Stable reconstructions in Hilbert spaces and the resolution of the Gibbs phenomenon

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

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

July 7, 2011

Abstract

We introduce a simple and efficient method to reconstruct an element of a Hilbert space in terms of an arbitrary finite collection of linearly independent reconstruction vectors, given a finite number of its samples with respect to any Riesz basis. As we establish, provided the dimension of the reconstruction space is chosen suitably in relation to the number of samples, this procedure can be implemented in a completely numerically stable manner. Moreover, the accuracy of the resulting approximation is determined solely by the choice of reconstruction basis, meaning that reconstruction vectors can be readily tailored to the particular problem at hand.

An important example of this approach is the accurate recovery of a piecewise analytic function from its first few Fourier coefficients. Whilst the standard Fourier projection suffers from the Gibbs phenomenon, by reconstructing in a piecewise polynomial basis we obtain an approximation with root-exponential accuracy in terms of the number of Fourier samples and exponential accuracy in terms of the degree of the reconstruction. Numerical examples illustrate the advantage of this approach over other existing methods.

1 Introduction

Suppose that H is a separable Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and corresponding norm $\| \cdot \|$. In this paper, we consider following problem: given the first m samples $\{\langle f, \psi_j \rangle\}_{j=1}^{m}$ of an element $f \in \mathrm{H}$ with respect to some Riesz basis $\{\psi_j\}_{j=1}^{\infty}$ of H (the *sampling* basis), reconstruct f to high accuracy. Not only does such a problem lie at the heart of modern sampling theory [28, 29, 63], it also occurs in a myriad of applications, including image processing (in particular, Magnetic Resonance Imaging), and the numerical solution of hyperbolic partial differential equations (PDEs).

In practice, straightforward reconstruction of f may be achieved via orthogonal projection with respect to the sampling basis. Indeed, for an arbitrary $f \in \mathrm{H}$, this is the best possible strategy. However, in many important circumstances, this approximation converges only slowly in m, when measured in the norm on H, or not at all, if a stronger norm – the uniform norm, for example – is considered.

A prominent instance of this problem is the recovery of a function $f : [-1, 1] \to \mathbb{R}$ from its first m Fourier coefficients. In this instance, $\mathrm{H} = \mathrm{L}^2(-1, 1)$ is the space of all square-integrable functions. Provided f is analytic and periodic, it is well-known that its Fourier series (the orthogonal projection with respect to the Fourier basis) converges exponentially fast [20, chpt. 5]. However, whenever f has a jump discontinuity – in particular, if f is nonperiodic, or equivalently, has a jump discontinuity at $x = 1$ – its Fourier series suffers from the well-known Gibbs phenomenon [56, Part I]. Whilst convergence occurs in the L^2 norm, uniform convergence is lacking, and the approximation is polluted by characteristic $\mathcal{O}(1)$ oscillations near the discontinuity. Moreover, the rate of convergence is also slow: only $\mathcal{O}(m^{-\frac{1}{2}})$ when measured in the L^2 norm, and $\mathcal{O}\left(m^{-1}\right)$ pointwise away from the discontinuity. Needless to say, the Gibbs phenomenon is a significant blight of many practical applications of Fourier series [50]. It is a testament to its importance that the design of effective techniques for its removal remains an active area of inquiry [41, 61].

Returning to the general form of the problem, let us now suppose that some additional information is known about the element f. Specifically, suppose that we know that f can be well-represented in a particular basis. For example, in the Fourier setting, we may know that f is piecewise analytic with jump discontinuities at given locations in $[-1,1]$. In this circumstance, it seems plausible that a better approximation to f can be obtained by expanding in a different basis – a piecewise polynomial basis, for example. To this end, we introduce the so-called *reconstruction* space (of dimension n) and seek to approximate f by an element $f_{n,m}$ consisting of n linearly independent elements of this space.

As we will show in due course, provided reconstruction is carried out in a certain manner, a suitable approximation $f_{n,m}$ can always be found. Essential to this approach is that m (the number of samples) is chosen sufficiently large in comparison to n (equivalently, n is chosen sufficiently small in comparison to m). However, provided this is the case, the approximation $f_{n,m}$ inherits the principal features of the reconstruction space. In particular, $f_{n,m}$ is quasi-optimal in sense that the error $\|f - f_{n,m}\|$ can be bounded by a constant multiple of $\|f - \mathcal{Q}_n f\|$, where $\mathcal{Q}_n f$ is the orthogonal projection onto the reconstruction space – in other words, the best approximation to f from this space. Moreover, from a practical standpoint, this method can be implemented by solving a linear least squares problem. Whenever the reconstruction vectors are suitably chosen (e.g. if they form a Riesz basis), the corresponding linear system is well-conditioned and the least squares problem can be solved in $\mathcal{O}(mn)$ operations by standard iterative techniques.

Consider once more the example of Fourier series, and let $f : [-1,1] \to \mathbb{R}$ be an analytic, nonperiodic function. As mentioned, the Fourier series of f lacks uniform convergence. However, since f is analytic, it makes sense to seek to reconstruct f in a system of polynomials. It is well-known that the n^{th} degree polynomial expansion of an analytic function converges exponentially fast in n [12, chpt. 2]. With this in mind, a key theorem we prove in this paper is as follows:

Theorem 1.1. *Let the first m Fourier samples of a function $f \in \mathrm{L}^2(-1,1)$ be given. Then*

(i) With $n = \mathcal{O}(\sqrt{m})$ it is possible to compute a polynomial approximation $f_{n,m}$ of degree n that satisfies

$$\|f - \mathcal{Q}_n f\| \leq \|f - f_{n,m}\| \leq c\|f - \mathcal{Q}_n f\|,$$

where c is independent of f and m, and $\mathcal{Q}_n f$ is the best polynomial approximation to f of degree n. Furthermore, c can be made arbitrarily close to 1 by a suitably small choice of the constant c' in the scaling $n = c'\sqrt{m}$.

(ii) The approximation $f_{n,m}$ is completely independent of the choice of polynomial basis used to represent it. The particular basis can be chosen by the user, and such a choice only affects numerical stability and computational cost.

(iii) When implemented with Legendre polynomials, the numerical method is completely stable and $f_{n,m}$ can be computed in only $\mathcal{O}(m^{\frac{3}{2}})$ operations. Conversely, if Chebyshev polynomials are employed, for example, the condition number of the method is $\mathcal{O}(\sqrt{m})$ and the corresponding computational cost is $\mathcal{O}(m^{\frac{7}{4}})$.

An immediate consequence of this theorem is that $f_{n,m}$ converges exponentially fast in the polynomial degree n whenever f is analytic, and root-exponentially fast in terms of m, the number of Fourier samples. In addition, it transpires that both the method and this theorem can be generalised to the recovery of a piecewise analytic function of one variable (using piecewise polynomial bases), and to the case of multivariate functions defined in tensor-product regions. In particular, both the scaling $n = \mathcal{O}(\sqrt{m})$ and (root) exponential convergence are maintained in these more general settings.

The method developed in this paper was previously introduced by the authors in [4] within the context of sampling theory. Whilst this problem, in an abstract form, has been extensively studied in the last couple of decades (in particular, by Eldar et al [28, 29], see also [63]), to the best of our knowledge this method does not appear in any existing literature. For a more detailed discussion of the relation of this approach to existing schemes we refer the reader to [4]. Conversely, in this paper, after presenting the general version of the method in abstract terms, we focus primarily on its application to the Fourier coefficient reconstruction problem. On this topic, a similar approach, but only dealing with reconstructions in Legendre polynomials from Fourier samples of analytic functions, was discussed in [47]. This can be viewed as a special case of our general framework. Furthermore, by examining this example as part of the general framework, we are able to extend and improve the work of [47] in the following ways: (i) we derive a procedure allowing for reconstructions in any polynomial basis, not just Legendre polynomials, (ii) we extend this approach to reconstructions of piecewise smooth functions using (arbitrary) piecewise polynomial bases, (iii) we generalise this work to smooth functions of arbitrary numbers of variables and (iv) we obtain improved estimates for both the error and the necessary scaling $n = \mathcal{O}\left(\sqrt{m}\right)$ required for implementation.

Aside from these improvements, a great benefit of the general framework presented in this paper is that it is immediately applicable to a whole host of other reconstruction problems. To illustrate this generality, in the final part of this paper we consider its use in the accurate reconstruction of a piecewise analytic function from its orthogonal polynomial expansion coefficients. Such a problem is typical of that occurring in the application of polynomial spectral methods to hyperbolic PDEs [38, 46], where the shock formation inhibits fast convergence of the polynomial approximation. As we highlight, this issue can be overcome in a completely stable fashion by reconstructing in a piecewise polynomial basis.

There are numerous algorithms for the removal of the Gibbs phenomenon from Fourier series or expansions in orthogonal polynomials. One of the most well-known and widely used is spectral reprojection [35, 41, 42]. As we discuss further in Section 3, the method developed in this paper has a number of key advantages over this technique. Numerical results also indicate its superior performance.

The outline of the remainder of this paper is as follows. In Section 2 we introduce the reconstruction procedure and establish both stability and error estimates. Section 3 is devoted to (piecewise) polynomial reconstructions from Fourier samples. In Section 4 we consider reconstructions in tensor-product spaces, and in Section 5 we discuss other recovery problems. Finally, in Section 6 we present open problems and challenges.

2 General theory of reconstruction

In this section, we describe the reconstruction procedure in its full generality. To this end, suppose that $\{\psi_j\}_{j=1}^{\infty}$ is a Riesz basis (the sampling basis) for a separable Hilbert space H over the field $\mathbb{C}$. Let $\langle \cdot, \cdot \rangle$ be the inner product on H, with associated norm $\|\cdot\|$. Recall that, by definition, $\mathrm{span}\{\psi_1, \psi_2, \ldots\}$ is dense in H and

$$c_1 \sum_{j=1}^{\infty} |\alpha_j|^2 \leq \left\| \sum_{j=1}^{\infty} \alpha_j \psi_j \right\|^2 \leq c_2 \sum_{j=1}^{\infty} |\alpha_j|^2, \quad \forall \alpha = \{\alpha_1, \alpha_2, \ldots\} \in l^2(\mathbb{N}), \tag{2.1}$$

for positive constants c_1, c_2. Equivalently, $\psi_j = \mathcal{B}(\Psi_j)$, where $\{\Psi_j\}_{j=1}^{\infty}$ is an orthonormal basis for H and $\mathcal{B} : \mathrm{H} \rightarrow \mathrm{H}$ is a bounded, bijective operator. Using this definition, it is easy to deduce that $\{\psi_j\}_{j=1}^{\infty}$ also satisfies the frame property

$$d_1 \|f\|^2 \leq \sum_{j=1}^{\infty} | \langle f, \psi_j \rangle |^2 \leq d_2 \|f\|^2, \quad \forall f \in \mathrm{H}, \tag{2.2}$$

for $d_1, d_2 > 0$, where the smallest possible value for d_2 is $\|\mathcal{B}\|^2_{\mathrm{H}\rightarrow\mathrm{H}}$ and the largest possible value for d_1 is $\|\mathcal{B}^{-1}\|^{-2}_{\mathrm{H}\rightarrow\mathrm{H}}$ [21].

Suppose now that the first m coefficients of an element $f \in \mathrm{H}$ with respect to the sampling basis are given:

$$\hat{f}_j = \langle f, \psi_j \rangle, \quad j = 1, \ldots, m. \tag{2.3}$$

Set $\mathrm{S}_m = \mathrm{span}\{\psi_1, \ldots, \psi_m\}$ and let $\mathcal{P}_m : \mathrm{H} \rightarrow \mathrm{S}_m$ be the mapping

$$f \mapsto \mathcal{P}_m f = \sum_{j=1}^{m} \langle f, \psi_j \rangle \psi_j. \tag{2.4}$$

We now seek to reconstruct f in a different basis. To this end, suppose that $\{\phi_1, \ldots, \phi_n\}$ are linearly independent reconstruction vectors and define $\mathrm{T}_n = \mathrm{span}\{\phi_1, \ldots, \phi_n\}$. Let $\mathcal{Q}_n : \mathrm{H} \rightarrow \mathrm{T}_n$ be the orthogonal projection onto T_n. Direct computation of $\mathcal{Q}_n f$, the best approximation to f from T_n, is not possible, since the coefficients $\langle f, \phi_j \rangle$ are unknown. Instead, we seek to use the values (2.3) to compute an approximation $f_{n,m} \in \mathrm{T}_n$ that is *quasi-optimal*, i.e. $\| f - \mathcal{Q}_n f \| \leq \| f - f_{n,m} \| \leq C \| f - \mathcal{Q}_n f \|$ for some constant $C > 0$ independent of f and n. To do this, we introduce the sesquilinear form $a_m : \mathrm{H} \times \mathrm{H} \to \mathbb{C}$, given by

$$a_m(g, h) = \langle \mathcal{P}_m g, h \rangle \,, \quad \forall g, h \in \mathrm{H}. \tag{2.5}$$

Note that, since

$$\langle \mathcal{P}_m g, h \rangle = \sum_{j=1}^{m} \langle g, \psi_j \rangle \overline{\langle h, \psi_j \rangle} = \overline{\langle \mathcal{P}_m h, g \rangle}, \quad \forall f, g \in \mathrm{H},$$

a_m is a Hermitian form on $\mathrm{H} \times \mathrm{H}$ (here $\overline{z}$ denotes the complex conjugate of $z \in \mathbb{C}$). With this to hand, we now define $f_{n,m}$ as the solution to

$$a_m(f_{n,m}, \phi) = a_m(f, \phi), \quad \forall \phi \in \mathrm{T}_n, \qquad f_{n,m} \in \mathrm{T}_n. \tag{2.6}$$

Upon setting $\phi = \phi_j$, $j = 1, \ldots, n$, this becomes an $n \times n$ linear system of equations for the coefficients $\alpha_1, \ldots, \alpha_n$ of $f_{n,m} = \sum_{j=1}^{n} \alpha_j \phi_j$. We shall defer a discussion of the computation of this approximation to Section 2.3: first we consider the analysis of $f_{n,m}$.

Let us at this stage observe that (2.6) is equivalent to the following linear least squares problem

$$\min_{\phi \in \mathrm{T}_n} \left\{ \sum_{j=1}^{m} |\langle f, \psi_j \rangle - \langle \phi, \psi_j \rangle|^2 \right\}. \tag{2.7}$$

As a result, the coefficients $\alpha_1, \ldots, \alpha_n$ are the least squares solution of a system of m linear equations. Although (2.7) appears very familiar, we are yet to find this particular formulation, or any pertinent analysis, in the literature relating to sampling and reconstruction.

2.1 Analysis of $f_{n,m}$

Before proving the main theorem regarding (2.6), let us first give an intuitive explanation as to why this approach works. As mentioned, the key to this technique is that the parameter m is sufficiently large in comparison to n. To this end, let n be fixed and suppose that $m \to \infty$. Due to (2.1), the mappings $\mathcal{P}_m$ converge strongly to a bounded, linear operator $\mathcal{P}$, the *frame* operator [21], given by

$$\mathcal{P} f = \sum_{j=1}^{\infty} \langle f, \psi_j \rangle \psi_j, \quad \forall f \in \mathrm{H}. \tag{2.8}$$

Hence, for large m, the equations (2.6) defining $f_{n,m}$ resemble the equations

$$a(\tilde{f}_n, \phi) = a(f, \phi), \quad \forall \phi \in \mathrm{T}_n, \qquad \tilde{f}_n \in \mathrm{T}_n, \tag{2.9}$$

where $a : \mathrm{H} \times \mathrm{H} \to \mathbb{C}$ is the Hermitian form $a(f, g) = \langle \mathcal{P} f, g \rangle$. Thus, it is reasonable to expect that $f_{n,m} \to \tilde{f}_n$ as $m \to \infty$, provided such a function $\tilde{f}_n$ exists. However,

Theorem 2.1. *For all $n \in \mathbb{N}$, the function $\tilde{f}_n$ exists and is unique. Moreover,*

$$\| f - \tilde{f}_n \| \leq \frac{d_2}{d_1} \| f - \mathcal{Q}_n f \|, \tag{2.10}$$

where d_1 and d_2 are as in (2.2).

This theorem can be established with a straightforward application of the Lax–Milgram theorem and its counterpart, Céa's lemma [36]. Indeed, due to (2.2) and (2.8),

$$d_1 \| g \|^2 \leq a(g, g) = \sum_{j=1}^{\infty} | \langle g, \psi_j \rangle |^2 \leq d_2 \| g \|^2, \quad \forall g \in \mathrm{H}. \tag{2.11}$$

Hence the form $a(\cdot, \cdot)$ defines an equivalent inner product on H. Nonetheless, we shall present a self-contained proof, since similar techniques will be used subsequently.

Proof. Let $\mathcal{U} : \mathrm{T}_n \to \mathbb{C}^n$ be the linear mapping $g \mapsto \{\langle \mathcal{P}g, \phi_j \rangle\}_{j=1}^n$. To prove existence and uniqueness of $\tilde{f}_n$ it suffices to show that $\mathcal{U}$ is invertible, upon which it follows that $\tilde{f}_n = \mathcal{U}^{-1}\{\langle \mathcal{P}f, \phi_j \rangle\}_{j=1}^n$. Suppose that $\mathcal{U}g = 0$. Then, by definition, $\langle \mathcal{P}g, \phi_j \rangle = 0$ for $j = 1, \ldots, n$. Using linearity, we deduce that $\langle \mathcal{P}g, g \rangle = 0$. Now, it follows from (2.11) that $0 = \langle \mathcal{P}g, g \rangle \geq d_1 \|g\|^2$, giving $g = 0$. Hence, $\mathcal{U}$ is invertible and $\tilde{f}_n$ exists and is unique.

Now consider the error estimate (2.10). Using (2.2) once more, we obtain

$$\|f - \tilde{f}_n\|^2 \leq \frac{1}{d_1} \sum_{j=1}^{\infty} \left| \langle f - \tilde{f}_n, \psi_j \rangle \right|^2 = \frac{1}{d_1} \left\langle \mathcal{P}(f - \tilde{f}_n), f - \tilde{f}_n \right\rangle.$$

By definition of $\tilde{f}_n$, $\langle \mathcal{P}(f - \tilde{f}_n), \phi \rangle = 0$, $\forall \phi \in \mathrm{T}_n$. In particular, setting $\phi = \tilde{f}_n - \mathcal{Q}_n f$, yields

$$\|f - \tilde{f}_n\|^2 \leq \frac{1}{d_1} \left\langle \mathcal{P}(f - \tilde{f}_n), f - \mathcal{Q}_n f \right\rangle = a(f - \tilde{f}_n, f - \mathcal{Q}_n f).$$

Since $a(\cdot, \cdot)$ gives an equivalent inner product on H, an application of the Cauchy–Schwarz inequality yields

$$\|f - \tilde{f}_n\|^2 \leq \frac{1}{d_1} \left[a(f - \tilde{f}_n, f - \tilde{f}_n) a(f - \mathcal{Q}_n f, f - \mathcal{Q}_n f) \right]^{\frac{1}{2}} \leq \frac{d_2}{d_1} \|f - \tilde{f}_n\| \|f - \mathcal{Q}_n f\|,$$

as required. □

This theorem establishes existence and quasi-optimality of $\tilde{f}_n \approx f_{n,m}$, thereby giving an intuitive argument for the success of this method. We now wish to fully confirm this observation. To this end, let

$$C_{n,m} = \inf_{\substack{\phi \in \mathrm{T}_n \\ \|\phi\|=1}} a_m(\phi, \phi). \tag{2.12}$$

Note that $C_{n,m}$ has the equivalent forms

$$C_{n,m} = \inf_{\substack{\phi \in \mathrm{T}_n \\ \|\phi\|=1}} \langle \mathcal{P}_m \phi, \phi \rangle = \inf_{\substack{\phi \in \mathrm{T}_n \\ \|\phi\|=1}} \left\{ \sum_{j=1}^{m} | \langle \phi, \psi_j \rangle |^2 \right\}.$$

The quantity $C_{n,m}$ plays a fundamental role in this paper. Its key properties are described in the following lemma:

Lemma 2.2. *For all $n, m \in \mathbb{N}$, $0 \leq C_{n,m} \leq d_2$. Moreover, for each n, $C_{n,m} \to C_n^* \geq d_1$ as $m \to \infty$, where*

$$C_n^* = \inf_{\substack{\phi \in \mathrm{T}_n \\ \|\phi\|=1}} a(\phi, \phi),$$

and d_1 is defined in (2.2).

