2017

Sampling and Reconstruction of Spatial Signals

Cheng Cheng
University of Central Florida

SAMPLING AND RECONSTRUCTION OF SPATIAL SIGNALS

by

CHENG CHENG
MS University of Central Florida, 2013

A dissertation submitted in partial fulfilment of the requirements
for the degree of Doctor of Philosophy
in the Department of Mathematics
in the College of Science
at the University of Central Florida
Orlando, Florida

Summer Term
2017

Major Professors: Xin Li and Qiyu Sun

ABSTRACT

Digital processing of signals f may start from sampling on a discrete set Γ, $f \longmapsto \big(f(\gamma_n)\big)_{\gamma_n\in\Gamma}$. The sampling theory is one of the most basic and fascinating topics in applied mathematics and in engineering sciences. The most well known form is the uniform sampling theorem for bandlimited/wavelet signals, that gives a framework for converting analog signals into sequences of numbers. Over the past decade, the sampling theory has undergone a strong revival and the standard sampling paradigm is extended to non-bandlimited signals including signals in reproducing kernel spaces (RKSs), signals with finite rate of innovation (FRI) and sparse signals, and to non-traditional sampling methods, such as phaseless sampling.

In this dissertation, we first consider the sampling and Galerkin reconstruction in a reproducing kernel space. The fidelity measure of perceptual signals, such as acoustic and visual signals, might not be well measured by least squares. In the first part of this dissertation, we introduce a fidelity measure depending on a given sampling scheme and propose a Galerkin method in Banach space setting for signal reconstruction. We show that the proposed Galerkin method provides a quasi-optimal approximation, and the corresponding Galerkin equations could be solved by an iterative approximation-projection algorithm in a reproducing kernel subspace of L^p.

A spatially distributed network contains a large amount of agents with limited sensing, data processing, and communication capabilities. Recent technological advances have opened up possibilities to deploy spatially distributed networks for signal sampling and reconstruction. We introduce a graph structure for a distributed sampling and reconstruction system by coupling agents in a spatially distributed network with innovative positions of signals. We split a distributed sampling and reconstruction system into a family of overlapping smaller subsystems, and we show that the stability of the sensing matrix holds if and only if its quasi-restrictions to those subsystems have

uniform stability. This new stability criterion could be pivotal for the design of a robust distributed sampling and reconstruction system against supplement, replacement and impairment of agents, as we only need to check the uniform stability of affected subsystems. We also propose an exponentially convergent distributed algorithm for signal reconstruction, that provides a suboptimal approximation to the original signal in the presence of bounded sampling noises.

Phase retrieval (Phaseless Sampling and Reconstruction) arises in various fields of science and engineering. It consists of reconstructing a signal of interest from its magnitude measurements. Sampling in shift-invariant spaces is a realistic model for signals with smooth spectrum. We consider phaseless sampling and reconstruction of real-valued signals in a shift-invariant space from their magnitude measurements on the whole Euclidean space and from their phaseless samples taken on a discrete set with finite sampling density. We find an equivalence between nonseparability of signals in a shift-invariant space and their phase retrievability with phaseless samples taken on the whole Euclidean space. We also introduce an undirected graph to a signal and use connectivity of the graph to characterize the nonseparability of high-dimensional signals. Under the local complement property assumption on a shift-invariant space, we find a discrete set with finite sampling density such that signals in shift-invariant spaces, that are determined by their magnitude measurements on the whole Euclidean space, can be reconstructed in a stable way from their phaseless samples taken on that discrete set. We also propose a reconstruction algorithm which provides a suboptimal approximation to the original signal when its noisy phaseless samples are available only.

To my dear parents

ACKNOWLEDGMENTS

First I would like to thank my advisors Professors Xin Li and Qiyu Sun. I am extremely fortunate to work under their supervisory. Without their support, countless hours of explaining and generosity, I don't think this work can ever come to fruition. Dr. Li always ask me to think problems intuitively and ask "why" when doing research. I have enjoyed Dr. Sun's lectures on Wavelet and its application in my second year at UCF, and working closely with Dr. Sun the recent 3 years. He gives me a lot of freedom on my research, yet pushes me whenever necessary, and always makes time for me, even during the weekends. I appreciate Dr. Li and Dr. Sun's patience and encouragement when I am being unproductive or stuck. I would not have gone this far without a lot of their time and sincere effort. I am grateful to thank them for their support.

I am grateful to Dr. Jiongmin Yong, Dr. Zhe Liu and Dr. Mengyu Xu for serving in my dissertation committee. I would like to express my sincere gratitude to Dr. Zuhair Nashed, Dr. Deguang Han and Dr. Zhisheng Shuai for their willingness to provide help whenever I ask. I would like to extend my appreciation to Dr. Yingchun Jiang, Dr. Yang Chen and Dr. Junzheng Jiang for offering very valuable comments towards this dissertation, as well as to my research in general, during their visits at UCF.

I would like to express my sincere and greatest regards to my parents and friends. Without their constant motivation and inspiration, this work would have never come to fruition.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

Digital processing of signals f may start from sampling on a discrete set Γ,

$$f \longmapsto \big(f(\gamma_n)\big)_{\gamma_n \in \Gamma} \tag{1.1}$$

[13, 131, 158, 159]. The sampling theory is one of the most basic and fascinating topics in applied mathematics and in engineering sciences. The celebrated Whittaker-Shannon-Kotelnikov's sampling theorem states that a bandlimited signal can be recovered from its samples taken at a rate greater than twice the bandwidth [131, 163]. In last two decades, that paradigm for bandlimited signals has been extended to represent signals in a shift-invariant space [13, 17, 158], signals with finite rate of innovation [62, 115, 121, 141, 143, 159], signals in a reproducing kernel space [46, 75, 87, 116, 117], and to non-traditional sampling methods, such as dynamic sampling, phaseless sampling, random sampling and mobile sampling [8, 17, 29, 127].

1.1 Sampling and galerkin reconstruction in reproducing kernel spaces

A fundamental problem in sampling theory is how to obtain a good approximation of the signal f when only the noisy sampling data $\big(f(\gamma_n) + \epsilon(\gamma_n)\big)_{\gamma_n \in \Gamma}$ is available [3, 13, 141, 158]. The above problem is well studied and many algorithms, such as the frame algorithm and the approximation-projection algorithm, have been proposed [11, 47, 64, 70, 116, 141, 148].

A conventional way to reconstruct signals f in a linear space V from their sampling data is to solve a minimization problem

$$Rf := \operatorname{argmin}_{h \in V} \|h - f\|, \tag{1.2}$$

where the fidelity measure $\|h-f\|$ depends only on the sampling data of $h-f$ on Γ. Typical examples of fidelity measures in the bandlimited setting are weighted sampling energy $\sum_{\gamma_n\in\Gamma} w_n|f(\gamma_n)-h(\gamma_n)|^2$ and weighted pre-reconstruction energy $\|\sum_{\gamma_n\in\Gamma} w_n(f(\gamma_n)-h(\gamma_n))\mathrm{sinc}(\cdot-\gamma_n)\|_2$, where w_n are positive weights appropriately selected.

The fidelity of perceptual signals, such as acoustic and visual signals, might not be well measured by some weighted square errors [42, 162]. In Chapter 2, we introduce a general fidelity measurement associated with a linear operator S on a Banach space V, that depends on the sampling scheme (1.1). Then the minimization problem (1.2) becomes

$$Rf := \mathrm{argmin}_{h\in V}\|Sh-Sf\|_V. \tag{1.3}$$

The operator S in the above minimization problem can be selected as

$$Sf := \sum_{\gamma_n\in\Gamma} w_n f(\gamma_n)\mathrm{sinc}(\cdot-\gamma_n)$$

for the bandlimited setting, and

$$Sf := \sum_{\gamma_n\in\Gamma} w_n f(\gamma_n)K(\cdot,\gamma_n)$$

for the reproducing kernel space setting.

The nonlinear minimization problem (1.3) does not give a tractable signal reconstruction. Observe that

$$\|Sh-Sf\|_V = \sup_{\|g\|_{V^*}=1, g\in V^*} |\langle Sh-Sf, g\rangle|,$$

where $\langle\cdot,\cdot\rangle$ is the standard dual product between elements in V and its dual V^*. So we propose a Galerkin method (2.3) and (2.4) in Banach space setting for signal reconstruction in Chapter 2.

We also apply the Galerkin reconstruction (2.3) and (2.4) for signals in a reproducing kernel space (RKS) of the form

$$V_{K,p} := \{T_0 f :\ f \in L^p\} = \{f \in L^p :\ T_0 f = f\},\ 1 \le p \le \infty, \tag{1.4}$$

where T_0 is an idempotent integral operator with kernel K,

$$T_0 f(x) := \int_{\mathbb{R}^d} K(x, y) f(y) dy,\ \ f \in L^p. \tag{1.5}$$

The RKS of the form (1.4) has rich geometric structure, lots of flexibility and technical suitability for sampling. It has been used for modeling bandlimited signals, wavelet (spline) signals, and signals with finite rate of innovation [13, 116, 117, 143, 158].

For the sampling scheme (1.1) on $V_{K,p}$, take a disjoint covering

$$\{I_n \subset B(\gamma_n, \delta) :\ \gamma_n \in \Gamma\}$$

of $B(\Gamma, \delta) := \cup_{\gamma\in\Gamma} B(\gamma, \delta) = \cup_{\gamma\in\Gamma}\{x :\ |x - \gamma| \le \delta\}$, and define

$$S_{\Gamma,\delta} f(x) := \sum_{\gamma_n\in\Gamma} |I_n| f(\gamma_n) K(x, \gamma_n),\ \ f \in V_{K,p}, \tag{1.6}$$

where $\delta > 0$. The operator $S_{\Gamma,\delta}$ depends only on the sampling scheme (1.1). We call it a *pre-reconstruction operator*, as $S_{\Gamma,\delta} f(x)$ is a good approximation to $f(x)$ when δ is sufficiently small and $x \in B(\Gamma, \delta)$ is away from the complement of $B(\Gamma, \delta)$, see Figure 1.1. Plotted on the left in Figure 1.1 is a bandlimited signal $f_0 = \sum_i \alpha_i \mathrm{sinc}(\cdot - i)$ with $\alpha_i \in [-1, 1]$ randomly selected. On the right is the difference between f_0 and its pre-reconstruction $h_0 = S_{\Gamma,\delta} f_0$, where $\delta = 1$ and $\Gamma := \{\gamma_k, k = 1, 2, \cdots, 80\}$ is a nonuniform sampling set with $\gamma_1 = -40$ and $\gamma_k - \gamma_{k-1} \in [0.9, 1.1], 2 \le$

$k \leq 80$, being randomly selected. In this figure, the maximal amplitude $\max_{-38 \leq t \leq 38} |f_0(t)|$ of the signal f_0 is 1.7498, while the maximal pre-reconstruction error $\max_{-38 \leq t \leq 38} |h_0(t) - f_0(t)|$ on $[-38, 38] \subset B(\Gamma, 1)$ is 0.6708.

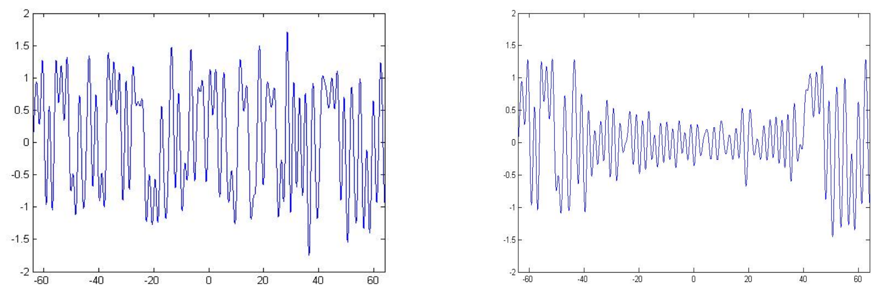

Figure 1.1: A bandlimited signal f_0 and the the difference between f_0 and its pre-reconstruction $h_0 = S_{\Gamma,\delta} f_0$

Due to the above approximation property of the pre-reconstruction operator $S_{\Gamma,\delta}$, we propose the following iterative approximation-projection algorithm

$$g_0 \in U \ \text{ and } \ g_{m+1} = g_m - P_{U,\tilde{U}} S_{\Gamma,\delta} g_m + g_0, \ m \geq 0, \tag{1.7}$$

to solve the Galerkin reconstruction (2.3) and (2.4), where $P_{U,\tilde{U}}$ is an oblique projection for the trial-test space pair $(U, \tilde{U})$. The above algorithm is shown in Theorem 2.12 to have exponential convergence, c.f. [11, 16, 69, 116, 151].

1.2 Spatially distributed sampling and reconstruction

A spatially distributed system (SDS) contains a large amount of agents with limited sensing, data processing, and communication capabilities [4, 44, 165, 166]. Recent technological advances have

opened up possibilities to deploy spatially distributed systems for signal sampling and reconstruction. Comparing with traditional centralized systems that have a powerful central processor and reliable communications between agents and the central processor, an SDS could give unprecedented capabilities especially when creating a data exchange network requires significant efforts (due to physical barriers such as interference), or when establishing a centralized processor presents the daunting challenge of processing all the information (such as big-data problems). In Chapter 3, we introduce a graph structure for a distributed sampling and reconstruction system. For a distributed sampling and reconstruction system, the robustness of signal reconstruction could be reduced to the stability of its sensing matrix. In Chapter 3, we split a distributed sampling and reconstruction system into a family of overlapping smaller subsystems, and we show that the stability of the sensing matrix holds if and only if its quasi-restrictions to those subsystems have uniform stability. Later in Chapter 3, we propose an exponential convergent distributed algorithm for signal reconstruction, that provides a suboptimal approximation to the original signal in the presence of bounded sampling noises.

1.3 Phaseless sampling and reconstruction in shift-invariant spaces

Phase retrieval plays important roles in signal/image/speech processing ([72, 73, 89, 91, 94, 101, 114, 128, 132]). It consists of reconstructing a signal of interest from its magnitude measurements. The underlying recovery problem is possible to solve only if we have additional information about the signal.

The phase retrieval problem of finite-dimensional signals has received considerable attention in recent years, see [19, 20, 34, 37, 161] and references therein, but there are still lots of open mathematical and engineering questions unanswered. In the finite-dimensional setting, a fundamental problem is whether and how a (sparse) vector $\mathbf{x} \in \mathbb{R}^d$ (or $\mathbb{C}^d$) can be reconstructed

from its magnitude measurements $\mathbf{y} = |\mathbf{A}\mathbf{x}|$, where $\mathbf{A}$ is a measurement matrix. The phase retrievability has been characterized via the measurement matrix $\mathbf{A}$ ([20, 23, 161]), and many algorithms have been proposed to reconstruct the vector $\mathbf{x}$ from its magnitude measurements $\mathbf{y}$ ([34, 35, 37, 72, 76, 119, 120, 123, 132]).

The phase retrieval problem in an infinite-dimensional space is different from a finite-dimensional setting. There are several papers devoted to that topic ([5, 6, 7, 33, 110, 125, 126, 133, 154, 164]). Thakur proved in [154] that real-valued bandlimited signals could be reconstructed from their phaseless samples taken at more than twice the Nyquist rate. The above result was extended to complex-valued bandlimited signals by Pohl, Yang and Boche in [126] with samples taken at more than four times the Nyquist rate. Recently, the phase retrievability of signals living in a principal shift-invariant space was studied by Shenoy, Mulleti and Seelamantula in [133] when only magnitude measurements of their frequency are available.

The concept of shift-invariant spaces arose in sampling theory, wavelet theory, approximation theory and signal processing, see [13, 17, 30, 54, 55, 98, 109, 158] and references therein. Sampling in shift-invariant spaces is well studied as it is a realistic model for modelling signals with smooth spectrum, and a suitable model for taking into account the real acquisition and reconstruction devices and the numerical implementation, see [10, 13, 17, 64, 145, 148, 158] and the extensive list of references therein.

Sampling in shift-invariant spaces is a realistic model for signals with smooth spectrum. In Chapter 4 and 5, we consider phaseless sampling and reconstruction of real-valued signals in a shift-invariant space from their magnitude measurements on the whole Euclidean space and from their phaseless samples taken on a discrete set with finite sampling density. We introduce an undirected graph to a signal and use connectivity of the graph to characterize whether the signal can be determined, up to a sign, by its magnitude measurements on the whole Euclidean space. Under the

local complement property assumption on a shift-invariant space, we find a discrete set with finite sampling density such that signals in the shift-invariant space, that are determined by their magnitude measurements on the whole Euclidean space, can be reconstructed in a stable way from their phaseless samples taken on that discrete set. In this paper, we also propose a reconstruction algorithm which provides a suboptimal approximation to the original signal when its noisy phaseless samples are available only. Finally, numerical simulations are performed to demonstrate the robust reconstruction of spline signals from their noisy phaseless samples.

CHAPTER 2: SAMPLING AND GALERKIN RECONSTRUCTION IN REPRODUCING KERNEL SPACES

In this chapter, we introduce a fidelity measure depending on a given sampling scheme and propose a Galerkin method in Banach space setting for signal reconstruction. We show that the proposed Galerkin method provides a quasi-optimal approximation, and the corresponding Galerkin equations could be solved by an iterative approximation-projection algorithm in a reproducing kernel subspace of L^p. Also we present detailed analysis and numerical simulations of the Galerkin method for reconstructing signals with finite rate of innovation.

2.1 Galerkin reconstruction in Banach spaces

In this section, we consider numerical stability and quasi-optimality of a (sub-)Galerkin reconstruction in a Banach space setting. First we introduce admissibility of operators for a trial-test space pair $(U, \tilde{U})$.

Definition 2.1. *Let (U, V, B) be a triple of Banach spaces with $U \subset V \subset B$, and let $\tilde{U} \subset B^*$. We say that a bounded linear operator $S : V \to V$ is* admissible *for the trial-test space pair $(U, \tilde{U})$ if there exist positive constants D_1 and D_2 such that*

$$\sup_{g\in\tilde{U}, \|g\|\le 1} |\langle Sf,\ g\rangle| \ge D_1\|f\| \ \text{ for all } \ f \in U, \tag{2.1}$$

and

$$\sup_{g\in\tilde{U}, \|g\|\le 1} |\langle Sf,\ g\rangle| \le D_2\|f\| \ \text{ for all } \ f \in V. \tag{2.2}$$

The above admissibility concept in a Hilbert space setting is a frame-like requirement, which was

introduced in [2, Definition 3.2]. In our model for sampling, S is the pre-reconstruction operator $S_{\Gamma,\delta}$ in (2.11), and the triple of Banach spaces contains the reconstruction space U, the reproducing kernel space $V_{K,p}$ in (1.4) and the space L^p.

Next we introduce a general notion of Galerkin reconstructions.

Definition 2.2. *Let $S: V \to V$ be a bounded linear operator, and $(U, \tilde{U})$ be a trial-test space pair. We say that a linear operator $R: V \to U$ is a* Galerkin reconstruction *if*

$$Rh = h,\ h \in U \tag{2.3}$$

and

$$\langle SRf, g\rangle = \langle Sf, g\rangle,\ f \in V \text{ and } g \in \tilde{U}; \tag{2.4}$$

and a sub-Galerkin reconstruction *if* (2.3) *holds and*

$$\sup_{g\in\tilde{U},\|g\|\le 1} |\langle SRf,\ g\rangle| \le D_3 \sup_{g\in\tilde{U},\|g\|\le 1} |\langle Sf,\ g\rangle|,\ f \in V, \tag{2.5}$$

for some $D_3 > 0$.

In the following theorem, we establish numerical stability and quasi-optimality of (sub-)Galerkin reconstructions associated with admissible operators.

Theorem 2.3. *Let $V, U, \tilde{U}$ be as in Definition 2.1, and S be admissible for the pair $(U, \tilde{U})$ with bounds D_1 and D_2. If $R: V \to U$ is a sub-Galerkin reconstruction with bound D_3, then*

(i) R is numerically stable,

$$\|Rf\| \le \frac{D_2 D_3}{D_1}\|f\|,\ f \in V;\ \text{and}$$

(ii) R *is quasi-optimal,*

$$\|Rf - f\| \leq \frac{D_1 + D_2 D_3}{D_1} \inf_{h \in U} \|f - h\|, \ f \in V.$$

Proof. (i) For $f \in V$, we obtain from (2.1), (2.2) and (2.5) that

$$D_1 \|Rf\| \leq \sup_{g \in \tilde{U}, \|g\| \leq 1} |\langle SRf, \ g\rangle| \leq D_3 \sup_{g \in \tilde{U}, \|g\| \leq 1} |\langle Sf, \ g\rangle| \leq D_2 D_3 \|f\|.$$

This proves numerical stability of the reconstruction operator R.

(ii) For $f \in V$ and $h \in U$,

$$\begin{aligned} \|f - Rf\| &\leq \|f - h\| + \|h - Rf\| \\ &= \|f - h\| + \|R(f - h)\| \leq \frac{D_1 + D_2 D_3}{D_1} \|f - h\|, \end{aligned}$$

where we have used the facts that R is a sub-Galerkin reconstruction and has numerical stability. Then quasi-optimality of the reconstruction operator R holds by taking infinimum over $h \in U$.

□

By Theorem 2.3, the existence of a quasi-optimal approximation reduces to finding a sub-Galerkin reconstruction. Now we show that such a sub-Galerkin reconstruction always exists when U and $\tilde{U}$ are finite-dimensional.

Theorem 2.4. *Let* $V, U, \tilde{U}$ *be as in Definition 2.1, and* S *be admissible for the pair* $(U, \tilde{U})$*. If* U *and* $\tilde{U}$ *are finite-dimensional, then there is a sub-Galerkin reconstruction.*

Proof. Let $\{f_i\}_{i=1}^m$ and $\{g_i\}_{i=1}^n$ be bases of U and $\widetilde{U}$ respectively. From the admissibility of S it follows immediately that $(\langle Sf_i, g_j\rangle)_{1 \leq i \leq m, 1 \leq j \leq n}$ has full rank m. Without loss of generality, we

assume that $B := (\langle Sf_i, g_j\rangle)_{1\le i,j\le m}$ is nonsingular. Let $\tilde U_*$ be the space spanned by $\{g_j\}_{j=1}^m$. By the non-singularity of the matrix B, there is a positive constant C_0 such that

$$C_0\|h\| \le \sup_{g\in\widetilde{U}_*, \|g\|\le 1} |\langle Sh,\ g\rangle|,\ h\in U. \tag{2.6}$$

Write $B^{-1} = (b_{ij})_{1\le i,j\le m}$, and define linear operator R by

$$Rf := \sum_{i,j=1}^m \langle Sf, g_i\rangle b_{ij} f_j,\ f\in V.$$

One may easily verify that R satisfies (2.3), and Rf solves the Galerkin equations

$$\langle SRf, g\rangle = \langle Sf, g\rangle,\ g\in \tilde U_* \tag{2.7}$$

for any $f\in V$. Therefore

$$\begin{aligned}
\sup_{g\in\widetilde{U}, \|g\|\le 1} |\langle SRf,\ g\rangle| &\le D_2\|Rf\| \le \frac{D_2}{C_0}\sup_{g\in\widetilde{U}_*, \|g\|\le 1} |\langle SRf,\ g\rangle| \\
&= \frac{D_2}{C_0}\sup_{g\in\widetilde{U}_*, \|g\|\le 1} |\langle Sf,\ g\rangle| \\
&\le \frac{D_2}{C_0}\sup_{g\in\widetilde{U}, \|g\|\le 1} |\langle Sf,\ g\rangle|,\ f\in V,
\end{aligned}$$

by (2.6), (2.7) and the admissibility of S. $\square$

For the case that U and $\tilde U$ have the same dimension, we have

Corollary 2.5. *Let $V, U, \tilde U$ be as in Definition 2.1, and S be admissible for the pair $(U, \tilde U)$. If dimensions of U and $\tilde U$ are the same, then for $f\in V$, the unique solution of Galerkin equations*

$$\langle SRf, g\rangle = \langle Sf, g\rangle,\ g\in\tilde U, \tag{2.8}$$

defines a Galerkin reconstruction.

In a Hilbert space setting, we can establish the following result for least squares solutions.

Corollary 2.6. *Let V be a Hilbert space, U and $\tilde{U}$ be linear subspaces of V, and let S be admissible for the pair $(U, \tilde{U})$. If U and $\tilde{U}$ are finite-dimensional, then the least squares solution of Galerkin equations* (2.8)*,*

$$Rf := \operatorname{argmin}_{h \in U} \sup_{g \in \tilde{U}, \|g\| \le 1} |\langle S(h-f), g\rangle|, \quad f \in V,$$

defines a sub-Galerkin reconstruction with bound $D_3 \le 1$.

