SPATIO-TEMPORAL TRADE-OFF FOR QUASI-UNIFORM SAMPLING OF SIGNALS IN EVOLUTIONARY SYSTEMS

By

Jacqueline Davis

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

August 2014
Nashville, Tennessee

Approved:
Professor Akram Aldroubi
Professor Doug Hardin
Professor Larry Schumaker
Professor Jerry Spinrad
For Penelope, who is a volcano of creative energy,

And for Bob, where intuition and spontaneity converge.
ACKNOWLEDGMENTS

My sincere appreciation goes to my advisor, Akram Aldroubi, for his endless encouragement and support during the last three years. Not only did he give invaluable mathematical advice but also helped me develop an identity as a mathematician. I am also indebted to each of my committee members, Doug Hardin, Ed Saff, and Alex Powell, for their advice and support at crucial times.

This work would have never been accomplished without the help of family and friends. I am infinitely thankful to my parents for being exemplary examples of a successful work-life balance and for their hands-on help during the last two years. I never could have done this without them.

Eternal gratitude goes to my husband for his support and willingness to accompany me on this mathematical adventure. His dedication to my achievements have made this work possible, and his encouragement has kept me moving forward throughout its creation.
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEDICATION</td>
<td>ii</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>iii</td>
</tr>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>I.1 Relation to Existing Fields</td>
<td>2</td>
</tr>
<tr>
<td>I.2 Application to Wireless Sensor Networks</td>
<td>3</td>
</tr>
<tr>
<td>I.3 Organization</td>
<td>3</td>
</tr>
<tr>
<td>II. DYNAMICAL SAMPLING IN FINITE DIMENSIONS</td>
<td>5</td>
</tr>
<tr>
<td>II.1 Problem Formulation</td>
<td>5</td>
</tr>
<tr>
<td>II.2 Dynamical Sampling in Invariant Evolution Systems</td>
<td>7</td>
</tr>
<tr>
<td>II.3 Stability in the Presence of Noise</td>
<td>11</td>
</tr>
<tr>
<td>III. DYNAMICAL SAMPLING IN INFINITE DIMENSIONS</td>
<td>16</td>
</tr>
<tr>
<td>III.1 Problem Formulation</td>
<td>16</td>
</tr>
<tr>
<td>III.2 Additional Sampling Sets for Low-Pass Filters</td>
<td>19</td>
</tr>
<tr>
<td>III.3 Stability in the Presence of Noise</td>
<td>29</td>
</tr>
<tr>
<td>III.4 Additional Sampling Sets for Generic Filters</td>
<td>36</td>
</tr>
<tr>
<td>IV. DYNAMICAL SAMPLING IN SHIFT-INvariant SPACES</td>
<td>46</td>
</tr>
<tr>
<td>IV.1 Background on Shift-Invariant Spaces</td>
<td>46</td>
</tr>
<tr>
<td>IV.2 Problem Formulation</td>
<td>47</td>
</tr>
<tr>
<td>IV.3 Reduction to the $\ell^2(\mathbb{Z})$ Case</td>
<td>51</td>
</tr>
<tr>
<td>IV.4 $\ell^2(\mathbb{Z})$-like Dynamical Systems</td>
<td>53</td>
</tr>
<tr>
<td>V. DYNAMICAL SAMPLING WITH A FORCING TERM</td>
<td>55</td>
</tr>
<tr>
<td>V.1 Problem Formulation</td>
<td>55</td>
</tr>
<tr>
<td>V.2 Assumptions on Forcing Term</td>
<td>58</td>
</tr>
<tr>
<td>VI. PROOFS OF LEMMAS AND FORMULAS</td>
<td>60</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>63</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION.

Dynamical sampling is a new type of sampling problem that results from sampling an evolving signal at various times and asks the question: when do coarse samplings taken at varying times contain the same information as a finer sampling taken at the earliest time? In other words, under what conditions on an evolving system, can time samples be traded for spatial samples?

Traditional sampling theory asks the question: when can a signal be reconstructed from partial knowledge of the signal? For example, the most famous result in sampling theory is the Shannon-Whitaker Sampling Theorem [34], which states that a bandlimited signal can be reconstructed exactly from evenly spaced samples of the signal. Specifically, if \( f(x) \) is \( T \)-bandlimited, i.e. the Fourier transform of \( f \) has support contained in \([-T, T]\), then

\[
f(x) = \sum_{n=-\infty}^{\infty} f \left( \frac{n}{2T} \right) \sin \frac{\pi(2Tx - n)}{\pi(2Tx - n)}.
\]

Because dynamical sampling uses samples from varying time levels for a single reconstruction, it departs from classical sampling theory in which a single signal is sampled and then reconstructed.

The dynamical sampling problem: Let \( x \in \ell^2(\Omega) \) be an initial state of a signal in a dynamical system with evolution rule given by the operator \( A : \ell^2(\Omega) \rightarrow \ell^2(\Omega) \) so that the signal at time \( t = n \) is given by \( A^n x \). The sampling scheme is defined by the sets \( \{ \Omega_n \}_{n=0}^N \), with \( \Omega_n \subset \Omega \). The signal is measured at location \( v_i \in \Omega \) at time \( t = n \) if and only if \( v_i \in \Omega_n \). If \( S(\Omega_n) \) represents the subsampling operator at time \( t = n \), then \( y_n = S(\Omega_n)A^n x \) is the measured signal at time \( t = n \). Under what conditions on \( A \) and \( \{ \Omega_n \}_{n=0}^N \) can the initial signal \( x \) can be recovered from the measurements \( y_n \) for \( n = 0, \ldots, N \)?

Some intuition about dynamical sampling can come from considering a diffusive process. The value of a signal at each location depends on the values of the signal at surrounding locations at an earlier time. Thus, it is logical to think that information from different times could allow us to conclude something about the values of the signal at locations where measurements are not taken. For instance, an increasing value at one location is an indication that there is a higher concentration of the signal nearby that is spreading to this location.

The general problem described above is very challenging, even in the finite dimensional setting. (See examples II.1.1 and II.1.2 below.) Thus, we focus the majority of
our effort in this dissertation on a few special cases of the general dynamical sampling problem. We assume that the family of operators $A_t$ acting on the initial state $f$ is spatially invariant, i.e., it is independent of (the absolute) position. This means that for each fixed $t$ we have $A_tf = a_t * f$, that is $A_t$ is a convolution operator. We also assume time invariance in the form $A_{t_1 + t_2} = A_{t_1}A_{t_2}$. Additionally, assumptions on the sampling sets $\Omega_n$ are made, which allow us to use Fourier techniques and simplify some of the calculations.

I.1 Relation to Existing Fields

Dynamical sampling is certainly not the first setting in which a signal $x$ is to be recovered from samples of related signals rather than from samples of $x$ itself. We discuss a few of these below and their relation to dynamical sampling.

In inverse problems, a single operator $B$ that often represents a physical process is to be inverted. The goal is to recover a signal $x$ from the observed signal $Bx$. If $B$ is not bounded below, the problem is considered an ill-posed inverse problem. In dynamical sampling, it is possible that the operator $A$ does not have a bounded inverse. Dynamical sampling is different because $A^n x$ is not known for any $n$; rather $x$ is to be recovered from partial knowledge of $A^n x$ for many values of $n$. In fact, the dynamical sampling problem can be phrased as an inverse problem when the operator $B$ is the operation of applying the operators $A, A^2, \ldots, A^N$ and then subsampling each of these signals accordingly.

In multisensor deconvolution, or multichannel sampling, a signal is to be recovered from various filtered versions of the signal [7, 14, 20]. The typical signals considered are images and the filters have a very specific structure - they are given by convolution with characteristic functions with relatively prime support. Philosophically, this strategy is close to dynamical sampling when $A$ is a low-pass filter. However, even if $A$ is defined by convolution with a characteristic function, certainly $A^2$ will not share this property, thus, violating one of the main assumptions in multi-sensor deconvolution. Furthermore, in dynamical sampling we only observe part of the filtered signals from each channel.

In wavelet theory, a high pass filter $H$ and a low pass filter $L$ are applied to the signal $x$. The goal is to design filters $H$ and $L$ so that reconstruction of $x$ from samples of $Hx$ and $Lx$ is feasible. In dynamical sampling there is only one filter $A$, and it is applied iteratively to the signal $x$. Furthermore, the filter $A$ may be high pass, low pass, or neither and is given in the problem formulation, not designed.

In filter bank theory, multiple filters are applied to a signal and the output of each filter is down-sampled [9,15,30,36,37,42]. The goal is to reconstruct the original signal from the subsampled filtered signals. While propositions II.2.1 and III.1.1 below can
be considered a special case of results in filter bank theory, filter bank theory focuses on designing sets of filters that satisfy these conditions. In contrast, the focus of our work is on expanding the sampling sets to ensure stable reconstruction when the conditions of these propositions are violated.

1.2 Application to Wireless Sensor Networks

A natural setting for dynamical sampling is wireless sensor networks (WSN). In WSN large numbers of physical sensors are distributed to gather information about the field to be monitored, such as temperature, pressure, or pollution. WSN are used in many industries, including the health, military, and environmental industries [2].

It was in this setting that Lu, Vetterli, and their collaborators proposed the idea of using temporal samples in combination with spatial samples for superresolution of heat distributions [24,27,31]. Their goal is to minimize the cost of reconstructing temperature fields in wireless sensor networks by reducing the number of spatial samples required to recover localized heat sources.

Other algorithms for reconstructing signals in WSN focus on minimizing energy consumption by limiting the number of wireless communications among sensors. One method given in [32,33] clusters nodes into regions and then reconstructs the physical field independently on each region. In these papers, the evolutionary nature of the system is mostly ignored.

The work in this dissertation is inspired by that of Lu, Vetterli, and their collaborators. Our goal is to exploit the evolutionary structure of the system to reduce the number of sensors required to reconstruct a field in a WSN. The idea is simple. The cost of a sensor is expensive relative to the cost of activating the sensor. If we are able to recover the same information with fewer sensors, each being activated more frequently, we will have made reconstruction cheaper, and in many cases plausible when it currently is not.

However, our work is not yet general enough to be applied directly to WSN. We consider only very structured sampling sets and cannot reconstruct signals from irregular or even random spatial samples. Additionally, while we provide stable sampling sets, the reconstruction operator has a norm too large to be considered robust for most applications. Further exploration of regularization techniques to increase robustness is needed.

1.3 Organization

In section II, we formulate the dynamical sampling problem in finite dimensions. We explore the general problem through two examples and then consider regular subsam-
pling in spatially invariant systems. We show that in many cases regular subsampling is not sufficient and show that with a few additional samples, the initial state of the signal can be recovered. An analysis of the stability of the reconstruction operator is provided. The work in this section is joint work with Akram Aldroubi and Ilya Krishtal and appears in [3].

In section III, we formulate the dynamical sampling problem in infinite dimensions. We consider regular subsampling in spatially invariant systems. A complete analysis of stable additional sampling sets is given. Bounds on the norm of the reconstruction operator are also provided. In section IV, the results of section III are extended to shift-invariant spaces. Portions of the work in these sections is joint work with Akram Aldroubi and Ilya Krishtal and Roza Aceska, Akram Aldroubi, and Armena Petrosyan and appears in [3] and [1].

In section V, dynamical sampling with a forcing term is introduced. In this problem, we remove an assumption made in dynamical sampling by allowing for an unknown source term to enter the system during the sampling period. Under strong assumptions, results similar to those in section III are given. This is independent work and can be found in [17].
CHAPTER II
DYNAMICAL SAMPLING IN FINITE DIMENSIONS

In this chapter, we study dynamical sampling in the case of finite dimensional signals. Such systems may arise as discretizations of inverse problems in partial differential equations. We introduce the general problem in finite dimensions and present a few examples illustrating how mathematically hard this problem can be, even though we formulate it in purely linear algebraic terms.

In section II.2, we concentrate on the particular case in which the underlying dynamical system is invariant in a certain sense. This allows us to use techniques from Fourier analysis. For this case, we show that we can trade-off spatial samples by time samples at essentially a one-to-one rate (only a small number of extra spatial samples is needed), and this trade-off is lossless. We also present in section II.3 an algorithm for recovering the initial state and study its stability and robustness to noise.

II.1 Problem Formulation

Let \( x \in \ell^2(\mathbb{Z}_d) \simeq \mathbb{C}^d \), where \( \mathbb{Z}_d \) is the cyclic group of order \( d \), and \( A \) be a \( d \times d \) invertible matrix with complex entries. In general, we seek to recover vector \( x \) from subsampled versions of the vectors \( A^T x, A^{T+1} x, A^{T+2} x, \) etc., for some non-negative integer \( T \) (due to invertibility of \( A \) we will assume that \( T = 0 \) without loss of generality). More precisely, we let \( S(\Omega_n) \) be diagonal idempotent matrices so that \( s_{ii} = 1 \) if and only if \( i \in \Omega_n \subseteq \{1, \ldots, d\} \), and

\[
y_n = S(\Omega_n)A^{n-1}x, \quad n = 1, \ldots, N.
\]  

(II.1)

We would like to know under which conditions we can recover \( x \) from \( y_n, n = 1, \ldots, N \), or, in other words, what information about \( x, Ax, \ldots, A^{N-1}x \) we need in order to make the recovery possible. By \( x \in \ell^2(\mathbb{Z}_d) \) we model an unknown spatial signal at time \( t = 0 \), and the matrix \( A \) represents an evolution operator so that \( A^nx \) is the signal at time \( t = n \). Then the vectors \( y_n, n = 1, \ldots, N \), give the samples of the evolving system at time \( t = n - 1 \) at a (possibly) reduced number of locations (given by the sum of the ranks of the matrices \( S(\Omega_n) \)). Clearly, we would like to keep the ranks of \( S(\Omega_n) \), \( n = 1, \ldots, N \), at a minimum to reduce the information we need to sample and store. Our motivation is to reduce the number of measuring devices and, thus, make the sampling process cheaper.
When $A$ is a permutation matrix, sampling the diagonal is not sufficient. When $A$ is the identity matrix, sampling the diagonal is necessary.

Figure II.1: Sampling the diagonal in $\mathbb{Z}_3$.

We write the problem in the following matrix form

$$y = Ax,$$  \hspace{1cm} (II.2)

where $y = (y_1, \ldots, y_N)$ and $A = \begin{pmatrix} S(\Omega_1) \\ S(\Omega_2)A \\ \vdots \\ S(\Omega_N)A^{N-1} \end{pmatrix}$ is an $Nd \times d$ matrix which we call the dynamical sampling matrix. The choice of the sets $\Omega_k$, $k = 1, \ldots, N$, will be referred to as the dynamical sampling procedure. We would like to establish conditions under which this procedure is admissible, i.e., which would ensure that the matrix $A$ has full rank $d$. In this case $A$ has a left inverse and the recovery of $x$ is possible.

Let us consider a few simple examples.

**Example II.1.1 (Sampling the diagonal).** In this example we assume that $N = d$, each $\Omega_k$, $k = 1, \ldots, d$, is a singleton and

$$\bigcup_{k=1}^{d} \Omega_k = \{1, \ldots, d\}.$$  \hspace{1cm} (II.3)

In other words, can we recover the signal $x$ if we sample at each node exactly once? Clearly, if the matrix $A$ is diagonal, condition (II.3) is necessary and sufficient due to invertibility of $A$. A similar result is expected if $A$ has strong diagonal dominance, i.e., when the time evolution of the signal at a point depends primarily on the current value at that point (there is little "mixing" in the system). Alternatively, if there is no diagonal dominance, sampling the diagonal may fail. The simplest example is provided by the $3 \times 3$ permutation matrix. See figure II.1.1 for an illustration of sampling diagonal.

Identifying matrices for which condition (II.3) is necessary and sufficient is an
interesting question.

Example II.1.2 (Sampling at one node). As another extreme choice we assume that \( \Omega_k = \{j\} \) for all \( k = 1, \ldots, d \), and some \( j \in \{1, \ldots, d\} \). In other words, we would like to recover the original signal \( x \) from its temporal samples at a single spatial location.

One would expect this to be possible only if the system is “well-mixed”, and, in fact, in some sense this is sufficient. To see this, let us assume that \( A = UDU^* \) is positive definite and \( U = (u_{jk}) \) is the unitary that diagonalizes \( A \), so that \( D \) is a diagonal matrix with eigenvalues \( \lambda_1, \ldots, \lambda_d \). Then the reduced dynamical sampling matrix \( \mathcal{A}_r \) obtained from \( A \) by eliminating the zero rows satisfies

\[
\mathcal{A}_r = \begin{pmatrix}
1 & 1 & \ldots & 1 \\
\lambda_1 & \lambda_2 & \ldots & \lambda_d \\
\vdots & \vdots & \ddots & \vdots \\
\lambda_1^{d-1} & \lambda_2^{d-1} & \ldots & \lambda_d^{d-1}
\end{pmatrix}
\begin{pmatrix}
u_{j1} & 0 & \ldots & 0 \\
0 & u_{j2} & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & u_{jd}
\end{pmatrix}
U^*.
\]

Since the first of the matrices in the above product is Vandermonde, \( \mathcal{A}_r \) is invertible if and only if all eigenvalues of \( A \) are distinct and the \( j \)-th row of the “mixing” matrix \( U \) has no zero entries.

We remark that the case of sampling at just two nodes already presents a non-trivial problem.

Remark II.1.3. Observe that one step of the wavelet cascade algorithm [6, 10, 16, 22, 23, 28, 38] is very similar to dynamical sampling. Indeed, if we allow two circular convolution matrices \( A \) and \( B \) in place of \( A \) and \( A^2 \), our mathematical set-up will cover the wavelet algorithm. On the other hand, the physical nature of our problem is fundamentally different – we are not free to choose \( A \) and \( B \) since they are determined by the dynamical system. Moreover, \( A \) and \( A^2 \) cannot be used in the wavelet method because of the Smith-Barnwell condition.

Remark II.1.4. Another mathematical approach to the study of sensor networks involves frames and fusion frames [12, 13, 39]. In the frame theoretic language the dynamical sampling problem consists of describing all subsets of the frame formed by the rows of the matrices \( I \), \( A \), \( A^2 \), etc., which are themselves frames for \( \ell^2(\mathbb{Z}_d) \). Yet, another related area of research is distributed sampling, see, e.g. [25].

II.2 Dynamical Sampling in Invariant Evolution Systems.

The general problem outlined in the previous section is very hard. This is clearly indicated by the examples above. In practice, however, one of the most important cases is represented by (spatially) invariant evolution systems in which the matrix \( A \) is circular and the subsampling is regular and independent of \( n \). In this case, the matrix
A represents the (circular) convolution operator with a fixed vector \( a \in \ell^2(\mathbb{Z}_d) \) and

\[
S(\Omega_n) = S_m : \ell^2(\mathbb{Z}_d) \rightarrow \ell^2(\mathbb{Z}_d), \quad n = 1, \ldots, N, \tag{II.4}
\]
is an operator of subsampling by some fixed factor \( m \in \mathbb{N} \). In this way, a vector \( x \in \ell^2(\mathbb{Z}_d) \) representing the signal at time \( t = 0 \) is sampled only at a fraction \( d/m \) of its components, and subsequently the vectors \( A^{n-1}x, n = 2, \ldots, N, \) are sampled at the same locations. Note that for this case, we would need a minimum of \( d \) “generalized samples” to recover \( x \). Thus, by choosing \( N = m \), we meet this minimum requirement. In effect, we have traded spatial samples for an equal number of “time samples”, thus reducing the number of measuring devices by a factor of \( m \) and activating each device \( m \) times more frequently.

To avoid unnecessary technicalities we let \( d = 2K + 1 \) and assume that \( J = d/m \) is an integer (so that \( d, m, \) and \( J \) are odd). Then the \((k, k)\) entry of the matrix \( S_m \) equals 1 if \( m \) divides \( K + 1 - k \) and is 0 otherwise. Clearly, in practice, any reasonable model can be tweaked to satisfy these conditions.

The following proposition is at the heart of the problem in this special case. In its formulation, we use the notation \( \hat{a} = \mathbf{F}_d a \) for the normalized (unitary) \( d \)-dimensional discrete Fourier transform (DFT) of \( a \) defined as

\[
\hat{a}(k) = \sqrt{\frac{1}{d}} \sum_{n=0}^{d-1} a(n) e^{-\frac{2\pi in k}{d}}.
\]

**Proposition II.2.1.** A vector \( x \in \ell^2(\mathbb{Z}_d) \) can be recovered from the measurements \( y_n = S_m A^{n-1}x, n = 1, \ldots, m, \) if and only if the \( J = d/m \) matrices

\[
\mathcal{A}_m(k) = \begin{pmatrix}
1 & 1 & 1 \\
\hat{a}(k) & \hat{a}(k+J) & \ldots & \hat{a}(k+(m-1)J) \\
\vdots & \vdots & \ddots & \vdots \\
\hat{a}^{(m-1)}(k) & \hat{a}^{(m-1)}(k+J) & \ldots & \hat{a}^{(m-1)}(k+(m-1)J)
\end{pmatrix}, \tag{II.5}
\]

\( k = 0, \ldots, J-1, \) are invertible.

**Proof.** Using the identities

\[
(S_mz)^\wedge(k) = \frac{1}{m} \sum_{\ell=0}^{m-1} \hat{z}(k+J\ell), \quad k = 1, \ldots, d, \ z \in \ell^2(\mathbb{Z}_d), \tag{II.6}
\]

\[
(Ax)^\wedge(k) = (a * x)^\wedge(k) = \hat{a}(k) \hat{x}(k), \tag{II.7}
\]
we get
\[
\begin{pmatrix}
\hat{y}_1(k) \\
\hat{y}_2(k) \\
\vdots \\
\hat{y}_m(k)
\end{pmatrix} = \mathcal{A}_m(k) \begin{pmatrix}
\hat{x}(k) \\
\hat{x}(k+J) \\
\vdots \\
\hat{x}(k+(m-1)J)
\end{pmatrix},
\tag{II.8}
\]
or, using a more compact notation,
\[
\bar{\mathbf{y}}(k) = \frac{1}{m} \mathcal{A}_m(k) \bar{\mathbf{x}}(k),
\tag{II.9}
\]
where \( \bar{\mathbf{y}}(k) = (\hat{y}_1(k), \hat{y}_2(k), \ldots, \hat{y}_m(k))^T \) and \( \bar{\mathbf{x}}(k) = (\hat{x}(k), \hat{x}(k+J), \ldots, \hat{x}(k+(m-1)J))^T \) \( k = 1, \ldots, J \). The proposition follows. Note that the rearrangement mapping \( R : \ell^2(\mathbb{Z}_d) \to (\ell^2(\mathbb{Z}_J))^m \) defined by \( R\hat{x} = \bar{\mathbf{x}} \) is an isometric isomorphism and the vectors \( \hat{y}_n \) are \( J \) periodic.

