

Generalized Sampling: Extension to Frames and Inverse and Ill-Posed Problems

Ben Adcock^{*} Anders C. Hansen[†] Evelyn Herrholz[‡] Gerd Teschke[‡]

Abstract

Generalized sampling is new framework for sampling and reconstruction in infinite-dimensional Hilbert spaces. Given measurements (inner products) of an element with respect to one basis, it allows one to reconstruct in another, arbitrary basis, in a way that is both convergent and numerically stable. However, generalized sampling is thus far only valid for sampling and reconstruction in systems that comprise bases. Thus, in the first part of this paper we extend this framework from bases to frames, and provide fundamental sampling theorems for this more general case. The second part of the paper is concerned with extending the idea of generalized sampling to the solution of inverse and ill-posed problems. In particular, we introduce two generalized sampling frameworks for such problems, based on regularized and non-regularized approaches. We furnish evidence of the usefulness of the proposed theories by providing a number of numerical experiments.

1 Introduction

A vital task in applied mathematics and engineering is the recovery of an object – a signal or image, for example – from a collection of its samples (or measurements). This problem, which lies at the heart of modern sampling theory, can be modelled in a Hilbert space \mathcal{H} , with the samples of the unknown $f \in \mathcal{H}$ being of the form

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

for some system $\{s_j\}_{j \in \mathbb{N}} \subset \mathcal{H}$. In many important cases, one encounters the situation where the system $\{s_j\}_{j \in \mathbb{N}}$ is *fixed*. Also, one usually only has access to a finite number of measurements, say

$$\langle f, s_j \rangle, \quad j = 1, \dots, m.$$

Thus, the computational problem is to recover f from a fixed, and finite, collection of its measurements.

Consider, for example, the Magnetic Resonance Imaging (MRI) problem of reconstructing an image from pointwise samples of its Fourier transform. Here $\mathcal{H} = L_2([-1, 1])$ (in the 1D setting) and $s_j(t) = e^{\pi i \rho(j)t}$, where $\rho : \mathbb{N} \rightarrow \mathbb{Z}$ is the standard re-indexing function. The particular design of the MRI scanner means that the sampling scheme cannot be easily altered. Moreover, it is infeasible in practice to take too many measurements, meaning that the parameter m may well be small. Of course, reconstructions from Fourier samples are precisely the setting of arguably the most classical result in sampling theory, the celebrated Shannon Sampling Theorem [21, 24]. However, whilst this theorem guarantees a reconstruction of f via its Fourier series, this is often an extremely poor way to proceed. Indeed, one typically requires an intolerably large number of samples m to recover f to any reasonable accuracy [13, 24]. Other effects, such as the Gibbs phenomenon, can also be highly problematic.

This concern aside, note that the MRI problem is fundamentally well posed: it involves the inversion of the Fourier transform. However, in applications, one may also encounter considerably more challenging situations. In particular, it may well be the case that f is the solution of $Af = g$, where A is some bounded linear operator, and, rather than measurements of f , we are actually given measurements of g . Thus, to recover f , one has to invert A , making the Shannon Sampling Theorem (for example) no longer directly

^{*}Department of Mathematics, Simon Fraser University, Burnaby, BC V5A 1S6, Canada.

[†]DAMTP, Centre for Mathematical Sciences, University of Cambridge, Wilberforce Rd, Cambridge CB3 0WA, United Kingdom.

[‡]Institute for Computational Mathematics in Science and Technology, Neubrandenburg University of Applied Sciences, Brodaer Straße 2, 17033 Neubrandenburg, Germany.

applicable. Another concern is that the operator A may well in practice be ill-posed. This is the case for the problem of X-Ray tomography, where one is tasked with inverting the famously ill-posed Radon transform.

With this in mind, the purpose of this paper is to study novel approaches for these types of reconstruction problems, based on the framework of *generalized sampling (GS)* introduced by Adcock & Hansen [6, 2, 3, 4]. Specifically, we provide two generalizations of this framework: in the first part of the paper we augment this framework to systems $\{s_j\}_{j \in \mathbb{N}}$ that comprise frames (as opposed to previously studied case of Riesz bases), and in the second part we present an extension of GS to the solution of inverse and ill-posed problems.

We explain GS in more detail in the forthcoming section. Before doing so, let us first mention one point that is essential to what follows. Given that we only have access to a finite number of measurements of f (or g), it is quite common to treat the reconstruction problem as finite dimensional, i.e. living in a finite-dimensional vector space. However, as we explain in due course, there are a number of potential pitfalls of ignoring the true infinite dimensionality of the problem. In particular, one may well end up with reconstructions that are neither stable nor convergent (in the sense $m \rightarrow \infty$). Fortunately, as we explain, GS allows one to work directly with the infinite-dimensional model, whilst giving stable, convergent numerical methods.

When introducing new sampling techniques (such as those we present in this paper) it is natural to ask the following question: can such one exploit sparsity in order to subsample? In many cases, the answer turns out to be yes. In [1, 19] (see also [5]), a new theory of compressed sensing for infinite-dimensional problems was introduced, based directly on ideas from GS. This development relates directly to the point made above: current compressed sensing theory (and techniques) are based largely on finite-dimensional models, which are not sufficient in all circumstances. Having said this, the focus of this paper will be on classical (i.e. non-sparse) sampling. In Section 5 we discuss how to extend and combine the sampling theorems introduced in this paper with infinite-dimensional compressed sensing techniques.

The outline of the remainder of this paper is as follows. In Section 2 we describe the problems we consider in more detail and recap GS. Section 3 extends GS to the setting of frames. In Section 4 we discussed inverse and ill-posed problems.

2 Generalized sampling

2.1 The Problems

The purpose of this paper is to study the following recovery problems. We are given a signal $f \in \mathcal{H}$, where \mathcal{H} is an infinite-dimensional separable Hilbert space. We are also given a *sampling system* $\{s_k\}_{k \in \mathbb{N}} \subset \mathcal{H}$ and a bounded linear operator A on \mathcal{H} , where A may be compact. The problems are as follows:

- (i) Recover a good approximation to f from the samples $\{\langle f, s_k \rangle\}_{k=1}^m$, for some $m \in \mathbb{N}$.
- (ii) Recover a good approximation to f from the samples $\{\langle g, s_k \rangle\}_{k=1}^m$, for some $m \in \mathbb{N}$, where $g = Af$.

We shall also consider the case where the samples are contaminated by noise.

As mentioned, a basic example of Problem (i) is the recovery of $f \in L_2([-1, 1])$ from its Fourier coefficients. In this case, the sampling system consists of the usual complex exponentials. Of course, one could construct the Fourier series approximation $f_m = \sum_{k=1}^m \langle f, s_k \rangle s_k$, knowing that $f_m \rightarrow f$ as $m \rightarrow \infty$. Yet, as discussed, such convergence is often in practice intolerably slow.

An alternative in this case is to proceed as follows. Suppose that we know that f has a ‘nice’ expansion in another system $\{w_j\}_{j \in \mathbb{N}} \subset \mathcal{H}$. In other words, there exists $\alpha_j \in \mathbb{C}$ such that

$$\sum_{j=1}^n \alpha_j w_j \rightarrow f,$$

rapidly as $n \rightarrow \infty$. In the above problem of Fourier sampling it is easy to think that this may well be the case. For example, the w_k could be wavelets of some variety, or more exotic objects such as curvelets, shearlets or contourlets. Now, had we access to $\{\alpha_j\}_{j=1}^n$ we could have recovered f to high accuracy. However, we only know $\{\langle f, s_j \rangle\}_{j=1}^m$. Nonetheless, given the additional information that f is well represented in $\{w_j\}_{j \in \mathbb{N}}$, we can now ask the following question: how do we obtain the values $\{\alpha_j\}_{j=1}^n$ (or some approximation thereof) from the given measurements?

This question is more subtle than it appears. As we explain in the next section, the most straightforward approach is not guaranteed to succeed.

2.2 Consistent Reconstructions

Suppose now that $\{w_k\}_{k \in \mathbb{N}}$ are orthonormal vectors in a Hilbert space \mathcal{H} , and that $\{s_k\}_{k \in \mathbb{N}}$ is an orthonormal basis of \mathcal{H} . Let $f \in \mathcal{H}$ be given by

$$f = \sum_{j=1}^m \alpha_j w_j, \quad \alpha_j = \langle f, w_j \rangle, \quad (2.1)$$

where $\langle \cdot, \cdot \rangle$ is the inner product on \mathcal{H} . Suppose that we can access samples $\langle f, s_k \rangle$ for $k \in \mathbb{N}$. It is then tempting to form

$$U_m = \begin{pmatrix} u_{11} & \dots & u_{1m} \\ \vdots & \ddots & \vdots \\ u_{m1} & \dots & u_{mm} \end{pmatrix}, \quad u_{ij} = \langle w_j, s_i \rangle,$$

and approach the problem of obtaining the α_j s by solving the linear system of equations

$$U_m x = y_m, \quad y_m = \{\langle f, s_1 \rangle, \dots, \langle f, s_m \rangle\}. \quad (2.2)$$

This approach is often referred to as a *consistent reconstruction* technique (i.e. the first m measurements of the reconstruction $\tilde{f} = \sum_{k=1}^m x_k w_k$ coincide with those of f). Introduced (in the context of sampling) by Unser & Aldroubi [25, 26, 7], and later generalized significantly by Eldar et al [10, 12, 11, 14], this technique is quite widely used in applications [24]. Note in the case above that f has only m nonzero coefficients in the basis $\{w_j\}_{j \in \mathbb{N}}$. Thus, ignoring any noise or numerical issues, the solution x of (2.2) will coincide with $\{\alpha_1, \dots, \alpha_m\}$, and thus f should, at least theoretically, be recovered *perfectly* by (2.2).

Let us now consider an example of this approach:

Example 2.1. Suppose that the w_k s are the Haar wavelets on $[-1/2, 1/2]$, and take $m = 769$. Let $\langle f, s_k \rangle$ be Fourier samples of f , i.e. $s_k = e^{2\pi i \rho(k)}$, where $\rho : \mathbb{N} \rightarrow \mathbb{N}$ is defined by $\rho(1) = 0, \rho(2) = 1, \rho(3) = -1, \rho(4) = 2, \rho(5) = -2 \dots$ (ρ is merely a re-indexing function).

If we form the matrix U_m , then, although U_m is invertible, and in particular the problem (2.2) has a unique solution, one finds that its condition number is intolerably large (in this instance it is $\approx 4.62 \times 10^{16}$). Thus, if the observations y_m are perturbed by $\delta_m \in \mathbb{C}^m$, and we solve $U_m x = y_m + \delta_m$, then we expect a catastrophic failure. Indeed, this is the case even in the simplest setting. For $f = \sum_{j=1}^m \alpha_j w_j$, where $\alpha_1 = \alpha_2 = \alpha_6 = \alpha_7 = 1, \alpha_{769} = -1/20$ and $\alpha_j = 0$ otherwise we get that when $\|\delta\| = 3.85 \times 10^{-13}$, the reconstruction $\tilde{f} = \sum_{j=1}^m x_j w_j$ satisfies $\|f - \tilde{f}\|_{L_2} = 146.85$. The disastrous error $f - \tilde{f}$ as well as \tilde{f} is visualized in Figure 1. Note that, even though the perturbation is small, the exceedingly large condition number leads to a complete failure in the reconstruction.

One may think that a remedy could be to increase m . In particular, maybe taking more Fourier samples would cure the problem. Let us increase m to 1351, but keeping f as before as well as the noise level. Unfortunately, this also gives a completely unsatisfactory result (even worse than before) with $\|f - \tilde{f}\|_{L_2} = 360.12$ as visualized in Figure 1.

There is, however, a remedy to the problem. In particular, by defining $\tilde{f}_{n,m}$ as in (3.1), with parameters $m = 1351$ and $n = 769$, we obtain a recovery error that is very close to the noise level. This approach was first introduced in [3, 4], and we will in this paper extend this framework even further.

This failure is somewhat surprising (for similar examples, see [2, 3]). Since U_m is nothing more than the change-of-basis matrix between two systems of m orthogonal functions, it is tempting to think that U_m is unitary (and therefore well-conditioned). However, herein lies a critical misunderstanding. There is no reason why U_m should be unitary (or even invertible). The trained eye of an operator theorist will immediately spot the mistake: U_m is unitary if and only if $\text{span}\{w_k\}_{k=1}^m = \text{span}\{s_k\}_{k=1}^m$. The problem is that this may very well not be the case since \mathcal{H} is infinite-dimensional. Indeed, it is straightforward to see that the two bases in Example 2.1 do not lead to a unitary matrix: the first m Haar wavelets span a space of piecewise constant functions, whereas the standard complex exponentials span a space of (smooth) trigonometric polynomials.

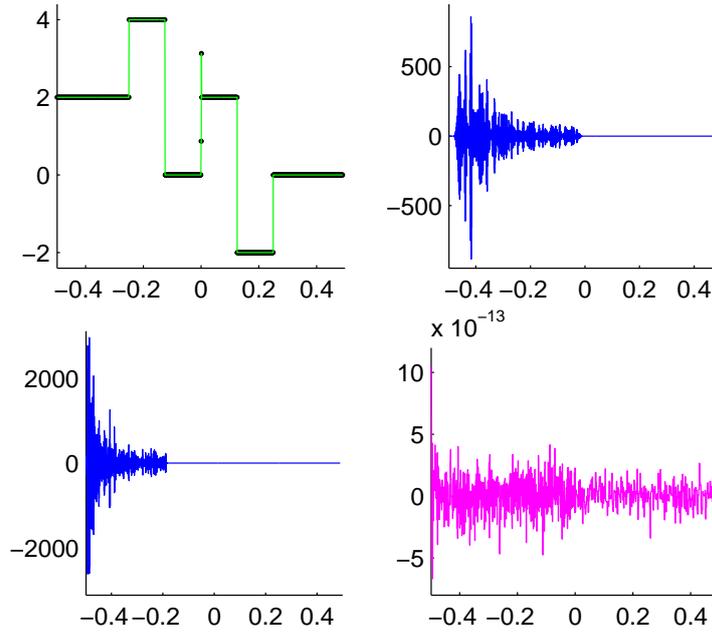


Figure 1: The figure displays f (upper left) as well as $f - \tilde{f}$ for two different values of m (upper right and lower left), and $f - \tilde{f}_{n,m}$ (lower right).

The approach considered above can be loosely described as ‘finite-dimensional’: one simply interpolates between two m -dimensional spaces. Herein lies the problem: the recovery of f , an element of the infinite-dimensional Hilbert space, is inherently infinite-dimensional. In particular, a typical $f \in \mathcal{H}$ will have an infinite expansion in $\sum_{k=1}^{\infty} \alpha_k w_k$ in the basis $\{w_k\}_{k \in \mathbb{N}}$, and a countably infinite collection of measurements $\{\langle f, s_k \rangle\}_{k \in \mathbb{N}}$. Thus, to tackle the reconstruction problem effectively, i.e. to avoid the pitfalls encountered above, it is necessary to pursue an infinite-dimensional approach.