Proof. Consider first the quantity

$$\epsilon_{n,m} = \sup_{\substack{\phi \in \mathrm{T}_n \\ \|\phi\|=1}} \langle \mathcal{P}\phi - \mathcal{P}_m \phi, \phi \rangle = \sup_{\substack{\phi \in \mathrm{T}_n \\ \|\phi\|=1}} \left\{ \sum_{j>m} | \langle \phi, \psi_j \rangle |^2 \right\}. \tag{2.13}$$

Due to (2.2), the infinite sum is finite for any fixed ϕ, and tends to zero as $m \to \infty$. Now let $\{\Phi_j\}_{j=1}^n$ be an orthonormal basis for T_n and set $\phi = \sum_{j=1}^n \alpha_j \Phi_j$. Two applications of the Cauchy–Schwarz inequality gives

$$\sum_{j>m} | \langle \phi, \psi_j \rangle |^2 \leq \|\phi\|^2 \sum_{k=1}^{n} \sum_{j>m} | \langle \Phi_k, \psi_j \rangle |^2.$$

Hence $\epsilon_{n,m} \leq \sum_{k=1}^n \sum_{j>m} | \langle \Phi_k, \psi_j \rangle |^2$, and we deduce that $\epsilon_{n,m}$ is both finite and $\epsilon_{n,m} \to 0$ as $m \to \infty$. Noticing that $|C_{n,m} - C_n^*| \leq \epsilon_{n,m}$, $\forall n, m \in \mathbb{N}$, gives the first part of the proof. For the second, we merely use (2.2). □

Aside from $C_{n,m}$, we also define the quantity

$$D_{n,m} = \sup_{\substack{f \in \mathrm{T}_n^\perp \\ \|f\|=1}} \sup_{\substack{g \in \mathrm{T}_n \\ \|g\|=1}} |a_m(f,g)| \,. \tag{2.14}$$

For this, we have the following lemma:

Lemma 2.3. *For all $m, n \in \mathbb{N}$, $0 \leq D_{n,m} \leq d_2$. Moreover, suppose that $\mathcal{P}$ is such that $\mathcal{P}(\mathrm{T}_n) \subseteq \mathrm{T}_n$ (for example, when $\mathcal{P} = \mathcal{I}$ is the identity), then $D^2_{n,m} \leq c_2 \epsilon_{n,m}$, where $\epsilon_{n,m}$ is as in (2.13). In particular, for fixed n, $D_{n,m} \to 0$ as $m \to \infty$.*

Proof. Let $f, g \in \mathrm{H}$. By definition,

$$a_m(f,g) = \langle \mathcal{P}_m f, g \rangle = \sum_{j=1}^{m} \langle f, \psi_j \rangle \overline{\langle g, \psi_j \rangle} \leq \left[\sum_{j=1}^{m} |\langle f, \psi_j \rangle|^2 \right]^{\frac{1}{2}} \left[\sum_{j=1}^{m} |\langle g, \psi_j \rangle|^2 \right]^{\frac{1}{2}} . \tag{2.15}$$

Hence (2.2) gives the first result. Now suppose that $f \in \mathrm{T}_n^\perp$ and $\mathcal{P}(\mathrm{T}_n) \subseteq \mathrm{T}_n$. Since $\mathcal{P}_m$ is self-adjoint, we have $\langle \mathcal{P}_m f, g \rangle = \langle f, \mathcal{P}_m g \rangle = \langle f, \mathcal{P}_m g - \mathcal{P} g \rangle$. Here the second equality is due to the fact that $f \perp \mathrm{T}_n$ and $\mathcal{P} g \in \mathrm{T}_n$ for $g \in \mathrm{T}_n$. By the Cauchy–Schwarz inequality, we obtain

$$D_{n,m} \leq \sup_{\substack{g \in \mathrm{T}_n \\ \|g\|=1}} \|\mathcal{P} g - \mathcal{P}_m g\|.$$

For $g \in \mathrm{T}_n$, we note from (2.1) that

$$\|\mathcal{P} g - \mathcal{P}_m g\|^2 \leq c_2 \sum_{j>m} |\langle g, \psi_j \rangle|^2 = c_2 \langle \mathcal{P} g - \mathcal{P}_m g, g \rangle \leq c_2 \epsilon_{n,m} \|g\|^2,$$

where the final equality follows from the definition of $\epsilon_{n,m}$. □

We are now able to state the main theorem of this section:

Theorem 2.4. *For every $n \in \mathbb{N}$ there exists an m_0 such that the approximation $f_{n,m}$, defined by (2.6), exists and is unique for all $m \geq m_0$, and satisfies the stability estimate*

$$\|f_{n,m}\| \leq \frac{d_2}{C_{n,m}} \|f\|.$$

Furthermore,

$$\|f - \mathcal{Q}_n f\| \leq \|f - f_{n,m}\| \leq K_{n,m} \|f - \mathcal{Q}_n f\|, \quad K_{n,m} = \sqrt{1 + D^2_{n,m} C^{-2}_{n,m}}. \tag{2.16}$$

Specifically, the parameter m_0 is the least value of m such that $C_{n,m} > 0$.

To prove this theorem, we first recall that a Hermitian form $a : \mathrm{H} \times \mathrm{H} \to \mathbb{R}$ is said to be *continuous* if, for some constant $\gamma > 0$, $|a(f,g)| \leq \gamma \|f\| \|g\|$ for all $f, g \in \mathrm{H}$. Moreover, a is *coercive*, provided $a(f,f) \geq \omega \|f\|^2$, $\forall f \in \mathrm{H}$, for $\omega > 0$ constant [36]. We now require the following lemma:

Lemma 2.5. *Suppose that $a_m : \mathrm{H} \times \mathrm{H} \to \mathbb{R}$ is the sesquilinear form $a_m(f,g) = \langle \mathcal{P}_m f, g \rangle$. Then a_m is continuous with constant $\gamma \leq d_2$. Moreover, for every $n \in \mathbb{N}$ there exists an m_0 such that the restriction of a_m to $\mathrm{T}_n \times \mathrm{T}_n$ is coercive for all $m \geq m_0$. Specifically, if $C_{n,m}$ is given by (2.12), then m_0 is the least value of m such that $C_{n,m} > 0$, and, for all $m \geq m_0$, $a_m(f,f) \geq C_{n,m} \|f\|^2$, $\forall f \in \mathrm{T}_n$. Finally, for all $f \in \mathrm{H}$ and $g \in \mathrm{T}_n$, we have $a_m(f - \mathcal{Q}_n f, g) \leq D_{n,m} \|f - \mathcal{Q}_n f\| \|g\|$.*

Proof. Continuity follows immediately from (2.15). For the second and final results, we merely use the definitions (2.12) and (2.14) of $C_{n,m}$ and $D_{n,m}$ respectively. □

Proof of Theorem 2.4. To establish existence and uniqueness, it suffices to prove that the linear operator $\mathcal{U} : \mathrm{T}_n \to \mathbb{C}^n$, $g \mapsto \{\langle \mathcal{P}_m g, \phi_j \rangle\}_{j=1}^n$ is invertible. Suppose that $g \in \mathrm{T}_n$ with $\mathcal{U} g = 0$. By definition, we have $\langle \mathcal{P}_m g, \phi_j \rangle = 0$ for $j = 1, \ldots, n$. Using linearity, it follows that $\langle \mathcal{P}_m g, g \rangle = 0$. Lemma 2.5 now gives $0 \leq C_{n,m} \|g\|^2 \leq 0$. Hence $g = 0$, and therefore $\mathcal{U}$ is invertible.

Stability of $f_{n,m}$ is easily established from the continuity and coercivity conditions. Setting $\phi = f_{n,m}$ in (2.6) gives

$$C_{n,m} \|f_{n,m}\|^2 \leq a_m(f_{n,m}, f_{n,m}) = a_m(f, f_{n,m}) \leq d_2 \|f\| \|f_{n,m}\|,$$

as required. Now consider the error estimate (2.16). Suppose that we define $e_{n,m} = f_{n,m} - \mathcal{Q}_n f \in \mathrm{T}_n$. Then, by definition of $f_{n,m}$, we have $a_m(e_{n,m}, \phi) = a_m(f - \mathcal{Q}_n f, \phi)$, $\forall \phi \in \mathrm{T}_n$. In particular, setting $\phi = e_{n,m}$ we obtain

$$\|e_{n,m}\| \leq C_{n,m}^{-1} D_{n,m} \|f - \mathcal{Q}_n f\|. \tag{2.17}$$

Since $\mathcal{Q}_n f$ is the orthogonal projection onto T_n, we have $\|f - f_{n,m}\|^2 = \|e_{n,m}\|^2 + \|f - \mathcal{Q}_n f\|^2$, which gives the full result. □

Let us sum up the key message of this theorem. It is possible to recover any element $f \in \mathrm{H}$ quasi-optimally from its samples in an arbitrary space T_n, provided the number of samples m is sufficiently large. Moreover, the condition that guarantees this recovery is known explicitly in terms of the quantity $C_{n,m}$.

In most practical situations, one has that $\mathrm{T}_1 \subseteq \mathrm{T}_2 \subseteq \ldots, \cup_{n=1}^{\infty} \mathrm{T}_n = \mathrm{H}$ (for example, when $\mathrm{T}_n = \mathbb{P}_n$). In this case, Theorem 2.4 states that both $f_{n,m}$ and $\mathcal{Q}_n$ will converge to f a precisely the same rate as $n \to \infty$, provided m is chosen sufficiently large for each n.

At this moment, we also mention one other important observation. The constants $C_{n,m}$, $D_{n,m}$ and C^*_n, as well as the approximation $f_{n,m}$, are determined only by the space T_n, not by the choice of reconstructions vectors $\phi_1, \ldots, \phi_n$ themselves. As we shall discuss later, the choice of such vectors only affects the stability of the scheme.

Remark 2.1 Given that a_m was shown to be continuous and coercive before proving Theorem 2.4, it may be tempting to seek to apply the Lax–Milgram theorem and Céa's lemma to obtain the result. However, the Hermitian form a_m, when considered as a mapping $\mathrm{H} \times \mathrm{H} \to \mathbb{C}$, will not, in general, be coercive. This is readily seen from the definition of $C_{n,m}$. The finite-dimensional operator $\mathcal{P}_m|_{\mathrm{T}_n}$ converges uniformly to $\mathcal{P}|_{\mathrm{T}_n}$, whereas its infinite-dimensional counterpart $\mathcal{P}_m : \mathrm{H} \to \mathrm{S}_m$ typically does not (for example, when $\mathcal{P}_m$ is the Fourier projection operator and $\mathrm{H} = \mathrm{L}^2(-1, 1)$). Hence, a_m only becomes coercive when restricted to $\mathrm{T}_n \times \mathrm{T}_n$, and these standard results do not automatically apply.

Although Theorem 2.4 establishes an estimate for the error $f - f_{n,m}$ measured in the natural norm on H, it is also useful to derive a result valid for any other norm defined on a suitable subspace of H (for example, this may be the uniform norm on $[-1, 1]$ in the case of Fourier series). To this end, let $|||\cdot|||$ be such a norm and define $\mathrm{G} = \{g \in \mathrm{H} : |||g||| < \infty\}$. We have

Corollary 2.6. *Suppose that $f \in \mathrm{G}$, $\mathrm{T}_n \subseteq \mathrm{G}$ and that $f_{n,m}$ is defined by (2.6). Then, for all $m \geq m_0$,*

$$|||f - f_{n,m}||| \leq |||f - \mathcal{Q}_n f||| + \frac{k_n D_{n,m}}{C_{n,m}} \|f - \mathcal{Q}_n f\|, \tag{2.18}$$

where $k_n = \sup_{\substack{\phi \in \mathrm{T}_n \\ \|\phi\|=1}} |||\phi|||$ and $C_{n,m}$, $D_{n,m}$ are given by (2.12) and (2.14) respectively.

Proof. Let $e_{n,m} = f_{n,m} - \mathcal{Q}_n f$ once more. Since $e_{n,m} \in \mathrm{T}_n$, it follows from the definition of k_n and the inequality (2.17) that

$$|||e_{n,m}||| \leq k_n \|e_{n,m}\| \leq \frac{k_n D_{n,m}}{C_{n,m}} \|f - \mathcal{Q}_n f\|.$$

The full result is obtained from the triangle inequality $|||f - f_{n,m}||| \leq |||e_{n,m}||| + |||f - \mathcal{Q}_n f|||$. □

This corollary verifies convergence of $f_{n,m}$ to f in $|||\cdot|||$, whenever $\mathcal{Q}_n f \to f$ in this norm and $k_n \|f - \mathcal{Q}_n f\| \to 0$ as $n \to \infty$. Note however that, although $f_{n,m}$ will converge at the same rate as $\mathcal{Q}_n f$ in this norm, this rate will in general be slower than that of best approximation in this norm, i.e.

$\phi = \arg\min_{\phi \in \mathrm{T}_n} |||f - \phi|||$. Having said this, the effect of this discrepancy is typically minimal, especially when $\mathcal{Q}_n f$ converges rapidly. See [20, chpt. 5] for a discussion of such differences in polynomial approximations.

Remark 2.2 In practice, it is useful to have an upper bound for the constant k_n. A simple exercise gives $k_n \leq \sum_{j=1}^{n} |||\Phi_j|||$, where $\{\Phi_j\}_{j=1}^{n}$ is any orthonormal basis for T_n.

Returning to main conclusion of Theorem 2.4 – namely, that guaranteed recovery can be obtained by allowing m to range independently of n – a natural question to ask is what happens if m is set equal to n. In abstract sampling theory this is known as the *consistent reconstruction* framework [29, 63]. This question was discussed in detail in [4], where it demonstrated that such an approach often leads to severe ill-conditioning as $n = m \to \infty$. Additionally, stringent restrictions are placed on the types of vectors f that can be reconstructed – see also Section 3.6. Conversely, by allowing m to vary independently of n, we obtain a reconstruction $f_{n,m}$ that is guaranteed to converge for any vector $f \in \mathrm{H}$. Moreover, as we discuss in Section 2.3, provided the reconstruction vectors are suitably chosen, the computation of $f_{n,m}$ is completely stable.

2.2 Oblique asymptotic optimality

Recall the intuitive argument of the previous section: namely, $f_{n,m} \approx \tilde{f}_n$ for all large m, where $\tilde{f}_n$ is defined by (2.9). We now wish to confirm this observation. Specifcally, we shall show that, for fixed $n \in \mathbb{N}$, $f_{n,m} \to \tilde{f}_n$ as $m \to \infty$, at a rate independent of the particular vector f.

Recall that the form $a(\cdot,\cdot)$ yields an equivalent inner product on H. Since $\tilde{f}_n$ is defined by the equations $a(\tilde{f}_n, \phi) = a(f, \phi)$, $\forall \phi \in \mathrm{T}_n$, the mapping $f \mapsto \tilde{f}_n$ is the orthogonal projection onto T_n with respect to this inner product. Letting $\|g\|_a = \sqrt{a(g,g)}$ be the corresponding norm on H, we now define the constants

$$\tilde{C}_{n,m} = \inf_{\substack{\phi \in \mathrm{T}_n \\ \|\phi\|_a = 1}} \langle \mathcal{P}_m \phi, \phi \rangle, \qquad \tilde{D}_{n,m} = \sup_{\substack{f \in \mathrm{T}_n^{\perp} \\ \|f\|_a = 1}} \sup_{\substack{g \in \mathrm{T}_n \\ \|g\|_a = 1}} |\langle \mathcal{P}_m f, g \rangle|. \tag{2.19}$$

In this instance, $\mathrm{T}_n^{\perp}$ is defined with respect to the a-inner product, i.e. $\mathrm{T}_n^{\perp} = \{f \in \mathrm{H} : a(f, \phi) = 0,\ \forall \phi \in \mathrm{T}_n\}$. Conversely, when considered with respect to the canonical inner product, this subspace is precisely $\mathcal{P}(\mathrm{T}_n)^{\perp} = \{f \in \mathcal{H} : \langle f, \phi \rangle = 0,\ \forall \phi \in \mathcal{P}(\mathrm{T}_n)\}$.

Note the similarity between $\tilde{C}_{n,m}$ and $\tilde{D}_{n,m}$ and the quantities $C_{n,m}$ and $D_{n,m}$ defined in (2.12) and (2.14) respectively. Roughly speaking, the former measure the deviation of $f_{n,m}$ from $\mathcal{Q}_n f$, whereas, as we will subsequently show, the latter determine the deviation of $f_{n,m}$ from $\tilde{f}_n$.

With these definitions to hand, identical arguments to those given in the proofs of Lemmas 2.2 and 2.3 now yield:

Lemma 2.7. *For all $m, n \in \mathbb{N}$, $\tilde{C}_{n,m} \geq \frac{1}{d_2} C_{n,m}$, where $C_{n,m}$ is as in (2.12). Moreover, for fixed n, $\tilde{C}_{n,m} \to 1$ as $m \to \infty$.*

Lemma 2.8. *For all $m, n \in \mathbb{N}$, $\tilde{D}_{n,m} \leq d_2$ and $\tilde{D}_{n,m}^2 \leq 1 - \tilde{C}_{n,m}$. In particular, for fixed n, $\tilde{D}_{n,m} \to 0$ as $m \to \infty$.*

Using these lemmas, we deduce

Corollary 2.9. *If $f_{n,m}$ and $\tilde{f}_n$ are given by (2.6) and (2.9) respectively, then*

$$\|f_{n,m} - \tilde{f}_n\|_a \leq \frac{\tilde{D}_{n,m}}{\tilde{C}_{n,m}} \|f - \tilde{f}_n\|_a,$$

and we have the error estimate

$$\|f - \tilde{f}_n\|_a \leq \|f - f_{n,m}\|_a \leq \tilde{K}_{n,m} \|f - \tilde{f}_n\|_a, \quad \tilde{K}_{n,m} = \sqrt{1 + \tilde{D}_{n,m}^2 \tilde{C}_{n,m}^{-2}}.$$

In particular, for any $f \in \mathrm{H}$, $f_{n,m} \to \tilde{f}_n$ as $m \to \infty$.

Proof. Since $(f_{n,m} - \tilde{f}_n) \in \mathrm{T}_n$, we have

$$\tilde{C}_{n,m} \| f_{n,m} - \tilde{f}_n \|_a^2 \leq \left\langle \mathcal{P}_m (f_{n,m} - \tilde{f}_n), f_{n,m} - \tilde{f}_n \right\rangle .$$

Moreover, because $\langle \mathcal{P}_m f_{n,m}, \phi \rangle = \langle \mathcal{P}_m f, \phi \rangle$, we deduce that

$$\tilde{C}_{n,m} \| f_{n,m} - \tilde{f}_n \|_a^2 \leq \left\langle \mathcal{P}_m (f - \tilde{f}_n), f_{n,m} - \tilde{f}_n \right\rangle \leq \tilde{D}_{n,m} \| f - \tilde{f}_n \|_a \| f_{n,m} - \tilde{f}_n \|_a ,$$

where the second inequality follows from the definition (2.19) of $\tilde{D}_{n,m}$ and the fact that $(f - \tilde{f}_n) \in \mathrm{T}_n^{\perp}$, the orthogonal complement of T_n with respect to the a-inner product. □

Note that the mapping $\mathcal{W}_n : f \mapsto \tilde{f}_n$ is an *oblique* projection with respect to the inner product $\langle \cdot, \cdot \rangle$ on H. In particular, $\mathcal{W}_n$ has range T_n and kernel $\mathcal{P}(\mathrm{T}_n)^{\perp}$, and we have the decomposition $\mathrm{H} = \mathrm{T}_n \oplus \mathcal{P}(\mathrm{T}_n)^{\perp}$. For this reason, we say that $f_{n,m}$ possesses *oblique asymptotic* optimality.

Whenever the sampling basis $\{\psi_j\}_{j=1}^{\infty}$ is orthonormal, we in fact witness so-called *asymptotic* optimality. In this setting, since $\mathcal{P} = \mathcal{I}$, the bilinear form $a(\cdot, \cdot)$ is precisely $\langle \cdot, \cdot \rangle$, and therefore $\tilde{f}_n = \mathcal{Q}_n f$ is the orthogonal projection. Hence, we can recover an approximation to f that is arbitrarily close to the error minimising approximation, which, as mentioned, cannot be computed directly from the given samples. Moreover, the rate of convergence of $f_{n,m}$ to $\mathcal{Q}_n f$ is completely independent of the particular vector f. Whatsmore, in this case the quantities $\tilde{C}_{n,m}$ and $\tilde{D}_{n,m}$ coincide with $C_{n,m}$ and $D_{n,m}$ respectively. It can also be shown that $D_{n,m} = 1 - C_{n,m}$. Thus, this rate of convergence depends only on $C_{n,m}$, and specifically, on the speed at which $C_{n,m} \to 1$.

Note that asymptotic optimality also occurs for general Riesz bases whenever $\mathcal{P}(\mathrm{T}_n) \subseteq \mathrm{T}_n$. The case of orthonormal sampling vectors presents the most obvious example of a basis satisfying this condition.

Remark 2.3 Whenever the vectors $\{\psi_j\}_{j=1}^{\infty}$ are not orthonormal, a natural question to ask is whether we can modify the method for computing $f_{n,m}$ to recover asymptotic optimality. This can be easily done, at least in theory, by replacing the operator $\mathcal{P}_m$ by some operator $\mathcal{P}'_m$ converging strongly to the identity on H (naturally, $\mathcal{P}'_m g$ must be also a function of $\hat{g}_1, \ldots, \hat{g}_m$).

One approach to do this is to let $\mathcal{P}'_m$ be the orthogonal projection $\mathrm{H} \to \mathrm{S}_m$. The downside of the approach is that it requires additional computational cost to compute $f_{n,m}$, as we explain at the end of the next section.

Another potential means to recover asymptotic optimality is to define $\mathcal{P}'_m g = \sum_{j=1}^{m} \langle g, \psi_j \rangle \psi_j^*$, where $\{\psi_j^*\}$ is the set of dual vectors to the sampling vectors $\{\psi_j\}$. In this case, $\mathcal{P}_m \to \mathcal{I}$ strongly, and asymptotic optimality follows. In practice, however, one may not have access to the dual vectors, thus this approach cannot necessarily be easily implemented.

