

Beyond consistent reconstructions: optimality and sharp bounds for generalized sampling, and application to the uniform resampling problem

Ben Adcock
Department of Mathematics
Purdue University
150 N. University Street
West Lafayette, IN 47907
USA

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

Clarice Poon
DAMTP, Centre for Mathematical Sciences
University of Cambridge
Wilberforce Rd, Cambridge CB3 0WA
United Kingdom

July 23, 2013

Abstract

Generalized sampling is a recently developed linear framework for sampling and reconstruction in separable Hilbert spaces. It allows one to recover any element in any finite-dimensional subspace given finitely many of its samples with respect to an arbitrary basis or frame. Unlike more common approaches for this problem, such as the consistent reconstruction technique of Eldar et al., it leads to numerical methods possessing both guaranteed stability and accuracy.

The purpose of this paper is twofold. First, we give a complete and formal analysis of generalized sampling, the main result of which being the derivation of new, sharp bounds for the accuracy and stability of this approach. Such bounds improve upon those given previously, and result in a necessary and sufficient condition, the stable sampling rate, which guarantees *a priori* a good reconstruction. Second, we address the topic of optimality. Under some assumptions, we show that generalized sampling is an optimal, stable method. Correspondingly, whenever these assumptions hold, the stable sampling rate is a universal quantity. In the final part of the paper we illustrate our results by applying generalized sampling to the so-called uniform resampling problem.

1 Introduction

A central theme in sampling theory is the recovery of a signal or an image from a collection of its measurements. Mathematically, this can be modelled in a separable Hilbert space H , with the samples

of the unknown signal $f \in H$ being of the form

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

where $\{\psi_j\}_{j=1}^\infty$ is a collection of elements in H and $\langle \cdot, \cdot \rangle$ is the inner product on H . Typically, the *sampling* system $\{\psi_j\}_{j=1}^\infty$ forms a frame for its span $S = \overline{\text{span}\{\psi_1, \psi_2, \dots\}}$.

One of the most common examples of this problem is the recovery of a function f with compact support from pointwise evaluations of its Fourier transform \hat{f} . In this case, $H = L^2(-1, 1)^d$, where $\text{supp}(f) \subseteq [-1, 1]^d$ without loss of generality, and $\psi_j(x) = e^{i\pi\omega_j \cdot x}$ for suitable values $\{\omega_j\}_{j \in \mathbb{N}} \subseteq \mathbb{R}^d$. This is precisely the type of sampling encountered in Magnetic Resonance Imaging (MRI), for example.

If the measurements $\{\omega_j\}_{j \in \mathbb{N}}$ are taken uniformly, both f and \hat{f} can be recovered via the Shannon Sampling Theorem [42, 55]. However, the slow convergence of the corresponding reconstructions (both infinite sums), as well as the appearance of the Gibbs phenomenon, means that this approach is often impractical [28, 47, 55]. In cases where measurements $\{\omega_j\}_{j \in \mathbb{N}}$ are not uniformly distributed, no simple reconstruction need exist. It is standard in this setting to use a gridding algorithm [40, 46, 49, 61]. However, this also typically leads to less than satisfactory accuracy [61].

The MRI problem serves to illustrate several key issues that are critical to this paper. First, although f is sampled via an infinite collection of elements $\{\psi_j\}_{j=1}^\infty$, in practice we only have access to a finite number. Thus, the problem we consider throughout this paper is that of recovering f from only its *first* n samples $\hat{f}_1, \dots, \hat{f}_n$. Key issues herein are those of *approximation* – namely, how well f can be recovered as $n \rightarrow \infty$ – and *robustness*; does increasing n lead to worse stability, and thus more sensitivity to noise and round-off error?

The MRI example also highlights another important consideration. Namely, the samples $\{\hat{f}_j\}_{j \in \mathbb{N}}$ of f are fixed, and cannot easily be altered. This situation occurs typically when the sampling scheme is specified by some physical device, e.g. the MR scanner in the above example. Although it is actually possible in some cases to modify MR scanners to acquire different types of measurements, such as in so-called *wavelet-encoded* MRI [36, 62], this is not without complications [44]. Thus, the question we consider in this paper is the following: given a finite number of *fixed* samples of an element f of a Hilbert space H , how can one obtain a good (i.e. accurate and robust) reconstruction?

This question is not new, and there has been much interest in the last several decades in alternative reconstructions to those given by the Shannon Sampling Theorem. This is typically based on the following principle: many signals that arise in practice can be much better represented in terms of a different collection of elements $\{\phi_j\}_{j \in \mathbb{N}} \subseteq H$ than by Shannon’s theorem [28, 55]. Common examples of such systems include wavelets, splines and polynomials, as well as more exotic objects such as curvelets [12, 14], shearlets [18, 19, 43] and contourlets [20, 48]. Thus, given this additional knowledge about f , the problem can now be restated as follows: how can we compute a reconstruction in the system $\{\phi_j\}_{j \in \mathbb{N}}$ from the measurements $\hat{f}_j = \langle f, \psi_j \rangle$?

Consistent sampling is a linear method for this problem, based on stipulating that the reconstruction of f agrees with the available measurements. Introduced by Unser & Aldroubi [56, 58] and later generalized significantly by Eldar et al. [23, 24, 25, 29], this technique has proved successful in a number of areas, and is quite widely used in practice [55]. However, there are several drawbacks. As discussed in [4, 3, 26, 38], consistent reconstructions need not be stable nor convergent as the number of measurements increases. Whilst stability and convergence can be guaranteed in certain shift-invariant spaces [25, 57], it is quite easy to devise examples outside of this setting for which consistent reconstructions either fail to converge, or are extremely unstable, or both [4].

Fortunately, it transpires that these issues can be overcome by using an recently-introduced alternative technique, known as *generalized sampling* [3, 4, 5]. This method forms the primary focus of our paper. Our main results are described in the next section.

1.1 Novelty of the paper and overview

The aim of this paper is to give a complete analysis of generalized sampling. First, we establish new, sharp bounds for its error and stability, which improve on those appearing previously in [3, 4, 5]. Second,

we provide several optimality results, which demonstrate that under appropriate conditions generalized sampling cannot be outperformed. To illustrate our results we consider the use of generalized sampling in so-called *uniform resampling* problem.

Let us now give an overview of the remainder of the paper. In §2–3 we provide a mathematical description of the general instability and nonconvergence of consistent reconstructions. For this, it is first necessary to present a formal description of the reconstruction problem (§2). Herein we also introduce two key constants to assess different methods: the *condition number* κ and the *quasi-optimality constant* μ . The former measures the sensitivity of a given reconstruction method to perturbations (e.g. noise and round-off), whilst the latter quantifies how close the reconstruction of an element f is to its best (i.e. optimal) reconstruction in the desired system of functions. For succinctness, we also introduce the reconstruction constant C of a method, defined as the maximum of κ and μ .

The focus of §3 is consistent sampling. By analyzing the reconstruction constant C in this instance, we provide a comprehensive answer as to when this approach will give poor (i.e. unstable or inaccurate) reconstructions. Moreover, we show how one can determine *a priori* an answer to this question by straightforward numerical computations. Finally, we discuss how this question can be reinterpreted in terms of the behaviour *finite sections* of infinite operators, and thus draw a connection between problems in sampling and computational spectral theory (such a connection was previously discussed in [4]. See also [32] for related work).

Generalized sampling was introduced in [3, 4, 5] to overcome the issues inherent to consistent reconstructions. In the second part of this paper (§4–5) we improve the previous analysis of [4, 5] by using the formal framework developed in §2. We explain conclusively how generalized sampling guarantees a stable and accurate reconstruction by deriving the exact values for μ and κ , and therefore C , as opposed to the nonsharp bounds given previously in [4, 5]. Moreover, we reinterpret generalized sampling using geometry of Hilbert spaces, and specifically, the notions of oblique projections and subspace angles. Next, we introduce a necessary and sufficient condition, the *stable sampling rate*, as well as its reciprocal, the *stable reconstruction rate*, which determines how to select the generalized sampling parameters so as to guarantee a good reconstruction. This improves on the previous sufficient conditions of [4, 5]. Once more, this condition is computable, as explained in §5. We also discuss the connections between generalized sampling and sections of infinite operators.

In §6 we consider optimality of generalized sampling. That is, we pose the following question: can another method outperform generalized sampling, and if so, in what sense? Using the sharp bounds derived in §4–5, we show that no method which is *perfect* (a definition is given later) can exhibit better stability than generalized sampling. Hence generalized sampling is an optimal, stable approach to the reconstruction problem amongst the class of perfect methods. Moreover, for problems where the stable sampling rate grows linearly, we show that no method (perfect or nonperfect) can outperform generalized sampling in terms of the reconstruction accuracy by more than a constant factor. Thus, although it is possible in theory to get a better approximation error with a different method, no method can converge at an asymptotically faster rate than generalized sampling.

In the final part of this paper, §7, we consider the application of generalized sampling to the so-called *uniform resampling problem*. This problem concerns the computation of the Fourier coefficients of a function from nonuniformly-spaced samples of its Fourier transform. We show that the standard approach to this problem is merely an instance of consistent sampling, and we explain how in general this will lead to an exponentially large reconstruction constant C . Next we consider the application of generalized sampling to this problem. We prove that the stable sampling rate is linear, and therefore generalized sampling is, in the senses defined in §6, an optimal, stable method for this problem. Finally, we consider alternatives to uniform resampling, and show how the incorporation of different reconstruction systems – splines or polynomials, for example – can lead to an improved reconstruction.

1.2 Relation to sparsity and compressed sensing

One of the most significant developments in signal and image reconstruction in the last several decades has been the introduction of sparsity-exploiting algorithms. Techniques such as compressed sensing

[13, 21, 27, 30], which use sparsity of the signal f in a particular basis (wavelets, for example) to reduce the number of measurements required, have recently become extremely popular.

Generalized sampling, in the form we discuss in this paper, does not exploit sparsity. It guarantees recovery of all signals, sparse or otherwise, from sufficiently many measurements. However, it transpires that generalized sampling can be combined with existing compressed sensing tools (randomization and convex optimization) to achieve subsampling, whenever the signal f is sparse (or compressible) in the basis $\{\phi_j\}_{j \in \mathbb{N}}$. See [1]. The importance of this development is that it permits compressed sensing of analog signals (i.e. functions in function spaces) which have sparse (or compressible) information content in some countable basis. Conversely, most compressed sensing techniques and theorems apply only to finite-dimensional signals, i.e. vectors in finite-dimensional vector spaces. For related research on sub-Nyquist sampling for analog signals, see the works of Vetterli et al. on finite rates of innovation [9, 22, 60] and, more recently, Candès & Fernandez-Granda on super-resolution [15].

2 The reconstruction problem

We now describe the reconstruction problem in more detail. To this end, suppose that $\{\psi_j\}_{j \in \mathbb{N}}$ is a collection of elements of a separable Hilbert space H (over \mathbb{C}) that forms a frame for a closed subspace S of H (the *sampling* space). In other words, $\text{span}\{\psi_j : j \in \mathbb{N}\}$ is dense in S and there exist constants $c_1, c_2 > 0$ (the frame constants) such that

$$c_1 \|f\|^2 \leq \sum_{j \in \mathbb{N}} |\langle f, \psi_j \rangle|^2 \leq c_2 \|f\|^2, \quad \forall f \in S, \quad (2.1)$$

where $\langle \cdot, \cdot \rangle$ and $\|\cdot\|$ are the inner product and norm on H respectively [17]. Suppose further that $\{\phi_j\}_{j \in \mathbb{N}}$ is a collection of *reconstruction* elements that form a frame for a closed subspace T (the *reconstruction* space), with frame constants $d_1, d_2 > 0$:

$$d_1 \|f\|^2 \leq \sum_{j \in \mathbb{N}} |\langle f, \phi_j \rangle|^2 \leq d_2 \|f\|^2, \quad \forall f \in T. \quad (2.2)$$

Let $f \in H$ be a given element we wish to recover, and assume that we have access to the samples

$$\hat{f}_j = \langle f, \psi_j \rangle, \quad j \in \mathbb{N}. \quad (2.3)$$

Note that the infinite vector $\hat{f} = \{\hat{f}_j\}_{j \in \mathbb{N}} \in \ell^2(\mathbb{N})$. Ignoring for the moment the issue of truncation – namely, that in practice we only have access to the first n measurements – the reconstruction problem can now be stated as follows: given $\hat{f} = \{\hat{f}_j\}_{j \in \mathbb{N}}$, find a reconstruction \tilde{f} of f from the subspace T .

2.1 Stability and quasi-optimality

There are two important conditions which a reconstruction, i.e. a mapping $\{\hat{f}_j\}_{j \in \mathbb{N}} \mapsto \tilde{f} \in T$, ought to possess. The first is so-called *quasi-optimality*:

Definition 2.1. *Let F be an operator on H_0 , where H_0 is a closed subspace of H , with range contained in T . The quasi-optimality constant of $\mu = \mu(F) > 0$ is the least number such that*

$$\|f - F(f)\| \leq \mu \|f - \mathcal{Q}_T f\|, \quad \forall f \in H_0,$$

where $\mathcal{Q}_T : H \rightarrow T$ is the orthogonal projection onto T . If no such constant exists, we write $\mu = \infty$. We say that F is quasi-optimal if $\mu(F)$ is small.

Note that $\mathcal{Q}_T f$ is the best, i.e. energy-minimizing, approximation to f from T . Thus, quasi-optimality states that the error committed by \tilde{f} is within a small and constant factor of that of the best approximation. The desire for quasi-optimality arises from the fact that typical images and signals are

known to be well represented in certain bases and frames, e.g. wavelets, splines or polynomials [55]. In other words, the error $\|f - \mathcal{Q}_T f\|$ is small. When reconstructing f in the corresponding subspace T from its measurements $\{\hat{f}_j\}_{j \in \mathbb{N}}$ it is therefore vital that $\mu \ll \infty$. Otherwise, the beneficial property of T for the signal f may be lost when computing the reconstruction \tilde{f} .

The second important consideration is that of stability. For this, we introduce a condition number (see, for example, [54]):

Definition 2.2. *Let H_0 be a closed subspace of H and suppose that $F : H_0 \rightarrow H$ is a mapping such that, for each $f \in H_0$, $F(f)$ depends only on the vector of samples $\hat{f} \in \ell^2(\mathbb{N})$. The (absolute) condition number $\kappa = \kappa(F)$ is given by*

$$\kappa = \sup_{f \in H_0} \lim_{\epsilon \rightarrow 0^+} \sup_{\substack{g \in H_0 \\ 0 < \|\hat{g}\|_{\ell^2} \leq \epsilon}} \left\{ \frac{\|F(f+g) - F(f)\|}{\|\hat{g}\|_{\ell^2}} \right\}. \quad (2.4)$$

We say that F is well-conditioned if κ is small. Otherwise it is ill-conditioned.

A well-conditioned mapping F is robust towards perturbations such as noise, and therefore this property is vital from a practical perspective.

It is worth noting at this stage that the condition number (2.4) does not assume linearity of F . If this is the case, then one has the much simpler form

$$\kappa(F) = \sup_{\substack{f \in H_0 \\ f \neq 0}} \left\{ \frac{\|F(f)\|}{\|f\|} \right\}.$$

We also remark that (2.4) is the *absolute* condition number, as opposed to the somewhat more standard *relative* condition number [54]. This is primarily for simplicity in the presentation: under some assumptions, it is possible to adapt the results we prove later in this paper for the latter.

Occasionally, we will also consider the absolute condition number at an element $f \in H_0$:

$$\kappa_f(F) = \lim_{\epsilon \rightarrow 0^+} \sup_{\substack{g \in H_0 \\ 0 < \|\hat{g}\|_{\ell^2} \leq \epsilon}} \left\{ \frac{\|F(f+g) - F(f)\|}{\|\hat{g}\|_{\ell^2}} \right\}, \quad f \in H_0. \quad (2.5)$$

This measures the local conditioning of F around f . Naturally, one has $\kappa(F) = \sup_{f \in H_0} \kappa_f(F)$.

For convenience, it is useful to introduce the notion of a reconstruction constant for F :

Definition 2.3. *Let F be as in Definition 2.2, and let $\mu(F)$ and $\kappa(F)$ be its quasi-optimality constant and condition number respectively. The reconstruction constant $C = C(F)$ is defined by $C(F) = \max\{\kappa(F), \mu(F)\}$. If F is not quasi-optimal or if $\kappa(F)$ is not defined, then we set $C(F) = \infty$.*

2.2 The computational reconstruction problem

As mentioned, in practice we do not have access to the infinite vector of samples \hat{f} . Thus, the computational reconstruction problem concerns the recovery of f from its first n measurements $\hat{f}_1, \dots, \hat{f}_n$. Since we only have access to these samples, it is natural to consider finite-dimensional subspaces of T . Thus, we let $\{T_n\}_{n \in \mathbb{N}}$ be a sequence of finite-dimensional subspaces satisfying

$$T_n \subseteq T, \quad \dim(T_n) < \infty, \quad (2.6)$$

and

$$\mathcal{Q}_{T_n} \rightarrow \mathcal{Q}_T, \quad n \rightarrow \infty, \quad (2.7)$$

strongly on H . In other words, the spaces $\{T_n\}_{n \in \mathbb{N}}$ form a sequence of finite-dimensional approximations to T . Strictly speaking, the second condition is not necessary. However, it is natural to make this assumption in order to guarantee a convergent approximation.

With this in hand, the computational reconstruction problem is now as follows: given the samples $\hat{f}_1, \dots, \hat{f}_n$, compute a reconstruction to f from the subspace T_n .