2.3 The new approach

To compensate for the problems that consistent reconstructions cause, a new type of theory, and associated technique, was introduced, known as generalized sampling (GS) [2, 3, 4]. This theory generalizes the Shannon Sampling Theorem [21], as well as some fundamental work by Hrycak & Gröchenig [20].

The main message of GS is the following: one can obtain samples via inner products with respect to one basis (e.g. Fourier coefficients) and reconstruct in another basis of choice, and this can always be done in a completely stable and convergent way. Hence, one can combat the problem of potentially slow convergence in the sampling basis by simply reconstructing in a basis more appropriate for the signal to be reconstructed. Stability also implies complete robustness with respect to noise.

2.4 Overview of the paper

GS was introduced in a series of papers [2, 6, 3, 4]. However, there are several important problems that have not yet been addressed. These form the content of this paper:

- **Frames:** GS is so far only valid for sampling and reconstruction systems $\{s_k\}_{k \in \mathbb{N}}$ and $\{w_k\}_{k \in \mathbb{N}}$ that comprise Riesz bases. Given the vast literature on frames and the wide variety of applications it is crucial that the GS is extended to frames. This is content of the first part of the paper.
- **Ill-posed problems:** GS is concerned with the problem of reconstructing an element f from samples $\langle f, s_j \rangle$. However, what if we cannot sample f , but rather $g := Af$, where A is a compact operator. The topic of ill-posed problems is vast and of great importance. Therefore it is vital to have an extension of GS that includes such problems. In the second part of this paper, we will introduce new techniques with GS using regularization, as well as a non-regularized technique.

Let us remark in passing that the ideas leading to GS originate computational spectral theory [17, 18]. In particular, the technique of uneven sections of infinite matrices. Spectral theory has been used in sampling before (although in a very different context than what will be presented here) to great success, with the fundamental paper [22] of Landau being an important example. The problems above will also be addressed through extensions and refinements of the uneven section technique (see also [8] and [16] for other results on even and uneven section techniques).

2.5 Background and Notation

We will let \mathcal{H} denote a separable Hilbert space. Let $\{w_k\}_{k \in \mathbb{N}}$ be a frame for $\mathcal{W} = \overline{\text{span}}\{w_k\}_{k \in \mathbb{N}} \subset \mathcal{H}$ and $\{s_k\}_{k \in \mathbb{N}}$ be a frame for $\mathcal{S} = \overline{\text{span}}\{s_k\}_{k \in \mathbb{N}} \subset \mathcal{H}$. In particular, this means that for all $f \in \mathcal{W}$ there exist constants $0 < A \leq B < \infty$ such that

$$A\|f\|^2 \leq \sum_{k \in \mathbb{N}} |\langle f, w_k \rangle|^2 \leq B\|f\|^2, \quad (2.3)$$

and that for all $f \in \mathcal{S}$ there exist constants $0 < C \leq D < \infty$ such that

$$C\|g\|^2 \leq \sum_{k \in \mathbb{N}} |\langle g, s_k \rangle|^2 \leq D\|g\|^2. \quad (2.4)$$

Recall the so-called *synthesis operators* $S, W : l^2(\mathbb{N}) \rightarrow \mathcal{H}$ are defined by

$$Sx = x_1 s_1 + x_2 s_2 + \dots, \quad Wy = y_1 w_1 + y_2 w_2 + \dots,$$

and their adjoints (the *analysis operators*) $S^*, W^* : \mathcal{H} \rightarrow l^2(\mathbb{N})$ are easily seen to be

$$S^*g = \{\langle g, s_1 \rangle, \langle g, s_2 \rangle, \dots\}, \quad W^*h = \{\langle h, w_1 \rangle, \langle h, w_2 \rangle, \dots\}.$$

Define also the operators $S_m, W_m : \mathbb{C}^m \rightarrow \mathcal{H}$ by

$$S_m x = x_1 s_1 + \dots + x_m s_m, \quad W_m y = y_1 w_1 + \dots + y_m w_m, \quad (2.5)$$

with adjoints $S^*, W^* : \mathcal{H} \rightarrow \mathbb{C}^m$ given by

$$S_m^* g = \{\langle g, s_1 \rangle, \dots, \langle g, s_m \rangle\}, \quad W_m^* h = \{\langle h, w_1 \rangle, \dots, \langle h, w_m \rangle\}.$$

Define also the corresponding spaces

$$\mathcal{S}_m = \text{span}\{s_1, \dots, s_m\}, \quad \mathcal{W}_m = \text{span}\{w_1, \dots, w_m\}.$$

For any closed subspace $\mathcal{V} \subset \mathcal{H}$ we denote the orthogonal projection onto \mathcal{V} by $P_{\mathcal{V}}$. We shall also let $\{e_j\}_{j \in \mathbb{N}}$ denote the canonical basis of $l^2(\mathbb{N})$, and, for any $k \in \mathbb{N}$, write P_k for the orthogonal projection onto $\text{span}\{e_1, \dots, e_k\}$. Lastly, if $A : \mathcal{H} \rightarrow \mathcal{H}$ is bounded, where \mathcal{H} is a Hilbert space, we denote its pseudo-inverse by A^\dagger .

3 Generalized Sampling with Frames

Let $f \in \mathcal{H}$ and $\{w_k\}_{k \in \mathbb{N}}, \{s_k\}_{k \in \mathbb{N}}$ be as in the previous section. Suppose also that we can, for some $m \in \mathbb{N}$, access the samples $S_m^* f$. The task is to reconstruct an approximation to f from these samples. As mentioned, if $\{s_k\}_{k \in \mathbb{N}}$ was an orthonormal basis (and $f \in \mathcal{S}$) we could have formed the approximation $f_m = \sum_{k=1}^m \langle f, s_k \rangle s_k$. However, it may well be the case that $f_m \rightarrow f$ intolerably slowly and therefore f_m would be a rather poor approximation. Moreover, if $\{s_k\}_{k \in \mathbb{N}}$ is merely a frame, then f_m in general will not converge to f , and hence will be completely useless.

Suppose now that $P_{\mathcal{W}_m} f \rightarrow f$ rapidly – in particular, much faster than $f_m \rightarrow f$ (if valid). Given this information we are led to the following question: can we reconstruct $P_{\mathcal{W}_m} f$ (or at least an approximation) from the samples $S_m^* f$?

3.1 Reconstruction

Suppose now that we are given $f \in \mathcal{W}$ and that we can access the samples $S_m^* f$. Based on this information we will construct the GS approximation $\tilde{f}_{n,m} \in \mathcal{W}_n$ as follows:

$$\tilde{f}_{n,m} = W_n(W_n^* S_m S_m^* W_n)^\dagger W_n^* S_m S_m^* f. \quad (3.1)$$

The following theorem, which is proved in section 3.3, concerns the behaviour of $\tilde{f}_{n,m}$:

Theorem 3.1. *Let \mathcal{H} be a separable Hilbert space and $\mathcal{S}, \mathcal{W} \subset \mathcal{H}$ be closed subspaces such that $\mathcal{W} \cap \mathcal{S}^\perp = \{0\}$. Suppose that $\{s_k\}_{k \in \mathbb{N}}$ and $\{w_k\}_{k \in \mathbb{N}}$ are frames for \mathcal{S} and \mathcal{W} respectively with frame bounds as in (2.3) and (2.4) respectively. Then, for each $n \in \mathbb{N}$ and any $f \in \mathcal{W}$, there is an $M \in \mathbb{N}$ such that, for all $m \geq M$, the reconstruction $\tilde{f}_{n,m}$ (defined in (3.1)) exists and is unique. Moreover,*

$$\|P_{\mathcal{W}_n}^\perp f\|_{\mathcal{H}} \leq \|f - \tilde{f}_{n,m}\|_{\mathcal{H}} \leq (1 + K_{n,m}) \|P_{\mathcal{W}_n}^\perp f\|_{\mathcal{H}},$$

where $P_{\mathcal{W}_n}$ is the orthogonal projection onto \mathcal{W}_n , $P_{\mathcal{W}_n}^\perp = I - P_{\mathcal{W}_n}$ and

$$K_{n,m} = \|W_n(W_n^* S_m S_m^* W_n)^\dagger W_n^* S_m S_m^* P_{\mathcal{W}_n}^\perp\|. \quad (3.2)$$

For each fixed n we have that $K_{n,m} \rightarrow K_n$ as $m \rightarrow \infty$, where

$$K_n = \|W_n(W_n^* S S^* W_n)^\dagger W_n^* S S^* P_{\mathcal{W}_n}^\perp\| \leq \frac{D}{C}.$$

Moreover, when $\{s_k\}_{k \in \mathbb{N}}$ is an orthonormal basis and $\mathcal{W} \subset \mathcal{S}$ then, for fixed n , $K_{n,m} \rightarrow 0$ as $m \rightarrow \infty$.

Note that this theorem generalizes the results in [3, 4] to frames (we give a proof in the next section).

3.2 Stability

The stability of GS is very important to analyze. One may at first glance suspect that this method could be unstable, due to the fact that $\|(W_n^* S_m S_m^* W_n)^\dagger\|$ could blow up for large n , even when m is arbitrarily large. Indeed, this could very well happen as suggested in the next proposition.

Proposition 3.2. *There exists a Hilbert space \mathcal{H} with frames $\{w_k\}_{k \in \mathbb{N}}$ and $\{s_k\}_{k \in \mathbb{N}}$ such that*

$$\lim_{m \rightarrow \infty} \|(W_n^* S_m S_m^* W_n)^\dagger\| \rightarrow \infty, \quad n \rightarrow \infty.$$

Proof. Let $\mathcal{H} = l^2(\mathbb{N})$ and write $\{e_k\}_{k \in \mathbb{N}}$ for its canonical basis. Let $\{w_k\}_{k \in \mathbb{N}}$ be defined by $w_1 = e_1$ and $w_k = e_{k-1} + \frac{e_k}{k}$, $k \geq 2$, and let $\{s_k\}_{k \in \mathbb{N}}$ be given by $s_k = e_k$. Then, as shown in [9], $\{w_k\}_{k \in \mathbb{N}}$ is a frame with frame bounds $1 - \sqrt{\pi^2/6 - 1}$ and 3 respectively. Note that

$$W_n y = y_1 e_1 + \sum_{k=2}^n y_k \left(\frac{e_k}{k} + e_{k-1} \right), \quad y \in \mathbb{C}^n,$$

and therefore

$$S_m^* W_n = \begin{pmatrix} 1 & 1 & 0 & 0 & \dots & 0 \\ 0 & \frac{1}{2} & 1 & 0 & \ddots & 0 \\ 0 & 0 & \frac{1}{3} & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \ddots & \ddots & 1 \\ 0 & 0 & 0 & \dots & \dots & \frac{1}{n} \\ 0 & 0 & 0 & \dots & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \dots & 0 \end{pmatrix} \in \mathbb{C}^{m \times n}, \quad m \geq n.$$

In particular, we have that $\|(W_n^* S_m S_m^* W_n)^\dagger\| \geq n^2$, yielding the assertion. \square

Despite this rather pessimistic result, the method is perfectly stable. Indeed, although $\|(W_n^* S_m S_m^* W_n)^\dagger\|$ may blow up, it is always the case that $\|W_n(W_n^* S_m S_m^* W_n)^\dagger W_n^*\|$ stays bounded provided m is sufficiently large. Specifically, we have the following theorem (proved in section 3.3).

Theorem 3.3. *Let \mathcal{H} , \mathcal{S} , \mathcal{W} , $\{s_k\}_{k \in \mathbb{N}}$, $\{w_k\}_{k \in \mathbb{N}}$, n and m be as in Theorem 3.1. Suppose that the samples of $f \in \mathcal{H}$ are contaminated by noise. In particular, suppose that we are given $S_m^*(f + h)$ for some $h \in \mathcal{H}$. If*

$$\tilde{f}_{n,m}^h = W_n(W_n^* S_m S_m^* W_n)^\dagger W_n^* S_m S_m^*(f + h),$$

then

$$\|f - \tilde{f}_{n,m}^h\|_{\mathcal{H}} \leq (1 + K_{n,m}) (\|P_{\mathcal{W}_n}^\perp f\|_{\mathcal{H}} + \|h\|_{\mathcal{H}}),$$

for all $m \geq M$ where M is as in Theorem 3.1.

3.3 Proofs

Proof of Theorem 3.1. Let us first observe that due to the fact that $\text{Ran}(W_n^* S_m S_m^* W_n)$ is finite dimensional, and hence closed, then the pseudo-inverse is unique and hence $\tilde{f}_{n,m}$ exists and is unique. Moreover, we observe that

$$W_n(W_n^* W_n)^\dagger W_n^* W_n(W_n^* W_n)^\dagger W_n^* = W_n(W_n^* W_n)^\dagger P_{\text{Ran}(W_n^* W_n)} W_n^* = W_n(W_n^* W_n)^\dagger W_n^*.$$

In particular, by self-adjointness, it follows that

$$P_{\mathcal{W}_n} = W_n(W_n^* W_n)^\dagger W_n^*. \quad (3.3)$$

Note that

$$\begin{aligned} W_n^* S_m S_m^* f &= W_n^* S_m S_m^* (P_{\mathcal{W}_n} f + P_{\mathcal{W}_n}^\perp f) \\ &= W_n^* S_m S_m^* W_n(W_n^* W_n)^\dagger W_n^* f + W_n^* S_m S_m^* P_{\mathcal{W}_n}^\perp f. \end{aligned} \quad (3.4)$$

Equation (3.4) now gives

$$\begin{aligned} W_n(W_n^* S_m S_m^* W_n)^\dagger W_n^* S_m S_m^* f &= W_n(W_n^* S_m S_m^* W_n)^\dagger W_n^* S_m S_m^* W_n(W_n^* W_n)^\dagger W_n^* f \\ &\quad + W_n(W_n^* S_m S_m^* W_n)^\dagger W_n^* S_m S_m^* P_{\mathcal{W}_n}^\perp f. \end{aligned} \quad (3.5)$$

We now claim that there exists an M such that

$$W_n(W_n^* S_m S_m^* W_n)^\dagger W_n^* S_m S_m^* W_n(W_n^* W_n)^\dagger W_n^* f = P_{\mathcal{W}_n} f, \quad \forall m \geq M.$$

To see this, we first observe that

$$W_n(W_n^* S_m S_m^* W_n)^\dagger W_n^* S_m S_m^* W_n(W_n^* W_n)^\dagger W_n^* f = W_n P_{\text{Ran}(W_n^* S_m S_m^* W_n)} (W_n^* W_n)^\dagger W_n^* f,$$

and it therefore suffices to show that

$$\text{Ran}(W_n^* S_m S_m^* W_n) = \text{Ran}(W_n^* W_n) \quad (3.6)$$

for all sufficiently large m . By assumption we have $\mathcal{W} \cap \mathcal{S}^\perp = \{0\}$. Thus, for all $n \in \mathbb{N}$ we deduce that $\mathcal{W}_n \cap \mathcal{S}^\perp = \{0\}$. It follows that

$$\inf\{\|S^* x\| : x \in \text{Ran}(W_n), \|x\| = 1\} > 0.$$

In particular, we have that $\text{Rank}(S^* W_n) = \text{Rank}(W_n^* W_n)$, and therefore $\text{Rank}(W_n^* S S^* W_n) = \text{Rank}(W_n^* W_n)$, which yields that $\text{Ran}(W_n^* S S^* W_n) = \text{Ran}(W_n^* W_n)$. Note that $S_m S_m^* \rightarrow S S^*$ strongly, however, by compactness of W_n it follows that $W_n^* S_m S_m^* W_n \rightarrow W_n^* S S^* W_n$ in norm as $m \rightarrow \infty$. Hence, we deduce (3.6) for all sufficiently large m . Thus,

$$\|f - \tilde{f}_{n,m}^h\|_{\mathcal{H}} \leq \|P_{\mathcal{W}_n}^\perp - W_n(W_n^* S_m S_m^* W_n)^\dagger W_n^* S_m S_m^* P_{\mathcal{W}_n}^\perp\|_{\mathcal{H}} \|P_{\mathcal{W}_n}^\perp f\|_{\mathcal{H}},$$

which gives the first part of the theorem.

