Dynamical Sampling and its Applications

By

Longxiu Huang

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 10, 2019

Nashville, Tennessee

Approved:

Akram Aldoubi, Ph.D.

Douglas Hardin, Ph.D.

Alexander Powell, Ph.D.

Larry Schumaker, Ph.D.

David Smith, Ph.D.

To the memory of my dear aunt and grandfather,

ACKNOWLEDGMENTS

First and foremost, I would like to express my deepest gratitude to my supervisor Prof. Akram Aldroubi, for his endless support, guidance, and encouragement. Honestly speaking, I think that I am the luckiest student in the world. For my research, my supervisor taught me how to ask questions and how to make baby steps when resolving research projects during my Ph.D. period. I thank him for carefully reading, commenting and polishing on countless revisions of manuscripts and writing works for publication. It is a precious privilege for me to have the experience of working with him. Besides mathematics, I have also learned from him how to deal with tough things in life and how to keep a positive attitude in life.

I am also profoundly grateful to Prof. Douglas Hardin, Prof. Alexander Powell, Prof. Larry Schumaker, and Prof. David Smith for serving as my dissertation committee members and their constructive comments and suggestions on my research work. Additionally, I would like to thank Prof. Yuesheng Xu and Dr. Weicai Ye for their help and encouragement when I was an undergraduate and later Ph.D student. In addition, I want to thank all of my collaborators for their great work and interesting discussions, epscially Prof. Ilya Krishtal, Prof. Keri Kornelson, Prof. Roy Lederman, Dr. Keaton Hamm, and Dr. Armenak Petrosyan.

Moreover, I want to thank the lovely HAGS (Harmonic Analysis Group) members at Vanderbilt, especially Prof. David Smith, Dr. Keaton Hamm, Dr. Sui Tang, and Dr. Armenak Petrosyan from whom I have learned a lot.

Special thanks go to my old friends Tingting Chen and Yachen Wang. Thank them for providing me companions and encouragement when I was in my hard time. I'm thankful to all my friends namely at Vanderbilt. They made the time I spent in Nashville enjoyable and meaningful. Special thanks go to my roommates Bin Gui and Bin Sun, who have helped me a lot in my research and my daily life.

Most importantly, none of my work would have been accomplished without the love of my family. I must thank my parents, elder sister, and younger brother. Their guidance, encouragement, and support always provided the onward motivation throughout my past journey in life. Meanwhile, I would like to thank my cousins Tao Wang and Bin Wang, for their patience, guidance, and encouragement. Especially, thank Tao Wang for carefully reading, commenting, and polishing on my writing works.

Finally, I would like to express my deepest gratefulness to my uncle's family for their kindness and providing me a positive learning environment. Special thanks to Xuanhao Huang for being a great cousin.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Motivation

A typical problem in sampling theory is to reconstruct a function (signal) f in a separable Hilbert space from its samples, for which a natural approach would be to sample the function f at as many accessible positions as possible. In general, one expects that, with some *a priori* information, f can be reconstructed from those samples. This idea motivates the classical sampling theory. Related results can be consulted in, e.g., [3, 4, 8, 16, 20, 53, 54]. However, there are various restrictions in real-world applications. For example, sensors may not be permitted to be installed at some required locations. Moreover, the spatial sampling density can be very limited, because sensors are often expensive and it is costly to achieve a high sampling density.

On the other hand, in many instances, the function f comes from an evolving system which is driven by a (partially) known operator, and the feature of the function can be exploited to compensate for the insufficiency in sampling locations. An intuitive example is provided by diffusion and modeled by the heat equation [43] which exemplifies a spatio-temporal trade-off. Aldroubi and his collaborators have developed a novel mathematical framework, called *dynamical sampling*, to study the spatio-temporal trade-off problem [5, 6, 12]. Dynamical sampling has potential applications in signal processing, medical imaging, wireless sensor networks, and to name a few.

1.2 Problem Formulation for General Dynamical Sampling

The general formulation can be stated as follows. Let f be a vector in a Hilbert space $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ and A be a bounded linear operator on $\mathcal{H}$. The initial signal f evolves, and at time t becomes

$$f_t = A^t f.$$

Given a countable (finite or countably infinite) set of vectors $\mathcal{G} \subset \mathcal{H}$, the samples (measurements) over a time set T are of the form

$$\mathcal{M} = \left\{\langle A^t f, g\rangle : g \in \mathcal{G}, t \in T\right\}.$$

The general dynamical sampling problem can be described as

Problem 1. *What are the conditions on the operator A, the sampling set $\mathcal{G}$, and the time set $T \subset [0, \infty)$ such that any function $f \in \mathcal{H}$ can be recovered or stably recovered from $\mathcal{M}$?*

Remark 1.2.1. *In general, the measurements $\left\{\langle A^t f, g\rangle : g \in \mathcal{G}\right\}$ at every single time point $t \in T$ are insufficient to recover $A^t f$. In other words, $A^t f$ is undersampled.*

By the recovery of f we mean that there exists an operator R from $\mathcal{G} \times T$ to $\mathcal{H}$ such that $R(\langle A^t f, g\rangle) = f$ for all $f \in \mathcal{H}$. If R is bounded, then we say that the recovery of f is stable. Using the relation between A and its adjoint operator A^*, Problem 1 is equivalent to

Problem 2. *What are the conditions on the operator A, $\mathcal{G}$, and the time set $T \subset [0, \infty)$ that ensure that the system $\{A^{*t} g : g \in \mathcal{G}, t \in T\}$ is complete or a (continuous) frame for $\mathcal{H}$?*

Because of this equivalence, we can investigate Problem 2 instead of Problem 1. For the notational simplicity, we drop the * and study the system of the form $\{A^t g : g \in \mathcal{G}, t \in T\}$ in Problem 2.

1.3 Relation to Other Fields

Dynamical sampling, as a new sampling theory, has relations with other areas of mathematics. For instance, it has similarities with wavelet theory [18, 29, 36, 38]. In wavelet theory, a high-pass operator H and a low-pass operator L are applied to the function f. The goal is to design operators H and L so that the reconstruction of f from the combined samples of Hf and Lf is possible. Similarly, in dynamical sampling the main purpose is to reproduce f via the samples $\mathcal{M}$ collected from different time points. However, in dynamical sampling there is only one operator A, which acts on the function f iteratively. There is no specific structural restrictions on A.

Dynamical sampling has close relation with inverse problems (see [47] and the references therein). An inverse problem is the process of finding the factors that result in a set of observations. The main goal of dynamical sampling is thus the inverse problem of finding the factor f from the knowledge of the driving operator A and the set $\mathcal{M} = \{\langle A^t f, g\rangle : t \in T, g \in \mathcal{G}\}$ of observations. If the full information of A is given, then the inverse problem is linear. Otherwise, the problem is non-linear and in this case we also want to recover the operator A. In other words, we ask the following question.

Question 1. *What are the conditions on A, T, and $\mathcal{G}$ such that A and f can be recovered from the observed data $\mathcal{M}$?*

In addition, methods to solve dynamical sampling problems have close relation with spectral theory, operator algebras, numerical linear algebra, frame theory, and complex analysis.

1.4 Overview and Organization

In the existing studies of dynamical sampling, only the case for discrete time sets T has been considered. However, time is continuous in the physical world, and thus it is natural to consider the dynamical sampling problem for continuous time intervals T. This work is presented in Chapter 2. In this work, we consider systems of the form $\{A^t g : g \in \mathcal{G}, t \in [0, L]\} \subset \mathcal{H}$, where $A \in \mathcal{B}(\mathcal{H})$. The goal is to study the frame property of such systems. To this end, we derive some other properties in the intermediate steps. In particular, we study the completeness and Besselness of these systems. These results are a joint work with Akram Aldroubi and Armenak Petrosyan and it is documented in the preprint entitled "Frames induced by the action of continuous powers of an operator", see [11].

In addition, noises are ubiquitous in real world and sampling applications which necessitates an investigation of the impacts of noise on dynamic sampling. The related results are reported in Chapter 3. The work is joint with Akram Aldroubi, Ilya Krishtal, Akos Ledeczi, Roy R. Lederman, and Peter Volgyesi and appears in [9, 10].

CHAPTER 2

Frames Induced by the Action of Continuous Powers of an Operator

2.1 Problem Formulation

In this chapter, we consider dynamical sampling with the time set belonging to a bounded interval. Specifically, we investigate systems of the form $\{A^t g : g \in \mathcal{G}, t \in [0, L]\}$, where A is a bounded linear normal operator in a separable Hilberst space $\mathcal{H}$, $\mathcal{G} \subset \mathcal{H}$ is a countable set, and L is a positive real number. The main goal is to investigate the following problems.

Problem 3. *What are the conditions on A, $\mathcal{G}$, and L that ensure that system $\{A^t g : g \in \mathcal{G}, t \in [0, L]\}$ is complete, Bessel, or a continuous frame in $\mathcal{H}$?*

The discretization of continuous frames [31, 32] is a central question. For systems of the form $\{A^t g : g \in \mathcal{G}, t \in [0, L]\}$, we ask

Problem 4. *Suppose $\{A^t g : g \in \mathcal{G}, t \in [0, L]\}$ is a continuous frame. Is there a partition $0 = t_1 < t_2 < \ldots < t_n < L$ of $[0, L]$ such that the system $\{A^{t_i} g : g \in \mathcal{G}, 1 \leqslant i \leqslant n$ and $i \in \mathbb{N}\}$ is a discrete frame (see inequality (2.1))?*

2.2 Recent Results on Dynamical Sampling and Frames

Existing studies on various aspect of the dynamical sampling problem and related frame theory grew out of the work in [1, 5, 6, 7, 42, 49], see, for example, [2, 21, 22, 24, 40, 45, 46, 48, 56, 57] and the references therein. However, except for a few, they all focus on uniform discrete-time sets $T \subset \{0, 1, 2, \ldots\}$, e.g., $T = \{1, \ldots, N\}$ or $T = \mathbb{N}$ (see e.g., [36]).

Even though the general dynamical sampling problem for discrete-time sets in finite dimensions (hence problems of systems and frames induced by iterations $\{A^n g : g \in \mathcal{G}, n \in T\}$) have been mostly resolved in [6], many problems and conjectures remain open for the infinite dimensional case. This state of affairs is not surprising because some of these problems take root in the deep theory of functional analysis and operator theory such as the Kadison Singer Theorem [44], some open generalizations of the Müntz-Szász Theorem [51], and the famous invariant subspace conjecture.

When $T = \mathbb{N}$ and $A \in \mathcal{B}(\mathcal{H})$, it is not difficult to show that

Theorem 2.2.1 ([13]). *If, for an operator $A \in B(\mathcal{H})$, there exists a countable set of vectors $\mathcal{G}$ in $\mathcal{H}$ such that $\{A^n g\}_{g\in\mathcal{G},\, n\geqslant 0}$ is a frame in $\mathcal{H}$, then for every $f \in \mathcal{H}$, $(A^*)^n f \to 0$ as $n \to \infty$.*

Thus, in particular it is not possible to construct frames using non-negative iterations when A is a unitary operator. For example, the right-shift operator S on $\mathcal{H} = \ell^2(\mathbb{N})$ generates an orthonormal basis for $\ell^2(\mathbb{N})$ by iterations over $\mathcal{G} = \{(1, 0, \dots,)\}$. Clearly, $(S^*)^n f \to 0$ as $n \to \infty$ for this case. However, if we change the space to $\mathcal{H} = \ell^2(\mathbb{Z})$, the right-shift operator S becomes unitary, and there exists no subset $\mathcal{G}$ of $\ell^2(\mathbb{Z})$ such that $\{S^n g\}_{g\in\mathcal{G},\, n\geqslant 0}$ is a frame for $\ell^2(\mathbb{Z})$.

On the other hand, for normal operators, it is possible to find frames of the form $\{A^n g\}_{g\in\mathcal{G},\, n\geqslant 0}$; however, no such a frame can be a basis [5].

Frames for $\mathcal{H}$ can be generated by the iterative action on a single vector g, i.e., there exist normal operators and associated cyclic vectors such that $\{A^n g\}_{\, n\geqslant 0}$ is a frame for $\mathcal{H}$ [6]. Specifically,

Theorem 2.2.2 ([5]). *Let A be a bounded normal operator on an infinite dimensional Hilbert space $\mathcal{H}$. Then, $\{A^n g\}_{n\geqslant 0}$ is a frame for $\mathcal{H}$ if and only if the following five conditions are satisfied: (i) $A = \sum_j \lambda_j P_j$, where P_j are rank one orthogonal projections; (ii) $|\lambda_k| < 1$ for all k; (iii) $|\lambda_k| \to 1$; (iv) $\{\lambda_k\}$ satisfies Carleson's condition $\inf_n \prod_{k\neq n} \frac{|\lambda_n-\lambda_k|}{|1-\bar{\lambda}_n\lambda_k|} \geqslant \delta$, for some $\delta > 0$; and (v) $0 < c \leqslant \frac{\|P_j g\|}{\sqrt{1-|\lambda_k|^2}} \leqslant C < \infty$, for some constants c, C.*

It turns out that if A is normal in an infinite dimensional Hilbert space $\mathcal{H}$, and $\{A^n g\}_{g\in\mathcal{G},\, n\geqslant 0}$ is a frame for some $\mathcal{G} \subset \mathcal{H}$ with $|\mathcal{G}| < \infty$, then A is necessarily of the form described in Theorem 2.2.2:

Theorem 2.2.3 ([13]). *Let A be a bounded normal operator in an infinite dimensional Hilbert space $\mathcal{H}$. If the system of vectors $\{A^n g\}_{g\in\mathcal{G},\, n\geqslant 0}$ is a frame for some $\mathcal{G} \subset \mathcal{H}$ with $|\mathcal{G}| < \infty$, then $A = \sum_j \lambda_j P_j$ where P_j are projections such that $\operatorname{rank}(P_j) \leqslant |\mathcal{G}|$ (i.e., the global multiplicity of A is less than or equal to $|\mathcal{G}|$) . In addition, (ii) and (iii) of Theorem 2.2.2 are satisfied.*

The necessary and sufficient conditions generalizing Theorem 2.2.2 for the case $1 < |G| < \infty$ have been derived in [21].

Viewing Theorem 2.2.2 from a different perspective, Christensen and Hasannasab ask whether a frame $\{h_n\}_{n\in I}$ has a representation of the form $h_n = A^n h_0$ for some operator A when $I = \mathbb{N} \cup \{0\}$ or $I = \mathbb{Z}$. This question is partially answered in [25] and gives rise to many new open problems and conjectures [24].

The set of self-adjoint operators is an important class of normal operators because it is often encountered in applications. For this class, one can rule out certain types of normalized frames.

Theorem 2.2.4 ([6])**.** *If A is a self-adjoint operator on $\mathcal{H}$, then the system $\left\{\frac{A^n g}{\|A^n g\|}\right\}_{g\in\mathcal{G},\ n\geqslant 0}$ is not a frame for $\mathcal{H}$.*

However, for normal operators, the following conjecture remains open:

Conjecture 2.2.5. *The statement of Theorem 2.2.4 holds for normal operators.*

Conjecture 2.2.5 does not hold if the operator is not normal. For example, the shift-operator S on $\ell^2(\mathbb{N})$ defined by $S(x_1, x_2, \dots) = (0, x_1, x_2, \dots)$, is not normal, and $\{S^n e_1\}$ is an orthonormal basis for $\ell^2(\mathbb{N})$, where $e_1 = (1, 0, \dots)$.

2.3 Notation and preliminaries

2.3.1 Frames

In a foundational paper, Duffin and Schaeffer introduced the theory of frames in the context of non-harmonic Fourier series [30]. Specifically, a frame $\{\phi_n\}_{n\in\mathbb{Z}}$ in a separable Hilbert space $\mathcal{H}$ is a sequence of vectors satisfying

$$c\,\|f\|^2 \leqslant \sum_{n\in\mathbb{Z}} |\langle f, \phi_n\rangle|^2 \leqslant C\|f\|^2, \text{ for all } f \in \mathcal{H}, \tag{2.1}$$

for some positive constants $c, C > 0$. Later, the notion in (2.1) is generalized to continuous frames [14, 15, 31, 33], where the definition is stated below:

Definition 2.3.1. *Let $\mathcal{H}$ be a complex Hilbert space and let (Ω, μ) be a measure space with positive measure μ. A mapping $F : \Omega \to \mathcal{H}$ is called a frame with respect to (Ω, μ), if*

1. *F is weakly-measurable, i.e.,* $\omega \to \langle f, F(\omega)\rangle$ *is a measurable function on* Ω *for all* $f \in \mathcal{H}$*;*

2. *there exist constants* c *and* $C > 0$ *such that*

$$c\|f\|^2 \leqslant \int_\Omega |\langle f, F(\omega)\rangle|^2 d\mu(\omega) \leqslant C\|f\|^2, \text{ for all } f \in \mathcal{H}. \tag{2.2}$$

Here the constants c *and* C *are called continuous frame (lower and upper) bounds. In addition, F is called a tight continuous frame if* $c = C$*. The mapping F is called Bessel if the second inequality in* (2.2) *holds. In this case, C is called a Bessel constant.*

The frame operator $S = S_F$ on $\mathcal{H}$ associated with F is defined in the weak sense by

$$S_F f = \int_\Omega \langle f, F(\omega)\rangle F(\omega) d\mu(\omega).$$

According to (2.2), S_F is well defined, invertible with bounded inverse (see [31]). Thus every $f \in \mathcal{H}$ has the representations

$$f = S_F^{-1} S_F f = \int_\Omega \langle f, F(\omega)\rangle S_F^{-1} F(\omega) d\mu(\omega),$$

$$f = S_F S_F^{-1} f = \int_\Omega \langle f, S_F^{-1} F(\omega)\rangle F(\omega) d\mu(\omega).$$

If μ is the counting measure and $\Omega = \mathbb{N}$, then one gets back the Duffin-Schaffer frame in (2.1).

In the sequel, $\Omega = \mathcal{G} \times [0, L]$, and μ is the product of the counting measure on $\mathcal{G}$ and the Lebesgue measure on $[0, L]$. In this case, F is called a *semi-continuous frame* and (2.2) becomes

$$c\|f\|^2 \leqslant \sum_{g\in\mathcal{G}} \int_0^L |\langle f, A^t g\rangle|^2 dt \leqslant C\|f\|^2, \text{ for all } f \in \mathcal{H}. \tag{2.3}$$

2.3.2 Normal operators

Let $\mathcal{B}(\mathcal{H})$ denote the space of bounded linear operators on a complex separable Hilbert space $\mathcal{H}$. In the sequel, all the operators are assumed to be normal. Normal operators have the following invertibility property (see [52, Theorem 12.12]).

Theorem 2.3.2. *If $A \in \mathcal{B}(\mathcal{H})$, then A is invertible (i.e., A has bounded inverse) if and only if there exists $c > 0$ such that $\|Af\| \geqslant c\|f\|$ for all $f \in \mathcal{H}$.*

For completeness, the spectral theorem with multiplicity is stated below, and the following notation is used in its statement.

For a non-negative regular Borel measure μ on $\mathbb{C}$, N_μ will denote the multiplication operator acting on $L^2(\mu)$, i.e., for a μ-measurable function $f : \mathbb{C} \to \mathbb{C}$ such that $\int_{\mathbb{C}} |f(z)|^2 d\mu(z) < \infty$,

$$N_\mu f(z) = zf(z).$$

We will use the notation $[\mu] = [\nu]$ to denote two mutually absolutely continuous measures μ and ν.

The operator $N_\mu^{(k)}$ will denote the direct sum of k copies of N_μ, i.e.,

$$(N_\mu)^{(k)} = \oplus_{i=1}^k N_\mu.$$

Similarly, the space $(L^2(\mu))^{(k)}$ will denote the direct sum of k copies of $L^2(\mu)$.

Theorem 2.3.3 (Spectral theorem with multiplicity). *For any normal operator A on $\mathcal{H}$ there are mutually singular non-negative Borel measures $\mu_j, 1 \leqslant j \leqslant \infty$, such that A is equivalent to the operator*

$$N_{\mu_\infty}^{(\infty)} \oplus N_{\mu_1} \oplus N_{\mu_2}^{(2)} \oplus \dots,$$

i.e., there exists a unitary transformation

$$U : \mathcal{H} \to (L^2(\mu_\infty))^{(\infty)} \oplus L^2(\mu_1) \oplus (L^2(\mu_2))^{(2)} \oplus \dots$$

such that

$$UAU^{-1} = N_{\mu_\infty}^{(\infty)} \oplus N_{\mu_1} \oplus N_{\mu_2}^{(2)} \oplus \dots. \tag{2.4}$$

Moreover, if $\tilde{A}$ is another normal operator with corresponding measures $\nu_\infty, \nu_1, \nu_2, \dots$, then $\tilde{A}$ is unitary equivalent to A if and only if $[\nu_j] = [\mu_j]$ for $j = 1, \dots, \infty$.

A proof of the theorem can be found in [28, Ch. IX, Theorem 10.16] and [27, Theorem 9.14].

Since the measures μ_j are mutually singular, there are mutually disjoint Borel sets $\{\mathcal{E}_j\}_{j=1}^{\infty} \cup \{\mathcal{E}_\infty\}$ such that μ_j is supported on $\mathcal{E}_j$ for every $1 \leqslant j \leqslant \infty$. The scalar-valued spectral measure μ associated with the normal operator A is defined as

$$\mu = \sum_{1 \leqslant j \leqslant \infty} \mu_j. \tag{2.5}$$

The Borel function $m_A : \mathbb{C} \to \mathbb{N}^* \cup \{0\}$ given by

$$m_A(z) = \infty \cdot \chi_{\mathcal{E}_\infty}(z) + \sum_{j=1}^{\infty} j \chi_{\mathcal{E}_j}(z) \tag{2.6}$$

is called the multiplicity function of the operator A, where $\mathbb{N}$ is the set of natural numbers starting with 1, $\mathbb{N}^* = \mathbb{N} \cup \{\infty\}$, $\chi_E(z)$ is the characteristic function on set E defined by $\chi_E(z) = \begin{cases} 1, z \in E \\ 0, \text{otherwise} \end{cases}$ and $\infty \cdot 0 = 0$.

From Theorem 2.3.3, every normal operator is uniquely determined, up to a unitary equivalence, by the pair $([\mu], m_A)$.

For $j \in \mathbb{N}$, let Ω_j be the set $\{1, ..., j\}$ and let Ω_∞ be the set $\mathbb{N}$. Then $\ell^2(\Omega_j) \cong \mathbb{C}^j$, for $j \in \mathbb{N}$, and $\ell^2(\Omega_\infty) = \ell^2(\mathbb{N})$. For $j = 0$, we use $\ell^2(\Omega_0)$ to represent the trivial space $\{0\}$.

Let $\mathcal{W}$ be the Hilbert space

$$\mathcal{W} = (L^2(\mu_\infty))^{(\infty)} \oplus L^2(\mu_1) \oplus (L^2(\mu_2))^{(2)} \oplus \cdots$$

associated with the operator A and let $U : \mathcal{H} \to \mathcal{W}$ be the unitary operator given by Theorem 2.3.3. If $g \in \mathcal{H}$, we denote by $\widetilde{g}$ the image of g under U. Since $\widetilde{g} \in \mathcal{W}$, one has $\widetilde{g} = (\widetilde{g}_j)_{j \in \mathbb{N}^*}$, where $\widetilde{g}_j$ is the restriction of $\widetilde{g}$ to $(L^2(\mu_j))^{(j)}$. Thus, for any $j \in \mathbb{N}^*$, $\widetilde{g}_j$ is a function from $\mathbb{C}$ to $\ell^2(\Omega_j)$ and

$$\sum_{j \in \mathbb{N}^*} \int_{\mathbb{C}} \|\widetilde{g}_j(z)\|^2_{\ell^2(\Omega_j)} d\mu_j(z) = \|g\|^2 < \infty.$$

Let P_j be the projection defined for every $\widetilde{g} \in \mathcal{W}$ by $P_j \widetilde{g} = \widetilde{f}$, where $\widetilde{f}_j = \widetilde{g}_j$ and $\widetilde{f}_k = 0$ for $k \neq j$.

Let E be the spectral measure for the normal operator A. Then, for every μ-measurable set

$G \subset \mathbb{C}$ and vectors f, g in $\mathcal{H}$, one has the following formula

$$\langle E(G)f, g\rangle_{\mathcal{H}} = \int_G \left[\sum_{1 \leqslant j \leqslant \infty} \chi_{\mathcal{E}_j}(z) \langle \widetilde{f}_j(z), \widetilde{g}_j(z)\rangle_{\ell^2(\Omega_j)} \right] d\mu(z),$$

which relates the spectral measure of A to the scalar-valued spectral measure of A.

