Nonuniform Generalized Sampling

Milana Gatarić

Robinson College

University of Cambridge

A thesis submitted for the degree of

Doctor of Philosophy

November 2015

Abstract

In this thesis we study a novel approach to stable recovery of unknown compactly supported L^2 functions from finitely many nonuniform samples of their Fourier transform, so-called Nonuniform Generalized Sampling (NUGS). This framework is based on a recently introduced idea of generalized sampling for stable sampling and reconstruction in abstract Hilbert spaces, which allows one to tailor the reconstruction space to suit the function to be approximated and thereby obtain rapidly-convergent approximations. While preserving this important hallmark, NUGS describes sampling by the use of weighted Fourier frames, thus allowing for highly nonuniform sampling schemes with the points taken arbitrarily close. The particular setting of NUGS directly corresponds to various image recovery models ubiquitous in applications such as magnetic resonance imaging, computed tomography and electron microscopy, where Fourier samples are often taken not necessarily on a Cartesian grid, but rather along spiral trajectories or radial lines.

Specifically, NUGS provides stable recovery in a desired reconstruction space subject to sufficient sampling density and sufficient sampling bandwidth, where the latter depends solely on the particular reconstruction space. For univariate compactly supported wavelets, we show that only a linear scaling between the number of wavelets and the sampling bandwidth is both sufficient and necessary for stable recovery. Furthermore, in the wavelet case, we provide an efficient implementation of NUGS for recovery of wavelet coefficients from Fourier data. Additionally, the sufficient relation between the dimension of the reconstruction space and the bandwidth of the nonuniform samples is analysed for the reconstruction spaces of piecewise polynomials or splines with a nonequidistant sequence of knots, and it is shown that this relation is also linear for splines and piecewise polynomials of fixed degree, but quadratic for piecewise polynomials of varying degree.

In order to derive explicit guarantees for stable recovery from nonuniform samples in terms of the sampling density, we also study conditions sufficient to ensure existence of a particular frame. Firstly, we establish the sharp and dimensionless sampling density that is sufficient to guarantee a weighted Fourier frame for the space of multivariate compactly supported L^2 functions. Furthermore, subject to non-sharp densities, we improve existing estimates of the corresponding frame bounds. Secondly, we provide sampling densities sufficient to ensure a frame, as well as, estimates of the corresponding frame bounds, when a multivariate bandlimited function and its derivatives are sampled at nonuniformly spaced points.

I dedicate this thesis to my family. To my parents Stana and Rajko, and my brother Zoran.

Acknowledgements

I would like to start by acknowledging the fantastic support that I was fortunate to have during my doctoral studies at the Cambridge Centre for Analysis (CCA) at the University of Cambridge. There are many people who helped me on my way towards the PhD and who made this part of my life truly amazing.

First and foremost, I would like to thank Ben Adcock for his immense support and guidance throughout my doctoral studies, without whom this research would not be possible. Although mostly at another continent and not officially my supervisor, Ben was nothing less than my true mentor who was constantly involved in my research. He is probably the person from whom I learned the most during my entire mathematical education. For his exceptional mentoring and help, I owe Ben an enormous amount of gratitude.

I also owe a massive amount of gratitude to Anders Hansen, my PhD supervisor at Cambridge. To Anders, I am deeply thankful for introducing me to interesting topics in the theory of information and for including me in the Applied Functional and Harmonic Analysis (AFHA) research group at the Department of Applied Mathematics and Theoretical Physics (DAMTP). He guided me through the first steps of my academic career and provided essential support from start to finish.

Without Ben and Anders, this research would certainly not be possible. It was a real pleasure and privilege to work with them and be their student.

I would like to thank Gil Ramos for his constant support and involvement in my PhD life. Over the past four years, Gil was an ideal officemate and a great friend always willing to offer a carefully thought-out advice. It was my pleasure to share the office with him and have frequent mathematical discussions which greatly benefited my mathematical development and my research.

My research also greatly benefited from collaborations and discussions with Clarice Poon and Jose Luis Romero, as well as other members of the AFHA research group at DAMTP, namely Alex Bastounis, Alex Jones and Bogdan Roman. Also, it was extremely constructive and a great pleasure to discuss my work with Akram Aldroubi and Karlheinz Gröchenig.

I would also like to thank my dear friends at CCA without whom my everyday life in the past four years would be unthinkable. Beside Gil and Clarice, these include: Luca Calatroni, Alberto Coca, Kevin Crooks, Eoin Devane, Vittoria Silvestri and Lukáš Vermach. I would also like to thank Nina Rismal for being a wonderful friend during the past four years.

I am grateful to the CCA directors Arieh Iserles and James Norris for giving me the opportunity to be a CCA student. In addition, I would like to thank Carola Schönlieb for her support especially at the beginning of my doctoral studies at CCA.

I acknowledge the financial support provided by the Engineering and Physical Sciences Research Council, the Cambridge Overseas Trust and the Serbian Fund for Young Talents. I am also thankful to Robinson College for providing excellent facilities during the past four years.

Prior to my doctoral studies at Cambridge, a lot of people at the University of Novi Sad also played a key role towards my mathematical development. I would particularly like to thank to my professors Nataša Krejić and Marko Nedeljkov for their enthusiastic mentoring and unwavering support.

Above all, I owe a great debt of gratitude to my family, my mother Stana, my father Rajko and my brother Zoran. Without them, their love and support all of this would be truly impossible. Volim vas puno.

Milana Gatarić
Cambridge
28 July 2015

Declaration

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University of similar institution.

Contents

List of Notation

$\mathbb{R}$, $\mathbb{R}^d$	The set of real numbers and its d-Cartesian product for $d \geq 1$
$\hat{\mathbb{R}}$	The set of real numbers in the frequency domain
$\mathbb{N}$	The set of positive integers
$\mathbb{N}_0$	The set of non-negative integers
$\mathbb{C}$	The set of complex numbers
$\lfloor x \rfloor$	The largest integer less than or equal to $x \in \mathbb{R}$
i	The imaginary unit
e	Euler's number
$\ln x$	The natural logarithm of $x > 0$, i.e. the logarithm to the base e
$x \cdot y$	The dot product $x_1 y_1 + \cdots + x_d y_d$ of variables $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$ and $y = (y_1, \ldots, y_d) \in \mathbb{R}^d$
$\lvert \cdot \rvert$	The Euclidean norm on $\mathbb{R}^d$
$\mathcal{B}_r$	The Euclidean ball of radius $r > 0$
$\lvert \cdot \rvert_p$	The ℓ^p-norm on $\mathbb{R}^d$
$\lvert \cdot \rvert_*$	An arbitrary norm on $\mathbb{R}^d$
D°	The polar set of compact, convex and symmetric set $D \subseteq \mathbb{R}^d$
$\lvert \cdot \rvert_D$	The norm on $\mathbb{R}^d$ induced by compact, convex and symmetric set $D \subseteq \mathbb{R}^d$
$\overline{D}$	The closure of set $D \subseteq \mathbb{R}^d$
$\mathrm{meas}(D)$	The Lebesgue measure of set $D \subseteq \mathbb{R}^d$
α	The multi-index $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}_0^d$
$\lvert \alpha \rvert_1$	The sum of components $\alpha_1 + \cdots + \alpha_d$ of multi-index α
$\alpha!$	The factorial $\prod_{j=1}^d \alpha_j!$ of multi-index α
x^α	The power $\prod_{j=1}^d x_j^{\alpha_j}$ of $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$ for multi-index α
D^α	The derivative operator $D^\alpha = \frac{\partial^{\lvert\alpha\rvert_1}}{\partial_{x_1}^{\alpha_1} \cdots \partial_{x_d}^{\alpha_d}}$

$\mathrm{L}^2(\mathbb{R}^d)$	Hilbert space of square-integrable functions defined on $\mathbb{R}^d$
$\mathrm{L}^2(D)$, H	The space of functions in $\mathrm{L}^2(\mathbb{R}^d)$ which are supported on compact set $D \subseteq \mathbb{R}^d$
$\mathrm{H}^r(D)$	Sobolev space of functions in $\mathrm{L}^2(D)$ such that the derivatives up to order r are also $\mathrm{L}^2(D)$ functions
$\mathrm{B}(\Omega)$	The space of bandlimited functions in $\mathrm{L}^2(\mathbb{R}^d)$ such that the Fourier transform is in $\mathrm{L}^2(\Omega)$
$\langle \cdot, \cdot \rangle$	The inner product on $\mathrm{L}^2(\mathbb{R}^d)$
$\\| \cdot \\|$	The norm in $\mathrm{L}^2(\mathbb{R}^d)$
χ_D	The indicator function of set $D \subseteq \mathbb{R}^d$
$\\| \cdot \\|_D$	The norm of a function in $\mathrm{L}^2(\mathbb{R}^d)$ multiplied by function χ_D
$\\| \cdot \\|_{\infty,D}$	The uniform norm on set $D \subseteq \mathbb{R}^d$
$\mathcal{I}$	The identity operator
$\mathcal{P}$	The projection operator
$\sigma_{\min}(A)$, $\sigma_{\max}(A)$	The minimal and maximal singular values of matrix $A \in \mathbb{R}^{N\times M}$
$\mathrm{cond}(A)$	The condition number $\frac{\sigma_{\max}(A)}{\sigma_{\min}(A)}$ of matrix $A \in \mathbb{R}^{N\times M}$
$\mathcal{O}$	The big-O notation
$\sim$	The asymptotic expansion symbol
$\psi_{j,k}$	j^{th} dilatation and k^{th} translation of the wavelet function (the mother wavelet) ψ
$\phi_{j,k}$	j^{th} dilatation and k^{th} translation of the scaling function (the father wavelet) ϕ
$\psi_{j,l}^{\mathrm{per}}$, $\phi_{j,l}^{\mathrm{per}}$	The periodic wavelet and scaling functions defined using the periodizing operation $f(\cdot) \mapsto f^{\mathrm{per}}(\cdot) = \sum_{k\in\mathbb{Z}} f(\cdot + k)$
$\psi_{j,l}^{\mathrm{fold}}$, $\phi_{j,l}^{\mathrm{fold}}$	The periodic wavelet and scaling functions defined using the folding operation $f(\cdot) \mapsto f^{\mathrm{fold}}(\cdot) = \sum_{k\in\mathbb{Z}} f(\cdot - 2k) + \sum_{k\in\mathbb{Z}} f(2k - \cdot)$
ψ_k^{left}, ϕ_k^{left}	The special boundary-corrected wavelet and scaling functions defined at the left boundary 0
ψ_k^{right}, ϕ_k^{right}	The special boundary-corrected wavelet and scaling functions defined at the right boundary 1
$\psi_{j,l}^{\mathrm{int}}$, $\phi_{j,l}^{\mathrm{int}}$	The wavelet and scaling functions defined on the interval $[0, 1]$ which include the special boundary corrected functions

Chapter 1

Introduction

The recovery of a function from pointwise measurements of its Fourier transform is a fundamental task in signal processing. It arises in numerous applications, ranging from Magnetic Resonance Imaging (MRI) [PWS+99, SNF03, LDSP08, GKHPU11] to Computed Tomography (CT) [Eps08], electron microscopy [LPE12], helium atom scattering [JHA+09, JCAHT15], reflection seismology [BCS00] and radar imaging [BC05]. In many of these applications, the case when the data is acquired nonuniformly, i.e., along a non-Cartesian sampling pattern, is of a particular interest. For instance, MR scanners commonly use spiral sampling geometries for fast data acquisition [AKC86, MHNM92, KPH+97, SNF03, DHC+10, GKHPU11]. Such sampling geometries are often preferable because of the higher resolution obtained in the Fourier domain and the lower magnetic gradients required to scan along such trajectories. Another important example is radial sampling of the Fourier transform, which is also used in MRI as well as in applications where the Radon transform is involved in the sampling process, such as CT for instance [Eps08]. For examples of different sampling schemes used in applications see Figure 1.1. Spurred by its practical importance, the past decades have witnessed the development of an extensive mathematical theory of nonuniform sampling, as evidenced by a vast body of literature. An in-exhaustive list includes the books of Marvasti [Mar01], Benedetto and Ferreira [BF01], Young [You01], Seip [Sei04] and others, as well as many excellent articles; see [AG01, Ben92, BW00, FG94, FGS95, GS01, Str00b] and references therein.

The main purpose of this thesis is to answer the following question. Given fixed measurements of an unknown compactly supported L^2 function f in the form of a *finite collection of samples* of its Fourier transform $\hat{f}$, not necessarily taken on a Cartesian grid, under what conditions is it possible to recover an approximation to f in an *arbitrarily chosen finite-dimensional reconstruction space* T, and how can this be achieved via a stable numerical algorithm? To this end, the main contributions of this thesis are:

(i) a theoretical framework for understanding when such stable reconstruction is possible in terms of the *sampling density* and the *sampling bandwidth*;

(ii) a stable numerical algorithm to achieve such reconstruction with an *efficient* implementation in the univariate and bivariate cases when the reconstruction space T consists of wavelets;

(iii) complete analysis for the univariate case when reconstruction space T consists of *wavelets*, *non-equispaced splines* or *piecewise polynomials*; and

(iv) understanding of the reduction in the sampling density if the set of measurements includes samples of *derivatives* of $\hat{f}$.

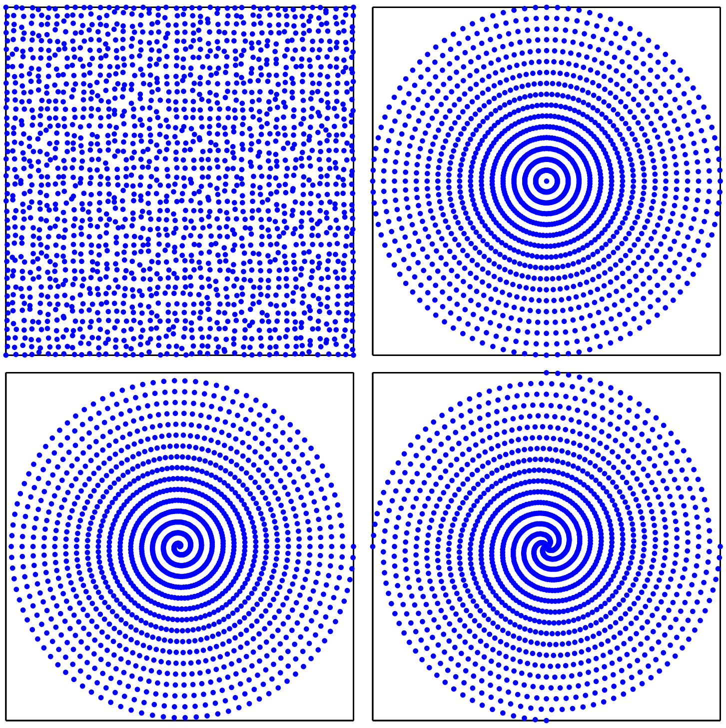

Figure 1.1: Different sampling schemes: (i) *jittered sampling scheme*, a standard model when a sensing devise does not acquire samples exactly on a uniform grid due to some measurement error, which is often used in MRI and geophysics [AG01, Mar01], (ii) *radial sampling scheme* used in MRI or in applications where Radon samples are acquired, such as CT for example [Eps08], (iii) *spiral* and (iv) *interleaving spiral* used for fast acquisition of data in MRI [DHC$^+$10]. All of these sampling schemes satisfy a (K, δ_{D°)-density condition (see Definition 3.3.1), for an appropriate Y, $D = [-1, 1]^2$, $\delta_{D^\circ} < 1/4$ and $K = 4$.

1.1 The generalized sampling approach

The approach we take in this work is based on recent developments in sampling and reconstruction in abstract Hilbert spaces, known as *generalized sampling* (GS). GS, in the form we consider in this paper, was introduced by Adcock & Hansen in [AH12a] (see also [AH12b, AH15b, AHP13, AHS14, AHP14, AHRT14]). Yet, its roots can be traced to earlier work of Unser & Aldroubi [UA94], Eldar [Eld04], Eldar & Werther [EW05], Gröchenig [Grö99, Grö01], Shizgal & Jung [JS04], Hrycak & Gröchenig [HG10], Aldroubi [Ald02], Gröchenig et al. [GRS10] and others.

GS addresses the following problem in sampling theory. Suppose that a finite number of samples of an element f of a Hilbert space are given as inner products with respect to a particular basis or frame. Suppose also that f can be efficiently represented in another basis or frame, for example, it has sparse or rapidly-decaying coefficients. GS obtains a stable reconstruction of f in this new system using only the original data. In the linear case, this is achieved by least-squares fitting [AH12b], but when sparsity is assumed, one can combine it with compressed sensing techniques to achieve substantial subsampling [AH15a]. By doing so, one obtains techniques for infinite-dimensional compressed sensing, known as GS–CS [AH15a, AHPR14].

The primary advantage of GS over most other approaches is that it allows one to take advantage of an efficient function representation by using a suitable reconstruction system. Namely, the free choice of the reconstruction basis or frame can be tailored to a specific application. In fact, it is well known that images are well represented using wavelets. Images may be sparse in wavelets, or have coefficients with rapid decay. Moreover, representing medical images in such systems has other benefits over classical Fourier series representations, such as improved compressibility, better feature detection and easier and more effective denoising [Lai00, Now98, WXHD91]. GS allows one to compute quasi-optimal reconstructions in wavelets from the given set of Fourier samples, and therefore exploit such beneficial properties. In the case of uniform Fourier samples, the use of GS/GS–CS with wavelets was studied in [AHP14, AHPR14].

1.2 Outline and contributions of the thesis

The main focus of this thesis is the case of GS with nonuniform Fourier measurements, which we refer to as *nonuniform generalized sampling* (NUGS). Since NUGS models sampling with weighted frames of exponentials, defined as in Chapter 2, it can be seen as a particular instance of GS corresponding to weighted Fourier frames.

Specifically, we assume the following setting. Let $d \geq 1$ denote dimension and $\mathbb{R}^d$ the d-dimensional Euclidean vector space. Following a standard convention, we use $\hat{\mathbb{R}}^d$ whenever $\mathbb{R}^d$ is considered as a frequency domain. Now, suppose that $\Omega = \{\omega_1, \dots, \omega_N\} \subseteq \hat{\mathbb{R}}^d$ is a set of N frequencies, and that we are given the measurements

$$\left\{ \hat{f}(\omega) : \omega \in \Omega \right\} \tag{1.2.1}$$

of an unknown signal $f \in \mathrm{L}^2(D)$, where $D \subseteq \mathbb{R}^d$ is compact. As typical in the aforementioned applications, set of frequencies Ω is fixed and cannot readily be altered. Let

$$\mathrm{T} \subseteq \mathrm{L}^2(D)$$

be a finite-dimensional space in which we wish to recover f. For example, T could consist of the first M functions in some wavelet basis.

In Chapter 3, we show how stable reconstruction in T can be achieved via NUGS using only given measurements (1.2.1), by presenting a general reconstruction framework together with guarantees for stable and accurate recovery. The guarantees are derived in terms of sampling *density* δ and sampling *bandwidth* K, where δ measures distance between the sampling points (see Definition 2.1.1) while K measures width of the sampling region. In particular, we show that if samples Ω have density $\delta < 1/4$ then stable reconstruction is possible provided the bandwidth K of samples Ω is sufficiently large. In the univariate case, the sufficient sampling bandwidth K depends solely on the properties of reconstruction space T. This later statement is also true in the multivariate case, but under more strict density condition. Furthermore, we address the case of critical density $\delta = 1/4$ in the univariate setting within the context of classical Fourier frames.

In order to develop such guarantees in the multivariate case, in Chapter 2, we first provide some novel results on weighted Fourier frames for spaces of multivariate compactly supported L^2 functions. By building upon the seminal work of Gröchenig [Grö92] and Beurling [Beu66], we allow for arbitrary clustering of sampling points subject to improved sampling densities. The results that address arbitrary clustering of sampling points Ω are of both theoretical and practical importance. Firstly, it is interesting to address the issue of arbitrary clustering, since it is natural to anticipate that adding more sampling points should not impair the recovery of a function. Secondly, this scenario often arises in applications. For example, consider Fourier measurements acquired on a radial sampling scheme. By increasing the number of radial lines along which samples are acquired, the sampling points cluster at low frequencies, which deteriorates the frame bounds of the corresponding Fourier frame. On the other hand, if we weight those points according to their relative densities, the resulting weighted Fourier frame has controllable frame bounds. To illustrate this last point, such a clustering of sampling points in a radial sampling scheme is depicted in Figure 1.2.

Specifically, in Chapter 2, our first result demonstrates how the separation condition can be successfully removed from Beurling's original result by using weights corresponding to the volumes of the Voronoi cells of the sampling points. Thereby we obtain the universal density condition $\delta < 1/4$, measured in a specific metric, sufficient to guarantee a weighted Fourier frame for the space of multivariate L^2 functions supported on a compact, convex and symmetric set, while simultaneously allowing arbitrary clustering of the sampling points. Unfortunately, this result does not lead to estimates of the frame

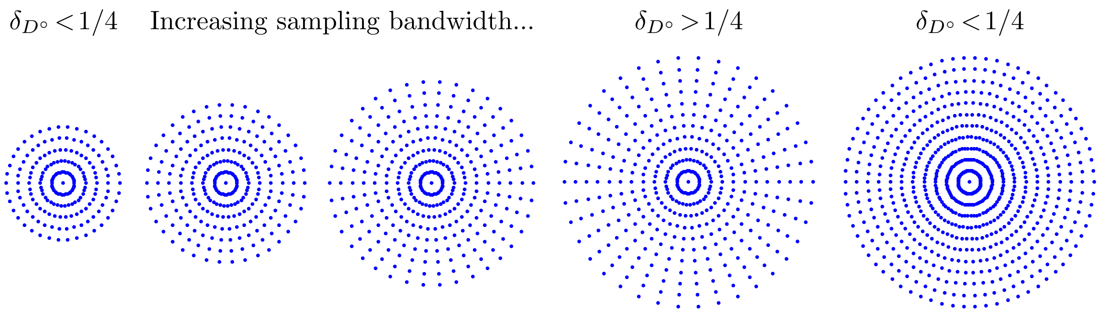

Figure 1.2: The first sampling scheme consists of points taken along 19 radial lines with the sampling bandwidth $K = 2$ such that the (K, δ_{D°)-density condition is satisfied with $\delta_{D^\circ} < 1/4$, $D = [-1,1]^2$, $Y = \mathcal{B}_1$ (see Definition 3.3.1). The following three sampling schemes are obtained by increasing the sampling bandwidth K, in order to see a larger fragment of the frequency domain, while the number of radial lines is kept the same. Thereby, one arrives to the forth sampling scheme which no longer satisfies the density condition. To have the desired density again, one may increase the number of radial lines from 19 to 35 and obtain the fifth sampling scheme. However, now the points at low frequencies become very close to each other, which deteriorates the corresponding frame bounds. In order to compensate for such clustering of sampling points, one may use appropriate weights.

bounds. Our second result, however, directly improves Gröchenig's explicit estimates of the frame bounds subject to a non-sharp density condition, which in certain cases also becomes dimension independent.

Next, having developed a general NUGS framework by using the results on weighted Fourier frames, in Chapter 4, we address the specific univariate case where the reconstruction space T corresponds to a wavelet basis. A result proved in [AHP14] for $d = 1$, and in [AHKM15] for $d = 2$, shows that when the sampling set Ω consists of the first N uniform frequencies one can recover the first $\mathcal{O}(N)$ coefficients in an arbitrary wavelet basis via GS. Thus wavelet bases are, up to constants, optimal bases in which to recover images from uniform Fourier samples. This is not true for example for algebraic polynomial bases, in which case one can stably recover only the first $\mathcal{O}(\sqrt{N})$ coefficients, see [HG10], as well as [AH12b, AHS14]. In Chapter 4 we extend the $d = 1$ result to the nonuniform case. Specifically, if the samples Ω have density $\delta < 1/4$ and bandwidth $K > 0$ then we prove that one can stably recover the first $\mathcal{O}(K)$ wavelet coefficients. Thus there is a linear relationship between the sampling bandwidth and the wavelet scale. As a corollary of this result, given samples of a smooth function $f \in \mathrm{H}^s$, NUGS obtains the convergence rate $\mathcal{O}(K^{-s})$ when recovering the coefficients of boundary-corrected wavelets. Furthermore, we show that any attempt—not restricted to NUGS—to reconstruct a fixed number of wavelet coefficients from a sampling bandwidth K below the critical threshold necessarily results in exponential ill-conditioning. This generalizes a result first proved in [AHP14] for uniform samples.

Additionally, in Chapter 4, we address implementation issues when recovering wavelet coefficients and demonstrate how the NUGS reconstruction can be computed efficiently by using Nonuniform Fast Fourier Transforms (NUFFTs) [FS03, PST01]. Namely, we show that NUGS requires only $\mathcal{O}(M \log N)$ operations when recovering N wavelet coefficients from M Fourier samples. At the end of this chapter, we provide a number of numerical examples simulated in MATLAB demonstrating aforementioned theoretical results. The code used to generate these examples presents the joint work with Clarice Poon and it is made publicly available at http://www.damtp.cam.ac.uk/user/mg617/GS_wavelets.zip.

In Chapter 5, we consider scaling of K and $\dim(\mathrm{T})$ sufficient for stable NUGS recovery of (piecewise) smooth functions in different polynomial spaces T from nonuniform samples of their Fourier transform. For this purpose, we derive guarantees for stable recovery in terms of two intrinsic quantities of the reconstruction space T, related to the maximal uniform growth of functions in T and the maximal growth of derivatives in T. For trigonometric polynomials, nonequidistant splines and piecewise algebraic polynomials with fixed polynomial degree, we show that this scaling is linear, and for piecewise algebraic polynomials with varying degree we show that it is quadratic.

In the final part, Chapter 6, we consider two different but related sampling scenarios for bandlimited functions. First, we provide sufficient density conditions for a set of nonuniform samples to give rise to a frame for the space of multivariate bandlimited functions when the measurements consist of pointwise evaluations of a function and its first k derivatives. This problem is motivated by applications in seismology, where certain recently developed detectors are able to measure both f and its spatial gradient. However, there are also various other applications and for different examples we direct the reader to [EO00, LSP$^+$03] and references therein. The second scenario considered in this chapter assumes that, instead of evaluating derivatives of f at $\{x_n\}_{n\in I}$, f is measured at an additional s sampling points around each x_n. One can think of this scenario as function f being evaluated at $s+1$ different nonuniform sampling sequences. When these sequences are uniform, the problem is known as bunched sampling or recurrent nonuniform sampling, and has been extensively studied in literature.

The purpose of Chapter 6 is to understand the gain one can expect by nonuniformly sampling derivatives or by nonuniformly sampling at bunched points. Although we do not discuss actual function recovery in this case, it also can be performed via NUGS. However, we do derive explicit sufficient conditions for stable recovery in terms of densities of sampling points as well as explicit estimates of the corresponding frame bounds. In particular, we show that the maximal allowed gap between sampling points (or bunches of sampling points) grows linearly in $k+1$ (or $s+1$) for large k (or s), which translates into increasing savings in cost and effort in practical acquisition of data. For a detailed description of the main results on this topic, we refer to Section 6.1.

1.2.1 List of publications

The novel results on weighted Fourier frames, which are presented in Chapter 2, were published in [AGH15b] as the joint work of the author with Ben Adcock and Anders Hansen. The NUGS framework from Chapter 3, along with the guarantees in the one-dimensional case, was published in [AGH14a], while the multidimensional case was considered in [AGH15b], both presenting the joint work of the author with Ben Adcock and Anders Hansen. The detail analysis of the wavelet case from Chapter 4 was also published in [AGH14a]. The efficient implementation of the wavelet reconstruction is the joint work of the author with Clarice Poon and this is published in [GP15]. Analysis of reconstruction in various polynomial spaces presented in Chapter 5 was published in [AGH14b] as the joint work of the author with Ben Adcock and Anders Hansen. The contributions of Chapter 6, on the topic of derivative and bunch sampling, were collected in [AGH15a] and submitted for publication as the joint work of the author with Ben Adcock and Anders Hansen.

1.3 Relation to previous work

Beside its key relation to generalized sampling, this work also relates to different aspects in nonuniform sampling theory and to a vast body of existing literature within this field.

1.3.1 Recovery from nonuniform samples

An algorithm commonly used for MRI reconstruction from nonuniform samples is so-called gridding [JMNM91, SN00, VGCR10, GS14], which simply discretizes the Fourier integral on a nonuniform mesh. Our work differs from this approach in that we assume an analog model for the image f, as opposed to viewing f as a finite-length Fourier series. Consequently, a key issue in NUGS is that of *approximation*. By using an appropriate reconstruction space T, we avoid the unpleasant artifacts (e.g. Gibbs ringing) associated with this algorithm.

Another popular method for MRI reconstruction is the iterative reconstruction algorithm [SNF03, MFK04], see also [PWS$^+$99, PWBB01]. This can be viewed as a special instance of NUGS, where T is a space of piecewise constant functions on a $M \times M$ grid (the term 'iterative' refers to the use of conjugate gradients to compute the reconstruction). Equivalently, when M is a power of 2, then T can be expressed as the space spanned by Haar wavelets up to some finite scale. Thus our work provides as a corollary theoretical guarantees for the stability and error of this algorithm. Importantly, we shall also show how NUGS allows one to obtain better reconstructions, by replacing the Haar wavelet choice for the subspace T with higher-order wavelets.

In [Grö99] (see also [Grö01, GS01, FGS95, Grö93]), the problem of recovering a bandlimited function from its own nonuniform samples was considered, where the arbitrary clustering is addressed by using weighted Fourier frames, exactly the same as we do in this work. Specifically, Gröchenig et al. developed an efficient algorithm for the nonuniform sampling problem, known as the ACT algorithm (Adaptive weights, Conjugate gradients, Toeplitz) where they consider reconstruction of bandlimited functions in a particular finite-dimensional space consisting of trigonometric polynomials. This corresponds to a specific instance of NUGS with a Dirac basis for T. Convergence and stability of the ACT algorithm [Grö01, Thm. 7.1] are guaranteed by the sufficient sampling density and the explicit weighted frame bounds given in [Grö01, Prop. 7.3]. Note that the focus of the present work is slightly different. Gröchenig et al. primarily consider the recovery of a bandlimited function from nonuniform pointwise samples, whereas we consider the recovery of a compactly supported function from pointwise samples of its Fourier transform. Although mathematically equivalent, the setup affects the choice of reconstruction space. In our setting for example, a Dirac basis would not be ideal for approximating an image f, whereas wavelet bases are typically well suited. Having said this, the results we prove here extend the work by Gröchenig et al. in two ways. First, we have a less stringent multi-dimensional density requirement based on the improved results for weighed frames of exponentials, which also directly improves the guarantees for ACT algorithm. Second, our framework allows arbitrary choices of T which can be tailored to the particular function f to be recovered.

Some of the earlier work in nonuniform sampling theory considers reconstruction of an unknown function based on an iterative inversion of the frame operator [Ben92, Ben93, BW00, FG94, Grö92, AG01]. These approaches would be fine if one would be given infinitely-many samples and infinite processing time, but since one has only finite data in practice, they typically lead to large truncation errors (similar to Gibbs phenomena). Additionally, such approaches are typically infeasible in more than one dimension due to their computational complexity.

1.3.2 Sets of sampling

In contrast to Cartesian sampling which leans on the celebrated Nyquist–Shannon sampling theorem as well as Parseval's identity, nonuniform sampling is typically studied within the concept of Fourier frames. Provided one has a Fourier frame, stable function recovery is possible and can be carried out via different algorithms. Therefore, it is crucial to understand conditions under which sampling points give rise to a Fourier frame. This topic has been extensively researched in the last several decades [AG01, Chr01, Ben93, BW00]. Sampling points that give rise to a Fourier frame for the space of L^2 functions supported on a compact domain, equivalently provide a frame for the space of functions bandlimited to the same compact domain [You01]. In nonuniform sampling literature, such a collection of sampling points is typically called a *set of sampling* [BW00]. In this thesis, we also study such sets.

The theory of Fourier frames was developed by Duffin and Shaeffer [DS52], more than half a century ago, and its roots can be traced back to earlier works of Paley and Wiener [PW87] and Levinson [Lev40]. In one dimension, there exists a near-complete characterization of Fourier frames in terms of the density of underlying samples, due primarily to Beurling [Beu66], Landau [Lan67], Jaffard [Jaf91] and Seip [Sei95a]. However, in higher dimensions, the situation becomes considerably more complicated [BW00, OU12]. Nevertheless, Beurling's seminal paper [Beu66] (see also [Beu89]) provides a sharp sufficient condition for sampling points in multiple dimensions to give rise to a Fourier frame for the space of square-integrable functions compactly supported on a sphere. This was generalized to the spaces of square-integrable functions compactly supported on any compact, convex and symmetric set by Benedetto and Wu [BW00] and also by Olevskii and Ulanovskii [OU12]. Regarding general bounded supports in $\mathbb{R}^d$, Landau [Lan67] provides a necessary density condition that fails to be sufficient in general. A recent result due to Matei and Meyer [MM10] proves this density condition to be sufficient in the special case of sampling on quasicrystals. Also, some of these density-type results where extended to shift-invariant spaces by Aldroubi and Gröchenig [AG00]. However, in our work, we focus only on compactly supported and square-integrable functions with supports in $\mathbb{R}^d$ which are compact, convex and symmetric. For a more detailed review on the theory of Fourier frames and nonuniform sampling, see [AG01, BW00, Chr01].

A limitation of the results mentioned above is that they require a minimal separation between the sampling points. In particular, clustering of sampling points deteriorates the associated upper frame bound. The result we present here removes the minimal separation restriction while it keeps the density condition sharp and dimensionless. Through the use of a weighted Fourier frame approach, based on Gröchenig's earlier work, we adapt Beurling's result to allow for arbitrary clustering of sampling points. The density condition given here is sharp in the sense that there exists a sampling set with smaller density and a function which could not be recovered from that set.

Weighted Fourier frames, which we also refer to as weighted frames of exponentials, were studied by Gröchenig [Grö92], and later also by Gabardo [Gab93]. In [Grö92], Gröchenig presents a density condition sufficient for a family of exponentials to constitute a weighted Fourier frame, and provides explicit frame bounds. This density condition is sharp in dimension $d = 1$, but fails to be sharp in higher dimensions, with the estimate on the density deteriorating linearly, and the estimates on the frame bounds, exponentially in d. The multi-dimensional result has been improved in [BG05], but under the assumption that the sampling set consists of a sequence of uniformly distributed independent random variables. In this setting, Bass and Gröchenig provide probabilistic estimates. Our work focuses on deterministic statements and provides two improvements of Gröchenig's result from [Grö92]. First, as discussed above, we provide a density condition which is both sharp *and* dimensionless. However, as in the original Beurling's result, this condition does not give rise to explicit frame bounds. Therefore, in our second result we present explicit frame bounds under a non-sharp, but at the same time, a less stringent density condition than previously known.

We note at this stage that, whilst Gröchenig was arguably the first to rigorously study weighted Fourier frames in sampling, the use of weights is commonplace in MRI reconstructions, where they are often referred to as 'density compensation factors' (see [DHC+10, SNF03] and references therein). However, such approaches are often heuristic. Building on Gröchenig's earlier work, our results provide further mathematical theory supporting for their use.

It is also of practical and theoretical interest to study conditions under which a set of nonuniform samples give rise to a set of sampling, once derivatives of a bandlimited functions are additionally evaluated at these points. Uniform sampling of derivatives is a classical topic in sampling theory, see [Fog56, JF56, LA60, Pap77a, Pap77b, Raw89, ZSCZ96] and references therein. It is known that one can exceed the Nyquist criterion by a factor of $k+1$ by sampling the function and its first k derivatives [Pap77b, Raw89]. On the other hand, relatively few papers have considered nonuniform sampling with derivatives. In the univariate setting, by extending Gröchenig's results [Grö92] for univariate nonuniform sampling to the case of derivatives, it was shown in [Raz95] that the maximum allowable spacing between sampling points grows asymptotically as a linear function of $k+1$, with constant of proportionality equal to $1/e$. Furthermore, multivariate nonuniform sampling with derivatives was addressed in [GR96], where necessary density conditions were derived. However, to the best of our knowledge, no work has addressed sufficient guarantees for stable sampling with derivatives in the multivariate setting, and this is the task addressed in the present work.

Lastly, we consider sets of sampling arising from nonuniform bunched sampling of bandlimited functions. Uniform bunched sampling—also known as recurrent nonuniform sampling since it assumes periodic groups of nonuniform samples—has been a topic of numerous papers, in both the one-dimensional [BH89, EO00, Koh53, Pap77a, Pap77b, SJ08, ST06, Yen56] and the multi-dimensional case [Far94, FG05]. Here, as in the derivative sampling, one is allowed to sample above the Nyquist rate. Namely, if the uniform bunched set of sampling points is interpreted as the union of $s+1$ uniform sequences, then each of these sequences can be taken at $s+1$ times the Nyquist rate. However, one might want to know what happens if these groups of nonuniform samples are not repeating periodically, but instead are distributed nonuniformly. This setting corresponds to nonuniform bunched sampling that we consider in our work. To the best of our knowledge, there is no earlier work which considers nonuniform bunched sampling.

Chapter 2

Weighted frames of exponentials

The subject of this chapter is conditions that ensure existence of a frame for the space of L^2 functions supported on a compact domain in $\mathbb{R}^d$. In this regard, we provide two novel results. Whilst keeping the density condition sharp and dimension independent, our first result, Theorem 2.3.1, removes the separation condition from Beurling's result [Beu66] and shows that density alone suffices for obtaining a frame for the space of L^2 functions supported on a compact symmetric and convex domain in $\mathbb{R}^d$. This is achieved by the use of appropriate weights, leading to a weighted Fourier frame. However, this result does not lead to estimates for the frame bounds. A result of Gröchenig [Grö92] provides explicit estimates, but only subject to a density condition that deteriorates linearly with dimension d. In our second result, Theorem 2.2.1, we improve these bounds by reducing the dimension dependence. In particular, we provide explicit frame bounds which are dimensionless for functions having compact support contained in a sphere.

The results of this chapter are collected from [AGH15b], which is the joint work of the author with Ben Adcock and Anders Hansen.

2.1 Background material and preliminaries

For $x = (x_1, \dots, x_d) \in \mathbb{R}^d$ and $y = (y_1, \dots, y_d) \in \mathbb{R}^d$ we write $x \cdot y = x_1 y_1 + \ldots + x_d y_d$ for the dot product of x and y, and for $p \geq 1$, we write $|x|_p$ for the ℓ^p-norm, i.e. $|x|_p = \left(\sum_{j=1}^d |x_j|^p\right)^{1/p}$. Let $\mathrm{L}^2(\mathbb{R}^d)$ be the space of square-integrable functions on $\mathbb{R}^d$ with inner product

$$\langle f, g \rangle = \int_{\mathbb{R}^d} f(x) \overline{g(x)} \, \mathrm{d}x,$$

and corresponding norm $\|f\| = \sqrt{\langle f, f \rangle}$. We denote the Fourier transform by

$$\hat{f}(\omega) = \int_{\mathbb{R}^d} f(x) \mathrm{e}^{-\mathrm{i}2\pi\omega \cdot x} \, \mathrm{d}x, \quad \omega \in \hat{\mathbb{R}}^d.$$

Let

$$\mathrm{H} = \left\{ f \in \mathrm{L}^2(\mathbb{R}^d) : \mathrm{supp}(f) \subseteq D \right\}$$

be the Hilbert space of square-integrable functions supported on a compact set $D \subseteq \mathbb{R}^d$. For a point in the frequency domain $\omega \in \hat{\mathbb{R}}^d$, we define

$$e_\omega(x) = \mathrm{e}^{\mathrm{i}2\pi\omega\cdot x}\chi_D(x), \quad x \in \mathbb{R}^d,$$

where χ_D is the indicator function of the set D. Note that, for a $f \in \mathrm{H}$, we have $\hat{f}(\omega) = \langle f, e_\omega \rangle$.

Let $|\cdot|_*$ denote an arbitrary norm on $\mathbb{R}^d$. Note that for every such norm the set $\{x \in \mathbb{R}^d : |x|_* \leq 1\}$ is convex, compact and symmetric. Since all norms on a finite-dimensional space are equivalent to the Euclidean norm, which we denote simply by $|\cdot|$, there exist (sharp) constants $c_*, c^* > 0$ such that

$$\forall x \in \mathbb{R}^d, \quad c_*|x|_* \leq |x| \leq c^*|x|_*.$$

Additionally, if $D \subseteq \mathbb{R}^d$ is a compact, convex and symmetric set, the function $|\cdot|_D : \mathbb{R}^d \to \mathbb{R}$ defined by

$$\forall x \in \mathbb{R}^d, \quad |x|_D = \inf\{a > 0 : x \in aD\},$$

is a norm on $\mathbb{R}^d$ [BW00]. Here, D is the unit ball with the respect to the norm $|\cdot|_D$, i.e.

$$D = \{x \in \mathbb{R}^d : |x|_D \leq 1\}.$$

For such set $D \subseteq \mathbb{R}^d$, its polar set is defined as

$$D^\circ = \{\hat{y} \in \hat{\mathbb{R}}^d : \forall x \in D,\ x \cdot \hat{y} \leq 1\}.$$

Note that D° is itself a convex, compact and symmetric set in $\hat{\mathbb{R}}^d$, which is the unit ball with respect to the norm $|\cdot|_{D^\circ}$. Also observe that, if D is the unit ball in the Euclidean norm, which we denote by $\mathcal{B}_1$, then $\mathcal{B}_1 = \mathcal{B}_1^\circ$ and $|\cdot|_{\mathcal{B}_1} = |\cdot|_{\mathcal{B}_1^\circ} = |\cdot|$.

Throughout, we denote ℓ^p-norm by $|\cdot|_p$, i.e. for $x \in \mathbb{R}^d$, $|x|_p = \left(\sum_{j=1}^d |x_j|^p\right)^{1/p}$. Hence $|\cdot|_2 = |\cdot|_{\mathcal{B}_1} = |\cdot|$. Also, we recall the well-know inequality

$$\forall x \in \mathbb{R}^d, \quad |x|_q \leq |x|_r \leq d^{1/r-1/q}|x|_q, \quad q > r > 0, \tag{2.1.1}$$

which we shall use later.

2.1.1 Classical Fourier frames

A family $\{\phi_n\}_{n\in\mathbb{N}} \subseteq \mathrm{H}$ is called a *frame* for the Hilbert space H if there exist constants $A, B > 0$ such that

$$\forall f \in \mathrm{H}, \quad A\|f\|^2 \leq \sum_{n\in\mathbb{N}} |\langle f, \phi_n\rangle|^2 \leq B\|f\|^2.$$

The optimal constants A and B are called the *upper* and the *lower frame bound*, respectively. For an excellent overview of frame theory see [Chr03] as well as [Chr01] and [BW00].

In what follows we will be interested in Fourier frames. For a countable set of sampling points $\Omega \subseteq \hat{\mathbb{R}}^d$, a family of functions $\{e_\omega\}_{\omega\in\Omega} \subseteq \mathrm{H}$ is called a *Fourier frame* for H if there exist constants $A, B > 0$ such that

$$\forall f \in \mathrm{H}, \quad A\|f\|^2 \leq \sum_{\omega\in\Omega} |\hat{f}(\omega)|^2 \leq B\|f\|^2. \tag{2.1.2}$$

We also refer to such a system as a *classical* Fourier frame. If $\{e_\omega\}_{\omega\in\Omega}$ is a frame, then the *frame operator* $\mathcal{S} : \mathrm{H} \to \mathrm{H}$ is defined by

$$\forall f \in \mathrm{H}, \quad \mathcal{S} : f \mapsto \mathcal{S}f = \sum_{\omega\in\Omega} \hat{f}(\omega) e_\omega. \tag{2.1.3}$$

Since the frame inequality (2.1.2) holds, the frame operator $\mathcal{S}$ is is linear, bounded, self-adjoint and invertible, with the inverse $\mathcal{S}^{-1} : \mathrm{H} \to \mathrm{H}$ satisfying

$$\forall f \in \mathrm{H}, \quad f = \sum_{\omega\in\Omega} \langle \mathcal{S}^{-1} f, e_\omega\rangle e_\omega. \tag{2.1.4}$$

Formula (2.1.4), with the appropriately truncated sum, is sometimes used for signal reconstruction [BW00]. However, for the types of sets Ω considered in practice, finding the inverse frame operator $\mathcal{S}^{-1}$ is often a nontrivial task. Typically, this renders such an approach infeasible in more than one dimension.

If the relation (2.1.2) holds with $A = B$, the family $\{e_\omega\}_{\omega\in\Omega}$ is called a tight frame, and if $A = B = 1$, this family forms an orthonormal basis for H. In these cases, the relation (2.1.2) is known as (generalized) Parseval's equality. Also, in these cases the frame operator becomes $\mathcal{S} = A\mathcal{I}$, where $\mathcal{I}$ is the identity operator on H, and the formula (2.1.4) represents the Fourier series of f, which, when appropriately truncated, converges strongly to f on H. This leads to a considerably simpler framework in the case when the samples are acquired uniformly, corresponding to an orthonormal basis or a tight frame for H.

Necessary and sufficient conditions

If $\{e_\omega : \omega \in \Omega\}$ is a classical Fourier frame, then Ω necessarily cannot have a clustering point, i.e. Ω must be (relatively) separated, or otherwise the upper frame bound blows up [Jaf91]. The set Ω is said to be *separated* with respect to the $|\cdot|_*$-norm if there exists a constant $\eta > 0$ such that

$$\forall \omega, \lambda \in \Omega, \quad \omega \neq \lambda, \quad |\omega - \lambda|_* \geq \eta,$$

and it is *relatively separated* if it is a finite union of separated sets. It is clear that, if Ω is separated in the $|\cdot|_*$-norm then it is separated in any norm on $\hat{\mathbb{R}}^d$ and vice-versa.

Beside separation, another characterizing property of Fourier frames is *density* of the underlying sampling points. The following definition of density originates in Beurling's work [Beu66] and it is used frequently in multi-dimensional nonuniform sampling literature.

Definition 2.1.1 (Sampling density)**.** *Let Ω be a sampling set contained in a closed, simply connected set $Z \subseteq \hat{\mathbb{R}}^d$. Let $|\cdot|_*$ be an arbitrary norm on $\mathbb{R}^d$ and let $\delta_* > 0$. We say that Ω is δ_*-dense in the domain Z if*

$$\delta_* = \sup_{\hat{y} \in Z} \inf_{\omega \in \Omega} |\omega - \hat{y}|_*.$$

If $|\cdot|_* = |\cdot|_S$ for a compact, convex and symmetric set $S \subseteq \mathbb{R}^d$, then we write δ_S. Also, to emphasise the sampling set, where necessary we use the notation $\delta_*(\Omega)$.

It is useful to note that δ_*-density condition from Definition 2.1.1 is equivalent to the δ_*-*covering condition*: for all $\rho \geq \delta_*$ it holds that

$$Z \subseteq \bigcup_{\omega \in \Omega} \left\{ x \in \mathbb{R}^d : |x - \omega|_* \leq \rho \right\}.$$

In other words, δ_* is the minimal radius of $|\cdot|_*$-balls described around the sampling points in Ω necessary to cover Z. In particular, the half distance between any two sampling points measured in the $|\cdot|_*$-norm cannot exceed δ_*. Moreover, in one dimension, δ_* is exactly the half length of the maximum gap between the sampling points of Ω.

In [Beu66], Beurling provides a sufficient density condition for a nonuniform set of sampling points to give a Fourier frame for H consisting of functions supported on the unit sphere in the Euclidean norm. In what follows, we use a variation of Beurling's result given by Benedetto & Wu in [BW00], and also by Olevskii & Ulanovskii [OU12], which is a generalization to arbitrary convex, compact and symmetric supports:

Theorem 2.1.2 (Beurling's theorem)**.** *Let $D \subseteq \mathbb{R}^d$ be compact, convex and symmetric* *set. If* $\Omega \subseteq \hat{\mathbb{R}}^d$ *is relatively separated and* δ_{D°*-dense in* $\hat{\mathbb{R}}^d$ *with*

$$\delta_{D^\circ} < \frac{1}{4},$$

then the family of functions $\{e_\omega\}_{\omega\in\Omega}$ *is a Fourier frame for* H.

Beurling [Beu66] also shows that this result is sharp in the sense that there exists a countable set with the density $\delta_{D^\circ} = 1/4$, where D is the unit ball in the Euclidean metric, which does not satisfy the lower frame condition in (2.1.2) (see also [OU12, Prop. 4.1]).

In the one-dimensional case, however, the list of existing results for nonuniform sampling is much more complete. Most notably, there exists a near-complete characterization of Fourier frames in terms of relative separation and the Beurling density. For a sequence $\Omega \subseteq \hat{\mathbb{R}}$, the lower Beurling density is defined by

$$\rho^- = \lim_{r\to\infty} \frac{n^-(r)}{r}, \qquad n^-(r) = \inf_{t\in\mathbb{R}} |\{\omega\in\Omega : \omega \in (t, t+r)\}| .$$

Note that by definition $1/\rho^- = 2\delta_{D^\circ}$, for $D = [-1, 1]$. The results of the following theorem are due to Duffin and Schaeffer [DS52], Landau [Lan67], Jaffard [Jaf91] and Seip [Sei95a].

Theorem 2.1.3 (One-dimensional characterization of Fourier frames)**.** *Let* $\Omega \subseteq \hat{\mathbb{R}}$ *be a sampling set and let* $\mathrm{H} = \{f \in \mathrm{L}^2(\mathbb{R}) : \mathrm{supp}(f) \subseteq [-1, 1]\}$.

(*i*) *If* Ω *is relatively separated and* $\rho^- > 2$ *then* $\{e_\omega\}_{\omega\in\Omega}$ *forms a frame for* H.

(*ii*) *Conversely, If* $\{e_\omega\}_{\omega\in\Omega}$ *forms a frame for* H *then* $\rho^- \geq 2$ *and* Ω *is relatively separated.*

Note that there exist both a relatively separated sequence with $\delta_{D^\circ} = 1/4$ which forms a frame and a relatively separated sequence with $\delta_{D^\circ} = 1/4$ which does not. We refer to [Chr01] and [BW00] for details. Instead, we only note that $\delta_{D^\circ} = 1/4$ is obtained exactly when sampling at the Nyquist rate in the uniform setting, and therefore $\delta_{D^\circ} = 1/4$ is allowed in this particular case. However, in general, nonuniform sampling requires sampling just above the Nyquist rate.

2.1.2 Weighted Fourier frames

To compensate for arbitrary clustering of sampling points, which often needs to be facilitated in practice, it is common to use weights, also known as density compensation factors.

Definition 2.1.4 (Weighted Fourier frames)**.** *A countable family of functions* $\{\sqrt{\mu_\omega} e_\omega\}_{\omega\in\Omega}$ *is a weighted Fourier frame for* H*, with weights* $\{\mu_\omega\}_{\omega\in\Omega}$*,* $\mu_\omega > 0$*, if there exist constants*

$A, B > 0$ *such that*

$$\forall f \in \mathrm{H}, \quad A\|f\|^2 \leq \sum_{\omega\in\Omega} \mu_\omega |\hat{f}(\omega)|^2 \leq B\|f\|^2. \tag{2.1.5}$$

In order to define appropriate weights $\{\mu_\omega\}_{\omega\in\Omega}$ corresponding to the varying density of the sampling set Ω, we use the Lebesgue measure of Voronoi regions. This is a standard practice in nonuniform sampling [AG01, RPS+99].

Definition 2.1.5 (Voronoi regions)**.** *Let Ω be a set of distinct points in a domain $Z \subseteq \hat{\mathbb{R}}^d$ and let $|\cdot|_*$ be an arbitrary norm on $\mathbb{R}^d$. The Voronoi region at $\omega \in \Omega$, with respect to the norm $|\cdot|_*$ and in domain Z, is given by*

$$V_\omega^* = \{\hat{y} \in Z : \forall\lambda \in \Omega,\ \lambda \neq \omega,\ |\omega - \hat{y}|_* \leq |\lambda - \hat{y}|_*\},$$

with the Lebesgue measure denoted as

$$\mathrm{meas}\,(V_\omega^*) = \int_Z \chi_{V_\omega^*}(\hat{y})\,\mathrm{d}\hat{y}.$$

For an example of Voronoi regions associated to a set of sampling points taken on a spiral see Figure 2.1. Note that as points get close to each other the associated Voronoi regions become smaller.

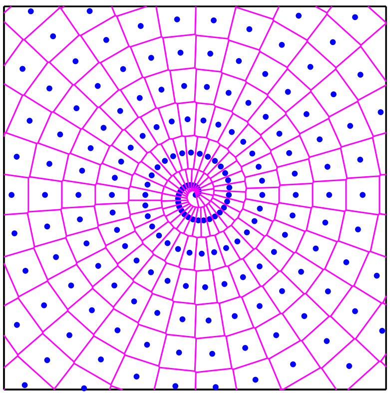

Figure 2.1: Edges of Voronoi regions (magenta) associated to a set of sampling points taken on a spiral (blue), with respect to the Euclidean norm.

In [Grö92], Gröchenig provides explicit frame bounds for weighted Fourier frames, provided the sample points Ω are sufficiently dense. In one dimension, the condition on the density is sharp and reads as follows:

Theorem 2.1.6 (Gröchenig's one-dimensional theorem)**.** *Let* $\mathrm{H} = \{f \in \mathrm{L}^2(\mathbb{R}) : \mathrm{supp}(f) \subseteq$

$D\}$, where $D = [-1, 1]$. If $\Omega \subseteq \hat{\mathbb{R}}$ is δ-dense in $\hat{\mathbb{R}}$ such that

$$\delta_{D^\circ} < \frac{1}{4},$$

then $\{\sqrt{\mu_\omega} e_\omega\}_{\omega \in \Omega}$ is a weighted Fourier frame for H, where the weights are defined as measures of the Voronoi intervals of the points Ω, with the frame bounds A, B satisfying

$$\sqrt{A} \geq 1 - 4\delta_{D^\circ} > 0, \qquad \sqrt{B} \leq 1 + 4\delta_{D^\circ} < 2.$$

However, the sharpness of this result is lost in higher dimensions. Here we state Gröchenig's multi-dimensional result [Grö01, Prop. 7.3], which is a more recent reformulation of [Grö92, Thm. 5]:

Theorem 2.1.7 (Gröchenig's theorem). *Let $\mathrm{H} = \{f \in \mathrm{L}^2(\mathbb{R}^d) : \mathrm{supp}(f) \subseteq D\}$, where $D = [-1, 1]^d$. If $\Omega \subseteq \hat{\mathbb{R}}^d$ is a $\delta_{\mathcal{B}_1}$-dense set in $\hat{\mathbb{R}}^d$ such that*

$$\delta_{\mathcal{B}_1} < \frac{\ln 2}{2\pi d}, \tag{2.1.6}$$

then $\{\sqrt{\mu_\omega} e_\omega\}_{\omega \in \Omega}$ is a weighted Fourier frame for H, where the weights are defined as measures of the Voronoi regions of the points Ω with respect to Euclidean norm. The frame bounds A, B satisfy

$$\sqrt{A} \geq 2 - e^{2\pi\delta_{\mathcal{B}_1} d} > 0, \qquad \sqrt{B} \leq e^{2\pi\delta_{\mathcal{B}_1} d} < 2.$$

Note that the bound (2.1.6) deteriorates linearly with the dimension d. Also, D can be any rectangular domain of the form $\prod_{i=1}^d [-s_i, s_i]$, since $\mathrm{supp}(f) \subseteq \prod_{i=1}^d [-s_i, s_i]$ implies that $\tilde{f}(x) = f(x_1/s_1, \ldots, x_d/s_d)$ has support in $[-1, 1]^d$. Hence, the result is stated for $D = [-1, 1]^d$ without loss of generality [Grö01]. Moreover, note that D may also be any compact set that is a subset of $[-1, 1]^d$ such as any ℓ^p unit ball, $p > 0$, for example.

2.2 Weighted Fourier frames with improved frame bounds

Much like Beurling's Theorem 2.1.2, it is expected that the density condition for weighted Fourier frames given in Theorem 2.1.7 does not depend on dimension. Unfortunately, Gröchenig's estimates deteriorate linearly with the dimension d, and thus cease to be sharp. Therefore, in the following theorem, we provide a modification of Gröchenig's theorem by presenting explicit bounds with slower, and sometimes, with no deterioration with respect to dimension.

Theorem 2.2.1. *Let $\mathrm{H} = \{f \in \mathrm{L}^2(\mathbb{R}^d) : \mathrm{supp}(f) \subseteq D\}$, where $D \subseteq \mathbb{R}^d$ is compact. Suppose that $|\cdot|_*$ is an arbitrary norm on $\mathbb{R}^d$ and $c^* > 0$ is the smallest constant for which*

$|\cdot| \leq c^* |\cdot|_*$, *where* $|\cdot|$ *denotes the Euclidean norm. Let* $\Omega \subseteq \hat{\mathbb{R}}^d$ *be* δ_**-dense in* $\hat{\mathbb{R}}^d$ *with*

$$\delta_* < \frac{\ln 2}{2\pi m_D c^*}, \tag{2.2.1}$$

where $m_D = \sup_{x\in D} |x|$. *Then* $\{\sqrt{\mu_\omega} e_\omega\}_{\omega\in\Omega}$ *is a weighted Fourier frame for* H *with the weights defined as the measures of Voronoi regions with respect to norm* $|\cdot|_*$. *The weighted Fourier frame bounds* A, B *satisfy*

$$\sqrt{A} \geq 2 - \exp(2\pi m_D \delta_* c^*) > 0, \qquad \sqrt{B} \leq \exp(2\pi m_D \delta_* c^*) < 2.$$

The estimates in Theorem 2.2.1 are presented in terms of the following quantity

$$m_D = \sup_{x\in D} |x|, \tag{2.2.2}$$

where $D \subseteq \mathbb{R}^d$ and $|\cdot|$ is Euclidean norm. Note that $m_{\mathcal{B}_1} = 1$ and therefore it is independent of dimension for spheres. Moreover, if D is the ℓ^p unit ball, i.e. $D = \{x : \mathbb{R}^d : |x|_p \leq 1\}$, $p > 0$, then

$$m_D = \max\{1, d^{1/2-1/p}\}, \tag{2.2.3}$$

due to inequality (2.1.1).

Remark 2.2.2 We first note that if the sampling density and Voronoi regions are defined in the Euclidean norm, i.e., if $|\cdot|_* = |\cdot|$, which is typically the case in practice, then $c^* = 1$. If additionally D is taken to be the unit Euclidean ball, which corresponds to Beurling's original setting, then $m_D = 1$. In this particular case, the dimension dependence is completely removed and the density condition (2.2.1) reads

$$\delta < \frac{\ln 2}{2\pi} \approx 0.11.$$

This is slightly stronger than the sharp condition $\delta < 0.25$ (see Theorem 2.3.1), but still, it is a dimension independent condition under which the *explicit* frame bounds are provided.

