Frames and Numerical Approximation II: Generalized Sampling

Ben Adcock[1] · Daan Huybrechs[2]

Abstract

In a previous paper (Adcock and Huybrechs in SIAM Rev 61(3):443–473, 2019) we described the numerical approximation of functions using redundant sets and frames. Redundancy in the function representation offers enormous flexibility compared to using a basis, but ill-conditioning often prevents the numerical computation of best approximations. We showed that, in spite of said ill-conditioning, approximations with regularization may still provide accuracy up to order $\sqrt{\epsilon}$, where ϵ is a small truncation threshold. When using frames, i.e. complete systems that are generally redundant but which provide infinite representations with coefficients of bounded norm, this accuracy can actually be achieved for all functions in a space. Here, we generalize that setting in two ways. We assume information or samples from f from a wide class of linear operators acting on f, rather than inner products associated with the best approximation projection. This enables the analysis of fully discrete approximations based, for instance, on function values only. Next, we allow oversampling, leading to least-squares approximations. We show that this leads to much improved accuracy on the order of ϵ rather than $\sqrt{\epsilon}$. Overall, we demonstrate that numerical function approximation using redundant representations may lead to highly accurate approximations in spite of having to solve ill-conditioned systems of equations.

1 Introduction

The approximation of functions in a Hilbert space typically assumes a basis for that space. The non-redundancy of a basis ensures that linear systems associated with approximation problems are non-singular. In addition, suitable structure—ideally the

---

✉ Daan Huybrechs
daan.huybrechs@cs.kuleuven.be

1 Department of Mathematics, Simon Fraser University, Burnaby, Canada

2 Department of Computer Science, KU Leuven, Leuven, Belgium

basis is orthonormal, more generally it may be a Riesz basis—renders these systems well-conditioned. There is a unique solution, it is stably computable, and there is a close correspondence between properties of the continuous function and of the coefficients in the representation, for example the Parseval identity.

Instead, for a redundant set of functions the corresponding linear systems may be ill-conditioned or even singular, and uniqueness may be lost. Still, good approximations may exist in the span of the set. It may even be much easier to ensure that this is the case than it is for a basis, and in fact this is a popular approach in a wide range of applications. For example, a basis can be 'enriched' by adding a few functions that capture a singularity [10]. A periodic Fourier basis can be augmented with a few polynomials to capture the non-periodicity of f [23]. Ill-conditioning and redundancy also frequently appear in solution methods for partial differential equations (PDEs). In Trefftz methods, solutions of a PDE are approximated using other solutions of the same PDE, and this is often successful yet notoriously ill-conditioned [16]. Several methods are based on embedding a domain with complicated geometry Ω, for which a basis is unknown, into a simple bounding box D. A basis for $L^2(D)$ yields a (redundant) frame for $L^2(\Omega)$. Examples include embedded/fictitious domain methods, immersed boundary methods and others [6,17,24].

As mentioned, redundant representations necessarily lead to ill-conditioning. To which extent are the corresponding function approximations computable? What convergence behaviour and accuracy can be expected? In this paper, we continue a line of investigation that commenced in [3] on numerical approximation of functions using redundant function sets in general and frames in particular. The main contribution of [3] was a detailed analysis of the accuracy and conditioning of the computation of best approximations with regularization, with a chosen threshold ϵ. We now briefly recall the main results of [3] in Sect. 1.1, followed by an overview of the theoretical results of this paper in Sect. 1.2.

1.1 Best Approximation With Regularization

The main concern of [3] was the computation of the best approximation, i.e. the orthogonal projection, in the space $\mathrm{H}_N := \mathrm{span}(\Phi_N)$ spanned by a set of N elements $\Phi_N := \{\phi_n\}_{n=1}^N$. This approximation is given by $\mathcal{P}_N f = \sum_{n=1}^N x_n \phi_n$, where $x = (x_n)_{n=1}^N$ is a solution of the linear system

$$G_N x = y, \tag{1.1}$$

where $y = \{\langle f, \phi_n \rangle\}_{n=1}^N$ and $G_N = \{\langle \phi_n, \phi_m \rangle\}_{m,n=1}^N$ is the Gram matrix of Φ_N.

If the elements of Φ_N are nearly or exactly linearly dependent, then G_N is ill-conditioned or even singular. Moreover, the coefficients x can also grow arbitrarily large, making them impossible to compute in floating point arithmetic for sufficiently large N. The remedy proposed in [3] was to regularize (1.1) by using a truncated Singular Value Decomposition (SVD) of G_N with a threshold parameter $\epsilon > 0$ below which all the singular values are discarded. This results in a new projection $\mathcal{P}_N^\epsilon f = \sum_{n=1}^N (x^\epsilon)_n \phi_n$, where x^ϵ is the regularized solution of (1.1).

The main result of [3] concerns the best approximation with regularization as follows:

Theorem 1.1 [3] *The truncated SVD projection* $\mathcal{P}_N^\epsilon$ *satisfies*

$$\|f - \mathcal{P}_N^\epsilon f\| \le \inf_{z \in \mathbb{C}^N} \left\{ \left\| f - \sum_{n=1}^{N} z_n \phi_n \right\| + \sqrt{\epsilon} \|z\| \right\}. \tag{1.2}$$

Moreover, the (absolute) condition number of the mapping $y \mapsto \mathcal{P}_N^\epsilon f$ *is at most* $1/\sqrt{\epsilon}$.

Observe that the right hand side of (1.2) contains two terms. Theorem 1.1 states that the regularized projection behaves like the best approximation to f in the span of Φ_N (the first term), as long as the coefficients have sufficiently small norm (second term). Furthermore, convergence can only be expected to an accuracy up to the order of $\sqrt{\epsilon}$. Whether or not this accuracy is achieved, depends on the existence of a representation $\sum_{n=1}^{N} z_n \phi_n$ in the span of Φ_N with that accuracy and with small norm $\|z\|$ of its coefficients. This question can be studied on a case-by-case basis, as done in [3] for a variety of examples.

To answer this question more generally, it is natural to impose additional structure on Φ_N. In [3], this was done via frames. Recall that an indexed family $\Phi := \{\phi_n\}_{n=1}^{\infty}$ is a frame for a Hilbert space H if it satisfies the *frame condition*

$$A\|f\|^2 \le \sum_{n=1}^{\infty} |\langle f, \phi_n \rangle|^2 \le B\|f\|^2, \qquad \forall f \in \mathrm{H}, \tag{1.3}$$

where $A, B > 0$ are positive constants and $\|\cdot\|$ is the norm on H. The frame condition ensures the existence of bounded representations, to any accuracy, for all functions in the space. In particular, this yields the following:

Corollary 1.2 *If* $\Phi := \{\phi_n\}_{n=1}^{\infty}$ *is a frame for* H, *and* $\Phi_N := \{\phi_n\}_{n=1}^{N}$, *then*

$$\limsup_{N \to \infty} \|f - \mathcal{P}_N^\epsilon f\| \le \sqrt{\frac{\epsilon}{A}} \|f\|, \qquad \forall f \in \mathrm{H}. \tag{1.4}$$

Unlike in the general setting, the frame condition imposes sufficient structure so that accuracy to order $\sqrt{\epsilon}$ is now guaranteed for all functions in H. For this reason, as well as the fact that frames occur in numerous computational problems, one can think of frames as an ideal setting in which to apply Theorem 1.1. Of course function approximation with redundancy can be successful without a frame property. For example, in the absence of a frame, one may still use Theorem 1.1 to show accuracy in a subspace of H consisting of functions with bounded-norm coefficient representations. But this raises the matter of whether functions of interest to a given problem belong to this space. In the absence of a frame structure, this question must then be answered on a case-by-case basis.

1.2 Main Results

In this paper, we generalize Theorem 1.1 using *oversampling* and allowing for *generalized samples* (or *indirect data*). In doing so, we not only allow a much broader class of samples, including, for example, pointwise evaluations, we also overcome the $\sqrt{\epsilon}$ bottleneck.

Approximation from generalized samples In Sect. 3, instead of inner products with the elements ϕ_n, as was used in Theorem 1.1 [see (1.1)], the 'data' about the function f is now given by bounded linear functionals $\ell_{m,M} : G \to \mathbb{C}$, $m = 1, \ldots, M$, which may depend on M and which may be only defined on a dense subspace G of H (e.g. in the case of pointwise evaluations when $H = L^2(\Omega)$ we consider $G = C(\overline{\Omega})$). Very much reminiscent of the frame condition (1.3), the strongest general statements can be made when this data is sufficiently 'rich' so as to stably recover f, in particular satisfying

$$A' \|f\|^2 \le \liminf_{M \to \infty} \sum_{m=1}^{M} |\ell_{m,M}(f)|^2 \le \limsup_{M \to \infty} \sum_{m=1}^{M} |\ell_{m,M}(f)|^2 \le B' \|f\|^2, \quad \forall f \in G,$$

$$(1.5)$$

for constants $A', B' > 0$. We refer to this as a *Marcinkiewicz-Zygmund* condition. A key ingredient is to allow oversampling, i.e. let $M > N$, and consider the $M \times N$ linear system

$$\boldsymbol{G}_{M,N} \boldsymbol{x} \approx \boldsymbol{y}, \qquad \boldsymbol{y} = \{\ell_{m,M}(f)\}_{m=1}^{M}, \tag{1.6}$$

where $\boldsymbol{G}_{M,N} = \{\ell_{m,M}(\phi_n)\}_{m,n=1}^{M,N}$. As we shall subsequently explain, this system generally remains ill-conditioned for large N, even when $M \gg N$. Hence we construct an approximation by singular value thresholding. This leads to a regularized approximation $\mathcal{P}_{M,N}^{\epsilon} f$ whose coefficients $\boldsymbol{x}^{\epsilon}$ are the solution of the SVD-regularized least-squares problem corresponding to (1.6). Our main result for this setup is the following:

Theorem 1.3 *The truncated SVD projection* $\mathcal{P}_{M,N}^{\epsilon} f$ *satisfies*

$$\left\| f - \mathcal{P}_{M,N}^{\epsilon} f \right\| \le \inf_{z \in \mathbb{C}^N} \left\{ \left\| f - \sum_{n=1}^{N} z_n \phi_n \right\| + \kappa_{M,N}^{\epsilon} \left\| f - \sum_{n=1}^{N} z_n \phi_n \right\|_M + \epsilon \lambda_{M,N}^{\epsilon} \|z\| \right\},$$

$$(1.7)$$

for constants $\kappa_{M,N}^{\epsilon}, \lambda_{M,N}^{\epsilon} > 0$. *The (absolute) condition number of the mapping* $\boldsymbol{y} \mapsto \mathcal{P}_{M,N}^{\epsilon} f$ *is precisely* $\kappa_{M,N}^{\epsilon}$. *Moreover, these constants satisfy*

$$\kappa_{M,N}^{\epsilon} \le \frac{\sqrt{B_N}}{\epsilon}, \qquad \lambda_{M,N}^{\epsilon} \le \frac{\sqrt{B_N}}{\epsilon},$$

for all M and N, $M \geq N$, where B_N is the Bessel constant of Φ_N over $H_N =$ span(Φ_N). *If the sampling functionals satisfy* (1.5) *then, for fixed N,*

$$\limsup_{M \to \infty} \kappa^\epsilon_{M,N} \leq \frac{1}{\sqrt{A'}}, \qquad \limsup_{M \to \infty} \lambda^\epsilon_{M,N} \leq \frac{1}{\sqrt{A'}}.$$

Here, $\|g\|^2_M = \sum_{m=1}^{M} |\ell_{m,M}(g)|^2$ is the discrete semi-norm defined by the data. We recall that Φ_N is a Bessel sequence, since it is finite, and therefore it has a finite Bessel constant $B_N > 0$, defined as the smallest constant for which $\sum_{n=1}^{N} |\langle f, \phi_n \rangle|^2 \leq B_N \|f\|^2$, $\forall f \in H$.

This result is very general, but its main conclusion is the following. Provided M is sufficiently large and the samples satisfy (1.5) then the approximation error depends on $f - \sum_{n=1}^{N} z_n \phi_n$ (measured in some norm) and $\epsilon \|z\|$. The constants $\kappa^\epsilon_{M,N}$ and $\lambda^\epsilon_{M,N}$ can be large when M is insufficiently large (behaving like $\sqrt{B_N}/\epsilon$ in the worse case), but they can be made arbitrarily close to $1/\sqrt{A'}$ by taking M large, i.e. by increasing oversampling. Furthermore, on comparison with Theorem 1.1, we notice that $\sqrt{\epsilon}$ in the error bound has been replaced by ϵ. Hence, under suitable assumptions on f, Φ_N and M, we expect order ϵ accuracy in the limit, as opposed to order $\sqrt{\epsilon}$ accuracy.

This result raises the question of how large M needs to be in comparison to N. Later, we quantify this via the so-called *stable sampling rate*. Unsurprisingly, the question 'how large is sufficiently large' depends completely on system Φ_N and samples $\{\ell_{m,M}\}$, and thus must be analyzed on a case-by-case basis. In Sect. 5, we illustrate an example for which the stable sampling rate is provably linear, i.e. there exists a $C \geq 1$ such that setting $M \geq CN$ implies that the constants $\kappa^\epsilon_{M,N}$ and $\lambda^\epsilon_{M,N}$ are bounded independently of ϵ. Alternatively, this rate can also be computed numerically, as we explain in Sect. 3.6.

Frame approximation from frame samples Theorem 1.3 applies for arbitrary Φ_N and any samples satisfying (1.5). In order to make statements about the limiting behaviour as $M, N \to \infty$ we need two ingredients. First, a sequence z for which $\sum_{n=1}^{N} z_n \phi_n$ converges to f in the Hilbert space norm and for which the coefficient norm $\|z\|$ does not blow up. Second, additional regularity of the samples and/or f so that the M-norm can be controlled by a suitable norm in which one also has $\sum_{n=1}^{N} z_n \phi_n \to f$.

Whether or not such conditions hold could be answered on a case-by-case basis for particular types of sampling and approximation systems. But instead, we now address them in a general scenario where both the sampling functionals and the approximation system Φ_N are endowed with a frame structure. Specifically, let $\Phi = \{\phi_n\}_{n=1}^{\infty}$ and $\Psi = \{\psi_m\}_{m=1}^{\infty}$ be frames for H and set $\Phi_N = \{\phi_n\}_{n=1}^{N}$ and $\ell_{m,M} = \langle \cdot, \psi_m \rangle$ for $m = 1, \ldots, M$. Note that (1.5) now automatically holds with G = H, with A' and B' being the frame bounds for Ψ.