Clearly, each matrix \( \mathcal{A}_m(k) \) in (II.5) is a Vandermonde matrix; thus it is invertible if and only if the values \( \{\hat{a}(k+\ell J) : \ell = 0, \ldots, m-1\} \) are distinct. If some of these values coincide, the signal \( x \) cannot be recovered unless we take extra spatial samples. It is not hard to outline the procedure that allows one to prescribe which extra spatial samples may be taken. It is easily seen that the kernels of the matrices \( \mathcal{A}_m(k) \), \( k = 1, \ldots, J \), are generated by vectors \( v_j \in \ell^2(\mathbb{Z}_d) \) such that each \( \hat{v}_j \) has exactly two non-zero components of which one is equal to 1 and the other is -1. Assuming that the nullity of the matrix \( \mathcal{A}_m(k) \) equals \( v_k \), \( k = 1, \ldots, J \), there are exactly \( v = \sum_{k=1}^{J} v_k \) of such linearly independent vectors \( v_j \). We then form a \( d \times v \) matrix \( V \) using these vectors as columns and call a subset \( \Omega_0 \subseteq \{1, \ldots, d\} \) an admissible extra sampling set if the rows of \( V \) indexed by \( \Omega_0 \) form a left invertible submatrix. The vector \( x \) can then be recovered from vectors \( y_n, n = 1, \ldots, m \), and the extra samples \( x(\ell), \ell \in \Omega_0 \). Observe that we need to take at least \( v \) extra samples and it is possible to take exactly \( v \) of them. We have essentially proved the following result.

**Theorem II.2.2.** Let \( a \in \ell^2(\mathbb{Z}_d) \) be such that the circular convolution matrix \( A \) defined by \( a \) is invertible. Consider the dynamical sampling procedure defined by the matrices \( S(\Omega_n) = S_m, n = 1, \ldots, N \), as in (II.4) and an admissible extra sampling set \( \Omega_0 \). Then any \( x \in \ell^2(\mathbb{Z}_d) \) can be recovered from the set of vectors \( \{y_0 = S(\Omega_0)x, y_n = S_mA^{n-1}x, n = 1, \ldots, m\} \).

In the case covered by the above theorem, we will write the dynamical sampling
matrix and the measurement vector \( y \) as

\[
\mathbf{A} = \begin{pmatrix}
S(\Omega_0) \\
S(\Omega_1) \\
S(\Omega_2) \\
\vdots \\
S(\Omega_m)A^{m-1}
\end{pmatrix}
\]

and

\[
y = \begin{pmatrix}
y_0 \\
y_1 \\
\vdots \\
y_m
\end{pmatrix}
\]

(II.10)

rather than \( \mathbf{A} = \begin{pmatrix}
S(\Omega_0 \cup \Omega_1) \\
S(\Omega_2) \\
\vdots \\
S(\Omega_m)A^{m-1}
\end{pmatrix} \) as we would in the general case. Recall that in this case we have \( S(\Omega_n) = S_m \) for \( n = 1, \ldots, m \).

**Remark II.2.3.** Note that the number of samples needed for the recovery in the above theorem is \( d + \nu \). Since generically \( \nu \ll d \), the oversampling factor is often negligible. It is also clear from the Vandermonde structure of the matrices (III.2) that adding more time samples at the same locations provides no additional information about \( x \), thus justifying our choice of \( N = m - 1 \). On the other hand, in the presence of noise and once an appropriate set \( \Omega_0 \) is chosen, additional time samples may be used to improve the estimation of \( x \).

The special case when \( a \) is a real symmetric kernel such that \( \hat{a} \) is monotonic (decreasing) on \( \{0, \ldots, K = \frac{d-1}{2} \} \) is encountered frequently in applications. The symmetry reflects the fact that there is often no preferential direction for physical kernels and monotonicity is a reflection of energy dissipation. Particularly interesting is the fact that in this case one can exactly specify the set \( \Omega_0 \).

**Theorem II.2.4.** Assume that in addition to the assumptions of Theorem II.2.2 the kernel \( a \) is real symmetric and \( \hat{a} \) is strictly monotonic on \( \{0, \ldots, K = \frac{d-1}{2} \} \). Then a set \( \Omega_0 \subseteq \{1, \ldots, d\} - m\mathbb{Z} \) is an admissible extra sampling set if and only if it contains a set of cardinality \( \frac{m-1}{2} \) such that no two elements of which are \( m \)-congruent or have a sum divisible by \( m \).

Before embarking on the proof of the above theorem we make a remark and state a corollary.

**Remark II.2.5.** A natural choice of \( \Omega_0 \) in the above theorem is

\[
\Omega_0 = \left\{ -K, -K + 1, \ldots, -K + \frac{m-1}{2}, K - \frac{m-1}{2}, \ldots, K - 1, K \right\}.
\]

Alternatively, we may assume that \( \text{supp} x \subseteq \left[ -K + \frac{m-1}{2}, K - \frac{m-1}{2} \right] \).

**Corollary II.2.6.** Assume that \( x \in \ell^2(\mathbb{Z}_d) \) satisfies

\[
\text{supp} x \subseteq \left[ -K + \frac{m-1}{2}, K - \frac{m-1}{2} \right].
\]
Then $x$ is completely recoverable from the samples $y_n$, $n = 1, \ldots, m$.

Proof of Theorem II.2.4. Consider the matrices $A_m(k)$, $k = 0, \ldots, J - 1$, defined by (III.2). The symmetry and monotonicity conditions imposed on $a$ imply that all these matrices except $A_m(0)$ are invertible and the rank of $A_m(0)$ is $\frac{m+1}{2}$. Indeed, the only way a $p$-th and an $\ell$-th columns of a matrix $A_m(k)$ coincide is when $\hat{a}(k + pJ) = \hat{a}(k + \ell J)$ for some integers $p, \ell \in [0, m - 1]$. This may only happen if $(k + pJ) + (k + \ell J)$ is divisible by $d = mJ$ and, hence, $2k$ is divisible by $J$. Since $J$ is odd and $0 \leq k < J$, this is only possible when $k = 0$ in which case the first column of the matrix $A_m(0)$ is distinct from all other columns and there is also $\frac{m-1}{2}$ different pairs of equal columns. This structure of $A_m(0)$ implies that the kernel vectors $v_j \in \ell^2(\mathbb{Z}_d)$ can be defined by

$$v_j(\ell) = \sin \frac{2\pi}{m} j\ell, \quad j = 1, \ldots, \frac{m-1}{2}, \ \ell = 1, \ldots, d. \quad (\text{II.11})$$

From the above formula it is clear that $m$-congruent rows of the matrix $V$ coincide and are opposite of each other if the sum of their indices is divisible by $m$. On the other hand, the linear independence of the vectors in (II.11) implies that $\Omega_0$ chosen according to the statement of the theorem is admissible.

Corollary II.2.7. Assume that the vector $x$ is $(J - 1)$-sparse, that is, has at most $J - 1$ non-zero components. Then $x$ is completely recoverable from the samples $y_n$, $n = 1, \ldots, m$.

Proof. Observe that invertibility of the matrices $A_m(k)$, $k = 1, \ldots, J - 1$, allows us to recover all Fourier coefficients of $x$ other than the ones divisible by $J$. From the structure of the kernel of $A_m(0)$ we see that we can also recover the coefficient $\hat{x}(0)$. This gives us $2J - 1$ consecutive Fourier coefficients of $x$. Therefore, using Prony’s method [8], we can recover $x$ provided that it is $J - 1$ sparse. Alternatively, if the sparsity level is sufficiently smaller than $J - 1$ we can use compressed sensing techniques [11] for recovery.

II.3 Stability in the Presence of Noise.

Let us now consider the recovery problem of Theorem II.2.4 in the presence of additive noise. More precisely, we consider the problem of recovering the vector $x$ from corrupted measurements $\tilde{y} = y + \eta$, where $\eta = (\eta_0, \ldots, \eta_m)^T$ and each $\eta_k$, $k = 0, \ldots, m$, is a random vector with independent identically distributed (iid) components that have mean 0 and variance $\sigma^2$ (in the case that a support condition is assumed on $x$, we have $\eta_0 = 0$). We write the problem in the form $y + \eta = A x$, where $A$ and $y$ are as in (II.10). Applying the DFT, combining all equations (III.8) into one, and adding the
extra samples we get

$$\vec{y}_e + \vec{\eta}_e = A \hat{x},$$

where $\hat{x} = R \hat{x}$ is as in (II.9),

$$\vec{\eta}_e = \left( \downarrow_{\Omega_0} [\eta_0] \right), \quad \vec{y}_e = \left( \downarrow_{\Omega_0} [y_0] \right),$$

the operator $\downarrow_{\Omega_0}: \ell^2(\mathbb{Z}_d) \to \ell^2(\mathbb{Z}_{|\Omega_0|})$ removes from its input vector the components that are not in $\Omega_0$, $\vec{\eta}$ is a $d$-dimensional random vector with iid components that have mean 0 and variance $\sigma^2$,

$$A = \frac{1}{m} \begin{pmatrix} V_0 & V_1 & \ldots & V_{J-1} \\ \mathcal{A}_m(0) & 0 & \ldots & 0 \\ 0 & \mathcal{A}_m(1) & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \mathcal{A}_m(J-1) \end{pmatrix},$$

and $R$ is the rearrangement operator defined after (II.9).

The matrix $A$ has full rank because the set $\Omega_0$ is chosen to be admissible. Thus, $A$ has the Moore-Penrose left inverse $A^\dagger = (A^* A)^{-1} A^*$. Note that

$$\mathbb{E} \| x - A^\dagger y \|^2 = \mathbb{E} \| \hat{x} - A^\dagger \vec{y}_e \|^2 \leq (d + |\Omega_0|) \sigma^2 \| A^\dagger \|^2 \leq \frac{(d + |\Omega_0|) \sigma^2}{\lambda_{\min}(A^* A)},$$

where $0 < \lambda_{\min}(A^* A)$ is the minimal eigenvalue of $A^* A$. Thus, estimating the expected value of the recovery error under these assumptions amounts to finding upper and lower bounds for the norm of the Moore-Penrose left inverse $A^\dagger$, or, equivalently, for the minimal singular value of $A$.

Observe that $A$ has the structure

$$A = \begin{pmatrix} P & Q \\ T & 0 \\ 0 & D \end{pmatrix},$$

where $P$ and $D$ are invertible $m \times m$ and $(d - m) \times (d - m)$ matrices, respectively. In fact, $D$ is the block diagonal matrix with $\mathcal{A}_m(k), k = 1, \ldots, J-1$, on the main diagonal. Note that a left inverse is then given by
$$A^\ell = \begin{pmatrix} P^{-1} & 0 & -P^{-1}QD^{-1} \\ 0 & 0 & D^{-1} \end{pmatrix}. \quad (\text{II.17})$$

Since the Moore-Penrose left inverse has the smallest operator norm among all left inverses we have

$$\|A^\dagger\| \leq \|A^\ell\| \leq \max\{\|P^{-1}\|, \|D^{-1}\|\} + \|P^{-1}\|\|D^{-1}\|\|Q\|. \quad (\text{II.18})$$

From (II.14) and (II.18) we see that given $a$, one can easily compute an upper bound on the $\|A^\dagger\|$ by finding maximal singular values of at most $J$ many $m \times m$ matrices.

For example, let us choose $\Omega_0 = \{0, 1, \ldots, m - 1\}$. In this case we have

$$A = \frac{1}{m}\begin{pmatrix} mF_m^* & mD_1F_m^* & \ldots & mD_JF_m^* \\ \mathcal{A}_m(0) & 0 & \ldots & 0 \\ 0 & \mathcal{A}_m(1) & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \mathcal{A}_m(J-1) \end{pmatrix}, \quad (\text{II.19})$$

where $D_k, k = 1, \ldots, J - 1$, are diagonal matrices with appropriate $d$-th roots of unity on the diagonal. The left inverse of $A$ in (III.49) is then given by

$$A^\ell = m\begin{pmatrix} \frac{1}{m}F_m & 0 & -F_mD_1F_m^* & \ldots & -F_mD_JF_m^* & \mathcal{A}_m^{-1}(1) & \ldots & \mathcal{A}_m^{-1}(J - 1) \\ 0 & 0 & \mathcal{A}_m^{-1}(1) & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & \mathcal{A}_m^{-1}(J-1) \end{pmatrix}, \quad (\text{II.20})$$

and (II.18) becomes

$$\|A^\dagger\| \leq \|A^\ell\| \leq (1 + \sqrt{J-1}) \max_{1 \leq k \leq J-1} \{1, m\|\mathcal{A}_m^{-1}(k)\|\}. \quad (\text{II.21})$$

Hence, if $\|\mathcal{A}_m^{-1}(k)\| \leq \frac{1}{m}, k = 1, \ldots, m$, we have

$$\|A^\dagger\| \leq 1 + \sqrt{J-1}. \quad (\text{II.21})$$

Otherwise we use the $\ell^\infty \to \ell^\infty$ operator norm estimate for inverses of Vandermonde matrices [21] and the inequality $\|\cdot\| \leq \sqrt{m}\|\cdot\|_{\infty \to \infty}$ to obtain

$$\|A^\dagger\| \leq m^\frac{3}{2}(1 + \sqrt{J-1})\mathcal{R}, \quad (\text{II.22})$$
where
\[
\mathcal{R} = \max_{0 \leq i \leq m-1} \prod_{1 \leq k \leq J-1}^{m-1} \frac{1 + |\hat{a}(k + jJ)|}{|\hat{a}(k + jJ) - \hat{a}(k + iJ)|}.
\] (II.23)

Finally, assuming that \( \parallel \hat{a} \parallel_{\infty} \leq 1 \) and letting \( \varepsilon = \min_{p,q} |\hat{a}(p) - \hat{a}(q)| \), where the minimum is taken over all \( p \neq q \in \mathbb{Z}_d \) such that \( J \) divides \( p - q \) and \( |\hat{a}(p) - \hat{a}(q)| > 0 \), we get (combining (II.21) and (II.22))
\[
\|A^+\| \leq (1 + \sqrt{J-1}) \max \left\{ 1, m^3 \left( \frac{2}{\varepsilon} \right)^{m-1} \right\}.
\] (II.24)

In particular, we see that dynamical sampling is robust for reasonable filters \( a \) when the subsampling factor \( m \) is small. As expected, robustness deteriorates quickly as \( m \) grows. To give some justification of the last statement we estimate \( \|A^+\| \) from below by obtaining an upper bound on the smallest singular value \( s_{\min}(A) \).

We will again use (III.47) for the general form of the matrix \( A \). This time, however, we partition \( A \) in (II.14) so that \( P = \frac{1}{m} V_0, T = \frac{1}{m} \mathcal{A} m(0) \), etc. Computing \( AA^* \) we get
\[
AA^* = \frac{1}{m^2} \begin{pmatrix}
* & * & * \\
* & TT^* & 0 \\
* & 0 & DD^*
\end{pmatrix},
\] (II.25)

where the matrices in the first row have \( |\Omega_0| \) rows. Observe that \( AA^* \) has rank \( d \) and, therefore, its smallest positive eigenvalue coincides with \( s_{\min}(A) \). Using a well-known generalization of Cauchy interlacing inequalities (see [18,26]) for the principal submatrix \( B = (TT^* \ 0 \ DD^*) \) we obtain
\[
s_{\min}^2(A) = \lambda_d(AA^*) \leq \lambda_{d-|\Omega_0|}(B),
\] (II.26)

where \( \lambda_j \) denotes the \( j \)-th largest eigenvalue of the corresponding matrix. Clearly, given \( a \), all one needs to do to find \( \lambda_{d-|\Omega_0|}(B) \) is compute at most \( J \) singular value decompositions of \( m \times m \) Vandermonde matrices which is an easy task of \( O(dm^2) \) operations.

If we take the minimal possible \( |\Omega_0| \), i.e. \( |\Omega_0| = m^{-1} \), we can obtain a more explicit estimate. In this case,
\[
s_{\min}(A) \leq \lambda_{d-\frac{m}{2}}^{\frac{1}{d-\frac{m}{2}}}(B) \leq \left[ \max_{1 \leq k \leq J-1} \left\{ m \|\mathcal{A}^{-1}_{m}(k)\| \right\} \right]^{-1}.
\] (II.27)

Reusing the estimates from [21] as we did in (II.22) and assuming that all \( \hat{a}(k), 1 \leq
\( k \leq d \), are positive we get
\[
\|A^\dagger\| = s_{\min}^{-1}(A) \geq m \hat{\kappa}, \tag{II.28}
\]
where \( \hat{\kappa} \) is given by (II.23). In particular, if 0 < \( \hat{a}(k) \leq 1 \), 1 \( \leq k \leq d \), and \( \delta = \max_{p,q} |\hat{a}(p) - \hat{a}(q)| \), where the maximum is taken over all \( p, q \in \mathbb{Z}_d \) such that \( J \) divides \( p - q \), we get
\[
\|A^\dagger\| = s_{\min}^{-1}(A) \geq m \delta^{1-m} > m. \tag{II.29}
\]
Clearly, increasing \(|\Omega_0|\) alleviates the growth of \( \|A^\dagger\| \).
CHAPTER III

DYNAMICAL SAMPLING IN INFINITE DIMENSIONS.

In this chapter, we study the dynamical sampling problem in infinite dimensions when the subsampling is regular and the evolutionary rule is given by convolution. In section III.1, we formulate the dynamical sampling problem in infinite dimensions. In section III.2, we give sufficient sampling sets for stable recovery of signals when the system is governed by a typical low pass filter, and in section III.3, we study the stability of the reconstruction operator. In section III.4, we extend the results of section III.2 for generic convolution operators. The results in sections III.2 and III.3 are similar to those in the previous chapter for the finite dimensional model; however, there are fundamental differences and the techniques have been amended by methods of functional analysis and topology. Additionally, this chapter contains results which have finite dimensional analogs not found in previous chapter.

III.1 Problem Formulation

By $x \in \ell^2(\mathbb{Z})$ we model an unknown spatial signal at time $t = 0$. Let $a \in \ell^2(\mathbb{Z})$ represent the kernel of an evolution operator $A$ so that $Ax = a * x$ and the signal at time $t = j$ is given by

$$a^j * x = (a * \ldots * a)_j * x.$$  

We also restrict ourselves to regular subsampling, so that $S = S_m$ is the operation of subsampling by a factor of $m$, i.e., $(S_m z)(k) = z(mk)$. The sampled signal $y_j$ at time $t = j$ is given by $y_j = S_m(a^j * x)$. The dynamical sampling procedure is written as

$$y = Ax \quad \text{(III.1)}$$

where $A$ is the dynamical sampling operator from $(\ell^2(\mathbb{Z}))$ to $(\ell^2(\mathbb{Z}))^N$ and

$$y = (y_0, y_1, \ldots, y_{N-1}) = (S_m x, S_m(a * x), \ldots, S_m(a^{N-1} * x)).$$

The goal of this paper is to study when $A$ has a bounded left inverse, giving a bounded reconstruction operator. The boundedness is necessary for stability when additive noise is present in the samples. The expected discrepancy, $\bar{x} - x$, between the recovered signal $\bar{x}$ and the original signal $x$ is controlled by the norm of the reconstruction operator.

If $A$ does not have a bounded left inverse, it may still be the case that $A$ is injective. In this case, we are often able to define a new sampling operator $\tilde{A}$ that does have
a bounded left inverse by expanding the dynamical sampling operator $A$ to include an additional sampling set. Describing such additional sampling sets is the focus of sections III.2 and III.4.

The main method of proof for the results in this chapter is as follows. Fourier techniques are used to reduce the study of the dynamical sampling operator $A$ to the study of an operator defined by pointwise matrix multiplication on the torus. Then a left inverse is defined in the Fourier domain by pointwise multiplication by the left inverse of such matrices, when such left inverses exist. For $z \in \ell^1(\mathbb{Z})$ the Fourier transform is defined on the torus $T \simeq [0, 1)$ by

$$\hat{z}(\xi) = \sum_{n \in \mathbb{Z}} z(n) e^{-2\pi i n \xi}, \xi \in T.$$ 

Below we give necessary and sufficient conditions for $A$ to have a bounded left inverse. This result can be derived from well known results in filterbank literature or from an earlier paper by Papoulis [30]. However, filterbank techniques are ill-suited for proving the remaining results in this section, which involve expanding the spatial sampling sets. In order to provide a cohesive theory, we prove the result below directly using the Poisson summation formula.

**Proposition III.1.1.** Assume that $\hat{a} \in L^\infty(T)$ and fix $m \in \mathbb{Z}^+$. Define

$$A(\xi) = \begin{pmatrix} 1 & \hat{a}(\xi/m) & \hat{a}(\xi+1/m) & \ldots & \hat{a}(\xi+m-1/m) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \hat{a}^{(m-1)}(\xi/m) & \ldots & \hat{a}^{(m-1)}(\xi+1/m) & \ldots & \hat{a}^{(m-1)}(\xi+m-1/m) \end{pmatrix}, \quad (III.2)$$

$\xi \in T$. Then $A$ in (III.1) has a bounded left inverse for some $N \geq m$ if and only if there exists $\alpha > 0$ such that the set $\{ \xi : |\det \mathcal{A}(\xi)| < \alpha \}$ has zero measure. Consequently, $A$ in (III.1) has a bounded left inverse for some $N \geq m$ if and only if $A$ has a bounded left inverse for all $N \geq m$.

The proof of proposition III.1.1 uses the identity $(a * z)\hat{}(\xi) = \hat{a}(\xi)\hat{z}(\xi)$, the Poisson summation formula

$$(S_m z)\hat{}(\xi) = \frac{1}{m} \sum_{l=0}^{m-1} \hat{z}(\frac{\xi+l}{m}) \quad (III.3)$$

and the lemma below. Lemma III.1.2 and equation (III.3) are proved in chapter VI.