To illustrate this density condition further, let $D = \{x \in \mathbb{R}^d : |x|_p \leq 1\}$, $p > 0$, and let $|\cdot|_*$ be the ℓ^q norm, $q \geq 1$. Then (2.2.1) becomes

$$\delta_q < \frac{\ln 2}{2\pi \max\{1, d^{1/2-1/p}\} \max\{1, d^{1/2-1/q}\}}, \tag{2.2.4}$$

due to (2.1.1) and (2.2.3). This bound attains its minimum for $p = q = \infty$, in which case it deteriorates linearly with the dimension d. However, in all other cases the deterioration of the bound on density, and also, the deterioration of weighted frame bounds estimations, is slower with the dimension. Moreover, they are independent of dimension whenever $p \leq 2$ and $q \leq 2$.

Finally, to directly compare this theorem with Gröchenig's multi-dimensional result given in Theorem 2.2.1, we set $p = \infty$ and $q = 2$ in (2.2.4). Thus, the density condition (2.2.4) becomes

$$\delta_2 < \frac{\ln 2}{2\pi\sqrt{d}}$$

whereas the density condition (2.1.6) in Gröchenig's theorem is

$$\delta_2 < \frac{\ln 2}{2\pi d}.$$

Hence Theorem 2.2.1 leads to an improvement by a factor of $\sqrt{d}$ and no deterioration in the constant $(\ln 2)/(2\pi)$.

Before we proceed to the proof of Theorem 2.2.1, let us recall the multinomial formula. For any $k \in \mathbb{N}_0$ and $x \in \mathbb{R}^d$, we have

$$\sum_{|\alpha|_1 = k} \frac{k!}{\alpha!} x^\alpha = (x_1 + \cdots + x_d)^k, \tag{2.2.5}$$

where $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}_0^d$, $|\alpha|_1 = |\alpha_1| + \ldots + |\alpha_d|$, $\alpha! = \prod_{j=1}^d \alpha_j!$ and $x^\alpha = \prod_{j=1}^d x_j^{\alpha_j}$. Regarding the multi-index notation, in what follows, we also use the derivative operator defined as

$$D^\alpha = \frac{\partial^{|\alpha|_1}}{\partial_{x_1}^{\alpha_1} \cdots \partial_{x_d}^{\alpha_d}}.$$

Now we are ready to prove our main result for weighted Fourier frames with explicit bounds:

Proof of Theorem 2.2.1. The proof is set up in the same manner as the proof of Gröchenig's original result, Theorem 2.1.7. For a function $f \in \mathrm{H}$, define

$$h(\hat{y}) = \sum_{\omega \in \Omega} \hat{f}(\omega) \chi_{V_\omega^*}(\hat{y}), \quad \hat{y} \in \hat{\mathbb{R}}^d.$$

Since the sets V_ω^*, $\omega \in \Omega$, make a disjoint partition of $\hat{\mathbb{R}}^d$, it holds that

$$\|h\| = \sqrt{\sum_{\omega \in \Omega} \mu_\omega |\hat{f}(\omega)|^2},$$

where $\mu_\omega = \mathrm{meas}(V_\omega^*)$. Note that

$$\|f\| - \|\hat{f} - h\| \leq \|h\| \leq \|\hat{f} - h\| + \|f\|. \tag{2.2.6}$$

Hence, we aim to estimate $\|\hat{f} - h\|$. Again, by using properties of Voronoi regions, it is possible to conclude that

$$\|\hat{f} - h\| = \sqrt{\sum_{\omega\in\Omega} \int_{V_\omega^*} |\hat{f}(\hat{y}) - \hat{f}(\omega)|^2 \, \mathrm{d}\hat{y}}.$$

In order to estimate $|\hat{f}(\hat{y}) - \hat{f}(\omega)|^2$, for all $\omega \in \Omega$ and all $\hat{y} \in V_\omega^*$, Taylor's expansion of the entire function $\hat{f}$ is used. Therefore, by the Cauchy–Schwarz inequality we get

$$\begin{aligned} |\hat{f}(\hat{y}) - \hat{f}(\omega)|^2 &\le \left(\sum_{\alpha\neq 0} \frac{|(\hat{y}-\omega)^\alpha|}{\alpha!} |D^\alpha \hat{f}(\hat{y})| \right)^2 \\ &\le \sum_{\alpha\neq 0} \frac{c^{|\alpha|_1}(\hat{y}-\omega)^{2\alpha}}{\alpha!} \sum_{\alpha\neq 0} \frac{c^{-|\alpha|_1}}{\alpha!} |D^\alpha \hat{f}(\hat{y})|^2, \end{aligned} \tag{2.2.7}$$

for some constant $c > 0$ to be determined later. The inequality (2.2.7) is where this proof starts to differ from Gröchenig's original proof. For the first term in (2.2.7), by the multinomial formula (2.2.5) we get

$$\begin{aligned} \sum_{\alpha\neq 0} \frac{c^{|\alpha|_1}(\hat{y}-\omega)^{2\alpha}}{\alpha!} &= \sum_{k=0}^\infty \frac{c^k}{k!} \sum_{|\alpha|_1=k} \frac{k!}{\alpha!}(\hat{y}-\omega)^{2\alpha} - 1 \\ &= \sum_{k=0}^\infty \frac{c^k}{k!} |\hat{y}-\omega|^{2k} - 1 \\ &\le \exp(c(\delta_* c^*)^2) - 1, \end{aligned}$$

where in the final inequality δ_*-density of the set Ω is used:

$$\forall \omega \in \Omega, \quad \forall \hat{y} \in V_\omega^*, \quad |\hat{y} - \omega| \le \delta_* c^*.$$

Now consider the other term in (2.2.7). If we integrate over the Voronoi region V_ω^* and sum over $\omega \in \Omega$ then

$$\begin{aligned} \sum_{\alpha\neq 0} \frac{c^{-|\alpha|_1}}{\alpha!} \sum_{\omega\in\Omega} \int_{V_\omega^*} |D^\alpha \hat{f}(\hat{y})|^2 \, \mathrm{d}\hat{y} &= \sum_{k=1}^\infty \frac{c^{-k}}{k!} \sum_{|\alpha|_1=k} \frac{k!}{\alpha!} \|D^\alpha \hat{f}\|^2 \\ &= \sum_{k=1}^\infty \frac{c^{-k}}{k!} \int_D \sum_{|\alpha|_1=k} \frac{k!}{\alpha!} (2\pi x)^{2\alpha} |f(x)|^2 \, \mathrm{d}x, \end{aligned}$$

since by Parseval's identity

$$\|D^\alpha \hat{f}\|^2 = \|\hat{F}\|^2 = \|F\|^2 = \int_D (2\pi x)^{2\alpha} |f(x)|^2 \, \mathrm{d}x,$$

where $F(x) = (-\mathrm{i}2\pi x)^\alpha f(x)$. Hence, again by the multinomial formula (2.2.5), we obtain

$$\begin{aligned}\sum_{\alpha\neq 0} \frac{c^{-|\alpha|_1}}{\alpha!} \sum_{\omega\in\Omega} \int_{V_\omega^*} |D^\alpha \hat{f}(\hat{y})|^2 \,\mathrm{d}\hat{y} &= \sum_{k=1}^{\infty} \frac{c^{-k}(2\pi m_D)^{2k}}{k!} \|f\|^2 \\ &= \left(\exp((2\pi m_D)^2/c) - 1\right) \|f\|^2.\end{aligned}$$

Therefore, from (2.2.7), we get

$$\|\hat{f} - h\|^2 \leq \left(\exp(c(\delta_* c^*)^2) - 1\right)\left(\exp((2\pi m_D)^2/c) - 1\right) \|f\|^2.$$

If we equate the two terms, then we set $c = 2\pi m_D/(\delta_* c^*)$ to get

$$\|\hat{f} - h\| \leq (\exp(2\pi m_D \delta_* c^*) - 1) \|f\|.$$

Thus (2.2.6) now gives

$$\sqrt{B} \leq \exp(2\pi m_D \delta_* c^*), \quad \sqrt{A} \geq 2 - \exp(2\pi m_D \delta_* c^*),$$

with the condition that

$$\delta_* < \frac{\ln 2}{2\pi m_D c^*},$$

as required. □

2.3 Sharp sufficient density for weighted Fourier frames

The relative separation of a sampling set Ω is necessary and sufficient for the existence of an upper frame bound [You01, Thm. 2.17], see also [Jaf91, Lem. 1]. However, if we introduce appropriate weights $\{\mu_\omega\}_{\omega\in\Omega}$ to compensate for the clustering of the sampling points Ω, and consider $\{\sqrt{\mu_\omega} e_\omega\}_{\omega\in\Omega}$ instead of $\{e_\omega\}_{\omega\in\Omega}$, then this condition ceases to be necessary, as it is evident from Gröchenig's Theorem 2.1.7 and the improved result given in Theorem 2.2.1. On the other hand, once nontrivial weights $\mu_\omega > 0$ are introduced, existence of a lower frame bound is no longer guaranteed by Beurling's result. Nevertheless, by the following result, we demonstrate how the separation condition from Beurling's result can be successfully removed by using weights corresponding to the volumes of the Voronoi cells of the sampling points.

Theorem 2.3.1. *Let* $\mathrm{H} = \{f \in \mathrm{L}^2(\mathbb{R}^d) : \mathrm{supp}(f) \subseteq D\}$, *where* $D \subseteq \mathbb{R}^d$ *is compact, convex and symmetric. If a countable set* $\Omega \subseteq \hat{\mathbb{R}}^d$ *has density*

$$\delta_{D^\circ} < \frac{1}{4} \tag{2.3.1}$$

in $\hat{\mathbb{R}}^d$, *then* $\{\sqrt{\mu_\omega} e_\omega\}_{\omega\in\Omega}$ *is a weighted Fourier frame for* H *with the weights* $\{\mu_\omega\}_{\omega\in\Omega}$

defined as the measures of Voronoi regions with respect to the $|\cdot|_{D^\circ}$ *norm. In other words, there exist constants* $A, B > 0$ *such that*

$$\forall f \in \mathrm{H}, \quad A\|f\|^2 \leq \sum_{\omega\in\Omega} \mu_\omega |\hat{f}(\omega)|^2 \leq B\|f\|^2.$$

As Theorem 2.2.1, without imposing separation, Theorem 2.3.1 gives density condition sufficient to yield a weighted Fourier frame. Although this result does not lead to explicit frame bounds, it provides the universal density condition (2.3.1) which is both dimension independent and sharp. The latter follows from the sharpness of Beurling's Theorem 2.1.2 and by [OU12, Prop. 4.1].

In order to prove Theorem 2.3.1, we need the following lemma.

Lemma 2.3.2. *If* Ω *is a sequence with the density* $\delta_{D^\circ}(\Omega) < 1/4$ *in* $\hat{\mathbb{R}}^d$*, then there exists a subsequence* $\tilde{\Omega} \subseteq \Omega$ *which is* η*-separated with respect to the norm* $|\cdot|_{D^\circ}$ *for some* $\eta > 0$*, and also has density* $\delta_{D^\circ}(\tilde{\Omega}) < 1/4$ *in* $\hat{\mathbb{R}}^d$*.*

Proof. To begin with, we introduce some notation. For the set D, we define $D(0,1) = D$, $D(0,r) = rD$ and $D(x,r) = x + rD$. Here, for δ_{D°, we simply write δ.

Let us choose $\eta > 0$ such that $\delta + \eta/2 < 1/4$ and set $\delta_1 = \delta + \eta$. Now define $\tilde{\Omega}$ inductively as follows. For an arbitrarily picked point $\omega_0 \in \Omega$, set $\tilde{\omega}_0 = \omega_0$. Given $\tilde{\omega}_0, \dots, \tilde{\omega}_N$, define $\tilde{\omega}_{N+1}$ by

$$\tilde{\omega}_{N+1} \in \Omega \cap D^\circ(x, \delta),$$

where

$$x \in \partial G = \partial \left(\bigcup_{\tilde{\omega}_n \in \tilde{\Omega}_N} D^\circ(\tilde{\omega}_n, \delta_1) \right) \quad \text{and} \quad \tilde{\Omega}_N = \{\tilde{\omega}_n\}_{n=0}^N.$$

Here, we picked any $x \in \partial G$ and then, for that x, any $\tilde{\omega}_{N+1} \in \Omega \cap D^\circ(x, \delta)$. Finally, we let $\tilde{\Omega} = \{\tilde{\omega}_n\}_{n=0}^\infty$.

Note that for any $x \in \hat{\mathbb{R}}^d$ there must exists a point $\omega \in \Omega$ in the set $D^\circ(x, \delta)$ such that x is covered by $D^\circ(\omega, \delta)$, since Ω is δ-dense in the norm $|\cdot|_{D^\circ}$ and $\hat{\mathbb{R}}^d$ can be covered by the sets $D^\circ(\omega, \delta)$, $\omega \in \Omega$. Moreover, for every $x \in \partial G$, every $\omega \in \Omega \cap D^\circ(x, \delta)$ must be different than any other $\omega \in \tilde{\Omega}_N$, since $\delta < \delta_1$. Also, note that for every such $\omega \in \Omega \cap D^\circ(x, \delta)$ it holds that

$$\eta = \delta_1 - \delta \leq \inf_{\tilde{\omega}_n \in \tilde{\Omega}_N} |\omega - \tilde{\omega}_n|_{D^\circ} \leq \delta_1 + \delta = 2\delta + \eta.$$

Therefore if we choose $\tilde{\omega}_{N+1}$ from $\Omega \cap D^\circ(x, \delta)$ arbitrarily, and continue the procedure until $G = \hat{\mathbb{R}}^d$, by the construction, $\tilde{\Omega}$ is $\tilde{\delta}$-dense in the norm $|\cdot|_{D^\circ}$ where $\tilde{\delta} = (2\delta + \eta)/2 < 1/4$. Also, it is η-separated in the norm $|\cdot|_{D^\circ}$. $\square$

Remark 2.3.3 In view of this lemma, it might be tempting to infer the following

$$\sum_{\omega\in\Omega} \mu_\omega |\hat{f}(\omega)|^2 \geq \sum_{\tilde{\omega}\in\tilde{\Omega}} \mu_{\tilde{\omega}} |\hat{f}(\tilde{\omega})|^2 \geq \text{meas}\left(\frac{\eta}{2}D^\circ\right) \sum_{\tilde{\omega}\in\tilde{\Omega}} |\hat{f}(\tilde{\omega})|^2 \geq \text{meas}\left(\frac{\eta}{2}D^\circ\right) A, \qquad (2.3.2)$$

and therefore seemingly obtain the lower frame bound for the weighted non-separated sequence Ω. However, note that the second inequality in (2.3.2) need not hold, since the weights at the very beginning are chosen as the Lebesgue measures of the Voronoi regions corresponding to Ω, which can be arbitrarily small due to clustering. Therefore, although the sequence $\tilde{\Omega}$ is separated, there might indeed exists $\tilde{\omega} \in \tilde{\Omega}$ such that its Voronoi region $V_{\tilde{\omega}}^{D^\circ}$ does not contain a ball of radius $\eta/2$ with respect to the D°-norm.

Proof of Theorem 2.3.1. First of all, for the upper bound we use Theorem 2.2.1. From the proof of Theorem 2.2.1, we can infer that the density condition (2.2.1) is imposed only to ensure $A > 0$, and that the estimate of the upper frame bound holds even if this density condition is not satisfied. Indeed, for any compact set $D \subseteq \mathbb{R}^d$, any norm $|\cdot|_*$ and any positive density $\delta_* < \infty$, the upper frame bound satisfies

$$B \leq \exp\left(4\pi m_D \delta_* c^*\right) < \infty.$$

In particular, if $\delta_{D^\circ} < 1/4$, then

$$B \leq \exp\left(\pi m_D c^\circ\right) < \infty,$$

where $c^\circ \in (0, \infty)$ is the smallest constant such that $|\cdot| \leq c^\circ \, |\cdot|_{D^\circ}$.

For the lower bound, we note that if Ω is separated, then everything follows easily. Namely, since Ω is η-separated with respect to the D°-norm, we get

$$\sum_{\omega\in\Omega} \mu_\omega |\hat{f}(\omega)|^2 \geq \text{meas}\left(\frac{\eta}{2}D^\circ\right) \sum_{\omega\in\Omega} |\hat{f}(\omega)|^2 \geq \text{meas}\left(\frac{\eta}{2}D^\circ\right) A' \|f\|^2,$$

where $A' > 0$ comes from application of Theorem 2.1.2. Thus we take $A = \text{meas}\left(\frac{\eta}{2}D^\circ\right) A'$.

However, if Ω is not separated, we proceed as follows. By Lemma 2.3.2, we know that there exists a subsequence $\tilde{\Omega} \subseteq \Omega$ with density $\delta_{D^\circ}(\tilde{\Omega}) < 1/4$ and separation $\eta = \eta_{D^\circ}(\tilde{\Omega}) > 0$. Let $\epsilon < \eta/2$. Then

$$\sum_{\omega\in\Omega} \mu_\omega |\hat{f}(\omega)|^2 \geq \sum_{\tilde{\omega}\in\tilde{\Omega}} \sum_{\omega\in D_\epsilon^\circ(\tilde{\omega})\cap\Omega} \mu_\omega |\hat{f}(\omega)|^2,$$

where $D_\epsilon^\circ(\tilde{\omega})$ denotes the ball with respect to the D°-norm of radius ϵ centred at $\tilde{\omega}$. Since $\hat{f}$ is continuous function, from the Extreme value theorem, for each $\tilde{\omega}$, we know there is a point $z_{\tilde{\omega}} \in \overline{D^\circ_\epsilon(\tilde{\omega})} = D^\circ_\epsilon(\tilde{\omega})$, such that

$$\forall \omega \in D^\circ_\epsilon(\tilde{\omega}), \quad |\hat{f}(\omega)| \geq |\hat{f}(z_{\tilde{\omega}})|.$$

Since also $\mu_\omega = \text{meas}\left(V_\omega^{D^\circ}\right)$ and the sets $V_\omega^{D^\circ}$ are disjoint, we get

$$\sum_{\omega\in\Omega} \mu_\omega |\hat{f}(\omega)|^2 \geq \sum_{\tilde{\omega}\in\tilde{\Omega}} \left(|\hat{f}(z_{\tilde{\omega}})|^2 \sum_{\omega \in D^\circ_\epsilon(\tilde{\omega})\cap\Omega} \mu_\omega \right) = \sum_{\tilde{\omega}\in\tilde{\Omega}} \left(|\hat{f}(z_{\tilde{\omega}})|^2 \text{meas} \left(\bigcup_{\omega \in D^\circ_\epsilon(\tilde{\omega})\cap\Omega} V_\omega^{D^\circ} \right) \right).$$

Now we claim the following:

$$\bigcup_{\omega \in D^\circ_\epsilon(\tilde{\omega})\cap\Omega} V_\omega^{D^\circ} \supseteq D^\circ_\rho(\tilde{\omega}), \quad \rho = \frac{\epsilon}{2}.$$

To see this, let $|\hat{y} - \tilde{\omega}|_{D^\circ} \leq \frac{\epsilon}{2}$. Since $\hat{y} \in V_\omega^{D^\circ}$ for some $\omega \in \Omega$, we have $|\hat{y} - \omega|_{D^\circ} \leq |\hat{y} - \tilde{\omega}|_{D^\circ}$. Therefore

$$|\hat{y} - \omega|_{D^\circ} \leq |\hat{y} - \tilde{\omega}|_{D^\circ} \leq \frac{\epsilon}{2},$$

and hence

$$|\omega - \tilde{\omega}|_{D^\circ} \leq |\hat{y} - \omega|_{D^\circ} + |\hat{y} - \tilde{\omega}|_{D^\circ} \leq \epsilon.$$

Thus $\omega \in D^\circ_\epsilon(\tilde{\omega}) \cap \Omega$ as required. Therefore, we get

$$\sum_{\omega\in\Omega} \mu_\omega |\hat{f}(\omega)|^2 \geq \text{meas}\left(\frac{\epsilon}{2} D^\circ\right) \sum_{\bar{\omega}\in\bar{\Omega}} |\hat{f}(\bar{\omega})|^2,$$

where $\bar{\Omega} = \{z_{\tilde{\omega}} \ : \ \tilde{\omega} \in \tilde{\Omega}\}$. To complete the proof, we only need to show that the set $\bar{\Omega}$ is separated and sufficiently dense, so that we can apply the Theorem 2.1.2. Consider $\bar{\omega}_1$ and $\bar{\omega}_2$. Then we clearly have

$$|\bar{\omega}_1 - \bar{\omega}_2|_{D^\circ} \geq \eta - 2\epsilon > 0,$$

since $\tilde{\Omega}$ is separated with the separation η and the $\bar{\omega}$'s lie in the ϵ-cover of this set. Moreover, it is straightforward to see that

$$\delta_{D^\circ}(\bar{\Omega}) \leq \delta_{D^\circ}(\tilde{\Omega}) + \epsilon.$$

Thus, since $\delta_{D^\circ}(\tilde{\Omega}) < 1/4$, we have the same for $\bar{\Omega}$ for sufficiently small $\epsilon > 0$. We set $A = \text{meas}\left(\frac{\epsilon}{2} D^\circ\right) A'$, where $A' > 0$ is as in Theorem 2.1.2, and finish the proof. □

Remark 2.3.4 From the proof of Theorem 2.3.1 and the proof of Lemma 2.3.2, we can conclude the following. If Ω has density $\delta_{D^\circ}(\Omega) < 1/4$, it yields a weighted Fourier frame with the lower weighted Fourier frame bound of the form

$$A = \text{meas}\left(\frac{\epsilon}{2} D^\circ\right) A',$$

where A' is the lower Fourier frame bound for sequence $\bar{\Omega} \subseteq \hat{\mathbb{R}}^d$ with separation $\eta_{D^\circ}(\bar{\Omega}) = \eta - 2\epsilon$ and density $\delta_{D^\circ}(\bar{\Omega}) \leq \delta_{D^\circ}(\Omega) + \eta/2 + \epsilon$, for some constants $\eta, \epsilon > 0$ chosen small enough so that existence of A' is ensured. However, this does not in general lead to an explicit estimate of A since we typically do not know an explicit estimate of A'. On the other hand, the upper weighted Fourier frame bound B is explicitly estimated by

$$B \leq \exp\left(\pi m_D c^\circ\right),$$

where $c^\circ \in (0, \infty)$ is the smallest constant such that $|\cdot| \leq c^\circ \, |\cdot|_{D^\circ}$.

Remark 2.3.5 Note that the density condition form Theorem 2.2.1 does not contradict the sharpness of the density condition from Theorem 2.3.1, i.e. note that

$$\frac{\ln 2}{2\pi m_D c^\circ} \leq \frac{1}{4},$$

where c° is the smallest constant such that $|\cdot| \leq c^\circ \, |\cdot|_{D^\circ}$ and D is a compact, convex and symmetric set. To see this, we now argue that $m_D c^\circ \geq 1$. Note that from the definition of a radial set, it follows that for all $y \in \mathbb{R}^d$ we have

$$|y|_{D^\circ} = \max_{x \in D} |x \cdot y|,$$

see for example [BW00]. Therefore $|\cdot|_{D^\circ} \leq m_D \, |\cdot|$, which implies $1/m_E \leq c_\circ$, where $c_\circ$ is the largest constant such that $c_\circ \, |\cdot|_{D^\circ} \leq |\cdot|$. Hence

$$m_D c^\circ \geq \frac{c^\circ}{c_\circ},$$

and since $c_\circ \leq c^\circ$, the claim follows.

To end this chapter, in order to illustrate differences between classical and weighted Fourier frames, as well as different uses of previously given results, let us consider the following two-dimensional example.

Example 2.3.6 Let $D = \mathcal{B}_1 \subseteq \mathbb{R}^2$ and let

$$\Lambda_1 = \tfrac{1}{8}\mathbb{Z}^2, \qquad \Lambda_2 = \left\{ \left(\frac{1}{n}, \frac{1}{m}\right) : (n, m) \in \mathbb{Z}^2, \min\{|n|, |m|\} > 8 \right\}.$$

Note that, for such D, $D^\circ = \mathcal{B}_1$ and therefore the D°-norm is the Euclidean norm $|\cdot|$.

The set of points Λ_1 is separated with the density

$$\delta_{\mathcal{B}_1}(\Lambda_1) = \frac{\sqrt{2}}{16} \approx 0.0884 < \frac{1}{4}.$$

Therefore, by Theorem 2.1.2, we conclude the family of functions $\{e_\lambda\}_{\lambda\in\Lambda_1}$ is a frame for $\mathrm{L}^2(\mathcal{B}_1)$. However, if we now consider the set

$$\Omega = \Lambda_1 \cup \Lambda_2,$$

for which $\delta_{\mathcal{B}_1}(\Omega) = \delta_{\mathcal{B}_1}(\Lambda_1) = \sqrt{2}/16$, Theorem 2.1.2 can not be used since Ω has infinitely many accumulation points at

$$\{0\} \cup \left\{ \left(\frac{1}{n}, 0\right) : n \in \mathbb{Z}, |n| > 8 \right\} \cup \left\{ \left(0, \frac{1}{m}\right) : m \in \mathbb{Z}, |m| > 8 \right\},$$

and therefore it is not separated. Moreover, it can be verified that the family $\{e_\omega\}_{\omega\in\Omega}$ fails in satisfying the right inequality of (2.1.2). To see this, we first note that

$$\int_{\mathcal{B}_1} \mathrm{e}^{-2\pi \mathrm{i}\omega\cdot x}\,\mathrm{d}x = \frac{J_1(2\pi|\omega|)}{|\omega|},$$

where J_1 is the Bessel function of the first kind and order 1. Therefore, there exists $c > 0$ such that

$$c \le \left| \int_{\mathcal{B}_1} \mathrm{e}^{-2\pi \mathrm{i}\left(\frac{1}{n}x_1 + \frac{1}{m}x_2\right)}\,\mathrm{d}x_1\,\mathrm{d}x_2 \right|^2 \le \pi^2, \tag{2.3.3}$$

for all $(n,m) \in \mathbb{Z}^2$ such that $\sqrt{1/n^2 + 1/m^2} < a j'_{1,1}/(2\pi) \approx 0.6098$, where a is some fixed constant from the interval $(0,1)$ and $j'_{1,1}$ is the first positive zero of the function J_1. Hence, it is enough to take the function $g(x) = \chi_{\mathcal{B}_1}(x)$ for which $\|g\|^2 = \pi$, whereas $\sum_{\omega\in\Omega} |\hat{g}(\omega)|^2$ is unbounded. Thus, we conclude that the set Ω does not give a Fourier frame.

On the other hand, if, for the same set of points $\Omega = \Lambda_1 \cup \Lambda_2$, we consider the weighted family $\{\sqrt{\mu_\omega} e_\omega\}_{\omega\in\Omega}$ with the weights defined as Voronoi regions in the ℓ^2-norm, this particular function g satisfies the weighted Fourier frame inequalities (2.1.5) with some $0 < A, B < \infty$. This can be easily proved by using the inequalities (2.3.3), and the fact that

$$\sum_{n=9}^{\infty} \sum_{m=9}^{\infty} \left(\frac{1}{n-1} - \frac{1}{n+1}\right)\left(\frac{1}{m-1} - \frac{1}{m+1}\right) = \left(\frac{17}{72}\right)^2.$$

which implies that the sum of the Voronoi regions corresponding to the points Λ_2 converges. Moreover, since $\delta_{\mathcal{B}_1}(\Omega) = \sqrt{2}/16$, by Theorem 2.3.1 we conclude that Ω gives rise to a weighted Fourier frame.

Note also, in order to verify that Ω forms a weighted Fourier frame, Gröchenig's original result could not be used since

$$\delta_{\mathcal{B}_1}(\Omega) = \frac{\sqrt{2}}{16} > \frac{\ln 2}{4\pi} \approx 0.0552.$$

However, since in this case $m_D = 1$ and $c^* = 1$ and since

$$\delta_{\mathcal{B}_1}(\Omega) = \frac{\sqrt{2}}{16} < \frac{\ln 2}{2\pi} \approx 0.1103,$$

we are able to use Theorem 2.2.1 to conclude that Ω generates a weighted Fourier frame with the weighted Fourier frame bounds $\sqrt{A} \geq 0.2574$ and $\sqrt{B} \leq 1.7426$.

Chapter 3

Generalized sampling for nonuniform Fourier samples

Having seen the conditions that ensure weighted Fourier frames, we are now interested in constructing a good—accurate and stable—approximation to an unknown function from nonuniform Fourier data. With this aim, in the present chapter, we introduce so-called Nonuniform Generalized Sampling (NUGS) that stably approximates a function in a desired reconstruction space from a finite collection of nonuniform samples. The results of this chapter are mainly published in [AGH14a], which is the joint work of the author with Ben Adcock and Anders Hansen.

Let $\Omega \subseteq \hat{\mathbb{R}}^d$ be a countable set of distinct (nonuniform) frequencies, henceforth referred to as a *sampling scheme*, and let $\mathrm{T} \subseteq \mathrm{H}$ be a finite-dimensional subspace of $\mathrm{H} = \{f \in \mathrm{L}^2(\mathbb{R}^d) : \mathrm{supp}(f) \subseteq D\}$ for a compact domain $D \subseteq \hat{\mathbb{R}}^d$, the so-called *reconstruction space*. We address the following reconstruction problem:

The reconstruction problem

Given a sampling scheme Ω and a reconstruction space $\mathrm{T} \subseteq \mathrm{H}$, compute an approximation $\tilde{f} \in \mathrm{T}$ to an unknown function $f \in \mathrm{H}$ via a mapping $F = F_{\Omega,\mathrm{T}} : f \mapsto \tilde{f}$, which depends only on the sampling data

$$\left\{ \hat{f}(\omega) : \omega \in \Omega \right\}$$

and satisfies the following critical properties:

(i) F is *quasi-optimal*: there exists a constant $\mu = \mu(F) \ll \infty$ such that

$$\forall f \in \mathrm{H}, \quad \|f - F(f)\| \leq \mu \|f - \mathcal{P}_{\mathrm{T}} f\|, \tag{$\star$}$$

where $\mathcal{P}_{\mathrm{T}}$ denotes the orthogonal projection onto T.

(ii) F is *well-conditioned*, i.e. *stable*: there exists a constant $\kappa = \kappa(F) \ll \infty$ such that
$$\kappa = \sup_{f\in \mathrm{H}} \lim_{\epsilon\to 0} \sup_{\substack{g\in \mathrm{H},\\ 0<\|\hat{g}|_\Omega\|_\times \le \epsilon}} \frac{\|F(f+g)-F(f)\|}{\|\hat{g}|_\Omega\|_\times}. \tag{$\star\star$}$$
where $\|\hat{g}|_\Omega\|_\times$ is a norm of the sampling data $\{\hat{g}(\omega) : \omega \in \Omega\}$.

In what follows, we will be interested in a linear mapping F. Note that if F is linear, then ($\star\star$) becomes
$$\kappa = \sup_{\substack{f\in \mathrm{H},\\ f\neq 0}} \frac{\|F(f)\|}{\|\hat{f}|_\Omega\|_\times}.$$
Also, in what follows, instead of ($\star$), we shall provide a stronger inequality:
$$\forall f, h \in \mathrm{H}, \quad \|f - F(f+h)\| \le \mu \left(\|f - \mathcal{P}_\mathrm{T} f\| + \|h\|\right).$$

Quasi-optimality of mapping F required by ($\star$) guarantees that the reconstruction $\tilde{f}$ inherits good approximation properties of the space T. Recall that the motivation for considering a particular reconstruction space T is that f is known to be well-represented in this space, i.e. it is known that the error $\|f - \mathcal{P}_\mathrm{T} f\|$ is small. On the other hand, well-conditioning of mapping F imposed by ($\star\star$) is vital to ensure that noisy data do not adversely affect the reconstruction. In particular, a well-conditioned mapping F is robust towards small perturbations in the input measurements $\left\{\hat{f}(\omega) : \omega \in \Omega\right\}$. We remark that the condition number defined by ($\star\star$) is typically referred to as the *absolute* condition number.

With this to hand, the main focus of this chapter is to answer the following questions:

(i) under what conditions on Ω and T, stable and quasi-optimal reconstruction via $F_{\Omega,\mathrm{T}}$ is possible, and

(ii) how large are the constants $\mu(F_{\Omega,\mathrm{T}})$ and $\kappa(F_{\Omega,\mathrm{T}})$.

We do this by analysing NUGS, which we introduce in §3.1. This provides a sufficient condition for (i) and an upper bound for (ii). We further refine the answers to these questions in §3.2 and §3.3, for the univariate and multivariate case respectively.

3.1 The nonuniform generalized sampling framework

In order to consider a sampling scheme of a general type, first we guarantee stability and accuracy of the GS reconstruction, as defined in [AH12a] (see (3.1.5)), for any so-called admissible sampling operator. By choosing a particular form of the admissible sampling operator, we then define the NUGS reconstruction.

3.1.1 Generalized sampling with admissible sampling operator

We commence with the definition of an admissible sampling operator.

Definition 3.1.1. *Let Ω be a sampling scheme, $\mathcal{S} : \mathrm{H} \to \mathrm{H}$ a bounded linear operator and let T a finite-dimensional subspace of a Hilbert space H. Suppose that $\mathcal{S}$ satisfies*

I for each $f \in \mathrm{H}$, $\mathcal{S}f$ depends only on the sampling data $\{\hat{f}(\omega) : \omega \in \Omega\}$,

II $\mathcal{S}$ is self-adjoint with respect to $\langle \cdot, \cdot \rangle$ and satisfies

$$\forall f, g \in \mathrm{H}, \quad |\langle \mathcal{S}f, g \rangle| \leq \sqrt{\langle \mathcal{S}f, f \rangle \langle \mathcal{S}g, g \rangle}, \tag{3.1.1}$$

III there exists a positive constant $C_1 = C_1(\Omega, \mathrm{T})$ such that

$$\forall f \in \mathrm{T} \setminus \{0\}, \quad \langle \mathcal{S}f, f \rangle \geq C_1 \|f\|^2. \tag{3.1.2}$$

Then $\mathcal{S}$ is said to be an admissible sampling operator for the pair (Ω, T).

Remark 3.1.2 As we show later in Theorem 3.2.5 in the univariate case, and in Theorems 3.3.2 and 3.3.3 in the multivariate case, this abstract definition is satisfied if $\mathcal{S}$ is defined as in (3.1.9), under appropriate conditions on the sampling density and the sampling bandwidth.

For convenience, throughout the remainder of the paper we shall assume that C_1 is the largest constant for which (3.1.2) holds. Given such an operator $\mathcal{S}$, we now also define the constants $C_2 = C_2(\Omega)$ and $C_3 = C_3(\Omega, \mathrm{T})$ by

$$\forall f \in \mathrm{H} \setminus \{0\}, \quad \langle \mathcal{S}f, f \rangle \leq C_2 \|f\|^2, \tag{3.1.3}$$

$$\forall f \in \mathrm{T} \setminus \{0\}, \quad \langle \mathcal{S}f, f \rangle \leq C_3 \|f\|^2. \tag{3.1.4}$$

Likewise, we assume these constants are the smallest possible. Note that C_2 and C_3 exist since $\mathcal{S}$ is bounded, and we also trivially have that $C_3 \leq C_2$.

Given a sampling scheme Ω, a finite-dimensional subspace T and an admissible sampling operator $\mathcal{S}$ for the pair (Ω, T), we now define the GS reconstruction $\tilde{f} \in \mathrm{T}$ such that

$$\forall g \in \mathrm{T}, \quad \langle \mathcal{S}\tilde{f}, g \rangle = \langle \mathcal{S}f, g \rangle, \tag{3.1.5}$$

and write $F = F_{\Omega,\mathrm{T}}$ for the mapping $f \mapsto \tilde{f}$. As we shall see next, the constants C_1 and C_2 arising from (3.1.2) and (3.1.3) determine the stability and quasi-optimality of the resulting reconstruction. We define the corresponding reconstruction constant $C(\Omega, \mathrm{T})$ as the ratio

$$C(\Omega, \mathrm{T}) = \sqrt{\frac{C_2}{C_1}}. \tag{3.1.6}$$

Now we present a result that is proved by exactly the same techniques as in [AHP13]. We include a simplified proof for completeness.

Theorem 3.1.3. *Let Ω be a sampling scheme and $\mathrm{T} \subseteq \mathrm{H}$ a finite-dimensional subspace, and suppose that $\mathcal{S}$ is an admissible sampling operator for the pair (Ω, T). Then the reconstruction $F(f) = \tilde{f}$ defined by (3.1.5) exists uniquely for any $f \in \mathrm{H}$ and we have the sharp bound*

$$\forall f, h \in \mathrm{H}, \quad \|f - F(f+h)\| \leq \tilde{C}\left(\|f - \mathcal{P}_{\mathrm{T}} f\| + \|h\|\right), \tag{3.1.7}$$

where the constant $\tilde{C}$ is given by

$$\tilde{C} = \tilde{C}(\Omega, \mathrm{T}) = \sup_{\substack{g \in \mathrm{T} \\ g \neq 0}} \frac{\|g\|}{\|\mathcal{P}_{\mathcal{S}(\mathrm{T})} g\|}$$

and it satisfies $\tilde{C} \leq C$, where $C = C(\Omega, \mathrm{T})$ is the corresponding reconstruction constant given by (3.1.6). Moreover, if

$$\kappa = \sup_{\substack{f \in \mathrm{H} \\ f \neq 0}} \frac{\|F(f)\|}{\sqrt{\langle \mathcal{S} f, f \rangle}}, \tag{3.1.8}$$

then $\kappa = 1/\sqrt{C_1}$ where $C_1 = C_1(\Omega, \mathrm{T})$ is as in (3.1.2). In particular, $\kappa \leq \max\left\{1/\sqrt{C_1}, C\right\}$.

Proof. Let us start from the end. To prove (3.1.8), first note that

$$\kappa(F) \geq \sup_{\substack{g \in \mathrm{T} \\ g \neq 0}} \frac{\|g\|}{\sqrt{\langle \mathcal{S} g, g \rangle}} = \frac{1}{\sqrt{C_1}},$$

where the equality follows from (3.1.2). For the upper bound, first note that $\langle \mathcal{S} \tilde{f}, \tilde{f} \rangle = \langle \mathcal{S} f, \tilde{f} \rangle \leq \sqrt{\langle \mathcal{S} f, f \rangle \langle \mathcal{S} \tilde{f}, \tilde{f} \rangle}$, by (3.1.5) and (3.1.1). Hence, since $\sqrt{\langle \mathcal{S} \tilde{f}, \tilde{f} \rangle} \leq \sqrt{\langle \mathcal{S} f, f \rangle}$ and since $F : f \mapsto \tilde{f}$ is a surjection, we have

$$\kappa(F) \leq \sup_{\substack{f \in \mathrm{H} \\ f \neq 0}} \frac{\|\tilde{f}\|}{\sqrt{\langle \mathcal{S} \tilde{f}, \tilde{f} \rangle}} = \sup_{\substack{g \in \mathrm{T} \\ g \neq 0}} \frac{\|g\|}{\sqrt{\langle \mathcal{S} g, g \rangle}} = \frac{1}{\sqrt{C_1}}.$$

Next we show that $\tilde{C} \leq C$, and in particular, that $\tilde{C} < \infty$. By definition

$$\frac{1}{\tilde{C}} = \inf_{\substack{g \in \mathrm{T} \\ g \neq 0}} \frac{\|\mathcal{P}_{\mathcal{S}(\mathrm{T})} g\|}{\|g\|} = \inf_{\substack{g \in \mathrm{T} \\ g \neq 0}} \sup_{\substack{g' \in \mathrm{T} \\ \mathcal{S} g' \neq 0}} \frac{|\langle g, \mathcal{S} g' \rangle|}{\|g\| \|\mathcal{S} g'\|}.$$

Let $g \in \mathrm{T} \setminus \{0\}$. If $\mathcal{S} g = 0$, then $\langle \mathcal{S} g, g \rangle = 0$ which contradicts the admissibility of $\mathcal{S}$. Hence $\mathcal{S} g \neq 0$. Therefore, we may set $g' = g$ above to get

$$\frac{1}{\tilde{C}} \geq \inf_{\substack{g \in \mathrm{T} \\ g \neq 0}} \frac{\langle \mathcal{S} g, g \rangle}{\|g\| \|\mathcal{S} g\|}.$$

Observe that

$$\|\mathcal{S}g\| = \sup_{\substack{h\in \mathrm{H} \\ \|h\|=1}} \langle \mathcal{S}g, h\rangle \leq \sqrt{C_2}\sqrt{\langle \mathcal{S}g, g\rangle},$$

where the inequality follows from (3.1.1) and (3.1.3). This now gives

$$\frac{1}{\tilde{C}} \geq \frac{1}{\sqrt{C_2}} \inf_{\substack{g\in \mathrm{T} \\ g\neq 0}} \frac{\sqrt{\langle \mathcal{S}g, g\rangle}}{\|g\|},$$

which, upon application of (3.1.2), yields $\tilde{C} \leq \sqrt{C_2/C_1} = C$ as required.

To prove the remainder of the theorem, we shall use the techniques of [AHP13] based on the geometric notions of subspace angles and oblique projections. Let $\mathrm{U} = \mathrm{T}$ and $\mathrm{V} = (\mathcal{S}(\mathrm{T}))^{\perp}$. Note that $1/\tilde{C} = \cos(\theta_{\mathrm{UV}^{\perp}})$ is cosine of the subspace angle between U and $\mathrm{V}^{\perp}$ defined by

$$\cos(\theta_{\mathrm{UV}^{\perp}}) = \inf_{\substack{u\in \mathrm{U} \\ \|u\|=1}} \|\mathcal{P}_{V^{\perp}} u\|.$$

Since $\tilde{C} < \infty$, the subspaces U and V satisfy the so-called subspace condition $\cos(\theta_{\mathrm{UV}^{\perp}}) > 0$. Thus [AHP13, Cor. 3.5] gives

$$\|\mathcal{W}_{\mathrm{UV}} f\| \leq \tilde{C}\|f\|, \quad \forall f \in \mathrm{H}_0,$$

and

$$\|f - \mathcal{W}_{\mathrm{UV}} f\| \leq \tilde{C}\|f - \mathcal{P}_{\mathrm{U}} f\|, \quad \forall f \in \mathrm{H}_0,$$

where $\mathrm{H}_0 = \mathrm{U} \oplus \mathrm{V}$ and $\mathcal{W}_{\mathrm{UV}} : \mathrm{H}_0 \to \mathrm{U}$ is the projection with range U and kernel V.

Hence, to establish (3.1.7) it remains to show the following: (i) $\mathrm{H}_0 = \mathrm{H}$ and (ii) $\tilde{f} = \mathcal{W}_{\mathrm{UV}} f, \forall f \in \mathrm{H}$. For (i), we note that $\mathrm{H}_0 = \mathrm{H}$ provided $\dim(\mathcal{S}(\mathrm{T})) = \dim(\mathrm{T})$ [AHP13, Lem. 3.10]. However, if not then there exists a nonzero $g \in \mathrm{T}$ such that $\mathcal{S}(g) = 0$. As previously observed, this implies that $g = 0$; a contradiction.

For (ii), we first note that

$$\langle \mathcal{W}_{\mathrm{UV}} f, \mathcal{S}g\rangle = \langle f, \mathcal{S}g\rangle, \quad \forall g \in \mathrm{T}.$$

Since $\mathcal{S}$ is self-adjoint, it follows that $\mathcal{W}_{\mathrm{UV}} f$ satisfies the same conditions (3.1.5) as $\tilde{f}$. Thus, it remains only to show that $\tilde{f}$ is unique. However, if not then we find that there is a nonzero $g \in \mathrm{T} \cap \mathcal{S}(\mathrm{T})^{\perp} = \mathrm{U} \cap \mathrm{V}$. But then $\cos(\theta_{\mathrm{UV}^{\perp}}) = 0$, and this contradicts the fact that U and V obey the subspace condition. □

This result confirms that admissibility of $\mathcal{S}$ is sufficient for quasi-optimality and stability of the reconstruction $\tilde{f}$ up to the magnitude of the reconstruction constant C. Although, this result is true under the slightly weaker assumption $\tilde{C} < \infty$ (which is of course implied by $C_1 > 0$ and $C_2 < \infty$), the constant $\tilde{C}$ is rather difficult to work with in practice [AHP13].

The NUGS reconstruction

So far it was not specified whether Ω has finite or infinite cardinality, but in practice we are always faced with a finite sampling set

$$\Omega_N = \{\omega_n : n = 1, \dots, N\}, \quad N \in \mathbb{N}.$$

For such a nonuniform sampling scheme, there are many potential ways to construct the operator $\mathcal{S}$. Here, we focus on the following simple construction

$$\mathcal{S}f(x) = \sum_{n=1}^{N} \mu_n \hat{f}(\omega_n) \mathrm{e}^{2\pi \mathrm{i} \omega_n \cdot x} \chi_D(x), \tag{3.1.9}$$

where $\mu_n > 0$ are particular weights specified later. For such $\mathcal{S}$, the GS reconstruction defined by (3.1.5) becomes equivalent to the weighted least-squares data fit

$$\tilde{f} = \operatorname*{argmin}_{g \in \mathrm{T}} \sum_{n=1}^{N} \mu_n \left| \hat{f}(\omega_n) - \hat{g}(\omega_n) \right|^2. \tag{3.1.10}$$

We shall refer to such $\tilde{f}$ as *nonuniform generalized sampling (NUGS)* reconstruction.

The operator $\mathcal{S}$ defined in this way automatically satisfies properties *I* and *II* in Definition 3.1.1 of admissible sampling operator. In what follows, by conveniently using results on weighted Fourier frames, we prove that $\mathcal{S}$, under suitable conditions, also satisfies property *III*, and thus, by Theorem 3.1.3, ensures a stable and quasi-optimal NUGS reconstruction, with a bounded reconstruction constant $C(\Omega_N, \mathrm{T})$ defined by (3.1.6).

Observe that for $\mathcal{S}$ defined as in (3.1.9), the condition number defined by (3.1.8) of the mapping $F_{\Omega_N,\mathrm{T}}(f) = F(f) = \tilde{f}$ given by (3.1.10) becomes

$$\kappa(F) = \sup_{\substack{f \in \mathrm{H} \\ f \neq 0}} \frac{\|F(f)\|}{\|\hat{f}\|_{\ell^2_\mu(\Omega_N)}}, \qquad \|\hat{f}\|^2_{\ell^2_\mu(\Omega_N)} = \sum_{n=1}^{N} \mu_n |\hat{f}(\omega_n)|^2. \tag{3.1.11}$$

Hence, by Theorem 3.1.3, we have

$$\kappa(F) = C_1(\Omega_N, \mathrm{T})^{-\frac{1}{2}}$$

where $C_1(\Omega_N, \mathrm{T})$ is as in (3.1.2) for $\mathcal{S}$ given by (3.1.9).

As shown in [AHP13] this constant $C_1(\Omega_N, \mathrm{T})$ is essentially an universal quantity. Namely, if $G = G_{\Omega_N,\mathrm{T}} : \mathrm{H} \to \mathrm{T}$ is any so-called *perfect* reconstruction method [AHP13, Def. 3.9], i.e. for each $f \in \mathrm{H}$, $G(f)$ depends only on the samples of $\hat{f}$ at Ω_N, and $G(f) = f$

whenever $f \in \mathrm{T}$, then [AHP13, Thm. 6.2] gives

$$\kappa(G) \geq C_1(\Omega_N, \mathrm{T})^{-\frac{1}{2}},$$

whenever $C_1(\Omega_N, \mathrm{T})^{-1/2} \neq 0$, where the condition number $\kappa(G)$ is defined by ($\star\star$) with $\|\cdot\|_\times = \|\cdot\|_{\ell^2_\mu(\Omega_N)}$. Furthermore, this generalizes to $\kappa(G) \geq (1-\lambda) C_1(\Omega_N, \mathrm{T})^{-1/2}$ for a larger class of methods, namely for any so-called *contractive* method which obeys $\|f - G(f)\| \leq \lambda \|f\|$, $f \in \mathrm{T}$, for some constant $\lambda \in [0,1)$. This universality of $C_1(\Omega_N, \mathrm{T})$ is the key property for the universality of the sampling rate shown for wavelets later in Chapter 4.

Now, let us argue why $\|\cdot\|_\times = \|\cdot\|_{\ell^2_\mu(\Omega)}$ presents a reasonable choice in the definition of the condition number ($\star\star$) whenever $\|\cdot\|_\times = \|\cdot\|_{\ell^2(\Omega)}$ is a reasonable choice. Here, for sampling points $\Omega \in \hat{\mathbb{R}}^d$ and a $f \in \mathrm{H}$ we define

$$\|\hat{f}\|^2_{\ell^2(\Omega)} = \sum_{\omega \in \Omega} |\hat{f}(\omega)|^2, \qquad \|\hat{f}\|^2_{\ell^2_\mu(\Omega)} = \sum_{\omega \in \Omega} \mu_\omega |\hat{f}(\omega)|^2,$$

with $\{\mu_\omega\}_{\omega \in \Omega}$ chosen as the Lebesgue measures of the Voronoi regions associated to Ω. Since

$$\inf_{\omega \in \Omega} \mu_\omega \|\hat{f}\|^2_{\ell^2(\Omega)} \leq \|\hat{f}\|^2_{\ell^2_\mu(\Omega)} \leq \sup_{\omega \in \Omega} \mu_\omega \|\hat{f}\|^2_{\ell^2(\Omega)},$$

the corresponding condition numbers are equivalent for any separated set Ω. However, if Ω is not separated then $\|\hat{f}\|^2_{\ell^2(\Omega)}$ blows up and the corresponding condition number is equal to zero. To prevent such scenario, one can use the Voronoi weights as in Chapter 2, and ensure $\kappa > 0$ by setting $\|\cdot\|_\times = \|\cdot\|_{\ell^2_\mu(\Omega)}$.

Remark 3.1.4 If $\{\mu_n e_{\omega_n}\}_{n \in \mathbb{N}}$ is a weighted Fourier frame, one might immediately notice that $\mathcal{S}$ chosen as (3.1.9) is just a truncated frame operator corresponding to frame $\{\mu_n e_{\omega_n}\}_{n \in \mathbb{N}}$. Therefore, in this case, Theorem 3.1.3 is just a particular instance of results shown in [AHP13], i.e. NUGS is equivalent to the case of GS where the sampling system is a weighed Fourier frame.

Remark 3.1.5 Although we assume throughout the remainder of the paper that $\mathcal{S}$ takes the form (3.1.9), the results of this section do not require this. They only assume that $\mathcal{S}$ is admissible in the sense of Definition 3.1.1. This allows one to consider more general forms for $\mathcal{S}$ than the diagonal choice (3.1.9), as has recently been considered in several works. In [GS14], Gelb & Song use banded operators $\mathcal{S}$ for nonuniform Fourier sampling, and in [BG13] Berger & Gröchenig consider improved choices for $\mathcal{S}$ within the setting of GS in general Hilbert spaces.

3.1.2 Computation of the reconstruction

The computation of the NUGS reconstruction is the same as the computation of the GS reconstruction from [AH12b, AHP13] with the additional computation of the corresponding weights. Namely, if $\{\phi_m\}_{m=1}^M$ spans T, and if the reconstruction $\tilde{f} \in \mathrm{T}$ is defined via (3.1.10) and written as

$$\tilde{f} = \sum_{m=1}^{M} a_m \phi_m,$$

then the vector of coefficients $a = (a_1, \ldots, a_M)^\top$ is the least-squares solution of the $N \times M$ linear system

$$Aa \approx b, \tag{3.1.12}$$

where $b = (b_1, \ldots, b_N)^\top$ and $A \in \mathbb{C}^{N \times M}$ have entries

$$b_n = \sqrt{\mu_n} \hat{f}(\omega_n), \quad A_{n,m} = \sqrt{\mu_n} \hat{\phi}_m(\omega_n), \quad n = 1, \ldots, N, \; m = 1, \ldots, M. \tag{3.1.13}$$

Thus, once a basis for T is specified, $\tilde{f}$ can be computed by solving the least-squares problem for (3.1.12). A least-squares problem is typically solved by an iterative scheme, such as the conjugate gradient method. The efficiency of such an iterative scheme is always dependent on the costs of performing matrix-vector operations with A and its adjoint A^*, which are in general $\mathcal{O}(NM)$. The computational cost is also proportional to the condition number of the matrix A, which determines the number of iterations required in an iterative solver. For the later, from [AH12b, Lem. 2.11], we have

$$\mathrm{cond}(A) \leq C_w(\Omega_N, \mathrm{T}) \mathrm{cond}(B),$$

where

$$C_w(\Omega_N, \mathrm{T}) = \sqrt{\frac{C_3(\Omega_N, \mathrm{T})}{C_1(\Omega_N, \mathrm{T})}} \tag{3.1.14}$$

and $B \in \mathbb{C}^{M \times M}$ is the Gram matrix for $\{\phi_m\}_{m=1}^M$. In particular, if $\{\phi_m\}_{m=1}^M$ is a Riesz basis for H with constants d_1 and d_2, then

$$\mathrm{cond}(A) \leq C_w(\Omega_N, \mathrm{T}) \frac{d_2}{d_1}.$$

Hence, provided a Riesz or orthonormal basis is chosen for T, the condition number of A is small precisely when $C_w(\Omega_N, \mathrm{T})$ is also small. In this case, the reconstruction $\tilde{f}$ can be computed using a correspondingly small number of iterations.

Next, we give a result that asserts that $C_1(\Omega_N, \mathrm{T})$ and $C_3(\Omega_N, \mathrm{T})$ can be computed. For the proof see [AH12b, Lem. 2.13] (see also [AHP13, Lem. 5.2]).

Lemma 3.1.6. *Let $\{g_m\}_{m=1}^M$ be a basis for* T *and suppose that A is defined by* (3.1.13).

Then the constants $C_3(\Omega_N, \mathrm{T})$ and $C_1(\Omega_N, \mathrm{T})$ are the maximal and minimal eigenvalues of the matrix pencil $\{A^*A, B\}$, where $B \in \mathbb{C}^{M\times M}$ is the Gram matrix for $\{\phi_m\}_{m=1}^M$. Moreover, if $\{g_m\}_{m=1}^M$ is an orthonormal basis then

$$C_3(\Omega_N, \mathrm{T}) = \sigma^2_{\max}(A), \quad C_1(\Omega_N, \mathrm{T}) = \sigma^2_{\min}(A),$$

and $\mathrm{cond}(A) = C_w(\Omega_N, \mathrm{T})$*, where $C_w(\Omega_N, \mathrm{T})$ is given by* (3.1.14)*.*

Unfortunately, $C_w(\Omega_N, \mathrm{T})$ provides only a lower bound for the reconstruction constant $C(\Omega_N, \mathrm{T})$, and thus computing $C_w(\Omega_N, \mathrm{T})$ does not give rise to an estimate for the constant in the error bound (3.1.7). Nevertheless, the fact that $C_w(\Omega_N, \mathrm{T})$ is computable means that $C(\Omega_N, \mathrm{T})$ can in fact be numerically approximated via the following limiting process:

Lemma 3.1.7. *Suppose that Ω_N is finite and let $\mathcal{S} : \mathrm{H} \to \mathrm{H}$ be a linear operator satisfying conditions (i) and (ii) of Definition* 3.1.1*. Let T_M, $M \in \mathbb{N}$, be a sequence of finite-dimensional reconstruction spaces such that the corresponding orthogonal projections $\mathcal{P}_M = \mathcal{P}_{\mathrm{T}_M}$ converge strongly to the identity on* H*. Then*

$$C_2(\Omega_N) = \lim_{M\to\infty} C_3(\Omega_N, \mathrm{T}_M).$$

In particular, $C_2(\Omega_N)$ can be approximated to arbitrary accuracy by taking M sufficiently large.

Proof. Note first that $C_3(\Omega_N, \mathrm{T}_M) \le C_2(\Omega_N)$. Let $f \in \mathrm{H}$, $\|f\| = 1$. Then

$$\begin{aligned} \langle \mathcal{S}f, f\rangle &= \langle \mathcal{S}\mathcal{P}_M f, \mathcal{P}_M f\rangle + \langle \mathcal{S}(f - \mathcal{P}_M f), \mathcal{P}_M f\rangle + \langle \mathcal{S}f, f - \mathcal{P}_M f\rangle \\ &\le C_3(\Omega_N, \mathrm{T}_M) + 2\sqrt{C_2(\Omega_N)}\sqrt{\langle \mathcal{S}(f - \mathcal{P}_M f), f - \mathcal{P}_M f\rangle}. \end{aligned}$$

Thus,

$$C_3(\Omega_N, \mathrm{T}_M) \le C_2(\Omega_N) \le C_3(\Omega_N, \mathrm{T}_M) + 2\sqrt{C_2(\Omega_N)} \sup_{\substack{f\in \mathrm{H} \\ \|f\|=1}} \sqrt{\langle \mathcal{S}(f - \mathcal{P}_M f), f - \mathcal{P}_M f\rangle}.$$

It suffices to prove that the final term tends to zero as $M \to \infty$.

The operator $\mathcal{S}$ is linear and, for any g, $\mathcal{S}g$ depends only on the finite set of values $\hat{g}(\omega)$, $\omega \in \Omega_N$. Therefore, $\mathcal{S}$ is bounded and has finite rank. The result now follows immediately from this and the strong convergence $\mathcal{P}_M \to \mathcal{I}$. □

Since $C_2(\Omega_N)$ can always be approximated for finite Ω_N, one can always numerically estimate the reconstruction constant $C(\Omega_N, \mathrm{T})$ and therefore guarantee stability and quasi-optimality of the reconstruction *a priori*. We note that this limiting process may be avoided altogether in the case when appropriate conditions on the sampling density and the weights are satisfied. This will become apparent in the following sections.

Remark 3.1.8 In our work, we chose weights as measures of Voronoi regions corresponding to the sampling points, see Definition 2.1.5. In order to compute Voronoi weights one might use MATLAB's function `voronoiDiagram` which is computed using the Delaunay triangulation; see for example [Kle89]. One can also use function `mri_density_comp` from the NUFFT package by Fessler et al. [FS03]

Remark 3.1.9 As mentioned, efficient computation of $\tilde{f}$ relies on a fast algorithm for performing matrix-vector computations with A and A^*. The existence of such algorithms depends critically on the choice of the reconstruction space T. Fortunately, in the important case of wavelets, fast algorithms, based on Nonuniform Fast Fourier Transforms (NUFFTs) and fast wavelet transforms, can be incorporated leading to computational cost $\mathcal{O}(M \log N)$. This is further discussed in Section 4.4.

Remark 3.1.10 Recall that the ACT algorithm [FGS95, Grö99, Grö01] can be viewed as an instance of NUGS where $\hat{\mathrm{T}} = \{\hat{g} : g \in \mathrm{T}\}$ is a space of trigonometric polynomials on a compact interval. Therein, efficient implementation in $\mathcal{O}(N \log N)$ time is carried out using fast Toeplitz solvers, although one could also use NUFFTs with the same overall complexity (see [KKP07]), as we shall do in the case of wavelet choices for T.