Our main result in this case is the following:

Theorem 1.4 *In the above setting, the truncated SVD projection $\mathcal{P}^\epsilon_{M,N} f$ satisfies*

$$\|f - \mathcal{P}^\epsilon_{M,N} f\| \leq \inf_{z \in \mathbb{C}^N} \left\{ (1 + \sqrt{B'} \kappa^\epsilon_{M,N}) \left\| f - \sum_{n=1}^{N} z_n \phi_n \right\| + \epsilon \lambda^\epsilon_{M,N} \|z\| \right\},$$

for certain constants $\kappa_{M,N}^{\epsilon}, \lambda_{M,N}^{\epsilon} > 0$. The (absolute) condition number of the mapping $\mathbf{y} \mapsto \mathcal{P}_{M,N}^{\epsilon} f$ is precisely $\kappa_{M,N}^{\epsilon}$. These constants satisfy

$$\kappa_{M,N}^{\epsilon}, \lambda_{M,N}^{\epsilon} \leq \begin{cases} \sqrt{B}/\epsilon & \Psi \neq \Phi \\ 1/\sqrt{\epsilon} & \Psi = \Phi, \end{cases}$$

for all M and N, $M \geq N$, and

$$\limsup_{M \to \infty} \kappa_{M,N}^{\epsilon} \leq \frac{1}{\sqrt{A'}}, \qquad \limsup_{M \to \infty} \lambda_{M,N}^{\epsilon} \leq \frac{1}{\sqrt{A'}}.$$

Moreover, for each $1 < \theta < \infty$ and $N \in \mathbb{N}$ there exists an integer $\Theta^{\epsilon}(N, \theta) \in \mathbb{N}$ such that

$$\limsup_{\substack{M,N \to \infty \\ M \geq \Theta^{\epsilon}(N,\theta)}} \|f - \mathcal{P}_{M,N}^{\epsilon} f\| \leq \epsilon \frac{\theta}{\sqrt{AA'}} \|f\|. \tag{1.8}$$

The quantity $\Theta^{\epsilon}(N, \theta)$ is what we term the stable sampling rate. As noted above, it can be computed numerically. Observe that in the special case $\Psi = \Phi$, the setting of Theorem 1.1 (i.e. sampling and approximating with the same functions) is restored. However, oversampling according to (1.8) overcomes the $\sqrt{\epsilon}$ bottleneck in (1.4), thus improving the limiting accuracy to order ϵ.

1.3 Relation to Other Work

This paper is a continuation of [3], in which the systematic study of numerical frame approximation was commenced. This study had its origins in earlier work on so-called *Fourier extensions* [5,15], which are particular frames arising as restrictions of the Fourier basis on a box to a subdomain.

Our use of oversampling here is inspired by earlier work on *generalized sampling* in Hilbert spaces by the first author and Hansen [1,2,4]. That work considered both sampling and approximation using orthonormal bases and frames, introducing the stable sampling rate as well, but did not address the ill-conditioning issue for approximation in the latter. Note that the matrices $\boldsymbol{G}_{M,N}$ (in the case of Theorem 1.4 with $\Psi = \Phi$) and $\boldsymbol{G}_N$ are so-called *uneven* and *finite* sections respectively of the infinite Gram matrix of the full frame Φ. Using uneven as opposed to finite sections is a well-known trick in computational spectral theory [13,14,18].

For a more in-depth discussion of relations between this work and standard frame theory, we refer to [3].

Our focus in this paper is accuracy and conditioning of the regularized frame approximations. We do not consider efficiency, i.e. computational time, which is very much dependent on the particular frame under consideration. There are efficient numerical methods for solving (1.6) for certain frames and sampling functionals [8,9,20–22]. In the absence of a more efficient method, the cost of computing the SVD of an $M \times N$ matrix with $M > N$ scales as $\mathcal{O}(MN^2)$. However, based in particular on [9], some of

the examples at the end of this paper can be implemented in $\mathcal{O}(N(\log N)^2)$ operations. We refer to [9] for more examples.

2 Preliminaries

We first introduce some notation and useful concepts from frame theory.

2.1 Bases and Frames

For the remainder of this paper, H is a separable Hilbert space over the field $\mathbb{C}$. We write $\langle \cdot, \cdot \rangle$ and $\|\cdot\|$ for the inner product and norm on H respectively. Recall that Φ is a *Riesz basis* if span(Φ) is dense in H, and there exist constants $A, B > 0$ such that

$$A\|\boldsymbol{x}\|^2 \leq \left\| \sum_{n \in I} x_n \phi_n \right\|^2 \leq B\|\boldsymbol{x}\|^2, \quad \forall \boldsymbol{x} = \{x_n\}_{n \in I} \in \ell^2(I). \tag{2.1}$$

Here and throughout, $\ell^2(I)$ denotes the space of square-summable sequences indexed over I, and $\|\cdot\|$ denotes its norm, i.e. $\|\boldsymbol{x}\| = \sqrt{\sum_{n \in I} |x_n|^2}$. In this paper, we assume that the constants appearing in (2.2) are the optimal constants such that the inequality holds. We note that a Riesz basis is an orthonormal basis if (2.1) holds with $A = B = 1$. In general, we may view (2.1) as a relaxed version of Parseval's identity.

An indexed family Φ is a *frame* if

$$A\|f\|^2 \leq \sum_{n \in I} |\langle f, \phi_n \rangle|^2 \leq B\|f\|^2, \quad \forall f \in \text{H}, \tag{2.2}$$

for positive constants $A, B > 0$. A frame is *tight* if $A = B$. We refer to (2.2) as the *frame condition*. Note that (2.2) implies that Φ is dense in H. It follows from (2.1) that a Riesz basis is also a frame with the same constants A, B [7, Prop. 3.6.4]. But, in general, a frame need not be a Riesz basis.

2.2 Operators on Frames

Associated to any frame Φ (and therefore any Riesz basis) is the so-called *synthesis* operator

$$\mathcal{T} : \ell^2(I) \to \text{H}, \quad \boldsymbol{y} = \{y_n\}_{n \in I} \mapsto \sum_{n \in I} y_n \phi_n.$$

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

$$\mathcal{T}^* : \text{H} \to \ell^2(I), \quad f \mapsto \{\langle f, \phi_n \rangle\}_{n \in I},$$

and the composition $S = TT^*$, known as the *frame* operator, is

$$S : H \to H, \quad f \mapsto \sum_{n \in I} \langle f, \phi_n \rangle \phi_n.$$

This operator is self-adjoint, bounded, invertible and positive with

$$A \|f\|^2 \le \langle S f, f \rangle \le B \|f\|^2. \tag{2.3}$$

See [7, Lemma 5.1.5]. Note that this inequality is equivalent to the frame condition (2.2). Note also that $S = \mathcal{I}$ is the identity operator for an orthonormal basis. Similarly, $S = A\mathcal{I}$ for a tight frame. However, for a general Riesz basis or frame, $S \ne \mathcal{I}$.

2.3 Dual Frames

A frame $\Psi = \{\psi_n\}_{n \in I} \subseteq H$ is a *dual frame* for a given frame Φ if

$$f = \sum_{n \in I} \langle f, \psi_n \rangle \phi_n = \sum_{n \in I} \langle f, \phi_n \rangle \psi_n, \quad \forall f \in H. \tag{2.4}$$

If a frame Φ is also a Riesz basis then it has a unique dual frame Ψ, which is also a Riesz basis. In this case, the pair (Φ, Ψ) is *biorthogonal*:

$$\langle \phi_n, \psi_m \rangle = \delta_{n,m}, \quad n, m \in I.$$

Note that an orthonormal basis is self-dual, i.e. $\Psi = \Phi$. Conversely, a frame and any of its duals are typically not biorthogonal. A frame may also have more than one dual. The so-called *canonical* dual frame of Φ is the frame $\Psi = \{S^{-1}\phi_n\}_{n \in I}$. This is a frame [7, Lemma 5.1.5], and its frame bounds are B^{-1} and A^{-1} respectively. In this case, (2.4) reads

$$f = \sum_{n \in I} \langle f, S^{-1}\phi_n \rangle \phi_n = \sum_{n \in I} \langle S^{-1}f, \phi_n \rangle \phi_n. \tag{2.5}$$

We refer to the coefficients $\boldsymbol{a} = \{\langle f, S^{-1}\phi_n \rangle\}_{n \in I}$ as the *canonical frame coefficients* of f. These coefficients have the property that, amongst all possible representations of f in Φ, they have the smallest norm [7, Lemma 5.4.2]. Specifically, if $f = \sum_{n \in I} a_n \phi_n = \sum_{n \in I} c_n \phi_n$ for some $\boldsymbol{c} = \{c_n\}_{n \in I}$, then $\|\boldsymbol{c}\| \ge \|\boldsymbol{a}\|$. Moreover, from the frame condition of the dual we have $\|\boldsymbol{a}\| \le \frac{1}{\sqrt{A}} \|f\|$.

3 Approximation from Indirect Data

In this section, we describe the general setup, which will lead eventually to Theorem 1.3. Throughout this section, the *approximation system* is defined by $\Phi_N =$

$\{\phi_n\}_{n \in I_N} \subset H$. It is an an indexed family of N elements in H, where I_N is an index set of cardinality N. For convenience, we now make a mild generalization over Sect. 1, allowing I_N to be an arbitrary index set rather than just $\{1, \ldots, N\}$. As noted, Φ_N is a Bessel sequence, since it is finite. We write $B_N > 0$ for its Bessel constant, i.e. the smallest constant for which $\sum_{n=1}^{N} |\langle f, \phi_n \rangle|^2 \leq B_N \|f\|^2, \forall f \in H$. We also define the operators

$$\mathcal{T}_N : \mathbb{C}^N \to H, \ \mathbf{z} = (z_n)_{n \in I_N} \mapsto \sum_{n \in I_N} z_n \phi_n$$

$$\mathcal{T}_N^* : H \to \mathbb{C}^N, \ f \mapsto (\langle f, \phi_n \rangle)_{n \in I_N}$$

$$\mathcal{S}_N = \mathcal{T}_N \mathcal{T}_N^* : H \to H, \ f \mapsto \sum_{n \in I_N} \langle f, \phi_n \rangle \phi_n.$$

3.1 Indirect Data

Let G be a dense subspace of the Hilbert space H endowed with a norm $\|\|\cdot\|\|$. Suppose that f, the function we seek to approximate, and Φ_N both belong to G. For each $M \in \mathbb{N}$, let J_M be an index set of cardinality $|J_M| = M$, and

$$\ell_{m,M} : G \to \mathbb{C}, \quad m \in J_M,$$

be a set of linear functionals which are bounded with respect to $\|\|\cdot\|\|$, i.e.

$$|\ell_{m,M}(f)| \leq c_{m,M} \|\|f\|\|, \qquad f \in G. \tag{3.1}$$

The data of f is given by

$$\mathbf{y} = \{\ell_{m,M}(f)\}_{m \in J_M}.$$

Write $\mathcal{M}_M : G \to \mathbb{C}^M$ for the mapping $\mathcal{M}_M f = \{\ell_{m,M}(f)\}_{m \in J_M}$. Our goal is to compute an approximation to f in Φ_N for some $N \leq M$ from this data.

In order to make meaningful general statements about the subsequent approximations we define, we require the data to be sufficiently rich. In analogy to the frame bounds (2.2), we shall often assume that there exist constants $A', B' > 0$ such that

$$A' \|f\|^2 \leq \liminf_{M \to \infty} \sum_{m \in J_M} |\ell_{m,M}(f)|^2 \leq \limsup_{M \to \infty} \sum_{m \in J_M} |\ell_{m,M}(f)|^2 \leq B' \|f\|^2, \quad \forall f \in G.$$

$$\tag{3.2}$$

We term this a *Marcinkiewicz–Zygmund* condition, because it is similar (although not identical) to the classical Marcinkiewicz–Zygmund inequality in approximation theory [19]; see also [12]. We comment further on this assumption and the constants involved in Sect. 3.5.

Before going any further, let us mention several examples of this framework:

Example 3.1 Suppose that $M = N$ and the samples of f are inner products with respect to the functions ϕ_m, i.e. $\ell_{m,M}(f) = \langle f, \phi_m \rangle$, $m = 1, \ldots, M$. Then this is precisely the setting of Theorem 1.1, and is a special case of the present setup with $G = H$, $\|\|\cdot\|\| = \|\cdot\|$. Note that (3.2) is precisely the frame condition (2.2); in particular, it holds with $A' = A$ and $B' = B$.

Example 3.2 Let $\Psi = \{\psi_m\}_{m \in J}$ be another frame (or Riesz/orthonormal basis) of H and consider samples of the form $\ell_{m,M}(f) = \langle f, \psi_m \rangle$, $m \in J_M$, where $\{J_M\}_{M \in \mathbb{N}}$ are nested index sets with $|J_M| = M$ and $\bigcup_{M=1}^{\infty} J_M = J$. This problem corresponds to sampling according to the frame Ψ and reconstructing in Φ, as in the original framework of generalized sampling [1,2]. We also have $G = H$ and $\|\|\cdot\|\| = \|\cdot\|$ in this case, and (3.2) holds with A', B' being the frame bounds for Ψ. In fact, it is straightforward to see that the the upper bound holds for any M in this case, i.e.

$$\|f\|_M \leq \sqrt{B'}\|f\|, \qquad \forall f \in H. \tag{3.3}$$

Example 3.3 Consider a frame for the Hilbert space $H = L^2(-1, 1)$ of square-integrable functions on $(-1, 1)$. Suppose that the samples are pointwise evaluations at equispaced points. In this case, we may take $G = C([-1, 1])$ with its usual norm, and $\ell_{m,M}(f) = \sqrt{2/M} f(-1 + 2(m - 1)/M)$, $m = 1, \ldots, M$. Then (3.2) holds with $A' = B' = 1$, since $\sum_{m=1}^{M} |\ell_{m,M}(f)|^2$ is a Riemann sum approximation to $\int_{-1}^{1} |f(x)|^2 \, dx$. More generally, if $-1 \leq t_{1,M} < \ldots < t_{M,M} \leq 1$ are (not necessarily equispaced) sample points, then (3.2) can be achieved with $A' = B' = 1$ if

$$h_M = \max_{0,\ldots,M} \{t_{m+1,M} - t_{m,M}\} \to 0, \qquad M \to \infty,$$

where $t_{0,M} = t_{M,M} - 2$ and $t_{M+1,M} = t_{1,M} + 2$. Indeed, in this case choosing the linear functionals as

$$\ell_{m,M}(f) = \sqrt{\frac{1}{2}(t_{m+1,M} - t_{m-1,M})} f(t_{m,M})$$

gives a convergent approximation $\sum_{m=1}^{M} |\ell_{m,M}(f)|^2$ to $\|f\|^2$ as $M \to \infty$.