**Lemma III.1.2.** Suppose an operator $B : (L^2(T))^m \rightarrow (L^2(T))^n$ is defined by $(Bx)(\xi) = \mathcal{B}(\xi)x(\xi)$ where the map $\xi \mapsto \mathcal{B}(\xi)$ from $T$ to the space of $n \times m$ matrices $\mathcal{M}^{nm}$ is measurable. Then $\|B\|_{op} = \text{ess sup}_T \|\mathcal{B}(\xi)\|_{op}$.
Proof of Proposition III.1.1. Using equation (III.3), the subsampled signal \( y_j \) at time \( t = j \) can be written as

\[
\hat{y}_j(\xi) = (S_m(a^j \ast x))(\xi) = \frac{1}{m} \sum_{l=0}^{m-1} \hat{a}^j\left(\frac{\xi + l}{m}\right) \hat{x}\left(\frac{\xi + l}{m}\right)
\]

(III.4)

Expressing this in matrix form, we have the equation

\[
\begin{pmatrix}
\hat{y}_0(\xi) \\
\hat{y}_1(\xi) \\
\vdots \\
\hat{y}_{N-1}(\xi)
\end{pmatrix} = \mathcal{A}_N(\xi) \begin{pmatrix}
\hat{x}(\frac{\xi}{m}) \\
\hat{x}(\frac{\xi + 1}{m}) \\
\vdots \\
\hat{x}(\frac{\xi + m-1}{m})
\end{pmatrix},
\]

where (III.5)

\[
\mathcal{A}_N(\xi) = \begin{bmatrix}
1 & \ldots & 1 \\
\hat{a}(\frac{\xi}{m}) & \ldots & \hat{a}(\frac{\xi + m-1}{m}) \\
\vdots & \ddots & \vdots \\
\hat{a}^{N-1}(\frac{\xi}{m}) & \ldots & \hat{a}^{N-1}(\frac{\xi + m-1}{m})
\end{bmatrix}.
\]

(III.6)

Define \( H : L^2(\mathbb{T}) \to (L^2(\mathbb{T}))^m \) to be the isometry given by

\[
(Hz)(\xi) = \frac{1}{\sqrt{m}} \left( z\left(\frac{\xi}{m}\right), z\left(\frac{\xi + 1}{m}\right), \ldots, z\left(\frac{\xi + m-1}{m}\right) \right)^T,
\]

and define \( \mathcal{F}_N : (\ell^2(\mathbb{Z}))^N \to (L^2(\mathbb{T}))^N \) to be the one-dimensional Fourier transform applied to each component of the product space \((\ell^2(\mathbb{Z}))^N\), and define \( \mathcal{F} = \mathcal{F}_N \). Then we can write (III.5) in more compact notation as

\[
\hat{y}(\xi) = \frac{1}{m} \mathcal{A}_N(\xi) \mathcal{F}(\xi),
\]

(III.8)

Define the operator \( A_N : (L^2(\mathbb{T}))^m \to (L^2(\mathbb{T}))^N \) by \( (A_N \mathcal{F})(\xi) = \mathcal{A}_N(\xi) \mathcal{F}(\xi) \). Then the dynamical sampling operator can be written as a product of operators as follows

\[
A = \frac{1}{\sqrt{m}} \mathcal{F}_N^{-1} A_N H \mathcal{F}. \quad \text{(III.9)}
\]

Since \( \mathcal{F}, \mathcal{F}_N \) and \( H \) are isometries, the operator \( A \) has a bounded left inverse if and only if the operator \( A_N \) has a bounded left inverse. By lemma III.1.2, it suffices to study the left invertibility of \( \mathcal{A}_N(\xi) \) for each \( \xi \in \mathbb{T} \).

The matrix \( \mathcal{A}_N(\xi) \) is created by adding rows to the matrix \( \mathcal{A}(\xi) \). Thus, if \( \mathcal{A}(\xi) \) has full rank, then \( \mathcal{A}_N(\xi) \) also has full rank. Alternatively, since \( \mathcal{A}(\xi) \) is a Vandermonde matrix, it is invertible if and only if no two columns coincide. When two columns of \( \mathcal{A}(\xi) \) coincide, the corresponding columns of \( \mathcal{A}_N(\xi) \) also coincide.
Thus, $A_N(\xi)$ has a left inverse if and only if $A(\xi)$ has full rank, or, equivalently, $\det A(\xi) \neq 0$. The conclusion follows from lemma III.1.2.

Under the conditions of proposition III.1.1 a vector $x \in \ell^2(\mathbb{Z})$ satisfying (III.1) can be recovered in a stable way from the measurements $y_n, n = 0, \ldots, N-1$, for any $N \geq m$. The proof shows that in the case $N = m$ the operator $A$ is, in fact, invertible and not just left invertible. For the case $N < m$, the operator $A$ is not injective and hence no recovery of $x$ is possible. We also note that if a signal $x$ cannot be recovered from the dynamical samples in proposition III.1.1 then taking additional samples at the same spatial locations will not help. The same phenomenon was observed in [27].

If $\hat{a}$ is continuous, then $\det A(\xi)$ is also a continuous function over the compact set $T$. Therefore, an $\alpha$ in proposition III.1.1 exists if and only if $\det A(\xi) \neq 0$ for all $\xi \in T$. This fact is captured in the corollary below.

**Corollary III.1.3.** Suppose $\hat{a} \in C(T)$. Then $A$ in (III.1) has a bounded left inverse for some (and, hence, all) $N \geq m - 1$ if and only if $\det A(\xi) \neq 0$ for all $\xi \in T$.

### III.2 Additional Sampling Sets for Low-Pass Filters

Although proposition III.1.1 gives necessary and sufficient conditions on convolution operators on $\ell^2(\mathbb{Z})$ for this special case of dynamical sampling problem to be solvable, many typical operators encountered in physical systems or in applications do not satisfy these conditions. For example, a typical low-pass filter is the convolution operator in which $\hat{a}$ is real, symmetric, continuous, and strictly decreasing on $[0, \frac{1}{2}]$. The following proposition shows that the dynamical sampling problem cannot be solved in this case without additional samples.

**Proposition III.2.1.** If $\hat{a}$ is real, symmetric, continuous, and strictly decreasing on $[0, \frac{1}{2}]$, then $A(\xi)$ is singular if and only if $\xi \in \{0, \frac{1}{2}\}$.

**Proof.** The Vandermonde matrix $A(\xi)$ in (III.2) is singular if and only if two of its columns coincide. Suppose the $j$-th and $l$-th columns coincide and $j < l$. Then $\hat{a}(\frac{\xi + j}{m}) = \hat{a}(\frac{\xi + l}{m})$. The symmetry and monotonicity conditions on $\hat{a}$ imply that $\frac{\xi + j}{m} = 1 - \frac{\xi + l}{m}$. Then $\xi = \frac{m - j - l}{2}$. Observing that $m - j - l \in \mathbb{Z}$ and $\xi \in T$, we conclude that $\xi \in \{0, \frac{1}{2}\}$.

Because $A(0)$ and $A\left(\frac{1}{2}\right)$ are not invertible, we cannot solve (III.1). To make recovery possible in this case, the sampling set needs to be expanded. In the following theorems, we describe additional sampling sets that enable stable reconstruction when the conditions of proposition III.2.1 are satisfied. It is assumed that these extra samples are taken in addition to the samples $y_0, y_1, \ldots, y_{N-1}$ for some $N \geq m$. 

19
The additional samples are taken in the following way. Let \( T_c \) be the operator that shifts a vector \( z \in \ell^2(\mathbb{Z}) \) to the right by \( c \) units so that \( T_c z(k) = z(k - c) \) so that \( S_{mn} T_c \) is the operator of shifting by \( c \) and then subsampling by a factor of \( mn \) for some \( n \in \mathbb{N} \).

In the theorem below, we describe conditions on the shifts \( c \) and the sampling factor \( nm \) such that additional samples taken at the initial time allow for stable recovery of the initial signal. An example of such a sampling set is illustrated in figure III.2.

**Theorem III.2.2.** Let \( m \in \mathbb{Z}^+ \) be odd. Suppose \( \hat{a} \) is real, symmetric, continuous, and decreasing on \( [0, \frac{1}{2}] \), and let \( \Omega = \{1, \ldots, \frac{m-1}{2}\} \). Then for any odd \( n \in \mathbb{Z}^+ \), the additional sampling given by \( \{S_{mn} T_c\}_{c \in \Omega} \) is sufficient to stably recover any \( x \in \ell^2(\mathbb{Z}) \), i.e. the reconstruction operator is bounded.

**Proof.** The dynamical sampling procedure with the additional samples is written as \( y = \tilde{A} x \) where \( \tilde{A} \) is the dynamical sampling operator from \( (\ell^2(\mathbb{Z}))^{m+\frac{m-1}{2}} \) and

\[
y = \left( y_0, y_1, \ldots, y_{m-1}, S_{mn} T_1 x, S_{mn} T_2 x, \ldots, S_{mn} T_{\frac{m-1}{2}} x \right)^T.
\]

It is shown below that \( \tilde{A} \) has a bounded left inverse. Note that we prove the result for the minimal case \( N = m \), which trivially extends to the case when \( N > m \). The techniques are similar to, but more complicated than, those in the proof of proposition III.1.1.

In the proof of proposition III.1.1, we used equation (III.3) to relate the Fourier transform of the subsampled signal \( y_j \) to the Fourier transform of the original signal \( x \) by writing each \( \hat{y}_j(\xi) \) as a linear combination of the \( m \)-unknowns

\[
\hat{x}\left(\frac{\xi}{m}\right), \hat{x}\left(\frac{\xi + 1}{m}\right), \ldots, \hat{x}\left(\frac{\xi + m - 1}{m}\right).
\]

Similarly, the additional samples are taken by subsampling by a factor of \( mn \) and so equation (III.3) expresses the Fourier transform of the additional samples as a linear combination of \( mn \)-unknowns. The goal is to write the linear combinations from both of these systems in such a way that they can be combined to create a system of \( mn \)-unknowns and \( mn + \frac{m-1}{2} \)-equations.

In order to choose equations with the same unknowns, we consider the formula below for \( (S_{mn} T_c z)^\wedge\left(\frac{n\xi}{n}\right) \) and note that right hand side contains the same variables given by equation (III.4) for \( \hat{y}\left(\xi + \frac{k}{n}\right) \) for \( k = 0, \ldots, n - 1 \). Using (III.3) and the identity \( (T_c z)^\wedge(\xi) = e^{-\frac{12\pi c \xi}{2}} z(\xi) \), we obtain the following formula.
\begin{align}
(S_{mnT_c})^\wedge(n\xi) &= \frac{1}{mn} e^{-i2\pi\xi m} \sum_{l=0}^{mn-1} e^{-i\frac{2\pi l}{m}} \hat{z}(\frac{\xi}{m} + \frac{l}{mn}) \quad \text{(III.10)}
&= \frac{1}{mn} e^{-i2\pi\xi m} \sum_{k=0}^{n-1} \sum_{j=0}^{m-1} e^{-i\frac{2\pi k}{m}} e^{-i\frac{2\pi j}{m}} \hat{z}(\frac{\xi}{n} + \frac{k}{n} + j).
\end{align}

Defining the row vector

\[ \bar{u}_c(k) = e^{-i\frac{2\pi k}{m}} (1, e^{-i\frac{2\pi}{m}}, \ldots, e^{-i\frac{2\pi(m-1)}{m}}), \quad \text{(III.11)} \]

we have

\[ (S_{mnT_c})^\wedge(n\xi) = \frac{1}{mn} e^{-i2\pi\xi m} \sum_{k=0}^{n-1} \bar{u}_c(k) \hat{y}(\xi + \frac{k}{n}). \quad \text{(III.12)} \]

where \( \hat{y}(\xi) = (\hat{x}(\xi), \hat{x}(\xi + \frac{1}{m}), \ldots, \hat{x}(\xi + \frac{m-1}{m}))^T \) as in (III.8).

The original and additional samples are related to the original signal in the Fourier domain by matrix multiplication:

\[ m \begin{pmatrix}
ne^{-i2\pi\xi m} (S_{mnT_1})^\wedge(n\xi) \\
\vdots \\
n e^{-i2\pi(m-1)\xi m} (S_{mnT_{(m-1)}}) \hat{y}(\xi) \\
\hat{y}(\xi + \frac{1}{n}) \\
\vdots \\
\hat{y}(\xi + \frac{n-1}{n})
\end{pmatrix}
= A_\Omega(\xi) \begin{pmatrix}
\hat{y}(\xi) \\
\hat{y}(\xi + \frac{1}{n}) \\
\vdots \\
\hat{y}(\xi + \frac{n-1}{n})
\end{pmatrix}, \quad \text{(III.13)} \]

where \( A_\Omega \) is given by

\[ A_\Omega(\xi) = \begin{pmatrix}
\bar{u}_1(0) & \bar{u}_1(1) & \ldots & \bar{u}_1(n-1) \\
\vdots & \vdots & \ddots & \vdots \\
\bar{u}_{m-1}(0) & \bar{u}_{m-1}(1) & \ldots & \bar{u}_{m-1}(n-1) \\
0 & 0 & \cdots & 0 \\
0 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 0
\end{pmatrix}, \quad \text{(III.14)} \]

and \( \hat{y}(\xi) \) and \( \hat{y}(\xi) \) are given in (III.8).

Similar to the technique of the proof of proposition III.1.1, we want to reduce the
study of the dynamical sampling operator $\tilde{A}$ to the study of the matrices $A_\Omega(\xi)$. To accomplish this, we define the invertible map $J : (L^2(\mathbb{T}))^m \to (L^2(\mathbb{T}/n))^{mn}$ by

$$
(J\tilde{z})(\xi) = \begin{pmatrix}
\tilde{z}(\xi)
\tilde{z}(\xi + \frac{1}{n}) \\
\vdots \\
\tilde{z}(\xi + \frac{m-1}{n})
\end{pmatrix}.
$$

Using $J$, the Fourier transform, and $H$ from (III.7), the dynamical sampling operator $\tilde{A}$ can be expressed as a product of operators so that finding a left inverse of $\tilde{A}$ reduces to finding a left inverse of the matrix $A_\Omega(\xi)$ in (III.14) for each $\xi \in \mathbb{T}/n$.

If $A_\Omega(\xi)$ has full column rank, then it has a left inverse. By Lemma III.1.2 and the fact that $\hat{a}$ is continuous, it suffices to show that the matrix $A(\xi)$ has full rank for every $\xi \in [0, \frac{1}{n}]$; it is not difficult to see that solving (III.13) for $\xi \in [0, \frac{1}{n}]$ is sufficient for the recovery of $\tilde{x}(\xi)$ for all $\xi \in [0, 1]$.

First, if $\xi + \frac{k}{n} \notin \{0, \frac{1}{2}\}, k = 0, \ldots n - 1$, the solvability is implied by Proposition III.2.1. Next, notice that for a fixed $\xi \in [0, \frac{1}{n}]$, we have $\xi + \frac{k}{n} \notin \{0, \frac{1}{2}\}$ for at most one $k = 0, \ldots n - 1$. This follows from the parity of $n$ (n is assumed to be odd). Therefore, for any $\xi \in [0, \frac{1}{n}]$ there is at most one singular block in $A_\Omega(\xi)$. This allows us to consider the singularities of $\mathcal{A}(0)$ and $\mathcal{A}'(\frac{1}{2})$ separately.

Because of the block diagonal structure of the lower portion of $A_\Omega(\xi)$, we can focus only on showing that the additional samples eliminate any singularities created by $\mathcal{A}(0)$ and $\mathcal{A}'(\frac{1}{2})$.

For a fixed $k = 0, \ldots, \frac{m-1}{2}$, we define the $\frac{m-1}{2} \times m$ matrix

$$
U_k = \begin{pmatrix}
\tilde{u}_1(k) \\
\vdots \\
\tilde{u}_\frac{m-1}{2}(k)
\end{pmatrix}.
$$

Since

$$
\langle U_k(c, \cdot), U_k(d, \cdot) \rangle = \sum_{j=0}^{m-1} e^{-\frac{i2\pi k}{m} c} e^{-\frac{i2\pi j}{m} d} e^{\frac{i2\pi j}{m} d} = e^{-\frac{i2\pi (c-d)k}{m}} \sum_{j=0}^{m-1} e^{-\frac{i2\pi (c-d)j}{m}}
$$

$$
= \begin{cases} 
me^{-\frac{i2\pi (c-d)k}{m}}, & (c - d) = 0 \mod m \\
0, & otherwise
\end{cases}
$$

the rows of the matrix $U_k$ form an orthogonal set, and we conclude that it has full rank.
Next, we show that $\left( \begin{array}{c} U_k \\
A(\xi + \frac{k}{n}) 
\end{array} \right)$ has a trivial kernel and, hence, full rank. A vector is in $\ker \left( \begin{array}{c} U_k \\
A(\xi + \frac{k}{n}) 
\end{array} \right)$ if and only if it is in the kernels of both $A(\xi + \frac{k}{n})$ and $U_k$. Therefore, we only need to consider $\xi + \frac{k}{n} \in \{0, \frac{1}{2}\}$.

Under the conditions of Proposition III.2.1, we can completely characterize the kernels of $A(0)$ and $A(\frac{1}{2})$. For simplicity, we assume $m$ is odd and begin indexing the columns of $A(\xi)$ at zero. When $\xi = 0$, the $l$-th column of the Vandermonde matrix $A(0)$ is found by evaluating $\hat{a}$ at $\frac{l}{m}$. By the symmetry and 1-periodicity of $\hat{a}$, we have $\hat{a}(\frac{j}{m}) = \hat{a}(\frac{m-j}{m})$. Therefore, the $j$-th and $(m-j)$-th columns of $A(0)$ coincide, and the kernel of $A(0)$ has dimension $\frac{m-1}{2}$. Similarly, the $j$-th and $(m-j-1)$-th columns of $A(\frac{1}{2})$ coincide for $j = 0, \ldots, \frac{m-3}{2}$ and and the kernel of $A(\frac{1}{2})$ also has dimension $\frac{m-1}{2}$.

The vector $\bar{v}_j$ with a 1 in the $j$-th position, a $(-1)$ in the $(m-j)$-th position, and zeros elsewhere is in the kernel of $A(0)$. Since there are exactly $\frac{m-1}{2}$ of such vectors, the kernel is their span:

$$\ker A(0) = \text{span} \left\{ \left( \begin{array}{c} 0 \\
1 \\
0 \\
\vdots \\
0 \\
-1 
\end{array} \right), \ldots, \left( \begin{array}{c} 0 \\
0 \\
-1 \\
\vdots \\
0 
\end{array} \right) \right\} = \text{span} \{ \bar{v}_j \}_{j=1}^{\frac{m-1}{2}}. \quad (\text{III.17})$$

Similarly, the $j$-th and $(m-j-1)$-th columns of $A(\frac{1}{2})$ coincide for $j = 0, \ldots, \frac{m-3}{2}$, and the vector $\bar{w}_j$ with a 1 in the $j$-th position, a $(-1)$ in the $(m-j-1)$-th position, and zeros elsewhere is in the kernel of $A(\frac{1}{2})$. Therefore,

$$\ker A(\frac{1}{2}) = \text{span} \left\{ \left( \begin{array}{c} 1 \\
0 \\
\vdots \\
0 \\
-1 
\end{array} \right), \ldots, \left( \begin{array}{c} 0 \\
1 \\
\vdots \\
0 \\
0 
\end{array} \right) \right\} = \text{span} \{ \bar{w}_j \}_{j=0}^{\frac{m-3}{2}}. \quad (\text{III.18})$$

Suppose $\bar{x} \in \ker A(0)$. Then $\bar{x} = \sum_{j=1}^{\frac{m-1}{2}} \alpha_j \bar{v}_j$, where $\bar{v}_j$ is defined in (III.26). We want to know if the equation $0 = U_k \bar{x} = \sum_{j=1}^{\frac{m-1}{2}} \alpha_j U_k \bar{v}_j$ has a unique (trivial) solution. This
Figure III.1: An example of a stable sampling scheme in Theorem III.2.2 with $m = 5$ and $n = 7$. The sampling locations are marked by crosses and the extra samples at $t = 0$ are marked as crosses inside squares.

happens if and only if the matrix $B = U_k \left( \bar{v}_1 \ldots \bar{v}_{m-1} \right)$ has full rank. Computing the $(c, j)$ entry of $B$, we have

$$B(c, j) = e^{-i\pi c k n} (e^{-i\pi c j} - e^{-i\pi c (m-j)})$$

(III.19)

Note that $\{1, \cos(\frac{2\pi}{m} c j), \sin(\frac{2\pi}{m} c j) : c = 1, \ldots, \frac{m-1}{2}\}$ is the Fourier basis for $\mathbb{C}^m$. Thus, $\{\sin(\frac{2\pi}{m} c j) : c = 0, \ldots, \frac{m-1}{2}\}$ are linearly independent in $\mathbb{C}^m$. Using the fact that $\{\sin(\frac{2\pi}{m} c j) : c = 0, \ldots, \frac{m-1}{2}\}$ form a basis of $\mathbb{C}^{\frac{m-1}{2}}$. Therefore, the $\frac{m-1}{2} \times \frac{m-1}{2}$ matrix $B$ does, indeed, have full rank.

Similarly, for the case $\mathscr{A}(\frac{1}{2})$, we consider the matrix $D = U_k \left( \bar{w}_0 \ldots \bar{w}_{m-1} \right)$. Its entries are

$$D(c, j) = 2e^{-i\pi c k n} e^{-i\pi c (2j+1)}$$

(III.20)

and, therefore, $D$ has full rank.

Thus, the matrix $A_{\Omega}$ has a bounded left inverse for every $\xi \in \mathbb{T}$ and the theorem is proved.

Remark III.2.3. Proposition III.1.1 and theorem III.2.2 parallel the finite dimensional results we obtained in chapter II. For example, one can use more complicated choices for $\Omega \subset \{1, \ldots, mn - 1\}$ in Theorem III.2.2, and the admissible choices are determined by the same equivalence relations as in the finite dimensional case. Moreover, some of the methods we use for obtaining stability results in Section III.3 are the same as in chapter II. There are, however, subtle but important differences in the infinite dimensional case. For example, in Theorem III.2.2, the dynamical samples without the samples in the extra sampling set $\Omega$ still form a uniqueness set (the operator $A$ has a trivial kernel). The latter was not the case in finite dimensions.

The additional sampling scheme given in theorem III.2.2 use samples taken only at the initial time. Here we present an additional sampling scheme that includes samples
Figure III.2: An example of a stable sampling scheme in Theorem III.2.4 with $m = 5, n = 7$, and $c = 1$. The sampling locations are marked by crosses and the extra samples are marked as crosses inside squares.

taken at each time. When $\hat{a}$ has the properties as in theorem III.2.2, the number of required shifts can be reduced from $\frac{m-1}{2}$ to just 1, thus, reducing the number of spatial samples required. An example of such a sampling set is illustrated in figure III.2.

**Theorem III.2.4.** Let $m \in \mathbb{Z}^+$ be odd. Suppose $\hat{a}$ is real, symmetric, continuous, and decreasing on $[0, \frac{1}{2}]$. Then for any odd $n \in \mathbb{Z}^+$ and any fixed $c$ relatively prime to $m$, the additional sampling given by $\{ (S_{mn} T_c)(a^j \ast x) \}_{j \in \{0, \ldots, m-1\}}$ is sufficient to stably recover any $x \in \ell^2(\mathbb{Z})$, i.e. the reconstruction operator is bounded.

**Proof.** The dynamical sampling procedure with the additional samples is written as $y = \tilde{A}x$ where $\tilde{A}$ is the dynamical sampling operator from $(\ell^2(\mathbb{Z}))$ to $(\ell^2(\mathbb{Z}))^{2m}$ and

$$y = (y_0, y_1, \ldots, y_{m-1}, S_{mn} T_c x, S_{mn} T_c (a \ast x), \ldots, S_{mn} T_c (a^{m-1} \ast x))^T.$$

It is shown below that $\tilde{A}$ has a left inverse.

The structure of this proof resembles that of theorem III.2.2, and the techniques are similar to, but more complicated than, those in the proof of proposition III.1.1. Note that we prove the result for the minimal case $N = m$, which trivially extends to the case when $N > m$.