Remark 2.4 It is quite common in practice to define $T_n = \text{span}\{\phi_1, \dots, \phi_n\}$ to be the space spanned by the first n elements of an infinite frame $\{\phi_j\}_{j \in \mathbb{N}}$ for T . Note that (2.6) and (2.7) automatically hold in this case. Moreover, one has the nesting property $T_1 \subseteq T_2 \subseteq \dots$. However, this is not necessary. The reconstructions we consider in this paper are actually independent of the spanning system for T_n . Such a system only needs to be specified in order to perform computations. Hence, we consider the more general setting outlined above.

When considering methods, i.e. mappings F_n , for this problem, it is desirable that the constants reconstruction $C(F_n)$ should not grow rapidly with n . If this is not the case, then increasing the number of measurements could, for example, lead to a worse approximation and increased sensitivity to noise. We shall see examples of this in §3.6. To avoid this scenario, we now make the following definition:

Definition 2.5. For each $n \in \mathbb{N}$, let F_n be such that, for each f , $F_n(f)$ belongs to a finite-dimensional reconstruction space T_n and depends only on the samples $\hat{f}^{[n]} = \{\hat{f}_1, \dots, \hat{f}_n\}$. We say that the reconstruction scheme $\{F_n\}_{n \in \mathbb{N}}$ is numerically stable and quasi-optimal if

$$C^* := \sup_{n \in \mathbb{N}} C(F_n) < \infty,$$

where $C(F_n)$ is the reconstruction constant of F_n . We refer to the constant C^* as the reconstruction constant of the reconstruction scheme $\{F_n\}_{n \in \mathbb{N}}$.

This definition incorporates the issue of *stable approximation* into a sequence of reconstruction schemes. Although in practice one only has access to a finite number of samples, it is natural to consider the behaviour of F_n as n – the number of samples – increases. Ideally we want $F_n(f)$ to behave like $\mathcal{Q}_n f$, the best approximation to f from T_n . Namely, $F_n(f)$ should converge to f at the same rate as $\mathcal{Q}_n f$. This is important from a practical standpoint. The premise for computing a reconstruction in T_n is the knowledge that f is well represented in terms of the reconstruction system $\{\phi_j\}_{j \in \mathbb{N}}$, i.e. its orthogonal projections $\mathcal{Q}_n f$ converge rapidly. Hence it is vital that the computed reconstruction $F_n(f)$ does not possess dramatically different convergence behaviour. Put simply, there is little point reconstructing in the basis $\{\phi_j\}_{j \in \mathbb{N}}$ if the good approximation properties of f in this basis are destroyed by reconstruction technique.

Remark 2.6 In some applications, one may wish to relax the above definition slightly to allow mild growth of $C(F_n)$. If $\theta_n \in (0, \infty)$ is an increasing sequence, we say that $\{F_n\}_{n \in \mathbb{N}}$ is stable and quasi-optimal with respect to $\{\theta_n\}_{n \in \mathbb{N}}$ if

$$C^* = \limsup_{n \rightarrow \infty} \frac{C(F_n)}{\theta_n} < \infty.$$

In other words, $C(F_n)$ can grow at worst like $\mathcal{O}(\theta_n)$ as $n \rightarrow \infty$.

3 Consistent reconstructions and oblique projections

We now consider the consistent sampling technique of [24, 25, 29, 56, 58].

3.1 Consistent sampling

Let us first return to the problem of recovering f from its infinite vector of samples \hat{f} . A simple way to obtain a reconstruction with small constant in this instance is by solving the so-called *consistency* conditions. Specifically, we let $\tilde{f} \in T$ (whenever it exists uniquely) be the solution of

$$\langle \tilde{f}, \psi_j \rangle = \langle f, \psi_j \rangle, \quad j = 1, 2, \dots, \quad \tilde{f} \in T. \quad (3.1)$$

Note that consistency means that the samples of \tilde{f} agree with those of f . Correspondingly, we say that \tilde{f} is a *consistent reconstruction* of f , and refer to the operator $F : f \mapsto \tilde{f}$, whenever defined, as *consistent sampling*. In Lemma 3.8 we shall provide explicit conditions under which this is the case.

An analysis of consistent reconstructions, which we shall recap and extend in §3.3, was given in [24, 25, 29]. Crucial to this is the notion of oblique projections in Hilbert spaces, which we discuss next.

3.2 Oblique projections and subspace angles

We commence with the definition of a subspace angle:

Definition 3.1. *Let U and V be closed subspaces of a Hilbert space H and $\mathcal{Q}_V : H \rightarrow V$ the orthogonal projection onto V . The subspace angle $\theta = \theta_{UV} \in [0, \frac{\pi}{2}]$ between U and V is given by*

$$\cos(\theta_{UV}) = \inf_{\substack{u \in U \\ \|u\|=1}} \|\mathcal{Q}_V u\|. \quad (3.2)$$

Note that there are a number of different ways to define the angle between subspaces [51, 53]. However, (3.2) is the most convenient for this paper. We shall also make use of the following equivalent expression for $\cos(\theta_{UV})$:

$$\cos(\theta_{UV}) = \inf_{\substack{u \in U \\ \|u\|=1}} \sup_{\substack{v \in V \\ \|v\|=1}} |\langle u, v \rangle|. \quad (3.3)$$

We are interested in subspaces for which the cosine of the associated angle is nonzero. The following lemma is useful in this regard:

Lemma 3.2. *Let U and V be closed subspaces of a Hilbert space H . Then $\cos(\theta_{UV^\perp}) > 0$ if and only if $U \cap V = \{0\}$ and $U + V$ is closed in H .*

Proof. See [53, Thm. 2.1]. □

We now make the following definition:

Definition 3.3. *Let U and V be closed subspaces of a Hilbert space H . Then U and V satisfy the subspace condition if $\cos(\theta_{UV^\perp}) > 0$, or equivalently, if $U \cap V = \{0\}$ and $U + V$ is closed in H .*

Subspaces U and V satisfying this condition give a decomposition $U \oplus V = H_0$ of a closed subspace H_0 of H . Equivalently, this ensures the existence of a projection of H_0 with range U and kernel V . We refer to such a projection as an *oblique projection* and denote it by \mathcal{W}_{UV} . Note that \mathcal{W}_{UV} will not, in general, be defined over the whole of H . However, this is true whenever $V = U^\perp$, for example, and in this case \mathcal{W}_{UV} coincides with the orthogonal projection \mathcal{Q}_U .

We shall also require the following results on oblique projections (see [11, 52]):

Theorem 3.4. *Let U and V be closed subspaces of H with $U \oplus V = H$. Then*

$$\|\mathcal{W}_{UV}\| = \|\mathcal{I} - \mathcal{W}_{UV}\| = \sec(\theta_{UV^\perp}),$$

where $\|\cdot\|$ is the standard norm on the space of bounded operators on H .

Corollary 3.5. *Suppose that U and V are closed subspaces of H satisfying the subspace condition, and let $\mathcal{W}_{UV} : H_0 \rightarrow U$ be the oblique projection with range U and kernel V , where $H_0 = U \oplus V$. Then*

$$\|\mathcal{W}_{UV} f\| \leq \sec(\theta_{UV^\perp}) \|f\|, \quad \forall f \in H_0, \quad (3.4)$$

and if $\mathcal{Q}_U : H \rightarrow U$ is the orthogonal projection,

$$\|f - \mathcal{Q}_U f\| \leq \|f - \mathcal{W}_{UV} f\| \leq \sec(\theta_{UV^\perp}) \|f - \mathcal{Q}_U f\|, \quad \forall f \in H_0. \quad (3.5)$$

Moreover, the upper bounds in (3.4) and (3.5) are sharp.

Proof. The sharp bound (3.4) is due to Theorem 3.4. For (3.5) we first note that $(\mathcal{I} - \mathcal{W}_{UV}) = (\mathcal{I} - \mathcal{W}_{UV})(\mathcal{I} - \mathcal{Q}_U)$, since \mathcal{W}_{UV} and \mathcal{Q}_U are both projections onto U . Hence, by Theorem 3.4,

$$\|f - \mathcal{W}_{UV}f\| = \|(\mathcal{I} - \mathcal{W}_{UV})(\mathcal{I} - \mathcal{Q}_U)f\| \leq \sec(\theta_{UV^\perp}) \|f - \mathcal{Q}_Uf\|,$$

with sharp bound. \square

Remark 3.6 Although arbitrary subspaces U and V need not obey the subspace condition, this is often the case in practice. For example, if $U \subseteq V^\perp$ then $\cos(\theta_{UV^\perp}) = 1$ by (3.3).

To complete this section, we present the following lemma which will be useful in what follows:

Lemma 3.7. *Let U and V be closed subspaces of H satisfying the subspace condition. Suppose also that $\dim(U) = \dim(V^\perp) = n < \infty$. Then $U \oplus V = H$.*

Proof. Note that $U \oplus V = H$ if and only if $\cos(\theta_{UV^\perp})$ and $\cos(\theta_{V^\perp U})$ are both positive [53, Thm. 2.3]. Since $\cos(\theta_{UV^\perp}) > 0$ by assumption, it remains to show that $\cos(\theta_{V^\perp U}) > 0$. Consider the mapping $\mathcal{Q}_{V^\perp}|_U : U \rightarrow V^\perp$. We claim that this mapping is invertible. Since U and V^\perp have the same dimension it suffices to show that $\mathcal{Q}_{V^\perp}|_U$ has trivial kernel. However, the existence of a nonzero $u \in U$ with $\mathcal{Q}_{V^\perp}u = 0$ implies that $\cos(\theta_{UV^\perp}) = 0$; a contradiction. Thus $\mathcal{Q}_{V^\perp}|_U$ is invertible, and in particular, it has range V^\perp . Now consider $\cos(\theta_{V^\perp U})$. By (3.3) and this result,

$$\cos(\theta_{V^\perp U}) = \inf_{\substack{w \in V^\perp \\ w \neq 0}} \sup_{\substack{u \in U \\ u \neq 0}} \frac{|\langle w, u \rangle|}{\|w\| \|u\|} = \inf_{\substack{u' \in U \\ u' \neq 0}} \sup_{\substack{u \in U \\ u \neq 0}} \frac{|\langle \mathcal{Q}_{V^\perp}u', u \rangle|}{\|\mathcal{Q}_{V^\perp}u'\| \|u\|} \geq \inf_{\substack{u' \in U \\ u' \neq 0}} \frac{\|\mathcal{Q}_{V^\perp}u'\|}{\|u'\|} = \cos(\theta_{UV^\perp}) > 0.$$

This completes the proof. \square

3.3 Quasi-optimality of consistent sampling

Oblique projections arise in many types of sampling [8], and are intimately related to consistent reconstructions. The following result was proved in [29, Thm. 2.1]:

Lemma 3.8. *Suppose that T and S^\perp satisfy the subspace condition. If $f \in H_0 := T \oplus S^\perp$ then there exists a unique $\tilde{f} \in T$ satisfying (3.1). Moreover, the mapping $F : H_0 \rightarrow T, f \mapsto \tilde{f}$ is well-defined and coincides with the oblique projection \mathcal{W}_{TS^\perp} .*

As a result of this lemma, consistent reconstructions are equivalent to oblique projections. However, we note one important distinction. When defining the consistent reconstruction (3.1), we assume that a frame $\{\psi_j\}_{j=1}^\infty$ of S is given. Indeed, this is natural in view of the sampling process and of course necessary in order to define the condition number $\kappa(F)$. However, the oblique projection \mathcal{W}_{TS^\perp} , being determined solely by the spaces T and S , is actually independent of this basis. In light of Lemma 3.8, the same must also be true for \tilde{f} .

Lemma 3.8, along with Corollary 3.5, gives the following sharp bounds for consistent sampling:

$$\|\tilde{f}\| \leq \sec(\theta_{TS}) \|f\|, \quad \forall f \in H_0, \quad (3.6)$$

$$\|f - \mathcal{Q}_T f\| \leq \|f - \tilde{f}\| \leq \sec(\theta_{TS}) \|f - \mathcal{Q}_T f\|, \quad \forall f \in H_0. \quad (3.7)$$

The latter establishes quasi-optimality of the reconstruction, whereas the former gives a continuous stability estimate, i.e. the norm of \tilde{f} is bounded by a constant multiple of the norm of the input signal. Note that (3.6) and (3.7) were derived previously in [58] and [56] respectively. However, the observation that they are also sharp does not, to the best of our knowledge, appear in the literature.

Another property of the consistent reconstruction is confirmed by (3.7). Namely, it is a *perfect* reconstruction:

Definition 3.9. *A mapping $F : H \rightarrow T$ is a perfect reconstruction if (i) for each $f \in H$, $F(f)$ depends only on the vector of samples \hat{f} , and (ii) $F(f) = f$ whenever $f \in T$.*

3.4 The condition number of consistent sampling

The sharp bound (3.7) shows that the quasi-optimality constant of consistent sampling is $\mu(F) = \sec(\theta_{\text{TS}})$. We now wish to determine the condition number. For this, it is useful to first recall several basic facts about frames [17]. Given the sampling frame $\{\psi_j\}_{j \in \mathbb{N}}$ for the subspace S , we define the *synthesis* operator $\mathcal{S} : \ell^2(\mathbb{N}) \rightarrow H$ by

$$\mathcal{S}\alpha = \sum_{j \in \mathbb{N}} \alpha_j \psi_j, \quad \alpha = \{\alpha_j\}_{j \in \mathbb{N}} \in \ell^2(\mathbb{N}).$$

Its adjoint, the *analysis* operator, is defined by

$$\mathcal{S}^* f = \hat{f} = \{\langle f, \psi_j \rangle\}_{j \in \mathbb{N}}, \quad f \in H.$$

The resulting composition $\mathcal{P} = \mathcal{S}\mathcal{S}^* : H \rightarrow H$, given by

$$\mathcal{P}f = \sum_{j \in \mathbb{N}} \langle f, \psi_j \rangle \psi_j, \quad \forall f \in H, \quad (3.8)$$

is well-defined, linear, self-adjoint and bounded. Moreover, the restriction $\mathcal{P}|_S : S \rightarrow S$ is positive and invertible with $c_1 \mathcal{I}|_S \leq \mathcal{P}|_S \leq c_2 \mathcal{I}|_S$, where c_1, c_2 are the frame constants appearing in (2.1).

We now require the following lemma:

Lemma 3.10. *Suppose that T and S^\perp satisfy the subspace condition, and let \mathcal{P} be given by (3.8). Then*

$$c_1 \cos^2(\theta_{\text{TS}}) \mathcal{I}|_T \leq \mathcal{P}|_T \leq c_2 \mathcal{I}|_T. \quad (3.9)$$

Proof. Let $f \in H$ be arbitrary, and write $f = \mathcal{Q}_S f + \mathcal{Q}_{S^\perp} f$. Then

$$\langle \mathcal{P}f, f \rangle = \sum_{j \in \mathbb{N}} |\langle f, \psi_j \rangle|^2 = \sum_{j \in \mathbb{N}} |\langle \mathcal{Q}_S f, \psi_j \rangle|^2 = \langle \mathcal{P} \mathcal{Q}_S f, \mathcal{Q}_S f \rangle. \quad (3.10)$$

Suppose now that $\phi \in T$. Using (3.10) and the frame condition (2.1) we find that

$$c_1 \|\mathcal{Q}_S \phi\|^2 \leq \langle \mathcal{P} \phi, \phi \rangle \leq c_2 \|\mathcal{Q}_S \phi\|^2 \leq c_2 \|\phi\|^2.$$

To obtain (3.9) we now use the definition of the subspace angle θ_{TS} . □

Corollary 3.11. *Suppose that S and T are as in Lemma 3.10, and let $F : H_0 := T \oplus S^\perp \rightarrow T$ denote the consistent reconstruction defined by (3.1). Then the condition number $\kappa(F)$ satisfies*

$$\frac{\sec(\theta_{\text{TS}})}{\sqrt{c_2}} \leq \kappa(F) \leq \frac{\sec(\theta_{\text{TS}})}{\sqrt{c_1}}.$$

Proof. Since the reconstruction $\tilde{f} = F(f) \in T$ is defined by (3.1), we have

$$\|\hat{f}\|_{\ell^2}^2 = \sum_{j \in \mathbb{N}} |\langle f, \psi_j \rangle|^2 = \sum_{j \in \mathbb{N}} |\langle \tilde{f}, \psi_j \rangle|^2 = \langle \mathcal{P} \tilde{f}, \tilde{f} \rangle.$$

Hence, by the previous lemma, $\|\hat{f}\|_{\ell^2}^2 \geq c_1 \cos^2(\theta_{\text{TS}}) \|\tilde{f}\|^2$. Since F is linear, this now gives

$$\kappa(F) = \sup_{\substack{f \in H_0 \\ \hat{f} \neq 0}} \left\{ \frac{\|F(f)\|}{\|\hat{f}\|_{\ell^2}} \right\} \leq \frac{\sec(\theta_{\text{TS}})}{\sqrt{c_1}}.$$

On the other hand, since the reconstruction F is perfect, and since $\hat{f} = 0$ if and only if $f = 0$ for $f \in T$,

$$\kappa(F) \geq \sup_{\substack{f \in T \\ \hat{f} \neq 0}} \left\{ \frac{\|f\|}{\|\hat{f}\|_{\ell^2}} \right\} = \sup_{\substack{f \in T \\ f \neq 0}} \left\{ \frac{\|f\|}{\|\hat{f}\|_{\ell^2}} \right\}.$$

By (3.10), we have $\|\hat{f}\|_{\ell^2}^2 \leq c_2 \|\mathcal{Q}_S f\|^2$. Hence

$$\kappa(F) \geq \frac{1}{\sqrt{c_2}} \sup_{\substack{f \in \mathbb{T} \\ f \neq 0}} \left\{ \frac{\|f\|}{\|\mathcal{Q}_S f\|} \right\} = \frac{\sec(\theta_{\mathbb{T}\mathbb{S}})}{\sqrt{c_2}},$$

as required. \square

Combining Corollaries 3.5 and 3.11, we now find that the reconstruction constant $C(F)$ of consistent sampling satisfies

$$\sec(\theta_{\mathbb{T}\mathbb{S}}) \max\{1, \frac{1}{\sqrt{c_2}}\} \leq C(F) \leq \sec(\theta_{\mathbb{T}\mathbb{S}}) \max\{1, \frac{1}{\sqrt{c_1}}\}. \quad (3.11)$$

Hence, if \mathbb{S} and \mathbb{T} are not close to perpendicular (i.e. if $\cos(\theta_{\mathbb{T}\mathbb{S}})$ is not too small), then consistent sampling is stable and quasi-optimal.

