DISSERTATION

Titel der Dissertation

Sampling and reconstruction in distinct subspaces using oblique projections

Verfasser

Peter Berger

angestrebter akademischer Grad

Doktor der Naturwissenschaften (Dr.rer.nat)

Wien, Juni 2015

Studienkennzahl lt. Studienblatt:	A 796 605 405
Studienrichtung lt. Studienblatt:	Mathematik
Betreuer:	Univ.-Prof. Dr. Karlheinz Gröchenig

Contents

Abstract **v**

Zusammenfassung **vii**

Acknowledgements **ix**

1 Introduction **1**

1.1 Main contributions of the first part of the thesis 8

1.2 The geometric idea of frame independent sampling 14

1.3 Short description of the hierarchical reconstruction algorithm . 15

1.4 List of used symbols 20

2 Frames and the Pseudoinverse **21**

2.1 Collection of Theorems from functional analysis 21

2.2 Properties of the pseudoinverse 23

2.3 Frames in Hilbert Spaces 25

3 Stability, quasi optimality and error estimates **33**

3.1 Problem formulation and notation 33

3.2 Stability, quasi optimality 35

3.3 Absolute and relative error 36

3.4 Oblique projections and subspace angles 37

4 Connection of quasi-optimal operators to dual frames **43**

5 Consistent, generalized and frame independent sampling **49**
5.1 Consistent sampling 50
5.2 Generalized sampling 52
5.3 Frame independent sampling 64
5.4 A combination of generalized and frame independent sampling 79
5.5 Connection to Statistics 84
5.6 Calculation of the approximations and conditions for them to coincide 93

6 Numerical experiments for reconstruction from Fourier measurements **99**
6.1 The noiseless case 100
6.2 Errors appearing after the sampling process 105
6.3 Errors appearing before the sampling process 108
6.4 Summary of the experiments 112

7 A hierarchical reconstruction algorithm **115**
7.1 Description of the setup 116
7.2 Solving the global least squares problem 116
7.3 Special case of a reproducing kernel Hilbert space 118
7.4 Preprocessed data 120
7.5 Hierarchical Algorithm 124
7.6 Stability of the algorithms 131
7.7 Numerical experiments 136
7.8 Operation count of the algorithm using orthonormal systems locally 146

8 Conclusions and future research **151**

Bibliography **153**

Curriculum Vitae **161**

Abstract

This PhD project is about sampling and reconstruction of signals in distinct subspaces. The reconstruction method presented in the thesis uses the optimal weighting of the measurements and therefore projects as close as possible to the orthogonal projection onto the reconstruction space. Furthermore, we prove that this reconstruction method is the most stable with respect to a systematic error appearing before the sampling process. Adcock, Gataric and Hansen weighted the point samples of the Fourier transform in order to obtain a projection direction closer to the orthogonal projection onto the reconstruction space. This theory can be used for example to approximate compactly supported functions from nonuniform point samples of the Fourier transform.

Weighting the measurements has a major drawback that the stability with respect to measurement errors is reduced. The approximation calculated from unweighted measurements is in some sense most stable with respect to error present in the measurements and is called generalized sampling by Adcock and Hansen. We investigate how to vary continuously between the two extreme reconstruction methods.

In the last chapter, we consider the reconstruction of a non-bandlimited function represented by a finite number of compactly supported generating functions in wireless sensor networks. Using the theory presented in the first part, we develop a novel hierarchical reconstruction. The idea is to preprocess the sensor measurements locally by taking inner products with suitable vectors and to send the resulting data (rather than sensor measurements) to a global fusion center for further processing. In oversampled regimes, this approach reduces communication workload.

Zusammenfassung

Diese Arbeit beschäftigt sich mit der Approximation von Signalen, wobei die Messungen dieser Signale in einem anderen Raum liegen als die Rekonstruktionen. Die in dieser Arbeit vorgestellte Rekonstruktionsmethode verwendet eine optimale Gewichtung der Messungen und projiziert daher so orthogonal wie möglich auf den Rekonstruktionsraum. Diese Herangehensweise ist auch optimal, wenn ein systematischer Fehler vor dem Messprozess auftritt.
Adcock, Gataric und Hansen gewichteten die Punktauswertungen der Fourier-Transformation, um eine bessere Projektionsrichtung auf den Rekonstruktionsraum zu erhalten. Diese Theorie kann zum Beispiel verwendet werden, um Funktionen mit kompaktem Träger von unregelmäßigen Punktauswertungen der Fourier-Transformation zu approximieren.

Das Gewichten der Messungen hat jedoch einen großen Nachteil: Die Stabilität bezüglich des Fehlers auf den Messwerten wird reduziert. Die Berechnung der Approximation ohne Gewichte ist in gewissem Sinne am stabilsten bezüglich des Fehlers auf den Messwerten und wird von Adcock und Hansen als "generalized sampling" bezeichnet. Wir zeigen, wie man stetig zwischen diesen beiden extremen Rekonstruktionsmethoden variieren kann.

Im letzten Kapitel beschäftigen wir uns mit der Approximation einer Funktion in drahtlosen Sensornetzwerken, welche durch eine Linearkombination endlich vieler Basisfunktionen mit kompaktem Träger modelliert werden. Wir verwenden die im ersten Teil der Arbeit präsentierte Theorie, um einen neuen hierarchischen Rekonstruktionsalgorithmus zu entwickeln. Die Absicht ist, die Messungen lokal zu verarbeiten (durch Berechnung innerer Produkte mit bestimmten Vektoren) und die resultierenden Daten anstatt der Sensormessungen zu übertragen. In überabgetasteten Systemen führt diese Herangehensweise zu einer Reduktion der Kommunikationsarbeit.

Acknowledgements

First of all, I would like to thank my advisor Professor Karlheinz Gröchenig. Without his great support this thesis would not exist. I am greatful for his excellent dedication as a scientific adviser. I want to thank Professor Hans Georg Feichtinger who invited me to join the Numerical Harmonic Analysis Group of the University of Vienna. I am also grateful to the members of the research network SISE, especially Gerald Matz for stimulating discussions and comments. I am grateful to Martin Ehler for his timely assistance. I want to thank also all other members of the Numerical Harmonic Analysis Group. It was great to work in such an inspiring and hospitable working environment. Special thanks to Andreas Klotz, Gero Fendler, Nicki Holighaus, Jose Luis Romero, Markus Faulhuber, Christoph Wiesmeyr, Sebastian Schmutzhard, Manuel Gräf and David Rottensteiner who gave me comments about my thesis.

My work on the thesis was partially funded by the FWF NFN S106 "Signal and Information Processing in Science and Engineering" (SISE) as well as by Vienna Science and Technology Fund (WWTF) through project VRG12-009. I am grateful for this support.

I want to thank especially my parents Peter and Sibylle for their big support and encouragement in anything I do. I would also like to thank my sister Maria and brother Paul. Ruth, thanks for the wonderful time we have and will have together. It is great to have such a strong family around me!

Chapter 1

Introduction

An important task in sampling theory is the reconstruction of a bandlimited signal f of finite energy from point samples $\{f(x_k)\}_{k\in\mathbb{N}}$. A function f is an element of the space of bandlimited functions B_Ω, if the Fourier transform

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x)e^{-2\pi i\xi x}\,\mathrm{d}x$$

has a compact support in the interval $\Omega = [-\omega, \omega]$, i.e., $\hat{f}(\xi) = 0$ for all $\xi \notin [-\omega, \omega]$. The classical result of Whittaker [67] states that a function $f \in L^2(\mathbb{R}) \cap B_{[-\frac{1}{2},\frac{1}{2}]}$ can be recovered exactly from its samples $\{f(k)\}_{k\in\mathbb{Z}}$ by the interpolation formula

$$f(x) = \sum_{k\in\mathbb{Z}} f(k)\mathrm{sinc}(k - x), \tag{1.1}$$

where $\mathrm{sinc}(x) = \frac{\sin(\pi x)}{\pi x}$. We observe that

$$f(k) = \int_{-\frac{1}{2}}^{\frac{1}{2}} \hat{f}(\xi)e^{2\pi i\xi k}\,\mathrm{d}\xi = \mathcal{F}\hat{f}(-k),$$

and therefore the reconstruction of a bandlimited signal from point evaluations is equivalent to the reconstruction of a compactly supported function $\hat{f}$ from point samples of its Fourier transform $\mathcal{F}\hat{f}$. This problem has a long list of applications, such as Magnetic Resonance Imaging (MRI), Computed

Tomography (CT), microscopy, seismology and geophysical imaging.

Taking the Fourier transform of (1.1), we obtain

$$\hat{f}(\xi) = \sum_{k\in\mathbb{Z}} f(k)e^{-2\pi ik\xi} = \sum_{k\in\mathbb{Z}} \langle \hat{f}, \chi_{[-\frac{1}{2},\frac{1}{2}]}\ e^{-2\pi ik\cdot}\rangle e^{-2\pi ik\xi}.$$

Therefore, (1.1) follows from the fact that the set $\{e^{2\pi ik\cdot}\}_{k\in\mathbb{N}}$ forms an orthonormal basis of $L^2([-\frac{1}{2}, \frac{1}{2}])$, the so called *harmonic Fourier basis*.

From now on we consider this problem on the Fourier side. This means that we reconstruct a compactly supported function $g \in L^2([-\frac{1}{2}, \frac{1}{2}])$ from point samples of the Fourier transform $\langle g, e^{2\pi i\omega_k\cdot}\rangle$, where $\omega_k \in \mathbb{R}$, $k \in \mathbb{N}$. If $\omega_k = k + \varepsilon_k$ and $\sup|\varepsilon_k| < \frac{1}{4}$, then $\{e^{2\pi i\omega_k\cdot}\}_{k\in\mathbb{Z}}$ is a *Riesz basis* of $L^2([-\frac{1}{2}, \frac{1}{2}])$, see [44]. A *Riesz basis* is the image of an orthonormal basis under a bounded and invertible operator. For a *Riesz basis* $\{e^{2\pi i\omega_k}\}_{k\in\mathbb{Z}}$ there exists a unique dual basis $\{f_k\}_{k\in\mathbb{Z}}$ such that for $g \in L^2([-\frac{1}{2}, \frac{1}{2}])$

$$g(\xi) = \sum_{k\in\mathbb{Z}} \langle g, e^{2\pi i\omega_k\cdot}\rangle f_k(\xi).$$

Let us assume that we are given the perturbed measurements

$$d_k = \langle g, e^{2\pi i\omega_k\cdot}\rangle + \delta_k,$$

where $\{\delta_k\}_{k\in\mathbb{N}} \in l^2(\mathbb{N})$ is the measurement error. The sampling frequencies are assumed to be fixed and given. Since $\{e^{2\pi i\omega_k\cdot}\}_{k\in\mathbb{Z}}$ is a Riesz basis, for any sequence in $\{d_k\}_{k\in\mathbb{N}} \in l^2(\mathbb{N})$, the moment problem

$$\left\langle \tilde{g}, e^{2\pi i\omega_k\cdot}\right\rangle = d_k, \quad k \in \mathbb{N}, \tag{1.2}$$

has the solution $\tilde{g} = \sum_{k\in\mathbb{Z}} d_k f_k$. For a detailed explanation see [68, Section 4]. In this case $\tilde{g}$ is called a consistent reconstruction, since the measurements of $\tilde{g}$ coincide with the given data. This concept is treated for example in [20, 22, 25–28]. In Section 5.1, we review some properties of consistent reconstructions, since the reconstruction operators discussed in this thesis are generalizations of this concept.

Since (1.2) is equivalent to

$$\mathcal{F}\tilde{g}(\omega_k) = d_k,$$

the function $\mathcal{F}\tilde{g}$ interpolates the data points $\{(\omega_k, d_k)\}_{k\in\mathbb{N}}$. Therefore the set $\{\omega_k\}_{k\in\mathbb{N}}$ is called a *set of interpolation*, see for example [38, 57].

When dealing with more general sampling sets than perturbation sets of $\mathbb{Z}$, instead of Riesz bases we need the more general notion of frames, introduced by Duffin and Schaeffer [23]. A sequence $\{e^{2\pi\omega_k\cdot}\}_{k\in\mathbb{N}}$ is a frame for $L^2([-\frac{1}{2},\frac{1}{2}])$ if there exist constants $A, B > 0$ such that for all $g \in L^2([-\frac{1}{2},\frac{1}{2}])$

$$A\|g\|^2 \leqslant \sum_{k\in\mathbb{Z}} |\langle g, e^{2\pi i\omega_k\cdot}\rangle|^2 \leqslant B\|g\|^2,$$

or equivalently for all $f \in L^2(\mathbb{R}) \cap B_{[-\frac{1}{2},\frac{1}{2}]}$

$$A\|f\|^2 \leqslant \sum_{k\in\mathbb{Z}} |f(\omega_k)|^2 \leqslant B\|f\|^2.$$

In this case $X := \{\omega_k \in \mathbb{R}\}_{k\in\mathbb{Z}}$ is called a *set of stable sampling*. For $\{e^{2\pi i\omega_k\cdot}\}_{k\in\mathbb{N}}$ to be a frame for $L^2([-\frac{1}{2},\frac{1}{2}])$ it is sufficient that the lower Beurling density

$$D(X) = \lim_{r\to\infty} \inf_{y\in\mathbb{R}} \frac{\operatorname{card}\big(X \cap (y + [0, r])\big)}{r}$$

satisfies $D(X) > 1$, see [16, 17]. Conversely, if $\{e^{2\pi i\omega_k\cdot}\}_{k\in\mathbb{N}}$ is a frame for $L^2([-\frac{1}{2},\frac{1}{2}])$, then $D(X) \geqslant 1$, see [46] and [33].

When $\{e^{2\pi\omega_k\cdot}\}_{k\in\mathbb{N}}$ is a frame for $L^2([-\frac{1}{2},\frac{1}{2}])$, then for $g \in L^2([-\frac{1}{2},\frac{1}{2}])$

$$g(\xi) = \sum_{k\in\mathbb{Z}} \langle g, e^{2\pi i\omega_k\cdot}\rangle f_k(\xi),$$

where $\{f_k\}_{k\in\mathbb{N}}$ is a so called *dual frame* of $\{e^{2\pi i\omega_k\cdot}\}_{k\in\mathbb{N}}$. Let

$$S : L^2\Big(\Big[-\frac{1}{2},\frac{1}{2}\Big]\Big) \to L^2\Big(\Big[-\frac{1}{2},\frac{1}{2}\Big]\Big), \quad Sg = \sum_{k\in\mathbb{Z}} \langle g, e^{2\pi i\omega_k\cdot}\rangle e^{2\pi i\omega_k\cdot}$$

denote the *frame operator* of the frame $\{e^{2\pi\omega_k \cdot}\}_{k\in\mathbb{N}}$. For any $g \in L^2([-\frac{1}{2}, \frac{1}{2}])$

$$g(\xi) = S^{-1}Sg(\xi) = \sum_{k\in\mathbb{Z}} \langle g, e^{2\pi i\omega_k \cdot}\rangle S^{-1}\left(e^{2\pi i\omega_k \cdot}\right)(\xi).$$

The frame $\{S^{-1}e^{2\pi i\omega_k \cdot}\}_{k\in\mathbb{N}}$ is called the *canonical dual frame.* In general, not every sequence $\{d_k\}_{k\in\mathbb{N}} \in l^2(\mathbb{N})$ can be realized as a sequence of inner products of an element $\tilde{g} \in L^2([-\frac{1}{2}, \frac{1}{2}])$ with frame vectors. If we choose as dual frame $\{f_j\}_{j\in\mathbb{N}}$ the canonical dual frame, then $\tilde{g} = \sum_{j\in\mathbb{Z}} d_j f_j$ is the least squares solution of

$$\tilde{g} = \underset{f\in L^2([-\frac{1}{2},\frac{1}{2}])}{\arg\min} \sum_{j\in\mathbb{Z}} \left|\langle f, e^{2\pi i\omega_j \cdot}\rangle - d_j\right|^2.$$

This least squares fit is also common when dealing with a finite number of samples $\{d_j\}_{j=-n}^n$. Clearly a function $g \in L^2([-\frac{1}{2}, \frac{1}{2}])$ is not determined by a finite number of samples. Therefore some system model

$$\mathcal{T} = \left\{\sum_{k=-m}^{m} c_k g_k : c_k \in \mathbb{C}\right\}$$

for the function to be approximated is used, followed by solving the (typically overdetermined) least squares problem

$$\tilde{g} = \underset{g\in\mathcal{T}}{\arg\min} \sum_{j=-n}^{n} \left|\langle g, e^{2\pi i\omega_j \cdot}\rangle - d_j\right|^2. \tag{1.3}$$

We observe that $\tilde{g}$ can be written in the form $\tilde{g} = \sum_{k=-m}^{m} \hat{c}_k g_k$. The vector $\hat{c} = [\hat{c}_{-m}, \ldots, \hat{c}_m]^T \in \mathbb{C}^{2m+1}$ containing the coefficients of the series expansion of $\tilde{g}$ is the solution of the overdetermined least squares problem

$$\hat{c} = \arg\min_{c} \|Ac - d\|, \tag{1.4}$$

where $A \in \mathbb{C}^{2n+1\times 2m+1}$ is defined by

$$A(j,k) = \langle g_k, e^{2\pi i\omega_j \cdot}\rangle,$$

and $d = [d_{-n}, \dots, d_n]^T \in \mathbb{C}^{2n+1}$. When reconstructing a bandlimited function from point samples, it is most common to use as system model trigonometric polynomials, see for example [29, 34, 35, 60].

On the Fourier side (when reconstructing a compactly supported function from samples of the Fourier transform), if the function to be reconstructed is smooth and periodic, the classical approach also uses trigonometric polynomials as a system model, i.e., it computes the truncated Fourier series representation. However, for non-periodic or discontinuous functions, the Fourier series representation suffers from the Gibbs phenomenon and slow convergence. The first approaches to a mitigation of the Gibbs phenomenon were based on projection and filtering, see [30, 32, 62, 64] for recent contributions. For reconstruction of compactly supported functions from non-uniform samples of their Fourier transform see [66]. A new approach to mitigate this problem is to expand the function in basis functions other than trigonometric polynomials. Of course, for different classes of functions different basis functions are needed, and a priori knowledge of the function to reconstruct should be taken into account. For example, smooth functions are efficiently approximated by algebraic polynomials. This motivates the Inverse Polynomial Reconstruction Method (IPRM), see [41–43, 58]. Given the first n Fourier coefficients, the IPRM approximates a compactly supported function g by an algebraic polynomial of degree at most $n-1$. In [39] a modified IPRM approach is presented, where the authors used n Fourier coefficients to determine the expansion coefficients of the first m Legendre polynomials P_k, $k = 0, \dots, m-1$, with $n > m$. This means that the authors approximate a compactly supported function g by a linear combination of the first m Legendre polynomials P_k, $k = 0, \dots, m-1$. Given the first n Fourier coefficients $\mathcal{F}g(l)$, $l = -\left\lfloor \frac{n-1}{2} \right\rfloor, \dots, \left\lfloor \frac{n}{2} \right\rfloor$, the approximation is calculated by solving the least squares problem (1.3) for $\mathcal{T}$ being the linear span of the first m Legendre polynomials. In [39] it is shown that the first m Legendre coefficients can be calculated in a stable way (by solving the least squares problem (1.4)) from the first n Fourier coefficients, provided that $n \geqslant m^2$. Furthermore, in [39] it is shown that for analytic functions the resulting algorithm has a root-exponential convergence rate in the number of the Fourier coefficients. The fact that there are more samples than reconstruction vectors is extremely important. For example, taking only $m = n$ measurements (the consistent approach) leads to exponentially growing condition numbers, see [11].

When solving the least squares problem (1.4) by an iterative method such as the conjugate gradient method applied to the normal equations, a condition number $\kappa(A)$ close to one results in fast convergence. In order to improve the convergence rate it is common practice to solve a preconditioned least squares problem

$$\hat{c} = \arg\min_{c} \|DAc - Dd\|$$

instead of (1.4) with a suitable preconditioner $D \in \mathbb{C}^{2n+1\times 2n+1}$. In sampling problems D is often chosen as a diagonal matrix, see [3, 5, 29, 34, 35, 60].

In [3–5] and [2] another positive effect of weighting the least squares problem is observed. By appropriate weights, a projection direction closer to the orthogonal projection onto the reconstruction space is obtained. Weighting the measurements has a major drawback that the stability with respect to measurement errors is reduced. The approximation calculated by solving the standard least squares problem (1.4) is most stable with respect to error present in the measurements. This will be explained in Section 1.1.

Let us formulate the least squares approach in an abstract setup. Let $\mathcal{H}$ denote a separable Hilbert space. Let $\{u_j\}_{j\in\mathbb{N}}$ be a frame sequence (i.e. a frame for its closed linear span $\mathcal{V}$) in $\mathcal{H}$. The operator $V^* : \mathcal{H} \to l^2(\mathbb{N})$ defined by the property

$$V^* g = \{\langle g, u_j\rangle\}_{j\in\mathbb{N}}$$

is called the *analysis operator* of the frame sequence $\{u_j\}_{j\in\mathbb{N}}$. The adjoint operator $V : l^2(\mathbb{N}) \to \mathcal{V}$,

$$Vc = \sum_{k\in\mathbb{N}} c_k u_j,$$

is called the *synthesis operator* of the frame sequence $\{u_j\}_{j\in\mathbb{N}}$, where $\mathcal{V} = \mathcal{R}(V) = \overline{\text{span}}\{u_j\}_{j\in\mathbb{N}}$. Let $S = VV^*$ denote the frame operator of $\{u_j\}_{j\in\mathbb{N}}$. Let $\mathcal{T}$ be a closed subspace of $\mathcal{H}$. Let the operator $Q_g : l^2(\mathbb{N}) \to \mathcal{T}$ be defined by the property

$$Q_g d = \arg\min_{g\in\mathcal{T}} \sum_{j\in\mathbb{N}} |\langle g, u_j\rangle - d_j|^2 = \arg\min_{g\in\mathcal{T}} \|V^* g - d\|^2. \qquad (1.5)$$

This means, given perturbed measurements $d = V^* g + \delta$ of an element $g \in \mathcal{H}$, the approximation $\tilde{g}$ of g is given by $\tilde{g} = Q_g d$. When we choose a frame $\{g_k\}_{k\in\mathbb{N}}$ for $\mathcal{T}$ with synthesis operator T,

$$Q_g d = \sum_{k=1}^{\infty} \hat{c}_k g_k$$

where $\hat{c} = \{\hat{c}_k\}_{k\in\mathbb{N}}$ is the minimal norm element of the set

$$\arg\min_{c} \|V^* T c - d\|. \qquad (1.6)$$

This operator is analyzed in detail in [10], and used in order to analyze the least squares approach with other basis functions than Legendre polynomials, such as wavelets, for reconstruction, see [1,6–9]. In [10] it is shown that $Q_g V^*$ is an oblique projection onto the *reconstruction space* $\mathcal{T}$. Specifically,

$$Q_g V^* = P_{\mathcal{T}, S(\mathcal{T})^\perp},$$

is the oblique projection with range $\mathcal{T}$ and null-space $S(\mathcal{T})^\perp$. The authors named the reconstruction method defined by this least squares approach *generalized sampling*. In Section 5.2, we review the concept of generalized sampling and supplement some additional aspects.

The authors of [10] use two quantities to measure the quality of a mapping $Q : l^2(\mathbb{N}) \to \mathcal{T}$. The first one is the *quasi-optimality* constant $\mu = \mu(Q) > 0$ which is defined as the smallest number μ, such that

$$\|f - QV^* f\| \leqslant \mu \|f - P_{\mathcal{T}} f\|, \quad \text{for all } f \in \mathcal{H},$$

where $P_{\mathcal{T}}$ denotes the orthogonal projection onto $\mathcal{T}$. The smaller the quasi-optimality constant, the closer the projection direction is to the orthogonal projection onto the reconstruction space, and for $\mu = 1$, $QV^* = P_{\mathcal{T}}$ is equal to the orthogonal projection onto the reconstruction space. As a measure of stability of the reconstruction, the absolute condition number [10, Definition 2.2] is used. Here we focus on linear reconstructions, in which case this quantity equals the operator norm $\|Q_{|\mathcal{R}(V^*)}\|$ of Q restricted to the subspace $\mathcal{R}(V^*)$. We show that for all reconstruction operators used throughout the thesis, the operator norm $\|Q\|$ coincides with the absolute condition number [10, Definition 2.2]. Therefore we use the operator norm $\|Q\|$ as a measure of the stability of an reconstruction operator Q.

1.1 Main contributions of the first part of the thesis

- Using the operator norm $\|Q\|$ as a measure of stability we derive the error estimate

$$\|f - Q(V^*f + c)\| \leqslant \mu(Q)\|f - P_{\mathcal{T}}f\| + \|Q\|\|c\|,$$

  for the approximations calculated by a particular reconstruction operator. This clarifies the meaning of the quasi-optimality constant and the absolute condition number introduced in [10].

- In Section 3.4, we collect some facts on oblique projections. Oblique projections are of great importance, because in Theorem 4.0.12 we show that any quasi-optimal and bounded reconstruction operator $Q : l^2(\mathbb{N}) \to \mathcal{T}$ has the property that the operator $P = QV^*$ is an oblique projection onto the reconstruction space. As shown in [10], in this case, the quasi-optimality constant is determined by the angle between the range and nullspace of this oblique projection.

- In [10, Theorem 6.2.] it is shown that the operator Q_g has the smallest possible absolute condition number among all operators Q with $QV^*g = g$ for all $g \in \mathcal{T}$ (i.e. Q recovers elements in the reconstruction space $\mathcal{T}$ in the absence of noise). As an easy consequence of this theorem, we show in Corollary 5.2.9 that the operator Q_g (generalized sampling) has the smallest possible operator norm among all operators Q with $QV^*g = g$ for all $g \in \mathcal{T}$. We also show that this implies that the operator Q_g has the smallest possible relative condition number for elements in the reconstruction space, see Corollary 5.2.11. Therefore this operator is well suited to reconstruct a function f lying inside the reconstruction space from noisy measurements. This is what we mean by the approximation calculated by solving the standard least squares problem (1.6) is most stable with respect to error present in the measurements.

- In Chapter 4, we point out connections of quasi-optimal and bounded reconstruction operators to dual frames. We show in Theorem 4.0.12 that any bounded and quasi-optimal operator $Q : l^2(\mathbb{N}) \to \mathcal{T}$ is the synthesis operator of a dual frame of $\{P_{\mathcal{T}} u_j\}_{j\in\mathbb{N}}$, where $P_{\mathcal{T}} u_j$ denotes the orthogonal projection of the jth sampling vector onto the reconstruction space.

- In Theorem 5.2.6, we present an alternative description of the reconstruction operator Q_g (generalized sampling) by using dual frames. Specifically, we show that Q_g is the synthesis operator of the canonical dual frame of $\{P_{\mathcal{T}} u_j\}_{j\in\mathbb{N}}$ of $\mathcal{T}$. By using the properties of the canonical dual frame we give an alternative proof to the statement that Q_g has the smallest possible operator norm among all operators Q with $QV^*g = g$ for all $g \in \mathcal{T}$, see Lemma 2.3.11 and Theorem 5.2.6.

- As already mentioned, in [2–5] it is shown that by weighting the least squares problem, a projection direction closer to the orthogonal projection onto the reconstruction space is obtained. We present a new approximation operator Q_f, defined by the properties that

$$Q_f V^* = P_{\mathcal{T}, P_{\mathcal{V}}(\mathcal{T})^\perp}, \tag{1.7}$$

and that Q_f is zero on the set $\mathcal{R}(V^*)^\perp$, the orthogonal complement of the range of V^*. Let us denote the square root of the pseudoinverse of a positive operator S by $S^{\frac{\dagger}{2}} := (S^{\dagger})^{\frac{1}{2}}$. Let $\{g_k\}_{k\in\mathbb{N}}$ be a frame for $\mathcal{T}$ with synthesis operator T. The operator Q_f can also be described by a preconditioned version of the least squares problem (1.4). For $d \in l^2(\mathbb{N})$

$$Q_f d = \sum_{k=1}^{\infty} \hat{c}_k g_k,$$

where $\hat{c} = \{\hat{c}_k\}_{k\in\mathbb{N}}$ is the minimal norm element of the set

$$\hat{c} = \arg\min_{c} \|(V^*V)^{\frac{\dagger}{2}} V^*Tc - (V^*V)^{\frac{\dagger}{2}} d\|.$$

In Section 5.3, the operator Q_f is analyzed in detail. In Theorem 5.3.8 we show that this reconstruction operator Q_f is the operator with the smallest possible quasi-optimality constant, namely

$$\mu(Q_f) = \frac{1}{\cos(\varphi_{\mathcal{T},\mathcal{V}})},$$

where $\varphi_{\mathcal{T},\mathcal{V}}$ is the angle between $\mathcal{T}$ and $\mathcal{V}$ (defined in (3.9)). This implies that the operator $P_{\mathcal{T},P_{\mathcal{V}}(\mathcal{T})^{\perp}}$ is as close to the orthogonal projection onto the reconstruction space $\mathcal{T}$ as possible. Therefore our reconstruction deals very well with the part of the function to reconstruct lying outside of the reconstruction space. This is tested in numerical experiments in Section 6, where we reconstruct compactly supported functions from nonuniform samples of the Fourier transform. If there are high irregularities in the sampling frequencies, the operator Q_f yields approximations of much higher accuracy than the operator Q_g (defined by the standard least squares fit (1.5)). This is illustrated in Figure 6.1, where we approximate the signum function from nonuniform samples of the Fourier transform.

We refer to the reconstruction operator Q_f as *frame independent sampling* to indicate that $Q_f V^* = P_{\mathcal{T},P_{\mathcal{V}}(\mathcal{T})^{\perp}}$ does not depend on the sequences $\{u_j\}_{j\in\mathbb{N}}$ and $\{g_k\}_{k\in\mathbb{N}}$ themselves, but only on their closed linear spans $\mathcal{V}$ and $\mathcal{T}$.

- In Section 5.5, we point out the connection of generalized sampling and frame independent sampling to unbiased estimators of minimal variance. Generalized sampling corresponds to an unbiased estimator of minimal variance for the reconstruction of an element $g \in \mathcal{T}$ from measurements perturbed by white noise, see Corollary 5.5.10. Frame independent sampling corresponds to an unbiased estimator of minimal variance for the reconstruction of an element $g \in \mathcal{T}$ from measurements of the perturbed object (white noise appearing before the sampling process), see Corollary 5.5.11. By determining an unbiased estimator of minimal variance for combinations of the two versions of noise, we show how to obtain mixtures between the two reconstructions, see Corollary 5.5.12. In Chapter 6, we compare the different reconstruction operators in numerical experiments for the reconstruction of compactly supported functions from nonuniform samples of the Fourier transform. We observe, that if we increase the standard deviation of the noise appearing before the sampling process (leaving the standard deviation of the noise appearing after the sampling process constant) we obtain a smaller quasi-optimality constant of the reconstruction at the cost of an increasing operator norm. This means that we investigate how to choose reconstruction operators corresponding to oblique projections closer to the orthogonal projection at the cost of a larger operator norm.

- In Section 5.4, we analyze these mixtures between the operators Q_g (generalized sampling) and Q_f (frame independent sampling). We show that they can be described by a single tuning parameter λ. Namely, for $A = V^*T$, $\Sigma_\lambda = \lambda I + V^*V$, and $d \in l^2(\mathbb{N})$,

  $$Q_\lambda d = \sum_{k=1}^{\infty} \hat{c}_k g_k,$$

  where $\hat{c} = \{\hat{c}_k\}_{k\in\mathbb{N}}$ is the solution of the last squares problem

  $$\hat{c} = \arg\min_c \|\Sigma_\lambda^{-\frac{1}{2}} Ac - \Sigma_\lambda^{-\frac{1}{2}} d\|. \tag{1.8}$$

We observe that $Q_0 = Q_f$ (frame independent sampling) and $Q_\infty = Q_g$ (generalized sampling).

In the presence of noise the operator Q_λ with a small value of λ yields more accurate reconstructions than both, $Q_f(= Q_0)$ and $Q_g(= Q_\infty)$. This is illustrated in Figure 6.2, where we reconstruct the signum functions from noisy, nonuniform samples of the Fourier transform.

- In Section 5.6, Lemma 5.6.1, we estimate the condition number of $A^*\Sigma_\lambda^{-1}A$, the matrix of the normal equations of the least squares problem (1.8), using an orthonormal system $\{g_k\}_{k\in\mathbb{N}}$ as reconstruction vectors. The larger the tuning parameter λ, the larger the upper bound of the condition number $\kappa(A^*\Sigma_\lambda^{-1}A)$. Therefore we expect

$$\kappa(A^*\Sigma_0^{-1}A) \leqslant \kappa(A^*\Sigma_\lambda^{-1}A) \leqslant \kappa(A^*A),$$

which is tested in numerical experiments in Section 6. This means that the matrix of the least squares problem corresponding to a reconstruction operator that projects closer to the orthogonal projection has a smaller condition number. Consequently, we expect faster convergence of an iterative algorithm when solving these least squares problems. As already mentioned, this is the motivation for preconditioning.

For $\lambda = 0$ we obtain the following bound:

$$\kappa(A^*\Sigma_0^{-1}A) \leqslant \frac{1}{\cos^2(\varphi_{\mathcal{T},\mathcal{V}})}.$$

We observe that this bound depends only on the angle between the reconstruction space and sampling space. If the reconstruction space is a subspace of the sampling space, then $\cos(\varphi_{\mathcal{T},\mathcal{V}}) = 1$. In this case $\kappa(A^*\Sigma_0^\dagger A) = 1$ and $\Sigma_0^\dagger = (V^*V)^\dagger$ is an "optimal" preconditioner.

In order to calculate approximations to the solution of these least squares problems at a low operation count, we first need to calculate approximations M_λ of Σ_λ^{-1}, followed by solving the normal equations

$$A^* M_\lambda A c = A^* M_\lambda d.$$

The theory of controlled frames, weighted frames and frame multipliers, see [13, 14, 18], may enable us to obtain approximations M_λ of Σ_λ^{-1} at a low operation count.

- While in general for different values of λ the reconstruction operators Q_λ are rather different, in Theorem 5.6.5 we show that they coincide whenever the sampling frame is tight.

- In [2–5] a different version of stability is considered. These papers consider systematic errors appearing before the sampling process, and the task is to reconstruct a function f from measurements of the form

$$s_j = \langle f + \Delta f, u_j \rangle.$$

The authors reconstruct a compactly supported function from nonuniform samples of the Fourier transform. They weight the least squares problem (1.4) in order to compensate for clustering of the sample points. The operator Q_f can be seen as an optimal solution for this problem in the following sense. In Theorem 5.3.10 we show that Q_f is the operator with the smallest possible constants $\mu > 0$ and $\beta > 0$, such that

$$\|f - QV^*(f + \Delta f)\| \leqslant \mu \|f - P_{\mathcal{T}} f\| + \beta \|\Delta f\|.$$

Therefore the operator Q_f deals very well with the part of the function to reconstruct lying outside of the reconstruction space, and also deals very well with systematic errors appearing before the sampling process. This is illustrated in Figure 6.3 and Figure 6.4, where we approximate the signum function and a trigonometric polynomial from nonuniform samples of the Fourier transform.

1.2 The geometric idea of frame independent sampling

We explain our reconstruction method in the following simple example. Let $\mathcal{H} = \mathbb{R}^3$, let $g \in \mathbb{R}^3$ be the reconstruction vector. Let $\mathcal{T}$ be the line spanned by g, the reconstruction space. Let $u_1, u_2 \in \mathbb{R}^3$ be two linearly independent vectors, the sampling vectors. Let $\mathcal{V}$ denote the plane spanned by the sampling vectors u_1, u_2, the sampling space. We intend to reconstruct an element $f \in \mathbb{R}^3$ from $\langle f, u_1 \rangle$ and $\langle f, u_2 \rangle$. From the measurements $\langle f, u_1 \rangle$ and $\langle f, u_2 \rangle$, we can calculate $P_{\mathcal{V}} f$, the orthogonal projection of f onto the plane $\mathcal{V}$. Conversely, $P_{\mathcal{V}} f$ determines $\langle f, u_1 \rangle$ and $\langle f, u_2 \rangle$.

Thus all the information we have about f is that f lies in the affine subspace $P_{\mathcal{V}} f + \mathcal{V}^{\perp}$, but we do not know the exact location of f in this affine subspace. Let $P_{\mathcal{T}, P_{\mathcal{V}}(\mathcal{T})^{\perp}}$ denote the oblique projection with range $\mathcal{T}$ and kernel $P_{\mathcal{V}}(\mathcal{T})^{\perp}$.

We assume that f, the element to be reconstructed, is close to the reconstruction space $\mathcal{T}$. Naturally, we now want to find $\tilde{f}$, the element of the reconstruction space $\mathcal{T}$ closest to $P_{\mathcal{V}} f + \mathcal{V}^{\perp}$. The two spaces $P_{\mathcal{V}} f + \mathcal{V}^{\perp}$ and $\mathcal{T}$ may, or may not intersect. In both cases, the element of $\mathcal{T}$ closest to $P_{\mathcal{V}} f + \mathcal{V}^{\perp}$ is exactly $P_{\mathcal{T}, P_{\mathcal{V}}(\mathcal{T})^{\perp}} f$. If they intersect, then $P_{\mathcal{T}, P_{\mathcal{V}}(\mathcal{T})^{\perp}} f = (P_{\mathcal{V}} f + \mathcal{V}^{\perp}) \cap \mathcal{T}$, and $\langle f, u_1 \rangle = \langle \tilde{f}, u_1 \rangle$ and $\langle f, u_2 \rangle = \langle \tilde{f}, u_2 \rangle$. In this case $\tilde{f}$ is a so called consistent reconstruction of f, a concept that is treated for example in [20, 22, 25–28].

1.3 Short description of the hierarchical reconstruction algorithm

In the last chapter of the thesis, we use the theory of sampling and reconstruction in distinct subspaces to develop a novel hierarchical reconstruction method from nonuniform point samples. The idea is to preprocess the sensor measurements locally and to send the resulting data (rather than sensor measurements) to a global fusion center for further processing. A similar reconstruction method is presented in [54]. Our reconstruction method can be seen as a refinement of that approach. The advantage of preprocessing the sensor measurements locally is that in oversampled regimes the communication workload can be reduced. Specifically, we treat the problem of reconstructing an unknown continuous function $f \in L^2(\mathbb{R}^d)$, $d \in \mathbb{N}$, from noisy point evaluations

$$d_i := f(x_i) + \delta_i, \quad i = 1, \ldots, n. \tag{1.9}$$

The value d_i can be viewed as the measurement of the ith sensor positioned at x_i. As a function model we choose a finite dimensional subspace $\mathcal{W}$ of $L^2(\mathbb{R}^d)$, generated by linearly independent continuous functions $g_k \in L^2(\mathbb{R}^d)$, $k = 1, \ldots, m$,

$$\mathcal{W} = \left\{ \sum_{k=1}^{m} c_k g_k : c_k \in \mathbb{C} \right\}.$$

If the sensor positions $\{x_i\}_{i=1}^n$ are known, a common approach to this problem is to approximate the function f by the linear combination

$$\tilde{f} = \sum_{k=1}^{m} \hat{c}_k g_k,$$

where the vector $\hat{\mathbf{c}} = [\hat{c}_1, \ldots, \hat{c}_m]^T$ is the solution of the least squares problem

$$\hat{\mathbf{c}} = \arg\min_{\mathbf{c}} \sum_{i=1}^{n} \left| \sum_{k=1}^{m} c_k g_k(x_i) - d_i \right|^2. \tag{1.10}$$

We denote by $\mathbf{h} = [f(x_1), \ldots, f(x_n)]^T \in \mathbb{C}^n$ the vector containing the point evaluations of f, and by $\boldsymbol{\delta} \in \mathbb{C}^n$

$$\boldsymbol{\delta} = [\delta_1, \ldots, \delta_n]^T \tag{1.11}$$

we denote the vector containing the measurement noise. We define the matrix $A \in \mathbb{C}^{n\times m}$ by prescribing its elements

$$A(i,k) = g_k(x_i), \quad i = 1, \ldots, n, \; k = 1, \ldots, m.$$

With this notation, the vector $\hat{\mathbf{c}} \in \mathbb{C}^{m\times 1}$, defined by (1.10), is the solution of the least squares problem

$$\hat{\mathbf{c}} = \arg\min_{\mathbf{c}} \|A\mathbf{c} - (\mathbf{h} + \boldsymbol{\delta})\|. \tag{1.12}$$

If the signal is oversampled by a factor of s, i.e. $n = sm$, then by (1.10) we calculate m coefficients from $n = sm$ noisy point evaluations. This means that the information sent is s times redundant. Using more point evaluations than generating functions often improves the stability of the reconstruction with respect to noise on the data points. Our intention is to get rid of this redundancy while maintaining the advantage of oversampling.

The idea of our algorithm is as follows. We assume that we have connected sets $B_j \subset \mathbb{R}^d$, $j = 1, \ldots, L$, such that $\bigcup_{j=1,\ldots,L} B_j = \mathrm{supp}(f)$. For each set B_j, $j = 1, \ldots L$, we cluster the sensors located in B_j. By C_j we denote the set of the indices of the sensors located in B_j, $C_j = \{i : x_i \in B_j\}$. This can be interpreted that sensors in the same cluster are close to each other, and each sensor is contained in at least one of the clusters. In every cluster the noisy point evaluations $d_i = f(x_i) + \delta_i$, $i \in C_j$, of the sensors are transferred to one sensor node, called the cluster head.

Let $\mathbf{d} = \mathbf{h} + \boldsymbol{\delta}$ denote the vector containing the noisy point evaluations of f. Let P_{C_j} denote the orthogonal projection onto the set $\mathrm{span}\{e_i : i \in C_j\}$, where $\{e_i\}_{i=1}^n$ denotes the canonical basis of $\mathbb{C}^n$. This means that in $P_{C_j}\mathbf{d}$ all point evaluations from sensors outside the jth cluster are set to zero. The idea is to send from each region B_j inner products

$$s_i = \langle P_{C_j}\mathbf{d}, \mathbf{v}_i\rangle = \langle \mathbf{d}, P_{C_j}\mathbf{v}_i\rangle = \langle \mathbf{h}+\boldsymbol{\delta}, P_{C_j}\mathbf{v}_i\rangle, \quad i = 1,\dots,r_j, \tag{1.13}$$

instead of the noisy point evaluations $\{d_i\}_{i\in C_j}$. If the number of sensors located in the jth cluster is larger than the number of transmitted numbers $\{s_i\}_{i=1}^{r_j}$, this approach requires fewer long distance transmissions. We observe that in (1.13), s_i are inner products of a perturbation $\mathbf{h}+\boldsymbol{\delta}$ of $\mathbf{h}$ with vectors $P_{C_j}\mathbf{v}_i$. Therefore the sampling vectors after the pre-processing (1.13) are

$$P_{C_j}\mathbf{v}_i, \quad j = 1,\dots,L,\ i = 1,\dots,r_j.$$

The strategy is to calculate an approximation $\tilde{\mathbf{h}}$ of $P_{\mathcal{R}(A)}\mathbf{h}$, the orthogonal projection onto the range of A. We then calculate the coefficients $\hat{c}_k$ for the approximation $\sum_{k=1}^{m}\hat{c}_k g_k$ of f from $\tilde{\mathbf{h}}$ by $\hat{\mathbf{c}} = \arg\min_{\mathbf{c}} \|A\mathbf{c}-\tilde{\mathbf{h}}\|$. Since $\tilde{\mathbf{h}}$ is an approximation to $P_{\mathcal{R}(A)}\mathbf{h}$, it is reasonable to choose $\mathcal{R}(A)$ as the reconstruction space, and the columns of A as reconstruction vectors. We recall that the operator Q_f is optimal when reconstructing from inner products of a perturbed function. Since the values s_i are inner products of a perturbation $\mathbf{h}+\boldsymbol{\delta}$ of $\mathbf{h}$ with vectors $P_{C_j}\mathbf{v}_i$, $i = 1,\dots,r_j$, $j = 1,\dots,L$, we use the operator Q_f to calculate the approximation $\tilde{\mathbf{h}}$ to $P_{\mathcal{R}(A)}\mathbf{h}$. If we use as sampling vectors $\{\mathbf{b}_j\}_{j=1}^{m}$ the columns of A, then this reconstruction algorithm simplifies to solving the normal equations

$$A^*A\hat{\mathbf{c}} = A^*\mathbf{d}. \tag{1.14}$$

The approximation $\tilde{f}$ to f is given by

$$\tilde{f} = \sum_{k=1}^{m}\hat{c}_k g_k.$$

This approach is explained in Section 7.5.1 in detail. With this setup we require a total number of m long distance transmissions, namely $\{\langle \mathbf{d}, \mathbf{b}_i\rangle\}_{i=1}^{m}$, to the fusion center. Therefore, if the signal is oversampled by a factor of

s, i.e. $n = sm$, then instead of n transmissions, this approach only requires $\frac{n}{s}$ transmissions. If we also assume transmission noise when sending the numbers s_i to a global fusion center, then calculating the coefficients for the reconstruction by (1.14) causes stability problems whenever the condition number $\kappa(A)$ is large.

