Generalized sampling and the stable and accurate reconstruction of piecewise analytic functions from their Fourier coefficients

Ben Adcock
Department of Mathematics
Purdue University
USA

Anders C. Hansen
DAMTP, Centre for Mathematical Sciences
University of Cambridge
United Kingdom

May 8, 2013

Abstract

Suppose that the first m Fourier coefficients of a piecewise analytic function are given. Direct expansion in a Fourier series suffers from the Gibbs phenomenon and lacks uniform convergence. Nonetheless, in this paper we show that, under very broad conditions, it is always possible to recover an n-term expansion in a different system of functions using only these coefficients. Such an expansion can be made arbitrarily close to the best possible n-term expansion in the given system. Thus, if a piecewise polynomial basis is employed, for example, exponential convergence can be restored. The resulting method is linear, numerically stable and can be implemented efficiently in only $\mathcal{O}(nm)$ operations.

A key issue is how the parameter m must scale in comparison to n to ensure such recovery. We derive analytical estimates for this scaling for large classes of polynomial and piecewise polynomial bases. In particular, we show that in many important cases, including the case of piecewise Chebyshev polynomials, this scaling is quadratic: $m = \mathcal{O}(n^2)$. Therefore, with a system of polynomials that the user is essentially free to choose, one can restore exponential accuracy in n and root-exponential accuracy in m. This generalizes a result proved recently for piecewise Legendre polynomials.

1 Introduction

The Fourier expansion of an analytic and periodic function f converges exponentially fast in the truncation parameter m. For this reason—as well as the fact that the expansion can be computed rapidly via the Fast Fourier Transform (FFT)—such approximations are extremely widely used in all areas of computational mathematics, with one important instance being the spectral solution of partial differential equations (PDEs) [17, 21]. However, rapid convergence is lost when the function is no longer analytic. Indeed, when f is only piecewise analytic, one encounters the familiar Gibbs phenomenon. This not only destroys local accuracy—characteristic $\mathcal{O}(1)$ oscillations are witnessed near each discontinuity—it also inhibits a good global approximation. The truncated expansion converges only very slowly in the L^2 norm, and not at all when the error is measured in the uniform norm. Notably, this issue arises when computing spectral approximations to PDEs that develop shocks—hyperbolic conservation laws, for example [31, 40].

This naturally leads to the following question: given the first m Fourier coefficients of a piecewise analytic function f, is it possible to compute a better approximation? The problem is not new, and there have been many different approaches developed for its solution (see, for example, [2, 12, 13, 14, 19, 20, 23, 24, 25, 26, 49] and references therein). Of this large collection, perhaps the most successful and commonly used is the method of spectral reprojection [30, 32, 33, 34]. Introduced by D. Gottlieb et al. [34], in this approach the m Fourier coefficients of f are used to approximate the first $n = \mathcal{O}(m)$ coefficients in a new basis of piecewise polynomials (the *reconstruction* basis). Provided this basis is chosen appropriately, one recovers exponential accuracy in m. To date, spectral reprojection has been successfully used in a range of applications, including image reconstruction [10, 11], and the postprocessing of spectral discretizations of hyperbolic PDEs [28, 31, 40].

Whilst spectral reprojection has been rather successful, there are unfortunately a number of drawbacks. Spectral reprojection achieves rapid convergence by employing a particular choice of reconstruction basis,

and herein lies a problem. Only very few bases (known as *Gibbs complementary bases* [30, 33]) have this property, with the two most commonly used being those consisting of ultraspherical (Gegenbauer) and Freud polynomials. In both cases, however, the parameter α defining the polynomials must scale linearly with n. This is not only computationally inconvenient—changing n requires recomputation of the whole reprojection basis—in addition a rather careful selection of parameters must be made to ensure such convergence and to avoid a Runge-type phenomenon [18].

For these reasons, spectral reprojection affords little flexibility to the user. At the very least, Gegenbauer and Freud polynomials are no where near as easy to use and manipulate as Chebyshev polynomials, for example, with the latter being amenable to the FFT. Another issue, as shown empirically in [6], is that even if exponential convergence occurs, the rate of such convergence in the Gibbs complementary basis may be quite slow in practice, meaning that many Fourier coefficients may be required to achieve high accuracy.

With this in mind, the purpose of this paper is to consider a different approach based on an alternative idea. Unlike spectral reprojection, where the reconstruction basis must be suitably chosen to ensure rapid convergence, the method we develop in this paper allows the user to employ an arbitrary basis. In other words, we address the following problem:

Problem 1.1. *Let the first m Fourier coefficients of a piecewise analytic function f be given, where m is sufficiently large. If $\mathcal{Q}_n f$ denotes the n-term expansion $\mathcal{Q}_n f$ in an arbitrary basis of piecewise polynomials orthogonal with respect to the inner product $\langle \cdot, \cdot \rangle$, recover an approximation $f_{n,m}$ to $\mathcal{Q}_n f$ from the given data that satisfies $\|f - f_{n,m}\| \asymp \|f - \mathcal{Q}_n f\|$ as $n, m \to \infty$.*

Here and elsewhere we use the notation $a_n \asymp b_n$ for nonnegative sequences $\{a_n\}_{n=1}^\infty$ and $\{b_n\}_{n=1}^\infty$ to mean that there exist constants $c_1, c_2 > 0$ such that $c_1 a_n \leq b_n \leq c_2 a_n$ for all sufficiently large n. In particular, the property $\|f - f_{n,m}\| \asymp \|f - \mathcal{Q}_n f\|$ stipulates that approximation $f_{n,m}$ is *quasi-optimal* to f from the set $\mathbb{P}_n$: as $n \to \infty$, $f_{n,m}$ converges to f at precisely the same rate as the best approximation $\mathcal{Q}_n f$. Since f is piecewise analytic, its polynomial expansion $\mathcal{Q}_n f$ converges exponentially fast in n. Hence quasi-optimality implies the same convergence for $f_{n,m}$ as $n \to \infty$.

As we show in this paper, Problem 1.1 can be solved under very general conditions by letting m vary with n in a suitable way. As a result, the restriction of having to use a Gibbs complementary basis can be avoided, and this allows more convenient polynomials bases to be used in reconstruction (such as the afore-mentioned Chebyshev basis). Note that the case of (piecewise) Legendre polynomial bases was developed in [6]. Therein several examples demonstrating the improved performance of the method over spectral re-projection were also given. The main contribution of this paper is to generalize this work to arbitrary bases of piecewise polynomials.

1.1 Reconstructions in Hilbert spaces

Our solution to Problem 1.1 is based on recasting it in terms of sampling and reconstruction in abstract Hilbert spaces. To this end, let H^S and H^R be subspaces of a vector space V that form Hilbert spaces with respect to the bilinear forms $\langle \cdot, \cdot \rangle_S$ and $\langle \cdot, \cdot \rangle_R$ respectively. Suppose that $\{\psi_j\}_{j=1}^\infty, \{\phi_j\}_{j=1}^\infty$ are orthonormal bases for $(H^S, \langle \cdot, \cdot \rangle_S)$ and $(H^R, \langle \cdot, \cdot \rangle_R)$ respectively (the *sampling* and *reconstruction* bases), and let $U \subseteq H^R$ be a subspace (U will consist of those functions f we wish to reconstruct). For $f \in U$, let

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

be the *samples* of f. The method we develop in this paper is designed to solve the following problem:

Problem 1.2. *Given the first m samples $\hat{f}_1, \ldots, \hat{f}_m$ of $f \in U$, recover an approximation $f_{n,m}$ to $\mathcal{Q}_n f$ satisfying $\|f - f_{n,m}\|_R \asymp \|f - \mathcal{Q}_n f\|_R$, where $\mathcal{Q}_n : (H^R, \langle \cdot, \cdot \rangle_R) \to T_n := \mathrm{span}\{\phi_1, \ldots, \phi_n\}$ is the orthogonal projection.*

Under certain assumptions, it transpires that this problem can be solved with a numerically stable, linear method. The key is that m must scale appropriately to n. The required scaling can be derived numerically, and, in the cases we consider in this paper, analytically.

It is straightforward to see that this abstract framework can be applied to the particular case of Problem 1.1. For example, consider the simple scenario where f is analytic, but nonperiodic (equivalently, $f : \mathbb{T} \to \mathbb{R}$ has a jump at $x = -1$, where $\mathbb{T} = [-1, 1)$ is the unit torus). Given the first m Fourier coefficients of f, we

would like to reconstruct using the orthonormal basis of Chebyshev polynomials $\phi_j(x) = c_j T_j(x)$ of the first kind (for example), where

$$T_j(x) = \cos(j \arccos x), \quad j = 0, 1, 2, \ldots,$$

and $c_0 = 1/\sqrt{\pi}$, $c_j = \sqrt{2/\pi}$, $j \neq 1$. Since the polynomials T_j are orthogonal with respect to the weight function $w(x) = (1 - x^2)^{-\frac{1}{2}}$, we thus let $\mathrm{H}^S = \mathrm{L}^2(-1, 1)$ and $\mathrm{H}^R = \mathrm{L}^2_w(-1, 1)$ be the spaces of square-integrable functions and weighted square-integrable functions with respect to the weight function w respectively, and define $\psi_j(x) = \frac{1}{\sqrt{2}} \exp(\mathrm{i} j \pi x)$ to be the standard Fourier basis on H^S.

As is no doubt apparent to the reader, Problem 1.2 is actually far more general than Problem 1.1. Unsurprisingly, the abstract framework introduced to solve Problem 1.2 can be applied far more widely than the problem of reconstructing in polynomial bases from Fourier data. We shall briefly discuss other applications of this framework at the end of this paper.

1.2 Orthogonal polynomial systems

Before we state our main results, it is useful to introduce some notation. In this paper we shall be principally concerned with Jacobi polynomials, with the corresponding weight function

$$w_{\alpha,\beta}(x) = (1 - x)^\alpha (1 + x)^\beta, \quad \alpha, \beta > -1.$$

We write $\mathrm{L}^2_{\alpha,\beta}(-1, 1)$ for the space of weighted square-integrable functions with respect to $w_{\alpha,\beta}$, and $\|\cdot\|_{\alpha,\beta}$ for the corresponding norm. Our main examples herein will be ultraspherical (or Gegenbauer) polynomials ($\alpha = \beta = \gamma$), and in particular, Legendre polynomials ($\gamma = 0$) and Chebyshev polynomials of the first ($\gamma = -\frac{1}{2}$) and second ($\gamma = \frac{1}{2}$) kinds. Whenever $\alpha = \beta = \gamma$ we shall use the slightly more succinct notation $\mathrm{L}^2_\gamma(-1, 1)$, $\|\cdot\|_\gamma$ and w_γ.

Without a doubt, Chebyshev and Legendre polynomials are the most commonly used in applications—in particular, spectral methods for PDEs [17, 21]. Chebyshev polynomials are chosen for reconstruction because of their aforementioned computational flexibility, whereas Legendre polynomials are desirable because of the simplicity of the Legendre weight function $w(x) \equiv 1$. However, it is of both theoretical and practical interest to maintain the generality of Jacobi polynomials. Indeed, whilst Legendre and Chebyshev polynomials are most frequently used in practice, there are a number of applications which employ more general ultraspherical and Jacobi polynomial systems [17, 21, 35, 36].

If f is piecewise analytic with jumps at $-1 < x_1 < \ldots < x_l < 1$ we shall seek to reconstruct it in a basis of piecewise polynomials. Thus, if $\alpha = \{\alpha_0, \ldots, \alpha_l\}$ and $\beta = \{\beta_0, \ldots, \beta_l\}$ where $\alpha_r, \beta_r > -1$ for $r = 0, \ldots, l$, we define the piecewise Jacobi weight function $w_{\alpha,\beta}(x)$ by

$$w_{\alpha,\beta}(x) = (x_{r+1} - x)^{\alpha_r}(x - x_r)^{\beta_r}, \quad x \in I_r := (x_r, x_{r+1}), \qquad r = 0, \ldots, l.$$

Here $x_0 = -1$ and $x_{l+1} = 1$. The corresponding orthonormal system of piecewise polynomials can be obtained by appropriately scaling from the standard interval $[-1, 1]$ to each subinterval of smoothness I_r (see §3.1). When considering a finite expansion in such functions, we shall write $n = (n_0, \ldots, n_l) \in \mathbb{N}^{l+1}$ for the vector of polynomial degrees in each subinterval I_r, and define

$$\mathrm{T}_n = \{\phi : \phi|_{I_r} \in \mathbb{P}_{n_r}, \ r = 0, \ldots, l\}. \tag{1.1}$$

The operator $\mathcal{Q}_n : \mathrm{L}^2_{\alpha,\beta}(-1, 1) \to \mathrm{T}_n$ will be the orthogonal projection with respect to the weight function $w_{\alpha,\beta}$. As in the no-jump (analytic and nonperiodic) setting, the main examples we consider in this paper involve reconstructions in piecewise Chebyshev or Legendre polynomials, i.e. $\alpha = \beta = \gamma$, where $\gamma = \{-\frac{1}{2}, \ldots, -\frac{1}{2}\}$ (first kind), $\gamma = \{\frac{1}{2}, \ldots, \frac{1}{2}\}$ (second kind) or $\gamma = \{0, \ldots, 0\}$ respectively.

Remark 1.3 Many of the results we prove in this paper extend quite trivially to the modified Jacobi weight

$$w_{\alpha,\beta}(x) = g(x)(1 - x)^\alpha(1 + x)^\beta, \quad \alpha, \beta > -1, \tag{1.2}$$

where $g(x)$ is analytic and positive on $[-1, 1]$. Likewise, one may define the piecewise modified Jacobi weight

$$w_{\alpha,\beta}(x) = g_r(x)(x_{r+1} - x)^{\alpha_r}(x - x_r)^{\beta_r}, \quad x \in I_r := (x_r, x_{r+1}), \tag{1.3}$$

where g_r is analytic and positive in $\overline{I_r}$. Although $w_{\alpha,\beta}$, as given by (1.3), is only unique (for given α and β) up to multiplication by a collection $g_0, \ldots, g_l$ of positive, analytic functions, we shall continue to write $\mathrm{L}^2_{\alpha,\beta}(-1, 1)$ for the corresponding space of square-integrable functions with respect to (1.3), and shall not make the dependence on the functions $g_0, \ldots, g_l$ explicit.

1.3 Key results

We devote the first part of this paper to the solution of Problem 1.2 and the properties of the resulting framework, including quasi-optimality and numerical stability. In the second half, we consider its application to Problem 1.1. In this regard, the key result we prove is as follows:

Theorem 1.4. *Let $\alpha = \{\alpha_0, \ldots, \alpha_l\}$ and $\beta = \{\beta_0, \ldots, \beta_l\}$, where $\alpha_r, \beta_r > -1$, $r = 0, \ldots, l$, be parameters corresponding to the piecewise (modified) Jacobi weight $w_{\alpha,\beta}$, and suppose that $\mathcal{Q}_n$ is the orthogonal projection onto T_n with respect to $w_{\alpha,\beta}$, as given by (1.1). Define*

$$\tilde{\alpha}_r = \left\{ \begin{array}{ll} \alpha_r & -1 < \alpha_r < 1 \\ 1 - \epsilon_r & \alpha_r \geq 1 \end{array} \right. , \qquad \tilde{\beta}_r = \left\{ \begin{array}{ll} \beta_r & -1 < \beta_r < 1 \\ 1 - \epsilon_r' & \beta_r \geq 1 \end{array} \right. ,$$

where $0 < \epsilon_r, \epsilon_r' < 2$ are arbitrary, and let $\mathrm{U} := \mathrm{L}^2_{\tilde{\alpha}, \tilde{\beta}}(-1, 1)$. Suppose that the first m Fourier coefficients of $f \in \mathrm{U}$ are given. Then there exists a $c > 0$ independent of m, $n = (n_0, \ldots, n_l)$ and f such that, whenever $m \geq c\bar{n}^r$, where $\bar{n} = \max\{n_0, \ldots, n_l\}$ and

$$r = \max\left\{ 2, p + 1, 2 + 2\frac{p - q}{q + 1} \right\},$$

$$p = \max\{\alpha_0, \ldots, \alpha_l, \beta_0, \ldots, \beta_l\}, \quad q = \min\{\alpha_0, \ldots, \alpha_l, \beta_0, \ldots, \beta_l\},$$

one may compute, in a completely numerically stable manner, a reconstruction $f_{n,m} \in \mathrm{T}_n$ satisfying

$$\|f - \mathcal{Q}_n f\|_{\alpha,\beta} \leq \|f - f_{n,m}\|_{\alpha,\beta} \leq c' \|f - \mathcal{Q}_n f\|_{\tilde{\alpha}, \tilde{\beta}}, \tag{1.4}$$

for some $c' > 0$ depending only on c. Moreover, excluding a preprocessing step of at most $\mathcal{O}\left(m^2\right)$ cost, the number of operations required to compute $f_{n,m}$ is at worst $\mathcal{O}\left(nm\right)$. If $w_{\alpha,\beta}$ is a piecewise ultraspherical weight, then such preprocessing involves only $\mathcal{O}\left(nm\right)$ operations.