To prove the second assertion note that

$$W_n(W_n^*SS^*W_n)^\dagger W_n^*SS^*P_{\mathcal{W}_n}^\perp = (SS^*)^{-1}SS^*W_n(W_n^*SS^*W_n)^\dagger W_n^*SS^*P_{\mathcal{W}_n}^\perp,$$

and therefore, since $S^*W_n(W_n^*SS^*W_n)^\dagger W_n^*S$ is the orthogonal projection $P_{\text{Ran}(S^*W_n)}$, it follows that

$$\begin{aligned} \|W_n(W_n^*SS^*W_n)^\dagger W_n^*SS^*P_{\mathcal{W}_n}^\perp\| &\leq \|(SS^*)^{-1}\|_{\mathcal{S} \rightarrow \mathcal{S}} \|S\|_{l^2(\mathbb{N}) \rightarrow \mathcal{S}} \|S^*\|_{\mathcal{S} \rightarrow l^2(\mathbb{N})} \\ &= \|(SS^*)^{-1}\|_{\mathcal{S} \rightarrow \mathcal{S}} \|SS^*\|_{\mathcal{S} \rightarrow \mathcal{S}} \leq \frac{D}{C}. \end{aligned}$$

Also, by a compactness argument, since $S_m S_m^* \rightarrow SS^*$ strongly, it follows that

$$W_n(W_n^*S_m S_m^* W_n)^\dagger W_n^* S_m S_m^* P_{\mathcal{W}_n}^\perp \longrightarrow W_n(W_n^*SS^*W_n)^\dagger W_n^*SS^*P_{\mathcal{W}_n}^\perp, \quad m \rightarrow \infty. \quad (3.7)$$

Therefore $K_{n,m} \rightarrow \|W_n(W_n^*SS^*W_n)^\dagger W_n^*SS^*P_{\mathcal{W}_n}^\perp\|$ as $m \rightarrow \infty$ and we have proved the second assertion.

To prove that $K_{n,m} \rightarrow 0$ as $m \rightarrow \infty$ when $\{s_k\}_{k \in \mathbb{N}}$ is an orthonormal basis, we start by making the following observation: since $\mathcal{W} \subset \mathcal{S}$ we certainly have that $\mathcal{W}_n \subset \mathcal{S}$. Therefore $P_{\mathcal{S}}P_{\mathcal{W}_n} = P_{\mathcal{W}_n} = P_{\mathcal{W}_n}P_{\mathcal{S}}$. Also, $W_n^* = W_n^*P_{\mathcal{W}_n}$, and therefore

$$W_n^*P_{\mathcal{S}}P_{\mathcal{W}_n}^\perp = W_n^*P_{\mathcal{S}}(I - P_{\mathcal{W}_n}) = 0.$$

We therefore immediately obtain

$$W_n(W_n^*SS^*W_n)^\dagger W_n^*SS^*P_{\mathcal{W}_n}^\perp = W_n(W_n^*P_{\mathcal{S}}W_n)^\dagger W_n^*P_{\mathcal{S}}P_{\mathcal{W}_n}^\perp = 0.$$

So, by (3.7) the assertion follows, and we are done. \square

Proof of Theorem 3.3. Note that (3.5) and the reasoning in the proof of Theorem 3.3 give

$$f - \tilde{f}_{n,m}^h = P_{\mathcal{W}_n}^\perp f - W_n(W_n^*S_m S_m^* W_n)^\dagger W_n^* S_m S_m^* P_{\mathcal{W}_n}^\perp (f + h) - h,$$

which yields the assertion. \square

3.4 Determining m : The Stable Sampling Rate

For Theorems 3.1 and 3.3 we require $m \geq M$ samples, where M depends only on n and the sampling and reconstruction systems. This leads to the question: how large must m be? This was first explored in [3, 4] in the context of Riesz bases. Continuing in the same way, let us define the *stable sampling rate*

$$\Phi(n, \theta) = \min \{m \in \mathbb{N} : K_{n,m} \leq \theta\}, \quad (n, \theta) \in \mathbb{N} \times \mathbb{R}_+, \quad (3.8)$$

where $K_{n,m}$ is as in (3.2). This quantity tells us how many samples one has to take, as a function of the number of coefficients to be computed, in order to secure a stable and convergent solution. It is of course very important to get analytical bounds for Φ . However, these may be very difficult to derive for arbitrary sampling and reconstruction systems. Fortunately, as we now explain, we can always obtain numerical estimates for Φ .

When confronted with the problem of computing Φ there are two obvious obstacles:

- (i) The operator $P_{\mathcal{W}_n}^\perp$ does not have finite rank (and this suggests therefore the need for an infinite amount of information to compute $K_{n,m}$).
- (ii) The operator $W_n(W_n^*S_m S_m^* W_n)^\dagger W_n^* S_m S_m^* P_{\mathcal{W}_n}^\perp$ acts on the abstract Hilbert space \mathcal{H} . In it therefore unclear how to obtain numerical norm estimates when the information we will (at best) be able to access are the numerical values of the inner products $\langle s_i, w_j \rangle$ and $\langle w_i, w_j \rangle$ for $i, j \in \mathbb{N}$.

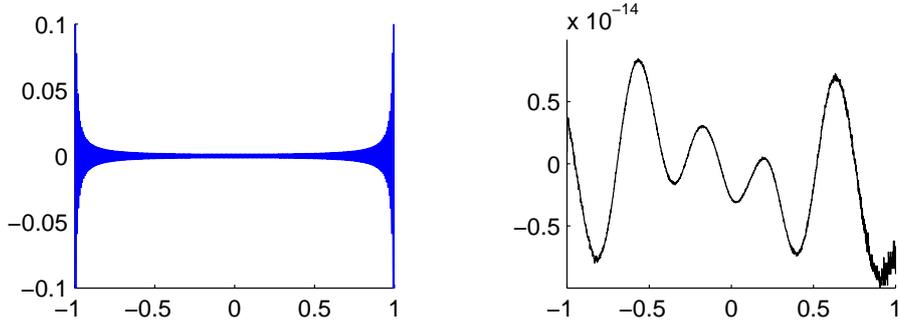


Figure 2: The left figure shows the error $f - f_m$ (the truncated Fourier series) for $m = 501$. The right figure shows the error $f - \tilde{f}_{n,m}$ for $m = 501, n = 45$. Note that both f_m and $\tilde{f}_{n,m}$ use exactly the same samples.

Fortunately, one can at least obtain a computable upper bound for $K_{n,m}$ (and therefore Φ). Indeed, if $\{s_k\}_{k \in \mathbb{N}}$ is an orthonormal basis for \mathcal{S} , $\mathcal{W} \subset \mathcal{S}$ and U and V are the infinite matrices

$$U = \begin{pmatrix} u_{11} & u_{12} & u_{13} & \dots \\ u_{21} & u_{22} & u_{23} & \dots \\ u_{31} & u_{32} & u_{33} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \quad V = \begin{pmatrix} v_{11} & v_{12} & v_{13} & \dots \\ v_{21} & v_{22} & v_{23} & \dots \\ v_{31} & v_{32} & v_{33} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad u_{ij} = \langle s_i, w_j \rangle, v_{ij} = \langle w_i, w_j \rangle, \quad (3.9)$$

then it is a straightforward exercise (along the same lines as arguments given in [3]) to show that $K_{n,m} \leq \tilde{K}_{n,m}$, where

$$\tilde{K}_{n,m} = \sqrt{B} \left\| (P_n U^* P_m U P_n)^\dagger P_n U^* P_m - (P_n V P_n)^\dagger P_n U^* P_m \right\| + \left\| (P_n V P_n)^\dagger (P_n U^* P_m U P_n - P_n V P_n) (P_n V P_n)^\dagger \right\|^{1/2}.$$

Thus, we have

$$\Phi(n, \theta) \leq \Psi(n, \theta) = \min \left\{ m \in \mathbb{N} : \tilde{K}_{n,m} \leq \theta \right\}, \quad \forall (n, \theta) \in \mathbb{N} \times \mathbb{R}_+, \quad (3.10)$$

where the right-hand side can be computed numerically, since $\tilde{K}_{n,m}$ involves only sections of the matrices U and V . Thus, in practice, one can always ensure, via a numerical computation, the bounds of Theorems 3.1 and 3.3, and therefore convergence and stability. In the next section we give an example of the behaviour of $\Psi(n, \theta)$ for a typical problem (further examples in the case of bases are given in [3, 4]).

3.5 Numerical Examples

In this section we will test the framework suggested in the previous section with the following frames: let $\{p_k\}_{k \in \mathbb{N}}$ denote the Legendre polynomials on $[-1, 1]$ and $\{h_k\}_{k \in \mathbb{N}}$ the Haar wavelets on $[0, 1]$. Define the frames

$$\{w_k\}_{k \in \mathbb{N}}, \quad w_{2k-1} = p_k, \quad w_{2k} = h_k, \quad \{\tilde{w}_k\}_{k \in \mathbb{N}}, \quad \tilde{w}_k = p_k. \quad (3.11)$$

Note that $\{\tilde{w}_k\}_{k \in \mathbb{N}}$ is actually an orthonormal basis. When it comes to indexing the frame we simply make the most natural choice. However, one should note that one has complete freedom to choose the indexing, although the synthesis and analysis operators will obviously depend on such a choice. We will denote the operators and spaces associated with this basis (as defined in Section 3) by simply adding a tilde, for example $\tilde{W}_n, \tilde{W}_n^*, \tilde{W}$ etc. Also, define the orthonormal basis for $L_2([-1, 1])$

$$\{s_k\}_{k \in \mathbb{Z}}, \quad s_k(t) = \frac{1}{\sqrt{2}} e^{\pi i k t}.$$

Thus, our samples of any function $f \in L_2([-1, 1])$ will be its Fourier coefficients. In particular,

$$\hat{f}(k) = \langle f, s_k \rangle = \frac{1}{\sqrt{2}} \int_{-1}^1 f(t) e^{\pi i k t} dt.$$

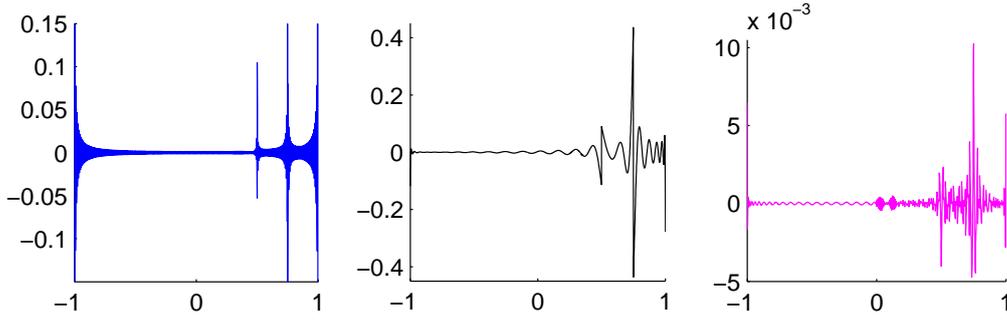


Figure 3: The figure shows the errors $f_d - f_{d,m}$ (left), $f_d - \tilde{f}_{d,n,m}$ (middle) and $f_d - f_{d,k,m}$ (right) for $m = 501$, $n = 45$, $k = 150$. Note that $f_{d,m}$, $\tilde{f}_{d,n,m}$ and $f_{d,k,m}$ use exactly the same samples.

Thus, in this case we have $\mathcal{W} = \tilde{\mathcal{W}} = \mathcal{S} = L_2([-1, 1])$. In the previous framework we have indexed the operators over the natural numbers, so to make it compatible with the setup in this section we will let (for odd m) $S_m : \mathbb{C}^m \rightarrow L_2([-1, 1])$ and $S_m^* : L_2([-1, 1]) \rightarrow \mathbb{C}^m$ be defined by

$$S_m x = x_1 s_{-(m-1)/2} + \dots + x_m s_{(m-1)/2}, \quad S_m^* g = \{\langle g, s_{-(m-1)/2} \rangle, \dots, \langle g, s_{(m-1)/2} \rangle\}.$$

Also, for $f \in L_2([-1, 1])$ and for odd $m \in \mathbb{N}$ we let f_m denote the truncated Fourier series of f . In particular,

$$f_m(t) = (S_m S_m^* f)(t) = \frac{1}{2} \sum_{j=-(m-1)/2}^{(m-1)/2} \hat{f}(j) e^{\pi i j t}.$$

Example 3.4. In order to illustrate the theory developed in the previous sections we have tested the functions

$$f(t) = t^3 e^t, \quad f_d(t) = t^3 e^t (1 - \chi_{[0.5, 0.75]})(t).$$

Note that the analyticity of f strongly favors reconstruction in a polynomial basis. In particular, the fact that f is non-periodic prevents rapid convergence of the Fourier series as shown in Figure 2, where we have displayed the error $f - f_m$ for $m = 501$. As an alternative we have defined the GS reconstruction

$$\tilde{f}_{n,m} = \tilde{W}_n (\tilde{W}_n^* S_m S_m^* \tilde{W}_n)^\dagger \tilde{W}_n^* S_m S_m^* f,$$

where $m = 501$ and $n = 45$. By using exactly the same samples as for f_m , namely $S_m^* f$, we simply reconstruct in a different basis (the Legendre basis), and dramatically reduce the error. This is illustrated in Figure 2. For more examples with reconstructions in different bases see [6, 2, 3, 4]. However, this paper is primarily about the use of frames. In particular, frames that can be used to capture specific characteristics of the function to be recovered.

