REDUCED CONSISTENCY SAMPLING IN HILBERT SPACES

Ben Adcock[1], *Anders C. Hansen*[2]

[1]Department of Mathematics, Simon Fraser University
[2]Department of Applied Mathematics and Theoretical Physics, University of Cambridge
Emails: ben_adcock@sfu.ca, a.hansen@damtp.cam.ac.uk

ABSTRACT

We present a new approach for the problem of reconstructing an element of a Hilbert space in an arbitrary reconstruction basis given its measurements with respect to certain sampling vectors. This procedure differs from more standard techniques in that it is not consistent: the samples of the reconstructed function are not equal to those of the original function. However, by dropping this requirement, we are able to obtain a method that is both numerically stable and which provides a reconstruction that is closer to optimal.

Keywords— Sampling theory, Riesz bases, finite sections

1. INTRODUCTION

Modern sampling concerns the problem of reconstructing an element $f \in \mathcal{H}$, where $\mathcal{H}$ is some separable Hilbert space, in a particular space $\mathcal{W}$, the *reconstruction space*, given its collection of samples with respect to a so-called *sampling space* $\mathcal{S}$. More specifically, if $s_1, s_2, \ldots$ are linearly independent sampling vectors with $\mathcal{S} = \overline{\operatorname{span}\{s_1, s_2, \ldots\}}$, the task is to compute an approximation $\tilde{f} \in \mathcal{W}$ from the measurements $c_j = \langle f, s_j \rangle$, $j = 1, 2, \ldots$, where $\langle \cdot, \cdot \rangle$ is the inner product on $\mathcal{H}$. If $w_1, w_2, \ldots$ are also linearly independent, with $\mathcal{W} = \overline{\operatorname{span}\{w_1, w_2, \ldots\}}$, then $\tilde{f}$ takes the form $\sum_{j=1}^{\infty} d_j w_j$ for some $d_j \in l^2(\mathbb{N})$.

An approach to compute $\tilde{f}$ was developed in [12], and later formalised by Eldar et al [4, 5]. Therein, the approximation $\tilde{f}$ is specified by enforcing the consistency conditions

$$\langle \tilde{f}, s_j \rangle = \langle f, s_j \rangle, \quad j = 1, 2, \ldots. \tag{1}$$

The approximation $\tilde{f}$ is referred to as *consistent* reconstruction of f. Under the additional assumption $\mathcal{H} = \mathcal{W} \oplus \mathcal{S}^{\perp}$ it is known that $\tilde{f}$ is unique. Moreover, if $\|\cdot\|$ is the norm associated with $\mathcal{H}$ and $P_{\mathcal{W}}$ is the orthogonal projection $\mathcal{H} \to \mathcal{W}$, then

$$\|f - P_{\mathcal{W}} f\| \leq \|f - \tilde{f}\| \leq (\cos \theta_{\mathcal{W}\mathcal{S}})^{-1} \|f - P_{\mathcal{W}} f\|,$$

where $\theta_{\mathcal{W}\mathcal{S}}$ is the angle between the subspaces $\mathcal{W}$ and $\mathcal{S}$:

$$\cos(\theta_{\mathcal{W}\mathcal{S}}) = \inf_{w \in \mathcal{W}, \|w\|=1} \|P_{\mathcal{S}} w\|.$$

Naturally, to implement such a scheme numerically we are limited to taking only a finite number of such samples, and to seeking a reconstruction in a finite dimensional subspace of $\mathcal{W}$. Hence, the finite dimensional version of this problem reads: compute $\tilde{f}_m \in \mathcal{W}_m = \operatorname{span}\{w_1, \ldots, w_m\}$ from the samples $c_j = \langle f, s_j \rangle$, $j = 1, 2, \ldots, m$. Once more, a reconstruction $\tilde{f}_m \in \mathcal{W}_m$ can be specified by imposing consistency

$$\langle \tilde{f}_m, s_j \rangle = \langle f, s_j \rangle, \quad j = 1, 2, \ldots, m, \tag{2}$$

leading to a linear system of equations $Ad = c$, where $c = (c_1, \ldots, c_m)^{\top}$, $d = (d_1, \ldots, d_m)^{\top}$,

$$A = \begin{pmatrix} \langle w_1, s_1 \rangle & \cdots & \langle w_m, s_1 \rangle \\ \vdots & \ddots & \vdots \\ \langle w_1, s_m \rangle & \cdots & \langle w_m, s_m \rangle \end{pmatrix}, \tag{3}$$

and

$$\tilde{f}_m = \sum_{k=1}^{m} d_k w_k. \tag{4}$$

Letting $f_m = P_{\mathcal{W}_m} f$, we find that

$$\|f - f_m\| \leq \|f - \tilde{f}_m\| \leq (\cos \theta_{\mathcal{W}_m \mathcal{S}_m})^{-1} \|f - f_m\|. \tag{5}$$