2.3 Computation of $f_{n,m}$

Recall that the computation of the approximation $f_{n,m}$ involves solving the system of equations (2.6). These can be interpreted as the normal equations of the least squares problem (2.7). Suppose now that $f_{n,m} = \sum_{j=1}^{n} \alpha_j \phi_j$, $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{C}^n$ and $\hat{f} = (\hat{f}_1, \ldots, \hat{f}_m)$. If U is the $m \times n$ matrix with $(j,k)^{\mathrm{th}}$ entry $\langle \phi_k, \psi_j \rangle$, then (2.6) is given exactly by $A\alpha = U^{\dagger} \hat{f}$, where $A = U^{\dagger} U$ and $U^{\dagger}$ is the adjoint of U. Equivalently, the vector α is the least squares solution of the problem $U\alpha \approx \hat{f}$.

This system can be solved iteratively by applying conjugate gradient iterations to the normal equations, for example. The number of required iterations is dependent on the condition number $\kappa(A)$ of the matrix A. Specifically, the number of iterations required to obtain numerical convergence (i.e. to within a prescribed tolerance) is proportional to $\sqrt{\kappa(A)}$ [37]. In particular, if $\kappa(A)$ is $\mathcal{O}(1)$ for all n and $m \geq m_0$, then the number of iterations is also $\mathcal{O}(1)$ for all n. Hence, the cost of computing $f_{n,m}$ is determined solely by the number of operations required to perform matrix-vector multiplications involving U. In other words, only $\mathcal{O}(mn)$ operations.

Naturally, aside from this consideration, the condition number of A is also important since it determines susceptibility of the numerical computation to both round-off error and noise. Specifically, an error of magnitude ϵ in the inputs (i.e. the samples $\hat{f}_j$, $j = 1, \ldots, m$) will yield an error of magnitude roughly $\kappa(A)\epsilon$ in the output $f_{n,m}$.

For these reasons it is of utmost importance to study the condition number of A. For this, we first introduce the Hermitian matrix $\tilde{A} \in \mathbb{C}^{n \times n}$ with $(j,k)^{\text{th}}$ entry $\langle \phi_j, \phi_k \rangle$. Note that $\tilde{A}$ is the Gram matrix of the vectors $\{\phi_1, \ldots, \phi_n\}$. In particular, $\kappa(\tilde{A})$ is a measure of the suitability of the particular vectors in which to compute $\mathcal{Q}_n f$. With $\tilde{A}$ to hand, we also introduce the related matrix $\tilde{A}_a \in \mathbb{C}^{n \times n}$ with $(j,k)^{\text{th}}$ entry $a(\phi_j, \phi_k) = \langle \mathcal{P} \phi_j, \phi_k \rangle$, i.e. the Gram matrix with respect to the inner product $a(\cdot, \cdot)$.

The following lemma comes as no surprise:

Lemma 2.10. *The matrices $\tilde{A}$ and $\tilde{A}_a$ are spectrally equivalent. In particular, for all $n \in \mathbb{N}$,*

$$\frac{d_1}{d_2} \kappa(\tilde{A}) \leq \kappa(\tilde{A}_a) \leq \frac{d_2}{d_1} \kappa(\tilde{A}).$$

Proof. For any Hermitian matrix B, the condition number is the ratio of the largest and smallest eigenvalues in absolute value. Moreover, if B is positive definite, then

$$\inf_{\substack{\alpha \in \mathbb{C}^n \\ \alpha \neq 0}} \left\{ \frac{\alpha^\dagger B \alpha}{\alpha^\dagger \alpha} \right\} = \lambda_{\min}(B), \quad \sup_{\substack{\alpha \in \mathbb{C}^n \\ \alpha \neq 0}} \left\{ \frac{\alpha^\dagger B \alpha}{\alpha^\dagger \alpha} \right\} = \lambda_{\max}(B). \tag{2.20}$$

If $\phi = \sum_{j=1}^n \alpha_j \phi_j$, then $\alpha^\dagger \tilde{A} \alpha = \|\phi\|^2$ and $\alpha^\dagger \tilde{A}_a \alpha = a(\phi, \phi)$. Hence, spectral equivalence now follows immediately from (2.11). □

Concerning the condition number of the matrix A, we now have the following:

Lemma 2.11. *Suppose that $m \geq m_0$, where m_0 is as in Theorem 2.4, and $\tilde{C}_{n,m}$ and $C_{n,m}$ are given by (2.12) and (2.19) respectively. Then*

$$\tilde{C}_{n,m} \kappa(\tilde{A}_a) \leq \kappa(A) \leq \frac{1}{\tilde{C}_{n,m}} \kappa(\tilde{A}_a), \qquad \frac{C_{n,m}}{d_2} \kappa(\tilde{A}) \leq \kappa(A) \leq \frac{d_2}{C_{n,m}} \kappa(\tilde{A}).$$

Moreover, for fixed n, $A \to \tilde{A}_a$ as $m \to \infty$, and, if $\mathcal{P} = \mathcal{I}$, $A \to \tilde{A} = \tilde{A}_a$.

Proof. The matrix A is Hermitian and, provided $m \geq m_0$, positive definite. Hence, its eigenvalues are given by (2.20). For $\phi = \sum_{j=1}^n \alpha_j \phi_j$, we have $\alpha^\dagger A \alpha = \langle \mathcal{P}_m \phi, \phi \rangle$. By definition of $C_{n,m}$, for example, we find that $\lambda_{\max}(A) \geq C_{n,m} \lambda_{\max}(\tilde{A})$ and $\lambda_{\min}(A) \geq C_{n,m} \lambda_{\min}(\tilde{A})$. Moreover, by (2.1) we have $\lambda_{\max}(A) \leq d_2 \lambda_{\max}(\tilde{A})$ and $\lambda_{\min}(A) \leq d_2 \lambda_{\min}(\tilde{A})$. The first result now follows immediately from (2.20). For the second, we merely note that each entry of A converges to the corresponding entry of $\tilde{A}_a$ as $m \to \infty$. □

Note the important conclusion of this lemma: computing $f_{n,m}$ from (2.6) is no more ill-conditioned than the computation of the orthogonal projection $\mathcal{Q}_n f$ or the oblique projection $\mathcal{W}_n f$ in terms of the vectors $\{\phi_1, \ldots, \phi_n\}$. In practice, it is often true that these vectors correspond to the first n vectors in a basis $\{\phi_j\}_{j=1}^\infty$ of H with additional structure. Whenever this is the case, as the following trivial corollary indicates, we can expect good conditioning:

Corollary 2.12. *Suppose that $\{\phi_j\}_{j=1}^\infty$ is a Riesz basis for H with respect to $\langle \cdot, \cdot \rangle$ with constants c_1' and c_2'. Then*

$$\kappa(A) \leq \frac{c_2' d_2}{c_1' C_{n,m}}.$$

Proof. This follows immediately follows from (2.1) and Lemma 2.11. □

Put together, the main conclusion of Theorem 2.4, Lemma 2.11 and Corollary 2.12 is the following: for a given reconstruction space T_n, the individual vectors $\phi_1, \ldots, \phi_n$ can be chosen arbitrarily, without altering either the approximation $f_{n,m}$ or its analysis. The choice of vectors only becomes important when considering the condition number of linear system to solve. Moreover, the quality of a system of vectors for the reconstruction problem is completely intrinsic, in that it is determined only by the corresponding Gram matrix. In particular, it is independent of the sampling vectors.

Corollary 2.12 confirms that the approximation $f_{n,m}$ can be readily computed in a stable manner for many choices of reconstruction basis. However, to fully implement this method, as we discuss further in the next section, it is useful to have numerical way of computing $C_{n,m}$. The following lemma provides such a means:

Lemma 2.13. *The quantity $C_{n,m}$ is given by $C_{n,m} = \lambda_{\min}(\tilde{A}^{-1}A)$. Moreover, if $\tilde{A}$ and A commute, then $C_{n,m} = 1 - \|I - \tilde{A}^{-1}A\|$. In particular, if $\{\phi_j\}_{j=1}^n$ is an orthonormal basis, then $C_{n,m} = \lambda_{\min}(A) = 1 - \|I - A\|$.*

Proof. By definition $C_{n,m} = \inf_{\substack{\phi\in\mathrm{T}_n \\ \|\phi\|=1}} \langle \mathcal{P}_m\phi, \phi\rangle$. Letting $\phi = \sum_{j=1}^n \alpha_j\phi_j$, we find that

$$C_{n,m} = \inf_{\substack{\alpha\in\mathbb{C}^n \\ \alpha\neq 0}} \left\{ \frac{\sum_{j,k=1}^n \alpha_j\overline{\alpha}_k \langle \mathcal{P}_m\phi_j, \phi_k\rangle}{\sum_{j,k=1}^n \alpha_j\overline{\alpha}_k \langle \phi_j, \phi_k\rangle} \right\} = \inf_{\substack{\alpha\in\mathbb{C}^n \\ \alpha\neq 0}} \frac{\alpha^\dagger A\alpha}{\alpha^\dagger \tilde{A}\alpha}.$$

We now claim that, for arbitrary Hermitian positive definite matrices B and C with B nonsingular, the following holds:

$$\inf_{\substack{\alpha\in\mathbb{C}^n \\ \alpha\neq 0}} \frac{\alpha^\dagger C\alpha}{\alpha^\dagger B\alpha} = \lambda_{\min}(B^{-1}C), \quad \sup_{\substack{\alpha\in\mathbb{C}^n \\ \alpha\neq 0}} \frac{\alpha^\dagger C\alpha}{\alpha^\dagger B\alpha} = \lambda_{\max}(B^{-1}C).$$

To do so, write $B = D^\dagger D$, with D nonsingular. Then, after rearranging, we obtain

$$\inf_{\substack{\alpha\in\mathbb{C}^n \\ \alpha\neq 0}} \frac{\alpha^\dagger C\alpha}{\alpha^\dagger B\alpha} = \inf_{\substack{\beta\in\mathbb{C}^n \\ \beta\neq 0}} \frac{\beta^\dagger D^{-\dagger}CD^{-1}\beta}{\beta^\dagger\beta} = \lambda_{\min}(D^{-\dagger}CD^{-1}),$$

for example. However, a trivial calculation confirms that the eigenvalues of $D^{-\dagger}CD^{-1}$ are identical to those of $B^{-1}C$, thus establishing the claim. Since $\tilde{A}$ is nonsingular, this confirms that $C_{n,m} = \lambda_{\min}(\tilde{A}^{-1}A)$. For the second result, we merely notice that $\lambda_{\min}(B) = 1 - \lambda_{\max}(I - B) = 1 - \|I - B\|$ whenever B is Hermitian. □

In Section 2.2 we briefly discussed a modified approach where the operator $\mathcal{P}_m$, usually given by (2.4), was replaced by the orthogonal projection operator. The advantage of this approach is that it guarantees asymptotic optimality. However, the downside is additional computational expense. Indeed, the corresponding matrix is of the form $A = U^\dagger V^{-1}U$, where $V \in \mathbb{C}^{m\times m}$ has $(j,k)^{\text{th}}$ entry $\langle \psi_j, \psi_k\rangle$. Hence, if conjugate gradients iterations are used, at each stage we are required to compute matrix-vector products involving the $m \times m$ matrix V^{-1} (assuming that V^{-1} had been precomputed). In general, this requires $\mathcal{O}\left(m^2\right)$ operations. Thus, we incur a cost of $\mathcal{O}\left(m^2\right)$, as opposed to $\mathcal{O}\left(mn\right)$ for the original algorithm. Hence, in practice it may be better settle for only quasi- and oblique asymptotic optimality, whilst retaining a lower computational cost.

Remark 2.4 One assumption made in this section when considering computational cost is that the matrix U has already been formed. In general, computing the entries $\langle \phi_k, \psi_j\rangle$, $j = 1,\ldots,m$, $k = 1,\ldots,n$, of U may be a difficult task. In Section 3 we show that this can always be done in only $\mathcal{O}\left(mn\right)$ operations when reconstructing in a (Gegenbauer) polynomial basis from Fourier samples. However, this need not be the case in general, and in Section 5.2 we discuss an instance for which we currently only have an $\mathcal{O}\left(m^2\right)$ algorithm for computing U. Potential remedies for improving this figure are discussed in Section 6.

Remark 2.5 As detailed in [4], the ideas for the framework of this paper originate with the question of how to discretise certain infinite-dimensional operators. In particular, the matrix $U \in \mathbb{C}^{m\times n}$ is an *uneven* section of the operator $\mathcal{U} : l^2(\mathbb{N}) \to l^2(\mathbb{N})$ corresponding to the infinite matrix $\{\langle \phi_k, \psi_j\rangle\}_{j,k=1}^\infty$. Uneven section techniques – as opposed to finite sections, which are not guaranteed to succeed – have recently gained prominence in the discretisation of non-self adjoint problems [43, 45]. In particular, they were employed in [44] to solve the long-standing computational spectral problem. The key idea is that, by allowing m to range independently of n, one can guarantee that the structure of $\mathcal{U}$ is preserved by its uneven section U. The beneficial features of the framework introduced herein, namely, numerical stability and accuracy, are direct consequences of this property.

2.4 Conditions for guaranteed, quasi-optimal recovery

Let us return to the standard form of the method once more. To implement this method, it is necessary to have conditions that guarantee nonsingularity, stability and quasi-optimal recovery. In other words, for given sampling and reconstruction bases, we wish to study the quantity

$$\Theta(n;\theta) = \min\{m \in \mathbb{N} : C_{n,m} \geq \theta\}, \quad \theta \in (0, d_2), \tag{2.21}$$

where $C_{n,m}$ is given by (2.12) and d_2 stems from (2.2) . Note that $C_{n,m} \leq d_2$ by (2.2), thereby explaining the stated range of θ. Also, by Lemma 2.2, we have that $\lim_{m\to\infty} C_{n,m} \geq d_1 > 0$, thus Θ is well-defined.

By definition, $\Theta(n;\theta)$ is the least m such that $\|f - f_{n,m}\| \leq c(\theta)\|f - \mathcal{Q}_n f\|$, where

$$c(\theta) = \sqrt{1 + d_2^2\theta^{-2}} \quad \text{or} \quad \sqrt{1 + (1-\theta)\theta^{-2}}, \tag{2.22}$$

whenever the sampling basis is orthonormal. In other words, the least m required for quasi-optimal recovery with constant $c(\theta)$. Thus, provided $m \geq \Theta(n;\theta)$, the approximation $f_{n,m}$ converges at the same rate as $\mathcal{Q}_n f$ as $n \to \infty$. In addition, $m \geq \Theta(n;\theta)$ guarantees that $\|f_{n,m}\| \leq d_2\theta^{-1}\|f\|$ and $\kappa(A) \leq d_2\theta^{-1}\kappa(\tilde{A})$, thus making the linear system for $f_{n,m}$ solvable in a number of operations proportional to $\sqrt{d_2\theta^{-1}\kappa(\tilde{A})}$.

Note that $\Theta(n;\theta)$ is determined only by the sampling vectors $\{\psi_j\}_{j=1}^m$ and reconstruction space T_n. Whilst $\Theta(n;\theta)$ can be numerically computed for any such pair via the expression given in Lemma 2.13, analytical bounds must be determined on a case-by-case basis. In the next section, where we consider the recovery of functions from their Fourier samples using (piecewise) polynomial bases, we are able to derive explicit forms for such bounds.

2.5 Summary

Let us sum up. Given the first m samples of any element $f \in \mathrm{H}$ with respect to any Riesz basis, it is possible to reconstruct f in an arbitrary finite-dimensional space T_n, provided the parameter m is sufficiently large in comparison to $n = \dim \mathrm{T}_n$. The resulting reconstruction $f_{n,m}$ is quasi-optimal, and can be computed in a completely numerically stable manner. Furthermore, the required scaling of m with n can be determined numerically by finding either the minimal eigenvalue or, in certain cases, the norm of an $n \times n$ matrix.

Remark 2.6 As mentioned, the framework developed in this section was first introduced by the authors in [4]. Whilst a result similar to Theorem 2.4 was proved, there are a number of important improvements offered by the theory presented in this paper:

1. In [4] it was assumed that the reconstruction vectors $\phi_1, \ldots, \phi_n$ were the first n in an infinite sequence of vectors that formed a Riesz basis for H. Conversely, Theorem 2.4 depends only on the subspace T_n, and thus the individual reconstruction vectors can be chosen arbitrarily.
2. The constants $K_{n,m}$ and $C_{n,m}$ are known exactly in terms of the sampling and reconstruction bases, and can be computed numerically.
3. Simple, explicit bounds for the condition number of the matrix A are known in terms of the constant $C_{n,m}$ and the Gram matrix $\tilde{A}$.
4. The behaviour of $f_{n,m}$ as $m \to \infty$ (for n fixed) can be fully explained in terms of oblique asymptotic optimality.

3 Polynomial reconstructions from Fourier samples

One of the most important examples of this procedure is the reconstruction of an analytic, but nonperiodic function f to high accuracy from its Fourier coefficients. Direct expansion in Fourier series converges only slowly in the L^2 norm, and suffers from the Gibbs phenomenon near the domain boundary. Hence, given the first m Fourier coefficients of f, we now seek to reconstruct f to high accuracy in another basis using the procedure developed in Section 2.

Let $\mathrm{H} = \mathrm{L}^2(-1,1)$, $f : (-1,1) \to \mathbb{R}$ and

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

be the standard Fourier basis. For $m \geq 2$, we assume that the coefficients

$$\hat{f}_j = \int_{-1}^{1} f(x) \overline{\psi_j(x)} \, \mathrm{d}x, \quad j = -\left\lfloor \frac{m}{2} \right\rfloor + 1, \ldots, \left\lfloor \frac{m}{2} \right\rfloor - 1,$$

are known (note that, whenever m is even, this means that the first $m-1$ Fourier coefficients of f are given. We will allow this minor discrepancy since it simplifies ensuing analysis). As a consequence of Theorem 2.4, we are free to choose the reconstruction space. The orthogonal projection of an analytic function onto the space $\mathbb{P}_{n-1}$ of polynomials of degree less than n is known to converge exponentially fast at rate ρ^{-n}, where $\rho > 1$ is determined from the largest Bernstein ellipse within which f is analytic [58]. Hence, we let $\mathrm{T}_n = \mathbb{P}_{n-1}$. Note that an orthonormal basis for T_n is given by the functions

$$\phi_j(x) = \sqrt{j + \tfrac{1}{2}} P_j(x), \quad j \in \mathbb{N}, \tag{3.1}$$

where P_j is the j^{th} Legendre polynomial. Moreover, if $\mathcal{Q}_n$ is the orthogonal projection onto T_n, then it is well-known that

$$\| f - \mathcal{Q}_n f \| \leq c_f \sqrt{n} \rho^{-n}, \tag{3.2}$$

where c_f depends only on the maximal value of f on the Bernstein ellipse indexed by ρ. Naturally, we could also assume finite regularity of f throughout, with suitable adjustments made to the various error estimates. However, for simplicity we shall not do this.

With this to hand, provided $m \geq \Theta(n;\theta)$, where $\Theta(n;\theta)$ is defined in (2.21), the approximation $f_{n,m}$ obtained from the reconstruction procedure satisfies $\| f - f_{n,m} \| \leq c(\theta) \| f - \mathcal{Q}_n f \|$ (see Theorem 2.4). In particular, $\| f - f_{n,m} \| \leq c(\theta) c_f \sqrt{n} \rho^{-n}$. Hence, we obtain exponential convergence of $f_{n,m}$. The key question remaining is how large m must be in comparison to n to ensure such behaviour. Resolving this question involves estimating the quantity $\Theta(n;\theta)$, a task we next pursue.

3.1 Estimates for $\Theta(n;\theta)$

Although $\Theta(n;\theta)$ is independent of the particular basis of $\mathbb{P}_{n-1}$ used, for both numerical and analytical estimates we need to select an appropriate basis. A natural choice is the orthonormal basis (3.1) of scaled Legendre polynomials. Fortunately, in this case, the inner products $\langle \phi_k, \psi_j \rangle$ (i.e. the entries of the matrix U) are known in closed form

$$\langle \phi_k, \psi_j \rangle = (-\mathrm{i})^k \sqrt{\frac{k + \frac{1}{2}}{j}} J_{k+\frac{1}{2}}(j\pi), \quad j \in \mathbb{Z},\ k \in \mathbb{N}, \tag{3.3}$$

where J_m is the Bessel function of first kind. This follows directly from the integral representation

$$j_m(z) = \frac{1}{2} (-\mathrm{i})^m \int_{-1}^{1} \mathrm{e}^{\mathrm{i} z x} P_m(x) \, \mathrm{d}x, \quad \forall z \in \mathbb{C}, \tag{3.4}$$

(see [1, 10.1.14]), where j_m is the spherical Bessel function of the first kind, given by

$$j_m(z) = \sqrt{\frac{\pi}{2z}} J_{m+\frac{1}{2}}(z).$$

With this to hand, we may compute $C_{n,m}$, and, in turn, $\Theta(n;\theta)$, via the expression given in Lemma 2.13. In Figure 1 we display the functions $\Theta(n; \frac{1}{2})$ and $\Theta(n; \frac{1}{4})$ against n. Immediately, quadratic growth of $\Theta(n;\theta)$ with n is apparent. We next verify this observation. In doing so, we derive an upper bound for $\Theta(n;\theta)$ in terms of n and θ. This gives an explicit, analytic condition for quasi-optimal recovery. Whilst such a bound is global, in that it holds for all n, we notice from Figure 1 that $\Theta(n;\theta)$, when scaled by n^{-2}, quickly converges to an asymptotic limit. In practice it is wasteful to use a larger value of m than necessary (or, conversely, for fixed m a overly pessimistic value of n). Hence, in the second part of this section, we will also derive an asymptotic bound for $\Theta(n;\theta)$.