To illustrate the effectiveness of the use of frames, consider the following. What happens if we introduce a discontinuity in f ? In particular, what happens if we replace f by the discontinuous function f_d ? This does have an effect on the truncated Fourier series (as documented in Figure 3) $f_{d,m} = S_m S_m^* f_d$, but not near the striking impact it has on

$$\tilde{f}_{d,n,m} = \tilde{W}_n (\tilde{W}_n^* S_m S_m^* \tilde{W}_n)^\dagger \tilde{W}_n^* S_m S_m^* f_d,$$

which uses the Legendre basis in the same way as for $\tilde{f}_{n,m}$. The quite dramatic effect of the discontinuity is visualized in Figure 3. The question is then: what to do? One idea that comes to mind is that one could try to mix two bases, one that would favor smoothness, and one that would favor discontinuities. This is exactly what the frame $\{w_k\}_{k \in \mathbb{N}}$ (defined in (3.11)) would do. By introducing

$$f_{d,k,m} = W_k (W_k^* S_m S_m^* W_k)^\dagger W_k^* S_m S_m^* f_d,$$

we can again dramatically reduce the error. This is documented in Figure 3. Note that $f_{d,m}$, $\tilde{f}_{d,n,m}$ and $f_{d,k,m}$ use exactly the same samples.

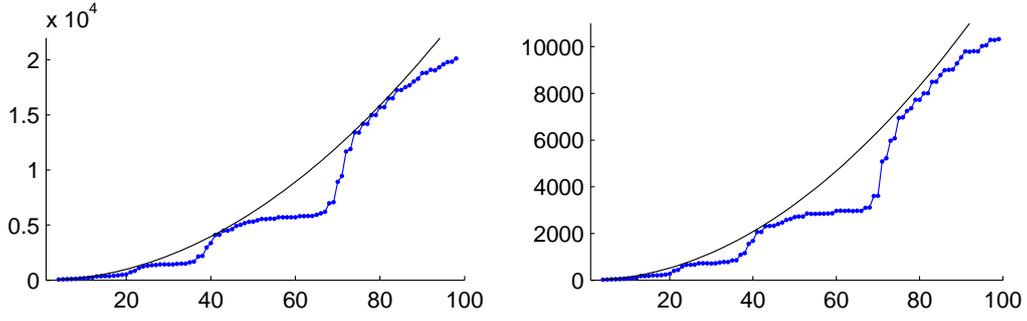


Figure 4: The left figure shows $\Psi(n, 1)$ for $n = 1, \dots, 100$ together with the mapping $t \mapsto 2.45t^2$. The right figure shows $\Psi(n, 2)$ for $n = 1, \dots, 100$ together with the mapping $t \mapsto 1.3t^2$

Example 3.5. In this example we consider numerical estimates of the function Φ defined in (3.8). However, as pointed out above, the function Ψ defined in (3.10) is much simpler to compute and we have

$$\Phi(S, W, n, \theta) \leq \Psi(S, W, n, \theta), \quad \forall (n, \theta) \in \mathbb{N} \times \mathbb{R}_+.$$

Sections of the graphs of $\Psi(\cdot, 1)$ and $\Psi(\cdot, 2)$ are displayed in Figure 4. Note that it is not a big surprise that

$$\Psi(S, W, n, \theta) = \mathcal{O}(n^2), \quad \theta = 1, 2.$$

In fact, it is not hard to use the fact that

$$\Phi(S, \widetilde{W}, n, \theta) = \mathcal{O}(n^2), \quad \forall \theta \in \mathbb{R}_+,$$

(see [4]) in order to prove that

$$\Phi(S, W, n, \theta) = \mathcal{O}(n^2), \quad \forall \theta \in \mathbb{R}_+.$$

4 Generalized Sampling and Inverse Problems

Suppose now that we are given Hilbert spaces \mathcal{X} and \mathcal{Y} and a linear and bounded operator A that maps between \mathcal{X} and \mathcal{Y} . Moreover, we restrict the following discussion to compact operators, i.e. to operators A that are equipped with a singular system $\{\sigma_k, v_k, u_k\}$, where the orthonormal system $\{v_k\}_{k \in \mathbb{N}}$ spans $\mathcal{V} = \mathcal{N}(A)^\perp$ and where the orthonormal system $\{u_k\}_{k \in \mathbb{N}}$ spans $\mathcal{U} = \mathcal{N}(A^*)^\perp$. We aim to solve

$$Af = g,$$

where we are typically faced with noisy data $g^\delta = g + z$ with $\|z\| \leq \delta$. We assume that we have a sampling system $\{s_k\}_{k \in \mathbb{N}}$ that spans $\mathcal{S} = \overline{\text{span}}\{s_k\}_{k \in \mathbb{N}} = \mathcal{N}(A^*)^\perp \subset \mathcal{Y}$ and a reconstruction system $\{w_k\}_{k \in \mathbb{N}}$ spanning $\mathcal{W} = \overline{\text{span}}\{w_k\}_{k \in \mathbb{N}} = \mathcal{N}(A)^\perp \subset \mathcal{X}$ at our disposal. With $\{s_k\}_{k \in \mathbb{N}}$ we are able to take samples of g ,

$$\eta = S^*g = \{\langle g, s_k \rangle\}_{k \in \mathbb{N}},$$

and with $\{w_k\}_{k \in \mathbb{N}}$ we expand f ,

$$f = W\beta = \sum_{k \in \mathbb{N}} \beta_k w_k.$$

As we cannot deal with infinitely many samples of g and infinite series expansions for f , the goal is to reconstruct the best possible approximation of the form $\tilde{f} = W_n \tilde{\beta} = \sum_{k=1}^n \tilde{\beta}_k w_k$ based on the finite subset $\{\langle g, s_k \rangle\}_{k=1}^r = S_r^*g$ of the full sampling information η . We shall do this by exploiting the ideas of generalized sampling. As we now allow the problem to be ill-posed and the data noisy, we rephrase the problem as follows: given n , is there some r (determine if possible) such that

$$\|S_r^*g^\delta - S_r^*AW_n\tilde{\beta}\|^2 \rightarrow \min_{\tilde{\beta}}$$

has a unique solution? The corresponding normal equation is given by

$$(W_n^* A^* S_r S_r^* A W_n) \begin{pmatrix} \tilde{\beta}_1 \\ \vdots \\ \tilde{\beta}_n \end{pmatrix} = W_n^* A^* S_r \begin{pmatrix} \langle g^\delta, s_1 \rangle \\ \vdots \\ \langle g^\delta, s_r \rangle \end{pmatrix} \quad (4.1)$$

and therefore

$$f_{n,r}^\delta = W_n \tilde{\beta} = W_n (W_n^* A^* S_r S_r^* A W_n)^\dagger W_n^* A^* S_r S_r^* g. \quad (4.2)$$

As already mentioned, the problem can be ill-posed and therefore the generalized inverse in (4.2) need not exist. Hence we are first faced with regularization issues. The second task is to analyze (4.2) by both establishing existence of approximations of the form $f_{n,r}^\delta$ and providing error bounds for $\|f - f_{n,r}^\delta\|$.

In what follows we discuss two different treatments of (4.2). Both proposed techniques heavily rely on the singular value decomposition of the operator A . This allows a splitting into a sampling and recovery step (which is quite natural). The sampling step in both algorithms is almost the same, whereas the recovery steps are rather different. In the first approach the recovery step relies on (classical) regularization principles, whereas in the second approach we seek to stabilize the ill-posed problem by adequate discretizations (which, of course, is also a regularization technique).

4.1 Regularized Reconstruction

Let us consider the normal equation $A^* A f = A^* g$ and let A^\dagger denote the generalized inverse of A . If $g \in \mathcal{D}(A^\dagger)$, we can define $f^\dagger := A^\dagger g$. If A is injective it makes sense to define $A^\dagger := (A^* A)^{-1} A^*$. A stabilized version of f^\dagger can be constructed through $f^\alpha = R_\alpha g$ with $R_\alpha = F_\alpha (A^* A) A^*$ with properly chosen F_α (for an extensive discussion on the choice of F_α see [15] or [23] and references therein).

Therefore, we can write

$$f^\alpha = W \beta^\alpha = \sum_k \beta_k^\alpha w_k = F_\alpha (A^* A) A^* g = \sum_l F_\alpha(\sigma_l^2) \sigma_l \langle g, u_l \rangle v_l.$$

For the singular system we associate the corresponding synthesis and analysis operators which we denote by V, U and V^*, U^* as well as its finite versions V_m, U_m and V_m^*, U_m^* . Therefore, for all $j \in \mathbb{N}$ we have

$$(V^* f^\alpha)_j = \sum_k \beta_k^\alpha \langle w_k, v_j \rangle = \sum_l F_\alpha(\sigma_l^2) \sigma_l \langle g, u_l \rangle \langle v_l, v_j \rangle,$$

which is nothing other than

$$V^* W \beta^\alpha = \Theta_\alpha \Sigma \gamma \quad \text{or equivalently,} \quad \Theta_\alpha^{-1} V^* W \beta^\alpha = \Sigma \gamma,$$

where

$$V^* W = \begin{pmatrix} \langle w_1, v_1 \rangle & \langle w_1, v_2 \rangle & \dots \\ \langle w_2, v_1 \rangle & \langle w_2, v_2 \rangle & \dots \\ \vdots & \vdots & \ddots \end{pmatrix}, \quad \Theta_\alpha = \begin{pmatrix} F_\alpha(\sigma_1^2) & 0 & \dots \\ 0 & F_\alpha(\sigma_2^2) & \dots \\ \vdots & \vdots & \ddots \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \sigma_1 & 0 & \dots \\ 0 & \sigma_2 & \dots \\ \vdots & \vdots & \ddots \end{pmatrix}.$$

The vector $\gamma = U^* g = \{\langle g, u_l \rangle\}_{l \in \mathbb{N}}$ is not accessible in practice and must be related to the samples of g with respect to our sampling frame $\{s_k\}_{k \in \mathbb{N}}$,

$$\eta = S^* g = S^* U U^* g = S^* U \gamma, \quad \text{where} \quad S^* U = \begin{pmatrix} \langle u_1, s_1 \rangle & \langle u_1, s_2 \rangle & \dots \\ \langle u_2, s_1 \rangle & \langle u_2, s_2 \rangle & \dots \\ \vdots & \vdots & \ddots \end{pmatrix}.$$

Consequently, in order to compute $f^\alpha = W \beta^\alpha$ we are faced with two infinite dimensional linear equations that need to be solved,

$$S^* U \gamma = \eta \quad (4.3)$$

$$\Theta_\alpha^{-1} V^* W \beta^\alpha = \Sigma \gamma. \quad (4.4)$$

We proceed in two steps. To find an approximation to the solution of (4.4) we assume in a first step that a finite length approximation $\gamma_m = U_m^* g$ to the solution of (4.3) is at our disposal (with m specified below). Then, in a second step, we derive an approximation $\gamma_{m,r}^\delta$ to γ_m where we also allow the data g to be noisy.

Suppose now that we have the samples $U_m^* g$ at our disposal. Based on this information we construct an approximation $f_{n,m}^\alpha = W_n \beta_{n,m}^\alpha \in \mathcal{W}_n$ as follows:

$$f_{n,m}^\alpha = W_n (W_n^* V_m \Theta_{\alpha,m}^{-1} \Theta_{\alpha,m}^{-1} V_m^* W_n)^\dagger W_n^* V_m \Theta_{\alpha,m}^{-1} \Sigma_m U_m^* g, \quad (4.5)$$

where $\Theta_{\alpha,m} = P_m \Theta_\alpha|_{P_m \ell^2(\mathbb{N})}$. As the following theorem indicates, $f_{n,m}^\alpha$ is uniquely defined for sufficiently large m :

Theorem 4.1. *Let \mathcal{X}, \mathcal{Y} be separable Hilbert spaces and $A : \mathcal{X} \rightarrow \mathcal{Y}$ be a linear and bounded operator with singular system $\{\sigma_k, v_k, u_k\}$. Moreover, let $\mathcal{W} = \mathcal{N}(A)^\perp \subset \mathcal{X}$ and $\mathcal{S} = \mathcal{N}(A^*)^\perp \subset \mathcal{Y}$. Suppose $\{s_k\}_{k \in \mathbb{N}}$ and $\{w_k\}_{k \in \mathbb{N}}$ are frames for \mathcal{S} and \mathcal{W} respectively. Then, for the generalized solution $f^\dagger \in \mathcal{W} \subset \mathcal{X}$ and $n \in \mathbb{N}$ and $\alpha > 0$, there is an $M \in \mathbb{N}$ such that, for all $m \geq M$ the approximation $f_{n,m}^\alpha$ exists and is unique. Moreover,*

$$\|f^\dagger - f_{n,m}^\alpha\|_{\mathcal{X}} \leq (2 + K_{n,m,\alpha}) \|f^\dagger - f^\alpha\|_{\mathcal{X}} + (1 + K_{n,m,\alpha}) \|P_{\mathcal{W}_n}^\perp f^\dagger\|_{\mathcal{X}},$$

where

$$K_{n,m,\alpha} := \|W_n (W_n^* V_m \Theta_{\alpha,m}^{-1} \Theta_{\alpha,m}^{-1} V_m^* W_n)^\dagger W_n^* V_m \Theta_{\alpha,m}^{-1} \Theta_{\alpha,m}^{-1} V_m^* P_{\mathcal{W}_n}^\perp\|.$$

In particular, for m large enough,

$$K_{n,m,\alpha} \leq \frac{\lambda_{\max}(\Theta_{\alpha,m}^{-2})}{\lambda_{\min}(\Theta_{\alpha,m}^{-2})},$$

where $\lambda_{\max}(\Theta_{\alpha,m}^{-2})$ and $\lambda_{\min}(\Theta_{\alpha,m}^{-2})$ denote the largest and smallest eigenvalue of $\Theta_{\alpha,m}^{-2}$ respectively.