This theorem guarantees solution of Problem 1.1 for a very large class of functions. In fact, if $-1 < \alpha_r, \beta_r < 1$, $r = 0, \ldots, l$, then we can recover any $f \in \mathrm{U} \equiv \mathrm{H}^R$, where $\mathrm{H}^R = \mathrm{L}^2_{\alpha,\beta}(-1, 1)$. In the general case, note that U, although not equal to H^R, is still a very large space. Therefore it is certainly sufficient for practical purposes. Recall that the principal concern of this paper is piecewise analytic functions. Such functions certainly belong to U, as in fact do any functions which are only piecewise continuous. Of course, higher regularity of f ensures fast convergence of $\mathcal{Q}_n f$, and therefore $f_{n,m}$. However, this theorem demonstrates that (almost) quasi-optimal, stable recovery is always possible regardless of the smoothness of the function being reconstructed.

It is also vitally important in practice to understand the constant c of Theorem 1.4. Fortunately, c can always be computed numerically (see §2.3). Also, empirical evidence suggests that a small value of c is usually suitable. In the numerical experiments in this paper we use $c = \frac{1}{4}$, which gives perfectly acceptable results (see §4). One could also use a smaller value, with the only difference being a slightly larger, but still $\mathcal{O}\left(1\right)$, condition number and constant c' in (1.4).

The proof of Theorem 1.4 relies on careful analysis of the behaviour of Fourier series in certain weighted spaces. Here a connection is made with the Helson–Szegö Theorem on positive measures [27]. Another feature of our analysis is that, in many cases, the best constants corresponding to the quasi-optimality $\|f - f_{n,m}\|_R \asymp \|f - \mathcal{Q}_n f\|_R$ are known, and can be explicitly computed. Specifically, one has the sharp bound

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

where the constant $C_{n,m}$ is bounded (provided $m \geq c\bar{n}^r$) and has a geometrical interpretation in terms of the angle between two particular finite-dimensional subspaces.

We note also the great flexibility guaranteed by Theorem 1.4. In many important cases, specifically whenever $p = q \leq 1$, $m = \mathcal{O}\left(n^2\right)$ samples of a piecewise analytic function allow one to recover close to the best n-term approximation in a given piecewise polynomial system. This includes (piecewise) Legendre ($p = q = 0$) and Chebyshev polynomials ($p = q = \pm\frac{1}{2}$), both of which are commonly used in applications. In particular,

Corollary 1.5. *Given the first m Fourier coefficients of a piecewise analytic function f, one can compute, in at worst $\mathcal{O}(m^{\frac{3}{2}})$ operations, a root-exponentially convergent piecewise polynomial approximation of f consisting of arbitrary piecewise ultraspherical polynomials with parameter $-1 < \gamma < 1$.*

In summary, the method we present in this paper improves on the more conventional approach of spectral reprojection by allowing reconstructions in arbitrary polynomial bases. Having said this, spectral reprojection is formally exponentially convergent, whereas this approach obtains only root-exponential convergence. As we shall discuss in §4.2, however, this rate of exponential convergence may be quite slow in practice, meaning that the approach of this paper, whilst formally less rapidly convergent, actually gives a better numerical approximation in many cases.

Finally, let us make the following remark. For the majority of this paper we shall assume that the discontinuity locations $x_1, \ldots, x_l$ are known exactly. In practice, however, they will usually be unknown, and hence must be computed in advance from the given Fourier data. There are a number of well-developed techniques for doing this (commonly referred to as *edge detection*), such as the method of concentration factors [29, 49]. However, this process always incurs an error, resulting in inexact locations $\tilde{x}_1, \ldots, \tilde{x}_l$. Although this error might be small, it is nonetheless essential that reconstruction methods be robust in its presence. In the final section of the paper, §5, we establish robustness of our approach in precisely this sense.

1.4 Relation to previous work

The special case $H^S = H^R = H$, $\langle \cdot, \cdot \rangle_R = \langle \cdot, \cdot \rangle_S = \langle \cdot, \cdot \rangle$ was first considered in [5, 6]. Therein an abstract framework, known as *generalized sampling (GS)*, was developed to solve Problem 1.2 in this more simple setting. The framework introduced in this paper, which addresses the significantly more general case $H^S \neq H^R$, $\langle \cdot, \cdot \rangle_S \neq \langle \cdot, \cdot \rangle_R$, is a direct extension of this work, and, for this reason, we shall continue to refer to the resulting framework as generalized sampling.

The original GS framework of [5, 6] can be applied successfully to Problem 1.1 whenever $\mathcal{Q}_n f$ is the Legendre polynomial expansion of f (this corresponds to $H^R = H^S = L^2(-1, 1)$), leading to a stable numerical method. This particular instance of GS was also is described [41]. Unfortunately this framework is not as flexible as one may hope. Although one can also reconstruct using arbitrary bases of polynomials, the resulting method is only numerically stable when Legendre polynomials are employed. Moreover, in the ultraspherical case, for example, this method also incurs an increased cost (over the value of $\mathcal{O}(m^{\frac{3}{2}})$ for the Legendre case) proportional to $\mathcal{O}(m^{\frac{3+|\gamma|}{2}})$, where γ is the corresponding polynomial parameter.

Besides spectral reprojection, reconstructions in orthogonal polynomials from Fourier coefficients were considered in the so-called *inverse polynomial reconstruction method* (IPRM) [43, 44]. This approach corresponds to setting $m = n$ in the GS framework we develop in this paper. Unfortunately, such an approach is exponentially ill-conditioned and suffers from the Runge phenomenon [9]. On the other hand, by allowing m to be larger than n, specifically $m = \mathcal{O}(n^2)$, on is able to restore stability and quasi-optimality. Note that such a regularization was first pursued in the context of the IPRM by Hrycak & Gröchenig [41] in the case of Legendre polynomial reconstructions. The extended GS framework we develop in this paper, which exploits the orthogonality of both the sampling and reconstruction bases, allows for stable reconstructions using significantly more general families of polynomials.

Whilst setting $m = n$ leads to unstable reconstructions, the reader may be wondering at this stage whether the scaling $m = \mathcal{O}(n^2)$ in Corollary 1.5, which corresponds to only root-exponential convergence in m, is optimal or not. In [9] it was proved that *any* stable method for recovering a function from its first m Fourier coefficients can converge at best root-exponentially fast in m. In particular, the scaling $m = \mathcal{O}(n^2)$ in generalized sampling is optimal, and any attempt to lower it will necessarily give rise to ill-conditioning. It was also shown in [9] that the scaling $m = \mathcal{O}(n^\tau)$, $1 \leq \tau < 2$, implies ill-conditioning of order $c^{m^{2-\tau}}$ for some $c > 1$. Hence, although this choice allows for formally faster convergence (specifically, exponential with index $1/\tau$) it will also render the reconstruction increasingly susceptible to noise. We refer to [9] for a more detailed discussion of this issue.

2 An extended generalized sampling framework

Let H^S and H^R be subspaces of a vector space V that form Hilbert spaces in their own right with respect to the bilinear forms $\langle \cdot, \cdot \rangle_S$ and $\langle \cdot, \cdot \rangle_R$ respectively. Let $U \subseteq H^R$ be a subspace (not necessarily closed) and assume that $f \in U$ and that $\{\phi_j\}_{j=1}^{\infty} \subseteq U$, where $\{\phi_j\}$ is some orthonormal basis for $(H^R, \langle \cdot, \cdot \rangle_R)$. The

subspace U consists of those functions f that we wish to recover and the subspace $\mathrm{T}_n := \mathrm{span}\{\phi_1, \ldots, \phi_n\}$ is the space in which we seek to reconstruct. If $\{\psi_j\}_{j=1}^{\infty} \subseteq \mathrm{U}$ is an orthonormal basis of $(\mathrm{H}^S, \langle \cdot, \cdot \rangle_S)$, we assume that we have access to the first m samples of $\hat{f}$

$$\hat{f}_j = \langle f, \psi_j \rangle_S, \quad j = 1, \ldots, m, \tag{2.1}$$

with respect to this basis (we shall assume that U is such that these values exist and are finite). In practice, the basis $\{\phi_j\}_{j=1}^{\infty}$ is chosen with some *a priori* knowledge about f. For the applications considered in this paper, where f is a piecewise analytic function on $[-1, 1]$, $\{\phi_j\}_{j=1}^{\infty}$ will consist of orthonormal piecewise polynomials with respect to some weight function and ψ_j will correspond to the complex exponential $\frac{1}{\sqrt{2}} \mathrm{e}^{ij\pi x}$ (in this case we shall enumerate the ψ_j over $\mathbb{Z}$ as opposed to $\mathbb{N}$).

Remark 2.1 The assumption of orthonormality in the sampling and reconstruction vectors $\{\psi_j\}_{j=1}^{\infty}, \{\phi_j\}_{j=1}^{\infty}$ is not necessary. It is reasonably straightforward to relax this condition to that of a Riesz basis or frame [6]. However, since all sampling and reconstruction vectors of interest in this paper will be orthonormal, we shall accept this slight loss of generality.

We seek to solve Problem 1.2: namely, compute an approximation $f_{n,m} \in \mathrm{T}_n$ from the samples (2.1) that is close to $\mathcal{Q}_n f$, where $\mathcal{Q}_n : (\mathrm{H}^R, \langle \cdot, \cdot \rangle_R) \to \mathrm{T}_n$ is the orthogonal projection. Note that $\mathcal{Q}_n f$ is defined by the equations

$$\langle \mathcal{Q}_n f, \phi_j \rangle_R = \langle f, \phi_j \rangle_R, \quad j = 1, \ldots, n, \quad \mathcal{Q}_n f \in \mathrm{T}_n. \tag{2.2}$$

Had we had access to the values $\langle f, \phi_j \rangle_R$, i.e. the coefficients of f in a particular polynomial basis, for example, we could have computed $\mathcal{Q}_n f$ immediately. However, in general we do not have this information. Rather, all we know is the values $\{\hat{f}_j\}_{j=1}^{m}$. Since the sampling basis $\{\psi_j\}_{j\in\mathbb{N}}$ is considered fixed in the setup of this paper, and cannot be altered, we are consequently unable to form $\mathcal{Q}_n f$ from the given data.

Nonetheless, let us write $\mathcal{P}_m : \mathrm{H}^S \to \mathrm{S}_m := \mathrm{span}\{\psi_1, \ldots, \psi_m\}$ for orthogonal projection onto $\mathrm{S}_m \subseteq \mathrm{U}$ with respect to $\langle \cdot, \cdot \rangle_S$, i.e.

$$\mathcal{P}_m g = \sum_{j=1}^{m} \langle g, \psi_j \rangle_S \psi_j, \quad g \in \mathrm{H}^S.$$

We now define $f_{n,m} \in \mathrm{T}_n$ as the solution to

$$\langle \mathcal{P}_m f_{n,m}, \mathcal{P}_m \phi_j \rangle_R = \langle \mathcal{P}_m f, \mathcal{P}_m \phi_j \rangle_R, \quad j = 1, \ldots, n, \quad f_{n,m} \in \mathrm{T}_n, \tag{2.3}$$

or equivalently (by linearity),

$$\langle \mathcal{P}_m f_{n,m}, \mathcal{P}_m \phi \rangle_R = \langle \mathcal{P}_m f, \mathcal{P}_m \phi \rangle_R, \quad \forall \phi \in \mathrm{T}_n, \quad f_{n,m} \in \mathrm{T}_n. \tag{2.4}$$

Suppose for a moment that $(\mathrm{H}^S, \langle \cdot, \cdot \rangle_S)$ and $(\mathrm{H}^R, \langle \cdot, \cdot \rangle_R)$ coincide. This is precisely the setting of the original GS framework of [6]. It is now quite simple to give an intuitive explanation as to how $f_{n,m}$ solves Problem 1.2. Indeed, since $\{\psi_j\}_{j=1}^{\infty}$ is an orthonormal basis for $(\mathrm{H}^S, \langle \cdot, \cdot \rangle_S)$, we have that $\mathcal{P}_m \to \mathcal{I}$ in the strong operator topology on $(\mathrm{H}^S, \langle \cdot, \cdot \rangle_S)$, where $\mathcal{I} : \mathrm{H}^S \to \mathrm{H}^S$ is the identity operator. Thus, for fixed $n \in \mathbb{N}$, if $m \to \infty$ the equations (2.3) resemble those equations (2.2) defining $\mathcal{Q}_n f$. Therefore, for sufficiently large m, we expect $f_{n,m}$ to inherit the key features of $\mathcal{Q}_n f$. In particular, $f_{n,m}$ should exist uniquely, and critically, possess the same approximation properties.

In the general case $(\mathrm{H}^S, \langle \cdot, \cdot \rangle_S) \neq (\mathrm{H}^R, \langle \cdot, \cdot \rangle_R)$, we need the additional assumption that $\mathcal{P}_m \to \mathcal{I}$ strongly on $(\mathrm{U}, \langle \cdot, \cdot \rangle_R)$:

$$\|g - \mathcal{P}_m g\|_R \to 0, \quad m \to \infty, \quad \forall g \in \mathrm{U}. \tag{A1}$$

Note that this assumption is not guaranteed in general, and must be verified for each particular problem. However, as we shall see in §3, it does hold in all cases of interest in this paper. We will also require one further assumption: namely, a uniform boundedness condition for $\{\mathcal{P}_m\}_{m\in\mathbb{N}}$ in terms of some norm $\|\cdot\|$ defined on U:

$$\|\mathcal{P}_m g\|_R \leq c \|g\|, \quad \forall g \in \mathrm{U}, \ m \in \mathbb{N}, \tag{A2}$$

where $c > 0$ is independent of m and g.

In some circumstances, but not all, it transpires that the following stronger version of (A2) holds:

$$\mathrm{U} = \mathrm{H}^R, \quad \mathcal{P}_m : (\mathrm{H}^R, \langle \cdot, \cdot \rangle_R) \to (\mathrm{H}^R, \langle \cdot, \cdot \rangle_R) \text{ bounded.} \tag{A2'}$$

To see that this is stronger than (A2), we note:

Lemma 2.2. *Suppose that (A1) and (A2′) hold. Then the family $\{\mathcal{P}_m\}_{m \in \mathbb{N}}$ is uniformly bounded. In other words, (A2) holds with $\|\!|\cdot|\!\| = \|\cdot\|_R$.*

Proof. (A1) and (A2′) imply that $\sup_{m \in \mathbb{N}} \|\mathcal{P}_m g\|_R < \infty$ for any $g \in \mathrm{U} = \mathrm{H}^R$. The uniform boundedness principle now gives the result. ☐

Whenever (A2′) holds instead of (A2) we may give a more precise analysis of the approximation $f_{n,m}$ (see §2.2). In §3 we detail the situations in which (A2′) holds as opposed to (A2) for the Fourier reconstruction problem.

2.1 Recovery from Fourier coefficients

Before analysing this framework, let us first discuss the main example of this paper. Suppose that $f : [-1,1] \to \mathbb{R}$ is piecewise analytic with jumps at $-1 < x_1 < \ldots < x_l < 1$. Let $\psi_j(x) = \frac{1}{\sqrt{2}} \exp(\mathrm{i}j\pi x)$, $j \in \mathbb{Z}$, and assume that the coefficients

$$\hat{f}_j = \int_{-1}^1 f(x)\overline{\psi_j}(x)\,\mathrm{d}x, \quad j = -\left\lfloor \frac{m}{2} \right\rfloor, \ldots, \left\lfloor \frac{m}{2} \right\rfloor, \tag{2.5}$$

are given. In particular, $\mathcal{P}_m f$ is the truncated Fourier series of f (we henceforth refer to $\mathcal{P}_m$ as the Fourier projection operator in this case). We wish to recover f in a basis of orthonormal piecewise polynomials corresponding to the piecewise (modified) Jacobi weight $w_{\alpha,\beta}$ given by (1.3), where $\alpha = \{\alpha_0, \ldots, \alpha_l\}$ and $\beta = \{\beta_0, \ldots, \beta_l\}$. To this end, let $x_0 = -1$, $x_{l+1} = 1$ and write

$$\mathrm{T}_n = \{\phi : \phi|_{I_r} \in \mathbb{P}_{n_r} : r = 0, \ldots, l\}, \quad n = (n_0, \ldots, n_l), \tag{2.6}$$

where $I_r = (x_r, x_{r+1})$. We now compute $f_{n,m} \in \mathrm{T}_n$ using (2.4). Thus, we let $\mathrm{V} = \mathrm{L}^1(-1,1)$ (the space of absolutely integrable functions), $\mathrm{H}^R = \mathrm{L}^2_{\alpha,\beta}(-1,1)$, $\|\cdot\|_R = \|\cdot\|_{\alpha,\beta}$, and set $\mathrm{H}^S = \mathrm{L}^2(-1,1)$, $\|\cdot\|_S = \|\cdot\|$.

