NON-LINEAR OPTIMAL SIGNAL MODELS AND STABILITY OF SAMPLING-RECONSTRUCTION

By

Ernesto Acosta Reyes

Dissertation
Submitted to the Faculty of the
Graduate School of Vanderbilt University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY
in
Mathematics

May, 2009
Nashville, Tennessee

Approved:
Professor Akram Aldroubi
Professor Doug Hardin
Professor Larry Schumaker
Professor Gieri Simonett
Professor Nilanjan Sarkar

To my parents

ACKNOWLEDGEMENTS

I am very thankful to my parents, whose love and encouragements are always with me in whatever I pursue. They have given me a lot and I can only return a little by dedicating my work to them.

I would like to express my sincere gratitude to my advisor, Professor Akram Aldroubi, for his invaluable guidance throughout my graduate studies, advice, and support. In addition to having been a fantastic mentor during the last three and a half years, he is a great person and a great mathematician to do research with. My thanks also go to Professors L. Schumaker, D. Hardin, G. Simonett, and N. Sarkar for their participation in my committee. I would also like to thank Professors I. Krishtal and A. Powell.

I am also thankful to my excellent friends in Havana and Miami: Gilberto Hilton, Alfredo Toral, Jose L. Aleman, Marta Barreto, Yara Marrase, Luis López, Roberto Carmona, Ignacio Pérez, Idalmis Granja, Gustavo Pérez, and Mr. Blondy.

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

Modern digital data processing of functions uses a discrete version of the original function f (it is also called signal) that is obtained by sampling f on a discrete set. The question then arises whether and how f can be recovered from its samples. Therefore, the objectives of research on the sampling problem are two fold. *First*: Given a class of functions V on $\mathbb{R}^d$, find conditions on a sampling set $X = \{x_j \in \mathbb{R}^d : j \in J\}$, where J is a countable index set, under which the function $f \in V$ can be reconstructed uniquely and stably from its samples $\{f(x_j) : j \in J\}$. *Second*: Find efficient and fast numerical algorithms that recover any function from its samples on X. In some applications, it is justified to assume that the sampling set X is uniform, i.e., that X forms a regular n-dimensional cartesian grid. For example, a digital image is often acquired by sampling light intensities on a uniform grid. Data acquisition requirements and the ability to process and reconstruct the data simply and efficiently often justify this type of uniform data collection. However, in many realistic situations the data are known only on a non-uniformly spaced sampling set. The following examples indicate that nonuniform sampling problems indeed are applied in science and engineering. *Communication theory*: When data from uniformly sampled signal (function) are lost, the result is generally a sequence of nonuniform samples. This scenario is usually referred as the missing data problem. Often, missing samples are due to the partial destruction of storage devices, e.g., scratches on a CD. *Astronomical measurements*: The measurement of star luminosity gives rise to extremely non-uniformly sampled series. Daylight periods and adverse nighttime weather conditions prevent regular data collection (see, e.g., [114] and the references therein). Other examples where nonuniform sampling sets are the following: geophysics [91], spectroscopy [98], general signal/image processing [21, 28, 104, 109], and biomedical imaging [27, 90, 98]. More information about modern techniques for nonuniform sampling and applications can be found in [24].

I.1 Sampling in Paley-Wiener spaces: Bandlimited Functions.

Infinitely many functions can have the same sampled values on $X = \{x_j\}_{j\in J} \subset \mathbb{R}^d$. For this reason some a priori conditions on f must be imposed. The standard assumption is that the function f on $\mathbb{R}^d$ belongs to the space of bandlimited functions B_Ω; i.e., the Fourier transform $\hat{f}(\xi) = \int_{\mathbb{R}^d} f(x)e^{-2\pi i\langle \xi,x\rangle}dx$ of f is such that $\hat{f}(\xi) = 0$ for all $\xi \in \mathbb{R}^d - \Omega$, where $\Omega = [-\omega, \omega]^d$ for some $0 < \omega < \infty$ (see, e.g., [23, 51, 58, 62, 72] and the review papers [33, 61, 66]). The reason for this assumption is a classical result of Cauchy rediscovered by Whittaker [116] in complex analysis which states that, for dimension $d = 1$, a function $f \in L^2(\mathbb{R}^d) \bigcap B_{[-1/2,1/2]}$ can be recovered exactly from its samples $\{f(k) : k \in \mathbb{Z}\}$ by the interpolation formula

$$f(x) = \sum_{k\in\mathbb{Z}} f(k)\,\mathrm{sinc}(x-k), \tag{1}$$

where $\mathrm{sinc}(x) = \frac{\sin(\pi x)}{\pi x}$. This series gave rise to the uniform sampling theory of Shannon [95], which is fundamental in engineering and digital signal processing because it allows us to convert analog signals into sequences of numbers which can be processed digitally and converted back to analog signals via (1).

If we take the Fourier transform of (1) and the fact that the Fourier transform of the sinc function is the characteristic function $\chi_{[-1/2,1/2]}$ it can be shown that for any $\xi \in [-1/2, 1/2]$

$$\hat{f}(\xi) = \sum_{k\in\mathbb{Z}} f(k) e^{2\pi k\xi} = \sum_{k\in\mathbb{Z}} \langle \hat{f}, e^{i2\pi\cdot} \rangle_{L^2(-1/2,1/2)} e^{i2\pi k\xi}.$$

Consequently, the reconstruction by means of the formula (1) is equivalent to the fact that the set $\{e^{i2\pi k\xi}, k \in \mathbb{Z}\}$ forms an orthonormal basis of $L^2(-1/2, 1/2)$. This equivalence between the above orthonormal basis and the reconstruction of a uniformly sampled bandlimited function has been extended to treat some special cases of nonuniformly sampled data. In particular, the results by Paley and Wiener [87], Kadec [71], and others on nonharmonic Fourier bases $\{e^{i2\pi x_k\xi}, k \in \mathbb{Z}\}$ can be translated into results about nonuniform sampling and reconstruction of bandlimited functions [23, 62, 89, 94]. Kadec's theorem [71] states that if $X = \{x_k \in \mathbb{R} : |x_k - k| \le L < 1/4\}_{k\in\mathbb{Z}}$, then $\{e^{i2\pi x_k\xi}, k \in \mathbb{Z}\}$ is a Riesz basis of $L^2(-1/2, 1/2)$; i.e., $\{e^{i2\pi x_k\xi}, k \in \mathbb{Z}\}$ is the image of an orthonormal basis of $L^2(-1/2, 1/2)$ under a bounded and invertible operator from $L^2(-1/2, 1/2)$ onto $L^2(-1/2, 1/2)$. Using the Fourier transform methods, this result implies that any bandlimited function $f \in L^2(\mathbb{R}) \bigcap B_{[-1/2,1/2]}$ can be completely recovered from its samples $f(x_k)$, $k \in \mathbb{Z}$, as long as the sampling set is of the form $X = \{x_k \in \mathbb{R} : |x_k - k| \le L < 1/4\}_{k\in\mathbb{Z}}$. The sampling set $X = \{x_k \in \mathbb{R} : |x_k - k| \le L < 1/4\}_{k\in\mathbb{Z}}$ in Kadec's theorem is just a perturbation of $\mathbb{Z}$. For more general sampling sets, the work of Beurling [29, 30], Landau [74], and others [25] provides a deep understanding of the one-dimensional theory of nonuniform sampling of bandlimited functions. Specifically, for the exact and stable reconstruction of a bandlimited function f from its samples $\{f(x_j) : x_j \in X\}$, it is sufficient that the *Beurling lower density*

$$D^-(X) = \lim_{r\to\infty} \inf_{y\in\mathbb{R}} \frac{Card\,(X \bigcap (y + [0, r]))}{r} \tag{2}$$

satisfies $D^-(X) > 1$, where $Card(A)$ denotes the cardinality of the set A. On the other hand, if f is uniquely and stably determined by its samples on $X \subset \mathbb{R}$, then $D^-(X) \ge 1$ [74]. We should emphasize that these results deal with stable reconstructions. This means that an inequality of the form

$$\|f\|_p \leq C \left(\sum_{j \in J} |f(x_j)|^p \right)^{1/p}$$

holds for all bandlimited functions $f \in L^p \bigcap B_\Omega$. A sampling set for which the reconstruction formula is stable in this sense is called a (*stable*) set of sampling. This terminology is used to contrast a set of sampling with the weaker notion of a *set of uniqueness*. X is a set of uniqueness for B_Ω if $f|_X = 0$ implies $f = 0$. Whereas a set of sampling for $B_{[-1/2,1/2]}$ has density $D \geq 1$, there are sets of uniqueness with arbitrarily small density. See [31, 73] for examples and characterizations of sets of uniqueness.

While the theorems of Paley, Wiener and Kadec about Riesz bases consisting of complex exponentials $e^{i2\pi x_k \xi}$ are equivalent to statements about sampling sets that are perturbation of $\mathbb{Z}$, the results about arbitrary sets of sampling are connected to the more general notion of frames introduced by Duffin and Schaeffer [48]. The concept of frames generalizes the notion of orthogonal bases and Riesz bases in Hilbert spaces and of unconditional bases in some Banach spaces [3, 13, 19, 22, 23, 27, 37, 53, 67].

I.2 Sampling in Shift-Invariant Spaces

The series (1) shows that the space of bandlimited functions $B_{[-1/2,1/2]}$ is identical with the space

$$V^2(\mathrm{sinc}) = \left\{ \sum_{k \in \mathbb{Z}} c_k \,\mathrm{sinc}(x - k) : (c_k) \in l^2 \right\}. \tag{3}$$

Since the sinc function has infinite support and slow decay, the space of bandlimited functions is often unsuitable for numerical implementations. For instance, the pointwise evaluation

$$f \longmapsto f(x_0) = \sum_{k \in \mathbb{Z}} c_k \,\mathrm{sinc}(x_0 - k)$$

is a nonlocal operation, because, as a consequence of the long-range behavior of sinc, many coefficients c_k will contribute to the value $f(x_0)$. In fact, all bandlimited functions have infinite support since they are analytic. Moreover, functions that are measured in applications tend to have frequency components that decay for higher frequencies, but these functions are not bandlimited in the strict sense. Thus, it has been advantageous to use non-bandlimited models that are more amenable to numerical implementation and more flexible for approximating real data [21, 64, 65, 82, 104, 105]. One such example are *shift-invariant spaces* which form the focus of the next two chapters and are defined in (4) (see page 9 below). Such spaces have been used in finite elements and approximation theory [68, 69, 70] and for the construction of multi-resolution approximations and wavelets [41, 42, 46, 55]. They have been extensively studied in recent years (see, for instance, [13, 68, 69, 70]).

Sampling in shift-invariant spaces that are not bandlimited is a suitable and realistic model for many applications, e.g., for taking into account real acquisition and reconstruction devices, for modeling signals with smoother spectrum than is the case with bandlimited functions, or for numerical implementations [15, 21, 28, 32, 81, 82, 104, 105, 112]. These requirements can often be met by choosing "appropriate" functions. This may mean that these functions have a shape corresponding to a particular "impulse response" of a device, or they are compactly supported, or they have a Fourier transform that decays smoothly to zero at infinity.

I.2.1 Uniform Sampling in Shift-Invariant Spaces

Early results on sampling in shift-invariant spaces concentrated on the problem of uniform sampling [14, 15, 16, 17, 65, 108, 110, 117, 118]. The problem of uniform sampling in shift-invariant spaces shares some similarities with Shannon's sampling theorem in that it requires only the Poisson summation formula and a few facts about Riesz bases [14, 15]. The connection between interpolation in spline spaces, filtering of signals, and Shannon's sampling theory was established in [17, 111]. These results imply that Shannon's sampling theory can be viewed as a limiting case of polynomial spline interpolation when the order of the spline tends to infinity [17, 111]. Moreover, Shannon's sampling theory is a special case for the interpolation in shift-invariant spaces [14, 15, 115, 118].

In applications, signals do not belong to a prescribed shift-invariant space in general. For instance, in engineering when using the bandlimited theory, the function f is often forced to become bandlimited before sampling. The idea of this procedure is to multiply the Fourier transform $\hat{f}$ of f by a characteristic function χ_Ω. The new function f_a with Fourier transform $\hat{f}_a = \hat{f}\chi_\Omega$ is then sampled and stored digitally for later processing or reconstruction. The multiplication by χ_Ω before sampling is called *pre-filtering with an ideal filter* and is used to reduce the errors in reconstructions called *aliasing errors*. It has been shown that the three steps of the traditional uniform sampling procedure, namely pre-filtering, sampling, and post-filtering for reconstruction, are equivalent to find the best L^2-approximation of a function in $L^2 \bigcap B_\Omega$ [15, 108]. This procedure generalizes to sampling in general shift-invariant spaces [14, 15, 16, 81, 109]. Indeed, the reconstruction from the samples of a function should be considered as an approximation in the shift-invariant space generated by the impulse response of the sampling device. This allows a reconstruction that optimally fits the available samples and can be done using fast algorithms [109, 110].

I.2.2 Nonuniform Sampling in Shift-Invariant Spaces

The theory of nonuniform sampling in general shift-invariant spaces is more recent [8, 38, 67, 75, 76, 99]. The earliest results in this subject [39, 76] concentrate on perturbation of regular shift-invariant spaces and are therefore similar in spirit to Kadec's result for bandlimited functions. For the L^2 case in dimension $d = 1$, and under some restrictions on the shift-invariant spaces, several theorems on nonuniform sampling can be found in [75, 99]. Moreover, a lower bound on the maximal distance between two sampling points needed for reconstructing a function from its samples was given for the case of polynomial splines and other special cases of shift-invariant spaces in [75]. For the general multivariate case L^p, the theory was developed in [8, 10], and for the case of polynomial spline shift-invariant spaces, the maximal allowable gap between samples was obtained in [9]. For general shift-invariant spaces, a Beurling density $D \geq 1$ is necessary for stable reconstruction . As in the case of bandlimited functions, the theory of frames is central in nonuniform sampling of shift-invariant spaces, and there is an equivalence between a certain type of frames and the problem of sampling in shift-invariant spaces [9, 67].

I.3 Compressive Sampling

A new theory for signal sampling and reconstruction which has been recently developed by Lu and Do (see [79]), starts from the point of view that signals live in some union of subspaces $\mathcal{M} = \bigcup_{i\in I} C_i$, instead of a single space $\mathcal{M} = C$, for instance, the space of band-limited functions. This new theory is general and extends the classical Shannon sampling theory [95], and sampling signals with finite rate of innovations (see [47, 85]). In the new framework, when there are more than one subspace, the signal model $\mathcal{M} = \bigcup_{i\in I} C_i$ is non-linear and the techniques for reconstructing a signal $f \in \bigcup_{i\in I} C_i$ from its samples are involved and the reconstruction operators are non-linear.

In this new paradigm for signal sampling and reconstruction, the starting point is the knowledge of the signal set $\mathcal{M} = \bigcup_{i\in I} C_i$. Therefore, the first step for implementing the theory is to find the appropriate signal model $\mathcal{M} = \bigcup_{i\in I} C_i$ from a set of observed data $\mathcal{F} = \{f_1, \dots, f_m\}$. This problem has been solved in [6] for the case of finding a shift-invariant space model $\mathcal{M} = C$ from a set of observed data. For the new sampling paradigm, the problem consists in providing the existence and finding subspaces $C_1, \dots, C_l$, of some Hilbert space $\mathcal{H}$ that minimize the expression $e(\mathcal{F}, \{C_1, \dots, C_l\}) = \sum_{j=1}^{m} \min_{1\leq i\leq l} \operatorname{dist}^2(f_j, C_i)$ over all closed subsets C_i, $1 \leq i \leq l$, belonging to a class $\mathcal{C}$ of subspaces of $\mathcal{H}$ (this problem will be referred as Problem 1). Here $\mathcal{F} = \{f_1, \dots, f_m\} \subset \mathcal{H}$ is a set of observed data. This problem has been studied in [7] where a necessary and sufficient condition is given in order to guarantee the existence of a solution to Problem 1, and an algorithm for finding a solution to this problem is also provided.

The problem of sampling signals with finite rate of innovation is closely related to the theory of *compressive sampling*, also known as *compressed sensing*. In compressive sampling it is proposed to find a vector $z \in \mathbb{C}^d$ from the knowledge of its values $\{y_j =< z, \alpha_j >, 1 \leq j \leq p\}$, when a small set of functionals $\{\alpha_j\}_{j=1}^{p}$ ($p \ll d$) is applied to z. Clearly, finding z from the set $\{y_j\}_{j=1}^{p}$ is ill-posed. Nevertheless, it becomes meaningful if z is assumed to be sufficiently sparse, that is, z has at most n non-zero components with respect to a given orthonormal basis of $\mathbb{C}^d$ ($\|z\|_0 \leq n$), where $n \leq 2p \leq d$. Consequently, from the sparsity of z it follows that z belongs to some union of subspaces, each of which is generated from exactly n vectors from the canonical basis of $\mathbb{C}^d$. This problem has the following matrix formulation: find $z \in \mathbb{C}^d$ with $\|z\|_0 \leq n$ from the matrix equation $y = Az$, where A is a $p \times d$ matrix and y is a given vector in $\mathbb{C}^p$. A related problem consists in finding an approximation to the vector y using a sparse vector z, that is, find $\min_{z \in \mathbb{C}^d} \|z\|_0$ subject to $\|Az - y\|_2 \leq \epsilon$, for some given ϵ. These problems, their analysis, extensions and algorithms for finding their solutions can be found, for instance, in [1, 2, 34, 35, 45, 57, 106].

If in the problems described in the previous paragraph the matrix A is also to be found together with the set of unknown vectors $\{z_i\}_{i=1}^m \subset \mathbb{C}^d$, then these problems become the problem of finding a *dictionary* A from the data $\{y_i\}_{i=1}^m \subset \mathbb{C}^p$ obtained by sampling the sparse vectors $\{z_i\}_{i=1}^m \subset \mathbb{C}^d$, see [1, 2, 57], and the references therein. In this context, the columns of A are called *atoms* of A. This problem has a unique solution up to a permutation of the atoms of A under certain assumptions on the dictionary and the data. Finding the solution to this problem is computationally intractable using exhaustive methods, however the K-SVD algorithm described in [1] provides an efficient search algorithm.

The signal modeling problem as described in Problem 1 is closed related to the dictionary design problem for sparse data described in the previous paragraph. To see this relation, let us formulate the dictionary design as follows: assume that we have a class of m signals, where m is a very large number. We want to know whether exists a dictionary such that every signal in the class is a linear combination of at most n atoms in the dictionary. Certainly, the length of the dictionary should be small compared with m in order to make this problem realistic and meaningful. Consequently, if for a given set of data such a dictionary exists, then the data can be partitioned into subsets each of which belongs to a subspace of dimension at most n, that is, the subspace generated by atoms that the signal uses in its representation. Conversely, if our class of signals can be partitioned into l subsets, such that the signals in each subset belong to a subspace of dimension at most n, then choosing a set of generators from each of the subspaces, we can construct a dictionary of length at most nl, with the property that each signal in the class can be represented using at most n atoms in the dictionary.

The subspace segmentation problem for a set of signals in $\mathbb{C}^d$ (see [83, 84]) is also related to Problem 1. This problem takes place in he context of segmentation clustering and classification, and consists in finding whether there exist l subspaces of dimension at most n, such that the signals in the class belong to the union of l subspaces. The subspace segmentation problem has applications in computer vision, image processing, and other areas of engineering. The method for solving it, known as the Generalized Principle Component Analysis, has been extended to deal with moderate noise in the data [113], and the uniqueness problem has been addressed in [84].

The above problems have connections to the geometry of Hilbert and Banach spaces, non-linear approximation, optimization, and functional analysis. In addition, these problems have applications to signal modeling and segmentation (e.g., face recognition, movement tracking, and DNA sequence comparison [47]).

I.4 Organization

This dissertation has been organized as follows: In Chapter II we consider a very general sampling model where the signal is assumed to belong to a finitely generated shift-invariant space and the sampling is performed on an irregular relatively separated set and is averaged by finite complex Borel measures. The main focus on Chapter II is on describing and quantifying admissible perturbations of the sampling model which may result from altering the sampling set (jitter) (see e.g. [12, 18, 54, 56, 80]), or the averaging sampling measures (measuring devices) or the generators of the underlying shift-invariant space (see e.g., [11, 54, 100]). In Chapter III we study the problem of recovering a signal f, which belongs to a shift-invariant space, from its samples values $\{g_{x_j}(f)\}_{j\in J}$ using an iterative algorithm which uses the density properties of the sampling set X, the support size conditions of the collection μ of measures used to average the signal, and the properties of the generator of the shift-invariant space. Here we show that the sequence of functions generated using the algorithm converges to f geometrically fast. In [59], [93] and [120], this method was used for iterative reconstruction of band-limited signals, in [4] and [10] it was used for reconstructing functions belonging to shift-invariant spaces, and in [119] it was used for reconstructing signals belonging to a weighted multiply generated shift-invariant spaces. Furthermore, if the sampling set is assumed to be a separated set, then we show that it is also a set of sampling for μ and the corresponding shift-invariant space, and we give explicit stability bounds. The stability of the sampling-reconstruction model presented in Chapter III is also analyzed when the samples of a signal are perturbed by noise, and we show that the reconstruction error is continuously controlled by the perturbation of the sampled data $\{g_{x_j}(f)\}_{j\in J}$. Finally, in Chapter IV, the main goal is to guarantee the existence of a solution (signal model) to Problem 1 in a more general setting, that is, when we are considering a complete metric space (X, d) instead of a separable Hilbert space $\mathcal{H}$. When X is a separable Hilbert space $\mathcal{H}$, we provide sufficient conditions in terms of the weak operator topology in order to guarantee the existence of a solution to Problem 1. Moreover, we also study the problem when we consider a class $\mathcal{C}$ which is defined in terms of a collection of unitary operators applied to a given convex subset of $\mathcal{H}$, and we obtain a procedure for constructing particular collections of closed subspaces of $\mathcal{H}$ for which we a priory know the existence of a minimizer to Problem 1. As a consequence, we obtain the well-known qualitative version of the Eckart-Young Theorem ([49]).

CHAPTER II

ON STABILITY OF SAMPLING-RECONSTRUCTION MODELS

Images are often modeled as real valued functions on $\mathbb{R}^2$. Typically, a class of images is a subspace of $L^2(\mathbb{R}^2)$ (or $L^p(\mathbb{R}^2)$) with some smoothness property, e.g., a bandlimited or spline subspace of $L^2(\mathbb{R}^d)$. Modern image processing schemes proceed in three steps. The first step consists in transforming the image (a function on $\mathbb{R}^2$) into a sequence (a function on $\mathbb{Z}^2$). This process is called sampling, and it uses measuring devices that act on the image to produce the corresponding sequence, e.g., an analog to digital converter, or an MR imaging device. One of the requirements of this conversion is that it must be invertible or nearly invertible, i.e., the image must be fully characterized by the sequence. The inverse process that transforms the sequence back into the image is called the reconstruction and is often the third step in the processing. Image processing algorithms, such as denoising, compression/decompression, etc., are performed between these two steps using digital devices (e.g., computers). In this paper, we concentrate on the sampling and reconstruction steps which are fundamental in any image processing scheme.

The sampling and reconstruction problem includes devising efficient methods for representing a signal (function) in terms of a discrete (finite or countable) set of its samples (values) and reconstructing the original signal from the samples (see e.g., [4, 9, 24, 26, 86, 107] and the reference therein). In this chapter we consider a very general sampling model where the signal is assumed to belong to a finitely generated shift invariant space and the sampling is performed on an irregular relatively separated set and is averaged by finite Borel measures. The main focus of this paper is on describing and quantifying "admissible" perturbations of the sampling model which may result from altering the sampling set (jitter) (see e.g. [12, 18, 54, 56, 80]), or the averaging sampling measures (measuring devices) or the generators of the underlying shift-invariant space (see e.g., [11, 54, 100]).

As recently became customary in sampling theory (see e.g. [4, 9, 38, 101, 102, 103, 107, 119]), we mesh operator theory techniques and those of shift invariant and Wiener amalgam spaces [52]. The latter provide us with relatively straight-forward proofs while the former allow us to keep in sight our objective. In Section II.2.1 we show that all the properties of our sampling model can be encoded in the sampling operator U. The sampling model admits reconstruction if its sampling operator is bounded both above and below. Our first goal is to show that any and all of the small perturbations mentioned above result in a small perturbation of U in the operator norm. This will prove the stability of sampling in our model with respect to those perturbations. Moreover, in some cases, the corresponding estimates we obtain will quantify this stability. Our second goal is to show how the dual frame method can be used to reconstruct signals in our sampling model. Finally, our last goal is to show that the reconstruction error due to the perturbations we describe is controlled continuously by the perturbation errors.

II.1 Description of the sampling model

This section is primarily devoted to introduction of the sampling model we use in this chapter. We also present most of the necessary notation and cite some of the preliminary results that will be used later.

The signals we are studying in this and next chapter are represented by functions $f \in L^p(\mathbb{R}^d)$, for some $p \in [1, \infty]$ and $d \in \mathbb{N}$. Moreover, we assume that f belongs to a shift invariant space

$$V^p(\Phi) = \left\{ \sum_{k \in \mathbb{Z}^d} C_k^T \Phi_k : C \in (\ell^p(\mathbb{Z}^d))^r \right\}. \tag{4}$$

Here the superscript T denotes the transpose operation, $\Phi = (\phi^1, \ldots, \phi^r)^T$ is a vector of functions, $\Phi_k = \Phi(\cdot - k)$, $C = (c^1, \ldots, c^r)^T \in (\ell^p(\mathbb{Z}^d))^r$ is a vector of sequences in $\ell^p(\mathbb{Z}^d)$, and $C_k^T = (c_k^1, \ldots, c_k^r)$. Among the equivalent norms in $(\ell^p(\mathbb{Z}^d))^r$ we choose

$$\|C\|_{(\ell^p(\mathbb{Z}^d))^r} = \sum_{i=1}^{r} \|c^i\|_{\ell^p(\mathbb{Z}^d)}. \tag{5}$$

In order to avoid convergence issues in (4) we assume that the set $\{\phi^1(\cdot - k), \ldots, \phi^r(\cdot - k); k \in \mathbb{Z}^d\}$ generates an unconditional basis for $V^p(\Phi)$. More specifically, we require that there exist constants $0 < m_p \leq M_p < \infty$, such that

$$m_p \|C\|_{(\ell^p(\mathbb{Z}^d))^r} \leq \left\| \sum_{k \in \mathbb{Z}^d} C_k^T \Phi_k \right\|_{L^p} \leq M_p \|C\|_{(\ell^p(\mathbb{Z}^d))^r}, \ \forall C \in (\ell^p(\mathbb{Z}^d))^r. \tag{6}$$

The basis assumption (6) implies that $\{\phi^1(\cdot - k), \ldots, \phi^r(\cdot - k); k \in \mathbb{Z}^d\}$ generates an unconditional basis for $V^p(\Phi)$ and that the space $V^p(\Phi)$ is a closed subspace of $L^p(\mathbb{R}^d)$(see [10]). For $p = 2$, a basis satisfying condition (6) is called a Riesz basis. A Riesz basic sequence is a Riesz basis for its closed span.

Traditionally, all the functions in $V^p(\Phi)$ are assumed to be continuous. This makes pointwise evaluations and, hence, ideal sampling meaningful. In our case, such an assumption is not always necessary, although, we retain it for technical reasons. More precisely, we add a slightly stronger assumption that all generators Φ belong to a Wiener amalgam space $(W_0^1)^r$ as defined below. For $1 \leq p < \infty$, a measurable function f belongs to W^p if it satisfies

$$\|f\|_{W^p} = \left(\sum_{k \in \mathbb{Z}^d} \operatorname*{ess\,sup}_{x \in [0,1]^d} |f(x+k)|^p \right)^{1/p} < \infty. \tag{7}$$

If $p = \infty$, a measurable function f belongs to W^∞ if it satisfies

$$\|f\|_{W^\infty} = \sup_{k \in \mathbb{Z}^d} \{ \operatorname*{ess\,sup}_{x \in [0,1]^d} |f(x+k)| \} < \infty. \tag{8}$$

Hence, W^∞ coincides with $L^\infty(\mathbb{R}^d)$. It is well known that W^p are Banach spaces [52], and clearly $W^p \subseteq L^p$. By $(W^p)^r$ we denote the space of vectors $\Psi = (\psi^1, \dots, \psi^r)^T$ of W^p-functions with the norm

$$\|\Psi\|_{(W^p)^r} = \sum_{i=1}^{r} \|\psi^i\|_{W^p}.$$

The closed subspace of (vectors of) continuous functions in W^p (respectively, $(W^p)^r$) will be denoted by W_0^p (or $(W_0^p)^r$).

In this chapter we are interested in average sampling performed by a vector of measures. We denote by $\mathcal{M}(\mathbb{R}^d) = \mathcal{M}_0(\mathbb{R}^d)$ the Banach space of finite complex Borel measures on $\mathbb{R}^d$. The norm on $\mathcal{M}(\mathbb{R}^d)$ is given by $\|\mu\| = \int_{\mathbb{R}^d} d|\mu|(y)$, i.e., the total variation of a measure μ. By $\mathcal{M}^t(\mathbb{R}^d)$ we denote the space of vectors $\overrightarrow{\mu} = (\mu^1, \dots, \mu^t)^T$ of measures from $\mathcal{M}(\mathbb{R}^d)$ with the norm $\|\overrightarrow{\mu}\|_{\mathcal{M}^t} = \sum_{j=1}^{t} \|\mu^j\|$. The symbols $\mathcal{M}_s(\mathbb{R}^d)$ $(\mathcal{M}_s^t(\mathbb{R}^d))$, $0 \le s < \infty$, will be used for the subspace of $\mathcal{M}(\mathbb{R}^d)$ $(\mathcal{M}^t(\mathbb{R}^d))$ of all (vectors of) measures $\mu \in \mathcal{M}(\mathbb{R}^d)$ such that $(1+|x|)^s \in L^1(\mathbb{R}^d, d|\mu|)$, i.e., $\int (1 + |x|)^s d|\mu|(x) < \infty$. By $\mathcal{M}_\infty(\mathbb{R}^d)$ $(\mathcal{M}_\infty^t(\mathbb{R}^d))$ we denote the space of all (vectors of) measures with compact support. Clearly $\mathcal{M}_s(\mathbb{R}^d) \subset \mathcal{M}_r(\mathbb{R}^d)$ for $0 \le r \le s \le \infty$.