Naturally, by taking m sufficiently large, we hope to recover an approximation arbitrarily close to the solution of the infinite dimensional problem $\tilde{f}$.

Whilst this approach has proved extremely popular since its original development, there are several drawbacks that serve to motivate the particular framework we develop in this paper. More precisely, the method we present is designed to tackle the following two issues:

1. What if $\mathcal{W}_m \cap \mathcal{S}_m^{\top} \neq \{0\}$, so that $\tilde{f}_m$ may not exist.
2. What if $\|A^{-1}\|$ is large? Even if A^{-1} exists, this may make the method unsuitable in practice, due to likely increased sensitivity to both round-off error and noise, as well as the increased cost required to solve the ill-conditioned system.

At first glance, it appears plausible that these issues may not have impact, since the infinite-dimensional problem (1) is well-posed. However, this intuition is incorrect; a fact that can be explained quite easily with an operator-theoretic perspective. Notice that the matrix A given by (3) is precisely an $m \times m$ *finite section* of the infinite matrix $A_{\infty} = \{\langle w_j, s_i \rangle\}_{i,j=1}^{\infty}$ of

m	25	50	100	200
(i)	3.4e03	1.0e08	6.2e16	3.2e34
(ii)	2.6e00	3.6e01	1.8e02	2.8e06
(iii)	1.1e-03	2.9e-06	1.2e-11	4.4e-18

Table 1. The quantities (i) $\|A^{-1}\|$, (ii) $\|f - \tilde{f}_m\|$ and (iii) $\|f - f_m\|$.

the infinite-dimensional problem (1). Finite sections of infinite matrices have been extensively studied in the last several decades (see, for example, [7] and the references therein). Unfortunately, the properties of the finite section may contrast starkly with those of the infinite-dimensional operator. For example, even if the infinite-dimensional problem is nonsingular, an $m \times m$ finite section A may not be invertible for any m. Moreover, even if A is nonsingular and $\|A_\infty^{-1}\| < \infty$, it may well be the case that $\|A^{-1}\| \to \infty$ as $m \to \infty$. In addition, if $A_\infty c = d$, where $c = (c_1, c_2, \ldots)^\top$, $d = (d_1, d_2, \ldots)^\top$, and $Ac_m = d_m$, where $d_m = (c_1, \ldots, c_m)^\top$, then it is not guaranteed that $c_m \to c$ as $m \to \infty$. Thus, in context of the sampling problem (2), even if $\tilde{f}_m$ exists, it may not be computable in a stable manner and may not converge to $\tilde{f}$.

To avoid these issues, stringent restrictions are required on the infinite matrix A_∞, such as positive self-adjointness. However, in many circumstances of interest, the sampling problem (1) for example, such conditions are not satisfied.

Let us illustrate these effects with an example. Suppose that $\mathcal{H} = L^2(-1,1)$ and $s_j(x) = \mathrm{e}^{\mathrm{i}j\pi x}$, so that $\mathcal{S}$ is the standard Fourier basis. This sampling basis represents the most commonly occurring type of reconstruction problem: namely, the recovery of a function from samples of its Fourier transform. Indeed, the classical Shannon Sampling Theorem [11], the foundation of modern information theory, deals precisely with this scenario.

Naturally, if f has high smoothness, it makes sense to reconstruct in a polynomial basis. To this end, suppose that w_j is the j^{th} orthogonal polynomial on $[-1,1]$ (i.e. the j^{th} Legendre polynomial). With this to hand, we now compute $\tilde{f}_m$ as in (4) via (2). In Table 1 we present the value $\|A^{-1}\|$ for various m. As is evident, this quantity grows geometrically in m (i.e. $\|A^{-1}\| \sim \gamma^m$ for some $\gamma > 1$), thus making the reconstruction procedure highly unstable even for moderate values of m. Moreover, not only does this make the method hard to implement numerically, it can also prohibit the convergence of $\tilde{f}_m$ as $m \to \infty$. In Table 1 we also display the errors $\|f - \tilde{f}_m\|$ and $\|f - f_m\|$, where $f(x) = \frac{1}{1+16x^2}$. Clearly, the reconstruction $\tilde{f}_m$ does not converge to f. Note that this issue is not due to the choice of reconstruction space $\mathcal{W}$–indeed, the orthogonal projection $f_m = P_{\mathcal{W}_m} f$ converges geometrically fast to f–but is caused solely by the particular method used to reconstruct.