We commence as follows:

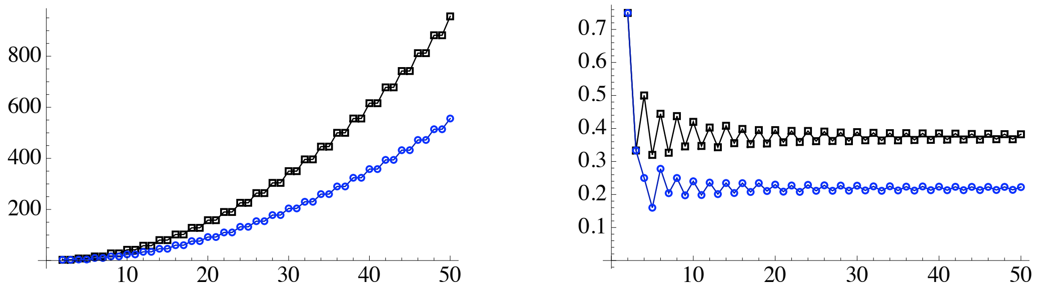

Figure 1: The functions $\Theta(n;\theta)$ (left) and $n^{-2}\Theta(n;\theta)$ (right) for $\theta = \frac{1}{2}$ (squares) and $\theta = \frac{1}{4}$ (circles).

Lemma 3.1. *Suppose that $\{\psi_j\}_{j\in\mathbb{Z}}$ is the Fourier basis, $\mathrm{T}_n = \mathbb{P}_{n-1}$ and $m \geq \max\{2, \frac{2}{\pi}n\}$. Then $C_{n,m}$ satisfies*

$$C_{n,m} \geq 1 - \frac{4(\pi-2)n^2}{\pi^2(2\lfloor \frac{m}{2} \rfloor - 1)}.$$

Proof. From the definition of $C_{n,m}$ and the fact that $\{\psi_j\}$ is an orthonormal basis we have

$$1 - C_{n,m} = 1 - \inf_{\substack{\phi\in\mathrm{T}_n \\ \|\phi\|=1}} \langle \mathcal{P}_m\phi, \phi\rangle = \sup_{\substack{\phi\in\mathrm{T}_n \\ \|\phi\|=1}} \langle \phi - \mathcal{P}_m\phi, \phi\rangle = \sup_{\substack{\phi\in\mathrm{T}_n \\ \|\phi\|=1}} \|\phi - \mathcal{P}_m\phi\|^2,$$

where $\mathcal{P}_m$ is the Fourier projection operator. It now follows that $1 - C_{n,m} \leq \sum_{k=0}^{n-1} \|\phi_k - \mathcal{P}_m\phi_k\|^2$, where ϕ_k is given by (3.1). By Parseval's theorem and the expression (3.3), we find that

$$\|\phi_k - \mathcal{P}_m\phi_k\|^2 = \sum_{|j|\geq\lfloor\frac{m}{2}\rfloor} \frac{k+\frac{1}{2}}{j}|J_{k+\frac{1}{2}}(j\pi)|^2.$$

Now, using a known result for Bessel functions [47], it can be shown that

$$\frac{k+\frac{1}{2}}{j}|J_{k+\frac{1}{2}}(j\pi)|^2 \leq \frac{2k+1}{j\pi\sqrt{j^2\pi^2 - (k+\frac{1}{2})^2}},$$

provided $j\pi > k + \frac{1}{2}$. Hence, for $m > \frac{2}{\pi}n$,

$$\|\phi_k - \mathcal{P}_m\phi_k\|^2 \leq \frac{2(2k+1)}{\pi^2} \sum_{j\geq\lfloor\frac{m}{2}\rfloor} \frac{1}{j\sqrt{j - \frac{(k+\frac{1}{2})^2}{\pi^2}}}. \tag{3.5}$$

Now, it was shown in [47] that $\sum_{j\geq m} \frac{1}{j\sqrt{j^2-c^2}} \leq \frac{1}{c} \arcsin \frac{c}{m-\frac{1}{2}}$, whenever $m \geq c + \frac{1}{2}$. This gives

$$\|\phi_k - \mathcal{P}_m\phi_k\|^2 \leq \frac{4}{\pi} \arcsin\left[\frac{2k+1}{(2\lfloor\frac{m}{2}\rfloor - 1)\pi}\right],$$

and

$$1 - C_{n,m} \leq \frac{4}{\pi} \sum_{k=0}^{n-1} \arcsin\left[\frac{2k+1}{(2\lfloor\frac{m}{2}\rfloor - 1)\pi}\right].$$

We estimate this sum by the integral of $\arcsin t$, thus giving

$$1 - C_{n,m} \leq 2(2\lfloor\tfrac{m}{2}\rfloor - 1) \int_0^{\frac{2n}{(2\lfloor\frac{m}{2}\rfloor-1)\pi}} \arcsin t \,\mathrm{d}t.$$

Now it can be shown that $F(x) = \int_0^x \arcsin t \,\mathrm{d}t \leq (\frac{\pi}{2} - 1)x^2$. Upon substituting $x = \frac{2n}{(2\lfloor\frac{m}{2}\rfloor-1)\pi}$, this completes the proof. □

Theorem 3.2. *Suppose that* T_n *and* S_m *are as in Lemma 3.1. Then, for* $n \geq 2$, $\Theta(n;\theta)$ *satisfies*

$$\Theta(n;\theta) \leq 2\left\lceil \frac{1}{2} + \frac{2(\pi-2)}{\pi^2(1-\theta)} n^2 \right\rceil, \quad \forall n \in \mathbb{N}.$$

Proof. Suppose that $m \geq \{2, \frac{2}{\pi}n\}$. Then, by Lemma 3.1, $C_{n,m} \geq \theta$ provided

$$1 - \frac{4(\pi-2)n^2}{\pi^2(2\lfloor \frac{m}{2} \rfloor - 1)} \geq \theta.$$

Rearranging, we find that

$$2\left\lfloor \frac{m}{2} \right\rfloor \geq 1 + \frac{4(\pi-2)n^2}{\pi^2(1-\theta)} \quad \Rightarrow \quad m \geq 2\left\lceil \frac{1}{2} + \frac{2(\pi-2)}{\pi^2(1-\theta)} n^2 \right\rceil$$

and the theorem is proved, provided the right-hand side exceeds $\max\{2, \frac{2}{\pi}n\}$. Since $n \geq 2$, the right-hand side is certainly greater than 2. Moreover,

$$1 + \frac{4(\pi-2)n^2}{\pi^2(1-\theta)} \geq \frac{8(\pi-2)n}{\pi} > \frac{2n}{\pi},$$

as required. □

Using a similar approach, we are also able to obtain an asymptotic bound for $\Theta(n;\theta)$, valid as $n \to \infty$, that is sharper than if were to use Theorem 3.2 directly:

Theorem 3.3. *Suppose that* $\{\psi_j\}_{j\in\mathbb{Z}}$ *and* T_n *are as in Lemma 3.1. Then the function* $\Theta(n;\theta)$ *satisfies*

$$n^{-2}\Theta(n;\theta) \leq \frac{4}{\pi^2(1-\theta)} + \mathcal{O}\left(n^{-2}\right), \quad n \to \infty.$$

Proof. Suppose that $m = cn^2$ and recall (3.5). Since $j < n$ and $k > \frac{1}{2}cn^2$, we deduce that

$$\|\phi_k - \mathcal{P}_m \phi_k\|^2 \leq \frac{2(2k+1)}{\pi^2} \sum_{j > \lfloor \frac{m}{2} \rfloor} \frac{1}{j^2} + \mathcal{O}\left(n^{-4}\right) = \frac{4(2k+1)}{c\pi^2 n^2} + \mathcal{O}\left(n^{-4}\right).$$

Hence

$$1 - C_{n,cn^2} \leq \frac{4}{c\pi^2 n^2} \sum_{k=0}^{n-1} (2k+1) + \mathcal{O}\left(n^{-2}\right) = \frac{4}{c\pi^2} + \mathcal{O}\left(n^{-2}\right).$$

Rearranging now gives the result. □

In Figure 2 we compare the function $n^{-2}\Theta(n;\theta)$ for $\theta = \frac{1}{2}, \frac{1}{4}$ and the global and asymptotic bounds of Theorems 3.2 and 3.3. Both bounds are reasonably sharp in comparison to the computed values. In particular, as $n \to \infty$, $n^{-2}\Theta(n;\frac{1}{2})$ quickly approaches the limiting value $c \approx 0.38$, whereas the global and asymptotic upper bounds are 0.93 and 0.81 respectively.

At this moment, we reiterate an important point. Whilst Legendre polynomials were used in the proof of Lemma 3.1, the constant $C_{n,m}$ is independent of the particular reconstruction basis, and is only determined by the space T_n. Hence, Theorems 3.2 and 3.3 provide *a priori* estimates regardless of the particular implementation of the reconstruction procedure. In the next section, we discuss the choice of polynomial basis and its effect on the numerical method. Note that Theorems 3.2 and 3.3 establish parts (i) and (ii) of Theorem 1.1. Parts (iii) and (iv) will be addressed in the next section.

Remark 3.1 In some applications, medical imaging, for example, oversampling is common. Formally speaking, this is the situation where we wish to recover a function f with support in $[-1,1]$ from its Fourier samples taken over an extended interval $K \supseteq [-1,1]$ (e.g. $K = [-\frac{1}{\epsilon}, \frac{1}{\epsilon}]$ for some $0 < \epsilon \leq 1$). In this case, proceeding in a similar manner to before, we let $\mathrm{H} = \mathrm{L}^2(K)$,

$$\psi_j(x) = \sqrt{\tfrac{c}{2}} \mathrm{e}^{\mathrm{i}cj\pi x}, \quad x \in K,$$

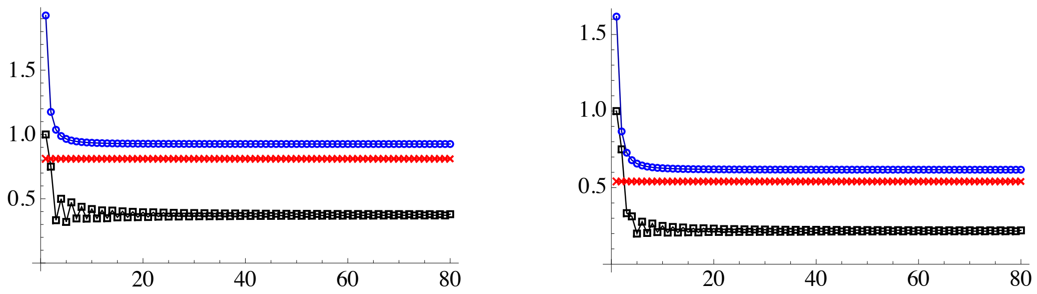

Figure 2: The function $n^{-2}\Theta(n;\theta)$ (squares), the global bound (circles) and the asymptotic bound (crosses), for $n=2,\ldots,80$ and $\theta=\frac{1}{2}$ (left), $\theta=\frac{1}{4}$ (right).

where $c=\frac{1}{2}|K|$ and $\mathrm{T}_n=\left\{\phi:\phi|_{[-1,1]}\in\mathbb{P}_{n-1},\ \mathrm{supp}(\phi)\subseteq[-1,1]\right\}$. Using similar arguments to those of Lemma 3.1, one can also derive estimates for $C_{n,m}$ and $\Theta(n;\theta)$ in this case. In fact,

$$C_{n,m}\geq 1-\frac{4(\pi-2)n^2}{c\pi^2(m-1)}, \tag{3.6}$$

and

$$\Theta(n;\theta)\leq 2\left[\frac{1}{2}+\frac{2(\pi-2)}{c\pi^2(1-\theta)}n^2\right],\ \forall n\in\mathbb{N}, \qquad n^{-2}\Theta(n;\theta)\leq\frac{4}{c\pi^2(1-\theta)}+\mathcal{O}\left(n^{-2}\right),\ n\to\infty.$$

In particular, we retain the scaling $m=\mathcal{O}\left(n^2\right)$, regardless of the of size of the interval K.

3.2 Choice of polynomial basis

The results proved in this section are independent of the polynomial basis used for implementation. In selecting such a basis, there are two questions which must be resolved. First, how stable is the resultant method, and second, how can the entries of the matrix U (as defined in Section 2.3) be computed? A straightforward choice is the orthogonal basis of Legendre polynomials (3.1). In this case, $\tilde{A}=I$, where $\tilde{A}$ is the Gram matrix for $\{\phi_0,\ldots,\phi_{n-1}\}$, making the method well-conditioned (Lemma 2.11). Moreover, the entries of U are known explicitly via (3.3).

Having said this, there is also interest in reconstructing in other polynomial bases. In many circumstances it may be advantageous to have an approximation $f_{n,m}$ that is easily manipulable. In this sense, an approximant composed of Legendre polynomials is not as convenient as one consisting of Chebyshev polynomials (of the first or second kind); the latter being easy to manipulate with the Fast Fourier Transform (FFT). To this end, the purpose of this section is to detail the implementation of this method in terms of general Gegenbauer polynomials.

Gegenbauer polynomials arise as orthogonal polynomials with respect to the inner product

$$\langle f,g\rangle_\lambda=\int_{-1}^{1}f(x)\overline{g(x)}(1-x^2)^{\lambda-\frac{1}{2}}\,\mathrm{d}x,\quad \lambda>-\frac{1}{2}.$$

For given λ, we denote the j^{th} such polynomial by $C_j^\lambda\in\mathbb{P}_j$. Important special cases are the Legendre polynomials ($\lambda=\frac{1}{2}$), and Chebyshev polynomials of the first ($\lambda=0$) and second ($\lambda=1$) kinds. By convention [10] (see also [42]), each polynomial C_j^λ is normalised so that

$$C_j^\lambda(1)=\frac{\Gamma(j+2\lambda)}{j!\Gamma(2\lambda)}, \tag{3.7}$$

where Γ is the Gamma function, in which case it is known that (see [10, p.174])

$$\|C_j^\lambda\|_\lambda^2=\sqrt{\pi}\frac{\Gamma(j+2\lambda)\Gamma(\lambda+\frac{1}{2})}{j!\Gamma(2\lambda)\Gamma(\lambda)(j+\lambda)}, \tag{3.8}$$

where $\|f\|_\lambda = \sqrt{\langle f, f\rangle_\lambda}$. With this to hand, we now define

$$\phi_j = \frac{1}{\|C_j^\lambda\|_\lambda} C_j^\lambda, \quad j = 0, 1, 2, \ldots, \tag{3.9}$$

and seek to reconstruct f in this basis.

Our first task is to compute the entries of the matrix U. For this, we need to compute integrals of the form

$$I_k(z) = \int_{-1}^{1} C_k^\lambda(x) \mathrm{e}^{\mathrm{i}zx} \,\mathrm{d}x, \qquad k = 0, 1, 2, \ldots,$$

where $z \in \mathbb{R}$. Fortunately, such integrals obey a simple recurrence relation:

Lemma 3.4. *For $z \neq 0$, the integrals $I_k(z)$ satisfy*

$$I_0(z) = 2C_0^\lambda(1)\frac{\sin z}{z}, \quad I_1(z) = 2\mathrm{i}C_1^\lambda(1)\frac{\sin z - z\cos z}{z^2},$$
$$I_{k+1}(z) = \frac{2\mathrm{i}(k+\lambda)}{z} I_k(z) + I_{k-1}(z) - \mathrm{i}\frac{\mathrm{e}^{\mathrm{i}z} + (-1)^k \mathrm{e}^{-\mathrm{i}z}}{z}\left[C_{k+1}^\lambda(1) - C_{k-1}^\lambda(1)\right], \quad k = 1, 2, \ldots.$$

When $z = 0$, we have

$$I_0(0) = 2C_0^\lambda(1), \qquad I_k(0) = \frac{1 + (-1)^k}{2(k+\lambda)}\left[C_{k+1}^\lambda(1) - C_{k-1}^\lambda(1)\right], \; k = 1, 2, \ldots.$$

Proof. Recall the identity (see [10, p.176])

$$C_j^\lambda(x) = \frac{1}{2(j+\lambda)}\frac{\mathrm{d}}{\mathrm{d}x}\left[C_{j+1}^\lambda - C_{j-1}^\lambda\right], \quad j = 1, 2, \ldots.$$

Substituting this into the expression for $I_k(z)$ and integrating by parts gives

$$I_k(z) = \frac{1}{2(k+\lambda)}\left[\left(C_{k+1}^\lambda(x) - C_{k-1}^\lambda(x)\right)\mathrm{e}^{\mathrm{i}zx}\right]_{x=-1}^{1} - \frac{\mathrm{i}z}{2(k+\lambda)}\left[I_{k+1}(z) - I_{k-1}(z)\right].$$

Rearranging now yields the general recurrence for $k \geq 1$. For $k = 0, 1$, we merely note that $C_0^\lambda(x) = C_0^\lambda(1)$, $C_1^\lambda(x) = C_1^\lambda(1)x$ and that

$$\int_{-1}^{1} \mathrm{e}^{\mathrm{i}zx}\,\mathrm{d}x = 2\frac{\sin z}{z}, \qquad \int_{-1}^{1} x\mathrm{e}^{\mathrm{i}zx}\,\mathrm{d}x = 2\frac{\sin z - z\cos z}{z^2}.$$

The result for $z = 0$ is derived in a similar manner. □

Using this recurrence formula, the matrix U can be formed in $\mathcal{O}(mn)$ operations.

Remark 3.2 In the case of Chebyshev polynomials, this iteration is well known. However, as discussed in detail in [22], this iteration is only stable for parameter values $k \leq |z|$. Fortunately, this iteration can be replaced by a two-phase algorithm in order to determine those integrals $I_k(z)$ with $k > |z|$. This hybrid algorithm has been shown to be stable, whilst at the same time maintaining the overall cost [22]. It is likely that a variant of this algorithm could also be used for arbitrary Gegenbauer polynomials.

Such considerations aside, we now turn our attention to the condition number of $\tilde{A}$:

Theorem 3.5. *Let $\tilde{A}$ be Gram matrix for the vectors $\{\phi_0, \ldots, \phi_{n-1}\}$, where ϕ_j is given by (3.9). Then, $\kappa(\tilde{A}) = \mathcal{O}\left(n^{|2\lambda-1|}\right)$ as $n \to \infty$. In particular, whenever $\phi_0, \ldots, \phi_{n-1}$ arise from Chebyshev polynomials (of the first or second kinds), then $\kappa(\tilde{A}) = \mathcal{O}(n)$.*

To prove this theorem, we first require the following two lemmas. For convenience, we will write $\mathrm{L}^2_\lambda(-1,1)$, $\lambda > -\frac{1}{2}$, for the space of square-integrable functions with respect to the Gegenbauer weight function $(1-x^2)^{\lambda-\frac{1}{2}}$.

Lemma 3.6. *Suppose that $-\frac{1}{2} < \lambda < \frac{1}{2}$. Then, for all $g \in \mathrm{L}^{\infty}(-1,1)$, we have $\|g\| \leq \|g\|_{\lambda}$ and*

$$\|g\|_{\lambda} \leq c_{\lambda} \|g\|^{\lambda+\frac{1}{2}} \|g\|_{\infty}^{\frac{1}{2}-\lambda}, \tag{3.10}$$

for some $c_{\lambda} > 0$ independent of g. Conversely, if $\lambda \geq \frac{1}{2}$ then, for all $g \in \mathrm{L}^{\infty}(-1,1)$, $\|g\| \leq \|g\|_{\lambda}$ and

$$\|g\| \leq c_{\lambda} \|g\|_{\lambda}^{\frac{1}{\lambda+\frac{1}{2}}} \|g\|_{\infty}^{\frac{\lambda-\frac{1}{2}}{\lambda+\frac{1}{2}}}. \tag{3.11}$$

Proof. Suppose first that $-\frac{1}{2} < \lambda < \frac{1}{2}$. Trivially, $\|g\| \leq \|g\|_{\lambda}$. Now consider the other inequality. For any $0 < \epsilon < 1$, we have

$$\begin{aligned}
\|g\|_{\lambda}^2 &= \int_{-1}^{1} |g(x)|^2 (1-x^2)^{\lambda-\frac{1}{2}} \,\mathrm{d}x \\
&= \int_{|x| \leq 1-\epsilon} |g(x)|^2 (1-x^2)^{\lambda-\frac{1}{2}} \,\mathrm{d}x + \int_{1-\epsilon<|x|\leq 1} |g(x)|^2 (1-x^2)^{\lambda-\frac{1}{2}} \,\mathrm{d}x \\
&\leq (1-(1-\epsilon)^2)^{\lambda-\frac{1}{2}} \|g\|^2 + 2\|g\|_{\infty}^2 \int_{1-\epsilon}^{1} (1-x^2)^{\lambda-\frac{1}{2}} \,\mathrm{d}x,
\end{aligned}$$

where $\|\cdot\|_{\infty}$ is the uniform norm on $[-1,1]$. Note that $(1-(1-y)^2)^{\lambda-\frac{1}{2}} < y^{\lambda-\frac{1}{2}}$, $\forall y \in (0,1)$. It follows that

$$\|g\|_{\lambda}^2 \leq \epsilon^{\lambda-\frac{1}{2}} \|g\|^2 + \frac{2}{\lambda+\frac{1}{2}} \epsilon^{\lambda+\frac{1}{2}} \|g\|_{\infty}^2, \quad 0 < \epsilon < 1.$$

Let $c > 2$ be arbitrary. Then $\|g\|^2 < c\|g\|_{\infty}^2$, so we may let $\epsilon = \frac{\|g\|^2}{c\|g\|_{\infty}^2} < 1$. Substituting this into the previous expression immediately gives (3.10).