For $\mu \in \mathcal{M}(\mathbb{R}^d)$ and $\phi \in L^1(\mathbb{R}^d, d\mu)$, the convolution of the function ϕ and the measure μ is defined by

$$(\phi * \mu)(x) = \int_{\mathbb{R}^d} \phi(x-y) d\mu(y), \quad x \in \mathbb{R}^d.$$

When we have a vector of finite complex Borel measures $\overrightarrow{\mu} = (\mu^1, \dots, \mu^t)^T$ and a vector $\Phi = (\phi^1, \dots, \phi^r)^T$ of functions integrable with respect to the corresponding measures, the convolution $\Phi * \overrightarrow{\mu}^T$ is the $r \times t$ matrix given by

$$\Phi * \overrightarrow{\mu}^T = \begin{pmatrix} \phi^1 * \mu^1 \dots \phi^1 * \mu^t \\ \vdots \qquad\quad \vdots \\ \phi^r * \mu^1 \dots \phi^r * \mu^t \end{pmatrix}.$$

Let J be a countable index set and $X = \{x_j : j \in J\}$ be a subset of $\mathbb{R}^d$. The reconstruction problem in our sampling model consists of finding the function $f \in V^p(\Phi)$ from the knowledge of its samples

$$(f * \overrightarrow{\mu})(X) = \{(f * \overrightarrow{\mu})(x_j) = \big((f * \mu^1)(x_j), \dots, (f * \mu^t)(x_j)\big)^T\}_{j \in J}.$$

When $t = 1$ and $\overrightarrow{\mu} = (\delta_0)$, i.e., $\overrightarrow{\mu}$ is the Dirac measure on $\mathbb{R}^d$ concentrated at zero, then $(f * \overrightarrow{\mu})(X) = \{f(x_j)\}_{j \in J}$ and we obtain the classical (ideal) sampling model. When $d\overrightarrow{\mu} = \Psi dx$, where $\Psi \in (L^1(\mathbb{R}^d))^t$ and dx is the Lebesgue measure on $\mathbb{R}^d$, i.e., $\overrightarrow{\mu}$ is absolutely continuous with respect to the Lebesgue measure, then we write $(f * \Psi)(X)$ instead of $(f * \overrightarrow{\mu})(X)$, and our model is reduced to the case analyzed in [11].

Definition 1. *Let $1 \le p \le \infty$ and $X = \{x_j : j \in J\}$ be a countable subset of $\mathbb{R}^d$. We say that X is a* set of sampling *for $V^p(\Phi)$ and $\overrightarrow{\mu}$ (or, simply, a $\overrightarrow{\mu}$-sampling set for*

$V^p(\Phi)$) *if there exist constants* $0 < A_p \le B_p < \infty$ *such that*

$$A_p\|f\|_{L^p} \le \|(f * \overrightarrow{\mu})(X)\|_{(\ell^p(J))^t} \le B_p\|f\|_{L^p}, \text{ for all } f \in V^p(\Phi). \tag{9}$$

If $d\overrightarrow{\mu} = \Psi dx$ then a $\overrightarrow{\mu}$-sampling set X will be called a Ψ-sampling set and, if $t = 1$ and $\mu = \delta_0$, then X will be called an *ideal* sampling set. To ensure that an upper bound B_p in (9) always exists we restrict our attention only to relatively separated sets X. On the other hand, if $p = 2$, then using the Riesz Representation Theorem, we get that the above definition is the definition of frame.

Definition 2. *We say that X is relatively separated if there exists $\mathcal{S} \in \mathbb{N}$ such that there are at most $\mathcal{S}$ sampling points in every d-dimensional cube $[0,1]^d + l$, $l \in \mathbb{Z}^d$.*

Remark 1. *In most of our estimates in the proofs we only use the quantity $\mathcal{N}_p = \mathcal{S}^{1/p}$, $p \in [1,\infty)$. In case $p = \infty$, we have $\mathcal{N}_\infty = 1$ and the assumption of relative separation may be unnecessary.*

Definition 3. *Let $\overrightarrow{\mu} \in \mathcal{M}^t(\mathbb{R}^d)$, $\Phi \in (W_0^1)^r$ satisfy (6), and $X = \{x_j, j \in J\} \subset \mathbb{R}^d$ be a relatively separated set. The* sampling model *is the triple $(X, \Phi, \overrightarrow{\mu})$. The sampling model $(X, \Phi, \overrightarrow{\mu})$ is called* p-stable *if X is a $\overrightarrow{\mu}$-sampling set for $V^p(\Phi)$, $p \in [1,\infty]$.*

Given a p-stable sampling model $(X, \Phi, \overrightarrow{\mu})$ we proceed to define its sampling operator.

Definition 4. *The* sampling operator $U = U_{(X,\Phi,\overrightarrow{\mu})} : (\ell^p(\mathbb{Z}^d))^r \to (\ell^p(J))^t$ *is defined by $UC = (f * \overrightarrow{\mu})(X)$, where $f = \sum\limits_{k\in\mathbb{Z}^d} C_k^T \Phi_k \in V^p(\Phi)$.*

We can think of U as a $t \times r$ block matrix

$$U = \begin{pmatrix} U^{1,1} \dots U^{1,r} \\ \vdots \quad\quad \vdots \\ U^{t,1} \dots U^{t,r} \end{pmatrix},$$

where for each $1 \le i \le r$ and $1 \le l \le t$ the operator $U^{l,i}$ is defined by an infinite matrix with entries $(U^{l,i})_{j,k} = (\phi^i * \mu^l)(x_j - k)$, $j \in J$, $k \in \mathbb{Z}^d$. The operator norm of U is given by $\|U\|_{p,op} = \sup\limits_{1\le i\le r} \sum_{l=1}^t \|U^{l,i}\|_{p,op}$.

The following proposition shows that all the interesting properties of a sampling model $(X, \Phi, \overrightarrow{\mu})$ are, indeed, encoded in the sampling operator U. The proof of this result follows immediately from (6) and (9).

Proposition 1. *The sampling model $(X, \Phi, \overrightarrow{\mu})$ is p-stable if and only if there exist $0 < \eta_p \le \beta_p < \infty$ such that for all $C \in (\ell^p(\mathbb{Z}^d))^r$ the sampling operator U satisfies*

$$\eta_p\|C\|_{(\ell^p(\mathbb{Z}^d))^r} \le \|UC\|_{(\ell^p(J))^t} \le \beta_p\|C\|_{(\ell^p(\mathbb{Z}^d))^r}. \tag{10}$$

Remark 2. *It is not hard to see how the norm bounds η_p and β_p in the above proposition are related to the sampling bounds A_p and B_p. Another important observation* *is that β_p is controlled by the norms of Φ and $\overrightarrow{\mu}$ and the separation constant $\mathcal{N}_p$. As will be apparent from the auxiliary results in Section II.3.1, we have*

$$\|U\|_{p,op} \leq 2^d \mathcal{N}_p \, \|\overrightarrow{\mu}\|_{\mathcal{M}^t} \, \|\Phi\|_{(W^1)^r} \, ,$$

and, hence, we may assume that $\beta_p \leq 2^d \mathcal{N}_p \, \|\overrightarrow{\mu}\|_{\mathcal{M}^t} \, \|\Phi\|_{(W^1)^r}$.

The next lemma is typical for perturbation arguments.

Lemma 1. *Let $(X, \Phi, \overrightarrow{\mu})$ be a p-stable sampling model and U be its sampling operator satisfying (10). Let also $(\widetilde{X}, \Theta, \overrightarrow{\alpha})$ be a sampling model such that its sampling operator U_Δ satisfies $\|U - U_\Delta\| < \eta_p$. Then $(\widetilde{X}, \Theta, \overrightarrow{\alpha})$ is also p-stable.*

Proof. Let $C \in (\ell^p(\mathbb{Z}^d))^r$. Then UC and $U_\Delta C$ are in $(\ell^p(J))^t$ and

$$\begin{aligned} \|U_\Delta C\| &\leq \|(U - U_\Delta)C\| + \|UC\| \\ &\leq \|U - U_\Delta\|\|C\| + \beta_p\|C\|. \end{aligned}$$

Therefore, since $\|U - U_\Delta\| < \eta_p$, then we have

$$\|U_\Delta C\| \leq (\eta_p + \beta_p)\, \|C\|. \tag{11}$$

On the other hand, since

$$\begin{aligned} \eta_p\|C\| \leq \|UC\| &\leq \|(U - U_\Delta)C\| + \|U_\Delta C\| \\ &\leq \|U - U_\Delta\|\|C\| + \|U_\Delta C\|. \end{aligned}$$

Hence,

$$(\eta_p - \|U - U_\Delta\|)\, \|C\| \leq \|U_\Delta C\|. \tag{12}$$

Since $\|U - U_\Delta\| < \eta_p$, the conclusion of the lemma follows from (11), (12), and Proposition 1. □

II.2 Main Results

In this section we collect the main results of this chapter.

II.2.1 Admissible perturbations of a sampling model

In practice, shift invariant spaces are used to model classes of signals that can occur (or that are allowed) in applications. However, often, the functions in a shift invariant space model only give approximations to the signals of interest. Moreover, signal samples are obtained using measuring devices with characteristics that are not fully known, and the measurements reflect local averages rather than exact sample values. Thus, a sampling measure $\overrightarrow{\mu}$ is a model that approximates the characteristics of a measuring device. Another source of uncertainty is the location of the sampling points $\{x_j\}$. This type of uncertainty is what is often called jitter error (see e.g., [12, 18] and the references therein). The jitter error can be modeled as a perturbation of the sampling set $X = \{x_j\}$. The true sampling set is $\widetilde{X} = X + \Delta = \{x_j + \delta_j\}_{j\in J}$, where $\Delta = \{\delta_j\}_{j\in J} \subset \mathbb{R}^d$.

The following result shows that the sampling operator varies continuously with respect to all of the above parameters. As a corollary, we conclude that sampling models considered in this paper remain stable under the three types of perturbations: 1) a perturbation due to a small change of the generators of the underlying shift invariant space; 2) a perturbation of the vector of measures $\overrightarrow{\mu}$ (uncertainty about the characteristics of the measuring devices); and 3) a perturbation $\Delta = \{\delta_j\}_{j\in J} \subset \mathbb{R}^d$ of the set of sampling X (jitter). We use the standard notation for $\|\Delta\|_\infty = \sup\{\|\delta_j\| : j \in J\}$.

Theorem 1. *Let $(X, \Phi, \overrightarrow{\mu})$ be a p-stable sampling model for some $p \in [1, \infty]$ and U be its sampling operator. Let also $(X + \Delta, \Theta, \overrightarrow{\alpha})$ be a perturbed sampling model with the sampling operator U_Δ. Then for every $\epsilon > 0$ there exists $\epsilon_0 > 0$ such that $\|U - U_\Delta\| < \epsilon$, whenever $\overrightarrow{\alpha} \in \mathcal{M}^t(\mathbb{R}^d)$, $\Theta \in (W_0^1)^r$, and*

$$\|\Delta\|_\infty + \|\Phi - \Theta\|_{(W^1)^r} + \|\overrightarrow{\mu} - \overrightarrow{\alpha}\|_{\mathcal{M}^t} < \epsilon_0.$$

Due to Lemma 1 we obtain the following corollary.

Corollary 1. *Let $(X, \Phi, \overrightarrow{\mu})$ be a p-stable sampling model for some $p \in [1, \infty]$. Then there exists $\epsilon_0 > 0$ such that the sampling model $(X + \Delta, \Theta, \overrightarrow{\alpha})$ is also p-stable, whenever $\overrightarrow{\alpha} \in \mathcal{M}^t(\mathbb{R}^d)$, $\Theta \in (W_0^1)^r$, and $\|\Delta\|_\infty + \|\Phi - \Theta\|_{(W^1)^r} + \|\overrightarrow{\mu} - \overrightarrow{\alpha}\|_{\mathcal{M}^t} < \epsilon_0$. That is, there exist $0 < A_p^{'} \le B_p^{'} < \infty$ such that*

$$A_p^{'}\|f\|_{L^p} \le \|(f * \overrightarrow{\alpha})(\widetilde{X})\|_{(\ell^p(J))^t} \le B_p^{'}\|f\|_{L^p}, \text{ for all } f \in V^p(\Theta), \tag{13}$$

where $\widetilde{X} = X + \Delta$.

Remark 3. *The proof of Theorem 1 will be established in Section II.3.2 by perturbing some of the parameters of the sampling models while leaving the other unchanged, and then combining the results. For some of these cases, we will provide explicit estimates for ϵ_0 and the bounds $A_p^{'}$ and $B_p^{'}$. For example, for the case where $\|\Delta\|_\infty = \|\overrightarrow{\mu} - \overrightarrow{\alpha}\|_{\mathcal{M}^t} = 0$, we get that*

$$\epsilon_0 = \frac{1}{2}\left(\sqrt{C_p^2 + \frac{4A_p m_p^2}{2^d \mathcal{N}_p \|\overrightarrow{\mu}\|_{\mathcal{M}^t}}} - C_p\right),$$

where $\mathcal{N}_p$ is as in Remark 1 and

$$C_p = \|\Phi\|_{(W^1)^r} + \frac{A_p m_p}{2^d \mathcal{N}_p \|\overrightarrow{\mu}\|_{\mathcal{M}^t}}.$$

Then, for $0 < \epsilon < \epsilon_0 < m_p$, we have

$$A_p' = \frac{A_p m_p}{\|\Phi\|_{(W^1)^r} + \epsilon} - \frac{2^d \mathcal{N}_p \|\overrightarrow{\mu}\|_{\mathcal{M}^t}}{m_p - \epsilon}\epsilon,$$

$$B_p' = \frac{2^d \mathcal{N}_p \|\overrightarrow{\mu}\|_{\mathcal{M}^t}(\|\Phi\|_{(W^1)^r} + \epsilon)}{m_p - \epsilon}.$$

The above result has many special cases and some of them have been obtained before. For example, Corollaries 2 and 3 below were first proved in [11]. They immediately follow from Corollary 1 with $d\overrightarrow{\mu} = \Psi dx$ for Corollary 2, and $\overrightarrow{\mu} = (\delta_0)$ for Corollary 3.

Corollary 2. *Let $\Psi \in (L^1(\mathbb{R}^d))^t$, $\Phi \in (W_0^1)^r$ satisfy (6), and $X = \{x_j, j \in J\} \subset \mathbb{R}^d$ be a relatively separated Ψ-sampling set for $V^p(\Phi)$. Then there exists $\epsilon_0 > 0$ such that X is a Ψ-sampling set for $V^p(\Theta)$, whenever $\Theta \in (W_0^1)^r$ and $\|\Phi - \Theta\|_{(W^1)^r} < \epsilon_0$.*

Corollary 3. *Let $\Phi \in (W_0^1)^r$ satisfying (6) and $X = \{x_j, j \in J\} \subset \mathbb{R}^d$ be a relatively separated ideal set of sampling for $V^p(\Phi)$. Then there exists $\epsilon_0 > 0$ such that X is an ideal set of sampling for $V^p(\Theta)$, whenever $\Theta \in (W_0^1)^r$ and $\|\Phi - \Theta\|_{(W^1)^r} < \epsilon_0$.*

Other special cases are analogous to Theorem 3.6 in [12], and [80, Theorem 3.4]. The amplitude error considered in [80, Theorem 3.3] can also be described in the above terms as a perturbation of $\overrightarrow{\mu} = (\delta_0)$.

II.2.2 Perfect reconstruction and localized frames

In this section we show that the standard dual frame method can be used to reconstruct $f \in V^2(\Phi)$ from its samples. We also obtain a useful modification of the above results using the theory of localized frames developed in [59] (see Definition 8). In the previous section, the number $p \in [1, \infty]$ was fixed, that is, we stated, for example, that if X is a $\overrightarrow{\mu}$-sampling set for $V^p(\Phi)$, then X is a $\overrightarrow{\mu}$-sampling set for $V^p(\Theta)$ for *the same* $p \in [1, \infty]$, as soon as Θ is sufficiently close to Φ in the appropriate norm. Here, we claim that if X is a $\overrightarrow{\mu}$-sampling set for $V^2(\Phi)$, then X is a $\overrightarrow{\mu}$-sampling set for $V^p(\Theta)$ for *all* $p \in [1, \infty]$, as soon as Θ is sufficiently close to Φ, Φ satisfies a mild decay condition, and $\overrightarrow{\mu}$ belongs to $\mathcal{M}_s(\mathbb{R}^d)$ for some $s > d$. It is natural to ask whether one can replace $V^2(\Phi)$ in the above statement with $V^q(\Phi)$, *for some* $q \in [1, \infty]$. Under certain assumptions the answer is "yes", but it turns out to be a much harder problem as shown in [5].

Definition 5. *Let $\mathcal{H}$ be a Hilbert space of functions and V a closed subspace of $\mathcal{H}$. Let $\{\Psi_{x_j} = (\psi_{x_j}^1, \dots, \psi_{x_j}^t)^T\}_{j \in J}$ be a countable collection of vectors of functions in V. We say that $\{\Psi_{x_j}\}_{j \in J}$ is a frame for V if there exist constants $0 < A \le B < \infty$ such that*

$$A\|f\|_{\mathcal{H}} \le \|\langle f, \Psi_{x_j}\rangle\|_{(\ell^2(J))^t} \le B\|f\|_{\mathcal{H}}, \text{ for all } f \in V,$$

where $\langle f, \Psi_{x_j}\rangle = (\langle f, \psi_{x_j}^1\rangle, \dots, \langle f, \psi_{x_j}^t\rangle)^T \in \mathbb{C}^t$.

Remark 4. *Notice that the above is not quite the standard definition of a frame in a Hilbert space. This is due to the way we defined the norm in* (5)*. Nevertheless, it is easily seen that* $\{\Psi_{x_j}\}_{j\in J}$ *is a frame for* V *according to the above definition if and only if* $\{\psi^i_{x_j},\ i=1,2,\dots,t,\ j\in J\}$ *is a frame for* V *according to the standard definition. The frame bounds, however, may be different.*

Definition 6. *Let* V *be a closed subspace of the Hilbert space* $\mathcal{H}$*. Let* $\{\Psi_{x_j} = (\psi^1_{x_j},\dots,\psi^t_{x_j})^T\}_{j\in J}$ *be a frame for* V*. The frame operator associated with the frame* $\{\Psi_{x_j}\}_{j\in J}$ *is the operator* $S : V \to V$ *defined by* $S(f) = \sum_{j\in J}\langle f, \Psi_{x_j}\rangle^T \Psi_{x_j}$*, for all* $f \in V$*. The (canonical) dual frame* $\{\widetilde{\Psi}_{x_j}\}_{j\in J}$ *of the frame* $\{\Psi_{x_j}\}_{j\in J}$ *is the sequence of vectors given by* $\{\widetilde{\Psi}_{x_j} = (\widetilde{\psi}^1_{x_j},\dots,\widetilde{\psi}^t_{x_j})^T\}_{j\in J}$*, where* $\widetilde{\psi}^s_{x_j} = S^{-1}\psi^s_{x_j}$, $1\le s\le t$.

Remark 5. *It is well know that a frame operator* S *is bounded, invertible, self-adjoint, and positive [48]. Hence, the canonical dual frame is well defined. There may exist other dual frames but we will refrain from defining the notion.*

The next well-known result shows that the dual frame method can be used to reconstruct a function from its samples. First, let us state the following definition.

Definition 7. *Let* J *be a countable set of index, and* $(\mathcal{B}, \|\cdot\|)$ *a Banach space. We shall say that the series* $\sum_{j\in J} g_j$ *converges unconditionally in* $\mathcal{B}$ *to* g*, if for every* $\epsilon > 0$ *there exists a finite set* $F = F(\epsilon) \subset J$ *such that*

$$\|g - \sum_{j\in I} g_j\| < \epsilon,$$

for all finite subsets $I \subset J$ *containing* F*.*

Proposition 2. *Let* $\Phi \in (W^1_0)^r$, $\overrightarrow{\mu} \in \mathcal{M}^t(\mathbb{R}^d)$*, and* X *be a* $\overrightarrow{\mu}$*-sampling set for* $V^2(\Phi)$*. Then there exists a sequence of vectors of functions* $\{\Psi_{x_j}\}_{j\in J}$*, which is a frame for* $V^2(\Phi)$ *and* $\langle f, \Psi_{x_j}\rangle = (f * \overrightarrow{\mu})(x_j)$ *for all* $f \in V^2(\Phi)$ *and* $j \in J$*. Moreover, every function* $f \in V^2(\Phi)$ *can be recovered from the sequence of its samples* $\{(f*\overrightarrow{\mu})(x_j)\}_{j\in J}$ *via*

$$f(x) = \sum_{j\in J}(f * \overrightarrow{\mu})^T(x_j)\widetilde{\Psi}_{x_j}(x), \tag{14}$$

where $\{\widetilde{\Psi}_{x_j}\}_{j\in J}$ *is the dual frame of* $\{\Psi_{x_j}\}_{j\in J}$ *and the series (14) converges unconditionally in* $V^2(\Phi)$*.*

The frame $\{\Psi_{x_j}\}_{j\in J}$ constructed in the previous proposition will be called a $(\overrightarrow{\mu}, X)$*-sampling frame* for $V^2(\Phi)$. The main idea of this section is to use the fact that if such a frame is localized then it is also a Banach frame [59] for $V^p(\Phi)$, $p \in [1,\infty)$.

Remark 6. *Observe that, in general, the frame operator* S *is the product of the* analysis operator $T : V \to (\ell^2(J))^t$*, defined by*

$$Tf = \{\langle f, \Psi_{x_j}\rangle\}_{j\in J} = \{(\langle f, \psi^1_{x_j}\rangle, \dots, \langle f, \psi^t_{x_j}\rangle)^T\}_{j\in J}$$

and its adjoint, that is $S = T^*T$*. Since* Φ *generates a Riesz basis, it is immediate that in case of a* $(\overrightarrow{\mu}, X)$*-sampling frame, the sampling operator* U *can be viewed as the matrix of the analysis operator* T *in the the basis generated by* Φ*.*

Recall that a sequence $\{\phi^1(\cdot - k), \ldots, \phi^r(\cdot - k); k \in \mathbb{Z}^d\}$ in a Hilbert space $\mathcal{H}$ is a Riesz basis for $\mathcal{H}$ if it is a basis for $\mathcal{H}$ and if it satisfies Condition (6) for $p = 2$. It is well known that the dual basis of a Riesz basis is itself a Riesz basis.

Definition 8. *Let V be a closed subspace of the Hilbert space $\mathcal{H}$. Let $\{\Psi_{x_j} = (\psi^1_{x_j}, \ldots, \psi^t_{x_j})^T\}_{j\in J}$ be a frame for V, and $\{G_k = (g^1_k, \ldots, g^r_k)^T\}_{k\in\mathbb{Z}^d}$ be a Riesz basis for V. We say that the frame $\{\Psi_{x_j}\}_{j\in J}$ is (polynomially) s-localized with respect to the Riesz basis $\{G_k\}_{k\in\mathbb{Z}^d}$, if*

$$|\langle G_k, \Psi^T_{x_j}\rangle| \le C_1(1 + |x_j - k|)^{-s}, \tag{15}$$

and

$$|\langle \widetilde{G}_k, \Psi^T_{x_j}\rangle| \le C_2(1 + |x_j - k|)^{-s}, \tag{16}$$

for all $j \in J$ and $k \in \mathbb{Z}^d$. Here, the constants $C_1, C_2 > 0$ are independent of j and k, $|\langle G_k, \Psi^T_{x_j}\rangle| = \sum_{i=1}^r \sum_{l=1}^t |\langle g^i_k, \psi^l_{x_j}\rangle|$, $\{\widetilde{G}_k\}_{k\in\mathbb{Z}^d}$ is the dual Riesz basis of $\{G_k\}_{k\in\mathbb{Z}^d}$, and $|\langle \widetilde{G}_k, \Psi^T_{x_j}\rangle|$ is defined similarly to $|\langle G_k, \Psi^T_{x_j}\rangle|$.

Remark 7. *Let V be a closed subspace of a Hilbert space $\mathcal{H}$. Assume that $\{G_k = (g^1_k, \ldots, g^r_k)^T\}_{k\in\mathbb{Z}^d}$ is a Riesz basis for V. The dual Riesz basis of the Riesz basis $\{G_k\}_{k\in\mathbb{Z}^d}$ is the sequence of vectors $\{\widetilde{G}_k = (\widetilde{g}^1_k, \ldots, \widetilde{g}^r_k)^T\}_{k\in\mathbb{Z}^d}$ satisfying $\langle \widetilde{G}_k, G^T_l\rangle = \delta_{kl} I$, where I is the $r \times r$ identity matrix, and δ_{kl} is the Kronecker delta. Since a Riesz basis $\{G_k\}$ is also a frame, $\{\widetilde{G}_k\}$ is, in fact, the canonical dual frame for $\{G_k\}$. In this case it is the unique dual frame.*

Definition 9. *Let $\Phi = (\phi^1, \ldots, \phi^r)^T \in (W^1_0)^r \subset (L^2(\mathbb{R}^d))^r$ and $s > d$. We say that Φ is an s-localized Riesz generator for $V^2(\Phi)$, denoted $\Phi \in \mathcal{W}_s$, if*

- *$\{\Phi_k = \Phi(\cdot - k)\}_{k\in\mathbb{Z}^d}$ generates a Riesz basis for $V^2(\Phi)$, i.e., condition (6) holds for $p = 2$;*
- *The components of Φ satisfy the decay condition*

$$|\phi^i(x)| \le C^i_0(1 + |x|)^{-s}, \tag{17}$$

for all $1 \le i \le r$ and some $C^i_0 > 0$ independent of $x \in \mathbb{R}^d$.

Remark 8. *If $\Phi \in \mathcal{W}_s$, then* (6) *holds for every $p \in [1, \infty]$ (see e.g., [9, 10]).*

The following two theorems are the main results of Section II.2.2.

Theorem 2. *Let $s > d$, $\Phi \in \mathcal{W}_s$, and $\overrightarrow{\mu} \in \mathcal{M}^t_s(\mathbb{R}^d)$. Assume that X is a $\overrightarrow{\mu}$-sampling set for $V^2(\Phi)$, and $\{\Psi_{x_j}\}_{j\in J}$ is the $(\overrightarrow{\mu}, X)$-sampling frame for $V^2(\Phi)$. Then*

- *X is a $\overrightarrow{\mu}$-sampling set for $V^p(\Phi)$ for all $p \in [1, \infty]$.*
- *If $\{\widetilde{\Psi}_{x_j}\}$ is the dual frame for $\{\Psi_{x_j}\}_{j\in J}$, then*

$$f = \sum_{j\in J} (f * \overrightarrow{\mu})^T(x_j)\widetilde{\Psi}_{x_j}, \text{ for all } f \in V^p(\Phi), \tag{18}$$

where the series converges unconditionally in $V^p(\Phi)$, $p \in [1, \infty)$.

Next, we combine Theorem 2 with the perturbation results of the previous section. The proof is immediate.

Theorem 3. *Let $s > d$, $\Phi \in \mathcal{W}_s$, and $\overrightarrow{\mu} \in \mathcal{M}_s^t(\mathbb{R}^d)$. Assume that X is a relatively separated $\overrightarrow{\mu}$-sampling set for $V^2(\Phi)$. Then there exists $\epsilon_0 > 0$ such that for every $\Delta = \{\delta_j,\ j \in J\}$, $\Theta \in \mathcal{W}_s$, and $\overrightarrow{\alpha} \in \mathcal{M}_s^t(\mathbb{R}^d)$ satisfying $\|\Delta\|_\infty + \|\Phi - \Theta\|_{(W^1)^r} + \|\overrightarrow{\mu} - \overrightarrow{\alpha}\|_{\mathcal{M}^t} < \epsilon_0$, there exists an $(\overrightarrow{\alpha}, X + \Delta)$-sampling frame $\{\Psi_{x_j}\}_{j \in J}$ for $V^2(\Theta)$. Moreover,*

- *$X + \Delta$ is an $\overrightarrow{\alpha}$-sampling set for $V^p(\Theta)$ for all $p \in [1, \infty]$.*
- *If $\{\widetilde{\Psi}_{x_j}\}$ is the dual frame for $\{\Psi_{x_j}\}_{j \in J}$, then*

$$f = \sum_{j \in J} (f * \overrightarrow{\alpha})^T (x_j + \delta_j) \widetilde{\Psi}_{x_j}, \ \textit{for all } f \in V^p(\Theta),$$

where the series converges unconditionally in $V^p(\Theta)$, $p \in [1, \infty)$.

Remark 9. *The crucial result for the proof of the theorems is Jaffard's non-commutative extension of the classical Wiener's Tauberian Lemma (see Theorem 5 in [59]). It states that if an invertible matrix has an off-diagonal decay defined by inequalities similar to* (15) *and* (16)*, then the inverse matrix has the same off-diagonal decay. There exist other extensions of Wiener's Lemma which deal with different types of off-diagonal decay (see, for example, [20, 60]). Many of those could be used to obtain results similar to Theorem 3.*

II.2.3 Imperfect reconstruction

In practice, we know that a perturbation exists because of imperfections of measuring devices, errors, etc. However, we can only estimate this perturbation and may not even know its nature. Here we show that even if we use a model $(X, \Phi, \overrightarrow{\mu})$ for reconstructing a signal from a perturbed model $(\widetilde{X}, \Theta, \overrightarrow{\alpha})$ (or vice versa), the reconstruction error depends continuously on the perturbation in the cases studied above.

As before, let U be the sampling operator for a p-stable sampling model $(X, \Phi, \overrightarrow{\mu})$ and U_Δ be the sampling operator for a perturbed model $(\widetilde{X}, \Theta, \overrightarrow{\alpha})$, where $\widetilde{X} = X + \Delta = \{x_j + \delta_j\}_{j \in J}$. The sampling operator U_Δ can be thought of as a $t \times r$ block matrix

$$U_\Delta = \begin{pmatrix} U_\Delta^{1,1} \dots U_\Delta^{1,r} \\ \vdots \qquad \vdots \\ U_\Delta^{t,1} \dots U_\Delta^{t,r} \end{pmatrix},$$

where for each $1 \le i \le r$ and $1 \le l \le t$ the operator $U_\Delta^{l,i}$ is defined by a bi-infinite matrix with entries $(U_\Delta^{l,i})_{j,k} = (\theta^i * \alpha^l)(x_j + \delta_j - k)$, $j \in J$, $k \in \mathbb{Z}^d$.