Proof. Solving the optimization problem

$$\|\Sigma_m U_m^* g - \Theta_{\alpha,m}^{-1} V_m^* W_n \beta_{n,m}^\alpha\|^2 \rightarrow \min_{\beta_{n,m}^\alpha} \quad (4.6)$$

results in

$$\begin{aligned} \beta_{n,m}^\alpha &= (W_n^* V_m \Theta_{\alpha,m}^{-2} V_m^* W_n)^\dagger W_n^* V_m \Theta_{\alpha,m}^{-1} \Sigma_m U_m^* g \\ &= (W_n^* V_m \Theta_{\alpha,m}^{-2} V_m^* W_n)^\dagger W_n^* V_m \Theta_{\alpha,m}^{-2} V_m^* (P_{\mathcal{W}_n} f^\alpha + P_{\mathcal{W}_n}^\perp f^\alpha) \end{aligned}$$

and therefore, as argued in the proof of Theorem 3.1, for $m \geq M$ and M large enough,

$$\begin{aligned} f_{n,m}^\alpha &= W_n (W_n^* V_m \Theta_{\alpha,m}^{-2} V_m^* W_n)^\dagger W_n^* V_m \Theta_{\alpha,m}^{-2} V_m^* (P_{\mathcal{W}_n} f^\alpha + P_{\mathcal{W}_n}^\perp f^\alpha) \\ &= P_{\mathcal{W}_n} f^\alpha + W_n (W_n^* V_m \Theta_{\alpha,m}^{-2} V_m^* W_n)^\dagger W_n^* V_m \Theta_{\alpha,m}^{-2} V_m^* P_{\mathcal{W}_n}^\perp f^\alpha, \end{aligned}$$

and $f_{n,m}^\alpha$ is unique. Consequently,

$$\begin{aligned} \|f^\dagger - f_{n,m}^\alpha\|_{\mathcal{X}} &\leq \|f^\dagger - f^\alpha\|_{\mathcal{X}} + \|f^\alpha - f_{n,m}^\alpha\|_{\mathcal{X}} \\ &\leq \|f^\dagger - f^\alpha\|_{\mathcal{X}} + \|P_{\mathcal{W}_n}^\perp f^\alpha - W_n (W_n^* V_m \Theta_{\alpha,m}^{-2} V_m^* W_n)^\dagger W_n^* V_m \Theta_{\alpha,m}^{-2} V_m^* P_{\mathcal{W}_n}^\perp f^\alpha\|_{\mathcal{X}} \\ &\leq \|f^\dagger - f^\alpha\|_{\mathcal{X}} + (1 + K_{n,m,\alpha}) \|P_{\mathcal{W}_n}^\perp f^\alpha\|_{\mathcal{X}} \\ &\leq (2 + K_{n,m,\alpha}) \|f^\dagger - f^\alpha\|_{\mathcal{X}} + (1 + K_{n,m,\alpha}) \|P_{\mathcal{W}_n}^\perp f^\dagger\|_{\mathcal{X}}. \end{aligned}$$

To prove that, for fixed n and $\alpha > 0$ and m large enough, the constant $K_{n,m,\alpha}$ can be bounded by $\lambda_{\max}(\Theta_{\alpha,m}^{-2}) \lambda_{\min}^{-1}(\Theta_{\alpha,m}^{-2})$, we first define $\tilde{v}_k := F_\alpha(\sigma_k^2)^{-1} v_k$. This results in

$$K_{n,m,\alpha} = \|W_n (W_n^* \tilde{V}_m \tilde{V}_m^* W_n)^\dagger W_n^* \tilde{V}_m \tilde{V}_m^* P_{\mathcal{W}_n}^\perp\|.$$

As $\{v_k\}_{k \in \mathbb{N}}$ is an orthonormal basis, the system $\{\tilde{v}_k\}_{k \in \mathbb{N}}$ is a Riesz basis and therefore we obtain

$$K_{n,m,\alpha} \leq \|(\tilde{V}_m \tilde{V}_m^*)^{-1}\|_{\mathcal{V}_m \rightarrow \mathcal{V}_m} \|\tilde{V}_m \tilde{V}_m^*\|_{\mathcal{V}_m \rightarrow \mathcal{V}_m}.$$

For $f \in \mathcal{V}_m$, it follows that

$$\min_{j=1,\dots,m} F_\alpha(\sigma_j^2)^{-2} \|f\|^2 \leq \langle \tilde{V}_m \tilde{V}_m^* f, f \rangle = \sum_{j=1}^m F_\alpha(\sigma_j^2)^{-2} |\langle f, v_j \rangle|^2 \leq \max_{j=1,\dots,m} F_\alpha(\sigma_j^2)^{-2} \|f\|^2$$

resulting in

$$K_{n,m,\alpha} \leq \frac{\max_{j=1,\dots,m} F_\alpha(\sigma_j^2)^{-2}}{\min_{j=1,\dots,m} F_\alpha(\sigma_j^2)^{-2}},$$

as required \square

As already mentioned, the samples to which we have access in practice are given by $\eta^\delta = S^* g^\delta$. Based on this (possible noisy) information we construct an approximation $\gamma_{m,r}^\delta$ to the solution of (4.3) as follows:

$$\gamma_{m,r}^\delta = (U_m^* S_r S_r^* U_m)^\dagger U_m^* S_r S_r^* g^\delta \text{ and hence, } g_{m,r}^\delta = U_m (U_m^* S_r S_r^* U_m)^\dagger U_m^* S_r S_r^* g^\delta. \quad (4.7)$$

Consequently, combining (4.5) and (4.7), we obtain an approximation to the solution of (4.4) that can indeed be realized in practice,

$$f_{n,m,r}^{\alpha,\delta} = W_n (W_n^* V_m \Theta_{\alpha,m}^{-1} \Theta_{\alpha,m}^{-1} V_m^* W_n)^\dagger W_n^* V_m \Theta_{\alpha,m}^{-1} \Sigma_m U_m^* g_{m,r}^\delta. \quad (4.8)$$

As a simple consequence of the Theorems 3.1 and 4.1 we have the following:

Theorem 4.2. *If the assumptions made in Theorems 3.1 and 4.1 hold true, then*

$$\|f^\dagger - f_{n,m,r}^{\alpha,\delta}\| \leq \|f^\dagger - f_{n,m}^\alpha\| + K_{n,m,\alpha}^1 (\sigma_{m+1} (1 + K_{m,r}^2) \|P_{\mathcal{V}_m}^\perp f\| + K_{m,r}^3 \delta)$$

with constants defined by

$$K_{n,m,\alpha}^1 := \|W_n (W_n^* \tilde{V}_m \tilde{V}_m^* W_n)^\dagger W_n^* \tilde{V}_m V_m^* A^*\|, \quad K_{m,r}^2 := \|U_m (U_m^* S_r S_r^* U_m)^\dagger U_m^* S_r S_r^* U P_m^\perp\|$$

$$K_{m,r}^3 := \|U_m (U_m^* S_r S_r^* U_m)^\dagger U_m^* S_r S_r^*\|,$$

where $K_{n,m,\alpha}^1 \leq \frac{\lambda_{\max}(\Theta_{\alpha,m}^{-1})}{\lambda_{\min}(\Theta_{\alpha,m}^{-2})} \|A\|$, and, for sufficiently large r , $K_{m,r}^2, K_{m,r}^3 \leq 2\frac{D}{C}$. Moreover, if $\{s_k\}_{k \in \mathbb{N}}$ is an orthonormal basis, then, for fixed m , $K_{m,r}^2 \rightarrow 0$, $K_{m,r}^3 \rightarrow 1$ as $r \rightarrow \infty$.

Proof. First, with $\Sigma_m U_m^* = V_m^* A^*$, we observe that

$$\begin{aligned} f_{n,m,r}^{\alpha,\delta} &= W_n (W_n^* V_m \Theta_{\alpha,m}^{-1} \Theta_{\alpha,m}^{-1} V_m^* W_n)^\dagger W_n^* V_m \Theta_{\alpha,m}^{-1} \Sigma_m U_m^* (g - g + g_{m,r}^\delta) \\ &= f_{n,m}^\alpha + W_n (W_n^* V_m \Theta_{\alpha,m}^{-1} \Theta_{\alpha,m}^{-1} V_m^* W_n)^\dagger W_n^* V_m \Theta_{\alpha,m}^{-1} V_m^* A^* (g_{m,r}^\delta - g). \end{aligned}$$

Moreover, with the help of Theorem 3.1,

$$g - g_{m,r}^\delta = P_{\mathcal{U}_m}^\perp g + U_m (U_m^* S_r S_r^* U_m)^\dagger U_m^* S_r S_r^* P_{\mathcal{U}_m}^\perp g + U_m (U_m^* S_r S_r^* U_m)^\dagger U_m^* S_r S_r^* (g^\delta - g),$$

and since

$$P_{\mathcal{U}_m}^\perp g = (I - U P_m U^*) g = U (I - P_m) U^* g = U P_m^\perp \gamma = U P_m^\perp \Sigma V^* W \beta$$

it follows that

$$\begin{aligned} \|g - g_{m,r}^\delta\| &\leq \sigma_{m+1} (1 + \|U_m (U_m^* S_r S_r^* U_m)^\dagger U_m^* S_r S_r^* U P_m^\perp\|) \|P_{\mathcal{V}_m}^\perp f\| \\ &\quad + \|U_m (U_m^* S_r S_r^* U_m)^\dagger U_m^* S_r S_r^*\| \delta, \end{aligned}$$

This implies, with $\tilde{V}_m = V_m \Theta_{\alpha,m}^{-1}$, that

$$\begin{aligned} \|f^\dagger - f_{n,m,r}^{\alpha,\delta}\| &\leq \|f^\dagger - f_{n,m}^\alpha\| + \|W_n (W_n^* \tilde{V}_m \tilde{V}_m^* W_n)^\dagger W_n^* \tilde{V}_m V_m^* A^*\| \|g_{m,r}^\delta - g\| \\ &\leq \|f^\dagger - f_{n,m}^\alpha\| + K_{n,m,\alpha}^1 (\sigma_{m+1} (1 + K_{m,r}^2) \|P_{\mathcal{V}_m}^\perp f\| + K_{m,r}^3 \delta). \end{aligned}$$

Now, for m sufficiently large,

$$\begin{aligned} K_{n,m,\alpha}^1 &= \|W_n(W_n^* \tilde{V}_m \tilde{V}_m^* W_n)^\dagger W_n^* \tilde{V}_m V_m^* A^*\| \\ &= \|(\tilde{V}_m \tilde{V}_m^*)^{-1} \tilde{V}_m \tilde{V}_m^* W_n (W_n^* \tilde{V}_m \tilde{V}_m^* W_n)^\dagger W_n^* \tilde{V}_m V_m^* A^*\| \\ &\leq \|(\tilde{V}_m \tilde{V}_m^*)^{-1}\| \|\tilde{V}_m\| \|V_m^*\| \|A^*\| \leq \frac{\max_{j=1,\dots,m} F_\alpha(\sigma_j^2)^{-1}}{\min_{j=1,\dots,m} F_\alpha(\sigma_j^2)^{-2}} \|A\|. \end{aligned}$$

Moreover, if r is also sufficiently large we obtain

$$K_{m,r}^2 \leq \|U_m(U_m^* S_r S_r^* U_m)^\dagger U_m^* S_r S_r^*\| \leq 2\|(SS^*)^{-1}\| \|SS^*\|, \quad K_{m,r}^3 \leq 2\|(SS^*)^{-1}\| \|SS^*\|.$$

To see that $K_{m,r}^2 \rightarrow 0$ as $r \rightarrow \infty$ we observe that, since $\{s_k\}_{k \in \mathbb{N}}$ and $\{u_k\}_{k \in \mathbb{N}}$ span the same space, we have that $P_S U = U$, where $P_S = SS^*$. Thus,

$$U_m(U_m^* SS^* U_m)^\dagger U_m^* SS^* U P_m^\perp = U_m(U_m^* SS^* U_m)^\dagger P_m U^* U P_m^\perp = 0.$$

Also,

$$U_m(U_m^* S_r S_r^* U_m)^\dagger U_m^* S_r S_r^* U P_m^\perp \longrightarrow U_m(U_m^* SS^* U_m)^\dagger U_m^* SS^* U P_m^\perp, \quad r \rightarrow \infty$$

by the strong convergence of $S_r S_r^*$ to SS^* and compactness of U_m . The assertion about $K_{m,r}^2$ follows. \square

4.2 Non-Regularized Reconstruction

The approach in the previous section was essentially based on the normal equation $A^* A f = A^* g$. As an alternative, we now propose an approach based on directly utilizing the singular value decomposition of A . In other words, since $A = U \Sigma V^*$, we have

$$\eta = S^* g = S^* U \Sigma V^* W \beta = S^* U \gamma$$

and, as in the previous section, this results in two linear equations

$$S^* U \gamma = \eta \tag{4.9}$$

$$V^* W \beta = \Sigma^{-1} \gamma. \tag{4.10}$$

Based on equation (4.9) we construct an approximation to g ,

$$g_{m,r}^\delta = U_m(U_m^* S_r S_r^* U_m)^\dagger U_m^* S_r S_r^* g^\delta, \tag{4.11}$$

and with the help of equation (4.10) we construct

$$f_{n,m} = W_n(W_n^* V_m V_m^* W_n)^\dagger W_n^* V_m \Sigma_m^{-1} U_m^* g. \tag{4.12}$$

Combining (4.11) and (4.12) suggests the following approximation to f :

$$f_{n,m,r}^\delta = W_n(W_n^* V_m V_m^* W_n)^\dagger W_n^* V_m \Sigma_m^{-1} U_m^* g_{m,r}^\delta. \tag{4.13}$$

Theorem 4.3. *Let \mathcal{X}, \mathcal{Y} be separable Hilbert spaces and $A : \mathcal{X} \rightarrow \mathcal{Y}$ be a linear and bounded operator with singular system $\{\sigma_k, v_k, u_k\}$ (such that $\{v_k\}_{k \in \mathbb{N}}$ is an orthonormal system for $\mathcal{N}(A)^\perp$ and $\{u_k\}_{k \in \mathbb{N}}$ is an orthonormal system for $\mathcal{N}(A^*)^\perp$). Moreover, let $\mathcal{W} = \mathcal{N}(A)^\perp \subset \mathcal{X}$ and $\mathcal{S} = \mathcal{N}(A^*)^\perp \subset \mathcal{Y}$. Suppose $\{s_k\}_{k \in \mathbb{N}}$ and $\{w_k\}_{k \in \mathbb{N}}$ are frames for \mathcal{S} and \mathcal{W} respectively. Then, for the solution $f \in \mathcal{W} \subset \mathcal{X}$ and $n \in \mathbb{N}$, there exist $M, R \in \mathbb{N}$ such that, for all $m \geq M$ and $r \geq R$ the approximation $f_{n,m,r}^\delta$ exists and is unique. Moreover,*

$$\|f - f_{n,m,r}^\delta\|_{\mathcal{X}} \leq (1 + K_{n,m}^1) \|P_{\mathcal{W}_n}^\perp f\| + K_{m,r}^2 \|P_{\mathcal{V}_m}^\perp f\| + K_{m,r}^3 \frac{\delta}{\sigma_m}, \tag{4.14}$$

with

$$K_{n,m}^1 = \|W_n(W_n^* V_m V_m^* W_n)^\dagger W_n^* V_m V_m^* P_{\mathcal{W}_n}^\perp\|, \quad K_{m,r}^2 = \|U_m(U_m^* S_r S_r^* U_m)^\dagger U_m^* S_r S_r^* U P_m^\perp\|,$$

$$K_{m,r}^3 = \|U_m(U_m^* S_r S_r^* U_m)^\dagger U_m^* S_r S_r^*\|$$

In particular, $K_{n,m}^1 \rightarrow 0$ as $m \rightarrow \infty$, and, for sufficiently large r , $K_{m,r}^2, K_{m,r}^3 \leq 2\frac{D}{\sigma}$. Moreover, if $\{s_k\}_{k \in \mathbb{N}}$ forms an orthonormal system, then $K_{m,r}^2 \rightarrow 0$ as $r \rightarrow \infty$ and $K_{m,r}^3 \rightarrow 1$ as $r \rightarrow \infty$.