3.2 Best Approximation with Regularization from Indirect Data

Given $f \in G$ and data $\mathcal{M}_M f$, we construct the approximation $\mathcal{P}_{M,N}^\epsilon f$ as follows. Let

$$\boldsymbol{G}_{M,N} = \mathcal{M}_M \mathcal{T}_N = \{\ell_{m,M}(\phi_n)\}_{m \in J_M, n \in I_N} \in \mathbb{C}^{M \times N}.$$

As we explain below, much like in the setting of Theorem 1.1, this matrix is generally ill-conditioned. This arises from the inherent redundancy of Φ_N, independently of the samples—in particular, it cannot be avoided by taking $M \gg N$. Hence we need to

regularize. Suppose that $G_{M,N}$ has singular value decomposition

$$G_{M,N} = U \Sigma V^*, \tag{3.4}$$

and let $\epsilon > 0$ be fixed. Then we set

$$x^\epsilon = (G_{M,N}^\epsilon)^\dagger y = V(\Sigma^\epsilon)^\dagger U^* y, \tag{3.5}$$

where $\dagger$ denotes the pseudoinverse and Σ^ϵ is the diagonal matrix with n^{th} entry σ_n if

$$\sigma_n > \epsilon, \tag{3.6}$$

and zero otherwise. The corresponding approximation to f is

$$f \approx \mathcal{P}_{M,N}^\epsilon f = \mathcal{T}_N x^\epsilon.$$

Observe that both the solution vector x^ϵ and the approximation $\mathcal{P}_{M,N}^\epsilon f$ are uniquely defined by construction, even if the family Φ_N happens to be linearly dependent. For convenience, we now define the mapping

$$\mathcal{L}_{M,N}^\epsilon : \mathbb{C}^M \to \mathrm{H}_N, \quad y \mapsto \mathcal{T}_N (G_{M,N}^\epsilon)^\dagger y,$$

so that

$$\mathcal{P}_{M,N}^\epsilon f = \mathcal{T}_N x^\epsilon = \mathcal{L}_{M,N}^\epsilon \mathcal{M}_M f. \tag{3.7}$$

Remark 3.4 To see why $G_{M,N}$ is generally ill-conditioned, consider the setting of Example 3.2 and suppose that Ψ is a tight frame, i.e. $A' = B'$. This assumption is made for convenience: the following argument readily extends to general frames. Let $\widetilde{S}_M : f \mapsto \sum_{m \in J_M} \langle f, \psi_m \rangle \psi_m$ be the partial frame operator for Ψ. Since Ψ is tight, $S_M \to A' \mathcal{I}$ strongly as $M \to \infty$, and therefore, for fixed N,

$$(G_{M,N})^* G_{M,N} \to A' G_N, \qquad M \to \infty,$$

where G_N is the Gram matrix of Φ_N. Hence, whenever G_N is ill-conditioned (i.e. Φ_N is near-redundant), we expect $G_{M,N}$ to inherit the same ill-conditioning for large M.

Remark 3.5 This argument gives some insight into the advantage of oversampling. For a tight frame, $(G_{M,N})^* G_{M,N}$ is an approximate factorization of G_N. Thus, solving the linear system (1.1) is akin to solving the normal equations of the least-squares problem $G_{M,N} x \approx y$. In this sense it is not surprising that oversampling yields $\mathcal{O}(\epsilon)$ accuracy, whereas solving (1.1) yields only $\mathcal{O}(\sqrt{\epsilon})$ accuracy. Indeed, this is reminiscent of the typical squaring of the condition number incurred when forming the normal equations of a least-squares problem [11, §5.3].

3.3 The Solution as an Orthogonal Projection

A key element of our subsequent analysis is the reinterpretation of the operator $\mathcal{P}_{M,N}^{\epsilon}$ as a projection with respect to a semi-definite sesquilinear form. Specifically, we now define the data-dependent sesquilinear form $\langle \cdot, \cdot \rangle_M$ on $G \times G$ as

$$\langle f, g \rangle_M = \langle \mathcal{M}_M f, \mathcal{M}_M g \rangle = \sum_{m \in J_M} \ell_{m,M}(f)\overline{\ell_{m,M}(g)}, \quad f, g \in G,$$

with corresponding discrete semi-norm $\|f\|_M = \sqrt{\langle f, f \rangle_M} = \|\mathcal{M}_M f\|$. Note that in general $\langle \cdot, \cdot \rangle_M$ is semi-definite on $G_N \times G_N$ as well, since

$$\langle g, g \rangle_M = \|g\|_M^2 = \sum_{m \in J_M} |\ell_{m,M}(g)|^2 \geq 0, \quad \forall g \in H_N, \quad g \neq 0.$$

In particular, for poorly chosen functionals it may be that $\|g\|_M = 0$ for some non-trivial function $g \in H_N$. However, with assumption (3.2) we do have the limiting behaviour

$$\liminf_{M \to \infty} \langle g, g \rangle_M = \liminf_{M \to \infty} \|g\|_M^2 \geq A' \|g\|^2 > 0, \quad \forall g \in H_N, g \neq 0.$$

This means that, ultimately, the sesquilinear form becomes an inner product on all of H_N.

Recall the singular value decomposition (3.4), and let $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_M, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_N$ and $\sigma_1, \ldots, \sigma_N$ be the left and right singular vectors and singular values of $\boldsymbol{G}_{M,N}$ respectively, with $\sigma_n \geq 0$, $n = 1, \ldots, N$. To the right singular vectors, we associate the functions

$$\xi_n = \mathcal{T}_N \boldsymbol{v}_n \in H_N, \quad n = 1, \ldots, N. \tag{3.8}$$

By construction, these functions are orthogonal with respect to $\langle \cdot, \cdot \rangle_M$. Indeed, by orthogonality of the singular vectors, we have

$$\langle \xi_m, \xi_n \rangle_M = \langle \mathcal{M}_M \mathcal{T}_N \boldsymbol{v}_n, \mathcal{M}_M \mathcal{T}_N \boldsymbol{v}_n \rangle = \sigma_n \sigma_m \langle \boldsymbol{u}_m, \boldsymbol{u}_n \rangle = \sigma_n \sigma_m \delta_{m,n}, \quad m, n \in I_N. \tag{3.9}$$

Here, too, it may be that $\|\xi_n\|_M = 0$. This is the case if $\sigma_n = 0$.

We shall, for convenience, let $\boldsymbol{x} = \boldsymbol{x}^0$ be the solution of the unregularized problem, given by

$$\boldsymbol{x} = (\boldsymbol{G}_{M,N})^{\dagger} \boldsymbol{y}.$$

We also write $\mathcal{P}_{M,N} = \mathcal{P}_{M,N}^0$ so that $\mathcal{P}_{M,N} f = \mathcal{T}_N \boldsymbol{x}$. Using the expression for the pseudoinverse in terms of the SVD, we can write both $\boldsymbol{x}$ and $\boldsymbol{x}^{\epsilon}$ in terms of the left

and right singular vectors:

$$x = \sum_{\sigma_n > 0} \frac{\langle y, u_n \rangle}{\sigma_n} v_n, \qquad x^\epsilon = \sum_{\sigma_n > \epsilon} \frac{\langle y, u_n \rangle}{\sigma_n} v_n. \tag{3.10}$$

Furthermore, for $\sigma_n > 0$ we also have

$$\langle y, u_n \rangle = \frac{\langle \mathcal{M}_M f, G_{M,N} v_n \rangle}{\sigma_n} = \frac{\langle \mathcal{M}_M f, \mathcal{M}_M \mathcal{T}_N v_n \rangle}{\sigma_n} = \frac{\langle f, \xi_n \rangle_M}{\sigma_n},$$

where in the last step we use (3.8). In particular, this gives

$$\mathcal{P}^\epsilon_{M,N} f = \mathcal{T}_N x^\epsilon = \sum_{\sigma_n > \epsilon} \frac{\langle y, u_n \rangle}{\sigma_n} \mathcal{T}_N v_n = \sum_{\sigma_n > \epsilon} \frac{\langle f, \xi_n \rangle_M}{\sigma_n^2} \xi_n. \tag{3.11}$$

Similarly, we have

$$\mathcal{P}_{M,N} f = \mathcal{T}_N x = \sum_{\sigma_n > 0} \frac{\langle f, \xi_n \rangle_M}{\sigma_n^2} \xi_n.$$

Finally, we define the regularized spaces

$$\mathrm{H}^\epsilon_{M,N} = \mathrm{span}\{\xi_n : \sigma_n > \epsilon\}.$$

Since $\{v_n\}_{n \in I_N}$ is a basis of $\mathbb{C}^N$, we see that the functions $\{\xi_n\}_{\sigma_n > 0}$ are an orthogonal basis of $\mathrm{H}^0_{M,N}$ with respect to $\langle \cdot, \cdot \rangle_M$. We can conclude that $\mathcal{P}_{M,N}$ is the orthogonal projection onto $\mathrm{H}^0_{M,N} \subseteq \mathrm{H}_N \subset G$. In turn, $\mathcal{P}^\epsilon_{M,N}$ is the orthogonal projection onto the subspace $\mathrm{H}^\epsilon_{M,N} \subseteq \mathrm{H}^0_{M,N}$.

A relevant property in the analysis that follows, is that these orthogonal projections imply a reduction in the M-norm:

Lemma 3.6 *For any $\epsilon \geq 0$, we have*

$$\|\mathcal{P}^\epsilon_{M,N} f\|_M \leq \|f\|_M, \qquad \forall f \in G. \tag{3.12}$$

3.4 Theoretical Results

We now define the constants

$$\kappa^\epsilon_{M,N} = \max_{\substack{y \in \mathrm{Ran}(\mathcal{M}_M) \\ \|y\|=1}} \|\mathcal{L}^\epsilon_{M,N} y\|, \quad \lambda^\epsilon_{M,N} = \epsilon^{-1} \max_{\substack{z \in \mathbb{C}^N \\ \|z\|=1}} \|\mathcal{T}_N z - \mathcal{P}^\epsilon_{M,N} \mathcal{T}_N z\|. \tag{3.13}$$

Note that $\kappa^\epsilon_{M,N}$ is precisely the operator norm of $\mathcal{L}^\epsilon_{M,N} : \mathrm{Ran}(\mathcal{M}_M) \to \mathrm{H}_N$. Since $\mathcal{L}^\epsilon_{M,N}$ is linear, it is also its absolute condition number, i.e. $\kappa^\epsilon_{M,N}$ measures the absolute

effect of perturbations in the data $\boldsymbol{y}$ on the final approximation. The constant $\lambda^\epsilon_{M,N}$ measures how close $\mathcal{P}^\epsilon_{M,N}$ is to being a projection on the subspace $\mathrm{H}_N = \mathcal{T}_N(\mathbb{C}^N)$.

Our first result concerns the approximation error of $\mathcal{P}^\epsilon_{M,N} f$:

Theorem 3.7 *Let $f \in$ G. The truncated SVD approximation $\mathcal{P}^\epsilon_{M,N} f$ satisfies*

$$\| f - \mathcal{P}^\epsilon_{M,N} f \| \leq \inf_{z \in \mathbb{C}^N} \left\{ \| f - \mathcal{T}_N z \| + \kappa^\epsilon_{M,N} \| f - \mathcal{T}_N z \|_M + \epsilon \lambda^\epsilon_{M,N} \| z \| \right\}. \tag{3.14}$$

This result differs from Theorem 1.1 in several respects. On the one hand, if the constants $\kappa^\epsilon_{M,N}$ and $\lambda^\epsilon_{M,N}$ are order one, the error depends on $\epsilon \| z \|$, not $\sqrt{\epsilon} \| z \|$, thus overcoming the $\sqrt{\epsilon}$ bottleneck. We will discuss when this occurs in the next subsection. On the other hand, the error bound involves the discrete data norm $\| f - \mathcal{T}_N z \|_M$. In general, this cannot be bounded by $\| f - \mathcal{T}_N z \|$. However, one clearly has

$$\| f - \mathcal{T}_N z \|_M \leq C_M \| | f - \mathcal{T}_N z | \|, \qquad C_M = \sqrt{\sum_{m \in J_M} (c_{m,M})^2},$$

where $c_{m,M}$ are the norms of the functionals $\ell_{m,M}$; recall (3.1). In particular, for Example 3.3 it follows that $\| f - \mathcal{T}_N z \|_M \leq \sqrt{2} \| f - \mathcal{T}_N z \|_{\mathrm{L}^\infty}$.