3.5 Consistent sampling for the computational reconstruction problem

The consistent reconstruction \tilde{f} solves the reconstruction problem of recovering f from the infinite vector $\hat{f} = \{\hat{f}_j\}_{j \in \mathbb{N}}$. Note that once a frame $\{\phi_j\}_{j \in \mathbb{N}}$ is specified for \mathbb{T} , this is equivalent to the infinite system of linear equations $U\alpha = \hat{f}$, where U is the infinite matrix

$$U = \begin{pmatrix} \langle \phi_1, \psi_1 \rangle & \langle \phi_2, \psi_1 \rangle & \cdots \\ \langle \phi_1, \psi_2 \rangle & \langle \phi_2, \psi_2 \rangle & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}, \quad (3.12)$$

and $\alpha = \{\alpha_j\}_{j \in \mathbb{N}} \in \ell^2(\mathbb{N})$ is such that $\tilde{f} = \sum_{j \in \mathbb{N}} \alpha_j \phi_j$. Observe that U , which we may view as a bounded operator on $\ell^2(\mathbb{N})$, coincides with $\mathcal{X} = \mathcal{S}^* \mathcal{T} : \ell^2(\mathbb{N}) \rightarrow \ell^2(\mathbb{N})$, where \mathcal{S} and \mathcal{T} are the synthesis operators for the frames $\{\psi_j\}_{j=1}^\infty$ and $\{\phi_j\}_{j=1}^\infty$ respectively.

Clearly this approach does not solve the computational reconstruction problem outlined in §2.2 since one cannot compute solutions to $U\alpha = \hat{f}$ in general. To overcome this, the standard approach [24, 25, 26, 38, 55] is to replace the infinite consistency conditions (3.1) by a finite version. That is, we seek a reconstruction $\tilde{f}_{n,n}$ defined by

$$\langle \tilde{f}_{n,n}, \psi_j \rangle = \langle f, \psi_j \rangle, \quad j = 1, \dots, n, \quad \tilde{f}_{n,n} \in \mathbb{T}_n, \quad (3.13)$$

(the use of the double index in $\tilde{f}_{n,n}$ is for agreement with subsequent notation). Note that if $\mathbb{T}_n = \text{span}\{\phi_1, \dots, \phi_n\}$ then this is equivalent to the finite linear system $U^{[n,n]} \alpha^{[n,n]} = \hat{f}^{[n]}$, where

$$U^{[n,n]} = \begin{pmatrix} \langle \phi_1, \psi_1 \rangle & \cdots & \langle \phi_n, \psi_1 \rangle \\ \vdots & \ddots & \vdots \\ \langle \phi_1, \psi_n \rangle & \cdots & \langle \phi_n, \psi_n \rangle \end{pmatrix}, \quad (3.14)$$

$\hat{f}^{[n]} = \{\hat{f}_1, \dots, \hat{f}_n\}$, $\alpha^{[n,n]} = \{\alpha_1^{[n]}, \dots, \alpha_n^{[n]}\}$ and $\tilde{f}_{n,n}$ is given by $\sum_{j=1}^n \alpha_j^{[n,n]} \phi_j$.

The condition (3.13) is completely natural enforce: it states that the available measurements of f agree with those of the reconstruction $\tilde{f}_{n,n}$. It is tempting to think that stability and quasi-optimality of the infinite-dimensional consistent reconstruction \tilde{f} should imply the same behaviour of $\tilde{f}_{n,n}$. In other words, $C^* = \sup_{n \in \mathbb{N}} C(F_{n,n})$ should be both finite and not too large. Unfortunately, as we explain in the next section, there is no guarantee that this will be the case in practice.

First, however, let us determine the reconstruction constant for this approach. As a direct consequence of Lemmas 3.7, 3.8 and Corollary 3.5, we have

Corollary 3.12. *Let $\mathbb{S}_n = \text{span}\{\psi_1, \dots, \psi_n\}$ and suppose that*

$$\cos(\theta_{n,n}) > 0, \quad (3.15)$$

where $\theta_{n,n} = \theta_{T_n S_n}$. Then, for each $f \in H_n := T_n \oplus S_n^\perp$ there exists a unique $\tilde{f}_{n,n} \in T_n$ satisfying (3.13). Moreover, the mapping $F_{n,n} : H_n \rightarrow T_n, f \mapsto \tilde{f}_{n,n}$ coincides with the oblique projection $\mathcal{W}_{T_n S_n^\perp}$, and we have the sharp bounds

$$\|\tilde{f}_{n,n}\| \leq \sec(\theta_{n,n}) \|f\|, \quad (3.16)$$

and

$$\|f - \mathcal{Q}_n f\| \leq \|f - \tilde{f}_{n,n}\| \leq \sec(\theta_{n,n}) \|f - \mathcal{Q}_n f\|, \quad (3.17)$$

where \mathcal{Q}_n is the orthogonal projection onto T_n . If $\dim(T_n) = \dim(S_n)$ (in particular, if both $\{\psi_j\}_{j \in \mathbb{N}}$ and $\{\phi_j\}_{j \in \mathbb{N}}$ are bases), then the above conclusions hold with $H_n = H$.

This theorem demonstrates that the quasi-optimality constant of consistent sampling satisfies

$$\mu(F_{n,n}) = \sec(\theta_{n,n}). \quad (3.18)$$

We now state the following result concerning the condition number (a proof is given in §4):

Corollary 3.13. *Suppose that $\cos(\theta_{n,n}) > 0$ and that $\dim(S_n) = \dim(T_n)$. Then the condition number of consistent sampling satisfies*

$$\kappa(F_{n,n}) \geq \frac{1}{\sqrt{c_2}} \sec(\theta_{n,n}).$$

This result, in combination with (3.18), implies that the reconstruction constant

$$C(F_{n,n}) \geq \max\{1, \frac{1}{\sqrt{c_2}}\} \sec(\theta_{n,n}).$$

Hence, good behaviour of the reconstructions (i.e. stability and quasi-optimality for all n) is only possible if the subspace angles $\theta_{n,n}$ remain bounded away from $\frac{\pi}{2}$ for all $n \in \mathbb{N}$. As we next explain, there is no reason for this to be the case in general.

3.6 Example

Consider the following simple problem. Let $H = L^2(-1, 1)$ and, for $0 < \delta < 1$ set

$$\psi_j(x) = \frac{1}{\sqrt{2}} e^{ij\pi\delta x}, \quad \phi_j(x) = \frac{1}{\sqrt{2}} e^{ij\pi x}, \quad j \in \mathbb{Z},$$

(for convenience we now index over \mathbb{Z} , as opposed to \mathbb{N}). Note that $\{\psi_j\}_{j \in \mathbb{Z}}$ is a tight frame for $S = H$ with frame bounds $c_1 = c_2 = \frac{1}{\delta}$, and $\{\phi_j\}_{j \in \mathbb{Z}}$ is an orthonormal basis for $T = H$. Since $S = T = H$ in this case, we have $\cos(\theta_{TS}) = 1$.

This example is a simple instance of the so-called *uniform resampling* (URS) problem. In this problem, Fourier samples are acquired nonuniformly and one wishes to compute standard harmonic Fourier coefficients. We shall consider the URS problem in more detail in §7. Note that the above example, where ‘nonuniformly’ merely means sampling at the rate $1/\delta$, is reasonably unrealistic. Yet, as we see next, it serves to illustrate the main point: namely, consistent reconstructions can be extremely unstable, even for apparently simple problems.

In Figure 1(a) we plot the behaviour of $\cos(\theta_{n,n})$ against n for $\delta = \frac{1}{2}$, where

$$S_n = \text{span}\{\psi_j : |j| \leq n\}, \quad T_n = \text{span}\{\phi_j : |j| \leq n\}.$$

As is evident, the quantity $\cos(\theta_{n,n})$ is exponentially small in n . In particular, when $n = 50$ Figure 1(a), in combination with Corollary 3.13, implies that the reconstruction constant for the consistent reconstruction based on these spaces is around 10^{14} in magnitude.

In this example, the reconstruction space T_n is the space of trigonometric polynomials of degree n , and as such, it is well-suited for recovering smooth and periodic functions. In particular, if f is smooth, then the error $\|f - \mathcal{Q}_n f\|$ faster than any algebraic power of n^{-1} . However, such good approximation

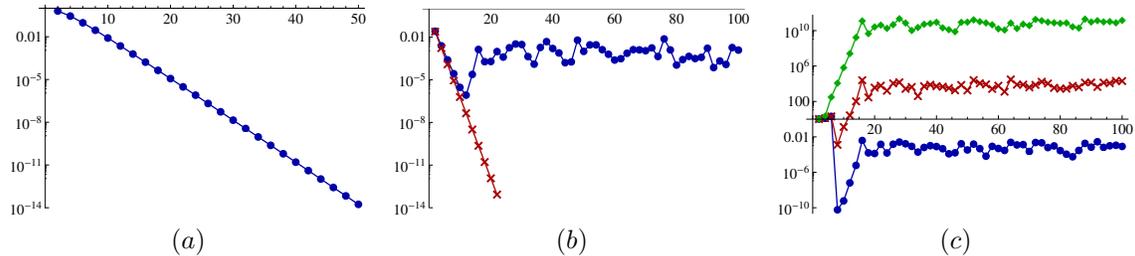


Figure 1: (a): $\cos(\theta_{n,n})$ against n . (b): the errors $\|f - \tilde{f}_{n,n}\|$ (circles) and $\|f - \mathcal{Q}_n f\|$ (crosses) against n , where $\tilde{f}_{n,n}$ is the consistent reconstruction of f , $\mathcal{Q}_n f$ is the orthogonal projection of f onto \mathbb{T}_n , and $f(x) = \frac{1}{2 + \cos \pi x}$. (c): the error $\|f - \tilde{f}_{n,n}\|$ against n , where $f(x) = \frac{1}{\sqrt{2}} e^{8i\pi x}$ and each sample \hat{f}_j is perturbed by an amount η_j chosen uniformly at random with $|\eta_j| \leq \eta$ for $\eta = 0, 10^{-9}, 10^{-2}$ (circles, crosses and diamonds respectively).

properties are destroyed when passing to $\tilde{f}_{n,n}$. In Figure 1(b) we plot the error $\|f - \tilde{f}_{n,n}\|$ for the consistent reconstruction, as well as that of the orthogonal projection $\mathcal{Q}_n f$. The latter decays rapidly with n , as expected. On the other hand, the maximal achievable accuracy of $\tilde{f}_{n,n}$ is limited to only around one or two digits due to the ill-conditioning and the effect of round-off errors.

The situation worsens significantly when the samples \hat{f}_j are corrupted by noise. The function used in Figure 1(c) actually lies in \mathbb{T}_n whenever $n \geq 8$, and therefore, in theory at least, should be recovered perfectly. However, this is completely obliterated by even moderate amounts of noise. For example, with noise at amplitude 10^{-9} the reconstruction error is around 10^4 (i.e. an amplification of 10^{15}), rendering such an approach useless for this particular problem.

Remark 3.14 The exponential blow-up of the reconstruction constant in the above example is by no means unique to this particular problem. In [4] several other problems, based on spaces \mathbb{T}_n consisting of polynomials or piecewise constant functions, were shown to exhibit similar exponential growth.

3.7 Operator-theoretic interpretation

The failure of the finite-dimensional consistent reconstruction $\tilde{f}_{n,n}$ can be traced to the poor behaviour of the finite subspace angles $\theta_{n,n} = \theta_{\mathbb{T}_n, \mathbb{S}_n}$ in relation to $\theta = \theta_{\mathbb{T}\mathbb{S}}$. As was discussed in [4], this failure can also be understood in terms of operators. If the sampling and reconstruction bases are orthonormal (this fact is not necessary, but simplifies what follows) then $\cos(\theta_{n,n})$ and $\cos(\theta)$ coincide with the minimal singular values of the matrices $U^{[n,n]}$ and U respectively. Hence, the fact that $\theta_{n,n}$ may behave wildly, even when θ is bounded away from $\frac{\pi}{2}$, demonstrates that the spectra of the matrices $U^{[n,n]}$ poorly approximate the spectrum of U .

This question – namely, how well does a sequence of finite-rank operators approximate the spectrum of a given infinite-rank operator – is one of the most fundamental in the field of spectral theory. Recalling the definition (3.14), we notice that $U^{[n,n]}$ is nothing more than the $n \times n$ *finite section* of U : if $\{e_j\}$ is the canonical basis for $\ell^2(\mathbb{N})$ and $P_n : \ell^2(\mathbb{N}) \rightarrow \text{span}\{e_1, \dots, e_n\}$ is the orthogonal projection, then $U^{[n,n]} = P_n U P_n$. 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 the equations $U\alpha = \hat{f}$.

The properties of finite sections have been extensively studied over the last several decades [10, 35, 45], and unfortunately there is no guarantee that they be well behaved. To put this in a formal perspective, suppose for the moment that we approximate the operator U with a sequence $U^{[n]}$ of finite-rank operators (which may or may not be finite sections), and instead of solving $U\alpha = \hat{f}$, we solve $U^{[n]}\alpha^{[n]} = \hat{f}^{[n]}$. For obvious reasons, it is vitally important that this sequence satisfies the three following conditions:

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

- (ii) *Stability*: $\|(U^{[n]})^{-1}\|$ is uniformly bounded for all $n = 1, 2, \dots$
- (iii) *Convergence*: the solutions $\alpha^{[n]} \rightarrow \alpha$ as $n \rightarrow \infty$.

Unfortunately, there is no guarantee that finite sections, and therefore the consistent reconstruction technique, possess any of these properties. In fact, one requires rather restrictive conditions, such as positive self-adjointness, for this to be the case. Typically operators U of the form (3.12) are not self-adjoint, thereby making this approach unsuitable in general for discretizing $U\alpha = \hat{f}$.

Generalized sampling, which we next discuss, overcomes these issues by obtaining a sequence of operators that possess the properties (i)–(iii) above. The key to doing this is to allow the number of samples m and the number of reconstruction vectors n to differ. When m is sufficiently large for a given n , or equivalently, n is sufficiently small for a given m , we obtain a finite-dimensional operator $U^{[n,m]}$ (which now depends on both m and n) that inherits the spectral structure of its infinite-dimensional counterpart U . This ensures a stable, quasi-optimal reconstruction.

4 Generalized sampling

Generalized sampling was introduced in [4], and applied to the resolution of the Gibbs phenomenon in [5]. Several extensions have also been pursued. These include infinite-dimensional compressed sensing [1], inverse and ill-posed problems [6], and problems where the sampling and reconstruction systems lie in different Hilbert spaces [2].

From now on we shall assume that the subspaces T and S^\perp obey the subspace condition. Without this, the infinite-dimensional reconstruction problem is itself ill-posed, and thus the computational reconstruction problem becomes substantially more difficult.

Let $S_m = \text{span}\{\psi_1, \dots, \psi_m\}$ and suppose that $\{T_n\}_{n \in \mathbb{N}}$ is a sequence of finite-dimensional reconstruction spaces satisfying (2.6) and (2.7). We seek a reconstruction $\tilde{f}_{n,m} \in T_n$ of f from the m samples $\hat{f}_1, \dots, \hat{f}_m$. Let $\mathcal{P}_m : H \rightarrow S_m$ be the finite rank operator given by

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

Note that the sequence of operators \mathcal{P}_m converge strongly to \mathcal{P} on H , where \mathcal{P} is given by (3.8), since $\{\psi_j\}_{j \in \mathbb{N}}$ is a frame [17]. With this to hand, the approach proposed in [4] is to define $\tilde{f}_{n,m} \in T_n$ as the solution of the equations

$$\langle \mathcal{P}_m \tilde{f}_{n,m}, \phi_j \rangle = \langle \mathcal{P}_m f, \phi_j \rangle, \quad j = 1, \dots, n, \quad \tilde{f}_{n,m} \in T_n. \quad (4.1)$$

We refer to the mapping $F_{n,m} : f \mapsto \tilde{f}_{n,m}$ as *generalized sampling*. Observe that $\mathcal{P}_m f$ is determined solely by the coefficients $\hat{f}_1, \dots, \hat{f}_m$ of f . Hence $F_{n,m}(f)$ is also determined solely by these values.

In what follows it will be useful to note that (4.1) is equivalent to

$$\langle \tilde{f}_{n,m}, \mathcal{P}_m \phi_j \rangle = \langle f, \mathcal{P}_m \phi_j \rangle, \quad j = 1, \dots, n, \quad \tilde{f}_{n,m} \in T_n, \quad (4.2)$$

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

Lemma 4.1. *Suppose that $\cos(\theta_{n,n}) > 0$ and that $\dim(S_n) = \dim(T_n)$. Then when $m = n$ the generalized sampling reconstruction $\tilde{f}_{n,m}$ of $f \in H$ defined by (4.1) is precisely the consistent reconstruction $\tilde{f}_{n,n}$ defined by (3.13).*

Proof. We first claim that \mathcal{P}_n is a bijection from T_n to S_n . Suppose that $\mathcal{P}_n \phi = 0$ for some $\phi \in T_n$. Then $0 = \langle \mathcal{P}_n \phi, \phi \rangle = \sum_{j=1}^n |\langle \phi, \psi_j \rangle|^2$ and therefore $\phi \in S_n^\perp$. Since $\phi \in T_n$, and $T_n \cap S_n^\perp = \{0\}$ by assumption, we have $\phi = 0$, as required.