Now suppose that $\lambda > \frac{1}{2}$. Once more, trivial arguments give that $\|g\|_{\lambda} \leq \|g\|$. For the other inequality, we proceed in a similar manner. We have $\|g\|^2 \leq \epsilon^{\frac{1}{2}-\lambda} \|g\|_{\lambda}^2 + 2\epsilon \|\phi\|_{\infty}^2$. For $c > 2$ we now set $\epsilon = \left(\frac{\|g\|_{\lambda}}{c\|g\|_{\infty}} \right)^{\frac{2}{\lambda+\frac{1}{2}}}$, which gives (3.11). □

Lemma 3.7. *Let $\lambda \geq \frac{1}{2}$. Then, for all $\phi \in \mathbb{P}_{n-1}$, $\|\phi\|_{\infty} \leq k_{n,\lambda} \|\phi\|_{\lambda}$, where $k_{n,\lambda}$ depends only on n and λ and satisfies $k_{n,\lambda} = \mathcal{O}\left(n^{\lambda+\frac{1}{2}}\right)$ as $n \to \infty$.*

Proof. Let $\{\phi_0, \ldots, \phi_{n-1}\}$ be given by (3.9), and write an arbitrary $\phi \in \mathbb{P}_{n-1}$ as $\phi = \sum_{j=0}^{n-1} a_j \phi_j$, where $\sum_{j=0}^{n-1} |a_j|^2 = \|\phi\|_{\lambda}^2$. By the Cauchy–Schwarz inequality,

$$\|\phi\|_{\infty}^2 \leq \|\phi\|_{\lambda}^2 \sum_{j=0}^{n-1} \|\phi_j\|_{\infty}^2 = k_{n,\lambda}^2 \|\phi\|_{\lambda}^2.$$

We now wish to estimate $\|\phi_j\|_{\infty}$. Recall that $\|C_j^{\lambda}\|_{\infty} = C_j^{\lambda}(1)$ [10, p.206]. Hence, by (3.7) and (3.8), we have

$$\|\phi_j\|_{\infty}^2 = \frac{\Gamma(j+2\lambda)(j+\lambda)}{j!\sqrt{\pi}\Gamma(2\lambda)\Gamma(\lambda+\frac{1}{2})}.$$

Consider the ratio $\frac{\Gamma(j+2\lambda)}{j!}$. By Stirling's formula,

$$\frac{\Gamma(j+2\lambda)}{j!} = \mathcal{O}\left(j^{2\lambda-1}\right), \quad j \to \infty.$$

Hence $\|\phi_j\|_{\infty}^2 = \mathcal{O}\left(j^{2\lambda}\right)$, which gives $k_{n,\lambda}^2 = \mathcal{O}\left(n^{2\lambda+1}\right)$, as required. □

Proof of Theorem 3.5. Since $\tilde{A}$ is Hermitian and positive definite, its condition number is the ratio of its maximum and minimum eigenvalues. By a simple argument, we find that

$$\lambda_{\max}(\tilde{A}) = \sup_{\substack{\phi \in \mathbb{P}_{n-1} \\ \phi \neq 0}} \frac{\|\phi\|^2}{\|\phi\|_{\lambda}^2}, \quad \lambda_{\min}(\tilde{A}) = \inf_{\substack{\phi \in \mathbb{P}_{n-1} \\ \phi \neq 0}} \frac{\|\phi\|^2}{\|\phi\|_{\lambda}^2}.$$

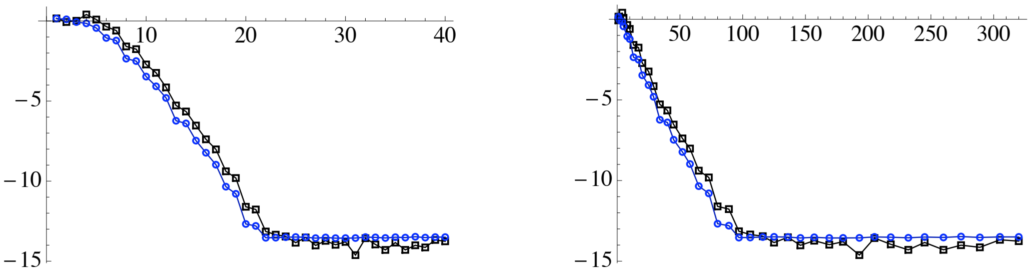

Figure 3: Error in approximating $f(x) = \mathrm{e}^{-x}\cos 4x$ by $f_{n,m}(x)$ for $n = 1, \ldots, 40$. Left: log error $\log_{10}\|f - f_{n,m}\|_\infty$ (squares) and $\log_{10}\|f - f_{n,m}\|$ (circles) against n. Right: log error against $m = 0.2n^2$.

Consider the case $\lambda > \frac{1}{2}$. By Lemma 3.6, we have $\lambda_{\min}(\tilde{A}) \geq 1$ and

$$\lambda_{\max}(\tilde{A}) \leq c_\lambda^2 \sup_{\substack{\phi\in\mathbb{P}_{n-1}\\ \phi\neq 0}} \left(\frac{\|\phi\|_\infty}{\|\phi\|_\lambda}\right)^{2\frac{\lambda-\frac{1}{2}}{\lambda+\frac{1}{2}}} .$$

Using Lemma 3.7, we deduce that $\lambda_{\max}(\tilde{A}) = \mathcal{O}\left(n^{2\lambda-1}\right)$, as required. For the case $-\frac{1}{2} < \lambda < \frac{1}{2}$, we proceed in a similar manner. □

This theorem confirms that the method can be implemented using Chebyshev polynomials whilst incurring only a mild growth in the condition number. Numerical results confirm the sharpness of the $\mathcal{O}(n)$ estimate for $\kappa(\tilde{A})$ in this case. It follows that, if conjugate gradients are used to compute the approximation, the total computational cost of forming $f_{n,m}$ is $\mathcal{O}(mn^{\frac{3}{2}})$, as opposed to $\mathcal{O}(mn)$ in the Legendre polynomial case. In the next section we present several examples of this implementation.

Remark 3.3 Whilst Theorem 3.5 provides an asymptotic estimate for $\kappa(\tilde{A})$ (and hence $\kappa(A)$), it may also be useful to derive global bounds. With effort, one could obtain versions of Lemmas 3.6 and 3.7 involving explicit bounds. For the sake of brevity, we shall not do this. However, whenever Chebyshev polynomials are used (arguably the most important case), it is possible to show that

$$\kappa(\tilde{A}) \leq 2\sqrt{2}n, \qquad \kappa(\tilde{A}) \leq 3^{\frac{2}{3}}\pi^{-\frac{1}{3}}\left[n(n+\tfrac{1}{2})(n+1)\right]^{\frac{1}{3}},$$

in the first and second kind cases respectively.

Observe that this theorem, in combination with Theorems 2.4, 3.2 and the arguments of this section, establish one of the main results of this paper: namely, Theorem 1.1.

3.3 Numerical examples

We now present several numerical examples of this method. All examples employ the value $m = 0.2n^2$, and the first series of examples consider the implementation using Legendre polynomials. In Figure 3 we consider the function $f(x) = \mathrm{e}^{-x}\cos 4x$. Since f is analytic in this case, we witness exponential convergence in terms of n and root exponential convergence in terms of m. Note the effectiveness of the method: using less than 100 Fourier coefficients, we obtain an approximation with 13 digits of accuracy.

As indicated by Theorem 2.4, the approximation $f_{n,m}$ is quasi-optimal. To highlight this feature of the method, Figure 4 displays both the error in approximating f by $f_{n,m}$ and the best approximation $\mathcal{Q}_n f$. Note the very close correspondence of the two graphs.

The example in Figures 3 and 4 is, in fact, entire. Hence, the approximation $f_{n,m}$ converges supergeometrically in n (as seen in Figure 3). For a meromorphic function, with complex singularity lying outside $[-1, 1]$, the convergence rate is truly exponential at a rate ρ. This is demonstrated in Figure 5, the approximated function being $f(x) = \frac{1}{1+x^2}$. Note that, despite the poles at $x = \pm\mathrm{i}$, the approximation $f_{n,m}$ still obtains 13 digits of accuracy using only 250 Fourier coefficients.

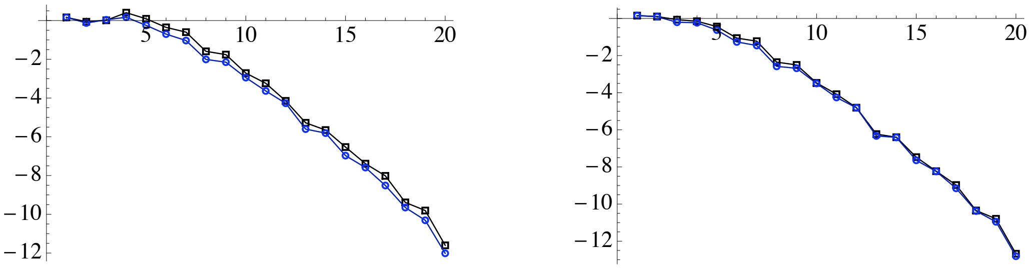

Figure 4: Error in approximating $f(x) = \mathrm{e}^{-x}\cos 4x$ by $f_{n,m}(x)$ (squares) and $\mathcal{Q}_n f(x)$ (circles) for $n = 1,\ldots,20$. Left: log uniform error. Right log L^2 error.

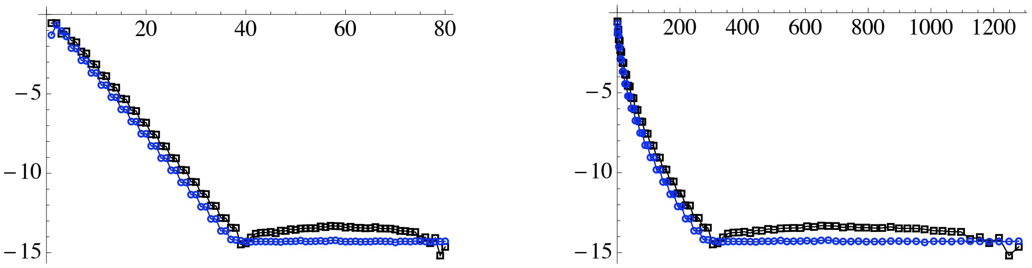

Figure 5: Error in approximating $f(x) = \frac{1}{1+x^2}$ by $f_{n,m}(x)$ for $n = 1,\ldots,80$. Left: log error $\log_{10}\|f - f_{n,m}\|_\infty$ (squares) and $\log_{10}\|f - f_{n,m}\|$ (circles) against n. Right: log error against $m = 0.2n^2$.

Next we consider reconstructions in other polynomials bases. In Table 1 we give the error in approximating the function $f(x) = \mathrm{e}^{-x}\cos 4x$ with Chebyshev polynomials of the first and second kinds. Note that the resulting uniform error is virtually identical to the case of the Legendre polynomial implementation. Since all three implementations compute exactly the same approximation $f_{n,m}$, up to numerical error, this comes as no surprise. Moreover, as evidenced by Table 2, the payoff is only mild growth in the condition number $\kappa(A)$.

3.4 Connections to earlier work

Rather than choosing m such that $C_{n,m} \geq \theta$, it may appear advantageous to find the minimum m such that $C_{n,m} > 0$. In other words, the smallest m such that $f_{n,m}$ is guaranteed to exist. Letting $\theta = 0$ in Theorems 3.2 and 3.3, we immediately obtain a sufficient condition of the form $m \geq cn^2$, for some $c > 0$. However, this result is far too pessimistic: it is known that reconstruction is always possible, provided $m \geq n$ [47]. For this reason, it may appear favourable to reconstruct using $m = n$. This results in a technique known as the *inverse polynomial reconstruction method* [52, 53]. Unfortunately, this approach is extremely unstable. The linear system has geometrically large condition number, making the procedure extremely sensitive to both noise and round-off error. Moreover, a continuous analogue of the Runge phenomenon occurs. In general, the approximation $f_{n,m}$ only converges to f if the geometric decay of $\|f - \mathcal{Q}_n f\|$ is faster than the geometric growth of $\|A^{-1}\|$, meaning that only functions analytic in sufficiently large complex regions can be approximated by this procedure (as discussed in detail in [4],

n	5	10	15	20	25	30	35	40
(a)	1.45e0	1.85e-3	3.03e-7	2.53e-12	1.06e-14	8.42e-14	4.06e-14	5.31e-14
(b)	1.45e0	1.85e-3	3.03e-7	2.53e-12	3.51e-14	1.16e-13	4.57e-14	7.70e-14
(c)	1.45e0	1.85e-3	3.03e-7	2.49e-12	6.76e-14	7.33e-14	6.40e-14	5.15e-14

Table 1: Comparison of the error $\|f - f_{n,m}\|_\infty$ with $m = 0.2n^2$, where $f_{n,m}$ is formed from (a) Legendre polynomials and Chebyshev polynomials of the (b) first and (c) second kinds.

n	5	10	15	20	25	30	35	40
(a)	3.57	5.55	4.21	5.20	4.40	5.06	4.50	6.77
(b)	13.74	49.99	52.63	91.89	92.89	133.02	133.49	191.19
(c)	3.90	5.67	7.25	9.33	11.91	13.96	16.56	18.92

Table 2: Comparison of the condition number $\kappa(A)$ with $m = 0.2n^2$, where A is formed from (a) Legendre polynomials and Chebyshev polynomials of the (b) first and (c) second kinds.

this behaviour can be understood in terms of the operator-theoretic properties of finite sections of certain non-Hermitian infinite matrices). On the other hand, by allowing m to range independently of n, we overcome all these difficulties, and obtain a stable method whose convergence is completely determined by the convergence of $\mathcal{Q}_n f$ to f.

The specific instance of Legendre polynomial reconstructions from Fourier samples using $m > n$ has also been considered in [47]. Therein, the estimate $m = \mathcal{O}\left(n^2\right)$ was derived, along with bounds for the error. Naturally, this problem is just one specific example of our general framework. However, within this context, our work improves and extends the results of [47] in the following ways:

1. Reconstruction is completely independent of the particular polynomial basis used. In particular, the estimates for $\Theta(n;\theta)$ and $\|f - f_{n,m}\|$ are determined only by the space T_n and the vectors $\{\psi_j\}_{j\in\mathbb{Z}}$. This allows for analysis of reconstructions in arbitrary polynomial bases, not just the Legendre polynomials used in [47].
2. The estimates for $\Theta(n;\theta)$ in Theorems 3.2 and 3.3 improve those given in [47]. In particular, it was shown in [47, Theorem 4.2] that

   $$C_{n,\alpha n^2} \geq 1 - \frac{8}{\pi} \arcsin \frac{1}{\pi\alpha}, \quad \forall n \in \mathbb{N}, \quad \alpha \geq 1, \tag{3.12}$$

   (our constant $C_{n,m}$ corresponds to the quantity $\sigma^2_{n,m}$ in [47]). Conversely, Lemma 3.1 leads to the improved bounds

   $$C_{n,\alpha n^2} \geq 1 - \frac{4(\pi-2)}{\pi^2(\alpha - n^{-2})} \quad \forall n \geq \max\left\{ \frac{2}{\pi\alpha}, \sqrt{\frac{2}{\alpha}} \right\}, \quad \alpha > 0, \tag{3.13}$$

   and

   $$C_{n,\alpha n^2} \geq 1 - \frac{4}{\pi^2\alpha} + \mathcal{O}\left(n^{-2}\right), \quad n \to \infty, \quad \alpha > 0. \tag{3.14}$$

   Not only are these bounds sharper, they also hold for a greater range of α, thus permitting reconstruction with $m = \alpha n^2$ for any $\alpha > 0$, as opposed to just $\alpha \geq 1$. This leads to savings in computational cost, and, in cases where m is fixed, allows larger values of n to be used, thereby increasing accuracy. To illustrate this improvement, note that (3.12) gives the estimate $C_{n,m} \geq 0.175$ when $m = n^2$. Conversely, our estimate (3.13) yields the improved bound 0.383 for $n \geq 2$, and (3.14) gives the asymptotic bound 0.595. To compare, direct computation of C_{n,n^2} indicates that $C_{n,n^2} \geq 0.68$ for all n, and $C_{n,n^2} \to 0.8$ as $n \to \infty$.
3. Piecewise analytic functions and function of arbitrary numbers of variables can be recovered in a analogous fashion, with similar analysis (see Sections 3.5 and 4 respectively).

3.5 Reconstruction of piecewise analytic functions

Naturally, whenever the approximated function is not analytic, the convergence rate of the polynomial approximant $f_{n,m}$ to f is not exponential. For example, consider the function

$$f(x) = \begin{cases} (2\mathrm{e}^{2\pi(x+1)} - 1 - \mathrm{e}^{\pi})(\mathrm{e}^{\pi} - 1)^{-1} & x \in [-1, -\frac{1}{2}) \\ -\sin(\frac{2\pi x}{3} + \frac{\pi}{3}) & x \in [-\frac{1}{2}, 1] \end{cases} \tag{3.15}$$

This function was put forth in [62] to test algorithms for overcoming the Gibbs phenomenon. Aside from the discontinuity, its sharp peak makes it a challenging function to reconstruct accurately. Since this

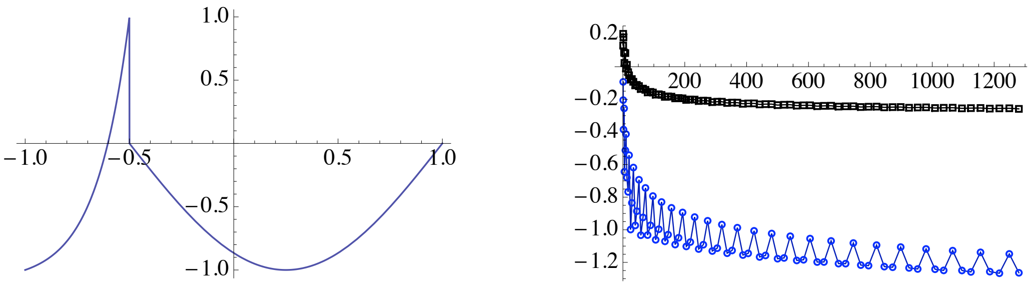

Figure 6: Error in approximating the function (3.15) by $f_{n,m}(x)$ for $n = 1, \ldots, 80$. Left: the function $f(x)$. Right: log error $\log_{10} \|f - f_{n,m}\|_\infty$ (squares) and $\log_{10} \|f - f_{n,m}\|$ (circles) against $m = 0.2n^2$.

function is discontinuous, we expect only low-order, algebraic convergence of $f_{n,m}$ in the L^2 norm, but no uniform convergence, an observation confirmed in Figure 6.

However, by reconstructing this function in a polynomial basis, we are not exploiting the known information about f: namely, the jump discontinuity at $x = -\frac{1}{2}$. The general procedure set out in Section 2 allows us to use such information in designing a reconstruction basis. Naturally, since f is analytic in the subintervals $[-1, -\frac{1}{2}]$ and $[-\frac{1}{2}, 1]$, a better choice is to reconstruct f in a piecewise polynomial basis. The aim of this section is to describe this procedure.

Seeking generality, suppose that $f : [-1, 1] \to \mathbb{R}$ is piecewise analytic with jump discontinuities at $-1 < x_1 < \ldots < x_l < 1$. Let $x_0 = -1$ and $x_{l+1} = 1$. We assume that f has been sampled via $\hat{f}_j = \langle f, \psi_j \rangle$, $j = 1, \ldots, m$, where $\langle \cdot, \cdot \rangle$ is the Euclidean inner product on $\mathrm{L}^2(-1, 1)$. In examples, these will be the Fourier samples of f, but the construction described below holds for arbitrary sampling bases consisting of functions defined on $[-1, 1]$.

Throughout we shall assume that the discontinuity locations $x_1, \ldots, x_l$ are known exactly. That is, we focus on *reconstruction*. Naturally, a fully-automated algorithm must also incorporate a scheme for singularity *detection*. There are numerous methods for this problem. We refer the reader to [34, 61] for further details.

Given the additional information about the location of the singularities of f, we now design a reconstruction basis to mirror this feature. We shall construct such a basis via local co-ordinate mappings. To this end, let $I_r = [x_r, x_{r+1}]$, $c_r = \frac{1}{2}(x_{r+1} - x_r)$ and define $\Lambda_r(x) = \frac{x - x_r}{c_r} - 1$, so that $\Lambda(I_r) = [-1, 1]$. Suppose now that T'_n is a space of functions defined on $[-1, 1]$ (e.g. the polynomial space $\mathbb{P}_{n-1}$). By convention, we assume that each $\phi \in \mathrm{T}'_n$ is extended by zero to the whole real line, i.e. $\phi(x) = 0$ for $x \in \mathbb{R} \backslash [-1, 1]$. Let $\mathrm{T}_{n,r}$ be the space of functions defined on I_r, given by $\mathrm{T}_{n,r} = \{\phi \circ \Lambda_r : \phi \in \mathrm{T}'_n\}$. We now define the new reconstruction space in the obvious manner:

$$\mathrm{T}_n = \{\phi : \phi|_{I_r} \in \mathrm{T}_{n_r,r},\ r = 0, \ldots, l\}, \quad n = \sum_{r=0}^{l} n_r,$$

and seek an approximation $f_{n,m} \in \mathrm{T}_n$ to f defined by (2.6). Suppose now that $\{\phi_1, \ldots, \phi_n\}$ is a collection of linearly independent reconstruction functions with $\mathrm{T}'_n = \operatorname{span}\{\phi_1, \ldots, \phi_n\}$. We construct a basis for T_n by scaling. To this end, we let $\phi_{r,j} = \frac{1}{\sqrt{c_r}} \phi_j \circ \Lambda_r$, and notice that $\mathrm{T}_n = \operatorname{span}\{\phi_{r,j} : j = 1, \ldots, n_r,\ r = 0, \ldots, l\}$. Note that, if $\{\phi_j\}$ are orthonormal, then so are $\{\phi_{r,j}\}$. With this basis in hand, the approximation $f_{n,m}$ is now given by

$$f_{n,m} = \sum_{r=0}^{l} \sum_{j=1}^{n_r} \alpha_{r,j} \phi_{r,j},$$

where the coefficients $\alpha_{r,j}$ are determined by the aforementioned equations. As before, this is equivalent to the least squares problem $U\alpha \approx \hat{f}$ with block matrix $U = [U_1, \ldots, U_l]$, where U_r is the $m \times n_r$ matrix with $(j, k)^{\text{th}}$ entry

$$\langle \phi_{r,k}, \psi_j \rangle = \frac{1}{\sqrt{c_r}} \int_{x_r}^{x_{r+1}} \phi_k(\Lambda_r(x)) \psi_j(x) \, \mathrm{d}x.$$

Here $\hat{f} = (\hat{f}_1, \ldots, \hat{f}_m)^\top$, $\alpha = [\alpha_0, \ldots, \alpha_l]$ and $\alpha_r = (\alpha_{r,1}, \ldots, \alpha_{r,n_r})^\top$.