Proof of Theorem 3.7 For any $z \in \mathbb{C}^N$,

$$\| f - \mathcal{P}^\epsilon_{M,N} f \| \leq \| f - \mathcal{T}_N z \| + \| \mathcal{P}^\epsilon_{M,N} (f - \mathcal{T}_N z) \| + \| \mathcal{T}_N z - \mathcal{P}^\epsilon_{M,N} \mathcal{T}_N z \|.$$

Consider the second term. We have

$$\| \mathcal{P}^\epsilon_{M,N} (f - \mathcal{T}_N z) \| = \| \mathcal{L}^\epsilon_{M,N} \mathcal{M}_M (f - \mathcal{T}_N z) \| \leq \kappa^\epsilon_{M,N} \| \mathcal{M}_M (f - \mathcal{T}_N z) \|$$
$$= \kappa^\epsilon_{M,N} \| f - \mathcal{T}_N z \|_M,$$

which gives the corresponding term in (3.14). Now consider the third term. It follows immediately from the definition of $\lambda^\epsilon_{M,N}$ that $\| \mathcal{T}_N z - \mathcal{P}^\epsilon_{M,N} \mathcal{T}_N z \| \leq \epsilon \lambda^\epsilon_{M,N} \| z \|$, as required. $\qquad\square$

We now consider the coefficients $\boldsymbol{x}^\epsilon$:

Theorem 3.8 *Let $f \in$ G. The coefficients $\boldsymbol{x}^\epsilon$ of the truncated SVD projection $\mathcal{P}^\epsilon_{M,N}$ satisfy*

$$\| \boldsymbol{x}^\epsilon \| \leq \inf_{z \in \mathbb{C}^N} \left\{ \frac{1}{\epsilon} \| f - \mathcal{T}_N z \|_M + \| z \| \right\}. \tag{3.15}$$

Moreover, if $\Phi = \{\phi_n\}_{n \in I}$ is a frame, $\Phi_N = \{\phi_n\}_{n \in I_N}$ and if $\boldsymbol{a} = \{\langle f, \mathcal{S}^{-1} \phi_n \rangle\}_{n \in I}$ are the canonical frame coefficients of f and $\boldsymbol{a}^\epsilon_{M,N} \in \ell^2(I)$ is the extension of $\boldsymbol{x}^\epsilon$ by

zero, then

$$\|\boldsymbol{a} - \boldsymbol{a}^\epsilon_{M,N}\| \le \sqrt{\sum_{n \in I \setminus I_N} |a_n|^2} + \frac{1}{\epsilon} \|(\mathcal{S} - \mathcal{S}_N)\mathcal{S}^{-1} f\|_M + \epsilon \frac{\lambda^\epsilon_{M,N}}{\sqrt{A}} \|\boldsymbol{a}\|. \quad (3.16)$$

For general measurements, (3.16) does not imply convergence of the coefficients $\boldsymbol{a}^\epsilon_{M,N}$ to the canonical frame coefficients $\boldsymbol{a}$ since the term $\|(\mathcal{S} - \mathcal{S}_N)\mathcal{S}^{-1} f\|_M$ cannot typically be bounded by $\|(\mathcal{S} - \mathcal{S}_N)\mathcal{S}^{-1} f\|$. There is also no guarantee that $\||(\mathcal{S} - \mathcal{S}_N)\mathcal{S}^{-1} f\|| \to 0$ as $N \to \infty$. This does hold, however, when the data arises from sampling with another frame $\{\psi_n\}_{n \in I}$, as in Example 3.2, due to (3.3). We will discuss this case further in Sect. 4.

Proof of Theorem 3.8 For the first part, we use (3.10) to write

$$x^\epsilon = \sum_{\sigma_n > \epsilon} \frac{\langle f, \xi_n \rangle_M}{\sigma_n^2} v_n = \sum_{\sigma_n > \epsilon} \frac{\langle f - \mathcal{T}_N z, \xi_n \rangle_M}{\sigma_n^2} v_n + \sum_{\sigma_n > \epsilon} \frac{\langle \mathcal{T}_N z, \xi_n \rangle_M}{\sigma_n^2} v_n. \quad (3.17)$$

Consider the first term on the right-hand side. Since the v_n are orthonormal, we have

$$\left\| \sum_{\sigma_n > \epsilon} \frac{\langle f - \mathcal{T}_N z, \xi_m \rangle_M}{\sigma_n^2} v_n \right\|^2 = \sum_{\sigma_n > \epsilon} \frac{|\langle f - \mathcal{T}_N z, \xi_m \rangle_M|^2}{\sigma_m^4} \le \frac{1}{\epsilon^2} \sum_{\sigma_n > \epsilon} \frac{|\langle f - \mathcal{T}_N z, \xi_m \rangle_M|^2}{\sigma_m^2}.$$

It follows from (3.9) and (3.11) that

$$\sum_{\sigma_n > \epsilon} \frac{|\langle g, \xi_m \rangle_M|^2}{\sigma_m^2} = \left\| \mathcal{P}^\epsilon_{M,N} g \right\|_M^2, \quad g \in \mathrm{G}.$$

Hence

$$\left\| \sum_{\sigma_n > \epsilon} \frac{\langle f - \mathcal{T}_N z, \xi_m \rangle_M}{\sigma_n^2} v_n \right\|^2 \le \frac{1}{\epsilon^2} \left\| \mathcal{P}^\epsilon_{M,N} (f - \mathcal{T}_N z) \right\|_M \le \frac{1}{\epsilon^2} \| f - \mathcal{T}_N z \|_M^2,$$

where in the second step we use the fact that $\mathcal{P}^\epsilon_{M,N}$ is the orthogonal projection with respect to $\langle \cdot, \cdot \rangle_M$ (recall Lemma 3.6). This gives the first term of (3.15). Next, consider the second term of the right-hand side of (3.17). Since

$$\langle \mathcal{T}_N z, \xi_m \rangle_M = \langle \mathcal{T}_N z, \mathcal{T}_N v_n \rangle_M = \sigma_n^2 \langle z, v_n \rangle, \quad (3.18)$$

it follows that

$$\left\| \sum_{\sigma_n > \epsilon} \frac{\langle \mathcal{T}_N z, \xi_m \rangle_M}{\sigma_n^2} v_n \right\|^2 = \sum_{\sigma_n > \epsilon} |\langle z, v_n \rangle|^2 \le \|z\|,$$

This gives the second term of (3.15).

For (3.16), of course the canonical frame coefficients are well defined since we now assume that Φ is a frame. We first note that

$$\|\boldsymbol{a} - \boldsymbol{a}_{M,N}^\epsilon\| \leq \sqrt{\sum_{n \in I \setminus I_N} |a_n|^2} + \|\boldsymbol{a}_N - \boldsymbol{x}^\epsilon\|,$$

where $\boldsymbol{a}_N = \{a_n\}_{n \in I_N}$. Therefore it suffices to consider $\|\boldsymbol{a}_N - \boldsymbol{x}^\epsilon\|$. Observe that

$$\boldsymbol{a}_N = \sum_{n \in I_N} \langle \boldsymbol{a}_N, \boldsymbol{v}_n \rangle \boldsymbol{v}_n = \sum_{n \in I_N} \langle \mathcal{S}^{-1} f, \xi_n \rangle \boldsymbol{v}_n.$$

Now $\langle f, \xi_n \rangle_M = \langle \mathcal{S} \mathcal{S}^{-1} f, \xi_n \rangle_M$ and therefore

$$\begin{aligned} \langle f, \xi_n \rangle_M &= \langle \mathcal{S}_N \mathcal{S}^{-1} f, \xi_n \rangle_M + \langle (\mathcal{S} - \mathcal{S}_N) \mathcal{S}^{-1} f, \xi_n \rangle_M \\ &= \sigma_n^2 \langle \mathcal{T}_N^* \mathcal{S}^{-1} f, \boldsymbol{v}_n \rangle + \langle (\mathcal{S} - \mathcal{S}_N) \mathcal{S}^{-1} f, \xi_n \rangle_M. \end{aligned}$$

Notice that $\mathcal{S}_N \mathcal{S}^{-1} f \in G$ and $(\mathcal{S} - \mathcal{S}_N) \mathcal{S}^{-1} f = f - \mathcal{S}_N \mathcal{S}^{-1} f \in G$. Therefore all the terms above are well defined. Hence, by (3.17),

$$\boldsymbol{x}^\epsilon = \sum_{\sigma_n > \epsilon} \langle \mathcal{S}^{-1} f, \xi_n \rangle \boldsymbol{v}_n + \sum_{\sigma_n > \epsilon} \frac{\langle (\mathcal{S} - \mathcal{S}_M) \mathcal{S}^{-1} f, \xi_m \rangle_N}{\sigma_n^2} \boldsymbol{v}_n,$$

which gives

$$\|\boldsymbol{a}_N - \boldsymbol{x}^\epsilon\| \leq \left\| \sum_{\sigma_n \leq \epsilon} \langle \mathcal{S}^{-1} f, \xi_n \rangle \boldsymbol{v}_n \right\| + \left\| \sum_{\sigma_n > \epsilon} \frac{\langle (\mathcal{S} - \mathcal{S}_M) \mathcal{S}^{-1} f, \xi_n \rangle_M}{\sigma_n^2} \boldsymbol{v}_n \right\|. \tag{3.19}$$

Consider the first term. Let $z \in \mathbb{C}^N$ be given by $z = \sum_{\sigma_n \leq \epsilon} \langle \mathcal{S}^{-1} f, \xi_n \rangle \boldsymbol{v}_n$, so that the first term is merely $\|z\|$. By the definition of $\lambda_{M,N}^\epsilon$, we have $\epsilon \lambda_{M,N}^\epsilon \|z\| \geq \|\mathcal{T}_N z - \mathcal{P}_{M,N}^\epsilon \mathcal{T}_N z\|$. Now, since $\mathcal{T}_N z \perp H_{M,N}^\epsilon$, we have that $\mathcal{P}_{M,N}^\epsilon \mathcal{T}_N z = 0$. Hence

$$\epsilon \lambda_{M,N}^\epsilon \|z\| \geq \|\mathcal{T}_N z\| = \sup_{\substack{g \in H \\ g \neq 0}} \frac{|\langle \mathcal{T}_N z, g \rangle|}{\|g\|}.$$

Set $g = \mathcal{S}^{-1} f$. Then $\langle \mathcal{T}_N z, g \rangle = \sum_{\sigma_n \leq \epsilon} |\langle \mathcal{S}^{-1} f, \xi_m \rangle|^2 = \|z\|^2$ and therefore we obtain $\epsilon \lambda_{M,N}^\epsilon \|z\| \geq \|z\|^2 / \|\mathcal{S}^{-1} f\|$. It follows that

$$\|z\| = \left\| \sum_{\sigma_n \leq \epsilon} \langle \mathcal{S}^{-1} f, \xi_m \rangle \boldsymbol{v}_n \right\| \leq \epsilon \lambda_{M,N}^\epsilon \|\mathcal{S}^{-1} f\| \leq \epsilon \lambda_{M,N}^\epsilon / \sqrt{A} \|\boldsymbol{a}\|. \tag{3.20}$$

Now consider the second term of (3.19). We have

$$
\left\| \sum_{\sigma_n > \epsilon} \frac{\langle (\mathcal{S} - \mathcal{S}_M)\mathcal{S}^{-1}f, \xi_n \rangle_M}{\sigma_n^2} v_n \right\|^2 = \sum_{\sigma_n > \epsilon} \frac{|\langle (\mathcal{S} - \mathcal{S}_N)\mathcal{S}^{-1}f, \xi_n \rangle_M|^2}{\sigma_n^4}
$$

$$
\leq \frac{1}{\epsilon^2} \| \mathcal{P}_{M,N}^{\epsilon} (\mathcal{S} - \mathcal{S}_N)\mathcal{S}^{-1}f \|_M^2
$$

$$
\leq \frac{1}{\epsilon^2} \| (\mathcal{S} - \mathcal{S}_N)\mathcal{S}^{-1}f \|_M^2.
$$

In the last line, we used Lemma 3.6 again. Combining this with (3.20) now gives the result. ☐

Remark 3.9 These results extend to the setting of noisy measurements. Suppose the measurements are $y + n$ where $y = \mathcal{M}_M f$ and n is a vector of noise. We assume that $n \in \text{Ran}(\mathcal{M}_M)$, in other words it takes the form $n = \mathcal{M}_M g$ for some $g \in G$. Then, by linearity, the reconstruction is

$$
\tilde{f} = \mathcal{P}_{M,N}^{\epsilon} f + \mathcal{L}_{M,N}^{\epsilon} n,
$$

where $\mathcal{P}_{M,N}^{\epsilon} f$ is the standard reconstruction from the noiseless data y. Hence, by Theorem 3.7, the error satisfies

$$
\| f - \tilde{f} \| \leq \inf_{z \in \mathbb{C}^N} \left\{ \| f - \mathcal{T}_N z \| + \kappa_{M,N}^{\epsilon} \| f - \mathcal{T}_N z \|_M + \epsilon \lambda_{M,N}^{\epsilon} \| z \| \right\} + \kappa_{M,N}^{\epsilon} \| n \|.
$$

$$
(3.21)
$$

In particular, when $\kappa_{M,N}^{\epsilon}$ is order one, the effect of the noise is proportional to its ℓ^2-norm.

3.5 Behaviour of the Constants

We now consider the behaviour of the constants $\kappa_{M,N}^{\epsilon}$ and $\lambda_{M,N}^{\epsilon}$. To do so, we define the constant $A'_{M,N}$ as follows:

$$
A'_{M,N} = \inf_{\substack{g \in H_N \\ \|g\|=1}} \| g \|_M^2.
\tag{3.22}
$$

In general, with a poor choice of sampling functionals, it may be that $A'_{M,N} = 0$. However, even in that case, with assumption (3.2) we have that $\liminf_{M \to \infty} A'_{M,N} \geq A'$ for any fixed N.

Proposition 3.10 *The constants* $\kappa_{M,N}^{\epsilon}$ *and* $\lambda_{M,N}^{\epsilon}$ *defined in* (3.13) *satisfy*

$$
\kappa_{M,N}^{\epsilon} \leq \frac{\sqrt{B_N}}{\epsilon}, \qquad \lambda_{M,N}^{\epsilon} \leq \frac{\sqrt{B_N}}{\epsilon},
\tag{3.23}
$$

for all M and N, M ≥ N, where B_N is the Bessel bound for Φ_N. Moreover,

$$\kappa_{M,N}^\epsilon \leq \frac{1}{\sqrt{A'_{M,N}}}, \qquad \lambda_{M,N}^\epsilon \leq \frac{1}{\sqrt{A'_{M,N}}}, \qquad (3.24)$$

and if the sampling functionals satisfy (3.2) then, for fixed N,

$$\limsup_{M\to\infty} \kappa_{M,N}^\epsilon \leq \frac{1}{\sqrt{A'}} \quad and \quad \limsup_{M\to\infty} \lambda_{M,N}^\epsilon \leq \frac{1}{\sqrt{A'}}.$$

Proof Let $y \in \mathrm{Ran}(\mathcal{M}_M)$ be given and write $y = \mathcal{M}_M f$ for some $f \in \mathrm{G}$. Then, by (3.11),

$$\|\mathcal{L}_{M,N}^\epsilon y\| = \|\mathcal{P}_{M,N}^\epsilon f\| = \left\| \mathcal{T}_N \sum_{\sigma_n > \epsilon} \frac{\langle y, u_n \rangle}{\sigma_n} v_n \right\|.$$

Notice that $\|\mathcal{T}_N x\| \leq \sqrt{B_N} \|x\|$. This follows since any frame automatically satisfies the upper Riesz basis condition with constant equal to the Bessel bound. Hence

$$\|\mathcal{P}_{M,N}^\epsilon f\|^2 \leq B_N \sum_{\sigma_n > \epsilon} \frac{|\langle y, u_n \rangle|^2}{\sigma_n^2} \leq \frac{B_N}{\epsilon^2} \|y\|^2,$$

which gives (3.23) for $\kappa_{M,N}^\epsilon$. For (3.24), we let $y \in \mathrm{Ran}(\mathcal{M}_M)$ and write $y = \mathcal{M}_M f$ for some $f \in \mathrm{G}$ once more. From the definition (3.22) of $A'_{M,N}$ and by Lemma 3.6 we find

$$\|\mathcal{P}_{M,N}^\epsilon f\| \leq \frac{1}{\sqrt{A'_{M,N}}} \|\mathcal{P}_{M,N}^\epsilon f\|_M \leq \frac{1}{\sqrt{A'_{M,N}}} \|f\|_M = \frac{1}{\sqrt{A'_{M,N}}} \|y\|.$$

This gives (3.24).