We let U^* be an operator defined by the following $r \times t$ matrix of operators from $(\ell^p(J))^t$ into $(\ell^p(\mathbb{Z}^d))^r$:

$$U^* = \begin{pmatrix} \overline{(U^{1,1})^T} \dots \overline{(U^{t,1})^T} \\ \vdots \qquad\quad \vdots \\ \overline{(U^{1,r})^T} \dots \overline{(U^{t,r})^T} \end{pmatrix},$$

where for each $1 \le i \le r$ and $1 \le l \le t$ the matrix $\overline{(U^{l,i})^T}$ is the complex conjugate transpose of $U^{l,i}$. If $p = 2$ and $t = r = 1$, then U^* is the Hilbert adjoint of U. The operator $(U_\Delta)^*$ is defined similarly. Notice that this definition implies $(U^*)^* = U$. Moreover, U^* is a bounded operator and $\|U^*\|_{p,op} = \|U\|_{q,op}$, $\frac{1}{p} + \frac{1}{q} = 1$. Hence, $\|U^*\|_{p,op} \le 2^d \mathcal{N}_q \|\overrightarrow{\mu}\|_{\mathcal{M}^t} \|\Phi\|_{(W^1)^r}$, $\quad \frac{1}{q} + \frac{1}{p} = 1$.

In the next two theorems we shall assume that the operator U^*U is invertible. Consequently, there is $0 < n_p < \infty$ such that $\|(U^*U)^{-1}\|_{p,op} \le n_p$. If the sampling model $(X, \Phi, \overrightarrow{\mu})$ is 2-stable, U^*U is a matrix of the frame operator S for the sampling frame $\{\Psi_{x_j}\}$, see Remark 6. Therefore, the bi-infinite matrix U^*U is invertible and positive definite. Moreover, the operator $(U^*U)^{-1}U^*$ is a left inverse for the sampling operator U and it can be viewed as the matrix of the synthesis operator used for the reconstruction. Hence, the importance of the following result.

Theorem 4. *Let $(X, \Phi, \overrightarrow{\mu})$ be a p-stable sampling model for some $p \in [1, \infty]$. Assume that its sampling operator U satisfies (10) and the operator U^*U is invertible. Let $0 < 2\epsilon < \left(-(\beta_p + \beta_q) + \sqrt{(\beta_p + \beta_q)^2 + \frac{4}{n_p}}\right)$ and $(\widetilde{X}, \Theta, \overrightarrow{\alpha})$ be a perturbed sampling model such that its sampling operator U_Δ satisfies $\|U - U_\Delta\| < \epsilon$. Define $\nu = \nu(\epsilon) = n_p\epsilon(\epsilon + \beta_p + \beta_q)$. Then $0 < \nu < 1$, the operator $U_\Delta^* U_\Delta$ is invertible, and*

$$\|(U^*U)^{-1}U^* - (U_\Delta^* U_\Delta)^{-1} U_\Delta^*\| < n_p \left(\epsilon + \frac{\nu(\beta_q + \epsilon)}{1 - \nu} \right),$$

for $\|\Delta\|_\infty$ sufficiently small.

Remark 10. *Observe that if $p = 2$ we do not need to require invertibility of U^*U. As we mentioned above, it follows automatically.*

Remark 11. *If in Theorem 4 we let $r = t = 1$, $p = 2$, and $\overrightarrow{\mu} = (\mu^1) = (\delta_0)$, then we obtain an analog of Theorem 3.3 in [12].*

Let $(X, \Phi, \overrightarrow{\mu})$ be a p-stable sampling model for some $p \in [1, \infty]$. Assume that its sampling operator U satisfies (10) and the operator U^*U is invertible. We define the *reconstruction operator* $R = R_{(X, \Phi, \overrightarrow{\mu})} : (\ell^p(J))^t \to V^p(\Phi)$ by

$$RD = \sum_{k \in \mathbb{Z}^d} [(U^*U)^{-1} U^* D]_k^T \Phi(\cdot - k),$$

$D = (d^1, \dots, d^t)^T$ in $(\ell^p(J))^t$.

Then as an immediate consequence of Theorems 1 and 4, we have the following result.

Theorem 5. *Let $(X, \Phi, \overrightarrow{\mu})$ be a p-stable sampling model for some $p \in [1, \infty]$. Assume that its sampling operator U is such that U^*U is invertible. Let R be the reconstruction operator. Then for every $\epsilon > 0$ there exists $\epsilon_0 > 0$ such that for every $\Delta = \{\delta_j,\ j \in J\}$, $\Theta \in (W_0^1)^r$, and $\overrightarrow{\alpha} \in \mathcal{M}^t(\mathbb{R}^d)$ satisfying*

$$\|\Delta\|_\infty + \|\Phi - \Theta\|_{(W^1)^r} + \|\overrightarrow{\mu} - \overrightarrow{\alpha}\|_{\mathcal{M}^t} < \epsilon_0,$$

we have

$$\|R((g * \overrightarrow{\alpha})(X + \Delta)) - f\|_{L^p} < \epsilon \|f\|_{L^p},\ f = \sum_{k \in \mathbb{Z}^d} C_k^T \Phi_k,\ g = \sum_{k \in \mathbb{Z}^d} C_k^T \Theta_k,$$

for all $C \in (\ell^p(\mathbb{Z}^d))^r$.

Theorem 5 tells us that the reconstruction error is, indeed, controlled in a continuous fashion by each and all of the perturbation errors studied in this paper.

Our final result is a combination of the above theorem with the results of Section II.2.1.

Theorem 6. *Let $(X, \Phi, \overrightarrow{\mu})$ be a 2-stable sampling model such that $\Phi \in \mathcal{W}_s$ and $\overrightarrow{\mu} \in \mathcal{M}_s^t(\mathbb{R}^d)$, with $s > d$. Let R be the reconstruction operator for $(X, \Phi, \overrightarrow{\mu})$. Then for every $\epsilon > 0$ there exists $\epsilon_0 > 0$ such that for every $\Delta = \{\delta_j,\ j \in J\}$, $\Theta \in \mathcal{W}_s$, and $\overrightarrow{\alpha} \in \mathcal{M}_s^t(\mathbb{R}^d)$ satisfying*

$$\|\Delta\|_\infty + \|\Phi - \Theta\|_{(W^1)^r} + \|\overrightarrow{\mu} - \overrightarrow{\alpha}\|_{\mathcal{M}^t} < \epsilon_0,$$

we have

$$\|R((g * \overrightarrow{\alpha})(X + \Delta)) - f\|_{L^p} < \epsilon \, \|f\|_{L^p},\ f = \sum_{k \in \mathbb{Z}^d} C_k^T \Phi_k,\ g = \sum_{k \in \mathbb{Z}^d} C_k^T \Theta_k,$$

for all $p \in [1, \infty]$ and all $C \in (\ell^p(\mathbb{Z}^d))^r$.

The proofs in the following section show implicitly how numerical estimates for ϵ_0 in Theorems 5 and 6 may be obtained.

II.3 Proofs

II.3.1 Auxiliary results

We begin with some technical results that are needed for the main proofs.

Lemma 2. *Let $\phi \in W_0^1$, and $\mu \in \mathcal{M}(\mathbb{R}^d)$. Then:*

$$\phi * \mu \in W_0^1, \text{ and} \tag{19}$$

$$\|\phi * \mu\|_{W^1} \le 2^d \|\phi\|_{W^1} \|\mu\|. \tag{20}$$

Proof. Note that if $\mu = 0$, the proof is immediate. Assume now $\mu \neq 0$, i.e. $\|\mu\| > 0$. Let $\epsilon > 0$ be given. Since $\phi \in W_0^1$, then ϕ is uniformly continuous in $\mathbb{R}^d$. Therefore, there exists $\delta = \delta(\epsilon) > 0$ such that

$$|\phi(w) - \phi(w_1)| < \frac{\epsilon}{\|\mu\|}, \quad \text{whenever} \quad \|w - w_1\| < \delta. \tag{21}$$

Let $z_0 \in \mathbb{R}^d$ be given, and let $z \in \mathbb{R}^d$ be such that $\|z - z_0\| < \delta$. Then we have

$$\begin{aligned} |(\phi * \mu)(z) - (\phi * \mu)(z_0)| &= \left| \int_{\mathbb{R}^d} \phi(z-y) d\mu(y) - \int_{\mathbb{R}^d} \phi(z_0 - y) d\mu(y) \right| \\ &\leq \int_{\mathbb{R}^d} |\phi(z-y) - \phi(z_0 - y)| d|\mu|(y). \end{aligned}$$

Since $\|(z-y) - (z_0 - y)\| = \|z - z_0\| < \delta$, for all $y \in \mathbb{R}^d$, then it follows from (21) that $\int_{\mathbb{R}^d} |(\phi(z-y) - \phi(z_0 - y))| \, d|\mu|(y) < \int_{\mathbb{R}^d} \frac{\epsilon}{\|\mu\|} d|\mu|(y) = \epsilon$. Since z_0 and $\epsilon > 0$ are arbitrary, we obtain the continuity of $\phi * \mu$ in $\mathbb{R}^d$.

Let us show (20). Let $\phi \in W^1$ and $\mu \in \mathcal{M}(\mathbb{R}^d)$ be given. Then

$$\begin{gathered} \|\phi * \mu\|_{W^1} = \sum_{k \in \mathbb{Z}^d} \operatorname*{ess\,sup}_{x \in [0,1]^d} \left| \int_{\mathbb{R}^d} \phi(x + k - y) \, d\mu(y) \right| \leq \\ \sum_{k \in \mathbb{Z}^d} \int_{\mathbb{R}^d} \operatorname*{ess\,sup}_{x \in [0,1]^d} |\phi(x + k - y)| \, d|\mu|(y) \leq \\ \int_{\mathbb{R}^d} \left(\sum_{k \in \mathbb{Z}^d} \operatorname*{ess\,sup}_{x \in [0,1]^d} |\phi(x + k - y)| \right) d|\mu|(y) = \int_{\mathbb{R}^d} \|\phi(\cdot - y)\|_{W^1} \, d|\mu|(y). \end{gathered}$$

Since $\|\phi(\cdot - y)\|_{W^1} \leq 2^d \|\phi\|_{W^1}$, for all $y \in \mathbb{R}^d$, we get

$$\int_{\mathbb{R}^d} \|\phi(\cdot - y)\|_{W^1} d|\mu|(y) \leq \int_{\mathbb{R}^d} 2^d \|\phi\|_{W^1} d|\mu|(y) = 2^d \|\phi\|_{W^1} \|\mu\|.$$

Therefore, we get (20). $\square$

The next proposition collects basic known facts about Wiener amalgam spaces, shift invariant spaces $V^p(\Phi)$, and relatively separated sets in $\mathbb{R}^d$.

Proposition 3. *Let* $\Phi \in (W_0^1)^r$, $\overrightarrow{\mu} \in \mathcal{M}^t(\mathbb{R}^d)$, $f = \sum_{k \in \mathbb{Z}^d} C_k^T \Phi_k$, *where* $C \in (\ell^p(\mathbb{Z}^d))^r$, *and* $\Phi_k = \Phi(\cdot - k)$, *for all* $k \in \mathbb{Z}^d$. *Let also* $X = \{x_j, j \in J\}$ *be a relatively separated set in* $\mathbb{R}^d$ *with* $\mathcal{S}$ *as is in Definition 2. Then*

$$\Phi * \overrightarrow{\mu}^T \in (W_0^1)^{r \times t}; \tag{22}$$

$$\|\Phi * \overrightarrow{\mu}^T\|_{(W^1)^{r \times t}} \leq 2^d \|\Phi\|_{(W^1)^r} \|\overrightarrow{\mu}\|_{\mathcal{M}^t}; \tag{23}$$

$$V^p(\Phi) \subset W_0^p, \ \text{for all } 1 \leq p \leq \infty; \tag{24}$$

$$\|f\|_{W^p} \leq \|C\|_{(\ell^p(\mathbb{Z}^d))^r} \|\Phi\|_{(W^1)^r}. \tag{25}$$

For $g \in W^p$ we have

$$\|g(X)\|_{\ell^p(J)} \leq \mathcal{N}_p \|g\|_{W^p},\ \mathcal{N}_p = \mathcal{S}^{1/p},\ p \in [1,\infty),\ \mathcal{N}_\infty = 1. \tag{26}$$

Proof. First, Lemma 2 immediately implies (22). Next, to prove (23) consider $\Phi = (\phi^1,\dots,\phi^r)^T \in (W_0^1)^r$ and $\overrightarrow{\mu} = (\mu^1,\dots,\mu^t)^T \in \mathcal{M}^t(\mathbb{R}^d)$. Then using (20), we obtain

$$\|\Phi * \overrightarrow{\mu}^T\|_{(W^1)^{r\times t}} = \sum_{j=1}^t \sum_{i=1}^r \|\phi^i * \mu^j\|_{W^1} \leq$$
$$\sum_{j=1}^t \sum_{i=1}^r 2^d \|\phi^i\|_{W^1} \|\mu^j\| = 2^d \|\Phi\|_{(W^1)^r} \|\overrightarrow{\mu}\|_{\mathcal{M}^t}.$$

Next, we prove (25). Consider $1 \leq p < \infty$ and $f = \sum_{k\in\mathbb{Z}^d} C_k^T \Phi_k$. For each $1 \leq s \leq r$ let $a^s(l) = \operatorname*{ess\,sup}_{x\in[0,1]^d} |\phi^s(x+l)|$, for all $l \in \mathbb{Z}^d$. Then $\|a^s\|_{\ell^1(\mathbb{Z}^d)} = \|\phi^s\|_{W^1}$. Consequently, $\|a\|_{(\ell^1(\mathbb{Z}^d))^r} = \|\Phi\|_{(W^1)^r}$, where $a = (a^1,\dots,a^r)^T$, and $\Phi = (\phi^1,\dots,\phi^r)^T$. Hence,

$$\operatorname*{ess\,sup}_{x\in[0,1]^d} |f(x+l)| \leq \sum_{s=1}^r \sum_{k\in\mathbb{Z}^d} |c^s(k)| \operatorname*{ess\,sup}_{x\in[0,1]^d} |\phi^s(x+l-k)| = \sum_{s=1}^r (a^s * |c^s|)(l),$$

where we have also used the notation $c^s(k)$ for c_k^s. By using the Young and Triangle inequalities, we have

$$\|f\|_{W^p} \leq \sum_{s=1}^r \|a^s * |c^s|\|_{\ell^p} \leq \sum_{s=1}^r \|a^s\|_{\ell^1} \|c^s\|_{\ell^p}.$$

Consequently, $\|f\|_{W^p} \leq \|C\|_{(\ell^p(\mathbb{Z}^d))^r} \|\Phi\|_{(W^1)^r}$. Let us prove (25) for $p = \infty$. Let $m \in \mathbb{Z}^d$ be given. Then

$$\begin{aligned} |f(x+m)| &\leq \sum_{s=1}^r \sum_{k\in\mathbb{Z}^d} |c^s(k)| |\phi^s(x+m-k)| \\ &\leq \|C\|_{(\ell^\infty(\mathbb{Z}^d))^r} \sum_{s=1}^r \sum_{k\in\mathbb{Z}^d} |\phi^s(x+m-k)|. \end{aligned}$$

Consequently,

$$\begin{aligned} \operatorname*{ess\,sup}_{x\in[0,1]^d} |f(x+m)| &\leq \|C\|_{(\ell^\infty(\mathbb{Z}^d))^r} \sum_{s=1}^r \sum_{k\in\mathbb{Z}^d} \operatorname*{ess\,sup}_{x\in[0,1]^d} |\phi^s(x+m-k)| \\ &= \|C\|_{(\ell^\infty(\mathbb{Z}^d))^r} \sum_{s=1}^r \|\phi^s\|_{W^1} \\ &= \|C\|_{(\ell^\infty(\mathbb{Z}^d))^r} \|\Phi\|_{(W^1)^r}. \end{aligned}$$

Therefore,

$$\|f\|_{W^\infty} = \sup_{m\in\mathbb{Z}^d} \{\operatorname*{ess\,sup}_{x\in[0,1]^d} |f(x+m)|\} \leq \|C\|_{(\ell^\infty(\mathbb{Z}^d))^r} \|\Phi\|_{(W^1)^r}.$$

Next, let us show (24). Let $f \in V^p(\Phi)$ be given. Then $f = \sum_{k\in\mathbb{Z}^d} C_k^T \Phi_k$, for some $C \in (\ell^p(\mathbb{Z}^d))^r$. Since (25) implies $f \in W^p$, it remains to show the continuity of f. Let us first consider the case $1 \leq p < \infty$. We observe that $W^p \subset W^\infty = L^\infty(\mathbb{R}^d)$ (see Theorem 2.1 in [10]), and, hence,

$$\|f\|_{L^\infty} \leq \|f\|_{W^p}. \tag{27}$$

Let $f_n = \sum_{|k|\leq n} C_k^T \Phi_k$ be a partial sum of f. Since $\Phi \in (W_0^1)^r$, then $\{f_n\}_{n\in\mathbb{N}}$ is a sequence of continuous functions, and from (25) and (27) we obtain

$$\|f - f_n\|_{L^\infty} \leq \|\Phi\|_{(W^1)^r} \left(\sum_{i=1}^r \left(\sum_{|k|>n} |c_k^i|^p \right)^{1/p} \right).$$

Therefore, the sequence of continuous functions $\{f_n\}_{n\in\mathbb{N}}$ converges uniformly to the function f. Thus, f is a continuous function as well. To treat the case $p = \infty$, we choose a sequence $\{\Phi^n\}_{n\geq 1}$ of continuous functions with compact support (see Theorem 3.1 in [10] for details) such that $\|\Phi^n - \Phi\|_{(W^1)^r} \to 0$ as $n \to \infty$. Set $f_n(x) = \sum_{k\in\mathbb{Z}^d} C_k^T \Phi^n(x-k)$. Since the sum is locally finite, then each f_n is continuous. Using (25) once again, we estimate

$$\|f_n - f\|_{L^\infty} \leq \|C\|_{(\ell^\infty(\mathbb{Z}^d))^r} \|\Phi^n - \Phi\|_{(W^1)^r}.$$

It follows that the sequence of continuous functions $\{f_n\}_{n\geq 1}$ converges uniformly to f. Hence, f is a continuous function as well.

Finally, (26) follows from

$$\sum_{j:x_j\in[0,1]^d+l} |g(x_j)|^p \leq \mathcal{S} \operatorname*{ess\,sup}_{x\in[0,1]^d+l} |g(x)|^p,$$

and the proof is complete. □

Using (22) and (23), we obtain the following result.

Corollary 4. *Let* $\Lambda : (W_0^1)^r \times \mathcal{M}^t(\mathbb{R}^d) \longrightarrow (W_0^1)^{r\times t}$ *be defined by* $\Lambda(\Phi, \overrightarrow{\mu}) = \Phi * \overrightarrow{\mu}^T$. *Then* Λ *is a bounded bilinear form, and* $\|\Lambda\| \leq 2^d$, *where*

$$\|\Lambda\| = \sup\{\|\Lambda(\Phi, \overrightarrow{\mu})\|_{(W_0^1)^{r\times t}} : \|\Phi\|_{(W_0^1)^r} \leq 1, \|\overrightarrow{\mu}\|_{\mathcal{M}^t} \leq 1\}.$$

Lemma 3. *Let* $\phi \in W_0^1$, $\mu \in \mathcal{M}$, *and* $c = \{c_k\}_{k\in\mathbb{Z}^d} \in \ell^p(\mathbb{Z}^d)$ *be given. Then the function defined on* $\mathbb{R}^d$ *by* $h(z) = ((\sum_{k\in\mathbb{Z}^d} c_k\phi_k) * \mu)(z)$ *belongs to* W^p.

Proof. It suffices to show that

$$h(z) = \sum_{k\in\mathbb{Z}^d} c_k(\phi * \mu)_k(z), \forall z \in \mathbb{R}^d, \tag{28}$$

because if (28) is takes place, then the conclusion of the Lemma follows as a consequence of Lemma 2, and (25) in Proposition 3. In order to show (28), we shall show that

$$\sum_{k\in\mathbb{Z}^d} \int_{\mathbb{R}^d} |c_k||\phi(z-w-k)|d|\mu|(w) < \infty, \forall z \in [0,1]^d.$$

Let $z \in [0,1]^d$ be given. Then

$$\begin{aligned}
\sum_{k\in\mathbb{Z}^d} \int_{\mathbb{R}^d} |c_k||\phi(z-w-k)|d|\mu|(w) &= \int_{\mathbb{R}^d} \sum_{k\in\mathbb{Z}^d} |c_k||\phi(z-w-k)|d|\mu|(w) \\
&\le \|c\|_{\ell^\infty(\mathbb{Z}^d)} \int_{\mathbb{R}^d} \sum_{k\in\mathbb{Z}^d} |\phi(z-w-k)|d|\mu|(w) \\
&\le \|c\|_{\ell^\infty(\mathbb{Z}^d)} \int_{\mathbb{R}^d} \sum_{k\in\mathbb{Z}^d} \operatorname*{ess\,sup}_{z\in[0,1]^d} |\phi(z-w-k)|\, d|\mu|(w) \\
&= \|c\|_{\ell^\infty(\mathbb{Z}^d)} \int_{\mathbb{R}^d} \|\phi(\cdot - w)\|_{W^1} d|\mu|(w) \\
&\le 2^d \|c\|_{\ell^\infty(\mathbb{Z}^d)} \|\phi\|_{W^1} \|\mu\| < \infty.
\end{aligned}$$

□

The following lemma, proved, for example, in [11], states that a small perturbation of a Riesz basic sequence remains a Riesz basic sequence.

Lemma 4. *Let $\Phi \in (W^1)^r$ satisfy (6). Then there exists $\epsilon_0 > 0$ such that every $\Theta \in (W^1)^r$ satisfying $\|\Phi - \Theta\|_{(W^1)^r} \le \epsilon < \epsilon_0$, also satisfies (6), for some $0 < m_p^{'} \le M_p^{'} < \infty$ and*

$$m_p^{'} \ge m_p - \epsilon \qquad \textit{and} \qquad M_p^{'} \le \|\Phi\|_{(W^1)^r} + \epsilon. \tag{29}$$

The following lemma justifies Remark 2.

Lemma 5. *Let $(X, \Phi, \overrightarrow{\mu})$ be a p-stable sampling model and U be its sampling operator. Then*

$$\|U\|_{p,op} \le 2^d \mathcal{N}_p \, \|\overrightarrow{\mu}\|_{\mathcal{M}^t} \, \|\Phi\|_{(W^1)^r} \,,$$

and

$$\|U^*\|_{p,op} \le 2^d \mathcal{N}_q \, \|\overrightarrow{\mu}\|_{\mathcal{M}^t} \, \|\Phi\|_{(W^1)^r} \,, \quad \frac{1}{p} + \frac{1}{q} = 1.$$

Proof. The first of the inequalities follows from

$$\|UC\|_{(\ell^p(J))^t} = \|(f * \overrightarrow{\mu})(X)\|_{(\ell^p(J))^t}$$

and (23), (25), (26). The other one is true since $\|U^*\|_{p,op} = \|U\|_{q,op}$, $\frac{1}{p} + \frac{1}{q} = 1$. □

II.3.2 Proofs for Section II.2.1

We will divide the proof of Theorem 1 into several lemmas, each of which is a separate perturbation result. The first lemma is concerned with perturbations of the generator for the space of signals.

Lemma 6. *Let $(X, \Phi, \overrightarrow{\mu})$ be a p-stable sampling model for some $p \in [1, \infty]$. Then there exists $\epsilon_0 > 0$ such that the sampling model $(X, \Theta, \overrightarrow{\mu})$ is also p-stable, whenever $\Theta \in (W_0^1)^r$ and $\|\Phi - \Theta\|_{(W^1)^r} < \epsilon_0$.*

Proof. Assume that $\overrightarrow{\mu} \in \mathcal{M}^t(\mathbb{R}^d)$, $\Phi \in (W_0^1)^r$ satisfies (6), and $X = \{x_j, j \in J\} \subset \mathbb{R}^d$ satisfies (9). We want to find $\epsilon_0 > 0$ such that whenever $\|\Phi - \Theta\|_{(W^1)^r} < \epsilon_0$, then

$$A_p' \|g\|_{L^p} \le \|(g * \overrightarrow{\mu})(X)\|_{(\ell^p(J))^t} \le B_p' \|g\|_{L^p}, \text{ for all } g \in V^p(\Theta).$$

for some $0 < A_p' \le B_p' < \infty$. Assume $0 < \epsilon < m_p$. Then, by Lemma 4, $\Theta \in (W^1)^r$ satisfies (6) and we let $g = \sum_{k \in \mathbb{Z}^d} C_k^T \Theta_k$ and $f = \sum_{k \in \mathbb{Z}^d} C_k^T \Phi_k$, $C \in (\ell^p(\mathbb{Z}^d))^r$. Consequently, we have

$$\begin{aligned}
\frac{1}{M_p'} \|g\|_{L^p} \le \|C\|_{(\ell^p(\mathbb{Z}^d))^r} \le \frac{1}{m_p} \left\| \sum_{k \in \mathbb{Z}^d} C_k^T \Phi_k \right\|_{L^p} &= \frac{1}{m_p} \|f\|_{L^p} \\
&\le \frac{A_p^{-1}}{m_p} \|(f * \overrightarrow{\mu})(X)\|_{(\ell^p(J))^t} \\
&= \frac{A_p^{-1}}{m_p} \left\| \left(\left(\sum_{k \in \mathbb{Z}^d} C_k^T \Phi_k \right) * \overrightarrow{\mu} \right) (X) \right\|_{(\ell^p(J))^t} \\
&= \frac{A_p^{-1}}{m_p} \sum_{l=1}^{t} \left\| \left(\left(\sum_{k \in \mathbb{Z}^d} C_k^T \Phi_k \right) * \mu^l \right) (X) \right\|_{\ell^p(J)} \\
&\le \frac{A_p^{-1}}{m_p} \sum_{l=1}^{t} \left\| \left(\sum_{k \in \mathbb{Z}^d} C_k^T \Xi_k^l \right) (X) \right\|_{\ell^p(J)} \\
&+ \frac{A_p^{-1}}{m_p} \|(g * \overrightarrow{\mu})(X)\|_{(\ell^p(J))^t},
\end{aligned}$$

where

$$\Xi_k^l := ((\phi_k^1 - \theta_k^1) * \mu^l, \dots, (\phi_k^r - \theta_k^r) * \mu^l)^T, \quad l = 1, \dots, t. \tag{30}$$

Since Φ and Θ are elements of $(W_0^1)^r$ and $\overrightarrow{\mu} \in \mathcal{M}^t(\mathbb{R}^d)$, then by (22), we have $\Xi^l = (\Phi - \Theta) * \mu^l \in (W_0^1)^r$, for $l = 1, \dots, t$. Hence, using (23), (25), (26), and condition (6) for $g = \sum_{k \in \mathbb{Z}^d} C_k^T \Theta_k$, we have

$$\begin{gathered}
\textstyle\sum_{l=1}^{t} \left\| (\sum_{k \in \mathbb{Z}^d} C_k^T \Xi_k^l)(X) \right\|_{\ell^p(J)} \le \\
2^d \mathcal{N}_p \|C\|_{(\ell^p(\mathbb{Z}^d))^r} \|\Phi - \Theta\|_{(W^1)^r} \|\overrightarrow{\mu}\|_{\mathcal{M}^t} \le \\
\tfrac{2^d \mathcal{N}_p \|\Phi - \Theta\|_{(W^1)^r} \|\overrightarrow{\mu}\|_{\mathcal{M}^t}}{m_p'} \|g\|_{L^p}.
\end{gathered}$$

Therefore,

$$\frac{1}{M_p'}\|g\|_{L^p} \le \frac{A_p^{-1}2^d\mathcal{N}_p\|\Phi-\Theta\|_{(W^1)^r}\|\overrightarrow{\mu}\|_{\mathcal{M}^t}}{m_p m_p'}\|g\|_{L^p} + \\ + \frac{A_p^{-1}}{m_p}\|(g*\overrightarrow{\mu})(X)\|_{(\ell^p(J))^t}.$$

Hence,

$$\begin{aligned}\left(\frac{A_p m_p}{M_p'} - \frac{2^d\mathcal{N}_p\|\Phi-\Theta\|_{(W^1)^r}\|\overrightarrow{\mu}\|_{\mathcal{M}^t}}{m_p'}\right)\|g\|_{L^p} \\ \le \|(g*\overrightarrow{\mu})(X)\|_{(\ell^p(J))^t}.\end{aligned} \tag{31}$$

On the other hand, since $\Theta \in (W_0^1)^r$ and $\overrightarrow{\mu} \in \mathcal{M}^t(\mathbb{R}^d)$, it follows from (22) that $(\theta^1 * \mu^l, \dots, \theta^r * \mu^l)^T \in (W_0^1)^r$, $l = 1, \dots, t$. Therefore, (23), (25), (26), the first of the estimates in (29), and Lemma 3 imply that

$$\begin{aligned}\|(g*\overrightarrow{\mu})(X)\|_{(\ell^p(J))^t} &= \left\|\left(\left(\sum_{k\in\mathbb{Z}^d} C_k^T\Theta_k\right) * \overrightarrow{\mu}\right)(X)\right\|_{(\ell^p(J))^t} \\ &\le \mathcal{N}_p\left\|\left(\left(\sum_{k\in\mathbb{Z}^d} C_k^T\Theta_k\right) * \overrightarrow{\mu}\right)\right\|_{(W^p)^t} \\ &= \mathcal{N}_p\left\|\sum_{k\in\mathbb{Z}^d} C_k^T(\Theta*\overrightarrow{\mu})_k\right\|_{(W^p)^t} \\ &\le \mathcal{N}_p\|C\|_{(\ell^p(\mathbb{Z}^d))^r}\,\|\Theta*\overrightarrow{\mu}\|_{(W^1)^{r\times t}} \\ &\le 2^d\mathcal{N}_p\|\overrightarrow{\mu}\|_{\mathcal{M}^t}\|C\|_{(\ell^p(\mathbb{Z}^d))^r}\|\Theta\|_{(W^1)^r} \\ &\le \frac{2^d\mathcal{N}_p\|\overrightarrow{\mu}\|_{\mathcal{M}^t}}{m_p'}(\|\Phi\|_{(W^1)^r}+\epsilon)\|g\|_{L^p} \\ &\le \frac{2^d\mathcal{N}_p\|\overrightarrow{\mu}\|_{\mathcal{M}^t}(\|\Phi\|_{(W^1)^r}+\epsilon)}{m_p-\epsilon}\|g\|_{L^p}.\end{aligned}$$

Hence,

$$\|(g*\overrightarrow{\mu})(X)\|_{(\ell^p(J))^t} \le \left(\frac{2^d\mathcal{N}_p\|\overrightarrow{\mu}\|_{\mathcal{M}^t}(\|\Phi\|_{(W^1)^r}+\epsilon)}{m_p-\epsilon}\right)\|g\|_{L^p}. \tag{32}$$

Using the estimates (29) and the left hand side of the inequality (31), we can obtain an explicit upper bound ϵ_0 for ϵ from

$$\frac{A_p m_p}{\|\Phi\|_{(W^1)^r}+\epsilon} - \frac{2^d\mathcal{N}_p\|\overrightarrow{\mu}\|_{\mathcal{M}^t}}{m_p-\epsilon}\epsilon = 0.$$

This is equivalent to the quadratic equation

$$\epsilon^2 + C_p\epsilon - \frac{A_p m_p^2}{2^d \mathcal{N}_p \|\overrightarrow{\mu}\|_{\mathcal{M}^t}} = 0,$$

where

$$C_p = \|\Phi\|_{(W^1)^r} + \frac{A_p m_p}{2^d \mathcal{N}_p \|\overrightarrow{\mu}\|_{\mathcal{M}^t}}.$$

Let ϵ_0 be the positive solution of the previous equation, i.e.,

$$\epsilon_0 = \frac{1}{2}\left(\sqrt{C_p^2 + \frac{4A_p m_p^2}{2^d \mathcal{N}_p \|\overrightarrow{\mu}\|_{\mathcal{M}^t}}} - C_p\right).$$

Then, for $0 < \epsilon < \epsilon_0 < m_p$, we use (31), (32), and (29) to obtain

$$A_p^{'} = \frac{A_p m_p}{\|\Phi\|_{(W^1)^r} + \epsilon} - \frac{2^d \mathcal{N}_p \|\overrightarrow{\mu}\|_{\mathcal{M}^t}}{m_p - \epsilon}\epsilon,$$

$$B_p^{'} = \frac{2^d \mathcal{N}_p \|\overrightarrow{\mu}\|_{\mathcal{M}^t}(\|\Phi\|_{(W^1)^r} + \epsilon)}{m_p - \epsilon},$$

and the proof is complete. □

The following lemma deals with perturbations of the sampling measure.