We remark that the above conclusion on least squares solutions with $\tilde{U} = U$ has been established by Adcock, Gataric and Hansen for non-uniform sampling [1, 2].

2.2 Admissible pre-reconstruction operator in reproducing kernel spaces

In this section, we discuss admissibility of the pre-reconstruction operator $S_{\Gamma,\delta}$ in (2.11). To do so, we introduce the *residue* $E(U, F)$ of signals in a linear space $U \subset L^p$ outside a measurable set F,

$$E(U, F) := \sup_{0 \ne f \in U} \frac{\|f\|_{L^p(\mathbb{R}^d \setminus F)}}{\|f\|_p},$$

where $\|\cdot\|_{L^p(E)}$ is the p-norm on a measurable set E. The reader may refer to [1, 90, 96] for some applications of residues of bandlimited signals.

2.2.1 Reproducing kernel spaces

We are interested in reproducing kernel spaces (RKSs) of the form (1.4), which has been introduced in Section 1.1. The RKS of the form (1.4) has rich geometric structure, lots of flexibility and technical suitability for sampling. It has been used for modeling bandlimited signals, wavelet (spline) signals, and signals with finite rate of innovation [13, 116, 117, 143, 158].

To consider sampling and reconstruction in $V_{K,p}$ of the form (1.4), we *always* assume that the kernel K of the space $V_{K,p}$ satisfies

$$\|K\|_{\mathcal{W}} := \max\left\{ \sup_{x\in\mathbb{R}^d} \|K(x,\cdot)\|_1, \sup_{y\in\mathbb{R}^d} \|K(\cdot,y)\|_1 \right\} < \infty \tag{2.9}$$

and

$$\lim_{\delta\to 0} \|\omega_\delta(K)\|_{\mathcal{W}} = 0, \tag{2.10}$$

where

$$\omega_\delta(K)(x,y) := \sup_{|x'|,|y'|\le\delta} |K(x+x', y+y') - K(x,y)|.$$

Under the above hypothesis, the integral operator T_0 in (1.5) is a bounded operator on L^p,

$$\|T_0 f\|_p \le \|K\|_{\mathcal{W}} \|f\|_p, \quad f \in L^p.$$

More importantly, its range space $V_{K,p}$ is a reproducing kernel space [116]. The model space (2.26) for FRI signals to live in is a reproducing kernel space of the form (1.4) with kernel K satisfying (2.9) and (2.10), see Theorem 2.16.

2.2.2 Admissible pre-reconstruction operator $S_{\Gamma,\delta}$

For the sampling scheme (1.1) on $V_{K,p}$ in (1.4), take a disjoint covering

$$\{I_n \subset B(\gamma_n, \delta) : \ \gamma_n \in \Gamma\}$$

of $B(\Gamma, \delta) := \cup_{\gamma \in \Gamma} B(\gamma, \delta) = \cup_{\gamma\in\Gamma}\{x : \ |x-\gamma| \le \delta\}$, and define

$$S_{\Gamma,\delta} f(x) := \sum_{\gamma_n \in \Gamma} |I_n| f(\gamma_n) K(x, \gamma_n), \quad f \in V_{K,p}, \tag{2.11}$$

where $\delta > 0$. The operator $S_{\Gamma,\delta}$ depends only on the sampling scheme (1.1). We call it a *pre-reconstruction operator*, as $S_{\Gamma,\delta} f(x)$ is a good approximation to $f(x)$ when δ is sufficiently small and $x \in B(\Gamma, \delta)$ is away from the complement of $B(\Gamma, \delta)$, see Figure 1.1.

Theorem 2.7. *Let $V_{K,p}$ and $S_{\Gamma,\delta}$ be as in* (1.4) *and* (2.11) *respectively. Assume that $U \subset V_{K,p}$ and $\tilde{U} \subset L^{p/(p-1)}$. If*

$$\sup_{g \in \tilde{U}, \|g\|_{p/(p-1)} \le 1} |\langle f, g\rangle| \ge D_4 \|f\|_p, \ f \in U \tag{2.12}$$

for some constant D_4 satisfying

$$r_0 := D_4^{-1}\big(E(U, B(\Gamma,\delta))\|K\|_{\mathcal{W}} + \|\omega_\delta(K)\|_{\mathcal{W}}\big(1 + \|K\|_{\mathcal{W}} + \|\omega_\delta(K)\|_{\mathcal{W}}\big)\big) < 1, \tag{2.13}$$

then $S_{\Gamma,\delta}$ is admissible for the pair $(U, \tilde{U})$.

To prove Theorem 2.7, we need the following technical lemma.

Lemma 2.8. *Let $V_{K,p}$ and $S_{\Gamma,\delta}$ be as in* (1.4) *and* (2.11) *respectively. Then*

$$\|S_{\Gamma,\delta} f\|_p \le \big(\|K\|_{\mathcal{W}} + \|\omega_\delta(K)\|_{\mathcal{W}}\big)\big(1 + \|\omega_\delta(K)\|_{\mathcal{W}}\big)\|f\|_p, \ f \in V_{K,p}.$$

Proof. Let $\{I_n\}$ be the disjoint covering of $B(\Gamma,\delta)$ in (2.11). For $f \in V_{K,p}$, write

$$\begin{aligned}
S_{\Gamma,\delta} f(x) &= \sum_n \int_{I_n} \int_{\mathbb{R}^d} K(x,\gamma_n) K(\gamma_n, z) f(z) dz dy \\
&= \sum_n \int_{I_n} \int_{\mathbb{R}^d} \Big\{ K(x,y)K(y,z) + (K(x,\gamma_n) - K(x,y)) \\
&\quad \times K(y,z) + K(x,y)(K(\gamma_n,z) - K(y,z)) \\
&\quad + (K(x,\gamma_n) - K(x,y))(K(\gamma_n,z) - K(y,z)) \Big\} f(z) dz dy \\
&=: I + II + III + IV. \qquad (2.14)
\end{aligned}$$

Observe that

$$\|I\|_p = \Big\| \int_{B(\Gamma,\delta)} K(\cdot,y) f(y) dy \Big\|_p \le \|K\|_{\mathcal{W}} \|f\|_p,$$

$$\|II\|_p \le \Big\| \int_{\mathbb{R}^d} \omega_\delta(K)(\cdot,y) |f(y)| dy \Big\|_p \le \|\omega_\delta(K)\|_{\mathcal{W}} \|f\|_p,$$

$$\begin{aligned}
\|III\|_p &\le \Big\| \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |K(\cdot,y)| \omega_\delta(K)(y,z) |f(z)| dz dy \Big\|_p \\
&\le \|K\|_{\mathcal{W}} \|\omega_\delta(K)\|_{\mathcal{W}} \|f\|_p,
\end{aligned}$$

and

$$\begin{aligned}
\|IV\|_p &\le \Big\| \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \omega_\delta(K)(\cdot,y) \omega_\delta(K)(y,z) |f(z)| dz dy \Big\|_p \\
&\le \|\omega_\delta(K)\|_{\mathcal{W}}^2 \|f\|_p.
\end{aligned}$$

Combining the above four estimates with (2.14) completes the proof. $\square$

Next, we continue our proofs of Theorems 2.7.

Proof. The upper bound estimate (2.2) for the operator $S_{\Gamma,\delta}$ follows immediately from Lemma 2.8.

Define

$$T_0^* g(x) := \int_{\mathbb{R}^d} K(y,x) g(y) dy,\ g \in L^{p/(p-1)}.$$

For $f \in U$ and $g \in \tilde{U} \subset L^{p/(p-1)}$ with $\|g\|_{p/(p-1)} \le 1$, we obtain

$$\begin{aligned}
|\langle S_{\Gamma,\delta} f, g\rangle - \langle f, g\rangle| &\le \Big| \int_{\mathbb{R}^d \setminus B(\Gamma,\delta)} f(x) T_0^* g(x) dx \Big| \\
&\quad + \Big| \sum_n \int_{I_n} f(\gamma_n)(T_0^* g)(\gamma_n) - f(x)(T_0^* g)(x) dx \Big| \\
&\le \|K\|_{\mathcal{W}} \|f\|_{L^p(\mathbb{R}^d \setminus B(\Gamma,\delta))} \\
&\quad + \|\omega_\delta(K)\|_{\mathcal{W}} \big(1 + \|K\|_{\mathcal{W}} + \|\omega_\delta(K)\|_{\mathcal{W}}\big) \|f\|_p, \qquad (2.15)
\end{aligned}$$

where $\{I_n\}$ is the disjoint covering of $B(\Gamma,\delta)$ in (2.11). This together with (2.12) and (2.13) proves the lower bound estimate (2.1) for the operator $S_{\Gamma,\delta}$. □

Given a sampling set Γ, we say that the sampling scheme (1.1) has *weighted ℓ^p-stability on U* if there exist positive constants C_1, C_2 and δ such that

$$C_1 \|f\|_p \le \Big(\sum_{\gamma_n \in \Gamma} |I_n| |f(\gamma_n)|^p \Big)^{1/p} \le C_2 \|f\|_p,\ f \in U,$$

if $1 \le p < \infty$, and

$$C_1 \|f\|_\infty \le \sup_{\gamma_n \in \Gamma} |f(\gamma_n)| \le C_2 \|f\|_\infty,\ f \in U,$$

if $p = \infty$, where $\{I_n \subset B(\gamma_n, \delta), \gamma_n \in \Gamma\}$ is a disjoint covering of the δ-neighborhood $B(\Gamma,\delta)$ of the sampling set Γ. Weighted stability of a sampling scheme is an important concept for the robustness and uniqueness of signal reconstructions, see [13, 16, 26, 67, 116, 141, 150, 151, 158] and references here.

In the next theorem, we show that weighted stability of the sampling scheme (1.1) follows from

admissibility of the pre-reconstruction operator in (2.11).

Theorem 2.9. *Let $V_{K,p}$ and $S_{\Gamma,\delta}$ be as in* (1.4) *and* (2.11) *respectively. Assume that $U \subset V_{K,p}$ and $\tilde{U} \subset L^{p/(p-1)}$. If $S_{\Gamma,\delta}$ is admissible for the pair $(U, \tilde{U})$, then the sampling scheme* (1.1) *on Γ has weighted ℓ^p-stability on U.*

Proof. Take $f \in V$. Following the argument used in Lemma 2.8, we obtain

$$\big(\|K\|_{\mathcal{W}} + \|\omega_\delta(K)\|_{\mathcal{W}}\big)^{-1}\|S_{\Gamma,\delta}f\|_p \le \Big(\sum_n |I_n||f(\omega_n)|^p\Big)^{1/p} \le \big(1 + \|\omega_\delta(K)\|_{\mathcal{W}}\big)\|f\|_p$$

for $1 \le p < \infty$, and

$$\big(\|K\|_{\mathcal{W}} + \|\omega_\delta(K)\|_{\mathcal{W}}\big)^{-1}\|S_{\Gamma,\delta}f\|_\infty \le \sup_n |f(\omega_n)| \le \|f\|_\infty$$

for $p = \infty$. The above two estimates together with admissibility of the operator $S_{\Gamma,\delta}$ complete the proof. □

By the regularity assumption (2.10) on the reproducing kernel K, the second requirement (2.13) in Theorem 2.7 is satisfied if δ is sufficiently small and $B(\Gamma, \delta)$ is the whole Euclidean space $\mathbb{R}^d$. For the case that $B(\Gamma, \delta)$ contains an open domain F_0 but not necessarily the whole space $\mathbb{R}^d$, we obtain the following samplability result from Theorems 2.7 and 2.9.

Corollary 2.10. *Let $U \subset V_{K,p}$ and D_4 be as in Theorem 2.7. Assume that F_0 is an open domain satisfying $E(U, F_0)\|K\|_{\mathcal{W}} < D_4$. If Γ is a sampling set with $B(\Gamma, \delta) \supset F_0$ for some sufficiently small $\delta > 0$, then signals in U are uniquely determined by their samples taken on Γ.*

The samplability of various signals is well-studied, see, e.g., [2, 69, 81] for band-limited signals, [13, 158] for signals in a shift-invariant space, [143, 141] for signals with finite rate of innovation, and [87, 116] for signals in a reproducing kernel space.

2.3 Galerkin reconstruction and iterative approximation-projection algorithm

The next topic of this section is how to solve the Galerkin reconstruction (2.3) and (2.4) for signals in a reproducing kernel space (RKS) of the form (1.4).

Due to the above approximation property of the pre-reconstruction operator $S_{\Gamma,\delta}$, we propose the following iterative approximation-projection algorithm

$$g_0 \in U \ \text{ and } \ g_{m+1} = g_m - P_{U,\tilde{U}} S_{\Gamma,\delta} g_m + g_0, \ m \geq 0, \tag{2.16}$$

to solve the Galerkin reconstruction (2.4) and (2.3), where $P_{U,\tilde{U}}$ is an oblique projection for the trial-test space pair $(U, \tilde{U})$. The above algorithm is shown in Theorem 2.12 to have exponential convergence, c.f. [11, 16, 69, 116, 151].

In this section, we apply the iterative approximation-projection algorithm (2.16) to define a unique Galerkin reconstruction associated with the pre-reconstruction operator $S_{\Gamma,\delta}$.

To define the iterative approximation-projection algorithm (2.16), we recall the oblique projection for a pair $(U, \tilde{U})$ of Banach spaces.

Definition 2.11. *Given $U \subset V_{K,p}$ and $\tilde{U} \subset L^{p/(p-1)}$, a bounded operator $P_{U,\tilde{U}} : V_{K,p} \to U$ is said to be an* oblique projection *for the pair $(U, \tilde{U})$ if*

$$P_{U,\tilde{U}} h = h, \ h \in U, \tag{2.17}$$

and

$$\langle P_{U,\tilde{U}} f, g \rangle = \langle f, g \rangle, \ f \in V_{K,p}, g \in \tilde{U}. \tag{2.18}$$

In Hilbert space setting, an oblique projection $P_{U,\tilde{U}}$ exists when cosine of the subspace angle

between U and $\tilde{U}^\perp$ is positive [3, 26, 67, 153]. Following the argument used in Theorem 2.4, we can show that if U and $\tilde{U}$ have the same dimension and satisfy the first requirement (2.12) of Theorem 2.7, then there is an oblique projection $P_{U,\tilde{U}}$ for the pair $(U, \tilde{U})$.

In the next theorem, we prove that the iterative approximation-projection algorithm (2.16) associated with the oblique projection $P_{U,\tilde{U}}$ has exponential convergence, c.f. Remark 2.20.

Theorem 2.12. *Let $V_{K,p}$, $S_{\Gamma,\delta}$ and $r_0 \in (0,1)$ be as in* (1.4)*,* (2.11) *and* (2.13) *respectively. Assume that $U \subset V_{K,p}$ and $\tilde{U} \subset L^{p/(p-1)}$ satisfy* (2.12) *and* (2.13)*, and an oblique projection $P_{U,\tilde{U}}$ associated with the pair $(U, \tilde{U})$ exists. Then for any $g_0 \in U$, the sequence $g_m, m \ge 0$, in the iterative algorithm* (2.16) *converges to some $g_\infty \in U$,*

$$\|g_m - g_\infty\|_p \le \frac{r_0^{m+1}}{1-r_0}\|g_0\|_p, \quad m \ge 0. \tag{2.19}$$

Moreover, if $g_0 = P_{U,\tilde{U}} S_{\Gamma,\delta} h + \tilde{g}$ for some $h, \tilde{g} \in U$, then

$$\|g_\infty - h\|_p \le \frac{\|\tilde{g}\|_p}{1-r_0}. \tag{2.20}$$

Proof. Combining (2.12), (2.18) and (2.15), we obtain

$$\begin{aligned} \|P_{U,\tilde{U}} S_{\Gamma,\delta} f - f\|_p &\le D_4^{-1} \sup_{g\in\tilde{U}, \|g\|_{p/(p-1)}\le 1} |\langle P_{U,\tilde{U}} S_{\Gamma,\delta} f - f, g\rangle| \\ &= D_4^{-1} \sup_{g\in\tilde{U}, \|g\|_{p/(p-1)}\le 1} |\langle S_{\Gamma,\delta} f - f, g\rangle| \\ &\le r_0\|f\|_p, \quad f \in U. \end{aligned} \tag{2.21}$$

Observe from (2.16) that

$$g_{m+1} - g_m = (I - P_{U,\tilde{U}} S_{\Gamma,\delta})(g_m - g_{m-1}), \ m \ge 1.$$

This together with (2.21) proves (2.19).

Now we prove (2.20). Taking limit in (2.16) leads to the following consistence condition

$$P_{U,\tilde{U}} S_{\Gamma,\delta} g_\infty = g_0. \tag{2.22}$$

Replacing g_0 in (2.22) by $P_{U,\tilde{U}} S_{\Gamma,\delta} h + \tilde{g}$ gives

$$P_{U,\tilde{U}} S_{\Gamma,\delta}(g_\infty - h) = \tilde{g}.$$

This together with (2.21) completes the proof.

□

The algorithm (2.16) has been widely used to reconstruct various signals. The reader may refer to [69, 151] for band-limited signals, [11, 16] for signals in a shift-invariant space, and [116] for signals in a reproducing kernel space.

Applying exponential convergence of the iterative approximation-projection algorithm (2.16), we can define a unique Galerkin reconstruction.

Theorem 2.13. *Let $V_{K,p}, S_{\Gamma,\delta}, U, \tilde{U}$ and $P_{U,\tilde{U}}$ be as in Theorem 2.12. Then Galerkin equations*

$$\langle S_{\Gamma,\delta} h, g\rangle = \langle S_{\Gamma,\delta} f, g\rangle,\ g \in \tilde{U}, \tag{2.23}$$

have a unique solution $h \in U$ for $f \in V_{K,p}$. Moreover, the mapping $f \to h$ defines a Galerkin reconstruction.

Proof. Take $f \in V_{K,p}$, set $g_0 = P_{U,\tilde{U}} S_{\Gamma,\delta} f$, and let $g_\infty \in U$ be the limit of $g_m, m \geq 0$, in the

iterative algorithm (2.16). The existence of such a limit follows from Theorem 2.12. Taking limit in (2.16) leads to

$$P_{U,\tilde{U}} S_{\Gamma,\delta} f = P_{U,\tilde{U}} S_{\Gamma,\delta} g_\infty. \tag{2.24}$$

Then for any $g \in \tilde{U}$,

$$\langle S_{\Gamma,\delta} g_\infty, g\rangle = \langle P_{U,\tilde{U}} S_{\Gamma,\delta} g_\infty, g\rangle = \langle P_{U,\tilde{U}} S_{\Gamma,\delta} f, g\rangle = \langle S_{\Gamma,\delta} f, g\rangle \tag{2.25}$$

by (2.18) and (2.24). This proves that g_∞ is a solution of Galerkin equations (2.23).

Next, we show that g_∞ is the unique solution of Galerkin equations (2.23). Let $h \in U$ be another solution. Then

$$\langle P_{U,\tilde{U}} S_{\Gamma,\delta}(h - g_\infty), g\rangle = \langle S_{\Gamma,\delta}(h - g_\infty), g\rangle = 0.$$

This together with (2.12) implies that

$$P_{U,\tilde{U}} S_{\Gamma,\delta}(h - g_\infty) = 0.$$

Recall from (2.21) that $P_{U,\tilde{U}} S_{\Gamma,\delta}$ is invertible on U. Then $h = g_\infty$ and the uniqueness follows.

Observe that any $f \in U$ satisfies Galerkin equations (2.23). This together with (2.25) proves that the unique solution of Galerkin equations (2.23) defines a Galerkin reconstruction.

□

We finish this section with a remark on the iterative approximation-projection algorithm (2.16).

Remark 2.14. Given $\delta > 0$, a sampling set Γ and probability measures μ_n supported on I_n, we

define

$$\tilde{S}_{\Gamma,\delta} f(x) = \sum_{\gamma_n \in \Gamma} |I_n| f(\gamma_n) \int_{I_n} K(x,y) d\mu_n(y), \ \ f \in V_{K,p},$$

where $\{I_n \subset B(\gamma, \delta), \ \gamma_n \in \Gamma\}$ is a disjoint covering of $B(\Gamma, \delta)$. The operator $\tilde{S}_{\Gamma,\delta}$ just defined becomes the pre-reconstruction operator $S_{\Gamma,\delta}$ in (2.11) when μ_n are point measures supported on γ_n, and the pre-reconstruction operator

$$S_{\Gamma,\delta} f(x) = \sum_{\omega_n \in \Gamma} f(\gamma_n) \int_{I_n} K(x,y) dy, \ \ f \in V_{K,p}$$

when μ_n are normalized Lebsegue measure supported on I_n. Following the argument used in Theorem 2.7 and Theorem 2.12, we can show that the approximation-projection algorithm (2.16) with $S_{\Gamma,\delta}$ replaced by $\tilde{S}_{\Gamma,\delta}$ has exponential convergence if

$$D_4^{-1}\big(E(U, B(\Gamma,\delta))\|K\|_{\mathcal{W}} + \|\omega_{2\delta}(K)\|_{\mathcal{W}}\big(1 + \|K\|_{\mathcal{W}} + \|\omega_{2\delta}(K)\|_{\mathcal{W}}\big)\big) < 1,$$

c.f., the second requirement (2.13) in Theorem 2.7.

2.4 Sampling signals with finite rate of innovation

A signal with *finite rate of innovation* (FRI) has finitely many degrees of freedom per unit of time [62, 115, 121, 143, 141, 159]. Define the *Wiener amalgam space* by

$$\mathcal{W}^1 := \Big\{\phi, \ \|\phi\|_{\mathcal{W}^1} := \sum_{k \in \mathbb{Z}} \sup_{0 \le x \le 1} |\phi(x+k)| < \infty\Big\}.$$

It is observed in [143] that lots of FRI signals live in a space of the form

$$V_2(\Phi) := \Big\{ \sum_{i\in\mathbb{Z}} c_i\phi_i(\cdot - i),\ \sum_{i\in\mathbb{Z}} |c_i|^2 < \infty \Big\}, \tag{2.26}$$

where the generator $\Phi := (\phi_i)_{i\in\mathbb{Z}}$ satisfies

$$\|\Phi\|_{\mathcal{W}^1} := \big\| \sup_{i\in\mathbb{Z}} |\phi_i| \big\|_{\mathcal{W}^1} < \infty \ \text{ and } \ \lim_{\delta\to 0} \big\| \sup_{i\in\mathbb{Z}} \omega_\delta(\phi_i) \big\|_{\mathcal{W}^1} = 0. \tag{2.27}$$

For $\Phi := (\phi_i)_{i\in\mathbb{Z}}$ and $\tilde\Phi := (\tilde\phi_j)_{j\in\mathbb{Z}}$ satisfying (2.27), define their correlation matrix by

$$A_{\Phi,\tilde\Phi} := \big(\langle \phi_i(\cdot - i), \tilde\phi_j(\cdot - j)\rangle\big)_{i,j\in\mathbb{Z}}. \tag{2.28}$$

In this section, we always assume that $A_{\Phi,\tilde\Phi}$ has bounded inverse on ℓ^2. Write $(A_{\Phi,\tilde\Phi})^{-1} = (b_{ij})_{i,j\in\mathbb{Z}}$. Applying Wiener's lemma for Baskakov-Gohberg-Sjöstrand class, one may verify that the space $V_2(\Phi)$ for FRI signals to live in is the range space $V_{K_{\Phi,\tilde\Phi},2}$ of an idempotent integral operator with kernel

$$K_{\Phi,\tilde\Phi}(x,y) := \sum_{i,j\in\mathbb{Z}} \phi_i(x-i) b_{ji} \tilde\phi_j(y-j) \tag{2.29}$$

satisfying (2.9) and (2.10), see Theorem 2.16.

Let $\mathcal{C}_1$ contain all infinite matrices $A := \big(a_{ij}\big)_{i,j\in\mathbb{Z}}$ with

$$\|A\|_{\mathcal{C}_1} := \sum_{k\in\mathbb{Z}} \Big(\sup_{i-j=k} |a_{ij}| \Big) < \infty.$$

To prove Theorem 2.16, we recall Wiener's lemma for the Baskakov-Gohberg-Sjöstrand class $\mathcal{C}_1$, see [24, 77, 83, 137, 142, 147] and references therein.

