Sharp bounds, optimality and a geometric interpretation for generalised sampling in Hilbert spaces

Ben Adcock
Department of Mathematics
Simon Fraser University
Burnaby, BC V5A 1S6
Canada

Anders C. Hansen
DAMTP, Centre for Mathematical Sciences
University of Cambridge
Wilberforce Rd, Cambridge CB3 0WA
United Kingdom

Abstract

Generalised sampling is a recently developed linear framework for sampling and reconstruction in separable Hilbert spaces. It allows one to recover any element in an arbitrary basis given its samples with respect to any other basis. Unlike more common approaches for this problem, such as the consistent reconstruction technique of Eldar et al, it leads to completely stable numerical methods possessing both guaranteed recovery and accuracy.

The purpose of this paper is twofold. First, we derive new bounds for the error committed by generalised sampling. These estimates improve existing results, and are shown to be sharp. In doing so, we also establish a new result concerning oblique projections in Hilbert spaces, with applications to other methods, such as the aforementioned consistent reconstruction technique. Second, we consider the topic of optimality. We prove a result demonstrating that no projection-type method can outperform generalised sampling in an asymptotic sense. Moreover, for the important example of the recovery of a smooth function to high accuracy from its Fourier coefficients, we illustrate that generalised sampling is optimally stable amongst all possible methods. That is to say, for this example at least, no other method can be both more stable and more accurate than generalised sampling. As we discuss, these results have potentially significant consequences for the design of practical reconstruction techniques in image and signal processing applications.

The key to obtaining these results is to first give an appropriate geometric interpretation of generalised sampling in terms of projections onto particular subspaces. As we demonstrate, not only can generalised sampling be interpreted in this way, the important constants which determine both numerical stability and accuracy can be identified with particular subspace angles. Such constants can therefore be easily computed, and hence stability and accuracy guaranteed *a priori*.

1 Introduction

A central theme in sampling theory is the recovery of a signal or an image from a collection of its samples. Mathematically, this is modelled in a separable Hilbert space H, with the samples of the unknown signal $f \in \mathrm{H}$ being of the form

$$\hat{f}_j = \langle f, \psi_j \rangle, \quad j = 1, 2, \ldots,$$

for some collection of vectors $\{\psi_j\}_{j=1}^{\infty}$ belonging to H (here $\langle \cdot, \cdot \rangle$ is the inner product on H). Typically, the *sampling* vectors $\{\psi_j\}_{j=1}^{\infty}$ form a Riesz basis for their span $\mathrm{S} = \mathrm{span}\{\psi_1, \psi_2, \ldots\}$.

One of the most common, and arguably one of the most important, examples of this type of sampling is the recovery of a function f with compact support and Fourier transform $\mathcal{F}f$ from the equispaced samples $\{\mathcal{F}f(j\epsilon)\}_{j\in\mathbb{Z}}$. In this case $\langle \cdot, \cdot \rangle$ is the standard $\mathrm{L}^2(\mathbb{R})$ inner product, $\psi_j(x) = \exp(2\pi \mathrm{i} \epsilon j x)$ and the values $\hat{f}_j$ coincide with $\mathcal{F}f(j\epsilon)$ (here $\epsilon \leq \frac{1}{2T}$, where $\mathrm{supp}(f) \subseteq [-T, T]$). This is precisely the type of sampling encountered in Magnetic Resonance Imaging (MRI), for example. The celebrated Shannon Sampling Theorem [29, 38] guarantees that both f and $\mathcal{F}f$ can be recovered exactly via the infinite sums

$$f(x) = \epsilon \sum_{j=-\infty}^{\infty} \mathcal{F}f(j\epsilon) \mathrm{e}^{2\pi \mathrm{i} \epsilon j x}, \qquad \mathcal{F}f(t) = \sum_{j=-\infty}^{\infty} \mathcal{F}f(j\epsilon) \mathrm{sinc}\left(\frac{t + j\epsilon}{\epsilon}\right). \tag{1.1}$$

Note that Shannon's sampling theorem is often stated in terms of $g = \mathcal{F}f$. However, in many applications, MRI included, it is f that is of primary interest, rather than $\mathcal{F}f$. Hence we shall retain this form throughout.

Shannon's Sampling Theorem is a mainstay of modern signal processing, and has become one of the most important theorems in the mathematics of information [38]. Whilst its influence cannot be understated, it transpires that Shannon's theorem is rarely used in practice [16, 38], the principle reason being that the infinite sum (1.1) typically converges intolerably slowly for real-world signals [31, 38]. In computations one can only process finite amounts of information, hence these sums must be truncated, leading to the approximations

$$f_n(x) = \sum_{j=-\lfloor n \rfloor}^{\lfloor n \rfloor - 1} \mathcal{F}f(j\epsilon) \mathrm{e}^{2\pi \mathrm{i} \epsilon j x}, \qquad \mathcal{F}f_n(t) = \sum_{j=-\lfloor n \rfloor}^{\lfloor n \rfloor - 1} \mathcal{F}f(j\epsilon) \mathrm{sinc}\left(\frac{t + j\epsilon}{\epsilon}\right). \tag{1.2}$$

The slow convergence of these truncated sums means that an infeasibly large number of samples $\mathcal{F}f(j\epsilon)$ must be taken to ensure that f_n and $\mathcal{F}f_n$ are good approximations to f and $\mathcal{F}f$ respectively. The former is also typically polluted by unsightly Gibbs oscillations.

It is important to note the following at this point: throughout this paper the samples of f will be fixed, and cannot be altered. Typically, this situation occurs when sampling scheme is specified by some physical device; Fourier samples in Magnetic Resonance Imaging (MRI), for example. Thus, the question is, given such fixed samples, how can one reconstruct f to better accuracy than (1.2)?

This question is not new, and there has been much interest in the last several decades in alternative approaches for reconstruction. In practice, many signals, whilst not well-represented by complex exponentials (i.e. by (1.1)–(1.2)), are well-represented in terms of a different collection of vectors $\{\phi_j\}$ [16, 38]. That is, for a given $f \in \mathrm{H}$, we have

$$f = \sum_{j=1}^{\infty} \alpha_j \phi_j, \tag{1.3}$$

where the coefficients $\alpha_j \to 0$ rapidly, or, in some special cases, $\alpha_j \neq 0$ only for a small number of values j (i.e. f is *sparse* in $\{\phi_j\}$). As a result, one needs only a comparatively small number of the values α_j to represent f to high accuracy. Given such additional knowledge, it is natural to ask how one can recover f in terms of this system of vectors using only the fixed samples $\hat{f}_j = \langle f, \psi_j \rangle$. Consistent reconstruction—introduced by Unser & Aldroubi [39, 40] and later generalised significantly by Eldar et al [12, 13, 14, 17]—is a linear technique designed specifically for this problem.

Whilst the consistent reconstruction technique has proved successful, and is widely used in practice, there are a number of issues. As discussed in [1, 3, 15, 27], consistent reconstructions have the significant drawback of being typically neither numerically stable or convergent as the number of samples increases. In particular, one can quite easily devise examples for which consistent reconstructions either fail to converge, or are extremely unstable, or both [1]. In other words, the approximate coefficients $\tilde{\alpha}_1, \tilde{\alpha}_2, \ldots$ computed by this approach need not bear any substantial resemblance to those in the exact representation (1.3), nor may their computation be stable.

To overcome these issues, a new approach, referred to as *generalised sampling*, was introduced in [1, 3, 4]. This method is based on a fundamentally new viewpoint on sampling and reconstruction, the so-called *stable sampling rate*, which conveniently combines notions of sampling with those of numerical stability. In §3 we recap this technique.

With this in mind, the purpose of this paper is twofold. First we provide sharp estimates for generalised sampling, thereby improving those given in [1, 3, 4]. In doing so, we also obtain a new result for oblique projections in Hilbert spaces, which, as a corollary, also confirms sharpness of a number of existing estimates for consistent reconstructions. Such estimates show that consistent reconstructions must fail in some cases, as we discuss in detail. Conversely, as we both prove and illustrate numerically, generalised sampling offers guaranteed recovery.

Implementation of generalised sampling relies on first determining the stable sampling rate. Roughly speaking, this quantity measures how many coefficients α_j in the basis $\{\phi_j\}$ can be stably and accurately computed from a given number of samples. As we demonstrate, the stable sampling rate can be interpreted completely in terms of subspace angles, and therefore computed numerically. Hence, both stability and accuracy of generalised sampling can in practice be guaranteed *a priori*.

In the second part of this paper we consider the topic of optimality. Specifically, we show that generalised sampling cannot be improved upon in a certain sense, and therefore is the optimal reconstruction technique amongst all so-called *projection-type* methods. Furthermore, for the important example of the recovery of a (piecewise) smooth function of one variable from its Fourier coefficients in a basis of (piecewise) polynomials, we give substantial evidence towards the conjecture that generalised sampling is an optimal stable method. Specifically, no stable method can produce an asymptotically better (i.e. more accurate) reconstruction.

Both the sharp bounds and optimality results of this paper rely on first deriving a new geometric interpretation of generalised sampling in terms of oblique projections onto specific subspaces. The interpretation of consistent reconstructions in terms of oblique projections is well understood [13, 14, 17, 39]. By extending this notion to generalised sampling, not only do we obtain the aforementioned results, we also give a geometric argument for how this framework overcomes the issues with stability and convergence inherent to consistent reconstructions. An equivalent argument can also be made in terms of finite sections of infinite-dimensional operators. This indicates how questions in sampling and reconstruction can be related to the issue of how to discretise certain operators (such a connection was first demonstrated in [1]).

It is worth noting that the optimality results of this paper can be viewed in the following alternative way. If no other method can be both more stable and more accurate than generalised sampling, then the stable sampling rate (the quantity which ensures stability and accuracy of generalised sampling) is not intrinsic, but in fact *universal*. That is, the number of coefficients in the basis $\{\phi_j\}$ that can be accurately and stably reconstructed from a given number of samples by *any* method is determined precisely by this quantity, and hence cannot be improved upon without compromising either stability or accuracy or both. This universality has potential ramifications for the design of practical reconstruction frameworks in image and signal processing applications. Having said this, we stress that such a notion remains a conjecture, and we currently have no proof of universality in a general context. In §5 we discuss future work in this direction.

Generalised sampling, the subject of this paper, is a linear technique designed to recover any signal from its samples. The development of compressed sensing techniques has had a huge impact in the reconstruction of sparse signals via subsampling. Generalised sampling can also be judiciously combined with compressed sensing techniques to achieve subsampling, whenever the signal f is sparse (or compressible) in the basis $\{\phi_j\}$. The importance of this development is that it allows for compressed sensing of analog (i.e. infinite-dimensional) signals which have sparse (or compressible) information content in some analog basis. Conversely, the standard compressed sensing techniques apply to inherently finite-dimensional signals, i.e. vectors in $\mathbb{C}^d$. We refer to [2] for details, and [6, 11, 41] for related methodologies based on analog, but finite information content, models for signals.

The outline of the remainder of this paper is as follows. In §2 we introduce linear consistent reconstructions and discuss their drawbacks. Generalised sampling, which overcomes these shortfalls, is introduced in §3 and sharp bounds given. Finally, in §4 we consider optimality.

2 Consistent reconstructions and oblique projections

We first review the consistent reconstruction technique of [13, 14, 17, 39, 40], as well as its connection to oblique projections. To this end, suppose that $\{\psi_j\}_{j=1}^{\infty}$ is a collection of vectors in a separable Hilbert space H (over $\mathbb{C}$) that form a Riesz basis for the closed *sampling space* S. In other words, there exist constants $d_1, d_2 > 0$ such that

$$d_1 \|\alpha\|_{l^2(\mathbb{N})} \leq \left\| \sum_{j=1}^{\infty} \alpha_j \psi_j \right\| \leq d_2 \|\alpha\|_{l^2(\mathbb{N})}, \quad \forall \alpha = \{\alpha_1, \alpha_2, \ldots\} \in l^2(\mathbb{N}), \tag{2.1}$$

where $\|\cdot\|$ is the norm on H [9]. Suppose further that $\phi_1, \phi_2, \ldots$ is a collection of *reconstruction* vectors that form a Riesz basis for a closed subspace T (the *reconstruction* space). We shall also assume that the spaces T and S satisfy the direct sum condition

$$\mathrm{T} \oplus \mathrm{S}^{\perp} = \mathrm{H}, \tag{2.2}$$

(for example, $\mathrm{T} = \mathrm{S} = \mathrm{H}$). Let $f \in \mathrm{H}$ and suppose that we have access to the samples

$$\hat{f}_j = \langle f, \psi_j \rangle, \quad j = 1, 2, \ldots.$$

Given $\{\hat{f}_j\}_{j=1}^{\infty}$, we seek a reconstruction $\tilde{f} \in \mathrm{T}$ of f. Naturally, this reconstruction ought to represent f well. This motivates the notion of *quasi-optimality*:

Definition 2.1. *A mapping* $\mathrm{H} \to \mathrm{T}$, $f \mapsto \tilde{f}$ *is quasi-optimal if there exists a constant* $c > 0$ *such that*

$$\|f - \tilde{f}\| \leq c\|f - \mathcal{Q}f\|, \quad \forall f \in \mathrm{H},$$

where $\mathcal{Q} : \mathrm{H} \to \mathrm{T}$ *is the orthogonal projection onto* T.

Note that $\mathcal{Q}f$ is the best, i.e. energy-minimising, approximation to f from T. Thus, quasi-optimality states that $\tilde{f}$ is within a constant factor of the energy-minimising approximation.

A simple and elegant way to obtain a quasi-optimal representation of $\tilde{f}$ from the samples $\{\hat{f}_j\}_{j=1}^{\infty}$ is via the so-called *consistency* conditions. That is, we specify $\tilde{f}$ by

$$\langle \tilde{f}, \psi_j \rangle = \langle f, \psi_j \rangle, \quad j = 1, 2, \ldots, \qquad \tilde{f} \in \mathrm{T}. \tag{2.3}$$

Consistency means that the samples of $\tilde{f}$ agree with those of f. We refer to $\tilde{f}$ as a *consistent reconstruction* of f.

A thorough analysis of consistent reconstructions, which we now recap, was given in [13, 14, 17]. For this, it is first useful to introduce the following two concepts:

Definition 2.2. *Let* U *and* V *be closed subspaces of* H *and* $\mathcal{Q}_V : \mathrm{H} \to \mathrm{V}$ *the orthogonal projection onto* V. *The subspace angle* $\theta = \theta_{\mathrm{UV}} \in [0, \frac{\pi}{2}]$ *between* U *and* V *is given by*

$$\cos(\theta_{\mathrm{UV}}) = \inf_{\substack{u \in \mathrm{U} \\ \|u\|=1}} \|\mathcal{Q}_V u\|.$$

Note that $\theta_{\mathrm{UV}} \neq \theta_{\mathrm{VU}}$ in general. However, it always holds that $\theta_{\mathrm{UV}} = \theta_{\mathrm{V}^{\perp}\mathrm{U}^{\perp}}$. Indeed, since $\mathcal{Q}_{\mathrm{V}}$ is an orthogonal projection,

$$\cos^2 \theta_{\mathrm{UV}} = 1 - \sup_{\substack{u \in \mathrm{U} \\ \|u\|=1}} \|\mathcal{Q}_{V^{\perp}} u\|^2 = 1 - \sup_{\substack{u \in \mathrm{U}, v \in \mathrm{V}^{\perp} \\ \|u\|, \|v\|=1}} \langle \mathcal{Q}_{\mathrm{V}^{\perp}} u, v \rangle = 1 - \sup_{\substack{u \in \mathrm{U}, v \in \mathrm{V}^{\perp} \\ \|u\|, \|v\|=1}} \langle u, v \rangle.$$

Symmetry of the right-hand side immediately implies the result. Moreover, in the case that $\mathrm{U} \oplus \mathrm{V} = \mathrm{H}$ it can be shown that $\theta_{\mathrm{UV}} = \theta_{\mathrm{VU}}$ [37].

Definition 2.3. *Let* U *and* V *be closed subspaces of* H. *A mapping* $\mathcal{W} := \mathcal{W}_{\mathrm{UV}} : \mathrm{H} \to \mathrm{U}$ *is an oblique projection onto* U *along* V *if* $\mathcal{W}^2 = \mathcal{W}$ *and* $\mathcal{W}(v) = 0$, $\forall v \in \mathrm{V}$.

Observe that $\mathrm{U} \oplus \mathrm{V} = \mathrm{H}$ if and only if there exists a unique oblique projection onto U along V [37]. Equivalently, the cosine of the subspace angle $\cos \theta_{\mathrm{UV}^{\perp}} \neq 0$. Indeed, $\mathcal{W}_{\mathrm{UV}}$ is defined by

$$\mathcal{W}_{\mathrm{UV}}(u + v) = u, \quad \forall u \in \mathrm{U}, v \in \mathrm{V}. \tag{2.4}$$

In the particular case that $\mathrm{V} = \mathrm{U}^{\perp}$ (i.e. $\cos \theta_{\mathrm{UV}} = 1$), $\mathcal{W}_{\mathrm{UV}} = \mathcal{Q}_{\mathrm{U}}$ is the orthogonal projection onto U.

Oblique projections arise in many types of sampling [5], and consistent reconstructions are intimately related with such mappings. Indeed, the following result was proved in [17]:

Theorem 2.4. *Let* T *and* S *satisfy (2.2). Then the mapping* $\mathrm{H} \to \mathrm{T}$, $f \mapsto \tilde{f}$, *where* $\tilde{f}$ *satisfies (2.3), is well-defined, and coincides with the oblique projection* $\mathcal{W}_{\mathrm{T},\mathrm{S}^{\perp}}$ *onto* T *along* $\mathrm{S}^{\perp}$.