Lemma 7. *Let $(X, \Phi, \overrightarrow{\mu})$ be a p-stable sampling model for some $p \in [1, \infty]$. Then there exists $\epsilon_0 > 0$ such that the sampling model $(X, \Phi, \overrightarrow{\alpha})$ is also p-stable, whenever $\overrightarrow{\alpha} \in \mathcal{M}^t(\mathbb{R}^d)$ and*

$$\|\overrightarrow{\mu} - \overrightarrow{\alpha}\|_{\mathcal{M}^t} < \epsilon_0.$$

Proof. Let $f = \sum_{k\in\mathbb{Z}^d} C_k^T \Phi_k \in V^p(\Phi)$, $C \in (\ell^p(\mathbb{Z}^d))^r$. We have

$$\begin{aligned}
A_p\|f\|_{L^p} &\le \|(f * \overrightarrow{\mu})(X)\|_{(\ell^p(J))^t} \\
&\le \|(f * (\overrightarrow{\mu} - \overrightarrow{\alpha}))(X)\|_{(\ell^p(J))^t} + \|(f * \overrightarrow{\alpha})(X)\|_{(\ell^p(J))^t} \\
&= \sum_{l=1}^{t} \|(f * (\mu^l - \alpha^l))(X)\|_{\ell^p(J)} + \|(f * \overrightarrow{\alpha})(X)\|_{(\ell^p(J)^t} \\
&= \sum_{l=1}^{t} \|((\sum_{k\in\mathbb{Z}^d} C_k^T \Phi_k) * (\mu^l - \alpha^l))(X)\|_{\ell^p(J)} + \|(f * \overrightarrow{\alpha})(X)\|_{(\ell^p(J))^t}.
\end{aligned}$$

Since $\overrightarrow{\mu}$ and $\overrightarrow{\alpha}$ are in $\mathcal{M}^t(\mathbb{R}^d)$, and $\Phi \in (W_0^1)^r$, then Proposition 3 implies $\Omega^l = (\phi^1 * (\mu^l - \alpha^l), \dots, \phi^r * (\mu^l - \alpha^l))^T \in (W_0^1)^r$, for $l = 1, \dots, t$. Using Proposition 3 once again we have:

$$\begin{aligned}
A_p\|f\|_{L^p} &\le \sum_{l=1}^{t} \mathcal{N}_p \|C\|_{(\ell^p(\mathbb{Z}^d))^r} \|\Omega^l\|_{(W^1)^r} + \|(f * \overrightarrow{\alpha})(X)\|_{(\ell^p(J))^t} \\
&\le 2^d \mathcal{N}_p \|C\|_{(\ell^p(\mathbb{Z}^d))^r} \|\Phi\|_{(W^1)^r} \|\overrightarrow{\mu} - \overrightarrow{\alpha}\|_{\mathcal{M}^t} + \|(f * \overrightarrow{\alpha})(X)\|_{(\ell^p(J))^t}.
\end{aligned}$$

Taking into account $\Phi \in (W^1)^r$, and f satisfies (6), then it follows

$$2^d \mathcal{N}_p \|C\|_{(\ell^p(\mathbb{Z}^d))^r} \|\Phi\|_{(W^1)^r} \|\overrightarrow{\mu} - \overrightarrow{\alpha}\|_{\mathcal{M}^t} \leq \frac{2^d \mathcal{N}_p \|f\|_{L^p} \|\Phi\|_{(W^1)^r} \|\overrightarrow{\mu} - \overrightarrow{\alpha}\|_{\mathcal{M}^t}}{m_p}.$$

Hence,

$$\left(A_p - \frac{2^d \mathcal{N}_p \|\Phi\|_{(W^1)^r} \|\overrightarrow{\mu} - \overrightarrow{\alpha}\|_{\mathcal{M}^t}}{m_p}\right) \|f\|_{L^p} \leq \|(f * \overrightarrow{\alpha})(X)\|_{(\ell^p(J))^t}. \tag{33}$$

On the other hand, using (9) again, we have

$$\begin{aligned}
\|(f * \overrightarrow{\alpha})(X)\|_{(\ell^p(J))^t} &\leq \|(f * (\overrightarrow{\alpha} - \overrightarrow{\mu}))(X)\|_{(\ell^p(J))^t} + \|(f * \overrightarrow{\mu})(X)\|_{(\ell^p(J))^t} \\
&= \sum_{l=1}^{t} \left\| \left(f * \left(\alpha^l - \mu^l\right)\right)(X) \right\|_{\ell^p(J)} + \|(f * \overrightarrow{\mu})(X)\|_{(\ell^p(J))^t} \\
&\leq \sum_{l=1}^{t} \left\| \left(\left(\sum_{k \in \mathbb{Z}^d} C_k^T \Phi_k \right) * \left(\alpha^l - \mu^l\right) \right)(X) \right\|_{\ell^p(J)} + B_p \|f\|_{L^p} \\
&\leq 2^d \mathcal{N}_p \|C\|_{(\ell^p(\mathbb{Z}^d))^r} \|\Phi\|_{(W^1)^r} \|\overrightarrow{\mu} - \overrightarrow{\alpha}\|_{\mathcal{M}^t} + B_p \|f\|_{L^p}.
\end{aligned}$$

Using condition (6), we obtain:

$$\|(f * \overrightarrow{\alpha})(X)\|_{(\ell^p(J))^t} \leq \left(\frac{2^d \mathcal{N}_p \|\Phi\|_{(W^1)^r} \|\overrightarrow{\mu} - \overrightarrow{\alpha}\|_{\mathcal{M}^t}}{m_p} + B_p\right) \|f\|_{L^p}. \tag{34}$$

From (33) and (34), by choosing

$$\epsilon_0 = \frac{A_p m_p}{2^d \mathcal{N}_p \|\Phi\|_{(W^1)^r}},$$

we obtain for $0 < \epsilon < \epsilon_0$,

$$A_p^{'} = A_p - \frac{2^d \mathcal{N}_p \|\Phi\|_{(W^1)^r}}{m_p} \epsilon, \quad \text{and}$$

$$B_p^{'} = B_p + \frac{2^d \mathcal{N}_p \|\Phi\|_{(W^1)^r}}{m_p} \epsilon.$$

□

From the two lemmas above, we obtain the following corollary:

Corollary 5. *Let $(X, \Phi, \overrightarrow{\mu})$ be a p-stable sampling model for some $p \in [1, \infty]$. Then there exists $\epsilon_0 > 0$ such that the sampling model $(X, \Theta, \overrightarrow{\alpha})$ is also p-stable, whenever $\overrightarrow{\alpha} \in \mathcal{M}^t(\mathbb{R}^d)$, $\Theta \in (W_0^1)^r$, and $\|\Phi - \Theta\|_{(W^1)^r} + \|\overrightarrow{\mu} - \overrightarrow{\alpha}\|_{\mathcal{M}^t} < \epsilon_0$.*

The above result is essentially obvious at this point. We proceed with a formal proof in order to obtain estimates for ϵ_0 and the bounds $A_p^{'}$ and $B_p^{'}$ of X as an $\overrightarrow{\alpha}$-sampling set for $V^p(\Theta)$.

Proof. Let $0 < \epsilon_1 < \frac{1}{2}\left(\sqrt{C_p^2 + \frac{4A_p m_p^2}{2^d \mathcal{N}_p \|\overrightarrow{\mu}\|_{\mathcal{M}^t}}} - C_p\right)$, where

$$C_p = \|\Phi\|_{(W^1)^r} + \frac{A_p m_p}{2^d \mathcal{N}_p \|\overrightarrow{\mu}\|_{\mathcal{M}^t}}.$$

Then, by Lemma 6, X is a $\overrightarrow{\mu}$-sampling set for $V^p(\Theta)$ as soon as

$$\|\Phi - \Theta\|_{(W^1)^r} \leq \epsilon_1.$$

Moreover,

$$A_p^{''}\|g\|_{L^p} \leq \|(g * \overrightarrow{\mu})(X)\|_{(\ell^p(J))^t} \leq B_p^{''}\|g\|_{L^p}, \text{ for all } g \in V^p(\Theta),$$

where

$$A_p^{''} = \frac{A_p m_p}{\|\Phi\|_{(W^1)^r} + \epsilon_1} - \frac{2^d \mathcal{N}_p \|\overrightarrow{\mu}\|_{\mathcal{M}^t}}{m_p - \epsilon_1}\epsilon_1$$

and

$$B_p^{''} = \frac{2^d \mathcal{N}_p \|\overrightarrow{\mu}\|_{\mathcal{M}^t}(\|\Phi\|_{(W^1)^r} + \epsilon_1)}{m_p - \epsilon_1}.$$

Assume now that

$$0 < \epsilon_2 \leq \frac{A_p^{''}(m_p - \epsilon_1)}{2^d \mathcal{N}_p(\|\Phi\|_{(W^1)^r} + \epsilon_1)}.$$

Then, by Lemma 7, X is an $\overrightarrow{\alpha}$-sampling set for $V^p(\Theta)$ as soon as

$$\|\Phi - \Theta\|_{(W^1)^r} \leq \epsilon_1 \text{ and } \|\overrightarrow{\mu} - \overrightarrow{\alpha}\|_{\mathcal{M}^t} \leq \epsilon_2.$$

Hence, if $0 < \epsilon < \epsilon_0 = \min\{\epsilon_1, \epsilon_2\}$, we obtain the sampling bounds

$$A_p^{'} = A_p^{''} - \frac{2^d \mathcal{N}_p(\|\Phi\|_{(W^1)^r} + \epsilon_1)}{m_p - \epsilon_1}\epsilon_2,$$

and

$$B_p^{'} = B_p^{''} + \frac{2^d \mathcal{N}_p(\|\Phi\|_{(W^1)^r} + \epsilon_1)}{m_p - \epsilon_1}\epsilon_2,$$

as soon as

$$\|\Phi - \Theta\|_{(W^1)^r} + \|\overrightarrow{\mu} - \overrightarrow{\alpha}\|_{\mathcal{M}^t} \leq \epsilon < \epsilon_0.$$

□

The following lemma is a result about jitter in the sampling set X:

Lemma 8. *Let $(X, \Phi, \overrightarrow{\mu})$ be a p-stable sampling model for some $p \in [1, \infty]$. Then there exists $\epsilon_0 > 0$ such that the sampling model $(X + \Delta, \Phi, \overrightarrow{\mu})$ is also p-stable, whenever $\|\Delta\|_\infty < \epsilon_0$.*

Its proof is immediately implied by Lemma 1 and the following result.

Lemma 9. *Let $(X, \Phi, \overrightarrow{\mu})$ be a p-stable sampling model for some $p \in [1, \infty]$ and $\widetilde{X} = X + \Delta$. Let U be the sampling operator for $(X, \Phi, \overrightarrow{\mu})$ and U_Δ be the sampling operator for $(\widetilde{X}, \Phi, \overrightarrow{\mu})$. Then $\|U - U_\Delta\| \to 0$ as $\|\Delta\|_\infty \to 0$.*

Proof. We recall that for any $\gamma > 0$, the function $\operatorname{osc}_\gamma g$ on $\mathbb{R}^d$ is defined by

$$\operatorname{osc}_\gamma g(x) = \sup_{|\Delta x| < \gamma} |g(x + \Delta x) - g(x)|.$$

From Lemma 8.1 in [10] it follows that if $g \in W_0^1$, then $\operatorname{osc}_\gamma g \in W^1$, and $\|\operatorname{osc}_\gamma g\|_{W^1} \to 0$ as $\gamma \to 0$. Therefore, by applying Proposition 3 we get

$$\operatorname{osc}_\gamma \Phi * \overrightarrow{\mu}^T \in (W^1)^{r \times t}, \text{ and } \|\operatorname{osc}_\gamma \Phi * \overrightarrow{\mu}^T\|_{(W^1)^{r\times t}} \to 0 \text{ as } \gamma \to 0,$$

where

$$\operatorname{osc}_\gamma \Phi * \overrightarrow{\mu}^T = \begin{pmatrix} \operatorname{osc}_\gamma \phi^1 * \mu^1 \ldots \operatorname{osc}_\gamma \phi^1 * \mu^t \\ \vdots \qquad\qquad \vdots \\ \operatorname{osc}_\gamma \phi^r * \mu^1 \ldots \operatorname{osc}_\gamma \phi^r * \mu^t \end{pmatrix}.$$

For any $m \in \mathbb{Z}^d$ there exist at most $\mathcal{S}$ sampling points in every hypercube $[0, 1]^d + m$. We set $X_m = X \bigcap ([0, 1]^d + m)$, $m \in \mathbb{Z}^d$, and, for each $1 \le i \le r$ and $1 \le l \le t$, define the sequence

$$b^{i,l}(m) := \operatorname{ess\,sup}_{x \in [0,1]^d} \{\operatorname{osc}_{\|\Delta\|_\infty} (\phi^i * \mu^l)(x + m)\}, \quad m \in \mathbb{Z}^d.$$

Then $\|b^{i,l}\|_{\ell^1(\mathbb{Z}^d)} = \|\operatorname{osc}_{\|\Delta\|_\infty} (\phi^i * \mu^l)\|_{W^1}$ and, hence,

$$\|b\|_{(\ell^1(\mathbb{Z}^d))^{(r\times t)}} = \|\operatorname{osc}_{\|\Delta\|_\infty} \Phi * \overrightarrow{\mu}^T\|_{(W^1)^{r\times t}}.$$

For $1 \le i \le r$ and $1 \le l \le t$ we have

$$\begin{aligned} \|(U^{l,i} - U^{l,i}_\Delta) c^i\|^p_{\ell^p(J)} &= \sum_{x_j \in X} \left| \sum_{k \in \mathbb{Z}^d} c^i_k \left((\phi^i * \mu^l)(x_j - k)) - (\phi^i * \mu^l)(x_j + \delta_j - k) \right) \right|^p \\ &\le \sum_{x_j \in X} \left(\sum_{k \in \mathbb{Z}^d} |c^i_k| \operatorname{osc}_{\|\Delta\|_\infty} (\phi^i * \mu^l)(x_j - k) \right)^p \\ &\le \sum_{m \in \mathbb{Z}^d} \mathcal{S} \left(\sum_{k \in \mathbb{Z}^d} |c^i_k| b^{i,l}(m - k) \right)^p \\ &= \mathcal{S} \left\| |c^i| * b^{i,l} \right\|^p_{\ell^p(\mathbb{Z}^d)}, \end{aligned}$$

where $\mathcal{S}$ is as in Definition 2. By using Young's inequality we obtain

$$\begin{aligned}\mathcal{S}\||c^i| * b^{i,l}\|^p_{\ell^p(\mathbb{Z}^d)} &\leq \mathcal{S}\|c^i\|^p_{\ell^p(\mathbb{Z}^d)}\|b^{i,l}\|^p_{l^1}\\ &= \mathcal{S}\|c^i\|^p_{\ell^p(\mathbb{Z}^d)}\left\|\mathrm{osc}_{\|\Delta\|_\infty}\ \phi^i * \mu^l\right\|^p_{W^1}.\end{aligned}$$

Consequently,

$$\|U^{l,i} - U^{l,i}_\Delta\| \leq \mathcal{N}_p \left\|\mathrm{osc}_{\|\Delta\|_\infty}\ \phi^i * \mu^l\right\|_{W^1}.$$

Hence,

$$\|U - U_\Delta\| \leq \mathcal{N}_p \left\|\mathrm{osc}_{\|\Delta\|_\infty}\ \Phi * \overrightarrow{\mu}^T\right\|_{(W^1)^{r\times t}} \to 0 \quad \text{as} \quad \|\Delta\|_\infty \to 0,$$

and the lemma is proved. □

Proof of Theorem 1.

Proof. The proof of Theorem 1 is hidden in the proofs of Lemmas 6, 7, and 8. In particular, keeping the notation of the proof of Lemma 6, we have

$$\|((f-g) * \overrightarrow{\mu})(X)\|_{(\ell^p(J))^t} \leq 2^d\mathcal{N}_p\|C\|_{(\ell^p(\mathbb{Z}^d))^r}\|\Phi - \Theta\|_{(W^1)^r}\|\overrightarrow{\mu}\|_{\mathcal{M}^t}.$$

Hence, $\left\|U_{(X,\Phi,\overrightarrow{\mu})} - U_{(X,\Theta,\overrightarrow{\mu})}\right\| \leq Const \cdot \|\Phi - \Theta\|_{(W^1)^r}$, where the constant is independent of Δ. Keeping the notation of the proof of Lemma 7, we have

$$\|(f * (\overrightarrow{\alpha} - \overrightarrow{\mu}))(X)\|_{(\ell^p(J))^t} \leq 2^d\mathcal{N}_p\|C\|_{(\ell^p(\mathbb{Z}^d))^r}\|\Phi\|_{(W^1)^r}\|\overrightarrow{\mu} - \overrightarrow{\alpha}\|_{\mathcal{M}^t}.$$

Hence, $\left\|U_{(X,\Phi,\overrightarrow{\mu})} - U_{(X,\Phi,\overrightarrow{\alpha})}\right\| \leq Const \cdot \|\overrightarrow{\mu} - \overrightarrow{\alpha}\|_{\mathcal{M}^t}$, where the constant is again independent of Δ. Finally, we can combine these estimates with Lemma 9 via the standard $\epsilon/3$ argument and complete the proof. □

II.3.3 Proofs for Section II.2.2

We begin with an auxiliary technical result for the convolution of functions with measures.

Lemma 10. *Let $\Phi = (\phi^1, \ldots, \phi^r)^T$ be a vector of continuous functions, $s > d$, and $\overrightarrow{\mu} \in \mathcal{M}^t_s(\mathbb{R}^d)$. If $|\phi^i(x)| \leq C^i_0(1+|x|)^{-s}$ for all $1 \leq i \leq r$, then*

$$\|(\Phi * \overrightarrow{\mu}^T)(x)\| \leq C_1(1+|x|)^{-s};$$

the constants $C^i_0 > 0$, $1 \leq i \leq r$, and $C_1 > 0$ are independent of $x \in \mathbb{R}^d$.

Proof. For $1 \leq i \leq r$ and $1 \leq j \leq t$ we have

$$\begin{aligned}|(\phi^i * \mu^j)(x)| &\leq \int_{\mathbb{R}^d} |\phi^i(x-y)| d|\mu^j|(y)\\ &\leq C^i_0 \int_{\mathbb{R}^d} (1+|x-y|)^{-s} d|\mu^j|(y).\end{aligned}$$

Since $(1+|u+w|)^{-l} \le (1+|u|)^l(1+|w|)^{-l}$, for all $u, w \in \mathbb{R}^d$, and $l \ge 0$, we have

$$\begin{aligned} |(\phi^i * \mu^j)(x)| &\le C_0^i \int_{\mathbb{R}^d} (1+|y|)^s (1+|x|)^{-s} d|\mu^j|(y) \\ &= C_0^i (1+|x|)^{-s} \int_{\mathbb{R}^d} (1+|y|)^s d|\mu^j|(y) \\ &\le C_1^{i,j}(1+|x|)^{-s}, \end{aligned}$$

where the last inequality follows from $\mu^j \in \mathcal{M}_s(\mathbb{R}^d)$. Therefore,

$$\|(\Phi * \overrightarrow{\mu}^T)(x)\| \le C_1(1+|x|)^{-s},$$

where $C_1 = \sum_{i=1}^r \sum_{j=1}^t C_1^{i,j}$. □

Remark 12. *If $\{\Phi_k\}_{k\in\mathbb{Z}^d}$ is an s-localized Riesz generator for $V^2(\Phi)$, as in Definition 9, then, by Lemma 14(a) in [59], we have that $\{\widetilde{\Phi}_k\}_{k\in\mathbb{Z}^d}$ is also an s-localized Riesz generator for $V^2(\Phi)$. Consequently, by Lemma 10 we have*

$$\|(\widetilde{\Phi} * \overrightarrow{\mu}^T)(x)\| \le D_1(1+|x|)^{-s}, \tag{35}$$

for some $D_1 > 0$ independent of $x \in \mathbb{R}^d$.

Proof of Proposition 2.

Proof. Let X be a $\overrightarrow{\mu}$-sampling set for $V^2(\Phi)$, $\overrightarrow{\mu} \in \mathcal{M}^t(\mathbb{R}^d)$. Then, by definition, there exist constants $0 < A_2 \le B_2 < \infty$ such that

$$A_2\|f\|_{L^2} \le \|(f * \overrightarrow{\mu})(X)\|_{(\ell^2(J))^t} \le B_2\|f\|_{L^2}, \text{for all } f \in V^2(\Phi). \tag{36}$$

Fix $x_j \in X$. Then, for each $1 \le i \le t$, the function $g^i_{x_j}\colon V^2(\Phi) \to \mathbb{C}$ given by $g^i_{x_j}(f) = (f * \mu^i)(x_j)$ is a bounded linear functional on the closed subspace $V^2(\Phi)$ of $L^2(\mathbb{R}^d)$ because $|g^i_{x_j}(f)| \le B_2\|f\|_{L^2}$ for all $f \in V^2(\Phi)$. Consequently, by Riesz representation theorem, there exists $\psi^i_{x_j} \in V^2(\Phi)$ such that $g^i_{x_j}(f) = \langle f, \psi^i_{x_j}\rangle$ for all $f \in V^2(\Phi)$. It follows immediately from (36) and Definition 5 that $\Psi_{x_j} = (\psi^1_{x_j}, \dots, \psi^t_{x_j})^T$ is a frame for $V^2(\Phi)$. Hence, every $f \in V^2(\Phi)$ can be recovered via $f = \sum_{j\in J}\langle f, \Psi_{x_j}\rangle^T \widetilde{\Psi}_{x_j}$, where $\{\widetilde{\Psi}_{x_j} = (\widetilde{\psi}^1_{x_j}, \dots, \widetilde{\psi}^t_{x_j})^T\}_{j\in J}$ is the dual frame of $\{\Psi_{x_j}\}_{j\in J}$ and the series converges unconditionally in $V^2(\Phi)$. Since $\langle f, \Psi_{x_j}\rangle = (f * \overrightarrow{\mu})(x_j)$ for all $j \in J$, we get (14). □

Next, we show that if the generator Φ and the measures $\overrightarrow{\mu}$ satisfy an appropriate decay condition then the $(\overrightarrow{\mu}, X)$-sampling frame $\{\Psi_{x_j}\}$ obtained above is s-localized.

Proposition 4. *Let $s > d$, $\Phi \in \mathcal{W}_s$, and $\overrightarrow{\mu} \in \mathcal{M}^t_s(\mathbb{R}^d)$. If X is a $\overrightarrow{\mu}$-sampling set for $V^2(\Phi)$, then the $(\overrightarrow{\mu}, X)$-sampling frame $\{\Psi_{x_j}\}$ is s-localized with respect to the Riesz basis $\{\Phi_k\}_{k\in\mathbb{Z}^d}$.*

Proof. Since $\{\Phi_k\}_{k\in\mathbb{Z}^d}$ is an s-localized Riesz generator for $V^2(\Phi)$, the components of Φ satisfy (17), and Lemma 10 implies

$$|\langle \Phi_k, \Psi^T_{x_j}\rangle| = |(\Phi * \overrightarrow{\mu}^T)(x_j - k)| \le C_1(1+|x_j - k|)^{-s},$$

for some $C_1 > 0$ independent of $j \in J$ and $k \in \mathbb{Z}^d$. On the other hand, it follows from Remark 12 that the dual Riesz basis $\{\widetilde{\Phi}_k\}_{k\in\mathbb{Z}^d}$ is also an s-localized Riesz generator for $V^2(\Phi)$, and its components also satisfy (17). Therefore, using Lemma 10 once again, we get

$$|\langle \widetilde{\Phi}_k, \Psi_{x_j}^T\rangle| = |(\widetilde{\Phi} * \overrightarrow{\mu}^T)(x_j - k)| \le D_1(1 + |x_j - k|)^{-s},$$

for some $D_1 > 0$ independent of $j \in J$ and $k \in \mathbb{Z}^d$. Hence, $\{\Psi_{x_j}\}$ satisfies all conditions of Definition 8. □

We conclude this subsection with the proof of the main result of Section II.2.2.

Proof of Theorem 2

Proof. Assume the hypotheses of Theorem 2. By Propositions 2 and 4, there exists a $(\overrightarrow{\mu}, X)$-sampling frame $\{\Psi_{x_j}\}_{j\in J}$ for $V^2(\Phi)$, which is s-localized with respect to the Riesz basis $\{\Phi_k\}_{k\in\mathbb{Z}^d}$ and satisfies

$$\langle f, \Psi_{x_j}\rangle = (f * \overrightarrow{\mu})(x_j), \text{ for all } f \in V^2(\Phi).$$

Moreover,

$$f = \sum_{j\in J}(f * \overrightarrow{\mu})^T(x_j)\widetilde{\Psi}_{x_j}, \text{ for all } f \in V^2(\Phi).$$

Consequently, applying Theorem 10(c) in [59], we get

$$f = \sum_{j\in J}(f * \overrightarrow{\mu})^T(x_j)\widetilde{\Psi}_{x_j}, \text{ for all } f \in V^p(\Phi),$$

where the series converges unconditionally in $V^p(\Phi)$, $1 \le p < \infty$. Moreover, since $\{\Psi_{x_j}\}_{j\in J}$ is an s-localized frame with respect to the Riesz basis $\{\Phi_k\}_{k\in\mathbb{Z}^d}$, then Theorem 10(d) in [59] implies that for each $1 \le p \le \infty$ there exist $0 < A_p \le B_p < \infty$ such that

$$A_p\|f\|_{L^p} \le \|(f * \overrightarrow{\mu})(X)\|_{(\ell^p(J))^t} \le B_p\|f\|_{L^p}, \text{ for all } f \in V^p(\Phi),$$

i.e., X is a $\overrightarrow{\mu}$-sampling set for $V^p(\Phi)$ and the theorem is proved. □

II.3.4 Proofs for Section II.2.3

For the proof of Theorem 4 we need the following two lemmas. To simplify notation, in this section all unspecified norms are the operator norm $\|\cdot\|_{p,op}$.

Lemma 11. *Let the assumptions of Theorem 4 hold. Then*

$$\|U^*U - U_\Delta^*U_\Delta\| < \epsilon\,(\epsilon + \beta_p + \beta_q)\,,$$

for $\|\Delta\|_\infty$ sufficiently small.