In the proof of proposition III.1.1, we used equation (III.3) to relate the Fourier transform of the subsampled signal $y_j$ to the Fourier transform of the original signal $x$ by writing each $\hat{y}_j(\xi)$ as a linear combination of the $m$-unknowns

$$\hat{\lambda}(\frac{\xi}{m}), \hat{\lambda}(\frac{\xi + 1}{m}), \ldots, \hat{\lambda}(\frac{\xi + m - 1}{m}).$$

Similarly, the additional samples are taken by subsampling by a factor of $mn$ and so equation (III.3) expresses the Fourier transform of the additional samples as a linear combination of $mn$-unknowns. The goal is to write the linear combinations from both of these systems in such a way that they can be combined to create a system of $mn$-unknowns and $(n+1)m$-equations.
In order to choose equations with the same unknowns, we consider the formula below for \((S_{mn}T_c)^\wedge(n\xi)\) and note that right hand side contains the same variables given by equation (III.4) for \(\hat{y}(\xi + \frac{k}{n})\) for \(k = 0, \ldots, n - 1\). Using equation (III.3) and the identity \((T_c)^\wedge(\xi) = e^{-i2\pi\xi}\hat{z}(\xi)\), we obtain the following formula.

\[
(S_{mn}T_c)^\wedge(n\xi) = \frac{1}{mn} e^{-\frac{\pi i n \xi}{m}} \sum_{l=0}^{mn-1} e^{-\frac{2\pi i n l}{mn}} \hat{z}(\frac{\xi}{m} + \frac{l}{mn}) \tag{III.21}
\]

\[
= \frac{1}{mn} e^{-\frac{\pi i n \xi}{m}} \sum_{k=0}^{n-1} e^{-\frac{2\pi i n k}{m}} \sum_{j=0}^{m-1} e^{-\frac{2\pi i n j}{m}} \hat{z}(\frac{\xi}{m} + \frac{k}{n} + \frac{j}{m}).
\]

The original and additional samples are related to the original signal in the Fourier domain by matrix multiplication:

\[
\begin{bmatrix}
    n e^{-\frac{\pi i n \xi}{m}} (S_{mn}T_c x)^\wedge(n\xi) \\
    n e^{-\frac{\pi i n \xi}{m}} (S_{mn}T_c (a * x))^\wedge(n\xi) \\
    \vdots \\
    n e^{-\frac{\pi i n \xi}{m}} (S_{mn}T_c(a^{m-1} * x))^\wedge(n\xi)
\end{bmatrix} = \mathcal{A}(\xi) \begin{bmatrix}
    \hat{y}(\xi) \\
    \hat{y}(\xi + \frac{1}{n}) \\
    \vdots \\
    \hat{y}(\xi + \frac{n-1}{n})
\end{bmatrix}, \tag{III.22}
\]

where \(\hat{y}(\xi)\) and \(\tilde{x}(\xi)\) are given in (III.8) and \(\mathcal{A}(\xi)\) is the block matrix

\[
\mathcal{A}(\xi) = \begin{bmatrix}
    \mathcal{A}^{c,0}(\xi) & \mathcal{A}^{c,1}(\xi + \frac{1}{n}) & \ldots & \mathcal{A}^{c,n-1}(\xi + \frac{n-1}{n}) \\
    \mathcal{A}(\xi) & 0 & \ldots & 0 \\
    0 & \mathcal{A}(\xi + \frac{1}{n}) & \ldots & 0 \\
    \vdots & \vdots & \ddots & \vdots \\
    0 & 0 & \ldots & \mathcal{A}(\xi + \frac{n-1}{n})
\end{bmatrix}, \text{ where } (III.23)
\]

\[
\mathcal{A}^{c,k}(\xi) = e^{-\frac{\pi i n \xi}{m}} \begin{bmatrix}
    1 & e^{-\frac{i2\pi \xi}{m}} & \ldots & e^{-\frac{i2\pi (m-1) \xi}{m}} \\
    \hat{d}(\frac{\xi}{m}) & e^{-\frac{i2\pi \xi}{m}} & \ldots & e^{-\frac{i2\pi (m-1) \xi}{m}} \\
    \vdots & \vdots & \ddots & \vdots \\
    \hat{d}(m-1)(\frac{\xi}{m}) & e^{-\frac{i2\pi \xi}{m}} & \ldots & e^{-\frac{i2\pi (m-1) \xi}{m}}
\end{bmatrix},
\]

Similar to the technique of the proof of proposition III.1.1, we want to reduce the study of the dynamical sampling operator \(\tilde{A}\) to the study of the matrices \(\mathcal{A}(\xi)\). To
accomplish this, we define the invertible map \( J : (L^2(\mathbb{T}))^m \to (L^2(\mathbb{T}/n))^m \) by

\[
(J\bar{z})(\xi) = \begin{pmatrix}
\bar{z}(\xi) \\
\bar{z}(\xi + \frac{1}{n}) \\
\vdots \\
\bar{z}(\xi + \frac{m-1}{n})
\end{pmatrix}.
\]

Using \( J \), the Fourier transform, and \( H \) from (III.7), the dynamical sampling operator \( \tilde{A} \) can be expressed as a product of operators so that finding a left inverse of \( \tilde{A} \) reduces to finding a left inverse of the matrix \( \mathcal{A}(\xi) \) in (III.22) for each \( \xi \in \mathbb{T}/n \). The remainder of the proof is showing that \( \mathcal{A}(\xi) \) has a left inverse.

In block form,

\[
\mathcal{A}(\xi) = \begin{pmatrix}
D(\xi) & F(\xi) \\
0 & G(\xi)
\end{pmatrix}
\]

where

\[
D(\xi) = \begin{pmatrix}
\mathcal{A}^{c,0}(\xi) \\
\mathcal{A}(\xi)
\end{pmatrix},
\]

\[
F(\xi) = \begin{pmatrix}
\mathcal{A}^{c,1}(\xi + \frac{1}{n}) & \ldots & \mathcal{A}^{c,n-1}(\xi + \frac{n-1}{n}) \\
0 & \ldots & 0
\end{pmatrix},
\]

and \( G(\xi) \) is the block diagonal matrix with \( \mathcal{A}(\xi + \frac{1}{n}), \ldots, \mathcal{A}(\xi + \frac{n-1}{n}) \) on the diagonal. If \( D'(\xi) \) and \( G'(\xi) \) are left inverses of \( D(\xi) \) and \( G(\xi) \), respectively, then a left inverse of \( \mathcal{A}(\xi) \) is given by

\[
\mathcal{A}^{-1}(\xi) = \begin{pmatrix}
D'(\xi) & -D'(\xi)F(\xi)G'(\xi) \\
0 & G'(\xi)
\end{pmatrix}.
\]

It remains to show that \( G(\xi) \) and \( D(\xi) \) have left inverses. Because \( n \) is odd and \( \mathcal{A}(\xi) \) is singular only when \( \xi \in \{0, \frac{1}{2}\} \), for any fixed \( \xi \), \( \mathcal{A}(\xi + \frac{k}{n}) \) is singular for at most one \( k = 0, \ldots, n-1 \). It is shown in the proof of theorem III.2.2 that without loss of generality, we can assume \( \mathcal{A}(\xi + \frac{1}{n}), \ldots, \mathcal{A}(\xi + \frac{n-1}{n}) \) are invertible. Thus, \( G(\xi) \) has a left inverse.

The proof that \( D(\xi) \) has a left inverse is more complicated and relies on the structure of \( \ker \mathcal{A}(0) \) and \( \ker \mathcal{A}(\frac{1}{2}) \). In the proof of theorem III.2.2, it is shown that

\[
\ker \mathcal{A}(0) = \text{span}\left\{ \begin{pmatrix}
0 \\
1 \\
0 \\
\vdots \\
0 \\
-1
\end{pmatrix}, \begin{pmatrix}
0 \\
0 \\
1 \\
\vdots \\
-1 \\
0
\end{pmatrix}, \ldots, \begin{pmatrix}
0 \\
0 \\
0 \\
\vdots \\
0 \\
0
\end{pmatrix} \right\} = \text{span}\{\psi_j\}_{j=1}^{m-1}
\]

and
\[
\text{ker } \mathcal{A}\left(\frac{1}{2}\right) = \text{span}\left\{ \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix} \right\} = \text{span}\{ \bar{w}_j \}_{j=0}^{m-3},
\]

Notice that

\[
\mathcal{A}^{c,k}(\xi) \begin{pmatrix} z_0 \\ z_1 \\ \vdots \\ z_{m-1} \end{pmatrix} = e^{-\frac{\lambda_{2\pi}}{m}k} \mathcal{A}(\xi) \begin{pmatrix} z_0 \\ e^{-\frac{\lambda_{2\pi}}{m}z_1} \\ \vdots \\ e^{-\frac{\lambda_{2\pi}}{m}(m-1)z_{m-1}} \end{pmatrix}, \quad \text{(III.26)}
\]

Define the map \( M_c : \text{ker } \mathcal{A}(\xi) \to \text{ker } \mathcal{A}^{c,k}(\xi) \) by \( M_c z = \begin{pmatrix} z_0 \\ e^{-\frac{\lambda_{2\pi}}{m}z_1} \\ \vdots \\ e^{-\frac{\lambda_{2\pi}}{m}(m-1)z_{m-1}} \end{pmatrix} \). We can now study \( \text{ker } \mathcal{A}^{c,k}(\xi) \) by looking at \( \text{ker } \mathcal{A}(\xi) \). This is summarized in the following claim.

**Claim 1**: The matrix \( D(\xi) \) has a left inverse if and only if

\[
\text{ker } \mathcal{A}(\xi) \cap M_c(\text{ker } \mathcal{A}(\xi)) = \{0\}.
\]

Since \( \text{ker } \mathcal{A}(\xi) = \{0\} \) for \( \xi \notin \{0, \frac{1}{2}\} \), we need only check the cases of \( \xi = 0 \) and \( \xi = \frac{1}{2} \).

Let \( v_1, \ldots, v_{m-1} \) be the basis for \( \text{ker } \mathcal{A}(0) \) given in (III.26). Using the equality for the dimension of subspaces

\[
\dim(U \cap W) = \dim(U) + \dim(W) - \dim(U \cup W),
\]

we see \( \text{ker } \mathcal{A}(0) \cap M_c(\text{ker } \mathcal{A}(0)) = \{0\} \) if and only if

\[
\dim\left( \text{span}\{v_1, \ldots, v_{m-1}, M_c v_1, \ldots, M_c v_{m-1}\} \right) = m - 1.
\]

Note that \( v_j \) has zeros everywhere except the \( j \)-th and \( (m - j) \)-th positions. Thus,
there is linear dependence among the set of vectors
\[ \{v_1, \ldots, v_{m-1}, M_cv_1, \ldots, M_cv_{m-1} \} \]
if and only if \( M_cv_j = \alpha v_j \) for some \( \alpha \in \mathbb{C} \) and \( j = 1, \ldots, \frac{m-1}{2} \). This condition is satisfied if and only if
\[ e^{\frac{j \pi i}{m}} = e^{\frac{j \pi (m-j)}{m}} = e^{-\frac{j \pi i}{m}}, \]
which is satisfied if and only if the product \( cj \) is a multiple of \( m \). Since \( (c, m) = 1 \) and \( j = 1, \ldots, \frac{m-1}{2} \), this can never be satisfied. Thus, \( M_c(\ker \mathcal{A}(0)) \cap \ker \mathcal{A}(0) = \{0\} \).

A similar argument shows that \( M_c(\ker \mathcal{A}(1/2)) \cap \ker \mathcal{A}(1/2) \) is non-trivial if and only if
\[ e^{\frac{j \pi i}{m}} = e^{\frac{j \pi (j+1)}{m}} = e^{-\frac{j \pi i}{m}}, \]
for some \( j = 0, \ldots, \frac{m-3}{2} \). This requires \( c(2j+1) \equiv 0 \mod m \). Since \( (c, m) = 1 \) and \( 0 \leq j \leq \frac{m-3}{2} \), equation (III.28) is never satisfied. Thus, \( M_c(\ker \mathcal{A}(1/2)) \cap \ker \mathcal{A}(1/2) = \{0\} \), and we have proved the theorem.

### III.3 Stability in the Presence of Noise

In this section, we assume that \( \hat{a} \) and \( \Omega \) satisfy the hypotheses of Theorem III.2.2 and consider the recovery of the signal \( f \) in the presence of additive noise. The minimal extra sampling set \( \Omega \) in Theorem III.2.2 allows us to stably recover any signal \( f \in \ell^2(\mathbb{Z}) \).

In the presence of additive Gaussian white noise, however, any linear recovery method does not generally reproduce the original function \( f \). Under the above hypotheses, the expected discrepancy, \( \tilde{f} - f \), between the recovered function \( \tilde{f} \) and the original function \( f \) is controlled by the norm of the operator \( A_\Omega^+ : (L^2(\mathbb{T}))^n \rightarrow (L^2(\mathbb{T}))^n \) defined by \( (A_\Omega^+y)(\xi) = A_\Omega^*(\xi)y(\xi) \), where \( A_\Omega^*(\xi) \) is the Moore-Penrose pseudoinverse of the matrix \( A_\Omega(\xi) \) in (III.50) below. An upper bound for \( \|A_\Omega^+\| \) is given in the following theorem.

**Theorem III.3.1.** If \( \Omega = \{0, \ldots, m-1\} \) and \( \hat{a} \) and \( n \) satisfy the hypotheses of Theorem III.2.2 then
\[ \|A_\Omega^+\| \leq m\beta_1(1 + m\sqrt{n-1}) \]
where \( \beta_1 = \max\{n, \text{ess sup}_{\xi \in J} \|\mathcal{A}^{-1}(\xi)\| \} < \infty \), \( J = \left[ \frac{1}{4n}, \frac{1}{2} \right] - \left[ \frac{1}{4n}, \frac{1}{4n} \right] \cup \left[ \frac{1}{2} + \frac{1}{4n}, 1 - \frac{1}{4n} \right] \), and \( \mathcal{A}(\xi) \) is defined by (III.2).
In the following corollaries we give more explicit bounds for the value of $\beta_1$. There, without loss of generality, we assume that $\sup |\hat{a}(\xi)| \leq 1$.

**Corollary III.3.2.** If $\Omega = \{0, \ldots, m - 1\}$, $\hat{a}$ and $n$ satisfy the hypotheses of Theorem III.2.2, and $\sup |\hat{a}(\xi)| \leq 1$ then

\[ \| A^\dagger_\Omega \| \leq m \beta_2 (1 + m \sqrt{n - 1}) \]

where $\beta_2 = \max \left\{ n, \left( \frac{2}{\delta} \right)^{m-1} \right\} < \infty$, $\delta = \min_{j=0,\ldots,m-1} |\hat{a}(\frac{j\xi + j}{m}) - \hat{a}(\frac{j\xi + j}{m})|$, and $J = \left\{ \frac{1}{m}, \frac{1}{2}, \frac{1}{4m} \right\} \cup \left\{ \frac{1}{2} + \frac{1}{4m}, 1 - \frac{1}{4m} \right\}$.

**Corollary III.3.3.** If $\Omega = \{0, \ldots, m - 1\}$, $\hat{a}$ and $n$ satisfy the hypotheses of Theorem III.2.2, $\sup |\hat{a}(\xi)| \leq 1$, and, in addition, $\hat{a} \in C^1(0, \frac{1}{2})$ and the derivative $\hat{a}'$ of $\hat{a}$ is nonzero (and, hence, negative) on $(0, \frac{1}{2})$, then

\[ \| A^\dagger_\Omega \| \leq m \beta_3 (1 + m \sqrt{n - 1}) , \]

where $\beta_3 = \max \left\{ n, \left( \frac{4mn}{\gamma} \right)^{m-1} \right\}$, $\gamma = \min_M |\hat{a}'(\xi)|$, and $M = \left[ \frac{1}{4mn}, \frac{1}{2} - \frac{1}{4mn} \right]$.

For a Gaussian i.i.d. additive noise $\mathcal{N}(0, \sigma^2)$ a reconstruction of $f$ using $A^\dagger_\Omega$ will result in an error estimated by $\| f - \hat{f} \| \leq \| A^\dagger_\Omega \| \sigma m^{-\frac{1}{2}}$. The theorem above provides an upper bound for the operator norm $\| A^\dagger_\Omega \|$. However, although the upper bound grows to infinity as $n$ or $m$ increases, it is not yet clear that $\| A^\dagger_\Omega \|$ deteriorates in this case. The following two results show that, indeed, as $m$ or $n$ increases $\| A^\dagger_\Omega \|$ is unbounded and the stability of reconstruction does in fact worsen.

**Theorem III.3.4.** Suppose $\hat{a}$, $n$, and $\Omega$ satisfy the hypotheses of Theorem III.2.2 with $|\Omega| = \frac{m-1}{2}$. Then $\| A^\dagger_\Omega \| \geq m \| \hat{a}'^{-1}(\frac{1}{4n}) \|.$

**Corollary III.3.5.** Suppose $\hat{a}$, $n$, and $\Omega$ satisfy the hypotheses of Theorem III.2.2 with $|\Omega| = \frac{m-1}{2}$. Then $\| A^\dagger_\Omega \| \to \infty$ as $n \to \infty$.

**Remark III.3.6.** The proof of the theorem shows that if $\Omega$ is some larger set, that is $|\Omega| > \frac{m-1}{2}$, then the growth of $\| A^\dagger \|$ may be alleviated. It should also be noted that in practice sampling on $\Omega$ will also likely to be performed at all times $n = 0, \ldots, m - 1$, rather than just when $n = 0$. This may also have the effect of decreasing $\| A^\dagger \|$.

In the proofs of the theorems and corollaries above, we use the following two lemmas, which are proved in chapter VI.

Before proving the theorem and corollaries above, we provide two well-known lemmas that we use in the proofs.
Lemma III.3.7. Let $A$ be an $m \times n$ matrix with $m > n$ so that the Moore-Penrose left inverse is given by $A^\dagger = (A^*A)^{-1}A^*$. If $A^\ell$ is any other left inverse of $A$, then $\|A^\dagger\| \leq \|A^\ell\|$.

Lemma III.3.8. Suppose $A$ is an $m \times n$ matrix with $m > n$, the maps $\sigma : \{1, \ldots, m\} \rightarrow \{1, \ldots, m\}$ and $\eta : \{1, \ldots, n\} \rightarrow \{1, \ldots, n\}$ are permutations, and $B$ is an $n \times m$ matrix such that $BA = I_n$. If the matrices $\tilde{A}$ and $\tilde{B}$ are given by $\tilde{A}(i,j) := A(\sigma(i), \eta(j))$ and $\tilde{B}(j,i) := B(\eta(j), \sigma(i))$, then $\tilde{B}\tilde{A} = I_n$ and $\|B\|_{op} = \|\tilde{B}\|_{op}$.

Proof of Theorem III.3.1

Similar to the matrix (III.14), the matrix obtained for the additional sampling on $\Omega = \{1, \ldots, m - 1\}$ is given by

$$A_\Omega(\tilde{\xi}) = \begin{pmatrix}
\frac{1}{mn} \tilde{u}_1(0) & \frac{1}{mn} \tilde{u}_1(1) & \cdots & \frac{1}{mn} \tilde{u}_1(n-1) \\
\vdots & \vdots & \ddots & \vdots \\
\frac{1}{mn} \tilde{u}_{m-1}(0) & \frac{1}{mn} \tilde{u}_{m-1}(1) & \cdots & \frac{1}{mn} \tilde{u}_{m-1}(n-1) \\
\frac{1}{m} \alpha'(\tilde{\xi}) & 0 & \cdots & 0 \\
0 & \frac{1}{m} \alpha'(\tilde{\xi} + \frac{1}{n}) & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \frac{1}{m} \alpha'(\tilde{\xi} + \frac{n-1}{n})
\end{pmatrix}.$$  \hspace{1cm} (III.29)

In light of Lemma III.1.2, a uniform upper bound for $\|A_\Omega^\dagger(\tilde{\xi})\|$, that is an upper bound independent of $\tilde{\xi}$, provides an upper bound for $\|A_\Omega^\dagger\|$. We choose $\Omega = \{0, 1, \ldots, m - 1\}$ and $n, m$ to be odd.

We will rearrange the rows and columns of the matrix $A_\Omega(\tilde{\xi})$ to create a matrix $\tilde{A}_\Omega(\tilde{\xi})$ for which we can explicitly give a left inverse. By Lemmas III.3.8 and III.3.7, it suffices to find an upper bound for any left inverse of $\tilde{A}_\Omega(\tilde{\xi})$.

For a fixed $\tilde{\xi} \in \left[0, \frac{1}{n}\right]$, let $k_0$ be such that $\tilde{\xi} + \frac{k_0}{n}$ is the closest point of $\{\tilde{\xi} + \frac{k}{n}\}_{k=0,\ldots,n-1}$ on the torus to a singularity of $\mathcal{A}$. Specifically, if $\tilde{\xi} \in \left[0, \frac{1}{4n}\right)$, then $k_0 = 0$; if $\tilde{\xi} \in \left[\frac{1}{4n}, \frac{3}{4n}\right)$, then $k_0 = \frac{n-1}{2}$; and if $\tilde{\xi} \in \left[\frac{3}{4n}, \frac{1}{n}\right]$, then $k_0 = n - 1$. We see that

$$\min_{k=0,\ldots,n-1, k \neq k_0} \left\{ \text{dist}(\tilde{\xi} + \frac{k}{n}, 0), \text{dist}(\tilde{\xi} + \frac{k}{n}, \frac{1}{2}), \text{dist}(\tilde{\xi} + \frac{k}{n}, 1) \right\} \geq \frac{1}{4n}. \hspace{1cm} (III.30)$$

In other words, for $k \neq k_0$, and $\tilde{\xi} \in \left[0, \frac{1}{n}\right]$, we have $\tilde{\xi} + \frac{k}{n} \in J$ where $J = J(n)$ is defined by

$$J = \left[\frac{1}{4n}, \frac{1}{2} - \frac{1}{4n}\right] \cup \left[\frac{1}{2} + \frac{1}{4n}, 1 - \frac{1}{4n}\right]. \hspace{1cm} (III.31)$$

By rearranging the columns and rows of the matrix $A_\Omega$ so that it has the form $\tilde{A}_\Omega$ below, we are able to explicitly define a left inverse that is independent of $\mathcal{A}(\tilde{\xi} + \frac{k_0}{n})$. 