3.2 The univariate guarantees

In this section, we provide a generalized sampling theorem in the univariate setting, which asserts that stable and quasi-optimal reconstruction is possible for any fixed finite-dimensional $\mathrm{T} \subseteq \mathrm{H} = \{f \in \mathrm{L}^2(\mathbb{R}) : \mathrm{supp}(f) \subseteq D\}$, where D is an interval on the real line, under appropriate conditions on nonuniform sampling scheme $\Omega_N \subseteq \hat{\mathbb{R}}$. We shall consider two scenarios in the next two subsections. First, sampling schemes Ω_N subject to appropriate density and bandwidth conditions. Second, sampling schemes arising from Fourier frames.

3.2.1 (K, δ)-dense sampling schemes

We commence with the following definition:

Definition 3.2.1. *Let $K > 0$ and $\omega_1 < \omega_2 < \ldots < \omega_N$. The sampling scheme $\Omega_N = \{\omega_1, \ldots, \omega_N\}$ has bandwidth K and density δ if*

(i) $\Omega_N \subseteq [-K, K]$, *and*

(ii) Ω_N *is δ-dense in the interval $[-K, K]$ in the sense of Definition 2.1.1.*

In this case, we say that Ω_N is (K, δ)-dense.

Our main result in this section is to show that, for an arbitrary fixed reconstruction space T, (K, δ)-density for suitably large K and small δ ensures stable reconstruction. This holds provided the weights μ_n in (3.1.9) are chosen as the lengths of the corresponding Voronoi intervals within $[-K, K]$. Namely, the weights are defined as

$$\mu_n = \frac{\omega_{n+1} - \omega_{n-1}}{2}, \; n = 2, \dots, N-1, \quad \mu_1 = \frac{\omega_1 + \omega_2}{2} + K, \quad \mu_N = K - \frac{\omega_{N-1} + \omega_N}{2} \tag{3.2.1}$$

Note that, since we are in dimension $d = 1$, (ii) in Definition 3.2.1 is equivalent to

$$\max_{n=0,\dots,N} \{\omega_{n+1} - \omega_n\} = 2\delta,$$

where $\omega_0 = -K - \delta$ and $\omega_{N+1} = K + \delta$.

We now require the following lemma.

Lemma 3.2.2. *Let $\Omega_N = \{\omega_1, \dots, \omega_N\}$ be (K, δ)-dense and suppose that $\mu_1, \dots, \mu_N$ are given by (3.2.1). Then for any nonzero function $f \in \mathrm{H}$ we have*

$$\left(\sqrt{1 - \|\hat{f}\|^2_{\mathbb{R} \setminus I} / \|f\|^2} - 4 m_D \delta \right)^2 \|f\|^2 \leq \sum_{n=1}^{N} \mu_n |\hat{f}(\omega_n)|^2 \leq (1 + 4 m_D \delta)^2 \|f\|^2,$$

where $m_D = \sup_{x \in D} |x|$, $I = (-K, K)$, and $\|\hat{f}\|^2_{\mathbb{R} \setminus I} = \int_{\mathbb{R} \setminus I} |\hat{f}(\omega)|^2 \, \mathrm{d}\omega$.

This lemma is an extension, with a similar proof, of the one-dimensional Gröchenig's result, Theorem 2.1.6, to the case where the number of samples N is finite. Gröchenig's result is obtained in the limit $N, K \to \infty$. Indeed, from Gröchenig's result, and as evident from the proof below, we have that the upper bound is less or equal to $(1 + 4m_D\delta)^2$ for any N and K. We also note that the lower bound is strictly less than $(1 - 4m_D\delta)^2$ for any nonzero f, since f is compactly supported and hence $\hat{f}$ cannot have compact support. However, the lower bound converges to $(1 - 4m_D\delta)^2$ as the bandwidth K is increased. In other words, N Fourier samples with density $\delta < 1/(4m_D)$ and appropriately large bandwidth K are sufficient to control $\|f\|$. This observation will lead to the main result in this section.

We briefly note that, in this lemma, D is an interval in $\mathbb{R}$, and that without loss of generality, it can be assumed that D is symmetric interval around zero so that m_D represents the half-length of the interval D. Namely, if D is not symmetric, then there exist a constant s such that $D_s = D - s$ is, and for all functions $F \in \mathrm{L}^2(D_s)$ defined by $F(\cdot) = f(\cdot + s)$, $f \in \mathrm{L}^2(D)$, we have $|\hat{F}(\omega)| = |\hat{f}(\omega)|$, and also $\|F\| = \|f\|$. Therefore, in this case, we can always consider $\mathrm{L}^2(D_s)$ instead of $\mathrm{L}^2(D)$.

Proof. Let $z_n = \frac{1}{2}(\omega_{n-1} + \omega_n)$, $n = 2, \ldots, N-1$ and $z_1 = -K$, $z_N = K$. Write

$$h(\omega) = \sum_{n=1}^{N} \hat{f}(\omega_n) \chi_{[z_n, z_{n+1})}(\omega)$$

so that

$$S^2 = \sum_{n=1}^{N} \mu_n |\hat{f}(\omega_n)|^2 = \int_{-K}^{K} |h(x)|^2 \,\mathrm{d}x = \|h\|_I^2,$$

where $I = (-K, K)$ and $\|\cdot\|_I$ denotes the L^2-norm over I. Hence

$$\|\hat{f}\|_I - \|\hat{f} - h\|_I \leq S \leq \|\hat{f}\|_{\mathbb{R}} + \|\hat{f} - h\|_I. \tag{3.2.2}$$

Using Wirtinger's inequality [Grö92, Lem. 1], we find that

$$\begin{aligned}
\|\hat{f} - h\|_I^2 &= \sum_{n=1}^{N} \int_{z_n}^{z_{n+1}} \left| \hat{f}(\omega) - \hat{f}(\omega_n) \right|^2 \,\mathrm{d}\omega \\
&= \sum_{n=1}^{N} \left(\int_{z_n}^{\omega_n} + \int_{\omega_n}^{z_{n+1}} \right) \left| \hat{f}(\omega) - \hat{f}(\omega_n) \right|^2 \,\mathrm{d}\omega \\
&\leq \sum_{n=1}^{N} \left(\frac{4(\omega_n - z_n)^2}{\pi^2} \int_{z_n}^{\omega_n} + \frac{4(z_{n+1} - \omega_n)^2}{\pi^2} \int_{\omega_n}^{z_{n+1}} \right) \left| \frac{\mathrm{d}}{\mathrm{d}\omega} \hat{f}(\omega) \right|^2 \,\mathrm{d}\omega \\
&\leq \frac{4\delta^2}{\pi^2} \int_I \left| \frac{\mathrm{d}}{\mathrm{d}\omega} \hat{f}(\omega) \right|^2 \,\mathrm{d}\omega,
\end{aligned}$$

where the final inequality follows from the (K, δ)-density of the samples. Since differentiation in Fourier space corresponds to multiplication by $(-2\pi \mathrm{i} x)$ in physical space, we conclude that

$$\|\hat{f} - h\|_I \leq 4\delta \|\widehat{f_1}\|_I \leq 4\delta \|\widehat{f_1}\| = 4\delta \|f_1\|,$$

where $f_1(x) = x f(x)$. Since f is supported in D, we deduce that

$$\|\hat{f} - h\|_I \leq 4 m_D \delta \|f\|. \tag{3.2.3}$$

Substituting this into the right-hand side of (3.2.2) gives $S \leq (1 + 4 m_D \delta)\|f\|$, and hence the upper bound. For the lower bound, by (3.2.2) and (3.2.3),

$$S \geq \|\hat{f}\|_I - 4 m_D \delta \|f\| \geq \sqrt{\|\hat{f}\|^2 - \|\hat{f}\|_{\mathbb{R}\setminus I}^2} - 4 m_D \delta \|f\|,$$

and the lower bound follows. $\square$

Definition 3.2.3. *Let* $\mathrm{T} \subseteq \mathrm{H}$*. For any* $z \in [0, \infty)$*, the* z*-residual of* T *is defined as*

$$E(\mathrm{T}, z) = \sup_{\substack{f \in \mathrm{T} \\ \|f\|=1}} \|\hat{f}\|_{\mathbb{R}\setminus(-z,z)}. \tag{3.2.4}$$

Note that $E(\mathrm{T}, z) \leq 1$, $\forall z$ and any T, since $\|\hat{f}\| = \|f\|$.

Lemma 3.2.4. *Let* $\mathrm{T} \subseteq \mathrm{H}$ *be a finite-dimensional subspace. Then* $E(\mathrm{T}, z) \to 0$ *monotonically as* $z \to \infty$*.*

Proof. Clearly $E(\mathrm{T}, z)$ is monotonically decreasing in z. Moreover, for any fixed $f \in \mathrm{T}$, we have $\|\hat{f}\|_{\mathbb{R}\setminus(-z,z)} \to 0$ as $z \to \infty$. The result now follows immediately from the fact that T is finite-dimensional. □

Combining the previous two lemmas, we immediately obtain our main result of this section:

Theorem 3.2.5. *Let* $\mathrm{T} \subseteq \mathrm{H}$ *be finite-dimensional and let* Ω_N *be* (K, δ)*-dense with*

$$\delta < \frac{1}{4m_D}.$$

Let $0 < \epsilon < \sqrt{1 - (4m_D\delta)^2}$*. For* $K > 0$ *large enough such that*

$$E(\mathrm{T}, K) \leq \epsilon,$$

the operator $\mathcal{S}$*, given by* (3.1.9) *with weights* (3.2.1)*, is admissible sampling operator with the reconstruction constant* $C(\Omega_N, \mathrm{T})$ *satisfying*

$$C(\Omega_N, \mathrm{T}) \leq \frac{1 + 4m_D\delta}{\sqrt{1 - \epsilon^2} - 4m_D\delta}. \tag{3.2.5}$$

Proof. The upper bound in Lemma 3.2.2 immediately gives $C_2(\Omega_N) \leq (1 + 4m_D\delta)^2$. For $C_1(\Omega_N, \mathrm{T})$ we set $f \in \mathrm{T}$ in Lemma 3.2.2, and then apply the definition of $E(\mathrm{T}, z)$ to get

$$C_1(\Omega_N, \mathrm{T}) \geq \left(\sqrt{1 - E(\mathrm{T}, K)^2} - 4m_D\delta\right)^2.$$

The result now follows from Lemma 3.2.4 and the definition of $C(\Omega_N, \mathrm{T})$. □

This theorem states the following. For a fixed reconstruction space T, the reconstruction constant $C(\Omega_N, \mathrm{T})$ can be made arbitrarily close to $(1 + 4m_D\delta)/(1 - 4m_D\delta)$ by taking K sufficiently large. Thus, even with highly nonuniform samples, we are guaranteed a stable reconstruction for large enough bandwidth K provided the density condition $\delta < 1/(4m_D)$ holds, with the precise level of stability controlled primarily by how close δ is to $1/(4m_D)$. As noted previously, in [Grö92] it was shown that infinite sequences

$\{\omega_n\}_{n\in\mathbb{N}}$ with bandwidth $K = \infty$ and density $\delta < 1/(4m_D)$ give rise to weighted Fourier frames $\{\sqrt{\mu_n}e_{\omega_n}\}_{n\in\mathbb{N}}$ for H. Therefore, based on arguments given in [Grö92], Theorem 3.2.5 shows that this condition also allows one to stably reconstruct from *finitely-many samples* in *any finite-dimensional subspace* T, provided the sampling bandwidth K is sufficiently large.

A key aspect of the Theorem 3.2.5 is the nature of the bound (3.2.5). The right-hand side separates geometric properties of the sampling scheme Ω_N, i.e. the density δ, from intrinsic properties of the reconstruction space T, i.e. the z-residual $E(\mathrm{T}, z)$. Hence, by analysing the z-residual for each particular choice of T, we can guarantee stable, quasi-optimal reconstruction for *all* sampling schemes Ω_N with $\delta < 1/(4m_D)$ and appropriate bandwidth K. This is how we shall proceed in Chapter 4 when we provide recovery guarantees for wavelet reconstruction spaces. We note in passing that a universal lower bound for $E(\mathrm{T}, z)$ for any subspace T of dimension M is provided by the M^{th} eigenvalue of the prolate spheroidal wavefunctions [LP62]. In particular, ensuring $E(\mathrm{T}, z) < c$ for some $c < 1$ necessitates at least a linear scaling of z with M, regardless of the choice of T. For wavelets, we show that a linear scaling is also sufficient.

Remark 3.2.6 In [Grö99], Gröchenig proves stability and convergence of the aforementioned ACT algorithm. As mentioned, this algorithm can be seen as a particular case corresponding to a trigonometric basis in frequency. The contribution of Theorem 3.2.5 is that it allows for arbitrary spaces T. Note that in Gröchenig's case (up to some minor differences in how the boundary is dealt with), $E(\mathrm{T}, K) = 0$ by construction of the space T. However, this is not true in general, and therefore it becomes important to estimate $E(\mathrm{T}, K)$ for particular choices of reconstruction space T.

3.2.2 Sampling at the critical density: the frame case

Unfortunately, the bound for $C(\Omega, \mathrm{T})$ declines as $\delta \to 1/(4m_D)$, and is infinitely large at the critical value $\delta = 1/(4m_D)$. This result is sharp in the sense that there are countable nonuniform sampling schemes $\Omega = \{\omega_n\}_{n\in\mathbb{Z}}$ (we now index over $\mathbb{Z}$ for convenience) with density $\delta = 1/(4m_D)$ which are not complete (see [Chr01] or [You01] for example), and for which one therefore cannot expect stable reconstructions. However, it is clear from considering uniform samples $\Omega = \{n/(2m_D)\}_{n\in\mathbb{Z}}$ that density $\delta = 1/(4m_D)$ is permissible in some cases since this is exactly the Nyquist rate. The standard approach to handle this "critical" density is to assume that the samples $\Omega = \{\omega_n\}_{n\in\mathbb{Z}}$ give rise to a (classical) Fourier frame $\{e_{\omega_n}\}_{n\in\mathbb{Z}}$ for H. As we show next, stable reconstruction with NUGS is also possible in this setting.

Let an ordered sequence $\{\omega_n : n \in \mathbb{Z}\}$ give rise to a Fourier frame and let $\Omega_N = \{\omega_n : |n| \leq N\}$. According to Theorem 3.1.3 stable reconstruction is possible provided an admissible sampling operator exists. Fortunately, this is always the case:

Theorem 3.2.7. *Let* T *be a finite-dimensional subspace of* H*, and suppose that* $\Omega_N = \{\omega_n : |n| \leq N\}$*, where* $\{\omega_n : n \in \mathbb{Z}\}$ *gives rise to a Fourier frame with* A *and* B *as the frame constants. Then the partial frame operator*

$$\mathcal{S}_N : f \mapsto \sum_{n=-N}^{N} \hat{f}(\omega_n) \mathrm{e}^{2\pi \mathrm{i} \omega_n \cdot}, \tag{3.2.6}$$

is admissible for all sufficiently large N*. Specifically, if*

$$\tilde{E}(\mathrm{T}, N)^2 = \sup_{\substack{f \in \mathrm{T} \\ \|f\|=1}} \sum_{|n|>N} |\hat{f}(\omega_n)|^2 \tag{3.2.7}$$

and for any $\epsilon \in (0, A)$*,* N *is large enough so that*

$$\tilde{E}(\mathrm{T}, N)^2 \leq \epsilon$$

then

$$C(\Omega, \mathrm{T}) \leq \frac{\sqrt{B}}{\sqrt{A - \epsilon}}. \tag{3.2.8}$$

Proof. The operator $\mathcal{S}_N$ trivially satisfies conditions (i) and (ii) of Definition 3.1.1. For the upper bound (3.1.3) we merely note that $\langle \mathcal{S}_N f, f \rangle \leq \langle \mathcal{S} f, f \rangle \leq B\|f\|^2$, where $\mathcal{S}$ is the frame operator (2.1.3). Moreover, since $\mathcal{S}_N \to \mathcal{S}$ strongly and T is finite-dimensional, (3.1.2) holds (with appropriate C_1) for all large N. Specifically, for $f \in \mathrm{T}$ we have

$$\langle \mathcal{S}_N f, f \rangle = \langle \mathcal{S} f, f \rangle - \langle (\mathcal{S} - \mathcal{S}_N) f, f \rangle \geq A\|f\|^2 - \sum_{|n|>N} |\hat{f}(\omega_n)|^2 \geq \left(A - \tilde{E}(\mathrm{T}, N)^2 \right) \|f\|^2,$$

which gives $C_1(\Omega, \mathrm{T}) \geq A - \tilde{E}(\mathrm{T}, N)^2$. We now apply the definition of $C(\Omega, \mathrm{T})$. $\square$

The result given here is a trivial adaptation of results for GS proved in [AHP13]. We include it and its proof for completeness. The novel results concerning classical Fourier frames come in Chapter 4 when we obtain estimates for the reconstruction constant $C(\Omega, \mathrm{T})$ for wavelets.

3.3 The multivariate guarantees

Next, by using the results of Chapter 2, we extend the work from the previous section to the multivariate case. For this, we use an analogous concept of (K, δ_*)-density.

Definition 3.3.1. *Let* $\Omega_N \subseteq \hat{\mathbb{R}}^d$ *be a set of sampling points,* $K > 0$ *and let* $|\cdot|_*$ *be an arbitrary norm on* $\mathbb{R}^d$*. If there exist a closed, simply connected set* $Z \subseteq \hat{\mathbb{R}}^d$ *with* 0 *in its interior such that* *(i)* $\max_{\hat{y}\in Z}|\hat{y}|_\infty = 1$,

(ii) $\Omega_N \subseteq Z_K$, *where* $Z_K = KZ$, *and*

(iii) Ω_N *is* δ_**-dense in the domain* Z_K *in the sense of Definition* 2.1.1,

then we say that the set Ω_N *is* (K,δ_*)*-dense with respect to* Z.

For a set of sampling points $\Omega_N \subseteq Z_K$, the weights $\mu_n > 0$ are chosen as measures of the Voronoi regions V_n^* within the area Z_K, i.e.

$$\mu_n = \mathrm{meas}(V_n^*), \quad V_n^* = \{\hat{y} \in Z_K \ : \ \forall m \neq n,\ |\hat{y}-\omega_n|_* \leq |\hat{y}-\omega_m|_*\}. \tag{3.3.1}$$

Also, analogously to (3.2.4), for a finite-dimensional subspace T we use the K-residuals

$$E(\mathrm{T},K) = \sup_{\substack{f\in\mathrm{T}\\ \|f\|=1}} \|\hat{f}-\hat{f}\chi_{Z_K}\| \tag{3.3.2}$$

We are ready to give the multivariate nonuniform generalized sampling theorem.

Theorem 3.3.2. *Let* $\mathrm{T} \subseteq \mathrm{H} = \{f \in \mathrm{L}^2(\mathbb{R}^d) : \mathrm{supp}(f) \subseteq D\}$ *be finite-dimensional,* $D \subseteq \mathbb{R}^d$ *compact, and let* $\Omega_N = \{\omega_n\}_{n=1}^N$ *be a sampling scheme. Let* Ω_N *be* (K,δ_*)*-dense with respect to* Z*, with*

$$\delta_* < \frac{\ln 2}{2\pi m_D c^*},$$

where $m_D = \sup_{x\in D}|x|$, $|\cdot|_*$ *is an arbitrary norm on* $\mathbb{R}^d$ *and* $c^* > 0$ *is the smallest constant such that* $|\cdot| \leq c^*|\cdot|_*$. *Let also* $0 < \epsilon < \sqrt{\exp(2\pi m_D\delta_* c^*)(2-\exp(2\pi m_D\delta_* c^*))}$. *If* $K > 0$ *is large enough so that*

$$E(\mathrm{T},K) \leq \epsilon,$$

then the operator $\mathcal{S}$ *given by* (3.1.9) *with the weights* (3.3.1) *is admissible and*

$$C(\Omega_N,\mathrm{T}) \leq \frac{\exp(2\pi m_D\delta_* c^*)}{\sqrt{1-\epsilon^2}+1-\exp(2\pi m_D\delta_* c^*)}. \tag{3.3.3}$$

Proof. The proof is similar as in the univariate case, but now we use the result of Section 2.2. Let

$$h(\hat{y}) = \sum_{\omega\in\Omega_N} \hat{f}(\omega)\chi_{V_{\omega,*}^K}(\hat{y}), \quad \hat{y}\in Z_K.$$

Hence

$$\|h\|_{Z_K}^2 = \sum_{\omega\in\Omega_N} \mu_\omega|\hat{f}(\omega)|^2.$$

Note that we have

$$\|f\|_{Z_K} - \|\hat{f}-h\|_{Z_K} \leq \|h\|_{Z_K} \leq \|\hat{f}-h\|_{Z_K} + \|f\|,$$

and also, by the same reasoning as in the proof of Theorem 2.2.1, we get

$$\|\hat{f} - h\|_{Z_K} \le (\exp(2\pi m_D \delta_* c^*) - 1)\, \|f\|.$$

Therefore for all $f \in \mathrm{H}\backslash\{0\}$

$$\begin{aligned}\left(\sqrt{1 - \|\hat{f}\|^2_{\hat{\mathbb{R}}^d \backslash Z_K}/\|f\|^2} + 1 - \exp\left(2\pi m_D \delta_* c^*\right)\right)^2 \|f\|^2 &\le \sum_{\omega\in\Omega_N} \mu_\omega |\hat{f}(\omega)|^2 \\ &\le \exp(4\pi m_D \delta_* c^*)\|f\|^2. \quad (3.3.4)\end{aligned}$$

Hence, we have $\sqrt{C_2(\Omega)} \le \exp(2\pi m_D \delta_* c^*)$ and

$$\sqrt{C_1(\Omega, \mathrm{T})} \ge \sqrt{1-\epsilon^2} + 1 - \exp\left(2\pi m_D \delta_* c^*\right) > 0,$$

due to (3.3.2) and the assumption that

$$E(\mathrm{T}, K) \le \epsilon < \sqrt{\exp\left(2\pi m_D \delta_* c^*\right)\left(2 - \exp\left(2\pi m_D \delta_* c^*\right)\right)}.$$

Now the statement follows by using the definition of the reconstruction constant $C(\Omega, \mathrm{T}) = \sqrt{C_2(\Omega)/C_1(\Omega, \mathrm{T})}$. □

Much like as we had in the univariate case, since T is finite-dimensional, the residual $E(\mathrm{T}, K)$ defined by (3.3.2) converges to zero when $K \to \infty$ and hence there always exists K such that $E(\mathrm{T}, K)$ is small enough. Therefore, this theorem guarantees stable and optimal recovery in an arbitrary finite-dimensional T, with the explicit bound on the reconstruction constant $C(\Omega, \mathrm{T})$, provided that the sampling scheme is sufficiently dense and wide in the frequency domain. However, the bound on density given by this result is not sharp.

Boundedness of the reconstruction constant $C(\Omega, \mathrm{T})$ under the sharp density condition can be provided by use of Theorem 2.3.1. However, the use of this theorem trades the explicitness of the bound, since it deploys non-explicit frame bounds A and B. Let $\Omega_N \subseteq Z_K$. We make use of the following K-residual

$$\tilde{E}(\mathrm{T}, K, \Omega_N) = \sup_{\substack{f\in \mathrm{T} \\ \|f\|=1}} \sqrt{\sum_{\omega\in\Omega\cap S_K} \mu_\omega |\hat{f}(\omega)|^2}, \tag{3.3.5}$$

where Ω is a sequence such that $\Omega_N \subseteq \Omega$ and such that it yields a weighted Fourier frame, $S_K = \hat{\mathbb{R}}^d \setminus D^\circ_{r(K)-1/2}$ and $D^\circ_{r(K)}$ is the largest ball with respect to the D°-norm inscribed into the set Z_K. Note that the existence of a sequence Ω is ensured if Ω_N has sufficient density. Also, note that the residual $\tilde{E}(\mathrm{T}, K, \Omega_N)$ again converges to zero as $K \to \infty$, but it now depends on both T and Ω_N.

Theorem 3.3.3. *Let* $\mathrm{T} \subseteq \mathrm{H} = \{f \in \mathrm{L}^2(\mathbb{R}^d) : \mathrm{supp}(f) \subseteq D\}$ *be finite-dimensional,* $D \subseteq \mathbb{R}^d$ *compact, convex and symmetric. Let* $\Omega_N = \{\omega_n\}_{n=1}^N$ *be* (K, δ_{D°)*-dense with respect to* Z*, with*

$$\delta_{D^\circ} < \frac{1}{4}.$$

Denote by A *and* B *the frame bounds corresponding to the weighed Fourier frame arising from* $\Omega = \{\omega_n\}_{n\in\mathbb{N}}$*,* $\Omega_N \subseteq \Omega$*, and let* $\epsilon \in (0, \sqrt{A})$*. If* $K > 0$ *is large enough so that*

$$\tilde{E}(\mathrm{T}, K, \Omega_N) \leq \epsilon,$$

then the operator $\mathcal{S}$ *given by* (3.1.9) *with the weights* (3.3.1) *is admissible sampling operator and*

$$C(\Omega_N, \mathrm{T}) \leq \sqrt{\frac{B}{A - \epsilon^2}}. \tag{3.3.6}$$

Proof. Due to (3.3.4), in this case we have $C_2 \leq \exp(\pi m_D c^\circ)$. However, for the lower bound C_1 we proceed as follows by using Theorem 2.3.1. Since Voronoi regions are taken with respect to Y_K instead of $\hat{\mathbb{R}}^d$, we need a subsequence $\Omega'_N \subseteq \Omega_N$ which has points sufficiently far from the boundary ∂Y_K so there is no any change in Voronoi regions. Since $\delta_{D^\circ} < 1/4$, we can take $\Omega'_N \subseteq D^\circ_{r(K)-1/2}$, where $D^\circ_{r(K)}$ is the largest inscribed ball with respect to D°-norm inside Y_K. Note that

$$\Omega \setminus \Omega'_N \subseteq \Omega \cap \left(\hat{\mathbb{R}}^d \setminus D^\circ_{r(K)-1/2}\right) = \Omega \cap S_K.$$

Therefore

$$\begin{aligned}
\sum_{\omega\in\Omega_N} \mu_\omega \left|\hat{f}(\omega)\right|^2 &\geq \sum_{\omega\in\Omega} \mu_\omega \left|\hat{f}(\omega)\right|^2 - \sum_{\omega\in\Omega\setminus\Omega'_N} \mu_\omega \left|\hat{f}(\omega)\right|^2 \\
&\geq A\|f\|^2 - \sum_{\omega\in\Omega\cap S_K} \mu_\omega \left|\hat{f}(\omega)\right|^2.
\end{aligned}$$

where the existence of $A > 0$ is provided by Theorem 2.3.1. Hence, by (3.3.5), for C_1 we have

$$C_1 \geq A - \tilde{E}(\mathrm{T}, K, \Omega_N)^2 \geq A - \epsilon^2 > 0.$$

Now since $C(\Omega, \mathrm{T}) = \sqrt{C_2/C_1}$, the result now follows immediately. □

Although the density conditions in theorems of this chapter are explicit, it is not yet stated how large the sampling bandwidth K needs to be. Nevertheless, this is possible to determine by analysing the residuals E or $\tilde{E}$. In particular, since the residual E depends only on a particular choice of the space T, once T is fixed, it is possible to determine scaling of K and dim(T) which gives sufficiently small E and therefore the stable and optimal recovery from any sufficiently dense sampling set Ω_K. This in return provides the so-called stable sampling rate, which we analyse in the one-dimensional case in the following chapter.

3.4 Examples of (K, δ)-dense sampling schemes

In this section, we construct some sampling schemes that are (K, δ)-dense in the sense of Definition 3.2.1 and 3.3.1 for the univariate and multivariate case respectively. In the following chapters, we shall illustrate NUGS for several different reconstruction spaces using (K, δ)-dense sampling schemes constructed here. Therein, we consider functions supported on $D = [-1, 1]^d$, $d = 1, 2$. According to Theorem 3.3.3, a sampling scheme Ω must satisfy the condition

$$\delta_{D^\circ}(\Omega) < \frac{1}{4}, \tag{3.4.1}$$

where D° is the unit ball in ℓ^1-norm, or, according to Theorem 3.3.2, if we choose the Euclidean norm to measure density, Ω must satisfy a more strict density condition

$$\delta_{\mathcal{B}_1}(\Omega) < \frac{\ln 2}{2\pi m_D}. \tag{3.4.2}$$

Recall that $m_D = \sqrt{2}$ if $D = [-1, 1]^2$. In this section, we construct some sampling schemes which satisfy these conditions. Note that for $D = [-1, 1]^2$ we have

$$\delta_{D^\circ}(\Omega) \leq \sqrt{2}\delta_{\mathcal{B}_1}(\Omega).$$

Hence, to have (3.4.1) it is enough to enforce

$$\delta_{\mathcal{B}_1}(\Omega) < \frac{1}{4\sqrt{2}}.$$

The condition

$$\delta_{\mathcal{B}_1}(\Omega) < c, \tag{3.4.3}$$

where $c > 0$ is a given constant, can be easily checked on a computer for an arbitrary nonuniform sampling scheme Ω. Moreover, as we shall show below, for special sampling schemes, e.g. radial and spiral, it is always possible to construct them so that they satisfy the condition (3.4.3). The advantage of considering the density condition in the Euclidean norm lies in its symmetry.

We mention that in [BW00], one can find a construction of a spiral sampling scheme satisfying condition (3.4.3). Here, we use a slightly different spiral scheme, one which has an accumulation point at the origin and cannot be treated without weights. More precisely, we use the constant angular velocity spiral, whereas Benedetto & Wu [BW00] use the constant linear velocity spiral (see [DHC+10, Fig 2]). Also, beside providing a sufficient condition for a spiral sampling scheme in order to satisfy (3.4.3), we provide both sufficient and necessary condition such that radial and jittered sampling schemes are appropriately dense.

3.4.1 Jittered sampling scheme

This sampling scheme is a standard model for jitter error, which appears when the measurement device is not scanning exactly on a uniform grid; see Figure 1.1. Due to its simplicity, we do not necessarily need to use the Euclidean norm in this case, therefore we consider directly the condition (3.4.1), and then, for completeness, we also consider (3.4.2). For a given $K > 0$ and parameters $\epsilon > 0$ and $\eta \geq 0$, we define the jittered sampling scheme as

$$\Omega = \{(n,m)\epsilon + \eta_{n,m} : n,m = -\lfloor K/\epsilon \rfloor, \dots, \lfloor K/\epsilon \rfloor\}, \tag{3.4.4}$$

where $\eta_{n,m} = (\eta^x_{n,m}, \eta^y_{n,m})$ with $\eta^x_{n,m}$ and $\eta^y_{n,m}$ such that $|\eta^x_{n,m}|, |\eta^y_{n,m}| \leq \eta$. Note that

$$\Omega \subseteq Z_{K'} = K'[-1,1]^2, \quad K' = \epsilon\lfloor K/\epsilon \rfloor + \eta.$$

Now, the following can easily be seen:

Proposition 3.4.1. *Let $D = [-1,1]^2$. Let also $K > 0$, $\epsilon > 0$ and $\eta \geq 0$ be given and define $K' = \epsilon\lfloor K/\epsilon \rfloor + \eta$. The sampling scheme Ω defined by (3.4.4) is*

(i) (δ_{E°, K')-dense with respect to $Z = [-1,1]^2$ and with $\delta_{E^\circ}(\Omega_K) < 1/4$ if and only if

$$\epsilon + 2\eta < \frac{1}{4}.$$

(ii) $(\delta_{\mathcal{B}_1}, K')$-dense with respect to $Z = [-1,1]^2$ and with $\delta_{\mathcal{B}_1}(\Omega_K) < (\ln 2)/(2\pi\sqrt{2})$ if and only if

$$\epsilon + 2\eta < \frac{\ln 2}{2\pi}.$$

Remark 3.4.2 For a given $K > 0$ and some $\eta, \epsilon > 0$, the one-dimensional jittered sampling scheme is defined by

$$\Omega = \{n\epsilon + \eta_n : n = -\lfloor K/\epsilon \rfloor, \dots, \lfloor K/\epsilon \rfloor\}, \tag{3.4.5}$$

where $\eta_n \in (-\eta, \eta)$ are chosen uniformly at random. Hence $\Omega \subseteq [-\epsilon\lfloor K/\epsilon \rfloor - \eta, \epsilon\lfloor K/\epsilon \rfloor + \eta]$ and the sampling density in this region is $\delta = \epsilon/2 + \eta$.

3.4.2 Radial sampling scheme

Here, we discuss an important type of sampling scheme used in MRI and also whenever the Radon transform is involved in sampling process, see Figure 1.1. For a given sampling bandwidth $K > 0$ and separation between consecutive concentric circles $r > 0$ we define a radial sampling scheme as

$$\Omega = \left\{ mre^{\mathrm{i}n\Delta\theta} \ : \ m = -\lfloor K/r \rfloor, \dots, \lfloor K/r \rfloor, \ n = 0, \dots, N-1 \right\}, \tag{3.4.6}$$

where $\Delta\theta = \pi/N \in (0,\pi)$ is the angle between neighbouring radial lines and $N \in \mathbb{N}$ is the number of radial lines in the upper half-plane. Note that

$$\Omega \subseteq \mathcal{B}_{r\lfloor K/r \rfloor} \subseteq \hat{\mathbb{R}}^2.$$

In what follows we shall assume that $K/r \in \mathbb{N}$ for simplicity.

Proposition 3.4.3. *Let $c > 0$, $K > c$, and $r \in (0, 2c)$ be given such that $K/r \in \mathbb{N}$. The sampling scheme Ω given by (3.4.6) is $(K, \delta_{\mathcal{B}_1})$-dense with respect to $Z = \mathcal{B}_1$ and with*

$$\delta_{\mathcal{B}_1}(\Omega) < c$$

if and only if

$$\Delta\theta < 2\min\left\{ \arctan\frac{\sqrt{c^2 - (r/2)^2}}{K - r/2}, \arccos\left(1 - \frac{c^2}{2K^2}\right) \right\}. \tag{3.4.7}$$

Proof. To prove this claim, we need to calculate

$$\sup_{\hat{y} \in \mathcal{B}_K} \inf_{\omega \in \Omega} |\hat{y} - \omega|_{\mathcal{B}_1}.$$

First note that, due to the definition of Voronoi regions 2.1.5, we have

$$\delta_{\mathcal{B}_1}(\Omega_K) = \sup_{\omega \in \Omega_K} \sup_{\hat{y} \in V_\omega} |\hat{y} - \omega|_{\mathcal{B}_1}, \tag{3.4.8}$$

where V_ω is the Voronoi region at ω with respect to the Euclidean norm and inside the domain $\mathcal{B}_K$. Therefore, we have to find the maximum radius of all Voronoi regions. Here, the radius of a Voronoi region V_ω is the radius of the Euclidean ball described around V_ω and centred at ω. Since the Voronoi regions are taken with respect to the Euclidean norm, they are convex polygons [Kle89], and hence, the Voronoi radius is always achieved at a vertex which is furthest away from the centre.

Since Ω_K is a radial sampling scheme with the uniform separation between consecutive concentric circles, the largest Voronoi radius is achieved at some of the vertices positioned between the two most outer circles of Ω_K, including the most outer circle. Note that, by the definition of Voronoi regions, a joint vertex of two adjacent Voronoi regions V_ω and $V_{\omega'}$ is equally distant from both points ω and ω'. Therefore, without loss of generality, in

(3.4.8), we may assume that the points form Ω_K are at the most outer circle.

Next, since $\mathcal{B}_1$ is symmetric with respect to any direction, and due to the symmetry of a radial sampling scheme, in (3.4.8), without loss of generality we may assume, that $\omega = K\mathrm{e}^{\mathrm{i}0}$, and $\hat{y} \in \left\{s\mathrm{e}^{\mathrm{i}\theta} : s \in (K-r, K],\ \theta \in [0, \Delta\theta/2]\right\}$. Denote $\omega' = (K-r)\mathrm{e}^{\mathrm{i}0}$. We now conclude that (3.4.8) is achieved at some of the following two vertices of V_ω, which are also the only vertices contained in the region $\left\{s\mathrm{e}^{\mathrm{i}\theta} : s \in (K-r, K],\ \theta \in [0, \Delta\theta/2]\right\}$:

(i) $v_1 = \frac{K-r/2}{\cos\theta_0}\mathrm{e}^{\mathrm{i}\Delta\theta/2}$, which is the joint vertex for adjacent V_ω and $V_{\omega'}$ lying on the radial line corresponding to angle $\Delta\theta/2$, at the equal distance $d(\Delta\theta)$ from both points ω and ω'. This point v_1 is easily calculated by equating the distances $|s\mathrm{e}^{\mathrm{i}\Delta\theta/2} - \omega|$ and $|s\mathrm{e}^{\mathrm{i}\Delta\theta/2} - \omega'|$. Also, one derives

$$d_1(\Delta\theta) = \sqrt{(r/2)^2 + ((K-r/2)\tan(\Delta\theta/2))^2}.$$

(ii) $v_2 = K\mathrm{e}^{\mathrm{i}\Delta\theta/2}$, which is a vertex of V_ω lying on the radial line corresponding to $\Delta\theta/2$ and at the most outer circle, at the distance

$$d_2(\Delta\theta) = K\sqrt{2 - 2\cos(\Delta\theta/2)}.$$

Hence, having $\delta_{\mathcal{B}_1}(\Omega_K) < c$ in the domain $\mathcal{B}_K$ is equivalent to

$$\max\{d_1(\Delta\theta), d_2(\Delta\theta)\} < c.$$

This is equivalent to

$$\Delta\theta < 2\min\left\{\arctan\frac{\sqrt{c^2 - (r/2)^2}}{K - r/2}, \arccos\left(1 - \frac{c^2}{2K^2}\right)\right\},$$

which proves our claim. □

This proposition asserts that δ-density of radial sampling scheme is satisfied if and only if the angle $\Delta\theta$ is sufficiently small and taken according to the formula (3.4.7). From (3.4.7), it is evident that the angle $\Delta\theta$ goes to zero linearly in $1/K$ when $K \to \infty$. Therefore, the condition $\delta_{\mathcal{B}_1}(\Omega) < c$ implies that the points Ω accumulate at the inner concentric circles as we increase K. Thus, the unweighted frame bounds for the frame sequence corresponding to Ω clearly blow up as $K \to \infty$, which can be prevented by using the weights.

3.4.3 Spiral sampling scheme

For a given $r > 0$,

$$S_r(\theta) = r\frac{\theta}{2\pi}\mathrm{e}^{\mathrm{i}\theta}, \quad \theta \geq 0, \tag{3.4.9}$$

is a spiral trajectory in $\hat{\mathbb{R}}^2$ with the constant separation r between the spiral turns. If $\theta \in [0, 2\pi k]$ for $k \in \mathbb{N}$, then the number of turns in the spiral is exactly k. For given $r > 0$ and $k \in \mathbb{N}$, let $Z_{rk} \subseteq \hat{\mathbb{R}}^2$ be defined as

$$Z_{rk} = \{S_\rho(\theta) : \rho \in [0, r],\ \theta \in [0, 2\pi k]\}\,. \tag{3.4.10}$$

Then $S_r(\theta) \subseteq Z_{rk} \subseteq \mathcal{B}_{rk}$, for $\theta \in [0, 2\pi k]$.

Now, let $K > 0$ and $r > 0$ be given, and for simplicity assume that they are such that

$$k = K/r \in \mathbb{N}.$$

We define a spiral sampling scheme as

$$\Omega = \left\{ r\tfrac{n\Delta\theta}{2\pi}\mathrm{e}^{\mathrm{i}n\Delta\theta} : n = 0, \dots, Nk \right\}. \tag{3.4.11}$$

where $\Delta\theta = 2\pi/N \in (0, \pi)$, $N \in \mathbb{N}$, is a discretization angle. Note that this Ω represents a discretization of the spiral trajectory (3.4.9), which consists of k turns with the constant separation r between them and with a constant angular distance $\Delta\theta$. Also, note that $\Omega \subseteq Z_K = KZ \subseteq \mathcal{B}_K \subseteq \hat{\mathbb{R}}^2$, where Z is

$$Z = \left\{ \rho\tfrac{\theta}{2\pi}\mathrm{e}^{\mathrm{i}\theta} : \rho \in [0, 1],\ \theta \in [0, 2\pi] \right\} \tag{3.4.12}$$

i.e. Z is given by (3.4.10) for $r = k = 1$.

Proposition 3.4.4. *Let $c > 0$, $K > (4/5)c$ and let $r \in (0, 2c)$ be given such that $K/r = k \in \mathbb{N}$. The sampling scheme Ω defined as (3.4.11) is $(\delta_{\mathcal{B}_1}, K)$-dense with respect to Z given by (3.4.12) and with*

$$\delta_{\mathcal{B}_1}(\Omega) < c$$

if $\Delta\theta < \tilde{\theta}$, where $\tilde{\theta}$ is such that $d_{r,k}(\tilde{\theta}) = c - r/2$, for $d_{r,k}(\cdot) = |S_r(2\pi k) - S_r(2\pi k - \cdot/2)|$ and S_r given by (3.4.9).

Proof. To prove this claim, we want to estimate $\delta_{\mathcal{B}_1}(\Omega)$. First note that the distance from any point inside region Y_{rk} to the spiral trajectory $S_r(\theta)$, $\theta \in [0, 2\pi k]$, is at most $r/2$, see [BW00, Eq. (18)]. Also, note that the distance from any point on the spiral trajectory $S_r(\theta)$, $\theta \in [0, 2\pi k]$, to a point from Ω_K is at most $|S_r(2\pi k) - S_r(2\pi k - \Delta\theta/2)|$. Hence, as in [BW00], by the triangle inequality we obtain

$$\delta_{\mathcal{B}_1}(\Omega) \leq \frac{r}{2} + |S_r(2\pi k) - S_r(2\pi k - \Delta\theta/2)|$$

where $S_r(\cdot)$ is given by (3.4.9). Therefore, the density condition is satisfied if $\Delta\theta$ is such that

$$d_{r,k}(\Delta\theta) < c - \frac{r}{2}.$$

Hence, it is enough to choose $\Delta\theta$ as

$$\Delta\theta < \tilde{\theta},$$

where $\tilde{\theta}$ is such that $d_{r,k}(\tilde{\theta}) = c - r/2$. Since function $d_{r,k}(\cdot)$ is continuous and strictly increasing on $(0,\pi)$ and also since

$$\lim_{\Delta\theta\to 0} d_{r,k}(\Delta\theta) = 0 < c - \frac{r}{2}, \quad \lim_{\Delta\theta\to \pi} d_{r,k}(\Delta\theta) = r\sqrt{k^2 + \left(k - \frac{1}{4}\right)^2} \geq \frac{5}{4}K > c - \frac{r}{2},$$

such $\tilde{\theta}$ exists and it is unique on interval $(0,\pi)$. $\square$

Let us mention here that in a similar manner an interleaving spiral sampling scheme can be analysed. An interleaving spiral consists of multiple single spirals. Both of these spirals are shown in Figure 1.1.

Remark 3.4.5 The one-dimensional analogue of a spiral sampling scheme is a *log* sampling scheme. For a sampling bandwidth $K > 0$ and some parameters $\nu, \delta > 0$ such that $2 \times 10^{-\nu} < \delta$, if $J = \left\lceil -\frac{\log_{10} K + \nu}{\log_{10}(1-\delta/K)} \right\rceil$, the log sampling scheme is defined by

$$\Omega = \{\pm\omega_j : \omega_j = 10^{-\nu + \frac{j}{J}(\log_{10} K + \nu)}, j = 0, \dots, J\}. \tag{3.4.13}$$

Note that this gives a (K,δ)-dense sampling sequence and $|\Omega| = 2(J+1)$.

Chapter 4

Reconstruction in wavelet spaces

In the previous chapter we established that stable and quasi-optimal reconstruction in arbitrary subspaces T is possible, provided one has the (K, δ)-dense sampling scheme with appropriate δ and large enough bandwidth K, or one has a frame sequence and the truncation parameter N sufficiently large. We now turn our attention to the question of precisely how large K (or N) needs to be for the important case where T consists of the first M terms of a wavelet basis in $\mathrm{L}^2(0,1)$. Our main result is to show that K (or N) needs to scale linearly in M to ensure stable and quasi-optimal reconstruction in this setting and this is presented in Section 4.2. Moreover, in Section 4.3, we show that the linear scaling is also necessary. These results are collected from [AGH14a], which is the joint work of the author with Ben Adcock and Anders Hansen.

In the last part of this chapter, Section 4.4, we describe how NUGS can be implemented in only $\mathcal{O}(N \log M)$ operations when recovering M wavelet coefficients from N Fourier samples. Due to the aforementioned linear correspondences, this leads to $\mathcal{O}(M \log M)$ operations in order to reconstruct M wavelet coefficients, provided that Fourier samples satisfy $N = \mathcal{O}(K)$. The material of this section is a joint work of the author with Clarice Poon [GP15]. The algorithm derived therein has been implemented in MATLAB and the code is available at http://www.damtp.cam.ac.uk/user/mg617/GS_wavelets.zip.

4.1 Preliminaries

Our interest lies in wavelet bases on the interval $[0, 1]$. Following [Mal09], we consider three standard constructions: periodic, folded and boundary-corrected wavelets. First, however, we recall the definition of a multiresolution analysis (MRA).

Definition 4.1.1. *A multiresolution analysis of* $\mathrm{L}^2(\mathbb{R})$ *generated by a scaling function* $\phi \in \mathrm{L}^2(\mathbb{R})$ *is a nested sequence of closed subspaces* $\{0\} \subseteq \cdots \subseteq V_{-1} \subseteq V_0 \subseteq V_1 \subseteq \cdots \subseteq \mathrm{L}^2(\mathbb{R})$ *such that*

(*i*) $\bigcup_{j\in\mathbb{Z}} V_j = \mathrm{L}^2(\mathbb{R})$ *and* $\bigcap_{j\in\mathbb{Z}} V_j = \{0\}$, *(ii) for all $j \in \mathbb{Z}$, $f(\cdot) \in V_j$ if and only if $f(2\cdot) \in V_{j+1}$,*

(iii) the collection $\{\phi(\cdot - k)\}_{k\in\mathbb{Z}}$ forms a Riesz basis for V_0.

Recall that a system $\{\phi(\cdot - k)\}_{k\in\mathbb{Z}}$ forms a Riesz basis for V_0 if and only if there exists constants $d_1, d_2 > 0$ such that

$$d_1 \sum_{k\in\mathbb{Z}} |\alpha_k|^2 \leq \left\| \sum_{k\in\mathbb{Z}} \alpha_k \phi(\cdot - k) \right\|^2 \leq d_2 \sum_{k\in\mathbb{Z}} |\alpha_k|^2, \quad \forall \{\alpha_k\}_{k\in\mathbb{Z}} \in l^2(\mathbb{Z}),$$

and $\{\phi(\cdot - k)\}_{k\in\mathbb{Z}}$ forms an orthonormal basis for V_0 if and only if $d_1 = d_2 = 1$. We recall also that this is equivalent to the condition

$$d_1 \leq \sum_{k\in\mathbb{Z}} \left| \hat{\phi}(k+\omega) \right|^2 \leq d_2, \quad a.e.\ \omega \in [0,1]. \tag{4.1.1}$$

In particular, the optimal Riesz basis constants are given by

$$d_1 = \operatorname*{essinf}_{\omega\in[0,1]} \sum_{k\in\mathbb{Z}} \left| \hat{\phi}(k+\omega) \right|^2, \qquad d_2 = \operatorname*{esssup}_{\omega\in[0,1]} \sum_{k\in\mathbb{Z}} \left| \hat{\phi}(k+\omega) \right|^2.$$

Periodic wavelets

Suppose that $\{\psi_{j,k}\}_{j,k\in\mathbb{Z}}$ is a wavelet basis of $\mathrm{L}^2(\mathbb{R})$ associated to an MRA with scaling function ϕ. Define the periodizing operation

$$f(x) \mapsto f^{\mathrm{per}}(x) = \sum_{k\in\mathbb{Z}} f(x+k), \tag{4.1.2}$$

and let $\psi_{j,k}^{\mathrm{per}}$ and $\phi_{j,k}^{\mathrm{per}}$ be the corresponding periodic wavelets and scaling functions. Define the periodized MRA spaces

$$V_j^{\mathrm{per}} = \operatorname{span}\left\{ \phi_{j,k}^{\mathrm{per}} : \ k = 0, \ldots, 2^j - 1 \right\}, \quad W_j^{\mathrm{per}} = \operatorname{span}\left\{ \psi_{j,k}^{\mathrm{per}} : \ k = 0, \ldots, 2^j - 1 \right\}.$$

Note that the maximal index k is finite, since $\phi_{j,k+2^j}^{\mathrm{per}} = \phi_{j,k}^{\mathrm{per}}$ and likewise for $\psi_{j,k}^{\mathrm{per}}$.

Now let $J \in \mathbb{N}_0$ be given. Then

$$\mathrm{L}^2(0,1) = \overline{V_J^{\mathrm{per}} \oplus W_J^{\mathrm{per}} \oplus W_{J+1}^{\mathrm{per}} \oplus \cdots},$$

and we may therefore introduce the finite-dimensional reconstruction space T by truncating the right-hand side:

$$\mathrm{T} = V_J^{\mathrm{per}} \oplus W_J^{\mathrm{per}} \oplus W_{J+1}^{\mathrm{per}} \oplus \cdots \oplus W_{R-1}^{\mathrm{per}}. \tag{4.1.3}$$

Note that $\dim(\mathrm{T}) = 2^R$. Since the original wavelets have an MRA, we also have that

$$\mathrm{T} = V_R^{\mathrm{per}} = \mathrm{span}\left\{\phi_{R,k}^{\mathrm{per}} : k = 0, \dots, 2^R - 1\right\}.$$

Our primary interest in this paper lies with wavelet bases having compact support. Without loss of generality, we now suppose that $\mathrm{supp}(\phi) \subseteq [-p+1, p]$ for $p \in \mathbb{N}$. Note the following: if $\mathrm{supp}(f) \subseteq [0,1]$ then $f(x) = f^{\mathrm{per}}(x)$ for $x \in [0,1]$. In particular, since

$$\mathrm{supp}(\phi_{R,k}) = [(k-p+1)/2^R, (k+p)/2^R],$$

we have that $\phi_{R,k}^{\mathrm{per}}(x) = \phi_{R,k}(x)$, $x \in [0,1]$, whenever $k = p, \dots, 2^R - p - 1$. Hence we may decompose the space T into

$$\mathrm{T} = \mathrm{T}^{\mathrm{left}} \oplus \mathrm{T}^{i} \oplus \mathrm{T}^{\mathrm{right}}, \tag{4.1.4}$$

where

$$\mathrm{T}^{i} = \mathrm{span}\left\{\phi_{R,k} : k = p, \dots, 2^R - p - 1\right\},$$

contains interior scaling functions with support in $(0,1)$ and

$$\begin{aligned} \mathrm{T}^{\mathrm{left}} &= \mathrm{span}\left\{\phi_{R,k}^{\mathrm{per}}\chi_{[0,1]} : k = 0, \dots, p-1\right\}, \\ \mathrm{T}^{\mathrm{right}} &= \mathrm{span}\left\{\phi_{R,k}^{\mathrm{per}}\chi_{[0,1]} : k = 2^R - p, \dots, 2^R - 1\right\}, \end{aligned}$$

contains the periodized scaling functions. Here $\chi_{[0,1]}$ is the indicator function of the interval $[0,1]$. Whilst not strictly necessary at this point, we add this function to the definitions of $\mathrm{T}^{\mathrm{left}}$ and $\mathrm{T}^{\mathrm{right}}$ so as to clarify that they are to be considered as subspaces of $\mathrm{H} = \{g \in \mathrm{L}^2(\mathbb{R}) : \mathrm{supp}(g) \subseteq [0,1]\}$ in our setting, and not $\mathrm{L}^2(\mathbb{R})$.

Remark 4.1.2 The stipulation that $\mathrm{supp}(\phi) \subseteq [-p+1, p]$ with $p \in \mathbb{N}$ makes little difference (besides affecting the constant) to the main result we establish in this section regarding $C(\Omega, \mathrm{T})$ with T as above. The key point is that ϕ should have compact support. In which case we can always find $p \in \mathbb{N}$ such that $\mathrm{supp}(\phi) \subseteq [-p+1, p]$.

Folded wavelets

Folded wavelets are defined via the folding operation

$$f(x) \mapsto f^{\mathrm{fold}}(x) = \sum_{k\in\mathbb{Z}} f(x-2k) + \sum_{k\in\mathbb{Z}} f(2k-x). \tag{4.1.5}$$

In this case, one obtains biorthogonal bases of wavelets for H. Note that we have

$$V_j^{\mathrm{fold}} = \mathrm{span}\left\{\phi_{j,k}^{\mathrm{fold}} :\ k = 0, \dots, 2^j - \iota\right\}, \quad W_j^{\mathrm{fold}} = \mathrm{span}\left\{\psi_{j,k}^{\mathrm{fold}} :\ k = 0, \dots, 2^j - 1\right\},$$

where ι takes value 0 if the wavelets are symmetric about $x = 1/2$ and 1 if they are antisymmetric. Much as before, we define the finite-dimensional reconstruction space

$$\mathrm{T} = V_J^{\text{fold}} \oplus W_J^{\text{fold}} \oplus W_{J+1}^{\text{fold}} \oplus \cdots \oplus W_{R-1}^{\text{fold}}, \tag{4.1.6}$$

and note that

$$\mathrm{T} = V_R^{\text{fold}} = \operatorname{span}\left\{\phi_{R,k}^{\text{fold}} : k = 0, \ldots, 2^R - \iota\right\},$$

As in the case of periodic wavelets, we can decompose T into three subspaces containing interior and boundary wavelets respectively. As before, suppose that $\operatorname{supp}(\phi) \subseteq [-p+1, p]$, $p \in \mathbb{N}$. Since $f(x) = f^{\text{fold}}(x)$ for $x \in [0, 1]$ whenever $\operatorname{supp}(f) \subseteq [0, 1]$, we have

$$\mathrm{T} = \mathrm{T}^{\text{left}} \oplus \mathrm{T}^{i} \oplus \mathrm{T}^{\text{right}},$$

where

$$\mathrm{T}^{i} = \left\{\phi_{R,k} : k = p, \ldots, 2^R - p - 1\right\},$$

and

$$\mathrm{T}^{\text{left}} = \left\{\phi_{R,k}^{\text{fold}}\chi_{[0,1]} : k = 0, \ldots, p-1\right\}, \quad \mathrm{T}^{\text{right}} = \left\{\phi_{R,k}^{\text{fold}}\chi_{[0,1]} : k = 2^R - p, \ldots, 2^R - \iota\right\}.$$

Boundary-corrected wavelets

We follow the boundary wavelet construction of Cohen, Daubechies & Vial [CDV93]. Let $p \in \mathbb{N}$ be given and denote the corresponding scaling and wavelet functions by ϕ and ψ. Note that the support of these functions is contained in $[-p+1, p]$. We define a new basis on $[0, 1]$ as follows. We set

$$\phi_{j,k}^{\text{int}}(x) = \begin{cases} 2^{j/2}\phi(2^j x - k) & p \le k < 2^j - p \\ 2^{j/2}\phi_k^{\text{left}}(2^j x) & 0 \le k < p \\ 2^{j/2}\phi_{2^j-k-1}^{\text{right}}(2^j(x-1)) & 2^j - p \le k < 2^j, \end{cases} \tag{4.1.7}$$

and similarly for the wavelet functions $\psi_{j,k}^{\text{int}}$. Here the functions ϕ_k^{left} and ϕ_k^{right} are particular boundary scaling functions. See [CDV93] for details. We may now define an MRA

$$V_j^{\text{int}} = \operatorname{span}\left\{\phi_{j,k}^{\text{int}} : \ k = 0, \ldots, 2^j - 1\right\}, \quad W_j^{\text{int}} = \operatorname{span}\left\{\psi_{j,k}^{\text{int}} : \ k = 0, \ldots, 2^j - 1\right\},$$

which, for $J \ge \log_2(2p)$ gives the reconstruction space

$$\mathrm{T} = V_J^{\text{int}} \oplus W_J^{\text{int}} \oplus \cdots \oplus W_{R-1}^{\text{int}} = V_R^{\text{int}}. \tag{4.1.8}$$

Note that, as before, we may decompose

$$\mathrm{T} = \mathrm{T}^{\text{left}} \oplus \mathrm{T}^{i} \oplus \mathrm{T}^{\text{right}},$$

where

$$\mathrm{T}^{i} = \operatorname{span}\left\{\phi^{\text{int}}_{R,k} : k = p, \ldots, 2^R - p - 1\right\},$$

contains the unmodified scaling functions with support in $[0,1]$ and

$$\begin{aligned} \mathrm{T}^{\text{left}} &= \operatorname{span}\left\{\phi^{\text{int}}_{R,k}\chi_{[0,1]} : k = 0, \ldots, p-1\right\}, \\ \mathrm{T}^{\text{right}} &= \operatorname{span}\left\{\phi^{\text{int}}_{R,k}\chi_{[0,1]} : k = 2^R - p, \ldots, 2^R - 1\right\}. \end{aligned}$$

Remark 4.1.3 Periodic wavelet bases on $[0,1]$ are widely used in standard implementations of wavelets, since their construction is extremely simple. However, the vanishing moments of the wavelet are lost due to the enforcement of periodic boundary conditions. This effectively introduces a discontinuity of the signal at the boundaries, and translates into lower approximation orders [Mal09]. Folded wavelets remove the artificial signal discontinuity introduced by periodization and allow for one vanishing moment to be retained. This approach is most commonly used for the CDF wavelets [CDF92]. However, since folded wavelets only retain one vanishing moment, they do not lead to high approximation orders for smooth functions. To obtain such orders, one may follow the boundary wavelet construction, due to Cohen, Daubechies & Vial [CDV93]. These boundary-corrected wavelets are particularly well suited for smooth functions. Indeed, if $f \in \mathrm{H}^s(0,1)$, where $\mathrm{H}^s(0,1)$ denotes the usual Sobolev space and $0 \leq s < p$, then the error

$$\|f - \mathcal{P}_{\mathrm{T}} f\| = \mathcal{O}(2^{-sR}), \quad R \to \infty, \tag{4.1.9}$$

where T is given by (4.1.8). Since NUGS is quasi-optimal, we obtain exactly the same approximation rates when reconstructing f from nonuniform Fourier samples, provided the bandwidth K (or N in the frame case) is chosen suitably large. Corollary 4.2.4 below establishes that K (or N) need only scale linearly in $M = 2^R$ to guarantee this.

Remark 4.1.4 Note that the wavelets introduced in this section—namely, periodic, folded or boundary-corrected—are considered as functions with support contained in $[0,1]$, even though they are actually defined over $\mathbb{R}$. In particular, their Fourier transforms are taken as integrals over $[0,1]$, as opposed to $\mathbb{R}$. Conversely, the scaling function ϕ is defined over the whole of $\mathbb{R}$, and thus its Fourier transform is also taken over $\mathbb{R}$.