In this case the following approach is a good strategy. In Section 7.5.2, we propose choosing sets $\{B_j\}_{j=1}^{L}$ as a partition of the support of f. From the jth cluster, we transmit the inner products $\{\langle \mathbf{d}, \mathbf{v}_{i,j}\rangle\}_{i=1}^{r_j}$, where the vectors $\{\mathbf{v}_{i,j}\}_{i=1}^{r_j}$ are an orthonormal system for $\mathcal{R}(P_{C_j}A)$. The coefficients $\hat{c}_k$ for the approximation $\sum_{k=1}^{m} \hat{c}_k g_k$ to f are calculated from $\{\langle \mathbf{d}, \mathbf{v}_{i,j}\rangle\}_{i=1}^{r_j}$, $j = 1, \ldots, L$, by solving the least squares problem

$$\hat{\mathbf{c}} = \arg\min_{\mathbf{c}} \|T^* A\mathbf{c} - T^*\mathbf{d}\|. \tag{1.15}$$

The columns of T consist of the orthonormal vectors $\{\mathbf{v}_{i,j}\}_{j=1,\ldots,L,\ i=1,\ldots,r_j}$ stacked together. In Section 7.5.2.1, we give an upper bound on the number of required long distance transmissions. This bound shows that if the generating functions g_k, $k = 1, \ldots, m$, have a compact support and the sets B_j are large enough in comparison to the support of the generating functions, then $r := \sum_{j=1}^{L} r_j$ (the number of required long distance transmissions required by this approach) is roughly m (the number of generating functions). Therefore, in this case, if the function is oversampled by a factor of s, i.e. $n = sm$, the number of required long distance transmissions is reduced by roughly a factor of s.

In Section 7.6, we prove stability of the three algorithms. The first is sending all sensor measurements and solving the global least squares problem (1.12). The second one is sending the inner products of $\mathbf{d}$ with the columns of A and solving the normal equations (1.14). The third one is sending the inner products of $\mathbf{d}$ with an orthonormal system $\{\mathbf{v}_{i,j}\}_{j=1,\ldots,L,\ i=1,\ldots,r_j}$, followed by solving the least squares problem (1.15). We show that from the analytical point of view all three reconstruction algorithms solve a least squares problem

$$\hat{\mathbf{c}} = \arg\min_{\mathbf{c}} \|A\mathbf{c} - (\mathbf{h} + \mathbf{a})\|$$

with different noise vectors $\mathbf{a}$. Using the same upper bound on the signal to noise ratio of the measurement noise and transmission noise, we determine an upper bound on the norm of the noise vector $\mathbf{a}$ for all three reconstruction methods. For the approach using an orthonormal system $\{\mathbf{v}_{i,j}\}_{j=1,\ldots,L,\ i=1,\ldots,r_j}$ we obtain the same upper bound as we obtain by sending all sensor measurements and solving the global least squares problem. Therefore we expect that these two algorithms have similar reconstruction accuracy. For the hierarchical reconstruction method which solves the normal equations (1.14), the condition number $\kappa(A)$ appears as an additional factor in the upper bound on $\|\mathbf{a}\|$. Therefore we expect worse reconstruction accuracy by this strategy, whenever the condition number $\kappa(A)$ is large. In Section 7.7 this is confirmed by numerical experiments, for the special case of integer translates of a basis spline of a certain order as generating functions $g_k,\ k = 1, \ldots, m$.

In Section 7.8 we determine the operation count of our algorithm, which uses an orthonormal system $\{\mathbf{v}_{i,j}\}_{j=1,\ldots,L,\ i=1,\ldots,r_j}$. We show that it has the same order as the complexity of solving the global least squares problem (1.10) (by a direct method such as the QR decomposition), whenever r (the number of required long distance transmissions required by our approach) is roughly m (the number of generating functions). As already mentioned, r is roughly m whenever the local regions $B_j,\ j = 1, \ldots, L$, are large enough in comparison to the support of the generating functions. In this case, if the function is oversampled by a factor of s, we reduce the communication workload roughly by a factor of s and obtain reconstructions of similar accuracy for operation count of the same order.

1.4 List of used symbols

$\mathcal{H}$:	Separable Hilbert space.
$\mathcal{T}, \mathcal{V}$:	Closed subspaces of $\mathcal{H}$.
$P_{\mathcal{V}}$:	Orthogonal projection onto $\mathcal{V}$.
$P_{\mathcal{V}}(\mathcal{T})$:	The set $\{P_{\mathcal{V}}g : g \in \mathcal{T}\}$.
$P_{\mathcal{T},\mathcal{V}}$:	Oblique projection with range $\mathcal{T}$ and kernel $\mathcal{V}$.
$\mathcal{R}(F)$:	The range of the operator F.
$\mathcal{N}(F)$:	The null-space of the operator F.
S :	The frame operator.
$\overline{A}$:	The closure of a set A.
$A^{\perp}$:	The orthogonal complement of a subset A of a Hilbert space.
$\oplus$:	Direct sum (not necessarily orthogonal).

Chapter 2

Frames and the Pseudoinverse

In this section we list several well known theorems concerning functional analysis, frames and the pseudoinverse, which are used throughout the thesis. This chapter can be skipped by an experienced reader. The only Lemma worth having a short look at is Lemma 2.3.18.

2.1 Collection of Theorems from functional analysis

The following lemma can be found in [55, 12.33 Theorem].

Lemma 2.1.1. *Let $\mathcal{H}$ be a Hilbert space. For every positive operator $S : \mathcal{H} \to \mathcal{H}$, there exists a unique positive operator $S^{\frac{1}{2}}$ such that*

$$\left(S^{\frac{1}{2}}\right)^2 = S.$$

If $S > 0$, then $S^{\frac{1}{2}} > 0$.

We call $S^{\frac{1}{2}}$ the *square root* of S.

The following Lemma is a part of [21, Lemma 2.4.1]

Lemma 2.1.2. *Let L and $\mathcal{H}$ be Hilbert spaces. If $A : L \to \mathcal{H}$ is a bounded operator, then*

1. *$\mathcal{R}(A)$ is closed in $\mathcal{H}$ if and only if $\mathcal{R}(A^*)$ is closed in L.*

2. $\|A\| = \|A^*\|$, *and* $\|AA^*\| = \|A\|^2$.

We also make use of the following well known formulas.

Lemma 2.1.3. *Let* L *and* $\mathcal{H}$ *be Hilbert spaces. For an operator* $A : L \to \mathcal{H}$, *the following identities hold,*

$$\mathcal{N}(AA^*) = \mathcal{N}(A^*), \qquad \mathcal{N}(A^*A) = \mathcal{N}(A).$$

If A *has a closed range, then the following identities hold,*

$$\mathcal{R}(AA^*) = \mathcal{R}(A), \qquad \mathcal{R}(A^*A) = \mathcal{R}(A^*).$$

Definition 2.1.4. *Let* X *and* Y *be normed vector spaces. We denote the set of all bounded linear operators from* X *to* Y *by* $\mathcal{L}(X, Y)$. *If* $X = Y$, *then we write* $\mathcal{L}(X)$ *instead of* $\mathcal{L}(X, X)$.

Definition 2.1.5. *Let* $A \in \mathcal{L}(X)$. *The spectrum of* A *is the set of all* $\lambda \in \mathbb{C}$ *for which the operator* $\lambda I - A$ *has no bounded inverse. We denote the spectrum of* A *by* $\sigma(A)$.

We use the continuous functional calculus, which can be found in [53, Thm VII.1].

Theorem 2.1.6. *If* A *is a self-adjoint bounded operator on a Hilbert space* $\mathcal{H}$, *then there exists a unique map* $\varphi : C(\sigma(A)) \to \mathcal{L}(\mathcal{H})$ *with the following properties.*

- φ *is an* $*$*-homomorphism, i.e.*
  $\varphi(fg) = \varphi(f)\varphi(g)$, $\varphi(\lambda f) = \lambda\varphi(f)$, $\varphi(1) = I$, $\varphi(\overline{f}) = \varphi(f)^*$,
- $\|\varphi(f)\| = \|f\|_\infty$.
- *If* f *is defined by* $f(x) = x$, *then* $\varphi(f) = A$.

Moreover, φ *has the additional properties:*

- *If* $A\psi = \lambda\psi$, *then* $\varphi(f)\psi = f(\lambda)\psi$,

- $\sigma[\varphi(f)] = \{f(\lambda) : \lambda \in \sigma(A)\} = f(\sigma(A))$,

- *If* $f \geqslant 0$, *then* $\varphi(f) \geqslant 0$,

The following Lemma can be found in [50, Theorem 5.17.2]

Lemma 2.1.7. *Let* $\mathcal{W}$ *and* $\mathcal{H}$ *be Hilbert spaces, and let* $V : \mathcal{H} \to \mathcal{W}$ *be a bounded operator. There exists an* $A > 0$, *such that*

$$A\|c\| \leqslant \|Vc\| \text{ for every } c \in \mathcal{N}(V)^{\perp}, \tag{2.1}$$

if and only if the operator V *has a closed range.*

2.2 Properties of the pseudoinverse

We need the definition of the pseudoinverse of an operator on Hilbert spaces.

Lemma 2.2.1. *Let* $\mathcal{H}$ *and* $\mathcal{W}$ *be Hilbert spaces. If* $A : \mathcal{W} \to \mathcal{H}$ *is a bounded operator with a closed range* $\mathcal{R}(A)$, *then there exists a unique bounded operator* $A^{\dagger} : \mathcal{H} \to \mathcal{W}$ *such that*

$$\mathcal{N}(A^{\dagger}) = \mathcal{R}(A)^{\perp} = \mathcal{N}(A^{*}), \tag{2.2}$$

$$\mathcal{R}(A^{\dagger}) = \mathcal{N}(A)^{\perp} = \mathcal{R}(A^{*}), \text{ and} \tag{2.3}$$

$$AA^{\dagger}x = x, \ x \in \mathcal{R}(A). \tag{2.4}$$

Definition 2.2.2. *Let* $\mathcal{H}$ *and* $\mathcal{W}$ *be Hilbert spaces and let* $A : \mathcal{W} \to \mathcal{H}$ *be a bounded operator with a closed range. We call the operator* $A^{\dagger}$ *satisfying* (2.2), (2.3) *and* (2.4) *the pseudoinverse of* A.

The following lemma is a part of [21, Lemma 2.5.2].

Lemma 2.2.3. *Let* $A : \mathcal{W} \to \mathcal{H}$ *be a bounded operator. If* A *has a closed range* $\mathcal{R}(A)$, *then the following holds:*

1. *The orthogonal projection of* $\mathcal{H}$ *onto* $\mathcal{R}(A)$ *is given by* $AA^{\dagger}$.

2. *The orthogonal projection of* $\mathcal{W}$ *onto* $\mathcal{R}(A^{\dagger})$ *is given by* $A^{\dagger}A$.

3. $(A^*)^\dagger = (A^\dagger)^*$.

The following properties are an easy consequence of Lemma 2.2.3, and often used to define the pseudoinverse.

Corollary 2.2.4. *Let $A : \mathcal{W} \to \mathcal{H}$ be a bounded operator. If A has a closed range $\mathcal{R}(A)$, then the following holds:*

$$AA^\dagger A = A, \tag{2.5}$$

$$A^\dagger AA^\dagger = A^\dagger. \tag{2.6}$$

The following lemma can be found in [15] as an exercise.

Lemma 2.2.5. *Let $A : \mathcal{W} \to \mathcal{H}$ be a bounded operator. If A has a closed range $\mathcal{R}(A)$, then the following holds:*

$$A^\dagger = (A^*A)^\dagger A^*$$

$$A^\dagger = A^*(AA^*)^\dagger.$$

Lemma 2.2.6 can be proven using the spectral theorem.

Lemma 2.2.6. *Let A be a bounded self-adjoint operator on $\mathcal{H}$ with a closed range. If $A \geqslant 0$ also $A^\dagger \geqslant 0$.*

The following Lemma can be found in [52].

Lemma 2.2.7. *Let $\mathcal{H}$ and $\mathcal{W}$ be Hilbert spaces and let $A : \mathcal{W} \to \mathcal{H}$ be a bounded operator with a closed range $\mathcal{R}(A)$, and let $b \in \mathcal{H}$. We define the set B by*

$$B := \arg\min_{x \in \mathcal{W}} \|Ax - b\|.$$

Then $x := A^\dagger b$ is the unique element of B of minimal norm.

2.3 Frames in Hilbert Spaces

Definition 2.3.1. *Let* $\{f_k\}_{k\in\mathbb{N}}$ *be a sequence in* $\mathcal{H}$. *If there exists a constant* $B > 0$, *such that*

$$\sum_{j\in\mathbb{N}} |\langle f, f_k\rangle|^2 \leqslant B\|f\|^2, \quad \text{for every } f \in \mathcal{H}, \tag{2.7}$$

then we call $\{f_k\}_{k\in\mathbb{N}}$ *a Bessel sequence. Any constant* $B > 0$ *satisfying* (2.7) *is called Bessel bound for* $\{f_k\}_{k\in\mathbb{N}}$.

Lemma 2.3.2. *If* $\{f_k\}_{k\in\mathbb{N}}$ *is a Bessel sequence in* $\mathcal{H}$, *then* $\sum_{k\in\mathbb{N}} c_k f_k$ *converges unconditionally for every* $\{c_k\}_{k\in\mathbb{N}} \in l^2(\mathbb{N})$.

Definition 2.3.3. *Let* $\{f_k\}_{k\in\mathbb{N}}$ *be a sequence in* $\mathcal{H}$. *We say that* $\{f_k\}_{k\in\mathbb{N}}$ *is a frame for* $\mathcal{H}$ *if there exist constants* $A, B > 0$ *such that*

$$A\|f\|^2 \leqslant \sum_{j\in\mathbb{N}} |\langle f, f_k\rangle|^2 \leqslant B\|f\|^2, \quad \text{for every } f \in \mathcal{H}. \tag{2.8}$$

The constant A *is called a lower frame bound, and the constant* B *is called an upper frame bound.*

Definition 2.3.4. *Let* $\{f_k\}_{k\in\mathbb{N}}$ *be a sequence in* $\mathcal{H}$. *We say that* $\{f_k\}_{k\in\mathbb{N}}$ *is a tight frame for* $\mathcal{H}$, *if* $A = B$ *in* (2.8).

Definition 2.3.5. *Let* $\{u_j\}_{j\in\mathbb{N}}$ *be a frame in* $\mathcal{H}$. *The operator*

$$V : l^2(\mathbb{N}) \to \mathcal{H}, \quad V\{c_j\}_{j\in\mathbb{N}} = \sum_{j=1}^{\infty} c_j u_j$$

is called the synthesis operator. The adjoint operator

$$V^* : \mathcal{H} \to l^2(\mathbb{N}), \quad V^* f = \{\langle f, u_j\rangle\}_{j\in\mathbb{N}}$$

is called the analysis operator. The composition

$$S : \mathcal{H} \to \mathcal{H}, \quad Sf = VV^* f = \sum_{j=1}^{\infty} \langle f, u_j\rangle u_j$$

is called the frame operator.

Theorem 2.3.6, which can be found in [21, Theorem 5.1.7], is an important result in frame theory. It states that every $f \in \mathcal{H}$ is an infinite linear combination of the frame elements.

Theorem 2.3.6. *If $\{f_k\}_{k\in\mathbb{N}}$ is a frame, then*

$$f = \sum_{k\in\mathbb{N}} \langle f, S^{-1} f_k\rangle f_k, \quad \text{for every } f \in \mathcal{H},$$

and

$$f = \sum_{k\in\mathbb{N}} \langle f, f_k\rangle S^{-1} f_k, \quad \text{for every } f \in \mathcal{H}. \tag{2.9}$$

Both series converge unconditionally for all $f \in \mathcal{H}$.

Theorem 2.3.7. *If $\{f_k\}_{k\in\mathbb{N}}$ is a tight frame with frame bound A, then*

$$f = \frac{1}{A} \sum_{k\in\mathbb{N}} \langle f, f_k\rangle f_k, \quad \text{for every } f \in \mathcal{H}.$$

The following theorem can be found in [21, Theorem 5.3.4]

Theorem 2.3.8. *If $\{f_k\}_{k\in\mathbb{N}}$ is a frame for $\mathcal{H}$ with frame operator S, then $\{S^{-\frac{1}{2}} f_k\}_{k\in\mathbb{N}}$ is a tight frame with frame bound equal to 1. Every $f \in \mathcal{H}$ can be written in the from*

$$f = S^{-\frac{1}{2}} S S^{-\frac{1}{2}} f = \sum_{k=1}^{\infty} \langle f, S^{-\frac{1}{2}} f_k\rangle S^{-\frac{1}{2}} f_k.$$

The following Lemma can be found in [21, Lemma 5.7.1]

Lemma 2.3.9. *Let $\{f_k\}_{k\in\mathbb{N}}$ and $\{g_k\}_{k\in\mathbb{N}}$ be Bessel sequences in $\mathcal{H}$. The following are equivalent*

1. $f = \sum_{k=1}^{\infty} \langle f, g_k\rangle f_k, \quad$ *for every $f \in \mathcal{H}$.*
2. $f = \sum_{k=1}^{\infty} \langle f, f_k\rangle g_k, \quad$ *for every $f \in \mathcal{H}$.*
3. $\langle f, g\rangle = \sum_{k=1}^{\infty} \langle f, f_k\rangle \langle g_k, g\rangle, \quad$ *for every $f, g \in \mathcal{H}$.*

If B denotes an upper frame bound for $\{f_k\}_{k\in\mathbb{N}}$, then B^{-1} is a lower frame bound for $\{g_k\}_{k\in\mathbb{N}}$.

Definition 2.3.10. *If one of the three equivalent conditions of Lemma 2.3.9 is satisfied, then $\{f_k\}_{k\in\mathbb{N}}$ and $\{g_k\}_{k\in\mathbb{N}}$ are called dual frames for $\mathcal{H}$.*

We observe that Theorem 2.3.6 shows that $\{f_k\}_{k\in\mathbb{N}}$ and $\{S^{-1}f_k\}_{k\in\mathbb{N}}$ are dual frames. The frame $\{S^{-1}f_k\}_{k\in\mathbb{N}}$ is called the canonical dual frame of $\{f_k\}_{k\in\mathbb{N}}$.

The following Lemma can be found in [21, Lemma 5.3.6]

Lemma 2.3.11. *Let $\{f_k\}_{k\in\mathbb{N}}$ be a frame for $\mathcal{H}$ and let $f \in \mathcal{H}$. If f has a representation $f = \sum_{k\in\mathbb{N}} c_k f_k$ for some coefficients $\{c_k\}_{k\in\mathbb{N}}$, then*

$$\sum_{k\in\mathbb{N}} |c_k|^2 = \sum_{k\in\mathbb{N}} |\langle f, S^{-1}f_k\rangle|^2 + \sum_{k\in\mathbb{N}} |c_k - \langle f, S^{-1}f_k\rangle|^2.$$

In our setup, sampling and reconstruction are done in distinct subspaces of a separable Hilbert space $\mathcal{H}$. We work with sequences which are frames for their closed linear spans, the so called frame sequences.

Definition 2.3.12. *Let $\{u_j\}_{j\in\mathbb{N}}$ be a sequence in a separable Hilbert space $\mathcal{H}$. We say that $\{u_j\}_{j\in\mathbb{N}}$ is a frame sequence, if there exist constants $A, B > 0$, such that for every $f \in \mathcal{V} := \overline{span}\{u_j\}_{j\in\mathbb{N}}$*

$$A\|f\|^2 \leqslant \sum_{j\in\mathbb{N}} |\langle f, u_j\rangle|^2 \leqslant B\|f\|^2. \tag{2.10}$$

Equivalently, $\{u_j\}_{j\in\mathbb{N}}$ is a frame for $\mathcal{V}$.

The following lemma can be found in [21, Lemma 5.4.5]

Lemma 2.3.13. *Let $\{u_j\}_{j\in\mathbb{N}}$ be a sequence in $\mathcal{H}$. $\{u_j\}_{j\in\mathbb{N}}$ is a frame sequence with frame bounds A and B if and only if the synthesis operator V is well defined on $l^2(\mathbb{N})$ and*

$$A\|c\|^2 \leqslant \|Vc\|^2 \leqslant B\|c\|^2 \quad \text{for all } c \in \mathcal{N}(V)^\perp \tag{2.11}$$

For a frame sequence $\{u_j\}_{j\in\mathbb{N}}$, the frame operator S is invertible on $\mathcal{V}$, its closed linear span, but not necessarily on the whole space $\mathcal{H}$. To state a counterpart of Theorem 2.3.6 for frame sequences, we need to use the pseudoinverse of S.

The following Theorem is stated in [37, Section 1] and proven similarly to Theorem 2.3.6.

Theorem 2.3.14. *If $\{u_j\}_{j\in\mathbb{N}}$ is a frame sequence for $\mathcal{V} := \overline{span}\{u_j\}_{j\in\mathbb{N}}$, then*

$$P_{\mathcal{V}} f = \sum_{j\in\mathbb{N}} \langle f, S^{\dagger} u_j \rangle u_j, \quad \text{for every } f \in \mathcal{H}, \tag{2.12}$$

and

$$P_{\mathcal{V}} f = \sum_{j\in\mathbb{N}} \langle f, u_j \rangle S^{\dagger} u_j, \quad \text{for every } f \in \mathcal{H}. \tag{2.13}$$

Both series converge unconditionally for all $f \in \mathcal{H}$.

Theorem 2.3.14 shows that given the measurements $\{\langle f, u_j \rangle\}_{j\in\mathbb{N}}$ of f with a frame sequence $\{u_j\}_{j\in\mathbb{N}}$, it is possible to calculate the orthogonal projection onto $\mathcal{V} := \overline{span}\{u_j\}_{j\in\mathbb{N}}$.

The following Theorem is the counterpart of Theorem 2.3.7 for frame sequences and can be proven almost the same way.

Theorem 2.3.15. *If $\{u_j\}_{j\in\mathbb{N}}$ is a tight frame sequence for $\mathcal{V}$ with frame bound A, then*

$$P_{\mathcal{V}} f = \frac{1}{A} \sum_{j\in\mathbb{N}} \langle f, u_j \rangle u_j, \quad \text{for every } f \in \mathcal{H}.$$

Lemma 2.3.16. *Let $\{u_j\}_{j\in\mathbb{N}}$ be a frame sequence in $\mathcal{H}$ and V be the corresponding synthesis operator. The operator $(VV^*)^{\dagger}$ and the operator $(V^*V)^{\dagger}$ are well defined and bounded, $(VV^*)^{\dagger} \geqslant 0$ and $(V^*V)^{\dagger} \geqslant 0$.*

Proof. From the upper frame bound of $\{u_j\}_{j\in\mathbb{N}}$ we infer that the positive operators VV^* and V^*V are bounded, and from the lower frame bound of $\{u_j\}_{j\in\mathbb{N}}$ and Lemma 2.1.7 we infer that they have a closed range. Therefore by

Lemma 2.2.6 the operator $(VV^*)^\dagger$ and the operator $(V^*V)^\dagger$ are well defined and bounded, $(VV^*)^\dagger \geqslant 0$ and $(V^*V)^\dagger \geqslant 0$. $\square$

Lemma 2.3.16 in combination with Lemma 2.1.1 shows that for the operator $A = (VV^*)^\dagger$ and $A = (V^*V)^\dagger$ the unique positive square root $A^{\frac{1}{2}}$ exists. From now on, we use for the square root of the pseudoinverse of an operator A the notation

$$A^{\frac{\dagger}{2}} := (A^\dagger)^{\frac{1}{2}}$$

We need a slightly modified version of Theorem 2.3.8 for frame sequences.

Lemma 2.3.17. *Let $\mathcal{V}$ be a closed subspace of $\mathcal{H}$, let $\{u_j\}_{j\in\mathbb{N}}$ be a frame for $\mathcal{V}$. By V we denote the corresponding analysis operator and by $S = VV^*$ the corresponding frame operator.*

The set

$$\{S^{\frac{\dagger}{2}} u_j\}_{j\in\mathbb{N}}$$

forms a tight frame for $\mathcal{V}$ with frame bound equal to 1. The synthesis operator of the tight frame sequence $\{S^{\frac{\dagger}{2}} u_j\}_{j\in\mathbb{N}}$ is given by

$$M := S^{\frac{\dagger}{2}} V,$$

and

$$P_{\mathcal{V}} = MM^* = S^{\frac{\dagger}{2}} S S^{\frac{\dagger}{2}} = S^\dagger S = SS^\dagger. \tag{2.14}$$

Lemma 2.3.18 proves the following. Suppose that we are given the inner products $\{\langle f, u_j\rangle\}_{j\in\mathbb{N}}$ of an element $f \in \mathcal{H}$ with a frame $\{u_j\}_{j\in\mathbb{N}}$ for $\mathcal{V}$ (a closed subspace of $\mathcal{H}$). Applying the operator $(V^*V)^{\frac{\dagger}{2}}$ to these measurements, we obtain the inner products of f with the tight frame $\{S^{\frac{\dagger}{2}} u_j\}_{j\in\mathbb{N}}$ for $\mathcal{V}$.

Lemma 2.3.18. *Let $\mathcal{V}$ be a closed subspace of $\mathcal{H}$. If $\{u_j\}_{j\in\mathbb{N}}$ is a frame for $\mathcal{V}$, V is the corresponding synthesis operator, V^* is the corresponding analysis operator, and S is the corresponding frame operator, then*

$$(V^*V)^{\frac{\dagger}{2}} V^* = V^*(VV^*)^{\frac{\dagger}{2}}. \tag{2.15}$$

Equivalently, the operator $(V^*V)^{\frac{\dagger}{2}} V^*$ is the analysis operator of the frame sequence $\{S^{\frac{\dagger}{2}} u_j\}_{j\in\mathbb{N}}$.

Proof. Obviously for $k \in \mathbb{N}$

$$(V^*V)^k V^* = V^*(VV^*)^k. \tag{2.16}$$

Therefore

$$p(V^*V)V^* = V^*p(VV^*)$$

for every polynomial p. We are going to prove that there exists a sequence of polynomials $\{p_k\}_{k\in\mathbb{N}}$, such that for $i = 1, 2$

$$\lim_{m\to\infty} \|p_m(A_i) - A_i^{\frac{\dagger}{2}}\| = 0$$

simultaneously for $A_1 := V^*V$ and $A_2 := VV^*$.

From the lower bound C_1 of the frame sequence we infer that for every $u \in \mathcal{V} = \mathcal{N}(VV^*)^\perp$

$$C_1\|u\|^2 \leqslant \langle VV^*u, u\rangle.$$

Consequently the set $\sigma(VV^*)\backslash\{0\}$ is bounded below by C_1. The upper frame bound C_2 ensures that the set $\sigma(VV^*)$ has the upper bound C_2. This shows that 0 is an isolated point of the spectrum, and that for $K := \{0\} \cup [C_1, C_2]$ the function $g : K \to \mathbb{R}$

$$g(x) = \begin{cases} \frac{1}{\sqrt{x}} & \text{for } x \in [C_1, C_2], \\ 0 & \text{for } x = 0 \end{cases}$$

is continuous on K. Since $\sigma(VV^*)\cup\{0\} = \sigma(V^*V)\cup\{0\}$, f is also continuous on $\sigma(V^*V)$.

By the Weierstrass approximation theorem there exists a sequence of polynomials $\{p_m\}_{m\in\mathbb{N}}$, such that

$$\lim_{m\to\infty} \|p_m - g\|_\infty = 0,$$

uniformly on K. By the continuous functional calculus (Theorem 2.1.6)

$$\lim_{m\to\infty} \|p_m(A_i) - g(A_i)\| = 0$$

simultaneously for $A_1 := V^*V$ and $A_2 := VV^*$.

It remains to prove that $g(A_i) = A_i^{\frac{\dagger}{2}}$ for $i = 1, 2$. We show that for $S = VV^*$ it holds $g(S) = S^{\frac{\dagger}{2}}$. By Lemma 2.2.1 we have to prove that $\mathcal{N}(g(S)) = \mathcal{R}(S)^\perp$, $\mathcal{R}(g(S)) = \mathcal{N}(S)^\perp$, $S^{\frac{1}{2}} g(S)x = x$, $x \in \mathcal{R}(S)$. We show that by restricting the operator $g(S)$ to the set $\mathcal{R}(S) = \mathcal{V}$ and the set $\mathcal{R}(S)^\perp = \mathcal{V}^\perp$.

On the set $\mathcal{V}$, the operator $S : \mathcal{V} \to \mathcal{V}$ is invertible, and $\sigma(S_{|\mathcal{V}}) \subset [C_1, C_2]$. Let the function f be defined by $f(x) = \sqrt{x}$. By the continuous functional calculus and the definition of g, we obtain for $u \in \mathcal{V}$

$$u = \varphi(1)u = \varphi(fg)u = \varphi(f)\varphi(g)u = \sqrt{S_{|\mathcal{V}}}\; g(S_{|\mathcal{V}})u.$$

This shows that on the set $\mathcal{V}$ it holds $g(S) = \sqrt{S}^{\,-1}$. Consequently for $x \in \mathcal{R}(S)$, $S^{\frac{1}{2}} g(S)x = x$, and $\mathcal{V} \subset \mathcal{R}(g(S))$.

On the set $\mathcal{V}^\perp$ it holds $S = 0$, and consequently $\sigma(S_{|\mathcal{V}^\perp}) = 0$. By the continuous functional calculus, on this set $g(S) = 0$. Consequently $\mathcal{V} = \mathcal{R}(g(S))$ and $\mathcal{N}(g(S)) = \mathcal{V}^\perp$. This proves that $g(S) = S^{\frac{\dagger}{2}}$. That $g(V^*V) = (V^*V)^{\frac{\dagger}{2}}$ is proven the same way. $\square$

Chapter 3

Stability, quasi optimality and error estimates

3.1 Problem formulation and notation

We now describe the reconstruction problem we treat in detail. Let $\mathcal{H}$ be a separable Hilbert space over $\mathbb{C}$. We assume that we are given the measurements

$$d_j = \langle f, u_j \rangle + \delta_j \quad j \in \mathbb{N}, \tag{3.1}$$

of an unknown function $f \in \mathcal{H}$, where $\{u_j\}_{j\in\mathbb{N}}$ is a fixed frame sequence. We call $\{u_j\}_{j\in\mathbb{N}}$ the sampling frame and

$$\mathcal{V} := \overline{\operatorname{span}}\{u_j\}_{j\in\mathbb{N}},$$

the sampling space.

We approximate the function f by a series expansion

$$\tilde{f} = \sum_{k\in\mathbb{N}} c_k g_k.$$

We call $\{g_k\}_{k\in\mathbb{N}}$ the reconstruction frame sequence and

$$\mathcal{T} := \overline{\operatorname{span}}\{g_k\}_{k\in\mathbb{N}},$$

the *reconstruction space*. Of course, for different classes of functions different basis functions $\{g_k\}$ are needed, and a priori knowledge of the function to reconstruct should be taken into account. Since the orthogonal projection $P_{\mathcal{T}}f$ is the point in $\mathcal{T}$ closest possible to f, the sequence $\{g_k\}_{k\in\mathbb{N}}$ should be such that

$$\rho(f) := \frac{\|f - P_{\mathcal{T}}f\|}{\|f\|} \tag{3.2}$$

is small. In this case we say f is well presented by the subspace $\mathcal{T}$.

Let $d \in l^2(\mathbb{N})$ denote the vector consisting of the measurements d_j, defined by (3.1). Our objective is to find a linear mapping

$$Q : l^2(\mathbb{N}) \to \mathcal{T},$$

such that

$$\tilde{f} := Qd$$

is a good approximation to f for every $f \in \mathcal{H}$ that is well presented by the reconstruction space $\mathcal{T}$.

By being $\tilde{f}$ a good approximation to f we mean that the relative error

$$\nu(f) := \frac{\|\tilde{f} - f\|}{\|f\|}$$

is small (depending on the relative measurement error).

We use the notation V for the synthesis operator, V^* for the analysis operator and $S = VV^*$ for the frame operator of the sampling frame sequence $\{u_j\}_{j\in\mathbb{N}}$. With this notation, $\{u_j\}_{j\in\mathbb{N}}$ being a frame for $\mathcal{V}$ is equivalent to the statement that there exist constants $A, B > 0$, such that

$$A\|f\|^2 \leqslant \|V^*f\|^2 \leqslant B\|f\|^2, \quad \text{for every } f \in \mathcal{V}. \tag{3.3}$$

We use the notation T for the synthesis operator and T^* for the analysis operator of the reconstruction frame sequence $\{g_k\}_{k\in\mathbb{N}}$, and C, D for the frame bounds. Consequently

$$C\|g\|^2 \leqslant \|T^*g\|^2 \leqslant D\|g\|^2, \quad \text{for every } g \in \mathcal{T}. \tag{3.4}$$

In general, to calculate $P_{\mathcal{T}}f$ from the vector V^*f containing the measurements $\{\langle f, u_j\rangle\}_{j\in\mathbb{N}}$, the information is not sufficient (to be able to do so, $\mathcal{T}$ has to be a subset of $\mathcal{V}$). Even if possible, in some cases calculating the orthogonal projection onto the reconstruction space $\mathcal{T}$ is numerically unstable. In other words, in general, there need not exist an operator Q, such that $QV^* = P_{\mathcal{T}}$. Certain oblique projections onto $\mathcal{T}$ can be calculated under much weaker assumptions. We discuss two such oblique projections. The first one is the oblique projection with range $\mathcal{T}$ and null-space $P_{\mathcal{V}}(\mathcal{T})^{\perp}$, denoted by $P_{\mathcal{T},P_{\mathcal{V}}(\mathcal{T})^{\perp}}$. The second one is $P_{\mathcal{T},S(\mathcal{T})^{\perp}}$, which was introduced in [6, 7, 10].

3.2 Stability, quasi optimality

We use two quantities to measure the quality of a reconstruction operator

$$Q : l^2(\mathbb{N}) \to \mathcal{H}.$$

In order to measure how well Q deals with the part of the function to reconstruct lying outside of the reconstruction space, following [10], we introduce the *quasi-optimality* constant.

Definition 3.2.1. *Let $Q : l^2(\mathbb{N}) \to \mathcal{H}$ be an operator. The quasi-optimality constant $\mu = \mu(Q) > 0$ is the smallest number μ, such that*

$$\|f - QV^*f\| \leqslant \mu\|f - P_{\mathcal{T}}f\|, \quad \text{for all } f \in \mathcal{H}. \tag{3.5}$$

If there does not exist a $\mu \in \mathbb{R}$ such that (3.5) is fulfilled, we set $\mu = \infty$.

If $\mu(Q) < \infty$, Q is called a quasi-optimal reconstruction operator.

We note that $P_{\mathcal{T}}f$ is the element of $\mathcal{T}$ closest to f. Thus the *quasi-optimality* constant μ is a measure of how well QV^* performs in comparison to $P_{\mathcal{T}}$.

As a measure of stability of the reconstruction, we use the operator norm $\|Q\|$ of Q. In [10], instead of the operator norm of Q, the absolute condition number [10, Definition 2.2] is used. Here we focus on linear reconstructions, in which case this quantity equals the operator norm $\|Q_{|\mathcal{R}(V^*)}\|$ of Q restricted to the subspace $\mathcal{R}(V^*)$.

If no additional information about the noise is given, it is reasonable to expect that Q annihilates everything in $\mathcal{R}(V^*)^\perp$, since the measurements (3.1) are assumed to form a noisy version of an element in $\mathcal{R}(V^*)$. This means that for $c \in l^2(\mathbb{N})$,

$$Qc = Q(P_{\mathcal{R}(V^*)}c + P_{\mathcal{R}(V^*)^\perp}c) = QP_{\mathcal{R}(V^*)}c.$$

In this case the following Lemma holds.

Lemma 3.2.2. *Let $Q : l^2(\mathbb{N}) \to \mathcal{H}$ be a bounded operator. If $Qc = 0$ for $c \in \mathcal{R}(V^*)^\perp$, then*

$$\|Q\| = \|Q_{|\mathcal{R}(V^*)}\|.$$

3.3 Absolute and relative error

For a quasi-optimal and bounded reconstruction operator we derive the following error estimates.

Lemma 3.3.1. *Let $Q : l^2(\mathbb{N}) \to \mathcal{H}$ be a bounded and quasi-optimal operator. For $f \in \mathcal{H}$ and $c \in l^2(\mathbb{N})$ it holds*

$$\|f - Q(V^*f + c)\| \leqslant \mu(Q)\|f - P_{\mathcal{T}}f\| + \|Q\|\|c\|. \tag{3.6}$$

Proof. Follows from Definition 3.2.1 and the triangle-inequality. □

Equation (3.6) bounds the absolute error of our reconstruction. We observe that $\|f - P_{\mathcal{T}}f\|$ is the norm of the part of f lying outside of the reconstruction space. This term is multiplied by the quasi-optimality constant. Therefore the quasi-optimality constant is a measure of stability of the reconstruction with respect to the part lying outside of the reconstruction space.

The norm of the measurement error c is multiplied by the operator norm of Q. This is the reason why we use the operator norm of Q as a measure of stability with respect to error present in the measurements.

The following Theorem gives an estimate for the relative error.

Lemma 3.3.2. *Let $Q : l^2(\mathbb{N}) \to \mathcal{H}$ be a bounded and quasi-optimal operator and let B denote the upper frame bound of the sampling frame. For $f \in \mathcal{H}$ and $c \in l^2(\mathbb{N})$*

$$\frac{\|f - Q(V^*f + c)\|}{\|f\|} \leqslant \mu \frac{\|f - P_{\mathcal{T}} f\|}{\|f\|} + \|Q\|\sqrt{B}\frac{\|c\|}{\|V^*f\|}. \tag{3.7}$$

Proof. From the frame equation (3.3) it follows that

$$\|f\| \geqslant \frac{1}{\sqrt{B}}\|V^*f\|. \tag{3.8}$$

Combining (3.6) and (3.8), we obtain (3.7). □

In Section 5.2 we discuss a reconstruction operator Q_g with the smallest possible operator norm in detail, which was introduced in [6, 7, 10]. The authors called this reconstruction generalized sampling. In Section 5.3 we discuss a reconstruction operator Q_f with the smallest possible quasi-optimality constant. We call this reconstruction frame independent sampling.

3.4 Oblique projections and subspace angles

3.4.1 Subspace angles

In Theorem 4.0.12 we show that any quasi-optimal and bounded operator $Q : l^2(\mathbb{N}) \to \mathcal{T}$ has the property that the operator $P = QV^*$ is an oblique projection onto the reconstruction space $\mathcal{T}$.

In order to be able to analyze these oblique projections, we need the concept of subspace angles. There are many different definitions of the angle between subspaces. The paper [61] is a good collection of different concepts of angles and points out the connections between them. See also [59].

Definition 3.4.1. *Let $\mathcal{T}$ and $\mathcal{V}$ be nonzero closed subspaces of a Hilbert space $\mathcal{H}$. We define the subspace angle $\varphi = \varphi_{\mathcal{T},\mathcal{V}} \in [0, \frac{\pi}{2}]$ between $\mathcal{T}$ and $\mathcal{V}$ by*

$$\cos(\varphi_{\mathcal{T},\mathcal{V}}) = \inf_{\substack{g\in\mathcal{T} \\ \|g\|=1}} \|P_{\mathcal{V}} g\| = \inf_{\substack{g\in\mathcal{T} \\ \|g\|=1}} \sup_{\substack{u\in\mathcal{V} \\ \|u\|=1}} |\langle g, u\rangle|. \tag{3.9}$$

It is important to realize that in general $\cos(\varphi_{\mathcal{T},\mathcal{V}}) \neq \cos(\varphi_{\mathcal{V},\mathcal{T}})$. For example let $\mathcal{H} = \mathbb{R}^3$, let $\mathcal{V}$ be the $x\,y$-plane and let $\mathcal{T}$ be the line spaned by the vector $[1, 0, 1]^T$. In this case $\varphi_{\mathcal{T},\mathcal{V}}$ is the angle between the two vectors $[1, 0, 1]^T$ and $[1, 0, 0]^T$ and $\cos(\varphi_{\mathcal{T},\mathcal{V}}) = \frac{1}{\sqrt{2}}$ is the cosine of the angle that is intuitively thought of. We observe that the vector $[0, 1, 0]^T \in \mathcal{V}$ and $[0, 1, 0]^T \perp \mathcal{T}$, and consequently $\cos(\varphi_{\mathcal{V},\mathcal{T}}) = 0$.

Most of our theorems use the assumption that the cosine of the angle between the reconstruction and sampling space fulfills $\cos(\varphi_{\mathcal{T},\mathcal{V}}) > 0$. The following theorem shows that this assumption is equivalent to the statement that the orthogonal projection of the sampling vectors onto the reconstruction space $\{P_{\mathcal{T}} u_j\}_{j\in\mathbb{N}}$ is a frame for $\mathcal{T}$.

Lemma 3.4.2. *Let $\mathcal{T}$ and $\mathcal{V}$ be closed subspaces of $\mathcal{H}$, and $\{u_j\}_{j\in\mathbb{N}}$ be a frame for $\mathcal{V}$. Let V denote the synthesis operator of the frame sequence $\{u_j\}_{j\in\mathbb{N}}$ and let V^* denote the corresponding analysis operator.*

The following are equivalent.

1. $\cos(\varphi_{\mathcal{T},\mathcal{V}}) > 0$.

2. $\{P_{\mathcal{T}} u_j\}_{j\in\mathbb{N}}$ *is a frame for* $\mathcal{T}$.

Proof. Let $\cos(\varphi_{\mathcal{T},\mathcal{V}}) > 0$. Let A and B denote the lower and upper frame bounds of $\{u_j\}_{j\in\mathbb{N}}$, respectively. From the assumption $\cos(\varphi_{\mathcal{T},\mathcal{V}}) > 0$ it follows that

$$\|g\| \cos(\varphi_{\mathcal{T},\mathcal{V}}) \leqslant \|P_{\mathcal{V}} g\| \quad \text{for all } g \in \mathcal{T}. \tag{3.10}$$

In particular, for $g \in \mathcal{T}$ we obtain with (3.10)

$$A\|g\|^2 \cos^2(\varphi_{\mathcal{T},\mathcal{V}}) \leqslant A\|P_{\mathcal{V}} g\|^2 \leqslant \|V^* g\|^2 = \sum_{j\in\mathbb{N}} |\langle g, P_{\mathcal{T}} u_j\rangle|^2 \leqslant B\|g\|^2. \tag{3.11}$$

This shows that $\{P_{\mathcal{T}} u_j\}_{j\in\mathbb{N}}$ is a frame for $\mathcal{T}$ with upper frame bound B and lower frame bound $A\cos^2(\varphi_{\mathcal{T},\mathcal{V}})$.

Let $\{P_{\mathcal{T}} u_j\}_{j\in\mathbb{N}}$ be a frame for $\mathcal{T}$ with lower frame bound C_1. Since $\langle g, P_{\mathcal{T}} u_j\rangle = \langle P_{\mathcal{V}} g, u_j\rangle$ for $g \in \mathcal{T}$, we obtain

$$C_1\|g\|^2 \leqslant \sum_{j\in\mathbb{N}} |\langle g, P_{\mathcal{T}} u_j\rangle|^2 = \sum_{j\in\mathbb{N}} |\langle P_{\mathcal{V}} g, u_j\rangle|^2 \leqslant B\|P_{\mathcal{V}} g\|^2.$$

This implies that

$$\cos(\varphi_{\mathcal{T},\mathcal{V}}) = \inf_{\substack{g\in\mathcal{T}\\ \|g\|=1}} \|P_{\mathcal{V}} g\| \geqslant \frac{\sqrt{C_1}}{\sqrt{B}} > 0.$$

□

If reconstruction space $\mathcal{T}$ is the whole Hilbert space $\mathcal{H}$, and $\{u_j\}_{j\in\mathbb{N}} \subset \mathcal{H}$ is a frame sequence with the closed linear span $\mathcal{V}$, then $\cos(\varphi_{\mathcal{T},\mathcal{V}}) > 0$ if, and only if $\{u_j\}_{j\in\mathbb{N}}$ is a frame for $\mathcal{H}$. For example, when reconstructing a bandlimited function $f \in L^2([-\frac{1}{2}, \frac{1}{2}])$ from point samples

$$f(x_j) = \langle \mathcal{F} f, e^{-2\pi i x_j \cdot}\rangle, \quad j \in \mathbb{N},$$

one needs to prove that $\{e^{2\pi i x_j \cdot}\}_{j\in\mathbb{N}}$ is a frame for $L^2([-\frac{1}{2}, \frac{1}{2}])$, in order to guarantee that a stable reconstruction from point evaluations $\{f(x_j)\}_{j\in\mathbb{N}}$ is possible. As mentioned in the introduction, in this case the set $\{x_j\}_{j\in\mathbb{N}}$ is called a set of stable sampling.