Naturally, estimation of the quantity $\Theta(n;\theta)$ is vital. The following lemma aids in this task:

Lemma 3.8. *The constant $C_{n,m}$ satisfies*

$$C_{n,m} \geq 1 - \sum_{r=0}^{l} \left(1 - C_{n_r,m,r}\right),$$

where $C_{n_r,m,r} = \inf_{\substack{\phi \in \mathrm{T}_{n_r,r} \\ \|\phi\|=1}} \langle \mathcal{P}_m \phi, \phi \rangle$.

Proof. For $\phi \in \mathrm{T}_n$, denote $\phi|_{I_r}$ by $\phi^{[r]}$. Assume that $\phi^{[r]}$ is extended to $[-1,1]$ by zero, so that $\phi = \sum_{r=}^{l} \phi^{[r]}$. Since $\phi^{[r]} \perp \phi^{[s]}$ for $r \neq s$, it follows that

$$1 - C_{n,m} = \sup \left\{ \frac{\sum_{r=0}^{l} \left\langle \phi^{[r]} - \mathcal{P}_m \phi^{[r]}, \phi^{[r]} \right\rangle}{\sum_{r=0}^{l} \|\phi^{[r]}\|^2} : \phi^{[r]} \in \mathrm{T}_{n_r,r},\ r = 0, \ldots, l,\ \sum_{r=0}^{l} \|\phi^{[r]}\|^2 \neq 0 \right\}.$$

Note that, for $a_r \geq 0$ and $b_r > 0$, $r = 0, \ldots, l$, the inequality

$$\sum_{r=0}^{l} a_r \leq \sum_{r=0}^{l} \frac{a_r}{b_r} \sum_{r=0}^{l} b_r,$$

holds. Setting $a_r = \left\langle \phi^{[r]} - \mathcal{P}_m \phi^{[r]}, \phi^{[r]} \right\rangle$ and $b_r = \|\phi^{[r]}\|^2$ and using this inequality gives

$$\begin{aligned} 1 - C_{n,m} &\leq \sup \left\{ \sum_{r=0}^{l} \frac{\left\langle \phi^{[r]} - \mathcal{P}_m \phi^{[r]}, \phi^{[r]} \right\rangle}{\|\phi^{[r]}\|^2} : \phi^{[r]} \in \mathrm{T}_{n_r,r},\ r = 0, \ldots, l,\ \sum_{r=0}^{l} \|\phi^{[r]}\|^2 \neq 0 \right\} \\ &\leq \sum_{r=0}^{l} \sup \left\{ \frac{\langle \phi - \mathcal{P}_m \phi, \phi \rangle}{\|\phi\|^2} : \phi \in \mathrm{T}_{n_r,r},\ \|\phi\| \neq 0 \right\}, \end{aligned}$$

and this is precisely $\sum_{r=0}^{l} (1 - C_{n_r,m,r})$. □

Let us now focus on piecewise polynomial reconstructions from Fourier samples, in which case

$$\mathrm{T}_n = \{\phi : \phi|_{I_r} \in \mathbb{P}_{n_r - 1},\ r = 0, \ldots, l\} \tag{3.16}$$

is the space of piecewise polynomials of total degree n. Regarding the rate of convergence of the resulting approximation $f_{n,m}$, it is a simple exercise to confirm that

$$\|f - f_{n,m}\| \leq c(\theta) c_f \sum_{r=0}^{l} \sqrt{n_r} \rho_r^{-n_r},$$

where $c(\theta)$ is defined in (2.22), c_f is a constant depending on f only and ρ_r is determined by the largest Bernstein ellipse (appropriately scaled) within which the function $f|_{I_r}$ is analytic. Hence, we expect exponential convergence of $f_{n,m}$ to f. The main question remaining is that of estimating the function $\Theta(n;\theta)$ for this reconstruction procedure. For this, we have the following result, which extends Theorems 3.2 and 3.3 to this more general case:

Theorem 3.9. *Let $\{\psi_j\}_{j \in \mathbb{Z}}$ be the Fourier basis and T_n be given by (3.16). Then the function $\Theta(n;\theta)$ satisfies*

$$\Theta(n;\theta) \leq 2 \left[\frac{1}{2} + \frac{2(\pi - 2)}{\pi^2 (1 - \theta)} \sum_{r=0}^{l} \frac{n_r^2}{c_r} \right], \quad \forall n = \sum_{r=0}^{l} n_r,\ n_0, \ldots, n_l \in \mathbb{N},$$

and

$$\Theta(n;\theta) \leq \frac{4}{\pi^2 (1 - \theta)} \sum_{0}^{l} \frac{n_r^2}{c_r} + \mathcal{O}(1), \quad n_0, \ldots, n_l \to \infty.$$

m	10	20	40	80	160	320	640	1280
$C_{n,m}$	0.05	0.34	0.33	0.44	0.44	0.47	0.49	0.50
$\kappa(A)$	19.88	2.92	3.06	2.27	2.27	2.11	2.03	1.98

Table 3: The quanties $C_{n,m}$ and $\kappa(A)$ against m, where n is as in Figure 7.

Proof. In view of Lemma 3.8, it suffices to consider $C_{n_r,m_r,r}$. To this end, let $J = [\alpha, \beta] \subseteq [-1, 1]$, $\mathrm{T}_{n,J}$ be the space of functions ϕ with supp$(\phi) \subseteq J$ and $\phi|_J \in \mathbb{P}_{n-1}$, and define

$$C^J_{n,m} = \inf_{\substack{\phi \in \mathrm{T}_{n,J} \\ \|\phi\|=1}} \langle \mathcal{P}_m \phi, \phi \rangle .$$

Let $\Lambda(x) = \frac{x-\alpha}{c} - 1$, where $c = \frac{1}{2}(\beta - \alpha)$, and write $\phi = \Phi \circ \Lambda$, where supp$(\Phi) \subseteq [-1, 1]$. Consider the quantity $\langle \phi, \psi_j \rangle$. By definition of ψ_j, we have

$$\langle \phi, \psi_j \rangle = \frac{1}{\sqrt{2}} \int_{-1}^{1} \phi(x) \mathrm{e}^{-\mathrm{i} j \pi x} \,\mathrm{d}x = \frac{c}{\sqrt{2}} \mathrm{e}^{-\mathrm{i} j \pi (\alpha + c)} \int_{\Lambda(-1)}^{\Lambda(1)} \Phi(y) \mathrm{e}^{-\mathrm{i} j \pi c y} \,\mathrm{d}y.$$

Let $K = [\Lambda(-1), \Lambda(1)] = \Lambda([-1, 1]) \supseteq [-1, 1]$ and let $\mathcal{P}_{m,K}$ be the Fourier projection operator based on the interval K. It now follows that

$$C^J_{n,m} = \inf \left\{ \langle \mathcal{P}_{m,K} \Phi, \Phi \rangle : \text{supp}(\Phi) \subseteq [-1, 1],\ \Phi|_{[-1,1]} \in \mathbb{P}_{n-1},\ \|\Phi\| = 1 \right\} .$$

This is now precisely the setup of Remark 3.1. Using (3.6), we therefore deduce that

$$1 - C^J_{n,m} \leq \frac{4(\pi - 2) n^2}{c \pi^2 (2 \lfloor \frac{m}{2} \rfloor - 1)}.$$

Letting $J = I_r$, $c = c_r$ and using Lemma 3.8, we now obtain

$$C_{n,m} \geq 1 - \frac{4(\pi - 2)}{\pi^2 (2 \lfloor \frac{m}{2} \rfloor - 1)} \sum_{r=0}^{l} \frac{n_r^2}{c_r}, \tag{3.17}$$

from which the result follows immediately. □

To implement this scheme, it is necessary to compute the values (3.5). By changing variables, it is easily seen that

$$\langle \phi_{r,k}, \psi_j \rangle = \sqrt{\frac{c_r}{2}} \mathrm{e}^{-\mathrm{i} j \pi d_r} \int_{-1}^{1} \phi_k(y) \mathrm{e}^{-\mathrm{i} j \pi c_r y} \,\mathrm{d}y,$$

where $d_r = \frac{1}{2}(x_{r+1} + x_r)$. Since (3.4) holds for all $z \in \mathbb{C}$, it follows that

$$\langle \phi_{r,k}, \psi_j \rangle = \mathrm{e}^{-\mathrm{i} j \pi d_r} (-\mathrm{i})^k \sqrt{\frac{k + \frac{1}{2}}{j}} J_{k+\frac{1}{2}}(j \pi c_r), \tag{3.18}$$

whenever the functions $\phi_{r,k}$ arise from scaled Legendre polynomials. Naturally, if the functions $\phi_{r,k}$ arise from arbitrary scaled Gegenbauer polynomials, computation of the values (3.5) can be carried out recursively via the algorithm described in Section 3.2.

In Figure 7 we apply this method to the function (3.15) using the orthonormal basis of scaled Legendre polynomials. The improvement over Figure 6 is dramatic: using only $m \approx 250$ (with $n_0 = n_1 = 16$) we obtain 13 digits of accuracy. Note that, as expected, root exponential convergence occurs. Moreover, as predicted by (3.17), and illustrated in Table 3, the condition number of the matrix A remains small.

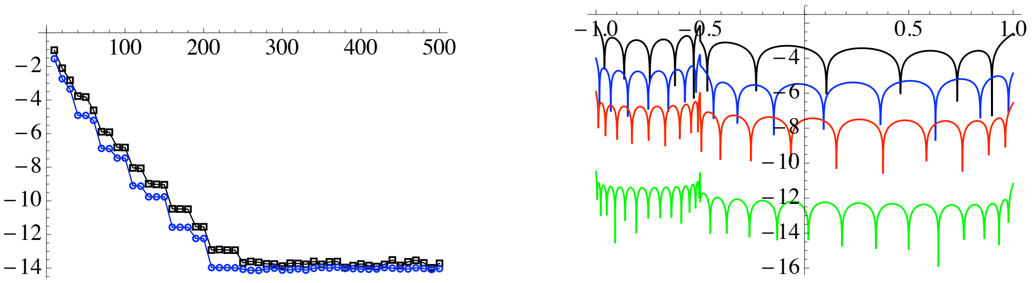

Figure 7: Error in approximating the function (3.15) by $f_{n,m}(x)$. Left: log error $\log_{10}\|f - f_{n,m}\|_\infty$ (squares) and $\log_{10}\|f - f_{n,m}\|$ (circles) against $m = 1,\ldots,500$ with $n_0 = n_1$ chosen so that $m = \frac{1}{5}\left(\frac{n_0^2}{c_0} + \frac{n_1^2}{c_1}\right)$. Right: the error $\log_{10}|f(x) - f_{n,m}(x)|$ against $x \in [-1,1]$ for $m = 20, 40, 80, 160$.

3.6 Comparison to existing methods

Numerous methods exist for recovering functions to high accuracy from their Fourier data. Applications are myriad, and range from medical imaging [8, 9] to postprocessing of numerical solutions of hyperbolic PDEs [38, 51]. Prominent examples which deliver high global accuracy (in contrast to standard filtering techniques, which only yield high accuracy away from the singularities of a function [41]) include spectral reprojection [40, 41, 42], techniques based on implicit matching of jump conditions [27], Padé methods [24], methods based on sequence extrapolation [18] and Fourier extension methods [13, 48] (for a more complete list, see [15] and references therein).

Whilst many of these methods deliver exponential convergence in terms of m (the number of Fourier coefficients), they all suffer from ill-conditioning. This comes as no surprise: the problem of reconstructing a function from its Fourier coefficients can be viewed as a continuous analogue of the recovery of a function from m equidistant pointwise values. As proved in [57], any method for this problem that converges exponentially fast in m will suffer from exponentially poor conditioning. We conjecture that a similar result also holds in the continuous case.

Aside from increased susceptibility to round-off error and noise, ill-conditioning often makes so-called inverse methods (e.g. extrapolation and Fourier extension methods) costly to implement. Conversely, the method proposed in the paper does not suffer from any ill-conditioning. This negative consequence of [57] is circumvented, precisely because we witness only root exponential convergence in m. However, an advantage of this approach is that it delivers exponential convergence in n, the degree of the final approximant $f_{n,m}$. In many applications it may be necessary to manipulate $f_{n,m}$, its relatively low degree making such operations reasonably cheap. Thus, this method has the advantage of compression, a feature not shared by the majority of the other methods mentioned previously.

A well-established and widely used alternative to this method is spectral reprojection, developed by Gottlieb et al [35, 41, 42]. Much like this approach, it computes a polynomial approximation. Yet it stands out as being direct, meaning that no linear system or least squares problem is required to be solved. Whilst the original method, based on Gegenbauer polynomials [41, 42] has been shown to suffer from a generalised Runge phenomenon [14], thereby severely affecting its applicability, an improved approach based on Freud polynomials was recently developed in [35]. Numerically at least, this method appears to overcome a number of the issues associated with the original Gegenbauer procedure.

Comparatively speaking, spectral reprojection delivers exponential convergence in $\mathcal{O}\left(m^2\right)$ operations. On the other hand, our method obtains root exponential convergence at a cost of $\mathcal{O}(m^{\frac{3}{2}})$ operations. However, despite being theoretically more efficient, the various constants involved in spectral reprojection tend to be rather large. Indeed, in Table 4 we compare the error in approximating the function (3.15) using both procedures (the data for spectral reprojection is taken from [35, Table 1]. Note that the parameter N used therein is such that $2N = m$ is the total number of Fourier coefficients). As is evident, the method proposed in this paper obtains an error of order 10^{-14} using less than 256 Fourier coefficients, whereas spectral reprojection does not reach this value until more than 1024 coefficients are used.

The most likely reason for this improvement is that the method of this paper is quasi-optimal, thereby

m	64	128	256	512	1024	2048	4096
(a)	8.90e-01	1.37e-01	1.84e-04	1.01e-07	9.33e-13	5.27e-13	5.23e-14
(b)	2.40e-04	8.36e-09	2.40e-14	1.38e-14	1.74e-14	2.26e-14	2.59e-14

Table 4: Comparison of the (a) spectral reprojection and (b) generalised reconstruction methods applied to (3.15). Here m is the total number of Fourier samples.

delivering a near-optimal polynomial approximation, whereas both the spectral reprojection does not possess this feature. In fact, although spectral reprojection formally converges exponentially fast in m, the corresponding rate may be substantially slower than that of the best polynomial approximation. Furthermore, for the Gegenbauer technique at least, there is the significant issue that various parameters have to be chosen in an essentially function-dependent manner to ensure convergence [35], and thereby avoid a Runge phenomenon. Numerical stability is also potentially a significant issue in both cases.

Aside from improved numerical performance, let us mention several other benefits. First, as discussed, the final approximation consists of only $\mathcal{O}(\sqrt{m})$ terms, as opposed to $\mathcal{O}(m)$. Second, the basis for the polynomial reconstruction space T_n can be chosen arbitrarily (in particular, independently of m) without affecting the convergence. The only downside is a mild increase in condition number if nonorthogonal polynomials are employed. In contrast, for the Freud/Gegenbauer spectral reprojection methods, only very specific types of polynomials can be used (which may not be simple to construct or manipulate [35]), and, whenever the number of samples m is varied, all polynomials employed for reconstruction must be changed.

One advantage of spectral reprojection is that it is local: the approximation in each subdomain of smoothness is computed separately and independently of any other subdomain. Conversely, with our approach, the computations are inherently coupled. Nevertheless, it may be possible to devise a local version of our approach, a question we intend to explore in future investigations.

4 Reconstructions in tensor-product spaces

Thus far, we have focused on the reconstruction of univariate functions from their Fourier samples. A simple extension of this approach, via tensor products, is to functions defined in cubes. The aim of this section is to detail this generalisation.

To formalise this idea, let us return to the general perspective of Section 2. Suppose that the Hilbert space H can be expressed as a tensor-product $\mathrm{H} = \mathrm{H}_1 \otimes \cdots \otimes \mathrm{H}_d$ of Hilbert spaces H_i, $i = 1, \ldots, d$, each having inner product $\langle \cdot, \cdot \rangle_i$. Note that, for $f = f_1 \otimes \cdots \otimes f_d \in \mathrm{H}$ and $g = g_1 \otimes \cdots \otimes g_d \in \mathrm{H}$, we have

$$\langle f, g \rangle = \prod_{i=1}^{d} \langle f_i, g_i \rangle_i .$$

Now suppose that the sampling basis consists of tensor-product functions. To this end, let

$$\psi_j = \psi_{1,j_1} \otimes \cdots \otimes \psi_{d,j_d}, \quad j = (j_1, \ldots, j_d) \in \mathbb{N}^d,$$

and, for $m = (m_1, \ldots, m_d) \in \mathbb{N}^d$, set

$$\mathcal{S}_m = \operatorname{span}\left\{ \psi_j : j = (j_1, \ldots, j_d),\ 1 \leq j_i \leq m_i,\ i = 1, \ldots, d \right\}.$$

We assume throughout that the collection $\{\psi_{i,j}\}_{j=1}^{\infty}$ is a Riesz basis for H_i for $i = 1, \ldots, d$. In particular, $\{\psi_j\}$ is a Riesz basis for H. With this to hand, we define the operator $\mathcal{P}_m : \mathrm{H} \to \mathcal{S}_m$ by

$$\mathcal{P}_m f = \sum_{j_1=1}^{m_1} \cdots \sum_{j_d=1}^{m_d} \langle f, \psi_j \rangle \psi_j .$$

Note that $\mathcal{P}_m(f_1 \otimes \cdots \otimes f_d) = (\mathcal{P}_{1,m_1} f_1) \otimes \cdots \otimes (\mathcal{P}_{d,m_d} f_d)$, where $\mathcal{P}_{i,m_i} : \mathrm{H}_i \to \mathcal{S}_{i,m_i} = \operatorname{span}\{\psi_{i,1}, \ldots, \psi_{i,m_i}\}$ is defined in the obvious manner. In a similar fashion, we introduce the reconstruction vectors $\phi_j = \phi_{1,j_1} \otimes \cdots \otimes \phi_{d,j_d}$, which form a basis for the reconstruction space

$$\mathrm{T}_n = \operatorname{span}\left\{ \phi_j : j = (j_1, \ldots, j_d),\ 1 \leq j_i \leq n_i,\ i = 1, \ldots, d \right\}, \quad n = (n_1, \ldots, n_d) \in \mathbb{N}^d.$$

Note that $\mathrm{T}_n = \mathrm{T}_{1,n_1} \otimes \cdots \otimes \mathrm{T}_{d,n_d}$, where $\mathrm{T}_{i,n_i} = \mathrm{span}\{\phi_{i,1}, \ldots, \phi_{i,n_i}\}$. As before, we construct the approximation $f_{n,m} \in \mathrm{T}_n$ via (2.6).

To cast this problem in a form suitable for computations, let $U^{[i]} \in \mathbb{C}^{m_i \times n_i}$ be the matrix with $(j,k)^{\mathrm{th}}$ entry $\langle \psi_{i,j}, \phi_{i,k} \rangle_i$. Let $U \in \mathbb{C}^{\bar{m},\bar{n}}$ be the matrix of the d-variate reconstruction method, where $\bar{m} = m_1 \ldots m_d$ and $\bar{n} = n_1 \ldots n_d$. It is easily shown that

$$U = \bigotimes_{i=1}^{d} U^{[i]}, \quad A = \bigotimes_{i=1}^{d} A^{[i]},$$

where $A = U^\dagger U$, and $A^{[i]} = (U^{[i]})^\dagger U^{[i]}$, and, in this case, $B_1 \otimes B_2$ denotes the Kronecker product of matrices the B_1 and B_2. By a trivial argument, we conclude that the number of operations required to compute $f_{n,m}$ is of order $(n_1 m_1) \ldots (n_d m_d) \sqrt{\kappa(A)}$.