As a result of this theorem, consistent reconstructions are equivalent to oblique projections. Thus a study of the latter informs the former. However, let us note one important distinction. When defining the consistent reconstruction (2.3), we assume that a basis $\{\psi_j\}_{j=1}^{\infty}$ of S is given. Indeed, this is natural in view of the sampling process. However, the oblique projection $\mathcal{W}_{\mathrm{T},\mathrm{S}^{\perp}}$ is independent of this basis, as it is determined solely by the spaces T and S. In light of Theorem 2.4, the same must also be true for $\tilde{f}$. In fact, as we detail in §2.2, specification of bases $\{\phi_j\}_{j=1}^{\infty}$ and $\{\psi_j\}_{j=1}^{\infty}$ is only necessary when writing (2.3) as a linear system of equations to be solved numerically.

We now wish to consider the error $\|f - \tilde{f}\|$ of the consistent reconstruction $\tilde{f}$, or equivalently, the oblique projection $\mathcal{W}_{\mathrm{T},\mathrm{S}^{\perp}} f$. We have

Theorem 2.5. *Suppose that* U *and* V *are closed subspaces of* H *satisfying* $\mathrm{U} \oplus \mathrm{V}^{\perp} = \mathrm{H}$*, and let* $\mathcal{W} : \mathrm{H} \to \mathrm{U}$ *be the oblique projection onto* U *along* $\mathrm{V}^{\perp}$*. Then*

$$\|\mathcal{W} f\| \leq \frac{1}{\cos\theta_{\mathrm{UV}}} \|f\|, \quad \forall f \in \mathrm{H}, \tag{2.5}$$

and

$$\|f - \mathcal{Q} f\| \leq \|f - \mathcal{W} f\| \leq \frac{1}{\cos\theta_{\mathrm{UV}}} \|f - \mathcal{Q} f\|, \quad \forall f \in \mathrm{H}, \tag{2.6}$$

where $\mathcal{Q} : \mathrm{H} \to \mathrm{U}$ *is the orthogonal projection. In particular, if* $f \in \mathrm{U}$*, then* $\mathcal{W} f = f$*.*

Proof. Whilst the bounds (2.5) and (2.6) were derived in [40] and [39] respectively, we shall include a proof, since similar techniques will be used subsequently.

If $f = u + v$ with $u \in \mathrm{U}$ and $v \in \mathrm{V}^{\perp}$, then $f - \mathcal{W} f = v \in \mathrm{V}^{\perp}$. Moreover, since $\mathcal{W} f \in \mathrm{U}$ we find that $(\mathcal{I} - \mathcal{Q}_{\mathrm{U}})\mathcal{W} f = 0$. Therefore

$$\|(\mathcal{I} - \mathcal{Q}_{\mathrm{U}}) f\| = \|\mathcal{Q}_{\mathrm{U}^{\perp}}(f - \mathcal{W} f)\| \geq \cos\theta_{\mathrm{V}^{\perp}\mathrm{U}^{\perp}} \|f - \mathcal{W} f\|.$$

Since $\theta_{\mathrm{V}^{\perp}\mathrm{U}^{\perp}} = \theta_{\mathrm{UV}}$, this gives (2.6). For (2.5), we first consider $\|\mathcal{Q}_{\mathrm{V}} \mathcal{W} f\|$. We have

$$\|\mathcal{Q}_{\mathrm{V}} \mathcal{W} f\| = \sup_{\substack{\psi \in \mathrm{V} \\ \|\psi\|=1}} \langle \mathcal{Q}_{\mathrm{V}} \mathcal{W} f, \psi \rangle = \sup_{\substack{\psi \in \mathrm{V} \\ \|\psi\|=1}} \langle \mathcal{W} f, \psi \rangle = \sup_{\substack{\psi \in \mathrm{V} \\ \|\psi\|=1}} \langle f, \psi \rangle = \sup_{\substack{\psi \in \mathrm{V} \\ \|\psi\|=1}} \langle \mathcal{Q}_{\mathrm{V}} f, \psi \rangle = \|\mathcal{Q}_{\mathrm{V}} f\|.$$

Since $\mathcal{Q}_{\mathrm{V}}$ is the orthogonal projection, we obtain $\|\mathcal{Q}_{\mathrm{V}} \mathcal{W} f\| \leq \|f\|$. By definition, $\cos\theta_{\mathrm{UV}} \|\mathcal{W} f\| \leq \|\mathcal{Q}_{\mathrm{V}} \mathcal{W} f\|$. Hence (2.5) now follows. ☐

Due to Theorem 2.4, equations (2.5) and (2.6) equivalently give bounds for the consistent reconstruction $\tilde{f}$ defined by (2.3). In particular, (2.6) illustrates quasi-optimality of the mapping $f \mapsto \tilde{f} \equiv \mathcal{W}_{\mathrm{T},\mathrm{S}^{\perp}} f$. Note that the first estimate (2.5) is a continuous stability estimate for $\tilde{f}$: the norm of the reconstruction is bounded by a constant multiple of the norm of the input signal.

As shown by (2.6), whenever $f \in \mathrm{T}$ its consistent reconstruction $\tilde{f} \equiv f$. In this case we say that f is recovered *perfectly*. We shall return to this concept subsequently.

2.1 Sharp bounds for oblique projections

Whilst Theorem 2.5 is a known result, it transpires that the bounds (2.5) and (2.6) are actually sharp. We now establish this claim. To the best of our knowledge, such a result does not currently exist in literature.

First, let us recall several properties of Riesz bases. Associated to any Riesz basis $\{\psi_j\}$ of a closed subspace S of H is a bounded linear operator $\mathcal{S} : l^2(\mathbb{N}) \to \mathrm{H}$ given by $\{\alpha_j\} \mapsto \sum_{j=1}^{\infty} \alpha_j \psi_j$. By (2.1),

$$d_1 \|\alpha\|_{l^2(\mathbb{N})} \leq \|\mathcal{S}\alpha\| \leq d_2 \|\alpha\|_{l^2(\mathbb{N})}, \quad \forall \alpha = \{\alpha_1, \alpha_2, \ldots\} \in l^2(\mathbb{N}).$$

The operator $\mathcal{S}$ referred to as the *synthesis operator* [9]. Its adjoint, the *analysis operator*, satisfies $\mathcal{S}^* f = \{\langle f, \psi_1 \rangle, \langle f, \psi_2 \rangle, \ldots\}$ for $f \in \mathrm{H}$. The composition $\mathcal{P} = \mathcal{S}\mathcal{S}^* : \mathrm{H} \to \mathrm{H}$, given by

$$\mathcal{P} f = \sum_{j=1}^{\infty} \langle f, \psi_j \rangle \psi_j, \quad \forall f \in \mathrm{H}, \tag{2.7}$$

is well-defined, linear, self-adjoint and bounded. Moreover, when considered as a mapping $\mathrm{S} \to \mathrm{S}$, $\mathcal{P}$ is positive and invertible.

Note that the consistent reconstruction $\tilde{f}$ can be expressed as $\mathcal{T}(\mathcal{S}^*\mathcal{T})^{-1}\mathcal{S}^* f$, where $\mathcal{S}$ and $\mathcal{T}$ are the synthesis operators for the Riesz bases $\{\psi_j\}$ and $\{\phi_j\}$ of S and T respectively. In particular, if $\operatorname{ran}(\mathcal{V})$ denotes the range of an operator $\mathcal{V}$, the bounded linear operator $\mathcal{X} = \mathcal{S}^*\mathcal{T}$ is continuously invertible as a mapping $\operatorname{ran}(\mathcal{T}^*) \to \operatorname{ran}(\mathcal{S}^*)$ [17].

We are now able to establish the following:

Theorem 2.6. *The bounds (2.5) and (2.6) are sharp.*

Proof. Let $\{u_j\}$ and $\{v_j\}$ be orthonormal bases for U and V respectively with synthesis operators $\mathcal{U}$ and $\mathcal{V}$. Observe that the composition $\mathcal{V}\mathcal{V}^* : \mathrm{H} \to \mathrm{V}$ is precisely the orthogonal projection onto V. In particular, $\mathcal{V}\mathcal{V}^*\psi = \psi$ for $\psi \in \mathrm{V}$. Moreover, $\mathcal{V}^*\mathcal{V} : l^2(\mathbb{N}) \to l^2(\mathbb{N})$ is nothing more than the identity operator $\mathcal{I}$ on $l^2(\mathbb{N})$.

With this to hand, write $\mathcal{A} = \mathcal{X}^*\mathcal{X} = \mathcal{U}^*\mathcal{V}\mathcal{V}^*\mathcal{U}$ and let $u \in \mathrm{U}$ be arbitrary with $u = \mathcal{U}\alpha$ for some $\alpha \in l^2(\mathbb{N})$. For reasons that will become clear, we are interested in writing the quantities

$$\|\mathcal{Q}_{\mathrm{V}} u\|, \quad \|\mathcal{Q}_{\mathrm{V}^\perp} u\|, \quad \|\mathcal{Q}_{\mathrm{U}^\perp}\mathcal{Q}_{\mathrm{V}^\perp} u\|,$$

in terms of $\mathcal{A}$. First, note that

$$\|\mathcal{Q}_{\mathrm{V}} u\|^2 = \langle \mathcal{V}\mathcal{V}^*\mathcal{U}\alpha, \mathcal{V}\mathcal{V}^*\mathcal{U}\alpha\rangle_{l^2(\mathbb{N})} = \langle \mathcal{U}^*\mathcal{V}\mathcal{V}^*\mathcal{V}\mathcal{V}^*\mathcal{U}\alpha, \alpha\rangle_{l^2(\mathbb{N})} = \langle \mathcal{A}\alpha, \alpha\rangle_{l^2(\mathbb{N})}. \tag{2.8}$$

Similarly,

$$\|\mathcal{Q}_{\mathrm{V}^\perp} u\|^2 = \langle (\mathcal{I} - \mathcal{A})\alpha, \alpha\rangle_{l^2(\mathbb{N})}. \tag{2.9}$$

Now consider $\|\mathcal{Q}_{\mathrm{T}^\perp}\mathcal{Q}_{\mathrm{S}^\perp} u\|$. Note that

$$\mathcal{Q}_{\mathrm{T}^\perp}\mathcal{Q}_{\mathrm{S}^\perp} u = (\mathcal{I} - \mathcal{U}\mathcal{U}^*)(\mathcal{I} - \mathcal{V}\mathcal{V}^*)\mathcal{U}\alpha = (\mathcal{U}\mathcal{A} - \mathcal{V}\mathcal{X})\alpha.$$

Moreover, since $\mathcal{A}$ is self-adjoint,

$$(\mathcal{U}\mathcal{A} - \mathcal{V}\mathcal{X})^*(\mathcal{U}\mathcal{A} - \mathcal{V}\mathcal{X}) = (\mathcal{A}\mathcal{U}^* - \mathcal{X}^*\mathcal{V}^*)(\mathcal{U}\mathcal{A} - \mathcal{V}\mathcal{X}) = \mathcal{A}\mathcal{U}^*\mathcal{U}\mathcal{A} - \mathcal{X}^*\mathcal{X}\mathcal{A} - \mathcal{A}\mathcal{X}^*\mathcal{X} + \mathcal{X}^*\mathcal{V}^*\mathcal{V}\mathcal{X},$$

and this is precisely $\mathcal{A}(\mathcal{I} - \mathcal{A})$. Thus, arguing as before, we obtain

$$\|\mathcal{Q}_{\mathrm{U}^\perp}\mathcal{Q}_{\mathrm{V}^\perp} u\|^2 = \langle \mathcal{A}(\mathcal{I} - \mathcal{A})\alpha, \alpha\rangle_{l^2(\mathbb{N})}. \tag{2.10}$$

Now consider the quantity

$$a = \cos^2\theta_{\mathrm{UV}} = \inf_{\substack{u \in \mathrm{U} \\ \|u\|=1}} \|\mathcal{Q}_{\mathrm{V}} u\|^2.$$

By (2.8) and the fact that $\mathcal{U}$ is an isometry onto U (since $\{u_j\}$ is an orthonormal basis), it follows that

$$a = \inf_{\substack{\alpha \in l^2(\mathbb{N}) \\ \|\alpha\|=1}} \langle \mathcal{A}\alpha, \alpha\rangle_{l^2(\mathbb{N})}.$$

The operator $\mathcal{A}$ is Hermitian positive-definite. Therefore a must belong to the spectrum of $\mathcal{A}$. Moreover, since the spectrum of $\mathcal{A}$ consists of approximate eigenvalues, there must be a sequence $\alpha^{[n]} \in l^2(\mathbb{N})$, $\|\alpha^{[n]}\|_{l^2(\mathbb{N})} = 1$, with $\|(\mathcal{A} - a\mathcal{I})\alpha^{[n]}\| \to 0$ as $n \to \infty$. Note that, if p is any polynomial, then

$$\langle p(\mathcal{A})\alpha^{[n]}, \alpha^{[n]}\rangle_{l^2(\mathbb{N})} \to p(a), \quad n \to \infty. \tag{2.11}$$

Let $u^{[n]} \in \mathrm{U}$, $\|u^{[n]}\| = 1$ be given by $\mathcal{U}\alpha^{[n]}/\|\mathcal{U}\alpha^{[n]}\|$. Define $w^{[n]} = \mathcal{Q}_{\mathrm{V}^\perp} u^{[n]} \in \mathrm{V}^\perp$. Since $w^{[n]} \in \mathrm{V}^\perp$, its image under the oblique projection $\mathcal{W} = \mathcal{W}_{\mathrm{U}\mathrm{V}^\perp}$ is $\mathcal{W}w^{[n]} \equiv 0$. Therefore, by (2.9),

$$\|w^{[n]} - \mathcal{W}w^{[n]}\|^2 = \|w^{[n]}\|^2 = \|\mathcal{Q}_{\mathrm{V}^\perp} u^{[n]}\|^2 = \langle (\mathcal{I} - \mathcal{A})\alpha^{[n]}, \alpha^{[n]}\rangle_{l^2(\mathbb{N})},$$

and we deduce that

$$\|w^{[n]} - \mathcal{W}w^{[n]}\|^2 \to 1 - a, \quad n \to \infty.$$

Moreover, by (2.10) and (2.11),

$$\|w^{[n]} - \mathcal{Q}_{\mathrm{U}} w^{[n]}\|^2 = \|\mathcal{Q}_{\mathrm{U}^\perp} \mathcal{Q}_{\mathrm{V}^\perp} u^{[n]}\|^2 = \langle (\mathcal{I} - \mathcal{A})\mathcal{A}\alpha^{[n]}, \alpha^{[n]} \rangle_{l^2(\mathbb{N})} \to (1-a)a.$$

Combining these results, we find that

$$\frac{\|w^{[n]} - \mathcal{Q}_{\mathrm{U}} w^{[n]}\|}{\|w^{[n]} - \mathcal{W} w^{[n]}\|} \to \sqrt{a} = \cos\theta_{\mathrm{UV}}, \quad n \to \infty,$$

and this indicates sharpness of the bound (2.6).

For (2.5) we write $w^{[n]} = \mathcal{Q}_{\mathrm{V}} u^{[n]}$, so that $\|w^{[n]}\|^2 = \langle \mathcal{A}\alpha^{[n]}, \alpha^{[n]} \rangle_{l^2(\mathbb{N})}$. Let $f \in \mathrm{H}$ be arbitrary and write $f = f_{\mathrm{U}} + f_{\mathrm{V}^\perp}$, where $f_{\mathrm{U}} \in \mathrm{U}$ and $f_{\mathrm{V}^\perp} \in \mathrm{V}^\perp$. Note that

$$\mathcal{W} f = f_{\mathrm{U}} = \mathcal{Q}_{\mathrm{V}} f,$$

and therefore $\mathcal{W} \circ \mathcal{Q}_{\mathrm{V}} = \mathcal{W}$. Hence, $\mathcal{W} w^{[n]} = \mathcal{W} u^{[n]}$. However, since $u^{[n]} \in \mathrm{U}$, we have $\mathcal{W} u^{[n]} = u^{[n]}$ and thus $\mathcal{W} w^{[n]} = u^{[n]}$. Therefore

$$\frac{\|w^{[n]}\|^2}{\|\mathcal{W} w^{[n]}\|^2} = \frac{\|w^{[n]}\|^2}{\|u^{[n]}\|^2} = \langle \mathcal{A}\alpha^{[n]}, \alpha^{[n]} \rangle_{l^2(\mathbb{N})} \to a, \quad n \to \infty,$$

as required. □

Note that the proof of this theorem can be significantly simplified if U is finite-dimensional. In that case, $\mathcal{A}$ can be considered as an operator $\mathbb{C}^n \to \mathbb{C}^n$, where $n = \dim \mathrm{U}$, with the value a being the minimum eigenvalue of $\mathcal{A}$. It particular, there exists $\alpha \in \mathbb{C}^n$ with $\mathcal{A}\alpha = a\alpha$. Thus, in this case, not only are the bounds (2.5) and (2.6) sharp, they are also attained by a specific elements of H.