If $\{e^{2\pi i x_j \cdot}\}_{j\in\mathbb{N}}$ is a frame for $\mathcal{V} = L^2([-1, 1])$, and we use $\mathcal{T} = L^2([-\frac{1}{2}, \frac{1}{2}])$ as reconstruction space, we are in the case of sampling and reconstruction in distinct subspaces where $\mathcal{T} \subset \mathcal{V} = \mathcal{H}$. In this paragraph, functions in $L^2([-\frac{1}{2}, \frac{1}{2}])$ are treated as elements of $L^2([-1, 1])$ which vanish outside $[-\frac{1}{2}, \frac{1}{2}]$. In this case $\cos(\varphi_{\mathcal{T},\mathcal{V}}) = 1$, and $\{P_{\mathcal{T}} e^{2\pi i x_j \cdot}\}_{j\in\mathbb{N}}$ is a frame for $\mathcal{T} = L^2([-\frac{1}{2}, \frac{1}{2}])$. Clearly a frame for $L^2([-\frac{1}{2}, \frac{1}{2}])$ is not a frame for $L^2([-1, 1])$, and therefore $\cos(\varphi_{\mathcal{V},\mathcal{T}}) = 0$.

3.4.2 Oblique projections

In Section 4, we show that any quasi-optimal and bounded reconstruction operator $Q : l^2(\mathbb{N}) \to \mathcal{T}$ has the property that the operator $P = QV^*$ is an oblique projection onto the reconstruction space. Therefore we collect some facts on oblique projections.

Definition 3.4.3. *Let $\mathcal{X}$ be a vector space and let $\mathcal{T}$ and $\mathcal{W}$ be subspaces of $\mathcal{X}$. We say that $\mathcal{X}$ is the direct sum of $\mathcal{T}$ and $\mathcal{W}$, if every $x \in \mathcal{X}$ can be written uniquely as $x = s + w$ with $s \in \mathcal{T}$ and $w \in \mathcal{W}$. If $\mathcal{X}$ is the direct sum of $\mathcal{T}$ and $\mathcal{W}$ write*

$$\mathcal{X} = \mathcal{T} \oplus \mathcal{W}.$$

Remark 3.4.4. *The decomposition $x = s + w$ with $x \in \mathcal{T}$ and $w \in \mathcal{W}$ is unique if and only if $\mathcal{T} \cap \mathcal{W} = \{0\}$.*

Definition 3.4.5. *Let $\mathcal{X}$ be a vector space and let $P : \mathcal{X} \to \mathcal{X}$ be a linear mapping. P is called a projection if and only if*

$$P^2 = P.$$

We also say oblique projection to a projection, to indicate that the projection need not be the orthogonal projection.

The following lemma collects the main properties of oblique projections. The first point of Lemma 3.4.6 is stated in [63, Theorem 2.1], the proof of the second point can be found in [19, Theorem 1], and for the third point see [10, Corollary 3.5], [61] and [19].

Lemma 3.4.6. *Let $\mathcal{T}$ and $\mathcal{W}$ are closed subspaces of a Hilbert space $\mathcal{H}$. Then*

1. $\cos(\varphi_{\mathcal{T},\mathcal{W}^\perp}) > 0$ *if and only if $\mathcal{T} \cap \mathcal{W} = \{0\}$ and $\mathcal{T} \oplus \mathcal{W}$ is closed in $\mathcal{H}$.*
2. *If $\mathcal{T} \cap \mathcal{W} = \{0\}$ and $\mathcal{H}_1 := \mathcal{T} \oplus \mathcal{W}$ is a closed subspace of $\mathcal{H}$, then the oblique projection $P_{\mathcal{T},\mathcal{W}} : \mathcal{H}_1 \to \mathcal{T}$ with range $\mathcal{T}$ and kernel $\mathcal{W}$ is well defined and bounded on $\mathcal{H}_1$.*

3. Let $\cos(\varphi_{\mathcal{T},\mathcal{W}^\perp}) > 0$ *and let* $\mathcal{H}_1 := \mathcal{T} \oplus \mathcal{W}$.

For the oblique projection $P_{\mathcal{T},\mathcal{W}} : \mathcal{H}_1 \to \mathcal{T}$ *it holds*

$$\|P_{\mathcal{T},\mathcal{W}}\| = \frac{1}{\cos(\varphi_{\mathcal{T},\mathcal{W}^\perp})} \tag{3.12}$$

and

$$\|f - P_{\mathcal{T}}f\| \leqslant \|f - P_{\mathcal{T},\mathcal{W}}f\| \leqslant \frac{1}{\cos(\varphi_{\mathcal{T},\mathcal{W}^\perp})}\|f - P_{\mathcal{T}}f\|, \tag{3.13}$$

for all $f \in \mathcal{H}_1$. *The upper bound in* (3.13) *is sharp.*

From Lemma 3.4.6, (3.), we obtain the following useful Corollary.

Corollary 3.4.7. *Let* $\mathcal{T}$ *and* $\mathcal{V}$ *be closed subspaces of* $\mathcal{H}$. *Let* $\{u_j\}_{j\in\mathbb{N}}$ *be a frame for* $\mathcal{V}$, *and let* V *and* V^* *be the corresponding synthesis and analysis operator. If* $Q : l^2(\mathbb{N}) \to \mathcal{T}$ *is an operator onto* $\mathcal{T}$ *with the property that* QV^* *is an oblique projection with range* $\mathcal{T}$ *and null-space* $\mathcal{W}^\perp$, *then*

$$\mu(Q) = \|QV^*\| = \frac{1}{\cos(\varphi_{\mathcal{T},\mathcal{W}})}.$$

Definition 3.4.8. *Let* $\{u_j\}_{j\in\mathbb{N}}$ *and* $\{g_k\}_{k\in\mathbb{N}}$ *be Bessel sequences in* $\mathcal{H}$. *We call the bounded operator* V^*T *the cross-Gramian of* $\{u_j\}_{j\in\mathbb{N}}$ *with respect to* $\{g_k\}_{k\in\mathbb{N}}$.

Lemma 3.4.9. *Let* $\mathcal{T}$ *and* $\mathcal{V}$ *be closed subspaces of a separable Hilbert space* $\mathcal{H}$. *Let* $\{u_j\}_{j\in\mathbb{N}}$ *be a frame for* $\mathcal{V}$, *and let* V *and* V^* *be the corresponding synthesis and analysis operator. Let* $\{g_k\}_{k\in\mathbb{N}}$ *be a Riesz basis for* $\mathcal{T}$, *and* T *the corresponding synthesis operator of* $\{g_k\}_{k\in\mathbb{N}}$.

Then $\mathcal{T} \cap \mathcal{V}^\perp = \{0\}$ *if and only if the cross-Gramian* V^*T *is injective.*

Proof. For $c \in l^2(\mathbb{N})$,

$$(V^*Tc)_j = \sum_{l\in\mathbb{N}} c_l \langle g_l, u_j\rangle = \langle Tc, u_j\rangle.$$

This means that

$$\mathcal{N}(V^*T) = \left\{ c \in l^2(\mathbb{N}) : g = \sum_{k\in\mathbb{N}} c_k g_k \in \mathcal{V}^\perp \right\}.$$

Since $\{g_k\}_{k\in\mathbb{N}}$ is a Riesz basis for $\mathcal{T}$, this means that $\mathcal{N}(V^*T) = \{0\}$ if and only if $\mathcal{T} \cap \mathcal{V}^\perp = \{0\}$.

□

Lemma 3.4.9 implies, that for finite sequences $\{u_j\}_{j\in J}$ and $\{g_k\}_{k\in K}$ with $\{g_k\}_{k\in K}$, linearly independent, $\cos(\varphi_{\mathcal{T},\mathcal{V}}) > 0$ if and only if the cross-Gramian V^*T is injective.

Chapter 4

Connection of quasi-optimal operators to dual frames

In this section, we point out the connections of quasi-optimal and bounded reconstruction operators to dual frames and pseudoframes.

Definition 4.0.10. *Let $\mathcal{T}$ be a closed subspace of $\mathcal{H}$ and let $\{u_j\}_{j\in\mathbb{N}}$ and $\{g_k\}_{k\in\mathbb{N}}$ be Bessel sequences in $\mathcal{H}$. We say that $\{g_k\}_{k\in\mathbb{N}}$ is a pseudoframe for the subspace $\mathcal{T}$ w.r.t. $\{u_j\}_{j\in\mathbb{N}}$ if for every $f \in \mathcal{T}$*

$$f = \sum_{k\in\mathbb{N}} \langle f, u_k \rangle g_k.$$

Let $\{u_j\}_{j\in\mathbb{N}}$ be a sequence in $\mathcal{H}$ and $\mathcal{V} = \overline{\text{span}}\{u_j\}_{j\in\mathbb{N}}$. In [48, Proposition 3] it is shown that if $\{g_k\}_{k\in\mathbb{N}}$ is a pseudoframe for the subspace $\mathcal{T}$ w.r.t. $\{u_j\}_{j\in\mathbb{N}}$, then $\mathcal{T} \cap \mathcal{V}^\perp = \{0\}$. Therefore $\mathcal{T} \cap \mathcal{V}^\perp = \{0\}$ is a necessary condition for the existence of a pseudoframe for the subspace $\mathcal{T}$ w.r.t. $\{u_j\}_{j\in\mathbb{N}}$. The following Lemma shows that if $\{u_j\}_{j\in\mathbb{N}}$ is a frame for $\mathcal{V}$, then $\cos(\varphi_{\mathcal{T},\mathcal{V}}) > 0$, (i.e. $\mathcal{T} \cap \mathcal{V}^\perp = \{0\}$ and $\mathcal{T} \oplus \mathcal{V}^\perp$ is closed in $\mathcal{H}$) is sufficient for the existence of a pseudoframe for the subspace $\mathcal{T}$ w.r.t. $\{u_j\}_{j\in\mathbb{N}}$. In addition the pseudoframes $\{g_k\}_{k\in\mathbb{N}} \subset \mathcal{T}$, (i.e. the reconstruction vectors are assumed to lie inside the reconstruction space) are exactly the dual frames of $\{P_\mathcal{T} u_k\}_{k\in\mathbb{N}}$.

The proof of the following theorem is similar to [48, Theorem 4]

Lemma 4.0.11. *Let $\mathcal{T}$ be a closed subspace of $\mathcal{H}$ and let $\{u_j\}_{j\in\mathbb{N}}$ be a Bessel sequence. There exists a pseudoframe $\{g_k\}_{k\in\mathbb{N}}$ for the subspace $\mathcal{T}$ w.r.t. $\{u_j\}_{j\in\mathbb{N}}$ if and only if $\{P_{\mathcal{T}}u_k\}_{k\in\mathbb{N}}$ is a frame for $\mathcal{T}$.*

The pseudoframes $\{g_k\}_{k\in\mathbb{N}} \subset \mathcal{T}$ for the subspace $\mathcal{T}$ w.r.t. $\{u_j\}_{j\in\mathbb{N}}$ are exactly the dual frames of $\{P_{\mathcal{T}}u_k\}_{k\in\mathbb{N}}$.

Proof. Let $\{g_k\}_{k\in\mathbb{N}}$ be a pseudoframe for the subspace $\mathcal{T}$ w.r.t. $\{u_j\}_{j\in\mathbb{N}}$. This is equivalent to the statement that for $g \in \mathcal{T}$

$$g = \sum_{k\in\mathbb{N}} \langle P_{\mathcal{T}}g, u_k\rangle g_k = \sum_{k\in\mathbb{N}} \langle f, P_{\mathcal{T}}u_k\rangle g_k. \tag{4.1}$$

Equation (4.1) implies that for $g \in \mathcal{T}$

$$g = P_{\mathcal{T}}(g) = P_{\mathcal{T}}\left(\sum_{k\in\mathbb{N}} \langle f, P_{\mathcal{T}}u_k\rangle g_k\right) = \sum_{k\in\mathbb{N}} \langle f, P_{\mathcal{T}}u_k\rangle P_{\mathcal{T}}g_k.$$

By assumption the sequences $\{u_j\}_{j\in\mathbb{N}}$ and $\{g_k\}_{k\in\mathbb{N}} \subset \mathcal{T}$ are both Bessel sequences. Therefore also the sequences $\{P_{\mathcal{T}}u_j\}_{j\in\mathbb{N}}$ and $\{P_{\mathcal{T}}g_k\}_{k\in\mathbb{N}} \subset \mathcal{T}$ are Bessel sequences, and by Lemma 2.3.9 $\{P_{\mathcal{T}}u_j\}_{j\in\mathbb{N}}$ and $\{P_{\mathcal{T}}g_k\}_{k\in\mathbb{N}} \subset \mathcal{T}$ are both frames for $\mathcal{T}$, dual frames of each other.

Conversely, if $\{P_{\mathcal{T}}u_k\}_{k\in\mathbb{N}}$ is a frame for $\mathcal{T}$, every dual frame $\{h_k\}_{k\in\mathbb{N}} \subset \mathcal{T}$ fulfills (4.1), and therefore $\{h_k\}_{k\in\mathbb{N}}$ is a pseudoframe for the subspace $\mathcal{T}$ w.r.t. $\{u_j\}_{j\in\mathbb{N}}$.

The second statement follows directly from (4.1). □

The reconstruction operators Q we treat in this thesis all have the property that $\mathcal{R}(Q) = \mathcal{T}$, which means that all reconstructions are located inside the reconstruction space $\mathcal{T}$. Among those operators we are interested in bounded and quasi-optimal operators. The following theorem gives several equivalent statements for an operator $Q : l^2(\mathbb{N}) \to \mathcal{T}$ being bounded and quasi-optimal.

Theorem 4.0.12. *Let $\mathcal{T}$ and $\mathcal{V}$ be closed subspaces of $\mathcal{H}$, and $\{u_j\}_{j\in\mathbb{N}}$ be a Bessel sequence with the closed linear span $\mathcal{V}$. For an operator $Q : l^2(\mathbb{N}) \to \mathcal{T}$ the following are equivalent.*

1. *There exist constants* $\mu, B \geqslant 0$*, such that for* $f \in \mathcal{H}$ *and* $c \in l^2(\mathbb{N})$

$$\|f - Q(V^* f + c)\| \leqslant \mu \|f - P_{\mathcal{T}} f\| + B\|c\|. \tag{4.2}$$

2. *The operator* Q *is quasi-optimal (with respect to* $\{u_j\}_{j\in\mathbb{N}}$*) and bounded.*

3. $QV^* g = g$ *for* $g \in \mathcal{T}$ *and* Q *is a bounded operator.*

4. *The sequence* $\{P_{\mathcal{T}} u_k\}_{k\in\mathbb{N}}$ *is a frame for* $\mathcal{T}$*. The operator* Q *is of the form*

$$Qc = \sum_{k\in\mathbb{N}} c_k h_k, \tag{4.3}$$

where $\{h_k\}_{k\in\mathbb{N}} \subset \mathcal{T}$ *is a dual frame of* $\{P_{\mathcal{T}} u_k\}_{k\in\mathbb{N}}$*, i.e.* Q *is the synthesis operator of a dual frame of* $\{P_{\mathcal{T}} u_k\}_{k\in\mathbb{N}}$*.*

5. *The sequence* $\{P_{\mathcal{T}} u_k\}_{k\in\mathbb{N}}$ *is a frame for* $\mathcal{T}$ *and* Q *is the synthesis operator of a pseudoframe* $\{g_k\}_{k\in\mathbb{N}} \subset \mathcal{T}$ *for the subspace* $\mathcal{T}$ *w.r.t.* $\{u_j\}_{j\in\mathbb{N}}$*.*

6. *The operator* Q *is bounded and* QV^* *is a bounded oblique projection onto* $\mathcal{T}$*.*

Proof. That (1) implies (2) is obtained by setting $c = 0$ and $f = 0$ in (4.2).

From (2) it follows that $QV^* g = g$ for $g \in \mathcal{T}$, since otherwise $\mu = \infty$. This implies (3).

Next we show that (3) implies (4). Let $Q : l^2(\mathbb{N}) \to \mathcal{T}$ be a bounded operator with $QV^* g = g$ for $g \in \mathcal{T}$. Let $\{e_k\}_{k\in\mathbb{N}}$ be the standard basis of $l^2(\mathbb{N})$ and let $h_k = Qe_k$. Then

$$Qc = \sum_{k\in\mathbb{N}} c_k h_k.$$

In particular for $g \in \mathcal{T}$,

$$QV^* g = \sum_{k\in\mathbb{N}} \langle g, P_{\mathcal{T}} u_k \rangle h_k = g. \tag{4.4}$$

Since Q is bounded also Q^* is bounded, and consequently $\{h_k\}_{k\in\mathbb{N}}$ is a Bessel sequence in $\mathcal{T}$. By assumption $\{u_k\}_{k\in\mathbb{N}}$ is a Bessel sequence in $\mathcal{V}$ with Bessel bound B and consequently

$$\sum_{k\in\mathbb{N}} |\langle f, P_{\mathcal{T}} u_k\rangle|^2 = \sum_{k\in\mathbb{N}} |\langle P_{\mathcal{T}} f, u_k\rangle|^2 \leqslant B\|P_{\mathcal{T}} f\|^2 \leqslant B\|f\|^2,$$

and $\{P_{\mathcal{T}} u_k\}_{k\in\mathbb{N}}$ is a Bessel sequence. From (4.4) and Lemma 2.3.9 we infer that $\{g_k\}_{k\in\mathbb{N}}$ is a dual frame of $\{P_{\mathcal{T}} u_k\}_{k\in\mathbb{N}}$.

The equivalence of (4) and (5) is shown in Lemma 4.0.11.

Next we show that (4) implies (6). Let the operator Q be defined by $Qc = \sum_{k\in\mathbb{N}} c_k h_k$, where $\{h_k\}_{k\in\mathbb{N}}$ is a dual frame of $\{P_{\mathcal{T}} u_k\}_{k\in\mathbb{N}}$. We define $P := QV^*$. Since $\mathcal{R}(Q) \subset \mathcal{T}$ and $QV^*g = g$ for $g \in \mathcal{T}$, it follows that $\mathcal{R}(P) = \mathcal{T}$ and that

$$P^2 = QV^*QV^* = QV^* = P.$$

Since both Q and V^* are bounded, P is bounded.

Finally we prove that (6) implies (1). Let Q be a bounded operator, and let $P := QV^*$ be a bounded oblique projection onto $\mathcal{T}$. Lemma 3.4.6, (3.), implies that

$$\|f - Pf\| \leqslant \|P\|\|f - P_{\mathcal{T}} f\|,$$

and consequently

$$\|f - Q(V^*f + c)\| \leqslant \|P\|\|f - P_{\mathcal{T}} f\| + \|Q\|\|c\|.$$

□

The following lemma shows that the synthesis operator of the canonical dual frame of $\{P_{\mathcal{T}} u_k\}_{k\in\mathbb{N}}$ is the operator with the smallest operator norm among all quasi-optimal and bounded operators $Q : l^2(\mathbb{N}) \to \mathcal{T}$.

Lemma 4.0.13. *Let $\mathcal{T}$ be a closed subspaces of $\mathcal{H}$. Let $\{u_j\}_{j\in\mathbb{N}}$ be a Bessel sequence in $\mathcal{H}$, let V be the corresponding synthesis operator, let V^* be the corresponding analysis operator and let $\{P_{\mathcal{T}} u_k\}_{k\in\mathbb{N}}$ be a frame for $\mathcal{T}$. Let $\{h_k\}_{k\in\mathbb{N}}$ denote the canonical dual frame of $\{P_{\mathcal{T}} u_k\}_{k\in\mathbb{N}}$, let T denote the synthesis operator of $\{h_k\}_{k\in\mathbb{N}}$, let T^* denote the corresponding analysis operator* *and let S denote the corresponding frame operator. If $Q : l^2(\mathbb{N}) \to \mathcal{H}$ is a bounded operator with the property that $QV^*g = g$ for $g \in \mathcal{T}$, then*

$$\|Q\| \geqslant \|T\|.$$

Proof. Let $Q : l^2(\mathbb{N}) \to \mathcal{T}$ be a bounded operator with the property that $QV^*g = g$ for $g \in \mathcal{T}$. Since for $g \in \mathcal{T}$

$$P_{\mathcal{T}}QV^*g = P_{\mathcal{T}}g = g,$$

and $\|P_{\mathcal{T}}Q\| \leqslant \|Q\|$ we may assume that $\mathcal{R}(Q) \subset \mathcal{T}$. By Theorem 4.0.12 Q is the synthesis operator of a dual frame of $\{P_{\mathcal{T}}u_k\}_{k\in\mathbb{N}}$. From Lemma 2.3.11 we infer that for $g \in \mathcal{T}$

$$\|Q^*g\|^2 = \|T^*g\|^2 + \sum_{k\in\mathbb{N}} \left|\langle g, h_k\rangle - \langle g, S^\dagger P_{\mathcal{T}}u_k\rangle\right|^2,$$

and consequently $\|Q^*_{|\mathcal{T}}\| \geqslant \|T^*_{|\mathcal{T}}\|$. For $g^\perp \in \mathcal{T}^\perp$ it holds $T^*g^\perp = 0$. Therefore $\|Q^*\| \geqslant \|T^*\|$. By Lemma 2.1.2 this implies that $\|Q\| \geqslant \|T\|$. □

Chapter 5

Consistent, generalized and frame independent sampling

We recall the problem of reconstructing a compactly supported function f from a finite number of perturbed Fourier coefficients

$$d_j = \mathcal{F}f(j) + \delta_j, \quad j = -\left\lfloor \frac{n-1}{2} \right\rfloor, \ldots, \left\lfloor \frac{n}{2} \right\rfloor.$$

When using the linear span of the first n Legendre polynomials P_k, $k = 0, \ldots, n-1$, as reconstruction space $\mathcal{T}$, the moment problem

$$\left\langle \tilde{f}, e^{2\pi i j \cdot} \right\rangle = d_j, \quad j = -\left\lfloor \frac{n-1}{2} \right\rfloor, \ldots, \left\lfloor \frac{n}{2} \right\rfloor, \tag{5.1}$$

has a unique solution $\tilde{f} \in \mathcal{T}$. We observe that the Fourier coefficients $\mathcal{F}\tilde{f}(j)$ of the reconstruction $\tilde{f}$ defined by (5.1), coincide with the given measurements d_j, $j = -\left\lfloor \frac{n-1}{2} \right\rfloor, \ldots, \left\lfloor \frac{n}{2} \right\rfloor$. Therefore $\tilde{f}$ is a so called consistent reconstruction. For general sampling vectors $\{u_j\}_{j\in\mathbb{N}}$ consistency means that

$$\langle \tilde{f}, u_j \rangle = d_j, \quad j \in \mathbb{N}. \tag{5.2}$$

The notion of consistent sampling is introduced in [25, 26] and further developed in [20, 22, 27, 28].

As already mentioned, calculating n Legendre coefficients from n Fourier coefficients yields an exponentially growing condition number of the matrix of the resulting linear problem. Therefore for a stable reconstruction more Fourier coefficients than Legendre polynomials are needed. If there are more measurements than reconstruction vectors, there exists no consistent reconstruction anymore. Therefore a generalization of the concept of consistent sampling is needed. In this chapter we present several generalizations of the concept of consistent sampling, but first we review some properties of the concept of consistent sampling.

5.1 Consistent sampling

The following Lemma can be found in [27, Lemma 1]

Lemma 5.1.1. *Let $\mathcal{T}$ and $\mathcal{V}$ be closed subspaces of a separable Hilbert space $\mathcal{H}$. The following are equivalent.*

1. $\mathcal{H} = \mathcal{T} \oplus \mathcal{V}^\perp$
2. $\mathcal{H} = \mathcal{V} \oplus \mathcal{T}^\perp$
3. $\cos(\varphi_{\mathcal{T},\mathcal{V}}) > 0$ *and* $\cos(\varphi_{\mathcal{V},\mathcal{T}}) > 0$.

The following Lemma can be found in [27, Lemma 4]

Lemma 5.1.2. *Let $\mathcal{T}$ and $\mathcal{V}$ be closed subspaces of a separable Hilbert space $\mathcal{H}$. If $\mathcal{H} = \mathcal{T} \oplus \mathcal{V}^\perp$, then the oblique projection $P_{\mathcal{T},\mathcal{V}^\perp} : \mathcal{H} \to \mathcal{T}$ with range $\mathcal{T}$ and kernel $\mathcal{V}^\perp$ is well defined and bounded. Let $\mathcal{H}_1$ be another separable Hilbert space, let $T : \mathcal{H}_1 :\to \mathcal{T}$ and $V : \mathcal{H}_1 \to \mathcal{V}$ be bounded operators with $\mathcal{R}(T) = \mathcal{T}$ and $\mathcal{R}(V) = \mathcal{V}$. Then the following hold.*

1. *The operator $(V^*T)^\dagger$ is a bounded operator on $\mathcal{H}_1$.*
2. *The oblique projection $P_{\mathcal{T},\mathcal{V}^\perp} : \mathcal{H} \to \mathcal{T}$ with range $\mathcal{T}$ and kernel $\mathcal{V}^\perp$ can be written in the form*

$$P_{\mathcal{T},\mathcal{V}^\perp} = T(V^*T)^\dagger V^* \tag{5.3}$$

3. The operator

$$Q_c := T(V^*T)^\dagger$$

is independent of the particular choice of the bounded operator T, as long as $\mathcal{R}(T) = \mathcal{T}$.

If we choose $\mathcal{H}_1 = l^2(\mathbb{N})$, V the synthesis operator of a frame $\{u_j\}_{j\in\mathbb{N}}$ for $\mathcal{V}$ and T the synthesis operator of a frame for $\mathcal{T}$, then clearly the assumptions of Lemma 5.1.2 are fulfilled. Since the following theory is developed for this setup, we use these assumptions instead of the more general of Lemma 5.1.2.

We use the notation

$$Q_c := T(V^*T)^\dagger. \tag{5.4}$$

The subscript c of Q_c is to indicate that by this operator we obtain a consistent reconstruction.

Lemma 5.1.3. *Let $\mathcal{T}$ and $\mathcal{V}$ be closed subspaces of $\mathcal{H}$. Let $\{u_j\}_{j\in\mathbb{N}}$ be a frame for $\mathcal{V}$ and let V be the corresponding synthesis operator. Let $\{g_k\}_{k\in\mathbb{N}}$ be a frame for $\mathcal{T}$ and let T be the corresponding synthesis operator. Let $\mathcal{H} = \mathcal{T} \oplus \mathcal{V}^\perp$ and let Q_c be the operator defined by* (5.4). *Then*

$$Q_{c|\mathcal{R}(V^*)^\perp} = 0, \tag{5.5}$$

and consequently

$$\|Q_{c|\mathcal{R}(V^*)}\| = \|Q_c\|. \tag{5.6}$$

Proof. Using that $\mathcal{N}(A^\dagger) = \mathcal{N}(A^*)$ we obtain

$$\begin{aligned} \mathcal{R}(V^*)^\perp &= \mathcal{N}(V) \subset \mathcal{N}(T^*V) = \mathcal{N}\left((V^*T)^*\right) \\ &= \mathcal{N}\left((V^*T)^\dagger\right) \subset \mathcal{N}(T(V^*T)^\dagger) = \mathcal{N}(Q_c). \end{aligned} \tag{5.7}$$

□

Corollary 5.1.4. *Let $\mathcal{T}$ and $\mathcal{V}$ be closed subspaces of $\mathcal{H}$ and Q_c be defined by* (5.4). *If $\mathcal{T} \oplus \mathcal{V}^\perp = \mathcal{H}$, then*

$$\mu(Q_c) = \|Q_cV^*\| = \|P_{\mathcal{T},\mathcal{V}^\perp}\| = \frac{1}{\cos(\varphi_{\mathcal{T},\mathcal{V}})}. \tag{5.8}$$

Proof. Equation (5.8) follows from Corollary 3.4.7 and Lemma 5.1.2. □

5.2 Generalized sampling

In this section we treat the concept of generalized sampling. In the finite dimensional case, for linearly independent sampling vectors and linearly independent reconstruction vectors, a unique consistent reconstruction typically exists if there are as many sampling vectors as reconstruction vectors. Generalized sampling is a generalization of the concept of consistent sampling, where more sampling vectors than reconstruction vectors are allowed.

Given the perturbed measurements

$$d_j = \langle f, u_j \rangle + \delta_j, \quad j = -n, \dots, n,$$

of an element $f \in \mathcal{H}$, the concept of generalized sampling chooses as approximation $\tilde{f} \in \mathcal{T}$ the solution of the standard least squares problem

$$\tilde{f} = \arg\min_{g \in \mathcal{T}} \sum_{j=-n}^{n} |\langle g, u_j \rangle - d_j|^2 .$$

The reconstruction operator $Q_g : l^2(\mathbb{N}) \to \mathcal{T}$ defined by this property is analyzed in detail in [10]. This reconstruction operator is of great importance, since it has the smallest operator norm among all quasi-optimal reconstructions. We review the most important properties of this reconstruction and supplement some additional aspects.

In order to prove Theorem 5.2.5 we need Lemma 5.2.1, Lemma 5.2.2, Lemma 5.2.3 and Lemma 5.2.4.

The proof of Lemma 5.2.1 is similar to the proof of [3, Theorem 3.4].

Lemma 5.2.1. *Let $\mathcal{T}$ and $\mathcal{V}$ be closed subspaces of $\mathcal{H}$. Let $\{u_j\}_{j\in\mathbb{N}}$ be a frame for $\mathcal{V}$, with lower and upper frame bound A and B, and corresponding frame operator S. If $\cos(\varphi_{\mathcal{T},\mathcal{V}}) > 0$, then $S(\mathcal{T})$ is closed, and*

$$\cos(\varphi_{\mathcal{T},S(\mathcal{T})}) \geqslant \frac{\sqrt{A}}{\sqrt{B}} \cos(\varphi_{\mathcal{T},\mathcal{V}}). \tag{5.9}$$

Proof. Since S is the frame operator of the frame sequence $\{u_j\}_{j\in\mathbb{N}}$, for every $u \in \mathcal{V}$ it holds

$$A\|u\| \leqslant \|Su\| \leqslant B\|u\|. \tag{5.10}$$

We recall inequality (3.10)

$$\|g\| \cos(\varphi_{\mathcal{T},\mathcal{V}}) \leqslant \|P_{\mathcal{V}} g\| \quad \text{for all } g \in \mathcal{T}.$$

Using (5.10) and the fact that $S = SP_{\mathcal{V}}$, we obtain

$$A \cos(\varphi_{\mathcal{T},\mathcal{V}})\|g\| \leqslant A\|P_{\mathcal{V}} g\| \leqslant \|SP_{\mathcal{V}} g\| = \|Sg\| \quad \text{for } g \in \mathcal{T}. \tag{5.11}$$

Equation (5.11) implies that $S_{|\mathcal{T}}$ has a closed range by virtue of Lemma 2.1.7. Since $\mathcal{R}(S_{|\mathcal{T}}) = S(\mathcal{T})$, the set $S(\mathcal{T})$ is closed.

By definition

$$\cos(\varphi_{\mathcal{T},S(\mathcal{T})}) = \inf_{\substack{g\in\mathcal{T}\\ g\neq 0}} \sup_{\substack{h\in\mathcal{T}\\ Sh\neq 0}} \frac{|\langle g, Sh\rangle|}{\|g\|\|Sh\|}. \tag{5.12}$$

Equation (5.11) implies that for $g \in \mathcal{T}$, $g \neq 0$, also $Sg \neq 0$. Combining this with (5.12) and (5.11) we obtain

$$\cos(\varphi_{\mathcal{T},S(\mathcal{T})}) \geqslant \inf_{\substack{g\in\mathcal{T}\\ g\neq 0}} \frac{\langle g, Sg\rangle}{\|g\|\|Sg\|}. \tag{5.13}$$

Using the positive definiteness of S and Cauchy-Schwarz inequality, we obtain

$$\begin{aligned}\|Sg\| &= \sup_{\substack{h\in\mathcal{H}\\ \|h\|=1}} |\langle Sg, h\rangle| = \sup_{\substack{h\in\mathcal{H}\\ \|h\|=1}} |\langle \sqrt{S}g, \sqrt{S}h\rangle| \leqslant \sup_{\substack{h\in\mathcal{H}\\ \|h\|=1}} \sqrt{\langle Sg, g\rangle}\sqrt{\langle Sh, h\rangle} \\ &= \sqrt{\langle Sg, g\rangle}\ \|\sqrt{S}h\| \leqslant \sqrt{\langle Sg, g\rangle}\sqrt{B}.\end{aligned} \tag{5.14}$$

Combining (5.14) and (5.13) we obtain

$$\cos(\varphi_{\mathcal{T},S(\mathcal{T})}) \geqslant \frac{1}{\sqrt{B}} \inf_{\substack{g\in\mathcal{T}\\ g\neq 0}} \frac{\sqrt{\langle g, Sg\rangle}}{\|g\|} \tag{5.15}$$

Using (5.10) and $\|g\| \cos(\varphi_{\mathcal{T},\mathcal{V}}) \leqslant \|P_{\mathcal{V}} g\|$, we obtain

$$\sqrt{\langle g, Sg\rangle} = \|\sqrt{S} g\| = \|\sqrt{S} P_{\mathcal{V}} g\| \geqslant \sqrt{A} \|P_{\mathcal{V}} g\| \geqslant \sqrt{A} \cos(\varphi_{\mathcal{T},\mathcal{V}}) \|g\|. \quad (5.16)$$

Combining (5.16) with (5.15) we deduce (5.9). □

Theorem 5.2.2. *Let $\mathcal{T}$ be a closed subspace of $\mathcal{H}$ and let $B : \mathcal{H} \to \mathcal{H}$ be a self-adjoint operator, such that $B(\mathcal{T})$ is closed. If $\cos(\varphi_{\mathcal{T},B(\mathcal{T})}) > 0$, then $\mathcal{H} = \mathcal{T} \oplus B(\mathcal{T})^{\perp}$, and the oblique projection $P_{\mathcal{T},B(\mathcal{T})^{\perp}} : \mathcal{H} \to \mathcal{T}$ is well defined and bounded on $\mathcal{H}$.*

Proof. By assumption $\cos(\varphi_{\mathcal{T},B(\mathcal{T})}) > 0$. By Lemma 3.4.6 the oblique projection $P_{\mathcal{T},B(\mathcal{T})^{\perp}}$ is well defined and bounded as a mapping from $\mathcal{T} \oplus B(\mathcal{T})^{\perp}$ onto $\mathcal{T}$.

It remains to prove that

$$\mathcal{T} \oplus B(\mathcal{T})^{\perp} = \mathcal{H}.$$

Lemma 3.4.6, (1.), implies that $\mathcal{T} \oplus B(\mathcal{T})^{\perp}$ is closed, so it is sufficient to show that $(\mathcal{T} \oplus B(\mathcal{T})^{\perp})^{\perp} = \{0\}$. By assumption, $B(\mathcal{T})$ is closed, and consequently

$$(\mathcal{T} \oplus B(\mathcal{T})^{\perp})^{\perp} = \mathcal{T}^{\perp} \cap \overline{B(\mathcal{T})} = \mathcal{T}^{\perp} \cap B(\mathcal{T}).$$

Since $\cos(\varphi_{\mathcal{T},B(\mathcal{T})}) > 0$, by (3.9)

$$\begin{aligned} \cos(\varphi_{\mathcal{T},B(\mathcal{T})}) &= \inf_{\substack{f \in \mathcal{T} \\ \|f\|=1}} \sup_{\substack{w \in B(\mathcal{T}) \\ \|w\|=1}} |\langle f, w\rangle| = \inf_{\substack{f \in \mathcal{T} \\ f \neq 0}} \sup_{\substack{s \in \mathcal{T} \\ Bs \neq 0}} \frac{|\langle f, Bs\rangle|}{\|f\|\|Bs\|} \\ &= \inf_{\substack{f \in \mathcal{T} \\ f \neq 0}} \sup_{\substack{s \in \mathcal{T} \\ Bs \neq 0}} \frac{|\langle Bf, s\rangle|}{\|f\|\|Bs\|} > 0. \end{aligned} \quad (5.17)$$

Equation (5.17) implies that for $f \in \mathcal{T} \backslash \{0\}$ it holds $Bf \notin \mathcal{T}^{\perp}$, and therefore $\mathcal{T}^{\perp} \cap B(\mathcal{T}) = \{0\}$. □

Lemma 5.2.3. *Let $\{u_j\}_{j \in \mathbb{N}}$ and $\{g_k\}_{k \in \mathbb{N}}$ be frames for $\mathcal{H}$. Let V denote the synthesis operator of $\{u_j\}_{j \in \mathbb{N}}$ and let V^* denote the corresponding analysis operator. Let T denote the synthesis operator of $\{g_k\}_{k \in \mathbb{N}}$.*

The cross-Gramian V^*T is bounded and has a closed range. Furthermore

$$\begin{aligned}\mathcal{N}(V^*T) &= \mathcal{N}(T) \text{ and}\\ \mathcal{R}(T^*V) &= \mathcal{R}(T^*).\end{aligned}$$

Proof. Since $\{u_j\}_{j\in\mathbb{N}}$ and $\{g_k\}_{k\in\mathbb{N}}$ are Bessel sequences, the cross-Gramian T^*V is bounded. Let $\{\tilde{u}_j\}_{j\in\mathbb{N}}$ be a dual frame of $\{u_j\}_{j\in\mathbb{N}}$. Let $f \in \mathcal{H}$. Setting $c_j = \langle f, \tilde{u}_j\rangle$ we obtain

$$(T^*Vc)_k = \sum_{j\in\mathbb{N}} \langle f, \tilde{u}_j\rangle\langle u_j, g_k\rangle = \langle f, g_k\rangle.$$

In other words, $\mathcal{R}(T^*V) = \mathcal{R}(T^*)$. Since $\{g_k\}_{k\in\mathbb{N}}$ is a frame, $\mathcal{R}(T^*)$ is closed and so is $\mathcal{R}(T^*V)$.

That $\mathcal{N}(V^*T) = \mathcal{N}(T)$ follows from

$$\mathcal{N}(T) = \mathcal{R}(T^*)^\perp = \mathcal{R}(T^*V)^\perp = \mathcal{N}(V^*T).$$

□

We need a slightly modified version of Lemma 5.2.3 for Bessel sequences.

Lemma 5.2.4. *Let $\mathcal{T}$ be a closed subspace of $\mathcal{H}$, let $\{g_k\}_{k\in\mathbb{N}}$ be a frame for $\mathcal{T}$ and let T denote the corresponding synthesis operator. Let $\{u_j\}_{j\in\mathbb{N}}$ be a Bessel sequence in $\mathcal{H}$. Let V denote the synthesis operator of $\{u_j\}_{j\in\mathbb{N}}$ and let V^* denote the corresponding analysis operator.*

If $\{P_\mathcal{T}u_j\}_{j\in\mathbb{N}}$ is a frame for $\mathcal{T}$, then the operator V^*T is bounded and has a closed range. Furthermore

$$\mathcal{N}(V^*T) = \mathcal{N}(T) \text{ and} \tag{5.18}$$

$$\mathcal{R}(T^*V) = \mathcal{R}(T^*) \tag{5.19}$$

Proof. Let $\{P_\mathcal{T}u_j\}_{j\in\mathbb{N}}$ be a frame for $\mathcal{T}$. Since $\{u_j\}_{j\in\mathbb{N}}$ and $\{g_k\}_{k\in\mathbb{N}}$ are both Bessel sequences, the operators V^* and T are bounded, and therefore also the composition V^*T is bounded. Since

$$(V^*T)_{j,k} = \langle g_k, u_j \rangle = \langle g_k, P_{\mathcal{T}} u_j \rangle,$$

the operator V^*T is a cross-Gramian of two frames for $\mathcal{T}$, and therefore by Lemma 5.2.3 V^*T has a closed range and $\mathcal{N}(V^*T) = \mathcal{N}(T)$. Since V^*T has a closed range, also T^*V has a closed range. As in Lemma 5.2.3 it is shown that $\mathcal{R}(T^*V) = \mathcal{R}(T^*)$. □

Equation (5.21) and (5.25) of Theorem 5.2.5 can be found in [10, Theorem 4.2.] and [10, Corollary 4.7.] for $\mathcal{T}$ finite dimensional. For the last point of Theorem 5.2.5 see [10, Section 4.1.].

Theorem 5.2.5. *Let $\mathcal{T}$ and $\mathcal{V}$ be closed subspaces of a separable Hilbert space $\mathcal{H}$. Let $\{u_j\}_{j\in\mathbb{N}}$ be a frame for $\mathcal{V}$ and let V be the corresponding synthesis operator and S be the corresponding frame operator. Let $\{g_k\}_{k\in\mathbb{N}}$ be a frame for $\mathcal{T}$ and let T be the corresponding synthesis operator. Furthermore let Q_g be defined by*

$$Q_g := T(V^*T)^{\dagger}. \tag{5.20}$$

If $\cos(\varphi_{\mathcal{T},\mathcal{V}}) > 0$, then $\mathcal{H} = \mathcal{T} \oplus S(\mathcal{T})^{\perp}$, the oblique projection $P_{\mathcal{T},S(\mathcal{T})^{\perp}}$ is well defined and bounded and the following holds.

- *The operator Q_g is defined by*

$$Q_g V^* = P_{\mathcal{T},S(\mathcal{T})^{\perp}} \tag{5.21}$$

  and

$$Q_{g|\mathcal{R}(V^*)^{\perp}} = 0, \tag{5.22}$$

  and consequently Q_g is independent of the particular choice of the frame $\{g_k\}_{k\in\mathbb{N}}$ for $\mathcal{T}$.

- $$Q_g = (V^*P_{\mathcal{T}})^{\dagger}. \tag{5.23}$$

- $$\|Q_{g|\mathcal{R}(V^*)}\| = \|Q_g\|. \tag{5.24}$$

- $$\mu(Q_g) = \|Q_g V^*\| = \|P_{\mathcal{T},S(\mathcal{T})^\perp}\| = \frac{1}{\cos(\varphi_{\mathcal{T},S(\mathcal{T})})} < \infty. \tag{5.25}$$

- *For* $d \in l^2(\mathbb{N})$, $\tilde{f} := Q_g d$ *is the unique least squares solution*

$$\tilde{f} = \arg\min_{\tilde{f}\in\mathcal{T}} \sum_{j\in\mathbb{N}} |\langle \tilde{f}, u_j\rangle - d_j|^2. \tag{5.26}$$

Furthermore it holds

$$Q_g d = \sum_{k=1}^{\infty} \hat{c}_k g_k$$

where $\hat{c} = \{\hat{c}_k\}_{k\in\mathbb{N}}$ *is the minimal norm element of the set*

$$\arg\min_{c} \|V^*Tc - d\|.$$

Proof. Since $\cos(\varphi_{\mathcal{T},\mathcal{V}}) > 0$, from Theorem 5.2.2 and Lemma 5.2.1 we infer that $\mathcal{T} \oplus S(\mathcal{T})^\perp = \mathcal{H}$, and that the oblique projection $P_{\mathcal{T},S(\mathcal{T})^\perp} : \mathcal{H} \to \mathcal{T}$ is well defined and bounded. Since $\cos(\varphi_{\mathcal{T},\mathcal{V}}) > 0$, by Lemma 3.4.2 the sequence $\{P_{\mathcal{T}} u_j\}_{j\in\mathbb{N}}$ is frame for $\mathcal{T}$. Therefore from Lemma 5.2.4 we infer that the operator V^*T is bounded and has a closed range. Lemma 2.2.1 implies that the operator $(V^*T)^\dagger$ is well defined and bounded, and consequently the operator $T(V^*T)^\dagger V^*$ is well defined and bounded. Next we prove (5.21).

We set $R := T(V^*T)^\dagger V^*$. We show

1. $R^2 = R$.

2. $\mathcal{R}(R) = \mathcal{T}$,

3. $\mathcal{N}(R) = S(\mathcal{T})^\perp$

The equality $R^2 = R$ follows from Corollary 2.2.4.