Recall that the spectrum of the Kronecker product matrix $B_1 \otimes B_2$ consists of the pairs $\lambda\mu$, where λ is an eigenvalue of B_1 and μ is an eigenvalue of B_2. From this, we deduce that

$$\kappa(A) = \prod_{i=1}^{d} \kappa(A^{[i]}).$$

Hence $\kappa(A)$ is completely determined by the matrices $A^{[i]}$, with the i^{th} such matrix corresponding to the univariate reconstruction problem with sampling basis $\{\psi_{i,j}\}_{j=1}^{m_i}$ and reconstruction basis $\{\phi_{i,j}\}_{j=1}^{n_i}$. Unsurprisingly, a similar observation also holds for the quantity $C_{n,m}$:

Lemma 4.1. *Let*

$$C_{n,m} = \inf_{\substack{\phi \in \mathrm{T}_n \\ \|\phi\|=1}} \langle \mathcal{P}_m \phi, \phi \rangle, \qquad C_{n_i,m_i} = \inf_{\substack{\phi \in \mathrm{T}_{i,n_i} \\ \|\phi\|_i=1}} \langle \mathcal{P}_{i,m_i} \phi, \phi \rangle_i, \quad i = 1, \ldots, d. \tag{4.1}$$

Then $C_{n,m} = \prod_{i=1}^{d} C_{i,n_i,m_i}$.

Proof. By Lemma 2.13, $C_{n,m} = \lambda_{\min}(\tilde{A}^{-1}A)$ and $C_{i,n_i,m_i} = \lambda_{\min}((\tilde{A}^{[i]})^{-1}A^{[i]})$, $i = 1, \ldots, d$, where $\tilde{A}$ and $\tilde{A}^{[i]}$ are defined in the obvious manner. Since $\tilde{A} = \tilde{A}^{[1]} \otimes \cdots \otimes \tilde{A}^{[d]}$, the matrix $\tilde{A}^{-1}$ is the Kronecker product of matrices $(\tilde{A}^{[i]})^{-1}$. The result now follows immediately. □

4.1 Reconstruction of piecewise smooth functions

Having presented the general case, we now turn our attention to the reconstruction of a piecewise smooth function $f : [-1,1]^d \to \mathbb{R}$. We shall make the significant assumption (see Remark 4.1) that f is smooth in hyper-rectangular subregions of $[-1,1]^d$. To this end, for $i = 1, \ldots, d$ let $l_i \in \mathbb{N}$ and suppose that

$$-1 = x_{0,i} < x_{1,i} < \ldots < x_{l_i,i} < x_{l_i+1,i} = 1.$$

Define $I_{r,i} = [x_{r,i}, x_{r+1,i}]$, $r = 0, \ldots, l_i$, and for $r = (r_1, \ldots, r_d)$ write $I_r = I_{r_1,1} \times \cdots \times I_{r_d,d}$, so that the collection

$$\{I_r : r = (r_1, \ldots, r_d),\ r_i = 0, \ldots, l_i,\ i = 1, \ldots, d\},$$

consists of disjoint sets whose union is $[-1,1]^d$. We assume that f is smooth within each subdomain I_r. In addition, for $r_i = 0, \ldots, l_i$ and $i = 1, \ldots, d$, let $c_{r_i,i} = \frac{1}{2}(x_{r_i+1,i} - x_{r_i,i})$ and set $\Lambda_{r_i,i}(x) = \frac{x - x_{r_i,i}}{c_{r_i,i}} - 1$, $x \in I_{r_i}$. Note that $\Lambda_{r_i,i}(I_{r_i,i}) = [-1,1]$.

We now design a reconstruction space. To this end, for $n \in \mathbb{N}$ let T'_n, $\dim \mathrm{T}'_n = n$, be a space of functions $\phi : \mathbb{R} \to \mathbb{C}$ with $\mathrm{supp}(\phi) \subseteq [-1,1]$. Define

$$T_{n,r,i} = \{\phi \circ \Lambda_{r,i} :\ \phi \in \mathrm{T}'_n\}, \quad n \in \mathbb{N}.$$

Now suppose that n is the vector $(n_1, \ldots, n_d)$, where

$$n_i = \sum_{r=0}^{l_i} n_{r,i}, \quad i = 1, \ldots, d,$$

for some $n_{r,i} \in \mathbb{N}$. We define the reconstruction space T_n by

$$\mathrm{T}_n = \bigotimes_{i=1}^{d} \bigoplus_{r=0}^{l_i} T_{n_{r,i},r,i}. \tag{4.2}$$

We require a basis for this space. Let $\{\phi_1, \ldots, \phi_n\}$, $n \in \mathbb{N}$, be a basis for T'_n, and set

$$\phi_{r,j,i} = \frac{1}{\sqrt{c_{r,i}}} \phi_j \circ \Lambda_{r,i}.$$

A basis for T_n is now given by

$$\left\{ \phi_{r_1,j_1,1} \otimes \cdots \otimes \phi_{r_d,j_d,d}, \ j = 1, \ldots, n_{r_i,i}, \ r_i = 0, \ldots, l_i, \ i = 1, \ldots, d \right\}.$$

This framework gives a general means in which to construct reconstruction bases in the tensor-product case for functions which are piecewise smooth with discontinuities parallel to the co-ordinate axes. Suppose now that we consider the recovery of such a function from its Fourier samples. Using the above framework, we construct a basis consisting of piecewise polynomials of several variables. The main question remaining is that of estimating the function $C_{n,m}$. However, in view of the Lemma 4.1 and the results derived in Section 3.1, a simple argument gives

Theorem 4.2. *Let $\{\psi_j\}_{j \in \mathbb{Z}}$ be the multivariate Fourier basis on $[-1,1]^d$ and suppose that T_n is defined by (4.2) for the choice $\mathrm{T}'_n = \{\phi : \phi|_{[-1,1]} \in \mathbb{P}_{n-1}\}$. Suppose further that $n = (n_1, \ldots, n_d)$, where $n_i = \sum_{r=0}^{l_i} n_{r,i}$, $i = 1, \ldots, d$, and let*

$$\Theta(n;\theta) = \min\{m = (m_1, \ldots, m_d) : C_{n,m} \geq \theta\}, \quad \theta \in (0,1),$$

where $C_{n,m}$ is as in (4.1). If $\theta_1, \ldots, \theta_d \in (0,1)$ satisfy $\theta = \theta_1 \ldots \theta_d$, then we may write

$$\Theta(n;\theta) = \left(\Theta_1(n_1;\theta_1), \ldots, \Theta_d(n_d;\theta_d)\right),$$

where

$$\Theta_i(n_i;\theta_i) \leq 2 \left\lceil \frac{1}{2} + \frac{2(\pi-2)}{\pi^2(1-\theta_i)} \sum_{r=0}^{l_i} \frac{n_{r,i}^2}{c_{r,i}} \right\rceil, \quad i = 1, \ldots, d,$$

and

$$\Theta_i(n_i;\theta_i) \leq \frac{4}{\pi^2(1-\theta_i)} \sum_{r=0}^{l_i} \frac{n_{r,i}^2}{c_{r,i}} + \mathcal{O}(1), \quad n_{0,i}, \ldots n_{l_i,i} \to \infty, \qquad i = 1, \ldots, d.$$

The main consequence of this theorem is the following: regardless of the dimension, the variables $m_1, \ldots, m_d$ must scale quadratically with $n_1, \ldots, n_d$ to ensure quasi-optimal recovery in a multivariate piecewise polynomial basis from Fourier samples. Consider now the most simple example of this approach: namely, where f is smooth in $[-1,1]^d$, so that T_n consists of multivariate polynomials. In Figure 8 we plot the error in approximating the functions $f(x,y) = \mathrm{e}^{x^2 y}$ and $f(x,y) = \sin 3xy$, using parameters $m_1 = m_2 = 0.5 n_1^2$ and $n_2 = n_1$. As in the univariate case, we observe the accuracy of this technique. For example, using only $m_1 = m_2 \approx 200$ and $n_1 = n_2 \approx 20$ we obtain an error of order 10^{-14}.

Remark 4.1 This approach (and many others based on tensor-product formulations) has the significant shortcoming that it requires the function to be singular in regions parallel to the co-ordinate axes. Naturally, this is a rather restrictive condition. For a function with singularities lying on a curve (in two dimensions, for example), one potential alternative is to apply the one-dimensional method along horizontal and vertical slices, and recover the two-dimensional function from the resulting one-dimensional reconstructions. However, the generality of the reconstruction framework presented in this paper allows one to potentially consider other multivariate reconstruction bases, better suited for functions not possessing such a simple singularity geometry. This is a topic for future investigation.

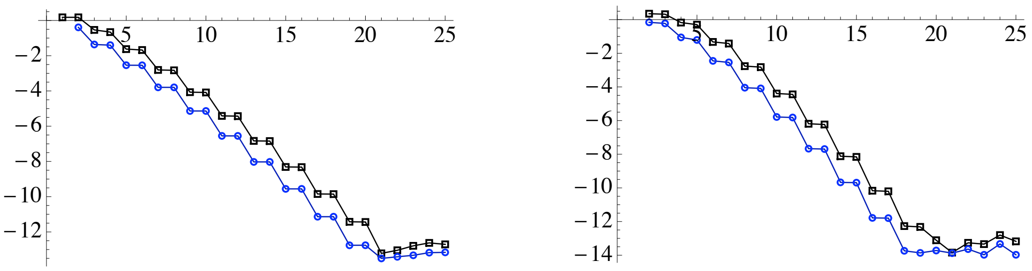

Figure 8: The errors $\log_{10}\|f-f_{n,m}\|$ (squares) and $\log_{10}\|f-f_{n,m}\|_\infty$ (circles) for $n_1=n_2=1,\ldots,25$, where $f(x)=\mathrm{e}^{x^2y}$ (left) and $f(x,y)=\sin 3xy$ (right).

5 Other sampling problems

Overcoming the Gibbs phenomenon in Fourier series is an obvious application of the general framework developed in Section 2. However, there is no reason to restrict to this case, and this framework can be readily applied to design effective methods for a variety of other problems. In this section we describe several related problems, and the application of this framework therein.

5.1 Modified Fourier sampling

Modified Fourier series were proposed in [49] as a minor adjustment of Fourier series. In the domain $[-1,1]$, rather than expanding a function f in the classical Fourier basis

$$\{\cos j\pi x : j\in\mathbb{N}\}\cup\{\sin j\pi x : j\in\mathbb{N}_+\},$$

we construct the modified Fourier expansion using the basis

$$\{\cos j\pi x : j\in\mathbb{N}\}\cup\{\sin(j-\tfrac{1}{2})\pi x : j\in\mathbb{N}_+\},$$

instead. Though this basis arises from only a minor adjustment of the Fourier basis, the result is an improved approximation: the modified Fourier series of a smooth, nonperiodic function converges uniformly at a rate of $\mathcal{O}\left(m^{-1}\right)$, whilst Fourier series suffers from the Gibbs phenomenon. Although the convergence rate remains slow, the improvement over Fourier series, whilst retaining many of their principal benefits, has given rise to a number of applications of such expansions. For a more detailed survey, we refer the reader to [2, 7].

We shall consider modified Fourier expansions in a somewhat different context. Given the similarity between the two bases, it is reasonable to assume that any sampling procedure (e.g. an MRI scanner) can be adapted to compute the modified Fourier coefficients of a given function (or image/signal), as opposed to the standard Fourier samples. Indeed, if

$$\mathcal{F}f(t)=\int_{-1}^{1}f(x)\mathrm{e}^{-\mathrm{i}\pi tx}\,\mathrm{d}x,$$

is the Fourier transform of f, then the modified Fourier coefficients are precisely

$$\begin{aligned}\hat{f}^C_j&=\int_{-1}^{1}f(x)\cos j\pi x\,\mathrm{d}x=\frac{1}{2}\left[\mathcal{F}f(j)+\mathcal{F}f(-j)\right],\\ \hat{f}^S_j&=\int_{-1}^{1}f(x)\sin(j-\tfrac{1}{2})\pi x\,\mathrm{d}x=\frac{\mathrm{i}}{2}\left[\mathcal{F}f(j-\tfrac{1}{2})+\mathcal{F}f(\tfrac{1}{2}-j)\right],\end{aligned}$$

and hence can be computed from samples of the Fourier transform. This raises the question: given that the general framework can handle sampling in either, is there an advantage gained from sampling in the modified Fourier basis, as opposed to the Fourier basis? As we shall show, provided the function is

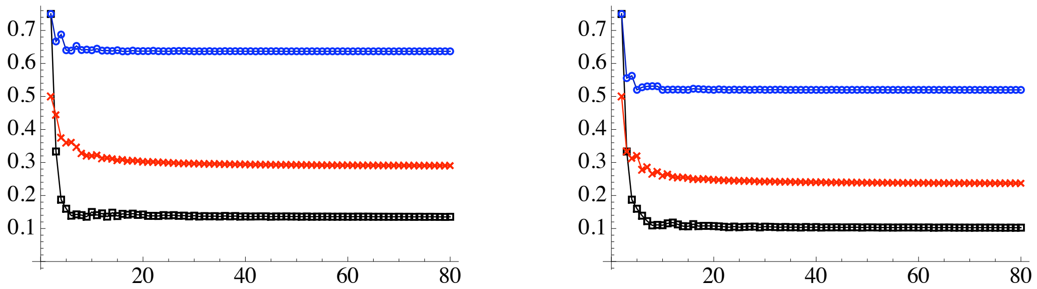

Figure 9: The function $n^{-2}\Theta(n;\theta)$ (squares), the global bound (circles) and the asymptotic bound (crosses) for $n = 1, \ldots, 80$ and $\theta = \frac{1}{2}$ (left), $\theta = \frac{1}{4}$ (right).

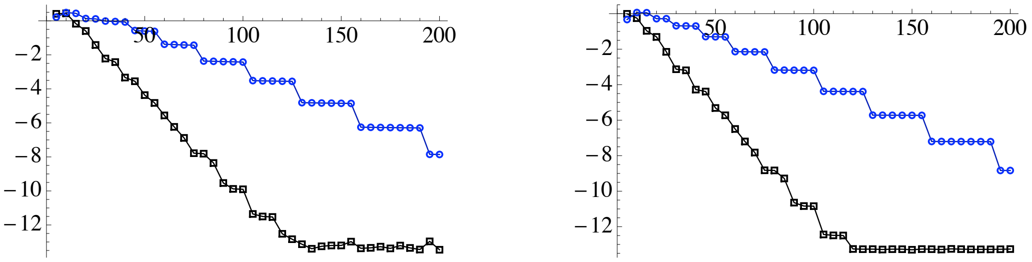

Figure 10: The errors $\log_{10} \|f - f_{n,m}\|_\infty$ (left) and $\log_{10} \|f - f_{n,m}\|$ (right) against $m = 5, 10, 15, \ldots, 200$, where $f_{n,m}$ is computed from modified Fourier (squares) or Fourier (circles) samples.

analytic and nonperiodic, this is indeed the case. Specifically, when we reconstruct in a polynomial basis, we require fewer samples to obtain stable, quasi-optimal recovery to within a prescribed tolerance.

Suppose that we carry out the reconstruction procedure as in Section 3 but using modified Fourier samples instead of Fourier samples. For this, we set

$$\mathcal{P}_m f(x) = \frac{1}{2}\hat{f}_0^C + \sum_{j=1}^{\lfloor \frac{m}{2} \rfloor} \left[\hat{f}_j^C \cos j\pi x + \hat{f}_j^S \sin j\pi x \right].$$

Naturally, we consider the function $\Theta(n;\theta)$ once more. In Figure 9 we plot the function $\Theta(n;\theta)$ for the modified Fourier basis. Upon comparison with Figure 2, we conclude the following: modified Fourier sampling, as opposed to standard Fourier sampling leads to a noticeable improvement. Specifically, the quantity $n^{-2}\Theta(n;\frac{1}{2})$ is approximately 0.15 for large n in the modified Fourier case, as opposed to 0.4 in the Fourier case.

This result means that, if the number of samples m is fixed, we are able to take a much larger value of n in the modified Fourier case, whilst retaining quasi-optimal recovery (with constant $c(\theta)$). To illustrate this improvement, in Figure 10 we compare the errors in approximating the function $f(x) = \mathrm{e}^{-x}\cos 8x$ from either its Fourier or modified Fourier data. In both cases the number of samples m was fixed, and n was chosen so that the parameter $C_{n,m} \geq \frac{1}{2}$. As is evident, the method based on modified Fourier samples greatly outperforms the other. For example, using only $m = 120$ samples, we obtain an error of order 10^{-14} for the former, in comparison to only 10^{-4} for the latter.

As in the Fourier case, to implement the modified Fourier-based approach it is necessary to have estimates for the function $\Theta(n;\theta)$. These are particularly simple to derive:

Lemma 5.1. *Let $\{\psi_j\}_{j\in\mathbb{N}}$ be the modified Fourier basis and $\mathrm{T}_n = \mathbb{P}_{n-1}$. Then the function $\Theta(n;\theta)$ satisfies*

$$\Theta(n;\theta) \leq \left\lceil \frac{2}{\pi\sqrt{1-\theta}} k_n n^2 \right\rceil, \quad \textit{where} \quad k_n = n^{-2} \sup_{\substack{\phi\in\mathbb{P}_{n-1} \\ \|\phi\|=1}} \|\phi'\|.$$

Proof. In [3] it was shown that $\|\phi - \mathcal{P}_m \phi\| \leq \frac{2}{m\pi} \|\phi'\|$ for all sufficiently smooth functions ϕ. The result now follows immediately from the definition of $C_{n,m}$. □

As a result of this lemma, analytical bounds for $\Theta(n;\theta)$ are dependent solely on the constant k_n of the *Markov inequality* $\|\phi'\| \leq k_n n^2 \|\phi\|$, $\forall \phi \in \mathbb{P}_{n-1}$. Markov inequalities and their constants are well understood. The question of determining k_n was first studied rigorously by Schmidt [59], in which the estimates

$$k_n \leq \frac{1}{\sqrt{2}}, \quad \forall n, \qquad \kappa_n \to \frac{1}{\pi}, \quad n \to \infty, \tag{5.1}$$

were derived. In [60] the following improved asymptotic estimate was also obtained:

$$k_n n^2 = \frac{(n+\frac{1}{2})^2}{\pi} \left[1 - \frac{\pi^2 - 3}{12(n+\frac{1}{2})^2} + \frac{R_n}{(n+\frac{1}{2})^4} \right]^{-1}, \quad n \geq 5, \tag{5.2}$$

where $-6 < R_n < 13$. We refer the reader to [11] for a more thorough discussion of both these results and more recent work on this topic. Returning to $\Theta(n;\theta)$ we now substitute the result of Lemma 5.1 into (5.1) and (5.2) to obtain the global and asymptotic bounds. In Figure 9 we compare these bounds to their numerically computed values. The relative sharpness of such estimates is once more observed.

5.2 Polynomial sampling

The primary concern of this paper has been reconstruction from Fourier (or Fourier-like) samples. However, in several circumstances, most notably the spectral approximation of PDEs with discontinuous solutions [32, 38], the problem arises where a piecewise analytic function has been sampled in an orthogonal polynomial basis. As previously noted, this approximation will converge slowly (and suffer from a Gibbs-type phenomenon near discontinuities), hence it is necessary to compute a new approximation with faster convergence. Whilst a version of the spectral reprojection using piecewise Gegenbauer polynomials has been developed for this task [39, 40], the advantages of the method proposed in this paper (see Section 3.6) make it a compelling alternative to this existing approach. Hence, the purpose of this section is to give a brief overview of this application.

It is beyond the scope of this paper to develop this example of the reconstruction procedure in its full generality. Instead, we consider only the recovery of a piecewise analytic function $f : [-1,1] \to \mathbb{R}$ from its first m Legendre polynomial coefficients $\hat{f}_j = \langle f, \psi_j \rangle$, $j = 0, \ldots, m-1$, where $\psi_j = (j+\frac{1}{2})^{\frac{1}{2}} P_j(x)$ is the j^{th} normalised Legendre polynomial. Proceeding as in Section 3.5, we assume that f has jump discontinuities as $-1 < x_1 < \ldots < x_l < 1$, and seek an approximation of the form

$$f_{n,m}(x) = \sum_{r=0}^{l} \sum_{j=0}^{n_r - 1} \alpha_{r,j} \phi_{r,j}(x), \quad n = \sum_{r=0}^{l} n_r,$$

where $\phi_{r,j}(x) = \frac{1}{\sqrt{c_r}} \phi_j(\Lambda_r(x))$, $\Lambda_r(x) = \frac{x - x_r}{c_r} - 1$, $c_r = \frac{1}{2}(x_{r+1} - x_r)$ and $\{\phi_0, \ldots, \phi_{n-1}\}$ is a system of polynomials on $[-1,1]$. Since f is piecewise analytic, we expect exponential convergence of $f_{n,m}$ to f, provided m is sufficiently large in comparison to n.