Lemma 2.15. *If $A \in \mathcal{C}_1$ has bounded inverse on ℓ^2, then its inverse A^{-1} belongs to $\mathcal{C}_1$ too.*

Theorem 2.16. *Let Φ and $\tilde{\Phi}$ satisfy* (2.27)*, and the correlation matrix $A_{\Phi,\tilde{\Phi}}$ in* (2.28) *have bounded inverse on ℓ^2. Then*

$$V_2(\Phi) = V_{K_{\Phi,\tilde{\Phi}},2}$$

for the kernel $K_{\Phi,\tilde{\Phi}}$ in (2.29)*, which satisfies* (2.9) *and* (2.10)*.*

Proof. By direct calculation, we have

$$\|A_{\Phi,\tilde{\Phi}}\|_{\mathcal{C}_1} \le \|\Phi\|_{\mathcal{W}^1}\|\tilde{\Phi}\|_{\mathcal{W}^1}.$$

Thus the inverse of the correlation matrix $A_{\Phi,\tilde{\Phi}}$ belongs to the Baskakov-Gohberg-Sjöstrand class by Lemma 2.15. One may then verify immediately that the kernel $K_{\Phi,\tilde{\Phi}}$ in (2.29) satisfies all requirements of the theorem. $\square$

Given a sampling set $\Gamma = \{\gamma_n\}_{n=1}^N$ ordered as $\gamma_1 < \gamma_2 < \cdots < \gamma_N$, define

$$S_{\Phi,\tilde{\Phi},\Gamma} f(x) := \sum_{n=1}^{N} \frac{\gamma_{n+1}-\gamma_{n-1}}{2} f(\gamma_n) K_{\Phi,\tilde{\Phi}}(x,\gamma_n),\ f \in V_2(\Phi), \tag{2.30}$$

where $\gamma_0 = \gamma_1$ and $\gamma_{N+1} = \gamma_N$. In the next theorem, we establish the equivalence between admissibility of the operator $S_{\Phi,\tilde{\Phi},\Gamma}$ and its corresponding Galerkin reconstruction in a finite-dimensional space, c.f. Corollary 2.5, and Theorems 2.7 and 2.13.

Theorem 2.17. *For $L \ge 1$, define*

$$V_{2,L}(\Phi) := \Big\{ \sum_{i=-L}^{L} c_i \phi_i(\cdot - i),\ \sum_{i=-L}^{L} |c_i|^2 < \infty \Big\} \tag{2.31}$$

and

$$V_{2,L}(\tilde{\Phi}) := \Big\{ \sum_{i=-L}^{L} d_i \tilde{\phi}_i(\cdot - i),\ \sum_{i=-L}^{L} |d_i|^2 < \infty \Big\}. \tag{2.32}$$

Assume that Φ, $\tilde{\Phi}$ satisfy (2.27)*, and the correlation matrix $A_{\Phi,\tilde{\Phi}}$ in* (2.28) *has bounded inverse on ℓ^2. Then the following statements are equivalent:*

(i) The $L \times L$ matrix

$$A_{\Phi,\tilde{\Phi},\Gamma} := \Big(\sum_{n=1}^{N} \frac{\gamma_{n+1} - \gamma_{n-1}}{2} \phi_i(\gamma_n - i) \tilde{\phi}_j(\gamma_n - j) \Big)_{-L \le i,j \le L} \tag{2.33}$$

is nonsingular.

(ii) $S_{\Phi,\tilde{\Phi},\Gamma}$ is admissible for the pair $(V_{2,L}(\Phi), V_{2,L}(\tilde{\Phi}))$.

(iii) For any $f \in V_2(\Phi)$, Galerkin equations

$$\langle S_{\Phi,\tilde{\Phi},\Gamma} h, g\rangle = \langle S_{\Phi,\tilde{\Phi},\Gamma} f, g\rangle,\ \ g \in V_{2,L}(\tilde{\Phi}) \tag{2.34}$$

have a unique solution h in $V_{2,L}(\Phi)$.

(iv) For any $g \in V_2(\tilde{\Phi})$, dual Galerkin equations

$$\langle S_{\Phi,\tilde{\Phi},\Gamma} f, \tilde{h}\rangle = \langle S_{\Phi,\tilde{\Phi},\Gamma} f, g\rangle,\ \ f \in V_{2,L}(\Phi)$$

have a unique solution $\tilde{h}$ in $V_{2,L}(\tilde{\Phi})$.

Proof. For $h = \sum_{i=-L}^{L} c_i \phi_i(\cdot - i) \in V_{2,L}(\Phi)$ and $g = \sum_{j=-L}^{L} d_j \tilde{\phi}_j(\cdot - j) \in V_{2,L}(\tilde{\Phi})$, we obtain

$$\begin{aligned}
\langle S_{\Phi,\tilde{\Phi},\Gamma} h, g\rangle &= \sum_{i,j=-L}^{L} \Big(\sum_{n=1}^{N} \frac{\gamma_{n+1} - \gamma_{n-1}}{2} \phi_i(\gamma_n - i) \langle K_{\Phi,\tilde{\Phi}}(t, \gamma_n), \tilde{\phi}_j(t - j)\rangle \Big) c_i d_j \\
&= \sum_{i,j=-L}^{L} \Big(\sum_{n=1}^{N} \frac{\gamma_{n+1} - \gamma_{n-1}}{2} \phi_i(\gamma_n - i) \tilde{\phi}_j(\gamma_n - j) \Big) c_i d_j \\
&= c^T A_{\Phi,\tilde{\Phi},\Gamma} d, \qquad (2.35)
\end{aligned}$$

where $c=(c_i)_{-L\le i\le L}$ and $d=(d_j)_{-L\le j\le L}$. By the invertibility assumption on $A_{\Phi,\tilde\Phi}$, $\{\phi_i(\cdot-i), -L\le i\le L\}$ and $\{\tilde\phi_i(\cdot-i), -L\le i\le L\}$ are Riesz bases of $V_{2,L}(\Phi)$ and $V_{2,L}(\tilde\Phi)$ respectively. This together with (2.35) proves the desired equivalent statements. $\square$

To solve the Galerkin equations (2.34) by the iterative approximation-projection algorithm (2.16), we need an oblique projection for the pair $(V_{2,L}(\Phi), V_{2,L}(\tilde\Phi))$.

Theorem 2.18. *Let $L\ge 1$, and let Φ and $\tilde\Phi$ satisfy* (2.27). *Assume that the correlation matrix $A_{\Phi,\tilde\Phi}$ in* (2.28) *has bounded inverse on ℓ^2. Then the principal submatrix*

$$A_{\Phi,\tilde\Phi,L} := \big(\langle \phi_i(\cdot-i), \tilde\phi_j(\cdot-j)\rangle\big)_{-L\le i,j\le L} \tag{2.36}$$

of the correlation matrix $A_{\Phi,\tilde\Phi}$ is nonsingular if and only if there exists a unique oblique projection for the pair $(V_{2,L}(\Phi), V_{2,L}(\tilde\Phi))$. Moreover, the oblique projection could be defined by

$$P_{\Phi,\tilde\Phi,L}f := \sum_{-L\le i,j\le L} \langle f, \tilde\phi_i(\cdot-i)\rangle \tilde b_{ij}\phi_j(\cdot-j),\ f\in V_2(\Phi), \tag{2.37}$$

where $(A_{\Phi,\tilde\Phi,L})^{-1}=(\tilde b_{ij})_{-L\le i,j\le L}$.

Proof. The sufficiency is obvious. Now we prove the necessity. Suppose, to the contrary, that $A_{\Phi,\tilde\Phi,L}$ in (2.36) is singular. Take a nonzero vector $e=(e_i)_{-L\le i\le L}$ in the null space $N((A_{\Phi,\tilde\Phi,L})^T)$ and a nonzero linear functional $\mathcal{J}$ on $V_2(\Phi)$ such that $\mathcal{J}(h)=0$ for all $h\in V_{2,L}(\Phi)$. Define

$$Q(f) := \mathcal{J}(f)\sum_{-L\le i\le L} e_i\phi_i(\cdot-i),\ f\in V_2(\Phi).$$

Then Q is a nonzero linear operator from $V_2(\Phi)$ to $V_{2,L}(\Phi)$,

$$Qh=0,\quad h\in V_{2,L}(\Phi)$$

and

$$\langle Qf, g\rangle = \mathcal{J}(f) \sum_{-L\le i,j\le L} e_i \langle \phi_i(\cdot - i), \tilde{\phi}_j(\cdot - j)\rangle d_j = 0,$$

where $g = \sum_{-L\le j\le L} d_j \tilde{\phi}_j(\cdot - j) \in V_{2,L}(\tilde{\Phi})$. This contradicts to the uniqueness of oblique projections. $\square$

We conclude this section by examining exponential convergence of an iterative algorithm for the recovery of signals with finite rate of innovation. Replacing $P_{U,\tilde{U}}$ and $S_{\Gamma,\delta}$ in the iterative algorithm (2.16) by $P_{\Phi,\tilde{\Phi},L}$ and $S_{\Phi,\tilde{\Phi},\Gamma}$ respectively, it becomes

$$g_{m+1} = g_m - \sum_{n=1}^{N} \sum_{i,j=-L}^{L} \frac{\gamma_{n+1} - \gamma_{n-1}}{2} g_m(\gamma_n) \tilde{\phi}_i(\gamma_n - i) \tilde{b}_{ij} \phi_j(\cdot - j) + g_0,\ m \ge 0, \tag{2.38}$$

with $g_0 \in V_{2,L}(\Phi)$. The above iterative algorithm has exponential convergence when

$$\|A_{\Phi,\tilde{\Phi},\Gamma}(A_{\Phi,\tilde{\Phi},L})^{-1} - I\| < 1. \tag{2.39}$$

Theorem 2.19. *Let* Φ *and* $\tilde{\Phi}$ *satisfy* (2.27)*. Assume that* $A_{\Phi,\tilde{\Phi},L}$ *is nonsingular. If* (2.39) *holds, then the iterative algorithm* (2.38) *has exponential convergence. Moreover, it recovers the original signal* $h \in V_{2,L}(\Phi)$ *when*

$$g_0 = \sum_{n=1}^{N} \sum_{i,j=-L}^{L} \frac{\gamma_{n+1} - \gamma_{n-1}}{2} h(\gamma_n) \tilde{\phi}_i(\gamma_n - i) \tilde{b}_{ij} \phi_j(\cdot - j).$$

Proof. Write $g_m = \sum_{-L\le i\le L} c_m(i)\phi_i(\cdot - i)$ and set $c_m = (c_m(i))_{-L\le i\le L}$. Then we can reformulate the iterative algorithm (2.38) as

$$c_{m+1}^T = c_m^T - c_m^T A_{\Phi,\tilde{\Phi},\Gamma}(A_{\Phi,\tilde{\Phi},L})^{-1} + c_0^T,\ m \ge 0.$$

This together with (2.39) proves the desired conclusions. $\square$

2.5 Numerical simulations

In this section, we present several examples to illustrate our Galerkin reconstruction of signals with finite rate of innovation.

Let $\Theta := \{\theta_i\}$ be either $\Theta_O := \{0\}$ (the identical zero set), or Θ_I with θ_i being randomly selected in $[-0.2, 0.2]$. Set

$$\Phi_0 = \{\phi_0(\cdot - \theta_i)\}_{i\in\mathbb{Z}},$$

where the generating function ϕ_0 is either (i) the sinc function $\mathrm{sinc}(t) := \frac{\sin \pi t}{\pi t}$, or (ii) the Gaussian function $\mathrm{gauss}(t) := \exp(-3t^2/2)$, or (iii) the cubic B-spline $\mathrm{spline}(t)$, see Figure 2.1 for examples of signals in $V_2(\Phi_0)$.

Plotted in Figure 2.1 are bandlimited signals $x(\mathrm{sinc}, 0) = \sum_i \alpha_i \mathrm{sinc}(t-i)$ with $(1+|i|)\alpha_i \in [-1, 1]$ randomly selected (top left), $x(\mathrm{sinc}, 1) = \sum_i \beta_i \mathrm{sinc}(t - i)$ with $\beta_i = (1 + |i|)^{-1} \cos(\pi i/8)$ (top right), $x(\mathrm{sinc}, 2) = \sum_i \alpha_i \mathrm{sinc}(t - i - \theta_i)$ (bottom left) and $x(\mathrm{sinc}, 3) = \sum_i \beta_i \mathrm{sinc}(t - i - \theta_i)$ with $\theta_i \in [-0.2, 0.2]$ randomly selected (bottom right).

In our numerical simulations, reconstructed signals live in the space

$$V_{2,L}(\Phi_0) = \Big\{ \sum_{i=-L}^{L} c_i\phi_0(t - i - \theta_i) : \sum_{i=-L}^{L} |c_i|^2 < \infty\Big\},\ L \geq 1,$$

and sampling schemes are

- Nonuniform sampling on $\Gamma_N := \{\gamma_k, |k| \leq L+2\}$, where $\gamma_{-L-3} = -L-2$ and $\gamma_k - \gamma_{k-1} \in$

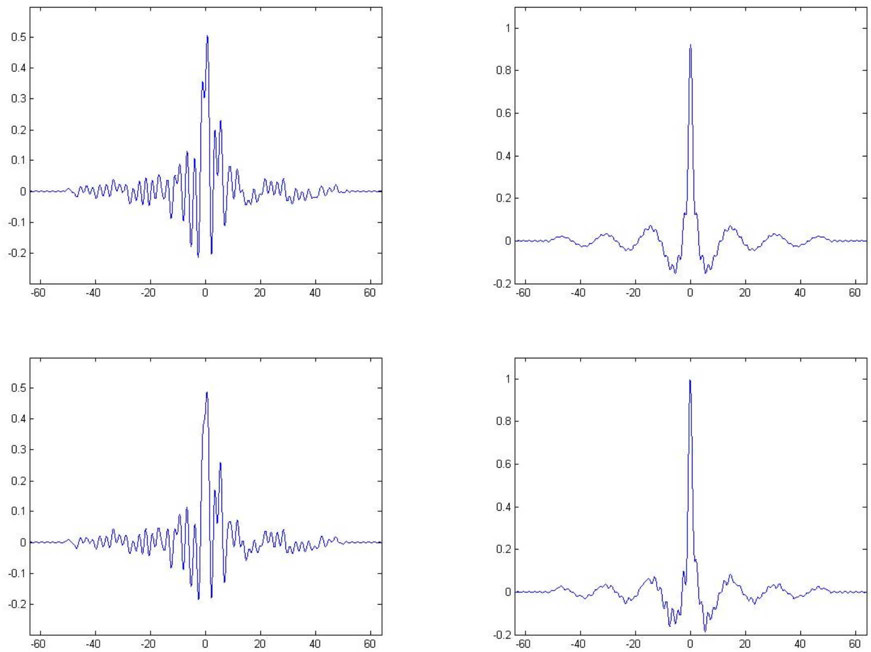

Figure 2.1: Original bandlimited signals

$[0.9, 1.1], |k| \leq L + 2$, are randomly selected.

- Jittered sampling on $\Gamma_J := \{\gamma_k := k+\delta_k, |k| \leq L+2\}$, where $\delta_k \in [-0.1, 0.1]$ are randomly selected.

- Adaptive sampling on $\Gamma_C := \{\gamma_k \in [-L-2, L+2]\}$ of a bounded signal $x \in V_2(\Phi)$ via crossing time encoding machine (C-TEM), where $x(t) \neq \|x\|_\infty \sin(\pi t)$ for all $t \in [-L-2, L+2]$ except $t = \gamma_k$ for some k, see Figure 2.2 [71, 78, 104].

Plotted in top of Figure 2.2 is the signal $x(\text{sinc}, 0)$ in Figure 2.1 and the crossing signal $\|x(\text{sinc}, 0)\|_\infty \sin \pi t$ on $[-L-2, L+2]$, and the bottom part of Figure 2.2 is the sampling data of $x(\text{sinc}, 0)$ on the sampling set $\Gamma_C \subset [-L-2, L+2]$, where $L = 30$.

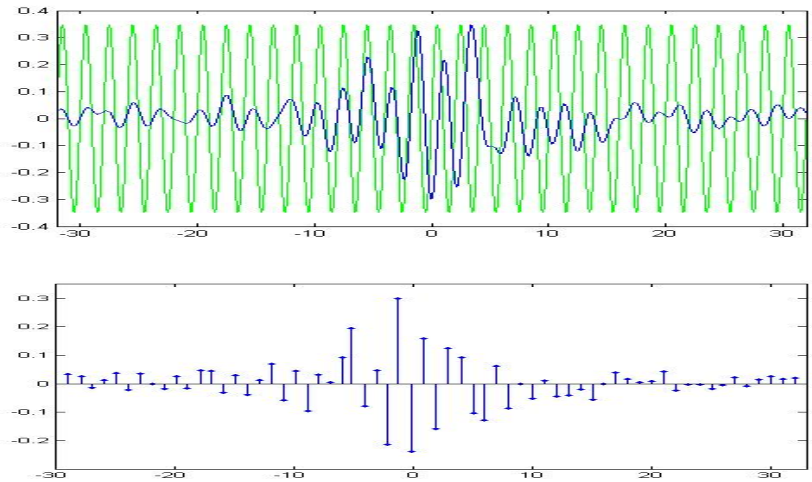

Figure 2.2: Signal $x(\mathrm{sinc}, 0)$ and sample data via C-TEM

To reconstruct signals via our Galerkin method, take

$$\tilde{\Phi}_0 = \{\tilde{\phi}_0\} \quad \text{with} \quad \tilde{\phi}_0 = \chi_{[-1/2,1/2)}.$$

Then the equation (2.34) to determine the Galerkin reconstruction

$$G_{\Phi_0,\tilde{\Phi}_0,\Gamma} f := \sum_{i=-L}^{L} c_i \phi_0(\cdot - i - \theta_i) \in V_{2,L}(\Phi_0)$$

can be reformulated as follows:

$$\begin{aligned} & \sum_{i=-L}^{L} \Big(\sum_{n=1}^{N} \frac{\gamma_{n+1} - \gamma_{n-1}}{2} \phi_0(\gamma_n - i - \theta_i) \tilde{\phi}_0(\gamma_n - j) \Big) c_i \\ = & \sum_{n=1}^{N} \frac{\gamma_{n+1} - \gamma_{n-1}}{2} f(\gamma_n) \tilde{\phi}_0(\gamma_n - j), -L \le j \le L, \end{aligned} \tag{2.40}$$

where $f \in V_2(\Phi_0)$ and $\Gamma := \{\gamma_n\}_{n=1}^{N}$ is either the nonuniform sampling set Γ_N, or the jittered sampling set Γ_J, or the adaptive C-TEM sampling set Γ_C. Considering the bandlimited signal

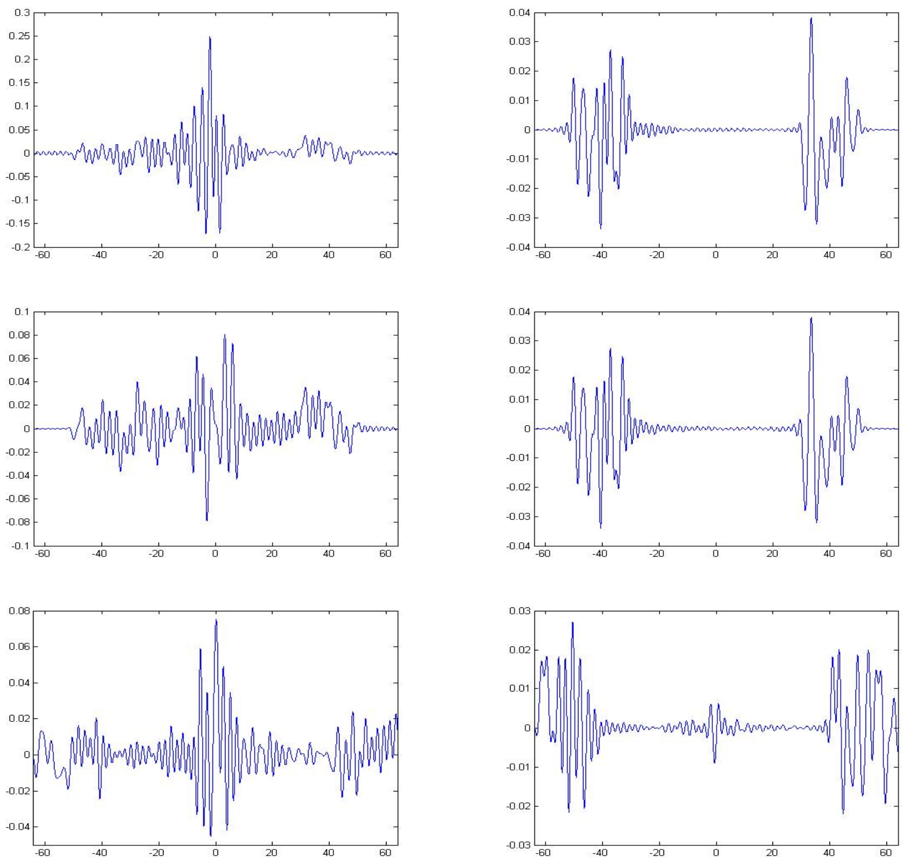

Figure 2.3: Comparison between pre-reconstructed signal and Galerkin reconstruction.

$x(\mathrm{sinc}, 0)$ described in Figure 2.1, we present some numerical results for its pre-reconstruction in $V_2(\Phi_0)$ and Galerkin reconstruction in $V_{2,L}(\Phi_0)$ in Figure 2.3. We see that a pre-reconstruction may provide a reasonable approximation, while a Galerkin reconstruction could recover the original signal almost perfectly in the sampling interval. In Figure 2.3, plotted on the top left is the difference between the signal $x(\mathrm{sinc}, 0)$ in Figure 2.1 and its pre-reconstructed signal $S_{\Phi_0,\tilde{\Phi}_0,\Gamma_N} x(\mathrm{sinc}, 0)$, while on the top right is the difference between $x(\mathrm{sinc}, 0)$ and its Galerkin reconstruction $G_{\Phi_0,\tilde{\Phi}_0,\Gamma_N} x(\mathrm{sinc}, 0)$. Shown in the middle are differences $x(\mathrm{sinc}, 0) - S_{\Phi_0,\tilde{\Phi}_0,\Gamma_J} x(\mathrm{sinc}, 0)$ (left) and $x(\mathrm{sinc}, 0) - G_{\Phi_0,\tilde{\Phi}_0,\Gamma_J} x(\mathrm{sinc}, 0)$ (right) associated with jittered sampling. Listed below are differ-

ences $x(\mathrm{sinc}, 0) - S_{\Phi_0,\tilde{\Phi}_0,\Gamma_C} x(\mathrm{sinc}, 0)$ (left) and $x(\mathrm{sinc}, 0) - G_{\Phi_0,\tilde{\Phi}_0,\Gamma_C} x(\mathrm{sinc}, 0)$ (right) associated with adaptive C-TEM sampling.

For $\Phi_0 = \{\phi_0(\cdot - \theta_i)\}$, let signals $x(\phi_0, l) \in V_2(\Phi_0), 0 \le l \le 3$, be as $x(\mathrm{sinc}, l)$ in Figure 2.1 with the sinc function replaced by the function ϕ_0. In Figure 2.4, we illustrate their best approximation in $V_{2,L}(\Phi_0)$ and solutions of the Galerkin system (2.40) with f replaced by $x(\phi_0, l), 0 \le l \le 3$, respectively. We observe that given a signal in $V_2(\Phi_0)$, its Galerkin reconstruction in $V_{2,L}(\Phi_0)$ could almost match its best approximation in $V_{2,L}(\Phi_0)$, except near the boundary of the sampling interval. The boundary effect is viewable especially when ϕ_0 has slow decay at infinity.