There are now three key questions:

(i) How do we choose the subspace U and the norm $\|\!|\cdot|\!\|$ so that (A1) and (A2) hold?
(ii) Does such a choice of U include all functions of interest? Specifically, U should at the very least contain all piecewise analytic functions.
(iii) In what circumstances does (A2′) hold instead of (A2)?

In §3 we provide answers to these questions.

2.2 Analysis of the extended framework

We now present two analyses of the generalized sampling framework introduced above. The first, for which the corresponding error estimates (i.e. bounds for $\|f - f_{n,m}\|_R$) are not necessarily sharp, is valid whenever (A1) and (A2) hold. The second, which assumes the stronger condition (A2′), leads to sharp bounds.

2.2.1 Version I

The definition (2.4) of $f_{n,m}$ motivates the introduction of the quantity

$$E_{n,m} = \sup\{\|\phi - \mathcal{P}_m \phi\|_R : \phi \in \mathrm{T}_n, \|\phi\|_R = 1\}. \tag{2.7}$$

This quantity measures how close the restriction $\mathcal{P}_m|_{\mathrm{T}_n}$ is to the identity operator $\mathcal{I}|_{\mathrm{T}_n}$. Since (A1) and (A2) hold and T_n is finite-dimensional, the following lemma comes as no surprise:

Lemma 2.3. *Suppose that (A1) and (A2) hold and let $E_{n,m}$ be defined by (2.7). Then $E_{n,m} \leq 1 + c d_n$ for all $m, n \in \mathbb{N}$, where c is as in (A2) and*

$$d_n = \sup\{\|\!|\phi|\!\| : \phi \in \mathrm{T}_n, \|\phi\|_R = 1\}. \tag{2.8}$$

Moreover, $E_{n,m} \to 0$ as $m \to \infty$ for fixed n.

Proof. Note that $E_{n,m} \leq 1 + \sup\{\|\mathcal{P}_m \phi\|_R : \phi \in \mathrm{T}_n, \|\phi\|_R = 1\}$. (A2) now gives the first result. For the second, observe that the set $\mathrm{B}_n := \{\phi \in \mathrm{T}_n : \|\phi\|_R = 1\}$ is compact. Since $\mathcal{P}_m \to \mathcal{I}$ strongly on $\mathrm{U} \supseteq \mathrm{B}_n$, this convergence is uniform on B_n. ☐

With this to hand, we are now able to prove the main result of this section:

Theorem 2.4. *Suppose that (A1) and (A2) hold. Then for each $n \in \mathbb{N}$ there exists an m_0 such that for $m \geq m_0$ the approximation $f_{n,m}$ defined by (2.4) exists and is unique and satisfies the estimates*

$$\|f_{n,m}\|_R \leq \frac{c}{1 - E_{n,m}} \|f\|, \tag{2.9}$$

and

$$\|f - f_{n,m}\|_R \leq \|f - \mathcal{Q}_n f\|_R + \frac{cE_{n,m}}{(1 - E_{n,m})^2} \|f - \mathcal{Q}_n f\| + \frac{1}{(1 - E_{n,m})^2} \|(\mathcal{I} - \mathcal{P}_m)(f - \mathcal{Q}_n f)\|_R, \tag{2.10}$$

where $\mathcal{Q}_n : \mathrm{H}^R \to \mathrm{T}_n$ is the orthogonal projection with respect to $\langle \cdot, \cdot \rangle_R$, $E_{n,m}$ is as in (2.7) and c is the constant of (A2). Specifically, m_0 is the least m such that $E_{n,m} < 1$.

Note that (2.9) is a continuous stabilty estimate for $f_{n,m}$. In particular, this implies that the coefficients of $f_{n,m}$ cannot grow large.

Proof. Let $\mathcal{U} : \mathrm{T}_n \to \mathbb{C}^n$ be the linear operator $g \mapsto \{\langle \mathcal{P}_m g, \mathcal{P}_m \phi_j \rangle_R\}_{j=1}^n$. It suffices to show that $\mathcal{U}$ is invertible; in other words, $\ker(\mathcal{U}) = \{0\}$. Suppose that $g \in \mathrm{T}_n$ with $\mathcal{U}g = 0$. Then, by linearity, $\langle \mathcal{P}_m g, \mathcal{P}_m \phi \rangle_R = 0, \forall \phi \in \mathrm{T}_n$. In particular, $0 = \|\mathcal{P}_m g\|_R \geq (1 - E_{n,m})\|g\|_R$. Thus, if $m \geq m_0$, where m_0 is the least m such that $E_{n,m} < 1$, then we must have that $g = 0$. Hence, $\mathcal{U}$ is nonsingular, and $f_{n,m}$ exists uniquely.

Now consider (2.9). Setting $\phi = f_{n,m}$ in (2.4) gives

$$\|\mathcal{P}_m f_{n,m}\|_R^2 = \langle \mathcal{P}_m f, \mathcal{P}_m f_{n,m} \rangle_R \leq \|\mathcal{P}_m f\|_R \|\mathcal{P}_m f_{n,m}\|_R.$$

The inequality (2.9) now follows from (A2) and the fact that $\|\mathcal{P}_m f_{n,m}\|_R \geq (1 - E_{n,m})\|f_{n,m}\|_R$.

For (2.10) we first write $e = f_{n,m} - \mathcal{Q}_n f \in \mathrm{T}_n$ and notice that

$$\|\mathcal{P}_m e\|_R^2 = \langle \mathcal{P}_m(f - \mathcal{Q}_n f), \mathcal{P}_m e \rangle_R.$$

Since $e \in \mathrm{T}_n$ and $f - \mathcal{Q}_n f \perp \mathrm{T}_n$ is the orthogonal projection, the right hand side may be written as

$$\langle \mathcal{P}_m(f - \mathcal{Q}_n f), \mathcal{P}_m e \rangle_R = -\langle (\mathcal{I} - \mathcal{P}_m)(f - \mathcal{Q}_n f), e \rangle_R - \langle \mathcal{P}_m(f - \mathcal{Q}_n f), (\mathcal{I} - \mathcal{P}_m)e \rangle_R.$$

This gives

$$(1 - E_{n,m})^2 \|e\|_R^2 \leq \|\mathcal{P}_m e\|_R^2 \leq \|(\mathcal{I} - \mathcal{P}_m)(f - \mathcal{Q}_n f)\|_R \|e\|_R + \|\mathcal{P}_m(f - \mathcal{Q}_n f)\|_R \|(\mathcal{I} - \mathcal{P}_m)e\|_R.$$

Hence, by (A2) and the definition of $E_{n,m}$, we obtain

$$(1 - E_{n,m})^2 \|e\|_R \leq \|(\mathcal{I} - \mathcal{P}_m)(f - \mathcal{Q}_n f)\|_R + cE_{n,m}\|f - \mathcal{Q}_n f\|. \tag{2.11}$$

The full result follows from the inequality $\|f - f_{n,m}\|_R \leq \|e\|_R + \|f - \mathcal{Q}_n f\|_R$. $\qquad \square$

This theorem confirms solution of Problem 1.2 using (2.4) whenever (A1) and (A2) hold, and provided m is sufficiently large in comparison to n. The question of how large will be discussed and quantified in §2.3. Observe that

$$\|g - \mathcal{P}_m g\|_R \leq \|g\|_R + c\|g\|, \quad \forall g \in \mathrm{U}, \quad \forall m \in \mathbb{N},$$

hence the term $\|(\mathcal{I} - \mathcal{P}_m)(f - \mathcal{Q}_n f)\|_R$ in right-hand side of (2.10) can be easily bounded independently of m to give the more convenient bound

$$\|f - f_{n,m}\|_R \leq \left[1 + \frac{1}{(1 - E_{n,m})^2}\right] \|f - \mathcal{Q}_n f\|_R + \frac{c(1 + E_{n,m})}{(1 - E_{n,m})^2} \|f - \mathcal{Q}_n f\|. \tag{2.12}$$

However, this bound obscures one interesting facet of $f_{n,m}$: namely, its *asymptotic optimality*. That is, by increasing the number of samples m, $f_{n,m}$ can be made arbitrarily close to the best approximation $\mathcal{Q}_n f$. Indeed,

Corollary 2.5. *Under the assumptions of Theorem 2.4, $f_{n,m} \to \mathcal{Q}_n f$ as $m \to \infty$ for fixed $n \in \mathbb{N}$.*

Proof. Since $E_{n,m} \to 0$ as $m \to \infty$ (Lemma 2.3), we have by (2.11),

$$\|f_{n,m} - \mathcal{Q}_n f\|_R \leq \frac{cE_{n,m}}{(1 - E_{n,m})^2} \|f - \mathcal{Q}_n f\| + \frac{1}{(1 - E_{n,m})^2} \|(\mathcal{I} - \mathcal{P}_m)(f - \mathcal{Q}_n f)\|_R \to 0,$$

as $m \to \infty$. Thus, $f_{n,m} \to \mathcal{Q}_n f$ as required. $\qquad\square$

The analysis of this section demonstrates (almost) quasi- and asymptotic optimality of $f_{n,m}$. However, it is clearly less than ideal to have an error estimate of the form (2.10) involving $\|f - \mathcal{Q}_n f\|$. In general, there is no guarantee that this term decays as $n \to \infty$ (although this is always true in the applications of this paper). Thus, one may ask: can a better analysis give an error bound involving only $\|f - \mathcal{Q}_n f\|_R$ (which must tend to zero, since $\mathcal{Q}_n : \mathrm{H}^R \to \mathrm{T}_n$ is the orthogonal projection)? As it transpires, whenever (A2$'$) holds instead of (A2), this is indeed the case. This is described in the next section. The resulting analysis also provides bounds for $\|f_{n,m}\|_R$ and $\|f - f_{n,m}\|_R$ which, unlike those given in Theorem 2.4, are sharp.

2.2.2 Version II

Suppose now that (A2$'$) holds instead of (A2). To derive an improved analysis, we shall proceed in a similar manner to the analysis of the original GS framework given in [8]. First, let us introduce some basic concepts from the geometry of Hilbert spaces (see [47, 48, 51] for further details):

Definition 2.6. *Let* U *and* V *be closed subspaces of a Hilbert space* H *and* $\mathcal{Q}_V : \mathrm{H} \to \mathrm{V}$ *the orthogonal projection onto* V. *The subspace angle* $\theta = \theta_{\mathrm{UV}} \in [0, \frac{\pi}{2}]$ *between* U *and* V *is given by*

$$\cos(\theta_{\mathrm{UV}}) = \inf \left\{ \|\mathcal{Q}_V u\| : u \in \mathrm{U}, \ \|u\| = 1 \right\}.$$

We are interested in conditions on U and V which imply that $\cos(\theta_{\mathrm{UV}})$ is nonzero. The following lemma (see [51, Thm. 2.1]) gives a full characterization:

Lemma 2.7. *Let* U *and* V *be closed subspaces of a Hilbert space* H. *Then* $\cos(\theta_{\mathrm{UV}^\perp}) > 0$ *if and only if* $\mathrm{U} \cap \mathrm{V} = \{0\}$ *and* $\mathrm{U} + \mathrm{V}$ *is closed* H.

Subspaces U and V for which $\cos(\theta_{\mathrm{UV}^\perp}) > 0$ give rise to a decomposition $\mathrm{U} \oplus \mathrm{V} = \mathrm{H}_0$ of a closed subspace H_0 of H. This ensures the existence of a projection of H_0 with range U and kernel V. We refer to such a projection as an *oblique projection* and denote it by $\mathcal{W}_{\mathrm{UV}}$. Note that $\mathcal{W}_{\mathrm{UV}}$ will not, in general, be defined over the whole of H, but rather the subspace H_0. However, this is true whenever $\mathrm{V} = \mathrm{U}^\perp$, for example, and in this case the projection $\mathcal{W}_{\mathrm{UV}}$ coincides with the orthogonal projection $\mathcal{Q}_{\mathrm{U}}$. Moreover,

Lemma 2.8 ([8]). *Suppose that* U *and* V *are closed subspaces of* H *satisfying* $\cos(\theta_{\mathrm{UV}^\perp}) > 0$. *Suppose also that* $\dim(\mathrm{U}) = \dim(\mathrm{V}^\perp) = n < \infty$. *Then* $\mathrm{U} \oplus \mathrm{V} = \mathrm{H}$.

We shall also require the following result:

Theorem 2.9 ([8]). *Suppose that* U *and* V *are closed subspaces of* H *satisfying* $\cos(\theta_{\mathrm{UV}^\perp}) > 0$, *and let* $\mathcal{W}_{\mathrm{UV}} : \mathrm{H}_0 \to \mathrm{U}$ *be the oblique projection with range* U *and kernel* V, *where* $\mathrm{H}_0 = \mathrm{U} \oplus \mathrm{V}$. *Then*

$$\|\mathcal{W}_{\mathrm{UV}} f\| \leq \sec(\theta_{\mathrm{UV}^\perp}) \|f\|, \quad \forall f \in \mathrm{H}_0, \tag{2.13}$$

and

$$\|f - \mathcal{Q}_{\mathrm{U}} f\| \leq \|f - \mathcal{W}_{\mathrm{UV}} f\| \leq \sec(\theta_{\mathrm{UV}^\perp}) \|f - \mathcal{Q}_{\mathrm{U}} f\|, \quad \forall f \in \mathrm{H}_0, \tag{2.14}$$

where $\mathcal{Q}_{\mathrm{U}} : \mathrm{H} \to \mathrm{U}$ *is the orthogonal projection. Moreover, the upper bounds in* (2.13) *and* (2.14) *are sharp.*

We now return to the problem at hand. Note that since (A2$'$) holds, the operator $\mathcal{P}_m$ is bounded on H^R. Hence it has a bounded adjoint $\mathcal{P}_m^* : (\mathrm{H}^R, \langle \cdot, \cdot \rangle) \to (\mathrm{H}^R, \langle \cdot, \cdot \rangle)$. Write $\mathcal{W}_m := \mathcal{P}_m^* \circ \mathcal{P}_m$. We now require the following lemma:

Lemma 2.10. *Assume that* (A1) *and* (A2$'$) *hold. Then the operators* $\mathcal{W}_m$ *converge strongly to the identity* $\mathcal{I}$ *on* $(\mathrm{H}^R, \langle \cdot, \cdot \rangle_R)$.

Proof. For $g \in \mathrm{H}^R$, we have

$$\|g - \mathcal{W}_m g\|^2 = \langle g - \mathcal{P}_m^* \mathcal{P}_m g, g - \mathcal{P}_m^* \mathcal{P}_m g \rangle$$
$$\leq |\langle \mathcal{P}_m^*(g - \mathcal{P}_m g), g - \mathcal{P}_m^* \mathcal{P}_m g \rangle| + |\langle g - \mathcal{P}_m^* g, g - \mathcal{P}_m^* \mathcal{P}_m g \rangle|.$$

Since $\{\mathcal{P}_m\}_{m \in \mathbb{N}}$ is uniformly bounded, so is $\{\mathcal{P}_m^*\}_{m \in \mathbb{N}}$. Moreover, the adjoint operation is strongly continuous on bounded sets. Hence $\mathcal{P}_m^* \to \mathcal{I}$ strongly on H^R and we obtain, for some $c > 0$ independent of g and m,

$$\|g - \mathcal{W}_m g\|^2 \leq c \left(\|g - \mathcal{P}_m g\| + \|g - \mathcal{P}_m^* g\| \right) \|g\| \to 0, \quad m \to \infty,$$

as required. $\qquad\square$

We are now ready to analyze $f_{n,m}$. First, note that we may rewrite (2.4) as

$$\langle f_{n,m}, \mathcal{W}_m \phi \rangle_R = \langle f, \mathcal{W}_m \phi \rangle_R, \quad \forall \phi \in \mathrm{T}_n.$$

In other words, the mapping $f \mapsto f_{n,m}$ (whenever defined) is precisely the oblique projection with range T_n and kernel $[\mathcal{W}_m(\mathrm{T}_n)]^{\perp}$ (here $\perp$ is taken with respect to $(\mathrm{H}^R, \langle \cdot, \cdot \rangle_R)$). Letting

$$C_{n,m} = \cos\left(\theta_{\mathrm{T}_n, \mathcal{W}_m(\mathrm{T}_n)} \right), \tag{2.15}$$

we now deduce the following:

Theorem 2.11. *Suppose that (A1) and (A2$'$) hold. Let $n, m \in \mathbb{N}$ be such that $C_{n,m} > 0$, where $C_{n,m}$ is given by (2.15). Then $f_{n,m}$, as defined by (2.4), exists uniquely for any $f \in \mathrm{H}^R$ and satisfies the sharp bounds*

$$\|f_{n,m}\|_R \leq \frac{1}{C_{n,m}} \|f\|_R, \tag{2.16}$$

and

$$\|f - f_{n,m}\|_R \leq \frac{1}{C_{n,m}} \|f - \mathcal{Q}_n f\|_R. \tag{2.17}$$

Proof. The result follows immediately from Theorem 2.9, provided $\mathrm{H} = \mathrm{H}_0 = \mathrm{T}_n \oplus [\mathcal{W}_m(\mathrm{T}_n)]^{\perp}$. By Lemma 2.8 it suffices to show that $\dim(\mathcal{W}_m(\mathrm{T}_n)) = \dim(\mathrm{T}_n) = n$. By definition $\dim(\mathcal{W}_m(\mathrm{T}_n)) \leq n$. Suppose that $\dim(\mathcal{W}_m(\mathrm{T}_n)) < n$. Then there exists a nonzero $\phi \in \mathrm{T}_n$ such that $\mathcal{W}_m \phi = 0$. Since $\mathcal{W}_m$ is self-adjoint, it follows that $\phi \in \mathrm{T}_n \cap [\mathcal{W}_m(\mathrm{T}_n)]^{\perp}$. Hence $C_{n,m} = 0$, a contradiction. $\qquad\square$

Naturally, to use this theorem we need to understand the quantity $C_{n,m}$. We have

Lemma 2.12. *Suppose that (A1) and (A2$'$) hold, and let $C_{n,m}$ be given by (2.15). Then*

$$c^{-1}(1 - E_{n,m}) \leq C_{n,m} \leq 1, \quad \forall n, m \in \mathbb{N}, \tag{2.18}$$

where $c = \sup_{m \in \mathbb{N}} \|\mathcal{P}_m\|_R$ and $\|\cdot\|_R$ in this case is the operator norm on the space of bounded linear maps $(\mathrm{H}^R, \langle \cdot, \cdot \rangle_R) \to (\mathrm{H}^R, \langle \cdot, \cdot \rangle_R)$. Moreover, $C_{n,m} \to 1$ as $m \to \infty$ for fixed n.