Definition 2.3.4. *Given a normal operator A, A^t is defined as follows:*

$$A^t : \mathcal{H} \to \mathcal{H}$$

by

$$\langle A^t f_1, f_2\rangle = \int_{z \in \sigma(A)} z^t \langle \tilde{f}_1(z), \tilde{f}_2(z)\rangle d\mu(z), \textit{ for all } f_1, f_2 \in \mathcal{H},$$

where $z^t = \exp(t(\ln(|z|) + i\arg(z)))$ *and* $\arg(z) \in [-\pi, \pi)$.

Using the fact that $\exp(i\arg(z) + i\arg(\bar{z})) = 1$, *it follows that* $(A^*)^t = (A^t)^*$ *for* $t \in \mathbb{R}$.

Section 2.5 will exploit the reductive operators which were introduced by P.Halmos and J.Wermer [37, 55]. For clarity, the definition is given as follows.

Definition 2.3.5. *A closed subspace $V \subset \mathcal{H}$ is called reducing for the operator A if both V and its orthogonal complement $V^{\perp}$ are invariant subspaces of A.*

Definition 2.3.6. *An operator A is called reductive if every invariant subspace of A is reducing.*

2.3.3 Holomorphic Function

The techniques of complex analysis, e.g., the properties of holomorphic functions (see [26, 51] and the references therein), are used extensively in the present work, including Montel's Theorem as stated below.

Definition 2.3.7 (Normal family). *A family $\mathfrak{F}$ of holomorphic functions in a region X of the complex plane with values in $\mathbb{C}$ is called normal if every sequence in $\mathfrak{F}$ contains a subsequence which converges uniformly to a holomorphic function on compact subsets of X.*

Theorem 2.3.8 (Montel's Theorem). *A uniformly bounded family of holomorphic functions defined on an open subset of the complex numbers is normal.*

2.4 Contributions and Organization

The present work concentrates on systems of the form $\{A^t g : g \in \mathcal{G}, t \in [0, L]\} \subset \mathcal{H}$, where $A \in \mathcal{B}(\mathcal{H})$. The goal is to study the frame property of such systems. To this end, we need to derive some other properties in the intermediate steps. In particular, we study the completeness and Besselness of these systems.

For the completeness of $\{A^t g : g \in \mathcal{G}, t \in [0, L]\}$, necessary and sufficient conditions are derived in Section 2.5. In light of the results in [5], the form of the necessary and sufficient conditions are not surprising. However, the proofs and reductions to the known cases are appealing due to the use of certain techniques of complex analysis, and they are useful for the analysis of frames in the subsequent sections.

The Bessel property of the system $\{A^t g : g \in \mathcal{G}, t \in [0, L]\}$ is investigated in Section 2.6. Specifically, if $\mathcal{H}$ is a finite dimensional space (e.g., $\mathbb{C}^d$) and A is a normal operator in $\mathcal{H}$, then the system $\{A^t g\}_{g\in\mathcal{G},t\in[0,L]}$ being Bessel is equivalent to the Besselness of $\mathcal{G}$ in the space $Range(A)$. On the other hand, if $\mathcal{H}$ is an infinite dimensional separable Hilbert space and A is a bounded invertible normal operator, then the only condition ensuring that $\{A^t g\}_{g\in\mathcal{G},t\in[0,L]}$ is Bessel is that $\mathcal{G}$ itself is a Bessel system in $\mathcal{H}$. In addition, an example is described to explain that the non-singularity of A is necessary for the equivalence between the Besselness of $\{A^t g\}_{g\in\mathcal{G},t\in[0,L]}$ and that of $\mathcal{G}$.

Section 5 is devoted to the relations between a semi-continuous frame $\{A^t g\}_{g\in\mathcal{G},t\in[0,L]}$ generated by the action of an operator $A \in \mathcal{B}(\mathcal{H})$ and the discrete systems generated by its time discretization. Specifically, we show that under some mild conditions, $\{A^t g\}_{g\in\mathcal{G},t\in[0,L]}$ is a semi-continuous frame if and only if there exists $T = \{t_i : i = I\} \subset [0, L)$ with $|I| < \infty$ such that $\{A^t g\}_{g\in\mathcal{G},t\in T}$ is a frame system in $\mathcal{H}$. Additionally, Theorem 2.7.5 shows that under proper conditions, the property that $\{A^t g\}_{g\in\mathcal{G},t\in[0,L]}$ is a semi-continuous frame is independent of L.

2.5 Completeness

In this section, we characterize the completeness of the system $\{A^t g\}_{g\in\mathcal{G},t\in[0,L]}$, where A is a (reductive) normal operator on a separable Hilbert space $\mathcal{H}$, $\mathcal{G}$ is a set of vectors in $\mathcal{H}$, and L is a finite positive number.

Theorem 2.5.1. *Let $A \in \mathcal{B}(\mathcal{H})$ be a normal operator, and let $\mathcal{G}$ be a countable set of vectors in $\mathcal{H}$ such that $\{A^t g\}_{g\in\mathcal{G},t\in[0,L]}$ is complete in $\mathcal{H}$. Let $\mu_\infty, \mu_1, \mu_2, \ldots$ be the measures in the representation* (2.4) *of the operator A. Then for every $1 \leqslant j \leqslant \infty$ and μ_j-a.e. z, the system of vectors $\{\tilde{g}_j(z)\}_{g\in\mathcal{G}}$ is complete in $\ell^2(\Omega_j)$.*

If A is also reductive, then $\{A^t g\}_{g\in\mathcal{G},t\in[0,L]}$ being complete in $\mathcal{H}$ is equivalent to $\{\tilde{g}_j(z)\}_{g\in\mathcal{G}}$ being complete in $\ell^2(\Omega_j)$ μ_j-a.e. z for every $1 \leqslant j \leqslant \infty$.

Particularly, if the evolution operator belongs to the following class $\mathcal{A}$ of bounded self-adjoint operators:

$$\begin{aligned} \mathcal{A} \quad = \quad & \{A \in \mathcal{B}(\ell^2(\mathbb{N})) : A = A^*, \\ & \text{and there exists a basis of } \ell^2(\mathbb{N}) \text{ of eigenvectors of } A\}, \end{aligned} \tag{2.7}$$

then, for $A \in \mathcal{A}$, there exists a unitary operator U such that $A = U^*DU$ with $D = \sum_j \lambda_j P_j$, where λ_j are the spectrum of A and P_j is the orthogonal projection to the eigenspace E_j of D associated with the eigenvalue λ_j. Since the operators in $\mathcal{A}$ are also normal and reductive, the following corollary holds.

Corollary 2.5.2. *Let $A \in \mathcal{A}$ with $A = U^*DU$, and let $\mathcal{G}$ be a countable set of vectors in $\ell^2(\mathbb{N})$. Then, $\{A^t g\}_{g\in\mathcal{G},t\in[0,L]}$ is complete in $\ell^2(\mathbb{N})$ if and only if $\{P_j(Ug)\}_{g\in\mathcal{G}}$ is complete in E_j.*

The proof of Theorem 2.5.1 below, also shows that, for normal reductive operators, completeness in $\mathcal{H}$ is equivalent to completeness of the system $\{N^t_{\mu_j}\tilde{g}_j\}_{g\in\mathcal{G},t\in[0,L]}$ in $(L^2(\mu_j))^{(j)}$ for every $1 \leqslant j \leqslant \infty$. In other words, the completeness of $\{A^t g\}_{g\in\mathcal{G},t\in[0,L]}$ is equivalent to the completeness of its projections onto the mutually orthogonal subspaces $U^*P_jU\mathcal{H}$ of $\mathcal{H}$. The following Theorem 2.5.3 summarizes the discussion above.

Theorem 2.5.3. *Let $A \in \mathcal{B}(\mathcal{H})$ be a normal reductive operator on the Hilbert space $\mathcal{H}$, and let $\mathcal{G}$ be a countable system of vectors in $\mathcal{H}$. Then, $\{A^t g\}_{g\in\mathcal{G},t\in[0,L]}$ is complete in $\mathcal{H}$ if only if the system $\{N^t_{\mu_j}\tilde{g}_j\}_{g\in\mathcal{G},t\in[0,L]}$ is complete in $(L^2(\mu_j))^{(j)}$ for every j, $1 \leqslant j \leqslant \infty$.*

2.5.1 Proofs

We begin this section by stating and proving a lemma used to prove Theorem 2.5.1 as well as other results in later sections.

Let A be a normal operator, L be a positive number, $f \in \mathcal{H}$, $\tilde{f} = Uf = (\tilde{f}_j)$, and $\tilde{g} = Ug = (\tilde{g}_j)$ (as in the notation section). Define $F(t)$ by

$$F(t) = \langle A^t g, f \rangle = \int_{\mathbb{C}} z^t \langle \tilde{g}(z), \tilde{f}(z) \rangle d\mu(z).$$

Then, the following lemma holds.

Lemma 2.5.4. *$F(t)$ is an analytic function of t in the domain $\Omega = \{t : \Re(t) > L/2\}$, where $\Re(t)$ stands for the real part of t.*

Proof. First, we aim to prove that $F(t)$ is a continuous function in Ω. Consider $t_0 \in \Omega$. For $|z| \leqslant M$, where $M = \|A\|$, and for $t \in \Omega$ with $|t - t_0| < L/4$, one has

$$\begin{aligned} |z^t \langle \tilde{g}(z), \tilde{f}(z) \rangle| &= |e^{t \ln(z)}| |\langle \tilde{g}(z), \tilde{f}(z) \rangle| \\ &\leqslant e^{(|\ln(M)| + \pi)|t|} |\langle \tilde{g}(z), \tilde{f}(z) \rangle| \\ &\leqslant e^{(|\ln(M)| + \pi)(|t_0| + \frac{L}{4})} |\langle \tilde{g}(z), \tilde{f}(z) \rangle|. \end{aligned}$$

Since the right hand side of the last inequality is an $L^2(\mu)$ function, we can use the dominated convergence theorem for $\Re(t) > L/2 > 0$, and get that for $t_0 \in \Omega$,

$$\lim_{t \to t_0} F(t) = \lim_{t \to t_0} \int_{\mathbb{C}} z^t \langle \tilde{g}(z), \tilde{f}(z) \rangle d\mu(z) = \int_{\mathbb{C}} \lim_{t \to t_0} z^t \langle \tilde{g}(z), \tilde{f}(z) \rangle d\mu(z) = F(t_0).$$

Therefore, $F(t)$ is a continuous function in Ω.

Next we show that for every closed piecewise C^1 curve γ in Ω,

$$\oint_{\gamma} F(t) dt = 0.$$

For fixed γ, there exists finite $M_1 > 0$ such that $L/2 < |t| < M_1$. Therefore, for $|z| \leqslant M$,

$$|z^t \langle \tilde{g}(z), \tilde{f}(z) \rangle| \leqslant e^{\tilde{M}} |\langle \tilde{g}(z), \tilde{f}(z) \rangle|,$$

with $\tilde{M} = M_1(|\ln M| + \pi)$. Then

$$\oint_\gamma \int_{\mathbb{C}} |z^t| |\langle \tilde{g}(z), \tilde{f}(z) \rangle| d\mu(z) dt \leqslant e^{\tilde{M}} \|f\|_2 \|g\|_2 \cdot m_1(\gamma) < \infty,$$

where $m_1(\gamma)$ stands for the length of γ.

By Fubini's theorem,

$$\begin{aligned} \oint_\gamma \int_{\mathbb{C}} z^t \langle \tilde{g}(z), \tilde{f}(z) \rangle d\mu(z) dt &= \int_{\mathbb{C}} \oint_\gamma z^t \langle \tilde{g}(z), \tilde{f}(z) \rangle dt d\mu(z) \\ &= \int_{\mathbb{C}} \langle \tilde{g}(z), \tilde{f}(z) \rangle \oint_\gamma z^t dt d\mu(z) = 0. \end{aligned}$$

where the last equality follows from the fact that for $z \in \mathbb{C}$, $h_z(t) = z^t$ is an analytic function of t in Ω and hence $\oint_\gamma z^t dt = 0$. Then, by Morera's Theorem [51, pp 208], $F(t)$ is analytic on Ω. $\square$

Proof of Theorem 2.5.1. Since $\{A^t g\}_{g \in \mathcal{G}, t \in [0,L]}$ is complete in $\mathcal{H}$,

$$U\{A^t g : g \in \mathcal{G}, t \in [0, L]\} = \{(N^t_{\mu_j} \tilde{g}_j)_{j \in \mathbb{N}^*} : g \in \mathcal{G}, t \in [0, L]\}$$

is complete in $\mathcal{W} = U\mathcal{H}$. Hence, for every $1 \leqslant j \leqslant \infty$, the system $\widetilde{\mathcal{S}}_j = \{N^t_{\mu_j} \tilde{g}_j\}_{g \in \mathcal{G}, t \in [0,L]}$ is complete in $(L^2(\mu_j))^{(j)}$.

To finish the proof of the first statement of Theorem 2.5.1 we use the following lemma, which is an adaptation of [39, Lemma 1] ([5, Lemma 3.5]).

Lemma 2.5.5. *Let $\mathfrak{S}$ be a complete countable set of vectors in $(L^2(\mu_j))^{(j)}$, then for μ_j-almost every z, $\{h(z) : h \in \mathfrak{S}\}$ is complete in $\ell^2(\Omega_j)$.*

Since $\mathcal{H}$ is separable, there exists a countable set $T = \{t_i\}_{i=1}^\infty \subset [0, L]$ with $t_1 = 0$ such that $\overline{span}\{A^t g\}_{g \in \mathcal{G}, t \in T} = \overline{span}\{A^t g\}_{g \in \mathcal{G}, t \in [0,L]}$. Hence, the fact that $\widetilde{\mathcal{S}}_j = \{N^t_{\mu_j} \tilde{g}_j\}_{g \in \mathcal{G}, t \in [0,L]}$ is complete in $(L^2(\mu_j))^{(j)}$ (together with Lemma 2.5.5) implies that $\{z^t \tilde{g}_j(z)\}_{g \in \mathcal{G}, t \in T}$ is complete in

$\ell^2(\Omega_j)$ for each $j \in \mathbb{N}^*$. Let $f \in \mathcal{H}$ and $F(t) = \langle A^t g, f\rangle = 0$ for all $g \in \mathcal{G}, t \in [0, L]$. Since $F(t) = 0$ for all $t \in [0, L]$, and F is analytic for $t \in \Omega = \{t : \Re(t) > L/2\}$, it follows that $F(t) = 0$, for all $t \in \Omega$ (see [51, Theorem 10.18]). Thus, $F(n) = 0$ for all $n \in \mathbb{N}$, i.e., for all $n \in \mathbb{N}$,

$$\int_{\mathbb{C}} z^n \langle \tilde{g}(z), \tilde{f}(z)\rangle d\mu(z) = \int_{\mathbb{C}} z^n \left[\sum_{1 \leqslant j \leqslant \infty} \chi_{\mathcal{E}_j}(z) \langle \tilde{g}_j(z), \tilde{f}_j(z)\rangle_{\ell^2(\Omega_j)} \right] d\mu(z) = 0. \tag{2.8}$$

To finish the proof, we need the following proposition from [55].

Proposition 2.5.6. *Let A be a normal operator on the Hilbert space $\mathcal{H}$ and let μ_j be the measures in the representation* (2.4) *of A. Let μ be as in* (2.5)*. Then, A is reductive if and only if, for any two vectors $f, g \in \mathcal{H}$,*

$$\int_{\mathbb{C}} z^n \left[\sum_{1 \leqslant j \leqslant \infty} \chi_{\mathcal{E}_j}(z) \langle \tilde{g}_j(z), \tilde{f}_j(z)\rangle_{\ell^2(\Omega_j)} \right] d\mu(z) = 0$$

for every $n \geqslant 0$ implies μ_j-a.e. $\langle \tilde{g}_j(z), \tilde{f}_j(z)\rangle_{\ell^2(\Omega_j)} = 0$ for every $j \in \mathbb{N}^*$.

Since A is reductive, it follows from Proposition 2.5.6 that $\langle \tilde{g}_j(z), \tilde{f}_j(z)\rangle_{\ell^2(\Omega_j)} = 0$ for every $j \in \mathbb{N}^*$. Finally, since $\{\tilde{g}_j(z)\}_{g\in\mathcal{G}}$ is complete in $\ell^2(\Omega_j)$ for μ_j-a.e. z, we get that $\tilde{f}_j(z) = 0, \mu_j$-a.e. z for every $j \in \mathbb{N}^*$. Thus, $\tilde{f} = 0$ μ-a.e. z, and hence $f = 0$. Therefore, $\{A^t g\}_{g\in\mathcal{G}, t\in[0,L]}$ is complete in $\mathcal{H}$. □

2.6 Bessel system

The goal of this section is to study the conditions for which the system $\{A^t g\}_{g\in\mathcal{G}, t\in[0,L]}$ is Bessel in $\mathcal{H}$. There are two main theorems that correspond to the finite dimensional case and the infinite dimensional case, respectively. The proofs of the results are relegated to the last subsection. We begin with the following proposition which is valid for both finite and infinite dimensional spaces.

Proposition 2.6.1. *Let $A \in \mathcal{B}(\mathcal{H})$ be normal, $\mathcal{G} \subset \mathcal{H}$ be a countable set of vectors, and let L be a positive finite number. If $\mathcal{G}$ is a Bessel system in $\mathcal{H}$, then $\{A^t g\}_{g\in\mathcal{G}, t\in[0,L]}$ is a Bessel system in $\mathcal{H}$.*

The fact that $\mathcal{G}$ is a Bessel system in $\mathcal{H}$ implies that $\{A^t g\}_{g\in\mathcal{G}, t\in[0,L]}$ is Bessel in $\mathcal{H}$ is not too surprising. However, the converse implication is not obvious. The next result characterizes the finite dimensional case.

Theorem 2.6.2 (Besselness in finite dimensional space)**.** *Let A be a normal operator on $\mathbb{C}^d$ and L be a positive finite number. Let $M = Range(A^*)$ and $P_M\mathcal{G} = \{P_M g\}_{g\in\mathcal{G}}$, where P_M is the orthogonal projection on M. Then, $\{A^t g\}_{g\in\mathcal{G},t\in[0,L]}$ is a Bessel system in $\mathbb{C}^d$ if and only if $P_M\mathcal{G}$ is a Bessel system in M.*

Under the appropriate restrictions on the spectrum $\sigma(A)$ of A, one can obtain a result similar to Theorem 2.6.2 for the infinite dimensional case. However, if $0 \notin \sigma(A)$, the main result for the infinite dimensional Hilbert space is stated in the following theorem.

Theorem 2.6.3 (Besselness in infinite dimensions)**.** *Let $A \in \mathcal{B}(\mathcal{H})$ be an invertible normal operator, and let $\mathcal{G}$ be a countable system of vectors in $\mathcal{H}$. Then, for any finite positive number L, $\{A^t g\}_{g\in\mathcal{G},t\in[0,L]}$ is a Bessel system in $\mathcal{H}$ if and only if $\mathcal{G}$ is a Bessel system in $\mathcal{H}$.*

The condition that A is invertible is necessary in Theorem 2.6.3 as can be shown by the following example.

Example 1. *Let $\mathcal{G} = \{ne_n\}_{n=1}^{\infty}$ with $\{e_n\}_{n=1}^{\infty}$ being the standard basis of $\ell^2(\mathbb{N})$, $f \in \ell^2(\mathbb{N})$ with $f(n) = 1/n$, and let D be the diagonal infinite matrix with diagonal entries $D_{n,n} = e^{-n^2}$. The operator D is injective but not an invertible operator on $\ell^2(\mathbb{N})$.*

Note that

$$\sum_{g\in\mathcal{G}} |\langle f, g\rangle|^2 = \infty.$$

Hence, $\mathcal{G}$ is not a Bessel system in $\ell^2(\mathbb{N})$. On the other hand,

$$\sum_{g\in\mathcal{G}} \int_0^1 |\langle f, D^t g\rangle| dt = \sum_{n=1}^{\infty} \frac{1-e^{-2n^2}}{2} |f_n|^2 \leqslant \|f\|^2/2. \tag{2.9}$$

Thus $\{D^t g\}_{g\in\mathcal{G},t\in[0,1]}$ is Bessel in $\ell^2(\mathbb{N})$.

2.6.1 Proofs for Section 2.6

Proof of Proposition 2.6.1. For all $f \in \mathcal{H}$,

$$\sum_{g\in\mathcal{G}} \int_0^L |\langle f, A^t g\rangle|^2 dt \quad = \quad \sum_{g\in\mathcal{G}} \int_0^L |\langle A^{*t} f, g\rangle|^2 dt$$

$$= \int_0^L \sum_{g\in\mathcal{G}} |\langle A^{*t}f, g\rangle|^2 dt \leqslant \int_0^L C_{\mathcal{G}} \|A^{*t}f\|^2 dt$$

$$\leqslant \int_0^L C_{\mathcal{G}} \|A\|^{2t}\|f\|^2 dt = \begin{cases} C_{\mathcal{G}} \frac{\|A\|^{2L}-1}{\ln \|A\|^2}\|f\|^2, \|A\| \neq 1 \\ C_{\mathcal{G}} L\|f\|^2, \|A\| = 1, \end{cases}$$

where $C_{\mathcal{G}}$ is a Bessel constant of the Bessel system $\mathcal{G}$. Therefore, $\{A^t g\}_{g\in\mathcal{G}, t\in[0,L]}$ is Bessel in $\mathcal{H}$. □

In order to prove Theorem 2.6.2, we need the following lemma:

Lemma 2.6.4. *Let $\mathcal{G} = \{g_j\}_{j\in J} \subset \mathbb{C}^d$ where J is a countable set. Then, $\mathcal{G}$ is a Bessel system if and only if $\sum_{j\in J} \|g_j\|^2 < \infty$.*

Proof of Lemma 2.6.4. ($\Longrightarrow$)Let $\{u_i\}_{i=1}^d$ be an orthonormal basis in $\mathbb{C}^d$. If $\{g_j\}_{j\in J}$ is a Bessel system with Bessel constant C, then, for $i = 1, \ldots, d$

$$\sum_{j\in J} |\langle u_i, g_j\rangle|^2 \leqslant C.$$

Since $\|g_j\|^2 = \sum_{i=1}^d |\langle u_i, g_j\rangle|^2$ for $j \in J$, one obtains

$$\sum_{j\in J} \|g_j\|^2 = \sum_{j\in J}\sum_{i=1}^d |\langle u_i, g_j\rangle|^2 \leqslant Cd < \infty.$$

($\Longleftarrow$) For any $f \in \mathcal{H}$, one has

$$\sum_{j\in J} |\langle f, g_j\rangle|^2 \leqslant \sum_{j\in J} \|f\|^2\|g_j\|^2 = \|f\|^2(\sum_{j\in J} \|g_j\|^2).$$

Therefore, $\{g_j\}_{j\in J}$ is Bessel in $\mathbb{C}^d$. □

Proof of Theorem 2.6.2. ($\Longleftarrow$) Since A is a normal operator on $\mathcal{H} = \mathbb{C}^d$, it is clear that $A = \sum_{i\in I} \lambda_i P_i$ where $P_iP_j = 0$ for $i \neq j$, $I = \{i : \lambda_i \neq 0\}$, and $(\sum_{i\in I} P_i)(\mathbb{C}^d) = M$, where $M = Range(A^*) = Null^{\perp}(A) = Null^{\perp}(A^*)$.

For $f \in \mathbb{C}^d$, one has

$$\begin{aligned}
\sum_{g\in G}\int_0^L |\langle f, A^t g\rangle|^2 dt &= \sum_{g\in G}\int_0^L \left|\langle A^{*t} f, g\rangle\right|^2 dt = \sum_{g\in G}\int_0^L \left|\sum_{i\in I}\overline{\lambda_i}^t \langle P_i f, P_i g\rangle\right|^2 dt \\
&\leqslant \sum_{g\in G}\int_0^L \|A\|^{2t}\left(\sum_{i\in I}|\langle P_i f, P_i g\rangle|\right)^2 dt \\
&\leqslant \int_0^L \|A\|^{2t}\sum_{g\in G}\left(\sum_{i\in I}\|P_i f\|^2\right)\left(\sum_{i\in I}\|P_i g\|^2\right) dt \\
&\leqslant \int_0^L \|A\|^{2t}\|P_M f\|^2 \sum_{g\in G}\|P_M g\|^2 dt \leqslant C_1 \cdot C_{P_M\mathcal{G}} \cdot \|f\|^2,
\end{aligned}$$

where $C_1 = \begin{cases} (\|A\|^{2L}-1)/\ln(\|A\|^2), \|A\| \neq 1 \\ L, \|A\| = 1 \end{cases}$ and $C_{P_M\mathcal{G}} = \sum_{g\in G}\|P_M g\|^2$.