By linearity, we now find that the conditions (4.2) are equivalent to (3.13). Since the consistent reconstruction $\tilde{f}_{n,n}$ satisfying (3.1) exists uniquely (Corollary 3.12), we obtain the result. \square

We conclude that generalized sampling contains consistent sampling as a special case corresponding to $n = m$, which explains our use of the same notation for both. However, as mentioned above, the key to generalized sampling is to allow m and n to vary independently. As we prove next, doing so leads to a small reconstruction constant.

4.1 An intuitive argument

Before providing a full analysis, we first give an intuitive explanation as to why letting n and m vary independently works. To this end, suppose that n is fixed and let $m \rightarrow \infty$. Equations (4.1) now read

$$\langle \mathcal{P}\tilde{f}_{n,\infty}, \phi \rangle = \langle \mathcal{P}f, \phi \rangle, \quad \forall \phi \in \mathbb{T}_n, \quad \tilde{f}_{n,\infty} \in \mathbb{T}_n,$$

for some $\tilde{f}_{n,\infty} \in \mathbb{T}_n$, where \mathcal{P} is given by (3.8). In [5] it was shown that $\tilde{f}_{n,\infty} = \lim_{m \rightarrow \infty} \tilde{f}_{n,m}$ for fixed $n \in \mathbb{N}$. Hence, we can understand the behaviour of $\tilde{f}_{n,m}$ for large m by first analyzing $\tilde{f}_{n,\infty}$.

Since \mathcal{P} is self-adjoint, $\tilde{f}_{n,\infty}$ is equivalently defined by

$$\langle \tilde{f}_{n,\infty}, \Phi \rangle = \langle f, \Phi \rangle, \quad \forall \Phi \in \mathcal{P}(\mathbb{T}_n), \quad \tilde{f}_{n,\infty} \in \mathbb{T}_n. \quad (4.3)$$

We now have

Theorem 4.2. *For each $f \in \mathbb{H}$, there exists a unique $\tilde{f}_{n,\infty} \in \mathbb{T}_n$ satisfying (4.3). Moreover, the mapping $f \mapsto \tilde{f}_{n,\infty}$ is precisely the oblique projection with range \mathbb{T}_n and kernel $(\mathcal{P}(\mathbb{T}_n))^\perp$, and we have the sharp bounds*

$$\|\tilde{f}_{n,\infty}\| \leq \sec(\theta_{n,\infty}) \|f\|, \quad (4.4)$$

and

$$\|f - \mathcal{Q}_n f\| \leq \|f - \tilde{f}_{n,\infty}\| \leq \sec(\theta_{n,\infty}) \|f - \mathcal{Q}_n f\|, \quad (4.5)$$

where $\theta_{n,\infty}$ is the angle between \mathbb{T}_n and $\mathcal{P}(\mathbb{T}_n)$.

Proof. We first claim that $\cos(\theta_{n,\infty}) > 0$, so that the oblique projection \mathcal{W} with range \mathbb{T}_n and kernel $(\mathcal{P}(\mathbb{T}_n))^\perp$ is well-defined on $\mathbb{H}_0 = \mathbb{T}_n \oplus (\mathcal{P}(\mathbb{T}_n))^\perp$. Suppose not. Since \mathbb{T}_n is finite dimensional, there exists a $\phi \in \mathbb{T}_n$, $\phi \neq 0$, satisfying $\mathcal{Q}_{\mathcal{P}(\mathbb{T}_n)}\phi = 0$. Thus

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

and, in particular, $\langle \phi, \mathcal{P}\phi \rangle = 0$. Thus $\phi = 0$ by Lemma 3.10, a contradiction. Hence \mathcal{W} is well-defined. Note that $\mathcal{W}f$ satisfies the equations (4.3). Arguing in the standard way, we can show that solutions to (4.3) are unique. Hence $\tilde{f}_{n,\infty} = \mathcal{W}f$, as required.

It remains to show that $\mathbb{H}_0 = \mathbb{H}$. The result follows immediately from Lemma 3.7 provided $\dim(\mathcal{P}(\mathbb{T}_n)) = \dim(\mathbb{T}_n)$. However, if not, then there is a nonzero $\phi \in \mathbb{T}_n$ with $\mathcal{P}\phi = 0$, which also contradicts Lemma 3.10. \square

Note that this theorem improves on [5, Thm. 2.1] by giving sharp bounds. We can also estimate the reconstruction constant of the mapping $F_{n,\infty} : f \mapsto \tilde{f}_{n,\infty}$:

Corollary 4.3. *Let $F_{n,\infty}$ be the mapping $f \mapsto \tilde{f}_{n,\infty}$, where $\tilde{f}_{n,\infty}$ is defined by (4.3). Then*

$$1 \leq \mu(F_{n,\infty}) \leq \sqrt{\frac{c_2}{c_1}} \sec(\theta_{\text{TS}}), \quad \frac{1}{\sqrt{c_2}} \leq \kappa(F_{n,\infty}) \leq \frac{\sec(\theta_{\text{TS}})}{\sqrt{c_1}}. \quad (4.6)$$

and therefore

$$\max \left\{ 1, \frac{1}{\sqrt{c_2}} \right\} \leq C(F_{n,\infty}) \leq \frac{\max \{1, \sqrt{c_2}\}}{\sqrt{c_1}} \sec(\theta_{\text{TS}}). \quad (4.7)$$

This corollary (we present the proof in the next section) confirms that the reconstruction scheme $\{F_{n,\infty}\}_{n \in \mathbb{N}}$ is stable and quasi-optimal in the sense of Definition 2.5, with constant $C^* \leq \frac{\max\{1, \sqrt{c_2}\}}{\sqrt{c_1}} \sec(\theta_{\text{TS}})$. Hence, unlike the consistent sampling scheme $\{F_{n,n}\}_{n \in \mathbb{N}}$, where the reconstruction constants $C(F_{n,n})$ can quite easily be exponentially large in n (see §3.6), this approach guarantees good approximation and robustness with respect to noise.

However, one cannot actually compute $\tilde{f}_{n,\infty}$, since it involves the infinite-rank operator \mathcal{P} . Nonetheless, since the generalized sampling reconstruction $\tilde{f}_{n,m} \approx \tilde{f}_{n,\infty}$ for large m , we expect the key properties of $F_{n,\infty}$ to be inherited whenever m is sufficiently large. In §4.2 we prove this to be the case.

Before doing so, however, let us relate $\tilde{f}_{n,\infty}$ to the discussion in §3.7. Recall first that we wish to solve $U\alpha = \hat{f}$. Since α satisfies these equations it also obeys the normal equations

$$U^*U\alpha = U^*\hat{f}. \quad (4.8)$$

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

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

where $\hat{f} = \{\hat{f}_1, \hat{f}_2, \dots\}$. Thus, $\alpha^{[n]}$ is precisely the result of the finite section method applied to the normal equations (4.8). Since the operator U^*U is self-adjoint and positive, its finite sections must possess properties (i)–(iii), and hence we are guaranteed a good reconstruction.

4.2 Analysis of generalized sampling

The analysis of $\tilde{f}_{n,m}$ is similar to that of $\tilde{f}_{n,\infty}$. Whilst such analysis was originally given in [4, 5], the estimates derived were not sharp. Our main result in this section is to present new, sharp bounds. We first require the following lemma:

Lemma 4.4. *Let $\theta_{n,m}$ and $\theta_{n,\infty}$ be the angles between T_n and the subspaces $\mathcal{P}_m(T_n)$ and $\mathcal{P}(T_n)$ respectively. Then, for fixed n , $\theta_{n,m} \rightarrow \theta_{n,\infty}$ as $m \rightarrow \infty$. In particular,*

$$1 \leq \lim_{m \rightarrow \infty} \sec(\theta_{n,m}) \leq \sqrt{\frac{c_2}{c_1}} \sec(\theta_{\text{TS}}).$$

Proof. From the definition (3.2), we have

$$\cos(\theta_{n,m}) = \inf_{\substack{\phi \in T_n \\ \|\phi\|=1}} \sup_{\substack{\phi' \in T_n \\ \mathcal{P}_m \phi' \neq 0}} \frac{|\langle \phi, \mathcal{P}_m \phi' \rangle|}{\|\mathcal{P}_m \phi'\|}.$$

Recall first that $\mathcal{P}_m \rightarrow \mathcal{P}$ strongly on H . Since T_n is finite-dimensional, this implies uniform convergence of $\mathcal{P}_m \rightarrow \mathcal{P}$ on T_n , i.e. $e_m = \|\mathcal{P}|_{T_n} - \mathcal{P}_m|_{T_n}\| \rightarrow 0$ as $m \rightarrow \infty$. In particular, for sufficiently large m , $\mathcal{P}_m \phi' \neq 0$ if and only if $\mathcal{P} \phi' \neq 0$. Thus, for large m ,

$$\cos(\theta_{n,m}) = \inf_{\substack{\phi \in T_n \\ \|\phi\|=1}} \sup_{\substack{\phi' \in T_n \\ \mathcal{P} \phi' \neq 0}} \frac{|\langle \phi, \mathcal{P}_m \phi' \rangle|}{\|\mathcal{P}_m \phi'\|}. \quad (4.9)$$

Now

$$\frac{\langle \phi, \mathcal{P}_m \phi' \rangle}{\|\mathcal{P}_m \phi'\|} = \left(\frac{\|\mathcal{P} \phi'\|}{\|\mathcal{P}_m \phi'\|} \right) \left(\frac{\langle \phi, \mathcal{P} \phi' \rangle}{\|\mathcal{P} \phi'\|} - \frac{\langle \phi, (\mathcal{P} - \mathcal{P}_m) \phi' \rangle}{\|\mathcal{P} \phi'\|} \right). \quad (4.10)$$

Note that

$$\left| \frac{\|\mathcal{P}_m \phi'\|}{\|\mathcal{P} \phi'\|} - 1 \right| \leq \frac{\|(\mathcal{P} - \mathcal{P}_m) \phi'\|}{\|\mathcal{P} \phi'\|} \leq e_m \|\phi'\|.$$

Moreover,

$$\|\mathcal{P}\phi'\| = \sup_{\substack{g \in \mathbb{H} \\ \|g\|=1}} |\langle \mathcal{P}\phi', g \rangle| \geq \frac{\langle \mathcal{P}\phi', \phi' \rangle}{\|\phi'\|} \geq c_1 \cos^2(\theta_{\text{TS}}) \|\phi'\|.$$

Thus,

$$\|\|\mathcal{P}_m\phi'\| - \|\mathcal{P}\phi'\|\| \leq \|(\mathcal{P} - \mathcal{P}_m)\phi'\| \leq \frac{e_m}{c_1} \sec^2(\theta_{\text{TS}}) \|\mathcal{P}\phi'\|. \quad (4.11)$$

Combining this with (4.10), we now obtain

$$\frac{|\langle \phi, \mathcal{P}_m\phi' \rangle|}{\|\mathcal{P}_m\phi'\|} \geq \frac{1}{1 + \frac{e_m}{c_1} \sec^2(\theta_{\text{TS}})} \left(\frac{|\langle \phi, \mathcal{P}\phi' \rangle|}{\|\mathcal{P}\phi'\|} - \|\phi\| \frac{e_m}{c_1} \sec^2(\theta_{\text{TS}}) \right),$$

and

$$\frac{|\langle \phi, \mathcal{P}_m\phi' \rangle|}{\|\mathcal{P}_m\phi'\|} \leq \frac{1}{1 - \frac{e_m}{c_1} \sec^2(\theta_{\text{TS}})} \left(\frac{|\langle \phi, \mathcal{P}\phi' \rangle|}{\|\mathcal{P}\phi'\|} + \|\phi\| \frac{e_m}{c_1} \sec^2(\theta_{\text{TS}}) \right),$$

Hence (4.9) now gives

$$\begin{aligned} \cos(\theta_{n,m}) &\geq \frac{1}{1 + \frac{e_m}{c_1} \sec^2(\theta_{\text{TS}})} \left(\cos(\theta_{n,\infty}) - \frac{e_m}{c_1} \sec^2(\theta_{\text{TS}}) \right) \\ \cos(\theta_{n,m}) &\leq \frac{1}{1 - \frac{e_m}{c_1} \sec^2(\theta_{\text{TS}})} \left(\cos(\theta_{n,\infty}) + \frac{e_m}{c_1} \sec^2(\theta_{\text{TS}}) \right), \end{aligned}$$

and the result now follows from the fact that $e_m \rightarrow 0$ as $m \rightarrow \infty$. \square

We now have:

Theorem 4.5. *For each $n \in \mathbb{N}$ and any $f \in \mathbb{H}$, there exists an m_0 , independent of f , such that the reconstruction $\tilde{f}_{n,m}$ defined by (4.1) exists and is unique for all $m \geq m_0$. In particular, m_0 is the least m such that $\cos(\theta_{n,m}) > 0$. Moreover, the mapping $f \mapsto \tilde{f}_{n,m}$ is precisely the oblique projection with range \mathbb{T}_n and kernel $[\mathcal{P}_m(\mathbb{T}_n)]^\perp$, and we have the sharp bounds*

$$\|\tilde{f}_{n,m}\| \leq \sec(\theta_{n,m}) \|f\|, \quad (4.12)$$

and

$$\|f - \mathcal{Q}_n f\| \leq \|f - \tilde{f}_{n,m}\| \leq \sec(\theta_{n,m}) \|f - \mathcal{Q}_n f\|. \quad (4.13)$$

Proof. The existence of an m_0 such that $\cos(\theta_{n,m}) > 0$ for all $m \geq m_0$ follows from Lemma 4.4. Thus, when $m \geq m_0$ the oblique projection \mathcal{W} with range \mathbb{T}_n and kernel $(\mathcal{P}_m(\mathbb{T}_n))^\perp$ is well-defined over $\mathbb{H}_0 := \mathbb{T}_n \oplus (\mathcal{P}_m(\mathbb{T}_n))^\perp$ and satisfies (4.1). Moreover, under this condition, solutions of (4.1) are unique, and thus $\tilde{f}_{n,m} = \mathcal{W}f$ whenever $f \in \mathbb{H}_0$. An application of Corollary 3.5 now gives (4.12) and (4.13). To complete the proof we need only show that $\mathbb{H}_0 = \mathbb{H}$. This follows immediately from Corollary 3.7, provided $\dim(\mathcal{P}_m(\mathbb{T}_n)) = \dim(\mathbb{T}_n)$. However, if not then there exists a nonzero $\phi \in \mathbb{T}_n \cap (\mathcal{P}_m(\mathbb{T}_n))^\perp$, which contradicts the fact that $\cos(\theta_{n,m}) > 0$. \square

Note that this theorem improves the bounds of [5, Thm. 2.4], and gives the exact value $\mu(F_{n,m}) = \sec(\theta_{n,m})$ for the quasi-optimality constant of generalized sampling.