Proof. Let $\epsilon > 0$ be given. Because $\|U^* - U_\Delta^*\|_{p,op} = \|U - U_\Delta\|_{q,op}$, $\frac{1}{p} + \frac{1}{q} = 1$, we have that $\|U^* - U_\Delta^*\|_{p,op} \to 0$ as $\|\Delta\|_\infty \to 0$ by Lemma 9. Since $\|U - U_\Delta\| \to 0$ and $\|U^* - U_\Delta^*\| \to 0$ as $\|\Delta\|_\infty \to 0$, then $\|U - U_\Delta\| < \epsilon$ and $\|U^* - U_\Delta^*\| < \epsilon$ for $\|\Delta\|_\infty$ sufficiently small. Therefore,

$$\begin{aligned}
\|U^*U - U_\Delta^* U_\Delta\| &= \|U^*U - U^*U_\Delta + U^*U_\Delta - U_\Delta^* U_\Delta\| \\
&= \|U^*(U - U_\Delta) + (U^* - U_\Delta^*)U_\Delta\| \\
&\le \|U^*\|\|U - U_\Delta\| + \|U^* - U_\Delta^*\|\|U_\Delta\| \\
&\le \|U^*\|\|U - U_\Delta\| + \|U^* - U_\Delta^*\|(\|U - U_\Delta\| + \|U\|) \\
&< \beta_q \epsilon + \epsilon(\epsilon + \beta_p) \\
&= \epsilon\left(\epsilon + \beta_p + \beta_q\right),
\end{aligned}$$

and the lemma is proved. $\square$

Lemma 12. *Let the assumptions of Theorem 4 hold. Then $0 < \nu < 1$, $(U_\Delta^* U_\Delta)^{-1}$ exists, and $\|(U^*U)^{-1} - (U_\Delta^* U_\Delta)^{-1}\| < \frac{\nu n_p}{1-\nu}$, for $\|\Delta\|_\infty$ sufficiently small.*

Proof. Since $(U^*U)^{-1}$ exists,

$$U_\Delta^* U_\Delta = U^*U\left(I + (U^*U)^{-1}\left(U_\Delta^* U_\Delta - U^*U\right)\right). \tag{37}$$

Taking into account that $\|(U^*U)^{-1}\| \le n_p$ and Lemma 11 we have

$$\begin{aligned}
\|(U^*U)^{-1}\left(U_\Delta^* U_\Delta - U^*U\right)\| &\le \|(U^*U)^{-1}\|\|U_\Delta^* U_\Delta - U^*U\| \\
&\le n_p \epsilon\left(\epsilon + \beta_p + \beta_q\right) \\
&< n_p \left(\frac{-(\beta_p+\beta_q) + \sqrt{(\beta_p+\beta_q)^2 + \frac{4}{n_p}}}{2}\right)\left(\frac{(\beta_p+\beta_q) + \sqrt{(\beta_p+\beta_q)^2 + \frac{4}{n_p}}}{2}\right) = 1.
\end{aligned}$$

Hence, $\nu = n_p \epsilon\left(\epsilon + \beta_p + \beta_q\right) \in (0, 1)$. To simplify the notation, we define

$$M := U^*U, \quad M_\Delta := U_\Delta^* U_\Delta, \text{ and } N := (U^*U)^{-1}\left(U_\Delta^* U_\Delta - U^*U\right).$$

Since $\|N\| \le \nu < 1$, then $(I + N)^{-1}$ exists and is given by the Neumann series

$$(I + N)^{-1} = \sum_{q=0}^{\infty} (-1)^q N^q.$$

From (37) we obtain

$$M_\Delta^{-1} = [M(I + N)]^{-1} = (I + N)^{-1} M^{-1}. \tag{38}$$

Therefore, $M_\Delta^{-1} = (U_\Delta^* U_\Delta)^{-1}$ exists.

Now we need to give an upper bound for $\|M^{-1} - M_\Delta^{-1}\|$. Using (38) we obtain

$$M^{-1} - M_\Delta^{-1} = N(I + N)^{-1} M^{-1}.$$

Consequently,

$$\begin{aligned}\|M^{-1} - M_\Delta^{-1}\| &\leq \|N\|\|(I+N)^{-1}\|\|M^{-1}\| \\ &\leq \frac{\|N\|}{1-\|N\|}\|M^{-1}\| \leq \frac{\nu n_p}{1-\nu},\end{aligned} \tag{39}$$

and the lemma is proved. □

Proof of Theorem 4.

Proof. Using the notation from Lemmas 11, 12, and the previous proofs, we get

$$\begin{aligned}\|(U^*U)^{-1}U^* - (U_\Delta^*U_\Delta)^{-1}U_\Delta^*\| &= \|M^{-1}U^* - M_\Delta^{-1}U_\Delta^*\| \\ &\leq \|M^{-1}U^* - M^{-1}U_\Delta^*\| + \|M^{-1}U_\Delta^* - M_\Delta^{-1}U_\Delta^*\| \\ &\leq \|M^{-1}\|\|U^* - U_\Delta^*\| + \|M^{-1} - M_\Delta^{-1}\|\|U_\Delta^*\| \\ &\leq n_p\epsilon + \frac{\nu n_p}{1-\nu}\left(\|U^* - U_\Delta^*\| + \|U^*\|\right) \\ &\leq n_p\epsilon + \frac{\nu n_p}{1-\nu}\left(\epsilon + \beta_q\right) \\ &= n_p\left(\epsilon + \frac{\nu(\epsilon+\beta_q)}{1-\nu}\right),\end{aligned}$$

for $\|\Delta\|_\infty$ sufficiently small. □

Proof of Theorem 5.

Proof. Let U_Δ be the sampling operator for a perturbed sampling model $(X + \Delta, \Theta, \overrightarrow{\alpha})$. Let also $C \in (\ell^p(\mathbb{Z}^d))^r$, $f = \sum_{k\in\mathbb{Z}^d} C_k^T \Phi_k$, and $g = \sum_{k\in\mathbb{Z}^d} C_k^T \Theta_k$. We have that $U_\Delta C = (g * \overrightarrow{\alpha})(X)$, $UC = (f * \overrightarrow{\mu})(X)$, and $RD = \sum_{k\in\mathbb{Z}^d}[(U^*U)^{-1}U^*D]_k^T \Phi(\cdot - k)$. Thus, $R(UC) = f$, and using (6) we get

$$\|R((g * \overrightarrow{\alpha})(X+\Delta)) - f\|_{L^p} \leq M_p \left\|((U^*U)^{-1}U^*U_\Delta C - C)\right\|_{\ell^p}.$$

Using the fact that $C = (U^*U)^{-1}U^*UC$ and (6) we get

$$\|R((g * \overrightarrow{\alpha})(X+\Delta)) - f\|_{L^p} \leq \frac{M_p}{m_p}\left\|(U^*U)^{-1}U^*\right\|\|U_\Delta - U\|\|f\|_{L^p},$$

where as before U_Δ is associated with the sampling model $(X + \Delta, \Theta, \alpha)$, and U is associated with (X, Φ, μ). The result is now immediate. □

Proof of Theorem 6.

Proof. Assume the hypotheses of Theorem 6. From Theorem 2 we know that, in this case, the sampling model $(X, \Phi, \overrightarrow{\mu})$ is p-stable for every $p \in [1, \infty]$. Hence, in view of Theorem 5, the only thing that we need to prove is that the operator U^*U is invertible for all $p \in [1, \infty]$ and not just for $p = 2$.

Taking into account that for each $1 \leq i \leq r$ and $1 \leq l \leq t$ the entries of the matrix of the operator $U^{l,i}$ satisfy

$$|(U^{l,i})_{j,k}| = |(\phi^i * \mu^l)(x_j - k)| \leq C_1(1 + |x_j - k|)^{-s},$$

for some $C_1 > 0$ independent of $j \in J$ and $k \in \mathbb{Z}^d$, it follows from Lemma 3 in [59] that the matrix of U defines a bounded linear operator from $(\ell^p(\mathbb{Z}^d))^r \to (\ell^p(J))^t$ for all $1 \leq p \leq \infty$. Hence, U^* is also well defined as a bounded linear operator from $(\ell^p(J))^t \to (\ell^p(\mathbb{Z}^d))^r$, and, therefore, $U^*U : (\ell^p(\mathbb{Z}^d))^r \to (\ell^p(\mathbb{Z}^d))^r$ is a well defined and bounded operator for all $1 \leq p \leq \infty$. On the other hand, since the operator U^*U is invertible on $(\ell^2(\mathbb{Z}^d))^r$ and its components $(M^{i,l})_{j,k}$, $1 \leq i \leq r$, $1 \leq l \leq r$, satisfy a decay condition

$$|(M^{i,l})_{j,k}| \leq C_2(1 + |x_j - k|)^{-s},$$

for some $C_2 > 0$ independent of $j \in J$ and $k \in \mathbb{Z}^d$, then Jaffard's Lemma (see Theorem 5 in [59]) implies that $(U^*U)^{-1} : (\ell^2(\mathbb{Z}^d))^r \to (\ell^2(\mathbb{Z}^d))^r$ is also a bounded linear operator defined by a matrix satisfying the same off-diagonal decay condition as U^*U. Consequently, using Lemma 3 in [59] once again, we get that the matrix of $(U^*U)^{-1}$ defines a bounded linear operator on $(\ell^p(\mathbb{Z}^d))^r$ for all $1 \leq p \leq \infty$. The theorem is proved. $\square$

CHAPTER III

ITERATIVE RECONSTRUCTION AND STABILITY BOUNDS FOR SAMPLING MODELS

In this chapter we study the reconstruction of a signal f belonging to a shift-invariant space $V^p(\Phi)$ for some $1 \le p \le \infty$ (see (4) on page 9) from the set of non-uniformly distributed local sampled values. We show that if the set of sampling $X = \{x_j\}_{j\in J} \subset \mathbb{R}^d$ satisfies a necessary density condition, then f can be recovered from its samples geometrically fast using an iterative algorithm. In addition, the algorithm is analyzed when the samples of the signal are perturbed by noise, and it is shown the reconstruction error is continuously controlled by the perturbation of the samples of the signal. Furthermore, if we assume that X is a separated set, then it is shown that X is a set of sampling and explicit stability bounds are given.

Let J be a countable index set and $X = \{x_j : j \in J\}$ be a subset of $\mathbb{R}^d$. The sampling-reconstruction problem in this chapter consists of recovering a function $f \in V^p(\Phi)$ from the knowledge of its samples

$$\left\{ g_{x_j}(f) = \int_{\mathbb{R}^d} f(x) d\mu_{x_j}(x) \right\}_{j\in J},$$

where $\boldsymbol{\mu}= \{\mu_{x_j}\}_{j\in J}$ is a countable collection of finite complex Borel measures on $\mathbb{R}^d$ satisfying the following properties:

1. There exists $a_1 > 0$ such that $\operatorname{supp} \mu_{x_j} \subset x_j + [-a_1, a_1]^d$, for all $j \in J$,
2. There exists $M > 0$ such that $\|\mu_{x_j}\| \le M$, for all $x_j \in X$; and
3. $\int_{\mathbb{R}^d} d\mu_{x_j} = 1$, for all $j \in J$.

The following definition is a slightly different version of Definition 1 adapted to the model we are working with in this chapter.

Definition 10. *Let $1 \le p \le \infty$ and $X = \{x_j : j \in J\}$ be a countable subset of $\mathbb{R}^d$. We say that X is a* set of sampling *for $V^p(\Phi)$ and $\boldsymbol{\mu}= \{\mu_{x_j}\}_{j\in J}$ if there exist constants $0 < A_p \le B_p < \infty$ such that*

$$A_p \|f\|_{L^p} \le \|\{g_{x_j}(f)\}\|_{\ell^p(J)} \le B_p \|f\|_{L^p}, \text{ for all } f \in V^p(\Phi). \tag{40}$$

A_p and B_p are called the stability bounds.

Remark 13. *If in the above definition we let $p = 2$, then applying the Riesz representation theorem, it follows that (40) is the definition of Frame. Thus, f can be reconstructed from its samples via dual frame expansion.*

Definition 11. *A set $X = \{x_j : j \in J\} \subset \mathbb{R}^d$ is $\gamma-$*dense *in $\mathbb{R}^d$ if*

$$\mathbb{R}^d = \bigcup_j B_r(x_j), \quad \forall r \ge \gamma,$$

where $B_r(x_j) = \prod_{l=1}^d [x_j^l - r, x_j^l + r)$.

Definition 12. *A set* $X = \{x_j : j \in J\} \subset \mathbb{R}^d$ *is called a* separated set *if there exists* $\delta > 0$ *such that* $\inf_{j \neq i} |x_j - x_i| \geq \delta$. *The positive constant* δ *is called a* separation constant.

Definition 13. *A bounded partition of unity adapted to* $\{B_\gamma(x_j)\}_{j\in J}$ *and associated with the sampling set* X *is a set of functions* $\{\beta_j\}_{j\in J}$ *that satisfies:*

1. $0 \leq \beta_j \leq 1$, $\forall j \in J$;
2. $\operatorname{supp} \beta_j \subset B_\gamma(x_j)$; *and*
3. $\sum_{j\in J} \beta_j = 1$.

Now we shall define two operators that will be used in the next section for reconstructing signals from their samples using an iterative algorithm.

Given a bounded partition of unity $\{\beta_j\}_{j\in J}$ associated with the sampling set X, we define the operator $\mathrm{A_X}$ on $V^p(\Phi)$ as follows

$$\mathrm{A_X}\, f = \sum_{j\in J} g_{x_j}(f)\beta_j. \tag{41}$$

The quasi-interpolant operator $\mathrm{Q_X}$ is defined on sequences $c = \{c_j\}_{j\in J} \in \ell^p(J)$ by

$$\mathrm{Q_X}\, c = \sum_{j\in J} c_j \beta_j. \tag{42}$$

If $f \in W_0^p$, we write

$$\mathrm{Q_X}\, f = \sum_{j\in J} f(x_j)\beta_j \tag{43}$$

for the quasi-interpolant of the sequence $c_j = f(x_j)$.

Remark 14. *Note that if* $\mu_{x_j} = \delta_{x_j}$, *for all* $j \in J$, *where* δ_{x_j} *is the Dirac measure on* $\mathbb{R}^d$ *concentrated at* x_j, *then* $\mathrm{A_X} = \mathrm{Q_X}$.

III.1 Main Results

In this section we collect the main results of this chapter.

Theorem 7. *Let* $\Phi \in (W_0^1)^{(r)}$, $1 \leq p \leq \infty$, *and* P *be a bounded projection from* $L^p(\mathbb{R}^d)$ *onto* $V^p(\Phi)$. *Then there exists a density* $\gamma_0 = \gamma_0(\Phi, \mathrm{P}, p) > 0$, *and* $a_0 = a_0(\Phi, \mathrm{P}, p) > 0$ *such that every* $f \in V^p(\Phi)$ *can be recovered from the data* $\{g_{x_j}(f)\}_{j\in J}$ *on any* $\gamma-$*dense set* $X = \{x_j\}_{j\in J}$ $(0 < \gamma \leq \gamma_0)$ *for any support size condition (for* μ*)* $0 < a \leq a_0$ *by the following iterative algorithm:*

$$f_1 = \mathrm{P\,A_X}\, f, \qquad f_{n+1} = \mathrm{P\,A_X}\,(f_n - f) + f_n. \tag{44}$$

In this case the sequence $\{f_n\}_{n\geq 1}$ converges to f in the W^p norm, hence both in the $L^p(\mathbb{R}^d)$, and uniformly. The convergence is geometric, that is,

$$\|f_n - f\|_{L^p(\mathbb{R}^d)} \leq \|f_n - f\|_{W^p} \leq c_p \alpha^n \|f\|_{W^p},$$

for some $\alpha = \alpha(\mathrm{P}, \gamma, a, \Phi, p) < 1$, and for some $0 < c_p < \infty$ independent of f and $n \in \mathbb{N}$.

Remark 15. *Notice that since $\Phi \in (W_0^1)^{(r)}$, then by Theorem 6.2 in [10] the existence of a bounded projection P is guaranteed for all $p \in [1, \infty]$, and in this case it is given by $\mathrm{P}\, f = \sum_{k\in\mathbb{Z}^d} \langle f, \widetilde{\Phi}(\cdot - k)\rangle \Phi(\cdot - k)$, where $\{\widetilde{\Phi}_k\}_{k\in\mathbb{Z}^d}$ is the canonical dual Riesz basis associated to $\{\Phi_k\}_{k\in\mathbb{Z}^d}$. Here $\langle f, \widetilde{\Phi}\rangle = (\langle f, \widetilde{\phi^1}\rangle, \ldots, \langle f, \widetilde{\phi^r}\rangle) \in \mathbb{C}^r$, $\langle f, \widetilde{\phi^i}\rangle = \int_{\mathbb{R}^d} f(z)\overline{\widetilde{\phi^i}(z)}dz$, for $1 \leq i \leq r$, and $\overline{z}$ denotes the complex conjugate of z.*

The next result shows that if the hypothesis of Theorem 7 holds, and X is also a separated set, then X is a set of sampling for $V^p(\Phi)$ and $\boldsymbol{\mu}$, and explicit stability bounds are given.

Theorem 8. *Let $\Phi \in (W_0^1)^{(r)}$ be given. If X is separated with separation constant $\delta > 0$, and P is a bounded projection from $L^p(\mathbb{R}^d)$ onto $V^p(\Phi)$, then X is a set of sampling for $V^p(\Phi)$ and $\boldsymbol{\mu}$ with stability bounds given by*

$$A_p = \frac{1-\alpha}{3^d \|\mathrm{P}\|_{op} \mathcal{N}^{1/p'}}, \tag{45}$$

and

$$B_p = \frac{M\mathcal{N}^{1/p} 3^{d/p} \|\Phi\|_{(W^1)^{(r)}}}{m_p}, \tag{46}$$

where $\mathcal{N} = \mathcal{N}(\delta, p, d) = ([\frac{\sqrt{d}}{\delta}] + 1)^d$, $\frac{1}{p} + \frac{1}{p'} = 1$, $[t]$ denotes the biggest integer lower than or equal to t, m_p is the lower bound constant in condition (6), $\|\mathrm{P}\|_{op}$ is the operator norm of P, and $M > 0$ is the uniform upper bound for the total variations of the elements in the collection $\boldsymbol{\mu}$.

III.1.1 Reconstructing in the presence of noise

Now we investigate the algorithm (44) in the case of noisy samples $\{f_j'\}_{j\in J} \in \ell^p(J)$, but we do not assume that and $\{f_j'\}_{j\in J}$ are samples of a function $f \in V^p(\Phi)$. Given $\{\beta_j\}_{j\in J}$, a bounded partition of unity associated with X, we use the initialization:

$$f_1 = \mathrm{P}\,\mathrm{Q_X}\,\{f_j'\}, \quad f_{n+1} = f_1 + (\mathrm{I} - \mathrm{P}\,\mathrm{A_X})f_n, \quad \forall n \geq 1, \tag{47}$$

and we have the following result.

Theorem 9. *Let $\Phi \in (W_0^1)^{(r)}$, $\{f_j'\}_{j\in J} \in \ell^p(J)$, and P a bounded projection from $L^p(\mathbb{R}^d)$ onto $V^p(\Phi)$ be given. Then the algorithm (47) converges to a function $f_\infty \in V^p(\Phi)$, which satisfies $\mathrm{P}\,\mathrm{A_X}\, f_\infty = \mathrm{P}\,\mathrm{Q_X}\,\{f_j'\}$.*

As a consequence of Theorems 7 and 9, the next result shows the stability of the sampling-reconstruction.

Theorem 10. *Let $\Phi \in (W_0^1)^{(r)}$, P a bounded projection from $L^p(\mathbb{R}^d)$ onto $V^p(\Phi)$ be given, and assume that X is a separated set. Let $\{f_j'\}_{j\in J} \in \ell^p(J)$, and $f \in V^p(\Phi)$ with sampled values $\{g_{x_j}(f)\}_{j\in J}$ be given. Then the following holds:*

$$\|f - f_\infty\|_{L^p} \leq \frac{3^d \mathcal{N}^{1/p'} \|\,\mathrm{P}\,\|_{op}}{1-\alpha} \|\{g_{x_j}(f) - f_j'\}\|_{\ell^p(J)}, \tag{48}$$

where $\mathcal{N} = ([\frac{\sqrt{d}}{\delta}] + 1)^d$, $\frac{1}{p} + \frac{1}{p'} = 1$, $\alpha = \|\,\mathrm{I} - \mathrm{P}\,\mathrm{A_X}\|_{op}$, $f_\infty \in V^p(\Phi)$ is the function given in Theorem 9, and $\delta > 0$ is the separation constant of the set X.

III.2 Proofs

III.2.1 Auxiliary results

We begin this section with two results that are needed for the main proofs.

The first lemma will be stated without proof. The proof of this lemma can be found in [4] Lemmas 5.1 and 5.2 , and in [10] Lemma 8.1.

Lemma 13. *Let $\Phi \in (W_0^1)^{(r)}$, and $f = \sum_{k\in\mathbb{Z}^d} C_k^T \Phi_k$, where $C \in (\ell^p(\mathbb{Z}^d))^{(r)}$. Then:*

1. *The oscillation $\operatorname{osc}_\gamma(f)$ belongs to W^p.*
2. *The oscillation $\operatorname{osc}_\gamma \Phi$ satisfies*

$$\|\operatorname{osc}_\gamma \Phi\|_{(W^1)^{(r)}} \leq ((1 + 2\lceil\gamma\rceil)^d + 1)\|\Phi\|_{(W^1)^{(r)}}, \tag{49}$$

and $\|\operatorname{osc}_\gamma \Phi\|_{(W^1)^{(r)}} \to 0$ as $\gamma \to 0$.

3. *If $|\bigtriangledown \Phi| \in (W_0^1)^{(r)}$, then*

$$\|\Phi\|_{(W^1)^{(r)}} \leq \gamma(2\lceil\gamma\rceil + 1)^d \||\bigtriangledown \Phi|\|_{(W^1)^{(r)}} \tag{50}$$

4. *The oscillation $\operatorname{osc}_\gamma(f)$ satisfies*

$$\|\operatorname{osc}_\gamma(f)\|_{W^p} \leq \|C\|_{(\ell^p(\mathbb{Z}^d))^{(r)}} \|\operatorname{osc}_\gamma \Phi\|_{(W^1)^{(r)}}, \quad \forall C \in (\ell^p(\mathbb{Z}^d))^{(r)}. \tag{51}$$

In particular, $\|\operatorname{osc}_\gamma(f)\|_{W^p} \to 0$ as $\gamma \to 0$. Moreover,

$$\|\,\mathrm{Q_X} f\|_{L^p} \leq \|\,\mathrm{Q_X} f\|_{W^p} \leq ((1+2\lceil\gamma\rceil)^d + 2)\|C\|_{(\ell^p(\mathbb{Z}^d))^{(r)}} \|\Phi\|_{(W^1)^{(r)}}, \quad \forall C \in (\ell^p(\mathbb{Z}^d))^{(r)}. \tag{52}$$

Lemma 14. *Let $\Phi \in (W_0^1)^{(r)}$ be given. Let P be a bounded projection from $L^p(\mathbb{R}^d)$ onto $V^p(\Phi)$. Then there exist $\gamma_0 = \gamma_0(\Phi, \mathrm{P}, p) > 0$, and $a_0 = a_0(\Phi, \mathrm{P}, p) > 0$ such* *that for any* $0 < a \le a_0$*, the operator* $\mathrm{I} - \mathrm{P}\,\mathrm{A_X}$ *is a contraction on* $V^p(\Phi)$ *for any* $\gamma-$*dense set* X *with* $0 < \gamma \le \gamma_0$.

Proof. Let P be a bounded projection from $L^p(\mathbb{R}^d)$ onto $V^p(\Phi)$, and $f = \sum\limits_{k\in\mathbb{Z}^d} C_k^T \Phi_k$, where $C \in (\ell^p(\mathbb{Z}^d))^{(r)}$ be given. Then

$$\begin{aligned}
|f(x) - (\mathrm{Q_X}\, f)(x)| &= \Big| f(x) - \sum_{j\in J} f(x_j)\beta_j(x) \Big| \\
&= \Big| \sum_{j\in J} (f(x) - f(x_j))\beta_j(x) \Big| \\
&\le \sum_{j\in J} |f(x) - f(x_j)|\beta_j(x) \\
&\le \operatorname{osc}_\gamma(f)(x) \sum_{j\in J} \beta_j(x) = \operatorname{osc}_\gamma(f)(x).
\end{aligned}$$

From this pointwise estimate and (51) we obtain

$$\begin{aligned}
\|f - \mathrm{Q_X}\, f\|_{L^p} &\le \|f - \mathrm{Q_X}\, f\|_{W^p} \\
&\le \|\operatorname{osc}_\gamma(f)\|_{W^p} \le \|C\|_{(\ell^p(\mathbb{Z}^d))^{(r)}} \|\operatorname{osc}_\gamma \Phi\|_{(W^1)^{(r)}} \\
&\le \frac{1}{m_p} \|f\|_{L^p} \|\operatorname{osc}_\gamma \Phi\|_{(W^1)^{(r)}},
\end{aligned}$$

where we have used condition (6) in the last inequality. Consequently,

$$\|f - \mathrm{Q_X}\, f\|_{L^p} \le \frac{1}{m_p} \|f\|_{L^p} \|\operatorname{osc}_\gamma \Phi\|_{(W^1)^{(r)}}. \tag{53}$$

On the other hand,

$$\begin{aligned}
|(\mathrm{Q_X}\, f - \mathrm{A_X}\, f)(x)| &= \Big| \sum_{j\in J} (f(x_j) - g_{x_j}(f))\beta_j(x) \Big| \\
&= \Big| \sum_{j\in J} \Big(\int_{\mathbb{R}^d} (f(x_j) - f(z))\, d\mu_{x_j}(z) \Big) \beta_j(x) \Big| \\
&\le \sum_{j\in J} \int_{\mathbb{R}^d} |f(x_j) - f(z)|\, d|\mu_{x_j}|(z) \beta_j(x) \\
&\le \sum_{j\in J} \operatorname{osc}_a(f)(x_j)\beta_j(x) \int_{\mathbb{R}^d} d|\mu_{x_j}|(z) \\
&\le M \sum_{j\in J} \operatorname{osc}_a(f)(x_j)\beta_j(x) \\
&\le M \sum_{j\in J} \Big(\sum_{i=1}^{r} \sum_{k\in\mathbb{Z}^d} |c_k^i| \operatorname{osc}_a(\phi^i)(x_j - k) \Big) \beta_j(x).
\end{aligned}$$

Using Lemma 13, condition (6), Triangular inequality, and the above pointwise estimate, we have

$$\| \mathrm{Q}_{\mathrm{X}} f - \mathrm{A}_{\mathrm{X}} f \|_{L^p} \leq \frac{M}{m_p}((1+2\lceil\gamma\rceil)^d + 2)\| \operatorname{osc}_a \Phi\|_{(W^1)^{(r)}} \|f\|_{L^p}. \tag{54}$$

Since $f \in V^p(\Phi)$, then $\mathrm{P} f = f$. Therefore,

$$\begin{aligned} \|f - \mathrm{P}\,\mathrm{A}_{\mathrm{X}} f\|_{L^p} &\leq \| \mathrm{P} f - \mathrm{P}\,\mathrm{Q}_{\mathrm{X}} f\|_{L^p} + \| \mathrm{P}\,\mathrm{Q}_{\mathrm{X}} f - \mathrm{P}\,\mathrm{A}_{\mathrm{X}} f\|_{L^p} \\ &\leq \| \mathrm{P} \|_{op} \|f - \mathrm{Q}_{\mathrm{X}} f\|_{L^p} + \| \mathrm{P} \|_{op} \| \mathrm{Q}_{\mathrm{X}} f - \mathrm{A}_{\mathrm{X}} f\|_{L^p}. \end{aligned}$$

Using now (53), (54), and the above inequality we get

$$\|f - \mathrm{P}\,\mathrm{A}_{\mathrm{X}} f\|_{L^p} \leq \frac{\| \mathrm{P} \|_{op}}{m_p}\Big(\| \operatorname{osc}_\gamma \Phi\|_{(W^1)^{(r)}} + M((1+2\lceil\gamma\rceil)^d+2)\| \operatorname{osc}_a \Phi\|_{(W^1)^{(r)}}\Big)\|f\|_{L^p}. \tag{55}$$

Let $0 < \epsilon < \frac{m_p}{\| \mathrm{P} \|_{op}}$ be given. Since $\| \operatorname{osc}_\gamma \Phi\|_{(W^1)^{(r)}} \to 0$ as $\gamma \to 0^+$, then there exists $\gamma_0 = \gamma_0(\epsilon, \Phi, \mathrm{P}, p) > 0$, and $a_0 = a_0(\epsilon, \Phi, \mathrm{P}, p) > 0$ such that

$$\| \operatorname{osc}_\gamma \Phi\|_{(W^1)^{(r)}} \leq \frac{\epsilon}{2}, \quad \text{for all} \quad 0 < \gamma \leq \gamma_0,$$

and

$$M((1+2\lceil\gamma\rceil)^d + 2)\| \operatorname{osc}_a \Phi\|_{(W^1)^{(r)}} \leq \frac{\epsilon}{2}, \quad \text{for all} \quad 0 < a \leq a_0.$$

Choosing γ_0 and a_0 so that for any $0 < \gamma \leq \gamma_0$, and $0 < a \leq a_0$ we have

$$\|f - \mathrm{P}\,\mathrm{A}_{\mathrm{X}} f\|_{L^p} \leq \frac{\epsilon \| \mathrm{P} \|_{op}}{m_p} \|f\|_{L^p},$$

then the conclusion of the lemma follows. □

III.2.2 Proofs for Section III.1

Now we are ready to prove the main results of this chapter.

Proof of Theorem 7.

Proof. Let $e_n = f - f_n$ be the error after n iterations of the algorithm (44). Then the sequence $\{e_n\}_{n\in\mathbb{N}}$ satisfies

$$\begin{aligned} e_{n+1} &= f - f_{n+1} = f - f_n - \mathrm{P}\,\mathrm{A}_{\mathrm{X}}(f - f_n) \\ &= (\mathrm{I} - \mathrm{P}\,\mathrm{A}_{\mathrm{X}})(f - f_n) = (\mathrm{I} - \mathrm{P}\,\mathrm{A}_{\mathrm{X}})(e_n). \end{aligned}$$

By using Lemma 14, there exist a density $\gamma_0 > 0$, and $a_0 > 0$ such that for any $0 < \gamma \leq \gamma_0$, and $0 < a \leq a_0$, $\mathrm{I} - \mathrm{P}\,\mathrm{A}_{\mathrm{X}}$ is a contraction on $V^p(\Phi)$. Therefore, by taking

$\alpha := \| \mathrm{I} - \mathrm{P}\,\mathrm{A_X}\|_{op} < 1$, we have

$$\|e_{n+1}\|_{L^p} \leq \alpha \|e_n\|_{L^p},$$

and by induction it follows that

$$\|e_{n+1}\|_{L^p} \leq \alpha^{n+1} \|f\|_{L^p}, \tag{56}$$

and $\|e_n\|_{L^p} \to 0$ geometrically fast. Since for $V^p(\Phi)$ the W^p norm, and the L^p norm are equivalent, then (56) also holds in the W^p norm and uniformly on $\mathbb{R}^d$, and Theorem 7 is proved. □

Proof of Theorem 8.

