 2012.

[32] W. E. Kyriakos, W. S. Hoge, and D. Mitsouras. Generalized encoding through the use of selective excitation in accelerated parallel MRI. *NMR Biomed.*, 19:379–392, 2006.

[33] A. F. Laine. Wavelets in temporal and spatial processing of biomedical images. *Annu. Rev. Biomed. Eng.*, 02:511–550, 2000.

[34] Y. M. Lu and M. N. Do. A theory for sampling signals from a union of subspaces. *IEEE Trans. Signal Process.*, 56(6):2334–2345, 2008.

[35] R. Nowak. Wavelet-based Rician noise removal for Magnetic Resonance Imaging. *IEEE Trans. Image Proc.*, 8:1408–19, 1998.

[36] L. P. Panych. Theoretical comparison of Fourier and wavelet encoding in Magnetic Resonance Imaging. *IEEE Trans. Med. Imaging*, 15(2):141–153, 1996.

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

[38] D. D.-Y. Po and M. N. Do. Directional multiscale modeling of images using the contourlet transform. *Trans. Img. Proc.*, 15(6):1610–1620, June 2006.

[39] M. Unser. Splines: A perfect fit for signal and image processing. *IEEE Signal Process. Mag.*, 16(6):22–38, 1999.

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

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

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

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

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

[45] J. B. Weaver, Y. Xu, D. M. Healy, and J. R. Driscoll. Filtering MR images in the wavelet transform domain. *Magn. Reson. Med.*, 21:288–295, 1991.

[46] J. B. Weaver, Y. Xu, D. M. Healy, and J. R. Driscoll. Wavelet-encoded MR imaging. *Magn. Reson. Med.*, 24:275–287, 1992.