Clearly $\mathcal{R}(R) \subset \mathcal{T}$. To prove the inverse implication we show that $\mathcal{R}(R_{|\mathcal{T}}) = \mathcal{T}$. Since V^*T has a closed range, by Lemma 2.1.2 also T^*V has a closed range. Using Lemma 2.2.3 and (5.18) we infer that

$$\begin{aligned} RT &= T(V^*T)^\dagger V^*T = TP_{\mathcal{R}(T^*V)} = TP_{\overline{\mathcal{R}(T^*V)}} \\ &= TP_{\mathcal{N}(V^*T)^\perp} = TP_{\mathcal{N}(T)^\perp} = T. \end{aligned}$$

Equation (2.3) implies that

$$\mathcal{R}((V^*T)^\dagger) = \mathcal{R}(T^*V) \subset \mathcal{R}(T^*).$$

Therefore by Lemma 2.1.3 and (2.2)

$$\mathcal{N}(R) = \mathcal{N}(T(V^*T)^\dagger V^*) = \mathcal{N}((V^*T)^\dagger V^*) = \mathcal{N}(T^*VV^*). \tag{5.27}$$

We observe that

$$S(\mathcal{T})^\perp = \mathcal{R}(VV^*T)^\perp = \mathcal{N}(T^*VV^*),$$

which finishes the proof of (5.21).

The proof of (5.22) and (5.24) is similar to the proof of Lemma 5.1.3 and hence omitted.

Next we prove (5.23). Using Lemma 2.2.5 we obtain

$$(V^*P_\mathcal{T})^\dagger = P_\mathcal{T}V(V^*P_\mathcal{T}V)^\dagger.$$

Since the synthesis operator of the frame $\{P_\mathcal{T}u_j\}_{j\in\mathbb{N}}$ of $\mathcal{T}$ is equal to $P_\mathcal{T}V$, and the operator Q_g is independent of the particular choice of the frame $\{g_k\}_{k\in\mathbb{N}}$ for $\mathcal{T}$, we can write Q_g in the form $Q_g = P_\mathcal{T}V(V^*P_\mathcal{T}V)^\dagger$, which proves (5.23).

Equation (5.25) follows from Lemma 3.4.6, (3.), (5.21) and (5.9).

From (5.18) it follows that $\mathcal{N}(V^*) \cap \mathcal{T} = \{0\}$. Otherwise there exists an element $g = Tc \in \mathcal{T}$, $g \neq 0$, with $V^*g = V^*Tc = 0$, a contradiction to (5.18). Therefore for $d \in l^2(\mathbb{N})$ there exists a unique least squares solution $\tilde{f} \in \mathcal{T}$

$$\tilde{f} = \arg\min_{g\in\mathcal{T}} \|V^*g - d\|^2. \tag{5.28}$$

Since $\mathcal{R}(T) = \mathcal{T}$, there exists an element $c \in l^2(\mathbb{N})$, such that $\tilde{f} = Tc$, and by (5.28), c is an element of the set

$$B := \arg\min_{c} \|V^*Tc - d\|. \tag{5.29}$$

From Lemma 2.2.7 we infer that $c = (V^*T)^\dagger V^* f$ is the unique element of B of minimal norm. Therefore $\tilde{f} = T(V^*T)^\dagger V^* f$. $\square$

We observe that the operators Q_g and Q_c coincide. In Lemma 5.1.2 we assume that $\mathcal{H} = \mathcal{T} \oplus \mathcal{V}^\perp$, in which case $Q_c V^* = P_{\mathcal{T},\mathcal{V}^\perp}$, whereas in Theorem 5.2.5 we assume that $\cos(\varphi_{\mathcal{T},\mathcal{V}}) > 0$, in which case $Q_g V^* = P_{\mathcal{T},S(\mathcal{T})^\perp}$. We use the distinct subscripts to clarify which assumption on the spaces $\mathcal{T}$ and $\mathcal{V}$ is used. To understand the difference between the assumptions $\mathcal{H} = \mathcal{T} \oplus \mathcal{V}^\perp$ and $\cos(\varphi_{\mathcal{T},\mathcal{V}}) > 0$, we choose again the example of reconstructing a bandlimited function $f \in L^2(\mathbb{R}) \cap B_{[-\frac{1}{2},\frac{1}{2}]}$ from point samples

$$f(x_j) = \langle \mathcal{F} f, e^{-2\pi i x_j \cdot} \rangle, \quad j \in \mathbb{N}.$$

Let $\{e^{2\pi i x_j \cdot}\}_{j\in\mathbb{N}}$ be a frame for $\mathcal{V} = L^2([-1,1])$, and let $\mathcal{T} = L^2([-\frac{1}{2},\frac{1}{2}])$ be the reconstruction space. In this case $\cos(\varphi_{\mathcal{T},\mathcal{V}}) = 1$, and $\{P_\mathcal{T} e^{2\pi i x_j \cdot}\}_{j\in\mathbb{N}}$ is a frame for $\mathcal{T} = L^2([-\frac{1}{2},\frac{1}{2}])$. A frame for $L^2([-\frac{1}{2},\frac{1}{2}])$ is not a frame for $L^2([-1,1])$, and therefore $\cos(\varphi_{\mathcal{V},\mathcal{T}}) = 0$. By Lemma 5.1.1 $\mathcal{H} = \mathcal{T} \oplus \mathcal{V}^\perp$ is equivalent to $\cos(\varphi_{\mathcal{T},\mathcal{V}}) > 0$ and $\cos(\varphi_{\mathcal{V},\mathcal{T}}) > 0$, and therefore not fulfilled.

We further observe that $\mu(Q_g)$ depends on the operator S, which means that the kernel of the projection $Q_g V^* = P_{\mathcal{T},S(\mathcal{T})^\perp}$ depends on the sampling frame sequence $\{u_j\}_{j\in\mathbb{N}}$.

In Theorem 5.2.6 we present an alternative description of $P_{\mathcal{T},S(\mathcal{T})^\perp}$ and Q_g using dual frames.

Theorem 5.2.6. *Let $\mathcal{T}$ and $\mathcal{V}$ be closed subspaces of $\mathcal{H}$ and let $\cos(\varphi_{\mathcal{T},\mathcal{V}}) > 0$. Let $\{u_k\}_{k\in\mathbb{N}}$ be a frame of $\mathcal{V}$, let V be the corresponding synthesis operator and let $S = VV^*$ be the corresponding frame operator.*

Let $\{\tilde{g}_k\}_{k\in\mathbb{N}}$ be the canonical dual frame of $\{P_\mathcal{T} u_k\}_{k\in\mathbb{N}}$ of $\mathcal{T}$. Then

$$P_{\mathcal{T},S(\mathcal{T})^\perp} f = \sum_{k\in\mathbb{N}} \langle f, u_k \rangle \tilde{g}_k \quad \text{for every } f \in \mathcal{H} \tag{5.30}$$

and

$$Q_g c = \sum_{k\in\mathbb{N}} c_k \tilde{g}_k \quad \text{for every } c \in l^2(\mathbb{N}). \tag{5.31}$$

Proof. Equation (5.30) follows from (5.31) and $Q_g V^* = P_{\mathcal{T},S(\mathcal{T})^\perp}$, see (5.21).

The frame operator $\tilde{S}$ of $\{P_\mathcal{T} u_k\}_{k\in\mathbb{N}}$ can be written in the form

$$\tilde{S}(f) = \sum_{k\in\mathbb{N}} \langle f, P_\mathcal{T} u_k\rangle P_\mathcal{T} u_k = P_\mathcal{T}\Big(\sum_{k\in\mathbb{N}} \langle P_\mathcal{T} f, u_k\rangle u_k\Big) = P_\mathcal{T} V V^* P_\mathcal{T} f$$

Consequently we obtain for the synthesis operator L of the canonical dual frame of $\{P_\mathcal{T} u_k\}_{k\in\mathbb{N}}$

$$Lc = \sum_{k\in\mathbb{N}} c_k (P_\mathcal{T} V V^* P_\mathcal{T})^\dagger P_\mathcal{T} u_k = (P_\mathcal{T} V V^* P_\mathcal{T})^\dagger P_\mathcal{T} V c = (V^* P_\mathcal{T})^\dagger c,$$

where we used $A^\dagger = (A^*A)^\dagger A^*$ (see Lemma 2.2.5) for the last equality. By equation (5.23), $Q_g = (V^* P_\mathcal{T})^\dagger$. This proves (5.31).

□

Theorem 5.2.7 is stated in [10, Section 4] for $\mathcal{T}$ finite dimensional.

Theorem 5.2.7. *Let $\mathcal{H}$, $\mathcal{T}$, $\mathcal{V}$, $\{u_j\}_{j\in\mathbb{N}}$, S and $\{g_k\}_{k\in\mathbb{N}}$ be as in Theorem 5.2.5. If $\cos(\varphi_{\mathcal{T},\mathcal{V}}) > 0$, the mapping $P_{\mathcal{T},S(\mathcal{T})^\perp}$ is the unique operator $F : \mathcal{H} \to \mathcal{T}$ that satisfies the equations*

$$\langle SFf, g_k\rangle = \langle Sf, g_k\rangle, \quad k \in \mathbb{N},\ f \in \mathcal{H}. \tag{5.32}$$

Proof. We first show that $P_{\mathcal{T},S(\mathcal{T})^\perp}$ fulfills (5.32). Since $\cos(\varphi_{\mathcal{T},\mathcal{V}}) > 0$, from Theorem 5.2.5 we infer that $\mathcal{T} \oplus S(\mathcal{T})^\perp = \mathcal{H}$. For $f \in \mathcal{T}$, $P_{\mathcal{T},S(\mathcal{T})^\perp} g = g$ and (5.32) is clearly fulfilled. For $f \in S(\mathcal{T})^\perp$, $P_{\mathcal{T},S(\mathcal{T})^\perp} f = 0$ and $\langle Sf, g_k\rangle = \langle f, Sg_k\rangle = 0$.

Finally we prove the uniqueness. We assume that there are two mappings $F_1, F_2 : \mathcal{H} \to \mathcal{T}$ that satisfy (5.32). This means for all $f \in \mathcal{H}$ and $\Phi \in S(\mathcal{T})$

$$\langle F_1 f, \Phi\rangle = \langle f, \Phi\rangle = \langle F_2 f, \Phi\rangle. \tag{5.33}$$

From (5.33), it follows that $\mathcal{R}(F_1 - F_2) \subset S(\mathcal{T})^\perp$. By assumption $\mathcal{R}(F_1) \subset \mathcal{T}$ and $\mathcal{R}(F_2) \subset \mathcal{T}$ and thus $\mathcal{R}(F_1 - F_2) \subset \mathcal{T} \cap S(\mathcal{T})^\perp$. Since $\mathcal{T} \cap S(\mathcal{T})^\perp = \{0\}$, it follows that $F_1 = F_2$. □

The following theorem is proven in [10, Theorem 6.2.] in a more general setup.

Theorem 5.2.8. *Let $\mathcal{H}$, $\mathcal{T}$, $\mathcal{V}$, $\{u_j\}_{j\in\mathbb{N}}$, V and Q_g be as in Theorem 5.2.5. Let $\cos(\varphi_{\mathcal{T},\mathcal{V}}) > 0$ and let $Q : l^2(\mathbb{N}) \to \mathcal{H}$ be a bounded operator.*

If $QV^*g = g$ for $g \in \mathcal{T}$, then

$$\|Q_{g|\mathcal{R}(V^*)}\| \leqslant \|Q_{|\mathcal{R}(V^*)}\|. \tag{5.34}$$

Corollary 5.2.9. *Let $\mathcal{H}$, $\mathcal{T}$, $\mathcal{V}$, $\{u_j\}_{j\in\mathbb{N}}$, V and Q_g be as in Theorem 5.2.5. Let $\cos(\varphi_{\mathcal{T},\mathcal{V}}) > 0$ and let $Q : l^2(\mathbb{N}) \to \mathcal{H}$ be a bounded operator.*

If $QV^*g = g$ for all $g \in \mathcal{T}$, then

$$\|Q\| \geqslant \|Q_g\|.$$

Proof. From (5.34) and (5.24) we infer that

$$\|Q_g\| = \|Q_{g|\mathcal{R}(V^*)}\| \leqslant \|Q_{|\mathcal{R}(V^*)}\| \leqslant \|Q\|.$$

□

It should be mentioned that Corollary 5.2.9 also follows from Theorem 5.2.6 and Lemma 4.0.13.

5.2.1 Relative condition number

Since for a quasi-optimal operator $Q : l^2(\mathbb{N}) \to \mathcal{T}$ the corresponding projection operator $P = QV^*$ has in general a kernel, we used the operator norm $\|Q\|$ as a measure of stability. When reconstructing only elements in the reconstruction space $\mathcal{T}$, the operator QV^* has a trivial kernel and we can use the relative condition number.

Definition 5.2.10. *Let $\mathcal{H}$, $\mathcal{T}$, $\mathcal{V}$, $\{u_j\}_{j\in\mathbb{N}}$, and V be as in Theorem 5.2.5. Furthermore let $Q : l^2(\mathbb{N}) \to \mathcal{H}$. We define $\kappa_{\mathcal{T}}(Q)$ as*

$$\kappa_{\mathcal{T}}(Q) := \sup_{\substack{c\in l^2(\mathbb{N}) \\ c\neq 0}} \sup_{\substack{g\in\mathcal{T} \\ g\neq 0}} \frac{\|g - Q(V^*g + c)\|}{\|g\|} \Big/ \frac{\|c\|}{\|V^*g\|}.$$

The value $\kappa_{\mathcal{T}}(Q)$ is the supremum of the ratio between the relative error of the reconstruction and the relative coefficient error. Therefore $\kappa_{\mathcal{T}}(Q)$ measures how stable the reconstruction $\tilde{g} = Q(V^*g + c)$ of g is for elements g inside the reconstruction space $\mathcal{T}$.

As an easy consequence of Corollary 5.2.9 (respectively [10, Theorem 6.2.]) we obtain the following corollary.

Corollary 5.2.11. *Let $\mathcal{H}$, $\mathcal{T}$, $\mathcal{V}$, $\{u_j\}_{j\in\mathbb{N}}$, V and Q_g be as in Theorem 5.2.5. If $\cos(\varphi_{\mathcal{T},\mathcal{V}}) > 0$, then for every operator $Q : l^2(\mathbb{N}) \to \mathcal{H}$*

$$\kappa_{\mathcal{T}}(Q) \geqslant \kappa_{\mathcal{T}}(Q_g). \tag{5.35}$$

Proof. Let $Q : l^2(\mathbb{N}) \to \mathcal{H}$. We may assume

$$QV^*g = g \quad \text{for } g \in \mathcal{T}. \tag{5.36}$$

Otherwise

$$\lim_{c\to 0} \frac{\|g - Q(V^*g + c)\|}{\|c\|} = \lim_{c\to 0} \frac{\|g - QV^*g - Qc\|}{\|c\|} = \infty. \tag{5.37}$$

Using that for $g \neq 0$ the fraction $\frac{\|V^*g\|}{\|g\|}$ is bounded below by $\sqrt{A}\cos(\varphi_{\mathcal{T},\mathcal{V}})$, where A is the lower frame bound of $\{u_j\}_{j\in\mathbb{N}}$ (see (3.11)), equation (5.37) implies that $\kappa_{\mathcal{T}}(Q) = \infty$.

Therefore $\kappa_{\mathcal{T}}(Q)$ can be written as

$$\kappa_{\mathcal{T}}(Q) = \sup_{\substack{c\in l^2(\mathbb{N})\\ c\neq 0}} \frac{\|Qc\|}{\|c\|} \cdot \sup_{\substack{g\in\mathcal{T}\\ g\neq 0}} \frac{\|V^*g\|}{\|g\|}. \tag{5.38}$$

Since the second term on the right hand side of (5.38) is independent of Q, it is sufficient to show that $\|Q\| \geqslant \|Q_g\|$. This follows from (5.36) and Corollary 5.2.9 □

Lemma 5.2.12. *Let $Q : l^2(\mathbb{N}) \to \mathcal{H}$ be a bounded operator. Let the reconstruction vectors $\{g_k\}_{k\in\mathbb{N}}$ form a Riesz sequence. If $QV^*g = g$ for $g \in \mathcal{T}$,* *then*

$$\kappa_{\mathcal{T}}(Q) = \|Q\| \; \|V^*T(T^*T)^{-\frac{1}{2}}\|. \tag{5.39}$$

Proof. By (5.38) we have

$$\kappa_{\mathcal{T}}(Q) = \sup_{\substack{c\in l^2(\mathbb{N})\\ c\neq 0}} \frac{\|Qc\|}{\|c\|} \cdot \sup_{\substack{g\in\mathcal{T}\\ g\neq 0}} \frac{\|V^*g\|}{\|g\|}.$$

Therefore it remains to prove that

$$\sup_{\substack{g\in\mathcal{T}\\ g\neq 0}} \frac{\|V^*g\|}{\|g\|} = \|V^*T(T^*T)^{-\frac{1}{2}}\|. \tag{5.40}$$

We observe that $T(T^*T)^{-\frac{1}{2}}$ is an isometry and $\mathcal{R}(T(T^*T)^{-\frac{1}{2}}) = \mathcal{T}$. Therefore

$$\sup_{\substack{g\in\mathcal{T}\\ g\neq 0}} \frac{\|V^*g\|}{\|g\|} = \sup_{\substack{g\in\mathcal{T}\\ \|g\|=1}} \|V^*g\| = \sup_{\substack{c\in l^2(\mathbb{N})\\ \|c\|=1}} \|V^*T(T^*T)^{-\frac{1}{2}}c\| = \|V^*T(T^*T)^{-\frac{1}{2}}\|.$$

□

Lemma 5.2.13 shows that if the reconstruction vectors $\{g_k\}_{k\in\mathbb{N}}$ are an orthonormal system, then $\kappa_{\mathcal{T}}(Q_g)$ is equal to the condition number of the cross-Gramian $\kappa(V^*T)$. From Corollary 5.2.11 we know that the operator Q_g has the smallest relative condition number $\kappa_{\mathcal{T}}(Q_g)$. Therefore, if we use an orthonormal system for reconstruction the condition number of the cross-Gramian $\kappa(V^*T)$ tells us if for elements inside the reconstruction space $\mathcal{T}$ a stable reconstruction is possible.

Lemma 5.2.13. *Let $\mathcal{H}$, $\mathcal{T}$, $\mathcal{V}$, $\{u_j\}_{j\in\mathbb{N}}$, V and Q_g be as in Theorem 5.2.5.*
If $\cos(\varphi_{\mathcal{T},\mathcal{V}}) > 0$ and $\{g_k\}_{k\in\mathbb{N}}$ is an orthonormal system, then

$$\kappa_{\mathcal{T}}(Q_g) = \|(V^*T)^{\dagger}\| \; \|V^*T\|. \tag{5.41}$$

Proof. From (5.38) and the definiton of Q_g we infer that

$$\kappa_{\mathcal{T}}(Q_g) = \|T(V^*T)^{\dagger}\|\|V^*T\| \tag{5.42}$$

Equation (5.41) follows from (5.42) and the fact that $\{g_k\}_{k\in\mathbb{N}}$ is an orthonormal system. □

5.3 Frame independent sampling

In this section we treat the concept of frame independent sampling. In the last section we have seen that the operator Q_g has the minimal operator norm among all reconstructions with the property that $QV^*g = g$ for $g \in \mathcal{T}$, and using this property we have seen in Corollary 5.2.11 that Q_g has the smallest possible relative condition number $\kappa_{\mathcal{T}}(Q_g)$. Therefore Q_g is well suited for reconstructing elements inside the reconstruction space. As mentioned in the introduction, the function to reconstruct typically has a part outside of the reconstruction space. In this section we present the reconstruction operator Q_f, defined by (5.44), which has the smallest possible quasi-optimality constant (see Theorem 5.3.8), and therefore deals very well with the part outside of the reconstruction space. Another important property of the operator Q_f is that it deals very well with systematic errors appearing before the sampling process, see Theorem 5.3.10.

Equation (5.45) points out the connection of the operator Q_f to the oblique projection $P_{\mathcal{T},P_{\mathcal{V}}(\mathcal{T})^\perp}$. It is shown that $Q_fV^* = P_{\mathcal{T},P_{\mathcal{V}}(\mathcal{T})^\perp}$. Since this oblique projection is independent of the particular choice of the frame sequences for the sampling and reconstruction space, but only dependent on their closed linear span $\mathcal{V}$ and $\mathcal{T}$, we refer to the reconstruction obtained by the operator Q_f as *frame independent sampling.*

Lemma 5.3.1. *Let $\mathcal{T}$ and $\mathcal{V}$ be closed subspaces of a Hilbert space $\mathcal{H}$. If* $\cos(\varphi_{\mathcal{T},\mathcal{V}}) > 0$*, then the subspace $P_{\mathcal{V}}(\mathcal{T})$ is closed. Furthermore,*

$$\cos(\varphi_{\mathcal{T},P_{\mathcal{V}}(\mathcal{T})}) = \cos(\varphi_{\mathcal{T},\mathcal{V}}). \tag{5.43}$$

Proof. We recall equation (3.10)

$$\|g\|\cos(\varphi_{\mathcal{T},\mathcal{V}}) \leqslant \|P_{\mathcal{V}}g\| \quad \text{for all } g \in \mathcal{T}.$$

Therefore by Lemma 2.1.7 the operator $(P_{\mathcal{V}})_{|\mathcal{T}}$ has a closed range and thus the subspace $P_{\mathcal{V}}(\mathcal{T})$ is closed. The second statement follows from

$$\begin{aligned}\cos(\varphi_{\mathcal{T},P_{\mathcal{V}}(\mathcal{T})}) &= \inf_{\substack{g\in\mathcal{T}\\ \|g\|=1}} \sup_{\substack{v\in P_{\mathcal{V}}(\mathcal{T})\\ \|v\|=1}} |\langle g,v\rangle| = \inf_{\substack{g\in\mathcal{T}\\ \|g\|=1}} \sup_{\substack{v\in P_{\mathcal{V}}(\mathcal{T})\\ \|v\|=1}} |\langle g,P_{\mathcal{V}}v\rangle| \\ &= \inf_{\substack{g\in\mathcal{T}\\ \|g\|=1}} \sup_{\substack{v\in P_{\mathcal{V}}(\mathcal{T})\\ \|v\|=1}} |\langle P_{\mathcal{V}}g,v\rangle| = \inf_{\substack{g\in\mathcal{T}\\ \|g\|=1}} \|P_{\mathcal{V}}g\| = \cos(\varphi_{\mathcal{T},\mathcal{V}}),\end{aligned}$$

using (3.9) for the first equality and last equality. □

Theorem 5.3.2. *Let $\mathcal{T}$ and $\mathcal{V}$ be closed subspaces of a separable Hilbert space $\mathcal{H}$. Let $\{u_j\}_{j\in\mathbb{N}}$ be a frame for $\mathcal{V}$ and let V be the corresponding synthesis operator and S be the corresponding frame operator. Let $\{g_k\}_{k\in\mathbb{N}}$ be a frame for $\mathcal{T}$ and let T be the corresponding synthesis operator. Furthermore let Q_f be defined by*

$$Q_f := T\left((V^*V)^{\frac{\dagger}{2}}V^*T\right)^{\dagger}(V^*V)^{\frac{\dagger}{2}}. \tag{5.44}$$

If $\cos(\varphi_{\mathcal{T},\mathcal{V}}) > 0$, then $\mathcal{H} = \mathcal{T}\oplus P_{\mathcal{V}}(\mathcal{T})^{\perp}$, the oblique projection $P_{\mathcal{T},P_{\mathcal{V}}(\mathcal{T})^{\perp}}$ is well defined and bounded and the following holds.

- *The operator Q_f is defined by*

$$Q_fV^* = P_{\mathcal{T},P_{\mathcal{V}}(\mathcal{T})^{\perp}}. \tag{5.45}$$

  and

$$Q_{f|\mathcal{R}(V^*)^{\perp}} = 0, \tag{5.46}$$

  and consequently Q_f is independent of the particular choice of the frame $\{g_k\}_{k\in\mathbb{N}}$ for $\mathcal{T}$.

- $$\|Q_{f|\mathcal{R}(V^*)}\| = \|Q_f\|. \tag{5.47}$$

- $$\mu(Q_f) = \|Q_fV^*\| = \|P_{\mathcal{T},P_{\mathcal{V}}(\mathcal{T})^{\perp}}\| = \frac{1}{\cos(\varphi_{\mathcal{T},\mathcal{V}})} \tag{5.48}$$

- *For $d\in l^2(\mathbb{N})$*

$$Q_fd = \sum_{k=1}^{\infty}\hat{c}_kg_k,$$

where $\hat{c} = \{\hat{c}_k\}_{k\in\mathbb{N}}$ *is the minimal norm element of the set*

$$\arg\min_c \|(V^*V)^{\frac{\dagger}{2}}V^*Tc - (V^*V)^{\frac{\dagger}{2}}d\|.$$

Proof. That $\mathcal{H} = \mathcal{T} \oplus P_{\mathcal{V}}(\mathcal{T})^\perp$ and that the oblique projection $P_{\mathcal{T},P_{\mathcal{V}}(\mathcal{T})^\perp}$ is well defined and bounded follows from Lemma 5.3.1 and Theorem 5.2.2.

From Lemma 2.3.18 we know that

$$L^* = (V^*V)^{\frac{\dagger}{2}}V^*. \tag{5.49}$$

is the analysis operator of the tight frame sequence $\{S^{\frac{\dagger}{2}}u_j\}_{j\in\mathbb{N}}$. Using the notation (5.49), we infer that

$$Q_fV^* = T(L^*T)^\dagger L^*.$$

From Lemma 2.3.17 we infer that $LL^* = P_{\mathcal{V}}$. Equation (5.45) follows now from Theorem 5.2.5, (5.21), applied to the frame $\{S^{\frac{\dagger}{2}}u_j\}_{j\in\mathbb{N}}$ instead of $\{u_j\}_{j\in\mathbb{N}}$ and L^* instead of V^*.

In order to prove (5.46), we show that $\mathcal{R}(V^*)^\perp = \mathcal{N}\left((V^*V)^{\frac{\dagger}{2}}\right)$. Using Lemma 2.1.3 and (2.2) we obtain

$$\begin{aligned}\mathcal{N}\left((V^*V)^{\frac{\dagger}{2}}\right) &= \mathcal{N}\left((V^*V)^\dagger\right) = \mathcal{N}\left((V^*V)^*\right) \\ &= \mathcal{N}(V^*V) = \mathcal{R}(V^*V)^\perp = \mathcal{R}(V^*)^\perp.\end{aligned}$$

Equation (5.47) is a direct consequence of (5.46).

By Corollary 3.4.7

$$\mu(Q_f) = \|Q_fV^*\| = \|P_{\mathcal{T},P_{\mathcal{V}}(\mathcal{T})^\perp}\| = \frac{1}{\cos(\varphi_{\mathcal{T},P_{\mathcal{V}}(\mathcal{T})})}.$$

Combining this with $\cos(\varphi_{\mathcal{T},P_{\mathcal{V}}(\mathcal{T})}) = \cos(\varphi_{\mathcal{T},\mathcal{V}})$ (see (5.43)) we deduce (5.48).

The last point can be shown similarly to the last point of Theorem 5.2.5. □

5.3.1 An abstract definition of $P_{\mathcal{T},P_{\mathcal{V}}(\mathcal{T})^{\perp}}$

The oblique projection $P_{\mathcal{T},P_{\mathcal{V}}(\mathcal{T})^{\perp}}$ is characterized as follows.

Theorem 5.3.3. *Let $\cos(\varphi_{\mathcal{T},\mathcal{V}}) > 0$. The mapping $P_{\mathcal{T},P_{\mathcal{V}}(\mathcal{T})^{\perp}}$ is the unique operator $F : \mathcal{H} \to \mathcal{T}$ that satisfies the equations*

$$\langle P_{\mathcal{V}} F f, g_k \rangle = \langle P_{\mathcal{V}} f, g_k \rangle, \quad k \in \mathbb{N}, \ f \in \mathcal{H}. \tag{5.50}$$

Proof. Let $\{u_j\}_{j\in\mathbb{N}}$ be a frame for $\mathcal{V}$ and let V be the corresponding synthesis operator and let S be the corresponding frame operator. From Lemma 2.3.18 we know that $(V^*V)^{\frac{\dagger}{2}} V^*$ is the analysis operator of the tight frame sequence $\{S^{\frac{\dagger}{2}} u_j\}_{j\in\mathbb{N}}$. Since the frame operator of $\{S^{\frac{\dagger}{2}} u_j\}_{j\in\mathbb{N}}$ is $P_{\mathcal{V}}$, Theorem 5.2.7 (stated in [10, Section 4]) applied to the frame $\{S^{\frac{\dagger}{2}} u_j\}_{j\in\mathbb{N}}$ instead of $\{u_j\}_{j\in\mathbb{N}}$ yields (5.50). □

In order to prove Theorem 5.3.5 we need the following well known lemma for distances of affine subspaces which can be found in [56, 5.2.5. Abstand affiner Unterräume].

Lemma 5.3.4. *Let $\mathcal{H}_1 = x_1 + \mathcal{V}_1$ and $\mathcal{H}_2 = x_2 + \mathcal{V}_2$ be affine subspaces of the Hilbert space $\mathcal{H}$, where $x_1, x_2 \in \mathcal{H}$ and $\mathcal{V}_1$ and $\mathcal{V}_2$ are subspaces of $\mathcal{H}$. If $\mathcal{V}_1 + \mathcal{V}_2$ is closed, then there exists exactly one $h \in \mathcal{H}$ with the properties:*

$$h = p_1 - p_2 \quad \text{for some } p_1 \in \mathcal{H}_1, \ p_2 \in \mathcal{H}_2, \tag{5.51}$$

$$h \perp \mathcal{V}_1 \text{ and } h \perp \mathcal{V}_2. \tag{5.52}$$

Furthermore the distance between $\mathcal{H}_1$ and $\mathcal{H}_2$ is given by

$$\operatorname{dist}(\mathcal{H}_1, \mathcal{H}_2) = \inf\{\|q_1 - q_2\| : q_1 \in \mathcal{H}_1 \text{ and } q_2 \in \mathcal{H}_2\} = \|h\|. \tag{5.53}$$

Proof. It is straightforward to show that

$$h := (I - P_{\mathcal{V}_1+\mathcal{V}_2})(x_1 - x_2) \in (\mathcal{V}_1 + \mathcal{V}_2)^{\perp}$$

has the properties (5.51), (5.52), and (5.53). To show the uniqueness, let

$$\begin{aligned} h &= p_1 - p_2, \quad p_1 \in \mathcal{H}_1,\ p_2 \in \mathcal{H}_2,\ h \perp \mathcal{V}_1, \mathcal{V}_2 \quad \text{and} \\ k &= q_1 - q_2, \quad q_1 \in \mathcal{H}_1,\ q_2 \in \mathcal{H}_2,\ k \perp \mathcal{V}_1, \mathcal{V}_2. \end{aligned}$$

From $h - k = (p_1 - q_1) - (p_2 - q_2)$ it follows that $h - k \in \mathcal{V}_1 + \mathcal{V}_2$. From $h \perp \mathcal{V}_1, \mathcal{V}_2$ and $k \perp \mathcal{V}_1, \mathcal{V}_2$, we know that $h - k \in (\mathcal{V}_1 + \mathcal{V}_2)^\perp$, which implies that $h - k = 0$. □

The following lemma shows that the mapping $P_{\mathcal{T},P_\mathcal{V}(\mathcal{T})^\perp}$ has the property, that for every $f \in \mathcal{H}$, $P_{\mathcal{T},P_\mathcal{V}(\mathcal{T})^\perp} f$ is the element in $\mathcal{T}$ closest to $P_\mathcal{V} f + \mathcal{V}^\perp$. This shows that $P_{\mathcal{T},P_\mathcal{V}(\mathcal{T})^\perp}$ fulfills the geometric properties described in the introduction.

Theorem 5.3.5. *Let $\mathcal{T}$ and $\mathcal{V}$ be closed subspaces of a Hilbert space $\mathcal{H}$ and let* $\cos(\varphi_{\mathcal{T},\mathcal{V}}) > 0$.

For every $f \in \mathcal{H}$

$$\operatorname{dist}(P_\mathcal{V} f + \mathcal{V}^\perp, \mathcal{T}) = \operatorname{dist}(P_\mathcal{V} f + \mathcal{V}^\perp, P_{\mathcal{T},P_\mathcal{V}(\mathcal{T})^\perp} f).$$

Proof. From Lemma 3.4.6 we infer that $\mathcal{T} + \mathcal{V}^\perp$ is closed in $\mathcal{H}$. We set $F := P_{\mathcal{T},P_\mathcal{V}(\mathcal{T})^\perp}$ and

$$h := P_\mathcal{V} f - P_\mathcal{V} F f.$$

We show that h fulfills (5.51) and (5.52). Writing h in the form

$$h = [P_\mathcal{V} f + (Ff - P_\mathcal{V} F f)] - Ff,$$

we see that h is of the form $h = p_1 - Ff$, with $p_1 \in P_\mathcal{V} f + \mathcal{V}^\perp$ and $p_2 = Ff \in \mathcal{T}$. It remains to prove that $h \perp \mathcal{V}^\perp$ and $h \perp \mathcal{T}$. The orthogonality $P_\mathcal{V} f - P_\mathcal{V} F f \perp \mathcal{T}$ follows from (5.50), and $h \perp \mathcal{V}^\perp$ follows from the definition of h. □

5.3.2 Illustration of the difference between generalized and frame independent sampling

Let us now consider a simple example in $\mathcal{H} = \mathbb{R}^2$ which illustrates the difference between the reconstruction operators Q_g (generalized sampling) and Q_f (frame independent sampling). Let $u_1, u_2 \in \mathbb{R}^2$ be two linearly independent vectors, and let $g \in \mathbb{R}^2$ be an arbitrary vector. With this choice, the reconstruction space $\mathcal{T} = \operatorname{span}(g)$ is the line spanned by the vector g. The sampling space is $\operatorname{span}(u_1, u_2) = \mathbb{R}^2$. From (1.7) we infer that $Q_f V^* = P_{\mathcal{T}}$, where $P_{\mathcal{T}}$ is the orthogonal projection onto $\mathcal{T}$. Both reconstruction operators $Q = Q_g$ (generalized sampling) and $Q = Q_f$ (frame independent sampling) fulfill $Q_{|\mathcal{R}(V^*)^\perp} = 0$, see Theorem 5.2.5 and Theorem 5.3.2, and therefore for $p \in \mathcal{H}$

$$\begin{aligned}\|Q\| = \|Q_{|\mathcal{R}(V^*)}\| &= \sup_{\|V^*f\|=1} \|QV^*f\| = \sup_{\|V^*f\|=1} \|QV^*p - Q_f V^*(p-f)\| \\ &= \sup_{x\in E} \|QV^*p - QV^*x\|,\end{aligned}$$

where E is the ellipse

$$E = \{x \in \mathbb{R}^2 : \|V^*p - V^*x\| \leqslant 1\}. \tag{5.54}$$

Since p is the center of the ellipse, for a projection operator P onto $\mathcal{T}$, $P(p)$ is the center of the set $P(E)$. This shows that half of the length of $Q_f V^*(E) = P_{\mathcal{T}}(E)$ is $\|Q_f\|$ and that half of the length of $Q_g V^*(E) = P_{\mathcal{T},S(\mathcal{T})^\perp}(E)$ is $\|Q_g\|$, see Figure 5.1.

For our figures we use

$$u_1 = (0,1),\ u_2 = \frac{(\frac{1}{2},1)}{\|(\frac{1}{2},1)\|},\ g = (1,0). \tag{5.55}$$

The ellipse in Figure 5.1 is the boundary of the set E defined by (5.54) for $p = (5,3)$. As we see in Figure 5.1, the length of $P_{\mathcal{T}}(E)$ is greater than the length of $P_{\mathcal{T},S(\mathcal{T})^\perp}(E)$, which shows that $\|Q_f\| > \|Q_g\|$.

The quasi-optimality constant of Q_f is

$$\mu(Q_f) = \cos(\varphi_{\mathcal{T},\mathcal{V}}))^{-1} = \cos(0)^{-1} = 1.$$

The quasi-optimality constant of Q_g is $\mu(Q_g) = \cos(\varphi_{\mathcal{T},S(\mathcal{T})}))^{-1}$. The angle $\alpha := \varphi_{\mathcal{T}S(\mathcal{T})}$ is plotted in Figure 5.1. Clearly, $\cos(\varphi_{\mathcal{T},S(\mathcal{T})}) < 1$, which shows that $\mu(Q_f) < \mu(Q_g)$. Therefore $Q_f V^*$ is closer to the orthogonal projection $P_\mathcal{T}$ than $Q_g V^*$ (in fact $Q_f V^*$ coincides with the orthogonal projection).

If we approximate $p \in \mathcal{H}$ from perturbed measurements V^*p+c with measurement errors satisfying $\|c\| \leqslant 1$, then the approximations $\tilde{p}$ to p calculated by Q ($Q = Q_f$ or $Q = Q_g$) are located in the set

$$\begin{aligned}\{Q(V^*p + c) : \|c\| \leqslant 1\} &= \{Q(V^*p + P_{\mathcal{R}(V^*)}c) : \|c\| \leqslant 1\} \\ &= \{Q(V^*p + V^*\Delta p) : \|V^*\Delta p\| \leqslant 1\} \\ &= \{QV^*x : x \in E\}.\end{aligned} \tag{5.56}$$

Therefore if the measurement errors are satisfying $\|c\| \leqslant 1$, then the approximations calculated by the operator Q_f are located in the set $P_\mathcal{T}(E)$, and the approximations calculated by Q_g are located in the set $P_{\mathcal{T},S(\mathcal{T})^\perp}(E)$.

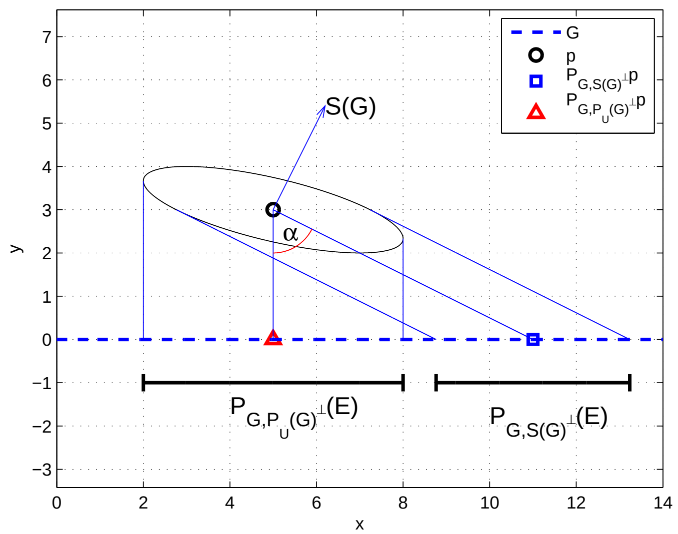

Figure 5.1: $P_{\mathcal{T},P_\mathcal{V}(\mathcal{T})^\perp}$ and $P_{\mathcal{T},S(\mathcal{T})^\perp}$ for $p = (5,3)$

We see in Figure 5.1 that for $p = (5, 3)$ any point in $P_{\mathcal{T}}(E)$ is closer to p than any point in $P_{\mathcal{T},S(\mathcal{T})^\perp}(E)$. Therefore for this position of p and this magnitude of noise, the operator Q_f (frame independent sampling) is preferable over Q_g (generalized sampling).

This changes when the element to reconstruct is closer the reconstruction space. We compare the two reconstructions for the same setup, with the only difference that we choose $p = (5, 0)$, i.e., inside the reconstruction space. Again the approximations calculated by the operator Q_f are located in the

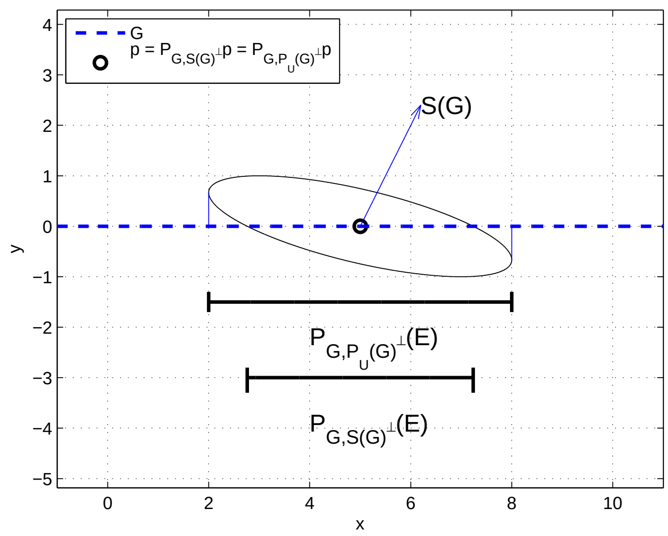

Figure 5.2: $P_{\mathcal{T},P_{\mathcal{V}}(\mathcal{T})^\perp}$ and $P_{\mathcal{T},S(\mathcal{T})^\perp}$ for $p = (5, 0)$

set $P_{\mathcal{T}}(E)$, and the approximations calculated by Q_g are located in the set $P_{\mathcal{T},S(\mathcal{T})^\perp}(E)$. We observe that the set $P_{\mathcal{T},S(\mathcal{T})^\perp}(E)$ is smaller than $P_{\mathcal{T}}(E)$, and that the centers of both sets lie at p. Therefore

$$\max_{\|c\|\leqslant 1} \|p - Q_f(V^*p + c)\| > \max_{\|c\|\leqslant 1} \|p - Q_g(V^*p + c)\|,$$

and consequently the operator Q_g is preferable.

If we are interested in reconstructing from measurements $QV^*(p + \Delta p)$ perturbed before the sampling process, according to Theorem 5.3.10, we expect to obtain more accurate reconstructions by the operator Q_f than by the operator Q_g independently of the position of the element to reconstruct. We illustrate this by taking the same setup as before, see (5.55). We first reconstruct the point $p = (5, 3)$ from the measurements $V^*(p + \Delta p)$, with $\|\Delta p\| \leqslant 1$. The circle in Figure 5.3 is the boundary of the set

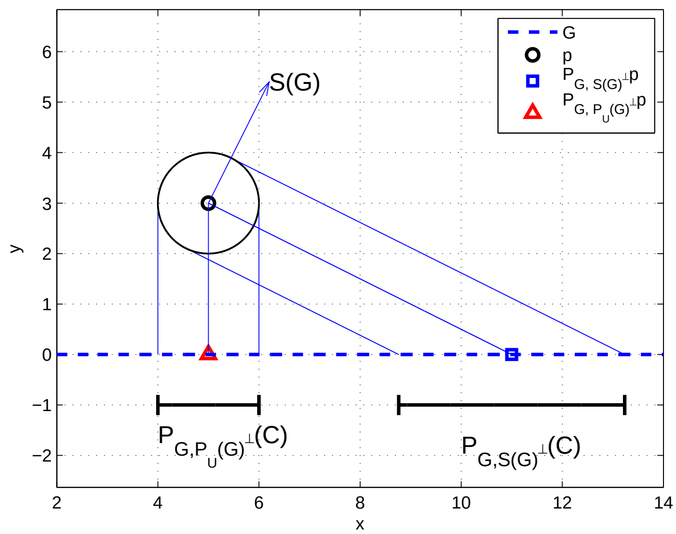

Figure 5.3: $P_{\mathcal{T},P_{\mathcal{V}}(\mathcal{T})^{\perp}}$ and $P_{\mathcal{T},S(\mathcal{T})^{\perp}}$ for $p = (5, 3)$

$$C := \{p + \Delta p : \|\Delta p\| \leqslant 1\}. \tag{5.57}$$

For every point $c \in C$, the approximation $\tilde{p} = Q_f V^* c$ is located in the set $P_{\mathcal{T},P_{\mathcal{V}}(\mathcal{T})^{\perp}}(C) = P_{\mathcal{T}}(C)$ and $\tilde{p} = Q_g V^* c$ is located in $P_{\mathcal{T},S(\mathcal{T})^{\perp}}(C)$. Since any point in $P_{\mathcal{T}}(C)$ is closer to p than any point in $P_{\mathcal{T},S(\mathcal{T})^{\perp}}(C)$, the operator Q_f is preferable over the operator Q_g.

In Figure 5.4, we reconstruct the point $p = (5, 0)$. The set $P_{\mathcal{T}}(E)$ is smaller than $P_{\mathcal{T},S(\mathcal{T})^{\perp}}(C)$, and the centers of both sets lie at p. Therefore

$$\max_{\|c\| \leqslant 1} \|p - Q_g V^*(p + \Delta p)\| > \max_{\|c\| \leqslant 1} \|p - Q_f V^*(p + \Delta p)\|,$$

and consequently the operator Q_f is preferable also for points inside the reconstruction space $\mathcal{T}$.