Proof. Following the same lines of the proof of Theorem 3.1, we obtain for (4.11),

$$\begin{aligned} g_{m,r}^\delta &= U_m(U_m^*S_rS_r^*U_m)^\dagger U_m^*S_rS_r^*(g+z) \\ &= P_{U_m}g + U_m(U_m^*S_rS_r^*U_m)^\dagger U_m^*S_rS_r^*P_{U_m}^\perp g + U_m(U_m^*S_rS_r^*U_m)^\dagger U_m^*S_rS_r^*z. \end{aligned}$$

Therefore, we have for (4.13) with $\Sigma_m^{-1}U_m^*P_{U_m}g = V_m^*f$,

$$\begin{aligned} f_{n,m,r}^\delta &= W_n(W_n^*V_mV_m^*W_n)^\dagger W_n^*V_mV_m^*(P_{W_n}f + P_{W_n}^\perp f) \\ &\quad + W_n(W_n^*V_mV_m^*W_n)^\dagger W_n^*V_m\Sigma_m^{-1}U_m^*U_m(U_m^*S_rS_r^*U_m)^\dagger U_m^*S_rS_r^*P_{U_m}^\perp g \\ &\quad + W_n(W_n^*V_mV_m^*W_n)^\dagger W_n^*V_m\Sigma_m^{-1}U_m^*U_m(U_m^*S_rS_r^*U_m)^\dagger U_m^*S_rS_r^*z. \end{aligned}$$

Note that we may argue exactly as in the proof of Theorem 3.1, by using the fact that $\{v_k\}_{k \in \mathbb{N}}$ is an orthonormal system for $\mathcal{N}(A)^\perp = \mathcal{W}$ and deduce that there is an $M \in \mathbb{N}$ such that

$$W_n(W_n^*V_mV_m^*W_n)^\dagger W_n^*V_mV_m^*P_{W_n}f = P_{W_n}f, \quad m \geq M. \quad (4.15)$$

Thus, by simply plugging in the expression for $f_{n,m,r}^\delta$, it follows that

$$\begin{aligned} \|f - f_{n,m,r}^\delta\| &\leq \|(I - W_n(W_n^*V_mV_m^*W_n)^\dagger W_n^*V_mV_m^*)P_{W_n}^\perp f\| \\ &\quad + \|W_n(W_n^*V_mV_m^*W_n)^\dagger W_n^*V_m\Sigma_m^{-1}U_m^*U_m(U_m^*S_rS_r^*U_m)^\dagger U_m^*S_rS_r^*P_{U_m}^\perp g\| \\ &\quad + \|W_n(W_n^*V_mV_m^*W_n)^\dagger W_n^*V_m\Sigma_m^{-1}U_m^*U_m(U_m^*S_rS_r^*U_m)^\dagger U_m^*S_rS_r^*z\| \quad m \geq M. \end{aligned} \quad (4.16)$$

To get the asserted bounds we shall bound separately the three terms on the right-hand side of (4.16). The first term is obvious, so we focus on the second and third. To bound the second term observe that

$$\|W_n(W_n^*V_mV_m^*W_n)^\dagger W_n^*V_m\Sigma_m^{-1}\| = \|(V_mV_m^*)^{-1}V_mP_{\text{Ran}(V_m^*W_n)}\Sigma_m^{-1}\| \leq \frac{1}{\sigma_m},$$

and that

$$P_{U_m}^\perp g = (I - UP_mU^*)g = U(I - P_m)U^*g = UP_m^\perp \gamma = UP_m^\perp \Sigma V^*W\beta.$$

In particular, since $\sigma_j \geq \sigma_{j+1}$ for all $j \in \mathbb{N}$, it follows that

$$\begin{aligned} &\|W_n(W_n^*V_mV_m^*W_n)^\dagger W_n^*V_m\Sigma_m^{-1}U_m^*U_m(U_m^*S_rS_r^*U_m)^\dagger U_m^*S_rS_r^*P_{U_m}^\perp g\| \\ &\leq \|W_n(W_n^*V_mV_m^*W_n)^\dagger W_n^*V_m\Sigma_m^{-1}\| \|U_m(U_m^*S_rS_r^*U_m)^\dagger U_m^*S_rS_r^*UP_m^\perp\| \|P_m^\perp \Sigma V^*W\beta\| \\ &\leq \frac{\sigma_{m+1}}{\sigma_m} \|U_m(U_m^*S_rS_r^*U_m)^\dagger U_m^*S_rS_r^*UP_m^\perp\| \|P_m^\perp V^*W\beta\| \\ &\leq \|U_m(U_m^*S_rS_r^*U_m)^\dagger U_m^*S_rS_r^*UP_m^\perp\| \|P_{V_m}^\perp f\|. \end{aligned} \quad (4.17)$$

As for the third term we use (4.15) and the assumption on z and get that

$$\begin{aligned} &\|W_n(W_n^*V_mV_m^*W_n)^\dagger W_n^*V_m\Sigma_m^{-1}U_m^*U_m(U_m^*S_rS_r^*U_m)^\dagger U_m^*S_rS_r^*z\| \\ &\leq \frac{1}{\sigma_m} \|U_m(U_m^*S_rS_r^*U_m)^\dagger U_m^*S_rS_r^*\| \delta. \end{aligned} \quad (4.18)$$

Hence, by (4.16), (4.17) and (4.18) we obtain

$$\begin{aligned} \|f - f_{n,m,r}^\delta\| &\leq (1 + \|W_n(W_n^*V_mV_m^*W_n)^\dagger W_n^*V_mV_m^*P_{W_n}^\perp\|) \|P_{W_n}^\perp f\| \\ &\quad + \|U_m(U_m^*S_rS_r^*U_m)^\dagger U_m^*S_rS_r^*UP_m^\perp\| \|P_{V_m}^\perp f\| \\ &\quad + \|U_m(U_m^*S_rS_r^*U_m)^\dagger U_m^*S_rS_r^*\| \frac{\delta}{\sigma_m}, \end{aligned}$$

which yields (4.14).

To see that $K_{n,m}^{-1} \rightarrow 0$ as $m \rightarrow \infty$ we can simply use Theorem 3.1 since, by assumption, $\{v_k\}_{k \in \mathbb{N}}$ and $\{w_k\}_{k \in \mathbb{N}}$ span the same space and $\{v_k\}_{k \in \mathbb{N}}$ is an orthonormal basis. Similarly, the statement about $K_{m,r}^3$ when $\{s_k\}_{k \in \mathbb{N}}$ is orthonormal follows from Theorem 3.1. To see that $K_{m,r}^2 \rightarrow 0$ as $r \rightarrow \infty$ we observe that, since $\{s_k\}_{k \in \mathbb{N}}$ and $\{u_k\}_{k \in \mathbb{N}}$ span the same space, we have $P_S U = U$. Thus,

$$U_m (U_m^* S S^* U_m)^\dagger U_m^* S S^* U P_m^\perp = U_m (U_m^* S S^* U_m)^\dagger P_m U^* U P_m^\perp = 0.$$

Also,

$$U_m (U_m^* S_r S_r^* U_m)^\dagger U_m^* S_r S_r^* U P_m^\perp \longrightarrow U_m (U_m^* S S^* U_m)^\dagger U_m^* S S^* U P_m^\perp, \quad r \rightarrow \infty$$

by the strong convergence of $S_r S_r^*$ to $S S^*$ and compactness of U_m . The assertion about $K_{m,r}^2$ follows. \square

4.3 Numerical Examples

In this section we will test the frameworks proposed in the previous subsections. First, we discuss a one-dimensional example for which we analyze the suggested regularized and non-regularized reconstruction methods. Thereafter, we consider a two-dimensional experiment. The goal is to verify that we can achieve, even in the presence of noise, a reasonable reconstruction by the proposed sampling-recovery technique.

Example 4.4. In order illustrate the proposed sampling theorems (Theorem 4.1, 4.2, and 4.3), we consider the linear operator $A : L_2([0, 1]) \rightarrow L_2([0, 1])$ defined by

$$g(t) = Af(t) = \int_0^t f(s) ds,$$

with singular system $\{\sigma_k, v_k, u_k\}$ given by

$$\sigma_k = \frac{1}{(k + 1/2)\pi}, \quad v_k = \sqrt{2} \cos(k + 1/2)\pi t, \quad u_k = \sqrt{2} \sin(k + 1/2)\pi t.$$

Note that $\{v_k\}_{k \in \mathbb{N}}$ and $\{u_k\}_{k \in \mathbb{N}}$ form orthonormal systems for $L_2([0, 1])$. To keep technicalities at a reasonable level, we choose the Fourier basis as both the recovery system $\{w_k\}_{k \in \mathbb{Z}}$ and sampling system $\{s_k\}_{k \in \mathbb{Z}}$, i.e.

$$w_k(t) = e^{2\pi i k t} \quad \text{and} \quad s_k(t) = e^{2\pi i k t}.$$

Let the signal f to be reconstructed be defined by $f(t) = \cos 2\pi t$. Consequently, f can be expanded as follows,

$$f(t) = W\beta = \sum_{k \in \mathbb{Z}} \beta_k e^{2\pi i k t} = \frac{1}{2} e^{2\pi i (-1)t} + \frac{1}{2} e^{2\pi i (+1)t}.$$

Consequently, $\beta_{-1} = 1/2$, $\beta_1 = 1/2$, and $\beta_k = 0$ for $k \in \mathbb{Z} \setminus \{-1, +1\}$. Moreover, the data g are given through $g(t) = Af(t) = 1/(2\pi) \sin 2\pi t$. In this particular example we also have explicit expression for all further required quantities,

$$\begin{aligned} \gamma &= U^* g = \{\gamma_l\}_{l \in \mathbb{N}} = \left\{ \frac{4\sqrt{2} \cos(l\pi)}{\pi^2(4l^2 + 4l - 15)} \right\}_{l \in \mathbb{N}} \\ \eta &= S^* g = \{\eta_k\}_{k \in \mathbb{Z}} \text{ with } \eta_{-1} = \frac{-i}{4\pi}, \eta_{+1} = \frac{i}{4\pi} \text{ and } \eta_k = 0, k \neq \pm 1 \\ V^* W &= \left(\frac{\sqrt{2}((l\pi + \pi/2) \cos(l\pi) - 2\pi i k)}{(2\pi i k)^2 + (l\pi + \pi/2)^2} \right)_{l \in \mathbb{N}, k \in \mathbb{Z}} \\ S^* U &= \left(\frac{(-1)^{l+1} \sqrt{2}((l\pi + \pi/2) \cos(l\pi) + 2\pi i k)}{(2\pi i k)^2 + (l\pi + \pi/2)^2} \right)_{k \in \mathbb{Z}, l \in \mathbb{N}}. \end{aligned}$$

The approximations to f from the r samples $S_r^* g^\delta$ may now be derived by

$$\begin{aligned} f_{n,m,r}^{\alpha,\delta} &= W_n (W_n^* V_m \Theta_{\alpha,m}^{-1} \Theta_{\alpha,m}^{-1} V_m^* W_n)^\dagger W_n^* V_m \Theta_{\alpha,m}^{-1} \Sigma_m (U_m^* S_r S_r^* U_m)^\dagger U_m^* S_r S_r^* g^\delta \quad \text{and} \\ f_{n,m,r}^\delta &= W_n (W_n^* V_m V_m^* W_n)^\dagger W_n^* V_m \Sigma_m^{-1} (U_m^* S_r S_r^* U_m)^\dagger U_m^* S_r S_r^* g^\delta, \end{aligned}$$

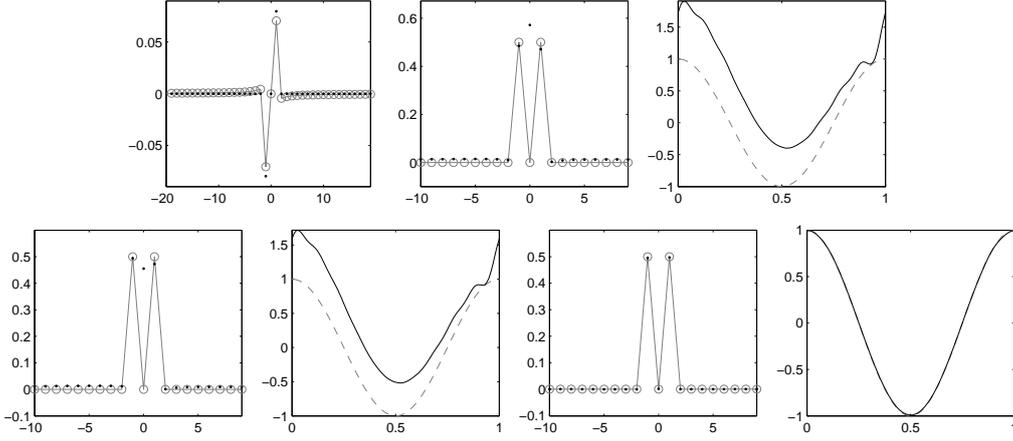


Figure 5: Recovery results for $\varepsilon_{rel} = 0\%$. Top (from left to right): $\eta^\delta = S_r^*(g+z)$ (\cdot) and $S_r^*U_m\gamma$ (\circ), $\beta_{20,30,40}^\delta$ (\cdot) and β (\circ), $f_{20,30,40}^\delta = W_n\beta_{20,30,40}^\delta$ ($-$) and f ($-$). Bottom (from left to right): $\beta_{20,30,40}^{0,\delta}$ (\cdot) and β (\circ), $f_{20,30,40}^{0,\delta} = W_n\beta_{20,30,40}^{0,\delta}$ ($-$) and f ($-$), $\beta_{20,30,40}^{\alpha_{opt},\delta}$ (\cdot) and β (\circ), $f_{20,30,40}^{\alpha_{opt},\delta} = W_n\beta_{20,30,40}^{\alpha_{opt},\delta}$ ($-$) and f ($-$).

where we focus here on Tikhonov regularization, i.e. the entries in $\Theta_{\alpha,m}$ are given by $F_\alpha(\sigma_k^2) = 1/(\alpha + \sigma_k^2)$. We discuss now several different recovery scenarios.