We now consider $\lambda_{M,N}^\epsilon$. Let $z \in \mathbb{C}^M$ be arbitrary. Using (3.11) and (3.18) we have

$$\mathcal{T}_N z - \mathcal{P}_{M,N}^\epsilon \mathcal{T}_N z = \mathcal{T}_N \sum_{\sigma_n \leq \epsilon} \langle z, v_n \rangle v_n.$$

Arguing as above, this implies that $\|\mathcal{T}_N z - \mathcal{P}_{M,N}^\epsilon \mathcal{T}_N z\|^2 \leq B_N \|z\|^2$, which gives (3.23). For (3.24), we again let $z \in \mathbb{C}^M$ be arbitrary. Then

$$\left\| \mathcal{T}_N z - \mathcal{P}_{M,N}^\epsilon \mathcal{T}_N z \right\|_M^2 = \sum_{\sigma_n \leq \epsilon} \sigma_n^2 |\langle z, v_n \rangle|^2 \leq \epsilon^2 \|z\|^2.$$

Moreover, since $\mathcal{T}_N z - \mathcal{P}^\epsilon_{M,N} \mathcal{T}_N z \in H_N$ we obtain

$$\|\mathcal{T}_N z - \mathcal{P}^\epsilon_{M,N} \mathcal{T}_N z\|^2 \leq \frac{1}{A'_{M,N}} \|\mathcal{T}_N z - \mathcal{P}^\epsilon_{M,N} \mathcal{T}_N z\|^2_M \leq \frac{\epsilon^2}{A'_{M,N}} \|z\|^2,$$

as required.   □

3.6 The Stable Sampling Rate

Suppose that the sampling functionals satisfy (3.2). Motivated by Proposition 3.10 we now introduce the following concept:

Definition 3.11 For $1 < \theta < \infty$ and $N \in \mathbb{N}$, the *stable sampling rate* is

$$\Theta^\epsilon(N, \theta) = \min \left\{ M \in \mathbb{N} : M \geq N, \; \kappa^\epsilon_{M,N} \leq \frac{\theta}{\sqrt{A'}}, \; \lambda^\epsilon_{M,N} \leq \frac{\theta}{\sqrt{A'}} \right\}.$$

For a fixed N, suppose that M is chosen so that $M \geq \Theta^\epsilon(N, \theta)$. Then this guarantees an error bound of the form

$$\|f - \mathcal{P}^\epsilon_{M,N} f\| \leq \inf_{z \in \mathbb{C}^N} \left\{ \|f - \mathcal{T}_N z\| + \frac{\theta}{\sqrt{A'}} \|f - \mathcal{T}_N z\|_M + \epsilon \frac{\theta}{\sqrt{A'}} \|z\| \right\}.$$

Hence, sampling according to the stable sampling rate, ensures that the error decays down to roughly ϵ as $N \to \infty$. This holds on the additional condition that the term $\|f - \mathcal{T}_N z\|_M \to 0$; see the discussion after Theorem 3.7. Furthermore, sampling according to $\Theta^\epsilon(N, \theta)$ means that the *rate* of decay of the error for finite N depends completely on how well f can be approximated by elements of H_N with bounded coefficients. As discussed, this depends completely on the frame Φ and the element f being approximated. For estimates in certain cases, see [3].

Remark 3.12 If the data is noisy as in Remark 3.9 and $M \geq \Theta^\epsilon(N, \theta)$ then (3.21) becomes

$$\|f - \tilde{f}\| \leq \inf_{z \in \mathbb{C}^N} \left\{ \|f - \mathcal{T}_N z\| + \frac{\theta}{\sqrt{A'}} \|f - \mathcal{T}_N z\|_M + \epsilon \frac{\theta}{\sqrt{A'}} \|z\| \right\} + \frac{\theta}{\sqrt{A'}} \|n\|.$$

Note that ϵ does not enter into the noise term. Recall that the first term will decrease down to a limiting accuracy of at best $\mathcal{O}(\epsilon)$. Hence, in the noisy case the limiting accuracy will depend on the maximum of ϵ and $\|n\|$. In particular, this yields a simple strategy for choosing ϵ in the noisy case, simply as proportional to the noise level.

The behaviour of $\Theta^\epsilon(N, \theta)$ as a function of N depends completely on Φ and the sampling functionals. Thus, theoretical estimates for this quantity can only be established on a case-by-case basis. We shall consider this issue further in Sect. 5 for a particular class of problems. However, there is no general recipe for providing such

estimates, and moreover, when possible, doing so typically only reveals the asymptotic growth rate of $\Theta^\epsilon(N, \theta)$ with N and not the precise constant.

On the other hand, $\Theta^\epsilon(N, \theta)$ can always be computed. To see this, we observe the following:

Lemma 3.13 *The constant $\kappa^\epsilon_{M,N}$ satisfies*

$$\kappa^\epsilon_{M,N} \leq \sqrt{\lambda_{\max}\left((\boldsymbol{B}^\epsilon_{M,N})^* \boldsymbol{G}_N \boldsymbol{B}^\epsilon_{M,N}\right)},$$

where $\boldsymbol{B}^\epsilon_{M,N} = (\boldsymbol{G}^\epsilon_{M,N})^\dagger$ and $\boldsymbol{G}_N \in \mathbb{C}^{N \times N}$ is the Gram matrix of Φ_N. If $\mathrm{Ran}(\mathcal{M}_M) = \mathbb{C}^M$ this holds with equality. The constant $\lambda^\epsilon_{M,N}$ satisfies

$$\lambda^\epsilon_{M,N} = \epsilon^{-1}\sqrt{\lambda_{\max}\left((\boldsymbol{C}^\epsilon_{M,N})^* \boldsymbol{G}_N \boldsymbol{C}^\epsilon_{M,N}\right)},$$

where $\boldsymbol{C}^\epsilon_{M,N} = \boldsymbol{V}(\boldsymbol{I} - \boldsymbol{I}^\epsilon)\boldsymbol{V}^*$ and $\boldsymbol{I}_\epsilon$ is the diagonal matrix with nth entry 1 if $\sigma_n \geq \epsilon$ and zero otherwise.

Proof Let $\boldsymbol{y} \in \mathrm{Ran}(\mathcal{M}_M)$ with $\|\boldsymbol{y}\| = 1$. Then

$$\left\|\mathcal{L}^\epsilon_{M,N}\boldsymbol{y}\right\|^2 = \left\|\mathcal{T}_N \boldsymbol{B}^\epsilon_{M,N}\boldsymbol{y}\right\|^2 = \boldsymbol{y}^*(\boldsymbol{B}^\epsilon_{M,N})^* \boldsymbol{G}_N \boldsymbol{B}^\epsilon_{M,N}\boldsymbol{y} \leq \lambda_{\max}\left((\boldsymbol{B}^\epsilon_{M,N})^* \boldsymbol{G}_N \boldsymbol{B}^\epsilon_{M,N}\right),$$

since $\boldsymbol{G}_N = \mathcal{T}_N^* \mathcal{T}_N$. This gives the first result.

Let $\boldsymbol{z} \in \mathbb{C}^N$, $\|\boldsymbol{z}\| = 1$. Then

$$\left\|\mathcal{T}_N \boldsymbol{z} - \mathcal{P}^\epsilon_{M,N}\mathcal{T}_N \boldsymbol{z}\right\|^2 = \left\|\mathcal{T}_N\left(\boldsymbol{I} - (\boldsymbol{G}^\epsilon_{M,N})^\dagger \boldsymbol{G}_{M,N}\right)\boldsymbol{z}\right\|^2 = \left\|\mathcal{T}_N\left(\boldsymbol{I} - \boldsymbol{V}\boldsymbol{I}^\epsilon \boldsymbol{V}^*\right)\boldsymbol{z}\right\|^2$$

$$= \left\|\mathcal{T}_N \boldsymbol{C}^\epsilon_{M,N}\boldsymbol{z}\right\|^2 = \boldsymbol{z}^*(\boldsymbol{C}^\epsilon_{M,N})^* \boldsymbol{G}_N \boldsymbol{C}^\epsilon_{M,N}\boldsymbol{z}.$$

Maximizing over $\boldsymbol{z}$ now gives the result. $\square$

This result implies that $\kappa^\epsilon_{M,N}$ and $\lambda^\epsilon_{M,N}$ can be computed, and therefore so can $\Theta^\epsilon(N; \theta)$, provided the matrix $\boldsymbol{G}_N$ has been computed.

Remark 3.14 In practice, it may be difficult to compute $\boldsymbol{G}_N$, since its entries are inner products which may for instance be integrals. This may be overcome by a further approximation, e.g. a quadrature. Specifically, let $K \geq 1$ and $J_{k,K}$ be a family of linear functionals such that

$$\lim_{K \to \infty} \sum_{k=1}^{K} J_{k,K}(f)\overline{J_{k,K}(g)} = \langle f, g \rangle, \qquad \forall f, g \in \mathrm{G}.$$

Let $\boldsymbol{H}_{K,N} = \{J_{k,K}(\phi_n)\}^K_{k=1, n \in I_N} \in \mathbb{C}^{K \times N}$. Then $(\boldsymbol{H}_{K,N})^* \boldsymbol{H}_{K,N} \approx \boldsymbol{G}_N$ for large K. Hence, by the previous lemma (assuming $\mathrm{Ran}(\mathcal{M}_M) = \mathbb{C}^M$ for ease of presentation),

we have

$$\kappa_{M,N}^{\epsilon} \approx \left\| H_{K,N} B_{M,N}^{\epsilon} \right\|_2 = \left\| H_{K,N} V (\Sigma^{\epsilon})^{\dagger} \right\|_2, \qquad \lambda_{M,N}^{\epsilon} \approx \left\| H_{K,N} V (I - I^{\epsilon}) \right\|_2,$$

for sufficiently large K. If, for instance, the functionals $J_{k,K}$ correspond to pointwise evaluations as part of a quadrature, this gives a means of numerically approximating $\kappa_{M,N}^{\epsilon}$ and $\lambda_{M,N}^{\epsilon}$.

4 Frame Approximation from Frame Samples

In this section, we discuss Example 3.2, in which both the approximation system and sampling functionals arise from frames of H. We write $\Phi = \{\phi_n\}_{n \in I}$ for the approximation frame (with bounds A and B) and $\Psi = \{\psi_m\}_{m \in J}$ for the sampling frame (with bounds A' and B'). We assume that $\{I_N\}_{N \in \mathbb{N}}$ is a sequence of nested index sets with

$$I_N \subset I, \ |I_N| = N, \ \forall N \in \mathbb{N}, \quad \bigcup_{N=1}^{\infty} I_N = I,$$

and similarly, we assume that $\{J_M\}_{M \in \mathbb{N}}$ is a sequence of nested index sets with

$$J_M \subset I, \ |J_M| = M, \ \forall M \in \mathbb{N}, \quad \bigcup_{M=1}^{\infty} I_M = J.$$

As before, we write $\mathcal{P}_{M,N}^{\epsilon} f$ for the truncated SVD approximation of $f \in$ H (note that G = H in this case, since the sampling functionals arise from a frame of H). We write x^{ϵ} for its coefficients.

4.1 Error and Coefficient Bounds

Theorem 4.1 *In the above setting, the truncated SVD projection $\mathcal{P}_{M,N}^{\epsilon} f$ of $f \in$ H satisfies*

$$\| f - \mathcal{P}_{M,N}^{\epsilon} f \| \leq \inf_{z \in \mathbb{C}^N} \left\{ \left(1 + \sqrt{B'} \kappa_{M,N}^{\epsilon} \right) \| f - \mathcal{T}_N z \| + \epsilon \lambda_{M,N}^{\epsilon} \| z \| \right\}. \quad (4.1)$$

Its coefficients x^{ϵ} satisfy

$$\| x^{\epsilon} \| \leq \inf_{z \in \mathbb{C}^N} \left\{ \frac{\sqrt{B'}}{\epsilon} \| f - \mathcal{T}_N z \| + \| z \| \right\}, \quad (4.2)$$

and, if $a^\epsilon_{M,N} \in \ell^2(I)$ the extension of x^ϵ by zero,

$$\|a - a^\epsilon_{M,N}\| \leq \left(1 + \frac{\sqrt{BB'}}{\epsilon}\right) \sqrt{\sum_{n \in I \setminus I_N} |a_n|^2} + \epsilon \frac{\lambda^\epsilon_{M,N}}{\sqrt{A}} \|a\|, \tag{4.3}$$

where $a = \{\langle f, \mathcal{S}^{-1}\phi_n\rangle\}_{n \in I}$ are the canonical frame coefficients of $f \in$ H.