2.2 Computing consistent reconstructions

To compute a consistent reconstruction $\tilde{f}$ we must first specify a basis $\{\phi_j\}_{j=1}^\infty$ for T. In this case, we write $\tilde{f} = \sum_{j=1}^\infty \alpha_j \phi_j$, and note that (2.3) is equivalent to the infinite system of linear equations $U\alpha = \hat{f}$, where U is the infinite matrix

$$U = \begin{pmatrix} \langle \phi_1, \psi_1 \rangle & \langle \phi_2, \psi_1 \rangle & \cdots \\ \langle \phi_1, \psi_2 \rangle & \langle \phi_2, \psi_2 \rangle & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}. \tag{2.12}$$

Observe that U is coincides with the operator $\mathcal{X} = \mathcal{S}^* \mathcal{T} : l^2(\mathbb{N}) \to l^2(\mathbb{N})$, where $\mathcal{S}$ and $\mathcal{T}$ are the synthesis operators for the Riesz bases $\{\psi_j\}_{j=1}^\infty$ and $\{\phi_j\}_{j=1}^\infty$ respectively.

Despite this formulation, it is still not possible to compute $\tilde{f}$, since we can neither access all samples of f, nor can we process infinite amounts of information. Thus, it is first necessary to discretise. A commonly used approach [13, 14, 15, 27, 38] is to replace the linear system $U\alpha = \hat{f}$ by a finite version $U^{[n]}\alpha^{[n]} = \hat{f}^{[n]}$, where

$$U^{[n]} = \begin{pmatrix} \langle \phi_1, \psi_1 \rangle & \cdots & \langle \phi_n, \psi_1 \rangle \\ \vdots & \ddots & \vdots \\ \langle \phi_1, \psi_n \rangle & \cdots & \langle \phi_n, \psi_n \rangle \end{pmatrix}, \tag{2.13}$$

$\hat{f}^{[n]} = \{\hat{f}_1, \ldots, \hat{f}_n\}$ and $\alpha^{[n]} = \{\alpha_1^{[n]}, \ldots, \alpha_n^{[n]}\}$. If $\tilde{f}_{n,n} = \sum_{j=1}^n \alpha_j^{[n]} \phi_j$, then this is equivalent to the following finite-dimensional analogue of the consistency conditions (2.3):

$$\langle \tilde{f}_{n,n}, \psi_j \rangle = \langle f, \psi_j \rangle, \quad j = 1, \ldots, n, \qquad \tilde{f}_{n,n} \in \mathrm{T}_n, \tag{2.14}$$

(the use of the double index in $\tilde{f}_{n,n}$ is for agreement with subsequent notation). Here $\mathrm{T}_n = \mathrm{span}\{\phi_1, \ldots, \phi_n\}$ is the finite-dimensional reconstruction space spanned by the first n reconstruction vectors.

The condition (2.14) is completely natural and reasonable to enforce. However, it turns out to be the wrong condition once considerations of stability and accuracy (as $n \to \infty$) are taken into account. To illustrate this, we first note that

Corollary 2.7. *Let* $\mathrm{S}_n = \mathrm{span}\{\psi_1, \ldots, \psi_n\}$ *and suppose that*

$$\cos\theta_{\mathrm{T}_n\mathrm{S}_n} \neq 0. \tag{2.15}$$

Then, for each $f \in \mathrm{H}$ *there exists a unique* $\tilde{f}_{n,n} \in \mathrm{T}_n$ *satisfying (2.14). In particular, the mapping* $f \mapsto \tilde{f}_{n,n}$ *is precisely the oblique projection onto* T_n *along* $\mathrm{S}_n^\perp$*, and we have the sharp bounds*

$$\|\tilde{f}_{n,n}\| \leq \frac{1}{\cos\theta_{\mathrm{T}_n\mathrm{S}_n}} \|f\|, \tag{2.16}$$

and

$$\|f - \mathcal{Q}_n f\| \leq \|f - \tilde{f}_n\| \leq \frac{1}{\cos\theta_{\mathrm{T}_n\mathrm{S}_n}} \|f - \mathcal{Q}_n f\|, \tag{2.17}$$

where $\mathcal{Q}_n$ *is the orthogonal projection onto* T_n*.*

Proof. Due to (2.15), we also have $\mathrm{T}_n \oplus \mathrm{S}_n^\perp = \mathrm{H}$, and therefore the oblique projection $\mathcal{W}_n = \mathcal{W}_{\mathrm{T}_n, \mathrm{S}_n^\perp}$ is well-defined. Clearly $\mathcal{W}_n f$ satisfies (2.14). Now suppose that (2.14) does not have a unique solution. Equivalently, there exists $\phi \in \mathrm{T}_n \backslash \{0\}$ with $\langle \phi, \psi_j \rangle = 0$, $j = 1, \ldots, n$. This implies that $\phi \in \mathrm{T}_n \cap \mathrm{S}_n^\perp = \{0\}$, and therefore $\phi = 0$ – a contradiction. Thus $\tilde{f}_{n,n}$ is unique, and coincides with the oblique projection $\mathcal{W}_n f$. The sharp bounds (2.16) and (2.17) now follow immediately from Theorem 2.5 and 2.6. □

Although one only has access to a finite number of samples, it is natural to consider the behaviour of $\tilde{f}_{n,n}$ as n – the number of samples – increases. This is a question of approximation. Ideally we would like $\tilde{f}_{n,n}$ to behave like $\mathcal{Q}_n f$, the best approximation to f from T_n. Namely, $\tilde{f}_{n,n}$ should be quasi-optimal, and thus converge to f at precise the same *rate* as $\mathcal{Q}_n f$. This is vitally important from a practical standpoint. The premise for computing a consistent reconstruction is the knowledge that f is well represented in terms of the vectors $\{\phi_j\}$. This is equivalent to the property that $\mathcal{Q}_n f$ converges rapidly. Hence it is vital that the computed reconstruction $\tilde{f}_{n,n}$ does not possess dramatically different behaviour as $n \to \infty$. Put simply, there is little point reconstructing in the basis $\{\phi_j\}$ if the good approximation properties of f in this basis are destroyed by reconstruction technique.

Whilst on the face of it (2.17) guarantees quasi-optimality, this is only true if the constant $C_{n,n} = \cos\theta_{\mathrm{T}_n\mathrm{S}_n}$ remains bounded away from zero as $n \to \infty$. As we show in the next section, there is no guarantee that this will be the case for arbitrary T_n and S_n. Indeed, there is actually no general guarantee that $\tilde{f}_{n,n}$ either exists, since (2.15) need not hold, or that the error $\|f - \tilde{f}_{n,n}\|$ does not diverge as $n \to \infty$. Furthermore, as we illustrate in §2.4, all these unpleasant features do actually occur in fairly straightforward examples.

Aside from convergence, or lack thereof, the behaviour of $C_{n,n}$ is also significant from a numerical point of view. Recall that, to compute $\tilde{f}_{n,n}$ we need to solve the linear system $U^{[n]}\alpha^{[n]} = \hat{f}^{[n]}$. The condition number $\kappa(U^{[n]}) = \|U^{[n]}\| \|(U^{[n]})^{-1}\|$ is therefore a critical consideration. Linear systems with large condition numbers are highly susceptible to both noise and round-off error, as well as being more costly to solve. In transpires that the quantity $C_{n,n}$ also determines $\kappa(U^{[n]})$. As we shall prove in Lemma 3.12, $\kappa(U^{[n]})$ satisfies

$$\kappa(U^{[n]}) \leq \frac{\sqrt{c_2}}{d_1 C_{n,n}} \sqrt{\kappa(A^{[n]})},$$

where d_1 is given by (2.1), c_2 is some absolute constant and $A^{[n]}$ is the Gram matrix for the vectors $\phi_1, \ldots, \phi_n$. Thus, small $C_{n,n}$ can potentially lead to a large condition number.

2.3 Operator-theoretic interpretation

To sum up, the properties of the finite-dimensional consistent reconstruction $\tilde{f}_{n,n}$ are determined critically by the quantity $C_{n,n}$ – or equivalently, by the behaviour of the matrices $U^{[n]}$ – as $n \to \infty$. It transpires that the general failure of consistent reconstructions can be explained quite easily using operator-theoretic arguments. Moreover, such arguments also indicate how to avoid these issues, leading to so-called *generalised sampling*.

Recall that the infinite-dimensional consistent reconstruction $\tilde{f}$, defined by (2.3), can be written as the solution of an infinite system of linear equations $U\alpha = \hat{f}$, with infinite-dimensional operator U

n	25	50	100	200
$1/C_{n,n}$	3.4e3	1.0e8	6.2e16	3.2e34
$\|f - \tilde{f}_{n,n}\|$	2.6e0	3.6e1	1.82	2.8e6
$\|f - \mathcal{Q}_n f\|$	1.1e-3	2.9-6	1.2-11	4.4-18

Table 1: Recovery of $f(x) = \frac{1}{1+16x^2}$ with the consistent reconstruction $\tilde{f}_{n,n}$.

given by (2.12) As discussed, we cannot solve these equations in finitely many operations. Hence we must discretise. Suppose that we do this in some way to give a sequence of finite-dimensional problems $U^{[n]}\alpha^{[n]} = \hat{f}^{[n]}$, $n = 1, 2, \ldots$. For obvious reasons, it is vitally important that this sequence of problems satisfies the three following conditions:

(i) *Invertibility:* $U^{[n]}$ is invertible for all $n = 1, 2, \ldots$.

(ii) *Stability:* $\|(U^{[n]})^{-1}\|$ is uniformly bounded for all $n = 1, 2, \ldots$.

(iii) *Convergence:* the solutions $\alpha^{[n]} \to \alpha$ as $n \to \infty$.

Suppose that $U^{[n]}$ is given by (2.13). Then $U^{[n]}$ is nothing more than the $n \times n$ *finite section* of U. In other words, if $\{e_j\}$ is the canonical basis for $l^2(\mathbb{N})$ and $P_n : l^2(\mathbb{N}) \to \mathrm{span}\{e_1, \ldots, e_n\}$ is the orthogonal projection, then $U^{[n]} = P_n U P_n$. Moreover, $\hat{f}^{[n]} = P_n \hat{f}$, and thus the finite-dimensional consistent reconstruction $\tilde{f}_{n,n}$ is precisely the result of the finite section method applied to $U\alpha = \hat{f}$.

The properties of finite sections of infinite matrices have been extensively studied over the last several decades [7, 25, 30]. Unfortunately, there is no guarantee that the finite sections $P_n U P_n$ of an arbitrary matrix U satisfy properties (i)–(iii). In fact, one requires rather restrictive conditions, such as positive self-adjointness, for this to be the case. Typically operators U of the form (2.12) are not self-adjoint, thereby making this approach unsuitable in general for discretising $U\alpha = \hat{f}$.

2.4 Failure of linear consistent reconstructions

Generalised sampling is based on alternatives to the finite section method for solving infinite-dimensional problems. Before discussing this, let us first numerically illustrate the failure of linear consistent reconstructions.

As mentioned in §1, consistent reconstructions will not in general possess invertibility, stability or convergence. This is made clear from (2.15) and (2.17). It may well be the case that T and S obey (2.2), yet $\mathrm{T}_n \cap \mathrm{S}_n^{\perp} \neq \{0\}$ for any n. This renders the finite-dimensional consistent reconstruction problem ill-posed. Such an eventuality has been well documented (see [1, 27] for examples). Potential remedies were suggested in [12, 27].

However, of arguably greater importance, even if (2.15) holds, it may be that the constant $C_{n,n} \to 0$ as $n \to \infty$. Consequently, consistent reconstructions may be neither stable nor convergent in the sense defined above. We remark in passing that, although the methods introduced in [12, 27] are described as *robust*, this is only a finite-dimensional consideration, and does not refer to the case $n \to \infty$.

Let us illustrate this effect with the following example, where we seek the reconstruction of a function $f : [-1, 1] \to \mathbb{R}$ in a polynomial basis from its Fourier samples (i.e. $\mathrm{T}_n = \mathbb{P}_{n-1}$ and $\psi_j(x) = \frac{1}{\sqrt{2}} \mathrm{e}^{\mathrm{i} j \pi x}$). As we shall see, this example succinctly demonstrates the potential failure of consistent reconstructions. Furthermore, it is also important in its own right. In the last several decades there has been intense study into the question of removing the Gibbs phenomenon from the Fourier expansion of a nonperiodic function by reconstructing in a polynomial basis [21]. This question is of particular importance in spectral methods for hyperbolic PDEs [20, 21, 36], where it is necessary to postprocess any numerical solution to obtain an accurate reconstruction in physical space (as opposed to Fourier space) of the solution of the given PDE. As considered in detail in [4], generalised sampling, which we discuss in the next section, is particularly well suited to this problem, and appears to outperform existing methods. Moreover, as we consider in §4.3, there is substantial evidence to support the conjecture that generalised sampling is in fact an optimal stable method for this problem.

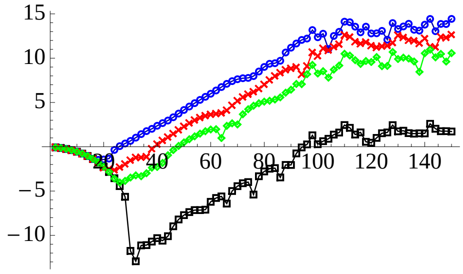

Figure 1: The quantity $\log_{10} \|f - \tilde{f}_{n,n}\|_\infty$ against n, where $f(x) = x^{30}$ and $\tilde{f}_{n,n}$ was computed from (i) noise-free data (squares) and (ii) noise at amplitude $\epsilon = 10^{-4}$ (circles), 10^{-6} (crosses) and 10^{-8} (diamonds). Noise was modelled by replacing $\hat{f}_j$ by $\hat{f}_j + \epsilon z_j$ for $j = -100, \ldots, 100$, where z_j is uniformly distributed in $[-1, 1]$.

Returning to consistent reconstructions, in the first row of Table 1 we consider the quantity $C_{n,n}$ for this problem. As is evident $C_{n,n}$ is exponentially small in n, with numerical results indicating that

$$C_{n,n} \sim c^{-n}, \tag{2.18}$$

where $c \approx 1.5$. Since the condition number $\kappa(U^{[n]})$ actually coincides with $C_{n,n}$ in this case, this means that with $n = 200$, for example, one is required to solve a linear system with condition number of order 10^{34}. As a result, the consistent reconstruction is highly susceptible to both round-off error and noise. To illustrate this, in Figure 1 we consider the reconstruction of $f(x) = x^{30}$ from its Fourier samples. Since f lies in the reconstruction space T_n whenever $n > 30$ we expect perfect recovery, i.e. $f = \tilde{f}_{n,n}$, for $n > 30$. Of course, perfect recovery cannot be realised in finite precision, or in the case that the samples are polluted by noise. However, it is highly desirable that a given reconstruction technique possess *numerically perfect* recovery. That is, perfect recovery up to an error proportional to the numerical precision used, or, in the noisy case, the magnitude of the noise. Unfortunately, the consistent reconstruction is not endowed with this property. As demonstrated in Figure 1, whilst initially obtaining close to numerically perfect recovery, once the parameter n is increased beyond 30 this error actually diverges.

Aside from stability issues, the exponential decay of $C_{n,n}$ means that there may well be functions f for which $\tilde{f}_{n,n} \nrightarrow f$ as $n \to \infty$. This problem is confirmed in Table 1 for the example $f(x) = \frac{1}{1+16x^2}$. In this case, the error $\|f - \tilde{f}_{n,n}\|$ increases at a rate exponentially fast in n.

This issue can actually be explained completely in terms of the sharp bound (2.17). Given that $C_{n,n}$ satisfies (2.18), the estimate (2.17) yields

$$\|f - \tilde{f}_{n,n}\| \leq c^n \|f - \mathcal{Q}_n f\|.$$

Thus, convergence can only be guaranteed when $\|f - \mathcal{Q}_n f\| \leq \rho^{-n}$ for some $\rho > c$. Now, the convergence rate of $\mathcal{Q}_n f$, a polynomial expansion of f, is completely understood in terms of subsets of the complex plane known as Bernstein ellipses

$$B(\rho) = \left\{ \tfrac{1}{2}(\rho^{-1} \mathrm{e}^{-\mathrm{i}\theta} + \rho \mathrm{e}^{\mathrm{i}\theta}) : \theta \in [-\pi, \pi] \right\} \subseteq \mathbb{C}, \qquad \rho > 1.$$

A classical result on polynomial expansions states that if f is analytic inside the Bernstein ellipse $B(\rho)$ and not analytic inside any Bernstein ellipse $B(\rho')$ with $\rho' > \rho$, then $\|f - \mathcal{Q}_n f\|$ converges exponentially fast at rate ρ [34]. Thus, we may conclude that

Corollary 2.8. *Let $\mathrm{T}_n = \mathbb{P}_{n-1}$ and $\psi_j(x) = \frac{1}{\sqrt{2}} \mathrm{e}^{\mathrm{i}j\pi x}$. Then the consistent reconstruction $\tilde{f}_{n,n}$ of f is guaranteed to converge to f when f is analytic inside the Bernstein ellipse $B(\rho)$, where $\rho > c$ and c is such that $C_{n,n} \geq c^{-n}$ for all sufficiently large n.*

This corollary does not explicitly state that the consistent reconstruction of an insufficiently regular function f will not converge. However, Table 1 indicates that this is indeed the case. Observe that for this example, the function $f(x) = \frac{1}{1+16x^2}$ has complex singularities at $x = \pm\frac{1}{4}\mathrm{i}$, meaning that the largest

Bernstein ellipse within which f is analytic has index $\rho = \frac{1}{4}(1+\sqrt{17}) \approx 1.28$. Since $c \approx 1.5$, we expect divergence of $\tilde{f}_{n,n}$.

To sum up, there is substantial evidence to suggest that consistent reconstructions only converge for a particularly small class of functions f. In contrast, as we shall see in §3, generalised sampling yields reconstructions that converge for all functions f, not just those of sufficient regularity.