In the first case we choose a fixed (and reasonable) setting for n , m and r and vary the noise level δ and compare the recovery quality of $f_{n,m,r}^{\alpha,\delta}$ and $f_{n,m,r}^\delta$ while experimentally tuning the regularization parameter α towards optimal recovery. This experiment shall show that for a fixed number of data samples and coefficients in the series expansion of the solution an optimal choice of regularization parameter induces a substantially improved recovery.

In the second case we fix the number n of coefficients in series expansion of the solution and try to find for different noise levels δ reasonable integers m and r to derive $f_{n,m,r}^\delta$. For the same numbers m and r we then experimentally determine an optimal α to compute $f_{n,m,r}^{\alpha,\delta}$. This experiment shall show that a reasonable choice of m and r may feasibly stabilize the recovery and providing approximations that cannot be significantly improved by a fine tuning of α .

First case: vary $z = g - g^\delta$ such the relative error $\varepsilon_{rel} = 100 \cdot \|z\|/\|g\|$ is 0%, 5% and 10% and let $n = 20$, $m = 30$ and $r = 40$. The numerical results are illustrated in the following table and visualized in Figures 5,6, and 7.

$\varepsilon_{rel}, \delta$	$\ f - f_{20,30,40}^\delta\ $	$\ f - f_{20,30,40}^{0,\delta}\ $	$\ f - f_{20,30,40}^{\alpha_{opt},\delta}\ $	α_{opt}	Fig.
0%, 0	0.6262	0.4995	0.0071	0.00017	5
5%, 0.0056	1.1738	0.9728	0.1536	0.00037	6
10%, 0.0113	1.7593	1.5268	0.2265	0.00061	7

Second case: we first fix $n = 10$ and ask then, for different relative errors $\varepsilon_{rel} \in \{0\%, 5\%, 10\%\}$, for an adequate choice (numerically determined) of m and r in order to derive an optimal approximation $f_{n,m,r}^\delta$. Then, we try by fine tuning α to obtain with $f_{n,m,r}^{\alpha,\delta}$ a comparable or possibly better approximation. The results are documented in the following table. The illustrations of this experiment are given in Figure 8 (the illustrations for $\varepsilon_{rel} = 0\%$ are not provided since there is no visual difference).

$\varepsilon_{rel}, \delta$	m	r	$\ f - f_{10,m,r}^\delta\ $	$\ f - f_{10,m,r}^{0,\delta}\ $	$\ f - f_{10,m,r}^{\alpha_{opt},\delta}\ $	α_{opt}	Fig.
0%, 0.0	10	1000	0.002114839173	0.002114839160	0.000112	0.0000035	-
5%, 0.0042	40	100	0.0303	0.0433	0.0371	0.000025	8
10%, 0.0075	40	80	0.1044	0.2990	0.2732	0.0001	8

Example 4.5. In the second example we discuss the Radon transform,

$$R(\tau, \omega) = \int_{-\sqrt{1-\tau^2}}^{+\sqrt{1-\tau^2}} f(\tau\omega + t\omega^\perp) dt, \quad (4.19)$$

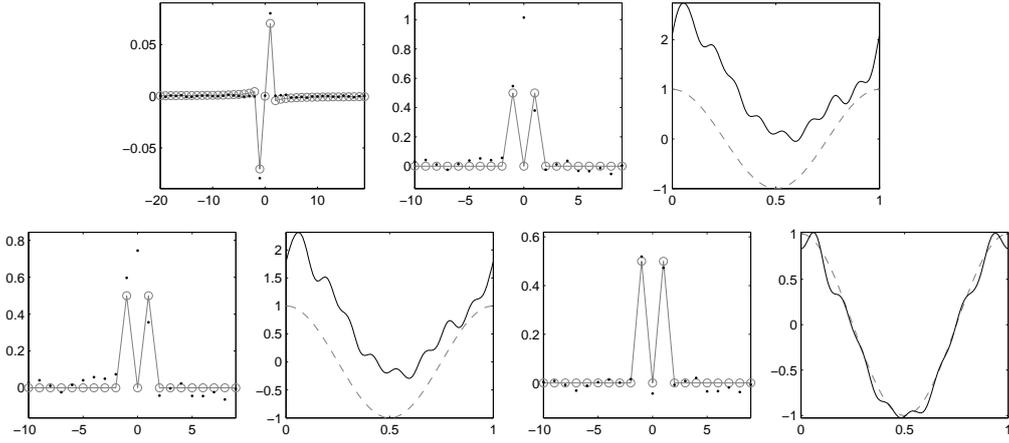


Figure 6: Recovery results for $\varepsilon_{rel} = 5\%$. Top (from left to right): $\eta^\delta = S_r^*(g+z)$ (\cdot) and $S_r^*U_m\gamma$ (\circ), $\beta_{20,30,40}^\delta$ (\cdot) and β (\circ), $f_{20,30,40}^\delta = W_n\beta_{20,30,40}^\delta$ ($-$) and f ($-$). Bottom (from left to right): $\beta_{20,30,40}^{0,\delta}$ (\cdot) and β (\circ), $f_{20,30,40}^{0,\delta} = W_n\beta_{20,30,40}^{0,\delta}$ ($-$) and f ($-$), $\beta_{20,30,40}^{\alpha_{opt},\delta}$ (\cdot) and β (\circ), $f_{20,30,40}^{\alpha_{opt},\delta} = W_n\beta_{20,30,40}^{\alpha_{opt},\delta}$ ($-$) and f ($-$).

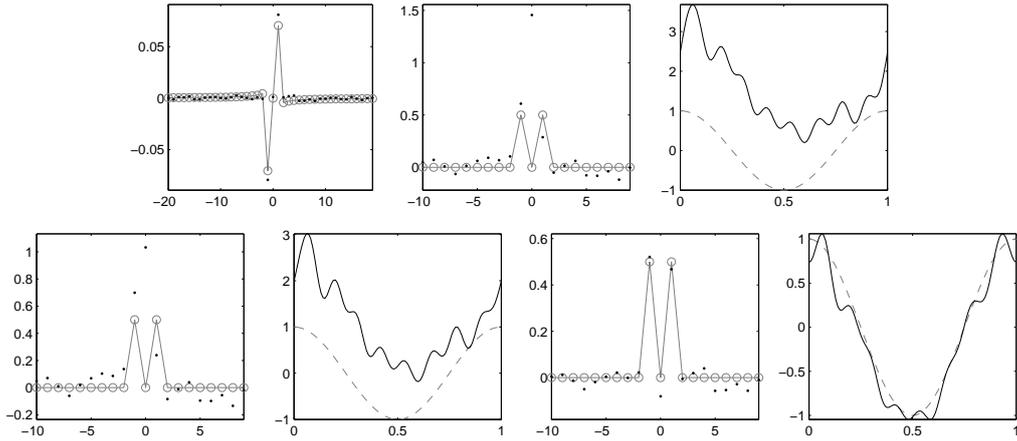


Figure 7: Recovery results for $\varepsilon_{rel} = 10\%$. Top (from left to right): $\eta^\delta = S_r^*(g+z)$ (\cdot) and $S_r^*U_m\gamma$ (\circ), $\beta_{20,30,40}^\delta$ (\cdot) and β (\circ), $f_{20,30,40}^\delta = W_n\beta_{20,30,40}^\delta$ ($-$) and f ($-$). Bottom (from left to right): $\beta_{20,30,40}^{0,\delta}$ (\cdot) and β (\circ), $f_{20,30,40}^{0,\delta} = W_n\beta_{20,30,40}^{0,\delta}$ ($-$) and f ($-$), $\beta_{20,30,40}^{\alpha_{opt},\delta}$ (\cdot) and β (\circ), $f_{20,30,40}^{\alpha_{opt},\delta} = W_n\beta_{20,30,40}^{\alpha_{opt},\delta}$ ($-$) and f ($-$).

where we have assumed that $\text{supp}(f) \subset D = \{x \in \mathbb{R}^2 : \|x\| \leq 1\}$, and $\omega \in S^1$, $\tau \in [-1, 1]$, see [23]. The map R is linear and continuous (with norm $\sqrt{4\pi}$) between $L_2(D)$ and $L_2([-1, 1] \times [0, 2\pi], g^{-1})$, with weight function $g(\tau) = \sqrt{1 - \tau^2}$. As a map between these spaces, the Radon transform has the following singular system (for details see again [23]),

$$\{(v_{ml}, u_{ml}, \sigma_{ml}) : m \geq 0, l \in \mathbb{Z} : |l| \leq m, m+l \text{ even}\},$$

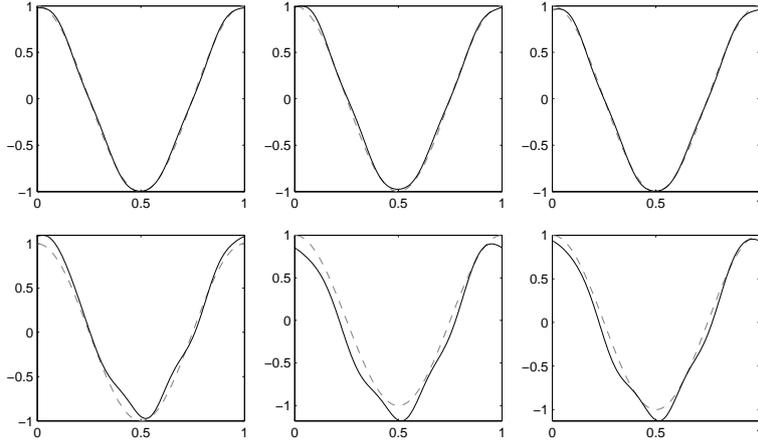


Figure 8: Experimental results for $\varepsilon_{rel} = 5\%$: $f_{10,40,100}^\delta$ (t.l.), $f_{10,40,100}^{0,\delta}$ (t.m.), $f_{10,40,100}^{0.000025,\delta}$ (t.r.), and for $\varepsilon_{rel} = 10\%$: $f_{10,40,100}^\delta$ (b.l.), $f_{10,40,100}^{0,\delta}$ (b.m.), $f_{10,40,100}^{0.0001,\delta}$ (b.r.). In all subfigures the dashed line (-) represents the true solution f .

$$v_{m,l}(x) = \begin{cases} \sqrt{\frac{m+1}{\pi}} \|x\|^{|l|} P_{(m-|l|)/2}^{(0,|l|)}(2\|x\|^2 - 1) Y_l(x/\|x\|) & \|x\| \leq 1 \\ 0 & \|x\| > 1 \end{cases}$$

$$u_{m,l}(\tau, \omega) = \begin{cases} \frac{1}{\pi} g(\tau) U_m(\tau) Y_l(\omega) & |\tau| \leq 1 \\ 0 & |\tau| > 1 \end{cases}$$

$$\sigma_{m,l} = 2\sqrt{\frac{\pi}{m+1}}$$

where

$$P_n^{(\alpha,\beta)}(x) = \frac{\Gamma(\alpha+n+1)}{n!\Gamma(\alpha+\beta+n+1)} \sum_{m=0}^n \binom{n}{m} \frac{\Gamma(\alpha+\beta+n+m+1)}{\Gamma(\alpha+m+1)} \left(\frac{x-1}{2}\right)^m,$$

$$U_m(\tau) = \frac{\sin((m+1)\arccos(\tau))}{\sin(\arccos(\tau))}.$$

Hence, for each $f \in L_2(D)$, we have $Rf = \sum_{m,l} \sigma_{ml} \langle f, v_{ml} \rangle_{L_2(D)} u_{ml}$. We choose as recovery system for $L_2(D) = L_2(rdrd\theta, [0, 1] \times [0, 2\pi])$ the separable Haar basis on $[0, 1] \times [0, 2\pi]$,

$$w_\lambda(r, \theta) = \psi_{\lambda_1}(r) \psi_{\lambda_2}(\theta), \quad \lambda_i = (q_i, j, k_i),$$

where q_i prescribes the species of the wavelet ($q_i = 0$ - generator, $q_i = 1$ - corresponding wavelets, $i = 1, 2$), $j \in \mathbb{Z}$ the scales, and $(k_1, k_2) \in I$ the translations. Then, we obtain

$$\begin{aligned} \langle f, v_{ml} \rangle_{L_2(D)} &= \int_D f(x) v_{ml}(x) dx = \int_0^{2\pi} \int_0^1 f(r \cos \theta, r \sin \theta) \bar{v}_{ml}(r \cos \theta, r \sin \theta) r dr d\theta \\ &= \int_0^{2\pi} \int_0^1 f(r \cos \theta, r \sin \theta) \sqrt{\frac{m+1}{\pi}} r^{|l|} P_{(m-|l|)/2}^{(0,|l|)}(2r^2 - 1) e^{-il\theta} r dr d\theta \\ &= \sum_\lambda \beta_\lambda \int_0^{2\pi} \psi_{\lambda_2}(\theta) e^{-il\theta} d\theta \sqrt{\frac{m+1}{\pi}} \int_0^1 \psi_{\lambda_1}(r) r^{|l|+1} P_{(m-|l|)/2}^{(0,|l|)}(2r^2 - 1) dr \\ &= \sum_\lambda \beta_\lambda (V^* W)_{\lambda, ml}. \end{aligned}$$

As sampling system, we choose an orthonormal Fourier-Mellin-type basis, $\{w_{n,k}\}_{(n,k) \in \mathbb{N} \times \mathbb{Z}}$, to span $L_2([-1, 1] \times [0, 2\pi], g^{-1})$, which we define by

$$s_{n,k}(\tau, \theta) = \frac{1}{4} \sqrt{\frac{\tau+1}{\alpha_n \pi}} Q_n((\tau+1)/2) e^{i\theta k} g^{1/2}(\tau), \quad (4.20)$$