Another key feature of (1) is so-called perfect reconstruction: if $f \in \mathcal{W}_m$, then f is reconstructed exactly by $\tilde{f}_m$. Of course, in the presence of noise this feature is lost. Having said this, it would be highly desirable to have a recovery procedure which gave perfect reconstruction up to an error determined by the amplitude of the noise in the input data. However, as shown in Figure 1, this is not the case for $\tilde{f}_m$. The combination of noise and ill-conditioning prohibits $\tilde{f}_m$ from being a good approximation to f. Note also that in the noise-free case, the approximation $\tilde{f}_m$ does initially obtain close to machine epsilon (i.e. it is numerically perfect), but once n is increased beyond $m = 30$, the error also diverges. Thus, one may not witness perfect reconstruction in practice, even in the absence of noise.

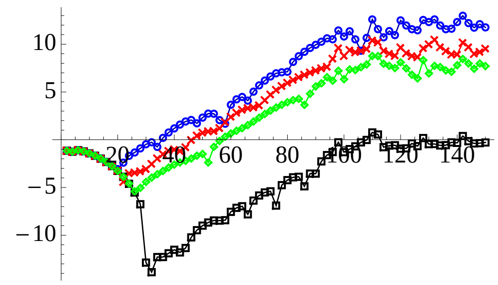

Fig. 1. The quantity $\log_{10} \|f - \tilde{f}_m\|$ against $m = 1, \ldots, 150$, where $f(x) = x^{30}$ and $\tilde{f}_m$ was computed from (i) noise-free data (squares) and (ii) noise at amplitude $\epsilon = 10^{-4}$ (circles), 10^{-6} (crosses) and 10^{-8} (diamonds). Noise was modelled by replacing $\hat{f}_j$ by $\hat{f}_j + \epsilon z_j$ for $j = -75, \ldots, 75$, where z_j is uniformly distributed in $[-1,1]$.

Whilst this example is predominantly for illustration, we mention that the problem of reconstructing a smooth function from its Fourier samples in a polynomial basis has been the subject of intense study in the last several decades, mainly in the field of spectral methods for PDEs [6]. Within this context, the approach (2) is known as the *inverse polynomial reconstruction method* (IPRM) [10], whose poor performance for the aforementioned examples is well documented [9].

Returning to the general case, it is tempting to think that there may be some way to avoid these pitfalls, and to obtain a reconstruction $\tilde{f}_m$, computable in a stable manner, that better resembles the projection f_m. Our method is designed to achieve precisely these goals. Such an approach was first proposed by the authors in [1], and later expanded upon in [2] (a similar idea, within the context of the IPRM, also appears in [9]).

2. REDUCED CONSISTENCY SAMPLING

The idea to circumvent the aforementioned problems is to allow the dimension m of the sampling basis $\mathcal{S}_m$ to differ from that of the reconstruction basis, which we now label $\mathcal{W}_n$, where $n = \dim \mathcal{W}_n$. Since $n \neq m$, we cannot enforce the consistency. Instead, we specify the approximation $\tilde{f}_{n,m}$ by the conditions

$$\langle S_m \tilde{f}_{n,m}, w_j \rangle = \langle S_m f, w_j \rangle, \quad j = 1, \ldots, n, \qquad (6)$$

where $S_m g = \sum_{j=1}^m \langle g, s_j \rangle s_j$ for $g \in \mathcal{H}$. We refer to (6) as a *reduced consistency condition*: rather than f and its reconstruction being identical on the space $\mathcal{S}_m$, they agree on the space $S_m(\mathcal{W}_n) \subseteq \mathcal{S}_m$. Note that if $c' = (\langle S_m f, w_1 \rangle, \ldots, \langle S_m f, w_n \rangle)^\top$ and $d = (d_1, \ldots, d_n)^\top$, then