4.2 Sufficiency of the linear scaling of K and $\dim(\mathrm{T})$

Here, we prove that a linear scaling of the sampling bandwidth K (or the frame truncation N) with $\dim(\mathrm{T})$ is sufficient for stable and quasi-optimal recovery when T consists of wavelets. First we state the main results, while the proofs are delayed until §4.2.3.

4.2.1 General wavelets

We commence with the (K, δ)-dense case:

Theorem 4.2.1. *Let Ω_N be a (K, δ)-dense sampling scheme with $\delta < 1/2$ and suppose that T is the reconstruction space (4.1.3) generated by the first 2^R elements of a periodic wavelet basis. Suppose that either of the following conditions holds:*

(i) the scaling function $\phi \in \mathrm{L}^2(\mathbb{R})$ and $\{\phi(\cdot - k)\}_{k\in\mathbb{Z}}$ forms an orthonormal basis of V_0,

(ii) the scaling function ϕ satisfies

$$|\hat{\phi}(\omega)| \leq \frac{c}{(1+|\omega|)^{\alpha}}, \quad \omega \in \mathbb{R}, \tag{4.2.1}$$

for some $\alpha > 1/2$, and the system $\{\phi(\cdot - k)\}_{k\in\mathbb{Z}}$ forms a Riesz basis of V_0.

Then for any $0 < \epsilon < \sqrt{1-2\delta}$, there exists a constant $c_0 = c_0(\epsilon)$ such that if

$$K \geq c_0(\epsilon) 2^R$$

then the reconstruction constant satisfies

$$C(\Omega_N, \mathrm{T}) \leq \frac{1+2\delta}{\sqrt{1-\epsilon^2-2\delta}}.$$

Theorem 4.2.2. *Let Ω_N be a (K, δ)-dense sampling scheme with $\delta < 1/2$ and suppose that either:*

(i) T *is generated by the first 2^R elements of the folded wavelets basis, given by (4.1.6), or*

(ii) T *is generated by the first 2^R elements of the boundary-corrected wavelets basis, given by (4.1.8).*

Suppose that $\{\phi(\cdot - k)\}_{k\in\mathbb{Z}}$ is a Riesz basis for V_0 and that ϕ satisfies (4.2.1) for some $\alpha > 1/2$. Then given $0 < \epsilon < \sqrt{1-2\delta}$ there exists a $c_0 = c_0(\epsilon)$ such that if

$$K \geq c_0(\epsilon) 2^R$$

then the reconstruction constant satisfies

$$C(\Omega_N, \mathrm{T}) \leq \frac{1+2\delta}{\sqrt{1-\epsilon^2-2\delta}}.$$

These theorems assert that it is sufficient that the bandwidth K scales linearly with the dimension of the reconstruction space T in the case of wavelets in order to ensure boundedness of the reconstruction constant. Note that the smoothness assumption (4.2.1) is extremely mild. For example, it holds if $\phi \in \mathrm{H}^{\alpha}(\mathbb{R})$ for $\alpha > 1/2$, and consequently includes all practical cases of interest. We remark also that the stipulation of a Riesz basis in these theorems is not necessary since this is implied by the MRA property. It is included merely for clarity.

We now give a similar result for the frame case:

Theorem 4.2.3. *Let $\Omega_N = \{\omega_n : |n| \leq N\}$, where $\{\omega_n : n \in \mathbb{Z}\}$ is a nondecreasing sequence that gives rise to a Fourier frame with frame bounds A and B. Let* T *be the reconstruction space of dimension 2^R consisting of either periodic* (4.1.3)*, folded* (4.1.6) *or boundary-corrected wavelets* (4.1.8)*, and suppose that ϕ satisfies* (4.2.1) *for some $\alpha > 1/2$. Then given $0 < \epsilon < A$ there exists a $c_0 = c_0(\epsilon)$ such that if*

$$N \geq c_0(\epsilon) 2^R$$

then the reconstruction constant satisfies

$$C(\Omega_N, \mathrm{T}) \leq \sqrt{\frac{B}{A-\epsilon}}.$$

As explained in Remark 4.1.3, boundary-corrected wavelets are an important case of these theorems. Due to (4.1.9), these results imply the following property of NUGS: up to constant factors, it obtains optimal convergence rates in terms of the sampling bandwidth when reconstructing smooth functions with boundary-corrected wavelets. Specifically,

Corollary 4.2.4. *Let* T *be the reconstruction space* (4.1.8) *consisting of the boundary-corrected wavelets with p vanishing moments. If $f \in \mathrm{H}^s(0,1)$, where $0 \leq s < p$, let $\tilde{f}$ denote the NUGS reconstruction based on a sampling scheme Ω_N. Then*

$$\|f - \tilde{f}\| = \mathcal{O}(K^{-s})$$

if Ω_N is as in Theorem 4.2.2*, and*

$$\|f - \tilde{f}\| = \mathcal{O}(N^{-s})$$

when Ω_N is as in Theorem 4.2.3*.*

4.2.2 Explicit estimates for Haar wavelets

Theorems 4.2.1–4.2.3 do not give explicit bounds for the constant $C(\Omega, \mathrm{T})$. In general, getting explicit estimates is difficult, due primarily to the contributions of the boundary subspaces $\mathrm{T}^{\mathrm{left}}$ and $\mathrm{T}^{\mathrm{right}}$. However, for the case of Haar wavelets, there are no such terms, and this means that explicit bounds are possible.

One motivation for studying the Haar wavelet case is that it corresponds to the situation of a digital model for the signal f. Specifically, the reconstruction space for Haar wavelets

$$\mathrm{T} = \mathrm{span}\left\{\phi \cup \{\psi_{j,k} : k = 0, \ldots, 2^j - 1,\ j = 0, \ldots, R-1\}\right\},$$

is a special case corresponding to $M = 2^R$ of reconstruction space

$$\mathrm{U} = \mathrm{U}_M = \left\{g \in \mathrm{L}^2(0,1) : g|_{[m/M,(m+1)/M)} = \text{constant},\ m = 0, \ldots, M-1\right\}, \tag{4.2.2}$$

consisting of piecewise constant functions (i.e. digital signals where $1/M$ is the pixel size). Note that

$$\mathrm{U}_M = \mathrm{span}\left\{\sqrt{M}\phi(M \cdot -m) : m = 0, \ldots, M-1\right\}, \tag{4.2.3}$$

is a subspace generated by shifts of the pixel indicator function $\phi(x) = \chi_{[0,1]}(x)$. This digital signal model is popular in imaging. In particular, it is the basis of the widely-used fast, iterative reconstruction technique for MRI [SNF03] (see Remark also 4.2.6).

Our next result gives an explicit upper bound for the reconstruction constant $C(\Omega, \mathrm{T})$ in this case, and demonstrates that $C(\Omega, \mathrm{T})$ is mild whenever M is at most $2K$.

Theorem 4.2.5. *Let Ω be a (K, δ)-dense sampling scheme with $\delta < 1/2$, and let $\mathrm{T} \subseteq \mathrm{U}_M$, where U_M is given by (4.2.3) for $\phi(x) = \chi_{[0,1]}(x)$ and $M \leq 2K$, such that $2K/M \in \mathbb{N}$. Then*

$$C(\Omega, \mathrm{T}) \leq \frac{\pi}{2}\left(\frac{1+2\delta}{1-2\delta}\right).$$

Remark 4.2.6 As noted previously, the well-known iterative reconstruction technique [SNF03] is a specific instance of NUGS corresponding to the choice (4.2.2) for T, where the term 'iterative' refers to the use of conjugate gradient iterations combined with NUFFTs to solve the least-squares problem. Thus, Theorem 4.2.5 provides an explicit guarantee for stable, quasi-optimal reconstruction with this method.

Remark 4.2.7 Theorem 4.2.5 can be easily generalized to the case where $\phi \in \mathrm{L}^2(\mathbb{R})$ is an arbitrary kernel such that (i) $\{\phi(\cdot - k)\}_{k\in\mathbb{Z}}$ forms a Riesz basis and (ii) $\mathrm{U}_M \subseteq \mathrm{H}$. Note that (ii) means that none of the shifted versions $\sqrt{M}\phi(M \cdot -m)$ can overlap with the interval endpoints $x = 0$ and $x = 1$. Thus such spaces have poor approximation properties for functions that do not themselves vanish at the endpoints. In such cases, it is preferable to consider the interval wavelet constructions based on periodic, folded or boundary-corrected wavelets, as described in the previous section, and whose reconstruction constants are addressed by Theorems 4.2.1 and 4.2.2 (albeit without explicit bounds).

4.2.3 Proofs of results from Sections 4.2.1 and 4.2.2

We first require the following lemma:

Lemma 4.2.8. *Let $I \subseteq \mathbb{N}$ be a finite index set and suppose that $\{\varphi_n : n \in I\} \subseteq \mathrm{H}$ is a Reisz basis for its span $\mathrm{T} = \mathrm{span}\{\varphi_n : n \in I\}$ with constants d_1 and d_2. Let I be partitioned into disjoint subsets $I_1, \ldots, I_r$, and write $\mathrm{T}_i = \mathrm{span}\{\varphi_n : n \in I_i\}$. Let $E(\mathrm{T}, z)$ and $\tilde{E}(\mathrm{T}, N)$ be given by (3.2.4) and (3.2.7) respectively. Then*

$$E(\mathrm{T}, z) \leq \sqrt{\frac{d_2}{d_1} \sum_{i=1}^{r} E(\mathrm{T}_i, z)^2}, \qquad \tilde{E}(\mathrm{T}, N) \leq \sqrt{\frac{d_2}{d_1} \sum_{i=1}^{r} \tilde{E}(\mathrm{T}_i, N)^2}$$

Proof. Let $f = \sum_{n \in I} \alpha_n \varphi_n \in \mathrm{T} \backslash \{0\}$ and write

$$f = \sum_{i=1}^{r} f_i, \qquad f_i = \sum_{n \in I_i} \alpha_n \varphi_n.$$

Note that

$$\|\hat{f}\|^2_{\mathbb{R}\backslash(-z,z)} \leq \left(\sum_{i=1}^{r} \|\widehat{f_i}\|_{\mathbb{R}\backslash(-z,z)} \right)^2 \leq \left(\sum_{i=1}^{r} E(\mathrm{T}_i, z) \|f_i\| \right)^2 \leq \sum_{i=1}^{r} E(\mathrm{T}_i, z)^2 \sum_{i=1}^{r} \|f_i\|^2.$$

Also, since $\{\varphi_n\}_{n \in I}$ forms a Riesz basis, we have $\sum_{i=1}^{r} \|f_i\|^2 \leq d_2/d_1 \|f\|^2$. Therefore

$$\frac{\|\hat{f}\|^2_{\mathbb{R}\backslash(-z,z)}}{\|f\|^2} \leq \frac{d_2}{d_1} \sum_{i=1}^{r} E(\mathrm{T}_i, z)^2.$$

Taking the supremum over f now gives the result for $E(\mathrm{T}, z)$. For $\tilde{E}(\mathrm{T}, N)$, we first note that $\sum_{|n|>N} |\hat{f}_i(\omega_n)|^2 < \infty$, $i = 1, \ldots, r$, since $\{\omega_n\}_{n \in \mathbb{Z}}$ gives rise to the Fourier frame and $f_i \in \mathrm{L}^2(0,1)$, $i = 1, \ldots, r$. Therefore, we can apply Minkowski's inequality to get

$$\sqrt{\sum_{|n|>N} |\hat{f}(\omega_n)|^2} \leq \sum_{i=1}^{r} \sqrt{\sum_{|n|>N} |\hat{f}_i(\omega_n)|^2}.$$

Thus,

$$\sum_{|n|>N} |\hat{f}(\omega_n)|^2 \leq \left(\sum_{i=1}^{r} \tilde{E}(\mathrm{T}^i, N) \|f_i\| \right)^2 \leq \frac{d_2}{d_1} \|f\|^2 \sum_{i=1}^{r} \tilde{E}(\mathrm{T}^i, N)^2,$$

as required. $\square$

Recall that all the wavelet reconstruction systems introduced in the previous section can be decomposed into interior wavelets having support in $[0,1]$ and boundary wavelets that intersect the endpoints $x = 0,1$. This lemma allows us to estimate the residuals $E(\mathrm{T}, z)$ and $\tilde{E}(\mathrm{T}, N)$ by considering each subspace separately. The next two propositions address the interior wavelets:

Proposition 4.2.9. *Let $\phi \in \mathrm{L}^2(\mathbb{R})$ have compact support and suppose that $\{\phi(\cdot - k)\}_{k\in\mathbb{Z}}$ forms a Riesz basis for its span with constants d_1 and d_2. Let $M \in \mathbb{N}$, $M_1, M_2 \in \mathbb{Z}$ and*

$$\mathrm{T} = \operatorname{span}\left\{\sqrt{M}\phi(M\cdot - m) : \ m = M_1, \ldots, M_2\right\},$$

and suppose that M, M_1, M_2 are such that $\mathrm{T} \subseteq \mathrm{H}$. Then the following hold:

(i) Given $\epsilon > 0$ there exists a $c_0 = c_0(\epsilon)$ such that for any $z \geq c_0 M$:

$$E(\mathrm{T}, z)^2 < 1 - \frac{d_1}{d_2} + \epsilon.$$

(ii) Suppose that ϕ satisfies (4.2.1) for some $\alpha > 1/2$. Then there exists a $c_0 = c_0(\epsilon)$ such that for any for $z \geq c_0 M$:

$$E(\mathrm{T}, z)^2 < \epsilon.$$

Proof. Let $f \in \mathrm{T}$ and write

$$f(x) = \sqrt{M} \sum_{k=M_1}^{M_2} a_k \phi(Mx - k).$$

Since $\{\phi(\cdot - k)\}_{k\in\mathbb{Z}}$ is a Riesz basis, we find that

$$d_1 \sum_{k=M_1}^{M_2} |a_k|^2 \leq \|f\|^2 \leq d_2 \sum_{k=M_1}^{M_2} |a_k|^2. \tag{4.2.4}$$

Moreover, a simple calculation gives that

$$\hat{f}(\omega) = \frac{1}{\sqrt{M}} \hat{\phi}\left(\frac{\omega}{M}\right) \Psi\left(\frac{\omega}{M}\right), \quad \omega \in \mathbb{R}, \tag{4.2.5}$$

where $\Psi(x) = \sum_{k=M_1}^{M_2} a_k \mathrm{e}^{-2\pi \mathrm{i} k x}$ is a trigonometric polynomial with $\|\Psi\|^2 = \sum_{k=M_1}^{M_2} |a_k|^2$. Thus

$$d_1 \|\Psi\|^2 \leq \|f\|^2 \leq d_2 \|\Psi\|^2, \tag{4.2.6}$$

by (4.2.4). We now estimate $\|\hat{f}\|^2_{(-z,z)}$. By (4.2.5), we have

$$\|\hat{f}\|^2_{(-z,z)} = \frac{1}{M}\int_{|\omega|<z} |\hat{\phi}(\omega/M)|^2\,|\Psi(\omega/M)|^2 \,\mathrm{d}\omega = \int_{|t|<z/M} |\hat{\phi}(t)|^2\,|\Psi(t)|^2 \,\mathrm{d}t.$$

Suppose that $z \geq M$ and write $\lfloor z/M \rfloor = n_0 + 1$, where $n_0 \in \mathbb{N}_0$. Then

$$\|\hat{f}\|^2_{(-z,z)} \geq \int_{t=-n_0}^{n_0+1} |\hat{\phi}(t)|^2\,|\Psi(t)|^2 \,\mathrm{d}t = \sum_{|n|\leq n_0}\int_0^1 |\hat{\phi}(t+n)|^2\,|\Psi(t+n)|^2 \,\mathrm{d}t.$$

Since Ψ is 1-periodic, and since (4.2.6) holds, we get

$$\|\hat{f}\|^2_{(-z,z)} \geq \left(\min_{t\in[0,1]}\sum_{|n|\leq n_0}|\hat{\phi}(n+t)|^2\right)\int_0^1 |\Psi(t)|^2\,\mathrm{d}t \geq \frac{1}{d_2}\left(\min_{t\in[0,1]}\sum_{|n|\leq n_0}|\hat{\phi}(n+t)|^2\right)\|f\|^2.$$

By [AHP14, Lem. 5.4], there exists an $n_0 \in \mathbb{N}$ sufficiently large such that the term in brackets is greater than $d_1 - \epsilon d_2$. Thus we get

$$\|\hat{f}\|^2_{(-z,z)} \geq \left(\frac{d_1}{d_2} - \epsilon\right)\|f\|^2.$$

We now use the definition of $E(\mathrm{T},z)^2$ to complete part 1. of the proof.

Our approach for part 2. is similar, where we estimate the tail $\|\hat{f}\|^2_{\mathbb{R}\setminus(-z,z)}$. Repeating the steps of the above proof, we find that

$$\|\hat{f}\|^2_{\mathbb{R}\setminus(-z,z)} \leq 1/d_1\left(\sup_{t\in[0,1]}\sum_{|n|\geq n_0}|\hat{\phi}(n+t)|^2\right)\|f\|^2.$$

Using the smoothness assumption (4.2.1), we find that

$$\sup_{t\in[0,1]}\sum_{|n|\geq n_0}|\hat{\phi}(n+t)|^2 \lesssim (n_0)^{1-2\alpha}.$$

Hence, if $z \geq c_0(\epsilon)M$ for some c_0, then

$$\|\hat{f}\|^2_{\mathbb{R}\setminus(-z,z)} \leq \epsilon\|f\|^2,$$

from which the result follows. $\square$

We are now ready to prove Theorem 4.2.1 and Theorem 4.2.2.

Proof of Theorems 4.2.1 and 4.2.2. By Theorem 3.2.5, it suffices to consider $E(\mathrm{T},z)$. Recall that in all three cases—periodic, folded or boundary-corrected wavelets—the reconstruction space T can be decomposed as $\mathrm{T} = \mathrm{T}^{\text{left}} \oplus \mathrm{T}^{i} \oplus \mathrm{T}^{\text{right}}$. Lemma 4.2.8 now gives

$$E(\mathrm{T}, z)^2 \leq \frac{d_2}{d_1}\left(E(\mathrm{T}^{\text{left}}, z)^2 + E(\mathrm{T}^{i}, z)^2 + E(\mathrm{T}^{\text{right}}, z)^2\right).$$

The subspace T^i contains wavelets supported in $[0,1]$, an application of Proposition 4.2.9 gives $E(\mathrm{T}^i, z)^2 < \epsilon$ in both case (i) and case (ii) of Theorem 4.2.1 (recall in case (i) that $\{\phi(\cdot - k)\}_{k\in\mathbb{Z}}$ is an orthonormal basis, and therefore $d_1 = d_2 = 1$), as well as in Theorem 4.2.2. Thus it remains to show in all cases that $E(\mathrm{T}^{\text{left}}, z)$ and $E(\mathrm{T}^{\text{right}}, z)$ can be made arbitrarily small with $z \gtrsim 2^R$

Consider the subspace T^{left} (the case of $\mathrm{T}^{\text{right}}$ is identical). For all three wavelet constructions, we may write

$$\mathrm{T}^{\text{left}} = \operatorname{span}\left\{\Phi_{R,k}\chi_{[0,1]} : k = 0, \ldots, p-1\right\},$$

where $\Phi_{R,k}$ is either $\phi^{\text{per}}_{R,k}$ (periodic), $\phi^{\text{fold}}_{R,k}$ (folded) or $\phi^{\text{int}}_{R,k}$ (boundary-corrected). The functions $\Phi_{R,k}\chi_{[0,1]}$ form a Riesz basis for T^{left} with bounds d_1 and d_2. Hence, if $f \in \mathrm{T}^{\text{left}}$ and

$$f = \sum_{k=0}^{p-1} \alpha_k \Phi_{R,k}\chi_{[0,1]},$$

then

$$d_1 \sum_{k=0}^{p-1} |\alpha_k|^2 \leq \|f\|^2 \leq d_2 \sum_{k=0}^{p-1} |\alpha_k|^2.$$

Now consider $\|\hat{f}\|_{\mathbb{R}\setminus(-z,z)}$. By the Cauchy–Schwarz inequality and the above inequality,

$$\begin{aligned}\|\hat{f}\|_{\mathbb{R}\setminus(-z,z)} &\leq \sum_{k=0}^{p-1} |\alpha_k| \|(\Phi_{R,k}\chi_{[0,1]})^\wedge\|_{\mathbb{R}\setminus(-z,z)} \\ &\leq \sqrt{p/d_1}\|f\| \max_{0\leq k\leq p-1}\left\{\|(\Phi_{R,k}\chi_{[0,1]})^\wedge\|_{\mathbb{R}\setminus(-z,z)}\right\}.\end{aligned}$$

Thus, to complete the proof, we only need to show that there exists a $c_0 = c_0(\epsilon)$ such that

$$\|(\Phi_{R,k}\chi_{[0,1]})^\wedge\|_{\mathbb{R}\setminus(-z,z)} < \epsilon, \qquad \forall k = 0, \ldots, p-1, \tag{4.2.7}$$

whenever $z \geq c_0(\epsilon)2^R$.

Assume now that $2^{R-1} > p$. Then one can determine the following:

(a) For periodic wavelets, $\Phi_{R,k}(x) = \phi_{R,k}(x) + \phi_{R,k}(x-1)$.

(b) For folded wavelets, $\Phi_{R,k}(x) = \phi_{R,k}(x) + \phi_{R,k}(-x)$.

(c) For boundary-corrected wavelets, $\Phi_{R,k}(x)$ can be written as a finite linear combination of the functions $\phi_{R,k}(x)$, where $k = -p+1, \ldots, p-1$.

Note that (a) and (b) follow by first writing $\phi^{\mathrm{per}}_{R,k}$ and $\phi^{\mathrm{fold}}_{R,k}$ in terms of infinite sums using the periodization and folding operations given by (4.1.2) and (4.1.5), and then by using the fact that $\mathrm{supp}(\phi) \subseteq [-p+1,p]$. Case (c) was shown in [CDV93]. Since in all cases $\Phi_{R,k}$ can be written as a finite sum with a number of terms independent of R, it therefore suffices to show that

$$\|(\phi_{R,k}\chi_{[0,1]})^\wedge\|_{\mathbb{R}\backslash(-z,z)}, \|(\phi_{R,k}(\cdot-1)\chi_{[0,1]})^\wedge\|_{\mathbb{R}\backslash(-z,z)}, \|(\phi_{R,k}(-\cdot)\chi_{[0,1]})^\wedge\|_{\mathbb{R}\backslash(-z,z)} < \epsilon, \tag{4.2.8}$$

where $k = -p+1,\ldots,p+1$ for the first term and $k = 0,\ldots,p-1$ for the second two terms, whenever $z \geq c_0(\epsilon)2^R$. Note that

$$\left|\left(\phi_{R,k}(\cdot+l)\chi_{[0,1]}\right)^\wedge(\omega)\right| = 2^{-R/2}\left|\int_{2^R l-k}^{2^R(l+1)-k} \phi(y)\mathrm{e}^{-2\pi\mathrm{i}\omega y/2^R}\,\mathrm{d}y\right|.$$

Suppose that $l=0$. Then the integration interval is $[-k, 2^R-k]$. Since $\mathrm{supp}(\phi) = [-p+1,p]$, we can replace this by $[-k,p]$ to give

$$\left|\left(\phi_{R,k}(\cdot)\chi_{[0,1]}\right)^\wedge(\omega)\right| = 2^{-R/2}\left|\widehat{\phi^{[-k,p]}}\left(\frac{\omega}{2^R}\right)\right|, \quad k=-p+1,\ldots,p-1,$$

where $\phi^{[a,b]}(x) = \phi(x)\chi_{[a,b]}(x)$ for $a<b$. Similarly, for $l=-1$ we have

$$\left|\left(\phi_{R,k}(\cdot-1)\chi_{[0,1]}\right)^\wedge(\omega)\right| = 2^{-R/2}\left|\widehat{\phi^{[-p+1,k]}}\left(\frac{\omega}{2^R}\right)\right|, \quad k=0,\ldots,p-1.$$

Likewise

$$\left|\left(\phi_{R,k}(-\cdot)\chi_{[0,1]}\right)^\wedge(\omega)\right| = 2^{-R/2}\left|\widehat{\phi^{[-p+1,k]}}\left(-\frac{\omega}{2^R}\right)\right|, \quad k=0,\ldots,p-1.$$

Thus, to establish (4.2.8), and therefore (4.2.7), it suffices to estimate the Fourier transforms of the functions $\phi^{[a,b]}$ for $(a,b) = (-k,p)$, $k=-p+1,\ldots,p-1$, and $(a,b) = (-p+1,k)$, $k=0,\ldots,p-1$. We now note the following:

$$\|2^{-R/2}f(\cdot/2^R)\|_{\mathbb{R}\backslash(-z,z)} = \|f\|_{\mathbb{R}\backslash(-z/2^R,z/2^R)}, \quad f \in \mathrm{L}^2(\mathbb{R}).$$

In particular, for any fixed f,

$$\|2^{-R/2}f(\cdot/2^R)\|_{\mathbb{R}\backslash(-z,z)} < \epsilon, \tag{4.2.9}$$

provided $z \geq c2^R$ for appropriately large $c>0$. Since the total number of functions $\phi^{[a,b]}$ is less than $2p$, and hence bounded independently of R, we obtain (4.2.8) and thus (4.2.7). $\square$

Having addressed the case of (K,δ)-dense samples, we now consider frame samples.

Recalling the setup of §3.2.2, let $\{\omega_n : n \in \mathbb{Z}\}$ be a nondecreasing sequence giving rise to a Fourier frame. Set $\Omega_N = \{\omega_n : |n| \le N\}$, and suppose that $\mathcal{S}_N$ is given by (3.2.6).

To prove our next result, we require the following two lemmas:

Lemma 4.2.10. *Let $\{\omega_n\}_{n\in\mathbb{Z}}$ be an increasing sequence of separated points with minimal separation $\eta = \inf_{n\in\mathbb{Z}}\{\omega_{n+1} - \omega_n\} > 0$. Then there exists a set of points $\{\tilde{\omega}_n\}_{n\in\mathbb{Z}}$ with minimal separation at least $\eta/2$ such that $\{\omega_n\}_{n\in\mathbb{Z}} \subseteq \{\tilde{\omega}_n\}_{n\in\mathbb{Z}}$ and*

$$\sup_{n\in\mathbb{Z}}\{\tilde{\omega}_{n+1} - \tilde{\omega}_n\} \le \eta.$$

Proof. Let $n \in \mathbb{Z}$. If $\omega_{n+1} - \omega_n = \eta$ then we do nothing. Otherwise, let $k \in \mathbb{N}$ be the smallest integer such that $\omega_{n+1} - \omega_n \le (k+1)\eta$. Introduce the new points

$$\omega_n + r\eta, \quad r = 1, \dots, k-1,$$

as well as

$$\frac{1}{2}\left(\omega_n + (k-1)\eta + \omega_{n+1}\right).$$

These new points are at least $\eta/2$ separated, and have maximal separation at most η. □

A variation of the following result was also proved in [Grö99, Lem. 1]. We include the proof for completeness.

Lemma 4.2.11. *Let $x_0 \le x_1 < x_2 < \dots < x_N \le x_{N+1}$ where $N \in \mathbb{N} \cup \{\infty\}$, and suppose that $\rho = \max_{n=0,\dots,N}\{x_{n+1} - x_n\} < \infty$. Let $f \in \mathrm{H}^1(a,b)$, where $a = \frac{1}{2}(x_1 + x_0)$, $b = \frac{1}{2}(x_{N+1} + x_N)$ and $\mathrm{H}^1(a,b)$ denotes the standard Sobolev space of first order on the interval (a,b). If $\mu_n = \frac{1}{2}(x_{n+1} - x_{n-1})$, $n = 1, \dots, N$, then the following inequalities hold:*

$$\left(\|f\|_{[a,b]} - \frac{\rho}{\pi}\|f'\|_{[a,b]}\right)^2 \le \sum_{n=1}^{N} \mu_n |f(x_n)|^2 \le \left(\|f\|_{[a,b]} + \frac{\rho}{\pi}\|f'\|_{[a,b]}\right)^2.$$

Proof. The proof of this lemma is similar to that of Lemma 3.2.2. Let $z_n = \frac{1}{2}(x_n + x_{n-1})$ and define $g(x) = \sum_{n=1}^{N} f(x_n) g_{[z_n, z_{n+1})}(x)$. Note that $z_1 = a$, $z_{N+1} = b$ and that

$$\sum_{n=1}^{N} \mu_n |f(x_n)|^2 = \|g\|^2_{[a,b]}.$$

We now have

$$\|f - g\|^2_{[a,b]} = \sum_{n=1}^{N} \int_{z_n}^{z_{n+1}} |f(x) - f(x_n)|^2 \, \mathrm{d}x,$$

and after an application of Wirtinger's inequality, we obtain

$$\|f - g\|^2_{[a,b]} \le \frac{\rho^2}{\pi^2}\|f'\|^2_{[a,b]}.$$

This gives the result. $\square$

Proposition 4.2.12. *Let $\{\omega_n\}_{n\in\mathbb{Z}} \subseteq \mathbb{R}$ be a nondecreasing sequence of frequencies that rise to a Fourier frame for H, and suppose that ϕ and T are as in Proposition 4.2.9. If ϕ satisfies (4.2.1) for some $\alpha > 1/2$, then given $\epsilon > 0$ there exists a $c_0 = c_0(\epsilon)$ such that for all $N \geq c_0 M$:*

$$\tilde{E}(\mathrm{T}, N) < \epsilon.$$

Proof. Recall from Theorem 2.1.3 that any sequence $\{\omega_n\}_{n\in\mathbb{Z}}$ that gives a frame is necessarily relatively separated, i.e. it is a finite union of separated sequences. Since we wish to obtain an upper bound for

$$\sum_{|n|>N} |\hat{f}(\omega_n)|^2,$$

for any $f \in \mathrm{T}$, we may therefore assume without loss of generality that $\{\omega_n\}_{n\in\mathbb{Z}}$ is a separated sequence with separation η. Moreover, after an application of Lemma 4.2.10, we may assume without loss of generality that $\{\omega_n\}_{n\in\mathbb{Z}}$ is $\eta/2$ separated with maximal spacing at most η.

As in the proof of Proposition 4.2.9 let $f = \sum_{k=M_1}^{M_2} a_k \sqrt{M}\phi(M\cdot -k) \in \mathrm{T}$ and write $\tilde{\Psi}(x) = \sum_{k=M_1}^{M_2} a_k \mathrm{e}^{-2\pi \mathrm{i}kx}$ so that

$$\hat{f}(\omega) = \frac{1}{\sqrt{M}}\hat{\phi}\left(\frac{\omega}{M}\right)\tilde{\Psi}\left(\frac{\omega}{M}\right). \tag{4.2.10}$$

Let

$$\Psi(x) = \mathrm{e}^{2\pi \mathrm{i}M_3 x}\tilde{\Psi}(x) = \sum_{k=M_1-M_3}^{M_2-M_3} a_{k+M_3}\mathrm{e}^{-2\pi \mathrm{i}kx}, \qquad M_3 = \left\lceil \frac{M_1+M_2}{2} \right\rceil, \tag{4.2.11}$$

so that $|\Psi(x)| = |\tilde{\Psi}(x)|$. By (4.2.10) we also have $|\hat{f}(\omega)| = \frac{1}{\sqrt{M}}|\hat{\phi}(\omega/M)||\Psi(\omega/M)|$, and therefore

$$\begin{aligned}\sum_{n>N} |\hat{f}(\omega_n)|^2 &\leq \frac{1}{M}\sum_{n>N}\left|\hat{\phi}\left(\frac{\omega_n}{M}\right)\right|^2\left|\Psi\left(\frac{\omega_n}{M}\right)\right|^2 \\ &\leq \frac{1}{M}\sum_{l=0}^{\infty}\sup_{\omega\in I_l}\left|\hat{\phi}\left(\frac{\omega}{M}\right)\right|^2\sum_{n:\omega_n\in I_l}\left|\Psi\left(\frac{\omega_n}{M}\right)\right|^2,\end{aligned}$$

where $I_l = [\omega_N + lM, \omega_N + (l+1)M)$. Since $\{\omega_n\}_{n\in\mathbb{Z}}$ is separated and increasing, we must have that $\omega_N \gtrsim N$ as $N \to \infty$. In particular $\omega_N > 0$ for sufficiently large N. By the assumption on ϕ, we therefore obtain

$$\sum_{n>N} |\hat{f}(\omega_n)|^2 \lesssim M^{2\alpha-1}\sum_{l=0}^{\infty}(\omega_N + 2lM)^{-2\alpha}\sum_{n:\omega_n\in I_l}\left|\Psi\left(\frac{\omega_n}{M}\right)\right|^2.$$

We now claim that the result follows, provided

$$\sum_{n:\omega_n \in I_l} \left| \Psi\left(\frac{\omega_n}{M}\right) \right|^2 \leq cM\|\Psi\|^2, \quad \forall l = 0, 1, 2, \ldots. \tag{4.2.12}$$

We shall prove that (4.2.12) holds in a moment. First, however, let us show how (4.2.12) implies the result. Substituting this bound into the previous expression gives

$$\sum_{n>N} |\hat{f}(\omega_n)|^2 \lesssim M^{2\alpha} \sum_{l=0}^{\infty} (\omega_N + 2lM)^{-2\alpha} \|\Psi\|^2 \lesssim \left(\frac{\omega_N}{M}\right)^{1-2\alpha} \|\Psi\|^2.$$

Similarly, we also get

$$\sum_{n<-N} |\hat{f}(\omega_n)|^2 \lesssim \left(\frac{|\omega_{-N}|}{M}\right)^{1-2\alpha} \|\Psi\|^2.$$

An application of (4.2.6) now gives

$$\tilde{E}(\mathrm{T}, N)^2 \lesssim \frac{1}{d_1} \left(\frac{\min\{\omega_N, |\omega_{-N}|\}}{M}\right)^{1-2\alpha}.$$

Since $\omega_N, |\omega_{-N}| \gtrsim N$ as $N \to \infty$, the result now follows.

It remains to establish (4.2.12). Write $\{\omega_n/M : \omega_n \in I_l\} = \{x_1, \ldots, x_L\}$ where

$$\omega_N/M + l \leq x_1 < x_2 < \ldots < x_L \leq \omega_N/M + l + 1,$$

and set $x_0 = x_1$ and $x_{L+1} = x_L$. Note that $\eta/(2M) \leq x_{n+1} - x_n \leq \eta/M$. Therefore

$$\sum_{n:\omega_n \in I_l} |\Psi(\omega_n/M)|^2 = \sum_{n=1}^{L} |\Psi(x_n)|^2 \leq \frac{2M}{\eta} \sum_{n=1}^{L} \mu_n |\Psi(x_n)|^2,$$

where $\mu_n = \frac{1}{2}(x_{n+1} - x_{n-1})$. Hence, by Lemma 4.2.11 we have

$$\sum_{n:\omega_n \in I_l} |\Psi(\omega_n/M)|^2 \leq \frac{2M}{\eta} \left[\|\Psi\|_{[a,b]} + \frac{\eta}{M\pi} \|\Psi'\|_{[a,b]}\right]^2,$$

where $a = \frac{1}{2}(x_1 + x_0) = x_1$ and $b = \frac{1}{2}(x_{L+1} + x_L) = x_L$. Note that $|b - a| \leq 1$. Hence since Ψ is periodic, we get

$$\sum_{n:\omega_n \in I_l} |\Psi(\omega_n/M)|^2 \leq \frac{2M}{\eta} \left[\|\Psi\| + \frac{\eta}{M\pi} \|\Psi'\|\right]^2.$$

To prove the result, we only need to show that $\|\Psi'\| \leq M\pi\|\Psi\|$. Since Ψ is a trigonometric polynomial given by (4.2.11), we have

$$\|\Psi'\| \leq 2\max\{M_2 - M_3, M_3 - M_1\}\,\pi\|\Psi\|.$$

Thus it remains to show that $M_2 - M_3, M_3 - M_1 \leq M/2$. Since $\mathrm{T} \subseteq \mathrm{H}$ by assumption, the function ϕ must have compact support. Let $\mathrm{supp}(\phi) \subseteq [a, b]$. Then we must also have that $-a \leq M_1 \leq M_2 \leq M - b$. In particular, $M_2 - M_1 \leq M - (b - a) < M$. Therefore

$$M_2 - M_3 \leq M_2 - \frac{M_1 + M_2}{2} < \frac{M}{2}, \quad M_3 - M_1 \leq \frac{M_1 + M_2}{2} + 1 - M_1 \leq \frac{M}{2} + 1 - \frac{b-a}{2}.$$

Since $M_3 - M_1 \in \mathbb{N}$ and $b - a > 0$ we obtain the result. □

We are now ready to prove our main result for frame samples.

Proof of Theorem 4.2.3. By Theorem 3.2.7, we may consider $\tilde{E}(\mathrm{T}, N)$. Proceeding in a similar manner to the proof of Theorems 4.2.1 and 4.2.2, we see from Lemma 4.2.8 that it suffices to estimate $\tilde{E}(\mathrm{T}^i, N)$, $\tilde{E}(\mathrm{T}^{\mathrm{left}}, N)$ and $\tilde{E}(\mathrm{T}^{\mathrm{right}}, N)$ separately. Since $\tilde{E}(\mathrm{T}^i, N)$ can be bounded using Proposition 4.2.12, it remains to derive bounds for $\tilde{E}(\mathrm{T}^{\mathrm{left}}, N)$ and $\tilde{E}(\mathrm{T}^{\mathrm{right}}, N)$ only. If we now argue in an identical way to the previous proof, i.e. by writing the spaces $\mathrm{T}^{\mathrm{left}}$ and $\mathrm{T}^{\mathrm{right}}$ as linear combinations of the functions $\phi^{[a,b]}$ whose total number is independent of R, then we see that it suffices to show the following: for an arbitrary function $f \in \mathrm{L}^2(0, 1)$,

$$2^{-R} \sum_{|n|>N} \left| \hat{f}\left(\frac{\omega_n}{2^R}\right) \right|^2 < \epsilon, \tag{4.2.13}$$

provided $N \geq c2^R$ for some $c > 0$ depending only on f (this replaces the condition (4.2.9) in the proof of Theorems 4.2.1 and 4.2.2). Recall from the proof of Proposition 4.2.12 that we may assume without loss of generality that the frame sequence $\{\omega_n\}_{n\in\mathbb{Z}}$ is separated with separation at least $\eta/2$ and maximal spacing at most η. Thus the points $\{\tilde{\omega}_n\}_{n\in\mathbb{Z}}$, where $\tilde{\omega}_n = \omega_n/2^R$, have maximal spacing at most $\eta/2^R$ and we find that

$$2^{-R} \sum_{|n|>N} \left| \hat{f}\left(\frac{\omega_n}{2^R}\right) \right| \leq \frac{2}{\eta} \sum_{|n|>N} \mu_n |\hat{f}(\tilde{\omega}_n)|^2,$$

where $\mu_n = \frac{\tilde{\omega}_{n+1} - \tilde{\omega}_{n-1}}{2}$. Since $f \in \mathrm{H}$ we may apply Lemma 4.2.11 to get

$$2^{-R} \sum_{|n|>N} \left| \hat{f}\left(\frac{\omega_n}{2^R}\right) \right| \leq \frac{2}{\eta}\left[\left(\|\hat{f}\|_{J_+} + \frac{\eta}{2^R\pi}\|\hat{f}'\|_{J_+} \right)^2 + \left(\|\hat{f}\|_{J_-} + \frac{\eta}{2^R\pi}\|\hat{f}'\|_{J_-} \right)^2 \right],$$

where $J_+ = (\tilde{\omega}_N, \infty)$ and $J_- = (-\infty, \tilde{\omega}_{-N})$. To obtain (4.2.13) we merely note that $\hat{f}' = \widehat{f_1} \in \mathrm{L}^2(\mathbb{R})$, where $f_1(x) = xf(x)$, and $\max\{\tilde{\omega}_N, -\tilde{\omega}_{-N}\} \gtrsim N/2^R$ for large N. □

Finally, we prove Theorem 4.2.5, which gives an explicit upper bound for the reconstruction constant in the case of reconstructing in Haar wavelets:

Proof of Theorem 4.2.5. Since we have already shown have $C_2(\Omega) \leq (1+2\delta)^2$, and since $\mathrm{T} \subseteq \mathrm{U}_M$, it is enough to estimate $C_1(\Omega, \mathrm{U}_M)$. For any $f \in \mathrm{U}_M$, we can write

$$f(x) = \sqrt{M} \sum_{m=0}^{M-1} a_m \phi(Mx - m).$$

Therefore, as before, we get

$$\hat{f}(\omega) = \frac{1}{\sqrt{M}} \hat{\phi}\left(\frac{\omega}{M}\right) \tilde{\Psi}\left(\frac{\omega}{M}\right), \tag{4.2.14}$$

where, for $M_0 = \lfloor M/2 \rfloor$,

$$\tilde{\Psi}(x) = \sum_{m=0}^{M-1} a_m \mathrm{e}^{-2\pi i m x} = \mathrm{e}^{-2\pi \mathrm{i} M_0 x} \sum_{m=-M_0}^{M-M_0-1} a_{m+M_0} \mathrm{e}^{-2\pi i m x} = \mathrm{e}^{-2\pi \mathrm{i} M_0 x} \Psi(x),$$

and $\Psi(x) = \sum_{m=-M_0}^{M-M_0-1} a_{m+M_0} \mathrm{e}^{-2\pi i m x}$. Note that Ψ is a trigonometric polynomial of degree at most M_0 and moreover, since $\{\phi(\cdot - k)\}_{k\in\mathbb{Z}}$ is an orthonormal basis, we have $\|\Psi\|^2 = \|f\|^2$. Set $x_n = \omega_n/M$ and $\nu_n = \mu_n/M$, for $n = 1, \dots, N$. Then, by (4.2.14), we have

$$\langle \mathcal{S} f, f \rangle = \sum_{n=1}^{N} \nu_n |\Psi(x_n)|^2 |\hat{\phi}(x_n)|^2. \tag{4.2.15}$$

Now, since $2K/M \in \mathbb{N}$, note that $\mathrm{U}_M \subseteq \mathrm{U}_{2K}$ in this case, and therefore it suffices to prove the result for $M = 2K$. Since $\{\omega_n\}_{n=1}^N$ are (K, δ)-dense, we have that $\{x_n\}_{n=1}^N$ are $(K/M, \delta/M)$-dense. In order to apply Lemma 4.2.11 to $\{x_n\}_{n=1}^N$, we set $x_0 = -x_1 - 2K/M$ and $x_{N+1} = 2K/M - x_N$, so that $a = -K/M = -1/2$ and $b = K/M = 1/2$, and so that $\max_{n=0,\dots,N}\{x_{n+1} - x_n\} = \delta/K$. Therefore, after an application of Lemma 4.2.11, we obtain

$$\langle \mathcal{S} f, f \rangle \geq \min_{n=1,\dots,N} |\hat{\phi}(x_n)|^2 \left(\|\Psi\|_{[a,b]} - \frac{\delta}{K\pi} \|\Psi'\|_{[a,b]} \right)^2 \geq d_0 \left(\|\Psi\|_{[a,b]} - \frac{\delta}{K\pi} \|\Psi'\|_{[a,b]} \right)^2,$$

where $d_0 = \min_{\omega \in [-1/2, 1/2]} |\hat{\phi}(\omega)|^2$. Since $b - a = 1$ and Ψ is periodic, we therefore have

$$\langle \mathcal{S} f, f \rangle \geq d_0 \left(\|\Psi\| - \frac{\delta}{K\pi} \|\Psi'\| \right)^2 \geq d_0 \left(1 - \frac{2\delta M_0}{K} \right)^2 \|\Psi\|^2 \geq d_0 \left(1 - 2\delta\right)^2 \|\Psi\|^2,$$

where we used $\|\Psi'\| \leq 2M_0\pi\|\Psi\|$ and $M_0 \leq K$. Finally, we note that $|\hat{\phi}(\omega)| = |\mathrm{sinc}(\omega\pi)|$ and that

$$|\mathrm{sinc}(\omega\pi)| \geq |\mathrm{sinc}(\pi/2)| = 2/\pi, \quad \omega \in [-1/2, 1/2],$$

which completes the proof.

□

4.3 Necessity of the linear scaling of K and dim(T)

Having shown that stable reconstruction is possible provided the bandwidth K scales linearly with the dimension $\dim(\mathrm{T}) = 2^R$ of the wavelet reconstruction space, we now consider the threshold of this scaling:

Theorem 4.3.1. *Let $\Omega_N = \{\omega_n : n = 1, \ldots, N\} \subseteq [-K, K]$ for some $K > 0$ and suppose that $\mathcal{S}$ is given by (3.1.9) with weights (3.2.1). Let* T *be the reconstruction space of dimension 2^R corresponding to either periodic, folded or boundary-corrected wavelets, where $2^{R-1} > K$. Then*

$$C_1(\Omega_N, \mathrm{T})^{-\frac{1}{2}} \geq \frac{c_1 \exp\left(c_2(1-z)2^R\right)}{\sqrt{K}},$$

where $z = \max\{1/2, K/2^{R-1}\} < 1$ and $c_1, c_2 > 0$ depend only on ϕ.

The constant $C_1(\Omega_N, \mathrm{T})^{-1/2}$ indicates stability of the NUGS reconstruction. Namely, recall that for the condition number of the NUGS mapping F we have

$$\kappa(F) = C_1(\Omega_N, \mathrm{T})^{-\frac{1}{2}},$$

when $\kappa(F)$ is defined as in (3.1.11). Moreover, a result in [AHP13] shows that constant $C_1(\Omega_N, \mathrm{T})^{-1/2}$ is essentially universal. Specifically, any reconstruction algorithm that is so-called *perfect* [AHP13, Def. 3.9] must have a condition number that is at least $C_1(\Omega_N, \mathrm{T})^{-1/2}$. In particular, noting Theorem 4.3.1, we see that to recover wavelet coefficients up to scale R stably, it is necessary to take samples from a bandwidth K that is at least 2^{R-1}, regardless of the method used.

Specifically, Theorem 4.3.1, which generalizes a result proved in [AHP14] to the case of nonuniform samples, establishes the following. Suppose that the size $M = 2^R$ of the reconstruction space is roughly $2\alpha K$. If $\alpha > 1$ then the condition number $C_1(\Omega_N, \mathrm{T})^{-1/2}$ blows up exponentially fast as $M \to \infty$. In other words, if the bandwidth K of the sampling is not sufficiently large in comparison to the wavelet scale R, then ill-conditioning is necessarily witnessed in the reconstruction. Therefore, stable recovery requires bandwidth K which is at least $M/2$. Note that this theorem does not assume density of the samples, just that their maximal bandwidth is K. In particular, even if $\hat{f}(\omega)$ were known for arbitrary $|\omega| \leq K$ one would still have the same result, i.e. insufficient sampling bandwidth implies ill-conditioning.

It is instructive to compare this result with Theorem 4.2.5, which estimates the reconstruction constant for Haar wavelets. If $M \approx 2\alpha K$ then Theorem 4.2.5 demonstrates that $C(\Omega_N, \mathrm{T})$ is bounded whenever α is less than or equal to the critical value $\alpha_0 = 1$. Conversely, if $\alpha > \alpha_0$ then exponential ill-conditioning necessarily results as a consequence of Theorem 4.3.1. For other wavelets, Theorems 4.2.1 and 4.2.2 show that stable reconstruction is possible for sufficiently small scaling α, but unlike the Haar wavelet case, they do not establish the exact value for α_0 that delineates the stability and instability regions.

Theorem 4.3.1 follows immediately from the following lemma:

Lemma 4.3.2. *Let Ω_N and $\mathcal{S}$ be as in Theorem 4.3.1. Let $\mathrm{T} \subseteq \mathrm{H}$ and suppose that $\mathrm{T} \supseteq \mathrm{U}$, where*

$$\mathrm{U} = \operatorname{span}\left\{\sqrt{M}\phi(M\cdot - m) : m = M_1, \ldots, M_2\right\}.$$

for some $M \in \mathbb{N}$, $M_1, M_2 \in \mathbb{Z}$ and $M > 2K$. If $\{\phi(\cdot - k)\}_{k\in\mathbb{Z}}$ is a Riesz basis for its span with bounds d_1 and d_2 then

$$C_1(\Omega_N, \mathrm{T})^{-1/2} \geq \sqrt{\frac{d_1}{d_2}} \frac{\exp\left[c(M_2 - M_1 - 2)(1-z)\right]}{\sqrt{2K+1}},$$

where $z = \max\{1/2, 2K/M\}$, and $c > 0$ depends only on ϕ.

Proof of Theorem 4.3.1. In each case, we merely set $\mathrm{U} = \mathrm{T}^i$ to be the space spanned by the interior wavelets. The result follows immediately from Lemma 4.3.2. □

To prove Lemma 4.3.2, we require the following result (see [AHP14, Prop. 6.2] for a proof):

Lemma 4.3.3. *Let $P \in \mathbb{N}$ and $z \in (0, 1/2)$. Then there exists a constant $c > 0$ independent of P and z such that, if $z' = \max\{1/4, z\}$, then*

$$\sup\left\{\frac{\sup_{|t|\leq 1/2}|\Psi(t)|}{\sup_{|t|\leq z}|\Psi(t)|} : \Psi(t) = \sum_{|n|\leq P} a_k \mathrm{e}^{\mathrm{i}2\pi kt},\ a_k \in \mathbb{C}\right\} \geq \exp\left(cP(1/2 - z')\right).$$

Proof of Lemma 4.3.2. Note that $C_1(\Omega_N, \mathrm{T}) \leq C_1(\Omega_N, \mathrm{U})$. Let $f \in \mathrm{U}$. Then

$$\langle \mathcal{S}f, f\rangle = \frac{1}{M}\sum_{n=1}^{N} \mu_n |\hat{\phi}(\omega_n/M)|^2 |\Psi(\omega_n/M)|^2,$$

where $\Psi(x) = \sum_{k=M_1}^{M_2} a_k \mathrm{e}^{-2\pi\mathrm{i}kx}$ satisfies $d_1\|\Psi\|^2 \leq \|f\|^2 \leq d_2\|\Psi\|^2$. Thus

$$\begin{aligned}
\langle \mathcal{S}f, f\rangle &\leq \sup_{|\omega|\leq K/M} |\hat{\phi}(\omega)|^2 \sup_{|t|\leq K/M} |\Psi(t)|^2 \left(\frac{1}{M}\sum_{n=1}^{N} \frac{\omega_{n+1} - \omega_{n-1}}{2}\right) \\
&= \frac{2K}{M} \sup_{|\omega|\leq K/M} |\hat{\phi}(\omega)|^2 \sup_{|t|\leq K/M} |\Psi(t)|^2 \\
&\leq \frac{2Kd_2}{M} \sup_{|t|\leq K/M} |\Psi(t)|^2,
\end{aligned}$$

where the final inequality follows from (4.1.1). The definition (3.1.2) of $C_1(\Omega_N, \mathrm{U})$, now gives

$$C_1(\Omega_N, \mathrm{U}) \leq \frac{2Kd_2}{Md_1} \inf_{\Psi \in \mathrm{V}} \left\{ \frac{\sup_{|t| \leq K/M} |\Psi(t)|^2}{\|\Psi\|^2} \right\},$$

where

$$\mathrm{V} = \left\{ \sum_{k=M_1-M_3}^{M_2-M_3} a_k \mathrm{e}^{2\pi \mathrm{i} kx} : a_k \in \mathbb{C} \right\}, \ M_3 = \left\lceil \frac{M_1 + M_2}{2} \right\rceil.$$

Since $M_2 - M_1 \leq M$ we have $|\Psi(t)|^2 \leq (M+1)\|\Psi\|^2$, and therefore

$$C_1(\Omega_N, \mathrm{T}) \leq \frac{d_2}{d_1}(2K+1) \inf_{\Psi \in \mathrm{V}} \left\{ \frac{\sup_{|t| \leq K/M} |\Psi(t)|^2}{\sup_{|t| \leq 1/2} |\Psi(t)|^2} \right\}. \tag{4.3.1}$$

To complete the proof, we first note that

$$\min\{M_2 - M_3, M_3 - M_1\} \geq (M_2 - M_1 - 1)/2.$$

Thus, V contains all trigonometric polynomials of degree $\lfloor \frac{M_2 - M_1 - 1}{2} \rfloor \geq \frac{M_2 - M_1}{2} - 1$. An application of Lemma 4.3.3 now gives the result. □

4.4 Efficient computation of wavelet coefficients

Recall from Section 3.1.2 that the NUGS reconstruction $\tilde{f}$ from the samples $\{\hat{f}(\omega_n) : n = 1, \ldots, N\}$, in the space T spanned by $\{\varphi_m : m = 1, \ldots, M\}$, can be written as $\tilde{f} = \sum_{m=1}^{M} a_m \varphi_m$ where the coefficients $a = (a_m)_{m=1}^{M}$ is the least-squares solution to the linear system

$$Aa = b, \tag{4.4.1}$$

which can be written as

$$\begin{pmatrix} \sqrt{\mu_1} & & \\ & \ddots & \\ & & \sqrt{\mu_N} \end{pmatrix} \begin{pmatrix} \hat{\varphi}_1(\omega_1) & \ldots & \hat{\varphi}_M(\omega_1) \\ \vdots & \ddots & \vdots \\ \hat{\varphi}_1(\omega_N) & \ldots & \hat{\varphi}_M(\omega_N) \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_N \end{pmatrix} = \begin{pmatrix} \sqrt{\mu_1} & & \\ & \ddots & \\ & & \sqrt{\mu_N} \end{pmatrix} \begin{pmatrix} \hat{f}(\omega_1) \\ \vdots \\ \hat{f}(\omega_N) \end{pmatrix}.$$

As mentioned before, in the general case, solving this system has a computational complexity of $\mathcal{O}(NM)$. In this section we show that in the case of recovering wavelet coefficients, the computational complexity is only $\mathcal{O}(N \log M)$, since in this case the cost of applying matrix A and its adjoint A^* is only $\mathcal{O}(N \log M)$. In fact, for nonuniform sampling where $N = \mathcal{O}(K)$, the computational complexity is simply $\mathcal{O}(M \log M)$, due to the linear correspondences between K and M derived in previous sections.

We describe the computational issues relating to the recovery in the space of boundary-corrected wavelets (4.1.8). NUGS may also be efficiently implemented with wavelets satisfying other boundary conditions such as periodic or symmetric boundary conditions—periodic and folded wavelets—however, here we consider the boundary-corrected wavelets, since such wavelets preserve vanishing moments at the domain boundaries and form unconditional bases on function spaces of certain regularity on bounded domains. Moreover, although we shall only address the reconstruction of coefficients for dimensions $d = 1$ and $d = 2$, the techniques described here can readily be applied to higher dimensional cases. Let us add that, while we mainly focus on the linear recovery model (4.4.1), the same computational aspects analysed here arise in various other nonlinear recovery schemes such as the ℓ_1-minimization schemes introduced in [AH15a, AHPR14, Poo14]. Namely, whenever one wants to recover wavelet coefficients from nonuniform Fourier measurements, one needs fast computations involving the same matrix as the one appearing in (4.4.1), as well as the fast computations involving its adjoint. Hence, the algorithms described here can readily be applied to yield efficient implementations of these other nonlinear recovery schemes.

4.4.1 The one-dimensional case

Let the reconstruction space T be generated by the first 2^R elements of the boundary-corrected wavelets basis defined on the interval $[0, 1]$ as in (4.1.8). We denote the dimension of T by $M = 2^R$. Note that the support of the corresponding scaling and wavelet functions is contained in $[-p + 1, p]$, for some $p \in \mathbb{N}$, and the finest wavelet scale R is chosen such that $R > \log_2(2p)$.

First of all, let us recall the following. For a function $f \in \mathrm{T}$, for T defined as in (4.1.8), we can write

$$f(x) = \sum_{k=0}^{2^J-1} c_{J,k}\phi^{\text{int}}_{J,k}(x) + \sum_{j=J}^{R-1}\sum_{k=0}^{2^j-1} d_{j,k}\psi^{\text{int}}_{j,k}(x)$$

and also

$$f(x) = \sum_{k=0}^{2^R-1} c_{R,k}\phi^{\text{int}}_{R,k}(x)$$

for some scaling coefficients $c_{j,k}$ and some detail coefficients $d_{j,k}$. Given the scaling coefficients $\{c_{R,k} : k = 0, \dots, 2^R - 1\}$, it is possible compute the scaling coefficients $\{c_{J,k} : k = 0, \dots, 2^J - 1\}$ and detail coefficients $\{d_{j,k} : k = 0, \dots, 2^j - 1,\ j = J, \dots, R-1\}$, and vice versa. This can be done by the discrete boundary-corrected Forward Wavelet Transform (FWT), which we denote by W and by $\boldsymbol{W}$ in two dimensions. The reverse operation is performed by the discrete boundary-corrected Inverse Wavelet Transform (IWT), denoted by W^{-1}, in one, and by $\boldsymbol{W}^{-1}$ in two dimensions.