We remark in passing that the aforementioned effect is similar to the classical Runge phenomenon arising in reconstructing a function $f : [-1,1] \to \mathbb{R}$ from its values at n equispaced nodes using a polynomial of degree n [18]. Indeed, rather than pointwise evaluations of f, here we consider equispaced samples of its Fourier transform. Corollary 2.8 can therefore be viewed as a continuous analogue of the Runge phenomenon. Note that in the classical Runge phenomenon it is possible to prove a much stronger version of Corollary 2.8 that stipulates precisely when the polynomial reconstruction will converge or diverge [19, chpt. 3].

The numerical instability and potential divergence of linear consistent reconstructions are not limited to this example problem. One also witnesses a similar effect when reconstructing in d^{th} order splines from Fourier samples. Examples involving Haar wavelets were also documented in [1]. Although there are certainly instances where success is not precluded – certain shift-invariant spaces being an example [39] – the situation is certainly less than ideal. However, returning to the narrative of §2.3, given that this issue arises purely from the type of discretisation applied to the operator U, there is potential for these problems to be avoided by means of a different approach. This turns out to be the case, and the result is generalised sampling.

3 Generalised sampling

The key to obtaining a discretisation of $U\alpha = \hat{f}$ obeying properties (i)–(iii) is to allow the number of samples m to vary independently of the number of reconstruction vectors n. When done in this way, we obtain a finite-dimensional operator $U^{[n,m]}$ (which now depends on both m and n) that inherits the structure and properties of its infinite-dimensional counterpart U, provided m is sufficiently large for a given n. It turn, this ensures properties (i)–(iii).

To this end, let $\mathrm{S}_m = \operatorname{span}\{\psi_1, \ldots, \psi_m\}$ and $\mathrm{T}_n = \operatorname{span}\{\phi_1, \ldots, \phi_n\}$ be finite-dimensional subspaces of H, the (finite-dimensional) *sampling* and *reconstruction* spaces respectively. We seek a reconstruction $\tilde{f}_{n,m} \in \mathrm{T}_n$ of f. Let $\mathcal{P}_m : \mathrm{H} \to \mathrm{S}_m$ be the finite rank operator given by

$$\mathcal{P}_m g = \sum_{j=1}^{m} \langle g, \psi_j \rangle \psi_j.$$

Note that, due to (2.1), the sequence of operators $\mathcal{P}_m$ converge strongly to $\mathcal{P}$ on H [9], where $\mathcal{P}$ is given by (2.7). With this to hand, the approach proposed in [1] is to define $\tilde{f}_{n,m} \in \mathrm{T}_n$ as the solution of the equations

$$\langle \mathcal{P}_m \tilde{f}_{n,m}, \phi_j \rangle = \langle \mathcal{P}_m f, \phi_j \rangle, \quad j = 1, \ldots, n, \qquad \tilde{f}_{n,m} \in \mathrm{T}_n, \tag{3.1}$$

We refer to the mapping $\{\hat{f}_j\}_{j=1}^m \mapsto \tilde{f}_{n,m}$ as *generalised sampling*. Observe that $\mathcal{P}_m f$ is determined solely by the coefficients $\hat{f}_1, \ldots \hat{f}_m$ of f. Thus, this mapping is well-defined in this sense.

In what follows it will be useful to note that (3.1) is equivalent to

$$\langle \tilde{f}_{n,m}, \mathcal{P}_m \phi_j \rangle = \langle f, \mathcal{P}_m \phi_j \rangle, \quad j = 1, \ldots, n, \qquad \tilde{f}_{n,m} \in \mathrm{T}_n, \tag{3.2}$$

due to the self-adjointness of the operator $\mathcal{P}_m$. An immediate consequence of this formulation is the following:

Lemma 3.1. *Suppose that* $\cos \theta_{\mathrm{T}_n \mathrm{S}_n} \neq 0$. *Then when* $m = n$ *the generalised sampling reconstruction* $\tilde{f}_{n,m}$ *defined by (3.1) is precisely the consistent reconstruction* $\tilde{f}_{n,n}$ *given by (2.3).*

Proof. We first claim that $\mathcal{P}_n(\mathrm{T}_n) = \mathrm{S}_n$. In other words, $\mathcal{P}_n$ is a bijection between T_n and S_n. Since T_n and S_n are both of dimension n, it suffices to show that if $\mathcal{P}_n \phi = 0$ for some $\phi \in \mathrm{T}_n$ then $\phi = 0$. Suppose that $\mathcal{P}_n \phi = 0$. Then, since $\psi_1, \ldots, \psi_n$ are linearly independent, we have that $\langle \phi, \psi_j \rangle = 0$ for

$j = 1, \dots, n$, and therefore $\phi \in \mathrm{S}_n^{\perp}$. Since $\phi \in \mathrm{T}_n$, it follows from the observation $\mathrm{T}_n \cap \mathrm{S}_n^{\perp} = \{0\}$ that $\phi = 0$, as required.

By linearity, the conditions (3.2) are equivalent to (2.3). Since the consistent reconstruction $\tilde{f}_{n,n}$ satisfying (2.3) is unique, we obtain the result. □

We conclude that generalised sampling contains the consistent reconstruction technique as a special case corresponding to $n = m$ (whenever the latter well-defined). However, as we next discuss, there is a significant advantage in allowing the parameter m to be independent of n.

3.1 An intuitive argument

As expounded in [1, 4], the approach (3.1) owes its success to allowing m and n to vary independently. To this end, suppose that n is fixed and let $m \to \infty$. Equations (3.1) now read

$$\langle \mathcal{P} \tilde{f}_n, \phi \rangle = \langle \mathcal{P} f, \phi \rangle, \quad \forall \phi \in \mathrm{T}_n, \qquad \tilde{f}_n \in \mathrm{T}_n,$$

for some $\tilde{f}_n \in \mathrm{T}_n$, where $\mathcal{P}$ is given by (2.7). Since $\mathcal{P}$ is self-adjoint, $\tilde{f}_n$ is equivalently defined by

$$\langle \tilde{f}_n, \Phi \rangle = \langle f, \Phi \rangle, \quad \forall \Phi \in \mathcal{P}(\mathrm{T}_n), \qquad \tilde{f}_n \in \mathrm{T}_n. \tag{3.3}$$

We have

Theorem 3.2. *For any $f \in \mathrm{H}$, there exists a unique $\tilde{f}_n \in \mathrm{T}_n$ satisfying (3.3). Moreover, the mapping $f \mapsto \tilde{f}_n$ is precisely the oblique projection onto T_n along $[\mathcal{P}(\mathrm{T}_n)]^{\perp}$.*

To prove this theorem it is useful to have the following lemma:

Lemma 3.3. *Suppose that $\mathrm{T} \oplus \mathrm{S}^{\perp} = \mathrm{H}$ and that $\mathcal{P}$ is given by (2.7). Then there are constants c_1 and c_2 such that*

$$c_1 \|g\|^2 \leq \langle \mathcal{P} g, g \rangle \leq c_2 \|g\|^2, \quad \forall g \in \mathrm{T}. \tag{3.4}$$

Moreover, $\mathcal{P}$ is continuously invertible as a mapping $\mathrm{T} \to \mathcal{P}(\mathrm{T})$.

Proof. Let $\mathcal{S}$ be the synthesis operator for the vectors $\{\psi_j\}_{j=1}^{\infty}$, and recall that $\mathcal{P} = \mathcal{S}\mathcal{S}^*$. Thus, if $\phi \in \mathrm{T}$ is arbitrary we have

$$\langle \mathcal{P}\phi, \phi \rangle = \langle \mathcal{S}\mathcal{S}^* \phi, \phi \rangle = \|\mathcal{S}^* \phi\|^2_{l^2(\mathbb{N})}.$$

Therefore, to prove (3.4) it suffices to show that $\mathcal{S}^*$ is continuously invertible as a mapping $\mathrm{T} \to \mathcal{S}^*(\mathrm{T})$. By the inverse mapping theorem, it is sufficient to show that $\mathcal{S}^*$ is bijective. By construction, $\mathcal{S}^*$ is surjective. Moreover, suppose that $\mathcal{S}^* \phi = 0$ for some $\phi \in \mathrm{T}$. Then $\langle \phi, \psi_j \rangle = 0$ for all j, and so $\phi \in \mathrm{S}^{\perp}$. However, $\mathrm{T} \cap \mathrm{S}^{\perp} = \{0\}$ and thus $\phi = 0$. Hence $\mathcal{S}^* : \mathrm{T} \to \mathcal{S}^*(\mathrm{T})$ is bijective.

For the final part, we merely note that (3.4) implies that $\mathcal{P}$ is bounded below on T, and therefore must be continuously invertible onto its range $\mathcal{P}(\mathrm{T})$. □

Proof of Theorem 3.2. We first claim that $\cos \theta_{\mathrm{T}_n, \mathcal{P}(\mathrm{T}_n)} \neq 0$, so that the oblique projection $\mathcal{W}$ onto T_n along $[\mathcal{P}(\mathrm{T}_n)]^{\perp}$ is well-defined and unique. Suppose not. Since T_n is finite dimensional, there exists a $\phi \in \mathrm{T}_n$, $\phi \neq 0$, satisfying $\mathcal{Q}_{\mathcal{P}(\mathrm{T}_n)} \phi = 0$. Thus

$$0 = \langle \mathcal{Q}_{\mathcal{P}(\mathrm{T}_n)} \phi, \mathcal{P} \phi' \rangle = \langle \phi, \mathcal{P} \phi' \rangle, \quad \forall \phi' \in \mathrm{T}_n,$$

and, in particular, $\langle \phi, \mathcal{P} \phi \rangle = 0$. Hence, $\phi = 0$ by Lemma 3.3 – a contradiction. Hence $\mathcal{W}$ exists and is unique. Moreover, it is clear that $\mathcal{W} f$ satisfies (3.3). Therefore $\tilde{f}_n = \mathcal{W} f$ by uniqueness, as required. □

As a result of this theorem, $\tilde{f}_n$ is precisely the oblique projection of f onto T_n along $[\mathcal{P}(\mathrm{T}_n)]^{\perp}$. Applications of Theorems 2.5 and 2.6 immediately give

Corollary 3.4. *Let $\tilde{f}_n$ be defined by (3.3). Then $\tilde{f}_n$ satisfies*

$$\|\tilde{f}_n\| \leq \frac{1}{C'_n}\|f\|, \tag{3.5}$$

and

$$\|f - \mathcal{Q}_n f\| \leq \|f - \tilde{f}_n\| \leq \frac{1}{C'_n}\|f - \mathcal{Q}_n f\|, \tag{3.6}$$

where $C'_n = \cos\theta_{\mathrm{T}_n,\mathcal{P}(\mathrm{T}_n)}$. Moreover, the bounds (3.5) and (3.6) are sharp.

Recall that the failure of consistent reconstruction was due to the potential decrease of the constant $C_{n,n} = \cos\theta_{\mathrm{T}_n,\mathrm{S}_n} \to 0$ as $n \to \infty$. This cannot happen with C'_n. Indeed

Lemma 3.5. *Let $C'_n = \cos\theta_{\mathrm{T}_n,\mathcal{P}(\mathrm{T}_n)}$. Then C'_n satisfies*

$$1 \geq C'_n \geq \inf_{\substack{\phi\in\mathrm{T}_n \\ \|\phi\|=1}} \left\{ \frac{\langle \mathcal{P}\phi, \phi\rangle}{\|\mathcal{P}\phi\|} \right\} \geq \frac{c_1}{c_3}. \tag{3.7}$$

where $c_3 = \|\mathcal{P}\|_{\mathrm{T}\to\mathrm{T}}$ and c_1 is given by (3.4). If $\mathcal{P}$ is a projection ($\mathcal{P}^2 = \mathcal{P}$), then the second inequality holds with equality, and has the simplified form

$$C'_n = \inf_{\substack{\phi\in\mathrm{T}_n \\ \|\phi\|=1}} \|\mathcal{P}\phi\|. \tag{3.8}$$

In particular, $C'_n = 1$ whenever $\mathcal{P} = \mathcal{I}$.

Proof. That $C'_n \leq 1$ follows directly from the definition. Also from the definition we have

$$\cos\theta_{\mathrm{T}_n,\mathcal{P}(\mathrm{T}_n)} = \inf_{\substack{\phi\in\mathrm{T}_n \\ \|\phi\|=1}} \|\mathcal{Q}_{\mathcal{P}(\mathrm{T}_n)}\phi\|.$$

Consider $\|\mathcal{Q}_{\mathcal{P}(\mathrm{T}_n)}\phi\|$ for some $\phi \in \mathrm{T}_n$. By the standard duality pairing,

$$\|\mathcal{Q}_{\mathcal{P}(\mathrm{T}_n)}\phi\| = \sup_{\substack{\phi'\in\mathrm{T}_n \\ \phi'\neq 0}} \left\{ \frac{\langle \mathcal{Q}_{\mathcal{P}(\mathrm{T}_n)}\phi, \mathcal{P}\phi'\rangle}{\|\mathcal{P}\phi'\|} \right\} = \sup_{\substack{\phi'\in\mathrm{T}_n \\ \phi'\neq 0}} \left\{ \frac{\langle \phi, \mathcal{P}\phi'\rangle}{\|\mathcal{P}\phi'\|} \right\}.$$

Now $\mathcal{P}$ is self-adjoint. Hence,

$$\|\mathcal{Q}_{\mathcal{P}(\mathrm{T}_n)}\phi\| \geq \frac{\langle \mathcal{P}\phi, \phi\rangle}{\|\mathcal{P}\phi\|},$$

which gives (3.7). The lower bound $C'_n \geq \frac{c_1}{c_3}$ follows immediately from properties of $\mathcal{P}$. For the final part, we merely notice that if $\mathcal{P}^2 = \mathcal{P}$ then

$$\langle \phi, \mathcal{P}\phi'\rangle = \langle \phi, \mathcal{P}^2\phi'\rangle = \langle \mathcal{P}\phi, \mathcal{P}\phi'\rangle, \quad \forall \phi' \in \mathrm{T}_n. \tag{3.9}$$

Thus

$$\langle \mathcal{P}\phi, \mathcal{P}\phi'\rangle = \langle \phi, \mathcal{P}\phi'\rangle = \langle \mathcal{Q}_{\mathcal{P}(\mathrm{T}_n)}\phi, \mathcal{P}\phi'\rangle, \quad \forall \phi' \in \mathrm{T}_n.$$

Hence $\mathcal{Q}_{\mathcal{P}(\mathrm{T}_n)}\phi = \mathcal{P}\phi$. Setting $\phi' = \phi$ in (3.9), we also note that $\langle \phi, \mathcal{P}\phi\rangle = \|\mathcal{P}\phi\|^2$. This gives (3.8). □

An important instance where $\mathcal{P}$ coincides with the identity occurs when the sampling vectors $\{\psi_j\}$ are orthonormal. In this case, $C'_n = 1$ and the sharp bound (3.6) immediately implies that $\tilde{f}_n$ is precisely the orthogonal projection $\mathcal{Q}_n f$. This is corroborated by Theorem 3.2: since $\mathcal{P}(\mathrm{T}_n) = \mathrm{T}_n$ in this case, the oblique projection $f \mapsto \tilde{f}_n$ coincides with the orthogonal projection onto T_n.

In general, although it need not coincide with $\mathcal{Q}_n f$, this lemma demonstrates that the reconstruction $\tilde{f}_n$ of f possesses the three aforementioned properties of invertibility, stability and convergence, as well as quasi-optimality. Unfortunately, $\tilde{f}_n$ cannot be computed in finitely many operations, since it involves the infinite-rank operator $\mathcal{P}$. Nonetheless, given that the equations defining $\tilde{f}_n$ are, roughly speaking, the limit of those defining $\tilde{f}_{n,m}$, we may expect the reconstruction $\tilde{f}_{n,m}$, which can be computed, to inherit all these properties, for all sufficiently large m. This is proved rigorously in the next section.

Before doing so, however, let us relate $\tilde{f}_n$ to the notion of finite sections. Recall first that we wish to solve $U\alpha = \hat{f}$. Since α satisfies these equations it also obeys the normal equations

$$U^* U \alpha = U^* \hat{f}. \tag{3.10}$$

Now write $\tilde{f}_n = \sum_{j=1}^{n} \alpha_j^{[n]} \phi_j$. It is easily shown that $\alpha^{[n]} = \{\alpha_1^{[n]}, \ldots, \alpha_n^{[n]}\}$ is defined by

$$P_n U^* U P_n \alpha = P_n U^* \hat{f},$$

where $\hat{f} = \{\hat{f}_1, \hat{f}_2, \ldots\}$. Thus, $\alpha^{[n]}$ is precisely the result of the finite section method applied to the normal equations (3.10). Since the operator U^*U is self-adjoint and positive, its finite sections must possess properties (i)–(iii) (see §2.3).

3.2 Analysis of generalised sampling

The analysis of $\tilde{f}_{n,m}$ is similar to that of $\tilde{f}_n$. Whilst an analysis was originally given in [1, 4], the bounds derived were not sharp. Conversely, the bounds we obtain in this section are sharp, and therefore improve those of [1, 4].