(6) corresponds to the linear system $Bd = c'$, where

$$B = \begin{pmatrix} \langle S_m w_1, w_1 \rangle & \cdots & \langle S_m w_n, w_1 \rangle \\ \vdots & \ddots & \vdots \\ \langle S_m w_1, w_n \rangle & \cdots & \langle S_m w_n, w_n \rangle \end{pmatrix},$$

and $\tilde{f}_n = \sum_{k=1}^{n} d_k w_k$. Additionally, B and c' can be expressed as $A^* A$ and $A^* c$ respectively, where $c = (c_1, \ldots, c_m)^\top$ and $A \in \mathbb{C}^{m \times n}$ has $(i,j)^{\text{th}}$ entry $\langle w_j, s_i \rangle$.

To connect this with the previous narrative, note that A is now an $m \times n$ finite section of the infinite matrix A_∞ (as opposed to an $m \times m$ section). Moreover, $B = A^* A$ can be viewed as an approximation (in m) to the $n \times n$ section of the positive self-adjoint matrix $(A_\infty)^* A_\infty$. For this reason, we expect, and it turns out to be the case, that the procedure (6) is both nonsingular and stable, provided m is sufficiently large in comparison to n. Note that if $m = n$ then this approach coincides with (1).

The simplest framework in which to study this procedure is when $\{s_j\}_{j=1}^\infty$ and $\{w_j\}_{j=1}^\infty$ are Riesz bases. The method (1) has also been studied within the more general setting of frames [5]. However, whilst it is probable that (6) could be extended to this case (a topic we intend to address in future work), for simplicity we shall remain within the Riesz basis setting. To this end, suppose that $a_1, a_2 > 0$ are constants such that

$$a_1 \sum_{j=1}^{\infty} |c_j|^2 \leq \Big\| \sum_{j=1}^{\infty} c_j s_j \Big\|^2 \leq a_2 \sum_{j=1}^{\infty} |c_j|^2,$$

for all $c = \{c_1, c_2, \ldots\} \in l^2(\mathbb{N})$. Observe that (see [3], for example) there also exist constants $a_1', a_2' > 0$ such that

$$a_1' \|g\|^2 \leq \sum_{j=1}^{\infty} |\langle g, s_j \rangle|^2 \leq a_2' \|g\|^2, \quad \forall g \in \mathcal{S}.$$

Write $b_1, b_2, b_1', b_2' > 0$ for the corresponding constants for the basis $\{w_j\}_{j=1}^\infty$.

We also introduce the continuous linear operator $S : \mathcal{H} \to \mathcal{H}$, given by $Sg = \sum_{j=1}^\infty \langle g, s_j \rangle s_j$. Note that the operators $S_m \to S$ strongly on $\mathcal{H}$. Moreover, the form $\langle f, g \rangle_S = \langle Sf, g \rangle$ yields an equivalent inner product on $\mathcal{W}$. With this to hand, we now state the following two theorems (see [1, 2] for proofs):

Theorem 1 *For each n, there exists and m_0 such that, for all $m \geq m_0$, the solution $\tilde{f}_{n,m}$ to (6) exists and is unique. Specifically, if $e_{n,m}$ is defined by*

$$e_{n,m} = \sup_{w \in \mathcal{W}_n, \|w\|_S = 1} \langle Sw - S_m w, w \rangle \leq 1,$$

then $e_{n,m} \to 0$ as $m \to \infty$ and m_0 is the least m such that $e_{n,m} < 1$.

Theorem 2 *For $m \geq m_0$ the reconstruction $\tilde{f}_{n,m}$ satisfies*

$$\|f - f_n\| \leq \|f - \tilde{f}_{n,m}\| \leq k_{n,m} \|f - f_n\|,$$

where $f_n = P_{\mathcal{W}_n} f$ is the orthogonal projection of f onto $\mathcal{W}_n$ and the constant $k_{n,m}$ is given by

$$k_{n,m} = \sqrt{1 + (a_2')^2 (a_1')^{-2} (1 - e_{n,m})^{-2}}.$$

Theorems 1 and 2 confirm the existence of $\tilde{f}_{n,m}$ and provide an estimate for the error. Note the following important point: for fixed m, we must choose n suitable small (conversely, for fixed n, we choose m suitably large) to ensure existence of of the reconstruction. For this reason, it is vital to be able to estimate $e_{n,m}$. Whilst analytical estimates may be available in some circumstances, in general this must be done numerically. Fortunately, this task is aided by the following lemma, which gives a computable expression for $e_{n,m}$ (for a proof, see [2]):