Having done this, we next determine the condition number $\kappa(F_{n,m})$, and as a result, the reconstruction constant $C(F_{n,m})$. For this, we introduce the following quantity:

$$D_{n,m} = \left(\inf_{\substack{\phi \in \mathbb{T}_n \\ \|\phi\|=1}} \langle \mathcal{P}_m\phi, \phi \rangle \right)^{-\frac{1}{2}}, \quad n, m \in \mathbb{N}, \quad (4.14)$$

(this is similar to the quantity $C_{n,m}$ of [5, Eqn. (2.12)]). Note that $D_{n,m}$ need not be defined for all $n, m \in \mathbb{N}$. However, we will show subsequently that this is the case provided m is sufficiently large (for a given n). We shall also let

$$D_{n,\infty} = \left(\inf_{\substack{\phi \in \mathbb{T}_n \\ \|\phi\|=1}} \langle \mathcal{P}\phi, \phi \rangle \right)^{-\frac{1}{2}}, \quad n \in \mathbb{N}.$$

We now have the following lemma:

Lemma 4.6. *For fixed $n \in \mathbb{N}$, $D_{n,m} \rightarrow D_{n,\infty}$ as $m \rightarrow \infty$. In particular,*

$$\frac{1}{\sqrt{c_2}} \leq \lim_{m \rightarrow \infty} D_{n,m} \leq \frac{\sec(\theta_{\text{TS}})}{\sqrt{c_1}}.$$

Proof. The first result follows from strong convergence of the operators $\mathcal{P}_m \rightarrow \mathcal{P}$ on \mathbb{H} and the fact that \mathbb{T}_n is finite-dimensional. The second result is due to Lemma 3.10. \square

With this to hand, our main result is as follows:

Corollary 4.7. *Let $n \in \mathbb{N}$ and suppose that $m \geq m_0$, where m_0 is as in Theorem 4.5. Let $F_{n,m}$ be the generalized sampling reconstruction $f \mapsto \tilde{f}_{n,m}$, where $\tilde{f}_{n,m}$ is defined by (4.1). Then*

$$\mu(F_{n,m}) = \sec(\theta_{n,m}), \quad \kappa(F_{n,m}) = D_{n,m}, \quad (4.15)$$

and the reconstruction constant $C(F_{n,m})$ satisfies

$$D_{n,m} \leq C(F_{n,m}) \leq \max\{1, \sqrt{c_2}\} D_{n,m}, \quad (4.16)$$

whenever $D_{n,m}$ is defined. In particular, for fixed n ,

$$1 \leq \lim_{m \rightarrow \infty} \mu(F_{n,m}) \leq \sqrt{\frac{c_2}{c_1}} \sec(\theta_{\text{TS}}), \quad \frac{1}{\sqrt{c_2}} \leq \lim_{m \rightarrow \infty} \kappa(F_{n,m}) \leq \frac{\sec(\theta_{\text{TS}})}{\sqrt{c_1}}, \quad (4.17)$$

and

$$\max\left\{1, \frac{1}{\sqrt{c_2}}\right\} \leq \lim_{m \rightarrow \infty} C(F_{n,m}) \leq \frac{\max\{1, \sqrt{c_2}\}}{\sqrt{c_1}} \sec(\theta_{\text{TS}}). \quad (4.18)$$

Proof. We claim that

$$\sec(\theta_{n,m}) \leq \sqrt{c_2} D_{n,m}. \quad (4.19)$$

Note first that $D_{n,m} < \infty$ implies that $\mathcal{P}_m|_{\mathbb{T}_n} : \mathbb{T}_n \rightarrow \mathcal{P}_m(\mathbb{T}_n)$ is invertible. Hence, by (3.2),

$$\cos(\theta_{n,m}) = \inf_{\substack{\phi \in \mathbb{T}_n \\ \phi \neq 0}} \sup_{\substack{\phi' \in \mathbb{T}_n \\ \phi' \neq 0}} \frac{|\langle \phi, \mathcal{P}_m \phi' \rangle|}{\|\phi\| \|\mathcal{P}_m \phi'\|} \geq \inf_{\substack{\phi \in \mathbb{T}_n \\ \phi \neq 0}} \frac{\langle \phi, \mathcal{P}_m \phi \rangle}{\|\phi\| \|\mathcal{P}_m \phi\|}. \quad (4.20)$$

Now consider $\|\mathcal{P}_m \phi\|$. Since $\mathcal{P}_m \phi \in \mathbb{S}_m$, basic properties of \mathcal{P}_m give that

$$\|\mathcal{P}_m \phi\| = \sup_{\substack{\psi \in \mathbb{S}_m \\ \|\psi\|=1}} |\langle \psi, \mathcal{P}_m \phi \rangle| \leq \sqrt{\langle \phi, \mathcal{P}_m \phi \rangle} \sup_{\substack{\psi \in \mathbb{S}_m \\ \|\psi\|=1}} \sqrt{\langle \psi, \mathcal{P}_m \psi \rangle} \leq \sqrt{\langle \phi, \mathcal{P}_m \phi \rangle} \sup_{\substack{\psi \in \mathbb{S} \\ \|\psi\|=1}} \sqrt{\langle \psi, \mathcal{P} \psi \rangle}, \quad (4.21)$$

and therefore

$$\|\mathcal{P}_m \phi\| \leq \sqrt{c_2} \sqrt{\langle \phi, \mathcal{P}_m \phi \rangle}.$$

Applying this to (4.20) now gives (4.19).

Note that (4.16) and (4.18) now follow immediately from (4.15), (4.17), (4.19) and the definition $C(F_{n,m}) = \max\{\mu(F_{n,m}), \kappa(F_{n,m})\}$. Moreover, (4.17) follows from (4.15) and Lemmas 4.4 and 4.6. Hence we only need to prove (4.15). The first part is due to Theorem 4.5. Therefore it remains to show that $\kappa(F_{n,m}) = D_{n,m}$. Since $F_{n,m}$ is linear and, due to (4.13), perfect on \mathbb{T}_n , we have

$$\kappa(F_{n,m}) = \sup_{\substack{f \in \mathbb{H} \\ \hat{f} \neq 0}} \left\{ \frac{\|F_{n,m}(f)\|}{\|\hat{f}\|_{\ell^2}} \right\} \geq \sup_{\substack{\phi \in \mathbb{T}_n \\ \hat{\phi} \neq 0}} \left\{ \frac{\|\phi\|}{\|\hat{\phi}\|_{\ell^2}} \right\}.$$

Since $\|\hat{\phi}\|_{\ell^2}^2 = \langle \mathcal{P}_m \phi, \phi \rangle$, this now gives

$$\kappa(F_{n,m}) \geq \left(\inf_{\substack{\phi \in \mathbb{T}_n \\ \|\phi\|=1}} \langle \mathcal{P}_m \phi, \phi \rangle \right)^{-\frac{1}{2}} = D_{n,m}.$$

We now wish to derive the upper bound. Let $F_{n,m}(f) = \tilde{f}_{n,m}$. Since $\tilde{f}_{n,m} \in \mathbb{T}_n$, (4.1) gives that

$$\left| \langle \mathcal{P}_m \tilde{f}_{n,m}, \tilde{f}_{n,m} \rangle \right| = \left| \langle \mathcal{P}_m f, \tilde{f}_{n,m} \rangle \right| \leq \sqrt{\langle \mathcal{P}_m f, f \rangle} \sqrt{\langle \mathcal{P}_m \tilde{f}_{n,m}, \tilde{f}_{n,m} \rangle}.$$

Thus $\|\hat{f}\|_{\ell^2} \geq \sqrt{\langle \mathcal{P}_m \tilde{f}_{n,m}, \tilde{f}_{n,m} \rangle}$. Since $f \mapsto \tilde{f}_{n,m}$ is a surjection onto \mathbb{T}_n , we therefore deduce that

$$\kappa(F_{n,m}) \leq \sup_{\substack{f \in \mathbb{H} \\ \hat{f} \neq 0}} \left\{ \frac{\|\tilde{f}_{n,m}\|}{\sqrt{\langle \mathcal{P}_m \tilde{f}_{n,m}, \tilde{f}_{n,m} \rangle}} \right\} = \sup_{\substack{\phi \in \mathbb{T}_n \\ \hat{\phi} \neq 0}} \left\{ \frac{\|\phi\|}{\sqrt{\langle \mathcal{P}_m \phi, \phi \rangle}} \right\} = D_{n,m},$$

as required. This completes the proof. \square

We are now in a position to establish the two results not proved previously:

Proof of Corollary 4.3. By replacing \mathcal{P}_m by \mathcal{P} in the proof of Corollary 4.7 we find that $\sec(\theta_{n,\infty}) \leq \sqrt{c_2} D_{n,\infty}$ and $\kappa(F_{n,\infty}) = D_{n,\infty}$. The result now follows from this and Lemmas 4.4 and 4.6. \square

Proof of Corollary 3.13. Under the assumption $\dim(S_n) = \dim(\mathbb{T}_n)$ the consistent reconstruction (3.13) coincides with the generalized sampling reconstruction (4.1) (Lemma 4.1). The result now follows immediately from (4.15) and (4.19). \square

Corollary 4.7 confirms the advantage of generalized sampling. Given $n \in \mathbb{N}$, one can always take m sufficiently large to guarantee a stable, quasi-optimal reconstruction with reconstruction constant asymptotically bounded by $\frac{\max\{1, \sqrt{c_2}\}}{\sqrt{c_1}} \sec(\theta_{\text{TS}})$. Conversely, if $m \in \mathbb{N}$ is fixed, one may take n sufficiently small to achieve the same goal.

The main issue remaining is to determine the nature of the scaling of m and n . This will be discussed in §5. First, however, let us connect generalized sampling to the discussion of §3.7. Observe that if $\tilde{f}_{n,m} = \sum_{j=1}^n \alpha_j^{[n,m]} \phi_j$, then the vector $\alpha^{[n,m]} = \{\alpha_1^{[n,m]}, \dots, \alpha_n^{[n,m]}\} \in \mathbb{C}^n$ 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 $\hat{f}^{[m]} = P_m \hat{f}$. The matrix $U^{[n,m]}$ is the leading $m \times n$ submatrix of U , and is sometimes referred to as an *uneven section* of U . Uneven sections have recently gained prominence as effective alternatives to the finite section method for discretizing non-self adjoint operators [32, 37]. In particular, in [34] 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

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

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 .

Remark 4.8 Finite (and uneven) sections have been extensively studied [10, 35, 45], and there exists a well-developed and intricate theory of their properties involving C^* -algebras [33]. However, these general results say little about the rate of convergence, nor do they provide explicit constants. Yet, as illustrated in Theorem 4.5, 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 practical importance, such constants can also be numerically computed (see §5).

4.3 The condition number and quasi-optimality constant

As shown, the condition number $\kappa(F_{n,m})$ coincides with $D_{n,m}$, and the quasi-optimality constant $\mu(F_{n,m}) = \sec(\theta_{n,m})$, where $\theta_{n,m}$ is the angle between T_n and $\mathcal{P}_m(T_n)$. In addition, since $\sec(\theta_{n,m}) \leq \sqrt{c_2} D_{n,m}$ by (4.19), one can control the behaviour of both quantities, and therefore also $C(F_{n,m})$, by controlling $D_{n,m}$. The advantage of this, as we discuss in §5, is that it is typically easier to compute $D_{n,m}$ than it is $\theta_{n,m}$.

However, it is in general possible for $\sec(\theta_{n,m})$ to be somewhat smaller than $D_{n,m}$. Thus, in numerical examples, one may see a better approximation than the bound (4.19) suggests. The following result clarifies the relationship between $\sec(\theta_{n,m})$ and $D_{n,m}$:

Lemma 4.9. *Let $c_{1,m}$ and $c_{2,m}$ be the frame bounds for the frame sequence $\{\psi_1, \dots, \psi_m\}$. Then*

$$\sqrt{c_{1,m}} D_{n,m} \leq \sec(\theta_{n,m}) \leq \sqrt{c_{2,m}} D_{n,m},$$

In particular, if $\{\psi_j\}_{j \in \mathbb{N}}$ is a Riesz basis, we have

$$d_1 D_{n,m} \leq \sec(\theta_{n,m}) \leq d_2 D_{n,m},$$

where $d_1, d_2 > 0$ are the Riesz basis constants for $\{\psi_j\}_{j \in \mathbb{N}}$, and when $\{\psi_j\}_{j \in \mathbb{N}}$ is an orthonormal basis it holds that $\sec(\theta_{n,m}) = D_{n,m}$.

Note that, by Riesz basis constants, we mean constants $d_1, d_2 > 0$ such that

$$d_1 \|\beta\|_{\ell^2} \leq \left\| \sum_{j \in \mathbb{N}} \beta_j \psi_j \right\| \leq d_2 \|\beta\|_{\ell^2}, \quad \forall \beta = \{\beta_j\}_{j \in \mathbb{N}} \in \ell^2(\mathbb{N}).$$

Proof of Lemma 4.9. By definition

$$\cos(\theta_{n,m}) = \inf_{\substack{\phi \in T_n \\ \|\phi\|=1}} \|\mathcal{Q}_{\mathcal{P}_m(T_n)} \phi\|.$$

Consider $\|\mathcal{Q}_{\mathcal{P}_m(T_n)} \phi\|$. Since $\mathcal{P}_m(T_n) \subseteq S_m$, we have

$$\|\mathcal{Q}_{\mathcal{P}_m(T_n)} \phi\| = \sup_{\substack{\psi \in S_m \\ \|\psi\|=1}} |\langle \phi, \psi \rangle|.$$

Recall that the operator \mathcal{P}_m is invertible on S_m . Hence

$$\|\mathcal{Q}_{\mathcal{P}_m(T_n)} \phi\| = \sup_{\substack{\psi \in S_m \\ \psi \neq 0}} \frac{|\langle \phi, \mathcal{P}_m \psi \rangle|}{\|\mathcal{P}_m \psi\|} \leq \sqrt{\langle \mathcal{P}_m \phi, \phi \rangle} \sup_{\substack{\psi \in S_m \\ \psi \neq 0}} \frac{\sqrt{\langle \mathcal{P}_m \psi, \psi \rangle}}{\|\mathcal{P}_m \psi\|}. \quad (4.22)$$

Consider the latter term. The operator $\mathcal{P}_m : S_m \rightarrow S_m$ is invertible, self-adjoint and positive-definite. Hence, it has a unique square root $(\mathcal{P}_m)^{\frac{1}{2}}$ with these properties [17, Lem. 2.4.4]. Thus

$$\sup_{\substack{\psi \in S_m \\ \psi \neq 0}} \frac{\sqrt{\langle \mathcal{P}_m \psi, \psi \rangle}}{\|\mathcal{P}_m \psi\|} = \sup_{\substack{\psi \in S_m \\ \psi \neq 0}} \frac{\|(\mathcal{P}_m)^{\frac{1}{2}} \psi\|}{\|\mathcal{P}_m \psi\|} = \sup_{\substack{\psi \in S_m \\ \psi \neq 0}} \frac{\|\psi\|}{\|(\mathcal{P}_m)^{\frac{1}{2}} \psi\|} = \sup_{\substack{\psi \in S_m \\ \psi \neq 0}} \frac{\|\psi\|}{\sqrt{\langle \mathcal{P}_m \psi, \psi \rangle}} = \frac{1}{\sqrt{c_{1,m}}}.$$

Combining this with (4.22) now gives $\sec(\theta_{n,m}) \geq \sqrt{c_{1,m}} D_{n,m}$ as required.

For the upper bound, we first notice that

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

Moreover, arguing as in (4.21) one finds that $\|\mathcal{P}_m\phi\| \leq \sqrt{c_{2,m}} \sqrt{\langle \mathcal{P}_m\phi, \phi \rangle}$. Combining this with the previous expression and using the definition of $\cos(\theta_{n,m})$ now gives the first result.

For the second part of the proof, we first recall that the Riesz basis bounds d_1, d_2 for $\{\psi_j\}_{j \in \mathbb{N}}$ are lower and upper bounds for the Riesz basis bounds $d_{1,m}, d_{2,m}$ for the finite subset $\{\psi_1, \dots, \psi_m\}$. Moreover, by [17, Thm. 5.2.1], the frame bounds $c_{1,m}, c_{2,m}$ for the Riesz basis $\{\psi_1, \dots, \psi_m\}$ are identical to the Riesz basis bounds $d_{1,m}, d_{2,m}$. This gives the second result. Finally, when $\{\psi_j\}_{j \in \mathbb{N}}$ is orthonormal we have $d_1 = d_2 = 1$, and thus we obtain the final result. \square

This lemma demonstrates that the difference in magnitudes between $\sec(\theta_{n,m})$ and $D_{n,m}$ is determined by $\sqrt{c_{1,m}}$ and $\sqrt{c_{2,m}}$. Note that $c_{2,m} \leq c_2$, where c_2 is the frame bound for the infinite frame $\{\psi_j\}_{j \in \mathbb{N}}$. However, $c_{1,m}$ can in general exhibit wild behaviour in comparison to c_1 : it is possible to construct simple frames for which $c_{1,m}$ is exponentially small in m , even though c_1 is finite [16]. On the other hand, if $\{\psi_j\}_{j \in \mathbb{N}}$ is a Riesz or orthonormal basis, we find that $D_{n,m}$ and $\sec(\theta_{n,m})$ are, up to a possible factor proportional to the constants d_1 and d_2 , equal.

4.4 Computing the generalized sampling reconstruction

Recall that the generalized sampling reconstruction $\tilde{f}_{n,m}$ depends only on \mathbb{T}_n , and not on the system of functions used to span \mathbb{T}_n . Let $\{\phi_j\}_{j=1}^{d_n}$ be a spanning set for \mathbb{T}_n , where $d_n \geq \dim(\mathbb{T}_n)$, and write

$$\tilde{f}_{n,m} = \sum_{j=1}^{d_n} \alpha_j^{[n,m]} \phi_j.$$

The vector $\alpha^{[n,m]} = \{\alpha_j^{[n,m]}\}_{j=1}^{d_n}$ is the least squares solution to the linear system

$$U^{[n,m]} \alpha^{[n,m]} = \tilde{f}^{[m]},$$

where $U^{[n,m]} \in \mathbb{C}^{m \times d_n}$ has $(j, k)^{\text{th}}$ entry $\langle \phi_k, \psi_j \rangle$. Thus, computing $\tilde{f}_{n,m}$ is equivalent to solving a least squares problem. From a numerical perspective, it is important to understand the condition number $\kappa(U^{[n,m]}) = \|U^{[n,m]}\| \|(U^{[n,m]})^\dagger\|$ of the matrix $U^{[n,m]}$, where \dagger denotes the pseudoinverse. The following lemma is similar to [5, Lem. 2.11] (for this reason we forgo the proof):

Lemma 4.10. *Let $\{\phi_j\}_{j=1}^{d_n}$ be a spanning set for \mathbb{T}_n , and write $G^{[n]} \in \mathbb{C}^{d_n \times d_n}$ for its Gram matrix. Then the condition number of the matrix $U^{[n,m]}$ satisfies*

$$\frac{1}{\sqrt{c_2} D_{n,m}} \sqrt{\kappa(G^{[n]})} \leq \kappa(U^{[n,m]}) \leq \sqrt{c_2} D_{n,m} \sqrt{\kappa(G^{[n]})}.$$

This lemma shows that the condition number of the matrix $U^{[n,m]}$ is no worse than that of the Gram matrix $G^{[n]}$ whenever m is chosen sufficiently large to ensure boundedness of $D_{n,m}$. In particular, if the vectors $\{\phi_1, \dots, \phi_n\}$ are a Riesz or orthonormal basis, then $\kappa(G^{[n]}) = \mathcal{O}(1)$ and hence the condition number of $U^{[n,m]}$ is completely determined by the magnitude of $D_{n,m}$. In this case, not only is the reconstruction $\tilde{f}_{n,m}$ numerically stable, but so is the computation of its coefficients $\alpha^{[n,m]}$. For further details on the computation of $\tilde{f}_{n,m}$, see [5].