We commence with the following lemma:

Lemma 3.6. *Let $C_{n,m} = \cos\theta_{\mathrm{T}_n, \mathcal{P}_m(\mathrm{T}_n)}$ be the cosine of the angle between the subspaces T_n and $\mathcal{P}_m(\mathrm{T}_n)$. Then $C_{n,m}$ satisfies*

$$1 \geq C'_n \geq C_{n,m} \geq \inf_{\substack{\phi \in \mathrm{T}_n \\ \|\phi\|=1}} \left\{ \frac{\langle \mathcal{P}_m \phi, \phi \rangle}{\|\mathcal{P}_m \phi\|} \right\},$$

where C'_n is as in Corollary 3.4. Moreover, $C_{n,m} \to C'_n$ as $m \to \infty$. In particular, given $n \in \mathbb{N}$ and $\theta \in (0, \frac{c_1}{c_3})$, there exists an $m_0 \in \mathbb{N}$ such that $C_{n,m} > \theta$ for all $m \geq m_0$.

Proof. Let $\mathrm{V} \subseteq \mathrm{W}$ be closed subspaces of H and $\mathcal{Q}_\mathrm{U}$, $\mathcal{Q}_\mathrm{V}$ the corresponding orthogonal projections. Then

$$\|\mathcal{Q}_\mathrm{V} g\|^2 = 1 - \|g - \mathcal{Q}_\mathrm{V} g\|^2 \leq 1 - \|g - \mathcal{Q}_\mathrm{W} g\|^2 = \|\mathcal{Q}_\mathrm{W} g\|^2, \quad \forall g \in \mathrm{H}.$$

In particular, for any closed subspace U,

$$\cos\theta_{\mathrm{UV}} = \inf_{\substack{u \in \mathrm{U} \\ \|u\|=1}} \|\mathcal{Q}_\mathrm{V} u\| \leq \inf_{\substack{u \in \mathrm{U} \\ \|u\|=1}} \|\mathcal{Q}_\mathrm{W} u\| = \cos\theta_{\mathrm{UW}}.$$

Therefore, the first result is obtained by setting $\mathrm{U} = \mathrm{T}_n$, $\mathrm{V} = \mathcal{P}_m(\mathrm{T}_n)$ and $\mathrm{W} = \mathcal{P}(\mathrm{T}_n)$.

The second part of the proof uses identical arguments to those given in the proof of Lemma 3.5. For the final part, we first write

$$\frac{\langle \mathcal{P}_m \phi, \phi \rangle}{\|\mathcal{P}_m \phi\|} = \frac{\|\mathcal{P}\phi\|}{\|\mathcal{P}_m \phi\|} \left[\frac{\langle \mathcal{P}\phi, \phi \rangle}{\|\mathcal{P}\phi\|} - \frac{\langle \mathcal{P}\phi - \mathcal{P}_m \phi, \phi \rangle}{\|\mathcal{P}\phi\|} \right],$$

from which it follows that

$$\begin{aligned} C_{n,m} &\geq \left[\sup_{\substack{\phi \in \mathrm{T}_n \\ \|\phi\|=1}} \frac{\|\mathcal{P}_m \phi\|}{\|\mathcal{P}\phi\|} \right]^{-1} \left[C'_n - \sup_{\substack{\phi \in \mathrm{T}_n \\ \|\phi\|=1}} \left\{ \frac{\langle \mathcal{P}\phi - \mathcal{P}_m \phi, \phi \rangle}{\|\mathcal{P}\phi\|} \right\} \right] \\ &\geq \left[1 + \frac{1}{c_4} \sup_{\substack{\phi \in \mathrm{T}_n \\ \|\phi\|=1}} \|\mathcal{P}\phi - \mathcal{P}_m \phi\| \right]^{-1} \left[C'_n - \frac{1}{c_4} \sup_{\substack{\phi \in \mathrm{T}_n \\ \|\phi\|=1}} \|\mathcal{P}\phi - \mathcal{P}_m \phi\| \right], \end{aligned}$$

where

$$c_4 = \inf_{\substack{g \in \mathrm{T} \\ \|g\|=1}} \|\mathcal{P} g\|,$$

(recall from Lemma 3.3 that $\mathcal{P} : \mathrm{T} \to \mathcal{P}(\mathrm{T})$ is continuously invertible). Therefore, it suffices to prove that

$$\sup_{\substack{\phi \in \mathrm{T}_n \\ \|\phi\|=1}} \|\mathcal{P}\phi - \mathcal{P}_m \phi\| \to 0, \quad m \to \infty,$$

for any fixed $n \in \mathbb{N}$. Let $\{\phi_j\}_{j=1}^n$ be an orthonormal basis for T_n. Write $\phi = \sum_{j=1}^n \alpha_j \phi_j$ for $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{C}^n$. Then,

$$\sup_{\substack{\phi \in \mathrm{T}_n \\ \|\phi\|=1}} \|\mathcal{P}\phi - \mathcal{P}_m \phi\| = \sup_{\substack{\alpha \in \mathbb{C}^n \\ \|\alpha\|=1}} \left\| \sum_{j=1}^n \alpha_j (\mathcal{P}\phi_j - \mathcal{P}_m \phi_j) \right\| \leq \left[\sum_{j=1}^n \|\mathcal{P}\phi_j - \mathcal{P}_m \phi_j\|^2 \right]^{\frac{1}{2}},$$

where $\|\alpha\|^2 = \sum_{j=1}^n |\alpha_j|^2$. Since $\mathcal{P}_m \to \mathcal{P}$ strongly on H, the result now follows. ☐

Recall that the issue with consistent reconstructions is that the subspaces T_n and S_n can become near perpendicular for large n. Provided m is sufficiently large, Lemma 3.6 shows that this cannot happen with T_n and $\mathcal{P}_m(\mathrm{T}_n)$.

Let us also note one further fact about $C_{n,m}$: it coincides with the constant $C_{n,n} = \cos\theta_{\mathrm{T}_n,\mathrm{S}_n}$ of consistent sampling whenever $m = n$. In view of Lemma 3.1, this result is of no surprise. However, it will be of use subsequently.

Lemma 3.7. *When $m = n$ the subspaces angles $\cos\theta_{\mathrm{T}_n,\mathcal{P}_m(\mathrm{T}_n)}$ and $\cos\theta_{\mathrm{T}_n,\mathrm{S}_n}$ are equal.*

Proof. As shown in the proof of Lemma 3.1, the subspaces $\mathcal{P}_n(\mathrm{T}_n)$ and S_n coincide. The result now follows immediately. ☐

We are now able to prove the main theorem concerning generalised sampling:

Theorem 3.8. *Let $f \in \mathrm{H}$. Then, for each $n \in \mathbb{N}$, there exists an m_0, independent of f, such that the approximation $\tilde{f}_{n,m}$ defined by (3.1) exists and is unique for all $m \geq m_0$. In particular, m_0 is the least m such that $C_{n,m} > 0$. Moreover, the mapping $f \mapsto \tilde{f}_{n,m}$ is precisely the oblique projection onto T_n along $\mathcal{P}_m(\mathrm{T}_n)$, and we have the sharp bounds*

$$\|\tilde{f}_{n,m}\| \leq \frac{1}{C_{n,m}} \|f\|, \tag{3.11}$$

and

$$\|f - \mathcal{Q}_n f\| \leq \|f - \tilde{f}_{n,m}\| \leq \frac{1}{C_{n,m}} \|f - \mathcal{Q}_n f\|, \tag{3.12}$$

where $C_{n,m} = \cos\theta_{\mathrm{T}_n,\mathcal{P}_m(\mathrm{T}_n)}$.

Proof. The existence of an m_0 such that $C_{n,m} > 0$ for all $m \geq m_0$ follows from Lemma 3.6. Thus, when $m \geq m_0$ the oblique projection $\mathcal{W}$ onto T_n along $[\mathcal{P}_m(\mathrm{T}_n)]^\perp$ is well-defined and unique. We now argue as in the proof of Theorem 3.2. Since the equations (3.1) are equivalent to (3.2), and since $\mathcal{W}f$ also satisfies (3.2), we must have that $\tilde{f}_{n,m} = \mathcal{W}f$. Applications of Theorems 2.5 and 2.6 now complete the proof. ☐

To connect generalised sampling to the narrative of §2.3, we note that if $\tilde{f}_{n,m} = \sum_{j=1}^n \alpha_j^{[n,m]} \phi_j$, then the vector $\alpha^{[n,m]} = \{\alpha_1^{[n,m]}, \ldots, \alpha_n^{[n,m]}\} \in \mathbb{C}^n$ is the unique solution to

$$(U^{[n,m]})^* U^{[n,m]} \alpha^{[n,m]} = (U^{[n,m]})^* \hat{f}^{[m]},$$

where $U^{[n,m]} \in \mathbb{C}^{m \times n}$ is precisely $P_m U P_n$ and $\hat{f}^{[m]} = P_m \hat{f}$. In other words, $\alpha^{[n,m]}$ is the solution to the overdetermined least squares problem $U^{[n,m]} \alpha^{[n,m]} \approx \hat{f}^{[m]}$, and thus can be computed in a completely straightforward manner.

Observe that $U^{[m,n]}$ is the leading $m \times n$ submatrix of U, and is consequently often referred to as an *uneven section* of U. Uneven section techniques have recently gained prominence as effective alternatives to the finite section method for discretising non-self adjoint operators [22, 26]. In particular, in [24] they were employed to solve the long-standing computational spectral problem. Their success is due to the observation that, under a number of assumptions (which are always guaranteed for the problem we consider in this paper), we have

$$(U^{[n,m]})^* U^{[n,m]} \to P_n U^* U P_n, \quad m \to \infty,$$

where $P_n U^* U P_n$ is the $n \times n$ finite section of the self-adjoint matrix $U^* U$. This guarantees properties (i)–(iii) for $U^{[n,m]}$, whenever m is sufficiently large in comparison to n.

Finite (and uneven) sections have been extensively studied [7, 25, 30], and there exists a well-developed and intricate theory of their properties involving C^*-algebras [23]. However, these general results say little about the rate of convergence, nor do they provide explicit constants. Yet, as illustrated in Theorem 3.8, the operator U in this case is so structured that its uneven sections admit both explicit constants and estimates for the rate of convergence. Moreover, of great importance for computations, such constants can also be numerically computed, as we demonstrate in §3.4.

3.3 Oblique asymptotic optimality

Recall the intuitive arguments of §3.1: namely, $\tilde{f}_{n,m} \approx \tilde{f}_n$ for all large m, where $\tilde{f}_n$ is the oblique projection onto T_n along $[\mathcal{P}(\mathrm{T}_n)]^\perp$. We refer to this property as *oblique asymptotic optimality*. Due to (3.4), the bilinear form $\langle \cdot, \cdot \rangle_{\mathcal{P}} = \langle \mathcal{P} \cdot, \cdot \rangle$ gives an inner product on T which is equivalent to $\langle \cdot, \cdot \rangle$. Hence, whenever $f \in \mathrm{T}$, one finds that $\tilde{f}_n$ is precisely the orthogonal projection of f onto T_n with respect to $\langle \cdot, \cdot \rangle_{\mathcal{P}}$, and thus $\tilde{f}_n$ is the best approximation to f from T_n in the corresponding norm $\| \cdot \|_{\mathcal{P}}$.

If the sampling vectors $\{\psi_j\}$ are orthonormal, $\mathcal{P} = \mathcal{I}$ is the identity and $\tilde{f}_n$ coincides with the orthogonal projection $\mathcal{Q}_n f$. In this case, Theorem 3.8 confirms oblique asymptotic optimality. Indeed, due to Lemma 3.14, the upper bound in (3.12) satisfies

$$\frac{1}{C_{n,m}} \| f - \mathcal{Q}_n f \| \to \| f - \mathcal{Q}_n f \|, \quad m \to \infty,$$

and therefore, since $\mathcal{Q}_n f$ is unique, $\tilde{f}_{n,m} \to \mathcal{Q}_n f$. It this case, we say that $\tilde{f}_{n,m}$ is *asymptotically* optimal.

On the other hand, when the sampling vectors $\{\psi_j\}$ are not orthonormal, $\tilde{f}_{n,m}$ need not be asymptotically optimal, since $C_{n,m} \nrightarrow 1$ as $m \to \infty$. Moreover, Theorem 3.8 does not confirm oblique asymptotic optimality either. For this, we first require an estimate for $\| f - \tilde{f}_{n,m} \|$ in terms of $\| f - \tilde{f}_n \|$. Fortunately, the interpretation in terms of oblique projections provides such an estimate.

Suppose that we consider the subspace T of H and the inner product $\langle \cdot, \cdot \rangle_{\mathcal{P}}$. Let us define the quantity

$$C^{\mathcal{P}}_{n,m} = \cos \theta_{\mathrm{T}_n, \mathcal{P}^{-1}(\mathcal{P}_m(\mathrm{T}_n))} = \inf_{\substack{\phi \in \mathrm{T}_n \\ \phi \neq 0}} \| \mathcal{Q}_{\mathcal{P}^{-1}(\mathcal{P}_m(\mathrm{T}_n))} \phi \|_{\mathcal{P}}. \tag{3.13}$$

Recall that $\mathcal{P} : \mathrm{T} \to \mathcal{P}(\mathrm{T})$ is invertible (Lemma 3.3), and that $\mathcal{P}_m \phi \in \mathcal{P}(\mathrm{T})$ for any m and $\phi \in \mathrm{T}$. Hence this quantity is well-defined. Moreover, we have

Lemma 3.9. *Let $C^{\mathcal{P}}_{n,m}$ be given by (3.13). Then, for fixed $n \in \mathbb{N}$, $C^{\mathcal{P}}_{n,m} \to 1$ as $m \to \infty$.*

Proof. As in Lemma 3.6, we notice that

$$1 \geq C^{\mathcal{P}}_{n,m} \geq \inf_{\substack{\phi \in \mathrm{T}_n \\ \|\phi\| = 1}} \left\{ \frac{\langle \mathcal{P}^{-1} \mathcal{P}_m \phi, \phi \rangle_{\mathcal{P}}}{\| \mathcal{P}^{-1} \mathcal{P}_m \phi \|_{\mathcal{P}}} \right\}.$$

Since $\mathcal{P}_m \to \mathcal{P}$ strongly on T, it follows that $\mathcal{P}^{-1} \mathcal{P}_m \phi \to \phi$ for any fixed $\phi \in \mathrm{T}$. We now argue as in the proof of Lemma 3.6. □

With this lemma to hand, we have

Theorem 3.10. *Suppose that* $f \in \mathrm{T}$ *and let* $\tilde{f}_{n,m}$ *be defined by (3.1), where* $m \geq m_0$ *and* m_0 *is as in Theorem 3.8. Then we have the sharp bounds*

$$\|\tilde{f}_{n,m}\|_{\mathcal{P}} \leq \frac{1}{C^{\mathcal{P}}_{n,m}} \|f\|_{\mathcal{P}}, \tag{3.14}$$

and

$$\|f - \tilde{f}_n\|_{\mathcal{P}} \leq \|f - \tilde{f}_{n,m}\|_{\mathcal{P}} \leq \frac{1}{C^{\mathcal{P}}_{n,m}} \|f - \tilde{f}_n\|_{\mathcal{P}}, \tag{3.15}$$

where $C^{\mathcal{P}}_{n,m}$ *is as in (3.13). In particular,* $\tilde{f}_{n,m} \to \tilde{f}_n$ *as* $m \to \infty$.

Proof. Let $\phi \in \mathrm{T}_n$ be arbitrary. Then

$$\|\mathcal{Q}_{\mathcal{P}^{-1}(\mathcal{P}_m(\mathrm{T}_n))}\phi\|_{\mathcal{P}} = \sup_{\substack{\phi' \in \mathrm{T}_n \\ \phi' \neq 0}} \left\{ \frac{\langle \phi, \mathcal{P}^{-1}\mathcal{P}_m \phi' \rangle_{\mathcal{P}}}{\|\mathcal{P}^{-1}\mathcal{P}_m \phi'\|_{\mathcal{P}}} \right\} = \sup_{\substack{\phi' \in \mathrm{T}_n \\ \phi' \neq 0}} \left\{ \frac{\langle \phi, \mathcal{P}_m \phi' \rangle}{\|\mathcal{P}^{-1}\mathcal{P}_m \phi'\|_{\mathcal{P}}} \right\}.$$

Since $\mathcal{P} : \mathrm{T} \to \mathcal{P}(\mathrm{T})$ is continuously invertible, we have

$$\|\mathcal{P}^{-1}\mathcal{P}_m \phi'\|_{\mathcal{P}} \leq c \|\mathcal{P}_m \phi'\|,$$

for some $c > 0$ independent of m and ϕ'. Hence

$$\|\mathcal{Q}_{\mathcal{P}^{-1}(\mathcal{P}_m(\mathrm{T}_n))}\phi\|_{\mathcal{P}} \geq \frac{1}{c} \sup_{\substack{\phi' \in \mathrm{T}_n \\ \phi' \neq 0}} \left\{ \frac{\langle \phi, \mathcal{P}_m \phi' \rangle}{\|\mathcal{P}_m \phi'\|} \right\} = \|\mathcal{Q}_{\mathcal{P}_m(\mathrm{T}_n)}\phi\|,$$

where $\mathcal{Q}_{\mathcal{P}_m(\mathrm{T}_n)}$ is the orthogonal projection onto $\mathcal{P}_m(\mathrm{T}_n)$ with respect to the canonical inner product $\langle \cdot, \cdot \rangle$. It now follows that $C^{\mathcal{P}}_{n,m} \geq c^{-1} C_{n,m}$, where $C_{n,m}$ is as in Lemma 3.6. Since $m \geq m_0$ we deduce that $C^{\mathcal{P}}_{n,m} > 0$. In particular, the oblique projection $\mathcal{W}$ onto T_n along $[\mathcal{P}^{-1}(\mathcal{P}_m(\mathrm{T}_n))]^{\perp}$ with respect to $\langle \cdot, \cdot \rangle_{\mathcal{P}}$ is well-defined and unique. Moreover, the equations (3.2) defining $\tilde{f}_{n,m}$ are equivalent to

$$\langle \tilde{f}_{n,m}, \Phi \rangle_{\mathcal{P}} = \langle f, \Phi \rangle_{\mathcal{P}}, \quad \forall \phi \in \mathcal{P}^{-1}\mathcal{P}_m(\mathrm{T}_n).$$

Since $\mathcal{W}f$ also satisfies these equations, we obtain $\mathcal{W}f = \tilde{f}_{n,m}$. The estimates (3.14) and (3.15) are now given by Theorems 2.5 and 2.6. For the final part of the proof, we apply Lemma 3.9 to (3.15) and use the fact that $\tilde{f}_n$ is the best approximation to f from T_n with respect to $\langle \cdot, \cdot \rangle_{\mathcal{P}}$. □

Note that oblique asymptotic optimality was first shown in [4]. This theorem improves on that work by providing sharp bounds.