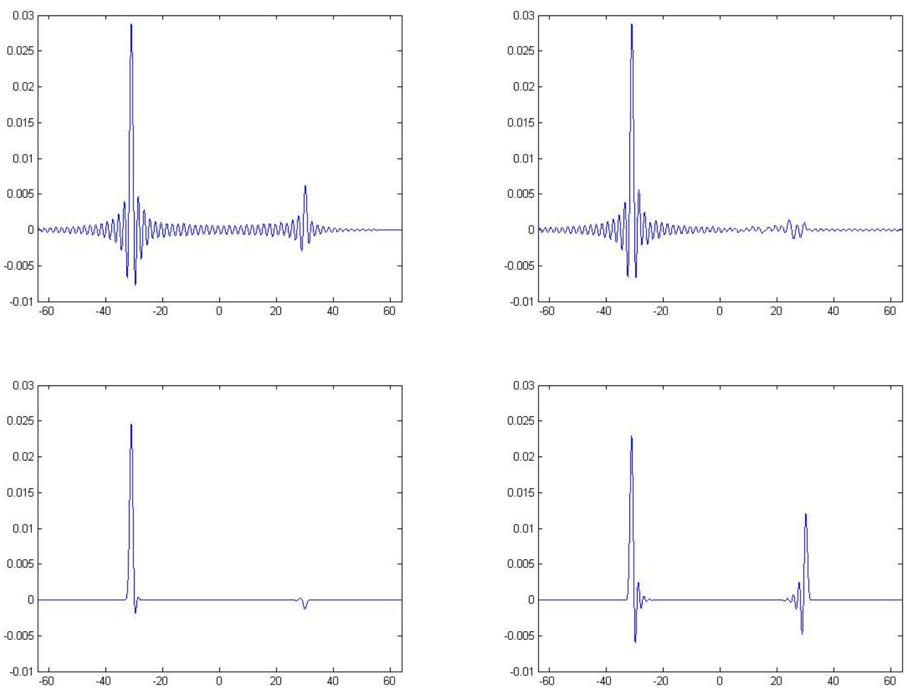

Figure 2.4: Comparison of best approximation and Galerkin reconstructions associated with operators $S_{\Phi_0,\tilde{\Phi}_0,\Gamma}$

Presented in Figure 2.4 are differences between best approximations of signals $x(\phi_0, 0)$ in $V_{2,30}(\Phi_0)$

and their Galerkin reconstructions associated with operators $S_{\Phi_0,\tilde{\Phi}_0,\Gamma}$, where on the above, $\phi_0 = \mathrm{sinc}$, $\Gamma = \Gamma_N$(left) and $\Gamma = \Gamma_J$ (right), while on the bottom $\Gamma = \Gamma_N$, $\phi_0 = \mathrm{gauss}$ (left) and $\phi_0 = \mathrm{spline}$ (right).

Given signals $x(\phi_0, l), 0 \le l \le 3$, let $y_L(\phi_0, l)$ be their best approximators in $V_{2,L}(\Phi_0)$, and denote by

$$e(\phi_0, l) = \|x(\phi_0, l) - y_L(\phi_0, l)\|$$

their best approximation error in $V_{2,L}(\Phi_0)$. For $\Gamma = \Gamma_N$ or Γ_J or Γ_C, set

$$\epsilon_\Gamma(\phi_0, l) = \|z_L(\Gamma, \phi_0, l) - y_L(\phi_0, l)\|,$$

where $z_L(\Gamma, \phi_0, l)$ is obtained from solving Galerkin system (2.40) with f replaced by $x(\phi_0, l)$. For signals $x(\phi_0, l), 0 \le l \le 3$, and sampling sets $\Gamma = \Gamma_N, \Gamma_J$ and Γ_C, Galerkin reconstruction (2.40) provides quasi-optimal approximation in $V_{2,L}(\Phi_0)$, and the quasi-optimal constant in Theorem 2.3 is well behaved,

$$\frac{\|z_L(\Gamma, \phi_0, l) - x(\phi_0, l)\|}{\|y_L(\phi_0, l) - x(\phi_0, l)\|} \le 1 + \frac{\epsilon_\Gamma(\phi_0, l)}{e(\phi_0, l)} \le \frac{3}{2},$$

see Table 2.1 for numerical results with abbrievated notations.

Table 2.1: Quasi-optimality of Galerkin reconstructions for bandlimited/Gauss/spline signals

L	10	15	20	25	30
$e(\mathrm{sinc}, 0)$	0.2176	0.1711	0.1388	0.1166	0.1024
$\epsilon_N(\mathrm{sinc}, 0)$	0.0795	0.0668	0.0197	0.0201	0.0294
$\epsilon_J(\mathrm{sinc}, 0)$	0.0770	0.0668	0.0201	0.0214	0.0290
$\epsilon_C(\mathrm{sinc}, 0)$	0.0789	0.0715	0.0239	0.0263	0.0325
$e(\mathrm{sinc}, 1)$	0.2600	0.2124	0.1816	0.1457	0.1303
$\epsilon_N(\mathrm{sinc}, 1)$	0.0344	0.0809	0.0370	0.0294	0.0431
$\epsilon_J(\mathrm{sinc}, 1)$	0.0353	0.0806	0.0372	0.0301	0.0433
$\epsilon_C(\mathrm{sinc}, 1)$	0.0363	0.0831	0.0379	0.0319	0.0442
$e(\mathrm{sinc}, 2)$	0.2095	0.1703	0.1365	0.1167	0.1007
$\epsilon_N(\mathrm{sinc}, 2)$	0.0619	0.0618	0.0256	0.0163	0.0281
$\epsilon_J(\mathrm{sinc}, 2)$	0.0596	0.0618	0.0260	0.0177	0.0275
$\epsilon_C(\mathrm{sinc}, 2)$	0.0608	0.0664	0.0284	0.0226	0.0308
$e(\mathrm{sinc}, 3)$	0.2655	0.2180	0.1863	0.1477	0.1322
$\epsilon_N(\mathrm{sinc}, 3)$	0.0461	0.0810	0.0374	0.0258	0.0406
$\epsilon_J(\mathrm{sinc}, 3)$	0.0446	0.0809	0.0375	0.0265	0.0401
$\epsilon_C(\mathrm{sinc}, 3)$	0.0474	0.0837	0.0392	0.0298	0.0418
$e(\mathrm{gauss}, 0)$	0.2055	0.1682	0.1398	0.1250	0.1086
$\epsilon_N(\mathrm{gauss}, 0)$	0.0437	0.0515	0.0270	0.0158	0.0093
$\epsilon_J(\mathrm{gauss}, 0)$	0.0439	0.0523	0.0259	0.0160	0.0096
$\epsilon_C(\mathrm{gauss}, 0)$	0.0433	0.0527	0.0270	0.0181	0.0108
$e(\mathrm{spline}, 0)$	0.1482	0.1325	0.1110	0.0924	0.0664
$\epsilon_N(\mathrm{spline}, 0)$	0.0405	0.0298	0.0204	0.0266	0.0176
$\epsilon_J(\mathrm{spline}, 0)$	0.0403	0.0299	0.0204	0.0281	0.0184
$\epsilon_C(\mathrm{spline}, 0)$	0.0407	0.0292	0.0209	0.0279	0.0181

Numerical stability of Galerkin reconstruction (2.40) could be reflected by the condition number $\mathrm{cond}_{\Gamma,\Theta}(\phi_0)$ of the square matrix

$$A_{\Phi_0,\tilde{\Phi}_0,\Gamma} = \Big(\sum_{n=1}^{N} \frac{\gamma_{n+1}-\gamma_{n-1}}{2}\phi_0(\gamma_n - i - \theta_i)\tilde{\phi}_0(\gamma_n - j)\Big)_{-L\le i,j\le L}.$$

Some numerical results of condition numbers $\mathrm{cond}_{\Gamma,\Theta}(\phi_0)$ with $\Gamma = \Gamma_N$ or Γ_J, and $\Theta = \Theta_O$ or Θ_I, are presented in Table 2.2 with abbreviated notations.

Table 2.2: Stability of Galerkin reconstructions for nonuniform/jittered sampling

L	10	15	20	25	30
$\mathrm{cond}_{N,O}(\mathrm{sinc})$	1.2059	1.2367	1.3458	1.4273	1.2904
$\mathrm{cond}_{N,I}(\mathrm{sinc})$	1.9190	1.8946	1.9828	2.0635	2.0421
$\mathrm{cond}_{N,O}(\mathrm{gauss})$	3.0162	2.7000	2.7908	3.3314	2.8362
$\mathrm{cond}_{N,I}(\mathrm{gauss})$	3.2850	3.1447	3.1421	4.0283	3.4391
$\mathrm{cond}_{N,O}(\mathrm{spline})$	3.7677	3.7534	3.0534	3.1400	4.1708
$\mathrm{cond}_{N,I}(\mathrm{spline})$	4.4768	5.2417	3.3507	3.5354	5.0292
$\mathrm{cond}_{J,O}(\mathrm{sinc})$	1.3737	1.4164	1.4105	1.4149	1.3763
$\mathrm{cond}_{J,I}(\mathrm{sinc})$	1.9723	1.9351	2.3328	2.2037	2.1744
$\mathrm{cond}_{J,O}(\mathrm{gauss})$	2.7066	2.7074	2.6936	2.6957	2.7190
$\mathrm{cond}_{J,I}(\mathrm{gauss})$	3.0847	3.1591	3.0696	3.0197	3.0878
$\mathrm{cond}_{J,O}(\mathrm{spline})$	3.1052	3.2109	3.2218	3.3257	3.2331
$\mathrm{cond}_{J,I}(\mathrm{spline})$	3.5570	3.7388	3.7140	3.9172	4.1830

For the robust (sub-)Galerkin reconstruction, the generating function $\tilde{\phi}_0$ of the test space $V_{2,L}(\tilde{\Phi}_0)$ should be so chosen that the corresponding matrice $A_{\Phi_0,\tilde{\Phi}_0,\Gamma}$ is well-conditioned, c.f. Theorem 2.3.

We conclude this sections with two more remarks.

Remark 2.20. The iterative approximation-projection algorithm (2.38) could have better performance on solving Galerkin equations (2.40), especially while matrices $A_{\Phi_0,\tilde{\Phi}_0,\Gamma}$ have large condition number, which is the case when the sampling set Γ and/or the shifting set Θ are not chosen appropriately.

Remark 2.21. For the admissibility of the pre-reconstruction operator $S_{\Gamma,\delta}$, the test space $\tilde{U}$ must have its dimension larger than or equal to the one of the reconstruction space U. For $U = V_{2,L}(\Phi_0)$ and $\tilde{U} = V_{2,\tilde{L}}(\tilde{\Phi}_0)$ with $\tilde{L} \geq L$, least square solutions of the linear system (2.40) with $-L \leq j \leq L$ replaced by $-\tilde{L} \leq j \leq \tilde{L}$ defines a sub-Galerkin reconstruction $\sum_{i=-L}^{L} c_i\phi_0(\cdot - i - \theta_i) \in V_{2,L}(\Phi_0)$ by Corollary 2.6, where $f \in V_2(\Phi_0)$ and $\Gamma := \Gamma_N, \Gamma_J, \Gamma_C$. Our numerical simulations show that the above sub-Galerkin reconstructions for different $\tilde{L} \geq L$ have comparable approximation errors.

CHAPTER 3: SPATIALLY DISTRIBUTED SAMPLING AND RECONSTRUCTION

A spatially distributed system contains a large amount of agents with limited sensing, data processing, and communication capabilities. Recent technological advances have opened up possibilities to deploy spatially distributed systems for signal sampling and reconstruction. In this chapter, we use a graph structure to describe a distributed sampling and reconstruction system by coupling agents in a spatially distributed system with innovative positions of signals. For a distributed sampling and reconstruction system, the robustness could be reduced to the stability of its sensing matrix. In a traditional centralized sampling and reconstruction system, the stability of the sensing matrix could be verified by its central processor, but the above procedure is infeasible in a distributed sampling and reconstruction system as it is decentralized. In this chapter, we split a distributed sampling and reconstruction system into a family of overlapping smaller subsystems, and we show that the stability of the sensing matrix holds if and only if its quasi-restrictions to those subsystems have uniform stability. This new stability criterion could be pivotal for the design of a robust distributed sampling and reconstruction system against supplement, replacement and impairment of agents, as we only need to check the uniform stability of affected subsystems. Here we also propose an exponentially convergent distributed algorithm for signal reconstruction, that provides a suboptimal approximation to the original signal in the presence of bounded sampling noises.

3.1 Spatially distributed systems

Spatially distributed systems (SDS) have been widely used in (underwater) multivehicle and multirobot networks, wireless sensor networks, smart grids, etc ([4, 44, 50, 165, 166]). Comparing with

traditional centralized systems that have a powerful central processor and reliable communication between agents and the central processor, an SDS could give unprecedented capabilities especially when creating a data exchange network requires significant efforts (due to physical barriers such as interference), or when establishing a centralized processor presents the daunting challenge of processing all the information (such as big-data problems). In this section, we consider SDSs for signal sampling and reconstruction, and we describe the topology of an SDS by an undirected (in)finite graph

$$\mathcal{G} := (G, S), \tag{3.1}$$

where a vertex represents an agent and an edge between two vertices means that a direct communication link exists.

Let $\mathcal{G}$ be the graph in (3.1) to describe our SDS. Let $\mathcal{G}$ be *connected* and *simple* (i.e., undirected, unweighted, no graph loops nor multiple edges), which can be interpreted as follows:

- Agents in the SDS can communicate across the entire network, but they have direct communication links only to adjacent agents.
- Direct communication links between agents are bidirectional.
- Agents have the same communication specification.
- The communication component is not used for data transmission within an agent.
- No multiple direct communication channels between agents exists.

In this section, we recall geodesic distance on the graph $\mathcal{G}$ to measure communication cost between agents. Then we consider doubling and polynomial growth properties of the counting measure on the graph $\mathcal{G}$, and we introduce Beurling dimension and sampling density of the SDS. For a discrete sampling set in the d-dimensional Euclidean space, the reader may refer to [51, 63] for its Beurling

dimension and to [13, 116, 141, 158] for its sampling density. Finally, we introduce a special family of balls to cover the graph $\mathcal{G}$, which will be used in Section 3.6 for the consensus of our proposed distributed algorithm.

3.1.1 Geodesic distance and communication cost

For a connected simple graph $\mathcal{G} := (G, S)$, let $\rho_{\mathcal{G}}(\lambda, \lambda) = 0$ for $\lambda \in G$, and $\rho_{\mathcal{G}}(\lambda, \lambda')$ be the number of edges in a shortest path connecting two distinct vertices $\lambda, \lambda' \in G$. The above function $\rho_{\mathcal{G}}$ on $G \times G$ is known as *geodesic distance* on the graph $\mathcal{G}$ ([48]). It is nonnegative and symmetric:

(i) $\rho_{\mathcal{G}}(\lambda, \lambda') \geq 0$ for all $\lambda, \lambda' \in G$;

(ii) $\rho_{\mathcal{G}}(\lambda, \lambda') = \rho_{\mathcal{G}}(\lambda', \lambda)$ for all $\lambda, \lambda' \in G$.

And it satisfies identity of indiscernibles and the triangle inequality:

(iii) $\rho_{\mathcal{G}}(\lambda, \lambda') = 0$ if and only if $\lambda = \lambda'$;

(iv) $\rho_{\mathcal{G}}(\lambda, \lambda') \leq \rho_{\mathcal{G}}(\lambda, \lambda'') + \rho_{\mathcal{G}}(\lambda'', \lambda')$ for all $\lambda, \lambda', \lambda'' \in G$.

Given two nonadjacent agents λ and $\lambda' \in G$, the distance $\rho_{\mathcal{G}}(\lambda, \lambda')$ can be used to measure the communication cost between these two agents if the communication is processed through their shortest path.

3.1.2 Counting measure, Beurling dimension and sampling density

For a connected simple graph $\mathcal{G} := (G, S)$, denote its *counting measure* by $\mu_{\mathcal{G}}$,

$$\mu_{\mathcal{G}}(F) := \sharp(F) \ \text{ for } F \subset G.$$

Definition 3.1. *The counting measure $\mu_{\mathcal{G}}$ is said to be a doubling measure if there exists a positive number $D_0(\mathcal{G})$ such that*

$$\mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda, 2r)) \le D_0(\mathcal{G})\mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda, r)) \ \text{ for all } \lambda \in G \text{ and } r \ge 0, \tag{3.2}$$

where

$$B_{\mathcal{G}}(\lambda, r) := \{\lambda' \in G, \ \ \rho_{\mathcal{G}}(\lambda, \lambda') \le r\}$$

is the closed ball with center λ and radius r.

The doubling property of the counting measure $\mu_{\mathcal{G}}$ can be interpreted as numbers of agents in r-neighborhood and $(2r)$-neighborhood of any agent are comparable. The doubling constant of $\mu_{\mathcal{G}}$ is the minimal constant $D_0(\mathcal{G}) \ge 1$ for (3.2) to hold ([49, 58]). It dominates the maximal vertex degree of the graph $\mathcal{G}$,

$$\deg(\mathcal{G}) \le D_0(\mathcal{G}), \tag{3.3}$$

because

$$\deg(\mathcal{G}) = \max_{\lambda \in G} \#\{\lambda' \in G, \ (\lambda, \lambda') \in S\} \le \max_{\lambda \in G} \#\big(B_{\mathcal{G}}(\lambda, 1)\big) \le D_0(\mathcal{G}).$$

We remark that for a finite graph $\mathcal{G}$, its doubling constant $D_0(\mathcal{G})$ could be much larger than its maximal vertex degree $\deg(\mathcal{G})$. For instance, a tree with one branch for the first L levels and two branches for the next L levels has 3 as its maximal vertex degree and $(2^{L+1} + L - 1)/(L + 1)$ as

its doubling constant, see Figure 3.1 with $L = 3$.

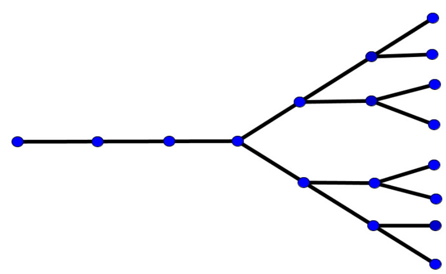

Figure 3.1: A tree with large doubling constant but limited maximal vertex degree.

The counting measure on an infinite graph is not necessarily a doubling measure. However, the counting measure on a finite graph is a doubling measure and its doubling constant could depend on the local topology and size of the graph, cf. the tree in Figure 3.1. In this paper, the graph $\mathcal{G}$ to describe our SDS is assumed to have its counting measure with the doubling property (3.2).

Assumption 1: *The counting measure $\mu_{\mathcal{G}}$ of the graph $\mathcal{G}$ is a doubling measure,*

$$D_0(\mathcal{G}) < \infty. \tag{3.4}$$

Therefore the maximal vertex degree of graph $\mathcal{G}$ is finite,

$$\deg(\mathcal{G}) < \infty,$$

which could be understood as that there are limited direct communication channels for every agent in the SDS.

Definition 3.2. *The counting measure $\mu_{\mathcal{G}}$ is said to have polynomial growth if there exist positive*

constants $D_1(\mathcal{G})$ and $d(\mathcal{G})$ such that

$$\mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda, r)) \le D_1(\mathcal{G})(1+r)^{d(\mathcal{G})} \ \text{ for all } \lambda \in G \text{ and } r \ge 0. \tag{3.5}$$

For the graph $\mathcal{G}$ associated with an SDS, we may consider minimal constants $d(\mathcal{G})$ and $D_1(\mathcal{G})$ in (3.5) as *Beurling dimension* and *sampling density* of the SDS respectively. We remark that

$$d(\mathcal{G}) \ge 1, \tag{3.6}$$

because

$$\sup_{\lambda \in G} \mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda, r)) \ge 1 + r \ \text{ for all } 0 \le r \le \operatorname{diam}(\mathcal{G}),$$

where $\operatorname{diam}(\mathcal{G}) := \sup_{\lambda, \lambda' \in G} \rho_{\mathcal{G}}(\lambda, \lambda')$ is the diameter of the graph $\mathcal{G}$.

Applying (3.2) repeatedly leads to the following general doubling property:

$$\mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda, sr)) \le (D_0(\mathcal{G}))^{\lceil \log_2 s \rceil} \mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda, r)) \le D_0(\mathcal{G}) s^{\log_2 D_0(\mathcal{G})} \mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda, r))$$

for all $\lambda \in G$, $s \ge 1$ and $r \ge 0$. Thus

$$\mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda, r)) \le D_0(\mathcal{G})(1+r)^{\log_2 D_0(\mathcal{G})} \mu_{\mathcal{G}}\Big(B_{\mathcal{G}}\Big(\lambda, \frac{r}{1+r}\Big)\Big) = D_0(\mathcal{G})(1+r)^{\log_2 D_0(\mathcal{G})}, \ r \ge 0.$$

This shows that a doubling measure has polynomial growth.

Proposition 3.3. *If the counting measure $\mu_{\mathcal{G}}$ on a connected simple graph $\mathcal{G}$ is a doubling measure, then it has polynomial growth.*

For a connected simple graph $\mathcal{G}$, its maximal vertex degree is finite if the counting measure $\mu_{\mathcal{G}}$ has polynomial growth, but the converse is not true. We observe that if the maximal vertex degree

$\deg(\mathcal{G})$ is finite, then the counting measure $\mu_{\mathcal{G}}$ has exponential growth,

$$\mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda, r)) \le \frac{(\deg(\mathcal{G}))^{r+1} - 1}{\deg(\mathcal{G}) - 1} \text{ for all } \lambda \in G \text{ and } r \ge 0. \tag{3.7}$$

3.1.3 Spatially distributed subsystems

For a connected simple graph $\mathcal{G} := (G, S)$, take a maximal N-disjoint subset $G_N \subset G, 0 \le N \in \mathbb{R}$, such that

$$B_{\mathcal{G}}(\lambda, N) \bigcap \big(\bigcup_{\lambda_m \in G_N} B_{\mathcal{G}}(\lambda_m, N) \big) \neq \emptyset \text{ for all } \lambda \in G, \tag{3.8}$$

and

$$B_{\mathcal{G}}(\lambda_m, N) \bigcap B_{\mathcal{G}}(\lambda_{m'}, N) = \emptyset \text{ for all } \lambda_m, \lambda_{m'} \in G_N. \tag{3.9}$$

For $0 \le N < 1$, it follows from (3.8) that $G_N = G$. For $N \ge 1$, there are many subsets G_N of vertices satisfying (3.8) and (3.9). For instance, we can construct $G_N = \{\lambda_m\}_{m \ge 1}$ as follows: take a $\lambda_1 \in G$ and define $\lambda_m,\ m \ge 2$, recursively by

$$\lambda_m = \operatorname*{argmin}_{\lambda \in A_m} \rho_{\mathcal{G}}(\lambda, \lambda_1),$$

where $A_m = \{\lambda \in G,\ B_{\mathcal{G}}(\lambda, N) \cap B_{\mathcal{G}}(\lambda_{m'}, N) = \emptyset, 1 \le m' \le m-1\}$.

For a set G_N satisfying (3.8) and (3.9), the family of balls $\{B_{\mathcal{G}}(\lambda_m, N'), \lambda_m \in G_N\}$ with $N' \ge 2N$ provides a finite covering for G.

Proposition 3.4. *Let $\mathcal{G} := (G, S)$ be a connected simple graph and $\mu_{\mathcal{G}}$ have the doubling property* (3.4) *with constant $D_0(\mathcal{G})$. If G_N satisfies* (3.8) *and* (3.9)*, then*

$$1 \le \inf_{\lambda \in G} \Sigma_{\lambda_m \in G_N} \chi_{B_{\mathcal{G}}(\lambda_m, N')}(\lambda) \le \sup_{\lambda \in G} \Sigma_{\lambda_m \in G_N} \chi_{B_{\mathcal{G}}(\lambda_m, N')}(\lambda) \le (D_0(\mathcal{G}))^{\lceil \log_2(2N'/N+1) \rceil} \tag{3.10}$$

for all $N' \geq 2N$.

Proof. For any $\lambda \in G$, take $\lambda_m \in G_N$ with $B_{\mathcal{G}}(\lambda, N) \cap B_{\mathcal{G}}(\lambda_m, N) \neq \emptyset$. Then

$$\rho_{\mathcal{G}}(\lambda, \lambda_m) \leq \rho_{\mathcal{G}}(\lambda, \lambda') + \rho_{\mathcal{G}}(\lambda', \lambda_m) \leq 2N,$$

where λ' is a vertex in $B_{\mathcal{G}}(\lambda, N) \cap B_{\mathcal{G}}(\lambda_m, N)$. This proves that for any $N' \geq 2N$, balls $\{B_{\mathcal{G}}(\lambda_m, N'), \lambda_m \in G_N\}$ provide a covering for G,

$$G \subset \bigcup_{\lambda_m \in G_N} B_{\mathcal{G}}(\lambda_m, N'), \tag{3.11}$$

and hence the first inequality in (3.10) follows.