---

31
We write

\[ \tilde{\Omega}(\xi) = \begin{pmatrix}
\frac{1}{mn} u_0(k_0) & \frac{1}{mn} u_0(k_1) & \ldots & \frac{1}{mn} u_0(k_{n-1}) \\
\vdots & \vdots & \ddots & \vdots \\
\frac{1}{mn} u_{m-1}(k_0) & \frac{1}{mn} u_{m-1}(k_1) & \ldots & \frac{1}{mn} u_{m-1}(k_{n-1}) \\
\frac{1}{m} \mathcal{A}(\xi + \frac{k_0}{n}) & 0 & \ldots & 0 \\
0 & \frac{1}{m} \mathcal{A}(\xi + \frac{k_1}{n}) & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & \frac{1}{m} \mathcal{A}(\xi + \frac{k_{n-1}}{n})
\end{pmatrix}. \quad (\text{III.32}) \]

in the block form

\[ \tilde{\Omega}(\xi) = \frac{1}{m} \begin{pmatrix}
mU_{k_0} & mQ \\
\mathcal{A}(\xi + \frac{k_0}{n}) & 0 \\
0 & D(\xi)
\end{pmatrix}, \quad (\text{III.33}) \]

where \( U_{k_0} = \begin{pmatrix}
\frac{1}{mn} u_0(k_0) \\
\frac{1}{mn} u_1(k_0) \\
\vdots \\
\frac{1}{mn} u_{m-1}(k_0)
\end{pmatrix} \), and \( D(\xi) \) is a \( m(n-1) \times m(n-1) \) block diagonal matrix with \( \mathcal{A}(\xi + \frac{k}{n}), k = 0, \ldots, n-1, k \neq k_0, \) on the main diagonal. Then, a left inverse is given by

\[ A^\ell_\Omega(\xi) = m \begin{pmatrix}
\frac{1}{m} U_{k_0}^{-1} & 0 & -U_{k_0}^{-1} Q D^{-1}(\xi) \\
0 & 0 & D^{-1}(\xi)
\end{pmatrix}, \quad (\text{III.34}) \]

and we easily compute that

\[ \|A^\ell_\Omega(\xi)\| \leq m(\max\{\|\frac{1}{m} U_{k_0}^{-1}\|, \|D^{-1}(\xi)\|\} + \|U_{k_0}^{-1}\| \|D^{-1}(\xi)\| \|Q\|). \quad (\text{III.35}) \]

Since \( D \) is a block diagonal matrix, we have

\[ \|D^{-1}(\xi)\| = \max_{k \neq k_0} \left\{ \|\mathcal{A}^{-1}(\xi + \frac{k}{n})\| \right\}. \quad (\text{III.36}) \]

The submatrix \( Q \) is an \( m \times m(n-1) \) matrix with entries of norm \( \frac{1}{mn} \). Thus, we have

\[ \|Q\| \leq \sqrt{n-1} \|Q\|_{\text{max}} = \frac{\sqrt{n-1}}{n}. \quad (\text{III.37}) \]

Observing that the columns of \( U_{k_0} \) are orthogonal, and we have \( \|U_{k_0}^{-1}\| = mn \). Our
estimate (III.35) becomes,
\[
\|A_\mathcal{A}(\xi)\| \leq m \max \left\{ n, \max_{k \neq k_0} \| A^{-1}(\xi + \frac{k}{n}) \| \right\} (1 + m\sqrt{n-1}). \tag{III.38}
\]

Taking the essential supremum over \( \xi \in [0, \frac{1}{n}] \), and noting that for \( k \neq k_0 \), \( \frac{k}{n} \in J \) as in (III.31), this last equation can be estimated by
\[
\|A_\mathcal{A}\| \leq \text{ess sup}_{\xi \in [0, \frac{1}{n}]} \left( m \max \left\{ n, \max_{k \neq k_0} \| A^{-1}(\xi + \frac{k}{n}) \| \right\} \right) (1 + m\sqrt{n-1}) 
\leq m \max \left\{ n, \text{ess sup}_{\eta \in J} \| A^{-1}(\eta) \| \right\} (1 + m\sqrt{n-1}). \tag{III.39}
\]

Since \( A(\eta) \) is invertible for all \( \eta \in J \) and \( J \) is a compact set, it follows that \( \text{ess sup}_{\eta \in J} \| A^{-1}(\eta) \| \) is finite, and Theorem III.3.1 follows.

To find the more explicit bound in Corollary III.3.2, we use the estimate for the norm of the inverse of a Vandermonde matrix [21]:
\[
\|A^{-1}(\xi)\| \leq \sqrt{m} \max_{0 \leq i \leq m-1} \sqrt{\frac{1}{2}} \prod_{j \neq i} \left| \frac{\hat{a}(\frac{\xi + j}{m}) - \hat{a}(\frac{\xi + i}{m})}{\hat{a}(\frac{\xi + i}{m}) - \hat{a}(\frac{\xi + j}{m})} \right|. \tag{III.40}
\]

To prove Corollary III.3.3, we find a uniform lower bound for \( |\hat{a}(\frac{\xi + j}{m}) - \hat{a}(\frac{\xi + i}{m})| \).

Note that when \( \xi \in J \), we have \( \frac{\xi + j}{m} \in [\frac{j}{m} + \frac{1}{m} J, j = 0, \ldots, m - 1 \). Then for \( \xi \in J \) and any \( j = 0, \ldots, m - 1 \), we have
\[
\frac{\xi + j}{m} \in \bigcup_{j=1}^{m-1} \left\{ \frac{j}{m} + \frac{1}{m} J \right\} \subset \left[ \frac{1}{4mn}, \frac{1}{2} - \frac{1}{4mn} \right] \cup \left[ \frac{1}{2} + \frac{1}{4mn}, \frac{1}{4mn} \right].
\]

Thus, defining \( M := [\frac{1}{4mn}, \frac{1}{2} - \frac{1}{4mn}] \), we have that \( \frac{\xi + j}{m} \in M \cup (M + \frac{1}{2}) \) for any \( j = 0, \ldots, m - 1 \).

Let \( \gamma = \min_{\xi \in M} |\hat{a}'(\xi)| > 0 \) where \( \hat{a}'(\xi) \) denotes the first derivative of \( \hat{a}(\xi) \). By the symmetry of \( \hat{a} \), we also have \( \gamma = \min_{\xi \in M + \frac{1}{2}} |\hat{a}'(\xi)| \). Without loss of generality, assume \( \frac{\xi + i}{m} > \frac{\xi + j}{m} \). If the interval \( [\frac{\xi + i}{m}, \frac{\xi + j}{m}] \) is contained in \( M \) or in \( M + \frac{1}{2} \), the Mean Value Theorem gives
\[
|\hat{a}(\frac{\xi + j}{m}) - \hat{a}(\frac{\xi + i}{m})| \geq \gamma \left| \frac{\xi + i}{m} - \frac{\xi + j}{m} \right| \geq \gamma \frac{1}{m}.
\]

If \( \frac{\xi + i}{m} \in M \) and \( \frac{\xi + j}{m} \in M + \frac{1}{2} \), we exploit the symmetry of \( \hat{a} \) and consider the interval between \( 1 - \frac{\xi + i}{m} \) and \( \frac{\xi + j}{m} \), which is contained in \( M \). Defining \( l = m - i - j \) and using
the Mean Value Theorem again, we have

$$\left| \hat{a}(1 - \frac{\xi + j}{m}) - \hat{a}(\frac{\xi + i}{m}) \right| \geq \gamma \left| \frac{l}{m} - 2(\frac{\xi}{m}) \right| = \frac{\gamma}{m} \left| \frac{l}{2} - \xi \right| \geq \frac{\gamma}{2mn},$$

where the last inequality follows from the fact that $l \in \mathbb{Z}$ and $\xi \in J$. This gives Corollary III.3.3. Notice that if $\hat{a}' \in C(\mathbb{T})$ then $\gamma \to 0$ as $n \to \infty$, due to the fact that the minimum is taken over a larger interval getting closer to the zeros of $\hat{a}'$.

**Proof of Theorem III.3.4**

Recall that $\|A^\dagger(\xi)\|$ is equal to the reciprocal of the smallest singular value of $A(\xi)$, denoted $s_{\text{min}}(A(\xi))$. We choose an extra sampling set $\Omega$ according to Theorem III.2.2. We claim that

**Claim 1**: There exists an interval $[0, r] \subset \left[0, \frac{1}{2n}\right]$, such that the smallest singular value of $s_{\text{min}}(A_\Omega(\xi))$ is bounded above on $[0, r]$, by

$$0 \leq s_{\text{min}}(A_\Omega(\xi)) \leq \frac{1}{m\|A^{-1}(\xi + \frac{1}{n})\|} < \infty, \quad \xi \in [0, r].$$

Using the claim, the theorem follows from

$$m\|A^{-1}(\xi + \frac{1}{n})\| \leq m \cdot \text{ess sup}_{\xi \in [0, r]} \|A^{-1}(\xi + \frac{1}{n})\| \leq \text{ess sup}_{\xi \in [0, \frac{1}{2n}]} \frac{1}{s_{\text{min}}(A_\Omega(\xi))} \leq \text{ess sup}_{\xi \in [0, \frac{1}{n}]} \frac{1}{s_{\text{min}}(A_\Omega(\xi))} = \|A^\dagger_\Omega\|.$$

**Proof of Claim 1.** We first show that $s_{\text{min}}^2(A_\Omega(\xi))$ is equal to the $mn$-th largest eigenvalue $\lambda_{mn}(A_\Omega(\xi)A^*_\Omega(\xi))$ of $A_\Omega(\xi)A^*_\Omega(\xi)$:

$$A_\Omega(\xi)A^*_\Omega(\xi) = \frac{1}{m^2} \left( \begin{array}{ccc} * & * & * \\ * & A(\xi)A^*(\xi) & 0 \\ * & 0 & D(\xi)D^*(\xi) \end{array} \right),$$

where the matrices in the first row have $|\Omega|$ rows and $D(\xi)D^*(\xi)$ is the block diagonal matrix with blocks $A_m(\xi + \frac{k}{n})A^*_m(\xi + \frac{k}{n}), k \neq 0$, as entries. The rank of the $(mn + |\Omega|) \times (mn + |\Omega|)$ matrix $A_\Omega(\xi)A^*_\Omega(\xi)$ is equal to the rank of $A_\Omega(\xi)$, which is $mn$. Thus, the smallest positive eigenvalue of $A_\Omega(\xi)A^*_\Omega(\xi)$ is the $mn$-th largest eigenvalue $\lambda_{mn}(A_\Omega(\xi)A^*_\Omega(\xi))$, and it is equal to $s_{\text{min}}^2(A_\Omega(\xi))$. Thus, to estimate $s_{\text{min}}^2(A_\Omega(\xi))$ from
above, we need to estimate $\lambda_{mn}(A_\Omega(\xi)A_\Omega^*(\xi))$.

In turn, the $mn$-th largest eigenvalue $\lambda_{mn}(A_\Omega(\xi)A_\Omega^*(\xi))$ can be estimated above using the eigenvalues of the $mn \times mn$ principal submatrix $B(\xi)$

$$B(\xi) = \begin{pmatrix} \mathcal{A}(\xi + \frac{k}{m})\mathcal{A}^*(\xi + \frac{k}{m}) & 0 \\ 0 & D(\xi)D^*(\xi) \end{pmatrix},$$

(III.43)

via the Cauchy Interlacing Theorem [18]:

$$s^2_{\min}(A_\Omega(\xi)) = \lambda_{mn}(A_\Omega(\xi)A_\Omega^*(\xi)) \leq \frac{1}{m^2} \lambda_{mn-|\Omega|}(B(\xi)), \quad \text{ (III.44)}$$

where we use $\lambda_j(M)$ to denotes the $j$-th largest eigenvalue of the a matrix $M$ counting the multiplicity.

We chose $\Omega$ to be a minimal extra sampling set so that $|\Omega| = \frac{m-1}{2}$. Observing that $B(\xi)$ is block diagonal so that the eigenvalues of $B(\xi)$ are the eigenvalues its $\mathcal{A}(\xi + \frac{k}{m})\mathcal{A}^*(\xi + \frac{k}{m})$, and using a continuity argument below we show that there exists $r$ with $0 < r < \frac{1}{4n}$ such that for all $\xi \in [0, r]$

$$\lambda_{mn-\frac{m-1}{2}}(B(\xi)) = \min_{k \neq 0} \left\{ \lambda_m \left( \mathcal{A}(\xi + \frac{k}{n})\mathcal{A}^*(\xi + \frac{k}{n}) \right) \right\} \quad \text{ (III.45)}$$

$$\leq \lambda_m \left( \mathcal{A}(\xi + \frac{1}{n})\mathcal{A}^*(\xi + \frac{1}{n}) \right)$$

$$= \frac{1}{\|\mathcal{A}^{-1}(\frac{1}{n})\|^2}.$$

In the last equality above, we used the relation between the minimum singular values of a matrix $M$ and the norm of its inverse: $s_{\min}^{-1}(M) = ||M^{-1}||$. Claim 1 then follows from (III.44) and (III.45).

We now use the continuity argument to prove (III.45). Let

$$\alpha := \inf_{|0, \frac{1}{4n}|} \lambda_{m(n-1)} \left( (D(\xi)D^*(\xi) \right) > 0.$$

Since $\lambda_{m-j}(\mathcal{A}(0)\mathcal{A}^*(0)) = 0$ for $j = 0, \ldots, (\frac{m-1}{2} - 1)$, continuity in $\xi$ implies that there exists $r$ with $0 < r < \frac{1}{4n}$ such that

$$\lambda_{m-j}(\mathcal{A}(\xi)\mathcal{A}^*(\xi)) < \alpha$$

for $\xi \in [0, r]$.

Thus, when $\xi \in [0, r]$, the smallest $\frac{m-1}{2}$ eigenvalues of $B(\xi)$ are precisely the smallest $\frac{m-1}{2}$ eigenvalues of $\mathcal{A}(\xi)\mathcal{A}^*(\xi)$, i.e., $\lambda_{mn-j}(B(\xi)) = \lambda_{mn-j}(\mathcal{A}(\xi)\mathcal{A}^*(\xi))$ for $j = 0, \ldots, (\frac{m-1}{2} - 1)$ and
\[
\lambda_{mn-\frac{m+1}{2}}(B(\xi)) = \min \left\{ \lambda_{m-\frac{m+1}{2}}(\mathcal{A}(\xi)\mathcal{A}^*(\xi)), \lambda_{m(n-1)}(D(\xi)D^*(\xi)) \right\} \\
\leq \lambda_{m(n-1)}(D(\xi)D^*(\xi)) \\
= \min_{k \neq 0} \left\{ \lambda_m(\mathcal{A}(\xi + \frac{k}{n})\mathcal{A}^*(\xi + \frac{k}{n})) \right\} \\
\leq \lambda_m\left(\mathcal{A}(\xi + \frac{1}{n})\mathcal{A}^*(\xi + \frac{1}{n})\right) \\
= \frac{1}{\|\mathcal{A}^{-1}(\xi + \frac{1}{n})\|^2},
\]

which is (III.45).

### III.4 Additional Sampling Sets for Generic Filters

In this section, we consider generic convolution operators \( A \) defined by \( Ax = a * x \). We remove the restrictions on the filter \( \hat{a} \) that are imposed in section III.2. The results parallel those in section III.2. We describe additional sampling sets that resolve the dynamical sampling problem when regular subsampling alone fails to allow for stable reconstruction, i.e. the conditions of proposition III.1.1 are not satisfied. We consider additional samples taken both at the initial time only and additional samples taken at each time.

The additional samples are taken in the following way. Let \( T_c \) be the operator that shifts a vector \( z \in \ell^2(\mathbb{Z}) \) to the right by \( c \) units so that \( T_c z(k) = z(k - c) \) so that \( S_{mn}T_c \) is the operator of shifting by \( c \) and then subsampling by a factor of \( mn \) for some \( n \in \mathbb{N} \). Below we describe an additional sample scheme that resolves the dynamical sampling problem in terms of the singularities of the family of matrices \( \mathcal{A}(\xi) \) given in (III.2) for \( \xi \in T \).

**Theorem III.4.1.** Let \( m \in \mathbb{Z}^+ \) be fixed. Suppose that \( \hat{a} \) is continuous and that \( \mathcal{A}(\xi) \) is singular only when \( \xi \in \{\xi_i\}_{i \in I} \). Let \( n \) be a positive integer such that \( |\xi_i - \xi_j| \neq \frac{k}{n} \) for any \( i, j \in I \) and \( k \in \{1, \ldots, n-1\} \). Then the extra samples given by \( \{(S_{mn}T_c)x\}_{c \in \{1, \ldots, m-1\}} \) provide enough additional information to stably recover any \( x \in \ell^2(\mathbb{Z}) \), i.e. the reconstruction operator is bounded.

**Proof.** The proof is very similar to the proof of theorem III.2.2 and we present here only the parts unique to this theorem. Using the techniques in the proof of theorem III.2.2, the problem is reduced to studying the left invertibility of a family of matrices over the torus. Here the matrices obtained for the additional sampling on
\( \Omega = \{1, \ldots, m-1\} \) are given by

\[
A_\Omega(\xi) = \begin{pmatrix}
\frac{1}{mn} \tilde{u}_1(0) & \frac{1}{mn} \tilde{u}_1(1) & \cdots & \frac{1}{mn} \tilde{u}_1(n-1) \\
\vdots & \vdots & \ddots & \vdots \\
\frac{1}{mn} \tilde{u}_{m-1}(0) & \frac{1}{mn} \tilde{u}_{m-1}(1) & \cdots & \frac{1}{mn} \tilde{u}_{m-1}(n-1) \\
\frac{1}{m} \mathcal{A}_m(\xi) & 0 & \cdots & 0 \\
0 & \frac{1}{m} \mathcal{A}_m(\xi + \frac{k}{n}) & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \frac{1}{m} \mathcal{A}_m(\xi + \frac{n-1}{n})
\end{pmatrix}, \quad (III.46)
\]

where \( \tilde{u}_c \) is defined in (III.11).

We will rearrange the rows and columns of the matrix \( A_\Omega(\xi) \) to create a matrix \( \tilde{A}_\Omega(\xi) \) for which we can explicitly give a left inverse. By Lemma III.3.8, it suffices to find an upper bound for any left inverse of \( \tilde{A}_\Omega(\xi) \).

The hypothesis on \( n \) guarantee that for any fixed \( \xi \), \( \mathcal{A}_m(\xi + \frac{k}{n}) \) is singular for at most one \( k = 0, \ldots, n-1 \). For a fixed \( \xi \in [0, \frac{1}{n}] \), let \( k_0 \) be such that \( \xi + \frac{k_0}{n} \) is a closest point of \( \{\xi + \frac{k}{n}\}_{k=0}^{n-1} \) on the torus to a singularity of \( \mathcal{A}_m \) so that for \( k \neq k_0 \), \( \mathcal{A}_m(\xi + \frac{k}{n}) \) is invertible.

By rearranging the columns and rows of the matrix \( A_\Omega \) so that it has the form \( \tilde{A}_\Omega \) below, we are able to explicitly define a left inverse that is independent of \( \mathcal{A}_m(\xi + \frac{k_0}{n}) \).

We write

\[
\tilde{A}_\Omega(\xi) = \begin{pmatrix}
\frac{1}{mn} \tilde{u}_0(0) & \frac{1}{mn} \tilde{u}_0(1) & \cdots & \frac{1}{mn} \tilde{u}_0(k-1) \\
\vdots & \vdots & \ddots & \vdots \\
\frac{1}{mn} \tilde{u}_{m-1}(0) & \frac{1}{mn} \tilde{u}_{m-1}(1) & \cdots & \frac{1}{mn} \tilde{u}_{m-1}(k-1) \\
\frac{1}{m} \mathcal{A}_m(\xi + \frac{k_0}{n}) & 0 & \cdots & 0 \\
0 & \frac{1}{m} \mathcal{A}_m(\xi + \frac{k_1}{n}) & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \frac{1}{m} \mathcal{A}_m(\xi + \frac{k_{n-1}}{n})
\end{pmatrix}. \quad (III.47)
\]

in the block form

\[
\tilde{A}_\Omega(\xi) = \frac{1}{m} mU_{k_0} mQ D(\xi), \quad (III.48)
\]

where \( U_{k_0} = \frac{1}{mn} \begin{pmatrix}
\tilde{u}_0(0) \\
\tilde{u}_1(0) \\
\vdots \\
\tilde{u}_{m-1}(0)
\end{pmatrix} \) is a scaled unitary matrix, and \( D(\xi) \) is a \( m(n-1) \times m(n-1) \) block diagonal matrix with \( \mathcal{A}_m(\xi + \frac{k}{n}), k = 0, \ldots, n-1, k \neq k_0 \), on the main
diagonal. Then, a left inverse of \( \tilde{A}_\Omega(\xi) \) is given by
\[
\tilde{A}_\Omega^\ell(\xi) = m \begin{pmatrix} \frac{1}{m} U_{k_0}^{-1} & 0 & -U_{k_0}^{-1}QD^{-1}(\xi) \\ 0 & 0 & D^{-1}(\xi) \end{pmatrix},
\]
(III.49)

\[\boxed{\}

Remark III.4.2. If the set \( I \) is finite, then the existence of \( n \) satisfying the conditions of theorem III.4.1 is guaranteed. Such an \( n \) can be chosen as follows. Let \( \{\mu_k\} \) be the pairwise distances between the singularities of \( \mathcal{A}(\xi) \), that is, \( \mu_k = |\xi_i - \xi_j| \) for some \( i, j \in I, i \neq j \). If \( \mu_k \) is rational, let \( \mu_k = \frac{p_k}{q_k} \) be irreducible. Then any \( n \) such that \( (n, \prod q_k) = 1 \), i.e. \( n \) is relatively prime to \( \prod q_k \), satisfies the hypotheses of theorem III.4.1. Furthermore, if \( \{q_{k_j}\} \) are the prime factors of \( q_k \), then any prime \( n \) such that \( n \geq \max_{k,j} q_{k_j} \) satisfies the hypotheses.

Theorem III.4.3. Let \( m, n \in \mathbb{Z}^+ \) be fixed and \( \hat{a} \) be continuous. Define \( s := \sup_\xi \{|k : \mathcal{A}(\xi + \frac{k}{n}) \text{ is singular}, k = 0, \ldots, n-1\}|. \) Then the extra samples given by \( \{(S_{mn}T_c)x\}_{c \in \Omega} \) for \( \Omega = \{1, \ldots, m-1, m+1, \ldots 2m-1, \ldots (s-1)m+1, \ldots sm-1\} \) provide enough additional information to stably recover any \( x \in \ell^2(\mathbb{Z}) \), i.e. the reconstruction operator is bounded.