3.4 The stable sampling rate

The key component of generalised sampling is that the parameter m (the number of samples) must be sufficiently large in comparison to n (the number of degrees of freedom). The purpose of this section is to precisely quantify this notion. This follows along similar lines to [1, 4].

The question of how large m must be motivates the following definition:

Definition 3.11. *For* $n \in \mathbb{N}$ *and* $a \in (0, \frac{c_1}{c_3})$, *where* c_1 *and* c_3 *are as in Lemma 3.5, the stable sampling rate is defined by*

$$\Theta(n; a) = \min \{ m \in \mathbb{N} : C_{n,m} > a \}.$$

Note that an equivalent definition of $\Theta(n; a)$ is that it is the minimum m such that the angle θ between T_n and $\mathcal{P}_m(\mathrm{T}_n)$ satisfies $\cos\theta > a$.

The stable sampling rate, as its name suggests, measures how large m must be (for a given n) to ensure guaranteed, stable and quasi-optimal recovery. Indeed, choosing $m \geq \Theta(n; a)$, we find that

$$\|\tilde{f}_{n,m}\| \leq \frac{1}{a} \|f\|,$$

and

$$\|f - \mathcal{Q}_n f\| \leq \|f - \tilde{f}_{n,m}\| \leq \frac{1}{a}\|f - \mathcal{Q}_n f\|.$$

Thus, $\tilde{f}_{n,m}$ is quasi-optimal to f from the set T_n (with constant a^{-1}), provided $m \geq \Theta(n;a)$. Moreover, the condition number of the matrix $U^{[n,m]}$ is also determined by this quantity:

Lemma 3.12. *Let $A^{[n]} \in \mathbb{C}^{n\times n}$ be the Gram matrix for the vectors $\{\phi_1,\ldots,\phi_n\}$. Then*

$$\frac{d_1 C_{n,m}}{\sqrt{c_2}}\sqrt{\kappa(A^{[n]})} \leq \kappa(U^{[n,m]}) \leq \frac{\sqrt{c_2}}{d_1 C_{n,m}}\sqrt{\kappa(A^{[n]})}, \tag{3.16}$$

where d_1 and c_2 are given by (2.1) and (3.4) respectively.

Proof. The condition number of $U^{[n,m]}$ is precisely the ratio of its maximal and minimal singular values. Let σ be a singular value of $U^{[n,m]}$ with corresponding normalised vector $\alpha = \{\alpha_1,\ldots,\alpha_n\}$ satisfying

$$(U^{[n,m]})^* U^{[n,m]}\alpha = \sigma^2\alpha.$$

By linearity, we deduce that

$$\langle \mathcal{P}_m\phi,\phi\rangle = \sigma^2\|\alpha\|^2 = \sigma^2.$$

First, note that

$$\langle \mathcal{P}_m\phi,\phi\rangle = \sum_{j=1}^{m}|\langle\phi,\psi_j\rangle|^2 \leq \sum_{j=1}^{\infty}|\langle\phi,\psi_j\rangle|^2 = \langle\mathcal{P}\phi,\phi\rangle \leq c_2\|\phi\|^2,$$

where c_2 is as in (3.4). Thus

$$\sigma^2 \leq c_2\|\phi\|^2 \leq c_2 \sup_{\substack{\alpha\in\mathbb{C}^n\\ \|\alpha\|=1}}\left\|\sum_{j=1}^{n}\alpha_j\phi_j\right\|^2 = c_2\lambda_{\max},$$

where $\lambda_{\max}$ is the maximal eigenvalue of $A^{[n]}$. This establishes an upper bound for σ.

Next we consider a lower bound. Observe that

$$\|\mathcal{Q}_{\mathcal{P}_m(\mathrm{T}_n)}\phi\| = \sup_{\substack{\psi\in\mathcal{P}_m(\mathrm{T}_n)\\ \|\psi\|=1}}\langle\phi,\psi\rangle \leq \sup_{\substack{\psi\in\mathrm{S}_m\\ \|\psi\|=1}}\langle\phi,\psi\rangle = \sup_{\substack{\alpha\in\mathbb{C}^n\\ \alpha\neq 0}}\frac{\sum_{j=1}^{m}\alpha_j\langle\phi,\psi_j\rangle}{\|\sum_{j=1}^{m}\alpha_j\psi_j\|}.$$

Using the Cauchy–Schwarz inequality, as well as the bound (2.1), we deduce that

$$\|\mathcal{Q}_{\mathcal{P}_m(\mathrm{T}_n)}\phi\| \leq \frac{1}{d_1}\left[\sum_{j=1}^{m}|\langle\phi,\psi_j\rangle|^2\right]^{\frac{1}{2}} = \frac{1}{d_1}\sqrt{\langle\mathcal{P}_m\phi,\phi\rangle}.$$

From this it follows that

$$\langle\mathcal{P}_m\phi,\phi\rangle \geq d_1^2 C_{n,m}^2\|\phi\|^2, \tag{3.17}$$

where $C_{n,m}$ is as in Lemma 3.6. Therefore

$$\sigma^2 \geq d_1^2 C_{n,m}\|\phi\|^2 \geq d_1^2 C_{n,m}\lambda_{\min},$$

where $\lambda_{\min}$ is the minimal eigenvalue of $A^{[n]}$. Since $A^{[n]}$ is self-adjoint, its condition number is precisely $\lambda_{\max}/\lambda_{\min}$. Therefore

$$\frac{d_1^2 C_{n,m}}{c_2}\kappa(A^{[n]}) = \frac{d_1^2 C_{n,m}\lambda_{\max}}{c_2\lambda_{\min}} \leq [\kappa(U^{[n,m]})]^2 \leq \frac{c_2\lambda_{\max}}{d_1^2 C_{n,m}\lambda_{\min}} = \frac{c_2}{d_1^2 C_{n,m}}\kappa(A^{[n]}),$$

which gives (3.16). □

As a result of this lemma, the matrix $U^{[n,m]}$ is no worse conditioned than the Gram matrix $A^{[n]}$, whenever $m \geq \Theta(n;a)$. In particular, if the reconstruction vectors $\{\phi_1,\ldots,\phi_n\}$ form a Riesz basis, then $\kappa(A^{[n]}) = \mathcal{O}(1)$ for all n and the computation of $\tilde{f}_{n,m}$ is consequently completely numerically stable.

At this point, it is also worth noting that $\tilde{f}_{n,m}$, being the image of f under an oblique projection, depends only on the subspace T_n, and is completely independent of the choice of reconstruction vectors. As the previous lemma demonstrates, such a choice only affects numerical stability.

The stable sampling rate is named for precisely the reason that it indicates when the computation of $\tilde{f}_{n,m}$ is guaranteed to be numerically stable. Note that it is completely at odds with the classical viewpoint of sampling theory. Sampling rates typically deal with how much can one obtain from the given samples, with the obvious notion being that if one has m samples then one has m corresponding degrees of freedom with which to form a reconstruction. However, whilst this viewpoint is perfectly valid, when the important considerations of stability and accuracy are taken into account, one needs to ask a somewhat different question.

Whilst the stable sampling rate is intimately related to generalised sampling, an important issue we raise in this paper is the possibility that it may be actually a *universal* quantity. In other words, any stable and accurate method that reconstructs in the space T_n from m samples, must have $m \geq \Theta(n;a)$. The optimality results of §4 indicate that this may well be the case in general, although we currently have no proof.

3.5 Determining the stable sampling rate

The stable sampling rate becomes a useful quantity for practical computations after noting that it can always be computed. This follows from the observation that $C_{n,m}$ is itself computable:

Lemma 3.13. *Let $A^{[n]}$ be as in Lemma 3.12, and suppose that $C^{[m]} \in \mathbb{C}^{m\times m}$ is the Gram matrix for the vectors $\{\psi_1,\ldots,\psi_m\}$. Then the quantity $C^2_{n,m} = \cos^2\theta_{\mathrm{T}_n,\mathcal{P}_m(\mathrm{T}_n)}$ is precisely the minimal generalised eigenvalue of the matrix pencil $\{B^{[n,m]},A^{[n]}\}$, where*

$$B^{[n,m]} = (U^{[n,m]})^* U^{[n,m]} \left((U^{[n,m]})^* C^{[m]} U^{[n,m]} \right)^{-1} (U^{[n,m]})^* U^{[n,m]}.$$

Proof. By definition

$$C^2_{n,m} = \inf_{\substack{\phi\in\mathrm{T}_n\\ \phi\neq 0}} \left\{ \frac{\|\mathcal{Q}_{\mathcal{P}_m(\mathrm{T}_n)}\phi\|^2}{\|\phi\|^2} \right\} = \inf_{\substack{\alpha\in\mathbb{C}^n\\ \alpha\neq 0}} \left\{ \frac{\sum_{j,k=1}^n \alpha_j \overline{\alpha_k} \langle \mathcal{Q}_{\mathcal{P}_m(\mathrm{T}_n)}\phi_j, \mathcal{Q}_{\mathcal{P}_m(\mathrm{T}_n)}\phi_k \rangle}{\sum_{j,k=1}^n \alpha_j \overline{\alpha_k}\langle \phi_j,\phi_k\rangle} \right\}. \tag{3.18}$$

The denominator coincides with $\alpha^* A^{[n]}\alpha$. Now consider the numerator. Suppose that we write

$$\mathcal{Q}_{\mathcal{P}_m(\mathrm{T}_n)}\phi_j = \sum_{l=1}^n \beta^{[j]}_r \mathcal{P}_m \phi_r,$$

for some vector $\beta^{[j]} = (\beta^{[j]}_1,\ldots,\beta^{[j]}_n)$, so that

$$\langle \mathcal{Q}_{\mathcal{P}_m(\mathrm{T}_n)}\phi_j, \mathcal{Q}_{\mathcal{P}_m(\mathrm{T}_n)}\phi_k \rangle = \sum_{r,s=1}^n \beta^{[j]}_r \overline{\beta^{[k]}_s} \langle \mathcal{P}_m\phi_r, \mathcal{P}_m\phi_s\rangle = (\beta^{[k]})^* \left((U^{[n,m]})^* C^{[m]} U^{[n,m]} \right) \beta^{[j]}.$$

By definition,

$$\sum_{l=1}^n \beta^{[j]}_r \langle \mathcal{P}_m\phi_r, \mathcal{P}_m\phi_k\rangle = \langle \phi_j, \mathcal{P}_m\phi_k\rangle, \quad k=1,\ldots,n.$$

Therefore

$$\beta^{[j]}_r = \left((U^{[n,m]})^* C^{[m]} U^{[n,m]} \right)^{-1}_{r,k} \left((U^{[n,m]})^* U^{[n,m]} \right)_{k,j}.$$

Substituting this into (3.18) and simplifying now gives

$$C^2_{n,m} = \inf_{\substack{\alpha\in\mathbb{C}^n\\ \alpha\neq 0}} \left\{ \frac{\alpha^* B^{[n,m]}\alpha}{\alpha^* A^{[n]}\alpha} \right\}.$$

It is straightforward to see that this coincides with the minimal eigenvalue of the pencil $\{B^{[n,m]}, A^{[n]}\}$. □

Whilst this lemma indicates that $C_{n,m}$ is always computable, the matrices involved, in particular $B^{[n,m]}$, may be somewhat complicated to evaluate. However, it is always possible to estimate $C_{n,m}$ above and below with a simpler quantity:

Lemma 3.14. *For $n, m \in \mathbb{N}$, ket $C'_{n,m}$ be given by*

$$C'_{n,m} = \inf_{\substack{\phi \in \mathrm{T}_n \\ \|\phi\|=1}} \sqrt{\langle \mathcal{P}_m \phi, \phi \rangle},$$

Then $C'_{n,m}$ is the minimal generalised eigenvalue of the matrix pencil $\{(U^{[n,m]})^* U^{[n,m]}, A^{[n]}\}$. Moreover, $C'_{n,m}$ satisfies

$$\frac{1}{d_2} C'_{n,m} \le C_{n,m} \le \frac{1}{d_1} C'_{n,m}, \tag{3.19}$$

where d_1 and d_2 are as in (2.1). In particular, whenever the sampling vectors $\{\psi_j\}_{j=1}^{\infty}$ are orthonormal, we have $C'_{n,m} = C_{n,m}$.

Proof. The first part of the proof is similar to the proof of Lemma 3.13. Moreover, the upper bound in (3.19) is given immediately by (3.17). For the lower bound, we first note that, for any $\phi \in \mathrm{T}_n$,

$$\|\mathcal{P}_m \phi\| = \left\| \sum_{j=1}^{m} \langle \phi, \psi_j \rangle \psi_j \right\| \le d_2 \left[\sum_{j=1}^{m} |\langle \phi, \psi_j \rangle|^2 \right]^{\frac{1}{2}} = d_2 \sqrt{\langle \mathcal{P}_m \phi, \phi \rangle}.$$

An application of Lemma 3.6 now gives the result. Finally, when $\{\psi_j\}_{j=1}^{\infty}$ are orthonormal we have $d_1 = d_2 = 1$, and equality therefore follows from (3.19). □

This lemma gives a somewhat simpler quantity to compute than $C_{n,m}$. Note that $C'_{n,m}$ was originally used in [4] to analyse generalised sampling.

3.6 Polynomial reconstructions from Fourier samples

It is easiest to illustrate the stable sampling rate with the concrete example of §2.4. As shown therein, consistent reconstruction of a function f from its first n Fourier samples in terms of a polynomial of degree n is both numerically unstable and not guaranteed to converge, unless the function is sufficiently regular. With generalised sampling, we know that stability and convergence, for *all* functions $f \in \mathrm{L}^2(-1,1)$, are guaranteed provided $m \ge \Theta(n; a)$. The question of how $\Theta(n; a)$ scales with n was studied in depth in [4]. Therein it was shown that $\Theta(n; a) = \mathcal{O}\left(n^2\right)$, meaning that to recover a polynomial approximation to f of degree n one needs $\mathcal{O}\left(n^2\right)$ Fourier samples. Numerical results confirm the sharpness of this estimate. Note that the application of generalised sampling to this specific problem is equivalent to a method first proposed in [28].

To demonstrate the improvement offered by generalised sampling, in Table 2 we consider the example $f(x) = \frac{1}{1+16x^2}$ once more. Observe that, unlike the case of the consistent reconstruction (see Table 1), the approximation $\tilde{f}_{n,m}$ converges to f. In particular, f is recovered to 10 digits of accuracy using only $m = 1250$ Fourier coefficients and a polynomial of degree $n = 100$. Moreover, for all n, the error $\|f - \tilde{f}_{n,m}\|$ is virtually indistinguishable from $\|f - \mathcal{Q}_n f\|$ – the best achievable error from the reconstruction space T_n. Note also the close correspondence between the sharp bound (3.12) and the true error: the upper bound is less than ten times larger than the true value in this case.

In Figure 2 we revisit the noisy example of §2.4. As is evident, generalised sampling is completely robust in the presence of noise, with the best achievable error being solely determined by the magnitude of the noise. Moreover, once f is recovered perfectly up to $\mathcal{O}(\epsilon)$, where ϵ is the noise parameter, there is no drift in the error, unlike the consistent reconstruction (see Figure 1).

Whilst generalised sampling is certainly effective in overcoming the issues discussed in §2.4, it is perhaps disappointing that one requires $m = \mathcal{O}\left(n^2\right)$ Fourier samples to reconstruct. An obvious question

n	20	40	60	80	100
$\|f - \tilde{f}_{n,m}\|$	5.09e-3	3.55e-5	2.49e-7	1.76e-9	1.24e-11
$\|f - \mathcal{Q}_n f\|$	4.87e-2	3.45e-5	2.45e-7	1.74e-9	1.23e-11
$\frac{1}{C_{n,m}}\|f - \mathcal{Q}_n f\|$	2.37e-1	1.60e-4	1.12e-6	7.85e-9	5.54e-11

Table 2: Reconstruction of $f(x) = \frac{1}{1+16x^2}$ by generalised sampling with $m = \frac{1}{8}n^2$.

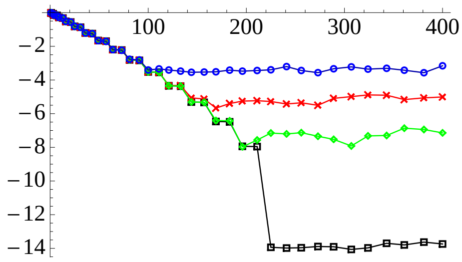

Figure 2: The quantity $\log_{10} \|f - \tilde{f}_{n,m}\|_\infty$ against m with $n = \lceil\sqrt{4m}\rceil$, where $f(x) = x^{30}$ and $\tilde{f}_{n,m}$ was computed from (i) noise-free data (squares) and (ii) noise at amplitude $\epsilon = 10^{-4}$ (circles), 10^{-6} (crosses) and 10^{-8} (diamonds).

to ask is whether another method can be both as stable and as accurate as generalised sampling whilst requiring fewer samples. In other words, is the stable sampling rate a property only of relevance for generalised sampling, or is it universal? This question is the topic of the next question.