Figures 5.1, 5.2, 5.3 and 5.4 can be interpreted as pictures in $\mathbb{R}^3$ by choosing

$$u_1 = (0,1,0), \; u_2 = \frac{(\frac{1}{2},1)}{\|(\frac{1}{2},1,0)\|}, \; g = (1,0,z). \tag{5.58}$$

for $z \in \mathbb{R}$. With this choice, the set E is an elliptic cylinder, and C is a ball with center p and radius 1 and the figures dipict the orthogonal projection onto the plane spanned by the first two coordinates.

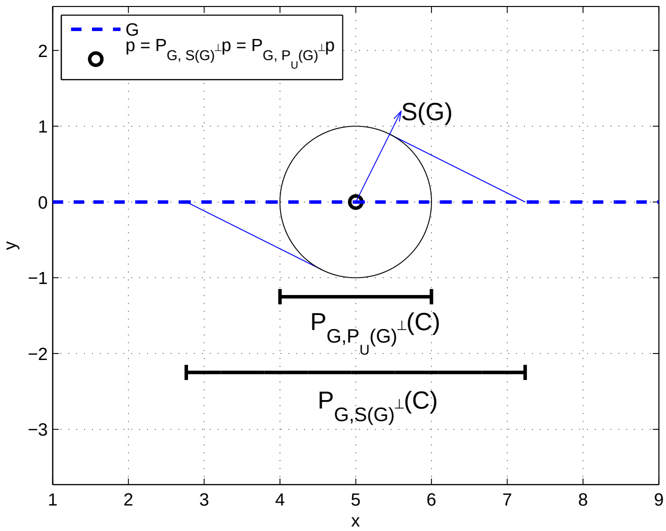

Figure 5.4: $P_{\mathcal{T},P_{\mathcal{V}}(\mathcal{T})^\perp}$ and $P_{\mathcal{T},S(\mathcal{T})^\perp}$ for $p=(5,0)$

5.3.3 Stability, Quasi-optimality constant and error estimates

It can be said that the operator Q_f relies more on the measurements V^*f, whereas the operator Q_g trusts more in the function to be reconstructed being close to the reconstruction space. Therefore, if the sampling frame is ill conditioned (a large ratio $\frac{B}{A}$ of the upper and lower frame bound), the operator Q_f can be very sensitive to noise. This can be seen in Theorem 5.3.7 from the term $\frac{\sqrt{B}}{\sqrt{A}} \frac{\|c\|}{\|V^*f\|}$, where the relative coefficient error is multiplied by the square root of the quotient of the upper and lower frame bound of the sampling frame sequence. Conversely, it can happen that the operator Q_g is very sensitive to the part outside of the reconstruction space of the function to reconstruct, whereas $Q_f V^*$ projects onto the reconstruction space almost orthogonally.

In order to prove Theorem 5.3.7 we need Lemma 5.3.6. The proof of Lemma 5.3.6 is similar to the proof of [10, Corollary 4.7].

Lemma 5.3.6. *Let $\mathcal{H}$, $\mathcal{T}$, $\mathcal{V}$, $\{u_j\}_{j\in\mathbb{N}}$, V, $\{g_k\}_{k\in\mathbb{N}}$ and T be as in Theorem 5.3.2. Let A and B denote the lower and the upper frame bound of $\{u_j\}_{j\in\mathbb{N}}$, and let Q_f be defined by* (5.44).

If $\cos(\varphi_{\mathcal{T},\mathcal{V}}) > 0$, *then*

$$\frac{1}{\sqrt{B}} \leqslant \|Q_f\| \leqslant \frac{1}{\sqrt{A}\cos(\varphi_{\mathcal{T},\mathcal{V}})}. \tag{5.59}$$

Proof. Since $\|Q_{f|\mathcal{R}(V^*)}\| = \|Q_f\|$, it suffices to prove (5.59) for $Q_{f|\mathcal{R}(V^*)}$ instead of Q_f. From the definition of $\cos(\varphi_{\mathcal{T},\mathcal{V}})$ we know that

$$\|g\|\cos(\varphi_{\mathcal{T},\mathcal{V}}) \leqslant \|P_{\mathcal{V}} g\| \quad \text{for every } g \in \mathcal{T}. \tag{5.60}$$

We define the operator F as

$$F := P_{\mathcal{T},P_{\mathcal{V}}(\mathcal{T})^\perp} = Q_f V^*.$$

From the Cauchy-Schwarz inequality, and (5.50), it follows that for $f \in \mathcal{H}$ it holds

$$\langle P_{\mathcal{V}} F f, F f\rangle = \langle P_{\mathcal{V}} f, F f\rangle \leqslant \langle P_{\mathcal{V}} f, f\rangle^{\frac{1}{2}} \langle P_{\mathcal{V}} F f, F f\rangle^{\frac{1}{2}}.$$

This yields

$$\|P_{\mathcal{V}} F f\| \leqslant \|P_{\mathcal{V}} f\|. \tag{5.61}$$

From the frame inequality (3.3), it follows that for $f \in H$ it holds

$$\sqrt{A}\, \|P_{\mathcal{V}} f\| \leqslant \|V^* f\| \leqslant \sqrt{B}\, \|P_{\mathcal{V}} f\|. \tag{5.62}$$

We combine (5.60),(5.61) and (5.62) and obtain

$$\|Ff\|\cos(\varphi_{\mathcal{T},\mathcal{V}}) \leqslant \|P_{\mathcal{V}}Ff\| \leqslant \|P_{\mathcal{V}}f\| \leqslant \frac{1}{\sqrt{A}}\ \|V^*f\|. \tag{5.63}$$

Equation (5.63) implies that

$$\|Q_{f|\mathcal{R}(V^*)}\| \leqslant \frac{1}{\sqrt{A}\cos(\varphi_{\mathcal{T},\mathcal{V}})}.$$

The lower bound of (5.59) follows from

$$\frac{1}{\sqrt{B}}\|V^*Ff\ \| \leqslant \|P_{\mathcal{V}}Ff\| \leqslant \|Ff\|,$$

where we use (5.62) for the first inequality. □

Theorem 5.3.7. *Let $\mathcal{H}$, $\mathcal{T}$, $\mathcal{V}$, $\{u_j\}_{j\in\mathbb{N}}$, V and Q_f be as in Theorem 5.3.2. Let A and B denote the lower and the upper frame bound of $\{u_j\}_{j\in\mathbb{N}}$. If* $\cos(\varphi_{\mathcal{T},\mathcal{V}}) > 0$ *then for $f \in \mathcal{H}$*

$$\frac{\|f - Q_f(V^*f + c)\|}{\|f\|} \leqslant \frac{1}{\cos(\varphi_{\mathcal{T},\mathcal{V}})}\left(\frac{\|f - P_{\mathcal{T}}f\|}{\|f\|} + \sqrt{\frac{B}{A}}\ \frac{\|c\|}{\|V^*f\|}\right). \tag{5.64}$$

Proof. We recall equation (3.7)

$$\frac{\|f - Q(V^*f + c)\|}{\|f\|} \leqslant \mu\frac{\|f - P_{\mathcal{T}}f\|}{\|f\|} + \|Q\|\sqrt{B}\frac{\|c\|}{\|V^*f\|}.$$

Combining this with $\mu(Q_f) = \frac{1}{\cos(\varphi_{\mathcal{T},\mathcal{V}})}$, see (5.48), and $\|Q_f\| \leqslant \frac{1}{\sqrt{A}\cos(\varphi_{\mathcal{T},\mathcal{V}})}$, see (5.59), we obtain (5.64). □

Theorem 5.3.8 shows that the operator Q_f has the smallest possible quasi-optimality constant. Therefore Q_fV^* projects as orthogonally as possible onto the reconstruction space.

Theorem 5.3.8. *Let $\mathcal{H}$, $\mathcal{T}$, $\mathcal{V}$, $\{u_j\}_{j\in\mathbb{N}}$, V and Q_f be as in Theorem 5.3.2. If* $\cos(\varphi_{\mathcal{T},\mathcal{V}}) > 0$ *then for any operator $Q : l^2(\mathbb{N}) \to \mathcal{H}$ it holds*

$$\mu(Q) \geqslant \mu(Q_f).$$

Proof. We recall that $\mu(Q)$ is the smallest α such that for every $f \in \mathcal{H}$

$$\|f - QV^* f\| \leqslant \alpha \|f - P_{\mathcal{T}} f\|. \tag{5.65}$$

Let $\mu(Q) < \infty$ and let $g \in \mathcal{T}$. Every element $f \in g + \mathcal{V}^\perp$ has the same value $V^* g$, and thus

$$QV^* f = QV^* g \quad \text{for all } f \in g + \mathcal{V}^\perp. \tag{5.66}$$

Assumption (5.65), implies that

$$QV^* g = g \qquad \text{for all } g \in \mathcal{T}, \tag{5.67}$$

since otherwise $\mu(Q) = \infty$. Therefore

$$QV^*(g + u^\perp) = g \quad \text{for all } g \in \mathcal{T} \text{ and } u^\perp \in \mathcal{V}^\perp.$$

This means that $(QV^*)_{|\mathcal{T} \oplus \mathcal{V}^\perp} = P_{\mathcal{T},\mathcal{V}^\perp}$. From (5.8) it follows that on this subspace $\mu(Q) = \mu(Q_c) = \frac{1}{\cos(\varphi_{\mathcal{T},\mathcal{V}})}$. This implies that for $f \in \mathcal{T} \oplus \mathcal{V}^\perp$ we have the sharp upper bound

$$\|f - QV^* f\| \leqslant \frac{1}{\cos(\varphi_{\mathcal{T},\mathcal{V}})} \|f - P_{\mathcal{T}} f\|.$$

which implies that $\alpha \geqslant \frac{1}{\cos(\varphi_{\mathcal{T},\mathcal{V}})}$. Since by (5.48) it holds $\mu(Q_f) = \frac{1}{\cos(\varphi_{\mathcal{T},\mathcal{V}})}$, this finishes the proof. $\square$

5.3.4 Error appearing before the sampling process

In [2–5] the stability to measurement error appearing before the sampling process is considered. The authors inspect the following two properties of a mapping $Q : l^2(\mathbb{N}) \to \mathcal{H}$. The first one is again the quasi-optimality constant $\mu(Q)$.

The measure of stability they use is the operator norm $\|QV^*\|$.

We obtain the error estimate

$$\|f - QV^*(f + \Delta f)\| \leqslant \mu(Q)\|f - P_{\mathcal{T}} f\| + \|QV^*\| \|\Delta f\|. \tag{5.68}$$

It is very important to realize a fundamental difference between (5.68) and

$$\|f - Q(V^*f + c)\| \leqslant \mu(Q)\|f - P_{\mathcal{T}}f\| + \|Q\|\|c\|, \tag{5.69}$$

which is considered in Lemma 3.3.1. In (5.69), the reconstruction error from perturbed measurements $V^*f + c$ is estimated.

By contrast, for (5.68) it is assumed, that the measurements of the perturbed function $f + \Delta f$ are available, and the corresponding reconstruction error is estimated. This means exact measurements of the perturbed function $f + \Delta f$ are available.

Corollary 5.3.9. *Let $\mathcal{H}$, $\mathcal{T}$, $\mathcal{V}$, $\{u_j\}_{j\in\mathbb{N}}$, V and Q_f be as in Theorem 5.3.2. Furthermore let* $\cos(\varphi_{\mathcal{T},\mathcal{V}}) > 0$.

If $Q : l^2(\mathbb{N}) \to \mathcal{H}$ is a bounded operator with $QV^*g = g$ for all $g \in \mathcal{T}$, then

$$\|QV^*\| \geqslant \|Q_fV^*\|.$$

Proof. The assumption $QV^*g = g$ implies that $(QV^*)_{|\mathcal{T}\oplus\mathcal{V}^\perp} = P_{\mathcal{T},\mathcal{V}^\perp}$, see the proof of Theorem 5.3.8. From Lemma 3.4.6, (3.), it follows that on this subspace $\|(QV^*)_{|\mathcal{T}\oplus\mathcal{V}^\perp}\| = \frac{1}{\cos(\varphi_{\mathcal{T},\mathcal{V}})}$, which implies that $\|QV^*\| \geqslant \frac{1}{\cos(\varphi_{\mathcal{T},\mathcal{V}})}$. Using (5.48) we obtain

$$\|Q_fV^*\| = \frac{1}{\cos(\varphi_{\mathcal{T},\mathcal{V}})} \leqslant \|QV^*\|$$

which completes the proof. □

Theorem 5.3.10 shows that the operator Q_f is optimal (in the following sense) for the problem consisdered in [2–5].

Theorem 5.3.10. *Let $\mathcal{H}$, $\mathcal{T}$, $\mathcal{V}$, $\{u_j\}_{j\in\mathbb{N}}$, V and Q_f be as in Theorem 5.3.2. Furthermore let* $\cos(\varphi_{\mathcal{T},\mathcal{V}}) > 0$.

If an operator $Q : l^2(\mathbb{N}) \to \mathcal{H}$ satisfies for $h \in \mathcal{H}$ and $\Delta h \in \mathcal{H}$

$$\|h - QV^*(h + \Delta h)\| \leqslant \beta_1\|h - P_{\mathcal{T}}h\| + \beta_2\|\Delta h\| \tag{5.70}$$

for some $0 < \beta_i < \infty$, then $\beta_i \geqslant \mu(Q_f)$, $i = 1, 2$.

Proof. Let $Q : l^2(\mathbb{N}) \to \mathcal{H}$ be an operator that satisfies (5.70) for some $0 < \beta_i < \infty$, $i = 1, 2$. Setting $\Delta h = 0$ in (5.70), Theorem 5.3.8 implies that $\beta_1 \geqslant \mu$.

Setting $\Delta h = 0$ in (5.70), it follows that $QV^*g = g$ for $g \in \mathcal{T}$. Otherwise $\beta_1 = \infty$. Setting $h = 0$ in (5.70), Corollary 5.3.9 implies that $\beta_2 \geqslant \mu$. □

In Chapter 7 we need the following slightly modified version of Theorem 5.3.10.

Lemma 5.3.11. *Let $\mathcal{H}$, $\mathcal{T}$, $\mathcal{V}$, $\{u_j\}_{j\in\mathbb{N}}$, V and Q_f be as in Theorem 5.3.2. Furthermore let $\cos(\varphi_{\mathcal{T},\mathcal{V}}) > 0$ and let $\mu = \mu(Q_f)$.*

For every $h \in \mathcal{H}$ and $\Delta h \in \mathcal{H}$ it holds

$$\|P_{\mathcal{T}}h - Q_f V^*(h + \Delta h)\| \leqslant \sqrt{\mu^2 - 1}\|h - P_{\mathcal{T}}h\| + \mu\|\Delta h\|. \tag{5.71}$$

If an operator $Q : l^2(\mathbb{N}) \to \mathcal{H}$ satisfies

$$\|P_{\mathcal{T}}h - QV^*(h + \Delta h)\| \leqslant \alpha\|h - P_{\mathcal{T}}h\| + \beta\|\Delta h\|.$$

for some $0 < \alpha, \beta < \infty$, then $\alpha \geqslant \sqrt{\mu^2 - 1}$ and $\beta \geqslant \mu$.

Proof. Since $Q_f V^*h \in \mathcal{T}$ for $h \in \mathcal{H}$, by the Pythagorean theorem

$$\|P_{\mathcal{T}}h - Q_f V^*h\| = \sqrt{\|h - Q_f V^*h\|^2 - \|h - P_{\mathcal{T}}h\|^2}. \tag{5.72}$$

Using (5.72), (5.48), the definition of the quasi-optimality constant and the triangle-inequality we obtain for $h \in \mathcal{H}$ and $\Delta h \in \mathcal{H}$

$$\begin{aligned}\|P_{\mathcal{T}}h - Q_f(V^*h + V^*\Delta h)\| &\leqslant \|P_{\mathcal{T}}h - Q_f V^*h\| + \|Q_f V^*\Delta h\| \\ &= \sqrt{\|h - Q_f V^*h\|^2 - \|h - P_{\mathcal{T}}h\|^2} + \|Q_f V^*\Delta h\| \\ &\leqslant \sqrt{\mu^2 - 1}\|h - P_{\mathcal{T}}h\| + \mu\|\Delta h\|.\end{aligned}$$

The second part of the Lemma follows as in Theorem 5.3.10 by setting $h = 0$ and $\Delta h = 0$. □

5.4 A combination of generalized and frame independent sampling

The operator Q_g has the smallest possible operator norm among all operators Q with the property that $QV^*g = Qg$ for $g \in \mathcal{T}$. Consequently, Q_g is the most stable operator among those operators. The operator Q_f has the smallest possible quasi-optimality constant, and therefore deals in this sense optimally with the part of the function to reconstruct lying outside of the reconstruction space. Naturally, there arises the question how to obtain mixtures between the two operators Q_f and Q_g, i.e., how to obtain operators Q_m with $\|Q_g\| < \|Q_m\| < \|Q_f\|$ and $\mu(Q_f) < \mu(Q_m) < \mu(Q_g)$. We see in Corollary 5.5.10 that the operator Q_g corresponds to an unbiased estimator of minimal variance for the reconstruction of an element in $g \in \mathcal{T}$ from measurements perturbed by a random vector ε with $\mathcal{E}(\varepsilon) = 0$ and $\Sigma = \mathrm{Cov}(\varepsilon) = \sigma_1^2 I_n$. We see in Corollary 5.5.11 that the operator Q_f corresponds to an unbiased estimator of minimal variance for the reconstruction of an element $g \in \mathcal{T}$ from measurements of the perturbed object (noise appearing before the sampling process). By determining an unbiased estimator of minimal variance for combinations of the two versions of noise, we obtain mixtures Q_m between the two operators. In this section we analyze these operators Q_m. We observe in numerical experiments in Chapter 6, that if we increase the standard deviation of the noise appearing before the sampling process (leaving the standard deviation of the noise appearing after the sampling process constant), we obtain a smaller quasi-optimality constant of the reconstruction at the cost of an increased operator norm. A proof for this statement is yet missing.

We set

$$\tilde{u}_j := ({\sigma_1}^2 I + {\sigma_2}^2 VV^*)^{\frac{\dagger}{2}} u_j, \tag{5.73}$$

where I denotes the identity operator on $\mathcal{H}$. Lemma 5.4.1 proves the following. Suppose that we are given the inner products $\{\langle f, u_j\rangle\}_{j\in\mathbb{N}}$ of an element $f \in \mathcal{H}$ with a frame $\{u_j\}_{j\in\mathbb{N}}$ for $\mathcal{V}$ (a closed subspace of $\mathcal{H}$). Applying the operator $({\sigma_1}^2 I + {\sigma_2}^2 V^*V)^{\frac{\dagger}{2}}$ to these measurements, we obtain the inner products of f with the frame $\{\tilde{u}_j\}_{j\in\mathbb{N}}$ for $\mathcal{V}$. Furthermore, Lemma 5.4.1 shows that by increasing ${\sigma_2}^2$ and leaving ${\sigma_1}^2$ constant, the frame $\{\tilde{u}_j\}_{j\in\mathbb{N}}$ for $\mathcal{V}$ becomes "tighter", i.e., the ratio $\frac{B}{A}$ of the upper and lower frame bound becomes smaller.

The case ${\sigma_1}^2 = 0$ of Lemma 5.4.1 is treated in Lemma 2.3.18. The case ${\sigma_1}^2 = 0$ of Theorem 5.4.2 is treated in Theorem 5.3.2. Therefore we we may assume that ${\sigma_1}^2 > 0$, in order to use the simpler notation $({\sigma_1}^2 I + {\sigma_2}^2 V^*V)^{-\frac{1}{2}}$ instead of $({\sigma_1}^2 I + {\sigma_2}^2 V^*V)^{\frac{\dagger}{2}}$.

Lemma 5.4.1. *Let $\mathcal{H}$, $\mathcal{T}$, $\mathcal{V}$, $\{u_j\}_{j\in\mathbb{N}}$ and V be as in Theorem 5.3.2. Furthermore, let $\sigma_1^2 > 0$, let $\sigma_2^2 \geqslant 0$, and let $\{\tilde{u}_j\}_{j\in\mathbb{N}}$ be defined by* (5.73).

If A and B denote the lower and the upper frame bound of $\{u_j\}_{j\in\mathbb{N}}$, then $\{\tilde{u}_j\}_{j\in\mathbb{N}}$ is a frame for $\mathcal{V}$ with lower and upper frame bound $\frac{A}{{\sigma_1}^2+{\sigma_2}^2 A}$ and $\frac{B}{{\sigma_1}^2+{\sigma_2}^2 B}$, respectively, i.e., for every $f \in \mathcal{V}$

$$\frac{A}{{\sigma_1}^2 + {\sigma_2}^2 A}\|f\|^2 \leqslant \sum_{j\in\mathbb{N}} |\langle f, \tilde{u}_j\rangle|^2 \leqslant \frac{B}{{\sigma_1}^2 + {\sigma_2}^2 B}\|f\|^2. \tag{5.74}$$

Furthermore

$$({\sigma_1}^2 I + {\sigma_2}^2 V^*V)^{-\frac{1}{2}} V^* = V^*({\sigma_1}^2 I + {\sigma_2}^2 VV^*)^{-\frac{1}{2}}, \tag{5.75}$$

i.e., the operator $\Sigma^{-\frac{1}{2}}V^*$ is the analysis operator of the frame $\{\tilde{u}_j\}_{j\in\mathbb{N}}$ for $\mathcal{V}$.

Proof. We observe that for $S = VV^*$ and $f \in \mathcal{V}$

$$\begin{aligned}
\sum_{j\in\mathbb{N}} |\langle f, \tilde{u}_j\rangle|^2 &= \sum_{j\in\mathbb{N}} |\langle f, ({\sigma_1}^2 I_n + {\sigma_2}^2 VV^*)^{-\frac{1}{2}} u_j\rangle|^2 \\
&= \sum_{j\in\mathbb{N}} |\langle ({\sigma_1}^2 I_n + {\sigma_2}^2 VV^*)^{-\frac{1}{2}} f, u_j\rangle|^2 \\
&= \|V^*({\sigma_1}^2 I_n + {\sigma_2}^2 VV^*)^{-\frac{1}{2}} f\|^2 \\
&= \langle V^*({\sigma_1}^2 I + {\sigma_2}^2 S)^{-\frac{1}{2}} f, V^*({\sigma_1}^2 I + {\sigma_2}^2 S)^{-\frac{1}{2}} f\rangle \\
&= \langle ({\sigma_1}^2 I + {\sigma_2}^2 S)^{-\frac{1}{2}} S ({\sigma_1}^2 I + {\sigma_2}^2 S)^{-\frac{1}{2}} f, f\rangle \\
&= \langle S({\sigma_1}^2 I + {\sigma_2}^2 S)^{-1} f, f\rangle.
\end{aligned} \tag{5.76}$$

From the frame equation $A\|f\|^2 \leqslant \langle Sf, f\rangle \leqslant B\|f\|^2$ we infer that on $\mathcal{V}$ $\min(\sigma(S)) = A$ and $\max(\sigma(S)) = B$ (assuming that the frame bounds A and B are optimal). Consequently, by the continuous functional calculus (Theorem 2.1.6), on $\mathcal{V}$

$$\min\left(\sigma\left(S({\sigma_1}^2 I + {\sigma_2}^2 S)^{-1}\right)\right) = \frac{A}{{\sigma_1}^2 + {\sigma_2}^2 A},$$

and

$$\max\left(\sigma\left(S({\sigma_1}^2 I + {\sigma_2}^2 S)^{-1}\right)\right) = \frac{B}{{\sigma_1}^2 + {\sigma_2}^2 B}.$$

Combining this with (5.76) we obtain (5.74)

The proof of (5.75) is similar to that of Lemma 2.3.18. □

Theorem 5.4.2. *Let $\mathcal{T}$ and $\mathcal{V}$ be closed subspaces of a separable Hilbert space $\mathcal{H}$. Let $\{u_j\}_{j\in\mathbb{N}}$ be a frame for $\mathcal{V}$, and let V be the corresponding synthesis operator. Let $\{g_k\}_{k\in\mathbb{N}}$ be a frame for $\mathcal{T}$, and let T be the corresponding synthesis operator. Furthermore, let $\sigma_1^2 > 0$, $\sigma_2^2 \geqslant 0$, and let*

$$\Sigma = {\sigma_1}^2 I + {\sigma_2}^2 V^* V.$$

Let L denote the synthesis operator of the frame $\{({\sigma_1}^2 I + {\sigma_2}^2 V V^*)^{\frac{\dagger}{2}} u_j\}_{j\in\mathbb{N}}$ for $\mathcal{V}$, L^* the corresponding analysis operator, and $S = LL^*$ the corresponding frame operator. Furthermore let Q_m be defined by

$$Q_m := T(\Sigma^{-\frac{1}{2}} V^* T)^\dagger \Sigma^{-\frac{1}{2}}. \tag{5.77}$$

If $\cos(\varphi_{\mathcal{T},\mathcal{V}}) > 0$, then $\mathcal{H} = \mathcal{T} \oplus S(\mathcal{T})^\perp$, the oblique projection $P_{\mathcal{T},P_{\mathcal{V}}(\mathcal{T})^\perp}$ is well defined and bounded and the following holds.

- *The operator Q_m is defined by*

$$Q_m V^* = P_{\mathcal{T},S(\mathcal{T})^\perp}. \tag{5.78}$$

and

$$Q_{m|\mathcal{R}(V^*)^\perp} = 0, \tag{5.79}$$

and consequently Q_m is independent of the particular choice of the frame $\{g_k\}_{k\in\mathbb{N}}$ for $\mathcal{T}$.

- $$\|Q_{m|\mathcal{R}(V^*)}\| = \|Q_m\|. \tag{5.80}$$

- *For $d \in l^2(\mathbb{N})$*
$$Q_m d = \sum_{k=1}^{\infty} \hat{c}_k g_k,$$
where $\hat{c} = \{\hat{c}_k\}_{k\in\mathbb{N}}$ is the minimal norm element of the set
$$\arg\min_c \|\Sigma^{-\frac{1}{2}} V^* T c - \Sigma^{-\frac{1}{2}} d\|.$$

Proof. Using that $L^* = \Sigma^{-\frac{1}{2}} V^*$ (see Lemma 5.4.1), we infer that

$$Q_m V^* = T(\Sigma^{-\frac{1}{2}} V^* T)^\dagger \Sigma^{-\frac{1}{2}} V^* = T(L^* T)^\dagger L^*.$$

Therefore equation (5.78) follows from Theorem 5.2.5, (5.21), applied to the frame $\{\tilde{u}_j = ({\sigma_1}^2 I_n + {\sigma_2}^2 V V^*)^{\frac{\dagger}{2}} u_j\}_{j\in\mathbb{N}}$ instead of $\{u_j\}_{j\in\mathbb{N}}$ and L^* instead of V^*.

Next we prove that $Q_{m|\mathcal{R}(V^*)^\perp} = 0$. Equation (2.3) implies that

$$\mathcal{R}((\Sigma^{-\frac{1}{2}} V^* T)^\dagger) = \mathcal{R}(T^* V \Sigma^{-\frac{1}{2}}) = \mathcal{R}(T^* V) = \mathcal{R}(T^*),$$

where we used (5.19) for the last equality. Therefore by Lemma 2.1.3

$$\mathcal{N}(T(\Sigma^{-\frac{1}{2}} V^* T)^\dagger \Sigma^{-\frac{1}{2}}) = \mathcal{N}((\Sigma^{-\frac{1}{2}} V^* T)^\dagger \Sigma^{-\frac{1}{2}}). \tag{5.81}$$

Consequently, using (2.2) and (5.81)

$$\mathcal{N}(Q_m) = \mathcal{N}((\Sigma^{-\frac{1}{2}} V^* T)^\dagger \Sigma^{-\frac{1}{2}}) = \mathcal{N}(T^* V \Sigma^{-1}).$$

We show that

$$\mathcal{N}(V) = \mathcal{N}(V \Sigma^{-1}). \tag{5.82}$$

Since Σ is invertible, (5.82) is equivalent to $\mathcal{N}(V\Sigma) = \mathcal{N}(V)$. We observe that

$$\begin{aligned} V\Sigma &= V({\sigma_1}^2 I + {\sigma_2}^2 V^*V) = {\sigma_1}^2 V + {\sigma_2}^2 VV^*V \\ &= ({\sigma_1}^2 I + {\sigma_2}^2 VV^*)V. \end{aligned}$$

Since ${\sigma_1}^2 > 0$ it holds $({\sigma_1}^2 I + {\sigma_2}^2 VV^*) > 0$, which implies that $\mathcal{N}(V\Sigma) = \mathcal{N}(V)$. This proves (5.82), and consequently

$$\mathcal{R}(V^*)^\perp = \mathcal{N}(V) = \mathcal{N}(V\Sigma^{-1}) \subset \mathcal{N}(T^*V\Sigma^{-1}) = \mathcal{N}(Q_m).$$

Equation (5.80) is a direct consequence of (5.79).

The last point can be shown similarly to the last point of Theorem 5.2.5. □

Lemma 5.4.3. *Let $\mathcal{H}$, $\mathcal{T}$, $\mathcal{V}$, $\{u_j\}_{j\in\mathbb{N}}$, $\{g_k\}_{k\in\mathbb{N}}$, V and T be as in Theorem 5.4.2. Let*

$$Q_{s_1,s_2} = T\left((s_1 I + s_2 V^*V)^{\frac{\dagger}{2}} V^*T\right)^\dagger (s_1 I + s_2 V^*V)^{\frac{\dagger}{2}}.$$

Furthermore let $Q_f = T\left((V^*V)^{\frac{\dagger}{2}} V^*T\right)^\dagger (V^*V)^{\frac{\dagger}{2}}$, let $Q_g = T(V^*T)^\dagger$ and let $\cos(\varphi_{\mathcal{T},\mathcal{V}}) > 0$.

Then

$$Q_{s_1,s_2} = Q_{\frac{s_1}{s_2},1} = Q_{1,\frac{s_2}{s_1}},$$

$$Q_{0,s_2} = Q_f, \quad \text{and} \quad Q_{s_1,0} = Q_g.$$

Proof. The proof is straightforward and hence omitted. □

Lemma 5.4.3 shows that we only need one parameter to describe all operators Q_{s_1,s_2}, namely

$$Q_\lambda = T\left((V^*V + \lambda I)^{\frac{\dagger}{2}} V^*T\right)^\dagger (V^*V + \lambda I)^{\frac{\dagger}{2}}. \tag{5.83}$$

Furthermore $Q_0 = Q_f$ and $Q_\infty = Q_g$.

5.5 Connection to Statistics

Let $\mathcal{H}$ be a Hilbert space over the real numbers and let $\{u_j\}_{j=1}^n$ and $\{g_k\}_{k=1}^m$ be finite sequences in $\mathcal{H}$. Let T denote the synthesis operator of the sequence $\{g_k\}_{k=1}^m$, and let T^* denote the corresponding analysis operator. By $\mathcal{T}$ we denote

$$\mathcal{T} := \operatorname{span}\{g_k\}_{k=1}^m = \mathcal{R}(T).$$

Let V denote the synthesis operator of the sequence $\{u_j\}_{k=1}^n$, and let V^* denote the corresponding analysis operator.

We discuss the problem of estimating an unknown element $g \in \mathcal{T}$, i.e., an element of the reconstruction space, from perturbed measurements

$$X = V^* g + \varepsilon,$$

where ε is a random vector in $\mathbb{R}^n$ with $\mathcal{E}(\varepsilon) = 0$ and $\operatorname{Cov}(\varepsilon) = \Sigma$. We treat different versions of measurement noise ε.

In Corollary 5.5.10, we treat the case where ε is a random vector in $\mathbb{R}^n$ with $\mathcal{E}(\varepsilon) = 0$ and $\Sigma = \operatorname{Cov}(\varepsilon) = \sigma_1^2 I_n$. We show that the estimator $Q_g X$ of g, Q_g defined by (5.20), has the minimum expected norm-squared deviation from g within the class of all linear estimators Q with the property that $QV^* g = g$ for $g \in \mathcal{T}$.

In Corollary 5.5.11 we treat the case where $\varepsilon = V^* \Delta g$, where Δg is a random vector in (the finite dimensional Hilbert space) $\mathcal{H}$ with $\mathcal{E}(\Delta g) = 0$ and $\operatorname{Cov}(\Delta g) = \sigma_2^2 I$. In this case $\Sigma = \operatorname{Cov}(\varepsilon) = \sigma_2^2 V^* V$. We show that the estimator $Q_f X$ of g, Q_f defined by (5.44), has the minimum expected norm-squared deviation from g within the class of all linear estimators Q with the property that $QV^* g = g$ for $g \in \mathcal{T}$.

Finally, in Corollary 5.5.12 we treat the combination $\varepsilon = V^* \Delta g + \delta$ of the two versions of noise. We assume that $\mathcal{E}(\delta) = 0$, $\operatorname{Cov}(\delta) = \sigma_1^2 I_n$, $\mathcal{E}(\Delta g) = 0$ and $\operatorname{Cov}(\Delta g) = \sigma_2^2 I$, and that Δg and δ are independent random vectors. In this case $\Sigma = \operatorname{Cov}(\varepsilon) = {\sigma_1}^2 I_n + {\sigma_2}^2 V^* V$. We show that the estimator $Q_m X$ of g, Q_m defined by (5.77), has the minimum expected norm-squared deviation from g within the class of all linear estimators Q with the property that $QV^*g = g$ for $g \in \mathcal{T}$.

Since every $g \in \mathcal{T}$ can be written in the form $g = Tc$ for some $c \in \mathbb{R}^m$, our intention is to reconstruct the unknown element $g = Tc \in \mathcal{T}$ from the random vector

$$X = V^*Tc + \varepsilon.$$

We observe that $V^*T \in \mathbb{R}^{n\times m}$.

In order to prove Corollary 5.5.10, Corollary 5.5.11 and Corollary 5.5.12 we need a general version of the Gauss-Markov theorem.

5.5.1 Gauss-Markov theorem

The following discussion can be found in [24, Chapter 4].

First, we formulate a vector space version of the Gauss-Markov theorem. In the following, $\mathcal{W}$ is a finite dimensional Hilbert space over the real numbers.

Definition 5.5.1. *Let X be a random vector on $\mathcal{W}$ with distribution Q and let f be a real, Borel measurable function defined on $\mathcal{W}$. If*

$$\int_{\mathcal{W}} |f(x)| Q(dx) < \infty,$$

then we say $f(X)$ has a finite expectation and we write

$$\mathcal{E}f(X) = \int_{\mathcal{W}} f(x)Q(dx).$$

Definition 5.5.2. *Let X be a random vector on $\mathcal{W}$ with distribution Q and let f and g be real, Borel measurable functions defined on $\mathcal{W}$. Let $\int_{\mathcal{W}} |f(x)|^2 Q(dx) < \infty$ and $\int_{\mathcal{W}} |g(x)|^2 Q(dx) < \infty$. We define the covariance between $f(X)$ and $g(X)$ by*

$$\begin{aligned}\operatorname{cov}(f(X), g(X)) &:= \int_{\mathcal{W}} f(x)g(x)Q(dx) - \int_{\mathcal{W}} f(x)Q(dx) \int_{\mathcal{W}} g(x)Q(dx) \\ &= \mathcal{E}[f(X)g(X)] - \mathcal{E}f(X)\mathcal{E}g(X).\end{aligned}$$

Definition 5.5.3. *Let X be a random vector on the Hilbert space $\mathcal{W}$, and let for each $x \in \mathcal{W}$ the random variable $\langle x, X\rangle$ have a finite expectation. The unique vector $\mu \in \mathcal{W}$ with the property that*

$$\mathcal{E}(\langle x, X\rangle) = \langle x, \mu\rangle$$

is called the mean vector of X and is denoted by $\mu = \mathcal{E}X$.

Definition 5.5.4. *Let X be a random vector on the Hilbert space $\mathcal{W}$ with $\mathcal{E}(\langle x, X\rangle)^2 < \infty$ for every $x \in X$. A unique non-negative definite operator Σ on $\mathcal{W}$ that satisfies*

$$\operatorname{cov}(\langle x, X\rangle, \langle y, X\rangle) = \langle x, \Sigma y\rangle$$

is called the covariance of X

Let M be a subspace of $\mathcal{W}$. Let Y be a random vector in $\mathcal{W}$ with $\mathcal{E}Y = \mu \in M$ and $\operatorname{Cov}(Y) = \Sigma$ for some $\Sigma \geqslant 0$, i.e.

$$Y = \mu + \varepsilon,$$

where ε is a random vector with $\mathcal{E}\varepsilon = 0$ and $\operatorname{Cov}(\varepsilon) = \Sigma$.

Let $\mathcal{H}$ be another Hilbert space over the real numbers (possibly infinite dimensional) and let $B \in \mathcal{L}(\mathcal{W}, \mathcal{H})$. Our goal is to estimate the unknown element $B\mu$, where $\mu \in M$, from the random vector $Y = \mu + \varepsilon$ with $\mathcal{E}\varepsilon = 0$ and $\operatorname{Cov}(\varepsilon) = \Sigma$.

Definition 5.5.5. *Let $\mu \in M$ and $Y = \mu + \varepsilon$, with $\mathcal{E}\varepsilon = 0$ and $\operatorname{Cov}(\varepsilon) = \Sigma$. Let $A \in \mathcal{L}(\mathcal{W}, \mathcal{H})$. We call AY a linear estimator of $B\mu$.*

A linear estimator AY is called an unbiased estimator of $B\mu$ if, and only if $\mathcal{E}A(\nu + \varepsilon) = B\nu$ for every $\nu \in M$.

The classical Gauss-Markov approach restricts the problem of estimating $B\mu$ to the class of unbiased linear estimators of $B\mu$. Within this class, the Gauss-Markov theorem determines the estimator with the minimum expected norm-squared deviation from $B\mu$.

Since $\mathcal{E}AY = A\mathcal{E}Y = A\mu$, the estimator AY is an unbiased estimator for $B\mu$ for $\mu \in M$, if, and only if $A\mu = B\mu$ for $\mu \in M$. We denote the set of all unbiased linear estimators of $B\mu$, for $\mu \in M$, by

$$\mathcal{K}_1(B) := \{A \in \mathcal{L}(\mathcal{W}, \mathcal{H}) : A\mu = B\mu \text{ for } \mu \in M\}.$$

Theorem 5.5.6. *Let Y be a random vector in $\mathcal{W}$ with $\mathcal{E}Y = \mu \in M$ and $\mathrm{Cov}(Y) = \Sigma$. Let N be a subspace of $\mathcal{W}$ with the property that, $N \oplus M = \mathcal{W}$ and $\Sigma(M^\perp) \subset N$. For every $A \in \mathcal{K}_1(B)$,*

$$\mathcal{E}\|AY - B\mu\|^2 \geqslant \mathcal{E}\|BP_{MN}Y - B\mu\|^2. \tag{5.84}$$

If $\mathrm{Cov}(\varepsilon) = \sigma^2 I$, then $N = M^\perp$ is the unique subspace with the desired properties, and there is equality in (5.84) if, and only if $A = BP_M$.

5.5.2 Application of the Gauss-Markov theorem to the reconstruction operators

We need Lemma 5.5.7 and Lemma 5.5.8 in order to prove Theorem 5.5.9, Corollary 5.5.10, Corollary 5.5.11 and Corollary 5.5.12.

Lemma 5.5.7. *Let $\mathcal{V}$ and $\mathcal{T}$ be closed subspaces of $\mathcal{H}$. Let $\{u_j\}_{j\in\mathbb{N}}$ be a frame for $\mathcal{V}$ and let V be the corresponding synthesis. Let $\{g_k\}_{k\in\mathbb{N}}$ be a frame for $\mathcal{T}$, let T be the corresponding synthesis operator. Furthermore let Σ be a bounded operator on $l^2(\mathbb{N})$ with a closed range, let $\Sigma \geqslant 0$, and let the operator $\Sigma^{\frac{\dagger}{2}} V^* T$ have a closed range.*

If $\mathcal{R}(T^* V \Sigma) = \mathcal{R}(T^*)$, then for the operator Q, defined by

$$Q_m := T(\Sigma^{\frac{\dagger}{2}} V^* T)^\dagger \Sigma^{\frac{\dagger}{2}},$$

and the operator P, defined by

$$P = V^* Q_m$$

the following holds.

1. *$Q_m P = Q_m$,*

2. *The operator $R := Q_m V^*$ is a bounded projection onto $\mathcal{T}$,*

3. *P is a bounded projection with $\mathcal{R}(P) = \mathcal{R}(V^*T)$,*

4. *$\mathcal{N}(P) = \mathcal{N}(T^*V\Sigma^\dagger)$.*

Proof. From Corollary 2.2.4 we infer that

$$\begin{aligned} Q_m P &= T(\Sigma^{\frac{\dagger}{2}} V^*T)^\dagger \Sigma^{\frac{\dagger}{2}} V^*T(\Sigma^{\frac{\dagger}{2}} V^*T)^\dagger \Sigma^{\frac{\dagger}{2}} \\ &= T(\Sigma^{\frac{\dagger}{2}} V^*T)^\dagger \Sigma^{\frac{\dagger}{2}} = Q_m. \end{aligned}$$

Next we show (2). Using again Corollary 2.2.4, we obtain

$$\begin{aligned} R^2 &= T(\Sigma^{\frac{\dagger}{2}} V^*T)^\dagger \Sigma^{\frac{\dagger}{2}} V^*T(\Sigma^{\frac{\dagger}{2}} V^*T)^\dagger \Sigma^{\frac{\dagger}{2}} V^* \\ &= T(\Sigma^{\frac{\dagger}{2}} V^*T)^\dagger \Sigma^{\frac{\dagger}{2}} V^* = R. \end{aligned}$$

Clearly $\mathcal{R}(R) \subset \mathcal{T}$. We show that $\mathcal{R}(R_{|\mathcal{T}}) = \mathcal{T}$, to prove the inverse implication,. Since Σ is selfadjoint we obtain

$$\Sigma = P_{\mathcal{R}(\Sigma)}\Sigma = P_{\mathcal{R}(\Sigma^*)}\Sigma = \Sigma^\dagger\Sigma\Sigma = \Sigma^{\frac{\dagger}{2}}\Sigma^{\frac{\dagger}{2}}\Sigma\Sigma,$$

and

$$\Sigma^{\frac{\dagger}{2}} = P_{\mathcal{R}(\Sigma^*)}\Sigma^{\frac{\dagger}{2}} = P_{\mathcal{R}(\Sigma)}\Sigma^{\frac{\dagger}{2}} = \Sigma\Sigma^\dagger\Sigma^{\frac{\dagger}{2}},$$

and consequently

$$\mathcal{R}(T^*V\Sigma^{\frac{\dagger}{2}}) = \mathcal{R}(T^*V\Sigma).$$

Since by assumption $\mathcal{R}(T^*V\Sigma) = \mathcal{R}(T^*)$, we infer that

$$\begin{aligned} RT &= T(\Sigma^{\frac{\dagger}{2}} V^*T)^\dagger \Sigma^{\frac{\dagger}{2}} V^*T = TP_{\mathcal{R}(T^*V\Sigma^{\frac{\dagger}{2}})} = TP_{\mathcal{R}(T^*)} \\ &= TP_{\mathcal{N}(T)^\perp} = T. \end{aligned}$$

Next we prove (3). Using Corollary 2.2.4 we obtain

$$\begin{aligned} P^2 &= V^*T(\Sigma^{\frac{\dagger}{2}}V^*T)^\dagger\Sigma^{\frac{\dagger}{2}}V^*T(\Sigma^{\frac{\dagger}{2}}V^*T)^\dagger\Sigma^{\frac{\dagger}{2}} \\ &= V^*T(\Sigma^{\frac{\dagger}{2}}V^*T)^\dagger\Sigma^{\frac{\dagger}{2}} = P. \end{aligned}$$

Clearly $\mathcal{R}(P) \subset \mathcal{T}$. We show that $P_{|\mathcal{R}(V^*T)} = \mathcal{R}(V^*T)$, to prove the inverse implication. This follows from

$$PV^*T = V^*T(\Sigma^{\frac{\dagger}{2}}V^*T)^\dagger\Sigma^{\frac{\dagger}{2}}V^*T = V^*RT = V^*T.$$

We observe that

$$\mathcal{R}((\Sigma^{\frac{\dagger}{2}}V^*T)^\dagger) = \mathcal{R}(T^*V\Sigma^{\frac{\dagger}{2}}) \subset \mathcal{R}(T^*V).$$

Combining this with Lemma 2.1.3 we infer that

$$\mathcal{N}(P) = \mathcal{N}(V^*T(\Sigma^{\frac{\dagger}{2}}V^*T)^\dagger\Sigma^{\frac{\dagger}{2}}) = \mathcal{N}((\Sigma^{\frac{\dagger}{2}}V^*T)^\dagger\Sigma^{\frac{\dagger}{2}}) \tag{5.85}$$

Using that $\mathcal{N}(\Sigma^{\frac{\dagger}{2}}V^*T)^\dagger) = \mathcal{N}(\Sigma^{\frac{\dagger}{2}}V^*T)^*) = \mathcal{N}(T^*V\Sigma^{\frac{\dagger}{2}})$ (see (2.2)), we infer from (5.85) that $\mathcal{N}(P) = \mathcal{N}(T^*V\Sigma^\dagger)$.