Proof. The proof is very similar to the proof of theorem III.2.2 and we present here only the parts unique to this theorem. Using the techniques in the proof of theorem III.2.2, the problem is reduced to studying the left invertibility of a family of matrices over the torus. Here the matrices obtained for the additional sampling on \( \Omega \) are given by
\[
A_\Omega(\xi) = \begin{pmatrix} U_{0,0} & U_{1,0} & \cdots & U_{n-1,0} \\ \vdots & \vdots & \ddots & \vdots \\ U_{0,s-1} & U_{1,s-1} & \cdots & U_{n-1,s-1} \\ \frac{1}{m} \mathcal{A}_m(\xi) & 0 & \cdots & 0 \\ 0 & \frac{1}{m} \mathcal{A}_m(\xi + \frac{1}{n}) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{m} \mathcal{A}_m(\xi + \frac{n-1}{n}) \end{pmatrix},
\]
(III.50)

where \( U_{k,v} = \frac{1}{mn} \begin{pmatrix} \hat{u}_{vm+1}(k) \\ \hat{u}_{vm+2}(k) \\ \vdots \\ \hat{u}_{vm+m-1}(k) \end{pmatrix} \) and \( \hat{u}_c \) is defined in (III.11). We will show that \( A_\Omega(\xi) \) is left invertible for every \( \xi \). Fix \( \xi \) and let \( k_0, k_1, \ldots, k_{n-1} \) be a reordering of
0, \ldots n - 1 \text{ such that } A(\xi + \frac{k}{n}) \text{ is nonsingular for } k \neq k_0, k_1, \ldots, k_{s-1}. \text{ By reordering the rows and columns, we have }

\begin{equation}
\tilde{A}_\Omega(\xi) = \begin{pmatrix}
U_{k_0,0} & U_{k_1,0} & \cdots & U_{k_{n-1},0} \\
\vdots & \vdots & \ddots & \vdots \\
U_{k_0,s-1} & U_{k_1,s-1} & \cdots & U_{k_{n-1},s-1} \\
\frac{1}{m} A_m(\xi + \frac{k_0}{n}) & 0 & \cdots & 0 \\
0 & \frac{1}{m} A_m(\xi + \frac{k_1}{n}) & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \frac{1}{m} A_m(\xi + \frac{k_{n-1}}{n}) 
\end{pmatrix}.
\end{equation} (III.51)

In block form

\begin{equation}
\tilde{A}_\Omega(\xi) = \frac{1}{m} \begin{pmatrix}
mU & mQ \\
D'(\xi) & 0 \\
0 & D(\xi)
\end{pmatrix},
\end{equation} (III.52)

where

\begin{equation}
U = \begin{pmatrix}
U_{k_0,0} & U_{k_1,0} & \cdots & U_{k_{n-1},0} \\
\vdots & \vdots & \ddots & \vdots \\
U_{k_0,s-1} & U_{k_1,s-1} & \cdots & U_{k_{n-1},s-1}
\end{pmatrix},
\end{equation} (III.53)

and $D'(\xi)$ is a square block diagonal matrix with $A_m(\xi + \frac{k}{n}), k = k_0, \ldots, k_{s-1}$ on the main diagonal, and $D(\xi)$ is a square block diagonal matrix with $A_m(\xi + \frac{k_i}{n}), k = k_s, \ldots, k_{n-1}$ on the main diagonal. Clearly, $D(\xi)$ is invertible and we show below that $U$ is invertible. A left inverse of $\tilde{A}_\Omega(\xi)$ is then given by

\begin{equation}
A_\Omega^\ell(\xi) = m \begin{pmatrix}
\frac{1}{m} U^{-1} & 0 & -U_{k_0}^{-1} QD^{-1}(\xi) \\
0 & 0 & D^{-1}(\xi)
\end{pmatrix},
\end{equation} (III.54)

We will show that the $(s-1)m \times (s-1)m$ matrix

\begin{equation}
U = \begin{pmatrix}
U_{k_0,0} & U_{k_1,0} & \cdots & U_{k_{n-1},0} \\
\vdots & \vdots & \ddots & \vdots \\
U_{k_0,s-1} & U_{k_1,s-1} & \cdots & U_{k_{n-1},s-1}
\end{pmatrix},
\end{equation} (III.55)

has full rank. Using equation (III.16), we see that the rows $c, d$ of $U$ are orthogonal if $c \neq d \mod m$:

\begin{equation}
\langle U(c, \cdot), U(d, \cdot) \rangle = \begin{cases}
\sum_{j=0}^{s-1} m e^{-\frac{2 \pi i (c-d) k_j}{m}}, & (c - d) = 0 \mod m \\
0, & \text{otherwise}
\end{cases},
\end{equation} (III.56)
It remains to show that sets of rows of the form

\[ \{U(c,.)U(c+m,.)\ldots U(c+(s-1)m,.)\} \]

are linearly independent. The rows form the matrix

\[
\begin{pmatrix}
\tilde{u}_c(k_0) & \tilde{u}_c(k_1) & \ldots & \tilde{u}_c(k_{s-1}) \\
\vdots & \vdots & \ddots & \vdots \\
\tilde{u}_{c+(s-1)m}(k_0) & \tilde{u}_{c+(s-1)m}(k_1) & \ldots & \tilde{u}_{c+(s-1)m}(k_{s-1})
\end{pmatrix}, \quad (III.57)
\]

whose 1-st, (m+1)-th, \ldots, ((s-1)m+1)-th columns form the matrix,

\[
\begin{pmatrix}
e^{-\frac{2\pi i k_0}{m}} & e^{-\frac{2\pi i k_1}{m}} & \ldots & e^{-\frac{2\pi i k_{s-1}}{m}} \\
e^{-\frac{2\pi i (c+m)k_0}{m}} & e^{-\frac{2\pi i (c+m)k_1}{m}} & \ldots & e^{-\frac{2\pi i (c+m)k_{s-1}}{m}} \\
& \vdots & \ddots & \vdots \\
e^{-\frac{2\pi i (c+(s-1)m)k_0}{m}} & e^{-\frac{2\pi i (c+(s-1)m)k_1}{m}} & \ldots & e^{-\frac{2\pi i (c+(s-1)m)k_{s-1}}{m}}
\end{pmatrix}.
\]

Multiplying the \( j \)-th column by \( e^{\frac{2\pi i c j}{m}} \) results in the following Vandermonde matrix

\[
\begin{pmatrix}
1 & e^{-\frac{2\pi i k_0}{n}} & e^{-\frac{2\pi i k_1}{n}} & \ldots & e^{-\frac{2\pi i k_{s-1}}{n}} \\
e^{-\frac{2\pi i (c+m)k_0}{n}} & e^{-\frac{2\pi i (c+m)k_1}{n}} & \ldots & e^{-\frac{2\pi i (c+m)k_{s-1}}{n}} \\
& \vdots & \ddots & \vdots \\
e^{-\frac{2\pi i (c+(s-1)m)k_0}{n}} & e^{-\frac{2\pi i (c+(s-1)m)k_1}{n}} & \ldots & e^{-\frac{2\pi i (c+(s-1)m)k_{s-1}}{n}}
\end{pmatrix},
\]

which is invertible since \( k_i \neq k_j \) for \( i \neq j \). Thus, the matrix in III.57 has rank \( s \), i.e., full row rank.

**Theorem III.4.** Let \( p = \max_{\xi \in \mathcal{A}} \{ \text{max number of columns of } \mathcal{A}(\xi) \text{ that coincide} \} \).

And suppose \( n \) satisfies the conditions of theorem III.4.1. Let \( \Omega \subset \mathbb{Z} \) contain elements \( c_1, \ldots, c_p \) such that \( c_1 = 1 \mod m, c_2 = 2 \mod m, \ldots, c_p = p \mod m \). Then the extra samples given by \( \{ S_{m \ell}T_{c}(a^j+x) \}_{c \in \Omega, j=0,\ldots,m-1} \) provide enough additional information to stably recover any \( x \in l^2(\mathbb{Z}) \), i.e., the reconstruction operator is bounded.

**Proof.** We consider the minimal case in which \( \Omega = \{ c_1, \ldots, c_p \} \). The proof is very similar to but more complicated than the proof of theorem III.2.4. Using the techniques in the proof of theorem III.2.4, the problem is reduced to studying the left invertibility of a family of matrices over the torus. The original and additional samples are related in the Fourier domain by the matrix equation below:
at most one of the matrices $A$ for generality, we assume $A_{\xi}$ and $\bar{A}_{\xi}$ where 

$$
\gamma_{\xi} = \begin{pmatrix}
ne^{-\frac{2\pi i}{m}}(S_{mn}T_{c_1}x)^{(n\xi)} \\
ne^{-\frac{2\pi i}{m}}(S_{mn}T_{c_1}(a^m*x))^{(n\xi)} \\
\vdots \\
ne^{-\frac{2\pi i}{m}}(S_{mn}T_{c_1}(a^{m-1}*x))^{(n\xi)} \\
\end{pmatrix} = \mathcal{A}(\xi)
$$

where $\mathcal{A}(\xi)$ and $\bar{\mathcal{A}}(\xi)$ are given in (III.8) and $\mathcal{A}(\xi)$ is the block matrix

$$
\mathcal{A}(\xi) = \begin{pmatrix}
\mathcal{A}^{c,0}(\xi) & \mathcal{A}^{c,1}(\xi + \frac{1}{n}) & \cdots & \mathcal{A}^{c,n-1}(\xi + \frac{n-1}{n}) \\
\vdots & \vdots & \ddots & \vdots \\
\mathcal{A}^{\alpha,0}(\xi) & \mathcal{A}^{\alpha,1}(\xi + \frac{1}{n}) & \cdots & \mathcal{A}^{\alpha,n-1}(\xi + \frac{n-1}{n}) \\
\mathcal{A}(\xi) & 0 & \cdots & 0 \\
0 & \mathcal{A}(\xi + \frac{1}{n}) & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \mathcal{A}(\xi + \frac{n-1}{n})
\end{pmatrix}, \text{ where (III.59)}
$$

$$
\mathcal{A}^{c,k}(\xi) = e^{-\frac{2\pi i}{m}}\begin{pmatrix}
1 & e^{-\frac{2\pi i}{m}} & \cdots & e^{-\frac{2\pi i(m-1)}{m}} \\
\hat{a}(\frac{\xi}{m}) & e^{-\frac{2\pi i}{m}}\hat{a}(\frac{\xi+1}{m}) & \cdots & e^{-\frac{2\pi i(m-1)}{m}}\hat{a}(\frac{\xi+m-1}{m}) \\
\vdots & \vdots & \ddots & \vdots \\
\hat{a}^{[m-1]}(\frac{\xi}{m}) & e^{-\frac{2\pi i}{m}}\hat{a}^{[m-1]}(\frac{\xi+1}{m}) & \cdots & e^{-\frac{2\pi i(m-1)}{m}}\hat{a}^{[m-1]}(\frac{\xi+m-1}{m})
\end{pmatrix}.
$$

The remainder of the proof is studying the left invertibility of the matrices $\mathcal{A}(\xi)$ for $\xi \in \mathbb{T}/n$. Because $n$ satisfies the hypotheses of theorem III.4.1, for any fixed $\xi$, at most one of the matrices $\mathcal{A}(\xi + \frac{k}{n})$, $k = 0, \ldots, n-1$ is singular. Without loss of generality, we assume $\mathcal{A}(\xi + \frac{k}{n})$ is nonsingular for $k = 1, \ldots, n-1$. 

41
In block form,

\[
\tilde{A}(\xi) = \begin{pmatrix} D(\xi) & F(\xi) \\ 0 & G(\xi) \end{pmatrix}
\]  

(III.60)

where

\[
D(\xi) = \begin{pmatrix} A_{c1}(\xi) & \ldots & A_{cp}(\xi) \end{pmatrix},
\]

\[
F(\xi) = \begin{pmatrix} A_{c1}(\xi + \frac{1}{n}) & \ldots & A_{c1}(\xi + \frac{n-1}{n}) \\ \vdots & \ddots & \vdots \\ A_{cp}(\xi + \frac{1}{n}) & \ldots & A_{cp}(\xi + \frac{n-1}{n}) \end{pmatrix},
\]

and

\[
G(\xi) = \text{block diagonal matrix with } A(\xi + \frac{1}{n}), \ldots, A(\xi + \frac{n-1}{n}) \text{ on the diagonal.}
\]

If \(D^l(\xi)\) and \(G^l(\xi)\) are left inverses of \(D(\xi)\) and \(G(\xi)\), respectively, then a left inverse of \(\tilde{A}(\xi)\) is given by

\[
\tilde{A}^l(\xi) = \begin{pmatrix} D^l(\xi) - D^l(\xi)F(\xi)G^l(\xi) \\ 0 & G^l(\xi) \end{pmatrix}
\]  

(III.61)

It is clear that \(G(\xi)\) has a left inverse, as it is the block diagonal matrix with the left inverses of the matrices on the diagonal of \(G(\xi)\). The fact that \(D(\xi)\) has a left inverse is non-trivial and relies on the Vandermonde structure of the matrices, which allows us to find explicit bases for \(\ker A(\xi)\).

First, note that \(D(\xi)\) has full rank, and thus a left inverse, if and only if

\[
\bigcap_{j=0}^{p} \ker A_{c_j}(\xi) = \{0\},
\]

where \(c_0 = 0\) so that \(A_{c0}(\xi) = A(\xi)\). We can relate \(\ker A(\xi)\) to \(\ker A_{c_j}(\xi)\) by using equation (III.26) and defining the maps \(M_{c_j} : \ker A(\xi) \rightarrow \ker A_{c_j}(\xi)\) by

\[
M_{c_j}z = \begin{pmatrix} z_0 \\ e^{\frac{2\pi i}{n}c_jz_1} \\ \vdots \\ e^{\frac{2\pi i}{n}(m-1)c_jz_{m-1}} \end{pmatrix}
\]

We can now study \(\ker A_{c_j}(\xi)\) by looking at the image of \(\ker A(\xi)\) under the map \(M_{c_j}\). This is summarized in the following claim.

**Claim 1:** The matrix \(D(\xi)\) has a left inverse if and only if

\[
\bigcap_{j=0}^{p} M_{c_j}(\ker A(\xi)) = \{0\}.
\]

In order to use the above claim, we find a useful characterization of \(\ker A(\xi)\). Let \(q\) be the number of distinct columns of \(A(\xi)\) so that \(q\) is the rank of \(A(\xi)\). We can split the columns of \(A(\xi)\) into \(q\) sets so that the columns in each set are coinciding. Let \(p_1, \ldots, p_q\) be the number of columns in each set of coinciding columns so that
\( m = p_1 + \cdots + p_q \). We will write \( \ker A(\xi) \) as a span of subspaces, where each subspace is determined by a set of coinciding columns.

For each set of coinciding columns, we define a basis for a subspace of \( \ker A(\xi) \) as follows. For \( 1 \leq k \leq q \) such that \( p_k > 1 \), let \( l_{k,1}, \ldots, l_{k,p_k} \) be the indices of the columns in one of the sets of coinciding columns. Define the vector \( v_{k,1} \) as the vector with \( v_{k,1}(l_{k,1}) = 1, v_{k,1}(l_{k,2}) = -1 \) and zeroes elsewhere. Similarly, define the vector \( v_{k,2} \) as the vector with \( v_{k,2}(l_{k,1}) = 1, v_{k,2}(l_{k,3}) = -1 \) and zeroes elsewhere. Continue in this way to create the vectors \( v_{k,1}, \ldots, v_{k,p_k-1} \).

Let \( V_k = \text{span}\{v_{k,1}, \ldots, v_{k,p_k-1}\} \).

Due to the construction of the vectors above, it is obvious that \( V_k \subseteq \ker A(\xi) \). It is clear from the linear independence of the spanning vectors that \( \text{rank} V_k = p_k - 1 \). If \( p_k = 1 \), then define \( V_k = \{0\} \).

It follows that \( \text{span}\{V_1, \ldots, V_q\} \subseteq \ker A(\xi) \). Observe that \( V_j \perp V_k \) for \( j \neq k \). This follows easily by looking at the positions of non-zero entries in the bases for \( V_j \) and \( V_k \).

Then

\[
\text{rank}(\text{span}\{V_1, \ldots, V_q\}) = \text{rank} V_1 + \text{rank} V_2 + \cdots + \text{rank} V_q
\]
\[
= p_1 + \cdots + p_q - q
\]
\[
= m - q
\]
\[
= \text{rank}(\ker A(\xi))
\]

Thus, \( \ker A(\xi) = \text{span}\{V_1, \ldots, V_q\} \).

We will use this characterization of \( \ker A(\xi) \) to study the co-dimension of the intersection in claim 1. Specifically, the conditions of claim 1 are satisfied if and only if

\[
\dim \left( \bigcap_{j=0}^{p} M_{\xi_j} (\ker A(\xi)) \right) \perp = m.
\]

Combining this fact with the characterization of \( \ker A(\xi) \) above, we have proved the following claim.

**Claim 2:** The matrix \( D(\xi) \) has a left inverse if and only if

\[
\dim \left( \bigcap_{j=0}^{p} \text{span}\{M_{\xi_j}V_1, \ldots, M_{\xi_j}V_q\} \right) \perp = m.
\]

The remainder of the proof is dedicated to showing that the conditions of claim 2 above are indeed satisfied. We use the following Lemma which is proved in chapter VI.
Lemma III.4.5. Let \( \{W_i\} \) be a finite family of subspaces in \( \mathbb{C}^d \). Then

\[
\left( \bigcap_{i} W_i \right)^\perp = \text{span}\{W_i^\perp\}. \tag{III.62}
\]

Using the lemma above, we are able to reduce the problem to studying the span of orthogonal compliments:

\[
\left( \bigcap_{j=0}^{p} \text{span} \{ M_j V_1, \ldots, M_j V_q \} \right)^\perp = \text{span}_{j=0}^{p} \left\{ \left( \text{span} \{ M_j V_1, \ldots, M_j V_q \} \right)^\perp \right\}
\]

For each \( j \), we explicitly find a basis for \( \left( \text{span} \{ M_j V_1, \ldots, M_j V_q \} \right)^\perp \) by finding normal vectors \( w_{j,k}, k = 1, \ldots, q \), for the space \( \text{span} \{ M_j V_1, \ldots, M_j V_q \} \). We will then study the dimension of the space spanned by all the normal vectors.

For each \( 0 \leq j \leq p \) and each \( 0 \leq k \leq q \) such that \( V_k \) is non-trivial, define \( w_{j,k} \) to be the vector with

\[
w_{j,k}(l_{k,1}) = e_{\frac{i2\pi l_{k,1}e_j}{m}} \equiv e_{\frac{i2\pi l_{k,1}e_j}{m}} \\
\vdots \\
w_{j,k}(l_{k,p_k}) = e_{\frac{i2\pi l_{k,p_k}e_j}{m}}
\]

and zeros elsewhere. Note that \( w_{j,k} \) has exactly \( p_k \) non-zero entries. We claim that \( w_{j,k} \perp M_j V_k \). Indeed, for \( 1 \leq n \leq p_k \),

\[
\langle w_{j,k}, M_j V_{k,n} \rangle = e^{-\frac{i2\pi l_{k,1}e_j}{m}} - e^{-\frac{i2\pi l_{k,1}e_j}{m}} = 0
\]

It follows that \( w_{j,k} \perp V_n, n \neq k \) by observing that there is no overlap in the positions of the non-zero entries of \( w_{j,k} \) and the non-zero entries in each of the basis vectors for \( V_n, n \neq k \). We have shown \( w_{j,k} \in \left( \text{span} \{ M_j V_1, \ldots, M_j V_q \} \right)^\perp \).

For each \( 0 \leq k \leq q \) such that \( V_k = 0 \), we define \( w_{j,k} \) to be the vector with zeros everywhere except at a location \( l_{k,1} \) such that \( l_{k,1} \neq l_{k',n} \) for any \( k \neq k' \) and \( 1 \leq n \leq p_{k'} \). The fact that \( m = p_1 + \cdots + p_q \) guarantees the existence of \( l_{k,1} \). Define \( w_{j,k} = e_{l_{k,1}} \), where \( \{e_n\} \) is the standard basis for \( \mathbb{C}^m \). We again have \( w_{j,k} \in \left( \text{span} \{ M_j V_1, \ldots, M_j V_q \} \right)^\perp \).

We claim that \( \text{span} \{ w_{j,1}, \ldots, w_{j,q} \} = \left( \text{span} \{ M_j V_1, \ldots, M_j V_q \} \right)^\perp \). By the above arguments, it is clear that \( \text{span} \{ w_{j,1}, \ldots, w_{j,q} \} \subseteq \left( \text{span} \{ M_j V_1, \ldots, M_j V_q \} \right)^\perp \). By considering the respective position of non-zero entries in \( w_{j,k} \) and \( w_{j,k'} \), we see that \( w_{j,k} \perp w_{j,k'} \) for \( k \neq k' \). Thus the dimension of \( \text{span} \{ w_{j,1}, \ldots, w_{j,q} \} \) is \( q \), which is equal to the dimension of \( \left( \text{span} \{ M_j V_1, \ldots, M_j V_q \} \right)^\perp \).
To finish the proof, we will show that
\[ \dim \text{span} \{ w_{1,1}, \ldots, w_{1,q}, \ldots, w_{p,1}, \ldots, w_{p,q} \} = m. \]

Define \( W_k = \text{span} \{ w_{1,k}, \ldots, w_{p,k} \} \) and observe that \( W_k \perp W_{k'} \) for \( k \neq k' \). The proof will be complete if we show \( \dim W_k = p_k \) for \( k = 1, \ldots, q \).

The \( \dim W_k \) is equal to the rank of the \( p_k \times p_k \) matrix below formed by deleting columns of zeros from the matrix of row vectors \( w_{j,k} \):

\[
\begin{pmatrix}
e^{\frac{2\pi i k_1}{m}} & e^{\frac{2\pi i k_2}{m}} & \cdots & e^{\frac{2\pi i k_{p_k}}{m}} \\
e^{\frac{2\pi i k_1}{m}} & e^{\frac{2\pi i k_2}{m}} & \cdots & e^{\frac{2\pi i k_{p_k}}{m}} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{2\pi i k_{p_k}}{m} & \frac{2\pi i k_{p_k}}{m} & \cdots & \frac{2\pi i k_{p_k}}{m}
\end{pmatrix}
\]

Applying the hypothesis that \( c_j = j \mod m \) to the above matrix, yields the following Vandermonde matrix

\[
\begin{pmatrix}
1 & 1 & \cdots & 1 \\
e^{\frac{2\pi i k_1}{m}} & e^{\frac{2\pi i k_2}{m}} & \cdots & e^{\frac{2\pi i k_{p_k}}{m}} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{2\pi i k_{p_k}}{m} & \frac{2\pi i k_{p_k}}{m} & \cdots & e^{\frac{2\pi i k_{p_k}}{m}}
\end{pmatrix},
\]

which has full rank. Thus, \( \dim W_k = p_k \), and the theorem is proved.
CHAPTER IV

DYNAMICAL SAMPLING IN SHIFT-INvariant SPACES

In shift-invariant spaces (SIS), analog functions can be represented by discrete sequences. By studying this connection, we hope to reduce the problem of dynamical sampling in shift-invariant spaces to that of dynamical sampling on $\ell^2(\mathbb{Z})$. Although all separable Hilbert spaces are isometrically isomorphic, the connection between dynamical sampling in shift-invariant spaces and that in $\ell^2(\mathbb{Z})$ is nontrivial. In fact, only under certain conditions, does the dynamical sampling problem in SIS reduce to the dynamical sampling problem in $\ell^2(\mathbb{Z})$.