Now we prove the last inequality in (3.10). Take $\lambda \in G$. For any $\lambda_m, \lambda_{m'} \in G_N \cap B_{\mathcal{G}}(\lambda, N')$,

$$\rho_{\mathcal{G}}(\lambda', \lambda_{m'}) \leq \rho_{\mathcal{G}}(\lambda', \lambda_m) + \rho_{\mathcal{G}}(\lambda_m, \lambda) + \rho_{\mathcal{G}}(\lambda, \lambda_{m'}) \leq 2N' + N$$

for all $\lambda' \in B(\lambda_m, N)$, which implies that

$$B_{\mathcal{G}}(\lambda_m, N) \subset B_{\mathcal{G}}(\lambda_{m'}, 2N' + N). \tag{3.12}$$

Hence

$$\begin{aligned}\sum_{\lambda_m \in G_N} \chi_{B_{\mathcal{G}}(\lambda_m, N')}(\lambda) &\leq \frac{\mu_{\mathcal{G}}(\cup_{\lambda_m \in G_N \cap B_{\mathcal{G}}(\lambda, N')} B_{\mathcal{G}}(\lambda_m, N))}{\inf_{\lambda_m \in G_N \cap B_{\mathcal{G}}(\lambda, N')} \mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda_m, N))} \\ &\leq \sup_{\lambda_m \in G_N \cap B_{\mathcal{G}}(\lambda, N')} \frac{\mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda_m, 2N' + N))}{\mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda_m, N))} \leq (D_0(\mathcal{G}))^{\lceil \log_2(2N'/N+1) \rceil},\end{aligned} \tag{3.13}$$

where the first inequality holds as $B_{\mathcal{G}}(\lambda_m, N), \lambda_m \in V_N$, are disjoint, the second one is true by (3.12), and the third inequality follows from the doubling assumption (3.2). $\square$

For $N' \geq 0$, define a family of spatially distributed subsystems

$$\mathcal{G}_{\lambda,N'} := (B_{\mathcal{G}}(\lambda, N'), S_{\lambda,N'})$$

with fusion agents $\lambda \in G_N$, where $(\lambda', \lambda'') \in S_{\lambda,N'}$ if $\lambda', \lambda'' \in B_{\mathcal{G}}(\lambda, N')$ and $(\lambda', \lambda'') \in S$. Then the maximal N-disjoint property of the set G_N means that the N-neighboring subsystems $\mathcal{G}_{\lambda_m,N}, \lambda_m \in G_N$, have no common agent. On the other hand, it follows from Proposition 3.4 that for any $N' \geq 2N$, every agent in our SDS is in at least one and at most finitely many of the N'-neighboring subsystems $\mathcal{G}_{\lambda_m,N'}, \lambda_m \in G_N$. The above idea to split the SDS into subsystems of small sizes is crucial in our proposed distributed algorithm in Section 3.6 for stable signal reconstruction.

3.2 Signals on the graph $\mathcal{V}$

The spatial signals with the following parametric representation,

$$f := \sum_{i \in V} c(i)\varphi_i, \tag{3.14}$$

where amplitudes $c(i), i \in V$, are bounded, and generators $\varphi_i, i \in V$, are essentially supported in a spatial neighborhood of the innovative position i. The above family of spatial signals appears in magnetic resonance spectrum, mass spectrometry, global positioning system, cellular radio, ultra wide-band communication, electrocardiogram, and many engineering applications, see [60, 143, 159] and references therein.

In our work, we associate every innovative position $i \in V$ with some anchor agents $\lambda \in G$, and denote the set of such associations (i, λ) by T. These associations can be easily understood as agents within certain (spatial) range of every innovative position. With the above associations, we describe our distributed sampling and reconstruction system (DSRS) by an undirected (in)finite graph

$$\mathcal{H} := (G \cup V,\ S \cup T \cup T^*), \tag{3.15}$$

where $T^* = \{(\lambda, i) \in G \times V,\ (i, \lambda) \in T\}$, see Figure 3.2. The above graph description of a DSRS plays a crucial role for us to study signal sampling and reconstruction.

Given a DSRS described by the above graph $\mathcal{H}$, set

$$E := \{(i, i') \in V \times V,\ \ i \neq i' \text{ and } (i, \lambda), (i', \lambda) \in T \text{ for some } \lambda \in G\}. \tag{3.16}$$

We then generate a graph structure

$$\mathcal{V} := (V, E) \tag{3.17}$$

for signals in (3.14), where an edge between two distinct innovative positions in V means that a common anchor agent exists. The above graph structure for signals is different from the conventional one in most of the literature, where the graph is usually preassigned. The reader may refer to [124, 130, 136] and Remark 3.12.

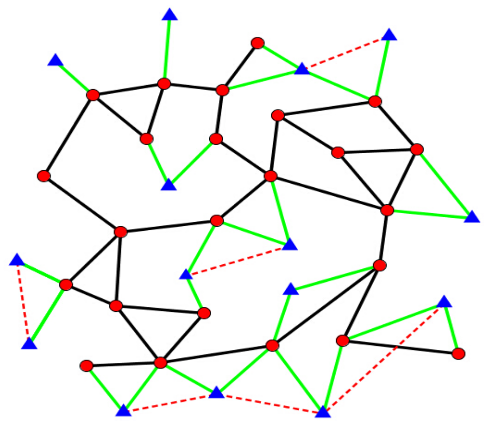

Figure 3.2: The graph structure $\mathcal{H}$ for a DSRS.

Presented in Figure 3.2 is the graph structure $\mathcal{H} = (G \cup V,\ S \cup T \cup T^*)$ in (3.15) to describe a DSRS, where vertices in G and V are plotted in red circles and blue triangles, and edges in S, T and E are in black solid lines, green solid lines and red dashed lines respectively.

Let V be the set of innovative positions of signals f in (3.14), and $\mathcal{G} = (G, S)$ be the graph in (3.1) to represent our SDS. We build the graph $\mathcal{H}$ in (3.15) to describe our DSRS by associating every innovative position in V with some anchor agents in G. In this paper, we consider DSRS with the following properties.

Assumption 2: *There is a direct communication link between distinct anchor agents of an innovative position,*

$$(\lambda_1, \lambda_2) \in S \ \text{ if } \ (i, \lambda_1) \ \text{ and } \ (i, \lambda_2) \in T \ \text{ for some } \ i \in V. \tag{3.18}$$

Assumption 3: *There are finitely many innovative positions for any anchor agent,*

$$L := \sup_{\lambda \in G} \#\{i \in V,\ (i, \lambda) \in T\} < \infty. \tag{3.19}$$

Assumption 4: *Any agent has an anchor agent within bounded distance,*

$$M := \sup_{\lambda \in G} \inf\{\rho_{\mathcal{G}}(\lambda, \lambda'),\ (i, \lambda') \in T \text{ for some } i \in V\} < \infty. \tag{3.20}$$

The graph $\mathcal{H}$ associated with the above DSRS is a connected simple graph. Moreover, we have the following important properties about shortest paths between different vertices in $\mathcal{H}$.

Proposition 3.5. *Let the graph $\mathcal{H}$ in* (3.15) *satisfy* (3.18)*. Then all intermediate vertices in the shortest paths in $\mathcal{H}$ to connect distinct vertices in $\mathcal{H}$ belong to the subgraph $\mathcal{G}$.*

Proof. By the structure of the graph $\mathcal{H}$, it suffices to show that the shortest path in $\mathcal{H}$ to connect distinct vertices $\lambda, \lambda' \in G$ must be a path in its subgraph $\mathcal{G}$. Suppose on the contrary that $\lambda u_1 \cdots u_{k-1} u_k u_{k+1} \cdots u_n \lambda'$ is a shortest path in $\mathcal{H}$ of length $\rho_{\mathcal{H}}(\lambda, \lambda')$ with vertex u_k along the path belonging to V. Then u_{k-1} and u_{k+1} are anchor agents of u_k in G.

For the case that u_{k-1} and u_{k+1} are distinct anchor agents of the innovative position u_k, $(u_{k-1}, u_{k+1}) \in S$ by (3.18). Hence $\lambda u_1 \cdots u_{k-1} u_{k+1} \cdots u_n \lambda'$ is a path of length $\rho_{\mathcal{H}}(\lambda, \lambda') - 1$ to connect vertices λ and λ', which is a contradiction.

Similarly for the case that u_{k-1} and u_{k+1} are the same, $\lambda u_1 \cdots u_{k-1} u_{k+2} \cdots u_n \lambda'$ is a path of length $\rho_{\mathcal{H}}(\lambda, \lambda') - 2$ to connect vertices λ and λ'. This is a contradiction. $\square$

By Proposition 3.5,

$$\rho_{\mathcal{H}}(\lambda, \lambda') = \rho_{\mathcal{G}}(\lambda, \lambda') \quad \text{for all } \lambda, \lambda' \in G, \tag{3.21}$$

and

$$\rho_{\mathcal{H}}(i, i') = 2 + \inf_{\lambda, \lambda' \in G} \{\rho_{\mathcal{G}}(\lambda, \lambda') : (i, \lambda), (i', \lambda') \in T\} \quad \text{for all distinct } i, i' \in V, \tag{3.22}$$

where $\rho_{\mathcal{H}}$ is the geodesic distance for the graph $\mathcal{H}$.

Let $\mathcal{V}$ be the graph in (3.17), where there is an edge between two distinct innovative positions if they share a common anchor agent. One may easily verify that the graph $\mathcal{V}$ is undirected and its maximal vertex degree is finite,

$$\deg(\mathcal{V}) \le L \sup_{i \in V} \#\{\lambda \in G,\ (i, \lambda) \in T\} \le L(\deg(\mathcal{G}) + 1) \tag{3.23}$$

by (3.3), (3.4), (3.18) and (3.19).

We cannot define a geodesic distance on $\mathcal{V}$ as in Subsection 3.1.1, since the graph $\mathcal{V}$ is unconnected in general. With the help of the graph $\mathcal{H}$ to describe our DSRS, we define a distance ρ on the graph $\mathcal{V}$.

Proposition 3.6. *Let $\mathcal{H}$ be the graph in* (3.15)*. Define a function $\rho : V \times V \longmapsto \mathbb{R}$ by*

$$\rho(i, i') = \begin{cases} 0 & \text{if } i = i' \\ \rho_{\mathcal{H}}(i, i') - 1 & \text{if } i \neq i'. \end{cases} \tag{3.24}$$

If the graph $\mathcal{H}$ satisfies (3.18)*, then ρ is a distance on the graph $\mathcal{V}$:*

(i) $\rho(i, i') \ge 0$ *for all* $i, i' \in V$*;*

(ii) $\rho(i,i') = \rho(i',i)$ *for all* $i,i' \in V$*;*

(iii) $\rho(i,i') = 0$ *if and only if* $i = i'$*; and*

(iv) $\rho(i,i') \leq \rho(i,i'') + \rho(i'',i')$ *for all* $i,i',i'' \in V$*.*

Proof. The non-negativity and symmetry is obvious, while the identity of indiscernibles holds since there is no edge assigned in $\mathcal{H}$ between two distinct vertices in V.

Now we prove the triangle inequality

$$\rho(i,i') \leq \rho(i,i'') + \rho(i'',i') \ \text{ for distinct vertices } i,i',i'' \in V. \tag{3.25}$$

Let $m = \rho(i,i'')$ and $n = \rho(i'',i')$. Take a path $iv_1 \dots v_m i''$ of length $m+1$ to connect i and i'', and another path $i''u_1 \dots u_n i'$ of length $n+1$ to connect i'' and i'. If $v_m = u_1$, then $iv_1 \dots v_m u_2 \cdots u_n i'$ is a path of length $m+n$ to connect vertices i and i', which implies that

$$\rho(i,i') \leq m+n-1 < \rho(i,i'') + \rho(i'',i'). \tag{3.26}$$

If $v_m \neq u_1$, then (v_m, u_1) is an edge in the graph $\mathcal{G}$ (and then also in the graph $\mathcal{H}$) by (3.18). Thus $iv_1 \dots v_m u_1 u_2 \cdots u_n i'$ is a path of length $m+n+1$ to connect vertices i and i', and

$$\rho(i,i') \leq m+n = \rho(i,i'') + \rho(i'',i'). \tag{3.27}$$

Combining (3.26) and (3.27) proves (3.25). $\square$

Clearly, the above distance between two endpoints of an edge in $\mathcal{V}$ is one. Denote the closed ball

with center $i \in V$ and radius r by

$$B(i,r) = \{i' \in V,\ \rho(i,i') \leq r\},$$

and the counting measure on V by μ. We say that μ is a *doubling measure* if

$$\mu(B(i,2r)) \leq D_0\mu(B(i,r)) \text{ for all } i \in V \text{ and } r \geq 0, \tag{3.28}$$

and it has *polynomial growth* if

$$\mu(B(i,r)) \leq D_1(1+r)^d \text{ for all } i \in V \text{ and } r \geq 0, \tag{3.29}$$

where D_0, D_1 and d are positive constants. The minimal constant D_0 for (3.28) to hold is known as the doubling constant, and the minimal constants d and D_1 in (3.29) are called *dimension* and *maximal rate of innovation* for signals on the graph $\mathcal{V}$ respectively. The concept of rate of innovation was introduced in [159] and later extended in [143, 150]. The reader may refer to [28, 29, 62, 111, 121, 135, 141, 143, 150, 159] and references therein for sampling and reconstruction of signals with finite rate of innovation.

In the next two propositions, we show that the counting measure μ on $\mathcal{V}$ has the doubling property (respectively, the polynomial growth property) if and only if the counting measure $\mu_{\mathcal{G}}$ on $\mathcal{G}$ does.

Proposition 3.7. *Let $\mathcal{G}$ and $\mathcal{H}$ satisfy Assumptions 1 – 4. If $\mu_{\mathcal{G}}$ is a doubling measure with constant $D_0(\mathcal{G})$, then*

$$\mu(B(i,2r)) \leq L(D_0(\mathcal{G}))^2\Big(\frac{(\deg(\mathcal{G}))^{2M+3}-1}{\deg(\mathcal{G})-1}\Big)\mu(B(i,r)) \text{ for all } i \in V \text{ and } r \geq 0. \tag{3.30}$$

Conversely, if μ is a doubling measure with constant D_0, then

$$\mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda, 2r)) \leq L D_0^2 \Big(\frac{(\deg(\mathcal{G}))^{2M+3}-1}{\deg(\mathcal{G})-1}\Big)^2 \mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda, r)) \ \text{ for all } \lambda \in G \text{ and } r \geq 0. \tag{3.31}$$

To prove Proposition 3.7, we need two lemmas comparing measures of balls in graphs $\mathcal{G}$ and $\mathcal{V}$.

Lemma 3.8. *If $\mathcal{H}$ satisfies* (3.18) *and* (3.19)*, then*

$$\mu(B(i, r)) \leq L \mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda, r)) \ \text{ for any } \lambda \in G \text{ with } (i, \lambda) \in T. \tag{3.32}$$

Proof. Let $i' \in B(i, r)$ with $i' \neq i$. By Proposition 3.5, there exists a path $\lambda_1 \ldots \lambda_n$ of length $\rho(i, i') - 1$ in the graph $\mathcal{G}$ such that $(i, \lambda_1), (i', \lambda_n) \in T$. Then

$$\rho_{\mathcal{G}}(\lambda, \lambda_n) \leq \rho_{\mathcal{G}}(\lambda, \lambda_1) + \rho_{\mathcal{G}}(\lambda_1, \lambda_n) \leq \rho(i, i') \leq r$$

as either $\lambda_1 = \lambda$ or (λ, λ_1) is an edge in $\mathcal{G}$ by (3.18). This shows that for any innovative position $i' \in B(i, r)$ there exists an anchor agent λ_n in the ball $B_{\mathcal{G}}(\lambda, r)$. This observation together with (3.19) proves (3.32). $\square$

Lemma 3.9. *If $\mathcal{H}$ satisfies* (3.4)*,* (3.18) *and* (3.20)*, then*

$$\mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda, r)) \leq \Big(\sup_{\lambda' \in G} \mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda', 2M+2))\Big) \mu(B(i, r+M+1)) \tag{3.33}$$

for any $\lambda \in G$ and $r \geq M+1$, where $(i, \lambda') \in T$ and $\lambda' \in B_{\mathcal{G}}(\lambda, M)$.

Proof. Let $\lambda_1 = \lambda$ and take $\Lambda = \{\lambda_m\}_{m \geq 1}$ such that (i) $B_{\mathcal{G}}(\lambda_m, M+1) \subset B_{\mathcal{G}}(\lambda, r)$ for all $\lambda_m \in \Lambda$; (ii) $B_{\mathcal{G}}(\lambda_m, M+1) \bigcap B_{\mathcal{G}}(\lambda_{m'}, M+1) = \emptyset$ for all distinct vertices $\lambda_m, \lambda_{m'} \in \Lambda$; and (iii) $B_{\mathcal{G}}(\tilde{\lambda}, M+1) \bigcap \big(\bigcup_{\lambda_m \in \Lambda} B_{\mathcal{G}}(\lambda_m, M+1)\big) \neq \emptyset$ for all $\tilde{\lambda} \in B_{\mathcal{G}}(\lambda, r)$. The set Λ could be

considered as a maximal $(M+1)$-disjoint subset of $B_{\mathcal{G}}(\lambda, r)$. Following the argument used in the proof of Proposition 3.4, $\{B_{\mathcal{G}}(\lambda_m, 2(M+1))\}_{\lambda_m \in \Lambda}$ forms a covering of the ball $B(\lambda, r)$, which implies that

$$\mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda, r)) \le \Big(\sup_{\lambda_m \in \Lambda} \mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda_m, 2M+2))\Big) \#\Lambda \le \Big(\sup_{\lambda' \in G} \mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda', 2M+2))\Big) \#\Lambda. \quad (3.34)$$

For $\lambda_m \in \Lambda$, define

$$V_{\lambda_m} = \{i' \in V, (i', \tilde{\lambda}) \in T \text{ for some } \tilde{\lambda} \in B_{\mathcal{G}}(\lambda_m, M)\}.$$

Then it follows from (3.20) that

$$\#V_{\lambda_m} \ge 1 \text{ for all } \lambda_m \in \Lambda. \quad (3.35)$$

Observe that the distance of anchor agents associated with innovative positions in distinct V_{λ_m} is at least 2 by the second requirement (ii) for the set Λ. This together with the assumption (3.18) implies that

$$V_{\lambda_m} \cap V_{\lambda_{m'}} = \emptyset \ \text{ for distinct } \lambda_m, \lambda_{m'} \in \Lambda. \quad (3.36)$$

Combining (3.34), (3.35) and (3.36) leads to

$$\mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda, r)) \le \Big(\sup_{\lambda' \in G} \mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda', 2M+2))\Big) \#\Big(\cup_{\lambda_m \in \Lambda} V_{\lambda_m}\Big). \quad (3.37)$$

Take $i \in V$ with $(i, \lambda') \in T$ for some $\lambda' \in B_{\mathcal{G}}(\lambda, M)$, and $i' \in V_{\lambda_m}, \lambda_m \in \Lambda$. Then

$$\rho_{\mathcal{H}}(i, \lambda) \le \rho_{\mathcal{H}}(i, \lambda') + \rho_{\mathcal{H}}(\lambda', \lambda) \le M+1,$$

and

$$\rho_{\mathcal{H}}(i', \lambda) \leq \rho_{\mathcal{H}}(i', \tilde{\lambda}) + \rho_{\mathcal{H}}(\tilde{\lambda}, \lambda) \leq r + 1,$$

where $\tilde{\lambda} \in B_{\mathcal{G}}(\lambda_m, M)$ and $(i', \tilde{\lambda}) \in T$. Thus

$$\rho(i, i') \leq r + M + 1. \tag{3.38}$$

Then the desired estimate (3.33) follows from (3.37) and (3.38). □

Here, we begin our proof of Proposition 3.7.

Proof. First we prove the doubling property (3.30) for the measure μ. Take $i \in V$. Then for $r \geq 2(M+1)$ it follows from Lemmas 3.8 and 3.9 that

$$\begin{aligned} \mu(B(i, 2r)) &\leq L\mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda, 2r)) \leq L(D_0(\mathcal{G}))^2 \mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda, r/2)) \\ &\leq KL(D_0(\mathcal{G}))^2 \mu(B(i, r/2 + M + 1)) \leq KL(D_0(\mathcal{G}))^2 \mu(B(i, r)), \end{aligned} \tag{3.39}$$

where $\lambda \in G$ is a vertex with $(i, \lambda) \in T$ and

$$K := \sup_{\lambda' \in G} \mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda', 2M + 2)) \leq \frac{((\deg(\mathcal{G}))^{2M+3} - 1}{\deg(\mathcal{G}) - 1} \tag{3.40}$$

by (3.7). From the doubling property (3.2) for the measure $\mu_{\mathcal{G}}$, we obtain

$$\mu(B(i, 2r)) \leq KLD_0(\mathcal{G}) \leq KLD_0(\mathcal{G})\mu(B(i, r)) \ \text{ for } 0 \leq r \leq 2(M+1). \tag{3.41}$$

Then the doubling property (3.30) follows from (3.39), (3.40) and (3.41).

Next we prove the doubling property (3.31) for the measure $\mu_{\mathcal{G}}$. Let $\lambda' \in B_{\mathcal{G}}(\lambda, M)$ with $(i, \lambda') \in$

T for some $i \in V$. The existence of such λ' follows from assumption (3.20). From Lemmas 3.8 and 3.9, we obtain

$$\begin{aligned}\mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda, 2r)) &\le K\mu(B(i, 2r+M+1)) \le D_0^2 K\mu\Big(B\Big(i, \frac{r}{2}+\frac{(M+1)}{4}\Big)\Big)\\ &\le D_0^2 LK\mu_{\mathcal{G}}\Big(B_{\mathcal{G}}\Big(\lambda', \frac{r}{2}+\frac{M+1}{4}\Big)\Big)\\ &\le D_0^2 LK\mu_{\mathcal{G}}\Big(B_{\mathcal{G}}\Big(\lambda, \frac{r}{2}+\frac{M+1}{4}+M\Big)\Big) \le D_0^2 LK\mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda, r)) \end{aligned} \tag{3.42}$$

for $r \ge 3M$, and

$$\begin{aligned}\mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda, 2r)) &\le K\mu(B(i, 7M)) \le D_0^2 K\mu(B(i, 2M))\\ &\le D_0^2 LK\mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda', 2M)) \le D_0^2 LK^2\mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda, r)) \end{aligned} \tag{3.43}$$

for $0 \le r \le 3M-1$. Combining (3.40), (3.42) and (3.43) proves (3.31). □

Proposition 3.10. *Let $\mathcal{G}$ and $\mathcal{H}$ satisfy Assumptions 1 – 4. If $\mu_{\mathcal{G}}$ has polynomial growth with Beurling dimension $d(\mathcal{G})$ and sampling density $D_1(\mathcal{G})$, then*

$$\mu(B(i, r)) \le LD_1(\mathcal{G})(1+r)^{d(\mathcal{G})} \text{ for all } i \in V \text{ and } r \ge 0. \tag{3.44}$$

Conversely, if μ has polynomial growth with dimension d and maximal rate of innovation D_1, then

$$\mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda, r)) \le 2^d\Big(\frac{(\deg(\mathcal{G}))^{2M+3}-1}{\deg(\mathcal{G})-1}\Big)D_1(1+r)^d \text{ for all } \lambda \in G \text{ and } r \ge 0. \tag{3.45}$$

Proof. The polynomial growth property (3.44) for the measure μ follows immediately from Lemma 3.8.

The polynomial growth property (3.45) for the measure $\mu_{\mathcal{G}}$ holds because

$$\mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda, r)) \le \frac{(\deg(\mathcal{G}))^M - 1}{\deg(\mathcal{G}) - 1},\ 0 \le r \le M - 1$$

by (3.7), and

$$\begin{aligned}\mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda, r)) \ &\le\ D_1\Big(\frac{(\deg(\mathcal{G}))^{2M+3} - 1}{\deg(\mathcal{G}) - 1}\Big)(r + M + 2)^d \\ &\le\ 2^d D_1\Big(\frac{(\deg(\mathcal{G}))^{2M+3} - 1}{\deg(\mathcal{G}) - 1}\Big)(r + 1)^d,\ r \ge M,\end{aligned}$$

by (3.40) and Lemma 3.9.

By (3.6), Propositions 3.7 and 3.10, we conclude that signals in (3.14) have their dimension d being the same as the Beurling dimension $d(\mathcal{G})$, and their maximal rate D_1 of innovation being approximately proportional to the sampling density $D_1(\mathcal{G})$. $\square$

Theorem 3.11. *Let $\mathcal{G}$ and $\mathcal{H}$ satisfy Assumptions 1 – 4. Then*

$$d(\mathcal{G}) = d \ge 1 \tag{3.46}$$

and

$$L^{-1} D_1 \le D_1(\mathcal{G}) \le 2^d\Big(\frac{(\deg(\mathcal{G}))^{2M+3} - 1}{\deg(\mathcal{G}) - 1}\Big) D_1. \tag{3.47}$$

We finish this section with a remark about signals on our graph $\mathcal{V}$, cf. [124, 130, 136].