In addition, one can use Lemma 2.6.4 to conclude that $C_{P_M\mathcal{G}} = \sum_{g\in G}\|P_M g\|^2 < \infty$. Therefore, $\{A^t g\}_{g\in\mathcal{G}, t\in[0,L]}$ is Bessel in $\mathbb{C}^d$.

($\Longrightarrow$) Since A is normal, A can be written as $A = \sum_{i\in I}\lambda_i P_i$, with $\operatorname{rank}(P_i) = 1$ (in this representation, we allow $\lambda_i = \lambda_j$ for $i \neq j$) and $I = \{i : \lambda_i \neq 0\}$, $P_i P_j = 0$ for $i \neq j$, and $(\sum_{i\in I} P_i)(\mathbb{C}^d) = M$. Specifically, by setting $f = u_i$, where u_i is a unit vector in the one dimensional space $P_i(\mathbb{C}^d)$ with $i \in I$, one has

$$\begin{aligned}
\sum_{g\in\mathcal{G}}\int_0^L |\langle u_i, A^t g\rangle|^2 dt &= \sum_{g\in\mathcal{G}}\int_0^L |\langle u_i, \lambda_i^t P_i g\rangle|^2 dt \\
&= \sum_{g\in\mathcal{G}}\int_0^L |\lambda_i|^{2t}\|P_i g\|^2 dt \\
&= \begin{cases} L\sum_{g\in\mathcal{G}}\|P_i g\|^2, & |\lambda_i| = 1 \\ \frac{|\lambda_i|^{2L}-1}{2\ln|\lambda_i|}\sum_{g\in\mathcal{G}}\|P_i g\|^2, & \text{otherwise.} \end{cases}
\end{aligned}$$

In addition, since by assumption $\{A^t g\}_{g\in\mathcal{G}, t\in[0,L]}$ is a Bessel system in $\mathbb{C}^d$ with Bessel constant C, then $\sum_{g\in\mathcal{G}}\int_0^L |\langle u_i, A^t g\rangle|^2 dt \leqslant C\|u_i\|^2 = C$. Hence, for each i,

$$\sum_{g\in\mathcal{G}}\|P_i g\|^2 < \infty.$$

Therefore, summing over (the finitely many) $i \in I$ we obtain

$$\sum_{g \in \mathcal{G}} \|P_M g\|^2 < \infty.$$

If $f \in M = Range(A^*)$, then

$$\begin{aligned}
\sum_{g \in \mathcal{G}} |\langle f, g \rangle|^2 &= \sum_{g \in \mathcal{G}} \left| \sum_{i \in I} \langle P_i f, P_i g \rangle \right|^2 \\
&\leqslant \sum_{g \in \mathcal{G}} \left(\sum_{i \in I} \|P_i f\|^2 \right) \left(\sum_{i \in I} \|P_i g\|^2 \right) \\
&= \|f\|^2 \sum_{g \in \mathcal{G}} \|P_M g\|^2.
\end{aligned}$$

Thus, $P_M \mathcal{G}$ is Bessel in M. $\square$

Before proving Theorem 2.6.3, we first state and prove the following lemmas.

Lemma 2.6.5. *Let $A \in \mathcal{B}(\mathcal{H})$ be an invertible operator in $\mathcal{H}$, then a countable set $\mathcal{G} \subset \mathcal{H}$ is a Bessel system in $\mathcal{H}$ if and only if $\tilde{\mathcal{G}} = A\mathcal{G}$ is a Bessel system in $\mathcal{H}$.*

Proof of Lemma 2.6.5. $(\Longrightarrow)$ For all $f \in \mathcal{H}$,

$$\begin{aligned}
\sum_{g \in \mathcal{G}} |\langle f, Ag \rangle|^2 &= \sum_{g \in \mathcal{G}} |\langle A^* f, g \rangle|^2 \\
&\leqslant C\|A^* f\|_2^2 \leqslant C\|A\|_2^2 \|f\|_2^2,
\end{aligned}$$

where C is a Bessel constant of the Bessel system $\mathcal{G}$. Therefore, $A\mathcal{G}$ is a Bessel system in $\mathcal{H}$.

$(\Longleftarrow)$ For all $f \in \mathcal{H}$,

$$\begin{aligned}
\sum_{g \in \mathcal{G}} |\langle f, g \rangle|^2 &= \sum_{g \in \mathcal{G}} \left|\langle (A^*)^{-1} f, Ag \rangle\right|^2 \\
&\leqslant C_1 \|(A^*)^{-1} f\|_2^2 \leqslant C_1 \|A^{-1}\|_2^2 \|f\|_2^2,
\end{aligned}$$

where C_1 is a Bessel constant of the Bessel system $A\mathcal{G}$. Therefore, $\mathcal{G}$ is a Bessel system in $\mathcal{H}$. $\square$

Proof of Theorem 2.6.3. $(\Longleftarrow)$ See Proposition 2.6.1.

$(\Longrightarrow)$ Since A is a normal operator in $\mathcal{H}$, by the Spectral Theorem, there exists a unitary operator

U such that

$$UAU^{-1} = N_{\mu_\infty}^{(\infty)} \oplus N_{\mu_1}^{(1)} \oplus N_{\mu_2}^{(2)} \oplus \dots$$

and μ is defined as by (2.5). Therefore, the task of proving that $\mathcal{G}$ is a Bessel system in $\mathcal{H}$ is equivalent to the task of showing that $U\mathcal{G}$ is a Bessel system in $\mathcal{W} = U\mathcal{H}$. Let $T : \mathcal{W} \to \mathcal{W}$ be the operator defined by:

$$T\tilde{f}(z) := \int_0^\ell z^t dt \tilde{f}(z), \text{ for all } \tilde{f} \in \mathcal{W} \text{ and } z \in \sigma(A) \text{ with } \ell = \min\{L, 1/2\}. \tag{2.10}$$

The condition that $\ell = \min\{L, 1/2\}$ ensures that T is an invertible operator as will be proved later.

By Lemma 2.6.5, $U\mathcal{G}$ is a Bessel system in $\mathcal{W}$ if and only if $T(U\mathcal{G})$ is a Bessel system in $\mathcal{W}$ as long as T is a bounded invertible normal operator. The fact that T is a bounded invertible operator is stated in the following lemma whose proof is postponed till after the completion of the proof of this theorem.

Lemma 2.6.6. *T is a bounded invertible operator in $\mathcal{W}$.*

So, to finish the proof of Theorem 2.6.3, it only remains to show that $T(U\mathcal{G})$ is a Bessel system in $\mathcal{W}$ which we do next.

Since $\{A^t g\}_{g\in\mathcal{G}, t\in[0,L]}$ is a Bessel system in $\mathcal{H}$, and $0 < \ell \leqslant L$, one has that, for all $f \in \mathcal{H}$,

$$\sum_{g\in\mathcal{G}} \int_0^\ell |\langle f, A^t g\rangle|^2 dt \leqslant C\|f\|^2.$$

Thus, using Hölder's inequality, we get

$$\sum_{g\in\mathcal{G}} \left|\int_0^\ell \langle f, A^t g\rangle dt\right|^2 \leqslant \ell \cdot \sum_{g\in\mathcal{G}} \int_0^\ell |\langle f, A^t g\rangle|^2 dt \leqslant \ell C\|f\|^2. \tag{2.11}$$

In addition,

$$\begin{aligned}
\sum_{g\in\mathcal{G}} \left|\int_0^\ell \langle f, A^t g\rangle dt\right|^2 &= \sum_{g\in\mathcal{G}} \left|\int_0^\ell \int_{\mathbb{C}} \overline{z}^t \langle \tilde{f}(z), \tilde{g}(z)\rangle d\mu(z) dt\right|^2 \\
&= \sum_{g\in\mathcal{G}} \left|\int_{\mathbb{C}} \int_0^\ell \overline{z}^t dt \langle \tilde{f}(z), \tilde{g}(z)\rangle d\mu(z)\right|^2
\end{aligned}$$

$$= \sum_{g\in\mathcal{G}} |\langle \tilde{f}, T\tilde{g}\rangle|^2. \tag{2.12}$$

Together, (2.11) and (2.12) induce the following inequality:

$$\sum_{g\in\mathcal{G}} |\langle \tilde{f}, T\tilde{g}\rangle|^2 \leqslant \ell C\|f\|^2 = \ell C\|\tilde{f}\|^2, \text{ for all } f \in \mathcal{H}.$$

This shows that $T(U\mathcal{G})$ is a Bessel system in $\mathcal{W}$.

In conclusion, by Lemma 2.6.6, T is bounded invertible. In addition, T is normal. Hence, $U\mathcal{G}$ is a Bessel system in $\mathcal{W}$ by Lemma 2.6.5. Consequently, $\mathcal{G}$ is a Bessel in $\mathcal{H}$. □

Proof of Lemma 2.6.6.

$$\begin{aligned}
\|T\tilde{f}\|^2 &= \langle T\tilde{f}, T\tilde{f}\rangle \\
&= \left\langle \int_0^\ell z^t dt \tilde{f}(z), \int_0^\ell z^\tau d\tau \tilde{f}(z)\right\rangle_{L^2(\sigma(A))} \\
&= \int_{\mathbb{C}} \int_0^\ell \int_0^\ell z^t \overline{z}^\tau \langle \tilde{f}(z), \tilde{f}(z)\rangle dt d\tau d\mu(z) \\
&= \int_{\mathbb{C}} |\phi(z)|^2 \|\tilde{f}(z)\|^2 d\mu(z),
\end{aligned}$$

where

$$\phi(z) = \begin{cases} \ell, & z = 1 \\ 0, & z = 0 \\ \frac{z^\ell - 1}{\ln(z)}, & \text{otherwise} \end{cases} . \tag{2.13}$$

Let $m = \inf\{|\phi(z)| : z \in \sigma(A)\}$ and $M = \sup\{|\phi(z)| : z \in \sigma(A)\}$. As shown below in claim 2.6.7, $m > 0$ and $M < \infty$. Thus

$$\begin{aligned}
\|T\tilde{f}\|^2 &\leqslant \int_{\mathbb{C}} M^2 \|\tilde{f}(z)\|^2 d\mu(z) = M^2\|\tilde{f}\|^2, \\
\|T\tilde{f}\|^2 &\geqslant \int_{\mathbb{C}} m^2 \|\tilde{f}(z)\|^2 d\mu(z) = m^2\|\tilde{f}\|^2, \text{ for all } \tilde{f} \in \mathcal{W}.
\end{aligned}$$

Since T is normal, it follows that T is a bounded invertible operator (see [52, Theorem 12.12]). We finish by proving the following fact that was used in the proof of this lemma.

Claim 2.6.7. *Let ϕ be the function defined in* (2.13). *Then* $M = \sup\{|\phi(z)| : z \in \sigma(A)\} < \infty$, *and* $m = \inf\{|\phi(z)| : z \in \sigma(A)\} > 0$.

Proof of Claim 2.6.7. Since A is a bounded invertible normal operator, it follows that $\|A^{-1}\|^{-1} \leqslant |z| \leqslant \|A\|$ for $z \in \sigma(A)$. Let $S = \{z \in \mathbb{C} : \|A^{-1}\|^{-1} \leqslant |z| \leqslant \|A\|\}$. Since $\sigma(A) \subset S$, $M \leqslant \sup\{|\phi(z)| : z \in S\}$ and $m \geqslant \{|\phi(z)| : z \in S\}$. Therefore, in order to prove Claim 2.6.7, it is sufficient to show that $\sup\{|\phi(z)| : z \in S\} < \infty$, $\inf\{|\phi(z)| : z \in S\} > 0$.

To prove that $\sup\{|\phi(z)| : z \in S\} < \infty$, it is noteworthy that

$$|\phi(z)| = \left|\int_0^\ell z^t dt\right| \leqslant \int_0^\ell |z^t| dt = \int_0^\ell |z|^t dt = \begin{cases} \ell, & z \in S \text{ and } |z| = 1 \\ \frac{|z|^\ell - 1}{\ln|z|}, & z \in S \text{ and } |z| \neq 1. \end{cases}$$

Let

$$\psi(x) = \begin{cases} \ell, x = 1 \\ \frac{x^\ell - 1}{\ln x}, x \in \mathbb{R}^+ \backslash \{1\}, \end{cases}$$

and note that (since $\lim_{x \to 1} \frac{x^\ell - 1}{\ln x} = \ell = \psi(1)$) ψ is continuous at $x = 1$. In addition, $\frac{x^\ell - 1}{\ln x}$ is a continuous function on $\mathbb{R}^+ \backslash \{1\}$. Hence, ψ is continuous on $\mathbb{R}^+$. Particularly, ψ is continuous on the closed interval $[\|A^{-1}\|^{-1}, \|A\|]$. Therefore,

$$\sup\{|\phi(z)| : z \in S\} = \max_{x \in [\|A^{-1}\|^{-1}, \|A\|]} \psi(x) < \infty.$$

Finally, it remains to show that $\inf\{|\phi(z)| : z \in S\} > 0$. First, we divide S into two sets with $S_1 = \{z \in S : \arg(z) \in [-\pi/2, \pi/2]\}$ and $S_2 = S \backslash S_1$. Since $|\phi(z)|$ is a continuous function on S_1 and S_1 is compact, there exists $z_0 \in S_1$ such that $|\phi(z_0)| = \inf\{|\phi(z)| : z \in S_1\}$. In addition, $|\phi(z)|$ has no root on S_1. Hence, $\inf\{|\phi(z)| : z \in S_1\} > 0$.

For $z \in S_2$, $\pi/2 \leqslant |\arg(z)| \leqslant \pi$. Therefore,

$$\begin{aligned} |z^\ell - 1| &= ||z|^\ell e^{i\ell \arg(z)} - 1| \\ &\geqslant |z|^\ell |\sin(\ell \arg(z))| \end{aligned}$$

$$\geqslant \min\{\|A^{-1}\|^{-\ell}\sin(\ell\pi), \|A^{-1}\|^{-\ell}\sin(\ell\pi/2)\} > 0,$$

where the last inequality follows from the fact that $0 < \ell < 1$ (in particular, we chose $\ell = \min\{L, 1/2\}$ as in Definition (2.10) for T). In addition, for $z \in S_2$, one has

$$|\ln(z)| \leqslant |\ln(|z|)| + |\arg(z)| \leqslant \max\{|\ln(\|A^{-1}\|^{-1})|, |\ln(\|A\|)|\} + \pi < \infty.$$

Hence, $\inf\{|\phi(z)| : z \in S_2\} > 0$. Combining the estimates on S_1 and S_2, we conclude that $\inf\{|\phi(z)| : z \in S\} > 0$. □

2.7 Frames generated by the action of bounded normal operators.

In this section, we study some properties of a semi-continuous frame of the form $\{A^t g\}_{g\in\mathcal{G},t\in[0,L]}$ generated by the continuous action of a normal operator $A \in \mathcal{B}(\mathcal{H})$ and relate them to the properties of the discrete systems generated by its time discretization. We also show that, under the appropriate conditions, if $\{A^t g\}_{g\in\mathcal{G},t\in[0,L_1]}$ is a semi-continuous frame for some positive number L_1, then $\{A^t g\}_{g\in\mathcal{G},t\in[0,L]}$ a semi-continuous frame for all $0 < L < \infty$. Before presenting the two main theorems, we first provide some necessary conditions for obtaining semi-continuous frames, and treat some special cases. The proofs are postponed to Subsection 2.7.1.

The following proposition (whose proof is obtained by direct calculation) provides a necessary condition to ensure the lower bound of the semi-continuous frame generated by $A \in \mathcal{B}(\mathcal{H})$.

Proposition 2.7.1. *Let $A \in \mathcal{B}(\mathcal{H})$ be an invertible normal operator, L be a finite positive number, and $\mathcal{G} \subset \mathcal{H}$ be a countable set of vectors. If, for all $f \in \mathcal{H}$,*

$$\sum_{g\in\mathcal{G}} |\langle f, g\rangle|^2 \geqslant c\|f\|^2, \tag{2.14}$$

where c is a positive constant, then there exists a finite positive constant C such that

$$\sum_{g\in\mathcal{G}} \int_0^L |\langle f, A^t g\rangle|^2 dt \geqslant C\|f\|^2, \text{ for all } f \in \mathcal{H}. \tag{2.15}$$

The converse of Proposition 2.7.1 is false, even in finite dimensional space as shown in Example

2. For the special case that A is equivalent to a diagonal operator on $\ell^2(\mathbb{N})$ we get:

Lemma 2.7.2. *Let $A \in \mathcal{A}$, where $\mathcal{A}$ is defined in* (2.7)*, and let $\mathcal{G} \subset \ell^2(\mathbb{N})$ be a countable set of vectors. If $\{A^t g\}_{g\in\mathcal{G},t\in[0,L]}$ satisfies* (2.15) *in $\ell^2(\mathbb{N})$, then*

$$\sum_{g\in\mathcal{G}} \|g\|^2 = \infty.$$

From Lemma 2.7.2, it follows that the cardinality of $\mathcal{G}$ must be infinite as stated in the following corollary.

Corollary 2.7.3. *If the assumptions of Lemma 2.7.2 hold then $|\mathcal{G}| = +\infty$. In particular, $|\mathcal{G}| = +\infty$ if $\{A^t g\}_{g\in\mathcal{G},t\in[0,L]}$ is a frame for $\ell^2(\mathbb{N})$.*

The discretization of continuous frames is a central question and has been studied extensively (see [31, 32] and the references therein). In particular, Freeman and Speegle have found necessary and sufficient conditions for the discretization of continuous frames [32]. In our situation, the systems $\{A^t g\}_{g\in\mathcal{G},t\in[0,L]}$ can be viewed as continuous frames and the theory in [32] may be applied to conclude that the system can be discretized. However, because of the particular structure of the systems $\{A^t g\}_{g\in\mathcal{G},t\in[0,L]}$, we can say more and obtain finer results for their discretization, as stated in the following theorem.

Theorem 2.7.4. *Let $A \in \mathcal{B}(\mathcal{H})$ be a normal operator on the Hilbert space $\mathcal{H}$ and let $\mathcal{G}$ be a Bessel system of vectors in $\mathcal{H}$. If $\{A^t g\}_{g\in\mathcal{G},t\in[0,L]}$ is a semi-continuous frame for $\mathcal{H}$, then there exists $\delta > 0$ such that for any finite set $T = \{t_i : i = 1, \ldots, n\}$ with $0 = t_1 < t_2 < \ldots < t_n < t_{n+1} = L$ and $|t_{i+1} - t_i| < \delta$, the system $\{A^t g\}_{g\in\mathcal{G},t\in T}$ is a frame for $\mathcal{H}$.*

If, in addition, A is invertible, then $\{A^t g\}_{g\in\mathcal{G},t\in[0,L]}$ is a semi-continuous frame for $\mathcal{H}$ if and only if there exists a finite set $T = \{t_i : i = 1, \ldots, n\}$ and $0 = t_1 < t_2 < \ldots < t_n < L$, such that $\{A^t g\}_{g\in\mathcal{G},t\in T}$ is a frame for $\mathcal{H}$.

Example 3 shows that the condition that A is invertible is necessary for the second statement of Theorem 2.7.4.

The next theorem shows that, under some appropriate conditions, if $\{A^t g\}_{g\in\mathcal{G},t\in[0,L_1]}$ is a semi-continuous frame for some finite positive number L_1, then $\{A^t g\}_{g\in\mathcal{G},t\in[0,L]}$ is a semi-continuous frame for any finite positive number L.

Theorem 2.7.5. *Let $A \in \mathcal{B}(\mathcal{H})$ be an invertible self-adjoint operator and $\mathcal{G}$ be a countable set in $\mathcal{H}$. Then, $\{A^t g\}_{g\in\mathcal{G},t\in[0,1]}$ is a semi-continuous frame in $\mathcal{H}$ if and only if $\{A^t g\}_{g\in\mathcal{G},t\in[0,L]}$ is a semi-continuous frame in $\mathcal{H}$ for all finite positive L.*

We postulate the following conjecture:

Conjecture 2.7.6. *Theorem 2.7.5 remains true if A is a normal reductive operator.*

This first example shows that the converse of Proposition 2.7.1 is false.

Example 2. *Let $A = \begin{bmatrix} \epsilon & 0 \\ 0 & 1 \end{bmatrix}$ with $0 < \epsilon < 1$ and $g = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.*

Note that for $L > 0$,

$$\mathcal{G}_1 = \left\{ g = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, A^{L/2} g = \begin{bmatrix} \epsilon^{L/2} \\ 1 \end{bmatrix} \right\}$$

is complete in $\mathbb{R}^2$. In addition, A is a bounded invertible normal operator in $\mathbb{R}^2$. Therefore, $\mathcal{G}_1$ is a frame in $\mathbb{R}^2$. By Theorem 2.7.4, $\{A^t g\}_{t\in[0,L]}$ is a semi-continuous frame in $\mathbb{R}^2$. However, the lower bound of (2.14) *does not hold for $\mathcal{G} = \{g\}$. For example, let $f = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, then $\langle f, g \rangle = 0$.*

This next example shows that the condition that A is invertible is required for the second statement of Theorem 2.7.4.

Example 3. *Let $\mathcal{G} = \{e_j\}_{j=1}^{\infty}$ be the standard basis of $\ell^2(\mathbb{N})$. Because $\mathcal{G}$ is an orthonormal basis, one has $\mathcal{G} \subset \{A^t g\}_{g\in\mathcal{G},t\in T}$, for any bounded operator A, and for any time steps $T = \{t_i : i = 1, \ldots, n\}$ with $0 = t_1 < t_2 < \ldots < t_n < L$. Thus, $\mathcal{G} \subset \{A^t g\}_{g\in\mathcal{G},t\in T}$ is a frame for $\ell^2(\mathbb{N})$.*

However, there exists a non-trivial bounded operator such that $\{A^t e_j\}_{j\in\mathbb{N},t\in[0,L]}$ is not a semi-continuous frame. For example, if D is a diagonal infinite matrix with diagonal entries $D_{j,j} = \frac{1}{j}$, then

$$\sum_{j=1}^{\infty} \int_0^L |\langle e_k, D^t e_j \rangle|^2 dt = \frac{1/k^{2L} - 1}{\ln(1/k^2)}. \tag{2.16}$$

Since $\lim\limits_{k\to\infty} \frac{1/k^{2L}-1}{\ln(1/k^2)} = 0$, it follows that $\{D^t e_j\}_{j\in\mathbb{N},t\in[0,L]}$ is not a semi-continuous frame for $\ell^2(\mathbb{N})$.

Additionally, a number of examples are available to illustrate that $\{A^t g\}_{g\in\mathcal{G},t\in[0,L]}$ is a semi-continuous frame for $\mathcal{H}$ does not require $\mathcal{G}$ to be a frame or even complete in $\mathcal{H}$. In fact, this is precisely why space-time sampling trade-off is feasible. The next two examples are toy examples to show this fact.

Example 4 ($\mathcal{G}$ is not a frame for $\mathcal{H}$)**.** *Let $\{e_n\}_{n=1}^{\infty}$ be the standard basis of $\ell^2(\mathbb{N})$ and $\mathcal{G} = \{g_n = e_n + e_{n+1} : n \in \mathbb{N}\}$, and let D be a diagonal operator with* $D_{n,n} = \begin{cases} 1, n \text{ is odd} \\ 3, n \text{ is even} \end{cases}$.

It can be shown that $\mathcal{G}$ is complete but that $\mathcal{G}$ is neither a basis nor a frame for $\ell^2(\mathbb{N})$ [23]. However, for all $f \in \ell^2(\mathbb{N})$, after a somewhat tedious computation, one gets

$$\frac{1}{2}\|f\|^2 \leqslant \sum_{n=1}^{\infty} \int_0^1 |\langle f, D^t g_n\rangle|^2 dt \leqslant \frac{16}{\ln(3)}\|f\|^2,$$

so that $\{D^t g_n\}_{n\in\mathbb{N}, t\in[0,1]}$ is a semi-continuous frame for $\ell^2(\mathbb{N})$.