Aside from determining how large m must be in comparison to n for recovery, the main question remaining is that of implementation, i.e. how to compute the entries of the matrix U. This requires evaluation of the integrals

$$\int_{x_r}^{x_{r+1}} \psi_j(x) \phi_{r,k}(x) \, \mathrm{d}x, \quad j = 0, \ldots, m-1, \; k = 0, \ldots, n-1.$$

Whenever the reconstruction functions $\phi_{r,k}$ arise from Gegenbauer polynomials, these calculations can be done iteratively. For the sake of brevity, we will not describe this computation in full generality. Instead, we consider only the situation where the functions $\phi_{r,k}$ are (appropriately scaled) Legendre polynomials, in which case we are required to compute the integrals

$$\int_{x_r}^{x_{r+1}} P_j(x) P_k(\Lambda_r(x)) \, \mathrm{d}x, \qquad j = 0, \ldots, m-1, \; k = 0, \ldots, n-1, \; r = 0, \ldots, l.$$

We have

Lemma 5.2. *Let*

$$u_{j,k} = \int_a^b P_j(x) P_k(cx+d)\,\mathrm{d}x, \quad j,k = 0,1,2,\dots, \tag{5.3}$$

where $ca + d = -1$ *and* $cb + d = 1$*. Then*

$$u_{0,0} = b - a, \quad u_{j,0} = \frac{1}{j}\left[P_{j+1}(x) - xP_j(x)\right]_{x=a}^{b}, \quad j = 1,2,\dots,$$

$$u_{0,k} = \frac{2}{c}\delta_{0,k}, \quad u_{1,k} = \frac{2}{3c}\delta_{1,k} - \frac{2d}{c}\delta_{0,k}, \quad k = 0,1,2,\dots,$$

and, for $j \geq 2$ *and* $k \geq 1$*,*

$$u_{j,k} = \frac{(2j-1)(k+1)}{cj(2k+1)} u_{j-1,k+1} + \frac{(2j-1)k}{cj(2k+1)} u_{j-1,k-1} - \frac{d(2j-1)}{cj} u_{j-1,k} - \frac{j-1}{j} u_{j-2,k}. \tag{5.4}$$

Proof. Recall the recurrence relation

$$jP_j(x) = (2j-1)xP_{j-1}(x) - (j-1)P_{j-2}(x), \quad j = 2,3,\dots, \tag{5.5}$$

for Legendre polynomials [1, chpt 22]. Substituting this into (5.3) gives

$$u_{j,k} = \frac{2j-1}{j}\int_a^b xP_{j-1}(x)P_k(cx+d)\,\mathrm{d}x - \frac{j-1}{j}u_{j-2,k}.$$

Letting $x \mapsto cx + d$ in (5.5) and rearranging, we find that

$$xP_k(x) = \frac{k+1}{c(2k+1)}P_{k+1}(cx+d) - \frac{d}{c}P_k(cx+d) + \frac{k}{c(2k+1)}P_{k-1}(cx+d).$$

The recurrence (5.4) now follows upon substituting this into the previous expression.

Next consider $u_{j,0}$. Since $P_0 \equiv 1$, we have $u_{0,0} = b - a$ and $u_{j,0} = \int_a^b P_j(x)\,\mathrm{d}x$ for $j \geq 1$. Recall that the j^{th} Legendre polynomial satisfies the Legendre differential equation [1, chpt 22]

$$\left[(1-x^2)P_j'(x)\right]' = -j(j+1)P_j(x).$$

Substituting for P_j in $\int_a^b P_j(x)\,\mathrm{d}x$ and integrating gives

$$u_{j,0} = \frac{1}{j(j+1)}\left[(x^2-1)P_j'(x)\right]_{x=a}^{b}.$$

The result now follows directly from the expression

$$(1-x^2)P_j'(x) = (j+1)(xP_j(x) - P_{j+1}(x)), \quad j = 0,1,2,\dots.$$

To complete the proof, we consider $u_{0,k}$ and $u_{1,k}$. By the assumptions on a, b, c, d, we find that

$$u_{0,k} = \frac{1}{c}\int_{-1}^{1} P_k(x)\,\mathrm{d}x.$$

Orthogonality now gives $u_{0,k} = \frac{2}{c}\delta_{0,k}$, as required. For $u_{1,k}$ we have

$$u_{1,k} = \frac{1}{c}\int_{-1}^{1}(y-d)P_k(y)\,\mathrm{d}y = \frac{1}{c}\left(\frac{2}{3}\delta_{1,k} - 2d\delta_{0,k}\right),$$

as required. □

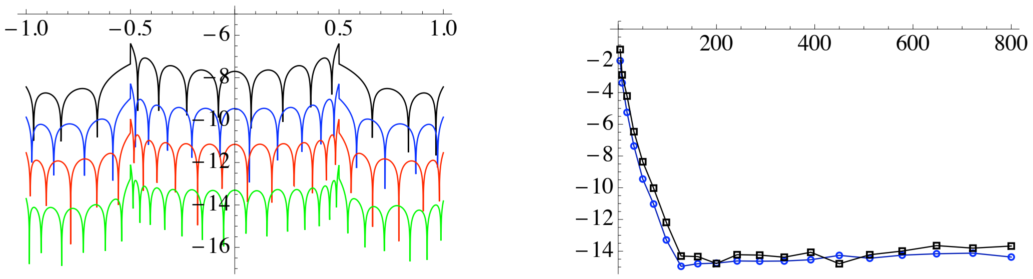

Figure 11: Left: the error $\log_{10}|f(x) - f_{n,m}(x)|$ for $-1 \leq x \leq 1$ and $m = 20, 40, 80, 160$. Right: log errors $\log_{10}\|f - f_{n,m}\|_\infty$ (squares) and $\log_{10}\|f - f_{n,m}\|$ (circles) against m.

n	8	16	24	32	40	48	56	64	72	80
$C_{n,m}$	0.98	0.87	0.85	0.85	0.85	0.84	0.84	0.84	0.84	0.84
$\kappa(U^*U)$	19.97	4.17	3.57	3.57	3.50	3.43	3.43	3.41	3.38	3.38

Table 5: The values $C_{n,m}$ and $\kappa(U^*U)$ against n, where $m = \frac{1}{8}n^2$.

In Figure 11 we consider the approximation of the function

$$f(x) = \begin{cases} \sin(\cos x) & -\frac{1}{2} \leq x < \frac{1}{2}, \\ 0 & \text{otherwise.} \end{cases} \tag{5.6}$$

by the aforementioned method, using parameter values $m = \frac{1}{8}n^2$, $n_0 = n_2 = \frac{1}{4}n$ and $n_1 = \frac{1}{2}n$. As shown, we obtain 13 digits of accuracy using only $m \approx 120$ Legendre coefficients of (5.6). Note that, although we have not shown it, the scaling $m = \mathcal{O}\left(n^2\right)$ appears to be sufficient for recovery. Numerical results demonstrating this hypothesis are given in Table 5.

The function (5.6) was introduced in [39] to test spectral reprojection when applied to this type of problem. As shown in Figure 11, we obtain a uniform error of roughly 10^{-8} using only $m = 40$ coefficients, and when $m = 120$, the corresponding value is 10^{-14}. In comparison, the spectral reprojection method of [39] gives errors of roughly 10^{-3} and 10^{-7} for these values of m (see [39, Fig. 3]), the latter being 10^7 times larger.

Whilst this method appears to be a promising alternative, it should be mentioned that the recursive scheme introduced to compute the entries of U requires $\mathcal{O}\left(m^2\right)$ operations. Since only $\mathcal{O}\left(mn\right)$ operations are required to compute the approximation $f_{n,m}$ once U has been computed, this is clearly less than ideal. Having said that, spectral reprojection method requires $\mathcal{O}\left(m^2\right)$ operations to compute each approximant, whereas with this scheme such higher cost is only incurred in a preprocessing step.

6 Conclusions and future work

We have presented a reconstruction procedure to recover any element of a Hilbert space using any collection of linearly independent vectors, given a finite number of its samples with respect to an arbitrary Riesz basis. This approach is both stable and quasi-optimal, provided the number of samples m is sufficiently large in comparison to the number of reconstruction vectors n. Moreover, this condition can be estimated numerically or, in certain circumstances, analytically.

A prominent example of this approach is the reconstruction of a piecewise analytic function from its Fourier samples. Using a piecewise polynomial basis, this results in an approximation that converges root-exponentially fast in terms of m, or exponentially fast in n.

The framework introduced in this paper is one of the first steps in the development of stable and accurate reconstruction techniques in Hilbert spaces, and their applications to a variety of different problems. We now detail a number of areas of current and future work:

1. Piecewise polynomial reconstructions from polynomial samples. In the penultimate section of this paper we detailed the recovery of a piecewise analytic function in a piecewise polynomial basis, given its Legendre polynomial expansion coefficients. Herein, an important open problem is verifying that the scaling $m = \mathcal{O}\left(n^2\right)$ is sufficient for reconstruction. Other challenges involve devising an iterative scheme for computing the entries of U valid for reconstructions in arbitrary Gegenbauer polynomials, and which involves only $\mathcal{O}\left(mn\right)$ operations, as opposed to $\mathcal{O}\left(m^2\right)$. Naturally, future work will also investigate the extension of this approach to reconstructions from arbitrary Gegenbauer polynomial expansion coefficients, as opposed to just Legendre polynomial expansion coefficients.

2. Gegenbauer polynomial reconstructions from Fourier samples. As shown, the reconstruction procedure can be implemented with arbitrary Gegenbauer polynomials. However, unless Legendre polynomials are used, the reconstruction is not completely stable. This problem arises because Gegenbauer polynomials do not form a Riesz basis for the space $\mathrm{L}^2(-1,1)$ unless $\lambda = \frac{1}{2}$. However, Gegenbauer polynomials do form an orthogonal basis for the weighted space $\mathrm{L}^2_{\omega}(-1,1)$, where $\omega(x) = (1-x^2)^{\lambda-\frac{1}{2}}$. Hence, it is natural to ask whether the reconstruction procedure can be adjusted to incorporate this additional structure, thereby yielding a stable method. It turns out that this can be done, with the first step being the derivation of an extended abstract framework along similar lines to Section 2. We are currently compiling results in this case, and will report the details in a future paper.

3. Fast methods. For practical, large-scale implementations of this framework, the computational cost figure of $\mathcal{O}\left(mn\right)$ may be too large. For this reason, the use of fast methods to compute the reconstruction $f_{n,m}$ is another topic of current investigation. Potential means for doing this include using so-called nonuniform FFTs [25, 26] to efficiently perform matrix-vector multiplications.

4. Applications. Aside from the obvious applications in image and signal processing, the are many other potential uses of the procedure. First, it may be applicable to the spectral discretisation of PDEs. Spectral methods are extremely efficient for solving problems with smooth solutions. However, for problems that develop discontinuities, e.g. hyperbolic conservation laws, a postprocessor is required to recover high accuracy [38]. Spectral reprojection is frequently used in such problems (see [32, 38] and references therein). Given the potential advantages of the method developed in this paper (see Section 3.6), it is of significant interest to apply this approach to these problems. Aside from high accuracy, a pertinent issue in the use of spectral approximations for nonsmooth problems is the question of stability [38]. Since the reconstruction procedure developed herein is numerically stable, we expect there to also be benefits in this regard. Outside of PDEs, the Gegenbauer reconstruction technique has also been extended to other types of series, including radial basis functions [54], Fourier–Bessel series [55] and spherical harmonics [30]. Future work will also consider generalisations along these lines.

5. Spline and wavelet-based reconstructions. Reconstructions in spline and wavelet bases are vitally important in numerous applications, including image and signal processing [63]. In [4], the authors gave a first insight into the application of such bases to the Fourier sampling problem. However, the theory is far from complete. In particular, good estimates for the corresponding quantity $\Theta(n;\theta)$ are currently lacking.

6. Recovery from other sampling data. The discrete analogue of the Fourier coefficient recovery problem involves the reconstruction of a function from m equispaced samples in $[-1,1]$. This problem has received more attention of late [16, 57] than the continuous case considered in this paper. In particular, the use of oversampling, similar to that done in this paper, has been considered in [17]. Future work will look to extend these ideas to related problems involving recovery from nonuniform grids (see [31] for a discussion of such problems). In particular, with application to spectral collocation schemes based on Chebyshev or Legendre polynomials, the recovery of piecewise analytic functions given their values at Gauss or Gauss–Lobatto nodes.

Returning to the continuous problem, there is also significant interest in reconstructions from nonharmonic Fourier samples. In this case, the sampling vectors are typically (but not always) assumed to constitute a frame. The extension of the framework of this paper to this problem is currently being investigated. For related discussions, as well as other methods for this problem, see [33, 64, 65].

7. A geometric interpretation of reconstruction. The abstract reconstruction framework developed in this paper can be viewed as a generalisation of the technique of consistent reconstructions (see Section 2.1 and [4]). As shown in [28, 29], consistent reconstructions can be interpreted geometrically in terms of oblique projections onto particular subspaces. It transpires that the framework developed in this paper also possesses such an interpretation. In turn, this viewpoint allows one to develop important notions of optimality for this approach. A detailed discussion of this topic can be found in [6].

8. Infinite-dimensional compressed sensing. An important question in modern sampling is that of sparsity. The recently-developed field of compressed sensing allows one to successfully reconstruct sparse signals with dramatic subsampling [19, 23]. However, although this has had a dramatic impact, compressed sensing is currently a finite-dimensional technique. Since real-world signal and images are typically infinite-dimensional (or analog), the need for a more comprehensive framework is apparent. It transpires that, by combining the techniques of this paper with those of compressed sensing, one can develop both a theory and a method for subsampling in infinite-dimensional sparse recovery problems. This presents a significant extension of compressed sensing to a large class of infinite-dimensional (i.e. analog) signal models. These recent developments are documented in [5].

References

[1] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions*. Dover, 1974.

[2] B. Adcock. *Modified Fourier expansions: theory, construction and applications*. PhD thesis, University of Cambridge, 2010.

[3] B. Adcock. Multivariate modified Fourier series and application to boundary value problems. *Numer. Math.*, 115(4):511–552, 2010.

[4] 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.

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

[6] B. Adcock and A. C. Hansen. Sharp bounds, optimality and a geometric interpretation for generalised sampling in Hilbert spaces. *Technical report NA2011/10, DAMTP, University of Cambridge*, 2011.

[7] B. Adcock and D. Huybrechs. Multivariate modified Fourier expansions. In E. Rønquist et al, editor, *Proceedings of the International Conference on Spectral and High Order Methods*, 2011.

[8] R. Archibald, K. Chen, A. Gelb, and R. Renault. Improving tissue segmentation of human brain MRI through preprocessing by the Gegenbauer reconstruction method. *NeuroImage*, 20(1):489–502, 2003.

[9] R. Archibald and A. Gelb. A method to reduce the Gibbs ringing artifact in MRI scans while keeping tissue boundary integrity. *IEEE Transactions on Medical Imaging*, 21(4):305–319, 2002.

[10] H. Bateman. *Higher Transcendental Functions*. Vol. 2, McGraw–Hill, New York, 1953.

[11] A. Böttcher and P. Dörfler. Weighted Markov-type inequalities, norms of Volterra operators, and zeros of Bessel functions. *Math. Nachr.*, 283(1):40–57, 2010.

[12] J. P. Boyd. *Chebyshev and Fourier Spectral Methods*. Springer–Verlag, 1989.

[13] J. P. Boyd. A comparison of numerical algorithms for Fourier Extension of the first, second, and third kinds. *J. Comput. Phys.*, 178:118–160, 2002.

[14] J. P. Boyd. Trouble with Gegenbauer reconstruction for defeating Gibbs phenomenon: Runge phenomenon in the diagonal limit of Gegenbauer polynomial approximations. *J. Comput. Phys.*, 204(1):253–264, 2005.

[15] J. P. Boyd. Acceleration of algebraically-converging Fourier series when the coefficients have series in powers of $1/n$. *J. Comput. Phys.*, 228:1404–1411, 2009.

[16] J. P. Boyd and J. R. Ong. Exponentially-convergent strategies for defeating the Runge phenomenon for the approximation of non-periodic functions. I. Single-interval schemes. *Commun. Comput. Phys.*, 5(2–4):484–497, 2009.

[17] J.P. Boyd and F. Xu. Divergence (Runge phenomenon) for least-squares polynomial approximation on an equispaced grid and mock-Chebyshev subset interpolation. *Appl. Math. Comput.*, 210(1):158–168, 2009.

[18] C. Brezinski. Extrapolation algorithms for filtering series of functions, and treating the Gibbs phenomenon. *Numer. Algorithms*, 36:309–329, 2004.

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

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

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

[22] V. Dominguez, I. G. Graham, and V. P. Smyshlyaev. Stability and error estimates for Filon-Clenshaw-Curtis rules for highly-oscillatory integrals. *IMA J. Num. Anal. (to appear)*, 2011.

[23] D. L. Donoho. Compressed sensing. *IEEE Trans. Inform. Theory*, 52(4):1289–1306, 2006.

[24] T. A. Driscoll and B. Fornberg. A Padé-based algorithm for overcoming the Gibbs phenomenon. *Numer. Algorithms*, 26:77–92, 2001.

[25] A. Dutt and V. Rokhlin. Fast Fourier Transforms for nonequispaced data. *SIAM J. Sci. Comput.*, 14(6):1368–1393, 1993.

[26] A. Dutt and V. Rokhlin. Fast Fourier Transforms for nonequispaced data, II. *Appl. Comput. Harmon. Anal.*, 2:85–100, 1995.

[27] K. S. Eckhoff. On a high order numerical method for functions with singularities. *Math. Comp.*, 67(223):1063–1087, 1998.

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

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

[30] A. Gelb. The resolution of the Gibbs phenomenon for spherical harmonics. *Math. Comp.*, 66(218):699–717, 1997.

[31] A. Gelb. Reconstruction of piecewise smooth functions from non-uniform grid point data. *J. Sci. Comput.*, 30(3):409–440, 2007.

[32] A. Gelb and S. Gottlieb. The resolution of the Gibbs phenomenon for Fourier spectral methods. In A. Jerri, editor, *Advances in The Gibbs Phenomenon*. Sampling Publishing, Potsdam, New York, 2007.

[33] A. Gelb and T. Hines. Detection of edges from nonuniform Fourier data. *J. Fourier Anal. Appl. (to appear)*, 2011.

[34] A. Gelb and E. Tadmor. Detection of edges in spectral data. *Appl. Comput. Harmon. Anal.*, 7(1):101, 1999.

[35] A. Gelb and J. Tanner. Robust reprojection methods for the resolution of the Gibbs phenomenon. *Appl. Comput. Harmon. Anal.*, 20:3–25, 2006.

[36] D. Gilbarg and N.S. Trudinger. *Elliptic Partial Differential Equations of Second Order*. Springer Verlag, 2001.

[37] G. H. Golub and C. F. Van Loan. *Matrix Computations*. John Hopkins University Press, Baltimore, 2nd edition, 1989.

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

[39] D. Gottlieb and C-W. Shu. On the Gibbs phenomenon IV: Recovering exponential accuracy in a subinterval from a Gegenbauer partial sum of a piecewise analytic function. *Math. Comp.*, 64(211):1081–1095, 1995.

[40] D. Gottlieb and C-W. Shu. On the Gibbs phenomenon III: Recovering exponential accuracy in a sub- interval from a spectral partial sum of a piecewise analytic function. *SIAM J. Num. Anal.*, 33(1):280–290, 1996.

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

[42] D. Gottlieb, C-W. Shu, A. Solomonoff, and H. Vandeven. On the Gibbs phenomenon I: Recovering exponential accuracy from the Fourier partial sum of a nonperiodic analytic function. *J. Comput. Appl. Math.*, 43(1–2):91–98, 1992.

[43] K. Gröchenig, Z. Rzeszotnik, and T. Strohmer. Quantitative estimates for the finite section method and Banach algebras of matrices. *Integral Equations and Operator Theory*, 67(2):183–202, 2011.

[44] 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.

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

[46] J. S. Hesthaven, S. Gottlieb, and D. Gottlieb. *Spectral Methods for Time-Dependent Problems*. Cambridge University Press, 2007.

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

[48] D. Huybrechs. On the Fourier extension of non-periodic functions. *SIAM J. Numer. Anal.*, 47(6):4326–4355, 2010.

[49] A. Iserles and S. P. Nørsett. From high oscillation to rapid approximation I: Modified Fourier expansions. *IMA J. Num. Anal.*, 28:862–887, 2008.

[50] A.J. Jerri, editor. *The Gibbs phenomenon in Fourier Analysis, Splines, and Wavelet Approximations*. Kluwer Academic, Kordrecht, The Netherlands, 1998.

[51] A.J. Jerri, editor. *Advances in the Gibbs Phenomenon*. Sampling Publishing, Potsdam, New York, 2007.

[52] J.-H. Jung and B. D. Shizgal. Towards the resolution of the Gibbs phenomena. *J. Comput. Appl. Math.*, 161(1):41–65, 2003.

[53] 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.

[54] J.H. Jung, S. Gottlieb, S.O. Kim, C.L. Bresten, and D. Higgs. Recovery of high order accuracy in radial basis function approximations of discontinuous problems. *J Sci Comput*, 45:359–381, 2010.

[55] J.R. Kamm, T.O. Williams, J.S. Brock, and S. Li. Application of Gegenbauer polynomial expansions to mitigate Gibbs phenomenon in Fourier–Bessel series solutions of a dynamic sphere problem. *Int. J. Numer. Meth. Biomed. Engng.*, 26(1276–1292), 2010.

[56] T. W. Körner. *Fourier Analysis*. Cambridge University Press, 1988.

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

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

[59] E. Schmidt. Die asymptotische Bestimmung des Maximums des Integrals über das Quadrat der Ableitung eines normierten Polynoms, dessen Grad ins Unendliche wächst. *Sitzungsber. Preuss. Akad. Wiss.*, page 287, 1932.

[60] E. Schmidt. Über die nebst ihren Ableitungen orthogonalen Polynomensysteme und das zugehörige Extremum. *Math. Ann.*, 119:165–204, 1944.

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

[62] E. Tadmor and J. Tanner. Adaptive mollifiers for high resolution recovery of piecewise smooth data from its spectral information. *Foundations of Computational Mathematics*, 2(2):155–189, 2002.

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

[64] A. Viswanathan, A. Gelb, D. Cochran, and R. Renaut. On reconstructions from non-uniform spectral data. *J. Sci. Comput.*, 45(1–3):487–513, 2010.

[65] L. Wang and Y. Wang. Reconstruction from irregular Fourier samples and Gaussian spectral mollifiers. *preprint*, 2011.