Proof Recall (3.3). The first two bounds now follow immediately from Theorems 3.7 and 3.8 respectively. For the third bound, we use use (3.16) and then observe that

$$\left\|(\mathcal{S} - \mathcal{S}_N)\mathcal{S}^{-1}f\right\|^2 = \left\|\sum_{n \in I \setminus I_N} a_n \phi_n\right\|^2 \leq B \sum_{n \in I \setminus I_N} |a_n|^2,$$

where for the final step we recall that a frame with upper frame bound B satisfies the upper Riesz basis condition with constant B. The result now follows immediately.
$\square$

4.2 Behaviour of the Coefficients and $\mathcal{O}(\epsilon)$ Accuracy

We now consider the constants $\kappa^\epsilon_{M,N}$ and $\lambda^\epsilon_{M,N}$ and the stable sampling rate:

Proposition 4.2 *The constants $\kappa^\epsilon_{M,N}$ and $\lambda^\epsilon_{M,N}$ satisfy*

$$\kappa^\epsilon_{M,N}, \lambda^\epsilon_{M,N} \leq \begin{cases} \sqrt{B}/\epsilon & \Psi \neq \Phi \\ 1/\sqrt{\epsilon} & \Psi = \Phi, \end{cases} \tag{4.4}$$

for all M and N, $M \geq N$. Moreover, for fixed N,

$$\limsup_{M \to \infty} \kappa^\epsilon_{M,N} \leq \frac{1}{\sqrt{A'}} \quad and \quad \limsup_{M \to \infty} \lambda^\epsilon_{M,N} \leq \frac{1}{\sqrt{A'}},$$

This result is essentially a special case of Proposition 3.10, except in the case where $\Psi = \Phi$ where we have a slightly improved worst-case behaviour, with the right-hand side of (4.4) scaling like $1/\sqrt{\epsilon}$ as opposed to $1/\epsilon$. As is made clear by the proofs, this discrepancy is due to the fact that in the latter case the measurements are just inner products with respect to the same frame. We prove this in a moment. First, however, we consider its implications for limiting accuracy:

Corollary 4.3 *For each $1 < \theta < \infty$ the truncated SVD approximation $\mathcal{P}^\epsilon_{M,N}f$ satisfies*

$$\limsup_{\substack{M,N \to \infty \\ M \geq \Theta^\epsilon(N,\theta)}} \|f - \mathcal{P}^\epsilon_{M,N}f\| \leq \epsilon \frac{\theta}{\sqrt{AA'}} \|f\|,$$

where $\Theta^\epsilon(N,\theta)$ *is as in Definition* 3.11. *Moreover, the coefficients* $\boldsymbol{x}^\epsilon$ *satisfy*

$$\limsup_{\substack{M,N\to\infty \\ M\geq\Theta^\epsilon(N,\theta)}} \|\boldsymbol{x}^\epsilon\| \leq \frac{1}{\sqrt{A}}\|f\|, \qquad \limsup_{\substack{M,N\to\infty \\ M\geq\Theta^\epsilon(N,\theta)}} \|\boldsymbol{a}-\boldsymbol{a}^\epsilon_{M,N}\| \leq \epsilon\frac{\theta}{\sqrt{AA'}}\|\boldsymbol{a}\|.$$

Proof The proof is based on the canonical frame coefficients $\boldsymbol{a} = \{\langle f, \mathcal{S}^{-1}\phi_n\rangle\}_{n\in I}$. Let $\boldsymbol{z} = \{a_n\}_{n\in I_N}$. Then $\|\boldsymbol{z}\| \leq \|\boldsymbol{a}\| \leq 1/\sqrt{A}\|f\|$ since the dual frame has upper frame bound A^{-1} (see Sect. 2.3). Therefore (4.1) and Proposition 4.2 gives

$$\|f - \mathcal{P}^\epsilon_{M,N}f\| \leq \left(1 + \sqrt{\frac{B}{A}}\theta\right)\left\|f - \sum_{n\in I_N}\langle f, \mathcal{S}^{-1}\phi_n\rangle\phi_n\right\| + \epsilon\frac{\theta}{A}\|f\|.$$

As $N \to \infty$ (2.5) gives that the first term vanishes. Hence we obtain the result for f. For the other results, we use (4.2) and (4.3) instead. □

In summary, provided M is chosen above the stable sampling rate $\Theta^\epsilon(N,\theta)$, the approximation $\mathcal{P}^\epsilon_{M,N}f$ converges to within roughly ϵ of f and the coefficients converge to within roughly ϵ of the frame coefficients $\boldsymbol{a}$, and in particular are small in norm for large N. As a consequence, in the setting of Theorem 1.1 where $\Psi = \Phi$, we overcome the $\sqrt{\epsilon}$ bottleneck, with the limiting accuracy bounded by $\epsilon\theta\|f\|/A$ as opposed to $\sqrt{\epsilon}\|f\|/\sqrt{A}$ (see Corollary 1.2).

This result also illuminates the role that the frame structure plays in both the approximation and the sampling. Indeed, the limiting error depends on both of the lower frame bounds A and A', while the limiting size of the coefficients depends only on A. This is as expected. The existence of small norm coefficients depends only on the approximation frame Φ, a small limiting error depends on both the sampling frame Ψ and the approximation frame Φ.

Proof of Proposition 4.2 All results follow immediately from Proposition 3.10, except for (4.4) in the case $\Psi = \Phi$ for which we require a different argument.

Consider $\kappa^\epsilon_{M,N}$ first. Let $\boldsymbol{y} \in \mathrm{Ran}(\mathcal{M}_M)$ be given and notice that we may write $\boldsymbol{y} = \mathcal{T}^*_M f$ for some $f \in \mathrm{H}$ so that $\|\mathcal{L}^\epsilon_{M,N}\boldsymbol{y}\| = \|\mathcal{P}^\epsilon_{M,N}f\|$. By (3.10) we have

$$\|\mathcal{P}^\epsilon_{M,N}f\|^2 = \sum_{\sigma_m,\sigma_n>\epsilon} \frac{\langle \boldsymbol{y}, \boldsymbol{u}_m\rangle\overline{\langle \boldsymbol{y}, \boldsymbol{u}_n\rangle}}{\sigma_m\sigma_n}\langle\xi_m,\xi_n\rangle.$$

Recall that $\mathcal{M}_M\mathcal{T}_N\boldsymbol{v}_m = \sigma_m\boldsymbol{u}_m$. Since $\mathcal{M}_M = \mathcal{T}^*_M$ in this case, we have $\mathcal{T}^*_N\mathcal{T}_N\boldsymbol{v}_m = \sigma_m\tilde{\boldsymbol{u}}_m$, where where $\tilde{\boldsymbol{u}}_m \in \mathbb{C}^N$ is the vector with entries $(\tilde{\boldsymbol{u}}_m)_k = (\boldsymbol{u}_m)_k$ for $k \in I_N$. Hence $\langle\xi_m,\xi_n\rangle = \langle\mathcal{T}^*_N\mathcal{T}_N\boldsymbol{v}_m, \boldsymbol{v}_n\rangle = \sigma_m\langle\tilde{\boldsymbol{u}}_m, \boldsymbol{v}_n\rangle$ and this gives

$$\|\mathcal{P}^\epsilon_{M,N}f\|^2 = \left\langle \sum_{\sigma_n>\epsilon}\langle \boldsymbol{y}, \boldsymbol{u}_m\rangle\tilde{\boldsymbol{u}}_m, \sum_{\sigma_n>\epsilon}\frac{\langle \boldsymbol{y}, \boldsymbol{u}_n\rangle}{\sigma_n}\boldsymbol{v}_n \right\rangle \leq \left\| \sum_{\sigma_n>\epsilon}\langle \boldsymbol{y}, \boldsymbol{u}_m\rangle\tilde{\boldsymbol{u}}_m \right\|\left\| \sum_{\sigma_n>\epsilon}\frac{\langle \boldsymbol{y}, \boldsymbol{u}_n\rangle}{\sigma_n}\boldsymbol{v}_n \right\|.$$

By orthogonality, the second term satisfies

$$\left\| \sum_{\sigma_n > \epsilon} \frac{\langle y, u_n \rangle}{\sigma_n} v_n \right\| = \sqrt{ \sum_{\sigma_n > \epsilon} \frac{|\langle y, u_n \rangle|^2}{\sigma_n^2} } \leq \frac{\|y\|}{\epsilon}.$$

Consider the first term. Let $\mathcal{Q}_N : \mathbb{C}^M \to \mathbb{C}^M$ be the projection defined by $(\mathcal{Q}_N x)_m = x_m$, $m \in I_N$ and $(\mathcal{Q}_N x)_m = 0$, $m \in I_M \backslash I_N$. Then

$$\left\| \sum_{\sigma_n > \epsilon} \langle y, u_m \rangle \tilde{u}_m \right\| = \left\| \mathcal{Q}_N \left(\sum_{\sigma_n > \epsilon} \langle y, u_m \rangle u_m \right) \right\| \leq \left\| \sum_{\sigma_n > \epsilon} \langle y, u_m \rangle u_m \right\| \leq \|y\|.$$

Therefore, we deduce that

$$\| \mathcal{P}^\epsilon_{M,N} f \|^2 \leq \|y\| / \epsilon,$$

and the result for $\kappa^\epsilon_{M,N}$ now follows from its definition.

Now consider $\lambda^\epsilon_{M,N}$. Let $z \in \mathbb{C}^N$ be arbitrary and recall that

$$\mathcal{T}_N z - \mathcal{P}^\epsilon_{M,N} \mathcal{T}_N z = \sum_{\sigma_n \leq \epsilon} \langle z, v_n \rangle \xi_n.$$

Hence

$$\left\| \mathcal{T}_N z - \mathcal{P}^\epsilon_{M,N} \mathcal{T}_N z \right\|^2 = \sum_{\sigma_m, \sigma_n \leq \epsilon} \langle z, v_n \rangle \overline{\langle z, v_n \rangle} \langle \xi_m, \xi_n \rangle.$$

As above, we note that $\langle \xi_m, \xi_n \rangle = \sigma_m \langle \tilde{u}_m, v_n \rangle$, and therefore

$$\left\| \mathcal{T}_N z - \mathcal{P}^\epsilon_{M,N} \mathcal{T}_N z \right\|^2 = \left\langle \sum_{\sigma_n \leq \epsilon} \sigma_m \langle z, v_m \rangle \tilde{u}_m, \sum_{\sigma_n \leq \epsilon} \langle z, v_n \rangle v_n \right\rangle \leq \epsilon \|z\|^2.$$

Since z was arbitrary, we now obtain the result for $\lambda^\epsilon_{M,N}$. $\square$

5 ONB+1 and ONB+K Frames

We conclude this paper with several examples to illustrate the stable sampling rate.

First, let $\{\varphi_n\}_{n \in \mathbb{N}}$ be an orthonormal basis of H and $\psi \in$ H, $\|\psi\| = 1$, be such that $\langle \psi, \varphi_n \rangle \neq 0$ for infinitely-many $n \in \mathbb{N}$. Then the indexed family

$$\Phi = \{\phi_0, \phi_1, \ldots\} = \{\psi, \varphi_1, \varphi_2, \ldots\},$$

is a frame for Φ with frame bounds $A = 1$ and $B = 2$. We refer to this frame as the *ONB+1 frame*. Note that it was previously used in [3] to show that the Gram matrix of a

frame can be arbitrarily badly conditioned. It is motivated by the idea of 'enriching' an orthonormal basis to better capture a certain feature of a function under approximation (e.g. a singularity or oscillation).

Throughout this section, we let $\mathcal{Q}_N$ denote the projection onto $\mathrm{span}\{\varphi_1, \ldots, \varphi_N\}$, i.e.

$$\mathcal{Q}_N f = \sum_{n=1}^{N} \langle f, \varphi_n \rangle \varphi_n.$$

5.1 The Stable Sampling Rate for the ONB+1 Frame

A problem of interest is that where the samples are inner products with respect to the orthonormal basis $\{\varphi_m\}_{m\in\mathbb{N}}$. That is,

$$\ell_{m,M}(f) = \ell_m(f) = \langle f, \varphi_m \rangle, \quad m = 1, \ldots, M, \ M \in \mathbb{N}. \tag{5.1}$$

For instance, these are Fourier coefficients if $\{\varphi_m\}_{m\in\mathbb{N}}$ is the Fourier basis, and hence the goal would be to compute a better approximation in the frame Φ from the given Fourier data. Note that this is an instance of Sect. 4 with $A' = B' = 1$. Note also that $\|g\|_M = \|\mathcal{Q}_M g\|$. Recalling Proposition 3.10, to determine the stable sampling rate we note that it suffices to estimate

$$A'_{M,N} = \inf_{\substack{g\in H_N \\ \|g\|=1}} \|\mathcal{Q}_M g\|^2 = 1 - \sup_{\substack{g\in H_N \\ \|g\|=1}} \|g - \mathcal{Q}_M g\|^2, \tag{5.2}$$

where $H_N = \mathrm{span}\{\psi, \varphi_1, \ldots, \varphi_{N-1}\}$.