Example 5 ($\mathcal{G}$ is not complete in $\mathcal{H}$)**.** *Let $\{e_n\}_{n=1}^{\infty}$ be the standard basis of $\ell^2(\mathbb{N})$ and $\mathcal{G} = \{g_n = e_n + 2e_{n+1} : n \in \mathbb{N}\}$. The set $\mathcal{G}$ is not complete in $\ell^2(\mathbb{N})$. For example $f = (f_k)$ with $f_k = (-1)^k \frac{1}{2^k}$ is orthogonal to $\overline{span}\,\mathcal{G}$. Thus, $\mathcal{G}$ is not a frame in $\ell^2(\mathbb{N})$. Let D be the diagonal operator with*

$$D_{n,n} = \begin{cases} 9, & n = 1 \\ 1 - \frac{1}{n}, & n \geqslant 2 \end{cases}.$$

A lengthy computation yields

$$\frac{1}{4}\|f\|^2 \leqslant \sum_{n=1}^{\infty} |\langle f, g_n\rangle|^2 + \sum_{n=1}^{\infty} |\langle f, Dg_n\rangle|^2 \leqslant 164\|f\|^2.$$

This implies that $\{D^t g\}_{g\in\mathcal{G}, t\in\{0,1\}}$ is a frame in $\ell^2(\mathbb{N})$. In addition, since D is a self-adjoint invertible operator, Theorem 2.7.4 implies that $\{D^t g_n\}_{n\in\mathbb{N}, t\in[0,2]}$ is a semi-continuous frame of $\ell^2(\mathbb{N})$.

2.7.1 Proofs of Section 2.7

Proof of Lemma 2.7.2. One can always assume that $A = \sum_{i=1}^{\infty} \lambda_i P_i$ with $\operatorname{rank}(P_i) = 1$, $P_i P_j = 0$ and $\sum_i P_i = Id_{\ell^2(\mathbb{N})}$ as long as $\lambda_i = \lambda_j$ for $i \neq j$ in the representation of A is allowed. Let e_i be a vector such that $\|e_i\| = 1$ and $span\{e_i\} = P_i(\mathcal{H})$. Then

$$\sum_{g\in\mathcal{G}} \int_0^L |\langle e_i, A^t g\rangle|_2^2 dt \quad = \quad \sum_{g\in\mathcal{G}} \int_0^L |\lambda_i|^{2t} |\langle e_i, P_i(g)\rangle|^2 dt.$$

Since $\{A^t g\}_{g\in\mathcal{G}, t\in[0,L]}$ satisfies (2.15), we have that $\lambda_i \neq 0$ for all $i \in \mathbb{N}$. Moreover, if $\sum_{g\in\mathcal{G}} \|g\|_2^2 = \sum_{i\in\mathbb{N}} \sum_{g\in\mathcal{G}} \|P_i g\|^2 < \infty$, then $\lim\limits_{i\to\infty} \sum_{g\in\mathcal{G}} \|P_i g\|^2 = 0$. In addition, since $\frac{\|A\|^{2L}-1}{2\ln(\|A\|)} \geqslant \frac{|\lambda_i|^{2L}-1}{2\ln(|\lambda_i|)} > 0$, we get that $\lim\limits_{i\to\infty} \frac{|\lambda_i|^{2L}-1}{2\ln(|\lambda_i|)} \sum_{g\in\mathcal{G}} \|P_i g\|^2 = 0$. This contradicts (2.15). Hence, $\sum_{g\in\mathcal{G}} \|g\|^2 = \infty$. □

Proof of theorem 2.7.4. From the assumption that $\mathcal{G}$ is a Bessel sequence in $\mathcal{H}$, there exists $K > 0$ such that $\sum_{g\in\mathcal{G}} |\langle f, g\rangle|^2 \leqslant K\|f\|^2$, for all $f \in \mathcal{H}$. Since A is a bounded normal operator, for any $0 \leqslant t < \infty$, one has

$$\sum_{g\in\mathcal{G}} |\langle f, A^t g\rangle|^2 = \sum_{g\in\mathcal{G}} |\langle A^{*t} f, g\rangle|^2 \leqslant K\|A^{*t} f\|^2 \leqslant K\|A\|^{2t}\|f\|^2. \tag{2.17}$$

Summing the inequalities (2.17) over $t \in T = \{t_i : i = 1, \ldots, n\}$, it immediately follows that $\{A^t g\}_{g\in\mathcal{G}, t\in T}$ is a Bessel sequence in $\mathcal{H}$.

Using (2.17), it follows that

$$\sum_{g\in\mathcal{G}} \int_0^L |\langle f, A^t g\rangle|^2 dt \leqslant K \int_0^L \|A\|^{2t} dt \|f\|^2. \tag{2.18}$$

Inequality (2.18) implies that for any $\epsilon > 0$, there exists an l with $L/2 > l > 0$, such that

$$\sum_{g\in\mathcal{G}} \int_0^l |\langle f, A^t g\rangle|^2 dt < \epsilon\|f\|^2. \tag{2.19}$$

Next, the goal is to find $\delta > 0$ such that for any finite set $T = \{t_i : i = 1, \ldots, n\}$ with $0 = t_1 < t_2 < \ldots < t_n < t_{n+1} = L$ and $|t_{i+1} - t_i| < \delta$, the system $\{A^t g\}_{g\in\mathcal{G}, t\in T}$ is a frame for $\mathcal{H}$, as long as $\{A^t g\}_{g\in\mathcal{G}, t\in[0,L]}$ is a semi-continuous frame for $\mathcal{H}$, i.e.,

$$c\|f\|^2 \leqslant \sum_{g\in\mathcal{G}} \int_0^L |\langle f, A^t g\rangle|^2 dt \leqslant C\|f\|^2, \quad \text{for all } f \in \mathcal{H}, \tag{2.20}$$

for some $c, C > 0$.

To finish the proof, we use the following lemma.

Lemma 2.7.7. *Let $A \in \mathcal{B}(\mathcal{H})$ be a normal operator and ℓ, L be positive numbers with $0 < \ell < L$. Then for any $\epsilon > 0$, there exists $\delta > 0$ such that whenever $s_1, s_2 \in [\ell, L]$ with $|s_1 - s_2| < \delta$, we have $\|A^{s_1} - A^{s_2}\| < \epsilon$.*

Proof of Lemma 2.7.7. For $s_1, s_2 \in [\ell, L]$,

$$\begin{aligned} |z^{s_1} - z^{s_2}|^2 &= |z|^{2s_1} - 2|z|^{s_1}|z|^{s_2}\cos((s_1 - s_2)arg(z)) + |z|^{2s_2} \\ &= ||z|^{s_1} - |z|^{s_2}|^2 + 2|z|^{s_1}|z|^{s_2}(1 - \cos((s_1 - s_2)arg(z))). \end{aligned}$$

For all $z \in \sigma(A)$, one has $0 \leqslant |z| \leqslant \|A\|$. Thus $|z|^s$ is uniformly bounded for all $s \in [\ell, L]$. In addition, the function $(t, r) \mapsto r^t$ is a continuous function on the compact set $[\ell, L] \times [0, \|A\|]$ and the function $t \mapsto \cos(t \cdot arg(z))$ is equicontinuous at $t = 0$ for $arg(z) \in [-\pi, \pi)$. The lemma then follows from the spectral theorem (i.e., Theorem 2.3.3).

□

By Lemma 2.7.7, there exists δ with $l/2 > \delta > 0$ such that whenever $|s_1 - s_2| < 2 \cdot \delta$ for $s_1, s_2 \in [l/2, L]$, then $\|A^{s_1} - A^{s_2}\| < \epsilon$. Assume that the set $T = \{t_i : i = 1, \dots, n\}$ satisfies $0 = t_1 < t_2 < \dots < t_n < t_{n+1} = L$ and $|t_{i+1} - t_i| < \delta$. Set $m = \min\{i : t_i > l/2\}$. Note that $l/2 > \delta > 0$. Therefore $t_m < l$. Then, using (2.19), the difference

$$\Delta = \left|\sum_{g \in \mathcal{G}} \int_0^L |\langle f, A^t g\rangle|^2 dt - \sum_{g \in \mathcal{G}} \sum_{i=m}^{n} \int_{t_i}^{t_{i+1}} |\langle f, A^{t_i} g\rangle|^2 dt\right|, \tag{2.21}$$

can be estimated as follows.

$$\begin{aligned} \Delta &= \left|\sum_{g \in \mathcal{G}} \int_0^L |\langle f, A^t g\rangle|^2 dt - \sum_{g \in \mathcal{G}} \sum_{i=m}^{n} \int_{t_i}^{t_{i+1}} |\langle f, A^{t_i} g\rangle|^2 dt\right| \\ &\leqslant \left(\sum_{g \in \mathcal{G}} \int_0^{t_m} |\langle f, A^t g\rangle|^2 dt\right) + \sum_{i=m}^{n} \int_{t_i}^{t_{i+1}} \sum_{g \in \mathcal{G}} ||\langle f, A^t g\rangle|^2 - |\langle f, A^{t_i} g\rangle|^2| dt \\ &= \left(\int_0^{t_m} \sum_{g \in \mathcal{G}} |\langle f, A^t g\rangle|^2 dt\right) + \sum_{i=m}^{n} \int_{t_i}^{t_{i+1}} \sum_{g \in \mathcal{G}} (|\langle f, A^t g\rangle| + |\langle f, A^{t_i} g\rangle|)(||\langle f, A^t g\rangle| - |\langle f, A^{t_i} g\rangle||) dt \\ &\leqslant \epsilon\|f\|^2 + \sum_{i=m}^{n} \int_{t_i}^{t_{i+1}} \sum_{g \in \mathcal{G}} (|\langle A^{*t} f, g\rangle| + |\langle A^{*t_i} f, g\rangle|)(|\langle A^{*t} f - A^{*t_i} f, g\rangle|) dt \end{aligned}$$

$$
\begin{aligned}
&\leqslant \quad \epsilon\|f\|^2 + \sum_{i=m}^{n} \int_{t_i}^{t_{i+1}} \left(\sum_{g\in\mathcal{G}} (|\langle A^{*t} f, g\rangle| + |\langle A^{*t_i} f, g\rangle|)^2 \right)^{1/2} \left(\sum_{g\in\mathcal{G}} (|\langle A^{*t} f - A^{*t_i} f, g\rangle|)^2 \right)^{1/2} dt \\
&\leqslant \quad \epsilon\|f\|^2 + \sum_{i=m}^{n} \int_{t_i}^{t_{i+1}} \left(2K(\|A^{*t} f\|^2 + \|A^{*t_i} f\|^2)\right)^{1/2} (K\|A^{*t} f - A^{*t_i} f\|^2)^{1/2} dt \\
&\leqslant \quad (\epsilon + 2C_1 KL\epsilon)\|f\|^2, \text{ where } C_1 = \max\{1, \|A\|^L\}.
\end{aligned}
$$

Using (2.21) and choosing ϵ so small that $(1 + 2C_1KL)\epsilon < c/2$, we find δ such that

$$
\delta \sum_{g\in\mathcal{G}} \sum_{i=m}^{n} |\langle f, A^{t_i} g\rangle|^2 \geqslant c\|f\|^2 - c/2\|f\|^2 = c/2\|f\|^2.
$$

Therefore, for any finite set $T = \{t_i : i = 1, \ldots, n\}$ with $0 = t_1 < t_2 < \ldots < t_n < t_{n+1} = L$ and $|t_{i+1} - t_i| < \delta$, the system $\{A^t g\}_{g\in\mathcal{G}, t\in T}$ is a frame in $\mathcal{H}$.

To prove the second statement, it is sufficient to prove that $\{A^t g\}_{g\in\mathcal{G}, t\in[0,L]}$ is a semi-continuous frame under the assumption that $\{A^t g\}_{g\in\mathcal{G}, t\in T}$ is a frame in $\mathcal{H}$ and A is an invertible normal operator. We already know by Theorem 2.6.3 that $\{A^t g\}_{g\in\mathcal{G}, t\in[0,L]}$ is Bessel since $\mathcal{G}$ is Bessel by assumption. Let $T = \{t_i : i = 1, \ldots, n\}$ with $0 = t_1 < t_2 < \ldots < t_n < L$ be such that $\{A^t g\}_{g\in\mathcal{G}, t\in T}$ is a frame for $\mathcal{H}$ with frame constants c, C i.e., for all $f \in \mathcal{H}$,

$$
c\|f\|^2 \leqslant \sum_{g\in\mathcal{G}} \sum_{i=1}^{n} |\langle f, A^{t_i} g\rangle| \leqslant C\|f\|^2.
$$

Let $m = \min\{t_{i+1} - t_i, 1 \leqslant i \leqslant n\}$ with $t_{n+1} = L$. Then,

$$
\begin{aligned}
\sum_{g\in\mathcal{G}} \int_0^L |\langle f, A^t g\rangle|^2 dt &= \sum_{g\in\mathcal{G}} \sum_{i=1}^{n} \int_{t_i}^{t_{i+1}} |\langle f, A^t g\rangle|^2 dt \\
&= \sum_{g\in\mathcal{G}} \sum_{i=1}^{n} \int_0^{t_{i+1}-t_i} |\langle (A^{*t} f, A^{t_i} g\rangle|^2 dt \\
&\geqslant \sum_{g\in\mathcal{G}} \sum_{i=1}^{n} \int_0^{m} |\langle A^{*t} f, A^{t_i} g\rangle|^2 dt \\
&\geqslant \int_0^m c\|A^{*t} f\|_2^2 dt.
\end{aligned}
$$

Since A is an invertible bounded normal operator, we have

$$\int_0^m c\|A^{*t}f\|_2^2 dt \quad \geqslant \quad c \cdot \frac{1-\|A^{-1}\|^{-2m}}{2\ln(\|A^{-1}\|)}\|f\|^2.$$

This concludes the proof that $\{A^t g\}_{g\in\mathcal{G},t\in[0,L]}$ is a semi-continuous frame for $\mathcal{H}$. $\square$

To prove Theorem 2.7.5, the following three lemmas, i.e., Lemmas 2.7.8, 2.7.9 and 2.7.10 are needed.

Lemma 2.7.8. *Let $\mathcal{G}$ be a countable Bessel sequence in $\mathcal{H}$ and let $A \in \mathcal{B}(\mathcal{H})$ be a normal operator. Let L be any positive real number, $\Omega_L = \{z : \Re(z) > L > 0\}$, and let $\{g_i\}_{i\in I}$ be any indexing of $\mathcal{G}$. Then, for fixed $f \in \mathcal{H}$, the partial sums $\sum\limits_{i=1}^{n} |\langle A^z g_i, f\rangle|^2$ converge uniformly on any compact subset of Ω_L.*

Proof of Lemma 2.7.8. Let $\overline{D_r}$ denote the closed disk of radius r. Then using the fact that $\mathcal{G}$ is Bessel with Bessel constant $C_{\mathcal{G}}$, for $z \in \overline{D_r} \cap \overline{\Omega_L}$, one gets,

$$\sum_{i=1}^{n} |\langle A^z g_i, f\rangle|^2 = \sum_{i=1}^{n} |\langle f, A^z g_i\rangle|^2 = \sum_{i=1}^{n} |\langle (A^z)^* f, g\rangle|^2 \leqslant C_{\mathcal{G}} \cdot e^{2\pi r} \cdot \|A\|^{2r}\|f\|^2,$$

from which the lemma follows. $\square$

Lemma 2.7.9. *Let $\mathcal{G}$ be a countable Bessel sequence in $\mathcal{H}$ and let $A \in \mathcal{B}(\mathcal{H})$ be a normal operator. Let L be any positive real number and let $\Omega_L = \{z : \Re(z) > L > 0\}$. Then, for fixed $f \in \mathcal{H}$,*

$$F(z) = \sum_{g\in\mathcal{G}} (\langle A^z g, f\rangle)^2,$$

is an analytic function of z in Ω_L.

Proof of Lemma 2.7.9. Since A is a normal operator on $\mathcal{H}$, by Lemma 2.5.4, $(\langle A^z g, f\rangle)^2$ is analytic in Ω_L. Since $\left|\sum_{g\in\mathcal{G}} (\langle A^z g, f\rangle)^2\right| \leqslant \sum_{g\in\mathcal{G}} |\langle A^z g, f\rangle|^2$, by Lemma 2.7.8, the series $\sum_{g\in\mathcal{G}} (\langle A^z g, f\rangle)^2$ converges absolutely and uniformly on any compact subset of Ω_L, and the partial sums of $\sum_{g\in\mathcal{G}} (\langle A^z g, f\rangle)^2$ are analytic in Ω_L and converge uniformly on any compact subset of Ω_L. It follows that the series $\sum_{g\in\mathcal{G}}(\langle A^z g, f\rangle)^2$ is an analytic function of z in Ω_L [51, Theorem 10.28]. $\square$

Let $A \in \mathcal{B}(\mathcal{H})$ be a normal operator, by the spectral theorem, there exists a unitary operator U such that

$$UAU^{-1} = N^{(\infty)}_{\mu_\infty} \oplus N^{(1)}_{\mu_1} \oplus N^{(2)}_{\mu_2} \oplus \ldots.$$

For every $f \in \mathcal{H}$, we define $\tilde{f} = Uf \in U\mathcal{H}$. Note that $\tilde{f} : \sigma(A) \to \ell^2(\Omega_\infty) \oplus \ell^2(\Omega_1) \oplus \ell^2(\Omega_2) \oplus \ldots$ is a function and hence it makes sense to talk about its real and imaginary parts. Set $f^{\Re} = U^{-1}\Re(\tilde{f})$ and $f^{\Im} = U^{-1}\Im(\tilde{f})$.

Lemma 2.7.10. *If $\mathcal{G}$ is a Bessel sequence in $\mathcal{H}$, then, $\{g^{\Re}\}_{g\in\mathcal{G}}$ and $\{g^{\Im}\}_{g\in\mathcal{G}}$ are also Bessel sequences in $\mathcal{H}$ for any given normal operator $A \in \mathcal{H}$.*

Proof of Lemma 2.7.10. Consider the subspace $S \subset \mathcal{H}$ defined by $S = \{f \in \mathcal{H} : Uf \text{ is real valued}\}$. Then, for $f \in S$, using the following identity

$$\sum_{g\in\mathcal{G}} |\langle f, g\rangle|^2 = \sum_{g\in\mathcal{G}} |\langle \tilde{f}, \tilde{g}\rangle| = \sum_{g\in\mathcal{G}} |\langle \tilde{f}, \Re(\tilde{g})\rangle|^2 + |\langle \tilde{f}, \Im(\tilde{g})\rangle|^2 = \sum_{g\in\mathcal{G}} |\langle f, g^{\Re}\rangle|^2 + |\langle f, g^{\Im}\rangle|^2,$$

it follows that $\{g^{\Re}\}_{g\in\mathcal{G}}$ and $\{g^{\Im}\}_{g\in\mathcal{G}}$ are Bessel sequences in S. For general $f \in \mathcal{H}$, we have $f^{\Re} \in S$, $f^{\Im} \in S$, and

$$\sum_{g\in\mathcal{G}} |\langle f, g^{\Re}\rangle|^2 = \sum_{g\in\mathcal{G}} |\langle f^{\Re}, g^{\Re}\rangle|^2 + \sum_{g\in\mathcal{G}} |\langle f^{\Im}, g^{\Re}\rangle|^2,$$

$$\sum_{g\in\mathcal{G}} |\langle f, g^{\Im}\rangle|^2 = \sum_{g\in\mathcal{G}} |\langle f^{\Re}, g^{\Im}\rangle|^2 + \sum_{g\in\mathcal{G}} |\langle f^{\Im}, g^{\Im}\rangle|^2.$$

It follows that $\{g^{\Re}\}_{g\in\mathcal{G}}$ and $\{g^{\Im}\}_{g\in\mathcal{G}}$ are Bessel sequences for $\mathcal{H}$.

$\square$

Proof of Theorem 2.7.5. Assume that $\{A^t g\}_{g\in\mathcal{G},t\in[0,1]}$ is a semi-continuous frame in $\mathcal{H}$ with frame bounds c, C. By Theorem 2.7.4, there exists a finite set T such that $\{A^t g\}_{g\in\mathcal{G},t\in T}$ is a frame for $\mathcal{H}$. Therefore, for $L \geqslant 1$, $\{A^t g\}_{g\in\mathcal{G},t\in[0,L]}$ is also a semi-continuous frame.

To prove that $\{A^t g\}_{g\in\mathcal{G},t\in[0,L]}$ is a semi-continuous frame for $L<1$, we note that the inequality

$$\sum_{g\in\mathcal{G}}\int_0^L |\langle f, A^t g\rangle|^2 dt \leqslant \sum_{g\in\mathcal{G}}\int_0^1 |\langle f, A^t g\rangle|^2 dt \leqslant C\|f\|_2^2$$

implies that $\{A^t g\}_{g\in\mathcal{G},t\in[0,L]}$ is a Bessel system in $\mathcal{H}$. Moreover, A is an invertible bounded self-adjoint operator. Therefore, by Theorem 2.6.3, $\mathcal{G}$ is Bessel in $\mathcal{H}$ with Bessel constant $C_{\mathcal{G}}$.

Suppose that $\{A^t g\}_{g\in\mathcal{G},t\in[0,L]}$ is not a frame. Then, there exists a sequence $\{f_n\}$ with $\|f_n\|=1$ such that $\sum_{g\in\mathcal{G}}\int_0^L |\langle f_n, A^t g\rangle|^2 dt \to 0$. It follows that $\sum_{g\in\mathcal{G}} |\langle f_n, A^t g\rangle|^2 \to 0$ in measure. Thus, there exists a subsequence $\{f_{n_k}\}$ of $\{f_n\}$ such that $\sum_{g\in\mathcal{G}} |\langle f_{n_k}, A^t g\rangle|^2 \to 0$, for a.e. $t\in[0,L]$. By passing to a subsequence, assume that $\sum_{g\in\mathcal{G}} |\langle f_n, A^t g\rangle|^2 \to 0$, for a.e. $t\in[0,L]$.

To finish the proof, we next prove that there exists a subsequence $\{f_{n_k}\}$ of $\{f_n\}$ such that

$$\sum_{g\in\mathcal{G}}\int_0^1 |\langle f_{n_k}, A^t g\rangle|^2 dt \to 0.$$

Since A is a self-adjoint operator, by the spectral theorem, there exists a unitary operator U such that A can be represented as (2.4) and $\sigma(A)\subset\mathbb{R}$. In addition, A is invertible. Then there exist $m, M>0$ such that $m\leqslant |z|\leqslant M$ for all $z\in\sigma(A)$. Set $\tilde{f}=Uf$ and $\tilde{g}=Ug$.

Case 1. *A is a positive self-adjoint operator, and $\{Ug\}_{g\in\mathcal{G}}$ and $\{Uf_n\}$ are real-valued, i.e., $Ug = \Re(\tilde{g})$ for all $g\in\mathcal{G}$ and $Uf_n = \Re(\tilde{f}_n)$:* In this case, one has $|\langle f_n, A^t g\rangle|^2 = (\langle A^t g, f_n\rangle)^2$, for all $t\in\mathbb{R}^+$. Therefore

$$\sum_{g\in\mathcal{G}} |\langle f_n, A^t g\rangle|^2 = \sum_{g\in\mathcal{G}} (\langle A^t g, f_n\rangle)^2, \text{ for all } t\in\mathbb{R}^+.$$

Moreover, since $\mathcal{G}$ is Bessel, by Lemma 2.7.9, the functions $F_n(t)=\sum_{g\in\mathcal{G}}(\langle A^t g, f_n\rangle)^2$ are analytic for $t\in\Omega_{L/4}\cap D_r\subset\mathbb{C}$ and satisfy

$$|F_n(t)| = \left|\sum_{g\in\mathcal{G}}(\langle A^t g, f_n\rangle)^2\right| \leqslant \sum_{g\in\mathcal{G}}\left|\langle g, (A^t)^* f_n\rangle\right|^2 \leqslant C_{\mathcal{G}}\|A\|^{2r}, \text{ for } t\in\Omega_{L/4}\cap D_r.$$

Thus, by Montel's theorem, there exists a subsequence $\{F_{n_k}\}$ of $\{F_n\}$ such that $\{F_{n_k}\}$ converge to an analytic function F on $\Omega_{L/4}\cap D_r$. Let $D_r\subset\mathbb{C}$ be a disk of radius r containing $[L/2,1]$. Since

F_n are analytic and $F_n(t) \to 0$, for all $t \in [L/2, L]$, it follows that $F(t) = 0$, for all $t \in [L/2, L]$. Moreover, since F is analytic, we conclude that $F(t) = 0$ for all $t \in \Omega_{L/4} \cap D_r$, and hence also on $[L/2, 1]$, i.e., $\lim_{n_k \to \infty} F_{n_k}(t) = 0$ for all $t \in [L/2, 1]$. Thus,

$$
\begin{aligned}
& \sum_{g \in \mathcal{G}} \int_0^1 |\langle f_{n_k}, A^t g \rangle|^2 dt \\
= & \sum_{g \in \mathcal{G}} \int_0^{L/2} |\langle f_{n_k}, A^t g \rangle|^2 dt + \sum_{g \in \mathcal{G}} \int_{L/2}^1 |\langle f_{n_k}, A^t g \rangle|^2 dt.
\end{aligned}
$$

Taking limits as n_k tends to infinity, one sees that $\lim\limits_{n_k \to \infty} \sum_{g \in \mathcal{G}} \int_0^1 |\langle f_{n_k}, A^t g \rangle|^2 dt = 0$. This contradicts the assumption that $\{A^t g\}_{g \in \mathcal{G}, t \in [0,1]}$ is a semi-continuous frame. Therefore, $\{A^t g\}_{g \in \mathcal{G}, t \in [0,L]}$ is a semi-continuous frame.