As explained previously, efficient implementation of NUGS leans on the efficient implementation of the forward and adjoint operations, A and A^*, which we now describe in detail. For this choice of the reconstruction space, given $\alpha \in \mathbb{C}^M$ and $\zeta \in \mathbb{C}^N$, the forward operation can be written as

$$\beta = A(\alpha) = \left(\sqrt{\mu_n} \langle \sum_{k=0}^{M-1} \alpha_m \varphi_k, e_{\omega_n} \rangle \right)_{n=1}^{N}, \tag{4.4.2}$$

and the adjoint operation as

$$\gamma = A^*(\zeta) = \left(\langle \sum_{n=1}^{N} \sqrt{\mu_n} e_{\omega_n} \zeta_n, \varphi_k \rangle \right)_{k=0}^{M-1}. \tag{4.4.3}$$

We describe how these operations can be computed efficiently by using the following operators:

i) For the set of frequencies Ω_N and the corresponding set of weights $\{\mu_n\}_{n=1}^N$, the diagonal weighting operator $V = V_{\Omega_N} : \mathbb{C}^N \to \mathbb{C}^N$ is given by

$$V(\gamma) = (\sqrt{\mu_n}\gamma_n)_{n=1}^N, \qquad \gamma \in \mathbb{C}^N. \tag{4.4.4}$$

ii) For the set of frequencies Ω_N, the operator $F = F_{\Omega_N} : \mathbb{C}^M \to \mathbb{C}^N$ is given by

$$F(\gamma) = \left(\frac{1}{\sqrt{M}} \sum_{k=p}^{M-p-1} \gamma_k e_{\omega_n} \left(-\frac{k}{M} \right) \right)_{n=1}^{N}, \qquad \gamma \in \mathbb{C}^M. \tag{4.4.5}$$

iii) For the set of frequencies Ω_N and the scaling function ϕ, the operator $D = D_{\Omega_N,\phi} : \mathbb{C}^N \to \mathbb{C}^N$ is given by

$$D(\zeta) = \left(\hat{\phi} \left(\frac{\omega_n}{M} \right) \zeta_n \right)_{n=1}^{N}, \qquad \zeta \in \mathbb{C}^N. \tag{4.4.6}$$

For the weighting operator we have $V^* = V$. The adjoint operator of F is $F^* : \mathbb{C}^N \to \mathbb{C}^M$ given by

$$(F^*(\zeta))_k = \begin{cases} \frac{1}{\sqrt{M}} \sum_{n=1}^N \zeta_n e_{\omega_n} \left(\frac{k}{M} \right) & k = p, \dots, M-p-1 \\ 0 & \text{otherwise} \end{cases}, \qquad \zeta \in \mathbb{C}^N \tag{4.4.7}$$

and the adjoint operator of D is $D^* : \mathbb{C}^N \to \mathbb{C}^N$ given by

$$D^*(\zeta) = \left(\overline{\hat{\phi} \left(\frac{\omega_n}{M} \right)} \zeta_n \right)_{n=1}^{N}, \qquad \zeta \in \mathbb{C}^N. \tag{4.4.8}$$

Now we can analyse the operations (4.4.2) and (4.4.3). We first consider the forward operation. Given $\alpha \in \mathbb{C}^M$, the equation $\beta = A(\alpha)$ is equivalent to

$$\beta_n = \sqrt{\mu_n} \sum_{k=0}^{M-1} \tilde{\alpha}_k \langle \phi^{\text{int}}_{R,k}, e_{\omega_n} \rangle, \quad n = 1, \ldots, N,$$

where $\tilde{\alpha} = W^{-1}(\alpha) \in \mathbb{C}^M$ and W^{-1} is discrete IWT. Since the Fourier transform of the internal scaling function $\langle \phi^{\text{int}}_{R,k}, e_\omega \rangle$ can be written as

$$\hat{\phi}^{\text{int}}_{R,k}(\omega) = \frac{1}{\sqrt{M}} \hat{\phi}\left(\frac{\omega}{M}\right) e_\omega \left(-\frac{k}{M}\right), \quad k = p, \ldots, M-p-1,$$

by using the definitions of operators F and D, we get

$$\begin{aligned}\tilde{\beta}_n &= \frac{1}{\sqrt{\mu_n}} \beta_n \\ &= \frac{1}{\sqrt{M}} \sum_{k=0}^{p-1} \tilde{\alpha}_k \hat{\phi}^{\text{left}}_k \left(\frac{\omega_n}{M}\right) + \left(D\left(F(\tilde{\alpha})\right)\right)_n + \frac{1}{\sqrt{M}} \sum_{k=M-p}^{M-1} \tilde{\alpha}_k \hat{\phi}^{\text{right}}_{M-k-1} \left(\frac{\omega_n}{M}\right), \ n = 1, \ldots, N.\end{aligned}$$

Once $\tilde{\beta}$ has been computed, it is left to apply the weighting operator and get $\beta = V(\tilde{\beta})$. For the adjoint operation, to compute $\gamma = A^*(\zeta)$ for given $\zeta \in \mathbb{C}^N$, we first apply the weighting operator and set $\tilde{\zeta} = V(\zeta)$. Then, similarly to the forward operation case, one can check that $\tilde{\gamma} = W^{-1}\gamma$ and ζ are related by the following equations

$$\begin{aligned}\tilde{\gamma}_k &= \frac{1}{\sqrt{M}} \sum_{n=1}^{N} \tilde{\zeta}_n \overline{\hat{\phi}^{\text{left}}_k \left(\frac{\omega_n}{M}\right)}, \ k = 0, \ldots, p-1, \\ \tilde{\gamma}_k &= \frac{1}{\sqrt{M}} \sum_{n=1}^{N} \tilde{\zeta}_m \overline{\hat{\phi}^{\text{right}}_{M-k-1} \left(\frac{\omega_n}{M}\right)}, \ k = M-p, \ldots, M-1,\end{aligned}$$

and

$$\tilde{\gamma}_k = \frac{1}{\sqrt{M}} \sum_{n=1}^{N} \overline{\hat{\phi}\left(\frac{\omega_n}{M}\right)} \tilde{\zeta}_n e_{\omega_n} \left(\frac{k}{M}\right), \quad k = p, \ldots, M-p-1.$$

Note that, by using adjoint operators D^* and F^*, this last part can be written as

$$\tilde{\gamma}_k = \left(F^* \left(D^*(\tilde{\zeta})\right)\right)_k, \quad k = p, \ldots, M-p-1.$$

These computational steps, that we summarize below, lead to the efficient algorithm for forward and adjoint operations, and therefore to to the efficient algorithm for solving the weighted least-squares system (4.4.1).

The one-dimensional algorithm

Precompute the weights $\{\mu_n\}_{n=1}^N$ and pointwise measurements of the Fourier transforms of the three scaling functions:

$$\left(\hat{\phi}\left(\frac{\omega_n}{M}\right)\right)_{n=1}^N, \quad \left(\hat{\phi}_k^{\text{left}}\left(\frac{\omega_n}{M}\right)\right)_{n=1}^N, \quad \left(\hat{\phi}_k^{\text{right}}\left(\frac{\omega_n}{M}\right)\right)_{n=1}^N, \quad k = 0, \dots, p-1.$$

The forward operation: Given $\alpha \in \mathbb{C}^M$, $\beta = A(\alpha)$ can be obtained by applying the following steps.

(i) Compute the scaling coefficients $\tilde{\alpha} = W^{-1}(\alpha)$, where W^{-1} is the one-dimensional discrete boundary-corrected IWT.

(ii) Compute contributions from the boundary scaling functions:

$$\tilde{\beta}^{\text{L}} = \left(\frac{1}{\sqrt{M}} \sum_{k=0}^{p-1} \tilde{\alpha}_k \hat{\phi}_k^{\text{left}}\left(\frac{\omega_n}{M}\right)\right)_{n=1}^N, \quad \tilde{\beta}^{\text{R}} = \left(\frac{1}{\sqrt{M}} \sum_{k=M-p}^{M-1} \tilde{\alpha}_k \hat{\phi}_{M-k-1}^{\text{right}}\left(\frac{\omega_n}{M}\right)\right)_{n=1}^N.$$

(iii) Compute contribution from the internal scaling functions:

(1) Apply F to $\tilde{\alpha}$ to get $\hat{\alpha} = F(\tilde{\alpha})$, where F is defined by (4.4.5).

(2) Apply D to $\hat{\alpha}$ to get $\tilde{\beta}^{\text{I}} = D(\hat{\alpha})$, where D is defined by (4.4.6).

(iv) Compute $\tilde{\beta} = \tilde{\beta}^{\text{L}} + \tilde{\beta}^{\text{R}} + \tilde{\beta}^{\text{I}}$.

(v) Apply V to compute $\beta = V(\tilde{\beta})$, where V is defined by (4.4.4).

The adjoint operation: Given $\zeta \in \mathbb{C}^N$, $\gamma = A^*(\zeta)$ can be computed as follows.

(i) Apply the weighting operator V and set $\tilde{\zeta} = V(\zeta)$.

(ii) Compute the coefficients of the boundary scaling functions:

$$\tilde{\gamma}_k = \frac{1}{\sqrt{M}} \sum_{n=1}^N \tilde{\zeta}_n \overline{\hat{\phi}_k^{\text{left}}\left(\frac{\omega_n}{M}\right)}, \quad \tilde{\gamma}_{M-k-1} = \frac{1}{\sqrt{M}} \sum_{n=1}^N \tilde{\zeta}_n \overline{\hat{\phi}_k^{\text{right}}\left(\frac{\omega_n}{M}\right)},$$

for $k = 0, \dots, p-1$.

(iii) Compute the coefficients of the internal scaling functions:

(1) Compute $\tilde{\zeta}_\phi = D^*(\tilde{\zeta})$, where D^* is defined by (4.4.8).

(2) Compute $\tilde{\gamma}_k = \left(F^*(\tilde{\zeta}_\phi)\right)_k$, $k = p-1, \dots, M-p-1$, where F^* is defined by (4.4.7).

(iv) Compute $\gamma = W(\tilde{\gamma})$, where W is discrete one-dimensional boundary-corrected FWT.

Remark 4.4.1 Regarding the computation of weights $\{\mu_n\}_{n=1}^N$, see Remark 3.1.8.

Remark 4.4.2 The above algorithm requires the precomputation of pointwise evaluations of the Fourier transform of the internal and boundary scaling functions. Note that for Daubechies wavelets, for the internal scaling function ϕ, we may use the approximation

$$\prod_{j=1}^{J} m_0(2^{-j}\xi) \to \hat{\phi}(\xi), \qquad J \to \infty$$

where m_0 is a trigonometric polynomial [Dau92]. A similar approximation may be used in the case of the boundary scaling functions. For more details see appendix in [GP15].

Remark 4.4.3 Recall that in solving (4.4.1) we obtain $a = (a_m)_{m=1}^M$ which is an approximation of the first M wavelet coefficients of f and the reconstructed signal is $\tilde{f} = \sum_{m=1}^M a_m \phi_m$. To evaluate the signal $\tilde{f}$ on the grid points $(j2^{-L})_{j=1}^{2^L}$ for $L \in \mathbb{N}$, it suffices to evaluate each ϕ on these grid points and we may do so by either implementing the cascade algorithm [Dau92] or the dyadic dilation algorithm [LMR97].

Computational cost of the one-dimensional algorithm. Let us analyse the computational cost of the forward operation. The adjoint operation can be analysed similarly leading to the same computational cost. The computational cost of step 1 and the discrete boundary-corrected IWT is $\mathcal{O}(M)$. The cost of step 2, involving boundary scaling functions, is $\mathcal{O}(pN)$. For step 3a, the key point is to observe that F is simply a restricted and shifted version of the discrete nonuniform Fourier transform, and thus its fast implementation NUFFT can be used when computing $F(\tilde{\alpha})$. Hence, the the cost of step 3a is $\mathcal{O}(L\log(M) + JN)$, where L is the length of underlying interpolating FFT for NUFFT, and J is the number of interpolating coefficients (typically $J = 7$) [FS03]. Finally, the cost of the diagonal operations in both steps 3b and 5 is $\mathcal{O}(N)$. Therefore, given that $J \sim p$ and $L \sim N$, the total cost is essentially $\mathcal{O}(pN + N\log(M))$.

4.4.2 The two-dimensional case

For the two-dimensional implementation of the NUGS reconstruction, we use wavelets obtained by applying the tensor product to the one-dimensional boundary-corrected wavelets, thereby defining a basis on $[0,1]^2$. Namely, we introduce the following two-dimensional functions

$$\Phi_{j,(k_1,k_2)}(x_1,x_2) = \phi^{\text{int}}_{j,k_1}(x_1)\phi^{\text{int}}_{j,k_2}(x_2), \quad \Psi^1_{j,(k_1,k_2)}(x_1,x_2) = \phi^{\text{int}}_{j,k_1}(x_1)\psi^{\text{int}}_{j,k_2}(x_2),$$
$$\Psi^2_{j,(k_1,k_2)}(x_1,x_2) = \psi^{\text{int}}_{j,k_1}(x_1)\phi^{\text{int}}_{j,k_2}(x_2), \quad \Psi^3_{j,(k_1,k_2)}(x_1,x_2) = \psi^{\text{int}}_{j,k_1}(x_1)\psi^{\text{int}}_{j,k_2}(x_2).$$

For $J \geq \log_2(2p)$,

$$\mathcal{W}_J^0 = \left\{\Phi_{J,(k_1,k_2)} : 0 \leq k_1, k_2 \leq 2^J - 1\right\}$$

and

$$\mathcal{W}_j^i = \left\{\Psi^i_{j,(k_1,k_2)} : 0 \leq k_1, k_2 \leq 2^j - 1\right\}, \quad j \in \mathbb{N},\ j \geq J,\ i = 1, 2, 3,$$

the set

$$\mathcal{W}_J^0 \cup \left\{ \bigcup_{j \geq J} \{\mathcal{W}_j^i : i = 1, 2, 3\} \right\} \tag{4.4.9}$$

forms a basis for $\mathrm{L}^2([0,1]^2)$. We now order the basis elements of (4.4.9) in increasing order of wavelet scales so that we can write

$$(\varphi_{m_1,m_2})_{m_1,m_2 \in \mathbb{N}} = \begin{pmatrix} \mathcal{W}_J^0 & \mathcal{W}_J^1 & \mathcal{W}_{J+1}^1 & \cdots \\ \mathcal{W}_J^2 & \mathcal{W}_J^3 & & \\ & & & \\ & \mathcal{W}_{J+1}^2 & \mathcal{W}_{J+1}^3 & \\ & \vdots & & \ddots \end{pmatrix}.$$

Let T be the space spanned by the first $M \times M$ wavelets via this ordering, so that

$$\mathrm{T} = \operatorname{span}\{\varphi_{m_1,m_2} : 1 \leq m_1, m_2 \leq M\}.$$

For $M = 2^R$ and $R > J \geq \log_2(2a)$, we have

$$\mathrm{T} = \operatorname{span} \mathcal{W}_J^0 \oplus \left(\oplus_{i=1}^3 \oplus_{j=J}^{R-1} \operatorname{span} \mathcal{W}_j^i\right) = \operatorname{span} \mathcal{W}_R^0, \tag{4.4.10}$$

which is the reconstruction space of dimension M^2 that we consider here. Additionally, for $N \geq M^2$, let $\Omega_N = \left\{\omega_n \,:\, \omega_n = (\omega_n^1, \omega_n^2),\ n = 1, \ldots, N\right\}$ be the set of sampling points in $\mathbb{R}^2$, which we write $\Omega_N = (\Omega_N^1, \Omega_N^2)$ correspondingly. In this case, the least-squares system (4.4.1) becomes

$$\left(\mu_n \langle \sum_{m_1,m_2=1}^{M} a_{m_1,m_2} \varphi_{m_1,m_2}, e_{\omega_n} \rangle \right)_{n=1}^{N} = (\mu_n \langle f, e_{\omega_n} \rangle)_{n=1}^{N}.$$

If we apply the two-dimensional boundary-corrected IWT, denoted by $\boldsymbol{W}^{-1}$, to the matrix of wavelet coefficients $a \in \mathbb{C}^{M\times M}$, so that $\tilde{a} = \boldsymbol{W}^{-1}(a) \in \mathbb{C}^{M\times M}$, we get

$$\left(\mu_n \langle \sum_{k_1,k_2=1}^{M} \tilde{a}_{k_1,k_2} \Phi_{R,(k_1,k_2)}, e_{\omega_n}\rangle\right)_{n=1}^{N} = (\mu_m \langle f, e_{\omega_n}\rangle)_{n=1}^{N}.$$

Since $\Phi_{R,(k_1,k_2)}(x_1,x_2) = \phi^{\text{int}}_{R,k_1}(x_1)\phi^{\text{int}}_{R,k_2}(x_2)$, we can write the following algorithm.

The two-dimensional algorithm

Precompute the vectors $\{\mu_n\}_{n=1}^N$ and

$$\left(\hat{\phi}\left(\frac{\omega_n^i}{M}\right)\right)_{n=1}^{N}, \quad \left(\hat{\phi}^{\text{left}}_k\left(\frac{\omega_n^i}{M}\right)\right)_{n=1}^{N}, \quad \left(\hat{\phi}^{\text{right}}_k\left(\frac{\omega_n^i}{M}\right)\right)_{n=1}^{N}, \quad k = 0,\dots,p-1, \; i = 1,2.$$

The forward operation: Given $\alpha \in \mathbb{C}^{M\times M}$, $\beta = A(\alpha) \in \mathbb{C}^N$ can be obtained by applying the following steps.

(i) Compute the scaling coefficients $\tilde{\alpha} = \boldsymbol{W}^{-1}(\alpha)$, where $\boldsymbol{W}^{-1}$ is the discrete two-dimensional boundary-corrected IWT.

(ii) Compute contributions from the corners (the boundary scaling functions in the both axis):

$$\beta^{\text{LL}} = \left(\frac{1}{M}\sum_{k_1=0}^{p-1}\sum_{k_2=0}^{p-1} \tilde{\alpha}_{k_1,k_2}\hat{\phi}^{\text{left}}_{k_1}\left(\frac{\omega_n^1}{M}\right)\hat{\phi}^{\text{left}}_{k_2}\left(\frac{\omega_n^2}{M}\right)\right)_{n=1}^{N}$$
$$\beta^{\text{LR}} = \left(\frac{1}{M}\sum_{k_1=0}^{p-1}\sum_{k_2=M-p}^{M-1} \tilde{\alpha}_{k_1,k_2}\hat{\phi}^{\text{left}}_{k_1}\left(\frac{\omega_n^1}{M}\right)\hat{\phi}^{\text{right}}_{M-k_2-1}\left(\frac{\omega_n^2}{M}\right)\right)_{n=1}^{N}$$
$$\beta^{\text{RL}} = \left(\frac{1}{M}\sum_{k_1=M-p}^{M-1}\sum_{k_2=0}^{p-1} \tilde{\alpha}_{k_1,k_2}\hat{\phi}^{\text{right}}_{M-k_1-1}\left(\frac{\omega_n^1}{M}\right)\hat{\phi}^{\text{left}}_{k_2}\left(\frac{\omega_n^2}{M}\right)\right)_{n=1}^{N}$$
$$\beta^{\text{RR}} = \left(\frac{1}{M}\sum_{k_1=M-p}^{M-1}\sum_{k_2=M-p}^{M-1} \tilde{\alpha}_{k_1,k_2}\hat{\phi}^{\text{right}}_{M-k_1-1}\left(\frac{\omega_n^1}{M}\right)\hat{\phi}^{\text{right}}_{M-k_2-1}\left(\frac{\omega_n^2}{M}\right)\right)_{n=1}^{N}$$

(iii) Compute contributions from the edges (the boundary scaling functions in only one of the axis):

$$\tilde{\beta}^{\mathrm{LI}} = \frac{1}{\sqrt{M}} \sum_{k_1=0}^{p-1} D_{\Omega_N^1 \phi_{k_1}^{\mathrm{left}}} D_{\Omega_N^2 \phi} F_{\Omega_N^2} \left(\tilde{\alpha}_{k_1,\cdot}\right)$$

$$\tilde{\beta}^{\mathrm{RI}} = \frac{1}{\sqrt{M}} \sum_{k_1=M-p}^{M-1} D_{\Omega_N^1 \phi_{M-k_1-1}^{\mathrm{right}}} D_{\Omega_N^2 \phi} F_{\Omega_N^2} \left(\tilde{\alpha}_{k_1,\cdot}\right)$$

$$\tilde{\beta}^{\mathrm{IL}} = \frac{1}{\sqrt{M}} \sum_{k_2=0}^{p-1} D_{\Omega_N^1 \phi} D_{\Omega_N^2 \phi_{k_2}^{\mathrm{left}}} F_{\Omega_N^1} \left(\tilde{\alpha}_{\cdot,k_2}\right)$$

$$\tilde{\beta}^{\mathrm{IR}} = \frac{1}{\sqrt{M}} \sum_{k_2=0}^{p-1} D_{\Omega_N^1 \phi} D_{\Omega_N^2 \phi_{M-k_2-1}^{\mathrm{right}}} F_{\Omega_N^1} \left(\tilde{\alpha}_{\cdot,k_2}\right)$$

where F and D are defined by (4.4.5) and (4.4.6), respectively.

(iv) Compute contribution from the internal scaling functions:

(1) $\hat{\alpha} = \mathbf{F}_{\Omega_N}(\tilde{\alpha})$, where $\mathbf{F}_{\Omega_N} : \mathbb{C}^{M \times M} \to \mathbb{C}^N$ is such that for each $\gamma \in \mathbb{C}^{M \times M}$

$$\mathbf{F}_{\Omega_N}(\gamma) = \left(\frac{1}{M} \sum_{k_1,k_2=p}^{M-p-1} \gamma_{k_1,k_2} e_{\omega_n} \left(-\frac{(k_1,k_2)}{M} \right) \right)_{n=1}^{N}.$$

(2) $\tilde{\beta}^{\mathrm{II}} = D_{\Omega_N^1,\phi} D_{\Omega_N^2,\phi}(\hat{\alpha})$.

(v) Compute $\tilde{\beta} = \tilde{\beta}^{\mathrm{LL}} + \tilde{\beta}^{\mathrm{LR}} + \tilde{\beta}^{\mathrm{RL}} + \tilde{\beta}^{\mathrm{RR}} + \tilde{\beta}^{\mathrm{LI}} + \tilde{\beta}^{\mathrm{RI}} + \tilde{\beta}^{\mathrm{IL}} + \tilde{\beta}^{\mathrm{IR}} + \tilde{\beta}^{\mathrm{II}}$.

(vi) Apply V to get $\beta = V(\tilde{\beta})$, where V is defined by (4.4.4).

The adjoint operation: Given $\zeta \in \mathbb{C}^N$, $\gamma = A^*(\zeta) \in \mathbb{C}^{M,M}$ can be computed as follows.

(i) Apply the weighting operator V and set $\tilde{\zeta} = V(\zeta)$.

(ii) Compute the scaling coefficients at the corners

$$\tilde{\gamma}_{k_1,k_2} = \frac{1}{M} \sum_{n=1}^{N} \zeta_n \overline{\hat{\phi}_{k_1}^{\mathrm{left}} \left(\frac{\omega_n^1}{M} \right) \hat{\phi}_{k_2}^{\mathrm{left}} \left(\frac{\omega_n^2}{M} \right)},$$

$$\tilde{\gamma}_{k_1,M-p+k_2} = \frac{1}{M} \sum_{n=1}^{N} \zeta_n \overline{\hat{\phi}_{k_1}^{\mathrm{left}} \left(\frac{\omega_n^1}{M} \right) \hat{\phi}_{p-k_2-1}^{\mathrm{right}} \left(\frac{\omega_n^2}{M} \right)}$$

$$\tilde{\gamma}_{M-p+k_1,k_2} = \frac{1}{M} \sum_{n=1}^{N} \zeta_n \overline{\hat{\phi}_{p-k_1-1}^{\mathrm{right}} \left(\frac{\omega_n^1}{M} \right) \hat{\phi}_{k_2}^{\mathrm{left}} \left(\frac{\omega_n^2}{M} \right)},$$

$$\tilde{\gamma}_{M-p+k_1,M-p+k_2} = \frac{1}{M} \sum_{n=1}^{N} \zeta_n \overline{\hat{\phi}_{p-k_1-1}^{\mathrm{right}} \left(\frac{\omega_n^1}{M} \right) \hat{\phi}_{p-k_2-1}^{\mathrm{right}} \left(\frac{\omega_n^2}{M} \right)}.$$

for $k_1, k_2 = 0, \dots, p-1$.

(iii) Compute the scaling coefficients at the edges

$$\tilde{\gamma}_{k_1,k_2} = \frac{1}{\sqrt{M}} \left(\left(F_{\Omega_N^1}\right)^* \left(D_{\Omega_N^1 \phi}\right)^* \left(D_{\Omega_N^2 \phi_{k_2}^{\text{left}}}\right)^* (\tilde{\zeta}) \right)_{k_1},$$
$$\tilde{\gamma}_{k_1,k_2} = \frac{1}{\sqrt{M}} \left(\left(F_{\Omega_N^1}\right)^* \left(D_{\Omega_N^1 \phi}\right)^* \left(D_{\Omega_N^2 \phi_{p-k_2-1}^{\text{right}}}\right)^* (\tilde{\zeta}) \right)_{k_1},$$

for $k_1 = p, \dots, M-p-1,\ k_2 = 0, \dots, p-1$ and

$$\tilde{\gamma}_{k_1,k_2} = \frac{1}{\sqrt{M}} \left(\left(F_{\Omega_N^2}\right)^* \left(D_{\Omega_N^1 \phi_{k_1}^{\text{left}}}\right)^* \left(D_{\Omega_N^2 \phi}\right)^* (\tilde{\zeta}) \right)_{k_2},$$
$$\tilde{\gamma}_{k_1,k_2} = \frac{1}{\sqrt{M}} \left(\left(F_{\Omega_N^2}\right)^* \left(D_{\Omega_N^1 \phi_{p-k_1-1}^{\text{right}}}\right)^* \left(D_{\Omega_N^2 \phi}\right)^* (\tilde{\zeta}) \right)_{k_2},$$

for $k_1 = 0, \dots, p-1,\ k_2 = p, \dots, M-p-1$, where F^* and D^* are defined by (4.4.7) and (4.4.8), respectively.

(iv) Compute the scaling coefficients of the internal wavelets

(1) $\tilde{\zeta}_{\phi,\phi} = \left(D_{\Omega_N^1,\phi}\right)^* \left(D_{\Omega_N^2,\phi}\right)^* (\tilde{\zeta})$.

(2) $\tilde{\gamma}_{k_1,k_2} = \left(\mathbf{F}^* \left(\tilde{\zeta}_{\phi,\phi}\right)\right)_{k_1,k_2}, \quad k_1, k_2 = p-1, \dots, M-p-1.$

(v) Compute $\gamma = \boldsymbol{W}(\tilde{\gamma})$, where $\boldsymbol{W}$ is the discrete two-dimensional boundary-corrected FWT.

Computational cost of the two-dimensional algorithm. Again, let us analyse the computational cost of the forward operation. The cost of step 1 is $\mathcal{O}(M^2)$ and of step 2 is $\mathcal{O}(p^2 N)$. Step 3 has $\mathcal{O}(p(N + L\log(M) + JN))$ computations. The cost of step 4a is basically the cost of the two-dimensional NUFFT, i.e. $\mathcal{O}(L^2 \log M^2 + J^2 N)$. The cost of step 4b as well as step 6 is $\mathcal{O}(N)$. Hence, if we assume $J \sim p$ and $L^2 \sim N$, the total cost is $\mathcal{O}(p^2 N + N \log M^2)$. The same cost holds for the adjoint operation.

4.5 Numerical examples

In this section, we illustrate theory developed so far. In particular, we demonstrate performance of NUGS reconstruction using different wavelets. The code used to generate most of these examples was developed in a collaboration with Clarice Poon and it is available at http://www.damtp.cam.ac.uk/user/mg617/GS_wavelets.zip.

Example 4.5.1 (Sufficiency of the sampling rate) The main result proved in the Section 4.2 is that one requires a linear scaling of the bandwidth K (or the truncation index N) with the dimension of the reconstruction subspace $M = 2^R$ for stable and quasi-optimal reconstruction in wavelet subspaces. We now illustrate this in Table 4.1 for the Haar and Daubechies wavelets of order 4 (DB4) on $[0, 1]$. We use two different sampling schemes: (i) a log sampling scheme defined by (3.4.13) with δ such that $\delta < 1/2$; and (ii) Seip's frame sequence defined as follows: for a given $N \in \mathbb{N}$, define

$$\Omega_N = \{\omega_n\}_{n=-1}^{-N} \cup \{\omega_n\}_{n=1}^{N}, \quad \omega_n = n(1 - |n|^{-1/2}); \tag{4.5.1}$$

in [Sei95b], it is shown that the corresponding infinite set of frequencies $\Omega = \Omega_\infty$ gives rise to a Fourier frame with density $\delta = 1/2$.

Namely, in Table 4.1, for a given reconstruction space, the smallest value of K (or N) is shown such that the reconstruction constant $C(\Omega, \mathrm{T})$ is upper-bounded by a constant, where $C(\Omega, \mathrm{T})$ is estimated by using the results given in §3.1.2. In order to have a well-conditioned and quasi-optimal reconstruction, note that the constant of the required scaling is roughly $1/2$, i.e. K (or N) behaves like $c_0 2^R$ with $c_0 \approx 1/2$. In the case of Haar wavelets, this is due to the explicit estimates of Theorem 4.2.5.

T	Ω	2^R	32	64	128	256	512	1024	T	Ω	2^R	32	64	128	256	512	1024
Haar	Log	K	16	32	64	128	256	512	DB4	Log	K	16	32	64	128	256	512
	Frame	N	20	38	72	139	272	535		Frame	N	20	38	72	139	272	535

Table 4.1: For a given number of reconstruction vectors 2^R, the smallest value of K (or N) is shown such that the reconstruction constant $C(\Omega, \mathrm{T})$ is at most 100. This is done for different reconstruction spaces T—Haar and DB4—and for different sampling schemes Ω: Seip's frame sequence and log sampling scheme, the later one with $\delta = 0.475$ and $\nu = 0.33$.

Example 4.5.2 (Necessity of the sampling rate) Theorem 4.3.1 provides a lower estimate for robust scaling of the sampling bandwidth K (or the truncation index N) with the dimension of the reconstruction subspace $M = 2^R$. In particular, if the scaling c_0 is less than $1/2$ then exponential instability necessarily results in the reconstruction, regardless of the wavelet basis used. This is shown in Table 4.2 for both Haar and DB4 wavelets. Note also that in the unstable regime, i.e. $c_0 < 1/2$, the reconstruction $\tilde{f}$ is also far from quasi-optimal. This is demonstrated by plotting $\|f - \tilde{f}\| / \|f - \mathcal{P}_{\mathrm{T}} f\|$ for function $f(x) = 1/2 \cos(4\pi x)$.

Example 4.5.3 (Convergence rates for boundary-corrected wavelets) High convergence rates given by Corollary 4.2.4 are depicted in Figure 4.1. Namely, using an example of a continuous, nonperiodic function $f(x) = x \cos(3\pi x) \chi_{[0,1]}(x)$, we compare the convergence rates of NUGS with boundary-corrected Daubechies wavelets to the sub-optimal convergence rates of the simple direct approaches based on the discretization of the Fourier integral called gridding [JMNM91, SN00, VGCR10, GS14].

T	c_0 / K	0.3125 / 20	0.3750 / 24	0.4375 / 28	0.5000 / 32	0.5625 / 36	0.6250 / 40
Haar	cond(A)	5.8569e15	2.9255e12	1.8347e05	1.7835	1.6474	1.5768
	$\frac{\Vert f-\tilde{f}\Vert}{\Vert f-\mathcal{P}_{\mathrm{T}} f\Vert}$	8.6294e04	7.3412e04	14.4886	1.0016	1.0016	1.0016
DB4	cond(A)	5.0079e15	2.6583e12	1.2918e05	1.6126	1.4744	1.4355
	$\frac{\Vert f-\tilde{f}\Vert}{\Vert f-\mathcal{P}_{\mathrm{T}} f\Vert}$	4.0459e06	3.2764e06	303.3421	1.0013	1.0009	1.0008

Table 4.2: The condition number cond(A) and the error $\|f-\tilde{f}\|/\|f-\mathcal{P}_{\mathrm{T}} f\|$ are shown for different bandwidths $K = c_0 2^R$ and different reconstruction spaces: Haar and DB4 wavelets, where $2^R = 64$ is taken. The jittered sampling scheme is used for $\epsilon = 0.6$ and $\eta = 0.15$.

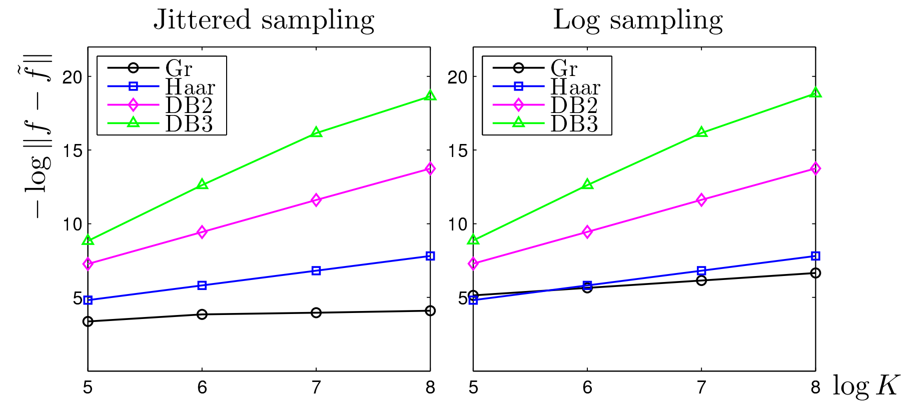

Figure 4.1: A nonperiodic continuous function $f(x) = x\cos(3\pi x)\chi_{[0,1]}(x)$ is reconstructed from pointwise samples of its Fourier transform taken on a jittered scheme with jitter 0.1 (left) and on a log scheme (right), where $\delta < 0.97$. Reconstruction is performed via NUGS using different types of boundary-corrected Daubechies wavelets: Haar, DB2 and DB3, and also via gridding.

Example 4.5.4 (Explicit estimates for Haar wavelets) Table 4.3 considers the case of Haar wavelet reconstructions more closely for the three different sampling schemes: jittered (3.4.5), log (3.4.13) and Seip's frame (4.5.1), and in particular, the magnitude of the reconstruction constant $C(\Omega, \mathrm{T})$ is considered. Recall that both the quasi-optimality constant μ and the condition number κ of NUGS are upper-bounded by $C(\Omega, \mathrm{T})$. The table suggests that this estimate is reasonably sharp. Recall the technique from §3.1.2 that $C(\Omega, \mathrm{T})$ can be approximated by a limiting process. The result of this is also shown in the table. Moreover, in the (K, δ)-dense case, we see that the estimate $C(\Omega, \mathrm{T}) \leq (1+\delta)/\sqrt{C_1(\Omega, \mathrm{T})}$ is also adequate. Finally, the table also shows that the explicit bound derived in Theorem 4.2.5 is also reasonably good.

Example 4.5.5 (Boundary vs periodic wavelets) We now wish to exhibit the advantage of NUGS: namely, it allows one to reconstruct in a subspace T that is well suited to the function to be recovered. In Figures 4.2 and 4.3 we consider the reconstruction of two functions using different wavelets from exactly the same set of measurements. The

Ω	K	$\lvert\Omega\rvert$	2^R	$\lVert f-\tilde{f}\rVert$	$\lVert f-\mathcal{P}_{\mathrm{T}}f\rVert$	$\frac{\lVert f-\tilde{f}\rVert}{\lVert f-\mathcal{P}_{\mathrm{T}}f\rVert}$	cond(A)	$\frac{\sigma_{\max}(A_{4096})}{\sigma_{\min}(A)}$	$\frac{1+\delta}{\sigma_{\min}(A)}$	$\frac{\pi}{2}\frac{1+\delta}{1-\delta}$
Jittered	32	108	64	6.1080e-2	6.0863e-2	1.003575	1.5507	3.7227	4.7892	14.1372
	64	215	128	3.0491e-2	3.0463e-2	1.000914	1.5687	3.8400	4.9401	
	128	428	256	1.5239e-2	1.5235e-2	1.000238	1.5960	3.9150	5.0365	
	256	855	512	7.6189e-3	7.6184e-3	1.000065	1.5916	4.1577	5.3488	
Log	32	350	64	6.1080e-2	6.0863e-2	1.003567	1.6591	3.4151	4.3935	14.1372
	64	814	128	3.0491e-2	3.0463e-2	1.000912	1.6825	3.4681	4.4616	
	128	1850	256	1.5239e-2	1.5236e-2	1.000237	1.6946	3.4899	4.4897	
	256	4146	512	7.6189e-3	7.6184e-3	1.000064	1.7007	3.5041	4.5079	
Frame	32	76	64	6.1080e-2	6.0863e-2	1.003568	2.5674	3.4455	×	×
	64	144	128	3.0492e-2	3.0463e-2	1.000932	2.5203	3.3188		
	128	278	256	1.5241e-2	1.5236e-2	1.000313	2.6211	3.5886		
	256	544	512	7.6189e-3	7.6184e-3	1.000067	2.5531	3.4046		

Table 4.3: The function $f(x) = \cos(6\pi x) + 1/2\sin(2\pi x)$ is reconstructed by NUGS with Haar wavelets for different sampling schemes Ω and different bandwidths K. Jittered sampling scheme is used for $\epsilon = 0.6$ and $\eta = 0.1$; and log sampling scheme is used for $\delta = 0.4$ and $\nu = 0.4$. In the last three columns, different estimates for the reconstruction constant are computed, by using the results from §3.1.2 and §4.2.2.

function from Figure 4.2 is periodic, hence we use periodic wavelets, and the function from Figure 4.3 is nonperiodic, and therefore we use boundary-corrected wavelets. Note that an inferior reconstruction is obtained if periodic wavelets are used for a nonperiodic function. Also, as is again evident, increasing the wavelet smoothness leads to a smaller error. This is due to the property of this approach described in Corollary 4.2.4: namely, since NUGS is quasi-optimal and since it requires only a linear scaling for wavelet bases, it obtains optimal approximation rates in terms of the sampling bandwidth.

Example 4.5.6 (Robustness of NUGS) Next we consider the effect of noise. In Table 4.4, the actual error $\|f - F(f + \eta h)\|$ when reconstructing a function f from noisy measurements, where the noise is described by the term ηh, is compared to the estimate $\tilde{C}(\Omega, \mathrm{T})\left(\|f - \mathcal{P}_{\mathrm{T}} f\| + \eta\|h\|\right)$, where $\tilde{C}(\Omega, \mathrm{T})$ is an approximation of the reconstruction constant computed by the techniques from §3.1.2. Note that the bound is reasonably close to the true value. We also note the robustness of the reconstruction with respect to noise level embedded in parameter η. This is further illustrated in Figure 4.4, where we plot the reconstruction of a function f from noisy measurements. Even in the presence of large noise with $\eta = 0.1$, we obtain a good approximation.

Example 4.5.7 (Numerical comparison) As mentioned, two common algorithms for MRI reconstruction are gridding [JMNM91, SN00, VGCR10] and iterative reconstructions

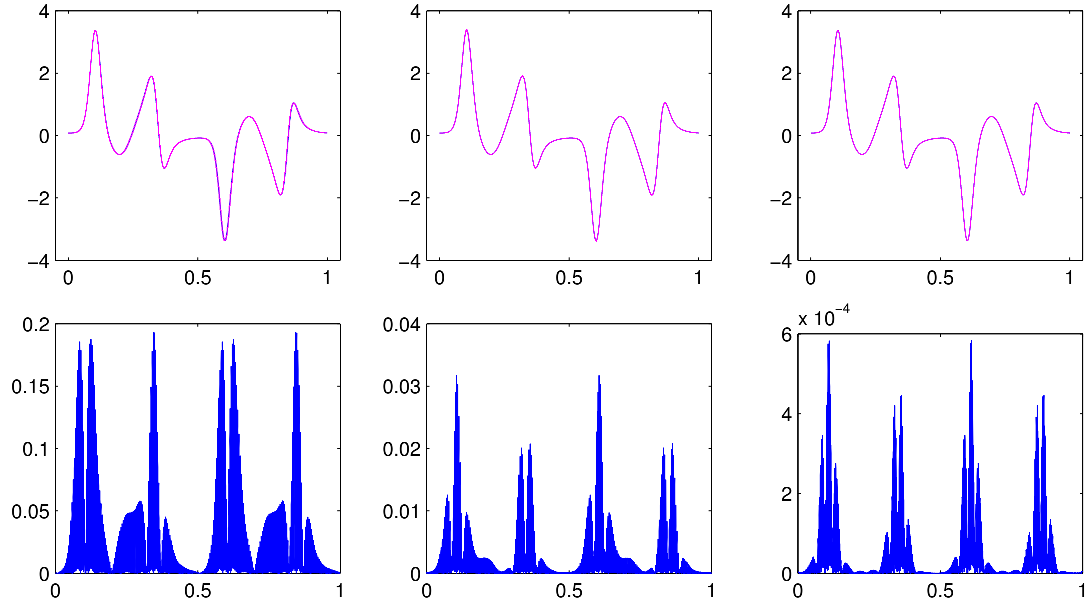

Figure 4.2: A smooth, periodic function reconstructed by $2^R = 256$ Haar, periodic DB2 and periodic DB4 wavelets, from left to right. Above is the reconstruction $\tilde{f}$ (magenta) and the original function f (blue), and below is the error $|f - \tilde{f}|$. In all experiments, the same jittered sampling scheme is used, with $K = 128$.

[SNF03]. We now compare these approaches with NUGS. Recall, however, that iterative reconstruction algorithm can be interpreted as a particular instance of NUGS corresponding to a Haar wavelet basis for T (see Remark 4.2.6). We therefore continue to refer to it as such in our numerics.

Gridding is a simple technique for MRI reconstruction. It is direct, as opposed to iterative, and can be computed with a single NUFFT. Unfortunately, this reconstruction is plagued by artefacts, even when the original function is periodic. This is shown in the left panels of Figures 4.5 and 4.6. Alternatively, one can use the NUGS reconstruction with wavelets. As shown in these figures, this gives a far superior reconstruction of f, even in the case of discontinuous functions with sharp peaks (see Figure 4.6). Recall also that the NUGS reconstruction can also be computed efficiently using NUFFTs (see Remark 3.1.9). Hence, using the same measurement data, and with roughly the same computational cost, we obtain a vastly improved reconstruction.

Figures 4.2–4.6 also show the clear advantage of changing the NUGS reconstruction space T from Haar wavelets (i.e. the iterative reconstructions) to higher-order wavelets. This improvement is justified by Corollary 4.2.4, following the discussion in Remark 4.1.3.

Remark 4.5.8 The reason why NUGS obtains an improvement by changing T is that it computes quasi-optimal approximations to the actual wavelet coefficients of f. In particular, it avoids the wavelet crime [SN96]. Let a^* be the vector of first M coefficients in

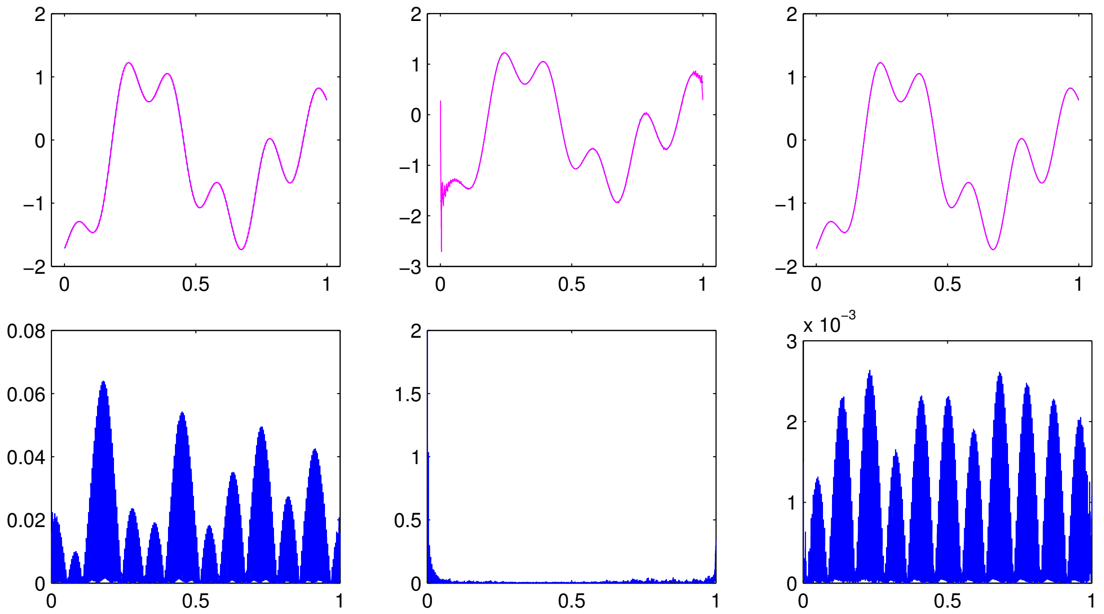

Figure 4.3: A smooth, nonperiodic function reconstructed by $2^R = 256$ Haar, periodic DB2 and boundary DB2, from left to right. Above is the reconstruction $\tilde{f}$ (magenta) and the original function f (blue), and below is the error $|f - \tilde{f}|$. In all experiments, the same jittered sampling scheme is used with $K = 128$.

some wavelet basis. NUGS solves the least-squares problem $Aa \approx b$, where A is the matrix of Fourier samples of wavelet basis functions and b is the vector of nonuniform Fourier samples of f (see (3.1.13)). The error estimates proved show that $\|a - a^*\| \equiv \|\tilde{f} - f\|$ is proportional to the best approximation error $\|f - \mathcal{P}_{\mathrm{T}} f\|$ of f in the wavelet subspace T.

As an alternative, to compute wavelet coefficients one may be tempted to construct the matrix $\tilde{A} = FW$, where $F \in \mathbb{C}^{N\times M}$ is the nonuniform discrete Fourier transform and $W \in \mathbb{C}^{M\times M}$ is the discrete wavelet transform, and solve the least-squares problem $\tilde{A}a \approx b$. Since W is orthogonal, this is equivalent to solving $Fc \approx b$ and then setting $a = W^T c$. However, c is a vector of pixel values of f, and is therefore equivalent to the solution of the iterative reconstruction algorithm (recall §4.2.2). Since W is orthogonal, we have $\|a - a^*\| = \|c - c^*\|$, where c is the vector of exact coefficients of f in the pixel basis. Thus, the accuracy of the computed wavelet coefficients $a = W^T c$ is not determined by how well f is approximated in the given wavelet basis, but how well f is approximated by a piecewise constant function. This accuracy is typically low, which means that one will not see the benefits of higher-order wavelets with this approach. In particular, the higher approximation orders—that is, faster decay of $\|f - \mathcal{P}_{\mathrm{T}} f\|$—offered by boundary-corrected wavelets (see Remark 4.1.3).

Example 4.5.9 (A two-dimensional example) In this example we reconstruct a two-dimensional function shown in Figure 4.7, which is continuous but nonperiodic. We use

T	η	error	estimate	T	η	error	estimate	T	η	error	estimate
Haar	0	4.48e-2	9.48e-2	DB2p	0	3.09e-3	6.55e-3	DB2b	0	4.70e-3	9.69-3
	0.05	6.66e-2	2.01e-1		0.05	4.92e-2	1.13e-1		0.05	6.97e-2	1.15e-1
	0.1	1.08e-1	3.06e-1		0.1	9.81e-2	2.19e-1		0.1	1.39e-1	2.21e-1
	0.2	2.02e-1	5.18e-1		0.2	1.96e-1	4.31e-1		0.2	2.78e-1	4.31e-1
	0.4	3.97e-1	9.42e-1		0.4	3.92e-1	8.52e-1		0.4	5.56e-1	8.54e-1

Table 4.4: The actual error $\|f-F(f+\eta h)\|$ and the error estimate $\tilde{C}(\Omega,\mathrm{T})\,(\|f-\mathcal{P}_{\mathrm{T}}f\|+\eta\|h\|)$ are computed for $f(x)=(\cos(8\pi x)-2\sin(2\pi x))\chi_{[0,1]}(x)$ and $h(x)=\sin(10\pi x)\chi_{[0,1]}(x)/\|\sin(10\pi x)\|$, where $\tilde{C}(\Omega,\mathrm{T})=C_3(\Omega,\mathrm{T}_{4096})/C_1(\Omega,\mathrm{T})$ (see the Section §3.1.2), and Ω is the log sampling scheme with $K=128$, $\delta=0.475$, $\nu=0.33$ and $N=1512$. The computation is done for different reconstruction spaces T with $2^R=128$ Haar, periodic DB2 and boundary-corrected DB2 functions.

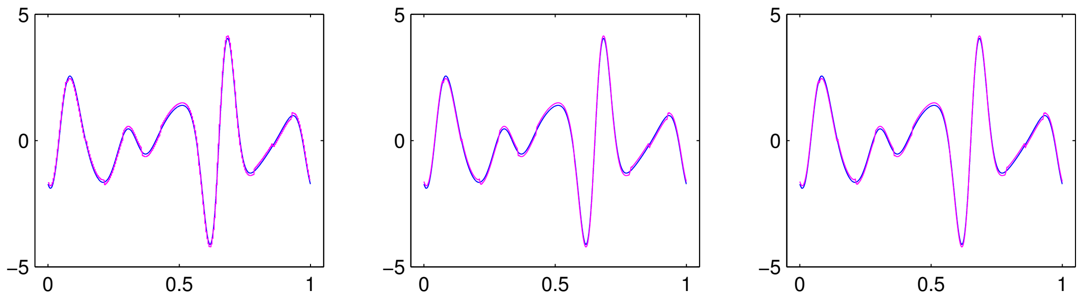

Figure 4.4: The function $f(x)=(-\exp((\cos(6\pi x))+\sin(4\pi x))\cos(10\pi x)+\cos(4\pi x))\chi_{[0,1]}(x)$ (blue) and the reconstruction $F(f+\eta h)$ (magenta), where $h(x)=\mathrm{sinc}(14\pi(x-0.5))\chi_{[0,1]}/\|\mathrm{sinc}(14\pi(x-0.5))\|$ and $\eta=0.1$. The log sampling scheme is used for $\delta=0.475$, $\nu=0.33$, $K=256$ and $N=3398$. From left to right different reconstruction basis are used: $2^R=256$ Haar, periodic DB3 and boundary DB3.

radial sampling scheme, which gains an accumulation point as the sampling bandwidth increases and which is taken in the Euclidean ball of radius $K=64$ with density $\delta_{\ell^1}<1/4$. This sampling scheme is constructed as in §3.4.2.

First, we demonstrate the use of weights when reconstructing from nonuniform Fourier measurements. Some of the advantages of using weights have been already reported earlier in the literature, see for example [FG94, FGS95, GS01] and also [JMNM91, SNF03]. In a different setting, in Figure 4.8, we provide further insight on the necessity of using weights. To this end, we perform function recovery using NUGS with boundary-corrected Daubechies wavelets of order 1, 2 and 3, as well as the direct recovery approach called gridding [JMNM91]. We perform function recovery with and without using weights, using the same number of NUGS iterations. As shown in Figure 4.8, the reconstruction error without using weights does not exceed order 10^{-2}. Hence, the advantages of higher order wavelets cannot be easily exploited in this case, as opposed to the case when reconstructing

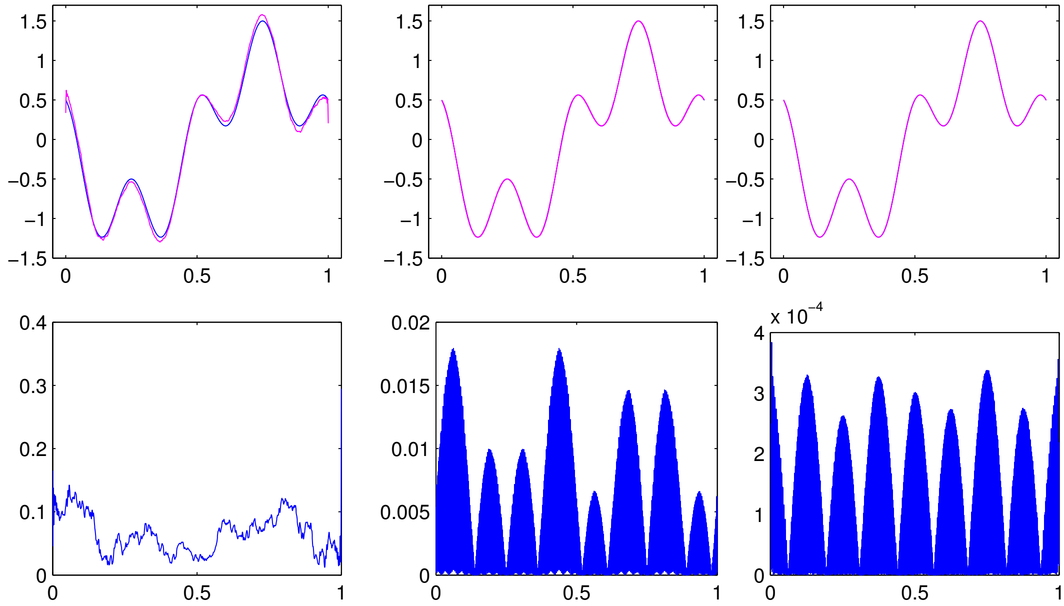

Figure 4.5: A periodic function $f(x) = (1/2\cos(8\pi x) - \sin(2\pi x))\chi_{[0,1]}(x)$ is reconstructed by gridding (left) and NUGS with Haar (middle) and DB2 (right) wavelets for $2^R = 512$. The lower pictures show the error $|f - \tilde{f}|$. The jittered sampling scheme is used for $\epsilon = 0.7$, $\eta = 0.14$ and $K = 256$.

with weights. Moreover, the gridding reconstruction obtained without using weights is distinctly inferior. Recall that gridding reconstruction is computed with only one iteration, i.e. with a single use of NUFFT.

Additionally, using the same example, in Figure 4.9, we demonstrate robustness of NUGS when white Gaussian noise is added to the Fourier samples.

Example 4.5.10 (Violation of the density condition) In our final example, in Figure 4.10, we examine how violation of the density condition $\delta_{D^\circ} < 1/4$ given in Theorems 2.3.1 and 3.3.3 influences reconstruction of a a high resolution test image with $D = [-1,1]^2$. We use radial sampling schemes with different number of radial lines n. Recall that the density condition from Theorem 2.3.1 is only sufficient, but not necessary to have a weighted Fourier frame, and that it is sharp in the sense that there exist a set of sampling points with $\delta_{D^\circ} = 1/4$ and a function which violate the frame condition. Yet for a fixed function and set of sampling points, a slight violation of the density condition may not worsen the recovery guaranteed by the II part of Theorem 3.3.2. As evident in the presented example from Figure 4.10, a slight violation of $\delta_{D^\circ} < 1/4$ does not impair the recovery noticeably therein. However, it is evident that further decreasing of number of radial lines n, i.e decreasing of sampling density, worsens the quality of the reconstructed image. Also, as illustrated in Table 4.5, this decreasing of sampling density, i.e. increasing of δ, causes blowing up of the condition number associated to (3.1.10).

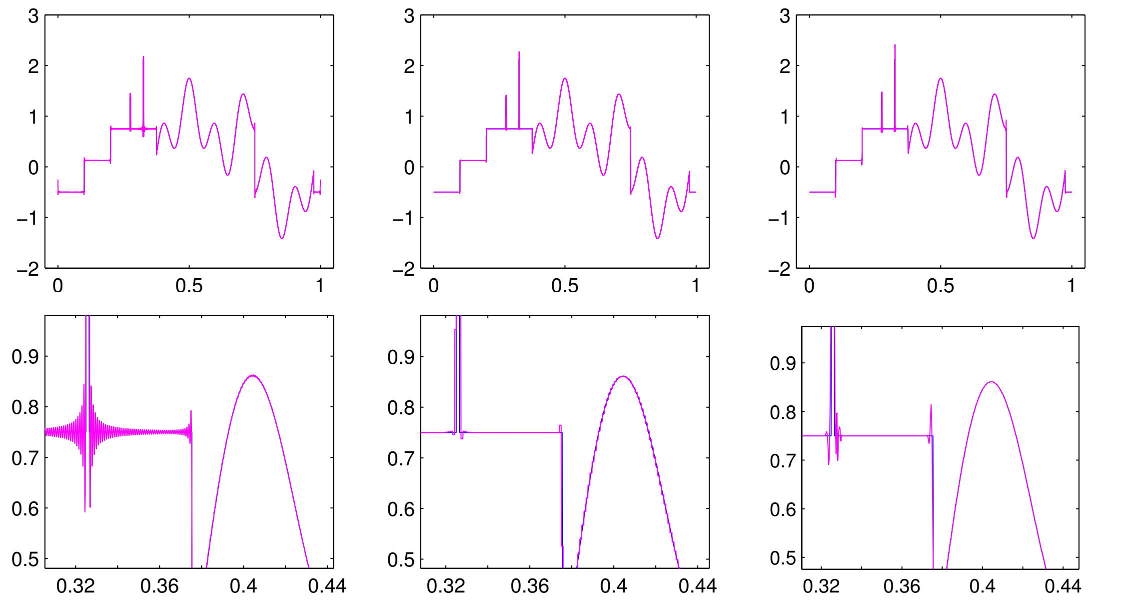

Figure 4.6: A discontinuous function reconstructed by gridding, and NUGS with Haar and DB4 wavelets (from left to right). The reconstruction is in magenta and original in blue. Below, a close-up is shown. The jittered sampling is used for $\epsilon = 0.75$, $\eta = 0.1$ and $K = 2^R = 1024$.

n	345	173	87	44	22	11
δ_2	0.1763	0.3064	0.5847	1.1437	2.2843	4.5547
cond(A)	1.6220	2.3821	1.4859×10^3	9.2459×10^{14}	5.3376×10^{16}	5.4891×10^{18}

Table 4.5: The condition number cond(A) of a reconstruction matrix arising from the least-squares system (3.1.10) is calculated when 88×88 indicator functions are used and samples are acquired on a radial sampling scheme contained in $[-K, K]^2$, $K = 32$, so that $\dim(\mathrm{T}) = (2.75K)^2$. The number of radial lines n of the radial scheme is varying, as well as the corresponding sampling density δ_2, which is measured with respect to the Euclidean norm.

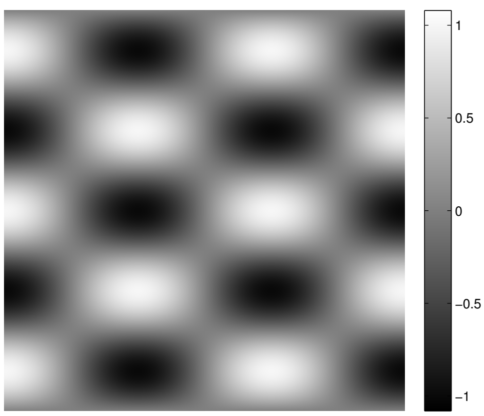

Figure 4.7: Function $f(x, y) = \sin(5/2\pi(x+1))\cos(3/2\pi(y+1))\chi_{[-1,1]^2(x,y)}$, ploted with the scale $[-1.08, 1.08]$ shown on the right.