Remark 3.12. Signals on the graph $\mathcal{V}$ are analog in nature, while signals on graphs in most of the literature are discrete ([124, 130, 136]). Let $\mathbf{p}_\lambda$ and $\mathbf{p}_i$ be the physical positions of the agent $\lambda \in G$

and innovative position $i \in V$, respectively. If there exist positive constants A and B such that

$$A\sum_{i\in V}|c(i)|^2 \le \sum_{i\in V}|f(\mathbf{p}_i)|^2+\sum_{\lambda\in G}|f(\mathbf{p}_\lambda)|^2 \le B\sum_{i\in V}|c(i)|^2$$

for all signals f with the parametric representation (3.14), then we can establish a one-to-one correspondence between the analog signal f and the discrete signal F on the graph $\mathcal{H}$, where

$$F(u)=f(\mathbf{p}_u),\ u\in G\cup V.$$

The above family of discrete signals F forms a linear space, which could be a Paley-Wiener space associated with some positive-semidefinite operator (such as Laplacian) on the graph $\mathcal{H}$. Using the above correspondence, our theory for signal sampling and reconstruction applies by assuming that the impulse response ψ_λ of every agent $\lambda \in G$ is supported on $\mathbf{p}_u, u \in G\cup V$.

3.3 Sensing matrices with polynomial off-diagonal decay

Let $\mathcal{H}$ be the connected simple graph in (3.15) to describe our DSRS. Define sensing matrix $\mathbf{S}$ of our DSRS by

$$\mathbf{S} := (\langle \varphi_i, \psi_\lambda\rangle)_{\lambda\in G, i\in V}. \tag{3.48}$$

The sensing matrix $\mathbf{S}$ is stored by agents in a distributed manner. Due to the storage limitation, each agent in our SDS stores its corresponding row (and perhaps also its neighboring rows) in the sensing matrix $\mathbf{S}$, but it does not have the whole matrix available. Agents in our SDS have limited acquisition ability and they could essentially catch signals not far from their physical locations. So the sensing matrix $\mathbf{S}$ has certain *polynomial off-diagonal decay*, i.e., there exist positive constants

D and α such that

$$|\langle \varphi_i, \psi_\lambda \rangle| \le D(1+\rho_{\mathcal{H}}(\lambda, i))^{-\alpha} \text{ for all } \lambda \in G \text{ and } i \in V, \tag{3.49}$$

where $\rho_{\mathcal{H}}$ is the geodesic distance on the graph $\mathcal{H}$. For most DSRSs in applications, such as multivehicle and multirobot networks and wireless sensor networks, the signal generated at any innovative position could be detected by its anchor agents and some of their neighboring agents, but not by agents in the SDS far away. Thus the sensing matrix $\mathbf{S}$ may have finite bandwidth $s \ge 0$,

$$\langle \varphi_i, \psi_\lambda \rangle = 0 \quad \text{if } \rho_{\mathcal{H}}(\lambda, i) > s. \tag{3.50}$$

The above global requirements (3.49) and (3.50) could be fulfilled in a distributed manner.

We assume in this paper that the sensing matrix $\mathbf{S}$ in (3.48) satisfies

$$\mathbf{S} \in \mathcal{J}_\alpha(\mathcal{G}, \mathcal{V}) \text{ for some } \alpha > d, \tag{3.51}$$

where

$$\mathcal{J}_\alpha(\mathcal{G}, \mathcal{V}) := \left\{ \mathbf{A} := (a(\lambda, i))_{\lambda \in G, i \in V},\ \|\mathbf{A}\|_{\mathcal{J}_\alpha(\mathcal{G}, \mathcal{V})} < \infty \right\} \tag{3.52}$$

is the *Jaffard class* $\mathcal{J}_\alpha(\mathcal{G}, \mathcal{V})$ of matrices with polynomial off-diagonal decay, and

$$\|\mathbf{A}\|_{\mathcal{J}_\alpha(\mathcal{G}, \mathcal{V})} := \sup_{\lambda \in G, i \in V} (1+\rho_{\mathcal{H}}(\lambda, i))^{\alpha} |a(\lambda, i)|, \quad \alpha \ge 0. \tag{3.53}$$

The reader may refer to [84, 85, 93, 140, 142, 147] for matrices with various off-diagonal decay.

We observe that a matrix in $\mathcal{J}_\alpha(\mathcal{G}, \mathcal{V}), \alpha > d$, defines a bounded operator from $\ell^p(V)$ to $\ell^p(G), 1 \le p \le \infty$.

Proposition 3.13. *Let $\mathcal{G}$ and $\mathcal{H}$ satisfy Assumptions 1 – 4, $\mathcal{V}$ be as in* (3.17)*, and let $\mu_{\mathcal{G}}$ have polynomial growth with Beurling dimension d and sampling density $D_1(\mathcal{G})$. If $\mathbf{A} \in \mathcal{J}_\alpha(\mathcal{G}, \mathcal{V})$ for some $\alpha > d$, then*

$$\|\mathbf{A}\mathbf{c}\|_p \leq \frac{D_1(\mathcal{G})L\alpha}{\alpha - d}\|\mathbf{A}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}\|\mathbf{c}\|_p \text{ for all } \mathbf{c} \in \ell^p, 1 \leq p \leq \infty. \tag{3.54}$$

To prove Proposition 3.13, we need a technical lemma.

Lemma 3.14. *Let $\mathcal{G}$ be a connected simple graph. If its counting measure has polynomial growth* (3.5)*, then*

$$\sup_{\lambda \in G} \sum_{\rho_{\mathcal{G}}(\lambda,\lambda') \geq s} (1 + \rho_{\mathcal{G}}(\lambda, \lambda'))^{-\alpha} \leq \frac{D_1(\mathcal{G})\alpha}{\alpha - d}(s+1)^{-\alpha+d} \tag{3.55}$$

for all $\alpha > d$ and nonnegative integers s, where d and $D_1(\mathcal{G})$ are the Beurling dimension and sampling density respectively.

Proof. Take $\lambda \in G$ and $\alpha > d$. Then

$$\begin{aligned}
&\sum_{\rho_{\mathcal{G}}(\lambda,\lambda')\geq s} (1+\rho_{\mathcal{G}}(\lambda,\lambda'))^{-\alpha} = \sum_{n\geq s}(n+1)^{-\alpha}\Big(\sum_{\rho_{\mathcal{G}}(\lambda,\lambda')=n} 1\Big) \\
&= \sum_{n\geq s}(n+1)^{-\alpha}\Big(\mu_{\mathcal{G}}\big(B_{\mathcal{G}}(\lambda,n)\big) - \mu_{\mathcal{G}}\big(B_{\mathcal{G}}(\lambda,n-1)\big)\Big) \\
&= \lim_{K\to\infty}\Big[(K+1)^{-\alpha}\mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda,K)) + \sum_{n=s}^{K-1}\mu_{\mathcal{G}}\big(B_{\mathcal{G}}(\lambda,n)\big)((n+1)^{-\alpha}-(n+2)^{-\alpha}) \\
&- \quad (s+1)^{-\alpha}\mu_{\mathcal{G}}\big(B_{\mathcal{G}}(\lambda,s-1)\big)\Big] \\
&\leq \sum_{n\geq s}\mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda,n))((n+1)^{-\alpha}-(n+2)^{-\alpha}\big) \\
&\leq D_1(\mathcal{G})\sum_{n=s}^{\infty}(n+1)^{d}((n+1)^{-\alpha}-(n+2)^{-\alpha}\big) \\
&= D_1(\mathcal{G})\Big((s+1)^{-\alpha+d} - (s+1)^{d}(s+2)^{-\alpha} + \sum_{n=s+1}^{\infty}(n+1)^{d}\big((n+1)^{-\alpha}-(n+2)^{-\alpha}\big)\Big) \\
&= D_1(\mathcal{G})\Big((s+1)^{-\alpha+d} + \sum_{n=s+1}^{\infty}(n+1)^{-\alpha}\big((n+1)^{d}-n^{d}\big)\Big) \\
&\leq D_1(\mathcal{G})\Big((s+1)^{-\alpha+d} + d\int_{s+1}^{\infty}t^{d-\alpha-1}dt\Big) = \frac{D_1(\mathcal{G})\alpha}{\alpha-d}(s+1)^{-\alpha+d},
\end{aligned} \tag{3.56}$$

where the fourth inequality follows from (3.5), and the seventh one is true as $(n+1)^d - n^d \leq d(n+1)^{d-1}$ for $n \geq 1$ and $d \geq 1$. $\square$

Now we continue our proof of Proposition 3.13.

Proof. Take $\mathbf{A} \in \mathcal{J}_\alpha(\mathcal{G}, \mathcal{V})$ and $\mathbf{c} := (c(i))_{i\in V} \in \ell^p, 1 < p < \infty$. Then

$$\begin{aligned}
\|\mathbf{A}\mathbf{c}\|_p^p &\le \|\mathbf{A}\|^p_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})} \sum_{\lambda\in G} \Big(\sum_{i\in V} (1+\rho_{\mathcal{H}}(\lambda, i))^{-\alpha} |c(i)|\Big)^p \\
&\le \|\mathbf{A}\|^p_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})} \sum_{\lambda\in G} \Big(\sum_{i\in V} (1+\rho_{\mathcal{H}}(\lambda, i))^{-\frac{\alpha}{q}} (1+\rho_{\mathcal{H}}(\lambda, i))^{-\frac{\alpha}{p}} |c(i)|\Big)^p \\
&\le \|\mathbf{A}\|^p_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})} \sum_{\lambda\in G} \Big(\sum_{i'\in V} (1+\rho_{\mathcal{H}}(\lambda, i'))^{-\alpha}\Big)^{p-1} \Big(\sum_{i\in V} (1+\rho_{\mathcal{H}}(\lambda, i))^{-\alpha} |c(i)|^p\Big) \\
&\le \|\mathbf{A}\|^p_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})} \|\mathbf{c}\|_p^p \Big(\sup_{\lambda'\in G} \sum_{i'\in V} (1+\rho_{\mathcal{H}}(\lambda', i'))^{-\alpha}\Big)^{p-1} \Big(\sup_{i'\in V} \sum_{\lambda'\in G} (1+\rho_{\mathcal{H}}(\lambda', i'))^{-\alpha}\Big),
\end{aligned} \tag{3.57}$$

where the first inequality follows by (3.53), and the second inequality follows by Hölder's inequality.

For any $\lambda' \in G$ and $i' \in V$, it follows from Proposition 3.5 that

$$\rho_{\mathcal{G}}(\lambda', \lambda'') + 1 \ge \rho_{\mathcal{H}}(\lambda', i') \ge \rho_{\mathcal{G}}(\lambda', \lambda'') \ \text{ for all } \lambda'' \in G \text{ with } (i', \lambda'') \in T. \tag{3.58}$$

By (3.19), (3.46), (3.58) and Lemma 3.14, we obtain

$$\begin{aligned}
\sum_{i'\in V} (1+\rho_{\mathcal{H}}(\lambda', i'))^{-\alpha} &\le \sum_{\lambda''\in G} \Big(\sum_{(i',\lambda'')\in T} 1\Big) (1+\rho_{\mathcal{G}}(\lambda', \lambda''))^{-\alpha} \\
&\le L \sum_{\lambda''\in G} (1+\rho_{\mathcal{G}}(\lambda', \lambda''))^{-\alpha} \le \frac{LD_1(\mathcal{G})\alpha}{\alpha - d} \ \text{ for any } \lambda' \in G,
\end{aligned} \tag{3.59}$$

and

$$\sum_{\lambda'\in G} (1+\rho_{\mathcal{H}}(\lambda', i'))^{-\alpha} \le \sum_{\lambda'\in G} (1+\rho_{\mathcal{G}}(\lambda', \lambda''))^{-\alpha} \le \frac{D_1(\mathcal{G})\alpha}{\alpha - d} \ \text{ for any } i' \in V, \tag{3.60}$$

where $\lambda'' \in G$ satisfies $(i', \lambda'') \in T$. Combining (3.57), (3.59) and (3.60) proves (3.54) for $1 <$

$p < \infty$.

We can use similar argument to prove (3.54) for $p = 1, \infty$. $\square$

For a DSRS with its sensing matrix in $\mathcal{J}_\alpha(\mathcal{G}, \mathcal{V})$, we obtain from (3.66) and Proposition 3.13 that a signal with bounded amplitude vector generates a bounded sampling data vector.

Define band matrix approximations of a matrix $\mathbf{A} = (a(\lambda, i))_{\lambda \in G, i \in V}$ by

$$\mathbf{A}_s := (a_s(\lambda, i))_{\lambda \in G, i \in V}, \ s \geq 0, \tag{3.61}$$

where

$$a_s(\lambda, i) = \begin{cases} a(\lambda, i) & \text{if } \ \rho_{\mathcal{H}}(\lambda, i) \leq s \\ 0 & \text{if } \ \rho_{\mathcal{H}}(\lambda, i) > s. \end{cases}$$

We say a matrix $\mathbf{A}$ has *bandwidth* s if $\mathbf{A} = \mathbf{A}_s$. Clearly, any matrix $\mathbf{A}$ with bounded entries and bandwidth s belongs to Jaffard class $\mathcal{J}_\alpha(\mathcal{G}, \mathcal{V})$,

$$\|\mathbf{A}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})} \leq (s+1)^\alpha \|\mathbf{A}\|_{\mathcal{J}_0(\mathcal{G},\mathcal{V})} \ \text{ for all } \alpha \geq 0.$$

In our DSRS, the sensing matrix $\mathbf{S}$ has bandwidth s means that any agent can only detect signals at innovative positions within their geodesic distance less than or equal to s. In the next proposition, we show that matrices in the Jaffard class can be well approximated by band matrices.

Proposition 3.15. *Let graphs $\mathcal{G}$, $\mathcal{H}$, $\mathcal{V}$, d and $D_1(\mathcal{G})$ be as in Proposition 3.13. If $\mathbf{A} \in \mathcal{J}_\alpha(\mathcal{G}, \mathcal{V})$ for some $\alpha > d$, then*

$$\|(\mathbf{A} - \mathbf{A}_s)\mathbf{c}\|_p \leq \frac{D_1(\mathcal{G}) L \alpha}{\alpha - d} (s+1)^{-\alpha+d} \|\mathbf{A}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})} \|\mathbf{c}\|_p \ \text{ for all } \mathbf{c} \in \ell^p, 1 \leq p \leq \infty, \tag{3.62}$$

where $\mathbf{A}_s, s \geq 1$, are band matrices in (3.61).

Proof. Following the proof of Proposition 3.13, we obtain

$$\|(\mathbf{A}-\mathbf{A}_s)\mathbf{c}\|_p \le \|\mathbf{A}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}\Big(\sup_{\lambda'\in G}\sum_{\rho_{\mathcal{H}}(\lambda',i')>s}(1+\rho_{\mathcal{H}}(\lambda',i'))^{-\alpha}\Big)^{1-1/p} \times\Big(\sup_{i'\in V}\sum_{\rho_{\mathcal{H}}(\lambda',i')>s}(1+\rho_{\mathcal{H}}(\lambda',i'))^{-\alpha}\Big)^{1/p}\|\mathbf{c}\|_p, \tag{3.63}$$

where $\mathbf{c}\in\ell^p, 1\le p\le\infty$. Applying similar argument used to prove (3.55), (3.59) and (3.60), we have

$$\sup_{\lambda'\in G}\sum_{\rho_{\mathcal{H}}(\lambda',i')>s}(1+\rho_{\mathcal{H}}(\lambda',i'))^{-\alpha}\le L\sup_{\lambda'\in G}\sum_{\rho_{\mathcal{G}}(\lambda',\lambda'')\ge s}(1+\rho_{\mathcal{G}}(\lambda',\lambda''))^{-\alpha}\le\frac{D_1(\mathcal{G})L\alpha}{\alpha-d}(s+1)^{-\alpha+d} \tag{3.64}$$

and

$$\sup_{i'\in V}\sum_{\rho_{\mathcal{H}}(\lambda',i')>s}(1+\rho_{\mathcal{H}}(\lambda',i'))^{-\alpha}\le\frac{D_1(\mathcal{G})\alpha}{\alpha-d}(s+1)^{-\alpha+d}. \tag{3.65}$$

Then the approximation error estimate (3.62) follows from (3.63), (3.64) and (3.65). □

The above band matrix approximation property will be used later in the establishment of a local stability criterion in Section 3.5 and exponential convergence of a distributed reconstruction algorithm in Section 3.6.

3.4 Robustness of distributed sampling and reconstruction systems

Let S be the sensing matrix associated with our DSRS that has the polynomial off-diagonal decay property satisfy (3.51). The sensing matrix S characterizes the sampling procedure of signals. Applying the sensing matrix S, we obtain the sample vector $\mathbf{y}=(\langle f,\psi_\lambda\rangle)_{\lambda\in G}$ of the signal f from its amplitude vector $\mathbf{c}:=(c(i))_{i\in V}$,

$$\mathbf{y}=\mathbf{S}\mathbf{c}. \tag{3.66}$$

Under the assumptions (3.49) and (3.50), it is shown in Proposition 3.13 that a signal f with bounded amplitude vector $\mathbf{c}$ generates a bounded sample vector $\mathbf{y}$. Thus there exists a positive constant C such that

$$\|\mathbf{y}\|_\infty \leq C\|\mathbf{c}\|_\infty \text{ for all } \mathbf{c} \in \ell^\infty,$$

where for $1 \leq p \leq \infty$, ℓ^p is the space of all p-summable sequences with norm $\|\cdot\|_p$.

A fundamental problem in sampling theory is the robustness of signal reconstruction in the presence of sampling noises ([28, 66, 111, 112, 116, 122, 138]). In this paper, we consider the scenario that the sampling data $\mathbf{y} = \mathbf{Sc}$ is corrupted by bounded deterministic/random noise $\boldsymbol{\eta} = (\eta(\lambda))_{\lambda \in G}$,

$$\mathbf{z} = \mathbf{Sc} + \boldsymbol{\eta} \tag{3.67}$$

([148, 162]). We say that a reconstruction algorithm Δ is a *perfect reconstruction* in noiseless environment if

$$\Delta(\mathbf{Sc}) = \mathbf{c} \ \text{ for all } \mathbf{c} \in \ell^\infty. \tag{3.68}$$

In this section, we first study robustness of the DSRS in term of the ℓ^∞-stability. For the robustness of our DSRS, one desires that the signal reconstructed by some (non)linear algorithm Δ is a suboptimal approximation to the original signal, in the sense that the differences between their corresponding amplitude vectors $\Delta(\mathbf{z})$ and $\mathbf{c}$ are bounded by a multiple of noise level $\delta = \|\boldsymbol{\eta}\|_\infty$, i.e.,

$$\|\Delta(\mathbf{z}) - \mathbf{c}\|_\infty \leq C\delta \tag{3.69}$$

for some absolute constant C ([3, 13, 42]).

Given the noisy sampling vector $\mathbf{z}$ in (3.67), solve the following nonlinear problem of maximal

sampling error ([31, 32]),

$$\Delta_\infty(\mathbf{z}) := \mathrm{argmin}_{\mathbf{d}\in\ell^\infty} \|\mathbf{S}\mathbf{d} - \mathbf{z}\|_\infty. \tag{3.70}$$

Observe from (3.67) and (3.70) that

$$\|\mathbf{S}\Delta_\infty(\mathbf{z}) - \mathbf{S}\mathbf{c}\|_\infty \le \|\mathbf{S}\Delta_\infty(\mathbf{z}) - \mathbf{z}\|_\infty + \|\boldsymbol{\eta}\|_\infty \le \|\mathbf{S}\mathbf{c} - \mathbf{z}\|_\infty + \|\boldsymbol{\eta}\|_\infty \le 2\|\boldsymbol{\eta}\|_\infty.$$

Thus the solution of the ℓ^∞-minimization problem (3.70) gives a suboptimal approximation to the true amplitude vector $\mathbf{c}$ if the sensing matrix $\mathbf{S}$ of the DSRS has ℓ^∞-stability ([13, 150, 158]).

Proposition 3.16. *Let $\mathcal{G}$ and $\mathcal{H}$ satisfy Assumptions 1 – 4, $\mathcal{V}$ be as in* (3.17)*, $\mu_{\mathcal{G}}$ have polynomial growth with Beurling dimension d, and let $\mathbf{S}$ satisfy* (3.51)*. Then there is a reconstruction algorithm Δ with the suboptimal approximation property* (3.69) *and the perfect reconstruction property* (3.68) *if and only if $\mathbf{S}$ has ℓ^∞-stability.*

The sufficiency in Proposition 3.16 holds by taking $\Delta = \Delta_\infty$ in (3.70), while the necessity follows by applying (3.69) to $\boldsymbol{\eta} = \mathbf{S}\mathbf{d}$ with $\mathbf{d} \in \ell^\infty$.

The ℓ^∞-stability of a matrix can not be verified in a distributed manner, up to our knowledge. In the next theorem, we circumvent such a verification problem by reducing ℓ^∞-stability of a matrix in Jaffard class to its ℓ^2-stability, for which a distributed verifiable criterion will be provided in Section 3.5.

Theorem 3.17. *Let $\mathcal{G}, \mathcal{H}, \mathcal{V}$ and d be as in Proposition 3.16, and let $\mathbf{A} \in \mathcal{J}_\alpha(\mathcal{G}, \mathcal{V})$ for some $\alpha > d$. If $\mathbf{A}$ has ℓ^2-stability, then it has ℓ^p-stability for all $1 \le p \le \infty$.*

To prove Theorem 3.17, we need Theorem 3.19 and the following lemma about families $\mathcal{J}_\alpha(\mathcal{G}, \mathcal{V})$ and $\mathcal{J}_\alpha(\mathcal{V})$ of matrices.

Lemma 3.18. *Let $\mathcal{G}, \mathcal{H}, \mathcal{V}$ and d be as in Proposition 3.16. Then*

(i) $\|\mathbf{AC}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})} \le \frac{2^{\alpha+1}LD_1(\mathcal{G})\alpha}{\alpha-d}\|\mathbf{A}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}\|\mathbf{C}\|_{\mathcal{J}_\alpha(\mathcal{V})}$ *for all* $\mathbf{A} \in \mathcal{J}_\alpha(\mathcal{G},\mathcal{V})$ *and* $\mathbf{C} \in \mathcal{J}_\alpha(\mathcal{V})$.

(ii) $\|\mathbf{A}^T\mathbf{B}\|_{\mathcal{J}_\alpha(\mathcal{V})} \le \frac{2^{\alpha+1}D_1(\mathcal{G})\alpha}{\alpha-d}\|\mathbf{A}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}\|\mathbf{B}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}$ *for all* $\mathbf{A}, \mathbf{B} \in \mathcal{J}_\alpha(\mathcal{G},\mathcal{V})$.

Proof. Take $\mathbf{A} \in \mathcal{J}_\alpha(\mathcal{G},\mathcal{V})$ and $\mathbf{C} \in \mathcal{J}_\alpha(\mathcal{V})$. Observe from (3.18) that

$$\rho_{\mathcal{H}}(\lambda, i) \le \rho_{\mathcal{H}}(\lambda, i') + \rho(i', i) \text{ for all } \lambda \in G \text{ and } i, i' \in V.$$

Similar to the argument used in the proof of Proposition 3.20, we obtain

$$\|\mathbf{AC}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})} \le 2^\alpha\|\mathbf{A}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}\|\mathbf{C}\|_{\mathcal{J}_\alpha(\mathcal{V})}\Big(\sup_{i\in V}\sum_{i'\in V}(1+\rho(i',i))^{-\alpha} + \sup_{\lambda\in G}\sum_{i'\in V}(1+\rho_{\mathcal{H}}(\lambda,i'))^{-\alpha}\Big).$$

This together with (3.47), (3.64) and (3.75) proves the first conclusion.

Recall that

$$\rho(i, i') \le \rho_{\mathcal{H}}(\lambda, i) + \rho_{\mathcal{H}}(\lambda, i') \text{ for all } \lambda \in G \text{ and } i, i' \in V. \tag{3.71}$$

Then for $\mathbf{A}, \mathbf{B} \in \mathcal{J}_\alpha(\mathcal{G},\mathcal{V})$, we obtain from (3.65) and (3.71) that

$$\begin{aligned}\|\mathbf{A}^T\mathbf{B}\|_{\mathcal{J}_\alpha(\mathcal{V})} &\le 2^{\alpha+1}\|\mathbf{A}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}\|\mathbf{B}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}\sup_{i\in V}\sum_{\lambda\in G}(1+\rho_{\mathcal{H}}(\lambda,i))^{-\alpha}\\ &\le \frac{2^{\alpha+1}D_1(\mathcal{G})\alpha}{\alpha-d}\|\mathbf{A}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}\|\mathbf{B}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}.\end{aligned}$$

This completes the proof of the second conclusion. □

Now we prove Theorem 3.17.