Case 2. *The general case*:

Let $\tilde{f}_n = \Re(\tilde{f}_n) + i\Im(\tilde{f}_n)$ and $\tilde{g} = \Re(\tilde{g}) + i\Im(\tilde{g})$. Define $f_n^{\Re} = U^{-1}\Re(\tilde{f}_n)$, $f_n^{\Im} = U^{-1}\Im(\tilde{f}_n)$, $g^{\Re} = U^{-1}\Re(\tilde{g})$, and $g^{\Im} = U^{-1}\Im(\tilde{g})$. Define A_+^t and A_-^t as

$$
\begin{aligned}
\langle A_+^t g, f \rangle &= \int_{z \in \sigma(A), z > 0} z^t \langle \tilde{g}, \tilde{f} \rangle d\mu(z), \\
\langle A_-^t g, f \rangle &= \int_{z \in \sigma(A), z < 0} (-z)^t \langle \tilde{g}, \tilde{f} \rangle d\mu(z).
\end{aligned}
$$

Then A_- and A_+ are positive operators, and $\langle A^t g, f \rangle = \langle A_+^t g, f \rangle + e^{i\pi t} \langle A_-^t g, f \rangle$.

For $t \in \mathbb{R}^+$, one has

$$
\sum_{g \in \mathcal{G}} |\langle f_n, A^t g \rangle|^2 = F_n(t) + G_n(t), \tag{2.22}
$$

where

$$
\begin{aligned}
F_n(t) = & \sum_{g \in \mathcal{G}} (\langle A_+^t g^{\Re}, f_n^{\Re} \rangle + \langle A_+^t g^{\Im}, f_n^{\Im} \rangle + \cos(\pi t) \cdot (\langle A_-^t g^{\Re}, f_n^{\Re} \rangle + \langle A_-^t g^{\Im}, f_n^{\Im} \rangle) + \\
& \sin(\pi t) \cdot (\langle A_-^t g^{\Re}, f_n^{\Im} \rangle - \langle A_-^t g^{\Im}, f_n^{\Re} \rangle))^2,
\end{aligned}
$$

and

$$
\begin{aligned}
G_n(t) &= \sum_{g\in\mathcal{G}}(\langle A_+^t g^{\Im}, f_n^{\Re}\rangle - \langle A_+^t g^{\Re}, f_n^{\Im}\rangle + \sin(\pi t)\cdot(\langle A_-^t g^{\Re}, f_n^{\Re}\rangle + \langle A_-^t g^{\Im}, f_n^{\Im}\rangle) + \\
&\quad \cos(\pi t)\cdot(\langle A_-^t g^{\Im}, f_n^{\Re}\rangle - \langle A_-^t g^{\Re}, f_n^{\Im}\rangle))^2.
\end{aligned}
$$

Note that for $t \in \Omega_{L/4} \cap D_r$, by Lemma 2.7.10, one has

$$
\begin{aligned}
|F_n(t)| &\leqslant 6\cdot\left(\sum_{g\in\mathcal{G}} |\langle f_n^{\Re}, A_+^t g^{\Re}\rangle|^2 + |\langle f_n^{\Im}, A_+^t g^{\Im}\rangle|^2 + \frac{3+e^{2\pi r}}{4}\cdot(|\langle f_n^{\Re}, A_-^t g^{\Re}\rangle|^2 + \right.\\
&\quad \left. |\langle f_n^{\Im}, A_-^t g^{\Im}\rangle|^2) + \frac{3+e^{2\pi r}}{4}\cdot(|\langle f_n^{\Re}, A_-^t g^{\Im}\rangle|^2 + |\langle f_n^{\Im}, A_-^t g^{\Re}\rangle|^2)\right) \\
&\leqslant 6\cdot\left(C_{\mathcal{G}}\|A\|^{2r} + \frac{3+e^{2\pi r}}{4}\cdot C_{\mathcal{G}}\|A\|^{2r} + \frac{3+e^{2\pi r}}{4}\cdot C_{\mathcal{G}}\|A\|^{2r}\right) \\
&= (15+3e^{2\pi r})\cdot C_{\mathcal{G}}\cdot\|A\|^{2r},
\end{aligned}
$$

and

$$
|G_n(t)| \leqslant (15+3e^{2\pi r})\cdot C_{\mathcal{G}}\cdot\|A\|^{2r}.
$$

Thus, (using a similar proof as in Lemma 2.7.9) F_n and G_n are uniformly bounded analytic functions in $\Omega_{L/4} \cap D_r$.

As in Case 1, one can find two subsequences $\{F_{n_k}\}$ and $\{G_{n_k}\}$ converging to analytic functions F and G, respectively. Moreover, since $G_n(t) \leqslant \sum_{g\in\mathcal{G}} |\langle f_n, A^t g\rangle|^2$, and $F_n(t) \leqslant \sum_{g\in\mathcal{G}} |\langle f_n, A^t g\rangle|^2$ for all $t \in \mathbb{R}^+$, and $\lim_{n\to\infty}\sum_{g\in\mathcal{G}} |\langle f_n, A^t g\rangle|^2 = 0, a.e.\ t \in [0, L]$, one can proceed as in the proof of Case 1 and get the contradiction that

$$
\lim_{n_{k_j}\to\infty}\sum_{g\in\mathcal{G}}\int_0^1 |\langle f_{n_{k_j}}, A^t g\rangle|^2 = 0.
$$

Thus, $\{A^t g\}_{g\in\mathcal{G}, t\in[0,L]}$ is a semi-continuous frame for $\mathcal{H}$. □

CHAPTER 3

Dynamical Sampling with Additive Random Noise

In this chapter, we study dynamical sampling in a finite dimensional space when the samples are corrupted by additive random noise. The main purpose of this work is to analyze the performance of the basic dynamical sampling algorithms (see [7, 12]) and study the impact of additive noise on the reconstructed signal. The general formulation is summarized in Section 3.1.

3.1 Problem Formulation

We consider a signal $f \in \mathbb{C}^d$ and a bounded linear operator $A \in \mathbb{C}^{d\times d}$. f evolves and becomes

$$f_n = A^n f \tag{3.1}$$

at time level $n \in \mathbb{N}$. Let $\Omega \subset \{1, \ldots, d\}$ be a set of spatial locations. The noiseless dynamical samples are then

$$\{f_n(j) : j \in \Omega, 0 \leqslant n \leqslant L \text{ and } n \in \mathbb{N}\}.$$

In [6], necessary and sufficient conditions for recovering $f \in \mathbb{C}^d$ have been derived in terms of A, Ω, and L. In the noisy case, we consider the corrupted dynamical samples of the form

$$\{f_n(j) + \eta_n(j),\ j \in \Omega, 0 \leqslant n \leqslant L \text{ and } n \in \mathbb{N}\}, \tag{3.2}$$

where η_n, $n \geqslant 0$ are independent identically distributed (i.i.d.) d-dimensional random variables with zero mean and covariance matrix $\sigma^2 I$, and $\eta_n(j)$ denotes the j-th component of η_n. Let $\tilde{y}_n = S_\Omega(f_n + \eta_n)$ be the vector of noisy samples at time level n, where the sub-sampling operator S_Ω is a $d \times d$ diagonal matrix such that $(S_\Omega)_{jj} = 1$ if $j \in \Omega$ and $(S_\Omega)_{jj} = 0$ otherwise. When A is given, the signal f can be approximately recovered by solving the least-square minimization problem

$$\tilde{f}_N = \arg\min_g \sum_{n=0}^{N} \|S_\Omega(A^n g) - \tilde{y}_n\|_2^2 \,. \tag{3.3}$$

The main question is

Problem 5. *How does the mean squared error (MSE) $E(\|\tilde{f}_N - f\|_2^2)$ behave, e.g., does it behave asymptotically?*

When A is not given but assumed to have some particular structure, the following question is considered:

Problem 6. *What is the performance of the algorithm in [12, Section 4.1](which is also stated in Algorithm 2) for estimating the spectrum of A? Can a denoising method be designed to effectively treat the corrupted data? How does the dynamical sampling theory perform on real data sets?*

3.2 Contribution and Organization.

In this study, an iterative algorithm for solving problem (3.3) is investigated. In addition, the mean squared error (MSE) $E(\|\epsilon_N\|^2)$ is estimated with $\epsilon_N \doteq \tilde{f}_N - f$ and the behavior of the MSE is analyzed as $N \to \infty$ for an unbiased linear estimator. The second problem of dynamical sampling deals with the case when the evolution operator A is unknown (or only partially known). In [12], an algorithm has been proposed for finding the spectrum of A from the dynamical samples. The present work delves deeper into this algorithm from both theoretical and numerical perspectives. From the theoretical perspective, an alternative proof is given for the fact that the algorithm in [12] can (almost surely) recover the spectrum of A from dynamical samples and also recover the operator A itself, in the case when it is known that A is given by circular convolution, i.e. $Af = a * f$ with some real symmetric filter a in $\mathbb{C}^d$. From a numerical point of view, this analytical result lays the theoretical foundation and paves the way toward recovering the operator A and the initial signal from the real data set. The nature of the spectrum recovery algorithm also motivates an integration of Cadzow-like denoising techniques [19, 34], which can be applied to both synthetic and real data.

In Section 3.3, we summarize the notation that is used throughout the chapter and present the algorithms for signal and filter recovery that work ideally in the noiseless case. To recover the signal, we borrow a least square updating technique from [17] and tailor it for dynamical sampling. To recover the driving operator (in the case of a convolution), we review the algorithm from [12] and provide its new derivation, which is more straightforward than the general proof in [12]. In Section 3.4, the Cadzow denoising method is sketched for a special case of uniform sub-sampling; it is validated to be numerically efficient in the context of dynamical sampling in Section 3.6. Section 3.5 is dedicated to the error analysis of the least square solutions for finding the original signal in the presence of additive white noise. It shows the relation between the MSE of the solution and the number of time levels considered. In Section 3.6, we outline the outcomes of the extensive tests performed for the algorithms discussed in Sections 3.3 and 3.4. More precisely, Section 3.6.1 demonstrates the consistency of the theory for the MSE of the least square solutions on synthetic data. Section 3.6.2 illustrates the effect of Cadzow denoising method on signal and filter recovery in the case of synthetic data. Finally, in Section 3.6.3, the recovery algorithms and denoising techniques are integrated together to process real data collected from cooling processes.

3.3 Notation and Preliminaries

3.3.1 Notation

Let $\mathbb{Z}$ be the set of all integers and $\mathbb{Z}_d$ be the cyclic group of order d. By $\mathbb{C}^d$ and $\mathbb{C}^{m\times d}$ we denote the linear space of all column vectors with d complex components and the space of complex matrices of dimension $m \times d$, respectively. Given a matrix $A \in \mathbb{C}^{m\times d}$, A_{ij} stands for the entry of the i-th row and j-th column of A, A^* represents the conjugate transpose of A, and the 2-norm of A is defined by

$$\|A\| = \sup_{f\in\mathbb{C}^d, \|f\|_2=1} \|Af\|_2,$$

where $\|f\|_2 = \sqrt{\sum_{i=1}^{d} |f(i)|^2}$ and $f(i)$ refers to i-th component of a vector $f \in \mathbb{C}^d$.

For a random variable x that is distributed normally with mean μ and variance σ^2, we may write $x \sim N(\mu, \sigma^2)$.

3.3.2 A general least squares updating technique for signal recovery

We borrow from [35] the following updating technique for adjusting a least squares solution when new equations are added. Consider the following least squares problem

$$f_L^\sharp = \arg\min_{g\in\mathbb{C}^d} \sum_{i=1}^{L} \|A_i g - b_i\|_2^2 \,, \tag{3.4}$$

where $A_i \in \mathbb{C}^{m_i\times d}$, and $\operatorname{rank}(A_1) = d$ (i.e., A_1 has full column rank).

We take the case of $L = 2$ as an example to explain the updating technique. Consider the QR decomposition $A_1 = Q_1 R_1$, where Q_1 is an $m_1 \times d$ matrix satisfying $Q_1^* Q_1 = I$ and R_1 is a $d \times d$ triangular matrix. Then

$$f_1^\sharp = \arg\min_g \|A_1 g - b_1\|_2^2 = \arg\min_g \|R_1 g - Q_1^* b_1\|_2^2 \,.$$

Let $\tilde{b}_1 = Q_1^* b_1$. Since A_1 has full rank, we have $f_1^\sharp = R_1^{-1}\tilde{b}_1$. Suppose that new information is added, then the least squares problem and its solution needs to be updated, i.e., $f_2^\sharp = \arg\min_{g\in\mathbb{C}^d} \sum_{i=1}^{2} \|A_i g - b_i\|_2^2 \,.$

To solve the new least squares problem, we note that

$$\begin{aligned}
\arg\min_{g\in\mathbb{C}^d} \sum_{i=1}^{2} \|A_i g - b_i\|_2^2 &= \arg\min_{g\in\mathbb{C}^d} \left\| \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} g - \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \right\|_2^2 \\
&= \arg\min_{g\in\mathbb{C}^d} \left\| \begin{pmatrix} Q_1 & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} R_1 \\ A_2 \end{pmatrix} g - \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \right\|_2^2 \\
&= \arg\min_{g\in\mathbb{C}^d} \left\| \begin{pmatrix} R_1 \\ A_2 \end{pmatrix} g - \begin{pmatrix} Q_1^* b_1 \\ b_2 \end{pmatrix} \right\|_2^2 \\
&= \arg\min_{g\in\mathbb{C}^d} \left\| \begin{pmatrix} R_1 \\ A_2 \end{pmatrix} g - \begin{pmatrix} \tilde{b}_1 \\ b_2 \end{pmatrix} \right\|_2^2 .
\end{aligned}$$

Therefore, the problem reduces to finding

$$f_2^\sharp = \arg\min_{g\in\mathbb{C}^d} \left\| \begin{pmatrix} R_1 \\ A_2 \end{pmatrix} g - \begin{pmatrix} \tilde{b}_1 \\ b_2 \end{pmatrix} \right\|_2^2 .$$

One further needs to calculate the QR decomposition

$$\begin{pmatrix} R_1 \\ A_2 \end{pmatrix} = Q_2 R_2,$$

where Q_2 is a unitary matrix and R_2 is a $d \times d$ triangular matrix. Denote

$$\tilde{b}_2 = Q_2^* \begin{pmatrix} \tilde{b}_1 \\ b_2 \end{pmatrix}.$$

It follows that $f_2^\sharp = R_2^{-1}\tilde{b}_2$.

The same process can be applied to the case $L \geqslant 3$ which leads to the iterated updating algorithm that is summarized in Algorithm 1.

This algorithm demonstrates that the recovery problem in dynamical sampling can be solved in a streaming setup, where the solution is updated as new measurements are collected over time,

1. without storing all the previous samples (b_j) or explicitly rewriting all the matrices (A_j) for all $j < i$ at the ith step,
2. and taking advantage of quantities that are stored from previous iterations to avoid the naive computation involving all the previous samples and matrices.

Observe that in the dynamical sampling framework we have $A_i = S_\Omega A^{i-1}$. Assume that at step i, the QR decomposition for

$$\mathcal{A}_i = \begin{pmatrix} S_\Omega I \\ S_\Omega A \\ \vdots \\ S_\Omega A^{i-1} \end{pmatrix}$$

is

$$\mathcal{A}_i = QR.$$

At step $i+1$, $\mathcal{A}_{i+1}$ can thus be written in the convenient form

$$\begin{pmatrix} S_\Omega I \\ QRA \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & Q \end{pmatrix} \begin{pmatrix} S_\Omega I \\ RA \end{pmatrix}.$$

Goal: Recover the original signal by processing time series data.

Input A_1, b_1

Set $A_1 = Q_1 R_1$, the economic QR decomposition of A_1 with the assumption that A_1 has full column rank (see Remark 3.5.4).

Set $\tilde{b}_1 = Q_1^* b_1$.

Set $f_1^\sharp = R^{-1}\tilde{b}_1$.

for $i = 2$ **to** L **do**

  Input A_i, b_i

  Compute the QR decomposition for $\begin{pmatrix} R_{i-1} \\ A_i \end{pmatrix} = Q_i R_i$ using the Householder transformation [35].

  Set $\tilde{b}_i = Q_i^* \begin{pmatrix} \tilde{b}_{i-1} \\ b_i \end{pmatrix}$.

  Set $f_i^\sharp = R_i^{-1}\tilde{b}_i$.

end

Output $f_L^\sharp$

Algorithm 1: Pseudo-code of the iterated updating algorithm.

3.3.3 Filter recovery for the special case of convolution operators and uniform subsampling

In this section, we recall from [12] an algorithm for recovering an unknown driving operator A that is defined via a convolution with a real symmetric filter i.e., A is a circulant matrix corresponding to a convolution with a: $Af = a * f$), and where the spatial sampling is uniform at every time-instant n. We also provide a new, direct proof of validity for the filter recovery algorithm for this case. Specifically, we consider samples of $A^\ell f = a^\ell * f$ at $m\mathbb{Z}_d$ where $m \geqslant 2$, and

$a^\ell = a * \cdots * a$ is the ℓ times convolution of the filter a. We also assume that the Fourier transform $\hat{a}$ of the filter a is real symmetric, and strictly decreasing on $[0, \frac{d-1}{2}]$. We will use the notation $S_m f_n$ to describe this uniform subsampling. In particular, for a vector $z \in \ell^2(\mathbb{Z}_d)$, $S_m z$ belongs to $\ell^2(\mathbb{Z}_J)$, and $S_m z(j) = z(mj)$ for $j = 1, \ldots, J$, where throughout we will assume that m is odd, and $d = Jm$ for some odd integer J.

Let

$$y_\ell = S_m(A^\ell f) = S_m(a^\ell * f),\ \ell \geqslant 0, \tag{3.5}$$

be the dynamical samples at time level ℓ. By Poisson's summation formula,

$$\widehat{(S_m z)}(j) = \frac{1}{m}\sum_{n=0}^{m-1} \hat{z}(j + nJ), \quad 0 \leqslant j \leqslant J-1,\ z \in \ell^2(\mathbb{Z}_d), \tag{3.6}$$

An application of the Fourier transform to (3.5) yields

$$\hat{y}_\ell(j) = \frac{1}{m}\sum_{n=0}^{m-1} \hat{a}^\ell(j+nJ)\hat{f}(j+nJ), \quad 0 \leqslant j \leqslant J-1. \tag{3.7}$$

For each fixed $j \in \mathbb{Z}_J$ and for some integer L with $L \geqslant 2m-1$ ($L = 2m-1$ is the minimum number of time levels that we need to recover the filter), we introduce the following notation:

$$\bar{\mathbf{y}}_\ell(j) = \left(\hat{y}_\ell(j), \hat{y}_{\ell+1}(j), \ldots, \hat{y}_{\ell+L}(j)\right)^T,$$

$$\bar{\mathbf{f}}(j) = \left(\hat{f}(j), \hat{f}(j+J), \ldots, \hat{f}(j+(m-1)J)\right)^T,$$

and

$$\mathcal{V}_m(j) = \begin{pmatrix} 1 & 1 & \ldots & 1 \\ \hat{a}(j) & \hat{a}(j+J) & \ldots & \hat{a}(j+(m-1)J) \\ \vdots & \vdots & \vdots & \vdots \\ \hat{a}^{L-1}(j) & \hat{a}^{L-1}(j+J) & \ldots & \hat{a}^{L-1}(j+(m-1)J) \end{pmatrix}, \tag{3.8}$$

where $0 \leqslant j \leqslant J-1$. From (3.7), it follows that

$$\bar{\mathbf{y}}_\ell(j) = \frac{1}{m}\mathcal{V}_m(j) D^\ell(j)\bar{\mathbf{f}}(j),\ \text{for } 0 \leqslant j \leqslant J-1,\ \ell \geqslant 0, \tag{3.9}$$

where $D(j)$ is the diagonal matrix $D(j) = \operatorname{diag}\big(\hat{a}(j), \hat{a}(j+J), \dots, \hat{a}(j+(m-1)J)\big)$. Let $p_j(x) = c_0(j) + c_1(j)x + \cdots + c_1(j)x^{n_j-1} + x^{n_j}$ be the minimal polynomial that annihilates $D(j)$. The degree of p_j is equal to the number of distinct diagonal values of $D(j)$. Since $L \geqslant 2m-1$, it follows from the assumptions on $\hat{a}$ ($\hat{a}$ is real symmetric, and strictly decreasing on $[0, \frac{d-1}{2}]$) that $\deg(p_j) = m$ for $j \neq 0$ and $\deg(p_0) = (m+1)/2$. Moreover, the rectangular Vandermonde matrix $\mathcal{V}_m(j)$ has rank $r_j = m$ if $j \neq 0$, and $r_0 = (m+1)/2$ if $j = 0$. Consequently, using (3.9), we have that for almost all $\hat{f}$,

$$\bar{\mathbf{y}}_{k+r_j}(j) + \sum_{\ell=0}^{r_j-1} c_\ell(j)\bar{\mathbf{y}}_{k+\ell}(j) = 0, \quad 0 \leqslant j \leqslant J-1, \tag{3.10}$$

where $c_\ell(j)$ are the coefficients of the polynomial p_j and $r_j = \deg p_j = \operatorname{rank} \mathcal{V}_m(j)$. The above discussion leads to the following Algorithm 2 for recovering the spectrum $\sigma(A)$.

Goal: Recover the spectrum $\sigma(A)$.

Set $J = d/m$.

for $j = 0$ **to** $J-1$ **do**

Find the minimal integer r_j for which the system (3.10) has a solution $c(j)$ and find the solution;

set $p_j(\lambda) = \lambda^{r_j} + \sum_{\ell=0}^{r_j-1} c_\ell(j)\lambda^\ell$ and find the set $R(j)$ of all roots of p_j.

end

Set $\sigma(A) = \bigcup_{j=0}^{J-1} R(j)$.

Algorithm 2: A spectrum recovery algorithm for convolution operators.

Remark 3.3.1. *The algorithm for spectrum recovery involves finding the roots of a set polynomials of degree m or $\frac{m+1}{2}$, where m is the subsampling factor. This problem becomes more and more difficult as m becomes larger and larger. However, in applications, one could expect m to be of moderate size ($m \leqslant 5$). Moreover, if some of the spectral values are too close to each other, then finding the coefficients of the minimal polynomials becomes unstable.*

Remark 3.3.2. *The recovery of both the filter and the signal from the measurements points to certain relations to the problem of Blind Deconvolution (see for example [41]); typically, Blind Deconvolution does not involve the difficulty arising from the sub-sampling (the operator S_m), but it is restricted to one time measurement, and uses other assumptions on the signal and filter.*

3.4 Cadzow Denoising Method

In this section, we describe a Cadzow-like algorithm (see Algorithm 3) [19, 34] which can be effectively applied to approximate the dynamical samples y_n in (3.5) from the noisy measurements $\tilde{y}_n = y_n + \eta_n$.