Proof. That $C_{n,m} \leq 1$ follows immediately from the definition (2.15). By the definition of subspace angles

$$C_{n,m} = \inf \left\{ \|\mathcal{Q}_{\mathcal{W}_m(\mathrm{T}_n)} \phi\|_R : \phi \in \mathrm{T}_n, \|\phi\|_R = 1 \right\}.$$

Consider the quantity $\|\mathcal{Q}_{\mathcal{W}_m(\mathrm{T}_n)} \phi\|_R$. By the standard duality pairing,

$$\|\mathcal{Q}_{\mathcal{W}_m(\mathrm{T}_n)} \phi\|_R = \sup \left\{ \frac{\langle \mathcal{Q}_{\mathcal{W}_m(\mathrm{T}_n)} \phi, \mathcal{W}_m \phi' \rangle_R}{\|\mathcal{W}_m \phi'\|_R} : \phi' \in \mathrm{T}_n, \phi' \neq 0 \right\}$$
$$= \sup \left\{ \frac{\langle \phi, \mathcal{W}_m \phi' \rangle_R}{\|\mathcal{W}_m \phi'\|_R} : \phi' \in \mathrm{T}_n, \phi' \neq 0 \right\}. \tag{2.19}$$

Setting $\phi' = \phi$, we obtain

$$\|\mathcal{Q}_{\mathcal{W}_m(\mathrm{T}_n)} \phi\|_R \geq \frac{\langle \phi, \mathcal{W}_m \phi \rangle_R}{\|\mathcal{W}_m \phi\|_R} = \frac{\|\mathcal{P}_m \phi\|_R^2}{\|\mathcal{W}_m \phi\|_R}.$$

Since $\mathcal{W}_m = \mathcal{P}_m^* \circ \mathcal{P}_m$ we have $\|\mathcal{W}_m \phi\|_R \leq c \|\mathcal{P}_m \phi\|_R$, and therefore

$$\|\mathcal{Q}_{\mathcal{W}_m(\mathrm{T}_n)} \phi\|_R \geq c^{-1} \|\mathcal{P}_m \phi\|_R \geq c^{-1} (1 - E_{n,m}) \|\phi\|,$$

which gives (2.18). The proof that $C_{n,m} \to 1$ as $m \to \infty$ is analogous to that of Lemma 2.3, and follows directly from the fact that $\mathcal{W}_m \to \mathcal{I}$ (Lemma 2.10). $\qquad\square$

The importance of (2.18) is that it allows $C_{n,m}$ to be estimated in terms of $E_{n,m}$; a somewhat easier task in practice. Thus, estimates obtained later for the asymptotic behaviour of $E_{n,m}$ automatically hold for $C_{n,m}$, and thus guarantee the conclusion of Theorem 2.11. We shall exploit this fact in §3. Note also that Lemma 2.12 implies asymptotic optimality of $f_{n,m}$. Since

$$\|f_{n,m} - \mathcal{Q}_n f\|^2 = \|f - f_{n,m}\|^2 - \|f - \mathcal{Q}_n f\|^2 \leq K_{n,m}\|f - \mathcal{Q}_n f\|^2 \leq K_{n,m}\|f\|^2,$$

where $K_{n,m} = \frac{1}{C_{n,m}^2} - 1$, we have $f_{n,m} \to \mathcal{Q}_n f$ as $m \to \infty$.

It is instructive to now relate this general framework to the example of §2.1. Suppose that we sample an analytic, nonperiodic f via its Fourier coefficients (so that $\mathcal{P}_m$ is the standard Fourier projection operator) and wish to reconstruct in Chebyshev polynomials. We thus let $V = L^1(-1,1)$, $H^S = L^2(-1,1)$ and $H^R = L_w^2(-1,1)$, where $w(x) = (1 - x^2)^{-\frac{1}{2}}$ is the Chebyshev weight on $[-1,1]$. It transpires that both (A1) and (A2′) hold in this case (see §3). Therefore Theorem 2.11 and Lemma 2.12 apply to this example. Note that in this case the operators $\mathcal{P}_m^*$ and $\mathcal{W}_m$ are given explicitly by

$$\mathcal{P}_m^* g = \frac{1}{w} \mathcal{P}_m(wg), \quad \mathcal{W}_m g = \frac{1}{w} \mathcal{P}_m\left(w \mathcal{P}_m g\right), \quad g \in L_w^2(-1,1). \tag{2.20}$$

2.3 Numerical implementation

Let us now consider the computation of $f_{n,m}$. Writing $f_{n,m} = \sum_{j=1}^n \alpha_j \phi_j$ for unknowns $\alpha_1, \ldots, \alpha_n$, it is easily seen that the vector $\alpha = (\alpha_1, \ldots, \alpha_n)$ is defined by

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

where $\hat{f} = \{\langle f, \psi_1 \rangle_S, \ldots, \langle f, \psi_m \rangle_S\}$, and $U \in \mathbb{C}^{m \times n}$ and $C \in \mathbb{C}^{m \times m}$ have $(j,k)^{\text{th}}$ entries $\langle \phi_k, \psi_j \rangle_S$ and $\langle \psi_j, \psi_k \rangle_R$ respectively. Thus the coefficients of $f_{n,m}$ are defined by an $n \times n$ linear system of equations with self-adjoint matrix $A = U^* C U$.

The condition number $\kappa(A)$ of the matrix A is critical from a numerical perspective. It determines both the stability of the numerical method—its susceptibility to noise and round-off errors, in particular—as well as the cost of any iterative solver, e.g. conjugate gradients, for computing the vector of unknowns α. In such case, the number of iterations required to compute α is proportional to $\sqrt{\kappa(A)}$. Thus, if $\kappa(A) = \mathcal{O}(1)$ for all n and sufficiently large m (which turns out to be the case), then the cost of computing $f_{n,m}$ is determined solely by the cost of performing matrix-vector multiplications involving A. Since the (typically dense) matrices U and C are of size $m \times n$ and $m \times m$ respectively, straightforward implementation of this method, that is, without any fast transforms, requires $\mathcal{O}(m^2)$ operations. However, as we shall explain in §4, C always has a Toeplitz structure when one considers Fourier samples. Thus, matrix-vector multiplications involving C require only $\mathcal{O}(m \log m)$ operations, giving a total figure of $\mathcal{O}(mn)$ for this method.

We now give an estimate for $\kappa(A)$:

Lemma 2.13. *Suppose that (A1) and (A2) hold and let $E_{n,m}$ be given by (2.7). Then the condition number of $A = U^* C U$ satisfies*

$$\kappa(A) \leq \left(\frac{1 + E_{n,m}}{1 - E_{n,m}} \right)^2.$$

Moreover, if n is fixed then $A \to I$ as $m \to \infty$, where $I \in \mathbb{C}^{n \times n}$ is the identity. In particular, $\kappa(A) \to 1$.

Proof. Let $\alpha = \{\alpha_1, \ldots, \alpha_n\} \in \mathbb{C}^n$ be a normalized eigenvector of A with corresponding eigenvalue λ. If $\phi = \sum_{j=1}^n \alpha_j \phi_j$ it follows that

$$\langle \mathcal{P}_m \phi, \mathcal{P}_m \phi_j \rangle_R = \lambda \alpha_j, \quad j = 1, \ldots, n,$$

and, in particular, $\|\mathcal{P}_m \phi\|_R^2 = \lambda$. Hence,

$$\lambda_{\min}(A) = \inf \left\{ \|\mathcal{P}_m \phi\|_R^2 : \phi \in T_n, \|\phi\|_R = 1 \right\}, \quad \lambda_{\max}(A) = \sup \left\{ \|\mathcal{P}_m \phi\|_R^2 : \phi \in T_n, \|\phi\|_R = 1 \right\}.$$

Using the definition $E_{n,m}$ we deduce that

$$(1 - E_{n,m})^2 \leq \lambda_{\min}(A) \leq \lambda_{\max}(A) \leq (1 + E_{n,m})^2,$$

which gives the first result. For the second, we merely note that $A \to I$ componentwise as $m \to \infty$ (since $\mathcal{P}_m \to \mathcal{I}$ strongly on U), and therefore in norm. $\qquad\square$

This lemma confirms good conditioning of A whenever m is sufficiently large in comparison to n.

Remark 2.14 The extended GS framework introduced in this section can actually be reinterpreted as a preconditioning of the original GS framework of [6]. In the original framework, one solves the linear system $U^*U\alpha = U^*\hat{f}$. However, unless $(\mathrm{H}^R, \langle\cdot,\cdot\rangle_R) = (\mathrm{H}^S, \langle\cdot,\cdot\rangle_S)$ the matrix U^*U is not well-conditioned, leading to a higher cost in computing α. The extended framework merely replaces this with the modified equations $U^*CU\alpha = U^*C\hat{f}$, where C is as above. This judicious choice of C preconditions the original equations by exploiting the orthogonality of both the sampling and reconstruction vectors, and therefore leads to the aforementioned improvement.

2.4 Scaling of m with n: the stable sampling rate

Observe that the same quantity $E_{n,m}$ which determines the magnitude of the error committed by $f_{n,m}$ also arises in the estimate for the condition number. Indeed, provided m is chosen so that $1 - E_{n,m}$ is bounded away from zero, we can ensure both quasi-optimality of $f_{n,m}$ and numerical stability. This motivates the following definition. Let

$$\Theta(n;\theta) = \min\left\{m \in \mathbb{N} : E_{n,m} < 1 - \theta\right\}, \quad \theta \in (0,1). \tag{2.21}$$

For a given n, choosing $m \geq \Theta(n;\theta)$ ensures that

$$\|f - f_{n,m}\|_R \leq \left(1 + \theta^{-2}\right)\|f - \mathcal{Q}_n f\|_R + c(2-\theta)\theta^{-2}\|f - \mathcal{Q}_n f\|, \tag{2.22}$$

and that

$$\kappa(A) \leq \left(2\theta^{-1} - 1\right)^2,$$

whenever (A1) and (A2) hold. In other words, $f_{n,m}$ is quasi-optimal to f from T_n, and its computation is numerically stable, uniformly in $n \in \mathbb{N}$.

The quantity $\Theta(n;\theta)$, originally introduced in [6] and named the *stable sampling rate* in [8], measures how large m (the number of samples) must be for a given n for stable, quasi-optimal reconstruction. Importantly, $\Theta(n;\theta)$ is computable. Indeed,

Lemma 2.15. *Let $E_{n,m}$ be given by (2.7). Then $E_{n,m} = \sqrt{\|B\|}$, where $B \in \mathbb{C}^{n \times n}$ is the Hermitian matrix with $(j,k)^{\mathrm{th}}$ entry $\langle \phi_j - \mathcal{P}_m \phi_j, \phi_k - \mathcal{P}_m \phi_k \rangle_R$.*

Proof. For $\phi \in \mathrm{T}_n$, write $\phi = \sum_{j=1}^n \alpha_j \phi_j$. Then

$$\begin{aligned}
E_{n,m}^2 &= \sup\left\{\frac{\sum_{j,k=1}^n \alpha_j\overline{\alpha_k}\langle\phi_j - \mathcal{P}_m\phi_j, \phi_k - \mathcal{P}_m\phi_k\rangle_R}{\sum_{j,k=1}^n \alpha_j\overline{\alpha_k}\langle\phi_j, \phi_k\rangle_R} : \alpha \in \mathbb{C}^n,\ \alpha \neq 0\right\} \\
&= \sup\left\{\frac{\alpha^*B\alpha}{\alpha^*\alpha} : \alpha \in \mathbb{C}^n,\ \alpha \neq 0\right\}.
\end{aligned}$$

This is precisely $\|B\|$. $\qquad\square$

Note that the matrix B can be expressed as $B = I - V^*U - U^*V + A$, where $A = U^*CU$ and $V \in \mathbb{C}^{m \times n}$ has $(j,k)^{\mathrm{th}}$ entry $\langle\phi_k, \psi_j\rangle_R$. Hence, given U, C and V, one can readily compute $E_{n,m}$ and therefore $\Theta(n;\theta)$.

In the case where (A2) is replaced by the stronger condition (A2′) it is preferable to replace the bound (2.22) by the improved estimate

$$\|f - f_{n,m}\|_R \leq \frac{c}{\theta}\|f - \mathcal{Q}_n f\|_R. \tag{2.23}$$

Note that this follows from the sharp bound (2.17), the inequality (2.18) for $C_{n,m}$, and the definition of θ. In fact, one could also use (2.17) directly by numerically computing $C_{n,m}$ (recall that $C_{n,m}$ is the angle between two finite-dimensional subspaces, and hence computable). This would give a sharp, and consequently slightly better, bound than (2.23). However, the computation of $C_{n,m}$ is more intensive than that of $E_{n,m}$, and therefore we shall not pursue this further (see [8] for a discussion on this issue in the case of the original GS framework).

The significance of $\Theta(n;\theta)$ is that it allows one to control *a priori* both the stability and accuracy of the numerical method. Whilst it is convenient that $\Theta(n;\theta)$ can always be computed numerically, it is also important to have analytical estimates for each particular application. Naturally, such estimates are completely

dependent on both the sampling and reconstruction spaces, and thus must be obtained on a case-by-case basis. We shall devote much of the second half of this paper to the derivation of such estimates for the case of Fourier sampling and (piecewise) polynomial reconstructions.

3 Recovery from Fourier coefficients

We now consider the specific application of recovering a piecewise analytic function from its Fourier coefficients using a piecewise polynomial basis. As we show, it is possible to reconstruct in a completely stable manner in an arbitrary basis of piecewise polynomials, orthogonal with respect to the (piecewise) modified Jacobi weight. Thus, the individual polynomials employed can be specified by the user, making the method extremely flexible in this sense.

3.1 Preliminaries

Recall the setup of §2.1. Suppose that $f : [-1, 1] \to \mathbb{R}$ is piecewise analytic with jumps at the known locations $-1 < x_1 < \ldots < x_l < 1$. Set $x_0 = -1$ and $x_{l+1} = 1$. We shall assume that the first m Fourier coefficients (2.5) of f are given. Thus, we let $\mathrm{H}^S = \mathrm{L}^2(-1, 1)$ with its standard inner product $\langle \cdot, \cdot \rangle$ and norm $\|\cdot\|$, and define $\psi_j(x) = \frac{1}{\sqrt{2}} \exp(\mathrm{i} j \pi x)$ (recall that in this case we enumerate $\{\psi_j\}$ over $\mathbb{Z}$ as opposed to $\mathbb{N}$).