Lemma 1 *Let $\tilde{B} \in \mathbb{C}^{n \times n}$ be the Hermitian matrix with $(j,k)^{\text{th}}$ entry $\langle w_j, w_k \rangle_S$. Then $e_{n,m} = 1 - \lambda_{\min}(\tilde{B}^{-1} B)$.*

Another important feature of this approach is its stability. To quantify this, we consider the condition number $\kappa(B) = \|B\| \|B^{-1}\|$. The following lemma is also found in [2]:

Lemma 2 *We have $\kappa(B) \leq \frac{1 + e_{n,m}}{1 - e_{n,m}} \kappa(\tilde{B}) \leq \frac{a_2' b_2 (1 + e_{n,m})}{a_1' b_1 (1 - e_{n,m})}$.*

Thus this framework is stable for sufficiently large m. Moreover, the condition required for stability is identical to that which guarantees existence and uniqueness of $\tilde{f}_{n,m}$.

As shown in Theorem 2, this framework also possesses so-called *quasi-optimality*: provided m is such that $e_{n,m}$ is bounded away from 1, the error in reconstructing f with $\tilde{f}_{n,m}$ is of the same order as $\|f - f_n\|$ (f_n being the best, i.e. energy-minimising approximation to f from $\mathcal{W}_n$). This is in stark contrast to the consistent reconstruction (2), as shown numerically in Table 1.

2.1. Oblique asymptotic optimality

This framework relies on allowing m to vary independently of n. A natural question to ask is what happens if $m \to \infty$ for fixed n? Intuitively, since $S_m \to S$ strongly, it appears that $\tilde{f}_{n,m}$ should converge to the element $\tilde{f}_n \in \mathcal{W}_n$ defined by

$$\langle S\tilde{f}_n, w_j \rangle = \langle Sf, w_j \rangle, \quad j = 1, \ldots, n, \tag{7}$$

(these equations being the 'limit' of the equations (6) defining $\tilde{f}_{n,m}$). The following theorem confirms this intuition:

Theorem 3 *Suppose that $\tilde{f}_{n,m}$ and $\tilde{f}_n$ are defined by (6) and (7) respectively. Then $\tilde{f}_{n,m} \to \tilde{f}_n$ as $m \to \infty$ and we have*

$$\|\tilde{f}_n - \tilde{f}_{n,m}\|^2 \leq \frac{a_2 e_{n,m}}{a_1' (1 - e_{n,m})^2} \|f - \tilde{f}_n\|^2.$$

Further insight is gained by noting that the mapping $\mathcal{H} \to \mathcal{W}_n$, $f \mapsto \tilde{f}_n$ is the *oblique projection* onto $\mathcal{W}_n$ along $(S(\mathcal{W}_n))^\perp$. Thus, we refer to the property that $\tilde{f}_{n,m} \approx \tilde{f}_n$ for large m as *oblique asymptotic optimality*. Additionally, whenever $f \in \mathcal{W}$, the quantity $\tilde{f}_n$ is precisely the orthogonal projection of f onto $\mathcal{W}_n$ with respect to the inner product $\langle \cdot, \cdot \rangle_S$. Thus, whilst $\tilde{f}_n$ may not be the energy-minimising approximation to f from $\mathcal{W}_n$ in the canonical norm $\|\cdot\|$, it is energy-minimising with respect to the equivalent norm defined by S.

Such a property is clearly highly desirable. Moreover, in many circumstances of interest (such as the reconstruction of a

m	25	50	100	200	400
(i)	4.1e00	3.5e00	3.5e00	3.3e00	3.3e00
(iii)	5.6e-02	2.1e-02	4.9e-03	6.7e-04	3.5e-05

Table 2. The quantities (i) $\|B^{-1}\|$ and (ii) $\|f - \tilde{f}_{\lfloor 2\sqrt{m} \rfloor, m}\|$.

function from its Fourier samples), $\tilde{f}_n$ actually coincides with the orthogonal projection $P_{\mathcal{W}_n} f = f_n$. Most notably, this occurs whenever $S = I$ is the identity operator on $\mathcal{H}$. In this case, we say that $\tilde{f}_{n,m}$ possesses *asymptotic optimality*.