Suppose data points y_n in (3.5) are such that m is an odd integer and A is a symmetric circulant matrix generated by a real symmetric filter a, i.e., the Fourier transform $\hat{a}$ of the filter a is real symmetric. In addition, we also assume that $\hat{a}$ is monotonic on $[0, \frac{d-1}{2}]$. Let L be the number of time levels as in (3.8). In particular, it is necessary that $L \geqslant 2m - 1$. Without loss of generailty we assume that L is even. From (3.7), (3.9), (3.10) in Section 3.3.3 (see also [7, 12]), it follows that the Hankel matrix

$$H(j) = \begin{pmatrix} \hat{y}_0(j) & \hat{y}_1(j) & \dots & \hat{y}_{\frac{L}{2}}(j) \\ \hat{y}_1(j) & \hat{y}_2(j) & \dots & \hat{y}_{\frac{L}{2}+1}(j) \\ \vdots & \vdots & \vdots & \vdots \\ \hat{y}_{\frac{L}{2}}(j) & \hat{y}_{\frac{L}{2}+1}(j) & \dots & \hat{y}_L(j) \end{pmatrix}, \tag{3.11}$$

has rank m for $j \neq 0$ and $(m+1)/2$ for $j = 0$. However, the matrices $\widetilde{H}(j)$ formed as in (3.11) using the noisy measurments $\tilde{y}_n$ will fail the rank conditions. Cadzow's Algorithm approximates $H(j)$ via iterative changes of $\widetilde{H}(j)$ that enforce the rank and the Hankel conditions successively.

Goal : Denoising measurements matrix $\widetilde{Y} = (\tilde{y}_0\ \tilde{y}_1\ \dots \tilde{y}_L)$.

Input : $\widetilde{Y}$ and $k_{\max}$ (maximal number of iterations).

Generate the matrix $(\widetilde{Y})^\wedge$ by taking the Fourier transform on $\widetilde{Y}$.

for $j = 0$ **to** $J - 1$ **do**

  if $j = 0$, **then**

    Set $r = \frac{m+1}{2}$

  end

  else Set $r = m$;

  Form Hankel matrix $X = \widetilde{H}(j)$ as in (3.11) from the jth row of $(\widetilde{Y})^\wedge$.

  for $k = 1$ **to** $k_{\max}$ **do**

    Compute the SVD of X: $X = U\Sigma V^*$, $\Sigma = diag(\sigma_1, \dots, \sigma_{\frac{L}{2}+1})$.

    Set $X = Udiag(\sigma_1, \dots, \sigma_r, 0, \dots, 0)V^*$.

    Generate a Hankel matrix H_{new} by averaging X across its anti-diagonals.

    Set $X = H_{new}$.

  end

  Update the jth row of $(\widetilde{Y})^\wedge$ by the vector obtained by averaging the anti-diagonals of X.

end

Update $\widetilde{Y}$ by taking inverse Fourier transform on $(\widetilde{Y})^\wedge$.

Output: Denoised data $\widetilde{Y}$.

Algorithm 3: The pseudo-code for the Cadzow denoising method.

For each $j \in \mathbb{Z}_J$, an application of the singular value decomposition (SVD) technique produces a decomposition $\widetilde{H}(j) = U\Sigma V^*$, where $\Sigma = diag(\sigma_1, \dots, \sigma_{\frac{L}{2}+1})$ and $\sigma_1 \geqslant \dots \geqslant \sigma_{\frac{L}{2}+1}$. Since the rank is known to be r_j, one can set $\sigma_i = 0$ for $i > r_j$ and obtain an amended matrix of singular values Σ_{r_j}. Then, one may proceed to compute the matrix $X_{new} = U\Sigma_{r_j}V^*$ and form a new Hankel matrix H_{new} by averaging X_{new} across its anti-diagonals. This procedure is applied iteratively. After several iterations, a better approximation of the Hankel matrix $H(j)$ is obtained and a vector of denoised data can be retrieved by applying the inverse Fourier transform.

3.5 Error Analysis

3.5.1 Error analysis for general least squares problems

We begin this section with the error analysis of a least squares problem that is more general than the first dynamical sampling problem. We let $A_i \in \mathbb{C}^{m_i \times d}$, $f \in \mathbb{C}^d$, and $\tilde{y}_i = A_i f + \eta_i$, where η_i are i.i.d. random variables with a zero mean and a variance matrix $\sigma^2 I$. The signal f can be approximately recovered via

$$f_L^\sharp = \arg\min_g \sum_{i=1}^{L} \|A_i g - \tilde{y}_i\|_2^2. \tag{3.12}$$

Denote the error $\epsilon_L = f_L^\sharp - f$. By (3.12) and the definition for $\tilde{y}_i$, it follows that

$$\epsilon_L = \arg\min_\epsilon \sum_{i=1}^{L} \|A_i \epsilon - \eta_i\|_2^2. \tag{3.13}$$

Let

$$\mathcal{A}_L = \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_L \end{pmatrix}, \tag{3.14}$$

and assume that for $L \geqslant N$, where N is some fixed number, $\mathcal{A}_L$, defined by (3.14) above, has full rank. By solving problem (3.13), we have

$$\epsilon_L = \left(\sum_{i=1}^{L} A_i^* A_i \right)^{-1} \sum_{i=1}^{L} A_i^* \eta_i, \text{ for all } L \geqslant N. \tag{3.15}$$

The following proposition can be derived from [50, Theorem B on p. 574]. For the convenience of the reader, however, we include the proof in the Appendix.

Proposition. *Assume that $\mathcal{A}_L$ is defined as in* (3.14) *and has full rank for $L \geqslant N$. Let $\lambda_j(L)$, $1 \leqslant j \leqslant d$, denote the eigenvalues of the matrix $\mathcal{A}_L^* \mathcal{A}_L = \sum_{i=1}^{L} A_i^* A_i, 1 \leqslant j \leqslant d$. Then, the following holds:*

$$E(\|\epsilon_L\|_2^2) = \sigma^2 \sum_{j=1}^{d} 1/\lambda_j(L), \tag{3.16}$$

where ϵ_L is obtained from (3.13) *and σ is the variance of the noise.*

To study the behavior of the MSE function in (3.16), we recall the well-known Courant-Fischer Minimax Theorem and one of its most useful corollaries.

Theorem 3.5.1. (***Courant-Fischer Minimax Theorem***) *Let A be a $d \times d$ Hermitian matrix with eigenvalues $\lambda_1 \geqslant \ldots \geqslant \lambda_k \geqslant \ldots \geqslant \lambda_d$. Then,*

$$\lambda_k = \max_U \left\{ \min_x \left\{ \frac{x^* A x}{x^* x} : \ x \in U \text{ and } x \neq 0 \right\} : \ \dim(U) = k \right\}.$$

Corollary 3.5.2. *Let $A \in \mathbb{C}^{d \times d}$ and $B \in \mathbb{C}^{d \times d}$ be self-adjoint positive semidefinite matrices. Then, $\lambda_i(A+B) \geqslant \lambda_i(A)$ and $\lambda_i(A+B) \geqslant \lambda_i(B)$.*

The following result is immediate from Corollary 3.5.2.

Proposition. *The function $E(\|\epsilon_L\|_2^2)$ defined by* (3.16) *is a non-negative non-increasing function of L for $L \geqslant N$ where N is some fixed number, such that $\mathcal{A}_N$ in* (3.14) *has full rank. Consequently, as L goes to ∞, it converges to a non-negative constant.*

The goal of the following example is to illustrate the above result in the context of dynamical sampling but without sub-sampling.

Example 6 (Special case: no sub-sampling)**.** *Suppose that A is a normal matrix and suppose that $A_i = A^{i-1}$ in* (3.13)*. Because A is normal, it can be written as $A = U^* D U$, where U is a unitary matrix and D is a diagonal matrix with the diagonal entries $s_1, s_2, \ldots, s_d$. Hence, $\mathcal{A}_L^* \mathcal{A}_L$ can be computed as*

$$\mathcal{A}_L^* \mathcal{A}_L \quad = \quad \sum_{k=1}^{L} (A^*)^{k-1} A^{k-1} = U^* \sum_{k=1}^{L} (D^* D)^{k-1} U. \tag{3.17}$$

Defining $\Lambda = \Lambda(L)$ by

$$\mathcal{A}_L^* \mathcal{A}_L = U^* \Lambda U,$$

and

$$\Lambda = \begin{pmatrix} \lambda_1(L) & & \\ & \ddots & \\ & & \lambda_d(L) \end{pmatrix},$$

we get from (3.17) *that*

$$\lambda_j(L) = \begin{cases} \frac{1-|s_j|^{2L}}{1-|s_j|^2}, & |s_j| \neq 1; \\ L, & |s_j| = 1. \end{cases}$$

The error ϵ_L can be represented as

$$\epsilon_L = \left(\sum_{i=1}^{L} A_i^* A_i\right)^{-1} \sum_{i=1}^{L} A_i^* \eta_i = U^* \Lambda^{-1} \sum_{i=1}^{L} (D^*)^{i-1} U \eta_i,$$

and (3.16) *follows immediately for this special case.*

To illustrate Proposition 3.5.1, note that, when $|s_j| < 1$, the expression $\frac{1}{\lambda_j(L)} = \frac{1-|s_j|^2}{1-|s_j|^{2L}}$ decreases and converges to $1 - |s_j|^2$ as n increases and tends to ∞.

When $|s_j| = 1$, then $\frac{1}{\lambda_j(L)} = \frac{1}{L}$ which decreases as L increases.

When $|s_j| > 1$, $\frac{1}{\lambda_j(L)} = \frac{1-|s_j|^2}{1-|s_j|^{2L}}$ decreases (as L increases) and converges to 0 as $L \to \infty$.

Thus, in all three cases the function $E(\|\epsilon_L\|_2^2)$ is decreasing as L increases. In addition,

$$E(\|\epsilon_L\|_2^2) \to \sigma^2 \sum_{\substack{1 \leqslant j \leqslant d \\ |s_j| < 1}} (1 - |s_j|^2), \text{ as } L \to \infty.$$

3.5.2 Error analysis for dynamical sampling

To derive a similar result for dynamical sampling, we replace the general operator A_i in (3.13) with the $i-1$ power A^{i-1} of a matrix A followed by a subsampling matrix S_Ω, i.e., we let $A_i = S_\Omega(A^{i-1})$. By Propositions 3.5.1 and 3.5.1, and using the fact that $S_\Omega^* S_\Omega = S_\Omega$, the following assertions hold.

Theorem 3.5.3. *Let $\lambda_j(L)$ denote the j-th eigenvalue of the matrix $\sum_{i=0}^{L-1} (A^*)^i S_\Omega A^i$. Then*

$$E(\|\epsilon_L\|_2^2) = \sigma^2 \sum_{j=1}^{d} 1/\lambda_j(L),$$

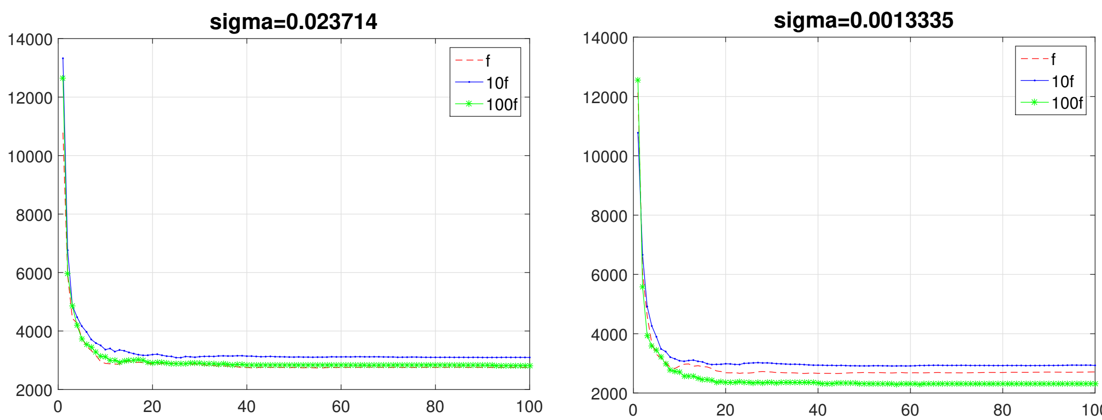
is non-increasing as a function of L. Hence, it converges to some constant as $L \to \infty$.

Remark 3.5.4. *The theorem above shows how the mean squared error depends on Ω, A and L. However, for a given A, not all choices of Ω are allowable: there are necessary and sufficient* *conditions on the choice of Ω that will allow us to reconstruct f by solving (3.12) when $A_i = S_\Omega(A^{i-1})$ and no noise is present [6] (i.e., A_1 is full rank and $\lambda_j(L) > 0$ for all j, L).*

3.6 Numerical Results

3.6.1 Error Analysis

In this section, we illustrate the performance of the least squares based method for signal recovery (i.e., Algorithm 1) in the case when the dynamical samples are corrupted by noise. We describe the numerical simulations that we conducted using synthetic data and examine the behavior of $E(\|\epsilon_L\|_2^2/\sigma^2)$ as a function of the number of time levels L.



(a) Noise with standard deviation $\sigma = 0.023714$. (b) Noise with standard deviate $\sigma = 1.3335 \times 10^{-3}$.

Figure 3.1: The behavior of $E(\|\epsilon_L\|_2^2/\sigma^2)$ for randomly generated signal. The signal f is randomly generated with norm 2.2914. Three signals f, $10f$, and $100f$ are used for the simulations, where x-axis stands for the time levels and y-axis represents the value of $E(\|\epsilon_L\|_2^2/\sigma^2)$.

To obtain synthetic data for the simulation, we use a random signal $f \in \ell^2(\mathbb{Z}_{18})$ and a convolution operator $Af = a * f$, determined by a real symmetric vector a with non-zero components given by $(\frac{1}{8}, \frac{1}{2}, 1, \frac{1}{2}, \frac{1}{8})$, i.e., $A \in \mathbb{R}^{18\times 18}$ is a circulant matrix with the first row $(1, 1/2, 1/8, 0, \ldots, 0, 1/8, 1/2)$. We generate the signals $f_i = A^i f$ at time levels $i = 0, 1, \ldots, L$. The non-uniform locations $\Omega = \{1, 5, 7, 10, 13, 15, 18\}$ are chosen to generate the samples $\{f_i(j) : j \in \Omega\}$. Independent and identically distributed Gaussian noise with zero mean is then added to the samples to obtain a set of noisy data $\{f_i(j) + \eta_i(j) : j \in \Omega\}$.

Figure 3.1 shows the relationship between $E(\|\epsilon_L\|_2^2/\sigma^2)$ and the time levels, where ϵ_L is defined

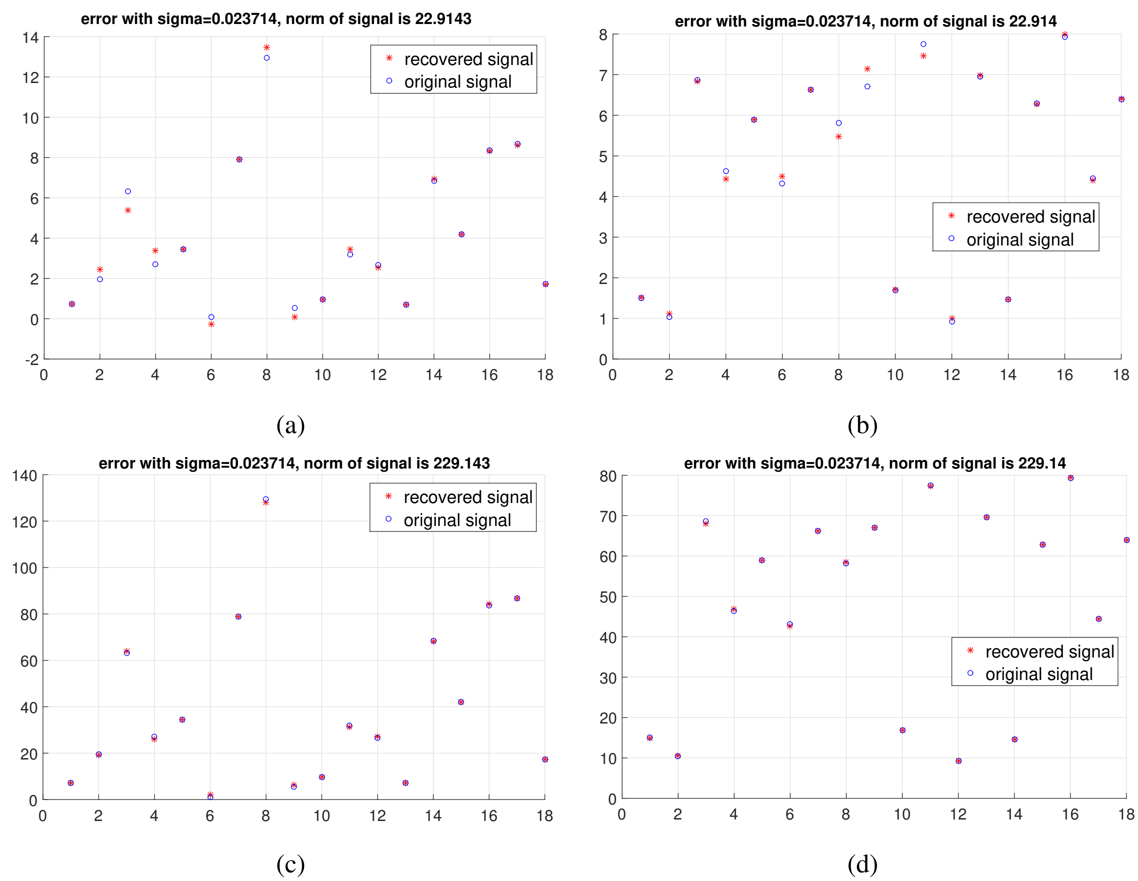
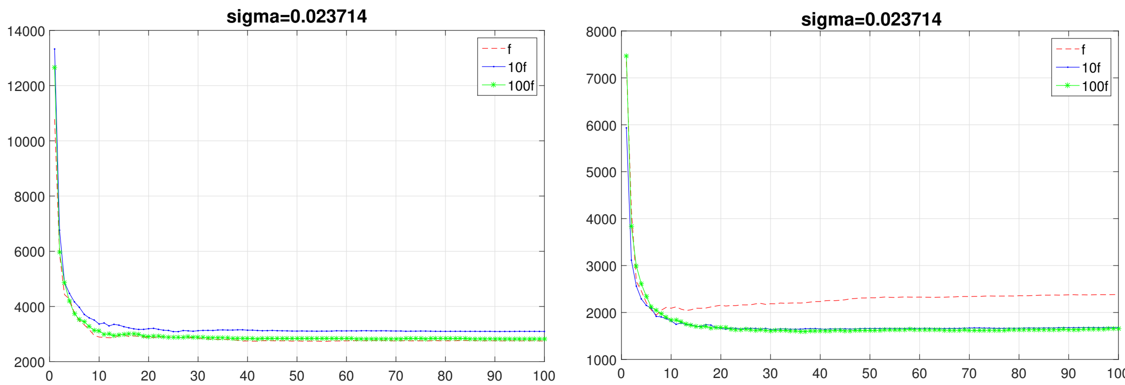


Figure 3.2: The original signals and the reconstructed signals are represented by blue circles and red stars, respectively. The signals in 3.2c and 3.2d are obtained from the signals in 3.2a and 3.2b, respectively, by multiplying by 10. The norm of the original signal in 3.2a equals the norm of the original signal in 3.2b, the same is true for the signals in 3.2c and 3.2d.

by (3.13). For each L, the simulation was repeated 100 times with the same distribution of noise, and $E(\|\epsilon_L\|_2^2/\sigma^2)$ was estimated by averaging the 100 values of $\|\epsilon_L\|_2^2/\sigma^2$. Figure 3.1a shows how $E(\|\epsilon_L\|_2^2/\sigma^2)$ changes as L varies for three different signals: f, $10f$, and $100f$, where the noise variance is $\sigma = 2.3714 \times 10^{-2}$ and the 2-norm of f approximately equals 2.2914. The graph of $10f$ is given in Figure 3.2a. Figure 3.1b shows the behavior of $E(\|\epsilon_L\|_2^2/\sigma^2)$ for the same signals as in Figure 3.2a, where the noise variance is $\sigma = 1.3335 \times 10^{-3}$. As shown in Figure 3.1, $E(\|\epsilon_L\|_2^2/\sigma^2)$ decreases as n increases and approaches the constant predicted by Theorem 3.5.3.

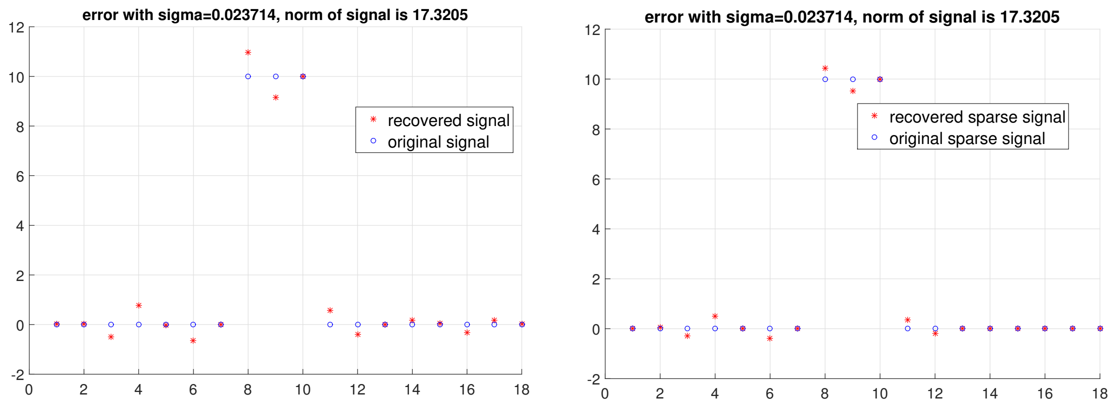


(a) (b)

Figure 3.3: The behavior of $E(\|\epsilon_L\|_2^2/\sigma^2)$ are shown in (3.3a) and (3.3b) for the sparsely supported signals without and with applying the threshold method, respectively, where the samples are corrupted by Gaussian noise with zero mean and standard deviation 2.3714×10^{-2}.

Figure 3.2 depicts the graphs of the reconstructed signals and the original signals $10f$, $100f$, $10g$, and $100g$ in Figures 3.2a, 3.2c, 3.2b, and 3.2d, respectively, where $f, g \in \mathbb{R}^{18}$ are randomly generated and scaled to the norm 2.2914. The reconstructed signals from the noisy data and the original signals are shown in Figure 3.2a for $10f$, in Figure 3.2b for $10g$, in Figure 3.2c for $100f$, and in Figure 3.2d for $100g$, respectively. The noisy data are corrupted by Gaussian noise with zero mean and standard deviation 2.3714×10^{-2}. As displayed in Figure 3.2, while the reconstructed signals in Figures 3.2a and 3.2b are clearly different from the original signals, it is hard to distinguish the reconstructed signals from the original signals in Figures 3.2c and 3.2d because the reconstructed signals are very close to the original signals.

Figures 3.3 and 3.4 are simulation results for the sparsely supported signals. For these special signals, a threshold method [50] is introduced for the samples and reconstructed signals. The method is implemented as follows. Let the threshold T be 2σ, let $\tilde{y}$ denote the sample vector, and let



(a) (b)

Figure 3.4: A comparison of the reconstruction results before and after applying the threshold method. (3.4a) and (3.4b) show the reconstruction results before and after applying the threshold method, respectively. In (3.4a) and (3.4b), the original signals have the same sparse support $\{8, 9, 10\}$. The samples are corrupted by the independent Gaussian noise with mean 0 and standard deviation 2.3714×10^{-2}.

$f_L^\sharp$ be the reconstructed signal. If $|\tilde{y}(i)| \leqslant T$, we set $\tilde{y}(i) = 0$, where $\tilde{y}(i)$ is the i-th component of $\tilde{y}$. Similarly, if $|f_L^\sharp(i)| \leqslant T$, we set $f_L^\sharp(i) = 0$. Then the reconstruction results before and after applying the threshold method are compared. Figures 3.3 and 3.4 illustrate the behavior of $E(\|\epsilon_L\|_2^2/\sigma^2)$ and the reconstructed signals before and after applying the threshold method, respectively. In the simulation, the samples are corrupted by Gaussian noise with zero mean and standard deviation 2.3714×10^{-2}. A sparsely supported signal $f \in \mathbb{R}^{18}$ is generated with support in the locations $\{8, 9, 10\}$ with $f(8) = f(9) = f(10) = 1$. The MSE $E(\|\epsilon_L\|_2^2/\sigma^2)$ are estimated for signals f, $10f$, and $100f$ separately. As shown in Figure 3.3, for n sufficiently large, $E(\|\epsilon_L\|_2^2/\sigma^2)$ is about 20% smaller after the threshold method is applied to the samples and reconstructed signals. Figure 3.4 shows the graphs of the original signal $10f$ and the reconstructed signal, which suggests that the signal reconstructed by applying the threshold method is more accurate than the one reconstructed without applying the threshold method in the locations outside the support of the original signal. These observations suggest that the threshold method can reduce $E(\|\epsilon_L\|_2^2/\sigma^2)$ by improving the accuracy of the zero sets.