5 The stable sampling and reconstruction rates

The key ingredient of generalized sampling is that the parameter m must be sufficiently large in comparison to n , or vice versa. The notion of how large was first quantified in [4, 5]. In this section we improve on this by using the sharp bounds of the previous section.

Definition 5.1. For $\theta \in \left(\frac{\max\{1, \sqrt{c_2}\}}{\sqrt{c_1}} \sec(\theta_{\text{TS}}), \infty \right)$, the stable sampling rate is given by

$$\Theta(n; \theta) = \min \{m \in \mathbb{N} : C(F_{n,m}) \leq \theta\}, \quad n \in \mathbb{N}. \quad (5.1)$$

The stable reconstruction rate is given by

$$\Psi(m; \theta) = \max \{n \in \mathbb{N} : C(F_{n,m}) \leq \theta\}, \quad m \in \mathbb{N}. \quad (5.2)$$

The stable sampling rate measures how large m must be for a fixed n to ensure guaranteed, stable and quasi-optimal recovery. Conversely, the stable reconstruction rate measures how large n can be for a fixed number of measurements m . Note that, by choosing either $m \geq \Theta(n; \theta)$ or $n \leq \Psi(m; \theta)$, we guarantee that the reconstruction $\tilde{f}_{n,m}$ is numerically stable and quasi-optimal, up to the magnitude of θ . Moreover, the condition $m \geq \Theta(n; \theta)$ (or $n \leq \Psi(m; \theta)$) is both sufficient and necessary to ensure stable, quasi-optimal reconstruction: if one were to sample at a rate below $\Theta(n; \theta)$ (or above $\Psi(m; \theta)$) then worse stability and convergence of the reconstruction would occur.

Recall that Remark 2.6 permits sequences of reconstruction schemes with mildly growing reconstruction constants. One can also readily define the stable sampling and reconstruction rates to reflect this. For a positive and increasing sequence $\theta = \{\theta_n\}_{n \in \mathbb{N}}$ with $\theta_n > \frac{\max\{1, \sqrt{c_2}\}}{\sqrt{c_1}} \sec(\theta_{\text{TS}})$, $\forall n$, we define

$$\Theta(n; \theta) = \min \{m \in \mathbb{N} : C(F_{n,m}) \leq \theta_n\}, \quad n \in \mathbb{N},$$

and

$$\Psi(m; \theta) = \max \{n \in \mathbb{N} : C(F_{n,m}) \leq \theta_n\}, \quad m \in \mathbb{N}.$$

Once more, one has the interpretation that sampling at the rate $m \geq \Theta(n; \theta)$ ensures stability and quasi-optimality up to the growth of θ_n .

A key property of the stable sampling and reconstruction rates is that they can be computed:

Lemma 5.2. Let $D_{n,m}$ and $\theta_{n,m}$ be as in (4.14) and Lemma 4.4 respectively, and suppose that $\{\phi_j\}_{j=1}^n$ is a basis for \mathbb{T}_n . Then the quantities $1/D_{n,m}^2$ and $\cos^2(\theta_{n,m})$ are the minimal generalized eigenvalues of the matrix pencils $\{(U^{[n,m]})^* U^{[n,m]}, A^{[n]}\}$ and $\{B^{[n,m]}, A^{[n]}\}$ respectively, where $A^{[n]}$ is the Gram matrix for $\{\phi_j\}_{j=1}^n$, $U^{[n,m]}$ is as in §4.4, $B^{[n,m]}$ is given by

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

and $C^{[m]}$ is the Gram matrix for $\{\psi_j\}_{j=1}^m$. In particular, if $\{\phi_j\}_{j=1}^n$ is an orthonormal basis for \mathbb{T}_n ,

$$D_{n,m} = \frac{1}{\sigma_{\min}(U^{[n,m]})}, \quad \sec(\theta_{n,m}) = \frac{1}{\sqrt{\lambda_{\min}(B^{[n,m]})}},$$

where $\sigma_{\min}(U^{[n,m]})$ and $\lambda_{\min}(B^{[n,m]})$ denote the minimal singular value and eigenvalue of the matrices $U^{[n,m]}$ and $B^{[n,m]}$ respectively.

Proof. The proof of this lemma is similar to that of [5, Lem. 2.13], and hence is omitted. \square

Although this lemma allows for computation of the reconstruction constant $C(F_{n,m})$ (recall that $C(F_{n,m}) = \max\{\sec(\theta_{n,m}), D_{n,m}\}$ as a result of Corollary 4.7), and therefore $\Theta(n; \theta)$ and $\Psi(m; \theta)$, it is somewhat inconvenient to have to compute both $D_{n,m}$ and $\sec(\theta_{n,m})$. The latter, in particular, can be

computationally intensive since it involves both forming and inverting the matrix $(U^{[n,m]})^* C^{[m]} U^{[n,m]}$. However, recalling the bound $C(F_{n,m}) \leq \max\{1, \sqrt{c_2}\} D_{n,m}$, we see that stability and quasi-optimality can be ensured, up to the magnitude of c_2 , by controlling the behaviour of $D_{n,m}$ only. This motivates the computationally more convenient alternative

$$\tilde{\Theta}(n; \theta) = \min \{m \in \mathbb{N} : D_{n,m} \leq \theta\}, \quad n \in \mathbb{N}, \theta \in \left(\frac{1}{\sqrt{c_1}} \sec(\theta_{\text{TS}}), \infty \right),$$

and likewise $\tilde{\Psi}(m; \theta)$. Note that setting $m \geq \tilde{\Theta}(n; \theta)$ or $n \leq \tilde{\Psi}(m; \theta)$ ensures a condition number of at worst θ and a quasi-optimality constant of at most $\max\{1, \sqrt{c_2}\} \theta$.

6 Optimality of generalized sampling

In the previous sections we gave an analysis of generalized sampling. This improved on [4, 5] by providing sharp bounds and establishing the connection between generalized sampling and certain oblique projections. In this section we address the question of optimality of generalized sampling; a topic which has not been addressed previously.

We shall consider the following problem:

Problem 6.1. *Given the m measurements $\{\langle f, \psi_j \rangle\}_{j=1}^m$ of an element $f \in \mathbb{H}$, compute a reconstruction \tilde{f} of f from the subspace \mathbb{T}_n .*

Generalized sampling provides a (perhaps the most) straightforward solution to this problem – namely, performing a least-squares fit of the data – with stability and quasi-optimality being determined by the quantities $D_{n,m}$ and $\sec(\theta_{n,m})$. An obvious question to pose is the following: can a different method outperform generalized sampling? A partial answer to this question is provided next.

6.1 Optimality amongst perfect methods

Theorem 6.2. *Suppose that $m, n \in \mathbb{N}$ are such that $D_{n,m} \neq 0$, where $D_{n,m}$ is given by (4.14). Let $G_{n,m}$ be a (linear or nonlinear) method taking measurements $\{\langle f, \psi_j \rangle\}_{j=1}^m$ and giving a reconstruction $G_{n,m}(f) \in \mathbb{T}_n$. Suppose that $G_{n,m}$ is perfect in the sense of Definition 3.9. Then, if the condition number $\kappa(G_{n,m})$ is defined as in (2.4), we have $\kappa(G_{n,m}) \geq D_{n,m}$. In particular, if $F_{n,m}$ is the generalized sampling reconstruction, then $\kappa(G_{n,m}) \geq \kappa(F_{n,m})$.*

Proof. Since $G_{n,m}$ is perfect, we have $G_{n,m}(0) = 0$. Setting $f = 0$ in (2.4), we notice that

$$\kappa(G_{n,m}) \geq \lim_{\epsilon \rightarrow 0^+} \sup_{\substack{g \in \mathbb{H} \\ \hat{g} \neq 0}} \frac{\|G_{n,m}(\epsilon g)\|}{\|\epsilon \hat{g}\|_{\ell^2}}.$$

Since $D_{n,m} \neq 0$, we have that $\hat{g} \neq 0$ for $g \in \mathbb{T}_n$ if and only if $g \neq 0$. Thus, using the perfectness of $G_{n,m}$ once more,

$$\kappa(G_{n,m}) \geq \lim_{\epsilon \rightarrow 0^+} \sup_{\substack{g \in \mathbb{T}_n \\ g \neq 0}} \frac{\|G_{n,m}(\epsilon g)\|}{\|\epsilon \hat{g}\|} = \lim_{\epsilon \rightarrow 0^+} \sup_{\substack{g \in \mathbb{T}_n \\ g \neq 0}} \frac{\|g\|}{\|\hat{g}\|} = D_{n,m},$$

as required. The second result follows from Corollary 4.7. \square

This theorem, which is embarrassingly simple to prove, states the following: any perfect method for Problem 6.1 must have a worse condition number than that of generalized sampling. We remark that perfectness is not an unreasonable assumption in practice. In particular, any method which is quasi-optimal, i.e. $\mu(G_{n,m}) < \infty$, is also perfect.

One can also generalize Theorem 6.2 somewhat to consider a larger class of methods. Indeed, let $G_{n,m}$ be a method such that

$$\|f - G_{n,m}f\| \leq \lambda \|f\|, \quad \forall f \in \mathbb{T}_n,$$

for some $\lambda \in (0, 1)$. We refer to such methods as *contractive*. Note that perfect methods are a particular example of contractive methods with $\lambda = 0$. Arguing as in the proof of Theorem 6.2, one can show that

$$\kappa(G_{n,m}) \geq (1 - \lambda)\kappa(F_{n,m}).$$

Hence, the condition number of generalized sampling can only possibly be improved by a factor of $(1 - \lambda)$ when using a contractive method.

Beside the imposition of perfectness (or contractiveness), it may appear at first sight that Theorem 6.2 is also restrictive because it deals with a worst case scenario taken over the whole of H . In practice, it may be the case that our interest does not lie with recovering all $f \in H$, but rather only those f belonging to some subspace U of H . For example, U could consist of functions with particular smoothness. One may reasonably ask: is it possible to circumvent this bound if one restricts ones interest to only a small class of functions? The answer is no. If $T_n \subseteq U$, then one can just redefine the condition number κ to be taken as a supremum over U , as opposed to H , and repeat the same argument.

It is also worth observing that Theorem 6.2 can be further weakened by considering the local condition number κ_f at a fixed $f \in H$ (we refer to (2.5) for the definition of κ_f). Indeed, it is clear from the proof that

$$\kappa_f(G_{n,m}) \geq D_{n,m} \geq \kappa_f(F_{n,m}), \quad \forall f \in T_n.$$

In some applications, typically where one's interest lies with recovering only one fixed signal f , the local condition number is arguably more important. Hence, the fact that condition numbers cannot be improved, even locally, demonstrates the importance of appropriately scaling m with n .

One can also reformulate the conclusions of Theorem 6.2 in terms of the stable sampling and reconstruction rates. To this end, suppose that $G_{n,m}$ is any quasi-optimal reconstruction method, and let $\kappa(G_{n,m})$ and $\mu(G_{n,m})$ be its condition number and quasi-optimality constant respectively. Define the reconstruction constant $C(G_{n,m}) = \max\{\kappa(G_{n,m}), \mu(G_{n,m})\}$ in the standard way, and let

$$\Theta_G(n; \theta) = \min\{m \in \mathbb{N} : C(G_{n,m}) \leq \theta\}, \quad n \in \mathbb{N},$$

be the stable sampling rate for $G_{n,m}$ (whenever defined). Theorem 6.2 trivially gives that

$$\Theta_G(n; \theta) \geq \Theta(n; \max\{1, \sqrt{c_2}\}\theta), \tag{6.1}$$

where Θ is the stable sampling rate for generalized sampling (one can make a similar statement in terms of the stable reconstruction rates Ψ_G and Ψ). This result implies the following: up to a constant on the order of $\max\{1, \sqrt{c_2}\}$, any reconstruction requires at least the same number of samples as generalized sampling to guarantee a stable quasi-optimal reconstruction with constant θ . In applications (see §7) one typically has that $\Theta(n; \theta) \sim c(\theta)g(n)$ as $n \rightarrow \infty$ for some increasing function $g(n)$ (typically $g(n) = n^\alpha$ for some $\alpha \geq 1$). Hence, (6.1) implies that no perfect method $G_{n,m}$ can have a stable sampling rate that grows at a slower asymptotic rate than that of generalized sampling, although the constant can potentially be improved whenever the sampling frame has $c_2 > 1$.

6.2 An optimality result for problems with linear stable sampling rates

Suppose that the stable sampling rate $\Theta(n; \theta)$ is linear in n for a particular example of Problem 6.1. This means that there is, up to a constant, a one-to-one correspondence between samples and reconstructed coefficients, which suggests that generalized sampling can only be outperformed by a constant factor in terms of the convergence of the reconstruction. Another method for the problem (perfect or otherwise) might obtain a slightly smaller error, but the asymptotic *rate* of convergence should be equal. This is formalized in the following theorem:

Theorem 6.3. *Let $\{\psi_j\}_{j \in \mathbb{N}}$ be a frame for H , and let $\{T_n\}_{n \in \mathbb{N}}$ a sequence of finite-dimensional subspaces satisfying (2.6) and (2.7). Suppose that the corresponding stable sampling rate $\Theta(n; \theta)$ is linear in n . Let $f \in H$ be fixed, and suppose that there exists a sequence of mappings*

$$G_m : \{\hat{f}_j\}_{j=1}^m \mapsto G_m(f) \in T_{\Psi_f(m)},$$

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

$$c_1(f)n^{-\alpha_f} \leq \|f - \mathcal{Q}_n f\| \leq c_2(f)n^{-\alpha_f}, \quad \forall n \in \mathbb{N}. \quad (6.2)$$

Then, given $\theta \in \left(\frac{\max\{1, \sqrt{c_2}\}}{\sqrt{c_1}} \sec(\theta_{\text{TS}}), \infty \right)$, there exists a constant $c_f(\theta) > 0$ such that

$$\|f - F_{\Psi(m;\theta),m}(f)\| \leq c_f(\theta)\|f - G_m(f)\|, \quad \forall m \in \mathbb{N}, \quad (6.3)$$

where $F_{n,m}$ corresponds to generalized sampling and $\Psi(m;\theta)$ is the stable reconstruction rate (5.2).

Proof. Since generalized sampling is quasi-optimal,

$$\|f - F_{\Psi(m;\theta),m}(f)\| \leq \theta \|f - \mathcal{Q}_{\Psi(m;\theta)} f\|.$$

Using (6.2) we deduce that

$$\|f - F_{\Psi(m;\theta),m}(f)\| \leq \theta \frac{c_2(f)}{c_1(f)} \left(\frac{\Psi_f(m)}{\Psi(m;\theta)} \right)^{\alpha_f} \|f - \mathcal{Q}_{\Psi_f(m)} f\|.$$

The orthogonal projection $\mathcal{Q}_n f$ is the best approximation to f from the subspace T_n . Therefore

$$\|f - F_{\Psi(m;\theta),m}(f)\| \leq \theta \frac{c_2(f)}{c_1(f)} \left(\frac{\Psi_f(m)}{\Psi(m;\theta)} \right)^{\alpha_f} \|f - G_m(f)\|.$$

The result now follows from the fact that $\Psi(m;\theta) = \mathcal{O}(m)$ and $\Psi_f(m) \leq cm$. \square

This theorem states that, in the case of a linear stable sampling rate, and for functions with algebraic decay of $\|f - \mathcal{Q}_n f\|$, generalized sampling can only be improved upon by a constant factor. As shown by (6.3), the error of generalized sampling decays at the same (or better) asymptotic rate as any other reconstruction method G_m . Note that the stipulation of algebraic convergence (6.2) is reasonable in practice. In the next section we shall see several examples for which this condition holds.

Unlike the case of generalized sampling, the method G_m in the above theorem can depend in a completely nontrivial manner on the function f . However, even with this added flexibility, this theorem shows that it is only possible to improve on generalized sampling by a constant factor. An example of such a method is an oracle. Suppose there was some method that, for a particular f satisfying (6.2), could recover the orthogonal projection $\mathcal{Q}_m f$ *exactly* (i.e. with no error) from m samples. The conclusion of the above corollary is that generalized sampling commits an error that is at worst a constant factor larger than that of this method.

Remark 6.4 The fact that the stable sampling rate is linear is key to Theorem 6.3. In situations where $\Theta(n;\theta)$ is superlinear (for an example, see the next section), it is possible to devise methods, albeit typically unstable methods, with asymptotically faster rates of convergence.

7 Uniform resampling with generalized sampling

§1–6 of this paper addressed generalized sampling in its abstract form. We now consider its application to a particular problem, the so-called *uniform resampling* (URS) problem. As we show, generalized sampling leads to an improvement over the standard approach to this problem, which results in an ill-posed discrete reconstruction.

7.1 The uniform resampling problem

In applications such as MRI, radio-astronomy and diffraction tomography [61, 49], the URS problem concerns the recovery of harmonic Fourier coefficients of a function $f \in L^2(-1, 1)^d$ from nonuniformly-spaced pointwise samples of its Fourier transform

$$\hat{f}(\omega) = 2^{-\frac{d}{2}} \int_{(-1,1)^d} f(x) e^{-i\omega\pi x} dx, \quad \omega \in \mathbb{R}.$$

This problem is important since typical sampling devices (such as MR scanners) are not best suited to acquire Fourier data in a uniform pattern (i.e. Fourier coefficients). Indeed, it is often more convenient to acquire samples along interlacing spirals or radial lines, for example (see [59] and references therein). In uniform resampling, one seeks to compute Fourier coefficients from these nonharmonic Fourier samples, and then recover the image via a standard DFT.