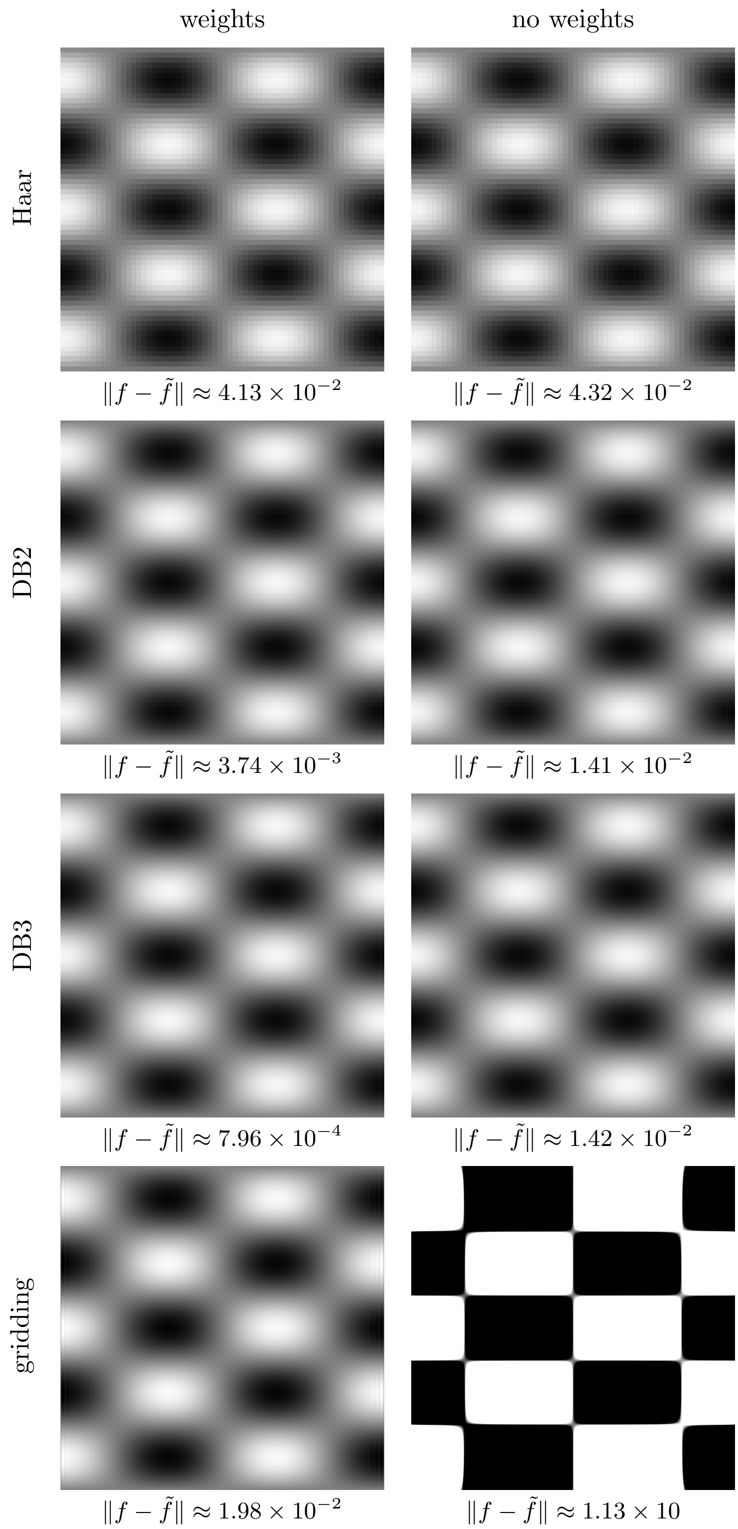

Figure 4.8: Reconstructions of the function from Figure 4.7, using the same scale $[-1.08, 1.08]$, from Fourier samples taken on the radial sampling scheme in the Euclidean ball of radius $K = 64$ with the density measured in ℓ^1-norm strictly less than $1/4$. The lower pictures are reconstructed without using weights and, as demonstrated, the error does not exceed order 10^{-2}. The NUGS reconstruction with with 64×64 Haar, DB2 and DB3 are also compared to gridding.

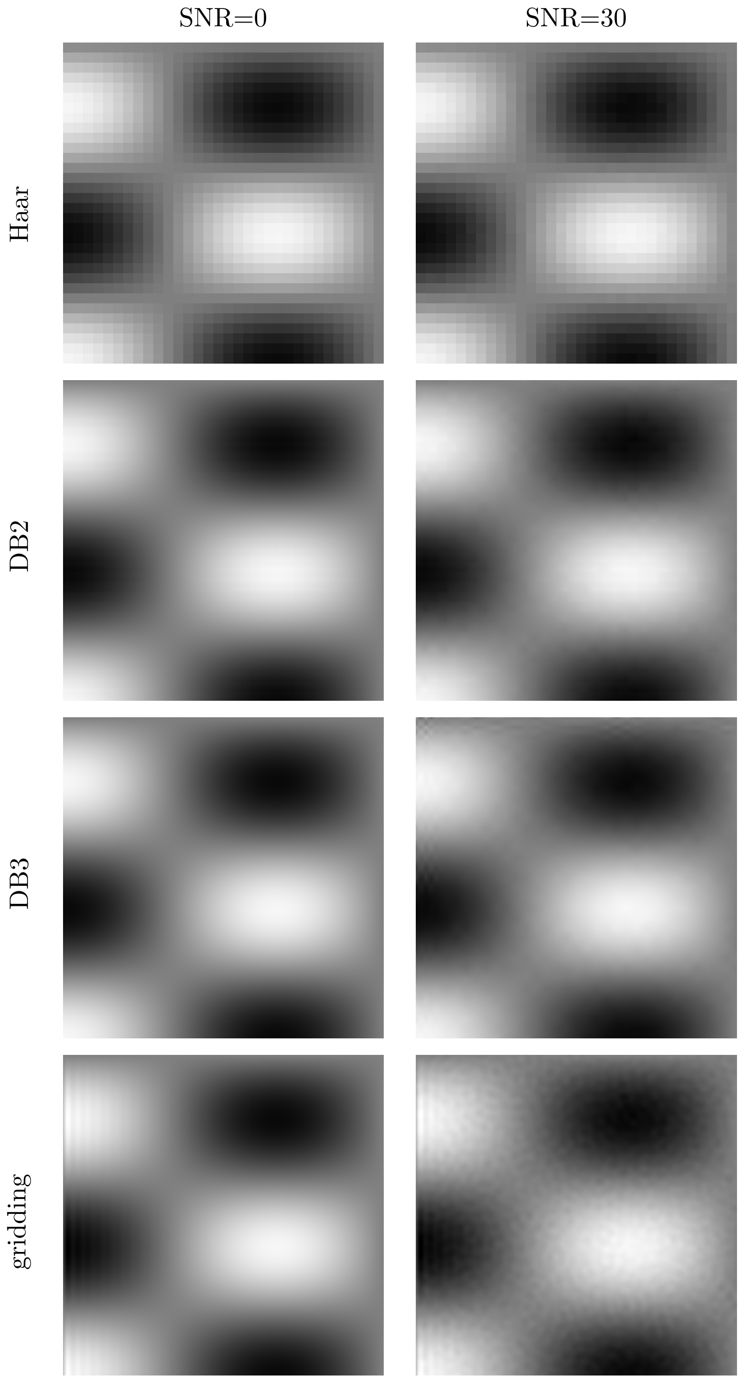

Figure 4.9: Top-left corner close-ups of the reconstructed function of f in Figure 4.7 using the same setting as in Figure 4.8. In the bottom row, the white Gaussian noise with SNR of 30dB is added to the samples. The L_2-error of bottom, noisy reconstructions is (from right to left): 4.28×10^{-2}, 1.07×10^{-2}, 1.08×10^{-2} and 2.93×10^{-2}.

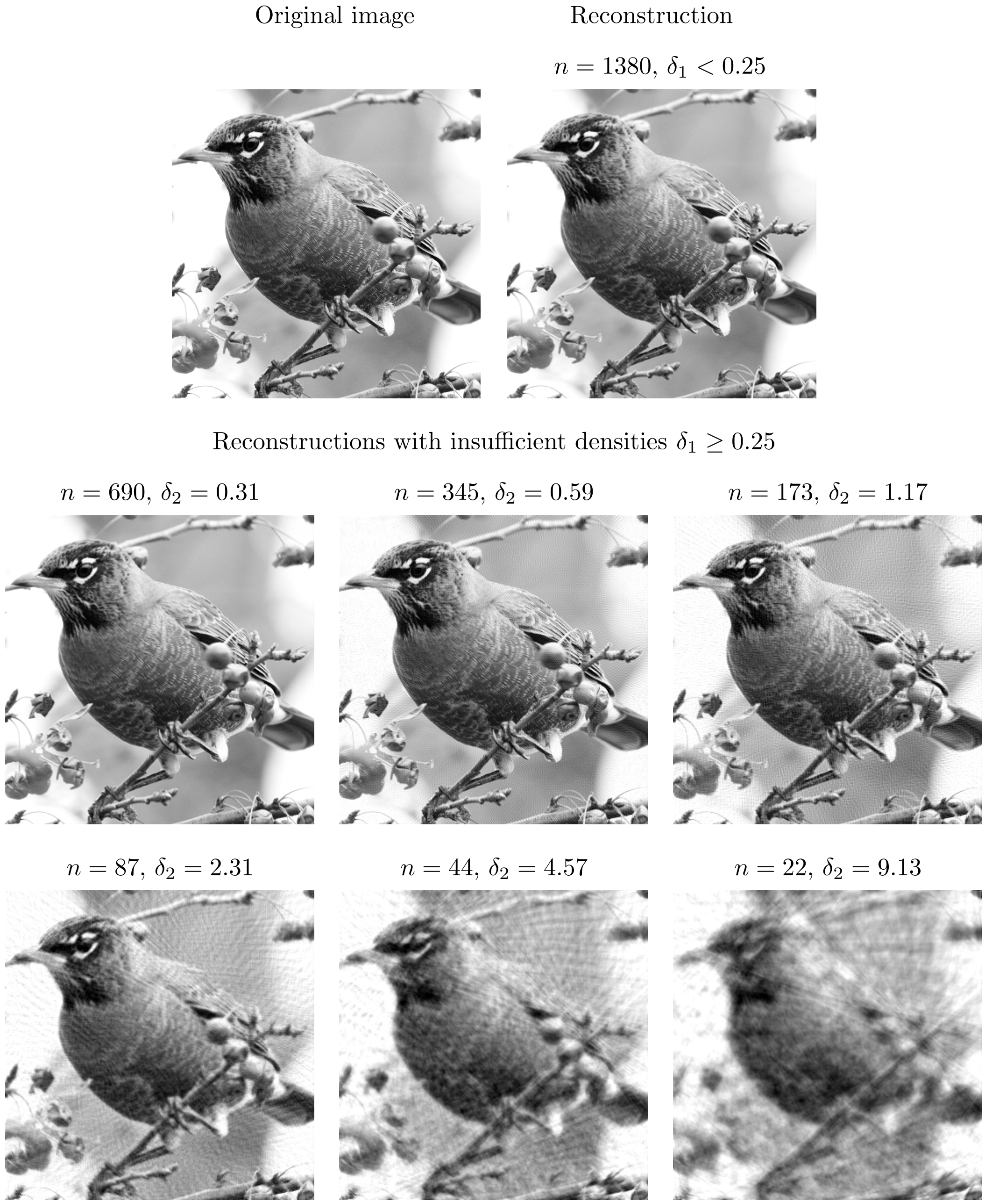

Figure 4.10: A high resolution image of 4500×4500 pixels is reconstructed by NUGS in the space T consisting of 352×352 indicator functions when samples are taken on a radial sampling scheme contained in $[-K, K]^2$, $K = 128$. The relation $\dim(\mathrm{T}) = (2.75K)^2$ is used. The reconstructions are shown for sampling schemes with different densities, i.e. different number of radial lines n. Here, the density in the Euclidean norm δ_2 was directly computed on a computer. Since $\delta_1 \geq \delta_2$, note that $\delta_2 \geq 0.25$ ensures that the density condition $\delta_1 < 0.25$ is violated.

Chapter 5

Reconstruction in piecewise polynomial spaces

In this chapter, we consider the problem of recovering (piecewise) smooth univariate functions to high accuracy from nonuniform samples of their Fourier transform using previously developed NUGS framework. In order to ensure high accuracy when recovering (piecewise) smooth functions, we employ reconstruction spaces consisting of splines or (piecewise) polynomials.

As we have seen so far, in NUGS the dimension of the reconstruction space T is allowed to vary in relation to the sampling bandwidth K. In order to obtain a reconstruction which is stable and quasi-optimal, the key issue prior to implementation of NUGS is to determine such scaling. In principle, this depends on both the nature of the nonuniform samples and the choice of reconstruction space. In this chapter we provide a general analysis which allows one to simultaneously determine such scaling for all possible nonuniform sampling schemes by scrutinizing two intrinsic quantities ζ and γ of the reconstruction space T, related to the *maximal uniform growth of functions* in T and the *maximal growth of derivatives* in T respectively. Provided these are known (as is the case for many choices of T), one can immediately estimate this scaling. As a particular consequence, for trigonometric polynomials, splines and piecewise algebraic polynomials (with fixed polynomial degree), we can show that this scaling is linear, and for piecewise algebraic polynomials with varying degree we show that it is quadratic.

Recently, a number of other works have investigated the problem of high-order reconstructions from nonuniform Fourier data. In [GH12, VGCR10] spectral reprojection techniques were used for this task, and a frame-theoretic approach was introduced in [GS14]. Recovering the Fourier transform to high accuracy was studied in [PGG12], and in [GH11, MGG14] the problem of high-order edge detection was addressed. We note again that the method based on NUGS can be shown to achieve optimal convergence rates amongst all stable, convergent algorithms [AHP13, AHS14].

The material of this chapter was published in [AGH14b], which is the joint work of the author with Ben Adcock and Anders Hansen.

5.1 Guarantees for piecewise smooth reconstruction spaces

Recall from Theorem 3.2.5 that for a finite-dimensional reconstruction space $\mathrm{T} \subseteq \mathrm{H} = \mathrm{L}^2(0,1)$, a sampling set Ω_N which is (K,δ)-dense with $\delta < 1/2$, and for a given $\epsilon \in (0, \sqrt{1-4\delta^2})$, if $K > 0$ is large enough such that

$$E(\mathrm{T}, K) \leq \epsilon, \tag{5.1.1}$$

then the NUGS reconstruction defined as in (3.1.10) with the weights (3.2.1) has the reconstruction constant $C(\Omega_N, \mathrm{T})$ satisfying

$$C(\Omega_N, \mathrm{T}) \leq \frac{1+2\delta}{\sqrt{1-\epsilon^2} - 2\delta}. \tag{5.1.2}$$

Instead of giving explicit scaling of $\dim(\mathrm{T})$ and K sufficient for stable and quasi-optimal recovery, Theorem 3.2.5 rather reinterprets the scaling of $\dim(\mathrm{T})$ and K in terms of the z-residual $E(\mathrm{T}, K)$, where as before, the z-residual of T is defined as

$$E(\mathrm{T}, z) = \sup_{\substack{f \in \mathrm{T} \\ \|f\|=1}} \|\hat{f}\|_{\mathbb{R}\setminus(-z,z)}, \quad z \in [0, \infty).$$

Note that, for a given reconstruction space T, this residual is independent of the geometry of the sampling points Ω_N, and depends solely on bandwidths $K = K(N)$. Hence, provided (5.1.1) holds, one ensures stable, quasi-optimal recovery for *any* sequence of sample points Ω_N with the same sampling bandwidth K.

Unsurprisingly, the behaviour of the z-residual depends completely on the choice of subspace T. Whilst one can often derive estimates for this quantity using ad-hoc approaches for each particular choice of T, as in Chapter 4 for the case of wavelet spaces, it is useful to have a more unified approach rising to an explicit scaling of $\dim(\mathrm{T})$ and K. We now present such an approach.

First, we recall the definition of the gap between two spaces.

Definition 5.1.1 ([Szy06]). *Let* U *and* V *be closed subspaces of* H *with corresponding orthogonal projections* $\mathcal{P}_{\mathrm{U}}$ *and* $\mathcal{P}_{\mathrm{V}}$ *respectively. The gap between* U *and* V *is the quantity*

$$G(\mathrm{U}, \mathrm{V}) = \|(\mathcal{I} - \mathcal{P}_{\mathrm{U}})\mathcal{P}_{\mathrm{V}}\|,$$

where $\mathcal{I} : \mathrm{H} \to \mathrm{H}$ *is the identity.*

Proposition 5.1.2. *Let* T *and* S *be finite-dimensional subspaces of* H *with* z*-residuals* $E(\mathrm{T}, z)$ *and* $E(\mathrm{S}, z)$ *respectively. Then*

$$E(\mathrm{T}, z) \leq E(\mathrm{S}, z) + G(\mathrm{S}, \mathrm{T})$$

for any $z \in [0, \infty)$.

Proof. Let $f \in \mathrm{T}$, $\|f\| = 1$. Then by Parseval's identity,

$$\begin{aligned}\|\hat{f}\|_{\mathbb{R}\setminus(-z,z)} &\leq \|\widehat{\mathcal{P}_{\mathrm{S}} f}\|_{\mathbb{R}\setminus(-z,z)} + \|\widehat{f - \mathcal{P}_{\mathrm{S}} f}\|_{\mathbb{R}\setminus(-z,z)} \\ &\leq \|\widehat{\mathcal{P}_{\mathrm{S}} f}\|_{\mathbb{R}\setminus(-z,z)} + \|f - \mathcal{P}_{\mathrm{S}} f\|.\end{aligned}$$

Since $f \in \mathrm{T}$, by definitions of z-residual and the gap, we get

$$\|\hat{f}\|_{\mathbb{R}\setminus(-z,z)} \leq E(\mathrm{S}, z)\|\mathcal{P}_{\mathrm{S}} f\| + G(\mathrm{S}, \mathrm{T})\|f\| \leq E(\mathrm{S}, z) + G(\mathrm{S}, \mathrm{T}),$$

as required. $\square$

This result implies the following: if the behaviour of z-residual $E(\mathrm{S}, z)$ and the gap $G(\mathrm{S}, \mathrm{T})$ are known, then one can immediately determine the required scaling of $\dim(\mathrm{T})$ with z to ensure that $E(\mathrm{T}, z)$ satisfies (5.1.1). We now make the following choice for S to allow us to exploit this result. For a given $L \in \mathbb{N}$, define

$$\mathrm{S}_L = \left\{ g \in \mathrm{H} : g|_{[l/L,(l+1)/L)} \in \mathbb{P}_0,\ l = 0, \ldots, L-1 \right\}, \tag{5.1.3}$$

where $\mathbb{P}_0$ is space of polynomials of degree zero. Note that $\dim(S_L) = L$. For such S_L, results from Section 4.2.2 show that for any $\epsilon > 0$ there exists a constant $c_0(\epsilon) > 0$ such that

$$E(\mathrm{S}_L, z) \leq \epsilon, \quad z \geq c_0(\epsilon) L.$$

Therefore, according to Proposition 5.1.2, in order to estimate $E(\mathrm{T}, z)$ for any finite-dimensional $\mathrm{T} \subseteq \mathrm{H}$, we now only need to determine $G(\mathrm{S}_L, \mathrm{T})$ for S_L as in (5.1.3).

From now on, we let $0 < w_1 < \ldots < w_k < 1$ be a fixed sequence of nodes, and define the space

$$\mathrm{H}^1_w(0,1) = \left\{ f : f|_{(w_j, w_{j+1})} \in \mathrm{H}^1(w_j, w_{j+1}),\ j = 0, \ldots, k \right\}$$

where $w_0 = 0$, $w_{k+1} = 1$ and $\mathrm{H}^1(I)$ is the usual Sobolev space of functions on an interval I. By convention, if $k = 0$ then $\mathrm{H}^1_w(0,1) = \mathrm{H}^1(0,1)$. Next, for

$$\mathrm{T} \subseteq \mathrm{H}^1_w(0,1)$$

and S_L as in (5.1.3), we derive a bound on $G(\mathrm{S}_L, \mathrm{T})$ in terms of the following quantities

$$\gamma_{\mathrm{T}} = \max_{j=0,\dots,k} \sup \left\{ \|f'\|_{(w_j,w_{j+1})} : f \in \mathrm{T},\ \|f\|_{(w_j,w_{j+1})} = 1 \right\}, \tag{5.1.4}$$

and

$$\zeta_{\mathrm{T}} = \max_{j=0,\dots,k} \sup \left\{ \|f\|_{\infty,(w_j,w_{j+1})} : f \in \mathrm{T},\ \|f\|_{(w_j,w_{j+1})} = 1 \right\}, \tag{5.1.5}$$

related to the maximal growth of function derivatives in T and the maximal uniform growth of functions in T respectively.

Proposition 5.1.3. *Suppose that* $\mathrm{T} \subseteq \mathrm{H}^1_w(0,1)$ *and let* S_L *be given by* (5.1.3)*. If* $L^{-1} \leq \eta = \min_{j=0,\dots,k}\{w_{j+1} - w_j\}$ *then*

$$G(\mathrm{S}_L, \mathrm{T}) \leq \sqrt{\left(\frac{\gamma_{\mathrm{T}}}{\pi L}\right)^2 + \frac{4\zeta_{\mathrm{T}}^2}{L}},$$

where γ_{T} *and* ζ_{T} *are given by* (5.1.4) *and* (5.1.5) *respectively, and, if* I *is an interval,* $\|f\|_I^2 = \int_I |f(x)|^2 \,\mathrm{d}x$ *and* $\|f\|_{\infty,I} = \operatorname{ess\,sup}_{x\in I}|f(x)|$*. Moreover, if* $k = 0$*, i.e.* $\mathrm{T} \subseteq \mathrm{H}^1(0,1)$*, then* $G(\mathrm{S}_L, \mathrm{T}) \leq \gamma_{\mathrm{T}}/(\pi L)$*.*

Proof. Since $L \geq 1/\eta$ there exist $l_j \in \mathbb{N}$ with $l_1 < l_2 < \ldots < l_k$ such that

$$0 \leq L w_j - l_j < 1, \quad j = 1, \dots, k.$$

In particular, $\frac{l_j}{L} \leq w_j < \frac{l_j+1}{L} \leq \frac{l_{j+1}}{L}$ for $j = 1, \dots, k$. For an interval $I \subseteq \mathbb{R}$, let us now write $f_I = \frac{1}{|I|}\int_I f$. Then

$$\|f - \mathcal{P}_{\mathrm{S}_L} f\|^2 = \sum_{l=0}^{L-1} \int_{I_l} |f - f_{I_l}|^2 = \sum_{\substack{l=0 \\ l \neq l_1,\dots,l_k}}^{L-1} \int_{I_l} |f - f_{I_l}|^2 + \sum_{j=1}^{k} \int_{I_{l_j}} \left| f - f_{I_{l_j}} \right|^2,$$

where $I_l = [l/L, (l+1)/L)$. Since $f \in \mathrm{H}^1(I_l)$ for $l \neq l_1, \dots, l_k$, an application of Poincaré's inequality gives that

$$\|f - \mathcal{P}_{\mathrm{S}_L} f\|^2 \leq \frac{1}{(L\pi)^2} \sum_{\substack{l=0 \\ l \neq l_1,\dots,l_k}}^{L-1} \|f'\|_{I_l}^2 + \sum_{j=1}^{k} \int_{I_{l_j}} \left| f - f_{I_{l_j}} \right|^2. \tag{5.1.6}$$

We now consider the second term. Write

$$I_{l_j} = (l_j/L, w_j) \cup (w_j, (l_j+1)/L) = A_j \cup B_j$$

and note that for an arbitrary interval I we have $\int_I |f - f_I|^2 = \|f\|_I^2 - |I||f_I|^2$. Hence

$$\begin{aligned}\int_{I_{l_j}} \left|f - f_{I_{l_j}}\right|^2 &= \|f\|_{A_j}^2 + \|f\|_{B_j}^2 - \frac{1}{|A_j| + |B_j|}\left||A_j| f_{A_j} + |B_j| f_{B_j}\right|^2 \\ &= \int_{A_j} |f - f_{A_j}|^2 + \int_{B_j} |f - f_{B_j}|^2 + \frac{|A_j||B_j|}{|A_j| + |B_j|}\left|f_{A_j} - f_{B_j}\right|^2 \\ &\leq \frac{1}{(\pi L)^2}\left(\|f'\|_{A_j}^2 + \|f'\|_{B_j}^2\right) + \frac{2|A_j||B_j|}{|A_j| + |B_j|}\left(\|f\|_{\infty,A_j}^2 + \|f\|_{\infty,B_j}^2\right),\end{aligned}$$

where in the final step we use Poincaré's inequality once more and the fact that f is H^1 within A_j and B_j. Since $|A_j|, |B_j| \leq L^{-1}$ and $|A_j| + |B_j| = |I_{l_j}| = L^{-1}$ we now get

$$\sum_{j=1}^{k} \int_{I_{l_j}} \left|f - f_{I_{l_j}}\right|^2 \leq \frac{1}{(\pi L)^2} \sum_{j=1}^{k} \left(\|f'\|_{A_j}^2 + \|f'\|_{B_j}^2\right) + \frac{4}{L} \sum_{j=0}^{k} \|f\|_{\infty,(w_j,w_{j+1})}^2.$$

Combining this with (5.1.6) gives

$$\|f - \mathcal{P}_{\mathrm{S}_L} f\|^2 \leq \left(\frac{\gamma_{\mathrm{T}}}{L\pi}\right)^2 \sum_{j=0}^{k} \|f\|_{(w_j,w_{j+1})}^2 + \frac{4\zeta_{\mathrm{T}}^2}{L} \sum_{j=0}^{k} \|f\|_{(w_j,w_{j+1})}^2.$$

Since $\|f\|^2 = \sum_{j=0}^{k} \|f\|_{(w_j,w_{j+1})}^2$ the result now follows. □

Using results of Propositions 5.1.3 and 5.1.2 in a combination with Theorem 3.2.5, we now obtain the following:

Corollary 5.1.4. *Let Ω_N be (K,δ)-dense with $\delta < 1/2$, $\mathrm{T} \subseteq \mathrm{H}_w^1(0,1)$ be a finite-dimensional subspaces of dimension $M \in \mathbb{N}$ such that*

$$\gamma_{\mathrm{T}} = \mathcal{O}(M^{\alpha}), \quad \zeta_{\mathrm{T}} = \mathcal{O}(M^{\beta}), \quad M \to \infty,$$

for some $\alpha, \beta > 0$. Then, for each $0 < \epsilon < \sqrt{1 - 4\delta^2}$ there exists a $c_0(\epsilon) > 0$ such that if

$$M \leq c_0(\epsilon) K^{\frac{1}{\tau}}, \quad \tau = \max\{\alpha, 2\beta\}$$

then then the NUGS reconstruction defined as in (3.1.10) with the weights (3.2.1) has the reconstruction constant $C(\Omega_N, \mathrm{T})$ satisfying (5.1.2). Moreover, if $\mathrm{T} \subseteq \mathrm{H}^1(0,1)$, the claim holds with $\tau = \alpha$.

This provides a unified approach to analysing the reconstruction. Given the bandwidth K stable reconstruction, for any sampling scheme Ω_N, can be ensured solely by estimating the quantities γ_{T} and ζ_{T}, which are intrinsic properties of the reconstruction space T completely unrelated to the sampling points.

5.2 Sufficient scaling of K and $\dim(\mathrm{T})$

To illustrate implications of Corollary 5.1.4, we now consider several different reconstruction spaces.

5.2.1 Trigonometric polynomials

Functions f that are smooth and periodic can be approximated in finite-dimensional spaces of trigonometric polynomials

$$\mathrm{T}_M = \left\{ \sum_{m=-M}^{M} a_m \mathrm{e}^{2\pi \mathrm{i} m x} : a_m \in \mathbb{C} \right\}.$$

If $f \in \mathrm{C}^\infty(\mathbb{T})$, where $\mathbb{T} = [0,1)$ is the unit torus, then the projection error $\|f - \mathcal{P}_{\mathrm{T}_M} f\|$ decay super-algebraically fast in M; that is, faster than any power of M^{-1}. If f is also real analytic then the error decays exponentially fast.

For this space, we have $\mathrm{T}_M \subseteq \mathrm{H}^1(0,1)$ and

$$\gamma_{\mathrm{T}_M} \leq 2\pi M$$

by Bernstein's inequality. Hence Corollary 5.1.4 gives that the NUGS reconstruction in T_M is stable and quasi-optimal provided M scales linearly with the sampling bandwidth K, namely provided that

$$M = \mathcal{O}(K), \quad K \to \infty.$$

This result extends a previous result of [AHP13] to the case of arbitrary nonuniform samples.

5.2.2 Algebraic polynomials

Functions that are smooth but nonperiodic can be approximated by algebraic polynomials. If

$$\mathrm{T}_M = \mathbb{P}_M$$

is the space of algebraic polynomials of degree at most M, then the projection error $\|f - \mathcal{P}_{\mathrm{T}_M} f\|$ decays super-algebraically fast in M whenever $f \in \mathrm{C}^\infty[0,1]$, and exponentially fast when f is analytic.

For this space we have $\mathrm{T}_M \subseteq \mathrm{H}^1(0,1)$ and the classical Markov inequality [BD10] gives

$$\gamma_{\mathrm{T}_M} \leq \sqrt{2} M^2, \quad \forall M \in \mathbb{N}.$$

Hence, from Corollary 5.1.4, we deduce stability and quasi-optimality of the NUGS reconstruction in T_M, but only subject to the square-root scaling

$$M = \mathcal{O}(\sqrt{K}), \quad K \to \infty.$$

This result extends previous results [HG10, AH12b, AH15b] to the case of nonuniform Fourier samples. On the face of it, this scaling is unfortunate since it means the approximation accuracy is limited to root-exponential in K, which is much slower than the exponential decay rate of the projection error. However, in the uniform case, such scaling is the best possible: as shown in [AHS14], any reconstruction algorithm (linear or nonlinear) that achieves faster than root-exponential accuracy for analytic functions must necessarily be unstable.

5.2.3 Splines with nonequidistant knots

Since the previous scaling is so severe, one may seek to choose a space with worse approximation properties but a better scaling. Spline spaces provide such a choice. Let $0 = y_0 < y_1 < \dots y_M < y_{M+1} = 1$ be a sequence of knots in $[0, 1]$, and suppose that s is a fixed integer greater than 1. We exclude the case $s = 1$ for now, since that requires a slightly different approach which will be presented in the next section. Consider the space

$$\mathrm{T}_{y,s} = \left\{ f \in \mathrm{C}^{s-1}[0,1] : f|_{[y_j, y_{j+1})} \in \mathbb{P}_s,\ j = 0, \dots, M \right\},$$

where $y = \{y_1, \dots, y_M\}$. This space is well-suited for approximating smooth functions. Indeed, for smooth f the projection error $\|f - \mathcal{P}_{\mathrm{T}_{y,s}} f\|$ decays like h^{-s-1}, where $h = \max_{j=0,\dots,M} |y_{j+1} - y_j|$.

Note that for $s > 1$, $\mathrm{T}_{y,s} \subseteq \mathrm{H}^1(0, 1)$. For $f \in \mathrm{T}_{y,s}$, we have the following

$$\|f'\|^2 = \sum_{j=0}^{M} \|f'\|^2_{[y_j, y_{j+1})} \leq \sum_{j=0}^{M} \left(\frac{\sqrt{2}s^2}{y_{j+1} - y_j} \right)^2 \|f\|^2_{[y_j, y_{j+1})} \leq \left(\frac{\sqrt{2}s^2}{\eta} \right)^2 \|f\|^2,$$

where $\eta = \min_{j=0,\dots,M} |y_{j+1} - y_j|$. Note that for the middle inequality we use the fact that Markov's inequality for an arbitrary interval $I \subseteq \mathbb{R}$ is given by

$$\|p'\|_I \leq \frac{\sqrt{2}s^2}{|I|} \|p\|_I, \quad \forall p \in \mathbb{P}_s,\ s \in \mathbb{N}, \tag{5.2.1}$$

where $|I|$ denotes the length of I, $\|\cdot\|_I$ is the L^2-norm over I. Therefore

$$\gamma_{\mathrm{T}_{y,s}} \leq \frac{\sqrt{2}s^2}{\eta}$$

and from Corollary 5.1.4 we deduce the sufficient condition for stable and quasi-optimal reconstruction

$$\frac{s^2}{\eta} = \mathcal{O}(K), \quad K \to \infty.$$

In particular, if $y_m = m/(M+1)$ are equispaced knots, then $\eta = 1/(M+1)$ and we obtain the scaling $M+1 = \mathcal{O}(K/s^2)$ as $K \to \infty$. Hence, up to a constant which depends on $1/s^2$, equispaced spline spaces possess a linear scaling of $M+1$ (number of knots) with sampling bandwidth K. This is substantially better than the case of polynomial spaces. Note that the polynomial result is actually a special case of this result corresponding to the case $M = 0$ and s being the polynomial degree.

5.2.4 Piecewise algebraic polynomials

The reconstruction spaces considered so far are not suitable for approximating piecewise smooth functions. For this reason, it may be useful to consider spaces of piecewise polynomials with possibly different degrees in each subinterval. These approximation spaces are appropriate if f is only piecewise smooth with known edges. Even if f is smooth over the whole interval $[0,1]$, in order to mitigate the severe scaling (5.2.2), one may wish to approximate it in a piecewise manner as in a spline space (i.e. by refining a sequence of knots rather than the polynomial degree), but without the additional effort of enforcing continuity as required in spline spaces.

We consider the space

$$\mathrm{T}_{w,M} = \{f \in \mathrm{H} : f|_{[w_j,w_{j+1})} \in \mathbb{P}_{M_j},\ j = 0,\dots,k\},$$

where $w = \{w_1,\dots,w_k\}$ for $0 = w_0 < w_1 < \dots w_k < w_{k+1} = 1$ and $M = \{M_0,\dots,M_k\} \in \mathbb{N}^{k+1}$. If f is piecewise smooth with jump discontinuities at known locations $0 = w_0 < w_1 < \dots w_k < w_{k+1} = 1$ then the projection error decays super-algebraically fast in powers of $(M_{\min})^{-1}$ as $M_{\min}$ increases, where $M_{\min} = \min\{M_0,\dots,M_k\}$, and exponentially fast if f is piecewise analytic. Alternatively, if f is smooth and the points w are varied whilst the degrees M are fixed, then the error decays like $h^{-M_{\min}-1}$, where $h = \max_{j=0,\dots,k} |w_{j+1} - w_j|$.

Since $\mathrm{T}_{M,w}$ is not in $\mathrm{H}^1(0,1)$, but rather in $\mathrm{H}^1_w(0,1)$, to apply Corollary 5.1.4, we need to determine $\gamma_{\mathrm{T}_{M,w}}$ and $\zeta_{\mathrm{T}_{M,w}}$. For the first we use the scaled Markov inequality

$$\|p'\|_I \le \frac{\sqrt{2}M^2}{|I|}\|p\|_I, \quad \forall p \in \mathbb{P}_M, M \in \mathbb{N}.$$

Hence, if $\eta = \min_{j=0,\dots,k}\{w_{j+1} - w_j\}$ and $M_{\max} = \max\{M_0,\dots,M_k\}$ then

$$\gamma_{\mathrm{T}_{M,w}} \le \frac{\sqrt{2}M_{\max}^2}{\eta},$$

For $\zeta_{\mathrm{T}_{M,w}}$, we recall the following inequality for polynomials

$$\|p\|_{\infty,I} \leq \frac{cM}{\sqrt{|I|}}\|p\|_I, \quad \forall p \in \mathbb{P}_M, M \in \mathbb{N},$$

where $c > 0$ is a constant. Hence

$$\zeta_{\mathrm{T}_{M,w}} \leq \frac{cM_{\max}}{\sqrt{\eta}}.$$

Due to Corollary 5.1.4, we now deduce the following sufficient condition for stable and quasi-optimal reconstruction

$$\frac{M_{\max}^2}{\eta} = \mathcal{O}(K), \quad K \to \infty.$$

In the first scenario, where η is fixed and $M_{\max}$ is varied, we attain the same square-root-type scaling for piecewise smooth functions when approximated by piecewise polynomials as with the polynomial space (5.2.2). In the second scenario, where $M_{\max}$ is fixed and η is varied, we see that this leads to a linear relation between K and $1/\eta$. Thus, by forfeiting the super-algebraic/exponential convergence of the polynomial space for only algebraic convergence, we obtain a better scaling with K.

5.3 Numerical example

We now demonstrate the results of this chapter on a numerical example.

In the upper two panels of Figure 5.1, using two common nonuniform sampling schemes Ω_N, jittered (3.4.5) and log (3.4.13), and using different reconstruction spaces T, we illustrate the scaling between the sampling bandwidth K and the space dimension $\dim(\mathrm{T})$. For a bandwidth K, we find $\dim(\mathrm{T})$ such that the reconstruction constant $C(\Omega_N, \mathrm{T})$ is bounded. The fact that the plotted scalings are bounded by a constant for large K aligns with our theoretical results.

Next, in the lower pair of panels of Figure 5.1, for such K and $\dim(\mathrm{T})$, we compute the L^2-error of the NUGS approximation $\tilde{f}$ for a continuous function

$$f(x) = x^2 + x\sin(4\pi x) - \mathrm{e}^{\frac{x}{2}}\cos(3\pi x)^2.$$

It is evident that the superb approximation orders are achieved when reconstructing with algebraic polynomials of high degree, however to have these one needs to pay by the severe scaling of $\dim(\mathrm{T})$ with the sampling bandwidth K. Hence, one may prefer to use spline spaces for smaller K and still attain relatively high approximation orders.

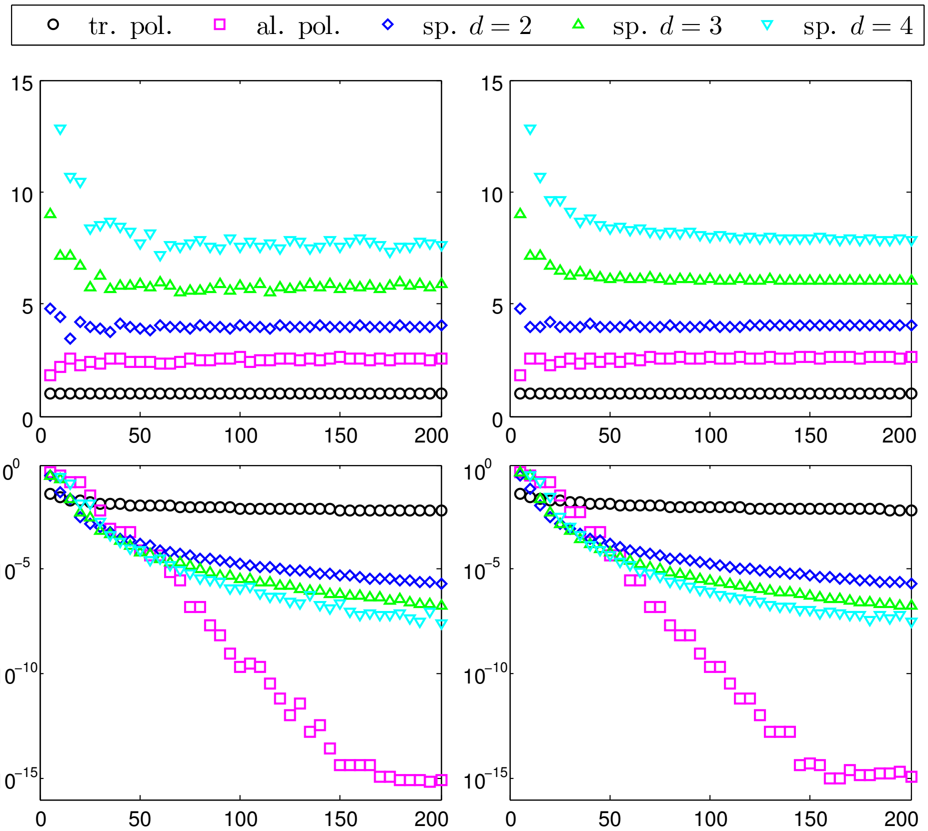

Figure 5.1: In the upper pair of panels, depending on the type of the reconstruction space, appropriate ratios are shown: M/K (for trigonometric polynomials), $M/\sqrt{K}$ (for algebraic polynomials) and $M/(K/d^2)$ (for splines of order d), where for a sampling bandwidth given $K \in [5, 200]$, we used $M = \max\{M \in \mathbb{N} : C(N, M) \leq 3\}$. In the lower pair of panels, for such K and M, the error $\|f - \tilde{f}\|$ is plotted where $f(x) = x^2 + x \sin(4\pi x) - \exp(x/2) \cos(3\pi x)^2$. The sampling schemes Ω_N is jittered for the left panels, and log for the right panel.

Chapter 6

Nonuniform sampling with derivatives or bunched points

As mentioned in Chapter 1, a set of sampling for the space of L^2 functions supported on a compact domain also constitutes a set of sampling for the space of functions bandlimited to the same compact domain. For a compact domain $\Omega \subseteq \hat{\mathbb{R}}^d$, the space of Ω-bandlimited functions is defined by

$$\mathrm{B}(\Omega) = \left\{ f \in \mathrm{L}^2(\mathbb{R}^d) : \mathrm{supp}(\hat{f}) \subseteq \Omega \right\}.$$

In Chapter 2 we studied sets of sampling in order to provide guarantees for stable recovery of a compactly supported L^2 function from pointwise measurements of its Fourier transform, which were studied in detail throughout Chapters 3–5. Similarly, one could analyse the recovery problem of a bandlimited function from its own samples. In this chapter we study sampling of bandlimited functions when measurements include some additional information. Specifically, in this chapter we address two different sampling scenarios of bandlimited functions that allow for a reduced sampling density: 1) nonuniform sampling of a function and its first k derivatives, and 2) nonuniform sampling of a function at bunched points. As before, in order to ensure a stable reconstruction, one is essentially concerned with conditions that ensure existence of frame for the corresponding function space. Deriving such condition in context of these two sampling scenarios is the topic of the present chapter.

As mentioned in introduction, the sampling scenarios considered here are motivated by applications in seismology. As it turns out, both of these scenarios allow sampling below the Nyquist rate, and hence bigger distances between sampling sensors, thereby reducing cost and effort in practical acquisition of data.

The results of this chapter are collected from [AGH15a], which is the joint work of the author with Ben Adcock and Anders Hansen.

6.1 Summary of main results

The first result of this chapter, Theorem 6.2.1, provides an upper bound on the maximum allowable sampling density δ, such that samples of derivatives give rise to a particular frame. The density bound as well as the explicit estimates of the corresponding frame bounds depend on the number of derivatives k, the norm used in specifying δ and a certain geometric property of the domain Ω. For large k, the maximum allowed δ grows linearly in $k+1$ with constant of proportionality $1/\mathrm{e}$. This extends the univariate result of [Raz95] to the multivariate setting, as well as the multivariate $k = 0$ (no derivative) results of [Grö92, Grö01] to the case of derivatives.

In our second result, Theorem 6.2.9, we present an univariate density condition that leads to a small improvement over [Raz95] for $k \geq 2$ derivatives. This follows the technique of [Grö92] for the univariate case based on Wirtinger inequalities. We provide an explicit calculation of the optimal constants in certain higher-order Wirtinger inequalities, which, replicating the techniques of [Grö92] for the case of derivatives, lead to modestly improved estimates for δ for finite k. Such improved bounds can be used to get better estimates for two-dimensional spatial-temporal sampling scenarios, as we consider in Proposition 6.2.11.

Next, we provide Theorem 6.2.12, which gives a perturbation estimate for nonuniform sampling with derivatives. We show that if $\{x_n\}_{n\in I}$ is a stable set of sampling for derivatives, then so is $\{\tilde{x}_n\}_{n\in I}$ whenever $\sup_{n\in I} |x_n - \tilde{x}_n|$ is sufficiently small. In particular, small perturbations of the uniform sampling points taken at k times Nyquist give rise to stable sets of sampling. This extends existing results given in [Bai10, SZ99] to the case of sampling with derivatives. Moreover, it improves those results since we provide a dimension independent bound for appropriate domains Ω.

In Section 6.3, we address univariate nonuniform bunched sampling and, in Theorem 6.3.1, we give density guarantees in order to obtain a particular fusion frame [CK04, CKL08]. Similarly as in derivatives sampling, we show that the density bound increases linearly with $s+1$ (the number of samples in each bunch) with constant of proportionality depending on the width of the bunches. The points within the same bunch are permitted to get arbitrarily close to each other, since we use appropriate weights.

Next, by Theorem 6.3.3, we obtain the same density condition as in Theorem 6.3.1, but now leading to a particular frame based on divided differences. Furthermore, in Corollary 6.3.4 we show that the corresponding density bound in the limit—for small width of bunches and for large s—gives the same density bound as the one we provide for the univariate sampling with s derivatives, i.e. the density bound grows linearly in $s+1$ with the constant of proportionality $1/\mathrm{e}$.

Lastly, by Theorem 6.3.5, we conclude that if a derivative sampling gives rises to a frame, then a bunched sampling that is a perturbation of the derivative sampling gives rise to a frame as well.

6.2 Nonuniform derivative sampling

Let $\{x_n\}_{n\in I} \subseteq \mathbb{R}^d$ be a set of sampling points, where I is a countable index set. Let $f \in \mathrm{B}(\Omega)$, and suppose that we are given the measurements

$$D^\alpha f(x_n), \quad n \in I, |\alpha|_1 \le k.$$

Stable recovery from these measurements is possible if there exist constants $A, B > 0$ such that

$$\forall f \in \mathrm{B}(\Omega), \quad A\|f\|^2 \le \sum_{n\in I} \sum_{|\alpha|_1 \le k} \mu_{n,\alpha} |D^\alpha f(x_n)|^2 \le B\|f\|^2, \tag{6.2.1}$$

holds for some weights $\mu_{n,\alpha} > 0$. Following [Raz95], let us now define the function

$$\Phi_\Omega(x) = \int_\Omega \mathrm{e}^{\mathrm{i}2\pi\omega\cdot x}\,\mathrm{d}\omega, \quad x \in \mathbb{R}^d. \tag{6.2.2}$$

For a given $f \in \mathrm{B}(\Omega)$, we have $\hat{f}(\omega) = \hat{f}(\omega)\chi_\Omega(\omega)$. If $g \in \mathrm{B}(\Omega)$ is such that $\hat{g}(\omega) = \chi_\Omega(\omega)$, then by the convolution theorem we can write

$$f(x) = \int_{\mathbb{R}^d} f(s)g(x-s)\,\mathrm{d}s = \int_{\mathbb{R}^d}\int_\Omega f(s)\mathrm{e}^{\mathrm{i}2\pi\omega\cdot(x-s)}\,\mathrm{d}\omega\,\mathrm{d}s = \langle f, \Phi_\Omega(\cdot - x)\rangle.$$

Therefore

$$D^\alpha f(x_n) = \langle D^\alpha f, \Phi_\Omega(\cdot - x_n)\rangle = (-1)^{|\alpha|_1}\langle f, D^\alpha \Phi_\Omega(\cdot - x_n)\rangle.$$

Hence (6.2.1) is equivalent to the condition that the set of functions

$$\left\{\sqrt{\mu_{n,\alpha}} D^\alpha \Phi_\Omega(\cdot - x_n) : n \in I, |\alpha|_1 \le k\right\},$$

forms a frame for $\mathrm{B}(\Omega)$ with frame bounds $A, B > 0$.

Similarly, after differentiation and using Parseval's identity, (6.2.1) becomes

$$\forall f \in \mathrm{B}(\Omega), \quad A\ \|\hat{f}\|^2 \le \sum_{n\in I}\sum_{|\alpha|_1\le k} \mu_{n,\alpha} |\langle \hat{f}, (-\mathrm{i}2\pi\omega)^\alpha \mathrm{e}^{-\mathrm{i}2\pi\omega\cdot x_n}\rangle|^2 \le B\ \|\hat{f}\|^2,$$

and therefore, (6.2.1) is equivalent to $\left\{\sqrt{\mu_{n,\alpha}}(-\mathrm{i}2\pi\omega)^\alpha \mathrm{e}^{-\mathrm{i}2\pi\omega\cdot x_n}\chi_\Omega(\omega) : n \in I, |\alpha|_1 \le k\right\}$ being a Fourier frame for $\mathrm{L}^2(\Omega) = \{f \in \mathrm{L}^2(\mathbb{R}^d) : \mathrm{supp}(f) \subseteq \Omega\}$ with the frame bounds $A, B > 0$; see for example [You01].

In what follows, we provide sufficient conditions for (6.2.1) for an appropriate choice of weights. Again, our weights shall be related to the Voronoi cells $\{V_n\}_{n\in I}$ of the sampling points $\{x_n\}_{n\in I}$ with respect to a norm $|\cdot|_*$. Namely, we define

$$\mu_{n,\alpha} = \frac{1}{\alpha!}\int_{V_n}(x-x_n)^{2\alpha}\,\mathrm{d}x, \quad \alpha\in\mathbb{N}_0^d,\ n\in I. \tag{6.2.3}$$

Also, our sufficient conditions for (6.2.1) with these weights will be in terms of the density of the sampling points, measured in the following sense:

$$\delta_* = \sup_{x\in\mathbb{R}^d}\inf_{n\in I}|x-x_n|_*. \tag{6.2.4}$$

Our aim is to find the maximal allowable density δ_* for which (6.2.1) holds. As in Chapter 2, our estimates will be derived in terms of the quantity

$$m_\Omega = \sup_{x\in\Omega}|x|, \tag{6.2.5}$$

and the sharp constant $c^* > 0$ such that

$$\forall x\in\mathbb{R}^d, \quad |x|\le c^*|x|_*. \tag{6.2.6}$$

6.2.1 The multivariate case

First, we need to define some functions. Let $k\in\mathbb{N}_0$, $d\in\mathbb{N}$, and define

$$h_k(z) = \exp(z)R_k(z), \tag{6.2.7}$$

$$g_{k,d}(z) = (1+2\sigma_d^*(z))^{\frac{d}{2}}\exp\left(\frac{z}{\sigma_d^*(z)}\right)R_k(z), \tag{6.2.8}$$

for $z\in(0,\infty)$, where

$$R_k(z) = \exp(z) - \sum_{r=0}^{k}\frac{1}{r!}z^r, \tag{6.2.9}$$

$$\sigma_d^*(z) = \frac{z+\sqrt{z(d+z)}}{d}, \tag{6.2.10}$$

for $z\in(0,\infty)$. Note that both h_k and $g_{k,d}$ have limiting value 0 as $z\to 0^+$, and both increase monotonically to infinity as $z\to\infty$. Hence they have well-defined inverse functions $H_k(w)$ and $G_{k,d}(w)$ for $w\in(0,\infty)$.

Our main result is now as follows:

Theorem 6.2.1. *Suppose that the weights $\mu_{n,\alpha}$ are given by (6.2.3) and let δ_* be as in (6.2.4). If*

$$\delta_* < \frac{C(k,d)}{2\pi m_\Omega c^*}, \qquad C(k,d) = \max\{H_k(1), G_{k,d}(1)\}, \tag{6.2.11}$$

then

$$\forall f \in \mathrm{B}(\Omega), \quad A\|f\|^2 \leq \sum_{n \in I} \sum_{|\alpha|_1 \leq k} \mu_{n,\alpha} |D^\alpha f(x_n)|^2 \leq B\|f\|^2,$$

where $A, B > 0$ *satisfy*

$$A \geq \mathrm{e}^{-d} \left(1 - \min\left\{h_k(2\pi m_\Omega c^* \delta_*), g_{k,d}(2\pi m_\Omega c^* \delta_*)\right\}\right)^2, \tag{6.2.12}$$

$$B \leq \exp\left(4\pi m_\Omega c^* \delta_* + (2\pi m_\Omega c^* \delta_*)^2\right). \tag{6.2.13}$$

Equivalently, the set $\{\sqrt{\mu_{n,\alpha}} D^\alpha \Phi_\Omega(\cdot - x_n) : n \in I, |\alpha|_1 \leq k\}$ *forms a frame for* $\mathrm{B}(\Omega)$ *with frame bounds* A *and* B*.*

The key part of this theorem—whose proof we defer to §6.2.1—is the condition (6.2.11). Note that an interesting facet of (6.2.11) is that it splits geometric terms depending on the domain Ω (the constant m_Ω) and the norm used (encapsulated by the term c^*), from the nongeometric constant $C(k, d)$.

Whilst values of $C(k, d)$ for fixed k and d are easily calculated and are presented in Table 6.1, to understand its behaviour it is interesting to consider the following two asymptotic regimes:

$$\text{(i) } k \text{ fixed, } d \to \infty, \qquad \text{(ii) } d \text{ fixed, } k \to \infty.$$

In (i) it is desirable for (6.2.11) to not decrease with d, i.e. the density bound does not worsen with increasing dimension; recall Chapter 2. For (ii), we desire linear increase in the bound with k, i.e. adding derivatives samples means that sampling points can be taken further apart, at as fast a rate as possible. As we show next, this is also the behaviour of $C(k, d)$.

Case (i)

As seen in Table 6.1, the constant $C(k, d)$ is independent of d for large d and fixed k. This follows from (6.2.11), where it is clear that for large d the maximum is achieved by $H_k(1)$, which is dimension independent, as opposed to $G_{k,d}(1)$ (it can be easily proved that $G_{k,d}(1)$ decreases with d). Hence for large d, the only possible dimension-dependence in (6.2.11) arises from the factors m_Ω and c^*, which are determined by the domain Ω and the norm $|\cdot|_*$ respectively. As in Chapter 2, for simplicity, suppose that Ω is the unit ℓ^p-ball, $p > 0$, and let $|\cdot|_* = |\cdot|_q$, $q \geq 1$, be the ℓ^q-norm. Then (6.2.11) reads

$$\delta < \frac{C(k, d)}{2\pi \max\{1, d^{1/2-1/p}\} \max\{1, d^{1/2-1/q}\}}.$$

In particular, if $p = q = 2$ for example (i.e. Ω is contained in the unit Euclidean ball and the δ-density is measured in the Euclidean metric), then (6.2.11) reduces to $\delta < C(k, d)/(2\pi)$.

k	0	1	2	3	4	...	8	9	...	13	14	...
$C(k,1)$	0.4812	0.8141	1.1268	1.4304	*1.7890*	...	*3.2501*	*3.6163*	...	*5.0828*	*5.4498*	...
$C(k,2)$	0.4812	0.8141	1.1268	1.4304	1.7290	...	2.8976	*3.2424*	...	*4.6462*	*5.0000*	...
$C(k,3)$	0.4812	0.8141	1.1268	1.4304	1.7290	...	2.8976	3.1862	...	4.3327	*4.6679*	...
$C(k,4)$	0.4812	0.8141	1.1268	1.4304	1.7290	...	2.8976	3.1862	...	4.3327	4.6180	...
$C(k,5)$	0.4812	0.8141	1.1268	1.4304	1.7290	...	2.8976	3.1862	...	4.3327	4.6180	...

k	...	17	18	19	20	21	22	23	24	25	26
$C(k,1)$	...	*6.5512*	*6.9184*	*7.2857*	*7.6531*	*8.0205*	*8.3879*	*8.7553*	*9.1228*	*9.4903*	*9.8578*
$C(k,2)$	...	*6.0660*	*6.4227*	*6.7799*	*7.1376*	*7.4958*	*7.8544*	*8.2134*	*8.5728*	*8.9325*	*9.2925*
$C(k,3)$	...	*5.7002*	*6.0466*	*6.3940*	*6.7424*	*7.0916*	*7.4415*	*7.7922*	*8.1435*	*8.4955*	*8.8480*
$C(k,4)$	...	5.4715	5.7553	*6.0813*	*6.4207*	*6.7614*	*7.1031*	*7.4457*	*7.7893*	*8.1338*	*8.4791*
$C(k,5)$	...	5.4715	5.7553	6.0389	6.3223	6.0654	6.8883	7.1711	*7.4879*	*7.8252*	*8.1636*

Table 6.1: The constant $C(k,d)$ in the multi-dimensional density bound (6.2.11). Italics indicate when $C(k,d) = G_{k,d}(1)$, and otherwise $C(k,d) = H_k(1)$.

For sufficiently large d, one therefore obtains the dimensionless bound $\delta < H_k(1)/(2\pi)$. On the other hand, if $\Omega = [-1,1]^d$ is the unit cube and $|\cdot|_* = |\cdot|_2$ is the ℓ^2-norm, then we get square-root decay of the corresponding bound, which reads $\delta < H_k(1)/(2\pi\sqrt{d})$ for large d.

Remark 6.2.2 The splitting of the bound (6.2.11) into the factors $C(k,d)$ and $m_\Omega c^*$ is an extension of Theorem 2.2.1 from Section 2.2 to the case $k \geq 1$. Therein the case $k = 0$ was considered and the bound $\delta < (\ln 2)/(2\pi m_\Omega c^*)$ was established. Conversely, (6.2.11) for $k = 0$ reduces to the somewhat stricter condition $\delta < \log\left(\frac{1}{2}(1+\sqrt{5})\right)/(2\pi m_\Omega c^*)$; note that $\ln\left(\frac{1}{2}(1+\sqrt{5})\right) \approx 0.4812 < 0.6931 \approx \ln 2$. This is due to the additional complications arising from a bound that holds for arbitrary many derivatives.

Case (ii)

We now discuss the case of fixed d and increasing k. Empirically, Table 6.1 and the left panel of Figure 6.1 show that, whilst $H_k(1)$ gives the better bound for small values of k, asymptotically for $k \to \infty$ the better bound is provided by $G_{k,d}(1)$. We confirm this with the following lemma:

Lemma 6.2.3. *Let W be the Lambert-W function* [CGH+96]. *We have*

(a) $R_k(z)^{1/(k+1)} \sim \mathrm{e}z/(k+1)$ *as* $k \to \infty$, *provided* $z \leq c(k+1)$ *for all large* k *and some* $c \in (0,1)$;

(b) $H_k(1) \sim W(1/\mathrm{e})\,(k+1)$ *as* $k \to \infty$;

(c) $G_{k,d}(1) \sim 1/\mathrm{e}\ (k+1)$ *as* $k \to \infty$;

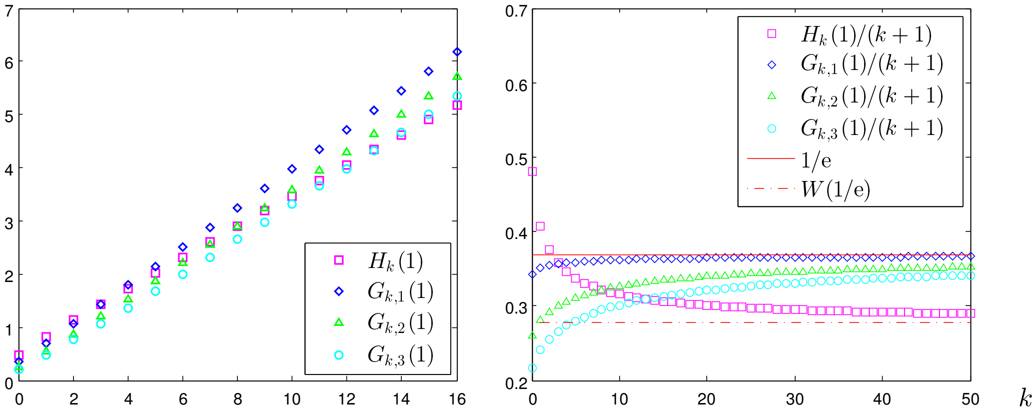

Figure 6.1: The constants in the multi-dimensional density bound (6.2.11) (left) and their asymptotic behaviour (right).

(*d*) $C(k, d) \sim 1/\mathrm{e}\ (k+1)$ *as* $k \to \infty$.

(*Note that* $W(1/\mathrm{e}) \approx 0.2785$, *while* $1/\mathrm{e} \approx 0.3679$.)