□

Lemma 5.5.8. *Let $\mathcal{H}$, $\mathcal{T}$, $\mathcal{V}$, $\{u_j\}_{j\in\mathbb{N}}$, V, $\{g_k\}_{k\in\mathbb{N}}$ and T be as in Theorem 5.4.2. Let $\cos(\varphi_{\mathcal{T},\mathcal{V}}) > 0$ and $\sigma^2 > 0$.*

For the operator $\Sigma = \sigma^2V^*V$ and for any strictly positive operator Σ on $l^2(\mathbb{N})$ it holds

$$\mathcal{R}(T^*V\Sigma) = \mathcal{R}(T^*).$$

Proof. Since $\cos(\varphi_{\mathcal{T},\mathcal{V}}) > 0$, $\{P_{\mathcal{T}}u_j\}_{j\in\mathbb{N}}$ is a frame for $\mathcal{T}$, and consequently by Lemma 5.2.4 it holds $\mathcal{R}(T^*V) = \mathcal{R}(T^*)$. Therefore it is sufficient to prove that $\mathcal{R}(T^*V\Sigma) = \mathcal{R}(T^*V)$. If $\Sigma > 0$, then Σ is invertible, and consequently $\mathcal{R}(T^*V\Sigma) = \mathcal{R}(T^*V)$. For $\Sigma = \sigma^2V^*V$ it holds $\mathcal{R}(T^*V\Sigma) = \mathcal{R}(T^*VV^*V)$. Clearly $\mathcal{R}(T^*VV^*V) \subset \mathcal{R}(T^*V)$. Using Lemma 2.1.3 and Lemma 2.2.3 we infer that

$$T^*VV^*V(V^*V)^\dagger = T^*VP_{\mathcal{R}(V^*V)} = T^*VP_{\mathcal{R}(V^*)} = T^*VP_{\mathcal{N}(V)^\perp} = T^*V.$$

Consequently also $\mathcal{R}(T^*V) \subset \mathcal{R}(T^*VV^*V)$. $\square$

Theorem 5.5.9. *Let $\mathcal{H}$ be a Hilbert space over the real numbers and let $\{u_j\}_{j=1}^n$ and $\{g_k\}_{k=1}^m$ be finite sequences in $\mathcal{H}$ with linear span $\mathcal{V}$ and $\mathcal{T}$ respectively. Let T denote the synthesis operator of the sequence $\{g_k\}_{k=1}^m$, and let T^* denote the corresponding analysis operator. Let V denote the synthesis operator of the sequence $\{u_j\}_{k=1}^n$, and let V^* denote the corresponding analysis operator. Let $g \in \mathcal{T}$ and let*

$$Y = V^*g + \varepsilon,$$

where ε is a random vector in $\mathbb{R}^n$ with $\mathcal{E}(\varepsilon) = 0$ and $\operatorname{Cov}(\varepsilon) = \Sigma$, with $\Sigma > 0$. Let

$$Q_m := T(\Sigma^{-\frac{1}{2}}V^*T)^\dagger\Sigma^{-\frac{1}{2}}.$$

If $\cos(\varphi_{\mathcal{T},\mathcal{V}}) > 0$, for every $Q \in \mathcal{L}(\mathbb{R}^n, \mathcal{H})$ with $QV^*h = h$ for $h \in \mathcal{T}$ it holds

$$\mathcal{E}\|QY - g\|^2 \geqslant \mathcal{E}\|Q_mY - g\|^2.$$

Proof. Let $g = Tc \in \mathcal{T}$. With this notation

$$Y = V^*Tc + \varepsilon.$$

For every $g \in \mathcal{T}$, $\mathcal{E}(Y) \in M$ with $M = \mathcal{R}(V^*T)$. Since $\Sigma > 0$, Lemma 5.5.8 implies that $\mathcal{R}(T^*V\Sigma) = \mathcal{R}(T^*)$, and consequently, using Lemma 5.5.7, (2), we obtain

$$\begin{aligned}
K_1(Q_m) &= \{Q \in \mathcal{L}(\mathbb{R}^n, \mathcal{H}) : Q\mu = Q_m\mu \text{ for } \mu \in \mathcal{R}(V^*T)\} \\
&= \{Q \in \mathcal{L}(\mathbb{R}^n, \mathcal{H}) : QV^*Td = Q_mV^*Td = Td \text{ for } d \in \mathbb{R}^m\} \\
&= \{Q \in \mathcal{L}(\mathbb{R}^n, \mathcal{H}) : QV^*h = h \text{ for } h \in \mathcal{T}\}.
\end{aligned}$$

Let P be defined by

$$P = V^*Q_m.$$

Lemma 5.5.7, (3), implies that P is a bounded projection onto $\mathcal{R}(V^*T)$ and consequently for $N = \mathcal{N}(P)$, $N \oplus M = \mathbb{R}^n$. Lemma 5.5.7, (4), implies that $\Sigma(\mathcal{R}(V^*T)^\perp) \subset \mathcal{N}(P)$, and consequently Theorem 5.5.6 implies that for every $Q \in K_1(Q_m)$,

$$\mathcal{E}\|QY - g\|^2 \geqslant \mathcal{E}\|Q_m PY - g\|^2.$$

Since $Q_m P = Q_m$ by Lemma 5.5.7, (1), this yields the desired result. $\square$

From Theorem 5.5.9 we obtain Corollary 5.5.10.

Corollary 5.5.10. *Let $\mathcal{H}$, $\{u_j\}_{j=1}^n$, $\{g_k\}_{k=1}^m$, $\mathcal{V}$, $\mathcal{T}$, V and T be as in Theorem 5.5.9. Let $g \in \mathcal{T}$ and let*

$$Y = V^*g + \varepsilon,$$

where ε is a random vector in $\mathbb{R}^n$ with $\mathcal{E}(\varepsilon) = 0$ and $\mathrm{Cov}(\varepsilon) = \sigma^2 I_n$, $\sigma^2 > 0$. Let

$$Q_g := T(V^*T)^\dagger.$$

If $\cos(\varphi_{\mathcal{T},\mathcal{V}}) > 0$, for every $Q \in \mathcal{L}(\mathbb{R}^n, \mathcal{H})$ with $QV^*h = h$ for $h \in \mathcal{T}$ it holds

$$\mathcal{E}\|QY - g\|^2 \geqslant \mathcal{E}\|Q_g Y - g\|^2,$$

with equality if, and only if $Q = Q_g$.

Corollary 5.5.11. *Let $\mathcal{H}$ be a finite dimensional Hilbert space over the real numbers and let $\{u_j\}_{j=1}^n$ and $\{g_k\}_{k=1}^m$ be finite sequences in $\mathcal{H}$ with linear span $\mathcal{V}$ and $\mathcal{T}$ respectively. Let T denote the synthesis operator of the sequence $\{g_k\}_{k=1}^m$, and let T^* denote the corresponding analysis operator. Let V denote the synthesis operator of the sequence $\{u_j\}_{k=1}^n$, and let V^* denote the corresponding analysis operator. Let $g \in \mathcal{T}$ and let*

$$Y = V^*g + V^*\Delta g,$$

where Δg is a random vector in $\mathcal{H}$ with $\mathcal{E}\Delta g = 0$ and $\mathrm{Cov}(\Delta g) = \sigma^2 I$. Let

$$Q_f := T\left((V^*V)^{\frac{\dagger}{2}} V^*T\right)^\dagger (V^*V)^{\frac{\dagger}{2}}.$$

If $\cos(\varphi_{\mathcal{T},\mathcal{V}}) > 0$, for every $Q \in \mathcal{L}(\mathbb{R}^n, \mathcal{H})$ with $QV^*h = h$ for $h \in \mathcal{T}$ it *holds*

$$\mathcal{E}\|QY - g\|^2 \geqslant \mathcal{E}\|Q_f Y - g\|^2.$$

Proof. We observe that $\Sigma := \mathrm{Cov}(V^*\Delta g) = \sigma^2 V^* V$. From Lemma 5.5.8 we obtain $\mathcal{R}(T^*V\Sigma) = \mathcal{R}(T^*)$, and consequently the assumptions of Lemma 5.5.7 are fulfilled. The proof follows now repeating the arguments of the proof of Theorem 5.5.9. □

Corollary 5.5.12. *Let* $\mathcal{H}$, $\{u_j\}_{j=1}^n$, $\{g_k\}_{k=1}^m$, $\mathcal{V}$, $\mathcal{T}$, V *and* T *be as in Theorem 5.5.11. Let* $g \in \mathcal{T}$ *and let*

$$Y = V^* g + V^* \Delta g + \varepsilon, \tag{5.86}$$

where Δg *and* ε *are independent random vectors with* $\sigma_1^2, \sigma_2^2 > 0$, $\mathcal{E}\varepsilon = 0$, $\mathrm{Cov}(\varepsilon) = \sigma_1^2 I_n$, $\mathcal{E}(\Delta g) = 0$ *and* $\mathrm{Cov}(\Delta g) = \sigma_2^2 I$. *Let*

$$\Sigma = {\sigma_1}^2 I_n + {\sigma_2}^2 V^* V,$$

and let

$$Q_m = T(\Sigma^{-\frac{1}{2}} V^* T)^{\dagger} \Sigma^{-\frac{1}{2}}.$$

If $\cos(\varphi_{\mathcal{T},\mathcal{V}}) > 0$, *for every* $Q \in \mathcal{L}(\mathbb{R}^n, \mathcal{H})$ *with* $QV^*h = h$ *for* $h \in \mathcal{T}$ *it holds*

$$\mathcal{E}\|QY - g\|^2 \geqslant \mathcal{E}\|Q_m Y - g\|^2. \tag{5.87}$$

Proof. Since Δg and ε are independent random vectors,

$$\Sigma = \mathrm{Cov}(V^*\Delta g + \varepsilon) = {\sigma_1}^2 I_n + {\sigma_2}^2 V^* V.$$

Since ${\sigma_1}^2 > 0$, $\Sigma > 0$ and the statement follows from Theorem 5.5.9. □

5.6 Calculation of the approximations and conditions for them to coincide

5.6.1 Calculation of the coefficients

We now discuss how to calculate the coefficients of the reconstructions in the finite dimensional setup. Let $\{u_j\}_{j=1}^n$ and $\{g_k\}_{k=1}^m$ be finite sequences in $\mathcal{H}$. We recall that our intention is to approximate an element $f \in \mathcal{H}$ from measurements

$$d_j = \langle f, u_j \rangle + \delta_j \quad j = 1, \ldots, n,$$

where δ_j is noise. Let $d \in \mathbb{C}^n$ denote the vector consisting of the measurements

$$d = [d_1, \ldots, d_n]^T.$$

We choose our approximations $\tilde{f}$ of f as a linear combination

$$\tilde{f} = \sum_{k=1}^{m} \hat{c}_k g_k, \tag{5.88}$$

with linearly independent reconstruction vectors $\{g_k\}_{k=1}^m$ and $n \geqslant m$.

In Section 5.2 we discussed the concept of generalized sampling where the vector $\hat{c} = [\hat{c}_1, \ldots, \hat{c}_m]^T$ containing the coefficients of the series expansion of $\tilde{f}$ is the solution of the overdetermined least squares problem

$$\hat{c} = \arg\min_{c} \|Ac - d\|, \tag{5.89}$$

where $A \in \mathbb{C}^{n \times m}$ is defined by

$$A(j,k) = (V^*T)(j,k) = \langle u_j, g_k \rangle.$$

In Section 5.3 we discussed the concept of frame independent sampling, where the vector $\hat{c} = [\hat{c}_1, \ldots, \hat{c}_m]^T$ containing the coefficients of the series expansion of $\tilde{f}$ is the solution of the overdetermined least squares problem

$$\hat{c} = \arg\min_{c} \|B^{\frac{\dagger}{2}} Ac - B^{\frac{\dagger}{2}} d\|, \tag{5.90}$$

where $B \in \mathbb{C}^{n \times n}$,

$$B(j,k) = (V^*V)(j,k) = \langle u_j, u_k \rangle.$$

In Section 5.4 we also mentioned combinations between those two extreme reconstruction methods, where $\hat{c} = [\hat{c}_1, \ldots, \hat{c}_m]^T$ is the solution of the overdetermined least squares problem

$$\hat{c} = \arg\min_{c} \|\Sigma_\lambda^{-\frac{1}{2}} Ac - \Sigma_\lambda^{-\frac{1}{2}} d\|, \tag{5.91}$$

with $\Sigma_\lambda = \lambda I + V^*V$.

We observe that calculating the coefficients $\{\hat{c}_k\}_{k=1}^m$ for the reconstruction $Q_g d = \sum_{k=1}^m \hat{c}_k g_k$ amounts to solving the least squares problem (5.89). For the reconstruction $Q_f d$, the least squares system (5.90) has to be solved, and for the reconstruction $Q_\lambda d$, the least squares system (5.91) has to be solved. We observe that (5.90) and (5.91) are (5.89) multiplied by the matrix $B^{\frac{\dagger}{2}}$ and $\Sigma^{\frac{\dagger}{2}}$, respectively.
Solving the overdetermined least squares problem (5.89) by a direct method as the QR decomposition with pivoting has an operation count of $\mathcal{O}(nm^2)$. Solving the least squares problem with a fixed precision $\varepsilon > 0$ can be accomplished iteratively in $\mathcal{O}(\log(\varepsilon)nm)$ flops. One can e.g. use the conjugate gradient method applied to the normal equations as an iterative solver. This can for example be realized by the LSQR algorithm, see [51]. A condition number $\kappa(A)$ close to one results in fast convergence of the conjugate gradient method. In [31, Theorem 10.2.6] it is shown that the conjugate gradient method applied to a matrix with condition number κ converges exponentially at the rate of $\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}$. In Lemma 5.6.1 we estimate the condition number $\kappa(A^*\Sigma_\lambda^{-1}A) = \|A^*\Sigma_\lambda^{-1}A\|\|(A^*\Sigma_\lambda^{-1}A)^{-1}\|$ of the matrix of the normal equations of the least squares problem (5.91).

We recall that for $\lambda \geqslant 0$ the reconstruction operators Q_λ (including $Q_0 = Q_f$ and $Q_\infty = Q_g$) are independent of the particular choice of the reconstruction vectors $\{g_k\}_{k=1}^m$, but only dependent on $\mathcal{T}$, the linear span. Therefore we may assume that $\{g_k\}_{k=1}^m$ are an orthonormal system.

Lemma 5.6.1. *Let $\mathcal{T}$ and $\mathcal{V}$ be closed subspaces of a separable Hilbert space $\mathcal{H}$. Let $\{u_j\}_{j\in\mathbb{N}}$ be a frame for $\mathcal{V}$ with lower and upper frame bound A and B respectively, and let V be the corresponding synthesis operator. Let $\{g_k\}_{k\in\mathbb{N}}$ be an orthonormal basis for $\mathcal{T}$, and let T be the corresponding synthesis operator. Furthermore let $\lambda \geqslant 0$, and let $\Sigma_\lambda = \lambda I + V^*V$. We set $B = T^*V\Sigma^{-1}V^*T$.*

If $\cos(\varphi_{\mathcal{T},\mathcal{V}}) > 0$, then

$$\kappa(B) \leqslant \frac{1}{\cos^2(\varphi_{\mathcal{T},\mathcal{V}})} \frac{B}{\lambda + B} \frac{\lambda + A}{A}. \tag{5.92}$$

Proof. From Lemma 5.4.1 we know that $\{\tilde{u}_j = (\lambda I + VV^*)^{\dagger \frac{1}{2}} u_j\}_{j\in\mathbb{N}}$ is a frame for $\mathcal{V}$ with lower and upper frame bounds $\frac{A}{\lambda+A}$, and $\frac{B}{\lambda+B}$, and the synthesis operator $L = \Sigma^{-\frac{1}{2}}V^*$. In the proof of Lemma 3.4.2 it is shown that for $g = Tc \in \mathcal{T}$

$$\left(\frac{A}{\lambda + A}\right)^{\frac{1}{2}} \cos(\varphi_{\mathcal{T},\mathcal{V}})\|Tc\| \leqslant \|L^*Tc\| \leqslant \left(\frac{B}{\lambda + B}\right)^{\frac{1}{2}} \|Tc\|. \tag{5.93}$$

Since $\{g_k\}_{k\in\mathbb{N}}$ is an orthonormal basis, $\|Tc\| = \|c\|$. This together with (5.93) implies (5.92). □

We observe that the larger the tuning parameter λ, the larger the upper bound of the condition number $\kappa(A^*\Sigma_\lambda^{-1}A)$ of the matrix of the normal equations of the least squares problem (5.91). Therefore we expect that

$$\kappa(A^*\Sigma_0^{-1}A) \leqslant \kappa(A^*\Sigma_\lambda^{-1}A) \leqslant \kappa(A^*A),$$

which is tested in numerical experiments in Section 6. Consequently we expect faster convergence of the conjugate gradient method applied to the matrix $A^*\Sigma_\lambda^{-1}A$ instead of A^*A. This is the general motivation for preconditioning, see [3, 5, 29, 34, 35, 60] for related examples.

In order to calculate approximations of the solution of the least squares problem (5.90) at low operation count, we first need to calculate approximations M_λ of Σ_λ^{-1} followed by solving the normal equations

$$A^* M_\lambda A c = A^* M_\lambda h.$$

We hope to obtain approximations M_λ of Σ_λ^{-1} at a low operation count from the theory of controlled frames, weighted frames and frame multipliers, see [13, 14, 18].

5.6.2 Conditions for the approximations to coincide

While in general the reconstructions Q_g, Q_f and Q_m are rather different, they coincide in several situations. Since by (5.22) $Q_{g|\mathcal{R}(V^*)^\perp} = 0$, by (5.46) $Q_{f|\mathcal{R}(V^*)^\perp} = 0$, and by (5.79) $Q_{m|\mathcal{R}(V^*)^\perp} = 0$, instead of studying Q_g, Q_f and Q_m we inspect the oblique projections Q_gV^*, Q_fV^*, and Q_mV^*. In Theorem 5.2.5 it is shown that $Q_gV^* = P_{\mathcal{T},S_u(\mathcal{T})^\perp}$, where S_u is the frame operator of the frame sequence $\{u_j\}_{j\in\mathbb{N}}$ of $\mathcal{V}$. In Theorem 5.3.2 it is shown that $Q_fV^* = P_{\mathcal{T},S_f(\mathcal{T})^\perp}$, where S_f is the frame operator of the tight frame sequence $\{S^{\frac{\dagger}{2}}u_j\}_{j\in\mathbb{N}}$ of $\mathcal{V}$. In Theorem 5.4.2 it is shown that $Q_mV^* = P_{\mathcal{T},S_m(\mathcal{T})^\perp}$, where S_m is the frame operator of the frame sequence $\left\{(\sigma_1{}^2 I_n + \sigma_2{}^2 VV^*)^{-\frac{1}{2}}u_j\right\}_{j\in\mathbb{N}}$ of $\mathcal{V}$. Therefore it is sufficient to determine when $S_f(\mathcal{T}) = S_g(\mathcal{T}) = S_m(\mathcal{T})$.

The following Lemma states that if $\mathcal{T} \oplus \mathcal{V}^\perp = \mathcal{H}$, then all operators $Q : l^2(\mathbb{N}) \to \mathcal{H}$, with $QV^*g = g$ for $g \in \mathcal{T}$, coincide on $\mathcal{R}(V^*)$. This property follows from [27, Theorem 1].

Lemma 5.6.2. *Let $\mathcal{T}$ and $\mathcal{V}$ be closed subspaces for $\mathcal{V}$. Let $\{u_j\}_{j=1}^n$ be a frame for $\mathcal{V}$. If $\mathcal{T} \oplus \mathcal{V}^\perp = \mathcal{H}$, then for any operator $Q : l^2(\mathbb{N}) \to \mathcal{H}$ with the property that $QV^*g = g$ for $g \in \mathcal{T}$ it holds*

$$QV^* = P_{\mathcal{T},\mathcal{V}^\perp}, \tag{5.94}$$

and consequently

$$Q_f = Q_g = Q_m.$$

Proof. As in Theorem it is shown that $QV^*g = g$ for $g \in \mathcal{T}$ implies that $(QV^*)_{|\mathcal{T}\oplus\mathcal{V}^\perp} = P_{\mathcal{T},\mathcal{V}^\perp}$. Since by assumption $\mathcal{T} \oplus \mathcal{V}^\perp = \mathcal{H}$, this proves (5.94).

□

Lemma 5.6.3 is the finite dimensional version of Lemma 5.6.2.

Lemma 5.6.3. *Let $\mathcal{T}$ and $\mathcal{V}$ be finite dimensional subspaces in $\mathcal{H}$ and let* $\dim(\mathcal{T}) = \dim(\mathcal{V})$. *If $\mathcal{T} \cap \mathcal{V}^{\perp} = \{0\}$, then for any operator $Q : l^2(\mathbb{N}) \to \mathcal{H}$ with the property that $QV^*g = g$ for $g \in \mathcal{T}$ it holds*

$$QV^* = P_{\mathcal{T},\mathcal{V}^{\perp}},$$

and consequently

$$Q_f = Q_g = Q_m.$$

In order to prove Theorem 5.6.5 we need Lemma 5.6.4.

Lemma 5.6.4. *Let $\mathcal{V}$ be a closed subspaces of $\mathcal{H}$. Furthermore let $\{u_j\}_{j\in\mathbb{N}}$ be a tight frame for $\mathcal{V}$, with corresponding synthesis operator V and corresponding analysis operator V^* and let ${\sigma_1}^2, {\sigma_2}^2 > 0$.*

Then also $\left\{({\sigma_1}^2 I + {\sigma_2}^2 VV^*)^{-\frac{1}{2}} u_j\right\}_{j\in\mathbb{N}}$ is a tight frame for $\mathcal{V}$.

Proof. Let A denote the frame bound of $\{u_j\}_{j\in\mathbb{N}}$. The synthesis operator of $\left\{({\sigma_1}^2 I + {\sigma_2}^2 VV^*)^{-\frac{1}{2}} u_j\right\}_{j\in\mathbb{N}}$ is given by

$$T := ({\sigma_1}^2 I + {\sigma_2}^2 VV^*)^{-\frac{1}{2}} V.$$

We have to prove that there exists a constant $C > 0$, such that

$$\|T^*u\|^2 = C\|u\|^2 \quad \text{for every } u \in \mathcal{V}.$$

By Theorem 2.3.15 $S = VV^* = AP_{\mathcal{V}}$, and consequently for $u \in \mathcal{V}$

$$({\sigma_1}^2 I + {\sigma_2}^2 VV^*)u = ({\sigma_1}^2 + {\sigma_2}^2 A)u.$$

Therefore for $u \in \mathcal{V}$

$$\begin{aligned}\|T^*u\|^2 = \|V^*({\sigma_1}^2 I + {\sigma_2}^2 VV^*)^{-\frac{1}{2}} u\|^2 &= \frac{1}{{\sigma_1}^2 + {\sigma_2}^2 A}\|V^*u\|^2 \\ &= \frac{A}{{\sigma_1}^2 + {\sigma_2}^2 A}\|u\|^2.\end{aligned}$$

□

Theorem 5.6.5 states that the reconstructions coincide whenever the sampling frame is tight. An important case of the sampling frame being tight is the reconstruction of a compactly supported function from its Fourier coefficients, as mentioned in the introduction.

Theorem 5.6.5. *Let $\mathcal{T}$ and $\mathcal{V}$ be closed subspaces of $\mathcal{H}$ and let $\cos(\varphi_{\mathcal{T},\mathcal{V}}) > 0$. If $\{u_j\}_{j\in\mathbb{N}}$ is a tight frame for $\mathcal{V}$, then*

$$Q_f = Q_g = Q_m. \tag{5.95}$$

Proof. If $\{f_j\}_{j\in\mathbb{N}}$ is a tight frame for $\mathcal{V}$ with frame operator S, then
$S(\mathcal{T}) = P_{\mathcal{V}}(\mathcal{T})$ and consequently

$$P_{\mathcal{T},S(\mathcal{T})^{\perp}} = P_{\mathcal{T},P_{\mathcal{V}}(\mathcal{T})^{\perp}}.$$

Equation (5.95) is now a direct consequence of Lemma 5.6.4 and the discussion at the beginning of this section. □

Chapter 6

Numerical experiments for reconstruction from Fourier measurements

In this chapter, we compare experimentally accuracy of various reconstruction methods. We consider a specific setup, where non-uniform samples of the Fourier transform of a compactly supported function are given. The reconstruction of compactly supported function from a non-uniform sampling pattern is common in practice. For example, radial sampling of the Fourier transform is used in MRI and CT, see [47] for a detailed explanation why this is the case. From the given data, we calculate the Fourier coefficients of the function, and construct a final approximation by a truncated Fourier series. This is called the uniform resampling problem, see [10, 66]. For smooth and periodic functions the Fourier series converges exponentially fast. For non-periodic or discontinuous functions, the Fourier series representation suffers from the Gibbs phenomenon and slow convergence. First we approximate the signum function. Due to the jumps of this function, trigonometric polynomials are a bad choice as reconstruction functions and for this function other basis functions are much better suited. Our objective in this section is not to choose optimal reconstruction functions, but rather to show how the different reconstruction operators deal with the part outside of the reconstruction space. For that purpose, using a function with a jump and using trigonometric polynomials for reconstruction is a good choice.

In this section we denote by $\langle\cdot,\cdot\rangle$ the standard inner product on the Hilbert space $L^2(\mathbb{R})$ defined by

$$\langle f, g\rangle = \int_{-\infty}^{\infty} f(x)\overline{g(x)}\,\mathrm{d}x,$$

and $\|\cdot\| = \sqrt{\langle\cdot,\cdot\rangle}$. We use the following definition of the Fourier transform on $L^2(\mathbb{R})$

$$\mathcal{F}f(\xi) = \int_{-\infty}^{\infty} f(x)e^{-2\pi i x\xi}\,\mathrm{d}x.$$

Let $\mathcal{H}$ be the subspace of $L^2(\mathbb{R})$ of functions with support in the interval $[-1/2, 1/2]$, i.e.,

$$\mathcal{H} = \left\{f \in L^2(\mathbb{R}) : \mathrm{supp}(f) \subset [-1/2, 1/2]\right\}.$$

6.1 The noiseless case

We approximate functions $f \in \mathcal{H}$ from a set of Fourier measurements

$$\mathcal{F}f(\omega_j) = \left\langle f(\cdot), e^{2\pi i\omega_j\cdot}\chi_{[-1/2,1/2]}(\cdot)\right\rangle, \quad j = -n, \dots, n,$$

where $\omega_j \in \mathbb{R}$. This means that the sampling vectors are

$$u_j = \mathrm{e}^{2\pi i\omega_j\cdot}\chi_{[-1/2,1/2]}(\cdot), \quad j = -n, \dots, n.$$

In the special case that we are given the Fourier coefficients

$$\mathcal{F}f(j), \quad j = -n, \dots, n,$$

the approximation $\tilde{f}$ of f is the Fourier series

$$\tilde{f}(x) = \sum_{k=-n}^{n} \mathcal{F}f(k)e^{2i\pi kx}\chi_{[-1/2,1/2]}(x).$$

This means that the reconstruction vectors are the complex exponentials

$$g_k = \mathrm{e}^{2\pi ik\cdot}\chi_{[-1/2,1/2]}, \quad k = -m, \dots, m$$

with $m = n$, i.e., we use as many reconstruction vectors as the number of given measurements.

If we are given irregular Fourier measurements $\mathcal{F}f(\omega_j)$, $j = -n, \dots, n$, then it is natural to calculate the coefficients c_k for the approximation

$$\tilde{f}(x) = \sum_{k=-n}^{n} c_k e^{2i\pi kx}\chi_{[-1/2,1/2]}(x) \tag{6.1}$$

of f such that for every $j = -n, \dots, n$

$$\mathcal{F}\left(\sum_{k=-n}^{n} c_k e^{2i\pi k\cdot}\chi_{[-1/2,1/2]}(\cdot)\right)(\omega_j) = \mathcal{F}f(\omega_j),$$

i.e., a consistent approximation. If the function f to be approximated has a part outside of the reconstruction space, then this consistent approach can lead to inaccurate approximations. We approximate the function

$$f(x) = \operatorname{sgn}(x) := \begin{cases} -1 & \text{if } x < 0, \\ 0 & \text{if } x = 0, \\ 1 & \text{if } x > 0, \end{cases}$$

given the Fourier measurements $\mathcal{F}f(\omega_j)$ at the sampling frequencies

$$\omega_j = j + \delta_j, \quad j = -n, \dots, n, \tag{6.2}$$

for $n = 40$, with $\delta_j \in [-1, 1]$ drawn from the uniform distribution in the interval $[-1, 1]$. We observe experimentally that the consistent approximation $\tilde{f} = Q_c V^* f$ has an average relative error of $\|\tilde{f} - f\|/\|f\| = 0.247$ (arithmetic mean over 5000 approximations).

We try to reduce the approximation error by taking fewer reconstruction vectors than measurements. We choose $m = 20$, i.e., a total amount of 41 reconstruction vectors and $n = 40$, i.e., a total amount of 81 measurements. We stay with the example in order to demonstrate the difference between various approximation procedures.

As mentioned in the introduction, a common choice is to calculate the coefficients for the approximation

$$\tilde{f}(x) = \sum_{k=-m}^{m} \hat{c}_k e^{2i\pi kx} \chi_{[-1/2,1/2]}(x)$$

of f by the standard least squares problem

$$\hat{c} = \arg\min_c \sum_{j=-n}^{n} \left| \mathcal{F}\Big(\sum_{k=-m}^{m} c_k e^{2i\pi kx} \chi_{[-\frac{1}{2},\frac{1}{2}]}(x) \Big)(j+\delta_j) - \mathcal{F}f(j+\delta_j) \right|^2.$$

We use the notation T for the synthesis operator of the reconstruction sequence, i.e.,

$$Tc = \sum_{k=-m}^{m} c_k e^{2i\pi k\cdot} \chi_{[-1/2,1/2]}(\cdot)$$

Using the notation V^* for the analysis operator of the sampling sequence $\{u_j\}_{j=-n}^{n}$, we have

$$h = [\mathcal{F}f(\omega_{-n}), \ldots, \mathcal{F}f(\omega_n)]^T = V^* f.$$

Thus the approximation $\tilde{f}$ of f can be written as

$$\tilde{f} = T(V^*T)^{\dagger} h = Q_g h,$$

which means that we calculate the approximation $\tilde{f}$ to f by means of generalized sampling, see (5.20).

We approximate by means of generalized sampling, and obtain an average relative error of $\|\tilde{f} - f\|/\|f\| = 0.465$ (arithmetic mean over 5000 approximations). We observe that the relative approximation error even becomes larger than in the consistent case. The reason for that is the following. The operator Q_g is constructed in such a way that it deals well with noise, but not made for dealing with the part outside of the reconstruction space of the function to be reconstructed (i.e. cannot cope with the jumps of the signum function).

The operator Q_f (defined by (5.44)) has the smallest possible quasi-optimality constant μ (see Theorem 5.3.8), and therefore is exactly designed to deal with the part outside of the reconstruction space (to deal with the jumps). If we run the same experiment, approximate by means of frame independent sampling, i.e. $\tilde{f} = Q_f h$, we obtain an average relative error of $\|\tilde{f} - f\|/\|f\| = 0.146$ (arithmetic mean over 5000 approximations). The orthogonal projection $P_\mathcal{T} f$ onto the reconstruction space

$$\mathcal{T} = \operatorname{span}\{\mathrm{e}^{2\pi ik\cdot}\chi_{[-1/2,1/2]}\}_{k=-m}^{m} \tag{6.3}$$

has a relative error $\frac{\|P_\mathcal{T} f - f\|}{\|f\|} = 0.142$, which means that $Q_f V^*$ projects almost orthogonally onto the reconstruction space.

In Section 5.4 we also discussed mixtures Q_λ (Q_λ defined by (5.83)) between the operators Q_g and Q_f. In Table 6.1 we list the relative reconstruction error $\frac{\|f - Q_\lambda h\|}{\|f\|}$, the operator norm $\|Q_\lambda\|$, the quasi-optimality constant $\mu(Q_\lambda)$ and the condition number $\kappa(\Sigma_\lambda V^* T)$ of the matrix of the least squares problem (5.91) for

$$\lambda = 0,\ 0.0001,\ 0.001,\ 0.01,\ 0.1,\ \infty.$$

We recall that $Q_0 = Q_f$ (frame independent sampling) and $Q_\infty = Q_g$ (generalized sampling). We average (arithmetic mean) these quantities over 5000 approximations of the signum function from the random sampling frequencies (6.2). The operator norm $\|Q_\lambda\|$, $\mu(Q_\lambda)$ and the condition numbers of matrices of the least squares problems are only dependent on the the sampling and reconstruction vectors (but neither on the noise level nor on the function to be approximated). Therefore we only list these quantities in Table 6.1. In Figure 6.1 we plot the approximations obtained by Q_f, $Q_{0.01}$ and Q_g for one realization of the sampling frequencies (6.2).

We see in Table 6.1 that the operator Q_f has the smallest quasi-optimality constant $\mu(Q)$. By increasing the magnitude of λ, the quasi-optimality con-

	error	$\mu(Q)$	$\|Q\|$	κ
$Q_f\ (= Q_0)$	0.146	1.311	784.0	1.311
$Q_{0.0001}$	0.150	1.790	135.1	2.285
$Q_{0.001}$	0.162	2.601	94.3	3.980
$Q_{0.01}$	0.199	4.568	70.4	8.360
$Q_{0.1}$	0.286	9.116	56.8	19.678
$Q_g\ (= Q_\infty)$	0.459	26.070	49.4	80.546

Table 6.1: The relative approximation error, $\mu(Q)$ and $\|Q\|$, averaged over 5000 approximations, approximating the signum function. We use $m = 20$ and $n = 40$.

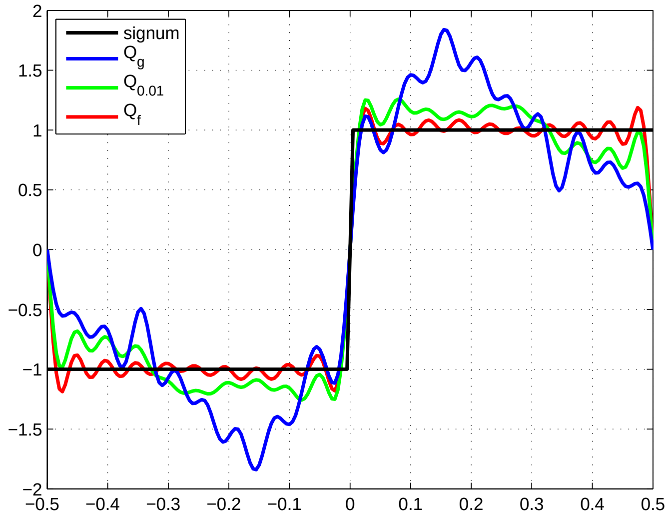

Figure 6.1: Approximation of the signum function. The approximations obtained by Q_f, $Q_{0.01}$ and Q_g have a relative error of 0.14, 0.21 and 0.41 respectively.

stant $\mu(Q_\lambda)$ is increasing. The quasi-optimality constant is a measure of how well the operator deals with the part of the function to reconstruct lying outside of the reconstruction space, and consequently, we expect a larger reconstruction error with increasing λ, which is in fact the case. We also observe that the condition number $\kappa((V^*V)^{\frac{\dagger}{2}}V^*T)$ of the matrix of the least squares problem (5.90) corresponding to frame independent sampling is equal to the quasi optimality constant $\mu(Q_f) = \frac{1}{\cos(\varphi_{\mathcal{T},\mathcal{V}})}$. The bound $\kappa((V^*V)^{\frac{\dagger}{2}} V^*T) \leqslant \frac{1}{\cos(\varphi_{\mathcal{T},\mathcal{V}})}$ is proven in Lemma 5.6.1. Therefore, if the value $\frac{1}{\cos(\varphi_{\mathcal{T},\mathcal{V}})}$ is close to one, preconditioning the least squares problem (5.89) by $(V^*V)^{\frac{\dagger}{2}}$ results in a least squares problem with condition number close to one. A small value $\frac{1}{\cos(\varphi_{\mathcal{T},\mathcal{V}})}$ is necessary for a stable reconstruction, and therefore the reconstruction space should always be chosen in such a way that this is the case.

We observe that with increasing λ, the operator norm of Q_λ is decreasing, and the approximation becomes less sensitive to noise present in the measurements. Since we used no noise in this experiment, the different magnitudes of the operator norm have no influence on the accuracy of the approximations, and therefore we obtain the most accurate approximations by the operator Q_f. This changes in the next setup.

6.2 Errors appearing after the sampling process

We repeat the same experiment with the only difference that we add some noise to the vectors

$$h = [\mathcal{F}f(\omega_{-n}), \ldots, \mathcal{F}f(\omega_n)]^T, \tag{6.4}$$

containing the Fourier samples, where ω_j, $j = -n, \ldots, n$ is as in (6.2). Specifically we reconstruct f from $d = h + v$, with vectors $v \in \mathbb{C}^{2n+1}$, where the real part and imaginary part of the vector entries of v are initially created with the standard normal distribution (that is with expectation 0 and standard deviation 1), and then the vectors are normalized so that

$$\frac{\|v\|}{\|h\|} = \varepsilon \tag{6.5}$$

for $\varepsilon = 0.01$.

6.2.1 Approximation of the signum function

First we approximate $f = \text{sign}$. In Table 6.2 we list the relative reconstruction error $\frac{\|Q_\lambda d - f\|}{\|f\|}$ for

$$\lambda = 0,\ 0.0001,\ 0.001,\ 0.01,\ 0.1,\ \infty,$$

averaged (arithmetic mean) over 5000 approximations calculated from the random vector $d = h + v$ containing the noisy measurements. For each realization of the sampling frequencies one realization of the noise vector is added.

	error
$Q_f\ (= Q_0)$	0.64
$Q_{0.0001}$	0.20
$Q_{0.001}$	0.19
$Q_{0.01}$	0.21
$Q_{0.1}$	0.29
$Q_g\ (= Q_\infty)$	0.46

Table 6.2: The relative approximation error averaged over 5000 approximations, approximating the signum function. We use $\varepsilon = 0.01$, $m = 20$ and $n = 40$.

Since the operator norm of Q_f is large, the small amount of noise destroys the accuracy of the approximations obtained by the operator Q_f. On average, the most accurate approximations are obtained by $Q_{0.001}$, a slightly regularized version of Q_f. The operator $Q_{0.001}$ has a quite small quasi-optimality constant μ (on average), such that this operator deals quite well with the jumps of $f = \text{sign}$, and due to the regularization parameter $\lambda = 0.001$ we obtain a smaller operator norm than the one of Q_f, and therefore the sensitivity to noise is reduced.

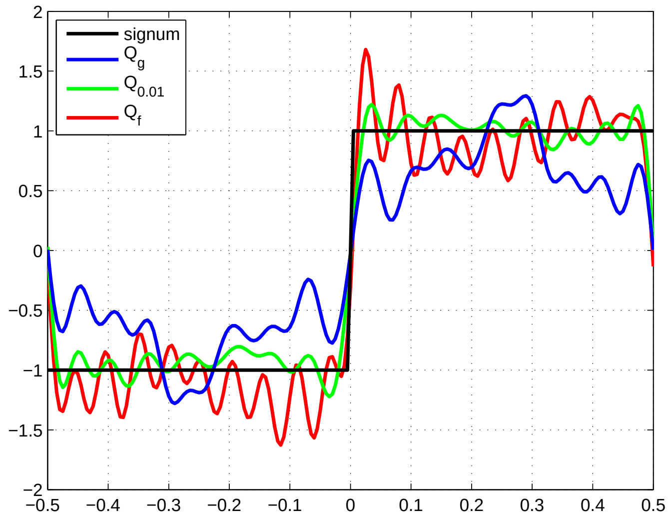

Figure 6.2: Approximation of the signum function. The approximations obtained by Q_f, $Q_{0.01}$ and Q_g have a relative error of 0.27, 0.17 and 0.42 respectively.

6.2.2 Approximation of an element lying inside the reconstruction space

When reconstructing a function f lying inside the reconstruction space there is no advantage in taking an operator with a small quasi-optimality constant. In this case, only the magnitude of the operator norm $\|Q\|$ is important. Therefore we expect to obtain more accurate reconstructions with the operator $Q_g(= Q_\infty)$ (having the smallest possible operator norm by Corollary 5.2.9) than with any of the others. As an example we reconstruct the trigonometric polynomial

$$f = \sum_{k=-m}^{m} \mathrm{e}^{2i\pi k \cdot} \chi_{[-1/2,1/2]} \tag{6.6}$$

from the random vector $d = h + v$, with h as in (6.4) and v as in (6.5) with $\varepsilon = 0.01$. In Table 6.3 we list the relative reconstruction error $\frac{\|Q_\lambda d - f\|}{\|f\|}$ for

$$\lambda = 0,\ 0.0001,\ 0.001,\ 0.01,\ 0.1,\ \infty,$$

averaged (arithmetic mean) over 5000 approximations calculated from the random vector d containing the noisy measurements.

	error
Q_f $(= Q_0)$	4.308
$Q_{0.0001}$	0.155
$Q_{0.001}$	0.116
$Q_{0.01}$	0.089
$Q_{0.1}$	0.073
Q_g $(= Q_\infty)$	0.064

Table 6.3: The relative approximation error averaged over 5000 approximations, approximating the trigonometric polynomial (6.6). We use $\varepsilon = 0.01$, $m = 20$ and $n = 40$.

We see in Table 6.3, that, as expected, the operator Q_g has the smallest relative error. Again, this is the case since the function to be approximated lies inside the reconstruction space.

6.3 Errors appearing before the sampling process

We now assume that we are given a set of Fourier measurements of a perturbation of $f \in \mathcal{H}$

$$\tilde{h} = [\mathcal{F}(f + \Delta f)(\omega_{-n}), \ldots, \mathcal{F}(f + \Delta f)(\omega_n)]^T. \tag{6.7}$$

According to Theorem 5.3.10, we expect to obtain more accurate reconstructions with the operator Q_f than with any of the others, independently of the position of the element to be approximated. The sampling frequencies ω_j are chosen as before, i.e., $\omega_j = j + \delta_j$, $j = -n, \ldots, n$, with δ_j drawn from the uniform distribution in the interval $[-1, 1]$. For each set of sampling frequencies we choose Δf as a trigonometric polynomial

$$\Delta f = \sum_{k=-n}^{n} a_k \mathrm{e}^{2i\pi k \cdot} \chi_{[-1/2,1/2]}.$$

Let $a \in \mathbb{C}^{2n+1}$ denote the vector

$$a := [a_{-n}, \ldots, a_n]^T.$$

The real part and imaginary part of the entries of a are initially created with the standard normal distribution (that is with expectation 0 and standard deviation 1), and then the vectors are normalized, so that

$$\frac{\|a\|}{\|f\|} = \frac{\|\Delta f\|}{\|f\|} = \varepsilon, \tag{6.8}$$

for $\varepsilon = 0.1$.

6.3.1 Approximation of the signum function

First we approximate the signum function $f = \mathrm{sgn}$. In Table 6.4 we list the relative approximation error $\frac{\|f - Q_\lambda \tilde{h}\|}{\|f\|}$ for

$$\lambda = 0,\ 0.0001,\ 0.001,\ 0.01,\ 0.1,\ \infty,$$

averaged (arithmetic mean) over 5000 approximations calculated from the random vector $\tilde{h}$ defined by (6.7). In Figure 6.3 we plot the approximations obtained by Q_f and Q_g for one realization of the Fourier measurements (6.7).

	error
$Q_f\ (= Q_0)$	0.154
$Q_{0.0001}$	0.159
$Q_{0.001}$	0.170
$Q_{0.01}$	0.206
$Q_{0.1}$	0.296
$Q_g\ (= Q_\infty)$	0.508

Table 6.4: The relative approximation error averaged over 5000 approximations, approximating the signum function. We use $\varepsilon = 0.1$ in (6.8), $m = 20$ and $n = 40$.