This chapter begins with a review of SIS. Then the dynamical sampling problem in SIS is formulated and conditions for stable reconstruction in the most general case are given. Next we give conditions under which the problem reduces to the $\ell^2(\mathbb{Z})$ case. We further provide a classification of systems which do not completely reduce to the $\ell^2(\mathbb{Z})$ case, but nonetheless make use of the results in the $\ell^2(\mathbb{Z})$ case.

IV.1 Background on Shift-Invariant Spaces

Shift-invariant spaces (SIS) are the typical spaces of functions considered in sampling theory [4, 5, 29, 35, 40, 41, 44]. Specifically, a shift invariant space $V$ has the form:

$$
V(\phi) = \{ \sum_{k \in \mathbb{Z}} c_k \phi(\cdot - k) : (c_k)_{k \in \mathbb{Z}} \in \ell^2(\mathbb{Z}) \}.
$$

The sum in (IV.1) can be viewed as the semi-discrete convolution between a sequence $c \in \ell^2(\mathbb{Z})$ and a generator $\phi \in W^1_0$ (see definition of $W^1_0$ below). In this paper, we use the notation $c *_{sd} \phi := \sum_{k \in \mathbb{Z}} c_k \phi(\cdot - k)$ to describe this semi-discrete convolution.

Certain assumptions must be imposed on the function $\phi$ in order for the sampling to make sense and for the space $V(\phi)$ to be well defined. Typically, the function $\phi$ is assumed to be continuous, to have sufficient decay, and to form a Riesz basis for $\text{span}\{\phi(\cdot - k) : k \in \mathbb{Z}\}$. The Riesz basis condition in Fourier domain states that there exist $m, M > 0$ such that

$$
m \leq \sum_{j \in \mathbb{Z}} |\hat{\phi}(\xi + j)|^2 \leq M \quad a.e. \, \xi
$$

where $\hat{\phi}(\xi) = \int_{\mathbb{R}} \phi(t)e^{-2\pi i \xi t}dt$ is the Fourier transform of $\phi$ (see e.g., [4]).

The local behavior and global decay of $\phi$ can be described in terms of the Wiener amalgam spaces [4, 19]. A measurable function $f$ belongs to the Wiener amalgam
space $W^p, 1 \leq p < \infty$, if it satisfies

$$\|f\|_{W^p}^p := \sum_{k \in \mathbb{Z}} \text{esssup}\{|f(x+k)|^p; x \in [0, 1]\} < \infty.$$  \hspace{1cm} (IV.3)

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

$$\|f\|_{W(L^\infty)} := \sup_{k \in \mathbb{Z}} \{\text{esssup}\{|f(x+k)|; x \in [0, 1]\}\} < \infty.$$  \hspace{1cm} (IV.4)

Because ideal sampling makes sense only for continuous functions, we work in the amalgam spaces $W^p_0 := W(L^p(\mathbb{R})) \cap C(\mathbb{R})$.

Now, if $\phi \in W^1_0$ and satisfies (IV.2), then $V(\phi)$ in (IV.1) is a subspace of $W^2_0$. Thus, under these conditions on $\phi$, any function $f \in V(\phi)$ is continuous and can be sampled at any $x \in \mathbb{R}$. Moreover, there exists $C > 0$ such that

$$\sum_{k \in \mathbb{Z}} |f(k)|^2 \leq C\|f\|^2_2, \quad \forall f \in V(\phi).$$

It is known that if $\sum_j \hat{\phi}(\xi + j) \neq 0$, then any function $f \in V(\phi)$ can be recovered from its samples on $\mathbb{Z}$, i.e., from $f(\mathbb{Z})$ [5, 43].

There are other conditions on $\phi$ that result in a useful sampling theory. One such condition is that the support of $\hat{\phi}$ is compact. In this case, $V(\phi)$ is a space of entire functions. For example, when $\phi = \text{sinc}$, where $\text{sinc}(x) = \frac{\sin(\pi x)}{\pi x}$, clearly $\phi \notin W^1_0$. However, the $\text{sinc}$ function belongs to $W_0(L^2)$, and the generated shift-invariant space can also be described as

$$V(\text{sinc}) = PW_{1/2} = \{f \in L^2 : \hat{f}(\xi) = 0, \quad \forall \xi \notin [-1/2, 1/2]\}. \hspace{1cm} (IV.5)$$

This space is sometimes called the Paley-Wiener space, or the space of 1/2-bandlimited functions. Clearly, ideal sampling of functions in $V(\text{sinc})$ is well-defined.

### IV.2 Problem Formulation

In this section, we explore the problem of dynamical sampling in a shift-invariant space in the general setting.

The dynamical sampling problem in shift-invariant spaces is to reconstruct the function $f \in V(\phi)$ from the coarse samplings $\{g_0 = S(\Omega_0)f, \ g_n = S_m A^{n-1} f, \ n = 1, \ldots, m\}$, where $\Omega_0$ is a "small" and possibly empty set. When the operator $A$ is a spatial convolution, i.e. $Af = a * f$, then even if $f$ is in $V(\phi)$, $f_1 = a * f$ is not necessarily in $V(\phi)$. For this reason, the dynamical sampling problem in SIS is not reducible to the one in $\ell^2(\mathbb{Z})$ in general. In fact, if $a \in W^1_0$ and $f \in V(\phi)$, then $a * f \in \ell^2(\mathbb{Z})$.
\(V(a * \phi). \) Moreover, \(a * \phi \in W^1_0\) whenever \(a \in W(L^1)\) and \(\phi \in W^1_0\) [4, 19].

Let \(f \in V(\phi).\) Then the series of samples of \(f\) on \(\mathbb{Z}, \quad f(l) = \sum_k c_k \phi(l - k),\) is the discrete convolution

\[
f|_\mathbb{Z} = c * d,
\]

where \(d = (\phi(k))_k.\) On the Fourier side, we then have \(\mathcal{F}(f|_\mathbb{Z}) = \hat{c} \cdot \hat{d}\) where \(\mathcal{F}\) is the Fourier transform operator and for any \(g \in L^2(\mathbb{Z}), \quad \hat{g}(\xi) = \sum_k g(k)e^{-2\pi ik\xi}.\)

Let \(A f := a * f.\) If \(a^j = a * a * ... * a\) and \(\phi_j = a^j * \phi\) for \(j = 1, ..., m - 1,\) we have

\[
A^j f = a^j * f = \sum_k c_k (a^j * \phi)(-k) = \sum_k c_k \phi_j(-k) \in V(\phi).
\]

Let \(V \in W^1_0\) and \(\phi \in W^1_0\) then \(\phi_j \in W^1_0\) and the samples of \(A^j f\) are well defined. The following lemma, whose proof is postponed until the end of this section, is useful for tackling the dynamical sampling problem in SIS.

**Lemma IV.2.1.** Let \(V(\phi)\) be a SIS and \(f = \sum_k c_k \phi(-k) \in V(\phi).\) Let

\[
\phi_j = a^j * \phi, \quad f_j = a^j * f, \quad h_j = f|_\mathbb{Z} \quad \text{and} \quad \Phi_j = \phi|_\mathbb{Z}
\]

for \(j = 0, 1, ..., m - 1.\) Then

\[
\mathcal{F}(S_m h_j)(\xi) = \frac{1}{m} \sum_{l=0}^{m-1} \hat{c}(\frac{\xi + l}{m}) \hat{\Phi}_j\left(\frac{\xi + l}{m}\right).
\]

Now letting \((y_j)(k) := (S_m(a^j * f))(k) = S_m h_l(k), k \in \mathbb{Z}, l = 0, ..., m - 1,\) using Lemma IV.2.1 we get

\[
\begin{pmatrix}
\hat{y}_0(\xi) \\
\hat{y}_1(\xi) \\
\vdots \\
\hat{y}_{m-1}(\xi)
\end{pmatrix} =
\begin{pmatrix}
\Phi_0(\frac{\xi}{m}) & \Phi_0(\frac{\xi+1}{m}) & \cdots & \Phi_0(\frac{\xi+m-1}{m}) \\
\Phi_1(\frac{\xi}{m}) & \Phi_1(\frac{\xi+1}{m}) & \cdots & \Phi_1(\frac{\xi+m-1}{m}) \\
\vdots & \vdots & \ddots & \vdots \\
\Phi_{m-1}(\frac{\xi}{m}) & \Phi_{m-1}(\frac{\xi+1}{m}) & \cdots & \Phi_{m-1}(\frac{\xi+m-1}{m})
\end{pmatrix}
\begin{pmatrix}
\hat{c}(\frac{\xi}{m}) \\
\hat{c}(\frac{\xi+1}{m}) \\
\vdots \\
\hat{c}(\frac{\xi+m-1}{m})
\end{pmatrix}
\]

In short notation, we have

\[
\hat{y}(\xi) = \mathcal{A}_m(\xi) \hat{c}_m(\xi).
\]

We can solve this equation with respect to \(\hat{c}_m(\xi)\) (which we use to produce \(f\) ) if \(\mathcal{A}_m(\xi)\) is invertible. To see how, we note that \(\hat{c}(\xi)\) is 1-periodic. Moreover, from (IV.8), it is not difficult to see that \(\hat{c}\) is \(\frac{1}{m}\)-periodic. Thus, by solving the above system for each \(\xi \in [0, 1/m],\) we can recover \(\hat{c}\) over \(\xi \in [0, 1].\)
Theorem IV.2.2. Let $\phi \in W_0^1$ and $a \in W(L^1)$, then $\Phi_j \in C(\mathbb{T})$ for $j = 1, \ldots, m$. Moreover, let

$$\hat{\Phi}_j(\xi) = \begin{pmatrix} \hat{\Phi}_0(\frac{\xi}{m}) & \hat{\Phi}_0(\frac{\xi+1}{m}) & \cdots & \hat{\Phi}_0(\frac{\xi+m-1}{m}) \\ \hat{\Phi}_1(\frac{\xi}{m}) & \hat{\Phi}_1(\frac{\xi+1}{m}) & \cdots & \hat{\Phi}_1(\frac{\xi+m-1}{m}) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\Phi}_{m-1}(\frac{\xi}{m}) & \hat{\Phi}_{m-1}(\frac{\xi+1}{m}) & \cdots & \hat{\Phi}_{m-1}(\frac{\xi+m-1}{m}) \end{pmatrix},$$

(IV.10)

$\xi \in \mathbb{T}$. Then a vector $f \in V(\phi)$ can be recovered in a stable way, i.e. the inverse is bounded, from the measurements $y_n$, for $n = 0, \ldots, m - 1$, if and only if $\det \mathcal{A}_m(\xi) \neq 0$ for any $\xi \in [0, 1]$.

Although there are cases in which the conditions of proposition III.1.1 are satisfied, there are many situations in practice for which the hypotheses of proposition III.1.1 are not satisfied. Thus, we need additional samples. The number of additional samples that we need and their locations in order to reconstruct the original signal $f$ will constitute the remainder of this section.

Since $\hat{\Phi}_j(\xi)$, $j = 0, \ldots, m - 1$, are 1-periodic, it is sufficient to study the behavior of $\mathcal{A}_m(\xi)$ for $|\xi| \leq \frac{1}{2}$. If we further assume that both $\hat{a}$ and $\hat{\phi}$ are real and symmetric, then the $\hat{\Phi}_j$s are also real and symmetric. However, the symmetry and periodicity of the $\hat{\Phi}_j$s cause $\mathcal{A}_m(0)$ and $\mathcal{A}_m(\frac{1}{2})$ to be singular. Note that these conditions together imply a symmetry about multiples of $\frac{1}{2}$. Writing the matrix $\mathcal{A}_m(0)$ explicitly, we have

$$\mathcal{A}_m(0) = \begin{pmatrix} \Phi_0(0) & \Phi_0(\frac{1}{m}) & \cdots & \Phi_0(\frac{m-1}{m}) \\ \Phi_1(0) & \Phi_1(\frac{1}{m}) & \cdots & \Phi_1(\frac{m-1}{m}) \\ \vdots & \vdots & \ddots & \vdots \\ \Phi_{m-1}(0) & \Phi_{m-1}(\frac{1}{m}) & \cdots & \Phi_{m-1}(\frac{m-1}{m}) \end{pmatrix}. \quad \text{(IV.11)}$$

It is easy to see that the second and last columns of $\mathcal{A}_m(0)$ coincide. In fact, the third and $(m - 1)$-th column coincide, and so on. Considering $\mathcal{A}_m(\frac{1}{2})$ similarly, we see that the first and last columns coincide, as do the second and $(m - 2)$-th columns and so on. Thus, we have proved the following theorem.

Theorem IV.2.3. Suppose that both $\hat{a}$ and $\hat{\phi}$ are real and symmetric, then $\mathcal{A}_m(0)$ and $\mathcal{A}_m(\frac{1}{2})$ are singular.

Thus, it is clear that the conditions of Theorem IV.2.4 are not satisfied whenever $\hat{a}$ and $\hat{\phi}$ are real and symmetric, which is a case of practical importance. Therefore,
in order to solve the dynamical sampling for this case, we need to take extra samples. But where do we take the extra samples and how large is this extra sampling set? We answer this last question when $A_m(\xi)$ has only finitely many singularities. For this case, it is possible to stably recover the original signal $f$ by taking some additional samples. Again, let $T_c$ be the operator that shifts a vector in $\ell^2(\mathbb{Z})$ to the right by $c$ units so that $T_c z(k) = z(k - c)$. Let $S_{mn} T_c$ represent shifting by $c$ and then sampling by $mn$ for some positive integer $n$. We have

**Theorem IV.2.4.** Suppose $A_m(\xi)$ is singular only when $\xi \in \{\xi_i\}_{i \in I}$ with $|I| < \infty$. Let $n$ be a positive integer such that $|\xi_i - \xi_j| \neq k_n$ for any $i, j \in I$ and $k \in \{1, \ldots, n - 1\}$. Then the additional sampling given by $\{S_{mn} T_c\}_{c \in \{1, \ldots, m - 1\}}$ provides enough additional information to stably recover any $f \in V(\phi)$.

Note that the finite nature of $I$ guarantees the existence of an $n$ satisfying the conditions of theorem IV.2.4. The proof is similar to that of theorem III.4.1 and will not be given here.

**Proof of Lemma IV.2.1**

In the light of (IV.7), basic convolution properties and using the Poisson summation formula

$$\mathcal{F}(S_m h)(\xi) = \frac{1}{m} \sum_{l=0}^{m-1} \hat{h}(\xi + l/m),$$

(IV.12)

where $\mathcal{F}$ is the Fourier transform operator and $\hat{h} = \sum_k h(k) e^{-2\pi ik\xi}$, we get that for all $j = 0, 1, \ldots, m - 1$

$$\hat{\Phi}_j(\xi) = \sum_k \hat{\phi}_j(\xi + k) \text{ is } 1\text{-periodic.}$$

(IV.13)

Let $\hat{h}_j(\xi) = (f_j|_{\xi})(\xi) = \sum_k \hat{f}_j(\xi + k)$ in (IV.12). Then, using the 1-periodicity of $\hat{c}$ we get

$$\mathcal{F}(S_m h_j)(\xi) = \frac{1}{m} \sum_{l=0}^{m-1} \sum_k \hat{f}_j\left(\frac{\xi + l}{m} + k\right)$$

(IV.14)

$$= \frac{1}{m} \sum_{l=0}^{m-1} \sum_{k=-\infty}^{\infty} \hat{c}\left(\frac{\xi + l}{m} + k\right) \hat{\phi}_j\left(\frac{\xi + l}{m} + k\right)$$

$$= \frac{1}{m} \sum_{l=0}^{m-1} \hat{c}\left(\frac{\xi + l}{m}\right) \sum_{k=-\infty}^{\infty} \hat{\phi}_j\left(\frac{\xi + l}{m} + k\right).$$

By (IV.13) it holds $\mathcal{F}(S_m h_j)(\xi) = \frac{1}{m} \sum_{l=0}^{m-1} \hat{c}(\xi + l/m) \hat{\Phi}_j(\xi + l/m)$. 

50
IV.3 Reduction to the $\ell^2(\mathbb{Z})$ Case

Under the appropriate conditions on $\phi$, the dynamical sampling in SIS reduces to the discrete case described in chapter III. To establish this connection we use the following theorem.

**Theorem IV.3.1.** Let $\phi \in L^2(\mathbb{R})$ be such that $\{\phi(-k) : k \in \mathbb{Z}\}$ is a Riesz basis for its closed span $V(\phi)$. For a distribution $a$ such that $\hat{a} \in L^\infty(\mathbb{R})$, the following are equivalent

1. $a \ast \phi \in V(\phi)$
2. $a \ast V(\phi) \subseteq V(\phi)$
3. there exists a convolutor $b$ with $\hat{b} \in L^\infty$ 1-periodic such that for any $c \in \ell^2(\mathbb{Z})$
   
   $$a \ast (c \ast_{sd} \phi) = (b \ast d c) \ast_{sd} \phi$$

   (IV.15)

4. for every $k \in \mathbb{Z}$ and a.e. $\xi \in [0,1]$
   
   $$\hat{a}(\xi + k) \hat{\phi}(\xi + k) = \hat{b}(\xi) \hat{\phi}(\xi + k)$$

   (IV.16)

**Proof.** (1) $\Rightarrow$ (4) If $a \ast \phi \in V(\phi)$, then there exists $(b_k)_{k \in \mathbb{Z}} \in \ell^2(\mathbb{Z})$ such that

$$a \ast \phi(x) = \sum_{k \in \mathbb{Z}} b_k \phi(x - k).$$

(IV.17)

Taking the Fourier transform of both sides of the (IV.17), for

$$\hat{b}(\xi) = \sum_{k \in \mathbb{Z}} b_k e^{-2\pi ik \xi}$$

we get

$$\hat{a}(\xi) \hat{\phi}(\xi) = \hat{b}(\xi) \hat{\phi}(\xi)$$

which is the same as (IV.16), since $\hat{b}$ is 1-periodic.

(4) $\Rightarrow$ (3) From (IV.16), we get

$$\sum_k |\hat{a}(\xi + k)|^2 |\hat{\phi}(\xi + k)|^2 = |\hat{b}(\xi)|^2 \sum_k |\hat{\phi}(\xi + k)|^2.$$  

Since $\hat{a} \in L^\infty$, using (IV.2) we get

$$|\hat{b}(\xi)|^2 \leq \|\hat{a}\|_{L^\infty}^2.$$
and taking the inverse Fourier transform.

(3) $\Rightarrow$ (2) Noting that $\hat{b}\hat{c} \in L^2[0,1]$ implies $b \ast c \in \ell^2$, we see clearly that the right hand side of (IV.15) is in $V(\phi)$.

The implication (2) $\Rightarrow$ (1) is straightforward. \hfill $\square$

Note that we can reduce the dynamical sampling problem in $V(\phi)$ to the one in $\ell^2(\mathbb{Z})$ using the theorem above. Specifically, if $\phi \in W_0^1$ and the condition that $\tilde{\Phi}_0(\xi) = \sum_k \hat{\phi}(\xi + k) \neq 0$, then for each $f = c \ast_{sd} \phi$ we associate $x \in \ell^2$ by $x = f(\mathbb{Z})$. The map $f \mapsto x$ from $V(\phi)$ to $\ell^2$ is well defined, since $\phi \in W_0^1$. Note that the convolution operator $a \ast f$ corresponds to the discrete convolution $b \ast_{sd} x$ where $b$ is obtained from $a$ as in Theorem IV.3.1. Hence $S_m(a^f) = S_m(b^n x)$. By solving the dynamical system on $\ell^2$ to obtain $x$, we can recover $f$ by finding $\hat{c} = \hat{x}/\hat{\Phi}_0$. Since $\tilde{\Phi}_0$ is continuous and nonzero, $\hat{c} \in L^2[0,1]$.

As a particular case of Theorem IV.3.1, if the sets $E_k = \{\text{supp } \hat{\phi}(\xi + k), \xi \in [-1/2, 1/2]\}$ are disjoint it is a sufficient conditions for the (IV.16) to hold, as we can take

$$
\hat{b}(\xi) = \sum_{k \in \mathbb{Z}} \hat{a}(\xi + k) \chi_{E_k} \quad \text{for } \xi \in [-1/2, 1/2].
$$

As an example, when $\phi$ is the sinc function, as discussed in Section IV, we get the following corollary.

**Corollary IV.3.2.** When the generating function $\phi$ is such that $\hat{\phi} = \chi_{[-1/2,1/2]}$, then the dynamical sampling in $V(\phi)$ can be reduced to that of the dynamical sampling in $\ell^2(\mathbb{Z})$ with $\hat{b}(\xi) = \hat{a}\chi_{[-1/2,1/2]}(\xi)$ for $\xi \in [-1/2, 1/2]$.

The condition under which the dynamical sampling problem in SIS can be reduced to that in $\ell^2(\mathbb{Z})$ can be further elucidated by the following theorem which can be proved by solving (IV.16).

**Theorem IV.3.3.** Let $\phi \in L^2$ be such that $\{\phi(\cdot - k) : k \in \mathbb{Z}\}$ is a Riesz basis for its closed span $V(\phi)$ with $E = \text{supp } \hat{\phi}$. For a convolutor a such that $\hat{a} \in L^\infty$, and any of the equivalent conditions (1)-(4) of Theorem IV.3.1 is satisfied, then there exists $g \in L^\infty$ such that

$$
\hat{a} = \hat{b}\chi_E + g\chi_{E^c}. \quad (IV.18)
$$

Conversely, if (IV.18) holds, for a 1-periodic $\hat{b} \in L^\infty$, some $g \in L^\infty$ and a measurable set $E$ such that $\sum_j \chi_E(\xi + j) \geq 1$ a.e. $\xi$ then clearly $\hat{a} \in L^\infty$. In addition, for any $\phi$
with $E = \text{supp} \hat{\phi}$ satisfying (IV.2) (i.e., $\{ \phi(\cdot - k) \; k \in \mathbb{Z} \}$ is a Riesz basis for $V(\phi)$), the four equivalent conditions of Theorem IV.3.1 are satisfied.

IV.4 $\ell^2(\mathbb{Z})$-like Dynamical Systems

We say that the matrix $A_m(\xi)$ in (IV.10) is a Vandermonde-like matrix if there exists a $\hat{b}(\xi) \in L^2[0,1]$ function such that

$$\hat{\Phi}_j(\xi) = \hat{b}^j(\xi) \hat{\Phi}_0(\xi), \; j = 0, 1, 2, \ldots$$

In case of Vandermonde-like matrices we have

$$A_m(\xi) = \mathcal{B}_m(\xi) \text{diag}\left( \hat{\Phi}_0(\frac{\xi}{m}), \hat{\Phi}_0(\frac{\xi+1}{m}), \ldots, \hat{\Phi}_0(\frac{\xi+m-1}{m}) \right). \quad (IV.19)$$

where

$$\mathcal{B}_m(\xi) = \begin{pmatrix} 1 & 1 & \ldots & 1 \\ \hat{b}(\frac{\xi}{m}) & \hat{b}(\frac{\xi+1}{m}) & \ldots & \hat{b}(\frac{\xi+m-1}{m}) \\ \hat{b}^{m-1}(\frac{\xi}{m}) & \hat{b}^{m-1}(\frac{\xi+1}{m}) & \ldots & \hat{b}^{m-1}(\frac{\xi+m-1}{m}) \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \end{pmatrix}. \quad (IV.20)$$

Thus $\det A_m(\xi) = \hat{\Phi}_0(\frac{\xi}{m}) \cdots \hat{\Phi}_0(\frac{\xi+m-1}{m}) \det \mathcal{B}_m(\xi)$.