4 Optimality

Recall that the mapping $f \mapsto \tilde{f}_{n,m}$ is the oblique projection onto T_n along $(\mathcal{P}_m(\mathrm{T}_n))^\perp$. An obvious question to ask is, can another oblique projection give a better reconstruction? That is to say, a reconstruction which is both more stable and more accurate. Under some natural assumptions, one finds that the answer is no.

4.1 Optimal projection-type method

Suppose that $\mathrm{L}_{n,m} \subseteq \mathrm{H}$ is a subspace, and assume that

$$\mathrm{L}_{n,m} \subseteq \mathrm{S}_m, \tag{4.1}$$

and that

$$\mathrm{T}_n \cap (\mathrm{L}_{n,m})^\perp = \{0\}. \tag{4.2}$$

Note that the former condition is natural to impose, since we only have access to the samples $\hat{f}_j = \langle f, \psi_j \rangle$ of f, where $\psi_j \in \mathrm{S}_m$, $j = 1, \ldots, m$. The second condition ensures that there exists a unique oblique projection $f \mapsto \tilde{f}^{\mathrm{L}}_{n,m}$ onto T_n along $(\mathrm{L}_{n,m})^\perp$ [13].

Definition 4.1. *A reconstruction method $\mathcal{G}_{n,m} : \mathrm{H} \to \mathrm{T}_n$, $f \mapsto \tilde{f}_{n,m}$ is of projection-type if $\mathcal{G}_{n,m}$ is the oblique projection onto T_n along $(\mathrm{L}_{n,m})^\perp$, where $L_{n,m}$ is a subspace of H satisfying (4.1) and (4.2).*

The optimality result we next prove concerns projection-type methods. First, it is useful to note the following:

Theorem 4.2. *Let $\mathcal{G}_{n,m}$ be a projection-type method and write*

$$C^{\mathrm{L}}_{n,m} = \cos \theta_{\mathrm{T}_n, \mathrm{L}_{n,m}}. \tag{4.3}$$

Suppose that $n, m \in \mathbb{N}$ are such that $C^{\mathrm{L}}_{n,m} > 0$. Then the projection-type method $\mathcal{G}_{n,m} : f \mapsto \tilde{f}^{\mathrm{L}}_{n,m}$ is well-defined, and we have the sharp bounds

$$\|\tilde{f}^{\mathrm{L}}_{n,m}\| \leq \frac{1}{C^{\mathrm{L}}_{n,m}} \|f\|,$$

and

$$\|f - \mathcal{Q}_n f\| \leq \|f - \tilde{f}^{\mathrm{L}}_{n,m}\| \leq \frac{1}{C^{\mathrm{L}}_{n,m}} \|f - \mathcal{Q}_n f\|.$$

Proof. Since any projection-type method is an oblique projection, the result follows from Theorem 2.5. □

As with generalised sampling, the constant $C^{\mathrm{L}}_{n,m}$ controls both the stability and the accuracy of any projection-type method (note that generalised sampling is the projection-type method with $\mathrm{L}_{n,m} = \mathcal{P}_m(\mathrm{T}_n)$). Therefore, a reasonable question to ask if whether, for given n and m, one can obtain a better, i.e. larger, constant $C^{\mathrm{L}}_{n,m}$ with a suitable alternative choice of $\mathrm{L}_{n,m}$. Our first result is as follows:

Theorem 4.3. *Suppose that $\{\psi_j\}$ are orthonormal and that $\mathrm{L}_{n,m} \subseteq \mathrm{H}$ satisfies (4.1) and (4.2). Then, if $C_{n,m} = \cos\theta_{\mathrm{T}_n, \mathcal{P}_m(\mathrm{T}_n)}$ and $C^{\mathrm{L}}_{n,m} = \cos\theta_{\mathrm{T}_n, \mathrm{L}_{n,m}}$, we have*

$$C^{\mathrm{L}}_{n,m} \leq C_{n,m}.$$

Consequently, no projection-type method can be both more stable and more accurate than generalised sampling.

Proof. By definition

$$(C^{\mathrm{L}}_{n,m})^2 = \inf_{\substack{\phi \in \mathrm{T}_n \\ \|\phi\|=1}} \|\mathcal{Q}_{\mathrm{L}_{n,m}} \phi\|^2 = 1 - \sup_{\substack{\phi \in \mathrm{T}_n \\ \|\phi\|=1}} \|\phi - \mathcal{Q}_{\mathrm{L}_{n,m}} \phi\|^2.$$

Now $\mathcal{Q}_{\mathrm{L}_{n,m}} \phi \in \mathrm{S}_m$, and therefore

$$\|\phi - \mathcal{Q}_{\mathrm{L}_{n,m}} \phi\| \geq \|\phi - \mathcal{Q}_{\mathrm{S}_m} \phi\| = \|\phi - \mathcal{P}_m \phi\|,$$

since the vectors $\{\psi_j\}_{j=1}^{\infty}$ are orthonormal. Thus we obtain

$$(C^{\mathrm{L}}_{n,m})^2 \leq 1 - \sup_{\substack{\phi \in \mathrm{T}_n \\ \|\phi\|=1}} \|\phi - \mathcal{P}_m \phi\|^2 = \inf_{\substack{\phi \in \mathrm{T}_n \\ \|\phi\|=1}} \|\mathcal{P}_m \phi\|^2 = (C_{n,m})^2,$$

as required. □

Although this theorem confirms that no projection-type method will outperform generalised sampling, it assumes an orthonormal sampling basis. When the sampling vectors constitute a Riesz basis, it may appear possible to outperform generalised sampling. As we show next, this is true, but only in a very limited sense.

Let us first recall the stable sampling rate (Definition 3.11)

$$\Theta(n; a) = \min\left\{m \in \mathbb{N} : C_{n,m} > a\right\}, \quad a \in (0, \tfrac{c_1}{c_3}).$$

We may define the same quantity for any projection-type method:

$$\Theta^{\mathrm{L}}(n; a) = \min\left\{m \in \mathbb{N} : C^{\mathrm{L}}_{n,m} > a\right\}, \quad a \in (0, c^*).$$

Here $0 < c^* < 1$ is such that $\Theta^{\mathrm{L}}(n; a)$ is well-defined for all $n \in \mathbb{N}$ and $0 < a < c^*$. We shall assume that such a c^* exists. Clearly if not, then one cannot choose a sequence $m = m(n)$ such that $C^{\mathrm{L}}_{n,m(n)}$ remains bounded away from zero, and therefore the projection-type method cannot be stable.

Recall that, for given $a > 0$, setting $m \geq \Theta(n; a)$ (respectively $\Theta^{\mathrm{L}}(n; a)$) ensures stability and accuracy of the method to within a constant factor of a^{-1}. Our interest lies in the asymptotic growth of $\Theta(n; a)$ as $n \to \infty$: this determines the precise scaling of the number of samples m with respect to n for stable, quasi-optimal recovery. Therefore, a natural question to pose is, can a suitable choice of $\mathrm{L}_{n,m}$ give a projection-type method with a stable sampling rate $\Theta^{\mathrm{L}}(n; a)$ possessing milder asymptotic growth than that of $\Theta(n; a)$? As we now show, this cannot occur. We first require the following lemma:

Lemma 4.4. *Let $C_{n,m}$ and $C^{\mathrm{L}}_{n,m}$ be as in Theorem 4.3. Then*

$$C_{n,m} \geq \frac{d_1}{d_2}\left[1-\sqrt{1-(C^{\mathrm{L}}_{n,m})^2}\right].$$

Proof. Recall from the proof of Theorem 4.3 that

$$(C^{\mathrm{L}}_{n,m})^2 \leq 1 - \sup_{\substack{\phi\in \mathrm{T}_n\\ \|\phi\|=1}} \|\phi - \mathcal{Q}_{\mathrm{S}_m}\phi\|^2.$$

Observe that $\|\phi - \mathcal{Q}_{\mathrm{S}_m}\phi\| \geq 1 - \|\mathcal{Q}_{\mathrm{S}_m}\phi\|$. Moreover, arguing as in the proof of Lemma 3.12, we find that

$$\|\mathcal{Q}_{\mathrm{S}_m}\phi\| \leq \frac{1}{d_1}\sqrt{\langle \mathcal{P}_m\phi,\phi\rangle}.$$

Hence

$$(C^{\mathrm{L}}_{n,m})^2 \leq 1 - \left[1-\frac{1}{d_1}\inf_{\substack{\phi\in \mathrm{T}_n\\ \|\phi\|=1}}\sqrt{\langle \mathcal{P}_m\phi,\phi\rangle}\right]^2.$$

Using Lemma 3.14, we obtain

$$(C^{\mathrm{L}}_{n,m})^2 \leq 1 - \left[1-\frac{d_2}{d_1}C_{n,m}\right]^2,$$

and rearranging now gives the result. □

Theorem 4.5. *For $a\in(0,1)$ suppose that $\Theta^{\mathrm{L}}(n;a)$ is well-defined for all n. Then*

$$\Theta(n;a') \leq \Theta^{\mathrm{L}}(n;a),$$

where $a' = \frac{d_1}{d_2}(1-\sqrt{1-a^2})$. Equivalently, if there exists a sequence $\{m(n)\}^{\infty}_{n=1}$ such that $C^{\mathrm{L}}_{n,m(n)} \geq a$ for all $n\in\mathbb{N}$, then $C_{n,m(n)} \geq a'$.

Proof. For $n\in\mathbb{N}$ let $m = \Theta^{\mathrm{L}}(n;a)$ so that $C^{\mathrm{L}}_{n,m} > a$. By the previous lemma, we find that $C_{n,m}\geq a'$. Hence $\Theta(n;a') \leq m = \Theta^{\mathrm{L}}(n;a)$, as required. □

As seen the polynomial example of §2.4, one typically finds that $\Theta(n;a) = c(a)n^{\alpha}$ for some $\alpha\geq 1$. Therefore, this theorem states that if $\Theta^{\mathrm{L}}(n;a) = \mathcal{O}(n^{\alpha})$, then $\Theta(n;a) = \mathcal{O}(n^{\alpha})$ also. In other words, the asymptotic growth of the stable sampling rate cannot be improved.

As shown, when sampling is carried out with an orthonormal basis, no projection-type method can improve generalised sampling. However, as evidenced by the previous theorem, there is room for improvement when sampling in a Riesz basis. In fact, it is straightforward to devise an 'optimal' projection-type method for such a problem:

Corollary 4.6. *Suppose that $\{\psi_j\}$ is a Riesz basis for* S*. Then the optimal projection-type method (in the sense of Theorems 4.3 and 4.5) is given by $\mathrm{L}_{n,m} = \mathcal{Q}_{\mathrm{S}_m}(\mathrm{T}_n)$.*

Proof. Let $\{\Psi_j\}^m_{j=1}$ be an orthonormal basis for S_m. Then the projection-type method with $\mathrm{L}_{n,m} = \mathcal{Q}_{\mathrm{S}_m}(\mathrm{T}_n)$ is equivalent to the generalised sampling scheme based on the orthonormal sampling vectors $\{\Psi_j\}^m_{j=1}$. Thus, by Theorem 4.3, no other projection-type method can outperform this method, including, of course, generalised sampling based on $\{\psi_j\}^m_{j=1}$. □

This importance of this corollary is that it indicates how to improve generalised sampling whenever the vectors $\{\psi_j\}$ are not orthonormal. Indeed, instead of specifying $\tilde{f}_{n,m}\in\mathrm{T}_n$ by (3.1), we instead define $\tilde{f}^o_{n,m}\in\mathrm{T}_n$ by

$$\langle \mathcal{Q}_{\mathrm{S}_m}\tilde{f}^o_{n,m},\phi_j\rangle = \langle \mathcal{Q}_{\mathrm{S}_m} f,\phi_j\rangle, \quad j=1,\ldots,n.$$

Suppose that $C^{[m]}\in\mathbb{C}^{m\times m}$ is the Gram matrix for the vectors $\{\psi_j\}^m_{j=1}$. Then this is equivalent to the system of equations

$$(U^{[n,m]})^*(C^{[m]})^{-1}U^{[n,m]}\alpha^{[n,m]} = (U^{[n,m]})^*(C^{[m]})^{-1}\hat{f}^{[m]},$$

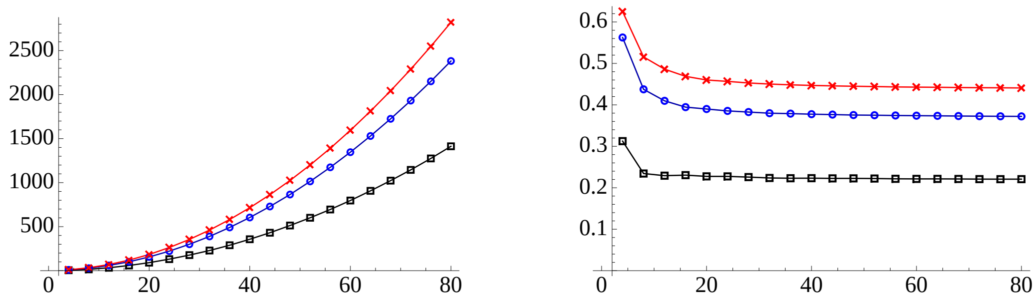

Figure 3: The quantities $\Theta(n; \frac{1}{2})$ (left) and $n^{-2}\Theta(n; \frac{1}{2})$ (right), using $\mathcal{P}_m$ (squares) and $\mathcal{P}_m^\sigma$, where σ is given by (4.4) (circles and crosses respectively).

for $\alpha^{[n,m]} \in \mathbb{C}^n$, where $\tilde{f}^o_{n,m} = \sum_{j=1}^n \alpha_j^{[n,m]} \phi_j$. Thus, at the extra expense of having to calculate $(C^{[m]})^{-1}$, this approach allows one to improve the constant of the stable sampling rate, although not the asymptotic growth. Note that, since $\{\psi_j\}_{j=1}^\infty$ is a Riesz basis, the Gram matrix $C^{[m]}$ is well-conditioned for all m.

Remark 1 There are several equivalent ways to look at Theorems 4.3 and 4.5.

1. One cannot replace the mapping $\mathcal{S}_m : \mathrm{ran}(\mathcal{S}_m^*) \to \mathrm{S}_m$ with another map $\mathcal{G}_{n,m} : \mathrm{ran}(\mathcal{S}_m^*) \to \mathrm{S}_m$ and obtain a better stable sampling rate.
2. One cannot precondition the matrix $A = U^*U$ via U^*PU with some matrix P and obtain a better stable sampling rate.

We shall discuss several consequences of this remark in the next section.

4.2 Numerical examples

To illustrate Theorem 4.3 let us consider the problem of reconstructing in a polynomial basis from Fourier samples. Since the samples are orthogonal in this case, we expect no projection-type method to offer a better stable sampling rate. This is shown in Figure 3, where we compare generalised sampling and two projection-type methods based on Fourier filters. Note that these methods are equivalent to replacing $\mathcal{P}_m$ by $\mathcal{P}_m^\sigma$ in the equations (3.1) defining $\tilde{f}_{n,m}$, where $\mathcal{P}_m^\sigma$ is the filtered Fourier operator

$$\mathcal{P}_m^\sigma = \sum_{|j| \leq \lfloor \frac{m}{2} \rfloor} \sigma_j \hat{f}_j \psi_j(x),$$

and σ is some filter. The two filters used were the Lanczos and the raised cosine filter [8, chpt. 2], given by

$$\sigma_j = \frac{m}{2j\pi} \sin\left(\frac{2j\pi}{m}\right), \qquad \sigma_j = \frac{1}{2}\left(1 + \cos\left(\frac{2j\pi}{m}\right)\right), \tag{4.4}$$

respectively. Figure 3 should come as no surprise. Whilst filters are extremely successful in improving pointwise accuracy of Fourier series [36], they do not decrease the error when measured in the L^2 norm $\|\cdot\|$, which is precisely the quantity that determines $C_{n,m}$.

An example of more interest is provided by replacing the standard Fourier series by a nonharmonic Fourier series [42]. In this case, instead of the standard exponentials $\{\mathrm{e}^{\mathrm{i}j\pi x}\}_{j\in\mathbb{Z}}$ we consider $\{\mathrm{e}^{\mathrm{i}\lambda_j \pi x}\}_{j\in\mathbb{Z}}$. For suitably chosen parameters $\lambda_j \in \mathbb{R}$, the latter constitutes a Riesz basis [42].

In Figure 4 we consider the stable sampling rate resulting from reconstructing in a polynomial basis whilst sampling with the nonharmonic Fourier basis corresponding to $\lambda_j = j + \frac{1}{3}(-1)^j$. We consider both generalised sampling in its standard form, as well as the optimal projection-type method described in Corollary 4.6. As predicted by this result, the optimal projection-type method offers a slightly reduced stable sampling rate, yet the asymptotic growth remains $\mathcal{O}\left(n^2\right)$. Thus, at the extra expense of having to invert an $m \times m$ Gram matrix, one can slightly reduce the number of samples required to reconstruct to within a prescribed accuracy.