Consider the case $d = 1$, and let $\omega_{-n} < \omega_{-n+1} < \dots < \omega_n$ be a set of $2n+1$ nonequispaced points at which $\hat{f}(\omega)$ is sampled. The derivation of the standard URS reconstruction follows from the Shannon Sampling theorem [49, 50, 61]. Using this theorem, we have

$$\hat{f}(\omega) = \sum_{k \in \mathbb{Z}} \hat{f}(k) \text{sinc}(\omega - k), \quad \omega \in \mathbb{R},$$

where the right-hand side converges uniformly, and therefore

$$\hat{f}(\omega_j) = \sum_{k \in \mathbb{Z}} \hat{f}(k) \text{sinc}(\omega_j - k), \quad |j| \leq n. \quad (7.1)$$

Let $\alpha_k, k = -n, \dots, n$ be the values $\alpha_k \approx \hat{f}(k)$ that we seek to compute from the samples $\{\hat{f}(\omega_j)\}_{|j| \leq n}$. It is natural to truncate (7.1) at level n , leading to

$$\hat{f}(\omega_j) \approx \sum_{|k| \leq n} \alpha_k \text{sinc}(\omega_j - k), \quad |j| \leq n.$$

Let $U^{[n,n]} \in \mathbb{C}^{2n+1, 2n+1}$ be the matrix with $(j, k)^{\text{th}}$ entry $\text{sinc}(\omega_j - k)$. The URS method determines the vector $\alpha^{[n,n]} = \{\alpha_k\}_{|k| \leq n}$ as the solution to the linear system

$$U^{[n,n]} \alpha^{[n,n]} = \hat{f}^{[n]}, \quad (7.2)$$

where $\hat{f}^{[n]} = \{\hat{f}(\omega_j)\}_{|j| \leq n}$.

Suppose now that the finite set $\{\omega_j\}_{|j| \leq n}$ extends to an infinite set $\{\omega_j\}_{j \in \mathbb{Z}}$ such that the system

$$\psi_j(x) = \frac{1}{\sqrt{2}} e^{i\omega_j \pi x}, \quad j \in \mathbb{Z},$$

is a frame for $L^2(-1, 1)$. Let $\phi_j(x) = \frac{1}{\sqrt{2}} e^{ij\pi x}$, so that

$$\mathbb{T}_n = \text{span} \{\phi_j : |j| \leq n\}, \quad (7.3)$$

is the space of trigonometric polynomials of degree n . Then the URS method (7.2) is nothing more than a specific instance of the consistent sampling framework described in §3.

It has been widely reported that the URS method (7.2) may be very ill-conditioned in practice [49, 61]. Various strategies have been applied to the linear system (7.4) to try to overcome this issue, with the most common involving first manually computing a singular value decomposition and then applying standard regularization techniques from the literature on discrete ill-posed problems [49, 61]. However, this approach is both computationally expensive and remains susceptible to noise (see [50] and references therein).

Although the effect of noise can be somewhat mitigated [50], it can never truly be removed, since the underlying discrete problem is ill-posed. However, the interpretation of URS as an example of consistent sampling means that this ill-posedness – which is merely another instance of that seen in the consistent reconstructions of §3 – is completely artificial. The key theorems presented in §4 demonstrate that by replacing (7.2) with an overdetermined the least squares (i.e. generalized sampling)

$$U^{[n,m]} \alpha^{[n,n]} \approx \hat{f}^{[m]}, \quad (7.4)$$

and by increasing m suitably (or decreasing n suitably) we will be able to obtain a stable reconstruction. Hence, rather than performing regularization on a discrete ill-posed problem, we discretize differently so as to obtain a well-posed discrete problem (recall the operator-theoretic interpretation of §3.7)

Remark 7.1 Of course, Theorem 6.2 states that if the uniform resampling (7.2) is ill-conditioned (for a given n) then so is any other perfect method. In other words, there is in essence no stable way to obtain n Fourier coefficients from n nonuniformly spaced Fourier samples. Hence, increasing m (or, if m is considered fixed, decreasing n) is not just a good way to proceed, in this sense it is the only possible way to obtain a stable reconstruction.

We also remark that overdetermined least squares of the form (7.4) has been used in the past for the uniform resampling problem. However, it is still reported as resulting in an ill-conditioned problem [49, 61]. This is unsurprising in the results of this paper: m needs to not only be larger than n to ensure stability, but also above the critical threshold of the stable sampling rate $\Theta(n; \theta)$.

Remark 7.2 There are a number of alternatives to uniform resampling, such as convolutional gridding techniques [40, 46, 49, 61], which is quite popular in practice. However, URS provides an optimal solution to the problem, and consequently often provides better results [49] (in particular, it can lead to a significant decrease in artifacts [59] over gridding). Gridding has the advantage of being more efficient [49] than the standard URS algorithm. However, modifications such as the *block uniform resampling* (BURS) [49] possess improved efficiency.

7.2 Generalized sampling for the URS problem

To ensure a stable, quasi-optimal reconstruction, it is critical to determine the stable sampling rate $\Theta(n; \theta)$, or equivalently, the stable reconstruction rate $\Psi(n; \theta)$. Our main result below demonstrates that $\Theta(n; \theta)$ is linear in n for (almost) all nonuniform sampling patterns arising as Fourier frames.

First we require the following definition [31]:

Definition 7.3. A sequence $\{\omega_j\}_{j \in \mathbb{Z}}$ is a *balanced sampling sequence* if the following conditions hold:

- (i) $\Omega = \{e^{i\omega_j \pi \cdot} : j \in \mathbb{Z}\}$ is a frame for $L^2(-1, 1)$,
- (ii) $\{\omega_j\}_{j \in \mathbb{Z}}$ is δ -separated, i.e. there exists $\delta > 0$ such that $|\omega_j - \omega_k| > \delta, \forall j \neq k$,
- (iii) $\{\omega_j\}_{j \in \mathbb{Z}}$ is increasing, i.e. $\omega_j < \omega_{j+1}, \forall j \in \mathbb{Z}$,
- (iv) $\{\omega_j\}_{j \in \mathbb{Z}}$ is balanced, i.e. $\omega_j \geq 0$ if $j \geq 0$ and $\omega_j < 0$ if $j < 0$.

Note conditions (iii) and (iv) can always be guaranteed by reordering. Condition (iv) is also reasonable in practice since sampling strategies are typically symmetric. Although (ii) does not hold for all Fourier frames, we shall assume it for simplicity in the presentation that follows. It is possible in what follows to derive a fully general result on the stable sampling rate for arbitrary Fourier frames using [41, Lem. 2] (see also [31, Thm. 3]). However, for simplicity we shall not do this.

Theorem 7.4. Suppose that $\{\omega_j\}_{j \in \mathbb{Z}}$ is a balanced sampling sequence. Then the stable sampling rate $\Theta(n; \theta) = \mathcal{O}(n)$. Specifically, let $\tau : \mathbb{N} \rightarrow (0, \infty)$ be given by $\tau(m) = \min\{\omega_m, -\omega_{-m}\}$, and define $\tau^{-1} : (0, \infty) \rightarrow \mathbb{N}$ by

$$\tau^{-1}(c) = \min\{m \in \mathbb{N} : \tau(m) > c\}.$$

Then $\tau^{-1}(c) < \lceil \frac{c}{\delta} \rceil$, $\forall c > 0$, and we have the upper bound

$$\Theta(n; \theta) \leq \tau^{-1} \left(\frac{g(\theta)}{g(\theta) - 1} + \frac{g(\theta) + 1}{g(\theta) - 1} n \right),$$

where $g(\theta) = \exp \left(\frac{\pi^2 \delta}{2} (c_1 - \max\{1, c_2\} \theta^{-2}) \right)$.

Proof. Let $\mathcal{P}_m g = \sum_{|j| \leq m} \langle g, \psi_j \rangle \psi_j$, and suppose that $\phi \in \mathbb{T}_n$ is arbitrary. Then

$$\langle \mathcal{P}_m \phi, \phi \rangle = \langle \mathcal{P} \phi, \phi \rangle - \langle (\mathcal{P} - \mathcal{P}_m) \phi, \phi \rangle \geq c_1 \|\phi\|^2 - \langle (\mathcal{P} - \mathcal{P}_m) \phi, \phi \rangle. \quad (7.5)$$

Let $\phi = \sum_{|j| \leq n} \alpha_j \phi_j$ so that $\|\phi\| = \|\alpha\|_{\ell^2}$. Since

$$|\langle (\mathcal{P} - \mathcal{P}_m) f, g \rangle| \leq \sqrt{\langle (\mathcal{P} - \mathcal{P}_m) f, f \rangle} \sqrt{\langle (\mathcal{P} - \mathcal{P}_m) g, g \rangle}, \quad \forall f, g \in \mathbb{L}^2(-1, 1),$$

it follows that

$$\begin{aligned} \langle (\mathcal{P} - \mathcal{P}_m) \phi, \phi \rangle &= \sum_{|j|, |k| \leq n} \alpha_j \bar{\alpha}_k \langle (\mathcal{P} - \mathcal{P}_m) \phi_j, \phi_k \rangle \\ &\leq \left(\sum_{|j| \leq n} |\alpha_j| \sqrt{\langle (\mathcal{P} - \mathcal{P}_m) \phi_j, \phi_j \rangle} \right)^2 \\ &\leq \|\phi\|^2 \sum_{|j| \leq n} \langle (\mathcal{P} - \mathcal{P}_m) \phi_j, \phi_j \rangle. \end{aligned} \quad (7.6)$$

Let us suppose that m is sufficiently large so that $|\omega_j| > n$ for $|j| > m$ (i.e. $m > \frac{n}{\delta}$). Note that

$$|\langle \phi_k, \psi_j \rangle| \leq \frac{1}{\pi |\omega_j - k|}, \quad |j| > m, \quad |k| \leq n.$$

Therefore

$$\langle (\mathcal{P} - \mathcal{P}_m) \phi_k, \phi_k \rangle \leq \sum_{|j| > m} \frac{1}{\pi^2 |\omega_j - k|^2} = \frac{1}{\pi^2} \sum_{j > m} \frac{1}{|\omega_j - k|^2} + \frac{1}{\pi^2} \sum_{j < -m} \frac{1}{|\omega_{-j} - k|^2}.$$

Consider the first sum. We have

$$\sum_{j > m} \frac{1}{|\omega_j - k|^2} \leq \frac{1}{\delta} \int_{\omega_m}^{\infty} \frac{1}{(\omega - k)^2} d\omega = \frac{1}{\delta(\omega_m - k)}.$$

Using a similar estimate for the other sum, we obtain

$$\langle (\mathcal{P} - \mathcal{P}_m) \phi_k, \phi_k \rangle \leq \frac{1}{\pi^2 \delta} \left(\frac{1}{\omega_m - k} + \frac{1}{k - \omega_{-m}} \right).$$

Substituting this into (7.6), we obtain

$$\langle (\mathcal{P} - \mathcal{P}_m) \phi, \phi \rangle \leq \frac{\|\phi\|^2}{\pi^2 \delta} \sum_{|k| \leq n} \left(\frac{1}{\omega_m - k} + \frac{1}{k - \omega_{-m}} \right).$$

Notice that

$$\sum_{|k| \leq n} \frac{1}{\omega_m - k} \leq \int_{-n}^{\omega_m + n} \frac{1}{\omega_m - x} dx = \ln \left(\frac{\omega_m + n}{\omega_m - n - 1} \right).$$

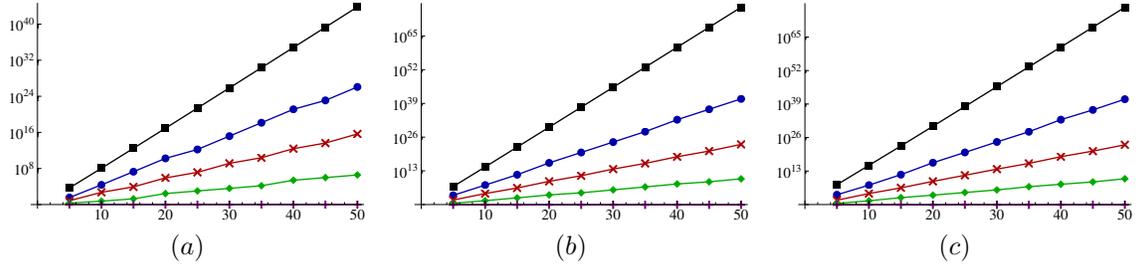


Figure 2: The quantity $D_{n,cn}$ against n for the generalized sampling applied to the uniform resampling problem for the frames (a)–(c). The values $c = 1, \frac{5}{4}, \frac{3}{2}, \frac{7}{4}, 2$ were used for (a) and $c = 1, \frac{7}{4}, \frac{5}{2}, \frac{13}{4}, 4$ for (b) and (c).

Likewise

$$\sum_{|k| \leq n} \frac{1}{k - \omega_{-m}} \leq \int_{-n-1}^n \frac{1}{x - \omega_{-m}} dx = \ln \left(\frac{n - \omega_{-m}}{-\omega_{-m} - n - 1} \right).$$

Hence

$$\langle (\mathcal{P} - \mathcal{P}_m)\phi, \phi \rangle \leq \frac{2\|\phi\|^2}{\pi^2\delta} \ln \left(\frac{\tau(m) + n}{\tau(m) - n - 1} \right).$$

Combining this with (7.5), we obtain

$$D_{n,m}^{-2} = \inf_{\substack{\phi \in \mathbb{T}_n \\ \|\phi\|=1}} \langle \mathcal{P}_m \phi, \phi \rangle \geq c_1 - \frac{2}{\pi^2\delta} \ln \left(\frac{\tau(m) + n}{\tau(m) - n - 1} \right).$$

Recall that $C(F_{n,m}) \leq \max\{1, \sqrt{c_2}\} D_{n,m}$. Hence $C(F_{n,m}) < \theta$ provided

$$c_1 - \frac{2}{\pi^2\delta} \ln \left(\frac{\tau(m) + n}{\tau(m) - n - 1} \right) > \max\{1, c_2\} \theta^{-2}.$$

Rearranging, we obtain

$$\frac{\tau(m) + n}{\tau(m) - n - 1} < g(\theta) \iff \tau(m) > \frac{g(\theta)}{g(\theta) - 1} + \frac{g(\theta) + 1}{g(\theta) - 1} n,$$

and this gives the result (note that this condition implies that $|\omega_j| > n$, $|j| > m$, which was the assumption made for the above analysis). Note also that $\tau(m) = \min\{\omega_m, -\omega_{-m}\} \geq \delta m$. Thus $\tau^{-1}(c) \leq \lceil \frac{c}{\delta} \rceil$. This completes the proof. \square

7.3 Numerical results

We now give numerical results for generalized sampling for applied to this problem. We shall consider the following three sequences

$$(a) : \omega_j = \frac{1}{2}j, \quad (b) : \omega_j = \frac{1}{4}j, \quad (c) : \omega_j = \frac{1}{4}j + \nu_j,$$

where in the last case $\nu_j \in (-\frac{1}{5}, \frac{1}{5})$ is chosen uniformly at random. Note that all three sequences are frames for $L^2(-1, 1)$ [17].

In Figure 2 we plot the quantity $D_{n,m}$ with various linear scalings of m with n . As is evident, this constant is exponentially large when $m = n$ (i.e. consistent sampling), and it remains exponentially large when $m = cn$ for small c below a certain threshold. However, as c increases the rate of exponential growth decreases, and once c is sufficiently large, there is no exponential growth at all.

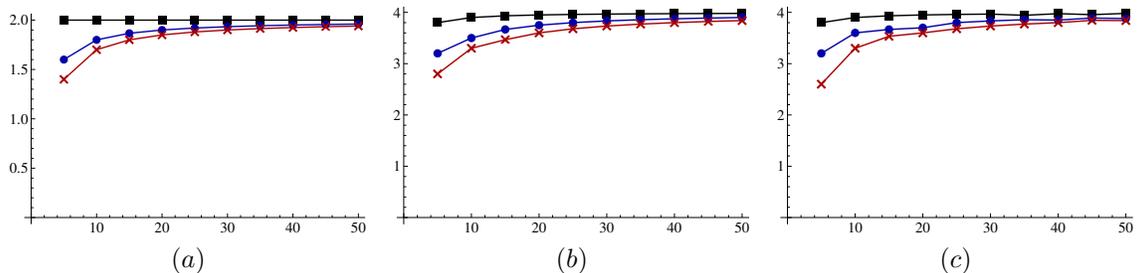


Figure 3: The stable sampling rate $\tilde{\Theta}(n; \theta)$, scaled by n^{-1} , for the frames (a)–(c), where $\theta = \frac{5}{4}, 10, 50$.

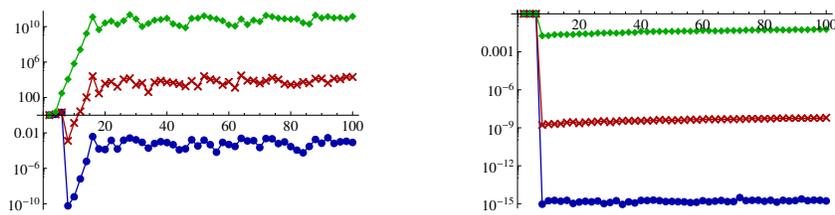


Figure 4: The error $\|f - \tilde{f}_{n,m}\|$ against m where $n = m$ (left) and $n = \frac{1}{2}m$ (right), $f(x) = \frac{1}{\sqrt{2}}e^{8i\pi x}$, and $\tilde{f}_{n,m}$ is computed from noisy data $\{\hat{f}_j + \eta_j\}_{|j| \leq n}$, where $|\eta_j| \leq \eta$ is chosen uniformly at random with $\eta = 0, 10^{-9}, 10^{-2}$ (circles, crosses and diamonds respectively). The sampling frame (a) was used.