Proof. To prove (a), observe that

$$1 - \exp(-z) \sum_{r=0}^{k} \frac{1}{r!} z^r = \frac{\gamma(k+1, z)}{\Gamma(k+1)} = P(k+1, z),$$

where $\gamma(\cdot, \cdot)$ is the lower incomplete Gamma function, and $\Gamma(\cdot)$ is the Gamma function [AS74]. We require an asymptotic expansion of $P(k+1, z)$ as $k \to \infty$ that is uniform in $z \leq c(k+1)$. Such an expansion was obtained by Temme [Tem75, Tem79]. Using the notation of [Tem79], it was shown that

$$P(a, x) = \frac{1}{2}\mathrm{erfc}\left[-\eta(a/2)^{1/2}\right] - S_a(\eta),$$

with

$$S_a(\eta) \sim (2\pi a)^{-1/2} \mathrm{e}^{-a\eta^2/2} \sum_{k=0}^{\infty} c_k(\eta) a^{-k},$$

as $a \to \infty$, uniformly with respect to $\eta \in \mathbb{R}$, where

$$\eta = \sqrt{2(\lambda - 1 - \log(\lambda))}, \quad \lambda = x/a, \quad \mu = \lambda - 1,$$

with the square root having the same sign as μ. Since $x/a < 1$, we have that $\eta < 0$. Here $\mathrm{erfc}(x) = 2\pi^{-1/2} \int_x^\infty \mathrm{e}^{-t^2}\,\mathrm{d}t$ is the complementary error function and $c_k(\eta)$ are functions of η only, with $c_0 = 1/\mu - 1/\eta$. We require only the first term in the asymptotic expansion of $P(a,x)$. Since $\operatorname{erfc}(x) \sim \frac{e^{-x^2}}{\sqrt{\pi}x}$ as $x \to \infty$,

$$P(a,x) \sim -\frac{\exp(-a\eta^2/2)}{\sqrt{2\pi a}\mu} = \frac{\exp(-\lambda a + a)\lambda^a}{\sqrt{2\pi a}(1-\lambda)}, \quad a \to \infty,$$

provided that $\lambda \le c$ for some $c < 1$. Set $k+1 = a$ and $z = x = a\lambda$. Then, since $z \le c(k+1)$ for some $c < 1$, we get

$$R_k(z)^{\frac{1}{k+1}} \sim e\lambda = e\frac{z}{k+1},$$

as $k \to \infty$ and (a) follows.

To prove (b), we shall use (a). Let $z = H_k(1)$, i.e. $h_k(z) = 1$. We first show that there exists a $0 < c < 1$ such that $z \le c(k+1)$ for all large k. Note that $R_k(z) \ge z^{k+1}/(k+1)!$. Thus z satisfies

$$\exp(z)\frac{z^{k+1}}{(k+1)!} \le 1.$$

Therefore

$$\exp\left(\frac{z}{k+1}\right)\frac{z}{k+1} \le \frac{((k+1)!)^{\frac{1}{k+1}}}{k+1}.$$

By Stirling's formula, the right-hand side is asymptotic to $1/e$ as $k \to \infty$. Hence for large k, $z/(k+1) \le W(1/e) < 1$, as required. We may now use (a). Since $h_k(z) = 1$ is equivalent to $\exp\left(\frac{z}{k+1}\right) R_k(z)^{\frac{1}{k+1}} = 1$, this now gives

$$\exp\left(\frac{z}{k+1}\right)\frac{z}{k+1} \sim \frac{1}{e}, \quad k \to \infty.$$

Since the last identity is equivalent to

$$\frac{z}{k+1} \sim W\left(\frac{1}{e}\right), \quad k \to \infty,$$

we get the result.

We use a similar approach to prove (c). Let $z = G_{k,d}(1)$, i.e. $g_{k,d}(z) = 1$. Then

$$R_k(z) \le (1 + 2\sigma_d^*(z))^{d/2}\, e^{z/\sigma_d^*(z)} R_k(z) = g_{k,d}(z) = 1,$$

and we deduce that $z \le \frac{1}{e}(k+1)$ as $k \to \infty$. Hence we may apply (a). Note also that $z \to \infty$ as $k \to \infty$. This follows from the fact that $\lim_{k\to\infty} g_{k,d}(z) = 0$ for fixed z and d. Therefore, the equation $g_{k,d}(z) = 1$ can be written as

$$\left(1 + \frac{4z}{d}\right)^{\frac{d}{2(k+1)}} \exp\left(\frac{d}{2(k+1)}\right)\frac{ez}{k+1} \sim 1, \quad k \to \infty,$$

which implies the result. Finally, we note that claim (d) follows directly from (b) and (c).

□

This lemma confirms that $G_{k,d}(1)$ gives a better bound asymptotically as $k \to \infty$ than $H_k(1)$. Illustration of this asymptotic behaviour is given in the right panel of Figure 6.1. More importantly, this lemma shows the overall advantage of sampling derivatives, i.e. we have the following:

Corollary 6.2.4. *For large k, the set $\{\sqrt{\mu_{n,\alpha}} D^\alpha \Phi_\Omega(\cdot - x_n) : n \in I, |\alpha|_1 \leq k\}$ forms a frame for $\mathrm{B}(\Omega)$ with frame bounds satisfying (6.2.12) and (6.2.13), provided*

$$\delta_* < \frac{1}{\mathrm{e}} \frac{k+1}{2\pi m_\Omega c^*}.$$

Hence, for all dimensions d, the maximum allowed density δ increases linearly with the number of derivatives k. Unfortunately the constant of proportionality $1/\mathrm{e} \approx 0.3679$ is rather small. Indeed, it is much smaller than in the case of equispaced samples, where the corresponding constant is $\pi/2 \approx 1.5708$. To ameliorate this gap, we will first prove an improved estimate in §6.2.2 for the case $d = 1$. Second, in §6.2.4 we will prove a perturbation result for nonuniform derivative sampling.

Remark 6.2.5 For the case $k = 0$, Beurling established the sharp, sufficient condition $\delta < 1/4$ when Ω is the unit Euclidean ball and $|\cdot|_* = |\cdot|$, provided the sampling points $\{x_n\}_{n\in I}$ are separated. In [BW00] and [OU12] this was extended to any compact, convex and symmetric domain Ω, where $|\cdot|_*$ is the norm induced by the radial set of Ω. The separation condition was removed in Theorem 2.3.1 by incorporating weights. To the best of our knowledge, it is an open problem to see if similar sharp results can be proved for the case of sampling with derivatives.

Proof of Theorem 6.2.1

The proof of this theorem uses the techniques of [Grö92, Grö99, Grö01], and more recently [AGH15b], which were applied to the $d \geq 1$ and $k = 0$ case, as well as the approach in [Raz95] for the $d = 1$ and $k \geq 0$ case. We first require the following three lemmas. In what follows, we denote Euclidean ball of radius r centred at v by $\mathcal{B}(v, r)$, and when centre does not matter, we write $\mathcal{B}_r$.

Lemma 6.2.6. *Let $\mu_n = \mu_{n,0}$, where $\mu_{n,0} = \mathrm{meas}(V_n)$ is the Lebesgue measure of Voronoi region V_n. Then*

$$\forall f \in \mathrm{B}(\Omega), \quad \sum_{n\in I} \mu_n |f(x_n)|^2 \leq \exp(4\pi c^* \delta_* r)\|f\|^2,$$

where c^* is as in (6.2.6) and $r > 0$ is radius of the smallest ball (with arbitrary centre) such that $\Omega \subseteq \mathcal{B}_r$.

Proof. Let $\mathcal{B}(\omega_t, r)$ be the minimal ball such that $\Omega \subseteq \mathcal{B}(\omega_t, r)$ and note that we can use the following shifting argument. For every $f \in \mathrm{B}(\Omega)$, if $F \in \mathrm{B}(\Omega - \omega_t)$ is defined so that $\hat{F}(\omega) = \hat{f}(\omega + \omega_t)$, then we have $|F| = |f|$ and also $\|F\| = \|f\|$. Therefore, without loss of generality, we may assume that $\Omega \subseteq \mathcal{B}(0, r)$. Since a bandlimited function is analytic, by Taylor's theorem we have

$$f(x_n) = \sum_{\alpha \in \mathbb{N}_0^d} \frac{(x_n - x)^\alpha}{\alpha!} D^\alpha f(x),$$

for any $n \in I$ and $x \in \mathbb{R}^d$. Let $c > 0$ be a constant. By the Cauchy–Schwarz inequality

$$|f(x_n)|^2 \leq \sum_{\alpha \in \mathbb{N}_0^d} \frac{|(x - x_n)^{2\alpha}| c^{|\alpha|_1}}{\alpha!} \sum_{\alpha \in \mathbb{N}_0^d} \frac{c^{-|\alpha|_1}}{\alpha!} |D^\alpha f(x)|^2.$$

By the multinomial formula (2.2.5),

$$\begin{aligned}
\sum_{\alpha \in \mathbb{N}_0^d} \frac{|(x - x_n)^{2\alpha}| c^{|\alpha|_1}}{\alpha!} &= \sum_{k=0}^{\infty} \frac{c^k}{k!} \sum_{|\alpha|_1 = k} \frac{k!}{\alpha!} |(x - x_n)^{2\alpha}| \\
&= \sum_{k=0}^{\infty} \frac{c^k}{k!} |x - x_n|_2^{2k} \\
&= \exp(c|x - x_n|_2^2).
\end{aligned}$$

By (6.2.6) and the definition of δ, we have $|x - x_n|_2 \leq c^* |x - x_n|_* \leq c^* \delta_*$. Hence we find that

$$|f(x_n)|^2 \leq \exp(c(c^*\delta_*)^2) \sum_{\alpha \in \mathbb{N}_0^d} \frac{c^{-|\alpha|_1}}{\alpha!} |D^\alpha f(x)|^2.$$

Using definition of μ_n and the fact that Voronoi cells form a partition of $\mathbb{R}^d$, this now gives

$$\begin{aligned}
\sum_{n \in I} \mu_n |f(x_n)|^2 &\leq \exp(c(c^*\delta_*)^2) \sum_{\alpha \in \mathbb{N}_0^d} \frac{c^{-|\alpha|_1}}{\alpha!} \sum_{n \in I} \int_{V_n} |D^\alpha f(x)|^2 \, \mathrm{d}x \\
&= \exp(c(c^*\delta_*)^2) \sum_{\alpha \in \mathbb{N}_0^d} \frac{c^{-|\alpha|_1}}{\alpha!} \|D^\alpha f\|^2.
\end{aligned}$$

Consider now the sum. First note that

$$D^\alpha f(x) = \int_\Omega (\mathrm{i}2\pi\omega)^\alpha \hat{f}(\omega) \mathrm{e}^{\mathrm{i}2\pi\omega x} \, \mathrm{d}\omega,$$

and hence

$$\|D^\alpha f\|^2 = \|\widehat{D^\alpha f}\|^2 = \int_\Omega (\mathrm{i}2\pi\omega)^{2\alpha}|\hat{f}(\omega)|^2\,\mathrm{d}\omega.$$

Therefore, by multinomial formula, we get

$$\begin{aligned}
\sum_{\alpha\in\mathbb{N}_0^d} \frac{c^{-|\alpha|_1}}{\alpha!}\|D^\alpha f\|^2 &= \sum_{\alpha\in\mathbb{N}_0^d} \frac{c^{-|\alpha|_1}}{\alpha!}\int_\Omega (2\pi\omega)^{2\alpha}|\hat{f}(\omega)|^2\,\mathrm{d}\omega \\
&= \sum_{k=0}^\infty \frac{c^{-k}}{k!}\int_\Omega \sum_{|\alpha|_1=k}\frac{k!}{\alpha!}(2\pi\omega)^{2\alpha}|\hat{f}(\omega)|^2\,\mathrm{d}\omega \\
&= \sum_{k=0}^\infty \frac{c^{-k}}{k!}\int_\Omega (2\pi)^{2k}(\omega_1^2+\ldots+\omega_d^2)^k|\hat{f}(\omega)|^2\,\mathrm{d}\omega \\
&= \int_\Omega \exp((2\pi|\omega|)^2/c)|\hat{f}(\omega)|^2\,\mathrm{d}\omega \\
&\leq \exp((2\pi r)^2/c)\|f\|^2.
\end{aligned}$$

Hence, we deduce that

$$\sum_{n\in I}\mu_n|f(x_n)|^2 \leq \exp(c(c^*\delta_*)^2 + (2\pi r)^2/c)\|f\|^2,$$

and setting $c = 2\pi r/(c^*\delta_*)$ gives the result. □

Later we will see that application of this lemma leads (after some additional work) to the bound $\delta_* < H_k(1)/(2\pi m_\Omega c^*)$. As discussed, this does not give the best scaling as $k\to\infty$, which can be traced to the exponential growth in δ of the bound obtained in this lemma. In order to mitigate this growth, and therefore eventually get a better density bound, we need the following result.

Lemma 6.2.7. *Let $\mu_n = \mu_{n,0}$, where $\mu_{n,0} = \mathrm{meas}(V_n)$ is the Lebesgue measure of Voronoi region V_n. Then for all $f\in \mathrm{B}(\Omega)$*

$$\sum_{n\in I}\mu_n|f(x_n)|^2 \leq (1+2\sigma_d^*(2\pi c^*\delta_* m_\Omega))^d \exp(4\pi c^*\delta_* m_\Omega/\sigma_d^*(2\pi c^*\delta_* m_\Omega))\|f\|^2,$$

where σ_d^*, c^* and m_Ω are as in (6.2.10), (6.2.6) and (6.2.5) respectively.

Proof. Let $\sigma > 0$ be fixed and let us cover Ω by $R = R(\Omega,\mathcal{B}_{m_\Omega/\sigma})$ Euclidean balls of radius m_Ω/σ. By using a classical result on covering numbers, see for example [FR13], we have

$$R \leq R(\mathcal{B}_{m_\Omega},\mathcal{B}_{m_\Omega/\sigma}) = R(\mathcal{B}_1,\mathcal{B}_{1/\sigma}) \leq (1+2\sigma)^d. \tag{6.2.14}$$

Let

$$\left\{\mathcal{B}^1_{m_\Omega/\sigma},\ldots,\mathcal{B}^R_{m_\Omega/\sigma}\right\}$$

be the prescribed cover of R balls for Ω. Using this cover, we form a partition of Ω as follows. Set $\Omega_1 = \mathcal{B}^1_{m_\Omega/\sigma} \cap \Omega$, and given $\Omega_1, \ldots, \Omega_r$, define

$$\Omega_{r+1} = \left(\mathcal{B}^{r+1}_{m_\Omega/\sigma} \cap \Omega\right) \setminus \bigcup_{j=1}^{r} \Omega_j.$$

This gives at most R nonempty sets $\Omega_1, \ldots, \Omega_R$ which make a disjoint cover of Ω. By construction, for each j, $\Omega_j \subseteq \mathcal{B}^j_{m_\Omega/\sigma}$. Due to Lemma 6.2.6, we know that

$$\sum_{n\in I} \mu_n |g(x_n)|^2 \leq \exp(4\pi c^* \delta_* m_\Omega/\sigma)\|g\|^2, \quad \forall g \in \mathrm{B}(\Omega_j), \quad j = 1, \ldots, R.$$

Since $\Omega_1, \ldots, \Omega_R$ are disjoint and $\bigcup_{j=1}^R \Omega_j = \Omega$, for each $f \in \mathrm{B}(\Omega)$ we have that $\hat{f} = \sum_{j=1}^R \hat{f}_j$, $f = \sum_{j=1}^R f_j$ and $\|f\|^2 = \sum_{j=1}^R \|f_j\|^2$, where $f_j \in \mathrm{B}(\Omega_j)$. Therefore we get

$$\sum_{n\in I} \mu_n |f(x_n)|^2 \leq R \sum_{j=1}^{R} \sum_{n\in I} \mu_n |f_j(x_n)|^2 \leq (1+2\sigma)^d \exp(4\pi c^* \delta_* m_\Omega/\sigma)\|f\|^2.$$

Now, if we minimize the right-hand side over $\sigma > 0$, denoting $z = 2\pi c^* \delta_* m_\Omega$, we obtain

$$\sum_{n\in I} \mu_n |f(x_n)|^2 \leq \left(1 + 2\frac{z + \sqrt{z(d+z)}}{d}\right)^d \exp\left(\frac{2zd}{z + \sqrt{z(d+z)}}\right) \|f\|^2$$

and the result follows. □

Lemma 6.2.8. *For for any $f \in \mathrm{B}(\Omega)$, we have*

$$\left\| f - \sum_{n\in I} \sum_{|\alpha|_1 \leq k} \frac{1}{\alpha!} D^\alpha f(x_n)(\cdot - x_n)^\alpha \chi_{V_n} \right\| \leq \min\left\{h_k(2\pi c^* \delta_* m_\Omega), g_{k,d}(2\pi c^* \delta_* m_\Omega)\right\} \|f\|$$

where h_k and $g_{k,d}$ are as in (6.2.7) and (6.2.8), and c^* and m_Ω are as in (6.2.6) and (6.2.5) respectively.

Proof. For $f \in \mathrm{B}(\Omega)$ let

$$g(x) = \sum_{n\in I} \sum_{|\alpha|_1 \leq k} \frac{1}{\alpha!} D^\alpha f(x_n)(x - x_n)^\alpha \chi_{V_n}(x), \quad x \in \mathbb{R}^d.$$

Since Voronoi cells form a partition of $\mathbb{R}^d$, we have

$$\|f - g\|^2 = \sum_{n\in I} \int_{V_n} \left| f(x) - \sum_{|\alpha|_1 \leq k} \frac{1}{\alpha!} D^\alpha f(x_n)(x - x_n)^\alpha \right|^2 \mathrm{d}x.$$

Let $x \in V_n$. By Taylor's theorem and the Cauchy–Schwarz inequality

$$\begin{aligned}
&\left| f(x) - \sum_{|\alpha|_1 \le k} \frac{1}{\alpha!} D^\alpha f(x_n)(x-x_n)^\alpha \right|^2 \\
&= \left| \sum_{|\alpha|_1 > k} \frac{1}{\alpha!} D^\alpha f(x_n)(x-x_n)^\alpha \right|^2 \\
&\le \sum_{|\alpha|_1 > k} \frac{c^{|\alpha|_1} |(x-x_n)^{2\alpha}|}{\alpha!} \sum_{|\alpha|_1 > k} \frac{c^{-|\alpha|_1}}{\alpha!} |D^\alpha f(x_n)|^2.
\end{aligned}$$

Note that

$$\sum_{|\alpha|_1 > k} \frac{c^{|\alpha|_1} |(x-x_n)^{2\alpha}|}{\alpha!} = \sum_{r>k} \frac{c^r}{r!} |x-x_n|_2^{2r} \le R_k(c(c^*\delta_*)^2),$$

where R_k is as in (6.2.9). Hence, by Lemma 6.2.6 applied to $D^\alpha f \in \mathrm{B}(\Omega)$,

$$\begin{aligned}
\|f-g\|^2 &\le R_k(c(c^*\delta_*)^2) \sum_{|\alpha|_1 > k} \frac{c^{-|\alpha|_1}}{\alpha!} \sum_{n \in I} \mu_n |D^\alpha f(x_n)|^2 \\
&\le R_k(c(c^*\delta_*)^2) \exp(4\pi c^* \delta_* m_\Omega) \sum_{|\alpha|_1 > k} \frac{c^{-|\alpha|_1}}{\alpha!} \|D^\alpha f\|^2.
\end{aligned}$$

Noting that

$$\sum_{|\alpha|_1 > k} \frac{c^{-|\alpha|_1}}{\alpha!} \|D^\alpha f\|^2 = \int_\Omega \sum_{|\alpha|_1 > k} \frac{c^{-|\alpha|_1}}{\alpha!} (2\pi\omega)^{2\alpha} |\hat{f}(\omega)|^2 \, \mathrm{d}\omega \le R_k((2\pi m_\Omega)^2/c) \|f\|^2$$

and setting $c = 2\pi m_\Omega / (c^* \delta_*)$ gives

$$\|f-g\| \le R_k(2\pi c^* \delta_* m_\Omega) \exp(2\pi c^* \delta_* m_\Omega) \|f\| = h_k(2\pi c^* \delta_* m_\Omega) \|f\|.$$

Similarly, if we apply Lemma 6.2.7 instead of Lemma 6.2.6, we get

$$\begin{aligned}
\|f-g\| &\le (1 + 2\sigma_d^*(2\pi c^* \delta_* m_\Omega))^{\frac{d}{2}} R_k(2\pi c^* \delta_* m_\Omega) \exp\left(\frac{2\pi c^* \delta_* m_\Omega}{\sigma_d^*(2\pi c^* \delta_* m_\Omega)} \right) \|f\| \\
&= g_{k,d}(2\pi c^* \delta_* m_\Omega) \|f\|,
\end{aligned}$$

and the result follows. $\square$

Now we are ready to prove Theorem 6.2.1.

Proof of Theorem 6.2.1. Fix $f \in \mathrm{B}(\Omega)$ and let

$$g(x) = \sum_{n \in I} \sum_{|\alpha|_1 \leq k} \frac{1}{\alpha!} D^\alpha f(x_n)(x - x_n)^\alpha \chi_{V_n}(x), \quad x \in \mathbb{R}^d.$$

Then for the upper bound on $\|g\|^2$ we have

$$\sum_{n \in I} \int_{V_n} \left| \sum_{|\alpha|_1 \leq k} \frac{1}{\alpha!} D^\alpha f(x_n)(x - x_n)^\alpha \right|^2 \mathrm{d}x \leq \left(\sum_{|\alpha|_1 \leq k} \frac{1}{\alpha!} \right) \sum_{n \in I} \sum_{|\alpha|_1 \leq k} \mu_{n,\alpha} |D^\alpha f(x_n)|^2.$$

By the multinomial formula

$$\sum_{|\alpha|_1 \leq k} \frac{1}{\alpha!} = \sum_{l=0}^{k} \frac{1}{l!} \sum_{|\alpha|_1 = l} \frac{l!}{\alpha!} = \sum_{l=0}^{k} \frac{d^l}{l!} \leq \mathrm{e}^d.$$

Using this we get

$$\sum_{n \in I} \sum_{|\alpha|_1 \leq k} \mu_{n,\alpha} |D^\alpha f(x_n)|^2 \geq \mathrm{e}^{-d} \|g\|^2 \geq \mathrm{e}^{-d} \left(\|f\| - \|f - g\| \right)^2. \tag{6.2.15}$$

Lemma 6.2.8 now gives the lower bound. Next, we address the upper bound. Note that

$$\mu_{n,\alpha} \leq \frac{1}{\alpha!} \sup_{x \in V_n} |(x - x_n)^{2\alpha}| \mu_{n,0}.$$

Moreover $|(x - x_n)^{2\alpha}| \leq |x - x_n|_\infty^{2|\alpha|_1} \leq |x - x_n|_2^{2|\alpha|_1} \leq (c^* \delta_*)^{2|\alpha|_1}$. Hence, Lemma 6.2.6 gives

$$\sum_{n \in I} \sum_{|\alpha|_1 \leq k} \mu_{n,\alpha} |D^\alpha f(x_n)|^2 \leq \exp(4\pi m_\Omega c^* \delta_*) \sum_{|\alpha|_1 \leq k} \frac{(c^* \delta_*)^{2|\alpha|_1}}{\alpha!} \|D^\alpha f\|^2.$$

Arguing in the same way now yields

$$\sum_{n \in I} \sum_{|\alpha|_1 \leq k} \mu_{n,\alpha} |D^\alpha f(x_n)|^2 \leq \exp(4\pi m_\Omega c^* \delta_* + (2\pi m_\Omega c^* \delta_*)^2) \|f\|^2,$$

as required. $\square$

6.2.2 The univariate case

In the one-dimensional setting it is possible to improve the bound derived in Theorem 6.2.1 somewhat using so-called Wirtinger inequalities. See [Grö92] for the case $k = 0$ and [Raz95] for $k = 1$.

Throughout this section $\Omega \subseteq \mathbb{R}$ is compact and $\{x_n\}_{n \in \mathbb{Z}}$ is a set of sampling points in $\mathbb{R}$, indexed over $\mathbb{Z}$. We assume the points are ordered so that $x_n < x_{n+1}$, $\forall n \in \mathbb{Z}$. As before, we let

$$\delta = \sup_{x\in\mathbb{R}} \inf_{n\in\mathbb{Z}} |x - x_n|, \tag{6.2.16}$$

where $|\cdot|$ denotes the absolute value. Note that the Voronoi cells V_n are the intervals

$$V_n = [z_n, z_{n+1}], \quad z_n = \frac{x_n + x_{n-1}}{2}, \qquad n \in \mathbb{Z}.$$

As stated above, we shall use Wirtinger inequalities to derive bounds for δ. Specifically, for $k \in \mathbb{N}$, let $c_k > 0$ be the minimal constant such that

$$\int_a^b |f(x)|^2 \,\mathrm{d}x \leq (c_k)^{2k}(b-a)^{2k} \int_a^b |f^{(k)}(x)|^2 \,\mathrm{d}x, \tag{6.2.17}$$

for all $f \in \mathrm{H}^k(a,b)$, the k^{th} Sobolev space, satisfying

$$f(a) = f'(a) = \ldots = f^{(k-1)}(a) = 0 \quad \text{or} \quad f(b) = f'(b) = \ldots = f^{(k-1)}(b) = 0.$$

Theorem 6.2.9. *Suppose that the weights are*

$$\mu_{n,l} = \frac{1}{l!}\int_{V_n} (x-x_n)^{2l} \,\mathrm{d}x = \frac{(z_{n+1}-x_n)^{2l+1} - (z_n - x_n)^{2l+1}}{l!(2l+1)}, \; l = 0,\ldots,k, \; n \in \mathbb{Z},$$

and let δ be as in (6.2.16). If

$$\delta < \frac{C(k)}{2\pi m_\Omega}, \qquad C(k) = \frac{1}{c_{k+1}}, \tag{6.2.18}$$

where c_k is as in (6.2.17), then

$$\forall f \in \mathrm{B}(\Omega), \quad A\|f\|^2 \leq \sum_{n\in\mathbb{Z}} \sum_{l=0}^{k} \mu_{n,l} |f^{(l)}(x_n)|^2 \leq B\|f\|^2,$$

where

$$A \geq \mathrm{e}^{-1}\left(1 - (2\pi m_\Omega c_{k+1}\delta)^{k+1}\right)^2, \quad B \leq (1 + 4m_\Omega\delta)^2 \exp((2\pi m_\Omega \delta)^2).$$

Equivalently, the set $\{\sqrt{\mu_{n,l}}\frac{\mathrm{d}^l}{\mathrm{d}x^l}\Phi_\Omega(\cdot - x_n) : n \in \mathbb{Z}, l = 0,\ldots,k\}$ forms a frame for $\mathrm{B}(\Omega)$ with the frame bounds A and B.

In the following section we examine the constants c_k and conclude by discussing the improvement offered by this theorem over the multivariate result Theorem 6.2.1.

Proof. We follow the arguments of [Grö99, Raz95]. Let

$$g(x) = \sum_{n\in\mathbb{Z}} \sum_{l=0}^{k} \frac{1}{l!} f^{(l)}(x_n)(x - x_n)^l \chi_{V_n}(x), \quad x \in \mathbb{R}.$$

Then

$$\begin{aligned} \|f - g\|^2 &= \sum_{n\in\mathbb{Z}} \int_{z_n}^{z_{n+1}} \left| f(x) - \sum_{l=0}^{k} \frac{1}{l!} f^{(l)}(x_n)(x - x_n)^l \right|^2 \mathrm{d}x. \\ &= \sum_{n\in\mathbb{Z}} \left(\int_{x_n}^{z_{n+1}} + \int_{z_n}^{x_n} \right) \left| f(x) - \sum_{l=0}^{k} \frac{1}{l!} f^{(l)}(x_n)(x - x_n)^l \right|^2 \mathrm{d}x. \end{aligned}$$

The function $f(x) - \sum_{l=0}^{k} \frac{1}{l!} f^{(l)}(x_n)(x - x_n)^l$ vanishes, along with its first k derivatives, at $x = x_n$. Applying (6.2.17) to each integral and noting that $|z_{n+1} - x_n| \le \delta$ and $|x_n - z_n| \le \delta$ gives

$$\|f - g\|^2 \le (c_{k+1}\delta)^{2k+2} \|f^{(k+1)}\|^2.$$

Observe that for all $f \in \mathrm{B}(\Omega)$ the Bernstein inequality reads

$$\|D^\alpha f\| \le (2\pi\bar{\omega})^\alpha \|f\|,$$

where $\bar{\omega} = (\bar{\omega}_1, \ldots, \bar{\omega}_d)^\top$ and $\bar{\omega}_j = \sup_{\omega\in\Omega} |\omega_j|$. Additionally, in one dimension, $\bar{\omega} = m_\Omega$. Therefore, by applying Bernstein's inequality we deduce that

$$\|f - g\| \le (2\pi m_\Omega c_{k+1}\delta)^{k+1} \|f\|,$$

and hence

$$\left(1 - (2\pi m_\Omega c_{k+1}\delta)^{k+1}\right) \|f\| \le \|g\| \le \left(1 + (2\pi m_\Omega c_{k+1}\delta)^{k+1}\right) \|f\|.$$

We now use this and (6.2.15) to get the estimate for A. For the bound on B, we argue similarly to the proof of Theorem 6.2.1. We have

$$\mu_{n,l} \le \frac{1}{l!} \sup_{x\in V_n} |x - x_n|^{2l} \mu_{n,0} \le \frac{1}{l!} \delta^{2l} \mu_{n,0}.$$

Hence

$$\sum_{n\in\mathbb{Z}} \sum_{l=0}^{k} \mu_{n,l} |f^{(l)}(x_n)|^2 \le \sum_{l=0}^{k} \frac{\delta^{2l}}{l!} \sum_{n\in\mathbb{Z}} \mu_{n,0} |f^{(l)}(x_n)|^2.$$

Gröchenig's result [Grö92] for $k = 0$ and $d = 1$, i.e. Theorem 2.1.6, gives that

$$\forall g \in \mathrm{B}(\Omega), \quad \sum_{n\in\mathbb{Z}} \mu_{n,0}|g(x_n)|^2 \leq (1+4\delta m_\Omega)^2\|g\|^2.$$

By this and Bernstein's inequality, we deduce that

$$\sum_{n\in\mathbb{Z}}\sum_{l=0}^{k} \mu_{n,l}|f^{(l)}(x_n)|^2 \leq (1+4\delta m_\Omega)^2 \sum_{l=0}^{k} \frac{(2\pi m_\Omega \delta)^{2l}}{l!}\|f\|^2.$$

Since

$$\sum_{l=0}^{k} \frac{(2\pi m_\Omega \delta)^{2l}}{l!} = \exp((2\pi m_\Omega\delta)^2),$$

the upper bound follows. □

Observe that for $k = 0$, i.e. the classical nonuniform sampling problem without derivatives, (6.2.18) reduces to $\delta < 1/(4m_\Omega)$ since $c_1 = 2/\pi$ [Grö92]. This is in agreement with the result of Gröchenig [Grö92], which is stated here in Theorem 2.1.6. This result is sharp, and says that one must sample at a rate just above the Nyquist rate $1/(4m_\Omega)$.

The magnitude of c_k

We now consider the constant c_k of Wirtinger's inequality (6.2.17) when $k \geq 1$. We first note the following:

Lemma 6.2.10. *Consider the polyharmonic eigenvalue problem*

$$(-1)^k g^{(2k)} = \lambda g, \quad g(0) = \ldots = g^{(k-1)}(0) = g^{(k)}(1) = \ldots g^{2k-1}(1) = 0. \tag{6.2.19}$$

This problem has a countable basis of positive eigenvalues $0 < \lambda_1^{(k)} < \lambda_2^{(k)} < \ldots$. Moreover, the best constant c_k in the inequality (6.2.17) is precisely $(\lambda_1^{(k)})^{-\frac{1}{2k}}$.

Proof. It is well known that (6.2.19) has a countable spectrum with eigenfunctions $\{\phi_n\}_{n=1}^\infty$ forming an orthonormal basis of $L^2(0,1)$ [Nai68]. It is straightforward to see that (6.2.19) has only strictly positive eigenvalues. Now let $f \in \mathrm{H}^k(0,1)$ satisfy $f(0) = \ldots = f^{(k-1)}(0) = 0$. Then

$$\langle f, \phi_n\rangle = \frac{(-1)^k}{\lambda_n^{(k)}}\langle f, \phi_n^{(2k)}\rangle = \frac{1}{\lambda_n^{(k)}}\langle f^{(k)}, \phi_n^{(k)}\rangle.$$

In particular, if $f = \phi_n$, then $\|\phi_n\|^2 = \frac{1}{\lambda_n^{(k)}}\|\phi_n^{(k)}\|^2$. Let $\psi_n = \frac{1}{\sqrt{\lambda_n^{(k)}}}\phi_n^{(k)}$, so that $\|\psi_n\| = 1$. The set $\{\psi_n\}_{n=1}^\infty$ is precisely the set of eigenfunctions of the problem

$$(-1)^k g^{(2k)} = \lambda g, \quad g^{(k)}(0) = \ldots = g^{(2k-1)}(0) = g(1) = \ldots g^{k-1}(1) = 0.$$

k	1	2	3	4	5	6	7	8	9	10
c_k	0.6366	0.5333	0.4495	0.3861	0.3376	0.2997	0.2694	0.2446	0.2240	0.2066
$1/c_k$	1.5708	1.8751	2.2248	2.5903	2.9621	3.3367	3.7125	4.0888	4.4652	4.8415

Table 6.2: The values c_k and $1/c_k$ for $k = 1, 2, \ldots, 10$. These values were calculated in high precision using *Mathematica*.

In particular, they form an orthonormal basis of $\mathrm{L}^2(0,1)$. Therefore, since $\langle f, \phi_n \rangle = \frac{1}{\sqrt{\lambda_n^{(k)}}} \langle f^{(k)}, \psi_n \rangle$, it follows from Parseval's identity that

$$\|f\|^2 = \sum_n |\langle f, \phi_n \rangle|^2 = \sum_n \frac{1}{\lambda_n^{(k)}} |\langle f^{(k)}, \psi_n \rangle|^2 \leq \frac{1}{\lambda_1^{(k)}} \sum_n |\langle f^{(k)}, \psi_n \rangle|^2 = \frac{1}{\lambda_1^{(k)}} \|f^{(k)}\|^2,$$

by completeness. Thus $\|f\|^2 \leq 1/\lambda_1^{(k)} \|f^{(k)}\|^2$, and this bound is sharp since we may set $f = \phi_1$. By a change of variables, we get that $(c_k)^{2k} = 1/\lambda_1^{(k)}$, as required. $\square$

This means we can determine the constant c_k by finding the eigenvalues of (6.2.19). When $k = 1$, the eigenvalues of (6.2.19) are $(\pi/2 + n\pi)^2$, $n \in \mathbb{N}_0$. Hence $\lambda_1^{(1)} = \pi^2/4$ and $c_1 = 2/\pi$, as stated. Unfortunately, for $k \geq 2$ no explicit expression exists for the eigenvalues, so we resort to numerical computation. For $k \geq 2$, write $\lambda = \tau^{2k}$ for $\tau > 0$. The general solution of (6.2.19) can be written as

$$g(x) = \sum_{s=0}^{2k-1} b_s \mathrm{e}^{\mathrm{i} z^s \tau x},$$

where $z = \mathrm{e}^{\mathrm{i}\pi/k}$ and $b_s \in \mathbb{C}$ are coefficients. Enforcing the boundary conditions results in a linear system of equations

$$\sum_{s=0}^{2k-1} (\mathrm{i} z^s \tau)^r b_s = 0, \quad \sum_{s=0}^{2k-1} (\mathrm{i} z^s \tau)^{k+r} \mathrm{e}^{\mathrm{i} z^s \tau} b_s = 0, \qquad r = 0, \ldots, k-1.$$

In matrix form, we have $A(\tau)b = 0$, where $A(\tau) \in \mathbb{C}^{2k \times 2k}, b = (b_0, \ldots, b_{2k-1})^\top$. Hence the minimal eigenvalue $\lambda_1^{(k)} = (\tau_1^{(k)})^{2k}$, and therefore $c_k = 1/\tau_1^{(k)}$, where $\tau_1^{(k)}$ is the first positive root of the function $D(\tau) = \det(A(\tau))$. In the case $k = 2$, we have $D(\tau) = 8\mathrm{i}\tau^6 (1 + \cos(\tau)\cosh(\tau))$, and numerical computation finds that $\tau_1^{(2)} = 1.8751$ (see also [Raz95]).

In Table 6.2 we compute $\tau_1^{(k)} = 1/c_k$ and c_k for $k = 1, \ldots, 10$. As is evident the values $1/c_k$, grow approximately linearly in k for large k. Linear regression on the computed values gives that $1/c_k \approx 1.1458 + 0.3674k$ for large k. Note that $1/\mathrm{e} = 0.3679$. We

k	0	1	2	3	4	5	6	7	8	9
(a)	0.4812	0.8141	1.1268	1.4304	1.7890	2.1535	2.5186	2.8842	3.2501	3.6163
(b)	1.5708	1.8751	2.2248	2.5903	2.9621	3.3367	3.7125	4.0888	4.4652	4.8415
(c)	1.4142	1.8612	2.2209	2.5886	2.9612	3.3361	3.7121	4.0885	4.4650	4.8413

Table 6.3: The constant $C(k)$ obtained from (a) Theorem 6.2.1 for the case $d = 1$, (b) Theorem 6.2.9 and (c) [Raz95, Thm. 1].

therefore conjecture that

$$\frac{1}{c_k} \sim \frac{1}{\mathrm{e}}(k+1), \quad k \to \infty. \tag{6.2.20}$$

We remark in passing that the large k asymptotics for the optimal constant in a variant of Wirtinger's inequality where f and its derivatives vanish at both endpoints has been derived by Böttcher & Widom [BW07]. We expect a similar approach can be applied to (6.2.17) to obtain (6.2.20).

We can now compare Theorem 6.2.9 with the multivariate result Theorem 6.2.1. In Table 6.3 we give the numerical values for the constant $C(k)$ arising from both theorems, where $\delta < C(k)/(2\pi m_\Omega)$ is the required condition on δ. The univariate bound is evidently superior for all values of k considered. However, the bounds behave the same asymptotically, since both Theorem 6.2.1 and Theorem 6.2.9 give $C(k) \sim 1/\mathrm{e}\ (k+1) \approx 0.3679\ (k+1)$ for large k (recall Corollary 6.2.4). In Table 6.3 we also compare Theorem 6.2.9 to the bound derived in [Raz95, Thm. 1] (note that the value 1.8751 for $k = 1$ was also provided in [Raz95] using Wirtinger's inequality arguments as we do above). Unfortunately, the improvement obtained from Theorem 6.2.9 is only marginal. In particular, both bounds are asymptotic to $1/\mathrm{e}\ (k+1)$ for large k, and therefore (we expect) a long way from being sharp (recall that the condition for equispaced samples is $\delta \leq (\pi/2\ (k+1))/(2\pi m_\Omega)$). We conclude that although Wirtinger's inequality obtains a sharp bound for $k = 0$, it is of little use in getting superior bounds for $k \geq 1$.

6.2.3 Line-by-line sampling

In some applications, not least seismology, the unknown function f depends on a spatial variable $z \in \mathbb{R}^{d-1}$ and a temporal variable $t \in \mathbb{R}$. Sensors are placed at fixed locations $\{z_n\}_{n\in I} \subseteq \mathbb{R}^{d-1}$, where $d = 2, 3$, in physical space, and measurements are taken at times $\{t_{m,n}\}_{m\in J}$. In particular, different sensors may take measurements at different times. This gives the set of samples

$$D_z^\alpha f(z_n, t_{m,n}), \quad n \in I, m \in J, |\alpha|_1 \leq k.$$

Note that $D_z^\alpha = \partial_{z_1}^{\alpha_1} \cdots \partial_{z_{d-1}}^{\alpha_{d-1}}$ is the partial derivative with respect to z only. We do not measure any temporal derivatives.

Let $x = (z,t) \in \mathbb{R}^d$ and write $f(z,t) = f(x)$. We shall assume that $f \in \mathrm{B}(\Omega)$ and moreover that $\Omega = \Omega_z \times \Omega_t$ for $\Omega_z \subseteq \hat{\mathbb{R}}^{d-1}$ and $\Omega_t \subseteq \hat{\mathbb{R}}$. Let

$$\delta_{z,*} = \sup_{z \in \mathbb{R}^{d-1}} \inf_{n \in I} |z - z_n|_*, \qquad \delta_t = \sup_{n \in I} \sup_{t \in \mathbb{R}} \inf_{m \in J} |t - t_{m,n}|,$$

and write $V_n \subseteq \mathbb{R}^{d-1}$ for the Voronoi cells of the sampling points $\{z_n\}_{n \in I}$ with respect to the $|\cdot|_*$ norm. We now have the following result. Note that this is a straightforward extension of a result of Strohmer [Str00a] (see also [Grö01]) to the case of derivatives and $d \geq 3$.

Proposition 6.2.11. *Suppose that the weights*

$$\mu_{m,n,\alpha} = \frac{t_{m+1,n} - t_{m,n}}{2\alpha!} \int_{V_n} (z - z_n)^{2\alpha} \, \mathrm{d}z.$$

If

$$\delta_t < \frac{1}{4 m_{\Omega_t}}, \quad \delta_{z,*} < \frac{C(k,d)}{2\pi m_{\Omega_z} c^*}, \qquad C(k,d) = \begin{cases} 1/c_{k+1} & d = 2 \\ \max\{H_k(1), G_{k,d}(1)\} & d \geq 3 \end{cases},$$

then for all $f \in \mathrm{B}(\Omega)$

$$\begin{aligned} (1 - 4\delta_t m_{\Omega_t})^2 A_z \|f\|^2 &\leq \sum_{m \in J} \sum_{n \in I} \sum_{|\alpha|_1 \leq k} \mu_{m,n,\alpha} |D_z^\alpha f(z_n, t_{m,n})|^2 \\ &\leq (1 + 4\delta_t m_{\Omega_t})^2 B_z \|f\|^2, \end{aligned}$$

where A_z and B_z satisfy

$$A_z \geq \mathrm{e}^{-1} \left(1 - (2\pi m_{\Omega_z} c_{k+1} \delta_{z,*})^{k+1}\right)^2, \quad B_z \leq (1 + 4 m_{\Omega_z} \delta_{z,*})^2 \, \mathrm{e}^{(2\pi m_{\Omega_z} \delta_{z,*})^2}, \quad d = 2,$$

with c_k as in (6.2.17), or

$$\begin{aligned} A_z &\geq \mathrm{e}^{-d} \left(1 - \min\{h_k(2\pi m_{\Omega_z} c^* \delta_{z,*}), g_{k,d}(2\pi m_{\Omega_z} c^* \delta_{z,*})\}\right)^2, \\ B_z &\leq \exp(4\pi m_{\Omega_z} c^* \delta_{z,*} + (2\pi m_{\Omega_z} c^* \delta_{z,*})^2), \end{aligned}$$

for $d \geq 3$, with h_k and $g_{k,d}$ as in (6.2.7) and (6.2.8) with inverse functions H_k and $G_{k,d}$ respectively. Equivalently, the set $\{\sqrt{\mu_{m,n,\alpha}} D_z^\alpha \Phi_\Omega(\cdot - x_{n,m}) : n \in I, m \in J, |\alpha|_1 \leq k\}$, where $x_{n,m} = (z_n, t_m)$, forms a frame for $\mathrm{B}(\Omega)$ with bounds

$$A \geq (1 - 4\delta_t m_{\Omega_t})^2 A_z, \qquad B \leq (1 + 4\delta_t m_{\Omega_t})^2 B_z.$$

Proof. Gröchenig's original, one-dimensional, derivative-free result from [Grö92] (see Theorem 2.1.6) gives that

$$(1-4\delta_t m_{\Omega_t})^2 \int_{\mathbb{R}} |f(z,t)|^2 \,\mathrm{d}t \leq \sum_{m\in J} \frac{t_{m+1,n}-t_{m,n}}{2} |f(z,t_{m,n})|^2$$
$$\leq (1+4\delta_t m_{\Omega_t})^2 \int_{\mathbb{R}} |f(z,t)|^2 \,\mathrm{d}t.$$

Hence, if $g(z) = \sqrt{\int_{\mathbb{R}} |f(z,t)|^2 \,\mathrm{d}t}$ and $\tilde{\mu}_{n,\alpha} = \frac{1}{\alpha!}\int_{V_n} (z-z_n)^{2\alpha} \,\mathrm{d}z$ then

$$(1-4\delta_t m_{\Omega_t})^2 \sum_{n\in I}\sum_{|\alpha|_1\leq k} \tilde{\mu}_{n,\alpha} |D_z^\alpha g(z_n)|^2 \leq \sum_{m\in J}\sum_{n\in I}\sum_{|\alpha|_1\leq k} \mu_{m,n,\alpha} |D_z^\alpha f(z_n,t_{m,n})|^2$$
$$\leq (1+4\delta_t m_{\Omega_t})^2 \sum_{n\in I}\sum_{|\alpha|_1\leq k} \tilde{\mu}_{n,\alpha} |D_z^\alpha g(z_n)|^2$$

and, to get the result, we now apply Theorem 6.2.1 ($d \geq 3$) or Theorem 6.2.9 ($d = 2$) to the sum and note that $\int_{\mathbb{R}^{d-1}} |g(z)|^2 \,\mathrm{d}z = \|f\|^2$. □

This proposition implies the following. With the above type of scheme, for stable sampling one requires (i) the usual derivative-free density for univariate nonuniform sampling in the time variable, i.e. $\delta_t < 1/(4m_{\Omega_t})$, and (ii) a density in the space variable depending on the number of derivatives.

6.2.4 A multivariate perturbation result with derivatives

The results proved thus far give explicit guarantees for nonuniform derivatives sampling. However, the conditions on the density δ are more stringent than those required for uniform samples. We now show that nonuniform sampling is possible with larger gaps under appropriate conditions.

Theorem 6.2.12. *Suppose that $\{x_n\}_{n\in I} \subseteq \mathbb{R}^d$ and $\mu_{n,\alpha} > 0$, $n \in I$, $|\alpha|_1 \leq k$, are such that (6.2.1) holds with constants $A, B > 0$. Let $\{\tilde{x}_n\}_{n\in I} \subseteq \mathbb{R}^d$ be such that*

$$\epsilon_* = \sup_{n\in I} |\tilde{x}_n - x_n|_* < \frac{\ln(1+\sqrt{A/B})}{2\pi m_\Omega c^*}, \tag{6.2.21}$$

then

$$\forall f \in \mathrm{B}(\Omega), \quad \tilde{A}\|f\|^2 \leq \sum_{n\in I}\sum_{|\alpha|_1\leq k} \mu_{n,\alpha} |D^\alpha f(\tilde{x}_n)|^2 \leq \tilde{B}\|f\|^2,$$

where

$$\tilde{A} \geq \left(\sqrt{A} - \sqrt{B}\,(\exp(2\pi m_\Omega c^* \epsilon_*) - 1)\right)^2, \qquad \tilde{B} \leq B\exp(4\pi m_\Omega c^* \epsilon_*).$$

That is, if the set $\{\sqrt{\mu_{n,\alpha}} D^\alpha \Phi_\Omega(\cdot - x_n) : n \in I, |\alpha|_1 \leq k\}$ *forms a frame for* $\mathrm{B}(\Omega)$ *with bounds* A *and* B*, then the set* $\{\sqrt{\mu_{n,\alpha}} D^\alpha \Phi_\Omega(\cdot - \tilde{x}_n) : n \in I, |\alpha|_1 \leq k\}$ *forms a frame for* $\mathrm{B}(\Omega)$ *with bounds* $\tilde{A}$ *and* $\tilde{B}$*.*

Proof. The proof is similar to those of the earlier results. Note first that by Minkowski inequality

$$\begin{aligned}
&\sqrt{\sum_{n \in I} \sum_{|\alpha|_1 \leq k} \mu_{n,\alpha} |D^\alpha f(\tilde{x}_n)|^2} \\
&\geq \sqrt{\sum_{n \in I} \sum_{|\alpha|_1 \leq k} \mu_{n,\alpha} |D^\alpha f(x_n)|^2} - \sqrt{\sum_{n \in I} \sum_{|\alpha|_1 \leq k} \mu_{n,\alpha} |D^\alpha f(x_n) - D^\alpha f(\tilde{x}_n)|^2}.
\end{aligned}$$

By identical arguments to those used in §6.2.1, we have

$$|g(x_n) - g(\tilde{x}_n)|^2 \leq \big(\exp(c(c^* \epsilon_*)^2) - 1\big) \sum_{|\beta|_1 > 0} \frac{c^{-|\beta|_1}}{\beta!} |D^\beta g(x_n)|^2,$$

for any function $g \in \mathrm{B}(\Omega)$. Using this, we deduce that

$$\begin{aligned}
&\sum_{n \in I} \sum_{|\alpha|_1 \leq k} \mu_{n,\alpha} |D^\alpha f(x_n) - D^\alpha f(\tilde{x}_n)|^2 \\
&\leq \big(\exp(c(c^* \epsilon_*)^2) - 1\big) \sum_{|\beta|_1 > 0} \frac{c^{-|\beta|_1}}{\beta!} \sum_{n \in I} \sum_{|\alpha|_1 \leq k} \mu_{n,\alpha} |D^\alpha D^\beta f(x_n)|^2 \\
&\leq B \big(\exp(c(c^* \epsilon_*)^2) - 1\big) \sum_{|\beta|_1 > 0} \frac{c^{-|\beta|_1}}{\beta!} \|D^\beta f\|^2 \\
&\leq B \big(\exp(c(c^* \epsilon_*)^2) - 1\big) \big(\exp((2\pi m_\Omega)^2 / c) - 1\big) \|f\|^2.
\end{aligned}$$

Setting $c = 2\pi m_\Omega / (c^* \epsilon_*)$ gives

$$\tilde{A} \geq \left(\sqrt{A} - \sqrt{B} \left(\exp(2\pi m_\Omega c^* \epsilon_*) - 1\right)\right)^2.$$

Hence, $\tilde{A} > 0$ provided that $\sqrt{A} - \sqrt{B} \left(\exp(2\pi m_\Omega c^* \epsilon_*) - 1\right) > 0$. Now, rearranging gives (6.2.21). The upper bound for $\tilde{B}$ follows similarly. □

As with the previous results, the right-hand side (6.2.21) is dimensionless whenever Ω is contained in the unit ball and $|\cdot|_* = |\cdot|_q$, $1 \leq q \leq 2$. Now suppose for simplicity that $\Omega \subseteq [-1,1]^d$. Then the points $x_n = (k+1)n/2$, $n \in \mathbb{Z}^d$, give rise to a stable set of sampling (this is due to the fact that they give rise to a Riesz basis for $\Omega = [-1,1]^d$, and therefore a frame when $\Omega \subseteq [-1,1]^d$). This theorem therefore allows for nonuniform samples with gaps roughly on the size of k, provided the sampling points $\tilde{x}_n$ are within ϵ_* of the x_n. An issue with this result is that the ratio A/B is liable to decrease with both

k and d. Hence, the maximal allowed ϵ_* may be rather small in practice. See [Raw89] for the one-dimensional case.

In [Bai10, Cor. 6.1], a multivariate perturbation result for the case $k = 0$ with $x_n = n/2$ was derived based on similar arguments. In our notation, the result proved therein corresponds to the case $p = q = \infty$. The precise condition given is $\epsilon_* < \ln 2/(2\pi d)$, which is equivalent to (6.2.21) with $k = 0$. Note that Sun & Zhou [SZ99] also prove a perturbation result in the same setting $p = q = \infty$, but based on expanding in Laplace–Neumann eigenfunctions, rather than Taylor series (this is similar to the proof of the original Kadec-1/4 theorem). Their constant is somewhat smaller than $\ln 2/(2\pi d)$ for finite d, but, as discussed in [Bai10], it is asymptotic to $\ln 2/(2\pi d)$ as $d \to \infty$. The generalizations of these results offered by Theorem 6.2.12 are:

(i) flexibility over the choice of domain Ω—in particular, a dimension-independent bound for appropriate Ω and $|\cdot|_*$, and

(ii) the case when derivatives are sampled, i.e. $k \neq 0$.

In [ARAK09], perturbation results are proved for a more general sampling model that includes derivatives sampling of bandlimited functions as a special case. However, for this particular case [ARAK09, Thm 3.8], the perturbation bound is not explicit, and additionally, it assumes separation of the sampling points.

6.3 Univariate nonuniform bunched sampling

We now consider nonuniform sampling with sampling points clustered in bunches. Given the difficulty of polynomial interpolation for $d \geq 2$ dimensions, we consider the univariate case only.

6.3.1 Problem statement

Assume that we are given samples at some nonuniform points $\{x_{n,0}\}_{n \in I} \subseteq \mathbb{R}$ which are δ-dense

$$\delta = \sup_{x \in \mathbb{R}} \inf_{n \in I} |x - x_{n,0}|, \tag{6.3.1}$$

and let V_n denotes the Voronoi region associated to $x_{n,0}$. Moreover, for each $n \in I$, we are given s additional samples inside each of the Voronoi region, namely s additional samples at distinct points

$$x_{n,m} \in [x_{n,0} - h_n, x_{n,0} + h_n] \subseteq V_n, \quad m = 1, \ldots, s, \tag{6.3.2}$$

which can be also nonuniform. If we denote

$$h = \sup_{n \in I} h_n, \tag{6.3.3}$$

then, by definition, there exists a positive constant $\tau \leq 1$ such that

$$h = \tau\delta.$$

Therefore, in each h-vicinity of $x_{n,0}$, there are s additional sampling points. We shall call such a sampling sequence

$$\{x_{n,m}\}_{n \in I, 0 \leq m \leq s}$$

a *bunched set* with the density δ defined by (6.3.1) and the bunch width h defined by (6.3.3). We are interested in a behaviour of the permitted density δ in terms of the bunch cardinality s and the bunch width h (or τ), while ensuring a (fusion) frame.

Much as in the case of derivatives sampling, in bunched sampling, we expect that a larger δ is possible if there are multiple sample points around each $x_{n,0}$. As discussed earlier, it is useful to have this type of sampling scheme in the situations where we must allow for bigger distances between sampling sensors due to some natural constraints.

6.3.2 Bunched sampling and fusion frames

In nonuniform derivative sampling, we showed the existence of a particular frame to establish stable sampling. In the case of bunched sampling, we will first show the existence of a particular *fusion* frame [CK04, CKL08]. We recall that a non-orthogonal fusion frame [CCL12] for a Hilbert space H is a set of positive scalars $\{v_n\}_{n \in I}$ and non-orthogonal projections $\{\mathcal{P}_n\}_{n \in I}$, each with closed range, satisfying

$$\forall f \in \mathrm{H}, \quad A\|f\|^2 \leq \sum_{n \in I} v_n \|\mathcal{P}_n f\|^2 \leq B\|f\|^2.$$

Much like a frame operator, the associated *fusion frame operator* $\mathcal{S} : \mathrm{H} \to \mathrm{H}$ given by

$$\mathcal{S}f = \sum_{n \in I} \mathcal{P}_n^* \mathcal{P}_n f$$

is linear, bounded, self-adjoint and invertible. Thus, any $f \in \mathrm{H}$ can be recovered stably from the data $\{\mathcal{P}_n f\}_{n \in I}$. In practice, if the projections have finite-dimensional ranges, using the results of Chapter 3, it can be easily seen that the reconstruction can be carried out via generalized sampling, for example.

Given the bunched set $\{x_{n,m}\}_{n \in I, 0 \leq m \leq s}$ and associated Voronoi regions $\{V_n\}_{n \in I}$, for each $n \in I$ we define the subspace

$$W_n = \{g \in \mathrm{L}^2(\mathbb{R}) : \operatorname{supp}(g) \subseteq V_n\}$$

and also for any $f \in \mathrm{B}(\Omega)$ we define the operator

$$\mathcal{P}_n(f) = p_n(f)\chi_{V_n} \tag{6.3.4}$$

where $p_n(f) \in \mathbb{P}_s$ is the unique interpolating polynomial of degree s such that

$$p_n(f)(x_{n,m}) = f(x_{n,m}), \quad m = 0, \ldots, s.$$

The bounded linear operator $\mathcal{P}_n : \mathrm{B}(\Omega) \to W_n$ is a non-orthogonal projection, i.e. $\mathcal{P}_n^2 = \mathcal{P}_n$ by uniqueness of the interpolating polynomial. Hence, if there exist $A, B > 0$ such that for all $f \in \mathrm{B}(\Omega)$

$$A\|f\|^2 \leq \sum_{n\in I} \|\mathcal{P}_n(f)\|^2 \leq B\|f\|^2,$$

then $\{\mathcal{P}_n\}_{n\in I}$ is a non-orthogonal fusion frame for $\mathrm{B}(\Omega)$ with weights $v_n = 1$. Our main result gives conditions for this to be the case:

Theorem 6.3.1. *Suppose that $\{x_{n,m}\}_{n\in I, 0\leq m\leq s} \subseteq \mathbb{R}$ is a bunched set with density δ and bunch width $h = \tau\delta$, where $\tau \in (0, 1]$. If*

$$\delta < \frac{\tilde{H}_{s,\tau}(1)}{2\pi m_\Omega}, \tag{6.3.5}$$

where $\tilde{H}_{s,\tau}$ is the inverse function of

$$\tilde{h}_{s,\tau}(z) = \frac{(1+\tau)^s z^{s+1}}{(s+1)!}\left(1 + \frac{4z}{\pi}\right), \quad z \in (0, \infty),$$

then

$$\forall f \in \mathrm{B}(\Omega), \quad A\|f\|^2 \leq \sum_{n\in I} \|\mathcal{P}_n(f)\|^2 \leq B\|f\|^2,$$

where $\mathcal{P}_n(f)$ are given by (6.3.4) and

$$\begin{aligned} A &\geq \left(1 - \frac{(1+\tau)^s(2\pi\delta m_\Omega)^{s+1}}{(s+1)!}(1 + 8\delta m_\Omega)\right)^2, \\ B &\leq \left(1 + \frac{(1+\tau)^s(2\pi\delta m_\Omega)^{s+1}}{(s+1)!}(1 + 8\delta m_\Omega)\right)^2. \end{aligned} \tag{6.3.6}$$

Equivalently, the family $\{\mathcal{P}_n\}_{n\in I}$ is a non-orthogonal fusion frame for $\mathrm{B}(\Omega)$ with weights $v_n = 1$.