Proof. Note that by hypothesis and Lemma 14 we have that there exists $\gamma_0 > 0$ such that $\mathrm{I} - \mathrm{P}\,\mathrm{A_X}$ is a contraction for any $\gamma-$dense set X with $0 < \gamma \leq \gamma_0$. Hence, $\alpha = \| \mathrm{I} - \mathrm{P}\,\mathrm{A_X}\|_{op} < 1$. Thus, the operator $\mathrm{P}\,\mathrm{A_X}$ is invertible on $V^p(\Phi)$. It is not hard to show that $\mathrm{P}\,\mathrm{A_X}$ and $(\mathrm{P}\,\mathrm{A_X})^{-1}$ satisfy:

$$1 - \alpha \leq \| \mathrm{P}\,\mathrm{A_X}\|_{op} \leq 1 + \alpha, \tag{57}$$

and

$$\frac{1}{1+\alpha} \leq \|(\mathrm{P}\,\mathrm{A_X})^{-1}\|_{op} \leq \frac{1}{1-\alpha}. \tag{58}$$

Let us show (45). Let $f \in V^p(\Phi)$ be given. From the definition of the operators $\mathrm{A_X}$ and $\mathrm{Q_X}$, it follows that $\mathrm{A_X}\, f = \mathrm{Q_X}\,\{g_{x_j}(f)\}$, and thus, $\mathrm{P}\,\mathrm{A_X}\, f = \mathrm{P}\,\mathrm{Q_X}\,\{g_{x_j}(f)\}$. Therefore, $f = (\mathrm{P}\,\mathrm{A_X})^{-1}\,\mathrm{P}\,\mathrm{Q_X}\,\{g_{x_j}(f)\}$. Consequently,

$$\begin{aligned}\|f\|_{L^p} &\leq \|(\mathrm{P}\,\mathrm{A_X})^{-1}\|_{op} \| \mathrm{P}\,\|_{op} \| \mathrm{Q_X}\,\|_{op} \|\{g_{x_j}(f)\}\|_{\ell^p(J)} \\ &\leq \frac{\| \mathrm{P}\,\|_{op} \| \mathrm{Q_X}\,\|_{op}}{1-\alpha} \|\{g_{x_j}(f)\}\|_{\ell^p(J)}.\end{aligned}$$

In order to complete the proof of (45), we need an upper estimate for $\| \mathrm{Q_X}\,\|_{op}$. Let χ be the characteristic function of the set $B_\gamma(0) + [0,1]^d$. Clearly, we may assume without loss of generality that $0 < \gamma < 1$. Since $0 \leq \beta_j \leq 1$, and $\operatorname{supp} \beta_j \subset B_\gamma(x_j)$, then for all $x_j \in k + [0,1]^d$, $\beta_j(x) \leq \chi(x - k)$. Therefore,

$$| \mathrm{Q_X}\, c| = \Big| \sum_{j \in J} c_j \beta_j(x) \Big| \leq \sum_{k \in \mathbb{Z}^d} \Big(\sum_{j : x_j \in k + [0,1]^d} |c_j| \Big) \chi(x - k),$$

and by (25) we have

$$\| \mathrm{Q_X}\, c\|_{W^p} \leq \left(\sum_{k \in \mathbb{Z}^d} \Big(\sum_{j : x_j \in k + [0,1]^d} |c_j| \Big)^p \right)^{1/p} \|\chi\|_{W^1}.$$

Since X is a separated set with separation constant $\delta > 0$, then there are at most $\mathcal{N} = \mathcal{N}(\delta, p, d) = ([\frac{\sqrt{d}}{\delta}] + 1)^d$ sampling points x_j in each cube $k + [0,1]^d$. By applying Hölder's inequality, we get

$$\Big(\sum_{j:x_j \in k+[0,1]^d} |c_j|\Big)^p \leq \mathcal{N}^{p/p'} \sum_{j:x_j \in k+[0,1]^d} |c_j|^p,$$

where $\frac{1}{p} + \frac{1}{p'} = 1$. Consequently,

$$\| \, \mathrm{Q_X} \, c \|_{W^p} \leq \mathcal{N}^{1/p'} \| \{c_j\} \|_{\ell^p(J)} \| \chi \|_{W^1}.$$

An easy computation using induction on $d \in \mathbb{N}$ shows that $\|\chi\|_{W^1} = 3^d$. Therefore,

$$\| \, \mathrm{Q_X} \, c \|_{W^p} \leq 3^d \mathcal{N}^{1/p'} \| \{c_j\} \|_{\ell^p(J)},$$

and

$$\| \, \mathrm{Q_X} \, \|_{op} \leq 3^d \mathcal{N}^{1/p'}. \tag{59}$$

Hence,

$$\frac{1-\alpha}{3^d \| \, \mathrm{P} \, \|_{op} \mathcal{N}^{1/p'}} \|f\|_{L^p} \leq \| \{g_{x_j}(f)\} \|_{\ell^p(J)}, \quad \text{for all} \quad f \in V^p(\Phi).$$

Let us show (46). Note that

$$\begin{aligned}
\sum_{x_j \in k+[0,1]^d} |g_{x_j}(f)|^p &= \sum_{x_j \in k+[0,1]^d} \Big| \int_{\mathbb{R}^d} f(z) d\mu_{x_j}(z) \Big|^p \\
&\leq \sum_{x_j \in k+[0,1]^d} \|\mu_{x_j}\|^p \Big(\int_{\mathbb{R}^d} |f(z)| \frac{d|\mu_{x_j}|(z)}{\|\mu_{x_j}\|} \Big)^p \\
&\leq \sum_{x_j \in k+[0,1]^d} \|\mu_{x_j}\|^p \int_{\mathbb{R}^d} |f(z)|^p \frac{d|\mu_{x_j}|(z)}{\|\mu_{x_j}\|} \\
&\leq M^p \sum_{x_j \in k+[0,1]^d} \operatorname*{esssup}_{z \in x_j + [-a,a]^d} |f(z)|^p.
\end{aligned}$$

Since X is a separated set, then there exist at most $\mathcal{N} = \mathcal{N}(\delta, p, d) = ([\frac{\sqrt{d}}{\delta}] + 1)^d$ sampling points in each cube $k + [0,1]^d$. Assuming without loss of generality that $0 < a \leq 1$, then

$$\sum_{x_j \in k+[0,1]^d} |g_{x_j}(f)|^p \leq M^p 3^d \mathcal{N} \operatorname*{esssup}_{z \in k+[0,1]^d} |f(z)|^p.$$

Consequently, by taking the sum over $k \in \mathbb{Z}^d$ in the above inequality, we obtain:

$$\begin{aligned}
\|\{g_{x_j}(f)\}\|_{\ell^p(J)} &\leq M\mathcal{N}^{1/p}3^{d/p}\|f\|_{W^p} \\
&\leq M\mathcal{N}^{1/p}3^{d/p}\|C\|_{(\ell^p(\mathbb{Z}^d))^{(r)}}\|\Phi\|_{(W^1)^{(r)}} \\
&\leq \frac{M\mathcal{N}^{1/p}3^{d/p}\|\Phi\|_{(W^1)^{(r)}}}{m_p}\|f\|_{L^p},
\end{aligned}$$

and Theorem 8 is proved. □

Proof of Theorem 9.

Proof. Assume the hypothesis of Theorem 9 holds. From Lemma 14, the operator $\mathrm{I} - \mathrm{P}\,\mathrm{A_X}$ is a contraction. Consequently, the sequence $\{f_n\}_{n\geq 1}$ defined by algorithm (47) converges to a function $f_\infty \in V^p(\Phi)$. Taking limits in both sides of (47) as $n \to \infty$, we have:

$$f_\infty = f_1 + (\mathrm{I} - \mathrm{P}\,\mathrm{A_X})f_\infty.$$

Therefore, $f_1 - \mathrm{P}\,\mathrm{A_X}\, f_\infty = 0$. Taking into account that $f_1 = \mathrm{P}\,\mathrm{Q_X}\,\{f_j^{'}\}$, then the conclusion of Theorem 9 follows. □

Proof of Theorem 10.

Proof. Assume that the hypothesis of Theorem 10 holds. By Lemma 14, there exists $\gamma_0 > 0$ such that the operator $\mathrm{I} - \mathrm{P}\,\mathrm{A_X}$ is a contraction on $V^p(\Phi)$ for any $\gamma-$dense set X with $0 < \gamma \leq \gamma_0$. Hence, $\alpha = \|\,\mathrm{I} - \mathrm{P}\,\mathrm{A_X}\|_{op} < 1$, the operator $\mathrm{P}\,\mathrm{A_X}$ is invertible, and (58) takes place. On the other hand, from the definition of the operators $\mathrm{A_X}$ and $\mathrm{Q_X}$, it follows that $\mathrm{A_X}\, f = \mathrm{Q_X}\,\{g_{x_j}(f)\}$, and thus, $\mathrm{P}\,\mathrm{A_X}\, f = \mathrm{P}\,\mathrm{Q_X}\,\{g_{x_j}(f)\}$. Therefore, $f = (\mathrm{P}\,\mathrm{A_X})^{-1}\,\mathrm{P}\,\mathrm{Q_X}\,\{g_{x_j}(f)\}$. Applying now Theorem 9, we have that there exists a function $f_\infty \in V^p(\Phi)$ such that $\mathrm{P}\,\mathrm{A_X}\, f_\infty = \mathrm{P}\,\mathrm{Q_X}\,\{f_j^{'}\}$. Hence, $f_\infty = (\mathrm{P}\,\mathrm{A_X})^{-1}\,\mathrm{P}\,\mathrm{Q_X}\,\{f_j^{'}\}$. Consequently,

$$\begin{aligned}
\|f - f_\infty\|_{L^p} &\leq \|(\mathrm{P}\,\mathrm{A_X})^{-1}\|_{op}\|\,\mathrm{P}\,\mathrm{Q_X}\|_{op}\|\{g_{x_j}(f) - f_j^{'}\}\|_{\ell^p(J)} \\
&\leq \frac{\|\,\mathrm{P}\,\mathrm{Q_X}\|_{op}}{1-\alpha}\|\{g_{x_j}(f) - f_j^{'}\}\|_{\ell^p(J)} \\
&\leq \frac{\|\,\mathrm{P}\,\|_{op}\|\,\mathrm{Q_X}\,\|_{op}}{1-\alpha}\|\{g_{x_j}(f) - f_j^{'}\}\|_{\ell^p(J)}.
\end{aligned}$$

Now the conclusion of Theorem 10 follows using (59). □

CHAPTER IV

ON THE CONSTRUCTION OF OPTIMAL NON-LINEAR SIGNAL MODELS

IV.1 Main Results

Recent research and new paradigms in mathematics, engineering, and science assume non-linear signal models of the form $\mathcal{M} = \bigcup_{i\in I} C_i \subset \mathcal{H}$ consisting of a union of closed subspaces $C_i \subset \mathcal{H}$ instead of a single closed subspace $\mathcal{M} = C \subset \mathcal{H}$, where $\mathcal{H}$ is a Hilbert space describing signals. The problem is to determine the signal model from a set of observed data $\mathcal{F} = \{f_1, \dots, f_m\}$. For the purposes of modeling, the subspaces C_i must be restricted and assumed to belong to some class of closed subspaces $\mathcal{C}$ (the assumption on $\mathcal{C}$ depends on the application), e.g., if $\mathcal{H} = \mathbb{C}^d$, then one possible choice of $\mathcal{C}$ is the set $\mathcal{L}_n$ of subspaces of $\mathbb{C}^d$ with dimension less than or equal to n. The optimal model compatible with the data $\mathcal{F} = \{f_1, \dots, f_m\}$ can then be obtained by minimizing the expression

$$e(\mathcal{F}, \{C_1, \dots, C_l\}) = \sum_{f\in\mathcal{F}} \min_{1\le j\le l} \operatorname{dist}^2(f, C_j),$$

where $C_j \in \mathcal{C}$ for $j = 1, \dots, l$. However, depending on $\mathcal{C}$ this expression may not have a minimum. In this chapter we will discuss this problem and produce sufficient conditions over $\mathcal{C}$ for the above non-linear minimization problem to have a solution.

IV.1.1 General Setting

We shall introduce the main notations and definitions that we shall use in this section.

Let (X, d) be a complete metric space. Given a finite set $\mathcal{F} \subset X$ and a closed subset C of X, we denote by $E(\mathcal{F}, C)$ the total distance of the data set $\mathcal{F}$ to the closed subset C, i.e.,

$$E(\mathcal{F}, C) = \sum_{f\in\mathcal{F}} d^2(f, C). \tag{60}$$

We set $E(\mathcal{F}, C) = 0$ for $\mathcal{F} = \emptyset$ and any closed subset C of X.

Definition 14. *Let $\mathcal{C}$ be a collection of closed subsets of X. We say that $\mathcal{C}$ has the Minimal Approximation Property(MAP) if for all finite subset $\mathcal{F} \subset X$ there exists $C_0 = C_0(\mathcal{F}) \in \mathcal{C}$ that minimize $E(\mathcal{F}, C)$ over all closed subsets $C \in \mathcal{C}$, that is,*

$$E(\mathcal{F}, C_0) = \min\{E(\mathcal{F}, C) : C \in \mathcal{C}\}. \tag{61}$$

Let us fix $l, m \in \mathbb{N}$ with $1 \le l \le m$, and let $\mathcal{F} = \{f_1, \dots, f_m\}$ be a finite set of points in X.

Define Ξ to be the set of sequences of elements in $\mathcal{C}$ of length l, i.e.,

$$\Xi = \Xi(l) = \{\{C_1, \ldots, C_l\} : C_i \in \mathcal{C}, 1 \leq i \leq l\}.$$

We shall call these finite sequences bundles. For $\mathbf{C} \in \Xi$ with $\mathbf{C} = \{C_1, \ldots, C_l\}$, we define

$$e(\mathcal{F}, \mathbf{C}) = \sum_{f \in \mathcal{F}} \min_{1 \leq j \leq l} d^2(f, C_j). \tag{62}$$

Remark 16. *Note that $e(\mathcal{F}, \mathbf{C})$ is computed as follows: For each $f \in \mathcal{F}$ find the closed subset $C_{j(f)} \in \mathbf{C}$ closest to f, compute $d^2(f, C_{j(f)})$, and then sum over all values found by letting f run through $\mathcal{F}$.*

Let us consider the following Problem:

Problem 1. *Given a finite set $\mathcal{F} \subset X$,*

1. *minimize $e(\mathcal{F}, \mathbf{C})$ over $\mathbf{C} \in \Xi$. That is, find*

$$\inf\{e(\mathcal{F}, \mathbf{C}) : \mathbf{C} \in \Xi\}. \tag{63}$$

2. *Find a bundle $\mathbf{C}_0 \in \Xi$, if it exists, such that*

$$e(\mathcal{F}, \mathbf{C}_0) = \inf\{e(\mathcal{F}, \mathbf{C}) : \mathbf{C} \in \Xi\}. \tag{64}$$

Any $C_0 \in \Xi$ satisfying (64) is called a solution to Problem 1.

Remark 17. *Let $\mathcal{C}$ be a collection of closed subsets of X. If Problem 1 can be solved for any $l \in \mathbb{N}$, then it can be solved for $l = 1$. Therefore, the collection $\mathcal{C}$ of closed subsets of X has the MAP.*

Our next task is to show that if $\mathcal{C}$ has the MAP, then we can solve Problem 1. So, for the rest of this section we shall assume that the collection $\mathcal{C}$ of closed subsets of X has the MAP.

Let us introduce some definitions and set some notations.

We shall denote by $\Pi = \Pi_l$ the set of all l-sequences $P = \{\mathcal{F}_1, \ldots, \mathcal{F}_l\}$ of subsets of $\mathcal{F}$ satisfying the property that for all $1 \leq i, j \leq l$,

$$\mathcal{F}_i \subset \mathcal{F}, \mathcal{F} = \cup_{s=1}^{l} \mathcal{F}_s, \text{and } \mathcal{F}_i \cap \mathcal{F}_j = \emptyset \text{ for } i \neq j.$$

Note that we allow some of the elements $P \in \Pi$ to be the empty set. We shall call the elements of Π_l partitions(of $\mathcal{F}$).

For $P \in \Pi_l$, $P = \{\mathcal{F}_1, \ldots, \mathcal{F}_l\}$, $\mathbf{C} \in \Xi$, $\mathbf{C} = \{C_1, \ldots, C_l\}$ we define

$$\Gamma(P, \mathbf{C}) = \sum_{i=1}^{l} E(\mathcal{F}_i, C_i). \tag{65}$$

Note that Γ measures the error between a fixed partition and a fixed bundle.

Given a bundle $\mathbf{C} \in \Xi$, $\mathbf{C} = \{C_1, \ldots, C_l\}$ there is a subset $\Omega_l(\mathbf{C}) \subset \Pi_l$ of *best partitions* in Π_l associated to $\mathbf{C}$ defined by $P = \{\mathcal{F}_1, \ldots, \mathcal{F}_l\} \in \Pi_l$ is a member of $\Omega_l(\mathbf{C})$ if it satisfies $f \in \mathcal{F}_j$ implies that $d(f, C_j) \leq d(f, C_r)$, $r = 1, \ldots, l$. Conversely, since $\mathcal{C}$ has the MAP, given a partition $P = \{\mathcal{F}_1, \ldots, \mathcal{F}_l\}$, we define a subset $\mathcal{W}(P) \subset \Xi$ of *best bundles* associated to $\mathcal{F}$ by $\mathbf{C}_l(P) = \{C_1, \ldots, C_l\} \in \Xi$ is a member of $\mathcal{W}(P)$ if C_i is an optimal closed subset for $\mathcal{F}_i$(in the sense of (61)) for each $i = 1, \ldots, l$.

In what follows when we refer to a best partition associated to a bundle $\mathbf{C}$ we will mean any element in $\Omega_l(\mathbf{C})$. Similarly, when we talk of a best bundle associated to a partition P, this will mean an element in $\mathcal{W}(P)$.

We also consider the set of pairs $(P, \mathbf{C}_l(P))$, where $P \in \Pi_l$ and $\mathbf{C}_l(P) \in \mathcal{W}(P)$. We shall say that a pair $(P_0, \mathbf{C}_l(P_0))$ is Γ-minimal if

$$\Gamma(P_0, \mathbf{C}_l(P_0)) \leq \Gamma(P, \mathbf{C}_l(P)) \tag{66}$$

for all such pairs.

Note that when we try to compute $e(\mathcal{F}, \mathbf{C})$, for each $f \in \mathcal{F}$, we first have to find the closed subset $C_{j(f)} \in \mathbf{C}$ that is closest to f and then compute $\mathrm{dist}^2(f, C_{j(f)})$. While for Γ, a partition is given and we compute the distance of each element in the partition to its corresponding closed subset(no the closest one necessarily). In Lemma 15 we shall show that e and Γ can indeed be compared.

The next result shows that if the collection $\mathcal{C}$ of closed subsets of X satisfies the MAP, then Problem 1 can be solved.

Theorem 11. *Let (X, d) be a complete metric space and $\mathcal{C}$ a collection of closed subsets of X satisfying the MAP. Let $l, m \in \mathbb{N}$ with $l \leq m$, and $\mathcal{F} = \{f_1, \ldots, f_m\}$ a set of points in X. Then*

1. *There exists a bundle $\mathbf{C}_0 \in \Xi$ that solves Problem 1 for the data $\mathcal{F}$, that is,*

$$e(\mathcal{F}, \mathbf{C}_0) = \inf\{e(\mathcal{F}, \mathbf{C}) : \mathbf{C} \in \Xi\}.$$

2. *If $(P_0, \mathbf{C}_l(P_0))$ is Γ-minimal pair, then all the elements of $\mathcal{W}(P_0)$ are solutions to Problem 1.*
3. *Furthermore, if $\mathbf{C}_0$ is a solution to Problem 1, then there exists $P_0 \in \Pi_l$ such that $\mathbf{C}_0 \in \mathcal{W}(P_0)$, i.e., $(P_0, \mathbf{C}_0)$ is a Γ-minimal pair.*

Remark 18. *From Remark 17 and Theorem 11 we have that Problem 1 can be solved if and only if the collection of closed subsets $\mathcal{C}$ satisfies the MAP.*

IV.1.2 The MAP in Hilbert Spaces. A sufficient condition

In this section we are interested in studying the MAP in Hilbert spaces. We shall show that in a separable Hilbert space, we can provide a sufficient condition so that the MAP takes place. Hence, Problem 1 can be solved.

Let $\mathcal{H}$ be a separable Hilbert space with inner product $< \cdot, \cdot >$ and induced norm $\| \cdot \|$. For $x, y \in \mathcal{H}$, we denote by $\mathrm{dist}(x, y) = \|x - y\|$ the distance form x to y.

We shall denote the linear space of bounded and linear operators on $\mathcal{H}$ by $\mathcal{B}(\mathcal{H})$. Next, we define the topologies on $\mathcal{B}(\mathcal{H})$ we will use in this chapter.

Definition 15. *Let $\{P_n\}_{n\geq 1} \subset \mathcal{B}(\mathcal{H})$, and $P \in \mathcal{B}(\mathcal{H})$.*

1. *We say that $P_n \to P$ in the norm operator topology(N.O.T.) if $\lim_{n\to\infty} \|P_n - P\|_{op} = 0$, where $\|P\|_{op} = \sup\{\|Px\| : \|x\| \leq 1\}$.*
2. *We say that $P_n \to P$ in the strong operator topology(S.O.T.) if $\lim_{n\to\infty} \|P_n x - Px\| = 0$, for all $x \in \mathcal{H}$.*
3. *We say that $P_n \to P$ in the weak operator topology(W.O.T.), or in the weak sense, or weakly, if $\lim_{n\to\infty} < P_n x, y >=< Px, y >$, for all $x, y \in \mathcal{H}$. Note that this is the weak-$*$ topology on $\mathcal{B}(\mathcal{H})$.*

Remark 19. *Note that from the above definition the following implications take place: N.O.T.$\Rightarrow$S.O.T.$\Rightarrow$W.O.T.*

The next result provides a sufficient condition in terms of orthogonal projections in order that the MAP takes place in a separable Hilbert space. Therefore, Problem 1 can be solved.

Theorem 12. *Let $\mathcal{C}$ be a collection of closed subspaces in $\mathcal{H}$. If the collection of orthogonal projections $\mathcal{P} = \{P_C\}_{C\in\mathcal{C}}$ is closed w.r.t. the W.O.T. on $\mathcal{B}(\mathcal{H})$, then $\mathcal{C}$ satisfies the MAP.*

Recall that for $1 \leq n \leq d$, we denote by $\mathcal{L}_n$ the collection of all subspaces of $\mathbb{C}^d$ with dimension smaller or equal than n. As a consequence of Theorem 12, we obtain the well-known qualitative version of the Eckart-Young Theorem, that is, $\mathcal{L}_n$ satisfies the MAP.

Corollary 6. *Let $\mathcal{H} = \mathbb{C}^d$, and $1 \leq n \leq d$. Then $\mathcal{L}_n$ satisfies the MAP.*

IV.1.3 A MAP related problem for Unitary Operators

In this section we study a MAP related problem when we consider a class $\mathcal{C}$ which is defined in terms of a collection of unitary operators applied to a convex subset of a separable Hilbert space $\mathcal{H}$, and we obtain an algorithm for constructing particular collections of closed subspaces of $\mathcal{H}$ for which we a priori know the existence of a minimizer to Problem 1.

Let us consider the following MAP related problem for unitary operators:

Problem 2. *Let $\mathcal{S}$ be a fixed and given convex subset of a separable Hilbert space $\mathcal{H}$ (here we are assuming that $\mathcal{S} \neq \emptyset$), $\mathcal{F}$ is a finite subset of $\mathcal{H}$, and $\mathcal{U} \subset \mathcal{B}(\mathcal{H})$ a collection of unitary operators. We want to find, if it exists, an operator $U_0 \in \mathcal{U}$ such that*

$$\widehat{H}(\mathcal{F}, U_0) = \inf\{\widehat{H}(\mathcal{F}, U) : U \in \mathcal{U}\}, \tag{67}$$

where $\widehat{H}(\mathcal{F}, U) = \sum_{f\in\mathcal{F}} \operatorname{dist}^2(f, U\mathcal{S})$.

Remark 20. *Any U_0 that satisfies (67) is called a solution to Problem 2. Moreover, if $\mathcal{S} = \mathcal{H}$, then any $U \in \mathcal{U}$ is a solution to Problem 2, and in this case we have $\widehat{H}(\mathcal{F}, U) = 0$ for all $U \in \mathcal{U}$.*

Our next result shows that if the collection $\mathcal{U}$ is weakly closed, then Problem 2 has a solution if $\mathcal{S}$ is either a bounded convex subset of $\mathbb{C}^d$ or a closed convex subset of a separable Hilbert space $\mathcal{H}$.

Theorem 13. *Assume that $\mathcal{U}$ is a collection of unitary operators which is closed w.r.t. the W.O.T. on $\mathcal{B}(\mathcal{H})$. If $\mathcal{S}$ is either a bounded convex subset of $\mathbb{C}^d$ or a closed convex subset of the separable Hilbert space $\mathcal{H}$, then Problem 2 has a solution, that is, there exists $U_0 \in \mathcal{U}$ such that*

$$\widehat{H}(\mathfrak{F}, U_0) = \inf\{\widehat{H}(\mathfrak{F}, U) : U \in \mathcal{U}\}.$$

Remark 21. *Note that if the collection $\mathcal{U}$ is not weakly closed, then the conclusion of Theorem 13 may fail. For example, if $\mathcal{H} = \mathbb{R}^2$, $\mathfrak{F} = \{(1,2)\}$, $\mathcal{S}$ is the x-axis, that is, $\mathcal{S} = span\{(1,0)\}$, and $\mathcal{U} = \{U_\theta : 0 < \theta < \frac{\pi}{4}\}$, where*

$$\mathbf{U}_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \tag{68}$$

then $\mathcal{U}$ is not weakly closed because the sequence $\{U_n\}_{n\geq 2}$, defined by

$$\mathbf{U_n} = \begin{pmatrix} \cos\theta_n & -\sin\theta_n \\ \sin\theta_n & \cos\theta_n \end{pmatrix},$$

with $\theta_n = \frac{1}{n}$, belongs to $\mathcal{U}$, and $\lim_{n\to\infty} U_n = I_2$, where I_2 is the 2×2 identity matrix, does not belong to $\mathcal{U}$. On the other hand, an easy computation shows that minimum of $\widehat{H}(\mathfrak{F}, U)$ on $\mathcal{U}$ does not exist.

As a consequence of Theorem 13 we obtain a result that allows us to construct collections of closed subspaces for which Problem 1 can be solved. First, let us introduce some definitions.

Definition 16. *A hyperplane in $\mathcal{H}$ is any set of the form*

$$\mathcal{S} = \{y \in \mathcal{H} :< y, x_0 >= c\}, \tag{69}$$

where $x_0 \in \mathcal{H} \setminus \{0\}$ and $c \in \mathbb{C}$.

Definition 17. *A subset $\mathcal{M}$ is called a closed affine subspace of $\mathcal{H}$ if $\mathcal{M}$ is a translate of a closed subspace of $\mathcal{H}$, that is,*

$$\mathcal{M} = v + \mathcal{S}, \tag{70}$$

for some $v \in \mathcal{M}$ and some closed subspace $\mathcal{S}$ of $\mathcal{H}$.

Definition 18. *A subset C of the inner product space X is called a convex cone if $\alpha x + \beta y \in C$ whenever $x, y \in C$ and $\alpha, \beta \geq 0$.*

Definition 19. *A half-space in a Hilbert space $\mathcal{H}$ is a set of the form*

$$\mathcal{S} = \{y \in \mathcal{H} :< y, x_0 >\leq c\}, \tag{71}$$

or

$$\mathcal{S} = \{y \in \mathcal{H} :< y, x_0 >\geq c\},$$

where $x_0 \in \mathcal{H} \setminus \{0\}$ and $c \in \mathbb{R}$.

Remark 22. *Every half-space of the form $\mathcal{S} = \{y \in \mathcal{H} :< y, x_0 >\geq c\}$ can be written in the form $\mathcal{S} = \{y \in \mathcal{H} :< y, x_1 >\leq c_1\}$, where $x_1 = -x_0$ and $c_1 = -c$. Thus, we only need to consider those half-spaces where the defining inequality is $\leq$.*

Corollary 7. *Assume that $\mathcal{U}$ is a collection of unitary operators which is closed w.r.t. the W.O.T. on $\mathcal{B}(\mathcal{H})$. If $\mathcal{S}$ is either a closed subspace, a closed convex cone, a hyperplane, a half-space, or a closed affine subspace, then Problem 2 has a solution, that is, there exists $U_0 \in \mathcal{U}$ such that*

$$\widehat{H}(\mathfrak{F}, U_0) = \inf\{\widehat{H}(\mathfrak{F}, U) : U \in \mathcal{U}\}.$$

As a consequence of Corollary 7 when $\mathcal{S}$ is a closed subspace of $\mathcal{H}$, we obtain the following result that allows us to construct collections of closed subspaces satisfying the MAP.