Proof. Take $\mathbf{A} \in \mathcal{J}_\alpha(\mathcal{G}, \mathcal{V})$ that has ℓ^2-stability. Then $\mathbf{A}^T\mathbf{A}$ has bounded inverse on ℓ^2. Observe that $\mathbf{A}^T\mathbf{A} \in \mathcal{J}_\alpha(\mathcal{V})$ by Lemma 3.18. Therefore $(\mathbf{A}^T\mathbf{A})^{-1} \in \mathcal{J}_\alpha(\mathcal{V})$ and $\mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1} \in \mathcal{J}_\alpha(\mathcal{G}, \mathcal{V})$ by Theorem 3.19 and Lemma 3.18. Hence for any $\mathbf{c} \in \ell^p$,

$$\|\mathbf{c}\|_p = \|(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{A}\mathbf{c}\|_p \le \frac{D_1(\mathcal{G})L\alpha}{\alpha - d}\|\mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}\|\mathbf{A}\mathbf{c}\|_p$$

and

$$\|\mathbf{A}\mathbf{c}\|_p \le \frac{D_1(\mathcal{G})L\alpha}{\alpha - d}\|\mathbf{A}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}\|\mathbf{c}\|_p$$

by Proposition 3.13 and the dual property between sequences ℓ^p and $\ell^{p/(p-1)}$. The ℓ^p-stability for the matrix $\mathbf{A}$ then follows. □

The reader may refer to [9, 134, 147] for equivalence of ℓ^p-stability of localized matrices for different $1 \le p \le \infty$. The lower and upper ℓ^p-stability bounds of the matrix $\mathbf{A}$ depend on its ℓ^2-stability bounds and local features of the graph $\mathcal{H}$. From the proof of Theorem 3.17, we observe that they depend only on the ℓ^2-stability bounds, $\mathcal{J}_\alpha(\mathcal{G}, \mathcal{V})$-norm of the matrix $\mathbf{A}$, maximal vertex degree $\deg(\mathcal{G})$, the Beurling dimension d, the sampling density $D_1(\mathcal{G})$, and the constants L and M in (3.19) and (3.20). So the sensing matrix of our DSRS may have its ℓ^p-stability bounds independent of the size of the DSRS.

For the graph $\mathcal{V}$ in (3.17) and the distance ρ in (3.24), define

$$\mathcal{J}_\alpha(\mathcal{V}) := \big\{\mathbf{A} := (a(i,i'))_{i,i' \in V},\ \|\mathbf{A}\|_{\mathcal{J}_\alpha(\mathcal{V})} < \infty\big\}, \tag{3.72}$$

where

$$\|\mathbf{A}\|_{\mathcal{J}_\alpha(\mathcal{V})} := \sup_{i,i' \in V}(1 + \rho(i,i'))^\alpha |a(i,i')|,\ \alpha \ge 0. \tag{3.73}$$

The proof of Theorem 3.17 depends highly on the following Wiener's lemma for the matrix algebra

$\mathcal{J}_\alpha(\mathcal{V}), \alpha > d$.

Theorem 3.19. *Let $\mathcal{V}$ be as in* (3.17) *and its counting measure μ satisfy* (3.29). *If $\mathbf{A} \in \mathcal{J}_\alpha(\mathcal{V}), \alpha > d$, and $\mathbf{A}^{-1}$ is bounded on ℓ^2, then $\mathbf{A}^{-1} \in \mathcal{J}_\alpha(\mathcal{V})$ too.*

To prove Wiener's lemma (Theorem 3.19) for $\mathcal{J}_\alpha(\mathcal{V}), \alpha > d$, we first show that it is a Banach algebra of matrices.

Proposition 3.20. *Let $\mathcal{V}$ be an undirected graph with the counting measure μ having polynomial growth* (3.29). *Then for any $\alpha > d$, $\mathcal{J}_\alpha(\mathcal{V})$ is a Banach algebra of matrices:*

(i) $\|\beta \mathbf{C}\|_{\mathcal{J}_\alpha(\mathcal{V})} = |\beta| \|\mathbf{C}\|_{\mathcal{J}_\alpha(\mathcal{V})}$;

(ii) $\|\mathbf{C} + \mathbf{D}\|_{\mathcal{J}_\alpha(\mathcal{V})} \leq \|\mathbf{C}\|_{\mathcal{J}_\alpha(\mathcal{V})} + \|\mathbf{D}\|_{\mathcal{J}_\alpha(\mathcal{V})}$;

(iii) $\|\mathbf{CD}\|_{\mathcal{J}_\alpha(\mathcal{V})} \leq \frac{2^{\alpha+1} D_1 \alpha}{\alpha - d} \|\mathbf{C}\|_{\mathcal{J}_\alpha(\mathcal{V})} \|\mathbf{D}\|_{\mathcal{J}_\alpha(\mathcal{V})}$; *and*

(iv) $\|\mathbf{Dc}\|_2 \leq \frac{D_1 \alpha}{\alpha - d} \|\mathbf{D}\|_{\mathcal{J}_\alpha(\mathcal{V})} \|\mathbf{c}\|_2$

for any scalar β, vector $\mathbf{c} \in \ell^2$ and matrices $\mathbf{C}, \mathbf{D} \in \mathcal{J}_\alpha(\mathcal{V})$.

Proof. The first two conclusions follow immediately from (3.72) and (3.73).

Now we prove the third conclusion. Take $\mathbf{C}, \mathbf{D} \in \mathcal{J}_\alpha(\mathcal{V})$. Then

$$\|\mathbf{CD}\|_{\mathcal{J}_\alpha(\mathcal{V})} \leq 2^\alpha \|\mathbf{C}\|_{\mathcal{J}_\alpha(\mathcal{V})} \|\mathbf{D}\|_{\mathcal{J}_\alpha(\mathcal{V})} \sup_{i,i' \in V} \Big(\sum_{\rho(i,i'') \geq \rho(i,i')/2} (1 + \rho(i'', i'))^{-\alpha} + \sum_{\rho(i'',i') \geq \rho(i,i')/2} (1 + \rho(i, i''))^{-\alpha} \Big). \tag{3.74}$$

Following the argument used in the proofs of Lemma 3.14, we have

$$\sup_{i\in V}\sum_{\rho(i,i')\geq s}(1+\rho(i,i'))^{-\alpha}\leq \frac{D_1\alpha}{\alpha-d}(s+1)^{-\alpha+d},\ 0\leq s\in\mathbb{Z}. \tag{3.75}$$

Combining (3.74) and (3.75) proves the third conclusion.

Following the proof of Proposition 3.13 and applying (3.75) instead of (3.59) and (3.60), we obtain the fourth conclusion. □

Now, we prove Theorem 3.19.

Proof. Following the argument in [140], it suffices to establish the following differential norm inequality:

$$\|\mathbf{C}^2\|_{\mathcal{J}_\alpha(\mathcal{V})}\leq 2^{\alpha+d/2+2}D_1^{1/2}(D_1\alpha/(\alpha-d))^{1-\theta}(\|\mathbf{C}\|_{\mathcal{J}_\alpha(\mathcal{V})})^{2-\theta}(\|\mathbf{C}\|_{\mathcal{B}^2})^{\theta} \tag{3.76}$$

holds for all $\mathbf{C}\in\mathcal{J}_\alpha(\mathcal{V})$, where $\theta=(2\alpha-2d)/(2\alpha-d)\in(0,1)$.

Write $\mathbf{C}=(c(i,i'))_{i,i'\in V}$. Then

$$\begin{aligned}\|\mathbf{C}^2\|_{\mathcal{J}_\alpha(\mathcal{V})}&\leq 2^\alpha\|\mathbf{C}\|_{\mathcal{J}_\alpha(\mathcal{V})}\Big(\sup_{i,i'\in V}\sum_{\rho(i,i'')\geq\rho(i,i')/2}|c(i'',i')|+\sup_{i,i'\in V}\sum_{\rho(i'',i')\geq\rho(i,i')/2}|c(i,i'')|\Big)\\&\leq 2^\alpha\|\mathbf{C}\|_{\mathcal{J}_\alpha(\mathcal{V})}\Big(\sup_{i'\in V}\sum_{i''\in V}|c(i'',i')|+\sup_{i\in V}\sum_{i''\in V}|c(i,i'')|\Big).\end{aligned} \tag{3.77}$$

Set

$$\tau:=\Big(\frac{D_1\alpha\|\mathbf{C}\|_{\mathcal{J}_\alpha(\mathcal{V})}}{(\alpha-d)\|\mathbf{C}\|_{\mathcal{B}^2}}\Big)^{2/(2\alpha-d)}\geq 1 \tag{3.78}$$

by Proposition 3.20. For $i' \in V$, we obtain

$$\begin{aligned}\sum_{i'' \in V} |c(i'', i')| &\le \Big(\sum_{\rho(i'', i') \le \tau} |c(i'', i')|^2 \Big)^{1/2} \Big(\sum_{\rho(i'', i') \le \tau} 1 \Big)^{1/2} + \|\mathbf{C}\|_{\mathcal{J}_\alpha(\mathcal{V})} \sum_{\rho(i'', i') > \tau} (1 + \rho(i'', i'))^{-\alpha} \\ &\le D_1^{1/2} \|\mathbf{C}\|_{\mathcal{B}^2} (1 + \lfloor \tau \rfloor)^{d/2} + D_1 \alpha (\alpha - d)^{-1} \|\mathbf{C}\|_{\mathcal{J}_\alpha(\mathcal{V})} (1 + \lfloor \tau \rfloor)^{-\alpha + d} \\ &\le 2^{d/2+1} D_1^{1/2} (D_1 \alpha / (\alpha - d))^{d/(2\alpha - d)} (\|\mathbf{C}\|_{\mathcal{J}_\alpha(\mathcal{V})})^{1-\theta} (\|\mathbf{C}\|_{\mathcal{B}^2})^{\theta}, \qquad (3.79)\end{aligned}$$

where the second inequality holds by (3.75) and the last inequality follows from (3.78). Similarly, for $i \in V$ we have

$$\sum_{i'' \in V} |c(i', i'')| \le 2^{d/2+1} D_1^{1/2} (D_1 \alpha / (\alpha - d))^{d/(2\alpha - d)} (\|\mathbf{C}\|_{\mathcal{J}_\alpha(\mathcal{V})})^{1-\theta} (\|\mathbf{C}\|_{\mathcal{B}^2})^{\theta}. \qquad (3.80)$$

Combining (3.77), (3.79) and (3.80) proves (3.76). This completes the proof of Theorem 3.19. □

Wiener's lemma has been established for infinite matrices, pseudodifferential operators, and integral operators satisfying various off-diagonal decay conditions ([18, 68, 82, 84, 85, 93, 140, 142, 144, 147]). It has been shown to be crucial for well-localization of dual Gabor/wavelet frames, fast implementation in numerical analysis, local reconstruction in sampling theory, local features of spatially distributed optimization, etc. The reader may refer to the survey papers [83, 103] for historical remarks, motivation and recent advances.

The Wiener's lemma (Theorem 3.19) is also used to establish the sub-optimal approximation property (3.69) for the "least squares" solution $\Delta_2(\mathbf{z})$ in (3.96), for which a distributed algorithm is proposed in Section 3.6.

Theorem 3.21. *Let $\mathcal{G}, \mathcal{H}$ and $\mathcal{V}$ be as in Proposition 3.16. Assume that the sensing matrix* $\mathbf{S}$

satisfies (3.51) *and it has* ℓ^2*-stability. Then there exists a positive constant* C *such that*

$$\|\Delta_2(\mathbf{z}) - \mathbf{c}\|_\infty \le C\|\boldsymbol{\eta}\|_\infty \quad \text{for all } \mathbf{c}, \boldsymbol{\eta} \in \ell^\infty, \tag{3.81}$$

where $\mathbf{z} = \mathbf{S}\mathbf{c} + \boldsymbol{\eta}$.

Proof. The conclusion (3.81) follows immediately from Proposition 3.13, Theorem 3.19 and Lemma 3.18. □

3.5 Stability criterion for distributed sampling and reconstruction system

Let $\mathcal{H}$ be the connected simple graph in (3.15) to describe our DSRS. Given $\lambda' \in G$ and a positive integer N, define *truncation operators* $\chi^N_{\lambda',G}$ and $\chi^N_{\lambda',V}$ by

$$\chi^N_{\lambda',G}: \ \ell^p(G) \ni (d(\lambda))_{\lambda\in G} \longmapsto \big(d(\lambda)\chi_{B_{\mathcal{H}}(\lambda',N)\cap G}(\lambda)\big)_{\lambda\in G} \in \ell^p(G)$$

and

$$\chi^N_{\lambda',V}: \ \ell^p(V) \ni (c(i))_{i\in V} \longmapsto \big(c(i)\chi_{B_{\mathcal{H}}(\lambda',N)\cap V}(i)\big)_{i\in V} \in \ell^p(V),$$

where $1 \le p \le \infty$ and

$$B_{\mathcal{H}}(u, r) := \{v \in G \cup V, \ \rho_{\mathcal{H}}(u, v) \le r\}$$

is the closed ball in $\mathcal{H}$ with center $u \in \mathcal{H}$ and radius $r \ge 0$.

For any matrix $\mathbf{A} \in \mathcal{J}_\alpha(\mathcal{G}, \mathcal{V})$ with ℓ^2-stability, we observe that its quasi-main submatrices $\chi^{2N}_\lambda \mathbf{A}\chi^N_\lambda, \lambda \in G$, of size $O(N^d)$ have uniform ℓ^2-stability for large N.

Theorem 3.22. *Let* $\mathcal{G}$ *and* $\mathcal{H}$ *satisfy Assumptions 1 – 4,* $\mathcal{V}$ *be as in* (3.17)*,* $\mu_{\mathcal{G}}$ *have polynomial growth with Beurling dimension* d *and sampling density* $D_1(\mathcal{G})$*, and let* $\mathbf{A} \in \mathcal{J}_\alpha(\mathcal{G}, \mathcal{V})$ *for some*

$\alpha > d$. If $\mathbf{A}$ has ℓ^2-stability with lower bound $A\|\mathbf{A}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}$, then

$$\|\chi_{\lambda,G}^{2N}\mathbf{A}\chi_{\lambda,V}^{N}\mathbf{c}\|_2 \ge \frac{A}{2}\|\mathbf{A}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}\|\chi_{\lambda,V}^{N}\mathbf{c}\|_2,\ \mathbf{c}\in\ell^2 \tag{3.82}$$

for all $\lambda \in G$ and all integers N satisfying

$$2D_1(\mathcal{G})N^{-\alpha+d}\sqrt{L\alpha/(\alpha-d)} \le A. \tag{3.83}$$

Proof. Observe from Proposition 3.5 that

$$B_{\mathcal{H}}(\gamma,r)\cap G = \{\gamma'\in G,\ \rho_{\mathcal{G}}(\gamma,\gamma')\le r\},\ \gamma\in G.$$

and

$$B_{\mathcal{H}}(i,r)\cap V = \{i'\in V,\ \rho(i,i')\le \max(r-1,0)\},\ i\in V.$$

Take $\mathbf{c}=(c(i))_{i\in V}$ supported in $B_{\mathcal{H}}(\lambda,N)\cap V$ and write $\mathbf{Ac}=(d(\lambda'))_{\lambda'\in G}$. Then

$$\|\mathbf{Ac}\|_2 \ge A\|\mathbf{A}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}\|\mathbf{c}\|_2 \tag{3.84}$$

and

$$\begin{aligned}\sum_{\rho_{\mathcal{H}}(\lambda',\lambda)>2N}|d(\lambda')|^2 &\le LD_1(\mathcal{G})N^{-\alpha+d}\|\mathbf{A}\|^2_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}\\ &\quad\times\sum_{\rho_{\mathcal{H}}(\lambda',\lambda)>2N}\ \sum_{i\in B_{\mathcal{H}}(\lambda,N)\cap V}(1+\rho_{\mathcal{H}}(\lambda',i))^{-\alpha}|c(i)|^2\\ &\le \big(D_1(\mathcal{G})\big)^2LN^{-2\alpha+2d}\alpha(\alpha-d)^{-1}\|\mathbf{A}\|^2_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}\|\mathbf{c}\|_2^2,\end{aligned} \tag{3.85}$$

where the first inequality holds as

$$\rho_{\mathcal{H}}(\lambda', i') \geq \rho_{\mathcal{H}}(\lambda', \lambda) - \rho_{\mathcal{H}}(i', \lambda) > N$$

for all $\lambda' \notin B_{\mathcal{H}}(\lambda, 2N)$ and $i' \in B_{\mathcal{H}}(\lambda, N)$, and the last inequality follows from (3.65). Combining (3.84) and (3.85) proves (3.82). $\square$

The above theorem provides a guideline to design a distributed algorithm for signal reconstruction, see Section 3.6. Surprisingly, the converse of Theorem 3.22 is true, cf. the stability criterion in [146, Theorem 2.1] for convolution-dominated matrices.

Theorem 3.23. *Let $\mathcal{G}, \mathcal{H}, \mathcal{V}$ be as in Theorem 3.22, and $\mathbf{A} \in \mathcal{J}_\alpha(\mathcal{G}, \mathcal{V})$ for some $\alpha > d$. If there exist a positive constant A_0 and an integer $N_0 \geq 3$ such that*

$$A_0 \geq 4(D_0(\mathcal{G}))^2 D_1(\mathcal{G}) L N_0^{-\min(\alpha-d,1)} \times \begin{cases} \left(\frac{4\alpha}{3(\alpha-d)} + \frac{2(\alpha-1)(\alpha-d)}{\alpha-d-1}\right) & \text{if } \alpha > d+1 \\ \left(\frac{10(d+1)}{3} + 2d \ln N_0\right) & \text{if } \alpha = d+1 \\ \left(\frac{4\alpha}{3(\alpha-d)} + \frac{4d}{d+1-\alpha}\right) & \text{if } \alpha < d+1, \end{cases} \tag{3.86}$$

and for all $\lambda \in G$,

$$\|\chi_{\lambda,G}^{2N_0} \mathbf{A} \chi_{\lambda,V}^{N_0} \mathbf{c}\|_2 \geq A_0 \|\mathbf{A}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})} \|\chi_{\lambda,V}^{N_0} \mathbf{c}\|_2, \ \mathbf{c} \in \ell^2, \tag{3.87}$$

then $\mathbf{A}$ has ℓ^2-stability,

$$\|\mathbf{A}\mathbf{c}\|_2 \geq \frac{A_0 \|\mathbf{A}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}}{12(D_0(\mathcal{G}))^2} \|\mathbf{c}\|_2, \ \mathbf{c} \in \ell^2. \tag{3.88}$$

Here, we will prove the following strong version of Theorem 3.23.

Theorem 3.24. *Let $\mathcal{G}, \mathcal{H}, \mathcal{V}$ and $\mathbf{A}$ be as in Theorem 3.23. If there exists a positive constant A_0, an integer $N_0 \geq 3$, and a maximal $\frac{N_0}{4}$-disjoint subset $G_{N_0/4}$ such that* (3.86) *is true and* (3.87) *hold*

for all $\lambda_m \in G_{N_0/4}$*, then* $\mathbf{A}$ *satisfies* (3.88).

Proof. Let ψ_0 be the trapezoid function,

$$\psi_0(t) = \begin{cases} 1 & \text{if } |t| \leq 1/2 \\ 2 - 2|t| & \text{if } 1/2 < |t| \leq 1 \\ 0 & \text{if } |t| > 1. \end{cases} \tag{3.89}$$

For $\lambda \in G$, define multiplication operators $\Psi^N_{\lambda,V}$ and $\Psi^N_{\lambda,G}$ by

$$\Psi^N_{\lambda,V} : \ (c(i))_{i \in V} \longmapsto \big(\psi_0(\rho_{\mathcal{H}}(\lambda, i)/N) c(i)\big)_{i \in V}, \tag{3.90}$$

$$\Psi^N_{\lambda,G} : \ (d(\lambda'))_{\lambda' \in G} \longmapsto \big(\psi_0(\rho_{\mathcal{H}}(\lambda, \lambda')/N) d(\lambda')\big)_{\lambda' \in G}. \tag{3.91}$$

Observe that

$$\mathbf{A}_N \Psi^N_{\lambda,V} = \mathbf{A}_N \chi^N_{\lambda,V} \Psi^N_{\lambda,V} = \chi^{2N}_{\lambda,G} \mathbf{A}_N \chi^N_{\lambda,V} \Psi^N_{\lambda,V}, N \geq 0,$$

where $\mathbf{A}_N$ is a band approximation of the matrix $\mathbf{A}$ in (3.61). Then for all $\lambda_m \in G_{N_0/4}$, it follows from Proposition 3.15 and our local stability assumption (3.87) that

$$\begin{aligned} \|\mathbf{A}_{N_0} \Psi^{N_0}_{\lambda_m,V} \mathbf{c}\|_2 &\geq \|\chi^{2N_0}_{\lambda_m,G} \mathbf{A} \chi^{N_0}_{\lambda_m,V} \Psi^{N_0}_{\lambda_m,V} \mathbf{c}\|_2 - \|\chi^{2N_0}_{\lambda_m,G} (\mathbf{A} - \mathbf{A}_{N_0}) \Psi^{N_0}_{\lambda_m,V} \mathbf{c}\|_2 \\ &\geq \Big(A_0 - \frac{D_1(\mathcal{G}) L \alpha}{\alpha - d} N_0^{-\alpha+d}\Big) \|\mathbf{A}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})} \|\Psi^{N_0}_{\lambda_m,V} \mathbf{c}\|_2, \ \ \mathbf{c} \in \ell^2. \end{aligned}$$

Therefore

$$
\begin{aligned}
& \Big(\sum_{\lambda_m \in G_{N_0/4}} \|\mathbf{A}_{N_0} \Psi_{\lambda_m,V}^{N_0} \mathbf{c}\|_2^2 \Big)^{1/2} \\
\geq & \ \Big(A_0 - \frac{D_1(\mathcal{G}) L \alpha}{\alpha - d} N_0^{-\alpha+d} \Big) \|\mathbf{A}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})} \Big(\sum_{\lambda_m \in G_{N_0/4}} \|\Psi_{\lambda_m,V}^{N_0} \mathbf{c}\|_2^2 \Big)^{1/2} \\
\geq & \ \Big(\frac{A_0}{3} - \frac{D_1(\mathcal{G}) L \alpha}{3(\alpha - d)} N_0^{-\alpha+d} \Big) \|\mathbf{A}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})} \|\mathbf{c}\|_2,
\end{aligned} \tag{3.92}
$$

where the last inequality holds because for all $i \in V$,

$$
\sum_{\lambda_m \in G_{N_0/4}} |\psi_0(\rho_{\mathcal{H}}(\lambda_m, i)/N_0)|^2 \geq \Big(\frac{N_0 - 2}{N_0} \Big)^2 \sum_{\lambda_m \in G_{N_0/4}} \chi_{B_{\mathcal{H}}(\lambda_m, N_0/2+1)}(i) \geq \frac{1}{9}
$$

by (3.89), Proposition 3.4 and the assumption that $N_0 \geq 3$.