We shall reconstruct in the space T_n defined by (2.6), with a basis of this space consisting piecewise polynomials orthogonal with respect to the piecewise (modified) Jacobi weight (1.3) with parameters $\alpha = \{\alpha_0, \ldots, \alpha_l\}$, $\beta = \{\beta_0, \ldots, \beta_l\}$, where $\alpha_r, \beta_r > -1$, $r = 0, \ldots, l$. We let $\mathrm{H}^R = \mathrm{L}^2_{\alpha,\beta}(-1, 1)$ with corresponding inner product $\langle \cdot, \cdot \rangle_{\alpha,\beta}$ and norm $\|\cdot\|_{\alpha,\beta}$.

We enumerate the orthonormal basis for T_n as follows. First, let $c_r = \frac{1}{2}(x_{r+1} - x_r)$ and define $\Lambda_r(x) = \frac{x - x_r}{c_r} - 1$, so that $\Lambda(I_r) = [-1, 1]$, where $I_r = [x_r, x_{r+1}]$. For given α_r, β_r, let ϕ_j be the corresponding orthonormal modified Jacobi polynomial of degree j on $[-1, 1]$. Assume that $\phi_j \equiv 0$ outside $[-1, 1]$ and define the local (modified) Jacobi polynomial on I_r by $\phi_{r,j} = \frac{1}{\sqrt{c_r}} \phi_j \circ \Lambda_r$. It follows that the set

$$\{\phi_{r,j} : j = 0, \ldots, n_r - 1, \ r = 0, \ldots, l\}, \tag{3.1}$$

forms an orthonormal basis for T_n, and hence we may write $f_{n,m}$ as

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

with unknowns $\alpha_{r,j} \in \mathbb{C}$. We now compute $f_{n,m}$ via (2.4).

We first wish to analyze $f_{n,m}$ (implementation will be discussed in §4). This means applying the analysis of §2, and for this we need to verify (A1) and (A2). Additionally, we also wish to determine when we may replace (A2) by the stronger condition (A2′), and hence when the sharp bounds of Theorem 2.11 apply.

Verifying these assumptions requires understanding the behaviour of the Fourier projection operators $\mathcal{P}_m$ in the weighted spaces $\mathrm{L}^2_{\alpha,\beta}(-1, 1)$. This is the content of the next section.

3.2 Convergence of Fourier series in weighted norms

Somewhat surprisingly, this question has been considered before. Convergence of Fourier series with respect to piecewise modified Jacobi weights are simple consequences of this more general theory, which we now recap. The reader is referred to [39, 42] or more recently [15], for further details.

Suppose that $w(x)$ is a weight function such that $0 < w(x) < \infty$ almost everywhere in $[-1, 1]$. We shall denote the space of weighted square integrable functions by $\mathrm{L}^2_w(-1, 1)$, with corresponding norm $\|\cdot\|_w$. We say that $w(x)$ is a *Helson–Szegö* weight function if and only if there exists a constant $c > 0$ such that

$$\left(\frac{1}{|I|} \int_I w \right) \left(\frac{1}{|I|} \int_I \frac{1}{w} \right) \le c,$$

for all subsets $I \subseteq [-1, 1]$ that are either subintervals or complements of subintervals in $[-1, 1]$ (here $|I|$ denotes the length of I).

There is a fundamental relationship between Helson–Szegö weight functions and Fourier series. This is summarized in the following theorem:

Theorem 3.1 (see [42] or [15])**.** *Let $\mathcal{P}_m f$ denote the truncated Fourier expansion of a function f on $[-1,1]$. Then following three statements are equivalent:*

(i) w is a Helson–Szegö weight function.
(ii) There exists a finite constant $c > 0$ such that, for all $f \in \mathrm{L}_w^2(-1,1)$, $\|\mathcal{P}_m f\|_w \leq c\|f\|_w$, $\forall m \in \mathbb{N}$.
(iii) For every $f \in \mathrm{L}_w^2(-1,1)$, $\|f - \mathcal{P}_m f\|_w \to 0$ as $m \to \infty$.

Note that, without further regularity assumptions of f, the convergence rate in (iii) can be arbitrarily slow. Using this theorem (specifically conditions (ii) and (iii)), we immediately obtain the following:

Corollary 3.2. *Let $\mathcal{P}_m$ be the Fourier projection operator. Then (A1) and (A2′) hold with $\mathrm{H}^R = \mathrm{L}_w^2(-1,1)$ if and only if the weight function w is of Helson–Szegö type.*

Remark 3.3 Recall that in §2.2.2 we introduced the operator $\mathcal{W}_m = \mathcal{P}_m^* \circ \mathcal{P}_m$, and Lemma 2.10 shows that $\mathcal{W}_m \to \mathcal{I}$ strongly on H^R. This can actually be proved quite easily in the case considered in this section. If w is a Helson–Szegö weight function then, recalling the explicit form (2.20) for $\mathcal{W}_m$,

$$\|g - \mathcal{W}_m g\|_w \leq \|wg - \mathcal{P}_m(wg)\|_{w^{-1}} + \|\mathcal{P}_m(wg - w\mathcal{P}_m g)\|_{w^{-1}}.$$

If w is a Helson–Szegö weight function then so is $w^{-1}(x) = \frac{1}{w(x)}$ (wherever defined). Also, $g \in \mathrm{L}^2(-1,1)$ if and only if $wg \in \mathrm{L}_{w^{-1}}^2(-1,1)$. Therefore Theorem 3.1 gives

$$\|g - \mathcal{W}_m g\|_w \leq \|wg - \mathcal{P}_m(wg)\|_{w^{-1}} + c\|g - \mathcal{P}_m g\|_w \to 0, \quad m \to \infty,$$

as required.

In view of Corollary 3.2, our interest lies in the case where (piecewise) modified Jacobi weight functions are of Helson–Szegö type. We have:

Lemma 3.4. *Suppose that $w_{\alpha,\beta}$ is a piecewise modified Jacobi weight function with parameters $\alpha = \{\alpha_0, \dots, \alpha_l\}$ and $\beta = \{\beta_0, \dots, \beta_l\}$. Then $w_{\alpha,\beta}$ is a Helson–Szegö weight function if and only if $-1 < \alpha_r, \beta_r < 1$ for all $r = 0, \dots, l$.*

Proof. Note that if $\tilde{w}$ is a Helson–Szegö weight function and $w \asymp \tilde{w}$ uniformly in x, then so is w. We now recall a result proved in [15]. Let $w(t) = \varphi(\mathrm{e}^{\mathrm{i}\pi t})$, where φ is finite, positive and continuous on the unit circle except at a finite set of points $z_1, \dots, z_l$. Moreover, suppose that

$$\varphi(z) = \mathcal{O}\left(|z - z_r|^{\gamma_r}\right), \quad \frac{1}{\varphi(z)} = \mathcal{O}\left(|z - z_r|^{-\gamma_r}\right), \quad z \to z_r.$$

Then w is a Helson–Szegö weight function if and only if $-1 < \gamma_r < 1$, $r = 1, \dots, l$. Noting that $1 - t^2 \asymp |1 + \mathrm{e}^{\mathrm{i}\pi t}|$ now gives the full result. $\qquad\square$

This lemma, in combination with Theorem 3.1, gives the following:

Corollary 3.5. *Let $\mathcal{P}_m$ be the Fourier projection operator and $\mathrm{H}^R = \mathrm{L}_{\alpha,\beta}^2(-1,1)$. Then assumptions (A1) and (A2′) hold if and only if $-1 < \alpha_r, \beta_r < 1$ for all $r = 0, \dots, l$.*

With this corollary in hand, the analysis of §2.2.2 immediately applies to this problem. In particular, we deduce stability, convergence, asymptotic optimality and the sharp bounds of Theorem 2.11. Note that this corollary includes the cases $\alpha = \beta = \{-\frac{1}{2}, \dots, -\frac{1}{2}, \}, \{\frac{1}{2}, \dots, \frac{1}{2}\}$ corresponding to Chebyshev weights (of the first and second kinds) that form the principal interest of this paper.

Let us now address the other case, where at least one α_j or β_j exceeds 1. As shown by Lemma 3.4, $w_{\alpha,\beta}$ is no longer a Helson–Szegö weight function. However, one can still verify (A1) and (A2) in this case, for a suitable choice of the subspace U:

Corollary 3.6. *Let $\mathcal{P}_m$ be the Fourier projection operator and $w_{\alpha,\beta}$ a piecewise Jacobi weight function with parameters $\alpha = \{\alpha_0, \dots, \alpha_l\}$ and $\beta = \{\beta_0, \dots, \beta_l\}$. For $r = 0, \dots, l$ define*

$$\tilde{\alpha}_r = \begin{cases} \alpha_r & -1 < \alpha_r < 1 \\ 1 - \epsilon_r & \alpha_r \geq 1 \end{cases}, \qquad \tilde{\beta}_r = \begin{cases} \beta_r & -1 < \beta_r < 1 \\ 1 - \epsilon_r' & \beta_r \geq 1 \end{cases},$$

where $0 < \epsilon_r, \epsilon_r' < 2$ are arbitrary. Then (A1) and (A2) hold with $\mathrm{U} = \mathrm{L}_{\tilde{\alpha},\tilde{\beta}}^2(-1,1)$ and $\|\cdot\| = \|\cdot\|_{\tilde{\alpha},\tilde{\beta}}$.

Proof. By construction, $w_{\tilde{\alpha},\tilde{\beta}}$ is a piecewise modified Jacobi weight function of Helson–Szegö type. Moreover, we have the continuous embedding $U \hookrightarrow H^R$. In other words, there exists $c > 0$ such that $\|g\|_{\alpha,\beta} \leq c\|g\|_{\tilde{\alpha},\tilde{\beta}}$, $\forall g \in U$. Therefore, by Theorem 3.1,

$$\|g - \mathcal{P}_m g\|_{\alpha,\beta} \leq c\|g - \mathcal{P}_m g\|_{\tilde{\alpha},\tilde{\beta}} \to 0, \quad \forall g \in U,$$

and

$$\|\mathcal{P}_m g\|_{\alpha,\beta} \leq c\|\mathcal{P}_m g\|_{\tilde{\alpha},\tilde{\beta}} \leq c\|g\|_{\tilde{\alpha},\tilde{\beta}}, \quad \forall g \in U.$$

This gives the result. □

As a result of this corollary, the analysis of §2.2.1 applies in this setting. Hence one verifies stability and convergence, albeit without the sharp bounds of §2.2.2. As commented, the space U of functions we can reconstruct with this approach is very large, and certainly includes all functions of interest in practice.

Observe that Corollaries 3.5 and 3.6, in combination with the analysis of §2, confirm the first part of the main theorem of this paper (Theorem 1.4). For the second part, we need to estimate the stable sampling rate $\Theta(n;\theta)$, as defined by (2.21). This is the content of the next section. The final part of Theorem 1.4, i.e. the computational cost estimates, will be addressed in §4.

3.3 Estimates for the stable sampling rate $\Theta(n;\theta)$

The key theorem we prove in this section is the following:

Theorem 3.7. *Let $\mathcal{P}_m$ be the Fourier projection operator, T_n be given by (2.6) for $n = (n_0, \ldots, n_l)$, and suppose that $H^R = L^2_{\alpha,\beta}(-1,1)$, where $\alpha = \{\alpha_0, \ldots, \alpha_l\}$, $\beta = \{\beta_0, \ldots, \beta_l\}$. Then, for $0 < \theta < 1$ there exists $c > 0$ depending only on θ such that*

$$\Theta(n;\theta) \leq c\bar{n}^r, \quad \forall n_0, \ldots, n_l \in \mathbb{N},$$

where $\Theta(n;\theta)$ is given by (2.21), $\bar{n} = \max\{n_0, \ldots, n_l\}$, and

$$r = \max\left\{2, p+1, 2+2\frac{p-q}{q+1}\right\},$$

$$p = \max\{\alpha_0, \ldots, \alpha_l, \beta_0, \ldots, \beta_l\}, \quad q = \min\{\alpha_0, \ldots, \alpha_l, \beta_0, \ldots, \beta_l\}.$$

A proof of this theorem is given later in this section. Recall that the most important case in this paper is when $\alpha_0 = \ldots = \alpha_l = \beta_0 = \ldots = \beta_l = \gamma$. In other words, a piecewise ultraspherical weight with the same parameter values in each of the intervals $I_0, \ldots, I_l$. In this case, $p = q = \gamma$, and Theorem 3.7 gives that

$$\Theta(n;\theta) \leq c\bar{n}^{\max\{2,\gamma+1\}}.$$

For example, whenever $-1 < \gamma < 1$, such as in the Legendre or Chebyshev cases, we obtain an $\mathcal{O}(n^2)$ scaling for $\Theta(n;\theta)$. Note that this result was first proved in [41] (see also [6]) for the Legendre case. In fact, quadratic scaling of $\Theta(n;\theta)$ with n is the best possible predicted by this theorem, and this is also a theoretical barrier for stability (recall the discussion in §1.4).

One negative consequence of this result is that if one chooses α and β so that either $p \neq q$ or $p > 1$, then the growth of $\Theta(n;\theta)$ is more severe. Thus, more Fourier coefficients are required to reconstruct in terms of the corresponding piecewise polynomial basis, and therefore such bases are suboptimal in view of the aforementioned theoretical barrier. Nevertheless, whenever α and β are chosen so that $p = q$ and $p \leq 1$, we retain quadratic growth.

The reason for this increased scaling whenever $p > 1$ is that the weighted norm $\|\cdot\|_{\alpha,\beta}$ is too weak in comparison to the standard L^2 norm to adequately control the error of the Fourier projection. On the other hand, note that the local convergence of the Fourier series of an arbitrary function ϕ at any point $x \in (-1,1)$ is determined by its global behaviour at *every* jump $x_0, x_1, \ldots, x_l$. Thus, convergence in the local norm over each I_r subinterval must balance this global behaviour, which leads to the higher scaling for $p \neq q$.

In Figure 1 we plot $\Theta(n;\theta)$ and $n^{-2}\Theta(n;\theta)$ for case where $l = 0$ (i.e. analytic and nonperiodic functions f) and $\alpha = \beta = \gamma$ (i.e. reconstructions in ultraspherical polynomials). As is evident, when $\gamma = \pm\frac{1}{2}$, $\Theta(n;\theta)$ scales quadratically with n. Conversely, this scaling is cubic when the value $\gamma = 2$ is used. These results confirm Theorem 3.7.
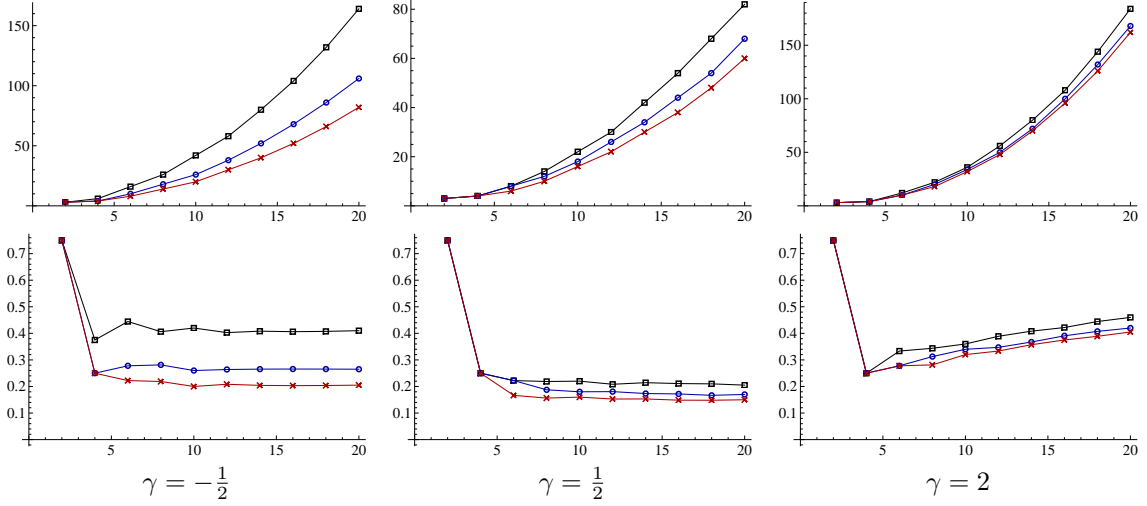
Figure 1: The functions $\Theta(n;\theta)$ (top) and $n^{-2}\Theta(n;\theta)$ (bottom) against $n = 1,\ldots,20$ for $\theta = \frac{1}{4}, \frac{1}{8}, \frac{1}{16}$ (squares, circles and crosses respectively).

The remainder of this section is devoted to a proof of Theorem 3.7. For this, we first require some general results about Fourier series, and in particular, Fourier series of arbitrary piecewise algebraic polynomials.