3.6.2 Cadzow Denoising

In this section, we describe the impact of the Cadzow denoising technique described in Section 3.4 on dynamical sampling using synthetic data.

3.6.2.1 Denoising of the sampled data

We use a symmetric convolution operator A with eigenvalues $\{1/8, 1/4, 3/8, 1/2, 5/8, 3/4, 7/8, 1\}$. We let A act on the normalized randomly generated signal $f = (0.2931, 0.3258, 0.04568, 0.3286, 0.2275, 0.0351, 0.1002, 0.1967, 0.3444, 0.34710, 0.0567, 0.3492, 0.3443, 0.1746, 0.2879)^T$ iteratively for 100 times. The iterated signals are stored in a matrix Π as

$$\Pi = \left(f \ Af \ A^2 f \ \ldots \ A^{100} f\right) = (f_0 \ f_1 \ f_2 \ldots \ f_{100})$$

where $A^k f$ is a column vector for each $0 \leqslant k \leqslant 100$ (see (3.1)). At each time level, the generated signals are perturbed by i.i.d. Gaussian noise with zero mean and standard deviation $\sigma \in \{10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}\}$; the noisy signals are denoted by

$$\widetilde{\Pi} = \Pi + H,$$

where $H_{i,j} \sim N(0, \sigma^2)$ and every two entries of H are independent (see (3.2)).

The samples are taken uniformly on $3\mathbb{Z}_{15}$ (i.e., $m = 3$), specifically at locations $\Omega = \{1, 4, 7, 10, 13\}$. The Cadzow algorithm (Algorithm 3) is applied to the data $\widetilde{Y} = S_m\widetilde{\Pi}$ where S_m is defined in the first paragraph of Section 3.3.3. The denoised data are denoted by Z which is compared to $S_m\Pi$ directly by computing

$$\frac{\|Z - S_m\Pi\|}{\|S_m\Pi\|}. \tag{3.18}$$

In addition, the relative difference between the noisy data $S_m\widetilde{\Pi}$ and $S_m\Pi$ is computed as

$$\frac{\|S_m\widetilde{\Pi} - S_m\Pi\|}{\|S_m\Pi\|}. \tag{3.19}$$

The same process is repeated for 80 times (with the same Π and different H). The numerical results are obtained by averaging the 80 values of (3.18) and (3.19), respectively.

The simulation results are shown in Figure 3.5. For $rank \geqslant 3$, the horizontal values depict the threshold ranks in the Cadzow algorithm. The corresponding vertical values are $\log_{10}$ of the values of (3.18) averaged over 80 repetitions. When $rank = 0$, (3.19) is used instead of (3.18). As shown in Figure 3.5, the Cadzow denoising technique works best for noise reduction when the rank of the Hankel matrix is chosen to be 3, which is consistent with the theory described in Section 3.4.

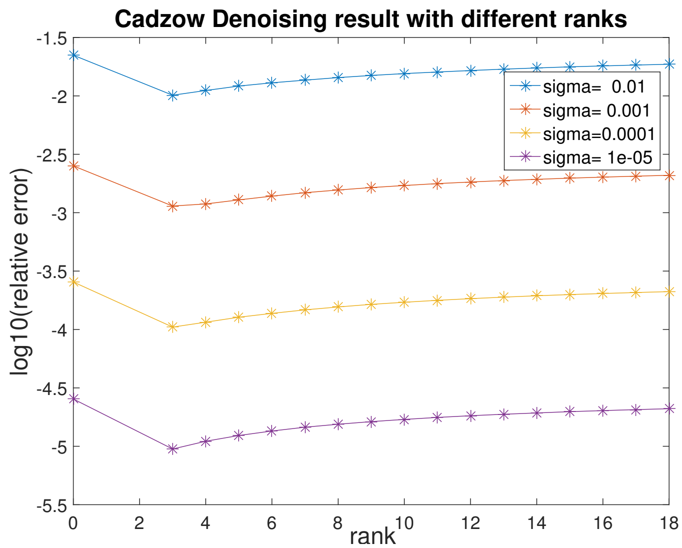


Figure 3.5: The relative errors using Cadzow denoising method. The vertical axis represents $\log_{10}$ of averaged (3.18) when the threshold rank in the Cadzow denoising Algorithm 3 is greater than or equal to 3. When the threshold rank is 0 (3.19) is used instead of (3.18).

In Figure 3.6, the curve labeled "rank $= 0$" shows the relationship between $\log_{10}$ of the averaged (3.19) and $\log_{10}$ of the noise standard deviations, while the curves labeled as "rank $= r$" for $r = 3, 7, 11, 15$ show the relationship between $\log_{10}$ of the averaged (3.18) and $\log_{10}$ of the noise standard deviation. As displayed in Figure 3.6, the curves are almost linear. For fixed noise standard deviation, the figure shows that, as predicted by the theory described in Section 3.4, the best denoising happens when "rank $= 3$" since the sub-sampling is 3.

3.6.2.2 Spectrum Reconstruction of the Convolution Operator

In order to evaluate the impact of the Cadzow denoising technique when reconstructing the spectrum of the convolution operator, we conducted a number of simulations on synthetic data. We repeated the same process as in Section 3.6.2.1 until the denoised data Z was generated. Then

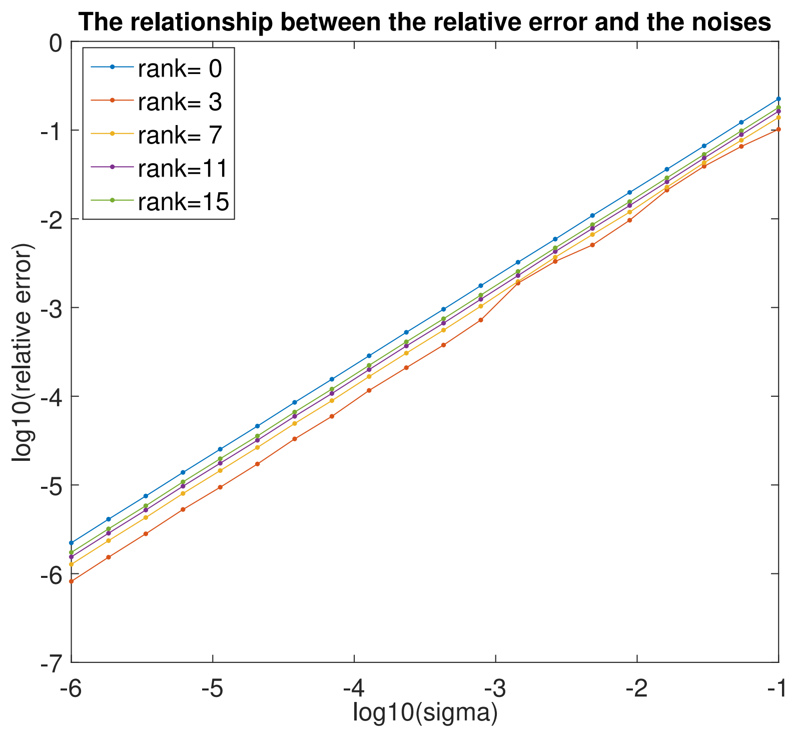


Figure 3.6: The relation between the relative errors and the noise standard deviations using the Cadzow denoising method with different threshold ranks. The curves labeled "rank $= r$", $r = 3, 7 \ldots$ reflect the relationship between $\log_{10}$ of the averaged relative errors and $\log_{10}$ of the noise standard deviations, where the relative errors are represented in (3.18) for $rank \geqslant 3$ and in (3.19) for $rank = 0$.

we used the results of Section 3.3.3 and Algorithm 2 to recover the spectrum of the convolution operator using separately denoised data Z and noisy data $\widetilde{Y} = S_m \widetilde{\Pi}$. The simulation results are shown in Figures 3.7, 3.8, and 3.9 for different noise standard deviations. Figure 3.7 shows the simulation results when the standard deviation of the noise is 10^{-5}. The curves in Figure 3.7 are simulation results for three different random choices of noise. For Figures 3.8 and 3.9, the noise has standard deviations 10^{-4} and 10^{-3}, respectively. As shown in Figures 3.7, 3.8, and 3.9, the Cadzow denoising technique can make a big difference for the spectrum recovery.

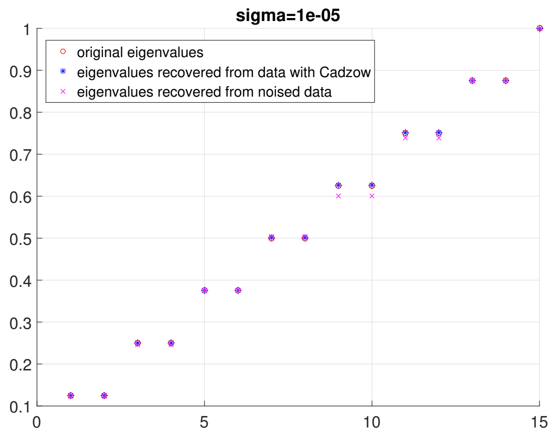
Using the estimated convolution operator and the denoised data, we also evaluated the effectiveness of the reconstruction algorithm, i.e., Algorithm 1, for which the simulation results are shown in Figure 3.10. The figure shows that if the noise is small, the recovered signals are extremely close to the original signals, which also verifies the effectiveness of the Cadzow denoising technique for dynamical sampling.

3.6.3 Real data

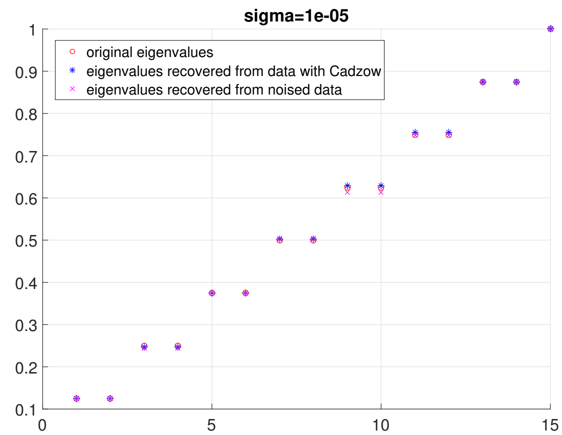
In this section, we describe numerical tests that we performed using two sets of real data. One data set documents a cooling process with a single heat source, and the other – a similar process with two heat sources. These data sets were labeled as “one hotspot” and “two hotspots”, respectively.

The set-up for the real data sets is shown in Figure 3.11. We used the bicycle (aluminum) wheel for the circular pattern. Fifteen (15) sensors are equidistantly placed around the perimeter of the wheel with 4.5 inches apart. The specified accuracy of the sensors is $0.5°C$ and the temperature samples are taken at 1.05Hz.

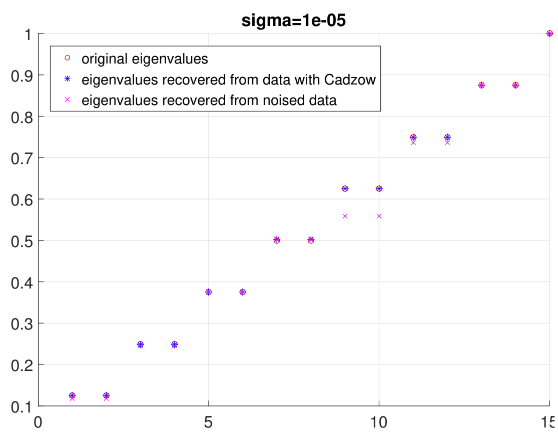
The goal was to estimate the dynamical operator and the original signals by using information from a subset of the thermometer measuring devices, while the totality of the measurements from all devices was used as control to assess the performance of our estimations. In our reconstructions, we did not use any a priori knowledge about the conducting material, its parameters, or the underlying operator driving the evolution of the temperature. Only raw, time-space subsamples of the temperatures was used to estimate the evolution operator, and the initial temperature distribution. The operator was assumed to be real, symmetric convolution operator whose Fourier transform consists of two monotonic pieces, so that recovery of the spectrum of the driving operator sufficed to recover



(a)

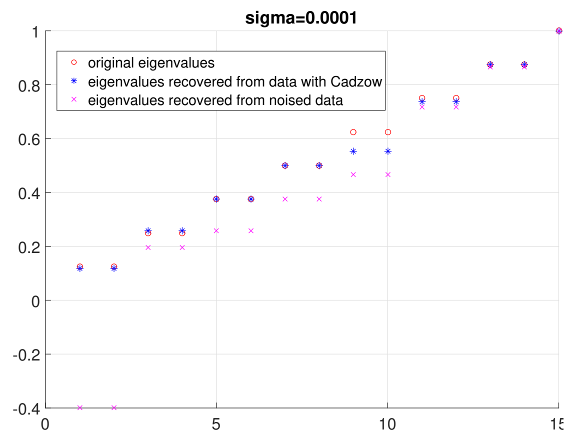


(b)

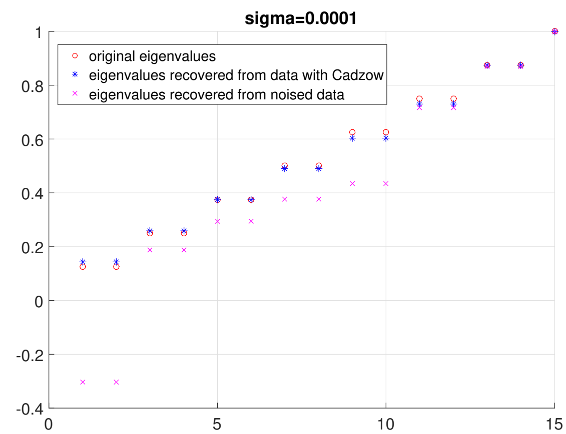


(c)

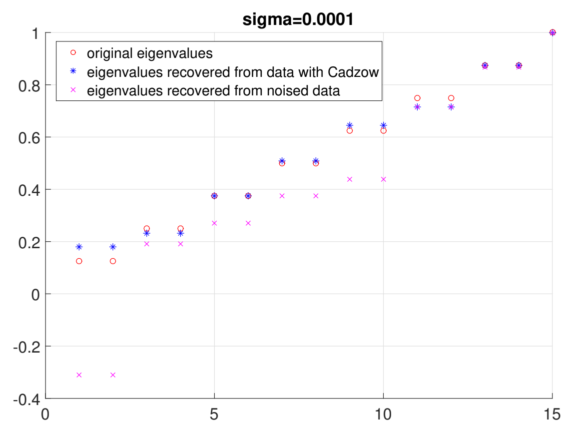
Figure 3.7: A comparison of the spectrum reconstruction with and without the Cadzow denoising technique for $\sigma = 10^{-5}$.



(a)

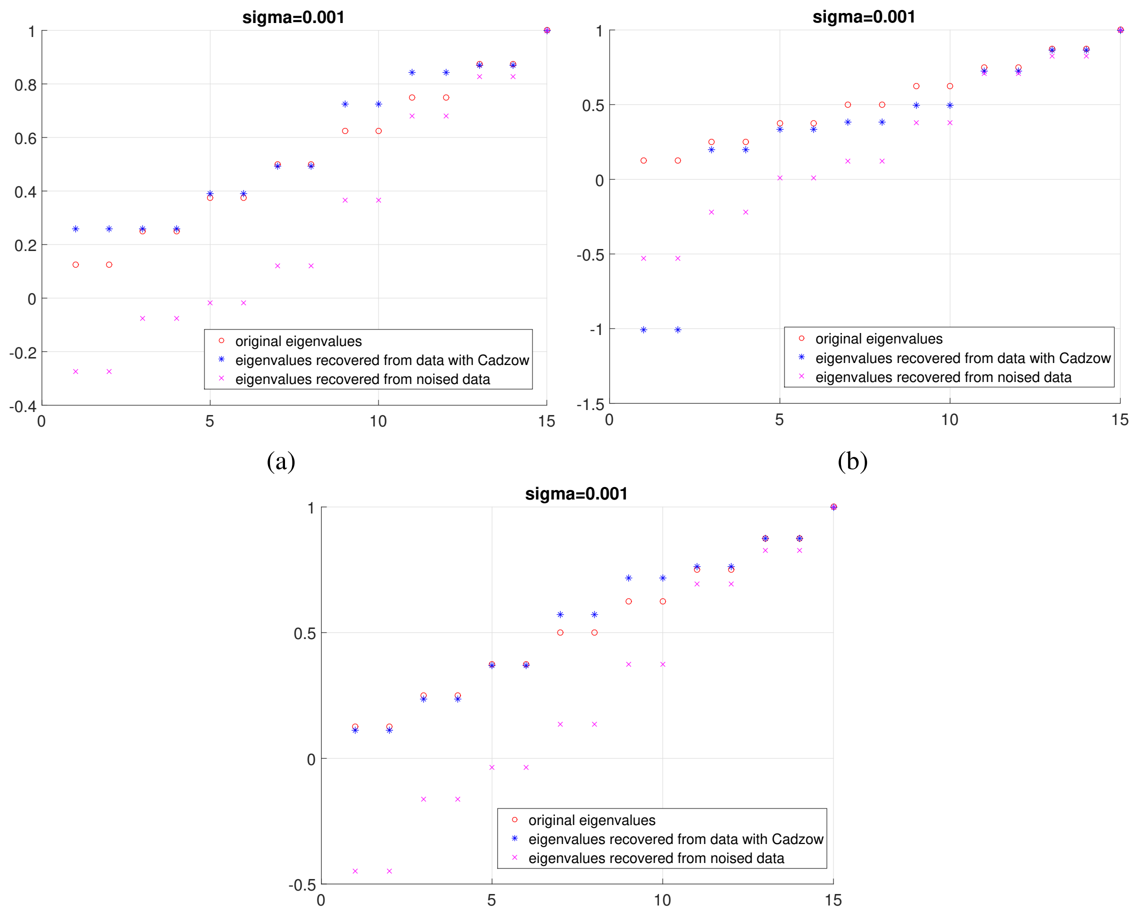


(b)



(c)

Figure 3.8: A comparison of the spectrum reconstruction with and without the Cadzow denoising technique for $\sigma = 10^{-4}$.



(a) (b)

(c)

Figure 3.9: A comparison of the spectrum reconstruction with and without the Cadzow denoising technique for $\sigma = 10^{-3}$.

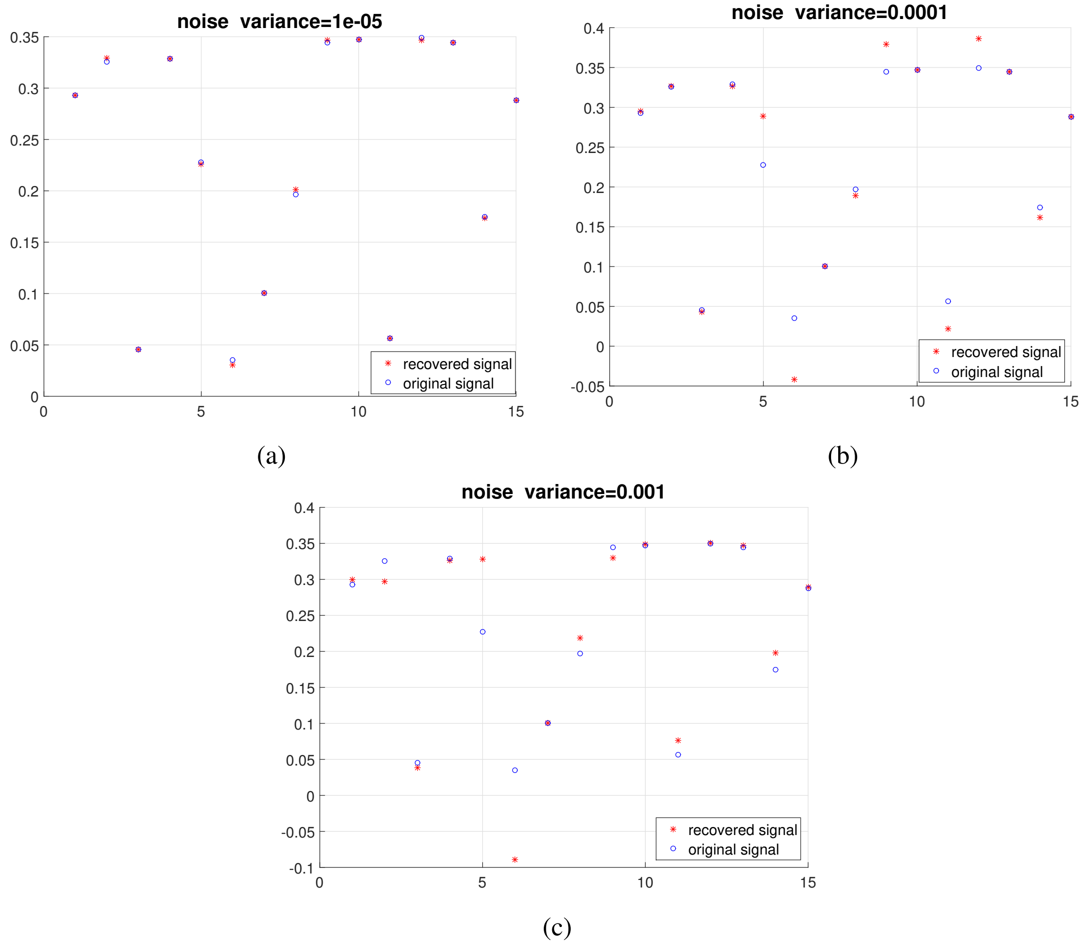


Figure 3.10: A comparison of the recovered signal and the original signal by using the estimated recovered convolution operator.

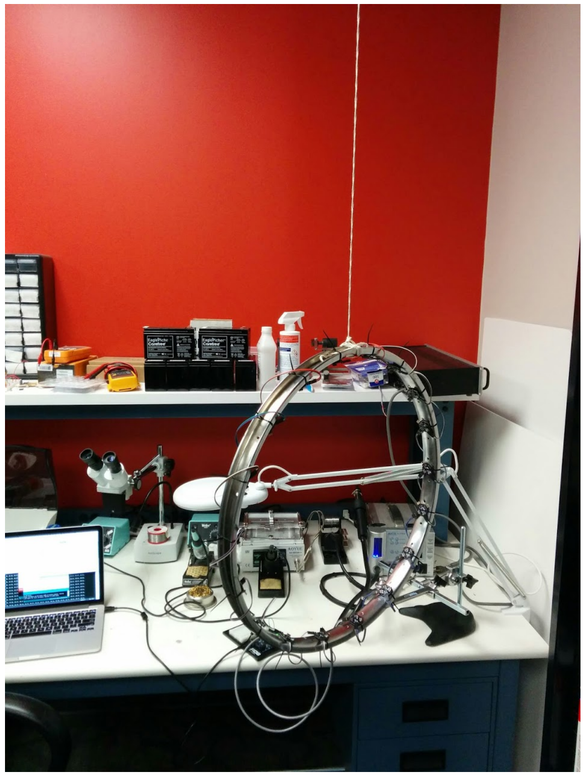
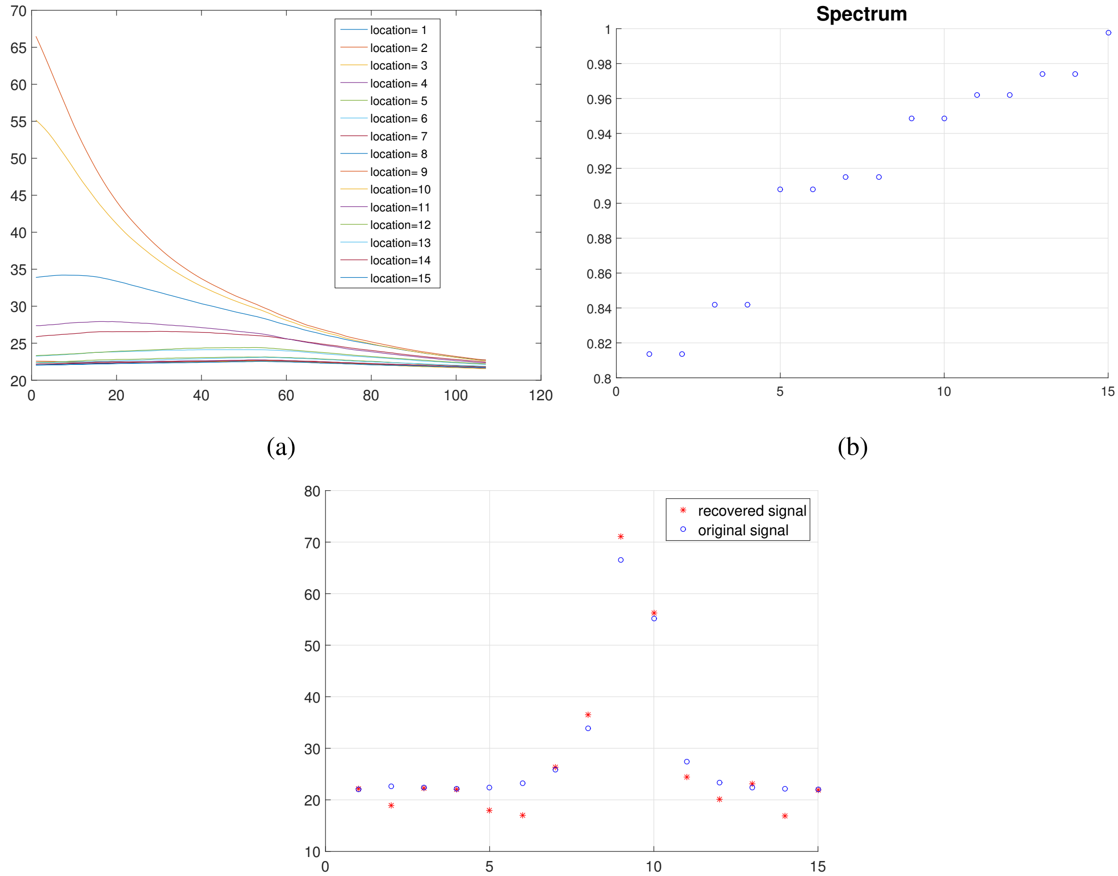


Figure 3.11: Set-up.

the filter.