To determine the critical c for which the reconstruction constant is bounded, we compute the stable sampling rate. This is shown in Figure 3. For the frame (a) this critical value is roughly 2, whereas for (b) and (c) it is approximately 4. Moreover, the closeness of the graphs indicates that one only needs to exceed this critical value by a very small amount to get an extremely good reconstruction constant. This is unsurprising. For (a) and (b) at least, the first n harmonic Fourier coefficients are contained within the given measurements precisely when $m = 2n$ for (a) and $m = 4n$ for (b).

To illustrate the effectiveness of generalized sampling for this problem, in Figure 4 we consider the reconstruction from noisy data. As is evident, when $m = n$, noise is amplified by around 10^{15} . However, double oversampling, as suggested in Figure 3, renders the reconstruction completely stable.

7.4 Alternatives to uniform resampling

The aim of uniform resampling is to recover the Fourier coefficients of the unknown function f from its nonuniform Fourier samples. However, it is well known that images and signals (which are typically nonperiodic) are poorly represented by their Fourier series. In particular, the finite series converges slowly and is polluted by Gibbs oscillations.

However, there is no reason besides familiarity to actually compute Fourier coefficients from nonuniform Fourier samples. With generalized sampling one is able to reconstruct in *any* subspace T_n ; in particular, one which is better suited to the particular function. Thus, provided such a subspace is known, we are able to obtain a better reconstruction over the classical Fourier series.

In this final section we consider briefly two alternative choices for T_n besides the Fourier space (7.3). The first is a spline space of piecewise polynomial functions of fixed degree $d \in \mathbb{N}$:

$$T_n = \left\{ \phi \in C^{d-1}[-1, 1] : \phi|_{\left[\frac{j}{n}, \frac{j+1}{n}\right)} \in \mathbb{P}^d, \quad j = -n, \dots, n-1 \right\}, \quad n \in \mathbb{N}. \quad (7.7)$$

Note that the sequence of orthogonal projections $\mathcal{Q}_n f$ of a function $f \in C^d[-1, 1]$ converge to f at the rate n^{-d-1} , without the assumption of periodicity of f . Conversely, the Fourier series (i.e. the URS

n	8	16	32	64	n	8	16	32	64
(a)	8.33e12	7.24e25	2.17e52	3.40e105	(a)	3.14e11	6.68e23	2.25e49	2.16e104
(b)	1.51e5	3.14e12	1.37e25	2.69e51	(b)	4.95e4	3.84e10	5.32e22	1.41e48
(c)	2.54e6	9.53e11	1.65e23	3.98e45	(c)	2.29e2	4.19e4	2.11e9	1.30e19

Table 1: The quantities $D_{n,n}$ (left) and $D_{n,2n}$ (right) for the spaces (a): (7.7) ($d = 2$), (b): (7.7) ($d = 4$) and (c): (7.8). The sampling frame (c) was used.

m	Fourier	splines ($d = 2$)	splines ($d = 4$)	polynomials
32	2.80e-2	3.64e-3	5.95e-4	1.42e-2
64	1.49e-4	4.15e-4	3.05e-5	3.23e-3
128	1.04e-11	4.89e-5	4.76e-7	3.01e-5
256	3.30e-15	6.04e-6	1.33e-8	5.43e-8
512	4.06e-15	7.55e-7	4.10e-10	5.02e-14
1024	4.79e-15	9.45e-8	1.27e-11	4.86e-14

Table 2: The error $\|f - \tilde{f}_{n,m}\|$ for the smooth and periodic function $f(x) = \sin 3\pi x + 2e^{\frac{20}{\pi^2}(\cos 2\pi x - 4 \cos \pi x - 5)}$, where the reconstruction space T_n is the Fourier space (7.3), the spline space (7.7) with $d = 2, 4$, or the polynomial space (7.8). The parameter $n = \Psi(m; 4)$, and the sampling frame (c) was used.

reconstruction with T_n given by (7.3) converges like $n^{-\frac{1}{2}}$ when f is nonperiodic. Hence, the spaces T_n are better suited for moderately smooth and nonperiodic functions.

The second choice for T_n is the polynomial space

$$T_n = \mathbb{P}_n. \tag{7.8}$$

Observe that if f is smooth, i.e. $f \in C^\infty[-1, 1]$, then $\mathcal{Q}_n f$ converges faster than any power of n^{-1} . Hence, this space is particularly well suited for smooth functions. Note that the use of this space for uniform Fourier samples $\omega_j = j$ was extensively discussed in [5].

As one might expect, both the spaces (7.7) and (7.8) lead to instability in the corresponding consistent reconstruction. This is shown in Table 1. In both cases, the constant $D_{n,n}$ is exponentially large in n . Nonetheless, such instability can be overcome by sampling according to the stable sampling rate. Although we shall not do it in this paper (for the sake of brevity), it is possible to prove that the stable sampling rate is linear $\Theta(n; \theta) = \mathcal{O}(n)$ for the spaces (7.7) for any fixed $d \in \mathbb{N}$, and quadratic $\Theta(n; \theta) = \mathcal{O}(n^2)$ for (7.8). Note that a similar result for the latter in the case of uniform Fourier samples was shown previously in [5] and [39].

Instead, we now illustrate the advantage gained from exploiting these different reconstruction spaces. In Tables 2 and 3 we give numerical results for the three different spaces considered. In each case, the parameter m (the number of samples) was fixed and n chosen according to the stable reconstruction rate $\Psi(m; \theta)$ with $\theta = 4$. Thus all three reconstruction use the same data, and are guaranteed to have the same condition number. As can be seen in Table 2, the Fourier space (7.3) is particularly well suited for periodic functions, and outperforms both the spline (7.7) and polynomial (7.8) spaces. However, the situation changes completely when the function to be reconstructed is not periodic. In Table 3 we see that the polynomial space (7.8) gives the best reconstruction, followed by the spline spaces (7.7). The URS reconstruction, which uses the Fourier space (7.3), suffers from the Gibbs phenomenon and thus exhibits only low accuracy.

7.5 On optimality

In view of the numerical results for $D_{n,m}$ (Figure 2), Theorem 6.2 demonstrates that any perfect method for the uniform resampling problem will be exponentially unstable, unless the stable sampling rate is adhered to. Moreover, since the stable sampling rate is linear in this case (Theorem 7.4), Theorem 6.3

m	Fourier	splines ($d = 2$)	splines ($d = 4$)	polynomials
32	3.13e-1	3.12e-2	4.72e-3	3.62e-2
64	1.53e-1	4.79e-3	1.80e-4	1.56e-3
128	8.65e-2	5.16e-4	1.31e-5	1.79e-6
256	9.27e-2	5.96e-5	4.84e-7	4.41e-11
512	6.21e-2	7.14e-6	1.32e-8	4.33e-14
1024	2.50e-2	8.82e-7	3.94e-10	4.19e-14

Table 3: The error $\|f - \tilde{f}_{n,m}\|$ for the smooth function $f(x) = \sin 10x + 2e^{20(x^2-1)}$, where the reconstruction space T_n is the Fourier space (7.3), the spline space (7.7) with $d = 2, 4$, or the polynomial space (7.8). The parameter $n = \Psi(m; 4)$, and the sampling frame (c) was used.

also applies. We conclude that, for periodic functions of finite smoothness (i.e. functions for which (6.2) holds), one cannot outperform generalized sampling by more than a constant factor.

The same conclusions also hold in the case of the spline spaces (7.7), in light of the numerics in Table 1. For the polynomial space (7.8), however, Theorem 6.3 does not apply, since the stable sampling rate is quadratic. Hence, it is in theory possible to outperform generalized sampling in terms of the asymptotic rate of convergence. Nonetheless, it transpires that this cannot be done in this case without compromising stability. For a more thorough analysis of stability and convergence for this reconstruction problem we refer the reader to [7].

Acknowledgements

The authors would like to thank Yonina Eldar, Hagai Kirschner, Nilima Nigam, Alexei Shadrin and Michael Unser for their helpful discussions and comments. AH acknowledges support from a Royal Society University Research Fellowship as well as UK Engineering and Physical Sciences Research Council (EPSRC) grant EP/L003457/1. CP acknowledges support from the EPSRC grant EP/H023348/1 for the University of Cambridge Centre for Doctoral Training, the Cambridge Centre for Analysis.

References

- [1] B. Adcock and A. C. Hansen. Generalized sampling and infinite-dimensional compressed sensing. *Technical report NA2011/02, DAMTP, University of Cambridge*, 2011.
- [2] B. Adcock and A. C. Hansen. Generalized sampling and the stable and accurate reconstruction of piecewise analytic functions from their Fourier coefficients. *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. A generalized sampling theorem for stable reconstructions in arbitrary bases. *J. Fourier Anal. Appl.*, 18(4):685–716, 2012.
- [5] B. Adcock and A. C. Hansen. Stable reconstructions in Hilbert spaces and the resolution of the Gibbs phenomenon. *Appl. Comput. Harmon. Anal.*, 32(3):357–388, 2012.
- [6] B. Adcock, A. C. Hansen, E. Herrholz, and G. Teschke. Generalized sampling: extension to frames and inverse and ill-posed problems. *Inverse Problems*, 29(1):015008, 2013.
- [7] B. Adcock, A. C. Hansen, and A. Shadrin. A stability barrier for reconstructions from Fourier samples. *Preprint*, 2012.
- [8] A. Aldroubi. Oblique projections in atomic spaces. *Proc. Amer. Math. Soc.*, 124(7):2051–2060, 1996.
- [9] 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.

- [10] 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.
- [11] D. Buckholtz. Hilbert space idempotents and involutions. *Proc. Amer. Math. Soc.*, 128:1415–1418, 1999.
- [12] E. Candès and D. L. Donoho. Recovering edges in ill-posed inverse problems: optimality of curvelet frames. *Ann. Statist.*, 30(3):784–842, 2002.
- [13] E. J. Candès. An introduction to compressive sensing. *IEEE Signal Process. Mag.*, 25(2):21–30, 2008.
- [14] E. J. Candès and D. Donoho. New tight frames of curvelets and optimal representations of objects with piecewise C^2 singularities. *Comm. Pure Appl. Math.*, 57(2):219–266, 2004.
- [15] E. J. Candès and C. Fernandez-Granda. Towards a mathematical theory of super-resolution. *Comm. Pure Appl. Math. (to appear)*, 2012.
- [16] O. Christensen. Frames and the projection method. *Appl. Comput. Harmon. Anal.*, 1:50–53, 1993.
- [17] O. Christensen. *An Introduction to Frames and Riesz Bases*. Birkhauser, 2003.
- [18] S. Dahlke, G. Kutyniok, P. Maass, C. Sagiv, H.-G. Stark, and G. Teschke. The uncertainty principle associated with the continuous shearlet transform. *Int. J. Wavelets Multiresolut. Inf. Process.*, 6(2):157–181, 2008.
- [19] S. Dahlke, G. Kutyniok, G. Steidl, and G. Teschke. Shearlet coorbit spaces and associated banach frames. *Appl. Comput. Harmon. Anal.*, 27(2):195–214, 2009.
- [20] M. N. Do and M. Vetterli. The contourlet transform: An efficient directional multiresolution image representation. *IEEE Transactions on Image Processing*, 14(12):2091–2106, 2005.
- [21] D. L. Donoho. Compressed sensing. *IEEE Trans. Inform. Theory*, 52(4):1289–1306, 2006.
- [22] 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.
- [23] T. Dvorkind and Y. C. Eldar. Robust and consistent sampling. *IEEE Signal Process. Letters*, 16(9):739–742, 2009.
- [24] Y. C. Eldar. Sampling with arbitrary sampling and reconstruction spaces and oblique dual frame vectors. *J. Fourier Anal. Appl.*, 9(1):77–96, 2003.
- [25] Y. C. Eldar. Sampling without input constraints: Consistent reconstruction in arbitrary spaces. *Sampling, Wavelets and Tomography*, 2003.
- [26] Y. C. Eldar and T. Dvorkind. A minimum squared-error framework for generalized sampling. *IEEE Trans. Signal Process.*, 54(6):2155–2167, 2006.
- [27] Y. C. Eldar and G. Kutyniok, editors. *Compressed Sensing: Theory and Applications*. Cambridge University Press, 2012.
- [28] Y. C. Eldar and T. Michaeli. Beyond Bandlimited Sampling. *IEEE Signal Process. Mag.*, 26(3):48–68, 2009.
- [29] Y. C. Eldar and T. Werther. General framework for consistent sampling in Hilbert spaces. *Int. J. Wavelets Multiresolut. Inf. Process.*, 3(3):347, 2005.
- [30] M. Fornasier and H. Rauhut. Compressive sensing. In *Handbook of Mathematical Methods in Imaging*, pages 187–228. Springer, 2011.
- [31] A. Gelb and T. Hines. Recovering exponential accuracy from non-harmonic fourier data through spectral reprojection. *J. Sci. Comput.*, 51(158–182), 2012.
- [32] K. Gröchenig, Z. Rzesotnik, and T. Strohmer. Quantitative estimates for the finite section method and Banach algebras of matrices. *Integral Equations and Operator Theory*, 67(2):183–202, 2011.
- [33] 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.
- [34] 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.
- [35] A. C. Hansen. On the approximation of spectra of linear operators on Hilbert spaces. *J. Funct. Anal.*, 254(8):2092–2126, 2008.

- [36] D. M. Healy and J. B. Weaver. Two applications of wavelet transforms in Magnetic Resonance Imaging. *IEEE Trans. Inform. Theory*, 38(2):840–862, 1992.
- [37] 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.
- [38] A. Hirabayashi and M. Unser. Consistent sampling and signal recovery. *IEEE Trans. Signal Process.*, 55(8):4104–4115, 2007.
- [39] T. Hrycak and K. Gröchenig. Pseudospectral Fourier reconstruction with the modified inverse polynomial reconstruction method. *J. Comput. Phys.*, 229(3):933–946, 2010.
- [40] J. I. Jackson, C. H. Meyer, D. G. Nishimura, and A. Macovski. Selection of a convolution function for Fourier inversion using gridding. *IEEE Trans. Med. Imaging*, 10:473–478, 1991.
- [41] S. Jaffard. A density criterion for frames of complex exponentials. *Mich. Math. J.*, 38(3):339–348, 1991.
- [42] A. J. Jerri. The Shannon sampling theorem – its various extensions and applications: A tutorial review. *Proc. IEEE*, 65(1565–1596), 1977.
- [43] G. Kutyniok, J. Lemvig, and W.-Q. Lim. Compactly supported shearlets. In M. Neamtu and L. Schumaker, editors, *Approximation Theory XIII: San Antonio 2010*, volume 13 of *Springer Proceedings in Mathematics*, pages 163–186. Springer New York, 2012.
- [44] A. F. Laine. Wavelets in temporal and spatial processing of biomedical images. *Annu. Rev. Biomed. Eng.*, 02:511–550, 2000.
- [45] M. Lindner. *Infinite Matrices and their Finite Sections*. Frontiers in Mathematics. Birkhäuser Verlag, Basel, 2006. An introduction to the limit operator method.
- [46] C. H. Meyer, B. S. Hu, D. G. Nishimura, and A. Macovski. Fast spiral coronary artery imaging. *Magn. Reson. Med.*, 28:202–213, 1992.
- [47] 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.
- [48] D. D.-Y. Po and M. N. Do. Directional multiscale modeling of images using the contourlet transform. *Trans. Img. Proc.*, 15(6):1610–1620, June 2006.
- [49] D. Rosenfeld. An optimal and efficient new gridding algorithm using singular value decomposition. *Magn. Reson. Med.*, 40(1):14–23, 1998.
- [50] D. Rosenfeld. New approach to gridding using regularization and estimation theory. *Magn. Reson. Med.*, 48(1):193–202, 2002.
- [51] J. Steinberg. Oblique projections in Hilbert spaces. *Integr. equ. oper. theory*, 38(1):81–119, 2000.
- [52] D. Szyld. The many proofs of an identity on the norm of oblique projections. *Numer. Algorithms*, 42:309–323, 2006.
- [53] W.-S. Tang. Oblique projections, biorthogonal Riesz bases and multiwavelets in Hilbert spaces. *Proc. Amer. Math. Soc.*, 128(2):463–473, 1999.
- [54] L. N. Trefethan and D. Bau III. *Numerical Linear Algebra*. SIAM, Philadelphia, 1997.
- [55] M. Unser. Sampling—50 years after Shannon. *Proc. IEEE*, 88(4):569–587, 2000.
- [56] M. Unser and A. Aldroubi. A general sampling theory for nonideal acquisition devices. *IEEE Trans. Signal Process.*, 42(11):2915–2925, 1994.
- [57] M. Unser and J. Zerubia. Generalized sampling: stability and performance analysis. *IEEE Trans. Signal Process.*, 45(12):2941–2950, 1997.
- [58] M. Unser and J. Zerubia. A generalized sampling theory without band-limiting constraints. *IEEE Trans. Circuits Syst. II.*, 45(8):959–969, 1998.
- [59] R. Van de Walle, H. Barrett, K. Myers, M. Altbach, B. Desplanques, A. Gmitro, J. Cornelis, and I. Lemahieu. Reconstruction of MR images from data acquired on a general noregular grid by pseudoinverse calculation. *IEEE Trans. Med. Imaging*, 19(12):1160–1167, 2000.
- [60] M. Vetterli, P. Marziliano, and T. Blu. Sampling signals with finite rate of innovation. *IEEE Trans. Signal Process.*, 50(6):1417–1428, 2002.
- [61] A. Viswanathan, A. Gelb, D. Cochran, and R. Renaut. On reconstructions from non-uniform spectral data. *J. Sci. Comput.*, 45(1–3):487–513, 2010.
- [62] J. B. Weaver, Y. Xu, D. M. Healy, and J. R. Driscoll. Wavelet-encoded MR imaging. *Magn. Reson. Med.*, 24:275–287, 1992.