Corollary 8. *Let $\mathcal{U}$ be a collection of unitary operators which is closed w.r.t. the W.O.T. on $\mathcal{B}(\mathcal{H})$. Let $\mathcal{S}$ be a fixed and given closed subspace of $\mathcal{H}$. Then the collection of closed subspaces $\mathcal{C}_{\mathcal{S}} := \{U\mathcal{S}\}_{U\in\mathcal{U}}$ has the MAP.*

Remark 23. *Note that Corollary 8 provides a "procedure" for constructing particular collections of closed subspaces in the Hilbert space $\mathcal{H}$ satisfying the MAP, that is, collections of closed subspaces for which we a priori know that Problem 1 can be solved:*

1. *Pick a closed subspace $\mathcal{S}$ in $\mathcal{H}$ (we are assuming that $\mathcal{S} \neq \{0\}$, $\mathcal{S} \neq \mathcal{H}$).*
2. *Pick a collection $\mathcal{U}$ of unitary operators which is closed w.r.t. the W.O.T. on $\mathcal{B}(\mathcal{H})$.*
3. *From Corollary 8 it follows that the collection of closed subspaces $\mathcal{C}_{\mathcal{S}} := \{U\mathcal{S}\}_{U\in\mathcal{U}}$ has the MAP, that is, Problem 1 can be solved.*

Remark 24. *At the end of this chapter we shall show that Corollary 6, that is, the qualitative version of the Eckart-Young Theorem, can be also obtained as a consequence of Corollary 8.*

IV.2 Proofs

IV.2.1 Proofs for Section IV.1.1

Lemma 15. *Let* $(P_0, \mathbf{C}_l(P_0))$ *be a* Γ*-minimal pair. Then we have*

$$e(\mathcal{F}, \mathbf{C}_l(P_0)) = \Gamma(P_0, \mathbf{C}_l(P_0)). \tag{72}$$

Proof. It is clear that $e(\mathcal{F}, \mathbf{C}_l(P_0)) \leq \Gamma(P_0, \mathbf{C}_l(P_0))$. Let us show the other inequality. If $\mathbf{C}_l(P_0) = \{C_1, \ldots, C_l\}$ then for any $P \in \Omega_l(\mathbf{C}_l(P_0))$ we have

$$e(\mathcal{F}, \mathbf{C}_l(P_0)) = \sum_{i=1}^{m} \min_{1 \leq j \leq l} \operatorname{dist}^2(f_i, C_j) = \Gamma(P, \mathbf{C}_l(P_0)). \tag{73}$$

In addition, $\Gamma(P, \mathbf{C}_l(P_0)) \geq \Gamma(P, \mathbf{C}_P)$, for $\mathbf{C}_P \in \mathcal{W}(P)$. Using the minimality of $\Gamma(P_0, \mathbf{C}_l(P_0))$ given in the hypothesis, we get that $\Gamma(P, \mathbf{C}_P) \geq \Gamma(P_0, \mathbf{C}_l(P_0))$, and the conclusion of the lemma follows. $\square$

Proof for Theorem 11

Proof. We shall prove that if $(P_0, \mathbf{C}_l(P_0))$ is a Γ-minimal pair, then

$$e(\mathcal{F}, \mathbf{C}_l(P_0)) \leq e(\mathcal{F}, \mathbf{C}), \forall \mathbf{C} \in \Xi.$$

For this, let us choose and arbitrary $\mathbf{C} \in \Xi$. We have that for each $P \in \Omega_l(\mathbf{C})$

$$e(\mathcal{F}, \mathbf{C}) = \Gamma(P, \mathbf{C}).$$

Clearly $\Gamma(P, \mathbf{C}) \geq \Gamma(P, \mathbf{C}_l(P))$, for each $\mathbf{C}_l(P) \in \mathcal{W}(P)$. On the other hand, using the minimality of $\Gamma(P_0, \mathbf{C}_l(P_0))$, we have

$$\Gamma(P, \mathbf{C}_l(P)) \geq \Gamma(P_0, \mathbf{C}_l(P_0)).$$

From Lemma 15 it follows that

$$e(\mathcal{F}, \mathbf{C}_l(P_0)) = \Gamma(P_0, \mathbf{C}_l(P_0)),$$

which proves

$$e(\mathcal{F}, \mathbf{C}_l(P_0)) \leq e(\mathcal{F}, \mathbf{C}).$$

This shows that if $(P_0, \mathbf{C}_l(P_0))$ is a Γ-minimal pair, then each bundle $\mathbf{C}_l(P_0)$ solves Problem 1 for the data $\mathcal{F}$. Since the total number of pairs is finite, then exist minimal pairs. This proves parts (1) and (2) of the Theorem.

For part (3) let $\mathbf{C}_0 \in \Xi$ be a solution of Problem 1, i.e., $e(\mathcal{F}, \mathbf{C}_0) \leq e(\mathcal{F}, \mathbf{C})$, for all $\mathbf{C} \in \Xi$. Consider $P_0 \in \Omega_l(\mathbf{C}_0)$ and let $\mathbf{C}_l(P_0) \in \mathcal{W}(P_0)$. Then since $P_0 \in \Omega_l(\mathbf{C}_0)$

and by the minimality of $\mathbf{C}_0$ we have

$$\Gamma(P_0, \mathbf{C}_0) = e(\mathcal{F}, \mathbf{C}_0) \leq e(\mathcal{F}, \mathbf{C}_l(P_0)) \leq \Gamma(P_0, \mathbf{C}_l(P_0)).$$

Therefore, $\Gamma(P_0, \mathbf{C}_0) \leq \Gamma(P_0, \mathbf{C}_l(P_0))$, but by definition of Γ, $\Gamma(P_0, \mathbf{C}_l(P_0)) \leq \Gamma(P_0, \mathbf{C})$, for any $\mathbf{C} \in \Xi$. So, $\Gamma(P_0, \mathbf{C}_0) = \Gamma(P_0, \mathbf{C}_l(P_0))$, and $\mathbf{C}_0 \in \mathcal{W}(P_0)$. Moreover, $(P_0, \mathbf{C}_0)$ is Γ-minimal since

$$\Gamma(P_0, \mathbf{C}_0) = e(\mathcal{F}, \mathbf{C}_0) \leq e(\mathcal{F}, \mathbf{C}_l(P)) \leq \Gamma(P, \mathbf{C}_l(P)).$$

This complete the proof of the Theorem. $\square$

IV.2.2 Auxiliary results for Section IV.1.2

Given a collection $\mathcal{C}$ of closed subspaces of $\mathcal{H}$, we associate the collection $\mathcal{P} = \{P_C\}_{C \in \mathcal{C}}$ of orthogonal projections from $\mathcal{H}$ onto C. Certainly, it is a 1-1 correspondence.

The following Problem is closed related related with Problem 1 when $X = \mathcal{H}$.

Problem 3. *Let $\mathcal{C}$ be a collection of closed subspaces in $\mathcal{H}$. Given a finite set of vectors $\mathcal{F} \subset \mathcal{H}$, and the collection $\mathcal{P}^\perp = \{P_{C^\perp}\}_{C \in \mathcal{C}}$ of orthogonal projections, find if it exists, $P_0 \in \mathcal{P}^\perp$ such that*

$$E_1(\mathcal{F}, P_0) = \inf\{E_1(\mathcal{F}, P) : P \in \mathcal{P}^\perp\}, \tag{74}$$

where $E_1(\mathcal{F}, P) = \sum_{f \in \mathcal{F}} \|Pf\|^2$.

Lemma 16. *If Problem 3 has a solution, then the collection of closed subspaces $\mathcal{C}$ satisfies the MAP.*

Proof. Let $\mathcal{C}$ be a collection of closed subspaces of $\mathcal{H}$. If $\mathcal{F}$ is a finite subset of $\mathcal{H}$, then

$$\begin{aligned} E(\mathcal{F}, C) &= \sum_{f \in \mathcal{F}} \operatorname{dist}^2(f, C) \\ &= \sum_{f \in \mathcal{F}} \inf\{\|f - g\|^2 : g \in C\} = \sum_{f \in \mathcal{F}} \|f - P_C f\|^2 \\ &= \sum_{f \in \mathcal{F}} \|P_{C^\perp} f\|^2 \end{aligned}$$

If we consider $\mathcal{F}$ and $\mathcal{P}^\perp = \{P_{C^\perp}\}_{C \in \mathcal{C}}$, then taking into account that Problem 3 has a solution, there exists $P_0 \in \mathcal{P}^\perp$ such that

$$E_1(\mathcal{F}, P_0) = \inf\{E_1(\mathcal{F}, P) : P \in \mathcal{P}^\perp\}.$$

By setting $C_0 = (P_0(\mathcal{H}))^\perp$, then $C_0 \in \mathcal{C}$, and we have

$$
\begin{aligned}
E(\mathcal{F}, C_0) &= \sum_{f\in\mathcal{F}} \|f - P_{C_0} f\|^2 = \sum_{f\in\mathcal{F}} \|P_{C_0^\perp} f\|^2 \\
&= E_1(\mathcal{F}, P_0) \\
&\le E_1(\mathcal{F}, P) = \sum_{f\in\mathcal{F}} \|P_{C^\perp} f\|^2 \\
&= \sum_{f\in\mathcal{F}} \|f - P_C f\|^2 = E(\mathcal{F}, C), \forall C \in \mathcal{C}.
\end{aligned}
$$

Now the conclusion follows. □

Lemma 17. *Let $f \in \mathcal{H}$ be given, and $\mathcal{P}$ a collection of orthogonal projections which is closed with respect to(w.r.t.) the W.O.T. on $\mathcal{B}(\mathcal{H})$. Define $E_f : \mathcal{P} \to \mathbb{R}_+$ by $E_f(P) =< Pf, f >$. Then E_f is continuous.*

Proof. Let $\{P_n\}_{n\ge1}$ be a sequence in $\mathcal{P}$ such that $P_n \to P$ in the weak sense. Clearly, since $\mathcal{P}$ is closed w.r.t the W.O.T., then $P \in \mathcal{P}$, and we have

$$
|E_f(P_n) - E_f(P)| = | < P_n f, f > - < Pf, f > | \to 0.
$$

Hence, E_f is continuous. □

Lemma 18. *Let $\mathcal{C}$ be a collection of closed subspaces of a Hilbert space $\mathcal{H}$. Then the collection $\mathcal{P} = \{P_C\}_{C\in\mathcal{C}}$ of orthogonal projections is closed w.r.t. the W.O.T. if and only if the collection of orthogonal projections $\mathcal{P}^\perp = \{P_{C^\perp}\}_{C\in\mathcal{C}}$ is closed w.r.t. the W.O.T.*

Proof. It suffices to show only one implication. Assume that $\mathcal{P} = \{P_C\}_{C\in\mathcal{C}}$ is closed w.r.t the W.O.T. Let $\{P_n\}_{n\ge1} \subset \mathcal{P}^\perp$ be such that $P_n \to P$ weakly. We want to show that $P \in \mathcal{P}^\perp$, that is, there exists $C \in \mathcal{C}$ such that $P = P_{C^\perp}$. Taking into account that $P_n = P_{C_n} \to P$ weakly, then $P_{C_n^\perp} \to I - P$ weakly because

$$
\begin{aligned}
\lim_{n\to\infty} < (P_{C_n^\perp} - (I - P))f, g > &= \lim_{n\to\infty} < (I - P_{C_n} - I + P)f, g > \\
&= -\lim_{n\to\infty} < (P_{C_n} - P)f, g > \\
&= 0, \forall f, g \in \mathcal{H}.
\end{aligned}
$$

Since $\mathcal{P}$ is closed w.r.t. the W.O.T., and $P_{C_n^\perp} \to I - P$, then there exists $C_\star \in \mathcal{C}$ such that $I - P = P_{C_\star}$. Hence, $P = I - P_{C_\star} = P_{C_\star^\perp}$. □

Proposition 5. *Let $\mathcal{C}$ be a collection of closed subspaces in $\mathcal{H}$, and $\mathcal{F}$ a finite subset of $\mathcal{H}$. If the collection of orthogonal projections $\mathcal{P}^\perp = \{P_{C^\perp}\}_{C\in\mathcal{C}}$ is closed w.r.t. the W.O.T. Then Problem 3 has a solution, that is, there exists $P_0 \in \mathcal{P}^\perp$ such that*

$$
E_1(\mathcal{F}, P_0) = \inf\{E_1(\mathcal{F}, P) : P \in \mathcal{P}^\perp\}.
$$

Proof. Taking into account that $\mathcal{P}^\perp$ is weakly closed and E_1 satisfies

$$E_1(\mathcal{F}, P) = \sum_{f \in \mathcal{F}} \|Pf\|^2 = \sum_{f \in \mathcal{F}} < Pf, f >$$
$$= \sum_{f \in \mathfrak{F}} E_f(P),$$

then from Lemma 17 it follows that E_1 is a continuous function on $\mathcal{P}^\perp$. On the other hand, since $\|P\|_{op} \leq 1$ for all $P \in \mathcal{P}^\perp$, and the unit ball in $\mathcal{B}(\mathcal{H})$ is compact w.r.t. the W.O.T., it follows that $\mathcal{P}^\perp$ is also compact w.r.t. the W.O.T. Now the conclusion is a straightforward consequence of the Extreme Value Theorem. □

IV.2.3 Proofs for Section IV.1.2

Proof of Theorem 12

Proof. Taking into account that $\mathcal{P}$ is weakly closed, it follows from Lemma 18 that $\mathcal{P}^\perp = \{P_{C^\perp}\}_{C \in \mathcal{C}}$ is also weakly closed. Thus, by applying Proposition 5 we have that Problem 3 has a solution, and from Lemma 16 we conclude that $\mathcal{C}$ satisfies the MAP. □

In order to prove Corollary 6, we need the following result.

Proposition 6. *Let $\mathcal{H} = \mathbb{C}^d$, $1 \leq n \leq d$, and $\mathcal{P}_n$ be the collection of orthogonal projections given by*

$$\mathcal{P}_n = \{P : \mathbb{C}^d \to \mathbb{C}^d : P^* = P = P^2, \dim P(\mathcal{H}) \leq n\}. \tag{75}$$

Then $\mathcal{P}_n$ is closed w.r.t. the W.O.T. on $\mathcal{B}(\mathbb{C}^d)$.

Proof. Let $\{P_k\}_{k \geq 1}$ be a sequence in $\mathcal{P}_n$ such that $P_k \to P$ weakly, we want to show: $P^* = P$, $P^2 = P$, and $\dim P(\mathcal{H}) \leq n$. Let $x, y \in \mathcal{H}$ be given. Then

$$< x, P^*y > = < Px, y > = \lim_k < P_kx, y > = \lim_k < x, P_k^*y >$$
$$= \lim_k < x, P_ky > = \lim_k \overline{< P_ky, x >} = \overline{\lim_k < P_ky, x >}$$
$$= \overline{< Py, x >} = < x, Py > .$$

So, $P^* = P$. Since $\mathcal{H} = \mathbb{C}^d$, then W.O.T.=S.O.T.=N.O.T. Therefore, $\|P_k - P\|_{op} \to 0$. Taking into account that

$$\begin{aligned}
|\|P_k^2 - P_k\|_{op} - \|P^2 - P\|_{op}| &\leq \|(P_k^2 - P_k) - (P^2 - P)\|_{op} \\
&\leq \|P_k^2 - P_k P\|_{op} + \|P_k P - P^2\|_{op} + \\
&\quad + \|P_k - P\|_{op} \\
&\leq \|P_k\|_{op}\|P_k - P\|_{op} + \|P\|_{op}\|P_k - P\|_{op} + \\
&\quad + \|P_k - P\|_{op} \\
&\leq (2 + \|P\|_{op})\|P_k - P\|_{op} \to 0,
\end{aligned}$$

then $\|P^2 - P\|_{op} = \lim_k \|P_k^2 - P_k\|_{op} = \lim_k \|P_k - P_k\|_{op} = 0$. Since $\|P^2 - P\|_{op} = 0$, then $P^2 = P$. Finally, let us show that $\dim P(\mathcal{H}) \leq n$. Assume for the sake of a contradiction that $\dim P(\mathcal{H}) > n$. We may assume that $\dim P(\mathcal{H}) = n + 1$. Since $\mathcal{H} = \mathbb{C}^d$, then

$$\mathbf{P_k} = \begin{pmatrix} q_{11}^k \dots q_{1d}^k \\ \vdots \qquad \vdots \\ q_{d1}^k \dots q_{dd}^k \end{pmatrix},$$

and

$$\mathbf{P} = \begin{pmatrix} q_{11} \dots q_{1d} \\ \vdots \qquad \vdots \\ q_{d1} \dots q_{dd} \end{pmatrix}.$$

Since $P_k \to P$ weakly, and on $\mathcal{B}(\mathcal{H})$ we have that W.O.T.=S.O.T., then

$$\lim_{k\to\infty} q_{ij}^k = q_{ij}, \forall 1 \leq i, j \leq d. \tag{76}$$

On the other hand, if $\dim P(\mathcal{H}) = n + 1$, then exists a $(n + 1) \times (n + 1)$ minor of P which is nonzero. From (76) and the continuity of the determinant, there exists $k_0 \in \mathbb{N}$ such that for each $k \geq k_0$, there is a $(n + 1) \times (n + 1)$ minor of P_k which is nonzero, contradicting the fact that $\dim P_k(\mathcal{H}) \leq n$, for all $k \geq 1$. □

Proof of Corollary 6

Proof. Let $1 \leq n \leq d$ be given. Note that the collection of orthogonal projections $\mathcal{P}_n = \{P_C\}_{C \in \mathcal{L}_n}$ is given in (75). By Proposition 6 we know that $\mathcal{P}_n$ is closed w.r.t. the W.O.T. on $\mathcal{B}(\mathbb{C}^d)$, and by applying Theorem 12, we have that $\mathcal{L}_n$ satisfies the MAP. □

IV.2.4 Auxiliary results for Section IV.1.3

Now we shall state and prove some auxiliary results that allow us to obtain the main results of Section IV.1.3.

Lemma 19. *Let $\{A_n\}$ be a sequence of unitary operators such that $A_n \to A$ weakly. If A is unitary, then $A_n \to A$ strongly.*

Proof. Let $x \in \mathcal{H}$ be given. Then we have

$$\begin{aligned}\|A_n x - Ax\|^2 &= < (A_n - A)x, (A_n - A)x > \\ &= < x, x > + < Ax, Ax > - < A_n x, Ax > - < Ax, A_n x > .\end{aligned}$$

Since $A_n \to A$ weakly, and A is unitary, then

$$\lim_{n\to\infty} < A_n x, Ax > = < Ax, Ax > = < x, x >,$$

and

$$\lim_{n\to\infty} < Ax, A_n x > = < Ax, Ax > = < x, x > .$$

Now the conclusion follows. $\square$

Lemma 20. *Let $\mathcal{S}$ be a closed subspace of the Hilbert space $\mathcal{H}$, and A a unitary operator defined on $\mathcal{H}$. Then*

$$\|P_{A\mathcal{S}}(x)\| = \|P_{\mathcal{S}}(A^* x)\|, \textit{for all } x \in \mathcal{H}.$$

Proof. Let $x \in \mathcal{H}$ be given. If $\{e_m\}_{m\geq 1}$ is an orthonormal basis(ONB) for $\mathcal{S}$, then by using that A is unitary it follows that $\{Ae_m\}_{m\geq 1}$ is an ONB for $A\mathcal{S}$. Therefore,

$$P_{A\mathcal{S}}(x) = \sum_{m\geq 1} < x, Ae_m > Ae_m,$$

and

$$P_{\mathcal{S}}(A^* x) = \sum_{m\geq 1} < A^* x, e_m > e_m.$$

Consequently,

$$\begin{aligned}\|P_{A\mathcal{S}}(x)\|^2 &= < P_{A\mathcal{S}}(x), x > \\ &= < \sum_{m\geq 1} < x, Ae_m > Ae_m, x > \\ &= \sum_{m\geq 1} | < x, Ae_m > |^2 = \sum_{m\geq 1} | < A^* x, e_m > |^2 \\ &= \|P_{\mathcal{S}}(A^* x)\|^2.\end{aligned}$$

So,

$$\|P_{A\mathcal{S}}(x)\| = \|P_{\mathcal{S}}(A^* x)\|,$$

and we are done. $\square$

Proposition 7. *Let $\mathcal{U}$ be a collection of unitary operators which is closed w.r.t. the W.O.T. on $\mathcal{B}(\mathcal{H})$, and $\mathcal{S}$ a fixed and given closed subspace of $\mathcal{H}$. Let $f \in \mathcal{H}$ be given, and define $L_f : \mathcal{U} \to \mathbb{R}_+$ by $L_f(U) = \mathrm{dist}^2(f, U\mathcal{S})$. Then L_f is continuous on $\mathcal{U}$.*

Proof. Taking into account that

$$L_f(U) = \operatorname{dist}^2(f, U\mathcal{S}) = \|f\|^2 - < P_{U\mathcal{S}}(f), f >,$$

it suffices to show that the function $M_f : \mathcal{U} \to \mathbb{R}_+$ defined by $M_f(U) =< P_{U\mathcal{S}}(f), f >$ is continuous on $\mathcal{U}$. Let $\{U_n\}_{n\geq 1} \subset \mathcal{U}$ be such that $U_n \to U$ weakly. Because $\mathcal{U}$ is weakly closed, then $U \in \mathcal{U}$. Moreover, $U_n^* \to U^*$ weakly, and from Lemma 19 we get $U_n^* \to U^*$ strongly. Therefore, with the aid of Lemma 20 we have

$$\begin{aligned}
|M_f(U_n) - M_f(U)| &= |\|P_{U_n\mathcal{S}}(f)\|^2 - \|P_{U\mathcal{S}}(f)\|^2| \\
&= |\|P_{\mathcal{S}}(U_n^* f)\|^2 - \|P_{\mathcal{S}}(U^* f)\|^2| \\
&\leq 2\|f\|\|P_{\mathcal{S}}(U_n^* f) - P_{\mathcal{S}}(U^* f)\| \\
&= 2\|f\|\|P_{\mathcal{S}}(U_n^* f - U^* f)\| \\
&\leq 2\|f\|\|U_n^* f - U^* f\| \to 0, \text{ as } n \to \infty.
\end{aligned}$$

Hence, M_f is continuous on $\mathcal{U}$, and we are done. ☐

In order to prove Theorem 13, we separately have to consider the cases when $\mathcal{S}$ is a bounded convex subset of $\mathbb{C}^d$ and when $\mathcal{S}$ is a closed convex subset of a separable Hilbert space $\mathcal{H}$.

Case 1: $\mathcal{S}$ is a bounded convex subset of $\mathbb{C}^d$.

We shall state without proof the Reduction Principle (see Theorem 5.14 on page 80 in [44]).

Lemma 21. *Let K be a convex subset of a Hilbert space $\mathcal{H}$, and let M be any closed subspace of $\mathcal{H}$ that contains K. Then:*

$$\operatorname{dist}^2(x, K) = \operatorname{dist}^2(x, M) + \operatorname{dist}^2(P_M(x), K), \tag{77}$$

for every $x \in \mathcal{H}$.

Proposition 8. *Assume that $\mathcal{S}$ is a fixed and given bounded convex subset of $\mathbb{C}^d$. Let $\mathcal{U}$ be a collection of unitary operators which is closed w.r.t. the W.O.T. on $\mathcal{B}(\mathbb{C}^d)$. Let $f \in \mathcal{H} = \mathbb{C}^d$ be given. If $\Theta_f : \mathcal{U} \to \mathbb{R}_+$ is defined by $\Theta_f(U) = \operatorname{dist}^2(f, U\mathcal{S})$, then Θ_f is continuous on $\mathcal{U}$.*

Proof. Taking into account that $\mathcal{S}$ is convex, and $\mathcal{S} \subset \Lambda := span(S) = \overline{span(S)}$, because $\mathcal{H} = \mathbb{C}^d$, then by using (77) in Lemma 21 we have

$$\operatorname{dist}^2(f, U\mathcal{S}) = \operatorname{dist}^2(f, U\Lambda) + \operatorname{dist}^2(P_{U\Lambda}(f), U\mathcal{S}),$$

for all $U \in \mathcal{U}$. From Proposition 7 we know that $\operatorname{dist}^2(f, U\Lambda)$ is a continuous function of U on $\mathcal{U}$. Consequently, we only need to show that the function $\Psi_f : \mathcal{U} \to \mathbb{R}_+$ defined by $\Psi_f(U) = \operatorname{dist}^2(P_{U\Lambda}(f), U\mathcal{S})$ is continuous on $\mathcal{U}$. Let $\{U_n\}_{n\geq 1} \subset \mathcal{U}$ be such that $U_n \to U$ weakly. Since $\mathcal{U}$ is closed w.r.t. the W.O.T. on $\mathcal{B}(\mathbb{C}^d)$, then $U \in \mathcal{U}$, and S.O.T.=W.O.T.=N.O.T. Let $\{e_i\}_{i=1}^J$ be an ONB for Λ, where $J = \dim \Lambda$, then

$\{Ue_i\}_{i=1}^J$ and $\{U_ne_i\}_{i=1}^J$ are ONB for $U\Lambda$ and $U_n\Lambda$ respectively. Therefore,

$$P_{U\Lambda}(f) = \sum_{i=1}^{J} < f, Ue_i > Ue_i,$$

$$P_{U_n\Lambda}(f) = \sum_{i=1}^{J} < f, U_ne_i > U_ne_i, \forall n \geq 1,$$

and since $U_n \to U$ weakly on $\mathcal{B}(\mathbb{C}^d)$, then

$$\lim_{n\to\infty} < U_ne_i, f >=< Ue_i, f >, \forall 1 \leq i \leq J,$$

and

$$\lim_{n\to\infty} \|U_ne_i - Ue_i\| = 0, \forall 1 \leq i \leq J.$$

Hence, $\lim_{n\to\infty} \|P_{U_n\Lambda}(f) - P_{U\Lambda}(f)\| = 0$. On the other hand, since $\lim_{n\to\infty} \|U_n - U\|_{op} = 0$, and $\mathcal{S}$ is a bounded set, then for $h \in \mathcal{S}$ we have

$$\lim_{n\to\infty} \|P_{U_n\Lambda}(f) - U_nh\| = \|P_{U\Lambda}(f) - Uh\|, \tag{78}$$

uniformly on $\mathcal{S}$. Let $\epsilon > 0$ be given. From (78) it follows that there exists $N_0 = N_0(\epsilon) \in \mathbb{N}$ such that for $n \geq N_0$, we have

$$\|P_{U\Lambda}(f) - Uh\|^2 - \epsilon \leq \|P_{U_n\Lambda}(f) - U_nh\|^2 \leq \|P_{U\Lambda}(f) - Uh\|^2 + \epsilon,$$

for all $h \in \mathcal{S}$. Taking the infimum in the above inequality over all $h \in \mathcal{S}$ yields

$$\text{dist}^2(P_{U\Lambda}(f), U\mathcal{S}) - \epsilon \leq \text{dist}^2(P_{U_n\Lambda}(f), U_n\mathcal{S}) \leq \text{dist}^2(P_{U\Lambda}(f), U\mathcal{S}) + \epsilon,$$

for all $n \geq N_0$. Thus, $\lim_{n\to\infty} \text{dist}^2(P_{U_n\Lambda}(f), U_n\mathcal{S}) = \text{dist}^2(P_{U\Lambda}(f), U\mathcal{S})$, Ψ_f is continuous on $\mathcal{U}$, and the conclusion follows. □

Case 2: $\mathcal{S}$ is a Closed Convex Subset.
We will denote the unit sphere of $\mathcal{H}$ by $S_1(\mathcal{H})$, that is, $S_1(\mathcal{H}) = \{z \in \mathcal{H} : \|z\| = 1\}$. The following Lemma will be stated without proof (see Theorem 7.1 in [44]).

Lemma 22. *Let K be a closed convex set in the Hilbert space $\mathcal{H}$, and $x \in \mathcal{H}$. Then*

$$\text{dist}(x, K) = \begin{cases} 0 & \textit{if } x \in K \\ \max\{< x, z > - \sup_{y\in K} < y, z > | z \in S_1(\mathcal{H})\} & \textit{if } x \notin K. \end{cases}$$

Moreover, the maximum is attained for a unique $z_0 \in S_1(\mathcal{H})$.