The matrix $\mathcal{B}_m(\xi)$ is the same type of matrix that appears when solving the dynamical sampling problem in $\ell^2(\mathbb{Z})$. If it is known that $\hat{\Phi}_0 = \sum_j \hat{\phi}(\xi + i) \neq 0$, then the invertibility of $A_m(\xi)$ is equivalent to the invertibility of $B_m(\xi)$. Moreover, if $\hat{\Phi}_0(\xi) = \sum_j \hat{\phi}(\xi + i) \neq 0$, we get $|\hat{\Phi}_0(\xi)| > \delta > 0$ for some positive $\delta$. Hence, the invertibility and stability of $B_m(\xi)$ will imply the invertibility and stability of $A_m(\xi)$. Thus, the results of chapter III can be used to determine invertibility and stability of $B_m(\xi)$.

Notice that if the condition (IV.16) holds then from the Poisson summation formula it follows that $A_m(\xi)$ is a Vandermonde-like matrix, but the inverse is not always true. For example, take

$$\hat{\phi}(\xi) = -1\chi_{[0,1]} + \chi_{[1,2]} + \chi_{[2,3]}$$

and

$$\hat{a}(\xi) = \chi_{[0,1]} + 2\chi_{[1,2]} + \chi_{[2,3]}.$$
Then
\[ \Phi_j(\xi) = \sum_k \hat{a}^j(\xi + k)\hat{\phi}(\xi + k) = 2^j \chi_{[0,1)}. \]

Hence, in this case \( A_m(\xi) \) is a Vandermonde-like matrix with \( \hat{b}(\xi) = 2\chi_{[0,1]} \), but obviously the condition (IV.16) fails.
CHAPTER V

DYNAMICAL SAMPLING WITH A FORCING TERM.

In this chapter, we explore the dynamical sampling problem when an unknown source term enters the system during the time period when samples are taken. For example, if the signal being measured is a pollutant, the dynamical sampling problem in the previous chapters assumes that the pollution was released at time \( t=0 \) and that no additional pollution enters the system while the samples are taken. Here we explore what happens when this assumption is removed and allow for pollution to enter the system during the sampling period.

Even under the strong assumption that the source is independent of time, that is the forcing term is constant in the time variable, the dynamical sampling with a forcing term problem is fundamentally different. In fact, the operator of regular subsampling is not injective, that is, it has a nontrivial kernel. However, if we impose additional constraints on the forcing term, we are able to reduce the problem to the one studied in chapter III.

V.1 Problem Formulation

If \( A \) is the evolution rule, \( x_0 \) is the unknown initial signal, and \( x \) is the unknown forcing term, then the signal \( y_j \) at time \( t = j \) is modeled by

\[
\begin{align*}
y_0 &= x_0 \\
y_1 &= Ay_0 + x = Ax_0 + x \\
y_2 &= Ay_1 + x = A^2x_0 + Ax + x \\
&\quad \vdots \\
y_N &= Ay_N + x = A^N x_0 + A^{N-1} x + A^{N-2} x + \ldots + Ax + x
\end{align*}
\]

Note that the notation in this section differs from that in chapter III, in which \( y_j \) represented the subsampled signal at time \( t = j \). Here \( y_j \) denotes the signal before it is subsampled. If for any \( j \in \mathbb{N} \), both \( y_j \) and \( y_{j+1} \) are completely known, then \( x \) can be easily recovered by the relationship \( x = y_{j+1} - Ay_j \). It’s quite simple, and of course, not the question we are interested in. Let \( S \) be some subsampling operator so that the measured signal at time \( j \) is given by \( Sy_j \). In general, \( S \) and \( A \) do not commute, so that

\[
Sy_{j+1} = S(Ay_j + x) \neq ASy_j + Sx \quad \text{(V.1)}
\]
This means that knowledge of $Sy_j$ and $Sy_{j+1}$ does not allow for the reconstruction of $Sx$ by taking differences as above.

In this section, we keep the assumptions of chapter III that $A$ is a convolution operator so that $Ax = a \ast x$ where $\hat{a} \in L^\infty(\mathbb{T})$ and that the subsampling is regular and constant in time. The measured signal at time $t = j$ is $S_my_j$. The dynamical sampling procedure in the presence of a forcing term can be written as

$$y = A_Fx$$

where $A_F$ is an operator from $(l^2(\mathbb{Z}))^2$ to $(l^2(\mathbb{Z}))^N$, $x = (x_0, x)$, and

$$y = (S_my_0, S_my_1, \ldots, S_my_{N-1}).$$

In general, the operator $A_F$ does not have a bounded left inverse. In fact, even under the best conditions on the filter $\hat{a}$, the operator $A_F$ is not injective.

**Theorem V.1.1.** Suppose $\hat{a}$ is such that the conditions of proposition III.1.1 are satisfied. Then $\ker A_F$ is the set of all $x = (x_0, x)$ such that for a.e. $\xi \in \mathbb{T}$

$$\sum_{l=0}^{m-1} \hat{x}_0 \left( \frac{\xi + l}{m} \right) = 0, \text{ and}$$

$$\hat{x}(\xi) = (1 - \hat{a}(\xi)) \hat{x}_0(\xi).$$

**Proof.** Using the techniques of the proof of proposition III.1.1 we can reduce the study of the operator $A_F$ to studying a family of matrices defined on $\mathbb{T}$. Here we derive and study the family of matrices that correspond to $A_F$.

Using the Poisson Summation formula (III.3), the subsampled signal at time $t_j$ can be written as

$$(S_my_j)^\wedge(\xi) = \frac{1}{m} \sum_{l=0}^{m-1} \hat{a}^j \left( \frac{\xi + l}{m} \right) \hat{x}_0 \left( \frac{\xi + l}{m} \right)$$

$$+ \frac{1}{m} \sum_{l=0}^{m-1} \left[ \hat{a}^{j+1} \left( \frac{\xi + l}{m} \right) + \ldots + \hat{a} \left( \frac{\xi + l}{m} \right) + 1 \right] \hat{x} \left( \frac{\xi + l}{m} \right)$$

We can write the problem in matrix form:

$$m\tilde{y}(\xi) = (B(\xi)C(\xi)) \begin{pmatrix} \hat{x}_0(\xi) \\ \hat{x}(\xi) \end{pmatrix},$$

(V.6)
where

\[ \hat{y}(\xi) = \begin{pmatrix} S_m y_0(\xi) \\ \vdots \\ S_m y_{2m}(\xi) \end{pmatrix}, \quad \hat{x}_0(\xi) = \begin{pmatrix} \hat{x}_0(\xi_m) \\ \vdots \\ \hat{x}_0(\xi_{m-1}+1) \end{pmatrix}, \quad \hat{x}(\xi) = \begin{pmatrix} \hat{x}(\xi_m) \\ \vdots \\ \hat{x}(\xi_{m-1}+1) \end{pmatrix}, \]

\[ B(\xi) = \begin{pmatrix} 1 & 1 & \ldots & 1 \\ \hat{a}(\xi_m) & \hat{a}(\xi_{m+1}) & \ldots & \hat{a}(\xi_{m-1}+1) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{a}^{(2m-1)}(\xi_m) & \hat{a}^{(2m-1)}(\xi_{m+1}) & \ldots & \hat{a}^{(2m-1)}(\xi_{m-1}+1) \end{pmatrix}, \]

\[ C(\xi) = \begin{pmatrix} 0 & \ldots & 0 \\ 1 & \ldots & 1 \\ \hat{a}(\xi_m) + 1 & \ldots & \hat{a}(\xi_{m-1}+1) + 1 \\ \vdots & \vdots & \vdots \\ \hat{a}^{(2m-2)}(\xi_m) + \ldots + \hat{a}(\xi_{m-1}+1) + 1 & \ldots & \hat{a}^{(2m-2)}(\xi_{m-1}+1) + \ldots + \hat{a}(\xi_{m-1}+1) + 1 \end{pmatrix}. \]

Note that 2m time samples are taken because the right hand side of (V.6) has 2m unknowns so there must be at least 2m equations.

Using the relation \( z^{n-1} + z^{n-2} + \ldots + z + 1 = z^m \) we can see that the matrix \((B(\xi)|C(\xi))\) has rank \( m + 1 \). Multiply the \( j \)-th column of \( C(\xi) \) by \( \hat{a}(\xi_m + j) - 1 \) and then from the resulting matrix subtract the \( j \)-th row of \( B(\xi) \) to obtain the following matrix

\[ \begin{pmatrix} -1 & \ldots & -1 \\ B(\xi) | & \ddots & \ddots \\ -1 & \ldots & -1 \end{pmatrix}, \]

which clearly has rank \( m + 1 \) when the hypotheses of the proposition are satisfied.

By inspection, we can see that vectors satisfying the two conditions below are in the kernel of \((B(\xi)|C(\xi))\):

\[ \hat{x}(\xi_m + l) = (1 - \hat{a}(\xi_m + l)) \hat{x}_0(\xi_m + l) \text{ for all } l = 0, \ldots, m - 1, \]

and \( \sum_{l=0}^{m-1} \hat{x}_0(\xi_m + l) = 0 \)

The above \( m + 1 \) equations define a subspace of \( \mathbb{C}^{2m} \) of dimension \( m - 1 \), and thus completely characterize \( \ker(B(\xi)|C(\xi)) \).
A signal \( x = (x_0, x) \) is in \( \ker A_F \) if and only if \( \left( \tilde{x}_0(\xi), \tilde{x}(\xi) \right) \in \ker \left( B(\xi) | C(\xi) \right) \) for a.e. \( \xi \in \mathbb{T} \). Thus, we have proved the theorem.

\[ \square \]

Remark V.1.2. If \( \|A\| < 1 \), \( x_0(mk) = 0 \) for all \( k \in \mathbb{Z} \), and \( x_0 = \sum_{n=0}^{\infty} A^n x \), then \( (x_0, x) \in \ker A_F \). It is not surprising that dynamical sampling fails in this case because the system is static, that is \( y_j = y_{j+1} = x_0 \). The spirit of dynamical sampling is to use the dynamics of a system to compensate for spatial undersampling by temporal oversampling. In a static system, temporal sampling does not provide the necessary additional information to offset spatial undersampling.

Remark V.1.3. Because the operator \( A_F \) is not injective, the additional sampling schemes given in section III are not enough to allow for the recovery of \( x \). However, these additional sampling schemes can be used to extend theorem V.1.1 when the conditions of proposition III.1.1 are not satisfied.

If we have some a priori information about the signal \( x \), we may be able to solve the dynamical sampling with a forcing term problem. One such example is discussed below.

V.2 Assumptions on Forcing Term

Perhaps the simplest case of dynamical sampling with a forcing term is when it is assumed that the initial signal is the forcing term, i.e. \( x_0 = x \). In the wireless sensor network setting, this may mean that sensors are in place and actively sampling a null field when a physical phenomena occurs, producing the forcing term. The phenomena is captured from the beginning. The results in this special case parallel those in chapter III.

Theorem V.2.1. Let \( m \in \mathbb{Z}^+ \) be fixed. Suppose it is known that \( x_0 = x \). Assume that \( \hat{a} \in L^\infty(\mathbb{T}) \) and define

\[
C(\xi) = \begin{pmatrix}
1 & \cdots & 1 \\
\hat{a}(\xi_0) + 1 & \cdots & \hat{a}(\xi + m-1_m) + 1 \\
\vdots & \vdots & \vdots \\
\hat{a}^{m-1}(\xi_0) + \cdots + \hat{a}(\xi) + 1 & \cdots & \hat{a}^{m-1}(\xi + m-1_m) + \cdots + \hat{a}(\xi + m-1_m) + 1
\end{pmatrix}, \quad \xi \in \mathbb{T}
\]

\( A_F \) in (V.2) has a bounded left inverse for some \( N \geq m \) if and only if there exists \( \alpha > 0 \) such that the set \( \{ \xi \in \mathbb{T} : |\det C(\xi)| < \alpha \} \) has zero measure. Consequently, \( A_F \) has a bounded left inverse for some \( N \geq m \) if and only if \( A_F \) has a bounded left inverse for all \( N \geq m \).
Proof. The assumption that $x_0 = x$ reduces the domain of $A_F$ to $\ell^2(\mathbb{Z})$, and the signal at time $t = j$ is given by $y_j = a^*_j x + a^{j-1}_* x + \ldots + a_* x + x$, where $a^*_j x = (a^*_j x + \ldots + a)_* x$.

The remainder of the proof is identical to the proof of proposition III.1.1. \qed

Remark V.2.2. The matrix $C(\xi)$ can be written as the matrix product

$$
C(\xi) = \begin{pmatrix}
1 & 0 & \ldots & 0 \\
1 & 1 & \ddots & 0 \\
\vdots & \vdots & \ddots & 0 \\
1 & 1 & \ldots & 1
\end{pmatrix} \mathcal{A}(\xi),
$$

(V.7)

where $\mathcal{A}(\xi)$ is defined in (III.2). Since the first matrix in the product of (V.7) is obviously invertible, $C(\xi)$ has an inverse if and only if $\mathcal{A}(\xi)$ has an inverse. Also $\mathcal{A}(\xi)$ and $C(\xi)$ have the same kernel. Thus, the additional sampling schemes defined in sections III.2 and III.4 to stabilize a left inverse of $A$ apply here and stabilize a left inverse of $A_F$. 

59
CHAPTER VI

PROOFS OF LEMMAS AND FORMULAS

Proof of Formula (III.3). Using the fact that \( \sum_{l=0}^{m-1} e^{\frac{i\pi j}{m}} = \begin{cases} m, & j = 0 \mod m \\ 0, & \text{otherwise} \end{cases} \), we compute the formula directly:

\[
(S_m z)^{\wedge} (\xi) = \sum_k (S_m z)(k) e^{-i 2\pi k \xi}
\]

\[
= \sum_k \frac{1}{m} \sum_{j=0}^{m-1} \left( z(mk + j) e^{-i 2\pi \frac{mk + j}{m}} \sum_{l=0}^{m-1} e^{\frac{i\pi l}{m}} \right)
\]

\[
= \frac{1}{m} \sum_{l=0}^{m-1} \sum_k z(mk + j) e^{-i 2\pi \frac{mk + j}{m}}
\]

\[
= \frac{1}{m} \sum_{l=0}^{m-1} z(k) e^{-i 2\pi \left( \frac{k}{m} + \frac{l}{m} \right)}
\]

\[
= \frac{1}{m} \sum_{l=0}^{m-1} \hat{z} \left( \frac{\xi}{m} + \frac{l}{m} \right)
\]

Proof of Lemma III.1.2. Suppose \( \text{ess sup}_{\xi} \|A(\xi)\|_{op} = \alpha < \infty \) and let \( z \in (L^2(\mathbb{T}))^m \) be such that \( \|z\|_{(L^2(\mathbb{T}))^m} = 1 \). Then

\[
\|A z\|_{(L^2(\mathbb{T}))^m}^2 = \int_{\mathbb{T}} |A(\xi) z(\xi)|^2 d\xi
\]

\[
\leq \int_{\mathbb{T}} \|A(\xi)\|_{op}^2 \|z(\xi)|^2 d\xi
\]

\[
\leq \alpha^2 \|z\|_{(L^2(\mathbb{T}))^m}^2
\]

\[
= \alpha^2.
\]

Thus \( \|A\|_{op} = \sup_{\|z\|_{(L^2(\mathbb{T}))^m} = 1} \|A z\|_{(L^2(\mathbb{T}))^m} \leq \alpha \).

For \( \varepsilon > 0 \), define \( B = \{ \xi : \|A(\xi)\|_{op} > \alpha - \varepsilon \} \). Using the singular value decomposition, we can write \( A(\xi) \) as the product

\[
A(\xi) = U(\xi) \Sigma(\xi) V^*(\xi),
\]

where \( U(\xi) \) is an \( n \times n \) unitary matrix, \( \Sigma(\xi) \) is a diagonal matrix with nonnegative, real entries on the diagonal, and \( V(\xi) \) is an \( m \times m \) unitary matrix. We assume the diagonal entries, \( \sigma_i(\xi) \) of \( \Sigma(\xi) \), called singular values of \( A(\xi) \), are listed in descending order. The singular value decomposition allows us to see clearly the operator norm of \( A(\xi) \).
In fact, \( \|A(\xi)\|_{op} = \sigma_1(\xi) \). To see this consider the product \( A(\xi)v_1(\xi) \), where \( v_1(\xi) \) is the first column vector of \( V(\xi) \). Then

\[
|A(\xi)v_1(\xi)| = |U(\xi)\Sigma(\xi)V(\xi)v_1(\xi)| = |U(\xi)\Sigma(\xi)| = |U(\xi)| |U(\xi)| |\Sigma(\xi)| = |\sigma_1(\xi)\|
\]

where the second and last equalities follow from the fact that \( V(\xi) \) and \( U(\xi) \) are unitary matrices, respectively. We define a function \( z \) in \( (L^2(T))^m \) by

\[
z(\xi) = \frac{1}{\sqrt{|B|}} \chi_B(\xi)v_1(\xi),
\]

where \( \chi_B \) is the characteristic function of the set \( B \). Measurability of \( z \) comes from being the composition of two measurable functions, the SVD (need reference) and the map \( \xi \mapsto A(\xi) \). Then

\[
\|Az\|^2 = \frac{1}{|B|} \int_B |A(\xi)z(\xi)|^2 d\xi = \frac{1}{|B|} \int_B |\sigma_1(\xi)|^2 d\xi > (\alpha - \varepsilon)^2 \quad (VI.6)
\]

Thus \( \|A\|_{op} = \sup_{\|z\|(L^2(T))^m=1} \|Az\|(L^2(T))^m \geq \alpha - \varepsilon \). This holds for any \( \varepsilon > 0 \), and so \( \|A\|_{op} \leq \alpha \).

Suppose \( \operatorname{esssup}_T \|A(\xi)\|_{op} = \infty \). Fix \( N > 0 \). Then the set \( B = \{\xi : \|A(\xi)\|_{op} > N\} \) has positive measure. Repeating the process above, we find a function \( z \) of unit norm in \( (L^2(T))^m \) such that \( \|Az\|(L^2(T))^m > N \). Since \( N \) was arbitrary, we conclude that \( \|A\| = \infty \).

\( \square \)

**Proof of Lemma III.3.7.** Recall that \( \ker A^* = (\ker A)^\perp \) and observe that \( \ker A^\perp = \ker A^* \). Since all left inverses of \( A \) must agree on \( \ker A \), we have \( A^\perp = A^\perp P_{\ker A} = A^\perp P_{\ker A} = A^\perp P_{\ker A} \), where \( P_{\ker A} \) is the orthogonal projection onto the range of \( A \). Then \( \|A^\perp\| = \|A^\perp P_{\ker A}\| \leq \|A^\perp\|\|P_{\ker A}\| = \|A^\perp\| \).

\( \square \)

**Proof of Lemma III.3.8.** Since \( BA = I_n \), we have

\[
\delta_{k,l} = \sum_{i=1}^m B(k,i)A(i,l) = \sum_{i=1}^m B(k,\sigma(i))A(\sigma(i),l) = \sum_{i=1}^m B(\eta(k),i)\tilde{A}(i,\eta(l))
\]

Since \( \eta \) is a permutation, we know \( \delta_{k,l} = \delta_{\eta(k),\eta(l)} \). Thus, \( \tilde{B} \tilde{A} = I_n \). Now, for any \( x \in C^m \)

\[
(Bx)(j) = \sum_{i=1}^m B(j,i)x(i) = \sum_{i=1}^m B(j,\sigma(i))x(\sigma(i)) = \sum_{i=1}^m \tilde{B}(\eta(j),i)\tilde{x}(i) = (\tilde{B}\tilde{x})(\eta(j))
\]

where \( \tilde{x}(i) := x(\sigma(i)) \). This shows \( \|Bx\| = \|\tilde{B}\tilde{x}\| \). Since \( \|x\| = 1 \) if and only if \( \|\tilde{x}\| = 1 \),
we have $\|B\|_{op} = \|\tilde{B}\|_{op}$. \hfill \Box

Proof of Lemma III.4.5. By replacing $W_i$ with $M_i^\perp$ and taking the orthogonal compliment of both sides of the equality in (III.62), it suffices to prove the following equality:

$$\left( \bigcap_i^N M_i^\perp \right) = (\operatorname{span}\{M_i, \ldots, M_N\})^\perp.$$  \tag{VI.9}

The proof is by induction. We first prove the base case.

Claim: For any subspaces $M, N$ of $\mathbb{C}^d$, the following holds

$$(\operatorname{span}\{M, N\})^\perp = M^\perp \cap N^\perp$$

Proof of claim: We first show $(\operatorname{span}\{M, N\})^\perp \subseteq M^\perp \cap N^\perp$. Suppose $v \in (\operatorname{span}\{M, N\})^\perp$ so that $v$ is perpendicular to every vector in $\operatorname{span}\{M, N\}$. Since $M \subseteq \operatorname{span}\{M, N\}$, it follows that $v$ is perpendicular to every vector in $M$. So $v \in M^\perp$, and similarly $v \in N^\perp$.

Thus, $v \in M^\perp \cap N^\perp$.

We next show $M^\perp \cap N^\perp \subseteq (\operatorname{span}\{M, N\})^\perp$. Suppose $v \in M^\perp \cap N^\perp$ and $u = ax + by$ with $a, b$ being scalars, $x \in M$, and $y \in N$ so that $u$ is any vector in $\operatorname{span}\{M, N\}$. Since $v$ is perpendicular to both $x$ and $y$, it follows from linearity that $v$ is perpendicular to $u$.

Thus, $v \in (\operatorname{span}\{M, N\})^\perp$. This proves the claim.

Now, assume (VI.9) holds for some $k$. By the claim above

$$\bigcap_{i}^{k+1} M_i^\perp = (\operatorname{span} \left\{ \left( \bigcap_{i}^{k} M_i^\perp \right)^\perp, M_{k+1} \right\} )^\perp = (\operatorname{span} \{ \operatorname{span}\{M_1, \ldots, M_k\}, M_{k+1} \} )^\perp = (\operatorname{span}\{M_1, \ldots, M_k, M_{k+1} \})^\perp.$$  

The first equality follows from the claim above, and the second equality follows from the induction hypothesis. This shows that if (VI.9) holds for $k$, it also hold for $k + 1$.

Thus, by induction, the lemma is proved. \hfill \Box
BIBLIOGRAPHY