We now introduce some additional notation. We write $A \lesssim B$ to mean that there exists a constant $c > 0$ independent of all relevant parameters such that $A \leq cB$. If $I \subseteq [-1,1]$ is an interval and $\alpha, \beta > -1$, we write $\|\cdot\|_{\infty,I}$ and $\|\cdot\|_{\alpha,\beta,I}$ for the norms corresponding to the spaces $\mathrm{L}^\infty(I)$ and $\mathrm{L}^2_{w_{\alpha,\beta}}(I)$ respectively, where $w_{\alpha,\beta}$ is a modified Jacobi weight on I. Finally, we let $[g](x) = g(x^+) - g(x^-)$ be the value of the jump of a function $g : [-1,1] \to \mathbb{R}$ at the point $x \in [-1,1]$, with the understanding that $[g](-1) = [g](1) = g(-1) - g(1)$.

Our first lemma is as follows:

Lemma 3.8. *Suppose that $f : [-1,1] \to \mathbb{R}$ is piecewise smooth with jump discontinuities at $-1 < x_1 < \ldots < x_l < 1$. Then we have the asymptotic expansion*

$$\hat{f}_j = \frac{1}{\sqrt{2}} \int_{-1}^1 f(x)\mathrm{e}^{-\mathrm{i}j\pi x}\,\mathrm{d}x \sim \frac{1}{\sqrt{2}} \sum_{r=0}^l \sum_{s=0}^\infty \frac{(-1)^{s+1}\mathrm{e}^{-\mathrm{i}\pi j x_r}}{(-\mathrm{i}j\pi)^{s+1}} \left[f^{(s)}\right](x_r), \quad |j| \gg 1. \tag{3.2}$$

In particular, if $\phi \in \mathrm{T}_n$, where T_n is as in (2.6), then

$$\hat{\phi}_j = \frac{1}{\sqrt{2}} \sum_{r=0}^l \sum_{s=0}^{n_r-1} \frac{(-1)^{s+1}\mathrm{e}^{-\mathrm{i}\pi j x_r}}{(-\mathrm{i}j\pi)^{s+1}} \left[\phi^{(s)}\right](x_r), \quad \forall j \in \mathbb{Z}\backslash\{0\}. \tag{3.3}$$

Proof. This follows immediately by repeated integration by parts. $\qquad\square$

Note that the symbol $\sim$ in (3.2) denotes an asymptotic expansion (in the usual Poincaré sense) in the parameter $j \to \pm\infty$. As is typical, the right-hand side of (3.2) does usually not converge for any fixed j. On the other hand, (3.3) holds for any finite $j \neq 0$ and arbitrary piecewise polynomials ϕ.

A convenient means to describe the Fourier series of an arbitrary function piecewise smooth function f is in terms of certain cardinal functions [3, 46]. For $r \in \mathbb{N}$, let $\tilde{C}_r(x)$ be the periodic extension of $-\frac{2^r}{(r+1)!}B_{r+1}(x)$ to the real line, where B_{r+1} is the $(r+1)^{\text{th}}$ Bernoulli polynomial [1, chpt. 23], and define $C_r(x) = \tilde{C}_r(\frac{1}{2}x + 1)$. Note that C_r is a piecewise polynomial of degree $r + 1$ with discontinuity at $x = 0$. Since the Bernoulli polynomials form an Appell sequence, we have that

$$\left[C_r^{(s)}\right](0) = \delta_{r-s}, \quad r, s \in \mathbb{N}.$$

Define the piecewise polynomials

$$q_{r,s}(x) = C_s(x - x_r), \quad r = 0,\ldots,l, \quad s \in \mathbb{N},$$

16

via translation. Substituting $\phi = q_{r,s}$ in (3.3), we note that

$$\widehat{q_{r,s}}_j = \frac{1}{\sqrt{2}} \frac{(-1)^{s+1} \mathrm{e}^{-\mathrm{i}j\pi x_r}}{(-\mathrm{i}j\pi)^{s+1}}, \quad j \in \mathbb{Z}\backslash\{0\}. \tag{3.4}$$

Thus (3.3) can be expressed more succinctly as

$$\hat{\phi}_j = \sum_{r=0}^{l} \sum_{s=0}^{n_r-1} \widehat{q_{r,s}}_j \left[\phi^{(r)} \right](x_r),$$

and we also have that

$$\phi(x) - \mathcal{P}_m\phi(x) = \sum_{r=0}^{l} \sum_{s=0}^{n_r-1} \left[\phi^{(s)} \right](x_r) \Big(q_{r,s}(x) - \mathcal{P}_m q_{r,s}(x) \Big), \tag{3.5}$$

Thus, in order to study the convergence of the Fourier series of an arbitrary $\phi \in \mathrm{T}_n$, it suffices to consider the Fourier series of the cardinal functions $q_{r,s}$. We shall use this approach when proving the next two results.

Lemma 3.9. *Suppose that $\phi \in \mathrm{T}_n$. Then, there exists $c > 0$ independent of n and ϕ such that $\|\phi - \mathcal{P}_m\phi\|_\infty \lesssim \|\phi\|_\infty$ whenever $m \geq c\bar{n}^2$.*

Proof. By (3.5) and Markov's inequality,

$$\|\phi'\|_\infty \lesssim n^2 \|\phi\|_\infty, \quad \forall \phi \in \mathbb{P}_{n-1}, \tag{3.6}$$

(see, for example, [16]) we have

$$\|\phi - \mathcal{P}_m\phi\|_\infty \leq \sum_{r=0}^{l} \sum_{s=0}^{n_r-1} \left| \left[\phi^{(s)} \right](x_r) \right| \|q_{r,s} - \mathcal{P}_m q_{r,s}\|_\infty$$

$$\lesssim \sum_{r=0}^{l} \sum_{s=0}^{n_r-1} \left\{ n_r^{2s} \|\phi\|_{\infty,I_r} + n_{r-1}^{2s} \|\phi\|_{\infty,I_{r-1}} \right\} \|q_{r,s} - \mathcal{P}_m q_{r,s}\|_\infty, \tag{3.7}$$

where, for convenience we let $I_{-1} = I_l$ and $n_{-1} = n_l$. Consider $q_{r,s}$. For $s = 1, 2, \ldots$, the function $q_{r,s} \in \mathrm{C}^1(\mathbb{T})$ and the derivative $q_{r,s}''$ is of bounded variation. Hence $\mathcal{P}_m q_{r,s}$ converges uniformly to $q_{r,s}$, the coefficients $\widehat{q_{r,s}}_j$ are absolutely summable and we may write $q_{r,s}(x) - \mathcal{P}_m q_{r,s}(x)$ as the infinite sum

$$q_{r,s}(x) - \mathcal{P}_m q_{r,s}(x) = \sum_{|j| > \frac{m}{2}} \widehat{q_{r,s}}_j \psi_j(x), \quad \forall x \in [-1, 1].$$

Using the expression (3.4) we deduce that

$$\|q_{r,s} - \mathcal{P}_m q_{r,s}\|_\infty \lesssim \sum_{|j| > \frac{m}{2}} |\widehat{q_{r,s}}_j| \lesssim \sum_{j > \frac{m}{2}} j^{-s-1} \lesssim m^{-s}.$$

Now consider $q_{r,0}$. By translation, it suffices to consider $q_{0,0}(x) = \frac{1}{4} - \frac{1}{2}x$. This function has uniformly bounded Fourier series, $\|\mathcal{P}_m q_{0,0}\|_\infty \leq c$. Substituting this and the previous result into (3.7) we obtain

$$\|\phi - \mathcal{P}_m\phi\|_\infty \lesssim \|\phi\|_\infty \sum_{r=0}^{l} \sum_{s=0}^{n_r-1} \left(\frac{\tilde{c}n_r^2}{m} \right)^s \lesssim \|\phi\|_\infty \sum_{s=0}^{\bar{n}-1} \left(\frac{\tilde{c}\bar{n}^2}{m} \right)^s,$$

for some $\tilde{c} > 0$. Therefore, provided $m \geq c\bar{n}^2$ with $c > \tilde{c}$, this sum is bounded independently of n, m and the result follows. $\square$

Lemma 3.10. *Suppose that $\phi \in \mathrm{T}_n$. Then there exists a $c > 0$ independent of n and ϕ such that*

$$|\phi(x) - \mathcal{P}_m\phi(x)| \lesssim m^{-1} \|\phi\|_\infty \sum_{r=0}^{l} \left(1 + |\sin \tfrac{\pi}{2}(x - x_r)|^{-1} \right), \quad \forall x \in (-1, 1)\backslash\{x_1, \ldots, x_l\},$$

whenever $m \geq c\bar{n}^2$.

Proof. Suppose that we can show that, for any $s \in \mathbb{N}$ and $r = 0, \ldots, l$,

$$|q_{r,s}(x) - \mathcal{P}_m q_{r,s}(x)| \lesssim m^{-s-1} \left(1 + |\sin \tfrac{\pi}{2}(x - x_r)|^{-1} \right), \quad x \in (-1, 1) \backslash \{x_r\}. \tag{3.8}$$

Then the result follows directly from (3.5) and Markov's inequality (3.6):

$$|\phi(x) - \mathcal{P}_m \phi(x)| \lesssim \sum_{r=0}^{l} \left(1 + |\sin \tfrac{\pi}{2}(x - x_r)|^{-1} \right) \sum_{s=0}^{n_r - 1} m^{-s-1} \left(n_r^{2s} + n_{r-1}^{2s} \right) \|\phi\|_\infty$$

$$\lesssim m^{-1} \|\phi\|_\infty \sum_{r=0}^{l} \left(1 + |\sin \tfrac{\pi}{2}(x - x_r)|^{-1} \right) \sum_{s=0}^{n_r - 1} \left(\frac{\bar{n}^2}{m} \right)^s.$$

Hence it remains to show (3.8). Without loss of generality, we consider $q_{0,s}(x)$. Using (3.4) and noticing that $\mathcal{P}_m q_{0,s}(x)$ converges to $q_{0,s}(x)$ whenever $-1 < x < 1$, we obtain

$$|q_{0,s}(x) - \mathcal{P}_m q_{0,s}(x)| = \frac{1}{2\pi^{s+1}} \left| \sum_{|j|>m} \frac{(-1)^j}{j^{s+1}} \exp(ij\pi x) \right| \lesssim \left| \sum_{j>m} \frac{z^j}{j^{s+1}} \right|,$$

where $z = -\exp(i\pi x)$. Note that we have replaced $\lfloor \frac{m}{2} \rfloor$ by m in the summation: this does not effect the result and slightly simplifies the notation. Thus it suffices to consider the sum

$$S_{s,m}(z) = \sum_{j>m} \frac{z^j}{j^{s+1}},$$

for arbitrary $|z| = 1$. We now use Abel summation:

$$S_{s,m}(z) = \sum_{j>m} \frac{z^j}{j^s} \left(\frac{1}{j} - \frac{1}{j+1} \right) + \sum_{j>m} \frac{z^j}{j^s(j+1)} = \sum_{j>m} \frac{z^j}{j^{s+1}(j+1)} + \sum_{j>m} \frac{z^j}{j^s(j+1)}.$$

The first term is $\mathcal{O}\left(m^{-s-1} \right)$ uniformly in $|z| = 1$. Now consider the second term. This is equal to

$$z^{-1} \sum_{j>m+1} \frac{z^j}{j(j-1)^s} = z^{-1} \sum_{j>m+1} \frac{z^j}{j^{s+1}} \left(1 + \mathcal{O}\left(j^{-1} \right) \right) = z^{-1} \sum_{j>m+1} \frac{z^j}{j^{s+1}} + \mathcal{O}\left(m^{-s-1} \right).$$

Note that the final equality holds uniformly in $|z| = 1$. Hence

$$S_{s,m}(z) = z^{-1} \left(S_{s,m}(z) - \frac{z^{m+1}}{(m+1)^{s+1}} \right) + \mathcal{O}\left(m^{-s-1} \right).$$

Rearranging, this gives

$$S_{s,m}(z) = \frac{z^{m+1}}{1-z} (m+1)^{-s-1} + \mathcal{O}\left(m^{-s-1} \right),$$

uniformly in $|z| = 1$. Note that $x \neq \pm 1$, therefore $z \neq 1$. Moreover, since $z = -\exp(i\pi x)$, we have that

$$|z - 1| = 2|\cos \tfrac{\pi}{2}x| = 2|\sin \tfrac{\pi}{2}(x - x_0)|,$$

and consequently

$$|S_{s,m}(z)| \lesssim \left(1 + |\sin \tfrac{\pi}{2}(x - x_0)|^{-1} \right) m^{-s-1}, \quad \forall z = -\exp(i\pi x) \neq 1,$$

which gives (3.8). $\qquad \square$

We are now able to prove the main result of this section:

Proof of Theorem 3.7. It suffices to show that, given $\theta \in (0, 1)$, there exists c such that $E_{n,m} < 1 - \theta$ for all $m \geq c\bar{n}^r$, where $E_{n,m}$ is given by (2.7). Let $\phi \in \mathrm{T}_n$ with $\|\phi\|_R = 1$. Note that

$$\|\phi - \mathcal{P}_m \phi\|_{\alpha,\beta}^2 = \sum_{r=0}^{l} \int_{I_r} |\phi(x) - \mathcal{P}_m \phi(x)|^2 w_{\alpha_r,\beta_r}(x) \, dx.$$

We consider each term of this sum separately. Write $I_r = I_r^+ \cup I_r^-$, where $I_r^+ = [d_r, x_{r+1})$, $I_r^- = (x_r, d_r]$ and $d_r = \frac{1}{2}(x_r + x_{r+1})$ is the midpoint of I_r. In a similar manner we also define, for arbitrary $0 < \epsilon < c_r = \frac{1}{2}(x_{r+1} - x_r)$,

$$I_{r,\epsilon}^+ = [d_r, x_{r+1} - \epsilon), \quad I_{r,\epsilon}^- = (x_r + \epsilon, d_r],$$
$$J_{r,\epsilon}^+ = [x_{r+1} - \epsilon, x_{r+1}), \quad J_{r,\epsilon}^- = (x_r, x_r + \epsilon].$$

Note that $I_r^+ = I_{r,\epsilon}^+ \cup J_{r,\epsilon}^+$ and $I_r^- = I_{r,\epsilon}^- \cup J_{r,\epsilon}^-$. We now write

$$\int_{I_r} |\phi(x) - \mathcal{P}_m\phi(x)|^2 w_{\alpha_r,\beta_r}(x)\,\mathrm{d}x$$

$$= \int_{I_{r,\epsilon}^+} |\phi(x) - \mathcal{P}_m\phi(x)|^2 w_{\alpha_r,\beta_r}(x)\,\mathrm{d}x + \int_{I_{r,\epsilon}^-} |\phi(x) - \mathcal{P}_m\phi(x)|^2 w_{\alpha_r,\beta_r}(x)\,\mathrm{d}x$$

$$+ \int_{J_{r,\epsilon}^+} |\phi(x) - \mathcal{P}_m\phi(x)|^2 w_{\alpha_r,\beta_r}(x)\,\mathrm{d}x + \int_{J_{r,\epsilon}^-} |\phi(x) - \mathcal{P}_m\phi(x)|^2 w_{\alpha_r,\beta_r}(x)\,\mathrm{d}x.$$

Consider the integral over $I_{r,\epsilon}^+$. By construction $w_{\alpha,\beta}(x) \lesssim (x_{r+1} - x)^{\alpha_r}$ uniformly for $x \in I_r^+$. Using this and Lemma 3.10 we find that

$$\int_{I_{r,\epsilon}^+} |\phi(x) - \mathcal{P}_m\phi(x)|^2 w_{\alpha_r,\beta_r}(x)\,\mathrm{d}x \lesssim m^{-2}\|\phi\|_\infty^2 \int_{I_{r,\epsilon}^+} \left[\sum_{s=0}^{l} \left(1 + |\sin\tfrac{\pi}{2}(x - x_s)|^{-1}\right)\right]^2 (x_{r+1} - x)^{\alpha_r}\,\mathrm{d}x.$$

Since $|\sin\frac{\pi}{2}(x - x_{r+1})| \gtrsim |x - x_{r+1}|$ and $|\sin\frac{\pi}{2}(x - x_s)| \gtrsim 1$, $s \neq r+1$, whenever $-1 < |x - x_{r+1}| < 1$ we deduce that

$$\int_{I_{r,\epsilon}^+} |\phi(x) - \mathcal{P}_m\phi(x)|^2 w_{\alpha_r,\beta_r}(x)\,\mathrm{d}x \lesssim m^{-2}\|\phi\|_\infty^2 \left(1 + \int_{I_{r,\epsilon}^+} (x_{r+1} - x)^{\alpha_r - 2}\,\mathrm{d}x\right)$$

$$\lesssim \max\{1, \epsilon^{\alpha_r - 1}\} m^{-2}\|\phi\|_\infty^2.$$

Similarly, for $I_{r,\epsilon}^-$, we have

$$\int_{I_{r,\epsilon}^-} |\phi(x) - \mathcal{P}_m\phi(x)|^2 w_{\alpha_r,\beta_r}(x)\,\mathrm{d}x \lesssim \max\{1, \epsilon^{\beta_r - 1}\} m^{-2}\|\phi\|_\infty^2.$$

We now require an estimate for $\|\phi\|_\infty$. First let $p \in \mathbb{P}_n$ be a polynomial and suppose that $\alpha, \beta > -1$. Then it can be shown that

$$\|p\|_\infty \lesssim n^{\max\{\alpha,\beta\}+1}\|p\|_{\alpha,\beta}.$$

(this follows directly from the decomposition $\|p\|_{\alpha,\beta}^2 = \sum_{j=0}^{n} |\langle p, p_j \rangle_{\alpha,\beta}|^2$, where $\{p_j\}$ are the orthonormal polynomials with respect to the weight function $w_{\alpha,\beta}$, and the fact that $\|p_j\|_\infty \sim j^{\max\{\alpha,\beta\}+\frac{1}{2}}$ [45]).