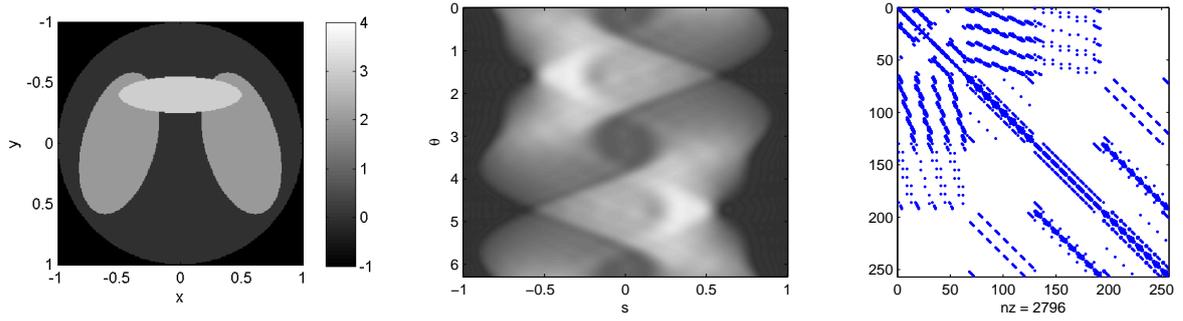


Figure 9: Left: phantom function f on D , middle: Radon transform Rf for $m \leq 30$ (resulting in 496 singular functions), right: matrix $W_{256}^* V_{496} \Theta_{\alpha, 496}^{-2} W_{256} V_{496}^* W_{256}$, where for the wavelet system the scale is limited to $2 \leq j \leq 3$ (resulting in 256 basis functions).

where

$$\alpha_n = \frac{1}{2(n+1)}, \quad Q_n(\tau) = \sum_{p=0}^n \alpha_{n,p} \tau^p, \quad \alpha_{n,p} = (-1)^{n+1} \frac{(n+p+1)!}{(n-p)! p! (p+1)!}.$$

Therefore,

$$(S^*U)_{nk,ml} = \delta_{lk} \cdot \frac{1}{2} \int_{-1}^1 U_m(\tau) \sqrt{\frac{\tau+1}{\alpha_n \pi}} Q_n((\tau+1)/2) g^{1/2}(\tau) d\tau.$$

The phantom function f to be recovered is now simulated on D by placing N ellipses,

$$E^k = \left\{ x \in \mathbb{R}^2 : \left\| \begin{pmatrix} a^k & 0 \\ 0 & b^k \end{pmatrix} \begin{pmatrix} \cos \nu^k & \sin \nu^k \\ -\sin \nu^k & \cos \nu^k \end{pmatrix} (x - x^k) \right\| \leq r^k \right\}, \quad k = 1, \dots, N,$$

through,

$$f^0(x) = 0, \quad f^{n+1}(x) = f^n(x) \chi_{D \setminus E^{n+1}}(x) + \xi^{n+1} \chi_{E^{n+1}}(x), \quad n = 0, \dots, N-1 \text{ and } f(x) := f^N(x).$$

The k -th ellipse is specified by a set of parameters $\Pi^k = (x^k, r^k, \nu^k, a^k, b^k, \xi^k)$, where x^k determines the localization, r^k the radius, ν^k the orientation, a^k, b^k the semi-axes, and ξ^k the plateau height.

In our particular example we selected three ellipses,

$$\begin{aligned} E^1 &: \Pi^1 = (0.5, 0.0, 0.3, -\pi/12, 1, 0.5, 2) \\ E^2 &: \Pi^2 = (-0.5, 0.0, 0.3, \pi/12, 1, 0.5, 2) \\ E^3 &: \Pi^3 = (0, -0.4, 0.3, \pi/2, 2, 0.6, 3), \end{aligned}$$

resulting in a phantom function $f(x) = f^3(x)$ which visualized in figure 9. The resolutions (which can be made as fine as desired) to represent f (on a cartesian and/or polar grid) as well as Rf are restricted in our computational experiments to equispaced grids of size 256×256 and 512×512 . This is of course not fine enough when significantly increasing the number of recovery, singular and sampling functions. In particular, the singular and sampling functions contain oscillatory components that indeed require a much finer resolution. But as we focus here on exemplarily documenting the applicability of the proposed approach, we restrict ourselves to problem dimensions that cause no extra sophistication when dealing with very large systems. By (4.8) an approximation to f is now obtained through

$$f_{n,m,r}^{\alpha,\delta} = W_n (W_n^* V_m \Theta_{\alpha,m}^{-1} \Theta_{\alpha,m}^{-1} V_m^* W_n)^\dagger W_n^* V_m \Theta_{\alpha,m}^{-1} \Sigma_m (U_m^* S_r S_r^* U_m)^\dagger U_m^* S_r S_r^* g^\delta.$$

We have derived $f_{n,m,r}^{\alpha,\delta}$ within the following scenarios, for visual inspection see figure 10,

	n (wavelet functions)	m (singular functions)	r (sampling functions)	$E(f, f_{n,m,r}^{\alpha,\delta})$, rel. recovery error
scenario 1	1024 ($2 \leq j \leq 4$)	1326	1681	22,03 %
scenario 2	4096 ($2 \leq j \leq 5$)	4186	4225	15,62 %
scenario 3	16384 ($2 \leq j \leq 6$)	16471	16641	11,40 %

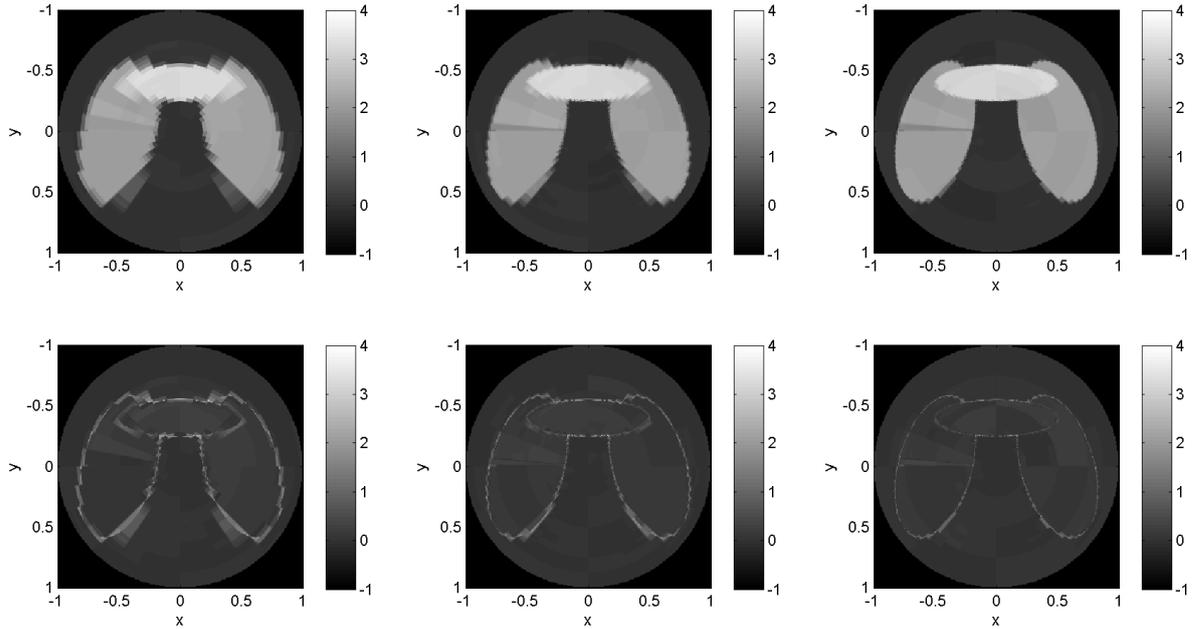


Figure 10: Top row (from left to right): recoveries of f by $f_{1024,1326,1681}^{\alpha,\delta}$ (scenario 1), $f_{4096,4186,4225}^{\alpha,\delta}$ (scenario 2), $f_{16384,16471,16641}^{\alpha,\delta}$ (scenario 3), where the relative error is $\varepsilon_{rel} = 5\%$ and the corresponding Tikhonov stabilization is fine tuned by $\alpha = 0.00001$. Bottom row (from left to right): modulus of difference between f and $f_{1024,1326,1681}^{\alpha,\delta}$, $f_{4096,4186,4225}^{\alpha,\delta}$, and $f_{16384,16471,16641}^{\alpha,\delta}$.

In our particular example the relative data error is $\varepsilon_{rel} = 5\%$ and the corresponding Tikhonov stabilization is fine tuned by $\alpha = 0.00001$. The relative recovery error is defined in this experiment by

$$E(f, f_{n,m,r}^{\alpha,\delta}) = \frac{\|f - f_{n,m,r}^{\alpha,\delta}\|_{L_2(D)}}{\|f\|_{L_2(D)}}.$$

5 Conclusions and challenges

The purpose of this paper was to extend generalized sampling to sampling and reconstructions in frames, and to inverse and ill-posed problems. The key component is to allow the various parameters (number of samples etc) to vary independently. When done appropriately, it is then possible to prove sampling theorems which give guaranteed error bounds for the reconstruction in terms of constants that can be estimated numerically.

An important theme of this work, as discussed earlier in the paper, is that finite-dimensional tools are not always well suited for formally infinite-dimensional problems. As such, this paper marks only one step in a much larger project on how to reconstruct in infinite dimensions, with many directions for future work. Several of these directions are now described.

All the key theorems in this paper rely on the existence of quantities (e.g. $K_{n,m}$ or $K_{n,m}^1$, $K_{m,r}^2$, $K_{m,r}^3$) which can be made small by an appropriate choice of parameters. In Section 3.4 an insight was given as to how to numerically estimate the quantity $K_{n,m}$ in the case of generalized sampling with frames, and thus how to compute the so-called stable sampling rate. It is a topic of ongoing investigation to extend this notion to the GS framework for inverse problems of Section 4. Note that the setting here is substantially more complicated – we have to deal with additional parameters, including possibly the regularization α . However, the aim of future work is to introduce explicitly computable sampling rates for such problems, which allow one to determine precisely how to balance such parameters. Much of this will be based on the more recent developments of [6] for GS with bases.

An important ill-posed problem is that of X-Ray tomography, which involves inverting the Radon transform. Although the GS framework introduced in Section 4, based on regularization of a compact operator, can be applied to this problem, this is not the only possible approach. Indeed, the Radon transform has a number of key features than one may hope to incorporate into the GS framework in order to get a better reconstruction for this particular problem. This is also a topic of future investigations.

As mentioned at the beginning of this paper, GS can be combined with convex optimization tools to give new techniques, and a completely new theory, for compressed sensing in infinite dimensions [1, 5]. One may therefore ask, is it possible to incorporate the developments in this paper to allow for sparse recovery for inverse and ill-posed problems? In the near future we intend to investigate this question, with one aim being to extend the previous work of [19] on this topic.

References

- [1] B. Adcock and A. C. Hansen. Generalized sampling and infinite dimensional compressed sensing. *Technical report NA2011/02, DAMTP, University of Cambridge*, 2011.
- [2] B. Adcock and A. C. Hansen. Reduced consistency sampling in Hilbert spaces. In *Proceedings of the 9th International Conference on Sampling Theory and Applications*, 2011.
- [3] B. Adcock and A. C. Hansen. A generalized sampling theorem for stable reconstructions in arbitrary bases. *J. Fourier Anal. Appl.*, 18(4):685–716, 2012.
- [4] B. Adcock and A. C. Hansen. Stable reconstructions in Hilbert spaces and the resolution of the Gibbs phenomenon. *Appl. Comput. Harmon. Anal.*, 32(3):357–388, 2012.
- [5] B. Adcock, A. C. Hansen, E. Herrholz, and G. Teschke. Generalized sampling, infinite-dimensional compressed sensing, and semi-random sampling for asymptotically incoherent dictionaries. *Preprint*, 2011.
- [6] B. Adcock, A. C. Hansen, and C. Poon. Beyond consistent reconstructions: optimality and sharp bounds for generalized sampling, and application to the uniform resampling problem. *Technical report NA2011/10, DAMTP, University of Cambridge*, 2012.
- [7] A. Aldroubi. Oblique projections in atomic spaces. *Proc. Amer. Math. Soc.*, 124(7):2051–2060, 1996.
- [8] 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.
- [9] O. Christensen. Frames and the projection method. *Appl. Comput. Harmon. Anal.*, 1(1):50–53, 1993.
- [10] T. Dvorkind and Y. C. Eldar. Robust and consistent sampling. *IEEE Signal Process. Letters*, 16(9):739–742, 2009.
- [11] Y. C. Eldar. Sampling with arbitrary sampling and reconstruction spaces and oblique dual frame vectors. *Journal of Fourier Analysis and Applications*, 9(1):77–96, 2003.
- [12] Y. C. Eldar. Sampling without input constraints: Consistent reconstruction in arbitrary spaces. In A. I. Zayed and J. J. Benedetto, editors, *Sampling, Wavelets and Tomography*, pages 33–60. Boston, MA: Birkhäuser, 2004.
- [13] Y. C. Eldar and T. Michaeli. Beyond Bandlimited Sampling. *IEEE Signal Process. Mag.*, 26(3):48–68, 2009.
- [14] Y. C. Eldar and T. Werther. General framework for consistent sampling in Hilbert spaces. *Int. J. Wavelets Multiresolut. Inf. Process.*, 3(3):347, 2005.
- [15] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of Inverse Problems*. Kluwer, Dordrecht, 1996.
- [16] K. Gröchenig, Z. Rzesotnik, and T. Strohmer. Quantitative estimates for the finite section method. *Integral Equations Operator Theory*, to appear.
- [17] A. C. Hansen. On the approximation of spectra of linear operators on Hilbert spaces. *J. Funct. Anal.*, 254(8):2092–2126, 2008.
- [18] A. C. Hansen. On the solvability complexity index, the n-pseudospectrum and approximations of spectra of operators. *J. Amer. Math. Soc.*, 24(1):81–124, 2011.
- [19] E. Herrholz and G. Teschke. Compressive sensing principles and iterative sparse recovery for inverse and ill-posed problems. *Inverse Problems*, 26(12):125012, 2010.
- [20] T. Hrycak and K. Gröchenig. Pseudospectral Fourier reconstruction with the modified inverse polynomial reconstruction method. *J. Comput. Phys.*, 229(3):933–946, 2010.
- [21] A. J. Jerri. The Shannon sampling theorem – its various extensions and applications: A tutorial review. *Proc. IEEE*, 65(1565–1596), 1977.

- [22] H. J. Landau. Necessary density conditions for sampling and interpolation of certain entire functions. *Acta Math.*, 117:37–52, 1967.
- [23] A. K. Louis. *Inverse und schlecht gestellte Probleme*. Teubner, Stuttgart, 1989.
- [24] M. Unser. Sampling—50 years after Shannon. *Proc. IEEE*, 88(4):569–587, 2000.
- [25] M. Unser and A. Aldroubi. A general sampling theory for nonideal acquisition devices. *IEEE Trans. Signal Process.*, 42(11):2915–2925, 1994.
- [26] M. Unser and J. Zerubia. A generalized sampling theory without band-limiting constraints. *IEEE Trans. Circuits Syst. II.*, 45(8):959–969, 1998.