In the experiment, the signal at time level 20 was set as the original state. First, we smoothed the data by averaging over time to obtain a new data set $\Gamma = (\gamma_1 \ \gamma_2 \ldots)$, where $\gamma_1 = \sum_{i=1}^{10} f_i$, $\gamma_2 = \sum_{i=11}^{20} f_i$, etc. Next, we extracted the information from the new data set at uniform locations Ω with gap $m = 3$ generating the data set $S_m(\Gamma)$. Cadzow Algorithm 3 is then used on $\widetilde{Y} = S_m(\Gamma)$ with the threshold rank close to 2 or 3 to obtain the denoised data Z. Using the data Z, Algorithm 2 was applied to estimate the filter. Finally, using the recovered filter, the original signals were estimated by repeating the computations as in Section 3.6.1.

The test results on the data set with one hotspot are shown in Figure 3.12. Figure 3.12a depicts the evolved signals at all 15 locations. Figure 3.12b shows the recovered spectrum of the evolution filter using the data from locations $\Omega = \{1, 4, 7, 10, 13\}$ to estimate the filter driving the system. Using the driving operator A recovered from Ω and the necessary extra sampling locations at $\{3, 15\}$ needed to recover the signal ($\Omega_e = \Omega \cup \{3, 15\}$) (see [7]), we reconstructed an approximation $f^\sharp$ of the signal that is displayed in Figure 3.12c; it has a relative error $\frac{\|\gamma_1 - f^\sharp\|_2}{\|\gamma_1\|_2}$ of 9.94% compared to the actual measurements at all 15 locations as the reference. This relative error shows that dynamical sampling also works reasonably well for a real data set.



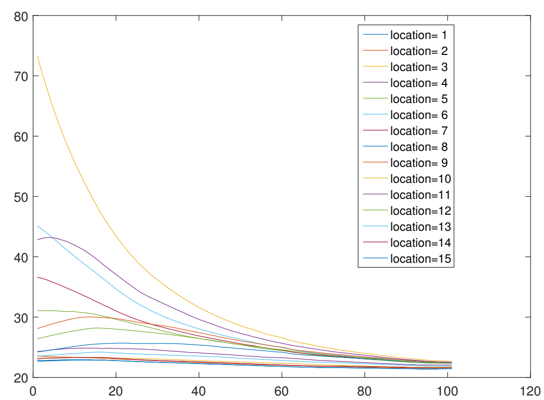
(a) (b)

(c)

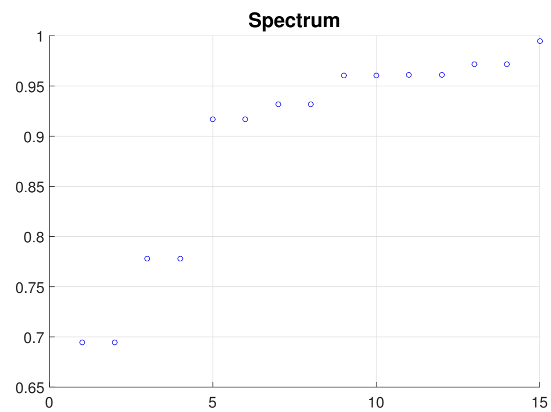
Figure 3.12: Simulation results for the data set with one hotspot. Here, (3.12a) plots the evolved signals, (3.12b) shows the recovered spectrum by using the data from partial locations, and (3.12c) sketches the recovered signal by using the recovered operator from partial locations and the sampled original signal. The partial locations for recovering the operator are $\Omega = \{1, 4, 7, 10, 13\}$. To recover the original signals, we use the data from locations $\Omega_e = \{1, 3, 4, 7, 10, 13, 15\}$.

The test results using the data set with two hotspots are shown in Figure 3.13. Figure 3.13a plots the evolved signals at the 15 locations. Figure 3.13b exhibits the recovered spectrum of the filter with $\Omega = \{2, 5, 8, 11, 14\}$. Using the driving operator A recovered from Ω and the data from locations $\Omega_e = \{2, 3, 5, 8, 10, 11, 14\}$, we recovered an approximation of the signal that is displayed in Figure 3.13c. In this case, the relative error was 12.45% compared to the actual measurements at all 15 locations. Such relative error is generally considered acceptable in this kind of real applications.

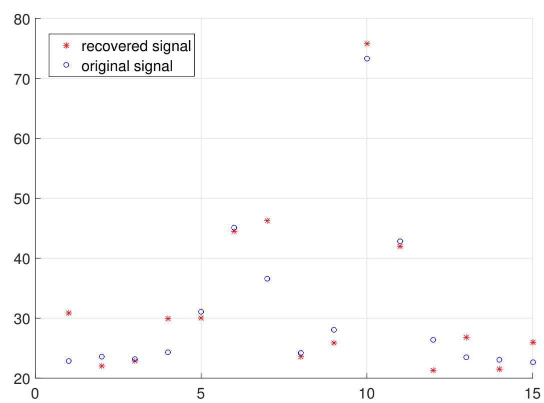
By making similar tests with different choices of Ω and Ω_e, we found that the relative errors



(a)

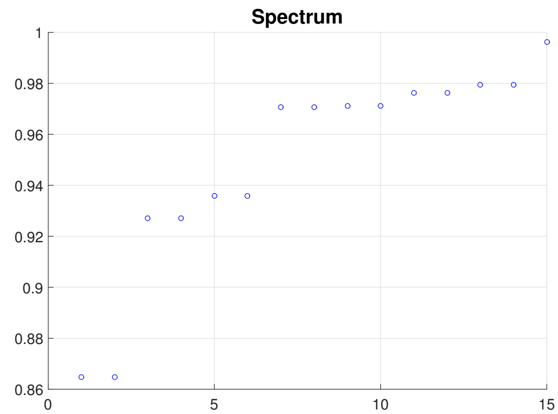


(b)

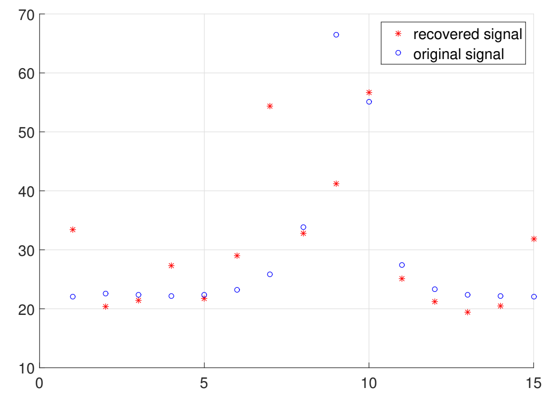


(c)

Figure 3.13: Simulation results for the data set with two hotspots. Here, (3.13a) plots the evolved signals, (3.13b) shows the recovered spectrum by using the data from partial locations, and (3.13c) sketches the recovered signal by using the recovered operator from partial locations and the sampled original signal. The partial locations for recovering the operator are $\Omega = \{2, 5, 8, 11, 14\}$. To recover the original signals, we use the data from locations $\Omega_e = \{2, 3, 5, 8, 10, 11, 14\}$.



(a)



(b)

Figure 3.14: Simulation results for the data set with one hotspot. Here, (3.14a) shows the recovered spectrum by using the data from partial locations, while (3.14b) plots the recovered signal by using the recovered operator from partial locations, where the partial locations for recovering the operator are $\Omega = \{2, 5, 8, 11, 14\}$. To recover the original signals, we use the samples from locations $\Omega_e = \{2, 3, 5, 8, 10, 11, 14\}$.

depend heavily on the choice of locations. The two pictures in Figure 3.14 are the results of the same process that was used to generate the last two pictures in Figure 3.12. In this case, however, we chose $\Omega = \{2, 5, 8, 11, 14\}$ and $\Omega_e = \{2, 3, 5, 8, 10, 11, 14\}$. This choice resulted in the relative error of 34.29% which is considerably larger than the 9.94% in Figure 3.12.

3.7 Concluding remarks

This chapter introduces the problem of noise into the modeling of dynamical sampling and discusses certain unbiased linear estimators for the recovery of signals from dynamical sampling. The addition of noise to the model highlights some of the difficulties in recovering a signal from measurements in dynamical sampling, and sets the stage for more detailed studies of the information theoretic bounds and other types of estimators.

In addition, this chapter studies a special case related to blind deconvolution, where the subsampling is uniform (to which extra samples are added for the recovery of the unknown signal), and the evolution operator is unknown, but is one dimensional, symmetric, real and decreasing in the frequency domain. The existence of multiple measurements over time, along with the assumptions on the properties of the filter, allow for the recovery of the unknown signal and unknown filter; we point to some of the factors that have an adverse effect on the stability of this procedure.

The basic algorithms and discussion of certain special cases are presented here with the intent of providing a starting point for future work on both the theoretical and algorithmic aspects of noisy instances of dynamical sampling and the case where the evolution operator is unknown.

Appendix A

Appendix for Chapter 3

A.1 Appendix for Proposition 3.5.1

Proof of Proposition 3.5.1. It is clear that

$$\begin{aligned}\|\epsilon_L\|_2^2 &= \left\|\left(\sum_{i=1}^{L} A_i^* A_i\right)^{-1}\left(\sum_{j=1}^{L} A_j^*\eta_j\right)\right\|_2^2 \\ &= \sum_{j=1}^{L}\left\|\left(\sum_{i=1}^{L} A_i^* A_i\right)^{-1} A_j^*\eta_j\right\|_2^2 + \sum_{j\neq k}\left\langle\left(\sum_{i=1}^{L} A_i^* A_i\right)^{-1} A_j^*\eta_j, \left(\sum_{i=1}^{L} A_i^* A_i\right)^{-1} A_k^*\eta_k\right\rangle.\end{aligned}$$

Since η_j and η_k for $j \neq k$ are independent and mean zero, the cross terms cancel out in expectation, and one has

$$E\left(\left\langle\left(\sum_{i=1}^{L} A_i^* A_i\right)^{-1} A_j^*\eta_j, \left(\sum_{i=1}^{L} A_i^* A_i\right)^{-1} A_k^*\eta_k\right\rangle\right) = 0.$$

Consequently,

$$E(\|\epsilon_L\|_2^2) = \sum_{j=1}^{L} E\left(\left\|\left(\sum_{i=1}^{L} A_i^* A_i\right)^{-1} A_j^*\eta_j\right\|_2^2\right). \tag{A.1}$$

Note that

$$\begin{aligned}&\left\|\left(\sum_{i=1}^{L} A_i^* A_i\right)^{-1} A_j^*\eta_j\right\|_2^2 = \left\|\left(\sum_{i=1}^{L} A_i^* A_i\right)^{-1}\sum_{l=1}^{m_j} A_j^{*(l)}\eta_j^l\right\|_2^2 \\ &= \sum_{l=1}^{m_j}\left\|\left(\sum_{i=1}^{L} A_i^* A_i\right)^{-1} A_j^{*(l)}\eta_j^l\right\|_2^2 + \sum_{l\neq p}\left\langle\left(\sum_{i=1}^{L} A_i^* A_i\right)^{-1} A_j^{*(l)}\eta_j^l, \left(\sum_{i=1}^{L} A_i^* A_i\right)^{-1} A_j^{*(p)}\eta_j^p\right\rangle,\end{aligned}$$

where $A_j^{*(l)}$ denotes the l-th column of matrix A_j^* and η_j^l is the l-th entry of η_j. Additionally, η_j^l and η_j^p are independent for $l \neq p$. It follows that

$$E\left(\left\langle \left(\sum_{i=1}^{L} A_i^* A_i\right)^{-1} A_j^{*(l)} \eta_j^l, \left(\sum_{i=1}^{L} A_i^* A_i\right)^{-1} A_j^{*(p)} \eta_j^p \right\rangle\right) = 0.$$

Thus,

$$\begin{aligned}
E(\|\epsilon_L\|_2^2) &= \sum_{j=1}^{L}\sum_{l=1}^{m_j} E\left(\left\| \left(\sum_{i=1}^{L} A_i^* A_i\right)^{-1} A_j^{*(l)} \eta_j^l \right\|_2^2\right) \\
&= \sigma^2 \sum_{j=1}^{L}\sum_{l=1}^{m_j} \left\| \left(\sum_{i=1}^{L} A_i^* A_i\right)^{-1} A_j^{*(l)} \right\|_2^2 \\
&= \sigma^2 \cdot \sum_{j=1}^{L} \text{trace}\left(A_j \left(\sum_{i=1}^{L} A_i^* A_i\right)^{-2} A_j^* \right) \\
&= \sigma^2 \cdot \text{trace}\left(\left(\sum_{j=1}^{L} A_j^* A_j\right) \left(\sum_{i=1}^{L} A_i^* A_i\right)^{-2} \right) \\
&= \sigma^2 \cdot \text{trace}\left(\left(\sum_{i=1}^{L} A_i^* A_i\right)^{-1} \right) = \sigma^2 \sum_{i=1}^{d} 1/\lambda_i(L) \qquad \text{(A.2)}
\end{aligned}$$

and the proposition is proved. □

BIBLIOGRAPHY

[1] R. Aceska, A. Aldroubi, J. Davis, and A. Petrosyan. Dynamical sampling in shift invariant spaces. In Azita Mayeli, Alex Iosevich, Palle E. T. Jorgensen, and Gestur Ólafsson, editors, *Commutative and Noncommutative Harmonic Analysis and Applications*, volume 603 of *Contemp. Math.*, pages 139–148. Amer. Math. Soc., Providence, RI, 2013.

[2] R. Aceska and Y.H. Kim. Scalability of frames generated by dynamical operators. *Frontiers in Applied Mathematics and Statistics*, 3, nov 2017.

[3] B. Adcock and A.C. Hansen. A generalized sampling theorem for stable reconstructions in arbitrary bases. *Journal of Fourier Analysis and Applications*, 18(4):685–716, feb 2012.

[4] A. Aldroubi, A. Baskakov, and I. Krishtal. Slanted matrices, Banach frames, and sampling. *J. Funct. Anal.*, 255(7):1667–1691, 2008.

[5] A. Aldroubi, C. Cabrelli, A.F. Cakmak, U. Molter, and A. Petrosyan. Iterative actions of normal operators. *Journal of Functional Analysis*, 272(3):1121–1146, 2017.

[6] A. Aldroubi, C. Cabrelli, U. Molter, and S. Tang. Dynamical sampling. *Applied and Computational Harmonic Analysis*, 42(3):378–401, 2017.

[7] A. Aldroubi, J. Davis, and I. Krishtal. Dynamical sampling: time-space trade-off. *Appl. Comput. Harmon. Anal.*, 34(3):495–503, 2013.

[8] A. Aldroubi and K. Gröchenig. Nonuniform sampling and reconstruction in shift-invariant spaces. *SIAM Rev.*, 43(4):585–620 (electronic), 2001.

[9] A. Aldroubi, L.X. Huang, I. Krishtal, and R. Lederman. Dynamical sampling with random noise. In *Sampling Theory and Applications (SampTA), 2017 International Conference on*, pages 409–412. IEEE, 2017.

[10] A. Aldroubi, L.X. Huang, I. Krishtal, R. Lederman, A. Ledeczi, and P. Volgyesi. Dynamical sampling with additive random noise. *Sampling Theory in Signal and Image Processing*, To appear.

[11] A. Aldroubi, L.X. Huang, and A. Petrosyan. Frames induced by the action of continuous powers of an operator. *arXiv preprint*, 2018.

[12] A. Aldroubi and I. Kryshtal. Krylov subspace methods in dynamical sampling. *Sampling Theory in Signal and Image Processing, Shannon Centennial Volume*, 15:9–20, 2016.

[13] A. Aldroubi and A. Petrosyan. Dynamical sampling and systems from iterative actions of operators. In *Frames and other bases in abstract and function spaces*, Appl. Numer. Harmon. Anal., pages 15–26. Birkhäuser/Springer, Cham, 2017.

[14] S.T. Ali, J.P. Antoine, and J.P. Gazeau. Continuous frames in hilbert space. *Annals of Physics*, 222(1):1–37, feb 1993.

[15] S.T. Ali, J.P. Antoine, and J.P. Gazeau. *Coherent States, Wavelets and Their Generalizations*. Springer New York, 2000.

[16] R.F. Bass and K. Gröchenig. Relevant sampling of bandlimited functions. *Illinois J. Math.*, to appear, 2012.

[17] Åke Björck. A general updating algorithm for constrained linear least squares problems. *SIAM J. Sci. Statist. Comput.*, 5(2):394–402, 1984.

[18] O. Bratteli and P. Jorgensen. *Wavelets through a looking glass*. Applied and Numerical Harmonic Analysis. Birkhäuser Boston Inc., Boston, MA, 2002. The world of the spectrum.

[19] J.A. Cadzow. High performance spectral estimation—a new ARMA method. *IEEE Trans. Acoust. Speech Signal Process.*, 28(5):524–529, 1980.

[20] E.J. Candès, J.K. Romberg, and T.Tao. Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics*, 59(8):1207–1223, 2006.

[21] C.Cabrelli, U.Molter, V.Paternostro, and F.Philipp. Dynamical Sampling on Finite Index Sets. *ArXiv e-prints*, February 2017.

[22] C. Cheng, Y. Jiang, and Q. Sun. Spatially distributed sampling and reconstruction. *CoRR*, abs/1511.08541, 2015.

[23] O. Christensen. *Frames and Bases*. Applied and Numerical Harmonic Analysis. Birkhäuser Boston Inc., Boston, MA, 2008.

[24] O. Christensen and M. Hasannasab. An open problem concerning operator representations of frames. *ArXiv e-prints*, May 2017.

[25] O. Christensen and M. Hasannasab. Operator representations of frames: boundedness, duality, and stability. *ArXiv e-prints*, April 2017.

[26] J.B. Conway. *Functions of One Complex Variable*. Springer US, 1973.

[27] J.B. Conway. *Subnormal operators*, volume 51 of *Research Notes in Mathematics*. Pitman (Advanced Publishing Program), Boston, Mass.-London, 1981.

[28] J.B. Conway. *A course in functional analysis*. Graduate Texts in Mathematics. Springer, 2 edition, 1994.

[29] I. Daubechies. *Ten Lectures on Wavelets*. Society for Industrial and Applied Mathematics, jan 1992.

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

[31] M. Fornasier and H. Rauhut. Continuous frames, function spaces, and the discretization problem. *Journal of Fourier Analysis and Applications*, 11(3):245–287, Apr 2005.

[32] D. Freeman and D. Speegle. The discretization problem for continuous frames. *ArXiv e-prints*, November 2016.

[33] J.P. Gabardo and D. Han. Frames associated with measurable spaces. *Advances in Computational Mathematics*, 18(2/4):127–147, 2003.

[34] J. Gillard. Cadzow's basic algorithm, alternating projections and singular spectrum analysis. *Stat. Interface*, 3(3):335–343, 2010.

[35] G.H. Golub and C.F. Van Loan. *Matrix computations*. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, Baltimore, MD, fourth edition, 2013.

[36] K. Gröchenig, J.L. Romero, J. Unnikrishnan, and M. Vetterli. On minimal trajectories for mobile sampling of bandlimited fields. *Appl. Comput. Harmon. Anal.*, 39(3):487–510, 2015.

[37] P.R. Halmos. Normal dilations and extensions of operators. *Summa Brasil. Math.*, 2:125–134, 1950.

[38] C. Heil. Wavelets and frames. In *Signal processing, Part I*, volume 22 of *IMA Vol. Math. Appl.*, pages 147–160. Springer, New York, 1990.

[39] T.L. Kriete III. An elementary approach to the multiplicity theory of multiplication operators. *Rocky Mountain J. Math.*, 16(1):23–32, 1986.

[40] P. Jorgensen and F. Tian. Von Neumann indices and classes of positive definite functions. *Journal of Mathematical Physics*, 55(9):093502, 2014.

[41] D. Kundur and D. Hatzinakos. Blind image deconvolution. *IEEE signal processing magazine*, 13(3):43–64, 1996.

[42] Y.M. Lu, P.L. Dragotti, and M. Vetterli. Localizing point sources in diffusion fields from spatiotemporal measurements. In *Proc. Int. Conf. Sampling Theory and applications (SampTA)*, Singapore, 2011.

[43] Y.M. Lu and M. Vetterli. Spatial super-resolution of a diffusion field by temporal oversampling in sensor networks. In *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing*, number LCAV-CONF-2009-009, pages 2249–2252, 2009.

[44] A.W. Marcus, D.A. Spielman, and N. Srivastava. Interlacing families II: Mixed characteristic polynomials and the Kadison-Singer problem. *Ann. of Math. (2)*, 182(1):327–350, 2015.

[45] J. Murray-Bruce and P.L. Dragotti. Estimating localized sources of diffusion fields using spatiotemporal sensor measurements. *IEEE Trans. Signal Process.*, 63(12):3018–3031, 2015.

[46] J. Murray-Bruce and P.L. Dragotti. A sampling framework for solving physics-driven inverse source problems. *IEEE Transactions on Signal Processing*, 65(24):6365–6380, Dec 2017.

[47] M. Zuhair Nashed. Inverse problems, moment problems, signal processing: un menage a trois. In *Mathematics in science and technology*, pages 2–19. World Sci. Publ., Hackensack, NJ, 2011.

[48] F. Philipp. Bessel orbits of normal operators. *Journal of Mathematical Analysis and Applications*, 448(2):767 – 785, 2017.

[49] J. Ranieri, A. Chebira, Y.M. Lu, and M. Vetterli. Sampling and reconstructing diffusion fields with localized sources. In *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*, pages 4016–4019, May 2011.

[50] J.A. Rice. *Mathematical Statistics and Data Analysis*. Duxbury Advanced Series. 3 edition, 2007. ISBN 0-534-39942-8.

[51] W. Rudin. *Real and Complex Analysis*. International Series in Pure and Applied Mathematics. McGraw-Hill Science/Engineering/Math, 3 edition, 1986.

[52] W. Rudin. *Functional Analysis*. International Series in Pure and Applied Mathematics. McGraw-Hill Science/Engineering/Math, 2 edition, 1991.

[53] T. Strohmer. Finite- and infinite-dimensional models for oversampled filter banks. In *Modern sampling theory*, Appl. Numer. Harmon. Anal., pages 293–315. Birkhäuser Boston, Boston, MA, 2001.

[54] Q. Sun. Nonuniform average sampling and reconstruction of signals with finite rate of innovation. *SIAM J. Math. Anal.*, 38(5):1389–1422 (electronic), 2006/07.

[55] J. Wermer. On invariant subspaces of normal operators. *Proc. Amer. Math. Soc.*, 3:270–277, 1952.

[56] Q. Zhang, R. Li, and B. Liu. Periodic nonuniform dynamical sampling in $\ell^2(\mathbb{Z})$ and shift-invariant spaces. *Numerical Functional Analysis and Optimization*, 38(3):395–407, 2017.

[57] Q. Zhang, B. Liu, and R. Li. Dynamical sampling in multiply generated shift-invariant spaces. *Applicable Analysis*, 96(5):760–770, 2017.