2.2. Orthogonal sampling bases and subspace angles

Suppose that $\{s_j\}_{j=1}^{\infty}$ is an orthonormal basis. It follows that $S = I$ and $S_m = P_{\mathcal{S}_m}$ is the orthogonal projection. In particular, $\tilde{f}_{n,m}$ possesses asymptotic optimality. Thus, by choosing m sufficiently large, we can make $\tilde{f}_{n,m}$ arbitrary close to the best approximation to f from $\mathcal{W}_n$.

This aside, we may actually give a geometric interpretation of the quantity $e_{n,m}$ in this case. Specifically,

$$e_{n,m} = \sup_{w \in \mathcal{W}_n, \|w\|=1} \|w - S_m w\|^2 = 1 - \cos^2 \theta = \sin^2 \theta,$$

where θ is the subspace angle $\theta_{\mathcal{W}_n \mathcal{S}_m}$. Furthermore, we also find that the constant $k_{n,m}$ of Theorem 2 can be expressed in terms of θ: $k_{n,m} = \sqrt{1 + \tan^2 \theta \sec^2 \theta}$.

2.3. Numerical examples

To illustrate this approach, consider the example of Section 1 once more: namely, recovery from Fourier samples using Legendre polynomials. In all cases the ratio $n = \lfloor 2\sqrt{m} \rfloor$ was used (this has shown to be a sufficient condition for quasi-optimal recovery [2]). In Table 2.3 we consider the error incurred by the approximation $\tilde{f}_{m, \lfloor 2\sqrt{m} \rfloor}$ for the example $f(x) = \frac{1}{1+16x^2}$. Upon comparison with Table 1, the improvement by offered the reduced consistency framework is evident. Moreover, as also illustrated, this approach is completely numerically stable.

This method also performs well in the presence of noise. In Figure 2 we consider the approximation of $f(x) = x^{30}$ with random noise once more. Upon comparison with Figure 1, we conclude that the stability of this procedure leads to a significant improvement over the consistent sampling scheme (2).

3. EXTENSIONS

There are a number of generalisations of this work. First, there are several important types of sampling problem not covered by the current framework. Specifically, suppose that, instead of the inner products $c_j = \langle f, s_j \rangle$, we had more general forms of measurements $\zeta_j(f)$, where ζ_j are linear functionals on $\mathcal{H}$. Moreover, suppose that $f \in \mathcal{H}_1$ is defined as the solution to the problem, $Lf = g$, where $L : \mathcal{H}_1 \to \mathcal{H}_2$ is linear and $g \in \mathcal{H}_2$. If we can only access fixed measurements of g, can we still

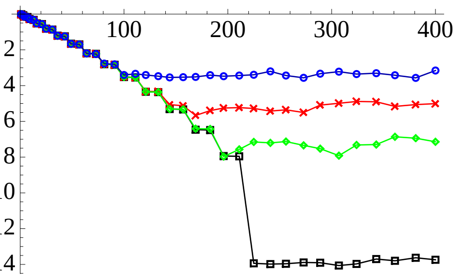

Fig. 2. The quantity $\log_{10} \|f - \tilde{f}_{\lfloor 2\sqrt{m} \rfloor, m}\|$ against $m = 1, \ldots, 400$, for noise-free data and noise at amplitudes $\epsilon = 10^{-4}, 10^{-6}, 10^{-8}$.

reconstruct f in an arbitrary basis? As we shall demonstrate in a future paper, both these problems can be solved by suitably generalising the reduced consistency framework.

Secondly, it transpires that a combination of the reduced consistency framework with techniques from compressed sensing allows one to stably and accurately reconstruct sparse signals with significant undersampling. This topic has recently been developed in [8].

4. REFERENCES

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

[2] B. Adcock and A. C. Hansen. Stable reconstructions in Hilbert spaces and the resolution of the Gibbs phenomenon. *Technical report NA2010/10, DAMTP, University of Cambridge*, 2010.

[3] O. Christensen. *An introduction to frames and Riesz bases*. Birkhauser, 2003.

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

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

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

[7] 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.

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

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

[10] J.-H. Jung and B. D. Shizgal. Generalization of the inverse polynomial reconstruction method in the resolution of the Gibbs phenomenon. *J. Comput. Appl. Math.*, 172(1):131–151, 2004.

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

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