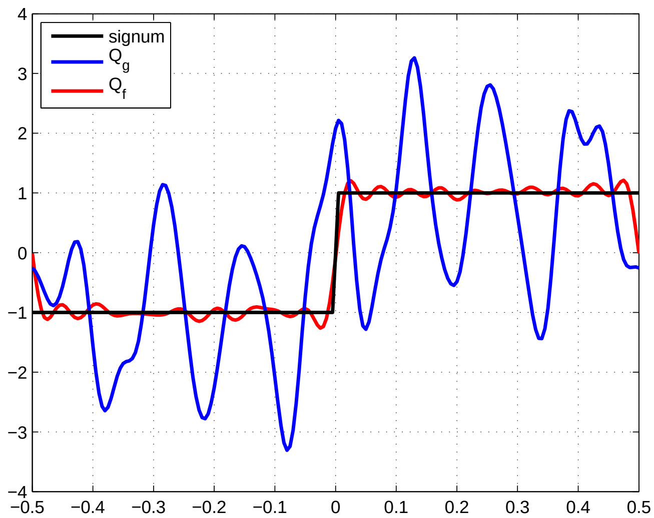

Figure 6.3: Approximation of the signum function. We use $\varepsilon = 0.1$ in (6.8), $m = 20$ and $n = 40$. The approximations obtained by Q_f and Q_g have a relative error of 0.15 and 1.28 respectively.

6.3.2 Approximation of an element lying inside the reconstruction space

Next we approximate the trigonometric polynomial

$$f = \sum_{k=-m}^{m} \mathrm{e}^{2i\pi k \cdot} \chi_{[-1/2,1/2]}. \tag{6.9}$$

In Table 6.5 we list the relative approximation error $\frac{\|f-Q_\lambda \tilde{h}\|}{\|f\|}$ for

$$\lambda = 0,\ 0.0001,\ 0.001,\ 0.01,\ 0.1,\ \infty,$$

averaged (arithmetic mean) over 5000 approximations calculated from the random vector $\tilde{h}$ defined by (6.7). In Figure 6.4 we plot the approximations obtained by Q_f and Q_g for one realization of the Fourier measurements (6.7).

	error
$Q_f\ (= Q_0)$	0.071
$Q_{0.0001}$	0.073
$Q_{0.001}$	0.077
$Q_{0.01}$	0.088
$Q_{0.1}$	0.121
$Q_g\ (= Q_\infty)$	0.285

Table 6.5: The relative approximation error approximating the trigonometric polynomial (6.9). We use $\varepsilon = 0.1$ in (6.8), $m = 20$ and $n = 40$.

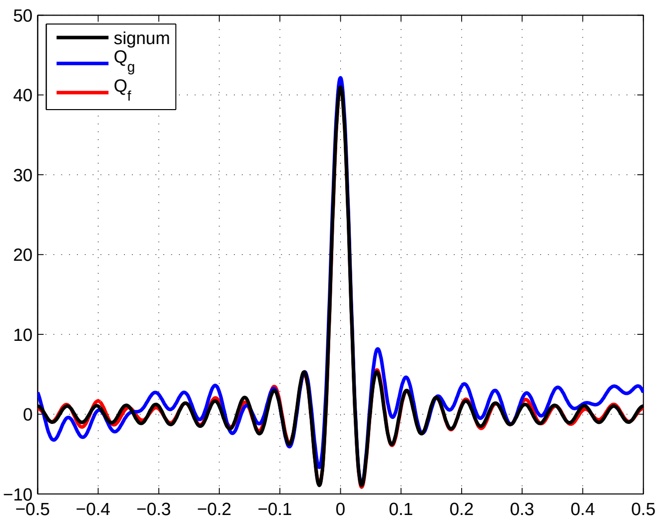

Figure 6.4: Approximation of the trigonometric polynomial (6.9). We use $\varepsilon = 0.1$ in (6.8), $m = 20$ and $n = 40$. The approximations obtained by Q_f and Q_g have a relative error of 0.08 and 0.45 respectively.

We observe that for the trigonometric polynomial (6.9) the operator Q_f yields the most accurate reconstruction, when reconstructing from Fourier measurements of the perturbed function. This is not the case when we reconstruct the same trigonometric polynomial from perturbed Fourier measurements. Therefore even for functions inside the reconstruction space, the operator Q_f is the optimal choice when systematic errors appear before the sampling process.

6.4 Summary of the experiments

Summarizing, when the sampling frame is tight or there are as many measurements as reconstruction vectors, then there is no difference between the reconstruction operators Q_λ.

If there are more measurements than reconstruction vectors and the sampling frame is not tight (the sampling frequencies are irregular), then the reconstructions differ and we have observed the following.

1. The operator Q_g gives the most accurate approximations if the function to be approximated is inside the reconstruction space, and we reconstruct from perturbed measurements (measurement error appearing after the sampling process).

2. The operator Q_f gives the most accurate approximations if the function to be approximated has a part outside of the reconstruction space and no measurement error is present.

3. The operator Q_f gives the most accurate approximations if we assume systematic errors appearing before the sampling process, independent of the amount of the function to reconstruct lying outside of the reconstruction space (no measurement error after the sampling process is assumed).

4. A mixture Q_λ between the two operators gives more accurate approximations than Q_g or Q_f, if the function to be reconstructed has a

part outside of the reconstruction space, and we approximate from perturbed measurements (measurement error appearing after the sampling process).

Chapter 7

A hierarchical reconstruction algorithm

In this chapter we consider the problem of approximating a non-bandlimited function represented by a finite number of compactly supported generating functions in wireless sensor networks. Typically in literature it is assumed that the function to be approximated is strictly bandlimited, see for example [40, 45, 49]. As explained in [54], this assumption is often unrealistic. For example electrostatic fields, gravitation fields and diffusion fields are in general non-bandlimited. Therefore the authors of [54] propose to use integer shifts of B-splines as a function model. Since B-splines have compact support, we can preprocess the sensor measurements locally, and send the resulting data (rather than sensor measurements) to a global fusion center for further processing. In [54] this is done by solving least squares problems locally, and sending the resulting expansion coefficients rather than the sensor measurements to a global fusion center.

We use the idea of preprocessing the sensor measurements locally to develop a new hierarchical algorithm. If the signal is oversampled by a factor s, then by our approach the number of required long distance transmissions is reduced roughly by a factor s (in comparison to sending the sensor measurements directly). This also holds true for the approach presented in [54]. Their approach has some drawbacks, the biggest being that their algorithm is not applicable if one of the local least squares problems is under-determined.

7.1 Description of the setup

We consider the problem of estimating a continuous function $f \in L^2(\mathbb{R}^d)$, $d \in \mathbb{N}$, from noisy point evaluations

$$d_i := f(x_i) + \delta_i, \quad i = 1, \dots, n,$$

where $\delta_i \in \mathbb{C}$ is measurement noise. The value d_i can be viewed as measurement of the ith sensor positioned at x_i.

As a function model we choose a finite dimensional subspace $\mathcal{W} \subset L^2(\mathbb{R}^d)$, generated by continuous, linearly independent and compactly supported generating functions $g_k \in L^2(\mathbb{R}^d, \mathbb{C})$, $k = 1, \dots, m$,

$$\mathcal{W} = \left\{ \sum_{k=1}^{m} c_k g_k : c_k \in \mathbb{C} \right\}. \tag{7.1}$$

This means that our intention is to calculate coefficients $\{\hat{c}_k\}_{k=1}^m$ from the point evaluations, such that

$$\tilde{f} = \sum_{k=1}^{m} \hat{c}_k g_k$$

is a good approximation to f. It should be mentioned that we do not assume that f is an element of $\mathcal{W}$.

7.2 Solving the global least squares problem

Let us describe the standard least squares approach to this problem, see for example [36, Algorithm 1].

Algorithm 7.2.1.

Input: *We assume that the noisy point evaluations*

$$\boldsymbol{d} = [f(x_1) + \delta_1, \dots, f(x_n) + \delta_n]^T,$$

and the sensor positions $\{x_i\}_{i=1}^n$ *are known.*

Step1: *Calculate the matrix* $A \in \mathbb{C}^{n\times m}$

$$A(i,k) = g_k(x_i), \quad i = 1,\dots,n, \; k = 1,\dots,m. \tag{7.2}$$

Step2: *Solve the least squares problem*

$$\hat{\boldsymbol{c}} = \arg\min_{\boldsymbol{c}} \|A\boldsymbol{c} - \boldsymbol{d}\|. \tag{7.3}$$

Step3: *The approximation* $\tilde{f}$ *of* f *is given by*

$$\tilde{f} = \sum_{k=1}^{m} \hat{c}_k g_k. \tag{7.4}$$

We observe that $\tilde{f}$ is the least squares solution

$$\tilde{f} = \arg\min_{g\in\mathcal{W}} \sum_{j\in\mathbb{N}} |g(x_j) - d_j|^2 . \tag{7.5}$$

We assume that the matrix A is injective. In this case the least squares problem (7.3) is overdetermined. Furthermore, in the absence of noise every element $f \in \mathcal{W}$ can be reconstructed from the point evaluations $\{f(x_i)\}_{i=1}^n$ by Algorithm 7.2.1. A necessary condition for A being injective is that $n \geqslant m$, i.e. we assume that there are at least as many measurements as generating functions. Another simple necessary condition is that for each k there must be at least one sensor located in the support of g_k. Otherwise the matrix A has a zero column and is consequently rank-deficient.

The equivalent statement for infinite dimensional sampling theory is that there exist constant $A, B > 0$ such that for every $f \in \mathcal{W}$

$$A\|f\|^2 \leqslant \sum_{i\in\mathbb{N}} |f(x_i)|^2 \leqslant B\|f\|^2.$$

In this case $X = \{x_i\}_{i\in\mathbb{N}}$ is called a set of stable sampling. In order to guarantee that a stable reconstruction is possible, this property has to be proven. As mentioned in the introduction, stable sets of sampling for the space $L^2(\mathbb{R}) \cap B_{[-\frac{1}{2},\frac{1}{2}]}$ are characterized in terms of the Beurling density. For our numerical experiments we use integer shifts of a one dimensional B-spline. For this type of generating functions, stable sampling sets are well understood, see [12, 36]. For a result in the two dimensional case, see [54].

It is important to realize that in this least squares approach all noisy point evaluations $\{d_i\}_{i=1}^n$ have to be sent to one point, called the fusion center. As already mentioned in the introduction, our intention is to get rid of this redundancy, maintaining the advantage of oversampling.

7.3 Special case of a reproducing kernel Hilbert space

This subsection is not needed for the description of our hierarchical algorithm, but it gives some insight about why it is a good strategy to calculate the coefficients $\hat{\mathbf{c}}$ for the reconstruction $\tilde{f}$ by (7.3).

What follows is the theory presented in [10], and some additional aspects worked out in Section 5.2, applied to the problem of reconstructing element f from point evaluations $\{f(x_j)\}_{j\in\mathbb{N}}$ in a reproducing kernel Hilbert space.

Definition 7.3.1. *Let $\mathcal{H}$ be a Hilbert space of complex-valued functions on a set X. We call $\mathcal{H}$ a reproducing kernel Hilbert space if for every $x \in X$ the point evaluation $\delta_x : \mathcal{H} \to \mathbb{C}$, $\delta_x(f) := f(x)$ is continuous.*

Definition 7.3.2. *Let $\mathcal{H}$ be a reproducing kernel Hilbert space. For $x \in X$*

we define $K_x \in \mathcal{H}$ as the unique element in $\mathcal{H}$, such that for every $f \in \mathcal{H}$

$$f(x) = \delta_x(f) = \langle f,\ K_x \rangle. \tag{7.6}$$

The function $K : X \times X \to \mathbb{C}$,

$$K(x, y) := \langle K_y, K_x \rangle = K_x(y),$$

is called the reproducing kernel for the Hilbert space $\mathcal{H}$.

If $\mathcal{H}$ is a reproducing kernel Hilbert space, by the Riesz representation theorem, for every $x \in X$ there exists a unique element $K_x \in \mathcal{H}$, such that (7.6) holds for every $f \in \mathcal{H}$. The function K_x defined by (7.6) is called the point-evaluation function at the point x.

Let us assume that the space $\mathcal{W} \subset L^2(\mathbb{R}^d)$ used as function model is a subspace of a reproducing kernel Hilbert space $\mathcal{H} \subset L^2(\mathbb{R}^d, \mathbb{C})$. In this case for every x_i, $i = 1, \ldots, n$, there exists a unique element $u_i = K_{x_i} \in \mathcal{H}$ with the property that for every $g \in \mathcal{H}$

$$g(x_i) = \langle g, u_i \rangle.$$

Let $\mathcal{V} = \text{span}\{u_i\}_{i=1}^n$. We denote by V the corresponding synthesis operator, by V^* the corresponding analysis operator, and by S the corresponding frame operator. By T we denote the synthesis operator of $\{g_k\}_{i=1}^m$, so that $\mathcal{W} = \text{span}\{g_k\}_{i=1}^m = \mathcal{R}(T)$. With this notation $A = V^*T$. We define the operator Q_g by $Q_g = T(V^*T)^\dagger = TA^\dagger$. If the function f to be approximated is an element of $\mathcal{H}$, then the approximation $\tilde{f}$ of f defined by (7.4), and the coefficients $\hat{c}_k$ defined by (7.5) can be written in the form

$$\tilde{f} = Q_g \mathbf{d} = TA^\dagger \mathbf{d}. \tag{7.7}$$

This shows that the approximation $\tilde{f}$ is obtained by means of generalized sampling. Since the functions $\{g_k\}_{k=1}^m$ are assumed to be linearly independent, the assumption A being injective is equivalent to $\cos(\varphi_{\mathcal{W},\mathcal{V}}) > 0$ (where the angle $\varphi_{\mathcal{W},\mathcal{V}}$ is defined by (3.9)). Consequently by (5.25) the bounded operator Q_g is quasi-optimal. By (3.6) for $f \in \mathcal{H}$

$$\|f - Q_g(V^* f + \boldsymbol{\delta})\| \leqslant \mu(Q_g)\|f - P_{\mathcal{W}} f\| + \|Q_g\|\|\boldsymbol{\delta}\|. \tag{7.8}$$

By Corollary 5.2.9, if $Q : \mathbb{C}^n \to \mathcal{H}$ is an operator with $QV^* f = f$ for all $f \in \mathcal{W}$ (i.e. Q recovers elements in $\mathcal{W}$ in the absence of noise), then

$$\|Q\| \geqslant \|Q_g\|.$$

Therefore the calculation of the coefficients by (7.3) gives the most stable reconstruction in the sense above.

7.4 Preprocessed data

In this section we describe how to calculate coefficients for the approximation $\tilde{f} = \sum_{k=1}^{m} \hat{c}_k g_k$ of f from inner products

$$s_j = \langle \mathbf{d}, \mathbf{v}_j \rangle, \quad j = 1, \ldots, r, \tag{7.9}$$

of the vector $\mathbf{d}$ containing the noisy point evaluations. This means that instead of estimating the coefficients $\{\hat{c}_k\}_{k=1}^m$ from noisy point evaluations $f(x_i) + \delta_i$, $i = 1, \ldots, n$, we estimate them from $\{s_j\}_{j=1}^r$. This is used in Section 7.5 to describe our hierarchical algorithm.

Algorithm 7.4.1.
Input: *We assume that the vectors* $\{\boldsymbol{v}_j \in \mathbb{C}^{n \times 1}\}_{j=1}^r$ *, the inner products*

$$\boldsymbol{s} = [\langle \boldsymbol{d}, \boldsymbol{v}_1 \rangle, \ldots, \langle \boldsymbol{d}, \boldsymbol{v}_r \rangle]^T$$

and the sensor positions $\{x_i\}_{i=1}^n$ *are known.*

Step1: *Calculate the matrix*

$$(T^* T)^{\frac{\dagger}{2}} T^* A$$

and the vector

$$(T^*T)^{\frac{\dagger}{2}}\boldsymbol{s},$$

where the matrix T *is defined by*

$$T(i,k) = \boldsymbol{v}_k(i), \quad i = 1,\dots,n, \;\; k = 1,\dots,r, \tag{7.10}$$

and A *is the matrix of the global least squares problem, defined by* (7.2)*.*

Step2: *Solve the least squares problem*

$$\hat{\boldsymbol{c}} = \arg\min_{\boldsymbol{c}} \|(T^*T)^{\frac{\dagger}{2}}T^*A\boldsymbol{c} - (T^*T)^{\frac{\dagger}{2}}\boldsymbol{s}\|. \tag{7.11}$$

Step3: *The approximation* $\tilde{f}$ *of* f *is given by*

$$\tilde{f} = \sum_{k=1}^{m} \hat{c}_k g_k.$$

Let us describe the idea behind Algorithm 7.4.1. Let $\mathbf{h}$ denote the vector containing the point evaluations

$$\mathbf{h} = [f(x_1),\dots,f(x_n)]^T, \tag{7.12}$$

and let $\boldsymbol{\delta}$ denote the vector containing the noise

$$\boldsymbol{\delta} = [\delta_1,\dots,\delta_n]^T,$$

such that $\mathbf{d} = \mathbf{h} + \boldsymbol{\delta}$ is the vector containing the noisy point evaluations of f. We first observe that the least squares solution

$$\hat{\mathbf{c}} = \arg\min_{\mathbf{c}} \|A\mathbf{c} - \mathbf{d}\|. \tag{7.13}$$

coincides with the least squares solution

$$\hat{\mathbf{c}} = \arg\min_{\mathbf{c}} \|A\mathbf{c} - P_{\mathcal{R}(A)}\mathbf{d}\| = \arg\min_{\mathbf{c}} \|A\mathbf{c} - (P_{\mathcal{R}(A)}\mathbf{h} + P_{\mathcal{R}(A)}\boldsymbol{\delta})\|$$

The idea is to calculate an approximation $\tilde{\mathbf{h}}$ to $P_{\mathcal{R}(A)}\mathbf{h}$ from s_j, and to estimate the coefficients $\hat{c}_k$ for the reconstruction $\tilde{f}$ by solving the least squares problem

$$\hat{\mathbf{c}} = \arg\min_{\mathbf{c}} \|A\mathbf{c} - \tilde{\mathbf{h}}\|. \tag{7.14}$$

Since $\tilde{\mathbf{h}}$ is an approximation to $P_{\mathcal{R}(A)}\mathbf{h}$, it is reasonable to choose $\mathcal{R}(A)$ as the reconstruction space, and the columns of A as reconstruction vectors. We denote the columns of A by $\{\mathbf{b}_k\}_{k=1}^m$. Consequently, A is the synthesis operator of the sequence $\{\mathbf{b}_k\}_{k=1}^m$. We observe that in (7.9), s_j are inner products of a perturbation $\mathbf{h}+\boldsymbol{\delta}$ of $\mathbf{h}$ with vectors $\mathbf{v}_j$. Therefore the sampling vectors are

$$\mathbf{v}_j, \quad j = 1, \ldots, r.$$

We recall that the operator Q_f is optimal when reconstructing from inner products of a perturbed function. Since the values s_i are inner products of a perturbation $\mathbf{h} + \boldsymbol{\delta}$ of $\mathbf{h}$ with vectors $P_{C_j}\mathbf{v}_i$, $i = 1, \ldots, r_j$, $j = 1, \ldots, L$, we use the operator Q_f to calculate the approximation $\tilde{\mathbf{h}}$ to $P_{\mathcal{R}(A)}\mathbf{h}$.

As in Algorithm 7.4.1, we denote the denote the synthesis operator of the sampling vectors by $T \in \mathbb{C}^{n\times r}$, and the corresponding analysis operator by T^*. With this notation the vector $\mathbf{s} \in \mathbb{C}^{r\times 1}$ can be written as

$$\mathbf{s} = T^*(\mathbf{h} + \boldsymbol{\delta}) = T^*\mathbf{d}. \tag{7.15}$$

We define the operator Q_f by

$$Q_f = A\left((T^*T)^{\frac{\dagger}{2}}T^*A\right)^{\dagger}(T^*T)^{\frac{\dagger}{2}},$$

and the reconstruction $\tilde{\mathbf{h}}$ of $P_R\mathbf{h}$ by

$$\tilde{\mathbf{h}} = Q_f\mathbf{s} = Q_fT^*(\mathbf{h} + \boldsymbol{\delta}) = P_{R,P_{W}(R)^{\perp}}(\mathbf{h} + \boldsymbol{\delta}). \tag{7.16}$$

We use the notation

$$R = \mathcal{R}(A) \quad \text{and} \quad W = \mathcal{R}(T) = \operatorname{span}\{\mathbf{v}_j\}_{j=1}^r.$$

We assume that $\cos(\varphi_{R,W}) > 0$. Since the columns of A are linearly independent, the assumption that T^*A is injective is equivalent to the assumption that $\cos(\varphi_{R,W}) > 0$, see Lemma 3.4.9. By Theorem 5.3.8, $Q_f T^*$ is the operator with the smallest possible quasi-optimality constant μ, namely $\mu(Q_f) = \frac{1}{\cos(\varphi_{R,W})}$. Lemma 5.3.11 shows that for $\mu = \mu(Q_f) = \frac{1}{\cos(\varphi_{R,W})}$ it holds

$$\|P_{\mathcal{R}(A)}\mathbf{h} - Q_f\mathbf{s}\| \leqslant \sqrt{\mu^2 - 1}\|\mathbf{h} - P_{\mathcal{R}(A)}\mathbf{h}\| + \mu\|\boldsymbol{\delta}\|, \tag{7.17}$$

and that for any other reconstruction operator $Q \in \mathbb{C}^{n\times r}$ with

$$\|P_{\mathcal{R}(A)}\mathbf{h} - Q\mathbf{s}\| \leqslant \alpha\|\mathbf{h} - P_{\mathcal{R}(A)}\mathbf{h}\| + \beta\|\boldsymbol{\delta}\|,$$

we have $\alpha \geqslant \sqrt{\mu^2 - 1}$ and $\beta \geqslant \mu$.

Solving the least squares problem (7.14) with $\tilde{\mathbf{h}}$ defined by (7.16), we obtain the coefficients

$$\hat{\mathbf{c}} = A^\dagger A\left((T^*T)^{\frac{\dagger}{2}}T^*A\right)^\dagger (T^*T)^{\frac{\dagger}{2}}\mathbf{s}. \tag{7.18}$$

for the approximation $\tilde{f}$ of f. From (7.16) we infer that

$$\hat{\mathbf{c}} = A^\dagger P_{R,P_W(R)^\perp}(\mathbf{h} + \boldsymbol{\delta}).$$

Therefore the coefficients for the approximation $\tilde{f}$ are independent of the particular choice of the sampling vectors, but only dependent on W, the linear span.

Since $A^\dagger A = P_{\mathcal{R}(A^*)}$ and

$$\mathcal{R}\left(\left((T^*T)^{\frac{\dagger}{2}}T^*A\right)^\dagger\right) = \mathcal{R}\left(A^*T(T^*T)^{\frac{\dagger}{2}}\right) \subset \mathcal{R}(A^*),$$

the matrix $A^\dagger A$ in (7.18) can be dropped, and

$$\hat{\mathbf{c}} = \left((T^*T)^{\frac{\dagger}{2}}T^*A\right)^\dagger (T^*T)^{\frac{\dagger}{2}}\mathbf{s}. \tag{7.19}$$

This is exactly the approximation $\tilde{f}$ of f as described in Algorithm 7.4.1. If

$$R = \mathcal{R}(A) \subset \text{span}\{\mathbf{v}_j : j = 1, \ldots r\} = R(T) = W, \tag{7.20}$$

then

$$\hat{\mathbf{c}} = A^\dagger P_R \mathbf{d} = A^\dagger \mathbf{d},$$

and the vector $\hat{\mathbf{c}}$ defined by (7.11) coincides with the solution of the global least squares problem (7.13). Consequently the vectors $\{\mathbf{v}_j\}_{j\in\mathbb{N}}$ should fulfill (7.20).

In Section 5.6.2 we discussed two cases when the operator $(T^*T)^{\frac{\dagger}{2}}$ in (7.11) can be dropped, and still the same reconstruction $\tilde{f}$ is obtained. The first one is when $\dim(\mathcal{R}(A)) = \dim(\mathcal{R}(T))$, i.e., in the case of linearly independent vectors $\mathbf{v}_j$, the number of vectors $\mathbf{v}_j$ coincides with the number of columns of A, i.e., $r = m$. The second one is when the vectors $\{\mathbf{v}_j\}_{j=1}^r$ form a tight frame sequence. In these two cases instead of (7.11) we can use the simpler formula

$$\hat{\mathbf{c}} = \arg\min_{\mathbf{c}} \|T^*A\mathbf{c} - \mathbf{s}\|. \tag{7.21}$$

7.5 Hierarchical Algorithm

In this section we describe our hierarchical algorithm in the absence of transmission noise. Let $f \in L^2(\mathbb{R}^d)$ denote the continuous function to be approximated from the noisy point evaluations

$$\mathbf{d}(i) := f(x_i) + \delta_i, \quad i = 1, \ldots, n.$$

We assume that we have connected sets $B_j \subset \mathbb{R}^d$, $j = 1, \ldots L$, such that

$$\bigcup_{j=1,\ldots,L} B_j = \text{supp}(f).$$

For each set B_j, $j = 1, \dots L$, we cluster the sensors located in B_j. By C_j we denote the set of the indices of the sensors located in B_j

$$C_j = \{i : x_i \in B_j\}.$$

This can be interpreted that sensors in the same cluster are close to each other, and each sensor is contained in at least one of the clusters. In every cluster the noisy point evaluations of the sensors are transferred to one sensor node, termed the cluster head. This means in every cluster C_j we have access to the data

$$\mathbf{d}(i) = f(x_i) + \delta_i, \quad i \in C_j.$$

Let $\{e_j\}_{j=1}^n$ denote the canonical basis of $\mathbb{C}^n$, and let P_{C_j} denote the orthogonal projection onto $\mathrm{span}\{e_j : j \in C_j\}$

$$(P_{C_j}\mathbf{b})_i = \begin{cases} \mathbf{b}_i & \text{if } i \in C_j, \\ 0 & \text{otherwise.} \end{cases}$$

In every cluster C_j, $j = 1, \dots, L$, we have access to the data $\mathbf{d}(i)$, $i \in C_j$. Therefore for every j and every $\mathbf{v} \in \mathbb{C}^n$ we can calculate

$$\langle P_{C_j}\mathbf{d}, \mathbf{v}\rangle = \langle \mathbf{d}, P_{C_j}\mathbf{v}\rangle.$$

The idea is to send processed data of the form $\langle \mathbf{d}, P_{C_j}\mathbf{v}\rangle$ from each cluster head to the fusion center. Precisely, assume that

$$\mathbf{v}_{i,j} \in \mathrm{span}\{e_j : j \in C_j\}, \quad i \in I_j,$$

and

$$s_{i,j} := \langle \mathbf{d}, \mathbf{v}_{i,j}\rangle, \quad i \in I_j, \tag{7.22}$$

then $s_{i,j}$ uses only point evaluations from sensors inside of the jth cluster. Therefore only the sensors inside each cluster need to transmit the measured data to the cluster head. Our aim is to send roughly m numbers $s_{i,j}$, where m is the number of generating functions. If the signal is oversampled by a factor of s, i.e. $n = sm$, then instead of n transmissions, this approach only requires roughly $\frac{n}{s}$ transmissions.

7.5.1 Approximation by solving the normal equations

We recall that if $\mathcal{R}(A) \subset \mathcal{R}(T)$ we obtain the same coefficients by (7.11) as by sending all the noisy point evaluations $\mathbf{d}(i) := f(x_i) + \delta_i$, $i = 1, \ldots, n$, and solving the least squares problem (7.13). If either $\dim(R) = \dim(W)$ or $\{\mathbf{v}_{i,j}\}$ forms a tight frame sequence, then we can use the simpler formula (7.21) for the calculation of the expansion coefficients of $\tilde{f}$.

Since the generating functions g_k, $k = 1, \ldots, m$ have a compact support, the reconstruction space $\mathcal{W}$ has a local nature. In this case it is possible to find reasonable sets B_j, and vectors $\mathbf{v}_{i,j} \in \mathbb{C}^n$, such that $\mathcal{R}(A) \subset \mathcal{R}(T)$ is fulfilled. An interesting choice for the sampling vectors $\mathbf{v}_{i,j}$, is $\{\mathbf{v}_{i,j}\} = \{\mathbf{b}_k\}_{k=1}^m$, the columns of A. In this case $T = A$. Therefore clearly $\mathcal{R}(A) \subset \mathcal{R}(T)$ and $\dim(\mathcal{R}(A)) = \dim(\mathcal{R}(T))$.

Having transmitted the inner products $\langle \mathbf{d}, \mathbf{b}_k \rangle$, $k = 1, \ldots m$, to the global fucion center, the reconstruction algorithm proceeds as follows.

Algorithm 7.5.1.
Input: *We assume that the inner products*

$$\boldsymbol{s} = [\langle \boldsymbol{d}, \boldsymbol{b}_1 \rangle, \ldots, \langle \boldsymbol{d}, \boldsymbol{b}_m \rangle]^T = A^* \boldsymbol{d}$$

and the sensor positions $\{x_i\}_{i=1}^n$ *are known.*

Step1: *Calculate the matrix* A^*A.

Step2: *Solve the normal equations*

$$A^* A \hat{\boldsymbol{c}} = A^* \boldsymbol{d}.$$

Step3: *The approximation $\tilde{f}$ of f is given by*

$$\tilde{f} = \sum_{k=1}^{m} \hat{c}_k g_k.$$

Let us now describe how to choose the sets $\{B_j\}_{j=1}^L$ in order to calculate the inner products $\langle \mathbf{d}, \mathbf{b}_k \rangle$, $k = 1, \ldots m$, locally. This can be done by choosing a partition $M_1, \ldots, M_L$ of $\{1, \ldots, m\}$, and setting

$$B_j = \bigcup_{i \in M_j} \operatorname{supp}(g_i), \quad j = 1, \ldots, L.$$

The supports of the generating functions g_i, g_k with i, k in the same set M_j should be close to each other, since otherwise for the local collection of the point evaluations long distance transmission are necessary.

We assume that at one point in B_j, the sensor locations $C_j = \{i : x_i \in B_j\}$ and measurements $f(x_i) + \delta_i$, $i \in B_j$ are known. It is important to realize that we do not assume that we know the sensor locations and measurements from sensors outside the region B_j. In each set B_j, $j = 1, \ldots, L$, we calculate the inner products

$$s_{i,j} = \langle P_{C_j} \mathbf{d}, \mathbf{b}_i \rangle = \langle \mathbf{d}, P_{C_j} \mathbf{b}_i \rangle = \langle \mathbf{d}, \mathbf{b}_i \rangle, \quad i \in M_j,$$

and transmit them to the fusion center.

Since $M_1, \ldots, M_L$ is a partition of $\{1, \ldots, m\}$, with this setup a total number of m numbers, namely $\{\langle \mathbf{d}, \mathbf{b}_i \rangle\}_{i=1}^m$, are transmitted to the fusion center. Therefore, if the signal is oversampled by a factor of s, i.e. $n = sm$, then instead of n transmissions, this approach only requires $\frac{n}{s}$ transmissions.

It should be mentioned that solving a least squares problem by first calculating the normal equations should be avoided, since $\kappa(A^*A) = \kappa(A)^2$. Therefore the condition number is squared, which causes stability problems whenever the condition number of A is large. A more stable way to solve an overdetermined least squares problem is solving it via a pivoted QR-factorization using Householder transformations, see [31].

7.5.2 Orthonormal systems locally

We see in our numerical experiments that if we assume transmission noise, when sending the inner products $\{\langle \mathbf{d}, \mathbf{b}_i \rangle\}_{i=1}^m$, then solving the normal equations causes stability problems whenever the condition number $\kappa(A)$ is large. This problem is solved by calculating an orthonormal system $\{\mathbf{v}_k\}_{k=1}^r$ locally, with the property that $\mathcal{R}(A) \subset \operatorname{span}\{\mathbf{v}_k : k = 1, \ldots, r\}$ followed by transmitting the inner products $\langle \mathbf{d}, \mathbf{v}_k \rangle$, $k = 1, \ldots r$. Before we discuss how this is accomplished, we first describe the reconstruction algorithm after transmission of the inner products $\langle \mathbf{d}, \mathbf{v}_k \rangle$, $k = 1, \ldots r$, to the fusion center.

Algorithm 7.5.2.
Input: *We assume that the vectors* $\{\boldsymbol{v}_j \in \mathbb{C}^{n\times 1}\}_{j=1}^r$ *, the inner products*

$$\boldsymbol{s} = [\langle \boldsymbol{d}, \boldsymbol{v}_1 \rangle, \ldots, \langle \boldsymbol{d}, \boldsymbol{v}_r \rangle]^T$$

and the sensor positions $\{x_i\}_{i=1}^n$ *are known. Furthermore let* $\{\boldsymbol{v}_k\}_{k=1}^r$ *be an orthonormal system with*

$$\mathcal{R}(A) \subset \operatorname{span}\{\boldsymbol{v}_k : k = 1, \ldots, r\}.$$

Step1: *Calculate the matrix* T^*A *where the matrix* T *is defined by*

$$T(i,k) = \boldsymbol{v}_k(i), \quad i = 1, \ldots, n, \; k = 1, \ldots, r.$$

Step2: *Solve the least squares problem*

$$\hat{\boldsymbol{c}} = \arg\min_{\boldsymbol{c}} \|T^*A\boldsymbol{c} - \boldsymbol{s}\|.$$

Step3: *The approximation* $\tilde{f}$ *of* f *is given by*

$$\tilde{f} = \sum_{k=1}^m \hat{c}_k g_k.$$

Next we describe how to calculate the orthonormal system $\{\mathbf{v}_k\}_{k=1}^r$. Let $B_1, \dots B_L$ be a partition of $B = \bigcup_{k=1,\dots,m} \text{supp}(g_k)$ and let the set

$$\{C_j\}_{j=1}^L = \{i : x_i \in B_j\}_{j=1}^L$$

be the corresponding partition of $\{1, \dots, n\}$. Furthermore let I_j be the indices of the supports of the generating function g_k containing a sensor of cluster B_j, i.e.,

$$I_j := \{k : \text{supp}(g_k) \cap \{x_i : x_i \in B_j\} \neq \emptyset\}. \tag{7.23}$$

We define the vectors $\mathbf{w}_{k,j} \in \mathbb{C}^n$ by

$$\mathbf{w}_{k,j} := P_{C_j} \mathbf{b}_k,$$

where $\mathbf{b}_k$ are the columns of A. We define the sets D_j, $j = 1, \dots, L$ by

$$D_j := \text{span}\{\mathbf{w}_{k,j} : k \in I_j\}.$$

Since the sets $C_1, \dots C_L$ are mutually disjoint

$$D_m \perp D_k \quad \text{for } m \neq k.$$

For each set D_j we choose an orthonormal basis $\{\mathbf{v}_{i,j}\}$, $i \in K_j$, and transmit

$$s_{i,j} = \langle \mathbf{d}, \mathbf{v}_{i,j} \rangle, \quad i \in K_j \tag{7.24}$$

to the global fusion center. Again, we do not use sensor measurements from sensors outside the region B_j.

The number of long distance transmissions is

$$r := \sum_{j=1}^{L} \text{card}(K_j).$$

Since $\mathcal{R}(A) \subset \text{span}(\{P_{C_j} \mathbf{b}_k\}_{j=1,\dots,L,\ k \in I_j})$, $\{\mathbf{v}_{i,j}\}_{j=1,\dots,L,\ i \in K_j}$ is an orthonormal system with

$$\mathcal{R}(A) \subset \operatorname{span}(\{\mathbf{v}_{i,j}\}_{j=1,\ldots,L,\ i\in K_j}). \tag{7.25}$$

Since the columns of A are assumed to be linearly independent, this implies that

$$r \geqslant m. \tag{7.26}$$

This procedure formulated in terms of the matrix A is as follows. We choose a partition $\{C_j\}_{j=1}^{L}$ of the indices of the rows of A. For each index set C_j, all rows of A whose index is not contained in C_j are set to zero. We denote the resulting matrices by A_j

$$A_j(i,k) = \begin{cases} A(i,k) & \text{for } i \in C_j, \\ 0 & \text{otherwise.} \end{cases} \tag{7.27}$$

The kth column of A_j equals $P_{C_j}\mathbf{b}_k$. Since the generating functions g_k, $k = 1,\ldots,m$, have a compact support, some of the columns of A_j are zero. The nonzero columns are exactly those whose index is contained in the set I_j. For each $j = 1,\ldots,L$ we calculate an orthonormal basis for $\mathcal{R}(A_j)$, for example by the singular value decomposition, and denote it by $\{\mathbf{v}_{i,j}\}$, $i \in K_j$.

7.5.2.1 Number of required long distance transmissions

We denote by $\operatorname{card}(I)$ the cardinality of a set I. Since at most $\operatorname{card}(I_j)$ columns and at most $\operatorname{card}(C_j)$ rows of A_j are nonzero,

$$\operatorname{card}(K_j) \leqslant \min(\operatorname{card}(I_j), \operatorname{card}(C_j)), \quad j = 1,\ldots,L.$$

We observe that

$$\begin{aligned}\sum_{j=1}^{L} \operatorname{card}(I_j) &= \sum_{j=1}^{L} \operatorname{card}(\{k : \exists\ x_i \in B_j \text{ such that } x_i \in \operatorname{supp}(g_k)\}) \\ &= \sum_{k=1}^{m} \operatorname{card}\{j : \exists x_i \in B_j \text{ such that } x_i \in \operatorname{supp}(g_k)\},\end{aligned}$$

and consequently

$$
\begin{aligned}
r &= \sum_{j=1}^{L} \operatorname{card}(K_j) \leqslant \sum_{j=1}^{L} \operatorname{card}(I_j) \\
&= \sum_{k=1}^{m} \operatorname{card}\{j : \exists x_i \in B_j \text{ such that } x_i \in \operatorname{supp}(g_k)\}.
\end{aligned}
$$

If the support of the generating function g_k has a non empty intersection with l sets B_j, then $\operatorname{card}(\{j : \exists x_i \in B_j \text{ such that } x_i \in \operatorname{supp}(g_k)\}) = l$. Consequently, if the support of the generating function g_k is contained completely in a set B_j, then $\operatorname{card}(\{j : \exists x_i \in B_j \text{ such that } x_i \in \operatorname{supp}(g_k)\}) = 1$. If the sets B_j are large enough in comparison to the support of the generating functions, most of the functions g_k are supported in one set B_j. In this case r, the number of required long distance transmissions, is roughly m, the number of generating functions. A concrete example for the count of the long distance transmissions required is given in Section 7.7.

7.6 Stability of the algorithms

Next we compare the different reconstruction strategies in the presence of transmission noise. We use for all three setups the same upper bound on the signal to noise ratio.

7.6.1 Solving the global least squares problem

In the presence of transmission noise, Algorithm 7.2.1 calculates the coefficients for the reconstruction by solving the least squares problem

$$\hat{\mathbf{c}} = \arg\min_{\mathbf{c}} \|A\mathbf{c} - (\mathbf{h} + \boldsymbol{\delta} + \boldsymbol{\varepsilon})\|,$$

where $\mathbf{h} = [f(x_1), \ldots, f(x_n)]^T$, $\boldsymbol{\delta} = [\delta_1, \ldots, \delta_n]^T$ is the measurement noise, and $\boldsymbol{\varepsilon} = [\varepsilon_1, \ldots, \varepsilon_n]^T$ is the transmission noise.

Lemma 7.6.1. *If the relative measurement error is bounded by*

$$\frac{\|\boldsymbol{\delta}\|}{\|\boldsymbol{h}\|} \leqslant C_1, \tag{7.28}$$

and the relative transmission error is bounded by

$$\frac{\|\boldsymbol{\varepsilon}\|}{\|\boldsymbol{h}+\boldsymbol{\delta}\|} \leqslant C_2, \tag{7.29}$$

then the noise vector $\boldsymbol{a} = \boldsymbol{\delta} + \boldsymbol{\varepsilon}$ *is bounded by*

$$\|\boldsymbol{a}\| \leqslant C_1\|\boldsymbol{h}\| + C_2(1 + C_1)\|\boldsymbol{h}\|. \tag{7.30}$$

Proof. The proof is straightforward and hence omitted. □

7.6.2 Solving the normal equations

Next we bound the noise vector of Algorithm 7.5.1 under the same upper bounds on the signal to noise ratio. Algorithm 7.5.1 calculates the coefficients for the reconstruction by solving the normal equations

$$A^*A\hat{\mathbf{c}} = A^*(\mathbf{h} + \boldsymbol{\delta}) + \boldsymbol{\varepsilon},$$

where again $\boldsymbol{\varepsilon}$ is the transmission noise.

We observe that the same solution is obtained when solving the least squares problem

$$\hat{\mathbf{c}} = \arg\min_{\mathbf{c}} \|A\mathbf{c} - (\mathbf{h} + \boldsymbol{\delta} + A(A^*A)^{-1}\boldsymbol{\varepsilon})\|.$$

Consequently, from the analytical point of view we calculate the solution of the global least squares

$$\hat{\mathbf{c}} = \arg\min_{\mathbf{c}} \|A\mathbf{c} - (\mathbf{h} + \mathbf{a})\|,$$

with the noise vector $\mathbf{a}$ defined by

$$\mathbf{a} := \boldsymbol{\delta} + A(A^*A)^{-1}\boldsymbol{\varepsilon}. \tag{7.31}$$

Lemma 7.6.2. *If the relative measurement error is bounded by*

$$\frac{\|\boldsymbol{\delta}\|}{\|\boldsymbol{h}\|} \leqslant C_1,$$

and the relative transmission error is bounded by

$$\frac{\|\boldsymbol{\varepsilon}\|}{\|A^*(\boldsymbol{h}+\boldsymbol{\delta})\|} \leqslant C_2,$$

then the noise vector $\boldsymbol{a}$ *defined by* (7.31) *is bounded by*

$$\|\boldsymbol{a}\| \leqslant C_1\|\boldsymbol{h}\| + \kappa(A)C_2(1+C_1)\|\boldsymbol{h}\|. \tag{7.32}$$

Proof. The proof is straightforward and hence omitted. □

We observe that we make the same assumption on the signal to noise ratio as imposed by (7.28) and (7.29). In contrast to (7.30) in (7.32), the term $C_2\|\mathbf{h}\|(1+C_1)$ is multiplied by the factor $\kappa(A)$. Therefore we expect a worse reconstruction by this strategy (in the case of transmission noise), than by sending all noisy point evaluations and solving the global least squares problem (7.3), whenever the condition number $\kappa(A)$ is large.

7.6.3 Orthonormal systems locally

Next we discuss the stability of Algorithm 7.5.2. We recall the notation $R = \mathcal{R}(A)$ and $W = \mathcal{R}(T)$. Since the columns of T form an orthonormal system, they also form a tight frame sequence, and consequently, using Theorem 5.6.5, equation (5.45) and the fact that $\mathcal{R}(A) \subset \mathcal{R}(T)$, we obtain

$$A(T^*A)^{\dagger}T^* = P_{R,P_W(R)^{\perp}} = P_R. \tag{7.33}$$

We observe that the least squares problem

$$\hat{\mathbf{c}} = \arg\min_{\mathbf{c}} \|T^*A\mathbf{c} - \big(T^*(\mathbf{h}+\boldsymbol{\delta}) + \boldsymbol{\varepsilon}\big)\|.$$

has the same solution $\hat{\mathbf{c}}$ as

$$\hat{\mathbf{c}} = \arg\min_{\mathbf{c}} \|A\mathbf{c} - (\mathbf{h}+\mathbf{a})\|, \tag{7.34}$$

with $\mathbf{a}$ defined by

$$\mathbf{a} := \boldsymbol{\delta} + A(T^*A)^\dagger \boldsymbol{\varepsilon}. \tag{7.35}$$

This can be seen from

$$\begin{aligned} A^\dagger(\mathbf{h}+\mathbf{a}) &= A^\dagger(\mathbf{h}+\boldsymbol{\delta}) + A^\dagger A(T^*A)^\dagger\boldsymbol{\varepsilon} \\ &= A^\dagger P_R(\mathbf{h}+\boldsymbol{\delta}) + (T^*A)^\dagger\boldsymbol{\varepsilon} \\ &= A^\dagger A(T^*A)^\dagger T^*(\mathbf{h}+\boldsymbol{\delta}) + (T^*A)^\dagger\boldsymbol{\varepsilon} \\ &= (T^*A)^\dagger\big(T^*(\mathbf{h}+\boldsymbol{\delta}) + \boldsymbol{\varepsilon}\big), \end{aligned}$$

using the fact that $\mathcal{R}((T^*A)^\dagger) = \mathcal{R}(A^*T) \subset \mathcal{R}(A^*)$, $A^\dagger A = P_{\mathcal{R}(A^*)}$ and (7.33).