Since $\phi \in \mathrm{T}_n$, and therefore $\phi|_{I_r} \in \mathbb{P}_{n_r-1}$, we have

$$\|\phi\|_\infty = \max_{r=0,\ldots,l} \|\phi\|_{\infty,I_r} \lesssim \max_{r=0,\ldots,l} \left\{n_r^{\max\{\alpha_r,\beta_r\}+1}\|\phi\|_{\alpha_r,\beta_r,I_r}\right\} \leq \bar{n}^{p+1}\|\phi\|_{\alpha,\beta},$$

where $p = \max\{\alpha_0, \ldots, \alpha_l, \beta_0, \ldots, \beta_l\}$. Hence we now find that

$$\sum_{r=0}^{l} \left\{\int_{I_{r,\epsilon}^+} |\phi(x) - \mathcal{P}_m\phi(x)|^2 w_{\alpha_r,\beta_r}(x)\,\mathrm{d}x + \int_{I_{r,\epsilon}^-} |\phi(x) - \mathcal{P}_m\phi(x)|^2 w_{\alpha_r,\beta_r}(x)\,\mathrm{d}x\right\}$$

$$\lesssim \max\{1, \epsilon^{p-1}\} \left(\frac{\bar{n}^{p+1}}{m}\right)^2 \|\phi\|_{\alpha,\beta}^2. \tag{3.9}$$

Now consider the integral over $J_{r,\epsilon}^\pm$. If $m \gtrsim \bar{n}^2$ an application of Lemma 3.9 gives

$$\int_{J_{r,\epsilon}^+} |\phi(x) - \mathcal{P}_m\phi(x)|^2 w_{\alpha_r,\beta_r}(x)\,\mathrm{d}x \lesssim \|\phi\|_\infty^2 \int_{1-\epsilon}^{1} (1 - x)^{\alpha_r}\,\mathrm{d}x \lesssim \|\phi\|_\infty^2 \epsilon^{\alpha_r+1},$$

19

and likewise for $J_{r,\epsilon}^-$ with the exponent α_r replaced by β_r. Since $\epsilon < 1$, this and (3.9) now give

$$\sum_{r=0}^l \left\{ \int_{J_{r,\epsilon}^+} |\phi(x) - \mathcal{P}_m\phi(x)|^2 w_{\alpha_r,\beta_r}(x)\,\mathrm{d}x + \int_{J_{r,\epsilon}^-} |\phi(x) - \mathcal{P}_m\phi(x)|^2 w_{\alpha_r,\beta_r}(x)\,\mathrm{d}x \right\} \lesssim \epsilon^{q+1} \bar{n}^{2(p+1)} \|\phi\|_{\alpha,\beta}^2. \tag{3.10}$$

Setting $\epsilon = cm^{-1}$ for $c < \min\{c_r : r = 0,\dots,l\}$ and combining (3.9) and (3.10), we obtain

$$\|\phi - \mathcal{P}_m\phi\|_{\alpha,\beta}^2 \lesssim \left[\max\{1, m^{1-p}\} \left(\frac{\bar{n}^{p+1}}{m} \right)^2 + \frac{\bar{n}^{2(p+1)}}{m^{q+1}} \right] \|\phi\|_{\alpha,\beta}^2.$$

Hence

$$E_{n,m}^2 = \sup_{\substack{\phi \in \mathrm{T}_n \\ \|\phi\|_{\alpha,\beta}=1}} \|\phi - \mathcal{P}_m\phi\|_{\alpha,\beta}^2 \lesssim \left[\max\{1, m^{1-p}\} \left(\frac{\bar{n}^{p+1}}{m} \right)^2 + \frac{\bar{n}^{2(p+1)}}{m^{q+1}} \right]. \tag{3.11}$$

The second term can be made arbitrarily small by selecting

$$m \gtrsim \bar{n}^{2\frac{p+1}{q+1}} = \bar{n}^{2+2\frac{p-q}{q+1}}.$$

Now consider the first term. If $p < 1$ then

$$\max\{1, m^{1-p}\} \left(\frac{\bar{n}^{p+1}}{m} \right)^2 = m^{1-p} \left(\frac{\bar{n}^{p+1}}{m} \right)^2 = \left(\frac{\bar{n}^2}{m} \right)^{p+1},$$

and therefore we require $m \gtrsim \bar{n}^2$. Conversely, when $p \geq 1$, we have

$$\max\{1, m^{1-p}\} \left(\frac{\bar{n}^{p+1}}{m} \right)^2 = \left(\frac{\bar{n}^{p+1}}{m} \right)^2,$$

meaning that $m \gtrsim \bar{n}^{p+1}$. Hence, the first term of the right-hand side of (3.11) can be made arbitrarily small by the choice $m \gtrsim \bar{n}^{\max\{2,p+1\}}$. This completes the proof. $\qquad\square$

Theorem 3.7 shows that in many cases, the scaling $m = \mathcal{O}\left(\bar{n}^2\right)$ is sufficient for a stable quasi-optimal reconstruction. One downside of our result is it does not give an explicit upper bound for the scaling constant c. In simpler case of Legendre polynomials explicit, and fairly sharp, bounds were derived in [6]. It is in theory possible to obtain explicit bounds by following the steps of the above proof and keeping track of constants. However, because the proof is rather involved, this is likely to be grossly pessimistic. On the other hand, since $\Theta(n;\theta)$ can always be computed, we may always numerically estimate c. As shown in Figure 1 one can take a fairly small value $c = \frac{1}{4}$ in implementations. We use this value in the next section.

4 Numerical implementation and examples

4.1 Implementation

As mentioned in §2.3, the main issue in implementation is the computation of the entries of the matrices U and C. Computation of the entries of U can be carried out using the iterative procedure of [6, Lemma 3.4]. The total cost incurred is $\mathcal{O}\left(mn\right)$ operations. Note that, as commented in [6], one may in practice need to use a two-phase algorithm (such as that described in [22]) to ensure stability of the computation of these entries. However, this does not increase the overall cost.

For the computation of C, recall that this matrix has $(j,k)^{\text{th}}$ entry $\langle \psi_j, \psi_k \rangle_R$. In other words, if $w_{\alpha,\beta}$ is the corresponding weight function, then

$$\langle \psi_j, \psi_k \rangle_R = \frac{1}{2} \int_{-1}^1 \mathrm{e}^{\mathrm{i}\pi(j-k)x} w_{\alpha,\beta}(x)\,\mathrm{d}x = c_{j-k}.$$
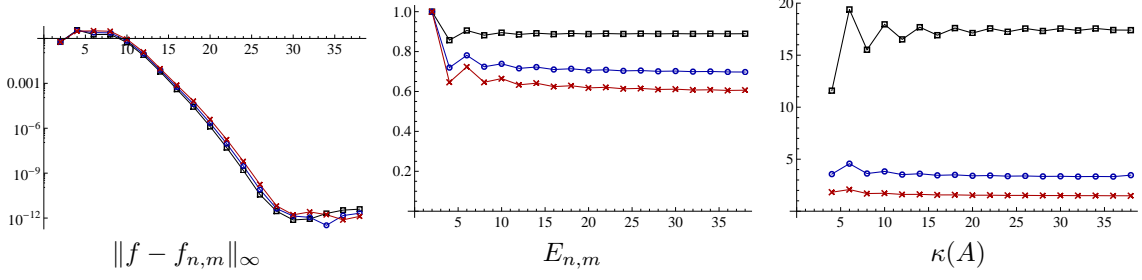
Figure 2: The quantities $\|f - f_{n,m}\|_\infty$, $E_{n,m}$ and $\kappa(A)$ (left to right) against $n = 1, 2, \ldots, 40$, for $f(x) = \mathrm{e}^x \cos 8x$, $m = \lceil \frac{1}{4} n^2 \rceil$ and $\gamma = -\frac{1}{2}, 0, \frac{1}{2}$ (squares, circles and crosses respectively).

Hence C is always a Toeplitz matrix, and to compute C we only need to determine c_j for $j = 0, \ldots, m$ (note that $c_{-j} = \overline{c_j}$). Fortunately, since

$$\int_{-1}^1 \mathrm{e}^{\mathrm{i}zx}(1 - x^2)^\alpha \, \mathrm{d}x = \sqrt{\pi} \frac{\Gamma(\alpha + 1)}{\Gamma(\alpha + \frac{1}{2})} (\alpha + \tfrac{1}{2}) \left(\frac{2}{z}\right)^{\alpha + \frac{1}{2}} J_{\alpha + \frac{1}{2}}(z), \quad z \in \mathbb{R},$$

where Γ is the Gamma function and J_ν is the Bessel function of the first kind [34], we can always determine c_j explicitly for the case where $w_{\alpha,\beta}$ is a piecewise ultraspherical weight function. Unfortunately, there is no such closed form expression when $w_{\alpha,\beta}$ is a piecewise Jacobi weight. However, in this case one can compute c_j using a Gaussian quadrature based on the given weight function, with the total cost of this approach being $\mathcal{O}(m^2)$. We shall not discuss this further.

For evaluation of the quantity $\Theta(n; \theta)$ it is also necessary to compute the matrix $V \in \mathbb{C}^{m \times n}$ with $(j, k)^{\text{th}}$ entry $\langle \phi_k, \psi_j \rangle_R$. Whenever w corresponds to a piecewise ultraspherical weight function, the entries of V are actually known explicitly. This follows from the fact that the integrals

$$\int_{-1}^1 \mathrm{e}^{\mathrm{i}zx} \phi_j(x)(1 - x^2)^\alpha \, \mathrm{d}x, \quad z \in \mathbb{R},$$

have an explicit expression whenever ϕ_j is the jth normalized ultraspherical polynomial corresponding to the parameter $\alpha > -1$ (see, for example, [34, eqn. (3.7)]).

Note that the above arguments establish the final part of Theorem 1.4 (the first parts were addressed in the previous sections) concerning the cost of the preprocessing step in the computation of $f_{n,m}$ (i.e. forming the matrices U and C).

4.2 Examples

In Figure 2 we consider the reconstruction of the function $f(x) = \mathrm{e}^x \cos 8x$ using Legendre polynomials and Chebyshev polynomials of the first and second kinds. As is evident, all three approaches give near identical approximation errors. In fact, regardless of the polynomial basis used, one attains roughly machine accuracy using only $m = 225$ Fourier samples (by means of comparison, direct expansion of this function in a Fourier series gives an error of order 10^{-1} for this value of m). Also, the condition number of the matrix $A = U^*CU$ remains bounded in n, as predicted by the analysis in §2.

In Figure 3 we consider the piecewise analytic function

$$f(x) = \begin{cases} (2\mathrm{e}^{2\pi(x+1)} - 1 - \mathrm{e}^\pi)(\mathrm{e}^\pi - 1)^{-1} & x \in [-1, -\frac{1}{2}) \\ -\sin(\frac{2\pi x}{3} + \frac{\pi}{3}) & x \in [-\frac{1}{2}, 1] \end{cases} \tag{4.1}$$

This function was put forth in [50] to test algorithms for overcoming the Gibbs phenomenon. Aside from the discontinuity, its sharp peak makes it a challenging function to reconstruct accurately. However, as shown in this figure, using a polynomial of degree 16 in each subinterval of smoothness, we recover the function to around 14 digits of accuracy. As in the previous example, all three polynomial bases used (Legendre and Chebyshev) give roughly the same approximation error.

Note that both functions considered above are piecewise analytic, and therefore their projections $\mathcal{Q}_n f$ converge geometrically fast in n. Since the GS reconstruction is quasi-optimal, it converges at the same
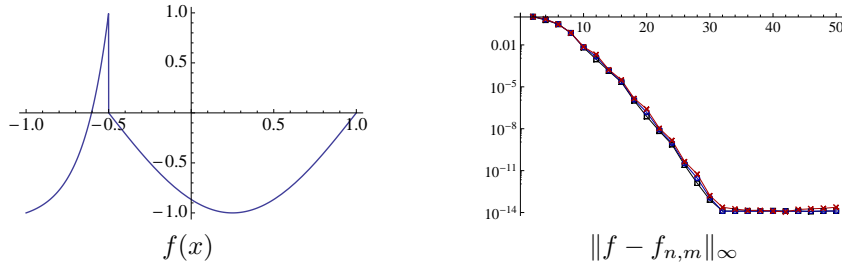
$$f(x) \qquad\qquad \|f - f_{n,m}\|_\infty$$

Figure 3: The function (4.1) (left) and the quantity $\|f - f_{n,m}\|_\infty$ (right) against $\bar{n}$, with $n_0 = n_1$, $m = \lceil \frac{1}{4}\bar{n}^2 \rceil$ and $\gamma = -\frac{1}{2}, 0, \frac{1}{2}$ (squares, circles and crosses respectively).

rate in terms of n, as shown in the figures. This translates into root-exponential convergence in terms of the number of Fourier coefficients m.

The function (2) was used in [30] to test spectral reprojection. Therein it was found that more than 1024 Fourier coefficients were required to obtain close to machine precision. Conversely, with the GS approach we achieve the same value using few than 256 such coefficients—a factor of four increase (a similar experiment was reported in [6]). Moreover, it is worth recalling that (see §1), with this approach one has near-complete freedom to choose the polynomial basis for reconstruction, whereas with spectral reprojection one must use a specific, Gibbs complementary, basis.

5 Stability

In this final section, we discuss the stability of our approach with respect to perturbations.

There are two main types of perturbations one encounters in this problem. The first, due to noise, is easily dealt with. Since our method has a bounded condition number, noise of magnitude ϵ results in an additional error in the final approximation proportional to ϵ.

The second perturbation involves the discontinuity locations $x_1, \ldots, x_l$ of the function f. Thus far, we have assumed that these are known exactly. However, as discussed in §1.3, in practice these values must be precomputed by an edge detection algorithm. This procedure incurs an error, and therefore it is vital that any reconstruction algorithm be robust in this regard. In the remainder of this section we analyse the robustness of the GS reconstruction (2.4) in the presence of such errors.

Suppose that approximate singularity locations $\tilde{x}_1, \ldots, \tilde{x}_l$ have been computed to within an accuracy of ϵ of the true locations $x_1, \ldots, x_l$, i.e. $|\tilde{x}_r - x_r| \leq \epsilon$ for $r = 1, \ldots, l$. We construct our piecewise polynomial approximation space $\tilde{T}_n$ based on these inexact values and compute the GS reconstruction $\tilde{f}_{n,m}$. Note that $\tilde{T}_n$ is a piecewise polynomial reconstruction space with the same parameters as T_n, where T_n is the space corresponding to the exact discontinuity locations $x_1, \ldots, x_l$. Hence $\tilde{T}_n$ and T_n asymptotically have the same stable sampling rate, and therefore the reconstruction $\tilde{f}_{n,m}$ is also quasi-optimal. Hence, to understand the effect of the inexact values $\tilde{x}_1, \ldots, \tilde{x}_l$ it suffices to only consider the orthogonal projection $\tilde{\mathcal{Q}}_n f$ onto the subspace $\tilde{T}_n$.

Let us briefly explain how such errors affect the approximation $\tilde{\mathcal{Q}}_n f$. Suppose that f is piecewise analytic with jumps at $x_1, \ldots, x_l$. The projection $\tilde{\mathcal{Q}}_n f$ is an approximation to f using piecewise polynomials defined in the *inexact* regions $\tilde{I}_r = [\tilde{x}_r, \tilde{x}_{r+1})$. When restricted to $\tilde{I}_r$, the function f is not analytic. In fact, it will have a jump discontinuity at any point in the set $\{x_1, \ldots, x_l\} \cap \tilde{I}_r$. Thus the error in the reconstruction stemming from inexact jump locations is due to approximating a function which is now only piecewise analytic in each prescribed region $\tilde{I}_r$.