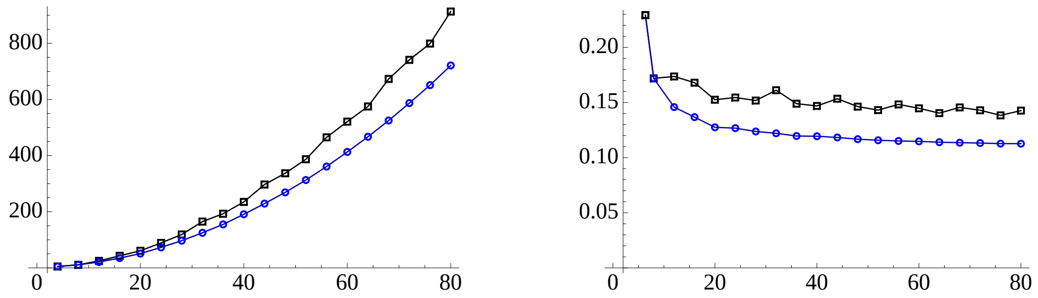

Figure 4: The quantities $\Theta(n; \frac{1}{2})$ (left) and $n^{-2}\Theta(n; \frac{1}{2})$ (right) for generalised sampling (squares) and the optimal projection-type method (circles).

4.3 Polynomial reconstructions from Fourier samples

The optimality results of the previous section are limited in that they only encompass one type of reconstruction method. Whether a broader, more powerful result holds – equivalently, whether the stable sampling rate is universal or not – is a matter for future research. However, in the important case of reconstructing a function in a polynomial basis from its Fourier samples, there is substantial evidence to conjecture that generalised sampling is optimal amongst all methods (subject to some minor restrictions). In other words, the stable sampling rate is universal for this particular pair of sampling and reconstruction vectors.

This argument is based on previous work by Platte et al [32]. Therein the authors consider the problem of recovering an analytic function $f : [-1, 1] \to \mathbb{R}$ from its values at m equispaced nodes in $[-1, 1]$. The principal result proved is as follows:

Theorem 4.7 ([32])**.** *Let a compact set $E \subseteq \mathbb{C}$ containing $[-1, 1]$ in its interior be fixed, and suppose that $\{\Phi_m\}$ is an approximation procedure based on equispaced m-grids such that for some $M < \infty$, $\sigma > 1$ and $\tau \in (\frac{1}{2}, 1]$,*

$$\|f - \Phi_m(f)\|_{\mathrm{L}^\infty[-1,1]} \leq M\sigma^{-m^\tau} \|f\|_{\mathrm{L}^\infty(E)}, \quad n = 1, 2, \ldots, \tag{4.5}$$

for all functions f that are continuous on E and analytic in its interior. Then the condition number κ_m of Φ_m satisfies

$$\kappa_m \geq c^{m^{2\tau-1}},$$

for some $c > 1$ and all sufficiently large m.

Here the condition number κ_m of Φ_m is defined in a particular way (see [32, eqn. (1.1)]) which reduces to the usual condition number of a matrix when Φ_m is linear. However, this theorem also holds for nonlinear methods, the only stipulations being that, for all functions f, $\Phi_m(f)$ depends only on the grid values of f and satisfies (4.5). This aside, note the main conclusion of this theorem: namely, any stable method can converge at best root-exponentially fast in m.

This key step in proving this theorem is to determine the nature of the constant

$$D_{n,m} = \sup_{\substack{p \in \mathbb{P}_n \\ p \neq 0}} \left\{ \frac{\|p\|_{\mathrm{L}^\infty[-1,1]}}{\|p\|_m} \right\}, \tag{4.6}$$

where $\|p\|_m$ is the maximum of $|p(x)|$ on the equispaced grid. This is a classical problem in approximation theory: how large can a polynomial of degree n be, given that it is bounded at m equispaced nodes in $[-1, 1]$? As shown by Coppersmith and Rivlin [10], there exist positive constants c_1 and c_2 such that

$$(c_1)^{\frac{n^2}{m}} \leq D_{n,m} \leq (c_2)^{\frac{n^2}{m}}. \tag{4.7}$$

In particular, for boundedness of $D_{n,m}$, we require $m = \mathcal{O}\left(n^2\right)$ (this result actually dates back to Schönhage [35], for more recent work on this problem see [33]).

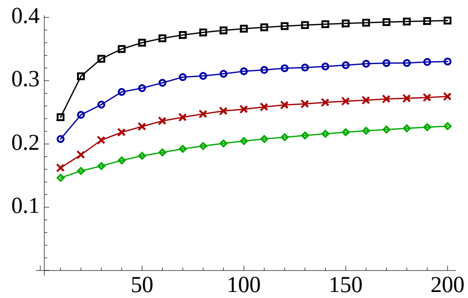

Figure 5: The quantity $-n^{\alpha-2}\log C_{n,n^\alpha}$ against $n = 1, 2, \ldots, 200$, where $\alpha = 1, \frac{5}{4}, \frac{4}{3}, \frac{3}{2}$ (squares, circles, crosses and diamonds respectively).

As discussed in §2.4, the problem of recovering a function from its values of an equispaced grid is very similar to that of reconstructing a function from its Fourier coefficients. It seems highly likely that an analogue of Theorem 4.7 ought to hold for this problem as well. To this end, suppose that $\| \cdot \|_m$ is the semi-norm defined by $\|f\|_m = \max_{j=-\lfloor \frac{m}{2} \rfloor, \ldots, \lfloor \frac{m}{2} \rfloor - 1} |\hat{f}_j|$, and let

$$\tilde{D}_{n,m} = \sup_{\substack{p \in \mathbb{P}_n \\ p \neq 0}} \left\{ \frac{\|p\|_{\mathrm{L}^\infty[-1,1]}}{\|p\|_m} \right\}. \tag{4.8}$$

Note that similarity between $\tilde{D}_{n,m}$ and the constant $D_{n,m}$ (as defined in (4.6)). Critically, the latter involves the semi-norm $\|p\|_m$ consisting of grid values, whereas the former consists of Fourier coefficients.

At this point, we make the following conjecture:

Conjecture 1. *If $\tilde{D}_{n,m}$ is as in (4.8), then there exist positive constants c_1, c_2 such that*

$$(c_1)^{\frac{n^2}{m}} \leq \tilde{D}_{n,m} \leq (c_2)^{\frac{n^2}{m}}, \quad \forall n, m \in \mathbb{N}. \tag{4.9}$$

Provided this conjecture holds, we have the following theorem:

Theorem 4.8. *Consider the problem of recovering a function $f : [-1, 1] \to \mathbb{R}$ from its Fourier coefficients $\{\hat{f}_j\}_{j \in \mathbb{Z}}$. Under the assumption that Conjecture 1 is true, then an anologue of Theorem 4.7 holds for this problem. In particular, generalised sampling is an optimal stable method in this instance.*

On the assumption that the Conjecture 1 holds, this theorem confirms optimality of generalised sampling for this problem (the proof is virtually identical to that of Theorem 4.7). Note the similarity of the quantity $D_{n,m}$ and $(C_{n,m})^{-1}$, where $C_{n,m}$ is the constant of generalised sampling. In fact, it is a simple exercise to show that

$$\frac{1}{n\sqrt{m}} D_{n,m} \leq (C_{n,m})^{-1} \leq \sqrt{2} D_{n,m}.$$

In particular, boundedness of $D_{n,m}$ implies boundedness of $(C_{n,m})^{-1}$, and therefore stability of generalised sampling.

At the moment, we have no proof of Conjecture 1. However, there is substantial numerical evidence. In Figure 5 we have plotted the quantity $(C_{n,n^\alpha})^{-1}$ for various values of $1 \leq \alpha \leq 2$. The straight lines in the diagrams indicate a bound of the form (4.9).

5 Conclusions

The purpose of this paper was to derive new estimates for generalised sampling, and to establish several results concerning its optimality. The latter topic warrants further study. In particular, whilst the results proved in Theorems 4.3 and 4.5 are interesting, they place rather stringent restrictions on the types of methods considered. An aim of future work is to prove a stronger result. Additionally, the example used throughout this paper – reconstructions in polynomials from Fourier samples – is of particular interest, due to its application in the resolution of the Gibbs phenomenon. A topic of ongoing work is a proof of Conjecture 1, and as a consequence, the removal of this as a condition of Theorem 4.8.

As mentioned, optimality of generalised sampling is equivalent to the question of universality of the stable sampling rate. This has important ramifications for the design of reconstruction methods in applications. Namely, given some sampling scheme and some desired reconstruction space, universality of the stable sampling rate places concrete limitations on the achievable accuracy of the reconstruction for a given number of samples. We shall explore this issue in more detail in future investigations.

On the other hand, although Theorems 4.3 and 4.5 demonstrate that no projection-type method can improve the stable sampling rate, there is a potential alternative based on a slight generalisation of projection-type class. This approach, whilst possibly not improving the asymptotic growth of $\Theta(n; a)$ (i.e. if universality holds), may potentially reduce the constant c of such growth. As commented in Remark 1, these theorems essentially preclude any improvement by replacing the mapping $\mathcal{S}_m : \mathrm{ran}(\mathcal{S}_m^*) \rightarrow \mathrm{S}_m$ with another map $\mathcal{G}_{n,m}$ with the same range and domain. However, although it is necessary that any alternative map $\mathcal{G}_{n,m}$ have the same domain (due to the nature of the samples), the restriction that $\mathcal{G}_{n,m}$ has the same range S_m as $\mathcal{S}_m$ is clearly not necessary. If we allow for mappings $\mathcal{G}_{n,m} : \mathrm{ran}(\mathcal{S}_m^*) \rightarrow \mathrm{H}$ with arbitrary ranges then it may be possible to decrease c. Note that, roughly speaking, the size of c is determined by the rate of convergence $\mathcal{P}_m \rightarrow \mathcal{I}$ uniformly on the reconstruction space T_n. Replacing $\mathcal{S}_m$ by $\mathcal{G}_{n,m}$ equates to replacing $\mathcal{P}_m = \mathcal{S}_m \circ \mathcal{S}_m^*$ by a mapping $\mathcal{P}'_{n,m} = \mathcal{G}_{n,m} \circ \mathcal{S}_m^*$. Thus, if a suitable sequence of mappings $\mathcal{P}'_{n,m}$ possessing faster convergence uniformly on T_n could be found, then it may be possible to decrease c. The specification of any such mappings would obviously depend on both the type of sampling and the reconstruction space T_n.

This aside, a central premise of generalised sampling is *guaranteed* recovery for all vectors f. Indeed, the stable sampling rate provides the exact condition for this property, and also, as a by-product, enforces numerical stability. However, suppose now that one was willing to compromise numerical stability, and consider the problem of recovering just one vector f as accurately as possible. For a given number of samples m is there a better way to choose n than that specified by the stable sampling rate? In particular, could one obtain a more accurate reconstruction with a larger (and typically vector-dependent) choice of n, and how could such a choice be made using only the given data (i.e. the samples of f)? Naturally, this question involves the delicate balance between the growth of $(C_{n,m})^{-1}$ and the decay of $\|f - \mathcal{Q}_n f\|$ in the error estimate

$$\|f - \mathcal{Q}_n f\| \leq \frac{1}{C_{n,m}} \|f - \mathcal{Q}_n f\|,$$

(see Theorem 3.8). In preliminary numerical experiments at least, it appears that choosing the best $n = n(f)$ for a given f (and m) can lead to substantial improvements in the accuracy of the reconstruction. However, the question of how to determine $n(f)$ for an unknown vector f is currently unanswered. An adaptive algorithm may be useful in this regard, as may be the use of nonlinear techniques.

On a different topic, one conclusion of this paper is that consistency is the incorrect condition to enforce for numerical reconstructions. However there are applications, most notably image resizing, for which consistency of the reconstruction with the original measurements is important. Thus, the question is, can one design a reconstruction method possessing both consistency and numerical stability? Given the results of this paper, it seems probable that this cannot be done without sacrificing numerical accuracy, at least with a linear method. However, in such applications, this may be acceptable. Moreover, the use of nonlinear techniques (e.g. l^1 minimisation) may potentially allow simultaneously for consistency, stability and accuracy. This is a topic for future research.

Consistent reconstructions are extremely popular in signal and imaging applications. There are some notable circumstances, in particular, when operating in certain shift-invariant or periodic spaces, for which such reconstructions do not suffer from any stability or convergence issues. However, there are also many examples, outside of the polynomial case of this paper, for which these issues do arise. These include reconstructions in both splines and wavelets (see [1] for the case of the latter). In a future paper we intend to present a longer list and detailed discussion of such important examples. That aside, let us note that the analysis of this paper actually has important consequence for consistent reconstructions themselves. Namely, it answers the question of when they do and do not work, and shows how this answer can be determined conclusively by a simple series of numerical computations.

Acknowledgements

The authors would like to thank Hagai Kirschner, Nilima Nigam, Alexei Shadrin and Michael Unser for their helpful discussions and comments.

References

[1] B. Adcock and A. C. Hansen. A generalized sampling theorem for stable reconstructions in arbitrary bases. *Technical report NA2010/07, DAMTP, University of Cambridge*, 2010.

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

[3] 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.

[4] B. Adcock and A. C. Hansen. Stable reconstructions in Hilbert spaces and the resolution of the Gibbs phenomenon. *Appl. Comput. Harmon. Anal. (accepted)*, 2011.

[5] A. Aldroubi. Oblique projections in atomic spaces. *Proc. Amer. Math. Soc.*, 124(7):2051–2060, 1996.

[6] T. Blu, P. L. Dragotti, M. Vetterli, P. Marziliano, and L. Coulout. Sparse sampling of signal innovations. *IEEE Signal Process. Mag.*, 25(2):31–40, 2008.

[7] A. Böttcher. Infinite matrices and projection methods. In *Lectures on operator theory and its applications (Waterloo, ON, 1994)*, volume 3 of *Fields Inst. Monogr.*, pages 1–72. Amer. Math. Soc., Providence, RI, 1996.

[8] C. Canuto, M. Y. Hussaini, A. Quarteroni, and T. A. Zang. *Spectral methods: Fundamentals in Single Domains*. Springer, 2006.

[9] O. Christensen. *An Introduction to Frames and Riesz Bases*. Birkhauser, 2003.

[10] D. Coppersmith and T. Rivlin. The growth of polynomials bounded at equally spaced points. *SIAM J. Math. Anal.*, 23:970–983, 1992.

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

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

[13] Y. C. Eldar. Sampling with arbitrary sampling and reconstruction spaces and oblique dual frame vectors. *Journal of Fourier Analysis and Applications*, 9(1):77–96, 2003.

[14] Y. C. Eldar. Sampling without input constraints: Consistent reconstruction in arbitrary spaces. *Sampling, Wavelets and Tomography*, 2003.

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

[16] Y. C. Eldar and T. Michaeli. Beyond Bandlimited Sampling. *IEEE Signal Process. Mag.*, 26(3):48–68, 2009.

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

[18] J. F. Epperson. On the Runge example. *Am. Math. Mon.*, 94:329–341, 1987.

[19] B. Fornberg. *A Practical Guide to Pseudospectral Methods*. Cambridge University Press, 1996.

[20] D. Gottlieb and J. S. Hesthaven. Spectral methods for hyperbolic problems. *J. Comput. Appl. Math.*, 128(1-2):83–131, 2001.

[21] D. Gottlieb and C.-W. Shu. On the Gibbs' phenomenon and its resolution. *SIAM Rev.*, 39(4):644–668, 1997.

[22] K. Gröchenig, Z. Rzeszotnik, and T. Strohmer. Quantitative estimates for the finite section method. *Integral Equations Operator Theory*, to appear.

[23] R. Hagen, S. Roch, and B. Silbermann. *C^*-Algebras and Numerical Analysis*, volume 236 of *Monographs and Textbooks in Pure and Applied Mathematics*. Marcel Dekker Inc., New York, 2001.

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

[25] A. C. Hansen. On the approximation of spectra of linear operators on Hilbert spaces. *J. Funct. Anal.*, 254(8):2092–2126, 2008.

[26] E. Heinemeyer, M. Lindner, and R. Potthast. Convergence and numerics of a multisection method for scattering by three-dimensional rough surfaces. *SIAM J. Numer. Anal.*, 46(4):1780–1798, 2008.

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

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

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

[30] M. Lindner. *Infinite Matrices and their Finite Sections*. Frontiers in Mathematics. Birkhäuser Verlag, Basel, 2006. An introduction to the limit operator method.

[31] J. A. Parker, R. V. Kenyon, and D. E. Troxel. Comparison of interpolating methods for image resampling. *IEEE Trans. Med. Imaging*, MI-2(1):31–39, 1983.

[32] R. Platte, L. N. Trefethen, and A. Kuijlaars. Impossibility of fast stable approximation of analytic functions from equispaced samples. *SIAM Rev. (to appear)*, 2010.

[33] E. A. Rakhmanov. Bounds for polynomials with a unit discrete norm. *Ann. Math.*, 165:55–88, 2007.

[34] T. J. Rivlin. *Chebyshev Polynomials: from Approximation Theory to Algebra and Number Theory*. Wiley New York, 1990.

[35] A. Schönhage. Fehlerfort pflantzung bei Interpolation. *Numer. Math.*, 3:62–71, 1961.

[36] E. Tadmor. Filters, mollifiers and the computation of the Gibbs' phenomenon. *Acta Numerica*, 16:305–378, 2007.

[37] W.-S. Tang. Oblique projections, biorthogonal Riesz bases and multiwavelets in Hilbert spaces. *Proc. Amer. Math. Soc.*, 128(2):463–473, 1999.

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

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

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

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

[42] R. M. Young. *An Introduction to Nonharmonic Fourier Series*. Academic Press Inc., San Diego, CA, first edition, 2001.