Lemma 5.1 *For $M \geq N$, we have*

$$\sup_{\substack{g\in H_N \\ \|g\|=1}} \|g - \mathcal{Q}_M g\| = \frac{\|\psi - \mathcal{Q}_M \psi\|}{\|\psi - \mathcal{Q}_{N-1}\psi\|}.$$

Proof Let $g \in H_N$ and write $g = x_0\psi + \sum_{n=1}^{N-1} x_n\varphi_n$. Then

$$\langle g, \varphi_n \rangle = x_0\langle \psi, \varphi_n \rangle + x_n, \quad n = 1, \ldots, N-1,$$

and hence

$$g = x_0\psi + \sum_{n=1}^{N-1} \left(\langle g, \varphi_n \rangle - x_0\langle \psi, \varphi_n \rangle \right) \varphi_n = x_0(\psi - \mathcal{Q}_{N-1}\psi) + \mathcal{Q}_{N-1}g.$$

Rearranging gives

$$g - \mathcal{Q}_{N-1}g = x_0(\psi - \mathcal{Q}_{N-1}\psi),$$

and taking the norm of both sides, we find that

$$|x_0| = \frac{\|g - \mathcal{Q}_{N-1}g\|}{\|\psi - \mathcal{Q}_{N-1}\psi\|} \leq \frac{\|g\|}{\|\psi - \mathcal{Q}_{N-1}\psi\|} = \frac{1}{\|\psi - \mathcal{Q}_{N-1}\psi\|}.$$

Also, we have

$$\langle g, \varphi_n \rangle = x_0 \langle \psi, \varphi_n \rangle, \quad n \geq N.$$

Therefore

$$\|g - \mathcal{Q}_M g\| = |x_0| \|\psi - \mathcal{Q}_M \psi\|.$$

Combining this with the bound for $|x_0|$ gives that

$$\sup_{\substack{g \in \mathsf{H}_N \\ \|g\|=1}} \|g - \mathcal{Q}_M g\| \leq \frac{\|\psi - \mathcal{Q}_M \psi\|}{\|\psi - \mathcal{Q}_{N-1}\psi\|}$$

To show equality, we divide into two cases. Suppose first that $\psi \perp \mathrm{span}\{\varphi_1, \ldots, \varphi_{N-1}\}$. Then $\mathcal{Q}_{N-1}\psi = 0$ and so we may take $g = \psi / \|\psi\|$ to obtain equality. On the other hand, if $\psi \not\perp \mathrm{span}\{\varphi_1, \ldots, \varphi_{N-1}\}$ then there exists a $g \in \mathsf{H}_N$, $\|g\| = 1$, with $\mathcal{Q}_{N-1}g = 0$. In this case, the above arguments give that $\|g - \mathcal{Q}_M g\| = \|\psi - \mathcal{Q}_M \psi\| / \|\psi - \mathcal{Q}_{N-1}\psi\|$, which implies the result. $\qquad\square$

This leads us to the following result:

Theorem 5.2 *Suppose that ψ is such that $|\langle \psi, \varphi_n \rangle| \sim cn^{-\alpha}$ as $n \to \infty$ for some $c > 0$ and $\alpha > 1/2$. Then the stable sampling rate*

$$\Theta^\epsilon(N, \theta) \leq CN,$$

for some constant $C > 0$ depending on c, α and θ only. Conversely, if $|\langle \psi, \varphi_n \rangle| \sim c\rho^{-n}$ as $n \to \infty$ for some $c > 0$ and $\rho > 1$ then

$$\Theta^\epsilon(N, \theta) \leq N + C,$$

where $C > 0$ depends on c, ρ and θ only.

Proof In the first case, the condition on the coefficients gives

$$\|\psi - \mathcal{Q}_N \psi\|^2 = \sum_{n \geq N} |\langle \psi, \varphi_n \rangle|^2 \sim c' N^{1-2\alpha}, \quad N \to \infty,$$

where c' depends on c and α. Hence Lemma 5.1 and the bound (5.2) give

$$A'_{M,N} \geq 1 - c'' \left(\frac{M^{1/2-\alpha}}{N^{1/2-\alpha}} \right)^2,$$

for some constant c'' depending on c and α. Recalling that $B' = A' = A = 1$ and $B = 2$ for this frame and using Proposition 3.10 gives the first result. For the second result, we notice that

$$\|\psi - \mathcal{Q}_N \psi\|^2 \sim c \frac{\rho^{-N}}{1 - \rho}.$$

We now argue as in the previous case.   □

This result shows that the stable sampling rate is linear when ψ has algebraically or exponentially-decaying coefficients in the orthonormal basis $\{\varphi_n\}_{n \in \mathbb{N}}$. Furthermore, the better ψ is approximated in this basis, the smaller the stable sampling rate is, as evidenced by the case of exponentially-decaying coefficients. In fact, Lemma 5.1 demonstrates the connection between the stable sampling rate and how well approximated ψ is in the orthonormal basis $\{\varphi_n\}_{n \in \mathbb{N}}$. Specifically, the faster the projection errors $\|\psi - \mathcal{Q}_M \psi\|$ decay, the smaller $M \geq N$ needs to be so that $\frac{\|\psi - \mathcal{Q}_M \psi\|}{\|\psi - \mathcal{Q}_{N-1}\psi\|} \leq \delta$ for constant $0 < \delta < 1$. This is intuitive. The better ψ is approximated by this basis, the more information the data, i.e. inner products with the φ_n, carries about the element g.

On the other hand, the worse g is approximated the higher the stable sampling rate. Indeed, if $\|\psi - \mathcal{Q}_M \psi\| \asymp (\log(N))^{-1}$ then it is a simple exercise to show that the stable sampling rate is algebraic in N with the power depending on θ, i.e. $\Theta^\epsilon(N, \theta) = \mathcal{O}\left(N^{h(\theta)}\right)$ for some function $h(\theta) \geq 1$ with $h(\theta) \to \infty$ as $\theta \to 1^+$.

Remark 5.3 One can also determine a bound on the stable sampling rate for special case $\Phi = \Psi$, also discussed in Sect. 4. In this case, the data is given by the inner products

$$\langle f, \psi \rangle, \quad \langle f, \varphi_m \rangle, \quad m = 1, \ldots, M - 1.$$

Indeed, observe that

$$A'_{M,N} = \inf_{\substack{g \in H_N \\ \|g\|=1}} \left(|\langle g, \psi \rangle|^2 + \|\mathcal{Q}_{M-1} g\|^2 \right) \geq \inf_{\substack{g \in H_N \\ \|g\|=1}} \|\mathcal{Q}_{M-1} g\|^2.$$

The right-hand side is precisely the constant in (5.2) with M replaced by $M-1$. Hence, up to an additive factor of one, the stable sampling rate for this problem satisfies the same bounds as those of Theorem 5.2.

5.2 The Approximation of Functions with Logarithmic Singularities

Let $H = L^2(0, 1)$. The scaled Legendre polynomials, $\varphi_n(x) = \sqrt{2n-1} P_{n-1}(2x-1)$, $n \in \mathbb{N}$, form an orthonormal basis for H. Here $P_n(x)$ is the usual Legendre polynomial, with normalization $P_n(1) = 1$. This basis is extremely good at approximating smooth functions. However, many functions that may arise in applications, such as Green's functions or solutions to PDEs on domains with corners, fail to be smooth at a point

x, yet posses a known type of singularity there. That is, in these applications we may want to approximate functions of the form

$$f(x) = g(x) + w(x)h(x), \quad x \in (0, 1) \tag{5.3}$$

where g, h are smooth functions, and $w \in L^2(0, 1)$ is a known function which may be singular at, say, $x = 0$. Such functions cannot generally be accurately approximated using polynomials alone. However, they can be more accurately captured by enriching the polynomial basis with the function w. This gives a frame

$$\Phi = \{\varphi_n\}_{n=1}^\infty \cup \{w\}, \tag{5.4}$$

for H. Indeed, since the φ_n are an orthonormal basis, it quickly follows that

$$\|f\|_2^2 = \sum_{n=1}^\infty |\langle f, \varphi_n \rangle|^2 \leq \sum_{n=1}^\infty |\langle f, \varphi_n \rangle|^2 + |\langle f, w \rangle|^2 \leq \|f\|^2 + \|f\|^2 \|w\|^2.$$

Hence this is a frame with bounds $A \geq 1$ and $B \leq 1 + \|w\|^2$.

The case of a logarithmic singularity, i.e. $w(x) = \log(x)$, is an important instance of the problem. Figure 1 gives an illustration of the benefits of this frame over just the polynomial basis for approximating the simple yet singular function

$$f(x) = e^x + \log(x)\cos(x). \tag{5.5}$$

The polynomial interpolation to f converges poorly, as expected. However, adding just the single element $w(x) = \log(x)$ to the basis results in significantly faster convergence rates, shown in Fig. 1b. Importantly, note that the approximation scheme does not evaluate the smooth parts of f separately. They are implicitly approximated
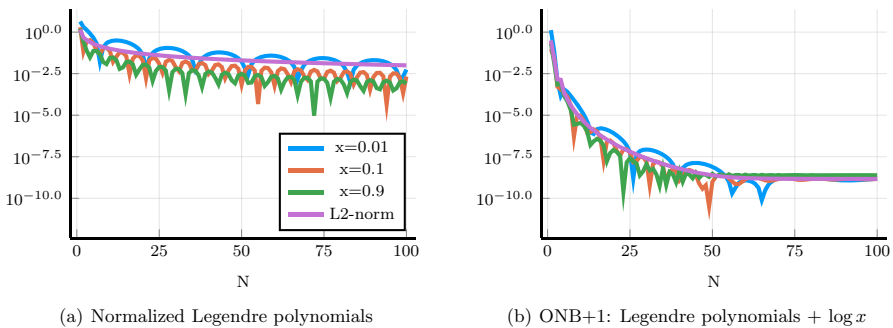

Fig. 1 Pointwise error as a function of the polynomial degree N for the approximation of the logarithmically singular function (5.5) on [0, 1] using Legendre polynomials (left panel) and Legendre polynomials augmented with $\log x$ (right panel). The error is shown in four points in the interval [0, 1]. In both cases, the approximation problem is solved using generalized sampling (5.1) with $M = 2N$. The generalized samples $(\langle f, \varphi_m \rangle)_{m=1}^M$ were evaluated using adaptive numerical integration. The regularization threshold is $\epsilon = 2e^{-13}$

simultaneously when approximating f from its samples. Indeed, if the smooth parts of f were known separately in an application, the approximation problem simplifies and there would be no need to construct a frame. Note also that the evaluation of the generalized samples (5.1) requires the evaluation of integrals, and this step is computationally demanding because the integrals are weakly singular. In subsequent examples, we shall consider a fully-discrete approximation based on function samples.

Using Theorem 5.2, we may estimate the stable sampling rate for this problem:

Proposition 5.4 *Let* $\mathrm{H} = L^2(0, 1)$, $w(x) = \log(x)$, $\{\varphi_n\}$ *be the Legendre basis on* H, Φ *be as in* (5.4) *and consider the sampling functionals* (5.1). *Then the stable sampling rate for this problem is linear in* N, *and specifically,*

$$\Theta^\epsilon(N, \theta) \leq \max \left\{ N, \frac{N-1}{\sqrt{1 - 1/\theta^2}} \right\}, \quad \forall \theta > 1, \ N \geq 2.$$

Proof The Legendre polynomials satisfy

$$\int_0^1 P_n(2x - 1) \log(x) \, dx = \frac{(-1)^{n+1}}{n(n+1)}, \quad n \geq 1.$$

This follows from the differential equation $((1 - x^2)P_n'(x))' + n(n+1)P_n(x) = 0$ after two integrations by parts. Let $\psi(x) = \log(x)$. Then, for $M \geq 1$,

$$\|\psi - \mathcal{Q}_M \psi\|^2 = \sum_{m > M} |\langle \psi, \varphi_m \rangle|^2 = \sum_{m \geq M} \frac{2m + 1}{m^2(m+1)^2} = \sum_{m \geq M} \left(\frac{1}{m^2} - \frac{1}{(m+1)^2} \right) = \frac{1}{M^2}.$$

Lemma 5.1 now gives

$$\sup_{\substack{g \in \mathrm{H}_N \\ \|g\|=1}} \|g - \mathcal{Q}_M g\| = \frac{N-1}{M}.$$

Therefore

$$\sup_{\substack{g \in \mathrm{H}_N \\ \|g\|=1}} \|g - \mathcal{Q}_M g\| \leq \sqrt{1 - 1/\theta^2},$$

provided

$$\frac{N-1}{M} \leq \sqrt{1 - 1/\theta^2} \quad \Leftrightarrow \quad M \geq \frac{N-1}{\sqrt{1 - 1/\theta^2}},$$

as required. □

5.3 ONB+K Frames

Functions with logarithmic singularities can be more accurately approximately using the frame (5.4) than the Legendre polynomial basis alone. However, the accuracy may be limited, due to the presence of weak logarithmic singularities. To increase the accuracy one may consider a frame of the type

$$\Phi = \{\varphi_n\}_{n=1}^{\infty} \cup \{\psi_k\}_{k=1}^{K}, \tag{5.6}$$
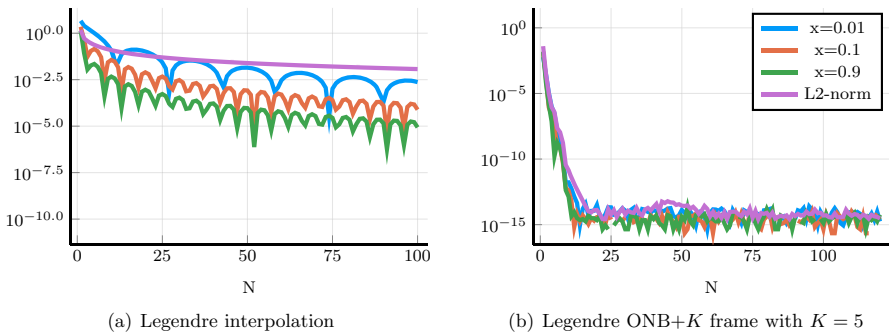
for fixed $K \geq 1$, where $\psi_k(x) = w(x)\varphi_k(x)$.