Proof. Let $g(x) = \sum_{n\in I} \mathcal{P}_n(f)(x)$. Then

$$\|g\|^2 = \int_{\mathbb{R}} \left| \sum_{n\in I} p_n(f)(x)\chi_{V_n}(x) \right|^2 \, \mathrm{d}x = \sum_{n\in I} \int_{V_n} |p_n(f)(x)|^2 \, \mathrm{d}x = \sum_{n\in I} \|\mathcal{P}_n(f)\|^2.$$

Since f is a bandlimited function, it is infinitely continuously differentiable. Also, since for each $n \in I$, $p_n(f)(x)$ is a polynomial of degree at most s that interpolates f at $s+1$ distinct points $\{x_{n,m} : m = 0, \dots, s\}$ in the closed interval V_n, a classical result gives that for each $n \in I$ and $x \in V_n$ there exists $\xi_n(x) \in V_n$ such that

$$f(x) - p_n(f)(x) = \frac{f^{(s+1)}(\xi_n(x))}{(s+1)!} \prod_{m=0}^{s} (x - x_{n,m}). \tag{6.3.7}$$

Let $\tilde{x}_n \in V_n$ be such that

$$|f^{(s+1)}(\tilde{x}_n)| = \max_{x\in V_n} |f^{(s+1)}(x)|,$$

which again exists because f is bandlimited. Note that, for all $x \in V_n$, $|x - x_{n,m}| \le \delta + h$ for $m \neq 0$ and $|x - x_{n,m}| \le \delta$ for $m = 0$. Thus, from (6.3.7), for all $x \in V_n$ we have

$$|f(x) - p_n(f)(x)| \le \frac{\left|f^{(s+1)}(\tilde{x}_n)\right|}{(s+1)!}(1+\tau)^s \delta^{s+1}.$$

Therefore

$$\|f-g\|^2 = \sum_{n\in I} \int_{V_n} |f(x) - p_n(f)(x)|^2 \, \mathrm{d}x \le \frac{(1+\tau)^{2s}\delta^{2(s+1)}}{((s+1)!)^2} \sum_{n\in I} \mathrm{meas}(V_n) |f^{(s+1)}(\tilde{x}_n)|^2.$$

By the construction, the points $\{\tilde{x}_n\}_{n\in I}$ are 2δ-dense and $\tilde{x}_n \in V_n$, $n \in I$. Hence, by adapting the proof of Gröchenig's one-dimensional result [Grö92] for $s = 0$ (to account for the fact that $\tilde{x}_n \neq x_{n,0}$), we get

$$\|f - g\| \le \frac{(1+\tau)^s \delta^{s+1}}{(s+1)!} (1 + 8\delta m_\Omega) (2\pi m_\Omega)^{s+1} \|f\|.$$

The result now follows immediately. □

The constant $\tilde{H}_{s,\tau}(1)$ in the density bound obtained by this theorem is explicitly calculated for different values of s and τ in Table 6.4. The asymptotic result is given in the following corollary:

Corollary 6.3.2. *For large s, if*

$$\delta < \frac{1}{(1+\tau)\mathrm{e}} \frac{s+1}{2\pi m_\Omega},$$

s	0	1	2	3	4	5	6	7	8	9
$\tilde{H}_{s,1}(1)$	0.5766	0.7218	0.8894	1.0626	1.2382	1.4151	1.5928	1.7710	1.9497	2.1287
$\tilde{H}_{s,1/2}(1)$	0.5766	0.8101	1.0458	1.2820	1.5187	1.7558	1.9934	2.2314	2.4696	2.7082
$\tilde{H}_{s,1/4}(1)$	0.5766	0.8710	1.1578	1.4426	1.7270	2.0115	2.2963	2.5815	2.8671	3.1531
$\tilde{H}_{s,1/8}(1)$	0.5766	0.9080	1.2275	1.5440	1.8597	2.1754	2.4914	2.8079	3.1248	3.4422
$\tilde{H}_{s,1/16}(1)$	0.5766	0.9287	1.2669	1.6017	1.9357	2.2696	2.6039	2.9387	3.2740	3.6099

Table 6.4: The constant in the bunched sampling density bound (6.3.5).

the set $\{\mathcal{P}_n\}_{n\in I}$ is a non-orthogonal fusion frame for $\mathrm{B}(\Omega)$ with weights $v_n = 1$ and frame bounds as in (6.3.6).

Proof. Let $z = \tilde{H}_{s,\tau}(1)$, i.e. $\tilde{h}_{s,\tau}(z) = 1$. This gives

$$\frac{z}{s+1}(1+\tau)^{1-\frac{1}{s+1}}(1+4z/\pi)^{\frac{1}{s+1}} = \frac{((s+1)!)^{\frac{1}{s+1}}}{s+1}.$$

Therefore

$$\tilde{H}_{s,\tau}(1) \sim \frac{s+1}{(1+\tau)\mathrm{e}}$$

as $s \to \infty$. □

By choosing a different form of the interpolation polynomial in (6.3.4), we get different families of fusion frames. In particular, for the *Lagrange* form of the interpolation polynomial the operator (6.3.4) becomes

$$\mathcal{P}_n(f)(x) = \sum_{m=0}^{s} f(x_{n,m})L_{n,m}(x)\chi_{V_n}(x),$$

where $L_{n,m}$ are Lagrange polynomials given by

$$L_{n,m}(x) = \frac{R_{n,m}(x)}{R_{n,m}(x_{n,m})}, \quad R_{n,m}(x) = \prod_{\substack{0\le j\le s\\ j\ne m}} (x - x_{n,j}), \tag{6.3.8}$$

and therefore, for the fusion frame operator we have

$$\mathcal{S}(f)(t) = \sum_{n\in I}\sum_{m=0}^{s}\sum_{l=0}^{s}\left(\int_{V_n} L_{n,m}(x)L_{n,l}(x)\,\mathrm{d}x\right) f(x_{n,l})\Phi_\Omega(t - x_{n,m}).$$

On the other hand, if we use the *Newton* form of the interpolation polynomial, we have

$$\mathcal{P}_n(f)(x) = \sum_{m=0}^{s} D_{x_{n,0},\ldots,x_{n,m}} f N_{n,m}(x)\chi_{V_n}(x), \tag{6.3.9}$$

where $D_{x_{n,0},\dots,x_{n,m}}f$ denotes divided difference of the function f at $x_{n,0},\dots,x_{n,m}$ and $N_{n,m}$ is Newton polynomial given by

$$N_{n,m}(x)=\prod_{l=0}^{m-1}(x-x_{n,l}). \tag{6.3.10}$$

The fusion frame operator in this case is

$$\mathcal{S}(f)(t)=\sum_{n\in I}\sum_{m=0}^{s}\sum_{l=0}^{s}\left(\int_{V_n}N_{n,m}(x)N_{n,l}(x)\,\mathrm{d}x\right)D_{x_{n,0},\dots,x_{n,l}}fD_{x_{n,0},\dots,x_{n,m}}\Phi_\Omega(t-\cdot).$$

Moreover, this approach allows us to consider the following more general sampling scenario. Suppose that we are additionally given k derivatives at the points of the bunched set $\{x_{n,m}\}_{n\in I,0\leq m\leq s}$, i.e. the given data is

$$f^{(j)}(x_{n,m}),\quad n\in I,\ m=0,\dots,s,\ j=0,\dots,k.$$

Now, for each $n\in I$, we can define the unique interpolation polynomial $p_n(f)$ such that

$$p_n^{(j)}(f)(x_{n,m})=f^{(j)}(x_{n,m}),\quad m=0,\dots,s,\ j=0,\dots,k.$$

In this case, we can use the *Hermite* form of the interpolation polynomial and set

$$\mathcal{P}_n(f)(x)=\sum_{j=0}^{k}\sum_{m=0}^{s}f^{(j)}(x_{n,m})c_{n,m,j}(x)\chi_{V_n}(x),$$

where

$$c_{n,m,j}(x)=L_{n,m}^{k+1}(x)\frac{(x-x_{n,m})^j}{j!}\sum_{i=0}^{k-j}\frac{(x-x_{n,m})^i}{i!}R_{n,m}^{k+1}(x_{n,m})\frac{d^i}{dx^i}R_{n,m}^{-(k+1)}(x_{n,m}),$$

and $L_{n,m}$, $R_{n,m}$ are as in (6.3.8), see [Tra64]. Since the error term (6.3.7) now reads as

$$f(x)-\mathcal{P}_n(f)(x)=\frac{f^{((s+1)(k+1))}(\xi_n(x))}{((s+1)(k+1))!}\prod_{m=0}^{s}(x-x_{n,m})^{k+1},$$

we obtain an additional $k+1$ factor in the density bound, i.e. the density condition now reads

$$\frac{(1+\tau)^{s(k+1)}(2\pi\delta m_\Omega)^{(s+1)(k+1)}}{(s+1)!(k+1)!}(1+8\delta m_\Omega)<1,$$

which for large s and large k leads to

$$\delta<\frac{1}{(1+\tau)\mathrm{e}}\frac{(s+1)(k+1)}{2\pi m_\Omega}.$$

Thus, a combination of bunched and derivative sampling increases the maximal allowed density by a multiplicative factor of $s+1$ (number of bunched points) and $k+1$ (number of derivatives).

6.3.3 Bunched sampling and frames

It transpires that the use of the Newton form of the interpolating polynomial also allows one to relate bunched sampling to a frame, as opposed to a fusion frame. Let us define $\mathcal{P}_n$ as in (6.3.9). Since the divided difference $D_{x_{n,0},\ldots,x_{n,m}} f$ is just a linear combination of the function f evaluated at the points $x_{n,0}, \ldots, x_{n,m}$ and since $f(x) = \langle f(t), \Phi_\Omega(t-x)\rangle$ with Φ_Ω defined by (6.2.2), we can write

$$D_{x_{n,0},\ldots,x_{n,m}} f = \langle f, \phi_{n,m}\rangle, \quad \phi_{n,m}(t) = D_{x_{n,0},\ldots,x_{n,m}} \Phi_\Omega(t-\cdot). \tag{6.3.11}$$

We now have the following:

Theorem 6.3.3. *Suppose that $\{x_{n,m}\}_{n\in I, 0\le m\le s} \subseteq \mathbb{R}$ is the bunched set with density δ and bunch width $h = \tau\delta$, where $\tau \in (0,1]$. Let $\{V_n\}_{n\in I}$ be the Voronoi regions corresponding to the points $\{x_{n,0}\}_{n\in I}$. If*

$$\delta < \frac{\tilde{H}_{s,\tau}(1)}{2\pi m_\Omega},$$

where $\tilde{H}_{s,\tau}$ is as in Theorem 6.3.1, then

$$\forall f \in \mathrm{B}(\Omega), \quad A\|f\|^2 \le \sum_{n\in I}\sum_{m=0}^{s} \mu_{n,m} \left|D_{x_{n,0},\ldots,x_{n,m}} f\right|^2 \le B\|f\|^2, \tag{6.3.12}$$

where $\mu_{n,m} = m! \int_{V_n} |N_{n,m}(x)|^2 \,\mathrm{d}x$, $N_{n,m}$ are given by (6.3.10) and

$$A \ge \frac{1}{\mathrm{e}}\left(1 - \frac{(1+\tau)^s (2\pi\delta m_\Omega)^{s+1}}{(s+1)!}(1+8\delta m_\Omega)\right)^2, \tag{6.3.13}$$

$$B \le \frac{(1+8(1+\tau)\delta m_\Omega)^2\, \mathrm{e}^{((1+\tau)2\pi m_\Omega \delta)^2}}{(1+\tau)^2}. \tag{6.3.14}$$

Equivalently, if $\phi_{n,m}$ is as in (6.3.11), the set $\left\{\sqrt{\mu_{n,m}}\phi_{n,m} : n \in I, m = 0,\ldots,s\right\}$ is a frame for $\mathrm{B}(\Omega)$.

Proof. As before, let

$$g(x) = \sum_{n\in I}\sum_{m=0}^{s} D_{x_{n,0},\ldots,x_{n,m}} f N_{n,m}(x)\chi_{V_n}(x).$$

Now we have

$$\begin{aligned}\|g\|^2 &= \sum_{n\in I}\int_{V_n}\left|\sum_{m=0}^{s} D_{x_{n,0},\ldots,x_{n,m}}f N_{n,m}(x)\right|^2 \,\mathrm{d}x\\ &\le \sum_{m=0}^{s}\frac{1}{m!}\sum_{n\in I}\sum_{m=0}^{s} m!\left(\int_{V_n}|N_{n,m}(x)|^2\,\mathrm{d}x\right)\left|D_{x_{n,0},\ldots,x_{n,m}}f\right|^2\end{aligned}$$

and hence

$$\sum_{n\in I}\sum_{m=0}^{s}\mu_{n,m}\left|D_{x_{n,0},\ldots,x_{n,m}}f\right|^2 \ge \mathrm{e}^{-1}\left(\|f\|-\|f-g\|\right)^2.$$

In the proof of Theorem 6.3.1 we obtained

$$\|f-g\| \le \frac{(1+\tau)^s(2\pi\delta m_\Omega)^{s+1}}{(s+1)!}\left(1+8\delta m_\Omega\right)\|f\|,$$

and therefore for the lower frame bound we get

$$A \ge \mathrm{e}^{-1}\left(1-\frac{(1+\tau)^s(2\pi\delta m_\Omega)^{s+1}}{(s+1)!}\left(1+8\delta m_\Omega\right)\right)^2.$$

For the upper frame bound first note that

$$\mu_{n,m} = m!\int_{V_n}\left|\prod_{l=0}^{m-1}(x-x_{n,l})\right|^2\,\mathrm{d}x \le m!(1+\tau)^{2(m-1)}\delta^{2m}\mathrm{meas}(V_n).$$

Since $f\in \mathrm{B}(\Omega)$ is infinitely differentiable, from the mean value theorem for divided differences, for any $n\in I$ and any $m\le s$, there exists $\tilde{x}_{n,m}\in\langle x_{n,0},\ldots,x_{n,m}\rangle$ such that

$$D_{x_{n,0},\ldots,x_{n,m}}f = \frac{1}{m!}f^{(m)}(\tilde{x}_{n,m})$$

where

$$\langle x_{n,0},\ldots,x_{n,m}\rangle = (\min\{x_{n,0},\ldots,x_{n,m}\},\max\{x_{n,0},\ldots,x_{n,m}\}) \subseteq [x_{n,0}-h, x_{n,0}+h]$$

Now, since for each m the points $\{\tilde{x}_{n,m}\}_{n\in I}$ are $(1+\tau)\delta$-dense, as before, by adapting Gröchenig's one-dimensional result, we obtain

$$\begin{aligned}\sum_{n\in I}\sum_{m=0}^{s}\mu_{n,m}\left|D_{x_{n,0},\ldots,x_{n,m}}f\right|^2 &\le \frac{1}{(1+\tau)^2}\sum_{m=0}^{s}\frac{((1+\tau)\delta)^{2m}}{m!}\sum_{n\in I}\mathrm{meas}(V_n)\left|f^{(m)}(\tilde{x}_{n,m})\right|^2\\ &\le \frac{(1+8m_\Omega\delta(1+\tau))^2}{(1+\tau)^2}\sum_{m=0}^{s}\frac{((1+\tau)\delta)^{2m}}{m!}\|f^{(m)}\|^2\end{aligned}$$

$$\leq \frac{\left(1+8m_\Omega\delta(1+\tau)\right)^2 \mathrm{e}^{((1+\tau)\delta 2\pi m_\Omega)^2}}{(1+\tau)^2}\|f\|^2,$$

and the estimate for the upper frame bound follows. □

In the limit, when the bunch width h becomes very small and the number of bunched points s very large, from this proposition we obtain precisely the one-dimensional derivative result given in Theorem 6.2.9 for large number of derivatives k:

Corollary 6.3.4. *For large s and small τ, if*

$$\delta < \frac{1}{\mathrm{e}}\frac{s+1}{2\pi m_\Omega},$$

then $\left\{\sqrt{\mu_{n,m}}\frac{d^m}{dx^m}\Phi_\Omega(\cdot - x_{n,0}) \ : \ \mu_{n,m} = \frac{1}{m!}\int_{V_n}(x-x_{n,0})^{2m}\,\mathrm{d}x, n\in I, m=0,\ldots,s\right\}$ *is a frame for* $\mathrm{B}(\Omega)$ *with the frame bounds satisfying* (6.3.13) *and* (6.3.14).

Proof. Consider the sum (6.3.12) as $\tau \to 0$. For $x_{n,0},\ldots,x_{n,m} \in [x_{n,0}-\tau\delta, x_{n,0}+\tau\delta]$

$$\begin{aligned}\lim_{\tau\to 0}\sum_{n\in I}\sum_{m=0}^{s}&\left(m!\int_{V_n}|N_{n,m}(x)|^2\,\mathrm{d}x\right)\left|D_{x_{n,0},\ldots,x_{n,m}}f\right|^2\\ &=\sum_{n\in I}\sum_{m=0}^{s}\left(\frac{1}{m!}\int_{V_n}(x-x_{n,0})^{2m}\,\mathrm{d}x\right)|f^{(m)}(x_{n,0})|^2.\end{aligned}$$

This holds due to dominated convergence theorem, since for any τ, n and m

$$m!\int_{V_n}|N_{n,m}(x)|^2\,\mathrm{d}x\left|D_{x_{n,0},\ldots,x_{n,m}}f\right|^2 \leq \mathrm{meas}(V_n)(2\delta)^{2m}|f^{(m)}(\tilde{x}_n)|^2,$$

where $\tilde{x}_n \in V_n$ is such that $|f^{(m)}(\tilde{x}_n)| = \max_{x\in V_n}|f^{(m)}(x)|$.

For the density condition, let $z = \tilde{H}_{s,\tau}(1)$. Since $1+\tau \sim 1$ as $\tau \to 0$, this gives

$$\frac{z}{s+1}(1+4z/\pi)^{\frac{1}{s+1}} \sim \frac{((s+1)!)^{\frac{1}{s+1}}}{s+1}, \quad \tau \to 0,$$

and hence $\tilde{H}_{s,\tau}(1) \sim (s+1)/\mathrm{e}$ as $\tau \to 0$ and $s \to \infty$. □

Therefore, for the large number of bunched sampling points s such that the width of all bunches is small, we obtain the same result as when sampling s derivatives.

6.3.4 Bunched sampling as a perturbation of derivative sampling

We have the following result:

Theorem 6.3.5. *Suppose that* $\{x_{n,0}\}_{n\in I} \subseteq \mathbb{R}$ *and* $\mu_{n,m} = \frac{1}{m!}\int_{V_n}(x - x_{n,0})^{2m}\,\mathrm{d}x$, $n \in I$, $m \leq s$, *are such that*

$$\forall f \in \mathrm{B}(\Omega), \quad A\|f\|^2 \leq \sum_{n\in I}\sum_{m=0}^{s} \mu_{n,m}\left|f^{(m)}(x_{n,0})\right|^2 \leq B\|f\|^2 \tag{6.3.15}$$

for some constants $A, B > 0$. *Let* $\{x_{n,m}\}_{n\in I, 0\leq m\leq s} \subseteq \mathbb{R}$ *be the bunched set with bunch width* h *such that*

$$h < \frac{\ln(1 + \sqrt{A/B})}{2\pi m_\Omega}, \tag{6.3.16}$$

then

$$\forall f \in \mathrm{B}(\Omega), \quad \tilde{A}\|f\|^2 \leq \sum_{n\in I}\sum_{m=0}^{s} \tilde{\mu}_{n,m}|D_{x_{n,0},\ldots,x_{n,m}} f|^2 \leq \tilde{B}\|f\|^2,$$

where $\tilde{\mu}_{n,m} = (m!)^2\,\mu_{n,m}$ *and*

$$\tilde{A} \geq \left(\sqrt{A} - \sqrt{B}\,(\exp(2\pi m_\Omega h) - 1)\right)^2, \qquad \tilde{B} \leq B\exp(4\pi m_\Omega h).$$

That is, if the family $\{\sqrt{\mu_{n,m}}\frac{d^m}{dx^m}\Phi_\Omega(\cdot - x_{n,0}) : n \in I, m \leq s\}$ *forms a frame for* $\mathrm{B}(\Omega)$ *with bounds* A *and* B, *then the family* $\left\{\sqrt{\tilde{\mu}_{n,m}}\phi_{n,m} : n \in I, m \leq s\right\}$ *is a frame for* $\mathrm{B}(\Omega)$ *with bounds* $\tilde{A}$ *and* $\tilde{B}$, *where* $\phi_{n,m}$ *is defined by* (6.3.11).

Proof. Since $f \in \mathrm{B}(\Omega)$ is infinitely differentiable, from the mean value theorem for divided differences, for any $n \in I$ and any $m \leq s$, there exists $\tilde{x}_{n,m} \in [x_{n,0} - h, x_{n,0} + h]$ such that

$$D_{x_{n,0},\ldots,x_{n,m}} f = \frac{1}{m!} f^{(m)}(\tilde{x}_{n,m}). \tag{6.3.17}$$

Since also $\tilde{\mu}_{n,m} = (m!)^2\,\mu_{n,m}$, we have

$$\sum_{n\in I}\sum_{m=0}^{s} \tilde{\mu}_{n,m}\left|D_{x_{n,0},\ldots,x_{n,m}} f\right|^2 = \sum_{n\in I}\sum_{m=0}^{s} \mu_{n,m}\left|f^{(m)}(\tilde{x}_{n,m})\right|^2.$$

Note that the sum on the right hand side is not in the scope of Theorem 6.2.12, since the point $\tilde{x}_{n,m}$ changes for every m. However, we can proceed as follows. Since

$$f^{(m)}(\tilde{x}_{n,m}) = f^{(m)}(x_{n,0}) + \sum_{l\geq 1}\frac{1}{l!} f^{(m+l)}(x_{n,0})(\tilde{x}_{n,m} - x_{n,0})^l,$$

by Minkowski's inequality we get

$$\sqrt{\sum_{n\in I}\sum_{m=0}^{s} \tilde{\mu}_{n,m}\left|D_{x_{n,0},\ldots,x_{n,m}} f\right|^2}$$

$$\geq \sqrt{\sum_{n\in I}\sum_{m=0}^{s}\mu_{n,m}\left|f^{(m)}(x_{n,0})\right|^2} - \sqrt{\sum_{n\in I}\sum_{m=0}^{s}\mu_{n,m}\left|\sum_{l\geq 1}\frac{1}{l!}f^{(m+l)}(x_{n,0})(\tilde{x}_{n,m}-x_{n,0})^l\right|^2}.$$

Applying Minkowski's inequality now to the second term and using $\sup_{n\in I}\sup_{m=0,\dots,s}|x_{n,0}-\tilde{x}_{n,m}|\leq h$ and (6.3.15), we get

$$\begin{aligned}
&\sqrt{\sum_{n\in I}\sum_{m=0}^{s}\mu_{n,m}\left|\sum_{l\geq 1}\frac{1}{l!}f^{(m+l)}(x_{n,0})(\tilde{x}_{n,m}-x_{n,0})^l\right|^2}\\
&\leq \sum_{l\geq 1}\sqrt{\sum_{n\in I}\sum_{m=0}^{s}\mu_{n,m}\left|\frac{1}{l!}f^{(m+l)}(x_{n,0})(\tilde{x}_{n,m}-x_{n,0})^l\right|^2}\\
&\leq \sum_{l\geq 1}\frac{h^l}{l!}\sqrt{\sum_{n\in I}\sum_{m=0}^{s}\mu_{n,m}\left|f^{(m+l)}(x_{n,0})\right|^2}\\
&\leq \sqrt{B}\sum_{l\geq 1}\frac{h^l}{l!}\|f^{(l)}\|
\end{aligned}$$

Now, by Cauchy-Schwarz and Bernstein's inequality we derive

$$\begin{aligned}
\sum_{n\in I}\sum_{m=0}^{s}\mu_{n,m}\left|\sum_{l\geq 1}\frac{1}{l!}f^{(m+l)}(x_{n,0})(\tilde{x}_{n,m}-x_{n,0})^l\right|^2 &\leq B\sum_{l\geq 1}\frac{c^l h^{2l}}{l!}\sum_{l\geq 1}\frac{c^{-l}}{l!}\|f^{(l)}\|^2\\
&\leq B\left(\exp\left(2\pi m_\Omega h\right)-1\right)^2\|f\|^2.
\end{aligned}$$

Therefore

$$\sum_{n\in I}\sum_{m=0}^{s}\tilde{\mu}_{n,m}\left|D_{x_{n,0},\dots,x_{n,m}}f\right|^2 \geq \left(\sqrt{A}-\sqrt{B}\left(\exp\left(2\pi m_\Omega h\right)-1\right)\right)^2\|f\|^2,$$

and similarly for the upper bound we obtain

$$\sum_{n\in I}\sum_{m=0}^{s}\tilde{\mu}_{n,m}\left|D_{x_{n,0},\dots,x_{n,m}}f\right|^2 \leq B\exp\left(4\pi m_\Omega h\right)\|f\|^2.$$

Hence, the sequence $\{x_{n,m}\}_{n\in I,m=0,\dots,s}$ gives rise to a frame if the width h satisfies (6.3.16). □

Note that, due to (6.3.17), this theorem implies the perturbation result given by Theorem 6.2.12, but only in the univariate setting. Moreover, this theorem allows a bunched set to be taken at the same density which is allowed for derivative sampling, as long as the width of bunches h satisfies condition (6.3.16).

Chapter 7

Conclusions

The main contribution of this thesis is a general framework for stable reconstruction in arbitrary reconstruction subspaces of multivariate compactly supported L^2 functions from nonuniform Fourier samples. We have shown that a stable reconstruction in any desired reconstruction space is always possible provided the samples are taken sufficiently dense and wide enough in the frequency domain. In general, the sampling scheme Ω needs to satisfy the universal density condition $\delta < 1/4$, whereas the sufficient sampling bandwidth K depends on the reconstruction space T as well as the sampling scheme Ω. For smaller δ's, we have shown that in fact the sampling bandwidth K depends solely on the reconstruction space T. This enabled us to analyse the sufficient scaling of K with $\dim(\mathrm{T})$ for specific choices of T. In particular, for the univariate case where T consists of wavelets or different types of polynomials, we have provided the explicit scaling of K with $\dim(\mathrm{T})$ sufficient for stable and quasi-optimal reconstruction via NUGS.

Closely related to these results, there are several topics left for future work. First, we expect that subject to the universal density condition $\delta < 1/4$, the magnitude of the sampling bandwidth K always depends solely on the reconstruction space T. Indeed, we have shown this to be true in the univariate case. However in the multivariate case, currently, we require a more stringent density condition $\delta < (\ln 2)/(2\pi m_D c_*)$. Improvement of this multi-dimensional δ-condition is left for future work. Associated to this issue is improvement of our results for weighted Fourier frames. Although the weighted Fourier frame bounds are explicitly estimated in the case of smaller densities than previously known, it remains an open problem to explicitly estimate the frame bounds for even smaller densities, closer to condition $\delta < 1/4$.

Second, there is a question of the sufficient sampling bandwidth K for specific reconstruction spaces T within the multivariate setting. We expect that our univariate results for wavelets and different polynomials extend to higher dimensions. In higher dimensions, it would be also important to analyse other reconstruction spaces, such as curvelets and shearlets. Moreover, it would be interesting to analyse the stability barrier for all these different reconstruction spaces in terms of the smallest necessary scaling of K with $\dim(\mathrm{T})$ required for stability. Note that, in the univariate setting, we have shown the stability barrier for wavelets: the linear scaling of K with $\dim(\mathrm{T})$ is necessary for stability via any reconstruction method from nonuniform samples. This is an extension of the result shown in [AHP14] for the special case of uniform samples. In the uniform case and within the univariate setting, in [AHS14], it was also shown that the quadratic scaling for polynomials is in fact necessary, providing the stability barrier for reconstruction in polynomials from uniform samples. We expect this to extend to the nonuniform case as well.

Recall that in this work the sampling scheme Ω is considered fixed. This situation arises in applications such as MRI, where Ω is often specified by physical constraints, e.g. magnetic gradients, noise etc. However, in many applications, one may have substantial flexibility to design Ω so as to optimize the reconstruction quality. That is, for a given subspace T, one seeks to design Ω as small as possible whilst keeping the reconstruction constant $C(\Omega, \mathrm{T})$ below a desired maximum value. This question is closely related to the existence of Marcinkiewicz–Zygmund inequalities (see [CZ99, Mar07, OCS07] and references therein), which have been well-researched for certain choices of $\hat{\mathrm{T}}$ (e.g. trigonometric polynomials, spherical harmonics,...). On the other hand, designing good (or perhaps even optimal) sampling schemes for families of wavelet subspaces, for example, remains an open problem, but one of practical interest.

This work does not address the issue of sparsity. Sparsity-exploiting algorithms are currently revolutionizing signal and image reconstruction. Since our main focus were wavelets, in which images are known to be sparse, it may at first sight appear strange not to seek to exploit such properties. For uniform samples this has indeed been done by using the aforementioned GS–CS framework, and the results are reported in [AH15a, AHPR14]. However, as was explained in [AH15a] (see also [AHRT14]), before one can exploit sparsity it is first necessary to understand the underlying linear mapping between the samples and coefficients in the reconstruction system, which is precisely what we do in this work. Exploiting sparsity by extending the work of [AHPR14] to the case of fully nonuniform Fourier samples is a topic of future investigations.

In this thesis, we have additionally presented several density bounds as sufficient guarantees for stable recovery of bandlimited functions when the measurement set includes samples of the first k derivatives. In particular, we have proved the linear growth of δ-density with $k+1$. However, the constant of proportionality $1/\mathrm{e}$ is rather small compared to the case of equispaced samples where the corresponding constant is $\pi/2$. Therefore, it would be of interest to see how these bounds can be improved in both the univariate and multivariate case.

As we have seen, a related problem to derivatives sampling is so-called bunched sampling. This sampling strategy also leads to increased δ-bound and, asymptotically, it approximates the derivatives sampling. Much as in the derivative case, it remains open to improve this density bound. Also, it would be important to generalize these results to the multivariate case and therefore broaden the range of their applications. Let us note that in higher dimensions, well-posedness of the bunched points and the possibility of constructing an unique multivariate interpolation polynomial complicates dramatically. Therefore, it is not trivial to extend the techniques used here to the multivariate case and we leave this problem for future investigations.

One might notice that in the last part of the thesis, Chapter 6, we have analysed two examples—derivatives and bunched sampling—both appearing at the end of Papoulis' paper [Pap77a]. Although these examples are of interest in applications by themselves, the remaining problem is to analyse a general setting given in Papoulis' paper in the context of nonuniform sampling. Namely, it remains open to see what happens with the sampling density when instead of $H_\alpha(\omega) = (-\mathrm{i}2\pi\omega)^\alpha$ one has more general functions H_α and a nonuniform set of sampling points.

Bibliography

[AG00] A. Aldroubi and K. Gröchenig. Beurling-Landau-type theorems for non-uniform sampling in shift invariant spline spaces. *J. Fourier Anal. Appl.*, 6(1):93–103, 2000. (Cited on page 9).

[AG01] A. Aldroubi and K. Gröchenig. Nonuniform sampling and reconstruction in shift-invariant spaces. *SIAM Rev.*, 43:585–620, 2001. (Cited on pages 1, 2, 8, 9, and 16).

[AGH14a] B. Adcock, M. Gataric, and A. C. Hansen. On stable reconstructions from nonuniform Fourier measurements. *SIAM J. Imaging Sci.*, 7(3):1690–1723, 2014. (Cited on pages 7, 29, and 53).

[AGH14b] B. Adcock, M. Gataric, and A. C. Hansen. Recovering piecewise smooth functions from nonuniform Fourier measurements. *Accepted in Proceedings of the 10th International Conference on Spectral and High Order Methods, and to appear in Springer Lecture Notes in Computational Science and Engineering*, 2014. (Cited on pages 7 and 96).

[AGH15a] B. Adcock, M. Gataric, and A. Hansen. Density theorems for nonuniform sampling of bandlimited functions using derivatives or bunched measurements. *Submitted*, 2015. (Cited on pages 7 and 105).

[AGH15b] B. Adcock, M. Gataric, and A. C. Hansen. Weighted frames of exponentials and stable recovery of multidimensional functions from nonuniform Fourier samples. *To appear in Applied and Computational Harmonic Analysis*, 2015. (Cited on pages 7, 11, and 113).

[AH12a] B. Adcock and A. C. Hansen. A generalized sampling theorem for stable reconstructions in arbitrary bases. *J. Fourier Anal. Appl.*, 18(4):685–716, 2012. (Cited on pages 2 and 30).

[AH12b] B. Adcock and A. C. Hansen. Stable reconstructions in Hilbert spaces and the resolution of the Gibbs phenomenon. *Appl. Comput. Harmon. Anal.*, 32(3):357–388, 2012. (Cited on pages 3, 5, 36, and 101).

[AH15a] B. Adcock and A. C. Hansen. Generalized Sampling and Infinite-Dimensional Compressed Sensing. *To appear in Found. Comput. Math.*, 2015. (Cited on pages 3, 74, and 140).

[AH15b] B. Adcock and A. 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. (Cited on pages 3 and 101).

[AHKM15] B. Adcock, A. C. Hansen, G. Kutyniok, and J. Ma. Linear stable sampling rate: Optimality of 2d wavelet reconstructions from fourier measurements. *SIAM Journal on Mathematical Analysis*, 47(2):1196–1233, 2015. (Cited on page 5).

[AHP13] B. Adcock, A. C. Hansen, and C. Poon. Beyond consistent reconstructions: optimality and sharp bounds for generalized sampling, and application to the uniform resampling problem. *SIAM J. Math. Anal.*, 45(5):3114–3131, 2013. (Cited on pages 3, 32, 33, 34, 35, 36, 43, 71, 95, and 100).

[AHP14] 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. (Cited on pages 3, 5, 63, 71, 72, and 140).

[AHPR14] B. Adcock, A. C. Hansen, C. Poon, and B. Roman. Breaking the coherence barrier: A new theory for compressed sensing. *arXiv:1302.0561*, 2014. (Cited on pages 3, 74, and 140).

[AHRT14] B. Adcock, A. C. Hansen, B. Roman, and G. Teschke. Chapter Four - Generalized Sampling: Stable Reconstructions, Inverse Problems and Compressed Sensing over the Continuum. volume 182 of *Advances in Imaging and Electron Physics*, pages 187–279. Elsevier, 2014. (Cited on pages 3 and 140).

[AHS14] B. Adcock, A. C. Hansen, and A. Shadrin. A stability barrier for reconstructions from Fourier samples. *SIAM J. Numer. Anal.*, 52(1):125–139, 2014. (Cited on pages 3, 5, 95, 101, and 140).

[AKC86] C. B. Ahn, J. H. Kim, and Z. H. Cho. High-speed spiral-scan echo planar NMR imaging-I. *IEEE Trans. Med. Imaging*, 5(1):2–7, 1986. (Cited on page 1).

[Ald02] A. Aldroubi. Non-uniform weighted average sampling and reconstruction in shift-invariant and wavelet spaces. *Appl. Comput. Harmon. Anal.*, 13:151–161, 2002. (Cited on page 3).

[ARAK09] E. Acosta-Reyes, A. Aldroubi, and I. Krishtal. On stability of sampling-reconstruction models. *Adv. Comput. Math.*, 31(1-3):5–34, 2009. (Cited on page 127).

[AS74] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions*. Dover, 1974. (Cited on page 111).

[Bai10] B. Bailey. Sampling and recovery of multidimensional bandlimited functions via frames. *J. Math. Anal. Appl.*, 367(2):374–388, 2010. (Cited on pages 106 and 127).

[BC05] B. Borden and M. Cheney. Synthetic-aperture imaging from high-Doppler-resolution measurements. *Inverse Problems*, 21(1):1–11, 2005. (Cited on page 1).

[BCS00] N. Bleistein, J.K. Cohen, and J.W.J. Stockwell. *Mathematics of Multidimensional Seismic Imaging, Migration, and Inversion*. Interdisciplinary Applied Mathematics. Springer New York, 2000. (Cited on page 1).

[BD10] A. Böttcher and P. Dörfler. Weighted Markov-type inequalities, norms of Volterra operators, and zeros of Bessel functions. *Math. Nachr.*, 283(1):40–57, 2010. (Cited on page 100).

[Ben92] J. J. Benedetto. Irregular sampling and frames. In Charles K. Chui, editor, *Wavelets: A Tutorial in Theory and Applications*, pages 445–507. Academic Press Professional, Inc., San Diego, CA, USA, 1992. (Cited on pages 1 and 8).

[Ben93] J. J. Benedetto. Frame decompositions, sampling, and uncertainty principle inequalities. In J. J. Benedetto and M. Frazier, editors, *Wavelets: Mathematics and Applications*, pages 247–304. Boca Raton, FL: CRC, 1993. (Cited on page 8).

[Beu66] A. Beurling. Local harmonic analysis with some applications to differential operators. In *Some Recent Advances in the Basic Sciences, Vol. 1 (Proc. Annual Sci. Conf., Belfer Grad. School Sci., Yeshiva Univ., New York, 1962–1964)*, pages 109–125. Belfer Graduate School of Science, Yeshiva Univ., New York, 1966. (Cited on pages 4, 9, 11, 14, and 15).

[Beu89] A. 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. (Cited on page 9).

[BF01] J.J. Benedetto and P.J.S.G. Ferreira. *Modern Sampling Theory: Mathematics and Applications*. Applied and Numerical Harmonic Analysis. Birkhäuser Boston, 2001. (Cited on page 1).

[BG05] R. F. Bass and K. Gröchenig. Random sampling of multivariate trigonometric polynomials. *SIAM J. Math. Anal.*, 36(3):773–795 (electronic), 2004/05. (Cited on page 9).

[BG13] P. Berger and K. Gröchenig. Sampling and reconstruction in different subspaces by using oblique projections. *arXiv:1312.1717*, 2013. (Cited on page 35).

[BH89] P. L. Butzer and G. Hinsen. Reconstruction of bounded signals from pseudo-periodic, irregularly spaced samples. *Signal Process.*, 17(1):1–17, 1989. (Cited on page 10).

[BW00] J. J. Benedetto and H. C. Wu. Non-uniform sampling and spiral MRI reconstruction. *Proc. SPIE*, 4119:130–141, 2000. (Cited on pages 1, 8, 9, 12, 13, 14, 15, 25, 47, 51, and 113).

[BW07] A. Böttcher and H. Widom. From Toeplitz eigenvalues through Green's kernels to higher-order Wirtinger–Sobolev inequalities. *Oper. Theory Adv. Appl.*, 171:73–87, 2007. (Cited on page 123).

[CCL12] J. Cahill, P. G. Casazza, and S. Li. Non-orthogonal fusion frames and the sparsity of fusion frame operators. *J. Fourier Anal. Appl.*, 18(2):287–308, 2012. (Cited on page 128).

[CDF92] A. Cohen, I. Daubechies, and J. C. Feauveau. Biorthogonal bases of compactly supported wavelets. *Comm. Pure Appl. Math.*, 45(5):485–560, 1992. (Cited on page 57).

[CDV93] A. Cohen, I. Daubechies, and P. Vial. Wavelets on the Interval and Fast Wavelet Transforms. *Appl. Comput. Harmon. Anal.*, 1(1):54 – 81, 1993. (Cited on pages 56, 57, and 65).

[CGH$^+$96] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth. On the Lambert W function. *Adv. Comput. Math.*, 5:329–359, 1996. (Cited on page 110).

[Chr01] O. Christensen. Frames, Riesz bases, and discrete Gabor/wavelet expansions. *Bull. Amer. Math. Soc*, 38(3):273–291, 2001. (Cited on pages 8, 9, 13, 15, and 42).

[Chr03] O. Christensen. *An Introduction to Frames and Riesz Bases*. Birkhauser, 2003. (Cited on page 13).

[CK04] P. G. Casazza and G. Kutyniok. Frames of subspaces. In *Wavelets, Frames, and Operator Theory*, volume 345 of *Contemp. Math.*, pages 87–113, Providence, RI, 2004. Amer. Math. Soc. (Cited on pages 106 and 128).

[CKL08] P. G. Casazza, G. Kutyniok, and S. Li. Fusion frames and distributed processing. *Appl. Comput. Harmon. Anal.*, 25:114–132, 2008. (Cited on pages 106 and 128).

[CZ99] C. K. Chui and L. Zhong. Polynomial interpolation and Marcinkiewicz–Zygmund inequalities on the unit circle. *J. Math. Anal. Appl.*, 233(1):387–405, 1999. (Cited on page 140).

[Dau92] I. Daubechies. *Ten Lectures on Wavelets*. Cbms-Nsf Regional Conference Series in Applied Mathematics. Society for Industrial and Applied Mathematics, 1992. (Cited on page 78).

[DHC+10] B. M. A. Delattre, R. M. Heidemann, L. A. Crowe, J.-P. Vallée, and J.-N. Hyacinthe. Spiral demystified. *Magn. Reson. Imaging*, 28(6):862–881, 2010. (Cited on pages 1, 2, 10, and 47).

[DS52] R. J. Duffin and A. C. Schaeffer. A class of nonharmonic Fourier series. *Trans. Amer. Math. Soc.*, 72:341–366, 1952. (Cited on pages 9 and 15).

[Eld04] Y. C. Eldar. Sampling without input constraints: Consistent reconstruction in arbitrary spaces. In A. I. Zayed and J. J. Benedetto, editors, *Sampling, Wavelets and Tomography*, pages 33–60. Boston, MA: Birkhäuser, 2004. (Cited on page 3).

[EO00] Y. C. Eldar and A. V. Oppenheim. Filterbank reconstruction of bandlimited signals from nonuniform and generalized samples. *IEEE Trans. Signal Process.*, 48(10):2864–2875, 2000. (Cited on pages 6 and 10).

[Eps08] C. L. Epstein. *Introduction to the mathematics of medical imaging*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, second edition, 2008. (Cited on pages 1 and 2).

[EW05] Y. C. Eldar and T. Werther. General framework for consistent sampling in Hilbert spaces. *Int. J. Wavelets Multiresolut. Inf. Process.*, 3(3):347, 2005. (Cited on page 3).

[Far94] A. Faridani. A generalized sampling theorem for locally compact abelian groups. *Math. Comp.*, 63(207):307–327, 1994. (Cited on page 10).

[FG94] H. G. Feichtinger and K. Gröchenig. Theory and practice of irregular sampling. In J. J. Benedetto and M. Frazier, editors, *Wavelets: Mathematics and Applications*, pages 305–363. Boca Raton, FL: CRC, 1994. (Cited on pages 1, 8, and 88).

[FG05] A. Feuer and G. C. Goodwin. Reconstruction of multidimensional bandlimited signals from nonuniform and generalized samples. *IEEE Trans. Signal Process.*, 53(11):4273–4282, 2005. (Cited on page 10).

[FGS95] H. G. Feichtinger, K. Gröchenig, and T. Strohmer. Efficient numerical methods in nonuniform sampling theory. *Numer. Math.*, 69:423–440, 1995. (Cited on pages 1, 7, 38, and 88).

[Fog56] L. J. Fogel. A note on the sampling theorem. *IRE Trans. Inform. Theory*, IT-1:47–48, 1956. (Cited on page 10).

[FR13] S. Foucart and H. Rauhut. *A mathematical introduction to compressive sensing*. Applied and Numerical Harmonic Analysis. Birkhäuser/Springer, New York, 2013. (Cited on page 115).

[FS03] J. A. Fessler and B. P. Sutton. Nonuniform fast Fourier transforms using min-max interpolation. *IEEE Trans. Signal Process.*, 51(2):560–574, 2003. (Cited on pages 6, 38, and 78).

[Gab93] J.-P. Gabardo. Weighted tight frames of exponentials on a finite interval. *Monatsh. Math.*, 116(3-4):197–229, 1993. (Cited on page 9).

[GH11] A. Gelb and T. Hines. Detection of edges from nonuniform Fourier data. *J. Fourier Anal. Appl.*, 17(6):1152–1179, 2011. (Cited on page 95).

[GH12] A. Gelb and T. Hines. Recovering exponential accuracy from non-harmonic Fourier data through spectral reprojection. *J. Sci. Comput.*, 51(158–182), 2012. (Cited on page 95).

[GKHPU11] M. Guerquin-Kern, M. Häberlin, K. P. Pruessmann, and M. Unser. A fast wavelet-based reconstruction method for Magnetic Resonance Imaging. *IEEE Trans. Med. Imaging*, 30(9):1649–1660, 2011. (Cited on page 1).

[GP15] M. Gataric and C. Poon. A practical guide to the recovery of wavelet coefficients from Fourier measurements. *Recommended for publication after minor revision in SIAM Journal on Scientific Computing*, 2015. (Cited on pages 7, 53, and 78).

[GR96] 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. (Cited on page 10).

[Grö92] K. Gröchenig. Reconstruction algorithms in irregular sampling. *Math. Comp.*, 59:181–194, 1992. (Cited on pages 4, 8, 9, 10, 11, 16, 17, 40, 41, 42, 106, 113, 118, 121, 125, and 130).

[Grö93] K. Gröchenig. A discrete theory of irregular sampling. *Linear Algebra Appl.*, 193:129–150, 1993. (Cited on page 7).

[Grö99] K. Gröchenig. Irregular sampling, Toeplitz matrices, and the approximation of entire functions of exponential type. *Math. Comp.*, 68(226):749–765, 1999. (Cited on pages 3, 7, 38, 42, 66, 113, and 120).

[Grö01] K. Gröchenig. Non-uniform sampling in higher dimensions: from trigonometric polynomials to bandlimited functions. In J. J. Benedetto, editor, *Modern Sampling Theory*, chapter 7, pages 155–171. Birkhöuser Boston, 2001. (Cited on pages 3, 7, 8, 17, 38, 106, 113, and 124).

[GRS10] K. Gröchenig, Z. Rzeszotnik, and T. Strohmer. Convergence analysis of the finite section method and Banach algebras of matrices. *Integral Equations Operator Theory*, 67(2):183–202, 2010. (Cited on page 3).

[GS01] K. Gröchenig and T. Strohmer. Numerical and theoretical aspects of non-uniform sampling of band-limited images. In F. Marvasti, editor, *Nonuniform Sampling: Theory and Applications*, chapter 6, pages 283–324. Kluwer Academic, Dordrecht, The Netherlands, 2001. (Cited on pages 1, 7, and 88).

[GS14] A. Gelb and G. Song. A frame theoretic approach to the nonuniform fast Fourier transform. *SIAM J. Numer. Anal.*, 52(3):1222–1242, 2014. (Cited on pages 7, 35, 83, and 95).

[HG10] T. Hrycak and K. Gröchenig. Pseudospectral Fourier reconstruction with the modified inverse polynomial reconstruction method. *J. Comput. Phys.*, 229(3):933–946, 2010. (Cited on pages 3, 5, and 101).

[Jaf91] S. Jaffard. A density criterion for frames of complex exponentials. *Mich. Math. J.*, 38(3):339–348, 1991. (Cited on pages 9, 14, 15, and 21).

[JCAHT15] A. Jones, I. Calvo-Almazán, A. C. Hansen, and A. Tamtögl. Compressed sensing of inelastic and quasielastic helium atom scattering spectra. *Preprint*, 2015. (Cited on page 1).

[JF56] D. L. Jagerman and L. J. Fogel. Some general aspects of the sampling theorem. *IRE Trans. Inform. Theory*, IT-2:139–146, 1956. (Cited on page 10).

[JHA+09] A. P. Jardine, H. Hedgeland, G. Alexandrowicz, W. Allison, and J. Ellis. Helium-3 spin-echo: Principles and application to dynamics at surfaces. *Progress in Surface Science*, 84(11–12):323–379, 2009. (Cited on page 1).

[JMNM91] J. I. Jackson, C. H. Meyer, D. G. Nishimura, and A. Macovski. Selection of a convolution function for Fourier inversion using gridding. *IEEE Trans. Med. Imaging*, 10:473–478, 1991. (Cited on pages 7, 83, 85, and 88).

[JS04] J.-H. Jung and B. D. Shizgal. Generalization of the inverse polynomial reconstruction method in the resolution of the Gibbs phenomenon. *J. Comput. Appl. Math.*, 172(1):131–151, 2004. (Cited on page 3).

[KKP07] T. Knopp, S. Kunis, and D. Potts. A note on the iterative MRI reconstruction from nonuniform k-space data. *Int. J. Bio. Imag.*, page 24727, 2007. (Cited on page 38).

[Kle89] R. Klein. *Concrete and abstract Voronoĭ diagrams*, volume 400 of *Lecture Notes in Computer Science*. Springer-Verlag, Berlin, 1989. (Cited on pages 38 and 49).

[Koh53] A. Kohlenberg. Exact interpolation of band-limited functions. *J. Appl. Phys.*, 24:1432–1436, 1953. (Cited on page 10).

[KPH+97] A. B. Kerr, J. M. Pauly, B. S. Hu, K. C. Li, C. J. Hardy, C. H. Meyer, A. Macovski, and D. G. Nishimura. Real-time interactive MRI on a conventional scanner. *Magn. Reson. Med.*, 38(3):355–367, 1997. (Cited on page 1).

[LA60] D. A. Linden and N. M. Abramson. A generalization of the sampling theorem. *Inform. Contr.*, 3:26–31, 1960. (Cited on page 10).

[Lai00] A. F. Laine. Wavelets in temporal and spatial processing of biomedical images. *Annu. Rev. Biomed. Eng.*, 02:511–550, 2000. (Cited on page 3).

[Lan67] H. J. Landau. Necessary density conditions for sampling and interpolation of certain entire functions. *Acta Math.*, 117:37–52, 1967. (Cited on pages 9 and 15).

[LDSP08] M. Lustig, D. L. Donoho, J. M. Santos, and J. M. Pauly. Compressed sensing MRI. *Signal Proc. Mag.*, 25(2):72–82, 2008. (Cited on page 1).

[Lev40] N. Levinson. *Gap and Density Theorems*. American Mathematical Society Colloquium Publications, v. 26. American Mathematical Society, New York, 1940. (Cited on page 9).

[LMR97] A. K. Louis, D. Maass, and A. Rieder. *Wavelets: Theory and Applications*. Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts. Wiley, 1997. (Cited on page 78).

[LP62] H. J. Landau and H. O. Pollak. Prolate spheroidal wave functions, Fourier analysis and uncertainty—III: The dimension of the space of essentially timeand bandlimited signals. *Bell System Tech J.*, 41(4):1295–1336, 1962. (Cited on page 42).

[LPE12] A. F. Lawrence, S. Phan, and M. Ellisman. Electron tomography and multiscale biology. In Manindra Agrawal, S.Barry Cooper, and Angsheng Li, editors, *Theory and Applications of Models of Computation*, volume 7287 of *Lecture Notes in Computer Science*, pages 109–130. Springer Berlin Heidelberg, 2012. (Cited on page 1).

[LSP+03] M. Lázaro, I. Santamaría, C. Pantaleón, J. Ibáñez, and L. Vielva. A regularized technique for the simultaneous reconstruction of a function and its derivatives with application to nonlinear transistor modeling. *Signal Processing*, 83(9):1859–1870, 2003. (Cited on page 6).

[Mal09] S. G. Mallat. *A Wavelet Tour of Signal Processing: The Sparse Way*. Academic Press, 3 edition, 2009. (Cited on pages 53 and 57).

[Mar01] F. Marvasti. *Nonuniform Sampling: Theory and Practice*. Number v. 1 in Information Technology Series. Springer US, 2001. (Cited on pages 1 and 2).

[Mar07] J. Marzo. Marcinkiewicz–Zygmund inequalities and interpolation by spherical harmonics. *J. Funct. Anal.*, 250(2):559–587, 2007. (Cited on page 140).

[MFK04] S. Matej, J. A. Fessler, and I. G. Kazantsev. Iterative tomographic image reconstruction using Fourier-based forward and back-projectors. *IEEE Trans. Med. Imaging*, 23:401–412, 2004. (Cited on page 7).

[MGG14] A. Martinez, A. Gelb, and A. Gutierrez. Edge detection from non-uniform Fourier data using the convolutional gridding algorithm. *J. Sci. Comput.*, 61(3):490–512, 2014. (Cited on page 95).

[MHNM92] C. H. Meyer, B. S. Hu, D. G. Nishimura, and A. Macovski. Fast spiral coronary artery imaging. *Magn. Reson. Med.*, 28:202–213, 1992. (Cited on page 1).

[MM10] B. Matei and Y. Meyer. Simple quasicrystals are sets of stable sampling. *Complex Var. Elliptic Equ.*, 55(8-10):947–964, 2010. (Cited on page 9).

[Nai68] M. A. Naimark. *Linear Differential Operators*. Harrap, 1968. (Cited on page 121).

[Now98] R. Nowak. Wavelet-based Rician noise removal for Magnetic Resonance Imaging. *IEEE Trans. Image Proc.*, 8:1408–19, 1998. (Cited on page 3).

[OCS07] J. Ortega-Cerda and J. Saludes. Marcinkiewicz–Zygmund inequalities. *J. Approx. Theory*, 145:237–252, 2007. (Cited on page 140).

[OU12] A. Olevskii and A. Ulanovskii. On multi-dimensional sampling and interpolation. *Anal. Math. Phys.*, 2(2):149–170, 2012. (Cited on pages 9, 14, 15, 22, and 113).

[Pap77a] A. Papoulis. Generalized sampling expansion. *IEEE Trans. Circuits Syst.*, 24:652–654, 1977. (Cited on pages 10 and 141).

[Pap77b] A. Papoulis. *Signal Analysis*. McGraw–Hill, 1977. (Cited on page 10).

[PGG12] R. Platte, A. J. Gutierrez, and A. Gelb. Edge informed Fourier reconstruction from non-uniform spectral data with exponential convergence rates. *Preprint*, 2012. (Cited on page 95).

[Poo14] C. Poon. A consistent and stable approach to generalized sampling. *J. Fourier Anal. and Appl.*, 20(5):985–1019, 2014. (Cited on page 74).

[PST01] D. Potts, G. Steidl, and M. Tasche. Fast Fourier Transforms for nonequispaced data: a tutorial. In J. J. Benedetto and P. Ferreira, editors, *Modern Sampling Theory: Mathematics and Applications*, chapter 12, pages 249–274. Birkhäuser, 2001. (Cited on page 6).

[PW87] R. E. A. C. Paley and N. Wiener. *Fourier transforms in the complex domain*, volume 19 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 1987. Reprint of the 1934 original. (Cited on page 9).

[PWBB01] K. P. Pruessmann, M. Weiger, P. Börnert, and P. Boesiger. Advances in sensitivity encoding with arbitrary k-space trajectories. *Magnetic Resonance in Medicine*, 46(4):638–651, 2001. (Cited on page 7).

[PWS+99] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, P. Boesiger, et al. SENSE: sensitivity encoding for fast MRI. *Magnet. Reson. Med.*, 42(5):952–962, 1999. (Cited on pages 1 and 7).

[Raw89] M. D. Rawn. A stable nonuniform sampling expansion involving derivatives. *IEEE Trans. Inform. Theory*, 35(6):1223–1227, 1989. (Cited on pages 10 and 127).

[Raz95] H. Razafinjatovo. Iterative reconstructions in irregular sampling with derivatives. *J. Fourier Anal. Appl.*, 1(3):281–295, 1995. (Cited on pages 10, 106, 107, 113, 118, 120, 122, and 123).

[RPS+99] V. Rasche, R. Proksa, R. Sinkus, P. Bornert, and H. Eggers. Resampling of data between arbitrary grids using convolution interpolation. *IEEE Trans. Med. Imaging*, 18(5):385–392, 1999. (Cited on page 16).

[Sei95a] K. Seip. On the connection between exponential bases and certain related sequences in $L^2(-\pi,\pi)$. *J. Funct. Anal.*, 130(1):131–160, 1995. (Cited on pages 9 and 15).

[Sei95b] K. Seip. On the connection between exponential bases and certain related sequences in $L^2(-\pi,\pi)$. *J. Funct. Anal.*, 130:131–160, 1995. (Cited on page 83).

[Sei04] K. Seip. *Interpolation and sampling in spaces of analytic functions*, volume 33 of *University Lecture Series*. American Mathematical Society, Providence, RI, 2004. (Cited on page 1).

[SJ08] P. Sommen and K. Janse. On the relationship between uniform and recurrent nonuniform discrete-time sampling schemes. *IEEE Trans. Signal Process.*, 56(10, part 2):5147–5156, 2008. (Cited on page 10).

[SN96] G. Strang and T. Nguyen. *Wavelets and Filter Banks*. Wellesley-Cambridge Press, Wellesley, M. A., 1996. (Cited on page 86).

[SN00] H. Sedarat and D. G. Nishimura. On the optimality of the gridding reconstruction algorithm. *IEEE Trans. Med. Imaging*, 19(4):306–317, 2000. (Cited on pages 7, 83, and 85).

[SNF03] B. P. Sutton, D. C. Noll, and J. A. Fessler. Fast, iterative image reconstruction for MRI in the presence of field inhomogeneities. *IEEE Trans. Med. Imaging*, 22(2):178–188, 2003. (Cited on pages 1, 7, 10, 60, 86, and 88).

[ST06] T. Strohmer and J. Tanner. Fast reconstruction methods for bandlimited functions from periodic nonuniform sampling. *SIAM J. Numer. Anal.*, 44(3):1073–1094 (electronic), 2006. (Cited on page 10).

[Str00a] T. Strohmer. A Levinson–Galerkin algorithm for regularized trigonometric approximation. *SIAM J. Sci. Comput.*, 22(4):1160–1183, 2000. (Cited on page 124).

[Str00b] T. Strohmer. Numerical analysis of the non-uniform sampling problem. *J. Comput. Appl. Math.*, 122(1–2):297–316, 2000. (Cited on page 1).

[SZ99] W. Sun and X. Zhou. On the stability of multivariate trigonometric systems. *J. Math. Anal. Appl.*, 235:159–167, 1999. (Cited on pages 106 and 127).

[Szy06] D. Szyld. The many proofs of an identity on the norm of oblique projections. *Numer. Algorithms*, 42:309–323, 2006. (Cited on page 96).

[Tem75] N. M. Temme. Uniform asymptotic expansions of the incomplete gamma functions and the incomplete beta function. *Math. Comp.*, 29(132):1109–1114, 1975. (Cited on page 111).

[Tem79] N. M. Temme. The asymptotic expansion of the incomplete gamma functions. *SIAM J. Math. Anal.*, 10(4):757–766, 1979. (Cited on page 111).

[Tra64] J. F. Traub. On Lagrange-Hermite interpolation. *J. Soc. Indust. Appl. Math.*, 12:886–891, 1964. (Cited on page 132).

[UA94] M. Unser and A. Aldroubi. A general sampling theory for nonideal acquisition devices. *IEEE Trans. Signal Process.*, 42(11):2915–2925, 1994. (Cited on page 3).

[VGCR10] A. Viswanathan, A. Gelb, D. Cochran, and R. Renaut. On reconstructions from non-uniform spectral data. *J. Sci. Comput.*, 45(1–3):487–513, 2010. (Cited on pages 7, 83, 85, and 95).

[WXHD91] J. B. Weaver, Y. Xu, D. M. Healy, and J. R. Driscoll. Filtering MR images in the wavelet transform domain. *Magn. Reson. Med.*, 21:288–295, 1991. (Cited on page 3).

[Yen56] J. L. Yen. On Nonuniform Sampling of Bandwidth-Limited Signals. *Trans. Circuits Syst. I, Reg. Papers1*, 3:251–257, 1956. (Cited on page 10).

[You01] R. M. Young. *An Introduction to Nonharmonic Fourier Series.* Academic Press Inc., San Diego, CA, first edition, 2001. (Cited on pages 1, 8, 21, 42, and 107).

[ZSCZ96] M. Zibulski, V. A. Segalescu, N. Cohen, and Y. Y. Zeevi. Frame analysis of irregular periodic sampling of signals and their derivatives. *J. Fourier Anal. Appl.*, 2(5):453–471, 1996. (Cited on page 10).