Lemma 23. *Let K be a closed convex set in the Hilbert space $\mathcal{H}$, and U a unitary operator defined on $\mathcal{H}$. Then for $x \in \mathcal{H}$ we have*

$$\text{dist}(x, UK) = \text{dist}(U^*x, K). \tag{79}$$

Proof. Note that by Lemma 22 we have

$$\operatorname{dist}(U^*x, K) = \begin{cases} 0 & \text{if } U^*x \in K \\ \max\{< U^*x, z > - \sup_{y\in K} < y, z > |z \in S_1(\mathcal{H})\} & \text{if } U^*x \notin K, \end{cases}$$

and

$$\operatorname{dist}(x, UK) = \begin{cases} 0 & \text{if } x \in UK \\ \max\{< x, z > - \sup_{y\in K} < Uy, z > |z \in S_1(\mathcal{H})\} & \text{if } x \notin UK. \end{cases}$$

Taking into account that U^* is unitary, and $z \in S_1(\mathcal{H})$ then $w = U^*z \in S_1(\mathcal{H})$, $< x, z >=< U^*x, U^*z >$, and $x \in UK$ if and only if $U^*x \in K$. Consequently,

$$\operatorname{dist}(x, UK) = \begin{cases} 0 & \text{if } U^*x \in K \\ \max\{< U^*x, w > - \sup_{y\in K} < y, w > |w \in S_1(\mathcal{H})\} & \text{if } U^*x \notin K. \end{cases}$$

So, $\operatorname{dist}(x, UK) = \operatorname{dist}(U^*x, K)$. ☐

Proposition 9. *Let $\mathcal{U}$ be a collection of unitary operators which is closed w.r.t. the W.O.T. on $\mathcal{B}(\mathcal{H})$. Let $\mathcal{S}$ be a fixed and given closed convex subset of $\mathcal{H}$. Let $f \in \mathcal{H}$ be given, and define $\Upsilon_f : \mathcal{U} \to \mathbb{R}_+$ by $\Upsilon_f(U) = \operatorname{dist}(f, U\mathcal{S})$. Then Υ_f is continuous on $\mathcal{U}$.*

Proof. Note that from (79) we have $\Upsilon_f(U) = \operatorname{dist}(U^*f, \mathcal{S})$. So,

$$\Upsilon_f(U) = \begin{cases} 0 & \text{if } U^*f \in \mathcal{S} \\ \max\{< U^*f, z > - \sup_{y\in\mathcal{S}} < y, z > |z \in S_1(\mathcal{H})\} & \text{if } U^*f \notin \mathcal{S}, \end{cases}$$

for all $U \in \mathcal{U}$. Let $\{U_n\}_{n\geq 1} \subset \mathcal{U}$ be such that $U_n \to U$ weakly. Since $\mathcal{U}$ is weakly closed, then $U \in \mathcal{U}$. Moreover, $U_n^* \to U^*$ weakly, and from Lemma 19, $U_n^* \to U^*$ strongly, and since $\|z\| = 1$ for all $z \in S_1(\mathcal{H})$, it follows that

$$\lim_{n\to\infty} < U_n^*f, z >=< U^*f, z >, \tag{80}$$

uniformly on $S_1(\mathcal{H})$. Let $\epsilon > 0$ be given. From (80), there exists $N_1 = N_1(\epsilon) \in \mathbb{N}$ such that for $n \geq N_1$ we have

$$< U^*f, z > -\epsilon \quad < \quad < U_n^*f, z > \quad < \quad < U^*f, z > +\epsilon,$$

for all $z \in S_1(\mathcal{H})$. Consequently, for $n \geq N_1$ we have

$$\begin{aligned} < U^*f, z > - \sup_{y\in\mathcal{S}} < y, z > -\epsilon << U_n^*f, z > - \sup_{y\in\mathcal{S}} < y, z > \\ << U^*f, z > - \sup_{y\in\mathcal{S}} < y, z > +\epsilon. \end{aligned}$$

Taking the maximum in the above inequality over all $z \in S_1(\mathcal{H})$ yields

$$\max\{< U^*f, z > - \sup_{y\in\mathcal{S}} < y, z > |z \in S_1(\mathcal{H})\} - \epsilon$$

$$< \max\{< U_n^* f, z > - \sup_{y \in \mathcal{S}} < y, z > | z \in S_1(\mathcal{H})\}$$

$$< \max\{< U^* f, z > - \sup_{y \in \mathcal{S}} < y, z > | z \in S_1(\mathcal{H})\} + \epsilon,$$

for all $n \geq N_1$. Hence, $\lim_{n\to\infty} \max\{< U_n^* f, z > - \sup_{y\in\mathcal{S}} < y, z > | z \in S_1(\mathcal{H})\} = \max\{< U^* f, z > - \sup_{y\in\mathcal{S}} < y, z > | z \in S_1(\mathcal{H})\}$. Consequently,

$$\lim_{n\to\infty} \Upsilon_f(U_n) = \Upsilon_f(U),$$

Υ_f is continuous, and we are done. □

IV.2.5 Proofs for Section IV.1.3

Proof of Theorem 13

Proof. Let $\mathcal{F}$ be a given finite subset of $\mathcal{H}$. Taking into account that

$$\widehat{H}(\mathcal{F}, U) = \begin{cases} \sum_{f\in\mathcal{F}} \Theta_f(U) \text{ if } \mathcal{S} \subset \mathbb{C}^d \text{ is bounded} \\ \sum_{f\in\mathcal{F}} \Upsilon_f^2(U) \text{ if } \mathcal{S} \text{ is a closed convex set,} \end{cases}$$

then the continuity of $\widehat{H}$ on $\mathcal{U}$ follows by applying Proposition 8 if $\mathcal{S}$ is a bounded convex subset of $\mathbb{C}^d$, and Proposition 9 if $\mathcal{S}$ is a closed convex set. On the other hand, since $\|U\|_{op} = 1$, for all $U \in \mathcal{U}$, and since the unit ball of $\mathcal{B}(\mathcal{H})$ is weakly compact, then $\mathcal{U}$ is also weakly compact because it is weakly closed. Now the conclusion follows by applying the Extreme Value Theorem. □

Proof of Corollary 7

Proof. It is an easy exercise to verify that if $\mathcal{S} \subset \mathcal{H}$ is either a closed subspace of $\mathcal{H}$, a closed convex cone, a hyperplane, a half-space, or a closed affine subspace, then $\mathcal{S}$ is a closed convex subset of $\mathcal{H}$. Now the proof is a straightforward consequence Theorem 13. □

Proof of Corollary 8

Proof. Let $\mathcal{F}$ be a given finite subset of $\mathcal{H}$. Note that if $C \in \mathcal{C}_{\mathcal{S}}$, then $C = U\mathcal{S}$, for some $U \in \mathcal{U}$. Taking into account that

$$\begin{aligned} E(\mathcal{F}, C) &= E(\mathcal{F}, U\mathcal{S}) \\ &= \sum_{f\in\mathcal{F}} \operatorname{dist}^2(f, U\mathcal{S}) \\ &= \widehat{H}(\mathcal{F}, U), \end{aligned}$$

then since $\mathcal{U}$ is weakly closed, from Corollary 7 it follows that Problem 2 has a solution, that is, there exists $U_0 \in \mathcal{U}$ such that

$$\widehat{H}(\mathcal{F}, U_0) = \inf\{\widehat{H}(\mathcal{F}, U) : U \in \mathcal{U}\}.$$

Consequently, by setting $C_0 = U_0\mathcal{S}$ we have that $E(\mathcal{F}, C)$ attains its minimum on $\mathcal{C}_{\mathcal{S}}$ at C_0. Therefore, $\mathcal{C}_{\mathcal{S}}$ has the MAP. $\square$

As we said in Remark 24 at the end of Section IV.1.3, we shall show that Corollary 6 can be also obtained as a consequence of Corollary 8. In order to show that, let us first state and prove some auxiliary results.

For $1 \leq n \leq d$, we shall denote by $\mathfrak{D}_n$ the collection of all subspaces of $\mathbb{C}^d$ with dimension equal to n.

Lemma 24. *Let $1 \leq n \leq d$ be given. Then $\mathcal{D}_n$ satisfies the MAP.*

Proof. Let us pick $D_0 \in \mathcal{D}_n$, and let $\mathcal{U}_{\mathbb{C}^d}$ be the set of all unitary operators on $\mathbb{C}^d$. From Lemma 2.1.8 on page 69 in [63] it follows that $\mathcal{U}_{\mathbb{C}^d}$ is weakly closed. Taking into account that $\mathcal{D}_n = \{UD_0\}_{U\in\mathcal{U}_{\mathbb{C}^d}}$, then the conclusion follows by applying Corollary 8. $\square$

Lemma 25. *Let $1 \leq n \leq d$, and $f \in \mathbb{C}^d$ be given. If $C_1 \in \mathcal{L}_n$ is such that $\dim C_1 < n$, then there exists $C \in \mathcal{D}_n$ such that $\operatorname{dist}^2(f, C) \leq \operatorname{dist}^2(f, C_1)$. Therefore, if $\mathcal{F}$ is a finite subset of $\mathbb{C}^d$, then $E(\mathcal{F}, C) \leq E(\mathcal{F}, C_1)$.*

Proof. Let $C_1 \in \mathcal{L}_n$ be such that $\dim C_1 < n$. Let us consider the orthonormal vectors $\{e_i\}_{i=1}^{n-\dim C_1}$ in $C_1^{\perp}$, and let $M = span\{e_1, \ldots, e_{n-\dim C_1}\}$. Set $C := C_1 \oplus M$. Then $C \in \mathcal{D}_n$, $P_C = P_{C_1} + P_M$, and

$$\begin{aligned}\operatorname{dist}^2(f, C) &= \|f\|^2 - < P_C(f), f > \\ &= \|f\|^2 - < P_{C_1}(f), f > - < P_M(f), f > \\ &\leq \|f\|^2 - < P_{C_1}(f), f >= \operatorname{dist}^2(f, C_1).\end{aligned}$$

$\square$

Proposition 10. *Let $\mathcal{F}$ be a finite set of vectors in $\mathbb{C}^d$, and $1 \leq n \leq d$. Then*

$$\inf\{E(\mathcal{F}, C) : C \in \mathcal{L}_n\} = \inf\{E(\mathcal{F}, C) : C \in \mathcal{D}_n\}. \tag{81}$$

Proof. Let $\mathcal{F}$ be a finite subset of $\mathbb{C}^d$, and $1 \leq n \leq d$. Taking into account that $\mathcal{D}_n \subset \mathcal{L}_n$, then

$$\inf\{E(\mathcal{F}, C) : C \in \mathcal{L}_n\} \leq \inf\{E(\mathcal{F}, C) : C \in \mathcal{D}_n\}. \tag{82}$$

On the other hand, if $C \in \mathcal{L}_n$, then, by Lemma 25, there exists $C_1 \in \mathcal{D}_n$ such that

$$E(\mathcal{F}, C_1) \leq E(\mathcal{F}, C).$$

Since $\inf\{E(\mathcal{F}, C) : C \in \mathcal{D}_n\} \leq E(\mathcal{F}, C_1)$, then

$$\inf\{E(\mathcal{F}, C) : C \in \mathcal{D}_n\} \leq E(\mathcal{F}, C).$$

Taking the infimum in the above inequality over all $C \in \mathcal{L}_n$ yields

$$\inf\{E(\mathcal{F}, C) : C \in \mathcal{D}_n\} \leq \inf\{E(\mathcal{F}, C) : C \in \mathcal{L}_n\}. \tag{83}$$

From (82) and (83) we get (81). □

Another proof of Corollary 6.

Proof. The proof of Corollary 6 is a straightforward consequence of Proposition 10 and Lemma 24. □

BIBLIOGRAPHY

[1] M. Aharon, M. Elad, A.M. Bruckstein, *The k-svd: An algorith for designing of overcomplete dictionaries for sparse representation*, IEEE Trans. Signal Process 54(11), 2006, 4311-4322.

[2] M. Aharon, M. Elad, A.M. Bruckstein,, *On the uniqueness of overcomplete dictionaries, and a practical way to retrieve them*, Linear Algebra Appl. 416(1), 2006, 48-67.

[3] A. Aldroubi, *Portraits of frames*, Proc. Amer. Math. Soc., 123, 1995, 1661-1668.

[4] A. Aldroubi, *Non-uniform average sampling and reconstrusction in shift-invatiant and wavelet sapaces*, Appl. Comput. Harmon. Anal. 13, 2002, 151-161.

[5] A. Aldroubi, A. Baskakov, and I. Krishtal, Slanted matrices and their various applications, *submitted.*

[6] A. Adroubi, C. Cabrelli, D. Hardin, U. Molter, *Optimal shift-invariant spaces and their parseval frame generators*, Appl. Comput. Harmon. Anal. 23, 2007, 273-283.

[7] A. Adroubi, C. Cabrelli, U. Molter, *Optimal non-linear models for sparsity and sampling*, J. Fourier Anal., Appl. 14, 2008, 793-812.

[8] A. Aldroubi, H. Feichtinger, *Exact iterative reconstruction algorithm for multivariate irregularly sampled functions in spline-like spaces: The L^p theory*, Proc. Amer. Math. Soc., 126, 1998, 2677-2686.

[9] A. Aldroubi, K. Gröchenig, *Beurling-Landau-type theorems for non-uniform sampling in shift invariant spline spaces*, J. Fourier Anal. Appl. 6, 2000, 93-103.

[10] A. Aldroubi, K. Gröchenig, *Nonuniform sampling and reconstruction in shift-invariant spaces*, SIAM Rev. 43, 2001, 585-620.

[11] A. Aldroubi, I. Krishtal, *Robustness of sampling and reconstruction and Beurling-Landau-type theorems for shift-invariant spaces*, Appl. Comput. Harmon. Anal. 20, 2006, 250-260.

[12] A. Aldroubi, C. Leonetti, *Non-uniform sampling and reconstruction from sampling sets with unknown jitter*, to appear in STSIP.

[13] A. Aldroubi, Q. Sun, W.S. Tang, *p-frames and shift-invariant spaces of L^p*, J. Fourier Anal. Appl. 7, 2001, 1-21.

[14] A. Aldroubi, M. Unser, *Families of wavelet transforms in connection with Shannon's sampling theory and the Gabor transform*, in Wavelets: A Tutorial in Theory and Applications, C.K. Chui, ed., Academic Press, San Diego, CA, 1992,

509-528.

[15] A. Aldroubi, M. Unser, *Sampling procedure in function spaces and asymptotic equivalence with Shannon's sampling theory*, Numer. Funct. Anal. Optim. 15, 1994, 1-21.

[16] A. Aldroubi, M. Unser, M. Eden, *Asymptotic properties of least square spline filters and application to multi-scale decomposition of signals*, in Proceedings of International Conference on Information Theory and Its Applications, Waikiki, Hawaii, 1990, 271-274.

[17] A. Aldroubi, M. Unser, M. Eden, *Cardinal spline filters: Stability and convergence to the ideal sinc interpolator*, Signal Processing, 28, 1992, 127-138.

[18] N. Atreas, N. Bagis, and C. Karanikas, The information loss error and the jitter error for regular sampling expansions, *Sampl. Theory Signal Image Process.* **1** (2002), no. 3, 261–276.

[19] R. Balan, *Equivalence relations and distances between Hilbert frames*, Proc. Amer. Math. Soc., 127, 1999, 2353-2366.

[20] A.G. Baskakov, Asymptotic estimates for elements of matrices of inverse operators, and harmonic analysis, (Russian) *Sibirsk. Mat. Zh.*, 38, 1997, no. 1, 14–28; translation in *Siberian Math. J.*, 38, 1997, no. 1, 10–22.

[21] E. Beller, G. de Haan, *New algorithms for motion estimation on interlaced video*, in Proc. SPIE-Visual Communication and Image Processing, 3309, 1998, 111-121.

[22] J.J. Benedetto, *Irregular sampling and frames*, in Wavelets: A Tutorial in Theory and Applications, C.K. Chui, ed., Academic Press, San Diego, CA, 1992, 445-507.

[23] J.J. Benedetto, *Frame decomposition, sampling, and uncertainty principle inequalities*, in Wavelets-Mathematics and Applications, J.J. Benedetto and M.W. Frazier, eds., CRC, Boca Raton, FL, 1993, 247-304.

[24] J.J. Benedetto and P. J. S. G. Ferreira, *Modern Sampling Theory*, Appl. Numer. Harmon. Anal., Birkhäuser Boston, Boston, MA, 2001.

[25] J.J. Benedetto, C. Heil, D.F. Walnut, *Gabor systems and the Balian-Low theorem*, in Gabor Analysis and Algorithms, H.G. Feichtinger and T. Strohmer, eds., Birkhäuser, Boston, 1998, 85-122.

[26] J.J. Benedetto, S. Scott, Frames, irregular sampling, and a wavelet auditory model, *Nonuniform Sampling*, 585–617, Inf. Technol. Transm. Process. Storage, Kluwer/Plenum, New York, 2001.

[27] J.J. Benedetto, D.F. Walnut, *Gabor frames for L^2 and related spaces*, in Wavelets-Mathematics and Applications, J.J. Benedetto and M. W. Frazier, eds., CRC, Boca Raton, FL, 1993, 247-304.

[28] C.A. Berenstein, E.V. Patrick, *Exact deconvolution operators-an overview, plus performance characterizations for imaging sensors*, in Proceedings in Multidimensional Signal Processing, IEEE, Piscataway, NJ, 1990, 723-734.

[29] A. Beurling, *The Collected Works of Arne Beurling. Vol.1, Complex Analysis*, L. Carleson, P. Malliavin, J. Neuberger, and J. Wermer, eds., Birhuser, Boston, 1989.

[30] A. Beurling, *The Collected Works of Arne Beurling. Vol.2, Harmonic Analysis*, L. Carleson, P. Malliavin, J. Neuberger, and J. Wermer, eds., Birhuser, Boston, 1989.

[31] A. Beurling, P. Malliavin, *On the closure of characters and the zeros of entire functions*, Acta Math., 118, 1967, 79-93.

[32] T. Blu, M. Unser, *Quantitative Fourier analysis of approximation tecniques: Part I-interpolators and projectors*, IEEE Trans. Signal Process., 47, 1999, 2783-2795.

[33] P.L. Butzer, *A survey of the Whittaker-Shannon sampling theorem and some of its extensions*, J. Math. Res. Exposition, 3, 1983, 185-212.

[34] E. Candés, J. Romberg, *Quantitative robust uncertainty principles and optimally sparse decompositions*, Fund. Comput. Math. 6, 2006, 227-254.

[35] E. Candés, J. Romberg, T. Tao, *Robust uncertainly principles: Exact signal reconstruction from highly imcomplete frequence information*, IEEE Trans. Inf. Theory 52, 2006, 489-509.

[36] P.G. Casazza, O. Christensen, *Frames containing a Riesz basis and preservation of this property under perturbations*, SIAM J. Math. Anal., 29, 1998, 266-278.

[37] P. Casazza, D. Han, D. Larson, *Frames for Banach spaces*, in The Functional and Harmonic Analysis of Wavelets and Frames, Contemp. Math. 247, AMS, Providence, RI, 1999, 149-182.

[38] W. Chen, S. Itoh, and J. Shiki, On sampling in shift invariant spaces, *IEEE Trans. on Information Theory*, 48, 2002, no. 10, 2802–2809.

[39] O. Christensen, *Moment problems for frames and applications to irregular sampling and Gabor frames*, Appl. Comput. Harmon. Anal., 3, 1996, 82-86.

[40] O. Christensen, *An Introduction to Frames and Riesz Basis*, Applied and Numerical Harmonic Analysis, Birkh́auser, Basel, 2003.

[41] C.K. Chui, ED., *Wavelets: A Tutorial in Theory and Applications*, Academic Press, San Diego, CA, 1992.

[42] I. Daubechies, *Ten Lectures on Wavelets*, SIAM, Philadelphia, 1992.

[43] R.A. DeVore, *Deterministic constructions of compressed sensing matrices*, J.

Complex. 23, 2007, 918-925.

[44] F. Deutsch, Best approximation in Inner Product Spaces, *Springer, New York*, 2001.

[45] D.L. Donoho, *Compressed sensing*, IEEE Trans. Inf. Theory 52, 2006, 1289-1306.

[46] G.C. Donovan, J.S. Geronomo, D.P. Hardin, *Intertwining multiresolution analyses and the construction of piecewise-polynomial wavelets*, SIAM J. Math. Anal., 27, 1996, 1791-1815.

[47] P.L. Dragotti, M. Vetterli, T. Blu, *Sampling moments and reconstructing signals with finite rate of innnovation: Shannon meets strang-fix*, IEEE Trans. Signal Process. 55, 2007, 1741-1757.

[48] R. J. Duffin and A. C. Schaeffer, A class of nonharmonic Fourier series, *Trans. Amer. Math. Soc.*, 72, 1952, 341-366.

[49] C. Eckart, G. Young, *The approximation of one marix by another of lower rank*, Psychometrica 1, 1936, 211-218.

[50] S.J. Favier, R.A. Zalik, *On the stability of frames and Riesz bases*, Appl. Comput. Harmon. Anal. 2, 1995, 160-173.

[51] H.G. Feichtinger, *New results on regular and irregular sampling based on Wiener amalgams*, in Proc. Conf. Function Spaces, K. Jarosz, ed., Lecture Notes in Math. 136, Springer-Verlag, New York, 1991, 107-121.

[52] H. G. Feichtinger, Wiener amalgam over Euclidean spaces and some of their applications, in: K. Jarosz (Ed.) Proc. Conf. Function Spaces, Edwardsville, IL, USA, 1990, in: *Lecture Notes in Pure and Appl. Math.*, vol.136, 1992, 107–121.

[53] H.G. Feichtinger, K. Grchenig, *Banach spaces related to integrable group representations and their atomic decomposition*, J. Funct. Anal., 86, 1989, 307-340.

[54] H. Feichtinger, U. Molter, and J.L. Romero, Pertubation techniques in spline-type spaces, *Int. J. Wavelets Multiresolut. Inf. Process*, preprint 2007.

[55] T.N.T. Goodman, S.L. Lee, W.S. Tang, *Wavelets wandering subspaces*, Trans. Amer. Math. Soc., 338, 1993, 639-654.

[56] H. G. Feichtinger and W. Sun, Stability of Gabor frames with arbitrary sampling points, *Acta Math. Hungar.* **113** (2006), no. 3, 187–212.

[57] R. Gribonval, M. Nielsen, *Sparse decomposition in union of bases*, IEEE Trans. Inf. Theory 49, 2003, 3320-3325.

[58] K. Gröchenig, *Reconstruction algorithms in irregular sampling*, Math. Comp., 59, 1992, 181-194.

[59] K. Gröchenig, *Localization of Frames, Banach Frames, and the Invertibility of the Frame Operator*, J. Fourier Anal. Appl., 10, 2004, 105-132.

[60] K. Gröchenig and M. Leinert, Wiener's Lemma for twisted convolution and Gabor Frames, *J. Amer. Math. Soc.* 17, 2004, 1-18.

[61] J.R. Higgins, *Five short stories about cardinal series*, Bull. Amer. Math. Soc., 121, 1985, 45-89.

[62] J.R. Higgins, *Sampling theory for Paley-Wiener spaces in the Riesz basis settings*, Proc. Roy. Irish Acad. Sect., 2, 1994, 219-235.

[63] R. Horn and C. Johnson, Matrix Analysis, *Cambridge University Press*, Cambridge, 1985.

[64] H.S. Hou, H.C. Andrews, *Cubic splines for image interpolation and the digital*, IEEE Trans. Acoust. Speech Signal Process., 26, 1978, 508-517.

[65] A.J.E.M. Janssen, *The Zak transform and sampling theorems for wavelet subspaces*, IEEE Trans. Signal Process., 41, 1993, 3360-3364.

[66] A. Jerri, *The Shannon sampling theorem-its various extensions and applications: A tutorial review*, Proc. IEEE, 65, 1977, 1565-1596.

[67] K. Jetter, J. Stckler, *Topics in scattered data interpolation and non-uniform sampling*, in Surface Fitting and Multiresolution Methods, A. Le Mhaute, C. Rabut, and L.L. Schumaker, eds., Vanderbilt University Press, Nashville, TN, 1997, 191-207.

[68] R.Q. Jia, *Shift-invariant spaces and linear operator equations*, Israel Math. J., 103, 1998, 259-288.

[69] R.Q. Jia, *Stability of the shifts of a finite number of functions*, J. Approx. Theory, 95, 1998, 194-202.

[70] R.Q. Jia, C.A. Micchelli, *On linear independence of integer translates of a finite number of functions*, Proc. Edinburgh Math. Soc., 36, 1992, 69-85.

[71] M.I. Kadec, *The exact value of the Paley-Wiener constant*, Soviet Math. Dolk., 5, 1964, 559-561.

[72] H.P. Kramer, *A general sampling theorem*, J. Math. Phys., 38, 1959, 68-72.

[73] H. Landau, *A sparse regular sequence of exponentials closed on large sets*, Bull. Amer. Math. Soc., 70, 1964, 566-569.

[74] H. Landau, *Necessary density conditions for sampling and interpolation of certain entire functions*, Acta Math., 117, 1967, 37-52.

[75] Y. Liu, *Irregular sampling for spline in wavelets subspaces*, IEEE Trans. Inform.

Theory, 42, 1996, 623-627.

[76] Y. Liu, G.G. Walter, *Irregular sampling in wavelets subspaces*, J. Fourier Anal. Appl., 2, 1996, 181-189.

[77] Y. Lyubarski, W.R. Madych, *The recovery of irregularly sampled band limited functions via tempered splines*, J. Funct. Anal., 125, 1994, 201-222.

[78] Y. Lyubarski, K. Seip, *Convergence and summability of Gabor expansions at the Nyqyuist density*, J. Fourier Anal. Appl., 5, 1999, 127-157.

[79] Y. Lu, M.N. Do, *A theory for sampling signals from union of subspaces*, IEEE TRans. Signal Process. 56, 2008, 2334-2345.

[80] S. Luo, Error estimation for non-uniform sampling in shift invariant space, *Appl. Anal.* 86, no. 4, 2007, 483-496.

[81] J.P. Oakley, M.J. Cunnigham, G. Little, *A Fourier-domain formula for the least square projection of a function onto a repetitive basis in N-dimensional space*, IEEE Trans. Acoust. Speech Signal Process., 38, 1990, 114-120.

[82] J.L. Ostuni, A.K.S. Santhaand, V.S. Mattay, D.R. Weinberger, R.L. Levin, J.A. Frank, *Analysis of interpolation effects in the reslicing of functional MR-Images*, J. Computer Assisted Tomography, 21, 1997, 803-810.

[83] Y. Ma, H. Derksen, W. Hong, W. Wright, *Segmentation of multivariate mixed data via lossy coding and compression*, IEEE Trans. Pattern Anal. Mach. Intell. 29(9), 2007, 1546-1562.

[84] Y. Ma, H. Derksen, R. Fossum, *Estimation of subspace arrangements with applications in modeling and segmanting mixed data*, SIAM Rev. 50, 2008, 413-458.

[85] I. Maravic, M. Vetterli, *Sampling and reconstructing signals with finite rate of innovation in presence of noise*, IEEE Trans. Signal Process. 53, 2005, 2788-2805.

[86] Z.M. Nashed and G.G. Walter, General sampling theorems for functions in reproducing kernel Hilbert spaces, *Math. Control Signals Systems* **4** (1991), no. 4, 363-390.

[87] R.E.A.C. Paley, N. Wiener, *Fourier transform in the complex domain*, in Amer. Math. Colloq. Publ., AMS, Providence, RI, 1934.

[88] A. Papoulis, *Generalized sampling expansions*, Circuits Systems, 24, 1977, 652-654.

[89] B.S. Pavlov, *The basis property of a system of exponentials and the condition of Muckenhoupt*, Dolk. Akad. Nauk SSSR, 247, 1979, 37-40.

[90] D. Potts, G. Steidl, *New Fourier reconstruction algorithms for computerized tomography*, in Proc. Wavelet Applications in Signal Image Processing VIII, 4199,

2000, 13-23.

[91] M. Rauth, T. Strohmer, *Smooth approximation of potential fields from noisy scattered data*, Geophysics, 63, 1998, 85-94.

[92] H. Rauhut, K. Schass, P. Vandergheynst, *Compressed sensing and redundant dictionaries*, IEEE Trans. Inf. Theory 4, 2008, 2210-2219.

[93] K. D. Sauer and J. P. Allebach, Iterative reconstruction of band-limited images from nonuniformly spaced samples, *IEEE Trans. Circuits and Systems*, 34, 1987, 1497-1506.

[94] K. Seip, *On the connection between exponential bases and certain related sequences in* $L^2(-\pi, \pi)$, J. Funct. Anal., 130, 1995, 131-160.

[95] C.E. Shannon, *Communications in presence of noise*, Proc. IRE, 37, 1949, 10-21.

[96] J. Stckler, *Multivariate Affine Frames*, Shaker Verlag, Aachen, Germany, 1998.

[97] G. Strang, T. Nguyen, *Wavelets and Filter Banks*, Wellesley-Cambridge Press, Wellesley, MA, 1996.

[98] T. Strohmer, *Numerical analysis of the non-uniform sampling problem*, J. Comput. Appl. Math., 122, 2000, 297-316.

[99] W. Sun, X. Zhou, *On the stability of multivariate trigonometric systems*, J. Math. Anal. Appl., 235, 1999, 159-167.

[100] Q. Sun, Non-uniform average sampling and reconstruction of signals with finite rate of innovation, *SIAM Journal of Mathematical Analysis*, 38, 2006, 1389-1422.

[101] W. Sun and X. Zhou, Reconstruction of functions in spline subspaces from local averages, *Proc. Amer. Math. Soc.*, 131, no. 8, 2003, 2561-2571.

[102] W. Sun and X. Zhou, Average sampling in shift invariant subspaces with symmetric averaging functions, *J. Math. Anal. Appl.*, 287, no. 1, 2003, 279-295.

[103] W. Sun, Sampling theorems for multivariate shift invariant subspaces, *Sampl. Theory Signal Image Process.*, 4, no. 1, 2005, 73-98.

[104] P. Thevenaz, T. Blu, M. Unser, *Image interpolation and resampling*, in Handbook of Medical Image Processing, Processing and Analysis, I.N. Bankman, ed., Academic Press, San Diego, CA, 2000, 393-420.

[105] G. Thomas, *A comparison of motion-compensated interlace-to-progressive conversion methods*, Image Communication, 12, 1999, 209-229.

[106] J.A. Tropp, *Algorithmic results for sparse approximation*, IEEE Trans. Inf. Theory 50, 2004, 2231-2242.

[107] M. Unser, Sampling - 50 years after Shannon, *Proceedings of the IEEE*, 88, no. 4, 2000, 569-587.

[108] M. Unser, A. Aldroubi, *Polynomial splines and wavelets-a signal processing perspective*, in Wavelets: A Tutorial in Theory and Applications, C.K. Chui, ed., Academic Press, San Diego, CA, 1992, 543-601.

[109] M. Unser, A. Aldroubi, *Generalized sampling with application to the wavelet transform*, in Proceedings of the Conference on Information Sciences and Systems, The Johns Hopkins University Press, Baltimore, MD, 1993.

[110] M. Unser, A. Aldroubi, *A general sampling theory for non-ideal acquisition devices*, IEEE Trans. Process., 42, 1994, 2915-2925.

[111] M. Unser, A. Aldroubi, M. Eden, *Polynomial spline signal approximations: Filter design and asymptotic equivalence with Shannon's sampling theorem*, IEEE Trans. Image Process., 38, 1992, 95-103.

[112] M. Unser, J. Zerubia, *A generalized sampling theory without bandlimiting constraints*, Trans. Circuits and Systems-II: Analog and Digital Signal Processing, 45, 1998, 959-884.

[113] R. Vidal, Y. Ma, S. Sastry, *Generalized principal component analysis(gpca)*, IEEE Trans. Pattern Anal. Mach. Intell. 27, 2005, 1-15.

[114] R. Vio, T. Strohmer, W. Wamsteker, *On the reconstruction of irregular sampled time series*, Publ. Astronom. Soc. Pac., 112, 2000, 74-90.

[115] G.G. Walter, *A sampling theorem for wavelet subspaces*, IEEE Trans. Inform. Theory, 38, 1992, 881-884.

[116] J.M. Whittaker, *Interpolatory Function Theory*, Cambridge University Press, London, 1935.

[117] X.G. Xia, J.S. Geronimo, D.P. Hardin, B.W. Sutter, *Design of prefilters for discrete multiwavelet transforms*, IEEE Trans. Signal Process., 44, 1996, 25-35.

[118] X.G. Xia, Z.Z. Zhang, *On sampling theorem, wavelets, and wavelet transforms*, IEEE Trans. Signal Process., 41, 1993, 3524-3535.

[119] J. Xian and S. Li, Sampling set conditions in weighted multiply generated shift-invariant spaces and their applications, *Appl. Comput. Harmon. Anal.*, 23, no. 2, 2007, 171-180.

[120] S. Yeh and H. Stark, Iterative and one-step reconstruction from nonuniform samples by convex projections, *J. Opt. Soc. Amer.*, **A7**, 1990, 491-499.