(a) Legendre interpolation   (b) Legendre ONB+K frame with $K = 5$

Fig. 2 Pointwise error as a function of the polynomial degree N for the approximation of the logarithmically singular function (5.5) on $[0, 1]$, using Legendre interpolation (left panel) and the ONB+K frame Φ with Legendre polynomials, $K = 5$ and $w(x) = \log x$ (right panel). The error is shown in three points in the interval $[0, 1]$ as well as in the L^2-norm on $[0, 1]$. In both cases, the samples are function evaluations in the Legendre nodes. The left panel is based on interpolation, the right panel corresponds to a discrete least squares approximation with $M = 2N$ and regularization threshold $\epsilon = 2e^{-13}$
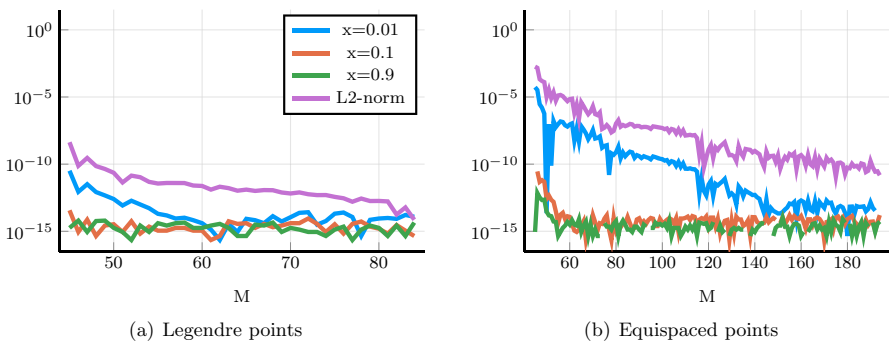

(a) Legendre points   (b) Equispaced points

Fig. 3 Pointwise error as a function of the number of samples M for the ONB+K frame based on Legendre polynomials, with degree $N = 40$, $K = 5$ and $w(x) = \log x$ as in Fig. 2b. Similar to regular polynomial approximation, using Legendre points yields better accuracy than using equispaced points. They also require less oversampling, i.e., smaller values of M, to achieve the best error
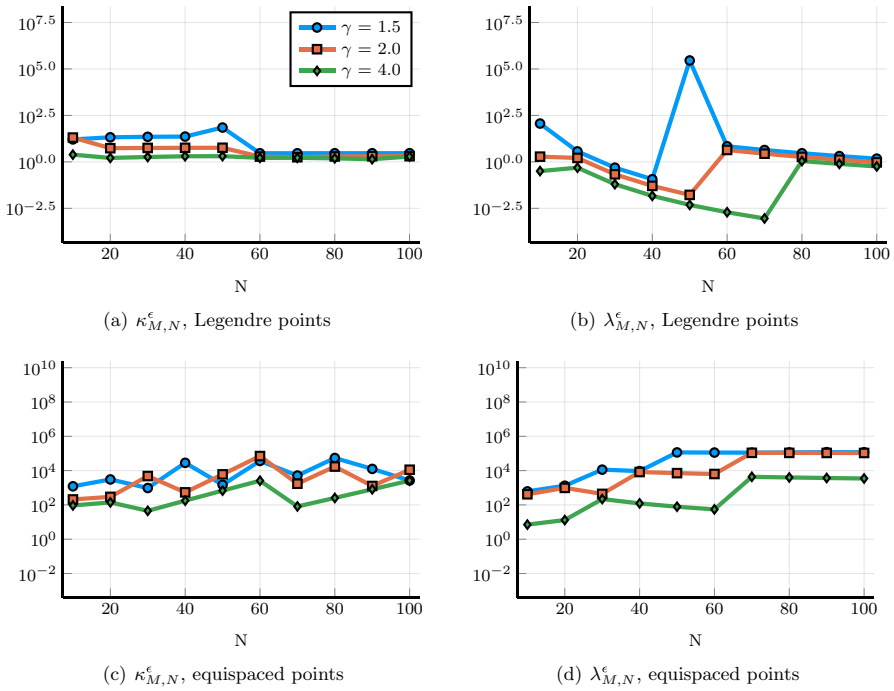
(a) $\kappa^{\epsilon}_{M,N}$, Legendre points

(b) $\lambda^{\epsilon}_{M,N}$, Legendre points

(c) $\kappa^{\epsilon}_{M,N}$, equispaced points

(d) $\lambda^{\epsilon}_{M,N}$, equispaced points

Fig. 4 The values of $\kappa^{\epsilon}_{M,N}$ and $\lambda^{\epsilon}_{M,N}$ are shown as a function of N with constant oversampling $M = \gamma N$ and varying factors γ. Legendre points (top row) and equispaced points (bottom row) are used for the ONB+K frame using Legendre polynomials, $w(x) = \log(x)$ and $K = 5$. The threshold used here is $\epsilon = 1e^{-5}$: the values are bounded for Legendre points but they approach $1/\epsilon$ for equispaced points. Equispaced points require more than linear oversampling

Proposition 5.5 *Let* $\{\varphi_n\}$ *be the Legendre basis on* H *and* $w \in L^2(0, 1)$. *Then* (5.6) *is a frame for any fixed $K \geq 1$, with frame bounds*

$$1 \leq A \leq B \leq 1 + \|w\|^2 K^2.$$

Proof First observe that $\psi_k \in L^2(0, 1)$ since $w \in L^2(0, 1)$ and $\varphi_k \in L^{\infty}(0, 1)$. Second, we have

$$\|f\|^2 = \sum_{n=1}^{\infty} |\langle f, \varphi_n\rangle|^2 \leq \sum_{n=1}^{\infty} |\langle f, \varphi_n\rangle|^2 + \sum_{k=1}^{K} |\langle f, \psi_k\rangle|^2,$$

and therefore the lower frame condition holds with $A \geq 1$. Moreover,

(a) $\kappa^{\epsilon}_{M,N}$, Legendre points

(b) $\lambda^{\epsilon}_{M,N}$, Legendre points

(c) $\kappa^{\epsilon}_{M,N}$, equispaced points

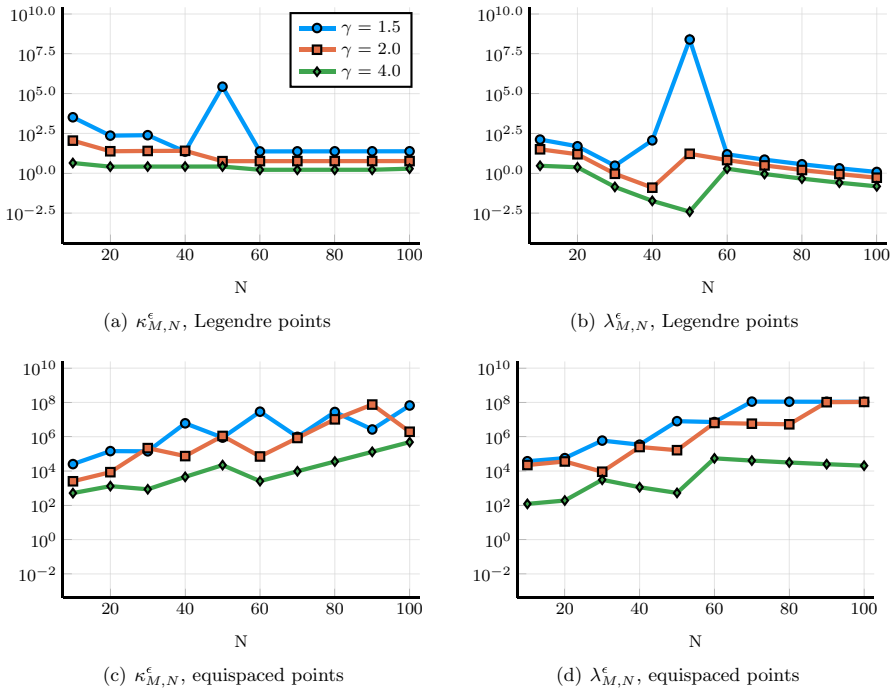(d) $\lambda^{\epsilon}_{M,N}$, equispaced points

Fig. 5 The same as Fig. 4 but with threshold $\epsilon = 1e^{-8}$. For Legendre nodes (top row), the intermediate peaks are higher, corresponding to the larger upper bounds in (3.23). However, the values settle down for increasing N, with a clear benefit for larger oversampling factors. This is in agreement with the limit (3.24). For equispaced samples, linear oversampling is not sufficient and, as in the previous figure, the values seem to approach $1/\epsilon$. This again corresponds to (3.23)

$$\sum_{n=1}^{\infty} |\langle f, \varphi_n \rangle|^2 + \sum_{k=1}^{K} |\langle f, \psi_k \rangle|^2 \leq \|f\|^2 + \|f\|^2 \|w\|^2 \sum_{k=1}^{K} \|\varphi_k\|_{L^{\infty}}^2$$

$$= \|f\|^2 \left(1 + \|w\|^2 \sum_{k=1}^{K} (2k-1) \right)$$

$$= \|f\|^2 \left(1 + \|w\|^2 K^2 \right).$$

Here, in the penultimate step, we use the fact that $|\varphi_n(x)| \leq \varphi_n(1) = \sqrt{2n-1}$ for $0 \leq x \leq 1$. This completes the proof. ☐

In Fig. 2 we demonstrate the benefits of this frame for approximating the singular function given by (5.5). Here, rather than the inner products (5.1), we take the sampling functionals to be pointwise evaluations at the Legendre nodes, i.e., at the roots of a high degree Legendre polynomial, mapped from $[-1, 1]$ to $[0, 1]$. We also compare these approximations with the polynomial interpolant at these nodes. As is evident, polynomial interpolation performs poorly. In contrast, the convergence for the ONB+K frame is significantly faster. Figure 2 also illustrates the stability of the

numerical approximation using oversampling with $M = 2N$. The accuracy reaches machine precision, in spite of it requiring the solution of an extremely ill-conditioned linear system of equations, and this high level of accuracy is maintained as N grows.

The influence of the oversampling factor M is illustrated in Fig. 3. Here, the error is shown as a function of M, for constant $N = 40$. Best accuracy is only achieved for $M > N$, i.e., when using some amount of oversampling. We have used discrete sampling in this figure using Legendre points (left panel) and equispaced points on [0, 1] (right panel). It is not unexpected that Legendre points are a better choice: less oversampling is needed to achieve the best accuracy for the given N.

The behaviour shown in Fig. 3 can be explained by computing the corresponding constants $\kappa^\epsilon_{M,N}$ and $\lambda^\epsilon_{M,N}$. This also serves to illustrate the stable sampling rate for this problem. Their values are shown in Fig. 4 for several choices of the oversampling factor γ, with M and N such that $M = \gamma N$. The convergence of both values to constants of modest size, in particular much smaller than $1/\epsilon$, suggests that the stable sampling rate is indeed linear when sampling in the Legendre points. For equispaced points, as could be expected, this does not seem to be the case.

The results in Fig. 4 correspond to the threshold $\epsilon = 1e^{-5}$. For comparison, the experiment is repeated in Fig. 5 for the smaller threshold $\epsilon = 1e^{-8}$. The latter figure illustrates the larger upper bound of (3.23), on the order of $1/\epsilon$, in the pre-asymptotic regime. Still, for the case of Legendre nodes, linear oversampling is sufficient to reach the ϵ-independent (and small) limit (3.24).

The constants were computed following the approach described in Remark 3.14. We have run this experiment in higher precision arithmetic, in order to exclude the possibility of inaccuracies in their computation. An exponentially converging composite hp-graded quadrature rule was used to approximate the singular integrals that arise in the elements of the Gram matrix. Furthermore, in this case we also weighted the discrete samples in Legendre points by the square roots of the corresponding Gauss–Legendre quadrature weights: this discrete normalization ensures that $A' = B' = 1$ in (3.2) and leads to slightly smaller values of the constants (and, correspondingly, smaller error in the approximation).

References

1. Adcock, B., Hansen, A.C.: A generalized sampling theorem for stable reconstructions in arbitrary bases. J. Fourier Anal. Appl. **18**(4), 685–716 (2012)
2. Adcock, B., Hansen, A.C.: Stable reconstructions in Hilbert spaces and the resolution of the Gibbs phenomenon. Appl. Comput. Harmon. Anal. **32**(3), 357–388 (2012)
3. Adcock, B., Huybrechs, D.: Frames and numerical approximation. SIAM Rev. **61**(3), 443–473 (2019)

4. Adcock, B., Hansen, A.C., Poon, C.: Beyond consistent reconstructions: optimality and sharp bounds for generalized sampling, and application to the uniform resampling problem. SIAM J. Math. Anal. **45**(5), 3114–3131 (2013)

5. Adcock, B., Huybrechs, D., Martín-Vaquero, J.: On the numerical stability of Fourier extensions. Found. Comput. Math. **14**(4), 635–687 (2014)

6. Boffi, D., Cavallini, N., Gastaldi, L.: The finite element immersed boundary method with distributed Lagrange multiplier. SIAM J. Numer. Anal. **53**(6), 2584–2604 (2015)

7. Christensen, O.: An Introduction to Frames and Riesz Bases. Applied and Numerical Harmonic Analysis, 2nd edn. Birkhäuser, Basel (2016)

8. Coppé, V., Huybrechs, D.: Efficient function approximation on general bounded domains using splines on a cartesian grid. Technical Report, KU Leuven. arXiv:1911.07894 (2020)

9. Coppé, V., Huybrechs, D., Matthysen, R., Webb, M.: The AZ algorithm for least squares problems with a known incomplete generalized inverse. SIAM J. Math. Anal. Appl. **41**(3), 1237–1259 (2020)

10. Fix, G.J., Gulati, S., Wakoff, G.I.: On the use of singular functions with finite elements approximations. J. Comput. Phys. **1**, 209–228 (1973)

11. Golub, G.H., van Loan, C.F.: Matrix Computations, 3rd edn. Johns Hopkins University Press, Baltimore (1996)

12. Gröchenig, K.: Sampling, Marcinkiewicz–Zygmund inequalities, approximation, and quadrature rules. J. Approx. Theory **257**, 105455 (2020)

13. Hansen, A.C.: On the solvability complexity index, the n-pseudospectrum and approximations of spectra of operators. J. Am. Math. Soc. **24**(1), 81–124 (2011)

14. Heinemeyer, E., Lindner, M., Potthast, R.: Convergence and numerics of a multisection method for scattering by three-dimensional rough surfaces. SIAM J. Numer. Anal. **46**(4), 1780–1798 (2008)

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

16. Huybrechs, D., Olteanu, A.-E.: An oversampled collocation approach of the wave based method for Helmholtz problems. Wave Motion **87**, 92–105 (2019)

17. Kasolis, F., Wadbro, E., Berggren, M.: Analysis of fictitious domain approximations of hard scatterers. SIAM J. Numer. Anal. **2015**(5), 2347–2362 (2015)

18. Lindner, M.: Infinite Matrices and their Finite Sections. Frontiers in Mathematics. Birkhäuser, Basel (2006)

19. Lubinsky, D.S.: Marcinkiewicz–Zygmund Inequalities: Methods and Results, pp. 213–240. Springer, Dordrecht (1998)

20. Lyon, M.: A fast algorithm for Fourier continuation. SIAM J. Sci. Comput. **33**(6), 3241–3260 (2012)

21. Matthysen, R., Huybrechs, D.: Fast algorithms for the computation of Fourier extensions of arbitrary length. SIAM J. Sci. Comput. **38**(2), A899–A922 (2016)

22. Matthysen, R., Huybrechs, D.: Function approximation on arbitrary domains using fourier frames. SIAM J. Numer. Anal. **56**, 1360–1385 (2018)

23. Roache, P.J.: A pseudo-spectral FFT technique for non-periodic problems. J. Comput. Phys. **27**, 204–220 (1978)

24. Shirokoff, D., Nave, J.-C.: A sharp-interface active penalty method for the incompressible Navier–Stokes equations. J. Sci. Comput. **62**(1), 53–77 (2015)