Since the approximation error decouples into separate subintervals, it suffices to consider the error in any one. By scaling, and by assuming that ϵ is small in comparison to $\min_r |I_r|$, we may therefore consider the following problem:

Problem 5.1. *Suppose that $f : [-1, 1] \to \mathbb{C}$ is such that the restrictions of f to the intervals $I_0 = [-1, -1 + \epsilon_1]$, $I_1 = [-1 + \epsilon_1, 1 - \epsilon_2]$ and $I_2 = [1 - \epsilon_2, 1]$ are analytic, where $0 < \epsilon_1, \epsilon_2 < 1$. Let $\tilde{T}_n$ be an orthonormal basis of Jacobi polynomials with parameters $\alpha, \beta > -1$ on the interval $[-1, 1]$. Determine the error of the orthogonal projection $\tilde{\mathcal{Q}}_n f$ of f onto $\tilde{T}_n$ in terms of n and $\epsilon = \max\{\epsilon_1, \epsilon_2\}$.*

22

The question now arises, which error metric on $[-1, 1]$ should one choose? Clearly the uniform norm is ill-suited for this purpose, since it would be $\mathcal{O}(1)$ regardless of the magnitude of ϵ. The L^2-norm is also a poor choice, because it would lead to the error behaving as $\mathcal{O}(\sqrt{\epsilon})$. To see this, note that the function g that is zero on $[-1 + \epsilon_1, 1 - \epsilon_2]$ and takes value 1 elsewhere has L^2-norm equal to $\sqrt{|I_0|} + \sqrt{|I_2|} = \mathcal{O}(\sqrt{\epsilon})$.

We shall use the L^1-norm instead:

$$\|g\|_{\mathrm{L}^1} = \int_{-1}^{1} |g(x)| \, \mathrm{d}x, \quad g \in \mathrm{L}^1(-1, 1).$$

We have

Lemma 5.2. *Let $f \in \mathrm{L}^\infty(-1, 1)$, and suppose that $\tilde{\mathcal{Q}}_n f$ is as in Problem 5.1. Then*

$$\|f - \tilde{\mathcal{Q}}_n f\|_{\mathrm{L}^1} \lesssim \|\tilde{f} - \tilde{\mathcal{Q}}_n \tilde{f}\|_{\mathrm{L}^1} + \epsilon \left(1 + \epsilon^q n^{2p+2}\right) \left(\|f\|_\infty + \|\tilde{f}\|_\infty\right), \tag{5.1}$$

where $p = \max\{\alpha, \beta\}$, $q = \min\{\alpha, \beta\}$ and $\tilde{f} \in \mathrm{L}^\infty(-1, 1)$ is any function such that $\tilde{f} = f$ a.e. in the interval $[-1 + \epsilon_1, 1 - \epsilon_2]$.

Proof. We have
$$\|f - \tilde{\mathcal{Q}}_n f\|_{\mathrm{L}^1} \leq \|\tilde{f} - \tilde{\mathcal{Q}}_n \tilde{f}\|_{\mathrm{L}^1} + \|f - \tilde{f}\|_{\mathrm{L}^1} + \|\tilde{\mathcal{Q}}_n(f - \tilde{f})\|_{\mathrm{L}^1}. \tag{5.2}$$
Since $f - \tilde{f}$ is supported only outside I_1, we have

$$\|f - \tilde{f}\|_{\mathrm{L}^1} = \int_{I_0 \cup I_2} |f(x) - \tilde{f}(x)| \, \mathrm{d}x \leq \epsilon \left(\|\tilde{f}\|_\infty + \|f\|_\infty\right). \tag{5.3}$$

Write $\{\phi_j\}_{j \in \mathbb{N}}$ for the orthonormal basis of Jacobi polynomials corresponding to the projection $\tilde{\mathcal{Q}}_n$. Then

$$\begin{aligned}
\|\tilde{\mathcal{Q}}_n(f - \tilde{f})\|_{\mathrm{L}^1} &\leq \sum_{j=0}^{n} \|\phi_j\|_{\mathrm{L}^1} |\langle f - \tilde{f}, \phi_j \rangle_{\alpha, \beta}| \\
&\leq \left(\|\tilde{f}\|_\infty + \|f\|_\infty\right) \sum_{j=0}^{n} \|\phi_j\|_{\mathrm{L}^1} \|\phi_j\|_\infty \int_{I_0 \cup I_2} w_{\alpha, \beta}(x) \, \mathrm{d}x \\
&\lesssim \left(\|\tilde{f}\|_\infty + \|f\|_\infty\right) \epsilon^{1+q} \sum_{j=0}^{n} \|\phi_j\|_\infty^2.
\end{aligned}$$

Recall from the proof of Theorem 3.7 that we have $\|\phi_j\|_\infty \sim j^{p+\frac{1}{2}}$. Hence

$$\|\tilde{\mathcal{Q}}_n(f - \tilde{f})\|_{\mathrm{L}^1} \lesssim \left(\|\tilde{f}\|_\infty + \|f\|_\infty\right) \epsilon^{1+q} \sum_{j=0}^{n} j^{2p+1} \lesssim \left(\|\tilde{f}\|_\infty + \|f\|_\infty\right) \epsilon^{1+q} n^{2p+2}.$$

Combining this with (5.2) and (5.3) now gives the result. $\qquad\square$

This lemma allows one to understand the convergence of $\tilde{\mathcal{Q}}_n f$. Suppose that f is analytic in $[-1 + \epsilon_1, 1 - \epsilon_2]$. For sufficiently small ϵ, f has an analytic continuation to $[-1, 1]$. Let $\tilde{f}$ be such continuation. Then the first term in (5.1) converges exponentially fast in n, and therefore for all moderately large n the error is determined by the second factor $\epsilon(1 + \epsilon^q n^{2p+2})$. From this we immediately deduce the following result:

Theorem 5.3. *Let f be piecewise analytic with jumps at $x_1, \ldots, x_l$. Suppose that $|x_r - \tilde{x}_r| \leq \epsilon$, $r = 1, \ldots, l$, and let $\tilde{\mathcal{Q}}_n f$ be the piecewise Jacobi polynomial expansion of f based on the inexact jump values $\tilde{x}_1, \ldots \tilde{x}_r$. Then, for sufficiently large n, the error*

$$\|f - \tilde{\mathcal{Q}}_n f\|_{\mathrm{L}^1} \lesssim \epsilon(1 + \epsilon^q n^{2p+2}),$$

where $p = \max\{\alpha_0, \ldots, \alpha_l, \beta_0, \ldots, \beta_l\}$ and $q = \min\{\alpha_0, \ldots, \alpha_l, \beta_0, \ldots, \beta_l\}$.
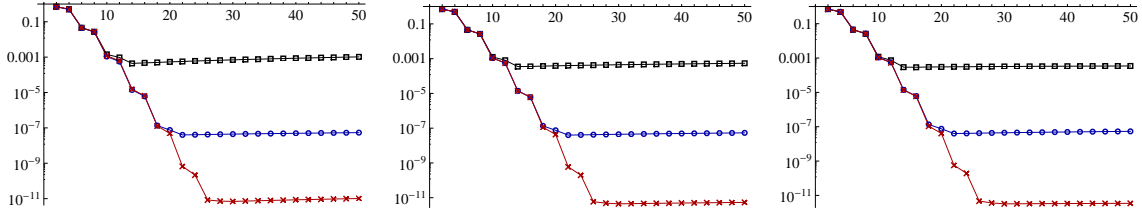
Figure 4: The error $\|f - \tilde{f}_{n,m}\|_{L^1}$ against $\bar{n}$ in approximating the function (4.1) with $n_0 = n_1$, $m = \lceil \frac{1}{4}\bar{n}^2 \rceil$ and $\gamma = -\frac{1}{2}, 0, \frac{1}{2}$ (left, middle, right). The edge $x_1 = -\frac{1}{2}$ was perturbed by an amount ϵ, giving the approximate edge $\tilde{x}_1 = -\frac{1}{2} + \epsilon$. The values of ϵ used were $10^{-4}, 10^{-8}, 10^{-12}$ (squares, circles and crosses respectively).

This theorem implies that whenever $q \geq 0$, such is the case for Legendre ($p = q = 0$) and second kind Chebyshev polynomials ($p = q = \frac{1}{2}$), errors of magnitude ϵ incurred from edge detection will, up to some mildly growing factor in n, correspond to errors of at worst $\mathcal{O}(\epsilon)$ in the overall reconstruction $\tilde{f}_{n,m}$. Thus the method is extremely robust in this respect. Such robustness is illustrated in Figure 4, which gives numerical results for the reconstruction $\tilde{f}_{n,m}$ from inexact edge data.

Recall that Chebyshev polynomials of the first kind correspond to the case $p = q = -\frac{1}{2}$. Theorem 5.3 suggests that it may be inadvisable to use such polynomials in applications where there are edge detection errors: the bound (5.1) implies that an error of magnitude ϵ in edge detection is translated into an error of magnitude $\sqrt{\epsilon}$ in the overall reconstruction. Note that the reason that the error worsens is that the weight function $w_{\alpha,\beta}$, become increasingly singular at the endpoints as $q \to -1^+$. This has the effect of amplifying any edge errors.

However, numerical results in Figure 4 suggest that theorem does not accurately predict the magnitude of edge detection errors in this case. In fact, much as in the case of Legendre or second kind Chebyshev polynomials, the error appears to remain of magnitude ϵ. Hence, according to these results, first kind Chebyshev polynomials can in fact be used for applications with edge detection errors.

The underlying reason for the disparity between these results and Theorem 5.3 is that the theorem concerns the orthogonal projection $\tilde{\mathcal{Q}}_n f$, and not the reconstruction $\tilde{f}_{n,m}$ itself. Numerical experiments (not shown here) suggest that the bound (5.1) is sharp for the orthogonal projection $\tilde{\mathcal{Q}}_n f$ in terms of the behaviour with respect to ϵ (the polynomial growth in n appears to be a nonphysical artefact of the proof). However, whilst quasi-optimality of the mapping $f \mapsto \tilde{f}_{n,m}$ guarantees that $\tilde{\mathcal{Q}}_n f$ and $\tilde{f}_{n,m}$ are close, it does so only in the sense of the $L^2_{w_{\alpha,\beta}}$-norm. But, as discussed, we measure edge detection error in the L^1-norm. This fact appears to save the approximation $\tilde{f}_{n,m}$ from the bad propagation of errors predicted by Theorem 5.3. In particular, it gives very much the same performance as in the case $q \geq 0$, as can be seen in Figure 4.

6   Conclusions and challenges

In this paper we have introduced a general framework for sampling and reconstruction in Hilbert spaces, and analyzed in detail the resulting numerical method for reconstructing a piecewise analytic function in a orthogonal basis of (piecewise) polynomials from its Fourier coefficients. The method can be implemented with a wide variety of different polynomial bases, including Chebyshev and Legendre polynomial bases. In all cases the numerical method is stable and exponentially convergent (in the polynomial degree n). Whenever Chebyshev or Legendre polynomials are employed, for example, the method is also root-exponentially convergent in the number of Fourier coefficients m. Finally, we have shown that this method is robust with respect to edge detection errors.

An important challenge for future work is to make a more detailed comparison between this and other approaches (such as spectral reprojection) for recovering piecewise analytic functions from their Fourier data. As mentioned, there are many approaches for this problem, yet few comparisons can be found in the literature. Of singular interest and importance is investigating the robustness of different methods to noise and edge detection errors. We remark at this stage that we have not seen any analysis of the type given in §5 for other reconstruction methods.

This aside, it is worth repeating that the GS framework developed in this paper is far more general than

the problem considered in §3. Namely, it allows one to reconstruct any element $f \in U$ from its samples in an arbitrary orthonormal basis, even in the case that the sampling and reconstruction vectors are orthogonal with respect to different inner products. In fact, this paper forms part of a much larger project on numerical methods for sampling and reconstruction (see [4, 5, 6, 7, 8]), with a long list of potential applications. Currently we are investigating a number of such applications, including spline and wavelet-based reconstructions of images, the solution of certain inverse and ill-posed problems (see [7]), and reconstructions from other types of integral transforms. It is also worth mentioning that the key ideas for GS originated from new tools for the computational spectral problem [37, 38]. Furthermore, similar ideas, when combined with convex optimization techniques, can be used to extended the current theory of compressed sensing to formally infinite-dimensional problems. This topic is described in [4].

Finally, we note that the general framework developed in this paper can be immediately applied to a very much related problem: namely, the recovery of a piecewise analytic function from its coefficients with respect to an orthogonal polynomial basis (as opposed to its Fourier coefficients). Much like the Fourier case, the problem arises in spectral discretizations of hyperbolic PDEs [31, 33, 40]. Investigating this application, and the critically, the analysis of the resulting stable sampling rate, is a topic for future work.

References

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

[2] B. Adcock. Convergence acceleration of modified Fourier series in one or more dimensions. *Math. Comp.*, 80(273):225–261, 2011.

[3] B. Adcock. Gibbs phenomenon and its removal for a class of orthogonal expansions. *BIT*, 51(1):7–41, 2011.

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

[5] B. Adcock and A. C. Hansen. A generalized sampling theorem for stable reconstructions in arbitrary bases. *J. Fourier Anal. Appl.*, 18(4):685–716, 2012.

[6] B. Adcock and A. C. Hansen. Stable reconstructions in Hilbert spaces and the resolution of the Gibbs phenomenon. *Appl. Comput. Harmon. Anal.*, 32(3):357–388, 2012.

[7] B. Adcock, A. C. Hansen, E. Herrholz, and G. Teschke. Generalized sampling: extension to frames and inverse and ill-posed problems. *Inverse Problems*, 29(1):015008, 2013.

[8] B. Adcock, A. C. Hansen, and C. Poon. Beyond consistent reconstructions: optimality and sharp bounds for generalized sampling, and application to the uniform resampling problem. *SIAM J. Math. Anal. (accepted)*, 2013.

[9] B. Adcock, A. C. Hansen, and A. Shadrin. A stability barrier for reconstructions from Fourier samples. *Preprint*, 2012.

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

[11] 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.

[12] D. Batenkov. Complete algebraic reconstruction of piecewise-smooth functions from fourier data. *Preprint*, 2012.

[13] D. Batenkov and Y. Yomdin. Algebraic Fourier reconstruction of piecewise smooth functions. *Math. Comp.*, 81:277–318, 2012.

[14] B. Beckermann, V. Kalyagin, A. Matos, and F. Wielonsky. How well does the Hermite–Padé approximation smooth the Gibbs phenomenon? *Math. Comp.*, 80:931–958, 2011.

[15] N. Borovykh and M. N. Spijker. Bounding partial sums of Fourier series in weighted L^2-norms, with applications to matrix analysis. *J. Comput. Appl. Math.*, 147(2):349–368, 2002.

[16] P. Borwein and T. Erdélyi. *Polynomials and Polynomial Inequalities*. Springer–Verlag, New York, 1995.

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

[18] 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.

[19] 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.

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

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

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

[24] K. S. Eckhoff. Accurate and efficient reconstruction of discontinuous functions from truncated series expansions. *Math. Comp.*, 61(204):745–763, 1993.

[25] K. S. Eckhoff. Accurate reconstructions of functions of finite regularity from truncated Fourier series expansions. *Math. Comp.*, 64(210):671–690, 1995.

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

[27] J. B. Garnett. *Bounded analytic functions*. Springer, New York, 2007.

[28] 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.

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

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

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

[32] 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.

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

[34] 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.

[35] B.-Y. Guo, J. Shen, and L.-L. Wang. Optimal spectral-Galerkin methods using generalized Jacobi polynomials. *J. Sci. Comput.*, 27(1–3):305–322, 2006.

[36] B.-Y. Guo, J. Shen, and L.-L. Wang. Generalized Jacobi polynomials/functions and their applications. *Appl. Numer. Math.*, 59:1011–1028, 2009.

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

[38] 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.

[39] H. Helson and Szegö. A problem in prediction theory. *Ann. Mat. Pura Appl.*, 51(107–138), 1960.

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

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

[42] R. Hunt, B. Muckenhoupt, and R. Wheeden. Weighted norm inequalities for the conjugate function and Hilbert transform. *Trans. Amer. Math. Soc.*, 176(227–251), 1973.

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

[44] 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.

[45] A. B. Kuijlaars, K. T.-R. McLaughlin, W. Van Assche, and M. Vanlessen. The Riemann–Hilbert approach to strong asymptotics of orthogonal polynomials on $[-1, 1]$. *Adv. Math.*, 188:337–398, 2004.

[46] J. N. Lyness. Computational techniques based on the Lanczos representation. *Math. Comp.*, 28(125):81–123, 1974.

[47] J. Steinberg. Oblique projections in Hilbert spaces. *Integr. equ. oper. theory*, 38(1):81–119, 2000.

[48] D. Szyld. The many proofs of an identity on the norm of oblique projections. *Numer. Algorithms*, 42:309–323, 2006.

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

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

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