Lemma 7.6.3. *If the relative measurement error is bounded by*

$$\frac{\|\boldsymbol{\delta}\|}{\|\boldsymbol{h}\|} \leqslant C_1,$$

and the relative transmission error is bounded by

$$\frac{\|\boldsymbol{\varepsilon}\|}{\|T^*(\boldsymbol{h}+\boldsymbol{\delta})\|} \leqslant C_2,$$

then the noise vector $\boldsymbol{a}$, defined by (7.35), *is bounded by*

$$\|\boldsymbol{a}\| \leqslant C_1\|\boldsymbol{h}\| + C_2(1+C_1)\|\boldsymbol{h}\|. \tag{7.36}$$

Proof. We observe that

$$\mathcal{R}(T^*)^\perp = \mathcal{N}(T) \subset \mathcal{N}(A^*T) = \mathcal{N}(T^*A)^\dagger.$$

Consequently

$$\|A(T^*A)^\dagger\| = \|A(T^*A)^\dagger T^*\| = \|P_{\mathcal{R}(A)}\| = 1, \tag{7.37}$$

where we used (7.33) for the second equality. Using (7.37), we infer that

$$\begin{aligned}
\|\mathbf{a}\| &= \|\boldsymbol{\delta} + A(T^*A)^\dagger \boldsymbol{\varepsilon}\| \\
&\leqslant C_1\|\mathbf{h}\| + \|A(T^*A)^\dagger\| \; C_2 \; \|T^*(\mathbf{h} + \boldsymbol{\delta})\| \\
&= C_1\|\mathbf{h}\| + C_2\|P_{\mathcal{R}(T)}(\mathbf{h} + \boldsymbol{\delta})\| \\
&\leqslant C_1\|\mathbf{h}\| + C_2(\|P_{\mathcal{R}(T)}\mathbf{h}\| + \|P_{\mathcal{R}(T)}\boldsymbol{\delta}\|) \\
&\leqslant C_1\|\mathbf{h}\| + C_2(1 + C_1)\|\mathbf{h}\|.
\end{aligned}$$

This finishes the proof. □

We observe that in Lemma 7.6.3 we make the same assumption on the signal to noise ratio as imposed by (7.28) and (7.29). We observe that (7.36) is the same bound as obtained in (7.30). Therefore we expect the accuracy of the approximation $\tilde{f}$ to f obtained by this hierarchical reconstruction algorithm to be similar to the accuracy of the approximation $\tilde{f}$ to f obtained by sending all noisy point evaluations and solving the global least squares problem (7.3).

It should also be mentioned that the matrix T^*A used for the calculation of the coefficients $\hat{\mathbf{c}}$ has the same condition number as A. This can be seen from the following observation. The columns of T are an orthonormal system, and $\mathcal{R}(A) \subset R(T)$. Therefore T^* is an isometry on $\mathcal{R}(A)$ and consequently $\kappa(T^*A) = \kappa(A)$. Therefore by calculating orthonormal systems locally we avoid squaring the condition number.

7.7 Numerical experiments

For the numerical experiments, we use integer translates of basis-splines of a certain order as generating functions for the reconstruction space $\mathcal{W}_p$. More detailed information on shift-invariant spaces generated by a one dimensional spline can be found in [36].

Definition 7.7.1. *Let the function* Π *be defined for* $x \in \mathbb{R}$ *by*

$$\Pi(x) = \begin{cases} 1, & -\frac{1}{2} \leqslant x < \frac{1}{2}, \\ 0, & \textit{otherwise.} \end{cases}$$

The B-spline of order $p - 1 \in \mathbb{N}$ *is defined by the p-fold convolution*

$$b_p(x) := \underbrace{\Pi(x) * \Pi(x) * \cdots * \Pi(x)}_{p \text{ factors}}.$$

As a function model we choose the $m+1$-dimensional space $\mathcal{W}_p \subset L^2(\mathbb{R})$

$$\mathcal{W}_p = \left\{ \sum_{k=0}^{m} c_k b_p(\cdot - k) : c_k \in \mathbb{R} \right\}.$$

For $m = 10$, in Figure 7.1, Figure 7.2 and Figure 7.3, we plot the generating functions of $\mathcal{W}_p$ for $p = 1$, $p = 2$ and $p = 3$ respectively.

For the experiment we generate random functions f_i, $i = 1, \ldots, P$, in $\mathcal{W}_p$

$$f_i = \sum_{k=0}^{m} c_k^i b_p(\cdot - k). \tag{7.38}$$

by choosing i.i.d. normally distributed coefficients c_k.

For a given oversampling factor $s \in \mathbb{N}$, i.e $n = s(m+1)$, the sensors are positioned equispaced in the interval $[0, m]$, which means that for

$$\delta = \frac{m}{n-1},$$

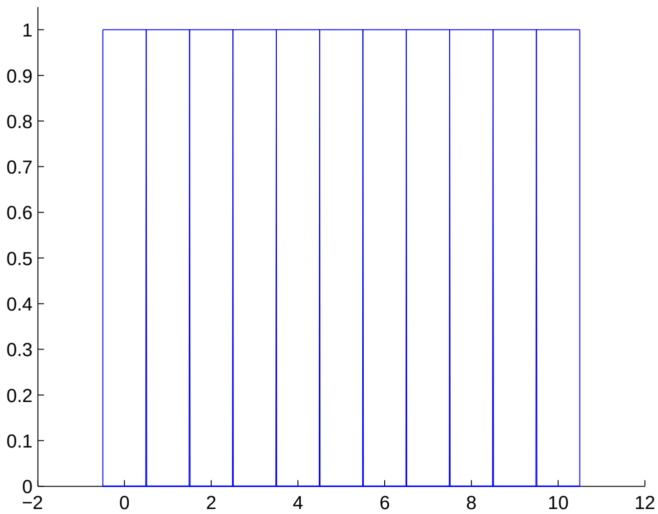

Figure 7.1: The generating function $b_1(\cdot - k)$, $k = 0, \ldots, 10$

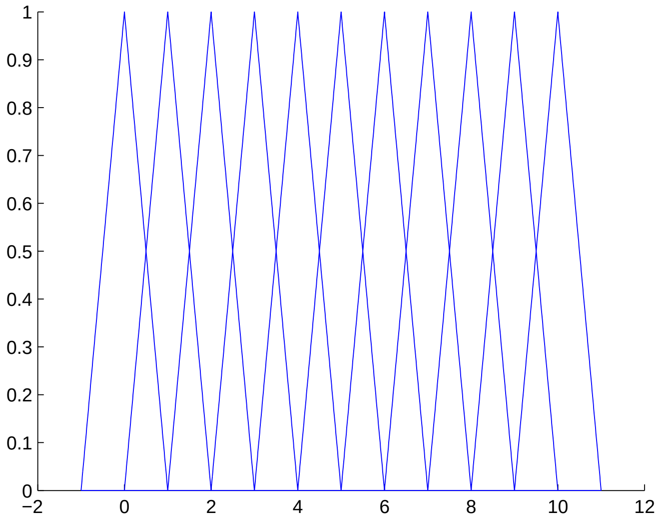

Figure 7.2: The generating function $b_2(\cdot - k)$, $k = 0, \ldots, 10$

the sensor positions are defined by

$$x_i = \delta(i-1) \quad i = 1, \ldots, n.$$

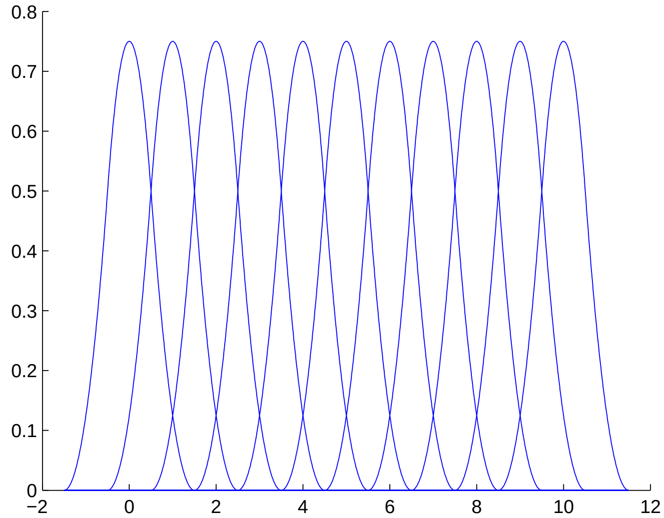

Figure 7.3: The generating function $b_3(\cdot - k)$, $k = 0, \dots, 10$

We compare the accuracy of the reconstruction obtained by sending all noisy point evaluations and solving the global least squares problem (as described in Section 7.2) with the reconstruction method described in Section 7.5.1 and the reconstruction method described in Section 7.5.2.

For all three setups, we choose $m = 80$, which means that the sensors are located in the interval $[0, 80]$.

7.7.0.1 Global least squares problem

The reconstruction method described in Section 7.2 calculates the coefficients for the approximation $\tilde{f}_i$ of f_i from noisy point evaluations

$$\tilde{\mathbf{d}}_i = \mathbf{h}_i + \boldsymbol{\delta}_i + \boldsymbol{\varepsilon}_i,$$

where $\mathbf{h}_i \in \mathbb{R}^n$,

$$\mathbf{h}_i = [f_i(x_1), \dots, f_i(x_n)]^T,$$

$\boldsymbol{\delta}_i \in \mathbb{R}^n$ is the vector containing the measurement noise and $\boldsymbol{\varepsilon}_i \in \mathbb{R}^n$ is the vector containing the transmission noise. The vector entries of $\boldsymbol{\delta}_i$ and $\boldsymbol{\varepsilon}_i$ are initially created with the standard normal distribution, and then the vectors are normalized, so that

$$\frac{\|\boldsymbol{\delta}_i\|}{\|\mathbf{h}_i\|} = C_1,$$

and the relative transmission error fulfills

$$\frac{\|\boldsymbol{\varepsilon}_i\|}{\|\mathbf{h}_i + \boldsymbol{\delta}_i\|} = C_2.$$

Let the matrix $A \in \mathbb{R}^{n\times(m+1)}$ be defined by

$$A(i,k) = b_p(x_i - k), \quad i = 1,\dots,n, \;\; k = 0,\dots,m.$$

For simplicity of notation, we index the first column of A by 0. The coefficients for the approximation

$$\tilde{f}_i = \sum_{k=0}^{m} \hat{c}_k^i b_p(\cdot - k)$$

of f_i are calculated as the least squares solution of

$$\hat{\mathbf{c}}^i := \arg\min_{\mathbf{c}} \|A\mathbf{c} - \tilde{\mathbf{d}}_i\|.$$

We average (arithmetic mean) the relative reconstruction error of this approach over the calculated approximations

$$r_0 := \frac{1}{P} \sum_{i=1,\dots,P} \frac{\|f_i - \tilde{f}_i\|}{\|f_i\|}.$$

Here the norm denotes the standard norm on $L^2([0,m])$, i.e.,

$$\|f\| = \left(\int_0^m |f(x)|^2 \mathrm{d}x \right)^{\frac{1}{2}}.$$

We choose the L^2 norm on the interval $[0,m]$ since the sensor positions $\{x_i\}_{i=1}^n$ are located in this region.

7.7.0.2 Normal equations

The reconstruction method described in Section 7.5.1, calculates the coefficients for the approximation $\tilde{f}_i$ of f_i from

$$\mathbf{s}_i = A^*(\mathbf{h}_i + \boldsymbol{\delta}_i) + \boldsymbol{\varepsilon}_i,$$

where we take the same vector $\boldsymbol{\delta}_i$ as in the previous setup (in order to get a fair comparison of the algorithms). We can not take the same vectors $\boldsymbol{\varepsilon}_i$ as before, since now $\boldsymbol{\varepsilon}_i \in \mathbb{R}^m$. The vector entries of $\boldsymbol{\varepsilon}_i$, containing the transmission noise, are initially created with the standard normal distribution, and then the vectors are normalized, so that

$$\frac{\|\boldsymbol{\varepsilon}_i\|}{\|A^*(\mathbf{h}_i + \boldsymbol{\delta}_i)\|} = C_2.$$

The coefficients for the approximation

$$\tilde{f}_i = \sum_{k=0}^{m} \hat{c}_k^i b_p(\cdot - k)$$

to f_i are calculated as the least squares solution of

$$\hat{\mathbf{c}}^i := \arg\min_{\mathbf{c}} \|A^*A\mathbf{c} - \mathbf{s}_i\|,$$

i.e. the solution of the normal equations. Again we average (arithmetic mean) the relative reconstruction error of this approach over the calculated approximations

$$r_1 := \frac{1}{P} \sum_{i=1,\ldots,P} \frac{\|f_i - \tilde{f}_i\|}{\|f_i\|}.$$

Again the norm denotes the standard norm on $L^2([0, m])$.

7.7.0.3 Orthonormal systems locally

For reconstruction strategy described in Section 7.5.2 we partition the interval $[0, m]$ containing the sensor positions $\{x_i\}_{i=1}^n$ into k Intervals B_j,

$j = 1, \dots, k$, of the same size. Except for the last interval, which is closed, all intervals are left-closed and right-open. We choose $k = 4$, and consequently

$$B_1 = [0, 20),\ B_2 = [20, 40),\ I_3 = [40, 60] \text{ and } B_4 = [60, 80]. \tag{7.39}$$

Let A_j be defined by (7.27), i.e., the matrix being formed by zeroing all rows of A with index not contained B_j. We denote the left endpoint of B_j by a_j and the right endpoint by e_j. Since the support of b_p is $[-p/2, p/2]$, the kth column of A_j is zero if $k \notin]a_j - p/2, e_j + p/2[$. For example for $p = 3$ the matrix A_1 and A_4 have at most 21 nonzero columns, A_2 and A_3 have at most 22 nonzero columns. For each matrix A_j we calculate an orthonormal basis $\{\mathbf{w}_i\}$, $i \in K_j$, for $\mathcal{R}(A_j)$, using the singular value decomposition. The amount of nonzero columns bounds the cardinality $\operatorname{card}(K_j)$. Consequently for $p = 3$, $\operatorname{card}(K_2), \operatorname{card}(K_3) \leqslant 23$ and $\operatorname{card}(K_1), \operatorname{card}(K_4) \leqslant 22$ and

$$\sum_{j=1}^{4} \operatorname{card}(K_j) \leqslant 90,$$

which is the number of generating functions plus 9, i.e., roughly an oversampling rate of 10 percent.

Let us count the number of long distance transmission for intervals B_j, $j = 1, \dots, k$, of length $R = e_j - a_j$ and generating functions $b_p(\cdot - l)$, $l = 0, \dots, kR$. For $k \geqslant 2$ we obtain the bound

$$\begin{aligned}\sum_{j=1}^{k} \operatorname{card}(K_j) &\leqslant 2\left(R + 1 + \left\lfloor \frac{p-1}{2} \right\rfloor\right) + (k-2)\left(R + 1 + 2\left\lfloor \frac{p-1}{2} \right\rfloor\right) \\ &= k(R+1) + 2(k-1)\left\lfloor \frac{p-1}{2} \right\rfloor\end{aligned}$$

Therefore if p is small in comparison to R, the first term $k(R+1)$ is dominant in relation to the term $2(k-1)\left\lfloor \frac{p-1}{2} \right\rfloor$, and the number of long distance transmissions required by this approach is roughly the number of generating functions, which is $kR + 1$ (whenever R is significantly larger than 1).

Let us return to the example of four intervals B_j, $j = 1, \dots, 4$, defined by (7.39). Let $T \in \mathbb{R}^{m \times r}$ denote the matrix whose columns consist of the local orthonormal systems, where r is $r = \sum_{j=1}^{4} \operatorname{card}(K_j)$. The reconstruction method described in Section 7.5.2, calculates the coefficients for the approximation $\tilde{f}_i$ of f_i from

$$\mathbf{s}_i = T^*(\mathbf{h}_i + \boldsymbol{\delta}_i) + \boldsymbol{\varepsilon}_i,$$

where we take the same vector $\boldsymbol{\delta}_i \in \mathbb{R}^n$ as in the previous two setups. The vector entries of $\boldsymbol{\varepsilon}_i \in \mathbb{R}^r$, containing the transmission noise, are initially created with the standard normal distribution, and then normalized, so that

$$\frac{\|\boldsymbol{\varepsilon}_i\|}{\|T^*(\mathbf{h}_i + \boldsymbol{\delta}_i)\|} = C_2.$$

The coefficients for the approximation

$$\tilde{f}_i = \sum_{k=0}^{m} \hat{c}_k^i b_p(\cdot - k)$$

of f_i are calculated as the least squares solution of

$$\hat{\mathbf{c}}^i := \arg\min_{\mathbf{c}} \|T^* A\mathbf{c} - \mathbf{s}_i\|.$$

Again we average (arithmetic mean) the relative reconstruction error of this approach over the calculated approximations

$$r_3 := \frac{1}{P} \sum_{i=1,\dots,P} \frac{\|f_i - \tilde{f}_i\|}{\|f_i\|}.$$

Again the norm denotes the standard norm on $L^2([0, m])$.

7.7.0.4 Tables

In Table 7.1 we list the relative reconstruction error, averaged over $P = 100$ reconstructions, of the three reconstruction methods, the number of required long distance transmissions, the condition number of the matrices used in the corresponding least squares problems for $C_1 = C_2 = 0$, $p = 3$ (spline order 2) and oversampling factor $s = 2$, and in Table 7.2 we list the same for $p = 3$ (spline order 2) and oversampling factor $s = 3$.

We label the reconstruction strategy of Section 7.7.0.1 by Global, the reconstruction strategy described in Section 7.7.0.2 by Normal equations, and the reconstruction strategy described in Section 7.7.0.3 by ON-locally.

	error	transmissions	κ
Global	3.9921e-16	162	$\kappa(A) = 2.7645$
ON-locally	6.7481e-16	90	$\kappa(T^*A) = 2.7645$
Normal equations	2.7133e-16	81	$\kappa(A^*A) = 7.6427$

Table 7.1: Comparison of the three reconstruction methods without noise, for $p = 3$ (spline order 2) and oversampling factor $s = 2$.

	error	transmissions	κ
Global	4.2020e-16	243	$\kappa(A) = 2.7385$
ON-locally	7.0249e-16	90	$\kappa(T^*A) = 2.7385$
Normal equations	2.8216e-16	81	$\kappa(A^*A) = 7.4995$

Table 7.2: Comparison of the three reconstruction methods without noise, for $p = 3$ (spline order 2) and oversampling factor $s = 3$.

In Table 7.3 we list the relative reconstruction error of the three reconstruction methods, averaged over $P = 100$ reconstructions, the number of required long distance transmissions, the condition number of the matrices used in the corresponding least squares problems for $C_1 = C_2 = 0.01$, $p = 3$ (spline order 2) and oversampling factor $s = 2$, and in Table 7.4 we list the same for $p = 3$ (spline order 2) and oversampling factor $s = 3$.

We observe that for the algorithm described in Section 7.7.0.2 (denoted by Normal equations) the amount of required long distance transmission is equal to the number of generating functions, which is a bit smaller than the amount of long distance transmission required by the algorithm described in Section 7.7.0.3 (denoted by ON-locally). We further observe that by all three reconstruction strategies we obtain accurate approximations. Therefore there

	error	transmissions	κ
Global	9.8537e-03	162	$\kappa(A) = 2.7645$
ON-locally	1.1834e-02	90	$\kappa(T^*A) = 2.7645$
Normal equations	1.5869e-02	81	$\kappa(A^*A) = 7.6427$

Table 7.3: Comparison of the three reconstruction methods with noise levels $C_1 = C_2 = 0.01$, for $p = 3$ (spline order 2) and oversampling factor $s = 2$.

	error	transmissions	κ
Global	8.0576e-03	243	$\kappa(A) = 2.7385$
ON-locally	1.1040e-02	90	$\kappa(T^*A) = 2.7385$
Normal equations	1.5292e-02	81	$\kappa(A^*A) = 7.4995$

Table 7.4: Comparison of the three reconstruction methods with noise levels $C_1 = C_2 = 0.01$, for $p = 3$ (spline order 2) and oversampling factor $s = 3$.

arises the question why should we use the more complicated algorithm described in Section 7.7.0.3 (denoted by ON-locally) if the algorithm described in Section 7.7.0.2 (denoted by Normal equations) requires less long distance transmissions, and we still obtain accurate approximations. The reason is that the algorithm described in Section 7.7.0.2 is very sensitive to transmission noise, whenever the condition number $\kappa(A)$ is large. This is shown in the next setup.

Next we compare the three reconstruction methods for a setup where the condition number of A is large. This is obtained by the following change of the setup. We leave everything unchanged, with the only difference that as function model we choose the $m+5$-dimensional space $\mathcal{W}_p \subset L^2(\mathbb{R}, \mathbb{R})$

$$\mathcal{W}_5 = \left\{ \sum_{k=-2}^{m+2} c_k b_5(\cdot - k) : c_k \in \mathbb{R} \right\}. \tag{7.40}$$

The support of the four splines $b_5(\cdot + 2)$, $b_5(\cdot + 1)$, $b_5(\cdot - (m+1))$ and $b_5(\cdot - (m+2))$ have only a small overlap with the interval $[0, m]$ containing the sensor positions. This increases the condition number of A. We reconstruct functions in $\mathcal{W}$ defined by (7.40).

Using as function model (7.40), in Table 7.5 we list the relative reconstruction error

$$\frac{1}{P} \sum_{i=1,\dots,P} \frac{\|f_i - \tilde{f}_i\|}{\|f_i\|}, \tag{7.41}$$

averaged over $P = 100$ reconstructions, of the three reconstruction methods, the number of required long distance transmissions, the condition number of the matrices used in the corresponding least squares problems for $C_1 = C_2 = 0$, $p = 5$ (spline order 4), and oversampling factor $s = 2$. In Table 7.6 we list the same for $p = 5$ (spline order 4) and oversampling factor $s = 3$. Again the norm denotes the standard norm on $L^2([0, m])$. This means the reconstruction error (7.41) is only measured in the region where the samples are located.

	error	transmissions	κ
Global	4.9379e-16	170	$\kappa(A)$ = 9.0039e+03
ON-locally	7.2133e-16	100	$\kappa(T^*A)$ = 9.0039e+03
Normal equations	1.6577e-15	85	$\kappa(A^*A)$ = 8.1071e+07

Table 7.5: Comparison of the three reconstruction methods without noise, for $p = 5$ (spline order 4), oversampling factor $s = 2$ and $\mathcal{W}$ defined by (7.40).

	error	transmissions	κ
Global	4.7793e-16	255	$\kappa(A)$ = 3.7388e+03
ON-locally	7.0442e-16	100	$\kappa(T^*A)$ = 3.7388e+03
Normal equations	5.8080e-16	85	$\kappa(A^*A)$ = 1.3979e+07

Table 7.6: Comparison of the three reconstruction methods without noise, for $p = 5$ (spline order 4), oversampling factor $s = 3$ and $\mathcal{W}$ defined by (7.40).

In Table 7.7 we list the same for $C_1 = C_2 = 0.01$, $p = 3$, and oversampling factor $s = 2$, and in Table 7.8 we list the same for $p = 3$ and oversampling factor $s = 3$.

We observe that now the approximations calculated by the algorithm described in Section 7.7.0.2 (denoted by Normal equations) have a huge approximation error when we use transmission noise. By using slightly more

	error	transmissions	κ
Global	1.0625e-02	170	$\kappa(A) = 9.0039\text{e}+03$
ON-locally	1.2212e-02	100	$\kappa(T^*A) = 9.0039\text{e}+03$
Normal equations	5.2123e+01	85	$\kappa(A^*A) = 8.1071\text{e}+07$

Table 7.7: Comparison of the three reconstruction methods with noise levels $C_1 = C_2 = 0.01$, for $p = 5$ (spline order 4), oversampling factor $s = 2$ and $\mathcal{W}$ defined by (7.40).

	error	transmissions	κ
Global	8.1496e-03	255	$\kappa(A) = 3.7388\text{e}+03$
ON-locally	1.0855e-02	100	$\kappa(T^*A) = 3.7388\text{e}+03$
Normal equations	5.4981e+00	85	$\kappa(A^*A) = 1.3979\text{e}+07$

Table 7.8: Comparison of the three reconstruction methods with noise levels $C_1 = C_2 = 0.01$, for $p = 5$ (spline order 4), oversampling factor $s = 3$ and $\mathcal{W}$ defined by (7.40).

long distance transmissions, by the algorithm described in Section 7.7.0.3 (denoted by ON-locally) the sensitivity to noisy is significantly reduced. We obtain by this algorithm (ON-locally) almost as accurate approximations as by sending all noisy point evaluations and solving the global least squares problem with a significantly (the reduction is dependent on the oversampling factor) reduced amount of long distance transmissions.

7.8 Operation count of the algorithm using orthonormal systems locally

In this section we determine the operation count of the algorithm described in Section 7.5.2. Let the matrix $A \in \mathbb{C}^{n\times m}$ be defined by (7.2). We recall that we assume $n \geqslant m$, i.e. the number of sensors, is larger or equal to the number of generating functions. In addition we assume the matrix A to be injective, i.e., to have full column rank.

We recall that the coefficients for the reconstruction are calculated by solving the least squares problem

$$\hat{\mathbf{c}} = \arg\min_{\mathbf{c}} \|T^* A\mathbf{c} - \mathbf{s}\|$$

with

$$\mathbf{s} = T^*\mathbf{d} + \boldsymbol{\varepsilon}. \tag{7.42}$$

Therefore we have to determine the operation count of the calculation of A, of T, of the multiplication T^*A, of the calculation of $T^*\mathbf{d}$ and of calculating the solution of the least squares problem.

Using De Boor's algorithm, the point evaluation of a B-spline of order $p-1$ has an operation count of $\mathcal{O}((p-1)^2)$. By definition (7.7.1), the support of $b_p(\cdot - k)$ is $[k - \frac{p}{2}, k + \frac{p}{2}]$. Therefore for a point $x \in \mathbb{R}$, only those generating functions $b_p(\cdot - k)$ with $x - \frac{p}{2} < k < x + \frac{p}{2}$ are nonzero. Therefore for each point x_i, $i = 1, \dots, n$ there are at most p spline evaluations necessary, which results in a total number of pn spline evaluations in order to create the matrix A. Therefore creating the matrix A has an operation count of $\mathcal{O}(p(p-1)^2 n)$.

Next we determine the operation count for the creation of the matrix T. At first, we choose a partition $\{C_j\}_{j=1}^L$ of the indices $\{1, \dots, n\}$ of the rows of A. For each index set C_j, all rows of A whose index is not contained in C_j are set to zero. We denote the resulting matrices by A_j, see (7.27). For each matrix A_j, we calculate an orthonormal system for the range $\mathcal{R}(A_j)$. We define

$$n_j := \mathrm{card}(C_j).$$

As in (7.23) we define

$$I_j := \{k : \mathrm{supp}(g_k) \cap \{x_i : x_i \in B_j\} \neq \emptyset\},$$

and we define m_j by

$$m_j = \mathrm{card}(I_j).$$

We denote by $\tilde{A}_j \in \mathbb{C}^{n_j \times m_j}$ the submatrix of A_j, where we delete the rows with index $\{1, \dots, n\} \backslash C_j$ and columns with index $\{1, \dots, m\} \backslash I_j$, which are zero by construction. We calculate an orthonormal system for the column range $\mathcal{R}(A_j)$ by calculating an orthonormal system for the column range of $\tilde{A}_j$, using for example the singular value decomposition. Calculating an orthonormal system for the column range $\mathcal{R}(\tilde{A}_j)$ has an operation count of $\mathcal{O}\big(n_j m_j \min(n_j, m_j)\big)$, see [65, Lecture 31]. Let $\{\mathbf{w}_{i,j}\}$, $i \in K_j$ denote the orthonormal system for $\mathcal{R}(A_j)$, with

$$\text{card}(K_j) \leqslant \min(n_j, m_j),$$

and

$$r := \sum_{j=1}^{L} \text{card}(K_j). \tag{7.43}$$

The calculation of the matrix $T \in \mathbb{C}^{n\times r}$ with columns consisting of the orthonormal systems $\{\mathbf{w}_{i,j}\}_{j=1,\dots,L,\ i\in K_j}$ stacked together, has an operation count of $\mathcal{O}\big(\sum_{j=1}^{L} n_j m_j \min(n_j, m_j)\big)$.

Next we determine the operation count of the calculation of T^*A and $T^*\mathbf{d}$. For $\mathbf{x} \in \mathbb{C}^{n\times 1}$, each entry of $T^*\mathbf{x} \in \mathbb{C}^{r\times 1}$ is the inner product of $\mathbf{x}$ with the vector $\mathbf{w}_{i,j} \in \mathbb{C}^{n\times 1}$ for some $j \in \{1, \dots, L\}$, $i \in K_j$. Since

$$\mathbf{w}_{i,j}(l) = 0 \quad \text{for } l \notin C_j,$$

the calculation of T^*A has an operation count of $\mathcal{O}\big(m \sum_{j=1}^{L} n_j \text{card}(K_j)\big)$, and the operation count of the calculation of $T^*\mathbf{d}$ is dominated by the operation count of the calculation of T^*A.

Finally we determine the operation count of solving the least squares problem (7.42). We recall that $T^*A \in \mathbb{C}^{r\times m}$ and that $r \geqslant m$, see (7.26). Therefore solving the overdetermined least squares problem (7.42) by a direct method as the QR decomposition has an operation count of $\mathcal{O}(rm^2)$.

Since $\{C_j\}_{j=1}^{L}$ is a partition of $\{1, \dots, n\}$ and $n_j = \text{card}(C_j)$, $\sum_{j=1}^{L} n_j = n$. Using this and $n_j \leqslant n$, $m_j \leqslant m$, $m \leqslant n$, $m \leqslant r$ and (7.43), we see that the operation count of the algorithm is

$$\mathcal{O}\Big(p(p-1)^2 n + rm^2 + \sum_{j=1}^{L} \big(n_j m_j \min(n_j, m_j) + mn_j \mathrm{card}(K_j)\big)\Big)$$
$$= \mathcal{O}\big(p^3 n + rm^2 + m^2 n + mnr\Big)$$
$$= \mathcal{O}\big(p^3 n + mnr).$$

We observe that solving the least squares problem (7.3) of the global setup, described in Section 7.2, by the QR decomposition has an operation count of $\mathcal{O}(nm^2)$. Consequently, if r (the number of long distance transmissions) is roughly m (the number of generating functions), then the global system has the same order of operation count as our algorithm.

Chapter 8

Conclusions and future research

We have treated the problem of sampling and reconstruction in distinct subspaces, a problem that arises when approximating a compactly supported function from point samples of the Fourier transform. The approximation calculated from the standard least squares problem is most stable with respect to errors of the measurements.

We have introduced a novel reconstruction operator which projects as orthogonally as possible onto the reconstruction space. Furthermore, we have shown that this reconstruction method is the most stable possible with respect to systematic error appearing before the sampling process.

This reconstruction operator is more sensitive to errors of the measurements. We have shown how to range continuously between the two extreme reconstruction methods by a regularization parameter, and therefore we have a tool to choose reconstruction operators corresponding to oblique projections closer to the orthogonal projection at the cost of a larger operator norm and vice versa. An open question is how to determine a good choice for the regularization parameter.

In Chapter 6 we applied the different reconstruction operators to the problem of reconstructing a compactly supported function from nonuniform samples of the Fourier transform. We have used trigonometric polynomials for reconstruction. It is important to realize that the developed theory is independent of the chosen reconstruction space. Besides trigonometric polynomials one can for example use wavelets, algebraic polynomials as well as important generalizations of wavelets, such as curvelets and shearlets as reconstruction vectors.

In Chapter 4, we have pointed out the connections of quasi-optimal and bounded reconstruction operators to dual frames. This opens the possibility of a new classification of dual frames.

In the last chapter, we have considered the reconstruction of a non-bandlimited function represented by a finite number of compactly supported generating functions in wireless sensor networks. Using the theory presented in the first part, we have developed a novel hierarchical reconstruction. We have shown that if the local regions are large enough in comparison to the support of the generating functions, then the number of required long distance transmissions required by our approach is roughly the number of generating functions. In this case, if the function is oversampled by a factor of s, we reduce the communication workload roughly by a factor of s and obtain reconstructions of similar accuracy for operation count of the same order.

Bibliography

[1] B. Adcock, A. C. Hansen, and C. Poon. On optimal wavelet reconstructions from Fourier samples: linearity and universality of the stable sampling rate. *Appl. Comput. Harmon. Anal.*, 36(3):387–415, 2014.

[2] Ben Adcock, Milana Gataric, and Anders Hansen. Density theorems for nonuniform sampling of bandlimited functions using derivatives or bunched measurements. *ArXiv e-prints*, nov 2014.

[3] Ben Adcock, Milana Gataric, and Anders Hansen. On stable reconstructions from nonuniform Fourier measurements. *SIAM J. Imaging Sci.*, 7(3):1690–1723, 2014.

[4] Ben Adcock, Milana Gataric, and Anders Hansen. Recovering piecewise smooth functions from nonuniform Fourier measurements. *ArXiv e-prints*, sep 2014.

[5] Ben Adcock, Milana Gataric, and Anders C. Hansen. Weighted frames of exponentials and stable recovery of multidimensional functions from nonuniform fourier samples. *ArXiv e-prints*, may 2014.

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

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

[8] Ben Adcock and Anders C. Hansen. Generalized sampling and the stable and accurate reconstruction of piecewise analytic functions from their Fourier coefficients. *Math. Comp.*, 84(291):237–270, 2015.

[9] Ben Adcock, Anders C. Hansen, Evelyn Herrholz, and Gerd Teschke. Generalized sampling: extension to frames and inverse and ill-posed problems. *Inverse Problems*, 29(1):015008, 27, 2013.

[10] Ben Adcock, Anders C. Hansen, and Clarice Poon. Beyond consistent reconstructions: optimality and sharp bounds for generalized sampling, and application to the uniform resampling problem. *SIAM J. Math. Anal.*, 45(5):3132–3167, 2013.

[11] Ben Adcock, Anders C. Hansen, and Alexei Shadrin. A stability barrier for reconstructions from Fourier samples. *SIAM J. Numer. Anal.*, 52(1):125–139, 2014.

[12] Akram Aldroubi and Karlheinz Gröchenig. Beurling-Landau-type theorems for non-uniform sampling in shift invariant spline spaces. *J. Fourier Anal. Appl.*, 6(1):93–103, 2000.

[13] Peter Balazs. Basic definition and properties of Bessel multipliers. *J. Math. Anal. Appl.*, 325(1):571–585, 2007.

[14] Peter Balazs, Jean-Pierre Antoine, and Anna Gryboś. Weighted and controlled frames: mutual relationship and first numerical properties. *Int. J. Wavelets Multiresolut. Inf. Process.*, 8(1):109–132, 2010.

[15] Adi Ben-Israel and Thomas N. E. Greville. *Generalized inverses*. CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, 15. Springer-Verlag, New York, second edition, 2003. Theory and applications.

[16] Arne Beurling. *The collected works of Arne Beurling. Vol. 1*. Contemporary Mathematicians. Birkhäuser Boston, Inc., Boston, MA, 1989. Complex analysis, Edited by L. Carleson, P. Malliavin, J. Neuberger and J. Wermer.

[17] Arne Beurling. *The collected works of Arne Beurling. Vol. 2.* Contemporary Mathematicians. Birkhäuser Boston, Inc., Boston, MA, 1989. Harmonic analysis, Edited by L. Carleson, P. Malliavin, J. Neuberger and J. Wermer.

[18] I. Bogdanova, P. Vandergheynst, J-P. Antoine, L. Jacques, and M. Morvidone. Stereographic wavelet frames on the sphere. *Appl. Comput. Harmon. Anal.*, 19(2):223–252, 2005.

[19] Don Buckholtz. Hilbert space idempotents and involutions. *Proc. Amer. Math. Soc.*, 128(5):1415–1418, 2000.

[20] O. Christensen and Y. C. Eldar. Oblique dual frames and shift-invariant spaces. *Appl. Comput. Harmon. Anal.*, 17(1):48–68, 2004.

[21] Ole Christensen. *Frames and bases.* Applied and Numerical Harmonic Analysis. Birkhäuser Boston, Inc., Boston, MA, 2008. An introductory course.

[22] Ole Christensen and Yonina C. Eldar. Generalized shift-invariant systems and frames for subspaces. *J. Fourier Anal. Appl.*, 11(3):299–313, 2005.

[23] R. J. Duffin and A. C. Schaeffer. A class of nonharmonic Fourier series. *Trans. Amer. Math. Soc.*, 72:341–366, 1952.

[24] Morris L. Eaton. Multivariate statistics: A vector space approach. *Wiley, New York*, 1983.

[25] Yonina C. Eldar. Sampling with arbitrary sampling and reconstruction spaces and oblique dual frame vectors. *J. Fourier Anal. Appl.*, 9(1):77–96, 2003.

[26] Yonina C. Eldar. Sampling without input constraints: consistent reconstruction in arbitrary spaces. In *Sampling, wavelets, and tomography*, Appl. Numer. Harmon. Anal., pages 33–60. Birkhäuser Boston, Boston, MA, 2004.

[27] Yonina C. Eldar and Ole Christensen. Characterization of oblique dual frame pairs. *EURASIP J. Appl. Signal Process.*, (Frames and overcomplete representations in signal processing, communications, and information theory):Art. ID 92674, 11, 2006.

[28] Yonina C. Eldar and Tobias Werther. General framework for consistent sampling in Hilbert spaces. *Int. J. Wavelets Multiresolut. Inf. Process.*, 3(3):347–359, 2005.

[29] Hans G. Feichtinger, Karlheinz Gröchenig, and Thomas Strohmer. Efficient numerical methods in non-uniform sampling theory. *Numer. Math.*, 69(4):423–440, 1995.

[30] Anne Gelb and Jared Tanner. Robust reprojection methods for the resolution of the Gibbs phenomenon. *Appl. Comput. Harmon. Anal.*, 20(1):3–25, 2006.

[31] Gene H. Golub and Charles F. Van Loan. *Matrix computations*. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, Baltimore, MD, third edition, 1996.

[32] David Gottlieb and Chi-Wang Shu. On the Gibbs phenomenon and its resolution. *SIAM Rev.*, 39(4):644–668, 1997.

[33] K. Gröchenig and H. Razafinjatovo. On Landau's necessary density conditions for sampling and interpolation of band-limited functions. *J. London Math. Soc. (2)*, 54(3):557–565, 1996.

[34] Karlheinz Gröchenig. Irregular sampling, Toeplitz matrices, and the approximation of entire functions of exponential type. *Math. Comp.*, 68(226):749–765, 1999.

[35] Karlheinz Gröchenig. Non-uniform sampling in higher dimensions: from trigonometric polynomials to bandlimited functions. In *Modern sampling theory*, Appl. Numer. Harmon. Anal., pages 155–171. Birkhäuser Boston, Boston, MA, 2001.

[36] Karlheinz Gröchenig and Harald Schwab. Fast local reconstruction methods for nonuniform sampling in shift-invariant spaces. *SIAM J. Matrix Anal. Appl.*, 24(4):899–913 (electronic), 2003.

[37] Christopher Heil, Yoo Young Koo, and Jae Kun Lim. Duals of frame sequences. *Acta Appl. Math.*, 107(1-3):75–90, 2009.

[38] J. R. Higgins. Sampling theory for Paley-Wiener spaces in the Riesz basis setting. *Proc. Roy. Irish Acad. Sect. A*, 94(2):219–236, 1994.

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

[40] Prakash Ishwar, Animesh Kumar, and Kannan Ramchandran. Distributed sampling for dense sensor networks: A "bit-conservation principle". In *Information Processing in Sensor Networks*, volume 2634 of *Lecture Notes in Computer Science*, pages 17–31. Springer Berlin Heidelberg, 2003.

[41] Jae-Hun Jung and Bernie 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.

[42] Jae-Hun Jung and Bernie D. Shizgal. Inverse polynomial reconstruction of two dimensional Fourier images. *J. Sci. Comput.*, 25(3):367–399, 2005.

[43] Jae-Hun Jung and Bernie D. Shizgal. On the numerical convergence with the inverse polynomial reconstruction method for the resolution of the Gibbs phenomenon. *J. Comput. Phys.*, 224(2):477–488, 2007.

[44] M.I. Kadets. The exact value of the Paley-Wiener constant. *Sov. Math., Dokl.*, 5:559–561, 1964.

[45] Animesh Kumar, Prakash Ishwar, and Kannan Ramchandran. High-resolution distributed sampling of bandlimited fields with low-precision sensors. *IEEE Trans. Inform. Theory*, 57(1):476–492, 2011.

[46] H. J. Landau. Necessary density conditions for sampling and interpolation of certain entire functions. *Acta Math.*, 117:37–52, 1967.

[47] R.M. Lewitt. Reconstruction algorithms: Transform methods. *Proceedings of the IEEE*, 71(3):390–408, 1983.

[48] Shidong Li and Hidemitsu Ogawa. Pseudoframes for subspaces with applications. *J. Fourier Anal. Appl.*, 10(4):409–431, 2004.

[49] A. Nordio, C. Chiasserini, and E. Viterbo. Performance of linear field reconstruction techniques with noise and uncertain sensor locations. *Signal Processing, IEEE Transactions on*, 56(8):3535–3547, Aug 2008.

[50] J. Tinsley Oden and Leszek F. Demkowicz. *Applied functional analysis*. CRC Press, Boca Raton, FL, second edition, 2010.

[51] Christopher C. Paige and Michael A. Saunders. LSQR: an algorithm for sparse linear equations and sparse least squares. *ACM Trans. Math. Software*, 8(1):43–71, 1982.

[52] M. Planitz. Inconsistent systems of linear equations. *The Mathematical Gazette*, 63(425):181–185, 1979.

[53] Michael Reed and Barry Simon. *Methods of modern mathematical physics. I*. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York, second edition, 1980. Functional analysis.

[54] Günter Reise, Gerald Matz, and Karlheinz Gröchenig. Distributed field reconstruction in wireless sensor networks based on hybrid shift-invariant spaces. *IEEE Trans. Signal Process.*, 60(10):5426–5439, 2012.

[55] Walter Rudin. *Functional analysis*. International Series in Pure and Applied Mathematics. McGraw-Hill, Inc., New York, second edition, 1991.

[56] Helmut Schwichtenberg. Vorlesungsskript lineare algebra und analytische geometrie. *Mathematisches Institut der Universität München*, 2004.

[57] Kristian Seip. On the connection between exponential bases and certain related sequences in $L^2(-\pi, \pi)$. *J. Funct. Anal.*, 130(1):131–160, 1995.

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

[59] Jacob Steinberg. Oblique projections in Hilbert spaces. *Integral Equations Operator Theory*, 38(1):81–119, 2000.

[60] Thomas Strohmer. Numerical analysis of the non-uniform sampling problem. *J. Comput. Appl. Math.*, 122(1-2):297–316, 2000. Numerical analysis 2000, Vol. II: Interpolation and extrapolation.

[61] Daniel B. Szyld. The many proofs of an identity on the norm of oblique projections. *Numer. Algorithms*, 42(3-4):309–323, 2006.

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

[63] Wai-Shing Tang. Oblique projections, biorthogonal Riesz bases and multiwavelets in Hilbert spaces. *Proc. Amer. Math. Soc.*, 128(2):463–473, 2000.

[64] Jared Tanner. Optimal filter and mollifier for piecewise smooth spectral data. *Math. Comp.*, 75(254):767–790 (electronic), 2006.

[65] Lloyd N. Trefethen and David Bau, III. *Numerical linear algebra*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997.

[66] Adityavikram Viswanathan, Anne Gelb, Douglas Cochran, and Rosemary Renaut. On reconstruction from non-uniform spectral data. *J. Sci. Comput.*, 45(1-3):487–513, 2010.

[67] J. M. Whittaker. *Interpolatory Function Theory*. Cambridge University Press, London, 1935.

[68] Robert M. Young. *An introduction to nonharmonic Fourier series*, volume 93 of *Pure and Applied Mathematics*. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London, 1980.

Curriculum Vitae of Peter Berger

Research interests

- sampling theory, harmonic analysis, inverse problems, numerical analysis.

Education

- Dipl. Ing. in Technical Mathematics, University of Innsbruck, 2011

Employment

- Research assistant, Faculty of Mathematics, University of Vienna, Numerical Harmonic Analysis Group January 2012 - April 2015.

Contact information

University of Vienna
Faculty of Mathematics
Oskar-Morgenstern-Platz 1
A-1090 Vienna, Austria
E-mail: p.berger@univie.ac.at