Next, we estimate commutators

$$
\mathbf{A}_{N_0} \Psi_{\lambda_m,V}^{N_0} - \Psi_{\lambda_m,G}^{N_0} \mathbf{A}_{N_0} = (\mathbf{A}_{N_0} \Psi_{\lambda_m,V}^{N_0} - \Psi_{\lambda_m,G}^{N_0} \mathbf{A}_{N_0}) \chi_{\lambda_m,V}^{2N_0}, \ \lambda_m \in G_{N_0/4}.
$$

Take $\mathbf{c} = (c(i))_{i \in V} \in \ell^2$. Then

$$
\begin{aligned}
& \sum_{\lambda_m \in G_{N_0/4}} \|(\mathbf{A}_{N_0} \Psi_{\lambda_m,V}^{N_0} - \Psi_{\lambda_m,G}^{N_0} \mathbf{A}_{N_0}) \mathbf{c}\|_2^2 \\
& \leq \|\mathbf{A}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}^2 \sum_{\lambda_m \in G_{N_0/4}} \sum_{\lambda \in G} \Big\{ \sum_{\rho_{\mathcal{H}}(\lambda,i) \leq N_0} (1 + \rho_{\mathcal{H}}(\lambda, i))^{-\alpha} \\
& \quad \times \Big| \psi_0\Big(\frac{\rho_{\mathcal{H}}(\lambda, \lambda_m)}{N_0} \Big) - \psi_0\Big(\frac{\rho_{\mathcal{H}}(i, \lambda_m)}{N_0} \Big) \Big| \chi_{B_{\mathcal{H}}(\lambda_m, 2N_0) \cap V}(i) |c(i)| \Big\}^2 \\
& \leq 4(D_0(\mathcal{G}))^4 N_0^{-2} \|\mathbf{A}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}^2 \Big(\sup_{i \in V} \sum_{\lambda \in B_{\mathcal{H}}(i, N_0) \cap G} (1 + \rho_{\mathcal{H}}(\lambda, i))^{-\alpha} \rho_{\mathcal{H}}(\lambda, i) \Big) \\
& \quad \times \Big(\sup_{\lambda \in G} \sum_{i \in B_{\mathcal{H}}(\lambda, N_0) \cap V} (1 + \rho_{\mathcal{H}}(\lambda, i))^{-\alpha} \rho_{\mathcal{H}}(\lambda, i) \Big) \|\mathbf{c}\|_2^2,
\end{aligned} \tag{3.93}
$$

where the last inequality follows from Propositions 3.4 and 3.5, and

$$|\psi_0(t) - \psi_0(t')| \le 2|t - t'| \ \text{ for all } t, t' \in \mathbb{R}.$$

Following the argument used in (3.55), we have

$$\begin{aligned}
&\sup_{i\in V} \sum_{\lambda\in B_{\mathcal{H}}(i,N_0)\cap G} (1+\rho_{\mathcal{H}}(\lambda,i))^{-\alpha}\rho_{\mathcal{H}}(\lambda,i)\\
&\le \sup_{\lambda'\in G} \sum_{\rho_{\mathcal{G}}(\lambda,\lambda')\le N_0} (1+\rho_{\mathcal{G}}(\lambda,\lambda'))^{-\alpha+1}\\
&\le D_1(\mathcal{G})(N_0+1)^{-\alpha+d+1} + (\alpha-1)D_1(\mathcal{G})\sum_{n=0}^{N_0-1}(n+1)^{-\alpha+d}\\
&\le D_1(\mathcal{G})(N_0+1)^{-\alpha+d+1} + D_1(\mathcal{G})(\alpha-1)\Big(1+\int_1^{N_0} t^{-\alpha+d}dt\Big)\\
&\le \begin{cases} \frac{D_1(\mathcal{G})(\alpha-1)(\alpha-d)}{\alpha-d-1} & \text{if } \alpha > d+1\\ D_1(\mathcal{G})(1+d+d\ln N_0) & \text{if } \alpha = d+1\\ \frac{2^{d+1-\alpha}D_1(\mathcal{G})d}{d+1-\alpha}N_0^{d+1-\alpha} & \text{if } \alpha < d+1\end{cases}
\end{aligned} \tag{3.94}$$

and

$$\begin{aligned}
&\sup_{\lambda\in G} \sum_{i\in B_{\mathcal{H}}(\lambda,N_0)\cap V} (1+\rho_{\mathcal{H}}(\lambda,i))^{-\alpha}\rho_{\mathcal{H}}(\lambda,i)\\
\le\ & L\sup_{\lambda\in G} \sum_{\lambda'\in B_{\mathcal{G}}(\lambda,N_0)} (1+\rho_{\mathcal{G}}(\lambda,\lambda'))^{-\alpha+1}\\
\le\ & \begin{cases} \frac{D_1(\mathcal{G})L(\alpha-1)(\alpha-d)}{\alpha-d-1} & \text{if } \alpha > d+1\\ D_1(\mathcal{G})L(1+d+d\ln N_0) & \text{if } \alpha = d+1\\ \frac{2^{d+1-\alpha}D_1(\mathcal{G})dL}{d+1-\alpha}N_0^{d+1-\alpha} & \text{if } \alpha < d+1.\end{cases}
\end{aligned} \tag{3.95}$$

Therefore,

$$
\begin{aligned}
(D_0(\mathcal{G}))^2\|\mathbf{A}_{N_0}\mathbf{c}\|_2 &\ge \Big(\sum_{\lambda_m\in G_{N_0/4}} \|\Psi^{N_0}_{\lambda_m,G}\mathbf{A}_{N_0}\mathbf{c}\|_2^2\Big)^{1/2}\\
&\ge \Big(\sum_{\lambda_m\in G_{N_0/4}} \|\mathbf{A}_{N_0}\Psi^{N_0}_{\lambda_m,V}\mathbf{c}\|_2^2\Big)^{1/2} - \Big(\sum_{\lambda_m\in G_{N_0/4}} \|(\mathbf{A}_{N_0}\Psi^{N_0}_{\lambda_m,V}-\Psi^{N_0}_{\lambda_m,G}\mathbf{A}_{N_0})\mathbf{c}\|_2^2\Big)^{1/2}\\
&\ge \frac{A_0\|\mathbf{A}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}}{3}\|\mathbf{c}\|_2 - D_1(\mathcal{G})L\|\mathbf{A}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}N_0^{-\min(\alpha-d,1)}\|\mathbf{c}\|_2\\
&\quad\times\begin{cases} \Big(\frac{\alpha}{3(\alpha-d)}+\frac{2(D_0(\mathcal{G}))^2(\alpha-1)(\alpha-d)}{\alpha-d-1}\Big) & \text{if } \alpha>d+1\\ \Big(\frac{d+1}{3}+2(D_0(\mathcal{G}))^2(1+d+d\ln N_0)\Big) & \text{if } \alpha=d+1\\ \Big(\frac{\alpha}{3(\alpha-d)}+\frac{4(D_0(\mathcal{G}))^2 d}{d+1-\alpha}\Big) & \text{if } \alpha<d+1,\end{cases}
\end{aligned}
$$

where the first inequality holds by Proposition 3.4, and the third inequality follows from (3.92) and (3.93). This together with Proposition 3.15 completes the proof. $\square$

Observe that the right hand side of (3.86) could be arbitrarily small when N_0 is sufficiently large. This together with Theorem 3.22 implies that the requirements (3.86) and (3.87) are necessary for the ℓ^2-stability property of any matrix in $\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})$. As shown in the example below, the term $N_0^{-\min(\alpha-d,1)}$ in (3.86) cannot be replaced by $N_0^{-\beta}$ with high order $\beta>1$ even if the matrix $\mathbf{A}$ has finite bandwidth.

Example 3.25. Let $\mathbf{A}_0=(a_0(i-j))_{i,j\in\mathbb{Z}}$ be the bi-infinite Toeplitz matrix with symbol $\sum_{k\in\mathbb{Z}} a_0(k)e^{-ik\xi}=1-e^{-i\xi}$. Then $\mathbf{A}_0$ belongs to the Jaffard class $\mathcal{J}_\alpha(\mathbb{Z},\mathbb{Z})$ for all $\alpha\ge 0$ and it does not have ℓ^2-stability. On the other hand, for any $\lambda\in G=V=\mathbb{Z}$ and $N_0\ge 1$,

$$
\begin{aligned}
\inf_{\|\chi^{N_0}_{\lambda,V}\mathbf{c}\|_2=1}\|\chi^{2N_0}_{\lambda,G}\mathbf{A}_0\chi^{N_0}_{\lambda,V}\mathbf{c}\|_2 &= \inf_{\|\chi^{N_0}_{\lambda,V}\mathbf{c}\|_2=1}\|\mathbf{A}_0\chi^{N_0}_{\lambda,V}\mathbf{c}\|_2\\
&= \inf_{|d_1|^2+\cdots+|d_{2N_0+1}|^2=1}\sqrt{|d_1|^2+|d_1-d_2|^2+\cdots+|d_{2N_0}-d_{2N_0+1}|^2+|d_{2N_0+1}|^2}\\
&= 2\sin\frac{\pi}{4N_0+4}\ge\frac{1}{2}N_0^{-1},
\end{aligned}
$$

where the last equality follows from [100, Lemma 1 of Chapter 9].

For our DSRS with sensing matrix $\mathbf{S}$ having the polynomial off-diagonal decay property (3.51), the uniform stability property (3.87) could be verified by finding minimal eigenvalues of its quasi-main submatrices $\chi_{\lambda,V}^{N_0}\mathbf{S}^T\chi_{\lambda,G}^{2N_0}\mathbf{S}\chi_{\lambda,V}^{N_0}$, $\lambda \in G$, of size about $O(N_0^d)$. The above verification could be implemented on agents in the DSRS via its computing and communication abilities. This provides a practical tool to verify ℓ^2-stability of a DSRS and to design a robust (dynamic) DSRS against supplement, replacement and impairment of agents.

3.6 Exponential convergence of a distributed reconstruction algorithm

In this section, we consider signal reconstructions in a distributed manner, under the assumption that the sensing matrix $\mathbf{S}$ of our DSRS has ℓ^2-stability. For centralized signal reconstruction systems, there are many robust algorithms, such as the frame algorithm and the approximation-projection algorithm, to approximate signals from their (non)linear noisy sampling data ([11, 42, 47, 64, 70, 116, 141, 148]). In this paper, we develop a distributed algorithm to find the suboptimal approximation

$$\Delta_2(\mathbf{z}) := (\mathbf{S}^T\mathbf{S})^{-1}\mathbf{S}^T\mathbf{z} \tag{3.96}$$

to the original signal f in (3.14). For the case that our DSRS has finitely many agents (which is the case in most of practical applications), the suboptimal approximation $\Delta_2(\mathbf{z})$ in (3.96) is the unique least squares solution,

$$\Delta_2(\mathbf{z}) = \mathrm{argmin}_{\mathbf{d}\in\ell^2} \|\mathbf{S}\mathbf{d}-\mathbf{z}\|_2^2 = \mathrm{argmin}_{\mathbf{d}\in\ell^2} \sum_{\lambda\in G} f_\lambda(\mathbf{d},\mathbf{z}), \tag{3.97}$$

where $\mathbf{d} = (d(i))_{i\in V}$, $\mathbf{z} = (z(\lambda))_{\lambda\in G}$, and

$$f_\lambda(\mathbf{d}, \mathbf{z}) = \Big|\sum_{i\in V}\langle\varphi_i, \psi_\lambda\rangle d(i) - z(\lambda)\Big|^2,\ \lambda \in G. \tag{3.98}$$

As our SDS has strict constraints in its data processing power and communication bandwidth, we need develop distributed algorithms to solve the optimization problem

$$\min \sum_{\lambda\in G} f_\lambda(\mathbf{d}, \mathbf{z}). \tag{3.99}$$

For the case that $G = V$ and the sensing matrix $\mathbf{S}$ is strictly diagonally dominant, the Jacobi iterative method,

$$\begin{cases} d_1(\lambda) = 0 \\ d_{n+1}(\lambda) = (\langle\varphi_\lambda, \psi_\lambda\rangle)^{-1}\big(\sum_{i\neq\lambda}\langle\varphi_i, \psi_\lambda\rangle d_n(i) - z(\lambda)\big) \\ \qquad\qquad = \operatorname{argmin}_{t\in\mathbb{R}} f_\lambda(\mathbf{d}_{n;t,\lambda}, \mathbf{z}), \lambda \in G = V,\ n \geq 1, \end{cases}$$

is a distributed algorithm to solve the minimization problem (3.99), where $\mathbf{d}_{n;t,\lambda}$ is obtained from $\mathbf{d}_n = (d_n(i))_{i\in V}$ by replacing its λ-component $d_n(\lambda)$ with t. The reader may refer to [27, 39, 102, 108, 118] and references therein for historical remarks, motivations, applications and recent advances on distributed algorithms, especially for the case that $G = V$.

In our DSRS, the set G of agents is not necessarily the same as the set V of innovative positions, and even for the case that the sets G and V are the same, the sensing matrix $\mathbf{S}$ need not be strictly diagonally dominant in general. In this paper, we introduce a distributed algorithm (3.132) and (3.133) to approximate $\Delta_2(\mathbf{z})$ in (3.96), when the sensing matrix $\mathbf{S}$ has ℓ^2-stability and satisfies the requirements (3.48) and (3.49). In the above distributed algorithm for signal reconstruction, each agent in the SDS collects noisy observations of neighboring agents, then interacts with its neighbors per iteration, and continues the above recursive procedure until arriving at an accurate

approximation to the solution $\Delta_2(z)$ in (3.96). More importantly, we show in Theorems 3.26 and 3.28 that the proposed distributed algorithm (3.132) and (3.133) converges exponentially to the solution $\Delta_2(z)$ in (3.96). The establishment for the above convergence is virtually based on Wiener's lemma for localized matrices ([84, 85, 93, 140, 142, 147]) and on the observation that our sensing matrices are quasi-diagonal block dominated.

In our DSRS, agents could essentially catch signals not far from their locations. So one may expect that a signal near any innovative position should substantially be determined by sampling data of neighboring agents, while data from distant agents should have (almost) no influence in the reconstruction. The most desirable method to meet the above expectation is local exact reconstruction, which could be implemented in a distributed manner without iterations ([12, 86, 145, 152]). In such a linear reconstruction procedure, there is a left-inverse $\mathbf{T}$ of the sensing matrix $\mathbf{S}$ with finite bandwidth,

$$\mathbf{TS} = \mathbf{I}.$$

For our DSRS, such a left-inverse $\mathbf{T}$ with finite bandwidth may not exist and/or it is difficult to find even it exists. We observe that

$$\mathbf{S}^\dagger := (\mathbf{S}^T\mathbf{S})^{-1}\mathbf{S}^T$$

is a left-inverse well approximated by matrices with finite bandwidth, and

$$\mathbf{d}_2 = \mathbf{S}^\dagger\mathbf{z} \tag{3.100}$$

is a suboptimal approximation, where $\mathbf{z}$ is given in (3.67). However, it is infeasible to find the pseudo-inverse $\mathbf{S}^\dagger$, because the DSRS does not have a central processor and it has huge amounts of agents and large number of innovative positions. In this section, we introduce a distributed algorithm to find the suboptimal approximation $\mathbf{d}_2$ in (3.100).

Let $\mathcal{H}$ be the connected simple graph in (3.15) to describe our DSRS, and the sensing matrix $\mathbf{S} \in \mathcal{J}_\alpha(\mathcal{G}, \mathcal{V}), \alpha > d$, have ℓ^2-stability. Then $\mathbf{d}_2$ in (3.100) is the unique solution to the "normal" equation

$$\mathbf{S}^T\mathbf{S}\mathbf{d}_2 = \mathbf{S}^T\mathbf{z}. \tag{3.101}$$

As principal submatrices $\chi_{\lambda,V}^N\mathbf{S}^T\mathbf{S}\chi_{\lambda,V}^N$ of the positive definite matrix $\mathbf{S}^T\mathbf{S}$ are uniformly stable, we solve localized linear systems

$$\chi_{\lambda,V}^N\mathbf{S}^T\mathbf{S}\chi_{\lambda,V}^N\mathbf{d}_{\lambda,N} = \chi_{\lambda,V}^N\mathbf{S}^T\mathbf{z},\ \lambda \in G, \tag{3.102}$$

of size $O(N^d)$, whose solutions $\mathbf{d}_{\lambda,N}$ are supported in the ball $B_{\mathcal{H}}(\lambda, N)\cap V$. One of **crucial** results of this paper is that for large integer N, the solution $\mathbf{d}_{\lambda,N}$ provides a reasonable approximation of the "least squares" solution $\mathbf{d}_2$ inside the half ball $B_{\mathcal{H}}(\lambda, N/2) \cap V$, see (3.105) in Proposition 3.26. However, the above local approximation can not be implemented distributedly in the DSRS, as only agents on the graph $\mathcal{G}$ have computing and telecommunication ability. So we propose to compute

$$\mathbf{w}_{\lambda,N} := \chi_{\lambda,G}^N\mathbf{S}\chi_{\lambda,V}^N(\chi_{\lambda,V}^N\mathbf{S}^T\mathbf{S}\chi_{\lambda,V}^N)^{-1}\mathbf{d}_{\lambda,N} = \chi_{\lambda,G}^N\mathbf{S}\chi_{\lambda,V}^N(\chi_{\lambda,V}^N\mathbf{S}^T\mathbf{S}\chi_{\lambda,V}^N)^{-2}\chi_{\lambda,V}^N\mathbf{S}^T\mathbf{z} \tag{3.103}$$

instead, which approximates

$$\mathbf{w}_{LS} := \mathbf{S}(\mathbf{S}^T\mathbf{S})^{-1}\mathbf{d}_2 \tag{3.104}$$

inside $B_{\mathcal{G}}(\lambda, N/2) \cap G$, see (3.106) in the proposition below.

Proposition 3.26. *Let $\mathcal{G}$ and $\mathcal{H}$ satisfy Assumptions 1 – 4, $\mathcal{V}$ be as in* (3.17)*, and let the sensing matrix $\mathbf{S} \in \mathcal{J}_\alpha(\mathcal{G}, \mathcal{V}), \alpha > d$, have ℓ^2-stability with lower stability bound $A\|\mathbf{S}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}$. Take an*

integer N satisfying (3.83)*, and set*

$$\theta = \frac{2\alpha - 2d}{2\alpha - d} \in (0,1) \text{ and } r_0 = 1 - \frac{A^2(\alpha-d)^2}{2^{\alpha+1}D_1 D_1(\mathcal{G})\alpha^2}.$$

Then

$$\|\chi_{\lambda,V}^{N/2}(\mathbf{d}_{\lambda,N} - \mathbf{d}_2)\|_\infty \le D_3(N+1)^{-\alpha+d}\|\mathbf{d}_2\|_\infty \tag{3.105}$$

and

$$\|\chi_{\lambda,G}^{N/2}(\mathbf{w}_{\lambda,N} - \mathbf{w}_{LS})\|_\infty \le D_4(N+1)^{-\alpha+d}\|\mathbf{d}_2\|_\infty, \tag{3.106}$$

where $D_3 = \frac{2^{2\alpha-d+1}\alpha D_1 D_2}{\alpha-d}$*,* $D_4 = \big(\frac{2^{3\alpha-d+3}\alpha L^2 D_1(\mathcal{G}) D_2^2}{\alpha-d} + LD_2\big)\|\mathbf{S}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}^{-1}$*, and*

$$D_2 = \sum_{n=0}^{\infty}\Big(\frac{2^{2\alpha+d/2+4}D_1^3\alpha^2}{r_0^{1-\theta}(\alpha-d)^2}\Big)^{\frac{2-\theta}{(1-\theta)^2}n^{\log_2^{(2-\theta)}}} r_0^n. \tag{3.107}$$

To prove Proposition 3.26, we need the following critical estimate.

Proposition 3.27. *Let $\mathcal{G}$, $\mathcal{H}$, $\mathcal{V}$ and* $\mathbf{S}$ *be as in Proposition 3.26. Then*

$$\|(\chi_{\lambda,V}^N \mathbf{S}^T\mathbf{S}\chi_{\lambda,V}^N)^{-1}\|_{\mathcal{J}_\alpha(\mathcal{V})} \le \frac{2^{-\alpha-1}(\alpha-d)^2 D_2}{\alpha^2 D_1 D_1(\mathcal{G})\|\mathbf{S}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}^2}, \tag{3.108}$$

where D_2 is the constant in (3.107).

Proof. Let $\mathbf{J}_{\lambda,N} := \chi_{\lambda,V}^N\mathbf{S}^T\mathbf{S}\chi_{\lambda,V}^N$. By Lemma 3.18, we have

$$\|\mathbf{J}_{\lambda,N}\|_{\mathcal{J}_\alpha(\mathcal{V})} \le \frac{2^{\alpha+1}D_1(\mathcal{G})\alpha}{\alpha-d}\|\mathbf{S}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}^2. \tag{3.109}$$

This together with Propositions 3.20 implies that

$$A^2\|\mathbf{S}\|^2_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}\|\chi^N_{\lambda,V}\mathbf{x}\|_2^2 \le \|\mathbf{S}\chi^N_{\lambda,V}\mathbf{x}\|_2^2 = \langle \mathbf{J}_{\lambda,N}\mathbf{x},\mathbf{x}\rangle \le \frac{2^{\alpha+1}\alpha^2 D_1 D_1(\mathcal{G})}{(\alpha-d)^2}\|\mathbf{S}\|^2_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}\|\chi^N_{\lambda,V}\mathbf{x}\|_2^2$$

for all $\mathbf{x}\in\ell^2$. Hence

$$\mathbf{J}_{\lambda,N} = \frac{2^{\alpha+1}\alpha^2 D_1 D_1(\mathcal{G})}{(\alpha-d)^2}\|\mathbf{S}\|^2_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}(\mathbf{I}_{B_{\mathcal{H}}(\lambda,N)\cap V}-\mathbf{B}_{\lambda,N}) \tag{3.110}$$

for some $\mathbf{B}_{\lambda,N}$ satisfying

$$\|\mathbf{B}_{\lambda,N}\|_{\mathcal{B}^2}\le r_0 \tag{3.111}$$

and

$$\|\mathbf{B}_{\lambda,N}\|_{\mathcal{J}_\alpha(\mathcal{V})} \le \|\mathbf{I}_{B_{\mathcal{H}}(\lambda,N)\cap V}\|_{\mathcal{J}_\alpha(\mathcal{V})} + \frac{2^{-\alpha-1}(\alpha-d)^2\|\mathbf{J}_{\lambda,N}\|_{\mathcal{J}_\alpha(\mathcal{V})}}{\alpha^2 D_1 D_1(\mathcal{G})\|\mathbf{S}\|^2_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}} \le 1+\frac{\alpha-d}{\alpha D_1}\le 2, \tag{3.112}$$

where $\mathbf{I}_{B_{\mathcal{H}}(\lambda,N)\cap V}$ is the identity matrix on $B_{\mathcal{H}}(\lambda,N)\cap V$. Then following the argument in [140] and applying (3.76) with $\mathbf{C}$ replaced by $\mathbf{B}_{\lambda,N}$ and V by $B_{\mathcal{H}}(\lambda,N)\cap V$, we obtain the following estimate

$$\|(\mathbf{B}_{\lambda,N})^n\|_{\mathcal{J}_\alpha(\mathcal{V})} \le \Big(\frac{D^{\frac{1}{1-\theta}}\|\mathbf{B}_{\lambda,N}\|_{\mathcal{J}_\alpha(\mathcal{V})}}{\|\mathbf{B}_{\lambda,N}\|_{\mathcal{B}^2}}\Big)^{\frac{2-\theta}{1-\theta}n^{\log_2(2-\theta)}}\|\mathbf{B}_{\lambda,N}\|^n_{\mathcal{B}^2} \ \text{ for all } n\ge 1,$$

where $D=2^{2\alpha+d/2+3}D_1^{1/2}(D_1\alpha/(\alpha-d))^{2-\theta}$. This together with (3.111) and (3.112) leads to

$$\|(\mathbf{B}_{k,N})^n\|_{\mathcal{J}_\alpha(\mathcal{V})} \le (2D^{\frac{1}{1-\theta}}/r_0)^{\frac{2-\theta}{1-\theta}n^{\log_2(2-\theta)}} r_0^n \ \text{ for all } n\ge 1. \tag{3.113}$$

Observe that

$$\|(\mathbf{J}_{\lambda,N})^{-1}\|_{\mathcal{J}_\alpha(\mathcal{V})} \le \frac{2^{-\alpha-1}(\alpha-d)^2}{\alpha^2 D_1 D_1(\mathcal{G})\|\mathbf{S}\|^2_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}}\Big(1+\sum_{n=1}^{\infty}\|(\mathbf{B}_{\lambda,N})^n\|_{\mathcal{J}_\alpha(\mathcal{V})}\Big) \tag{3.114}$$

by (3.110). Combining (3.113) and (3.114) completes the proof. □

Proof of Proposition 3.26. Observe from (3.101) and (3.102) that

$$\chi_{\lambda,V}^{N/2}(\mathbf{d}_{\lambda,N}-\mathbf{d}_2)=\chi_{\lambda,V}^{N/2}(\chi_{\lambda,V}^{N}\mathbf{S}^T\mathbf{S}\chi_{\lambda,V}^{N})^{-1}\chi_{\lambda,V}^{N}\mathbf{S}^T\mathbf{S}(\mathbf{I}-\chi_{\lambda,V}^{N})\mathbf{d}_2.$$

This together with (3.75), Lemma 3.18, and Propositions 3.20 and 3.27 implies that

$$\begin{aligned}
\|\chi_{\lambda,V}^{N/2}(\mathbf{d}_{\lambda,N}-\mathbf{d}_2)\|_\infty &\le \|(\chi_{\lambda,V}^{N}\mathbf{S}^T\mathbf{S}\chi_{\lambda,V}^{N})^{-1}\chi_{\lambda,V}^{N}\mathbf{S}^T\mathbf{S}\|_{\mathcal{J}_\alpha(\mathcal{V})}\times\\
&\quad\Big(\sup_{i\in B_{\mathcal{H}}(\lambda,N/2)\cap V}\sum_{j\notin B_{\mathcal{H}}(\lambda,N)\cap V}(1+\rho_{\mathcal{H}}(i,j))^{-\alpha}\Big)\|\mathbf{d}_2\|_\infty\\
&\le \frac{2^{\alpha+1}D_1\alpha}{\alpha-d}\|(\chi_{\lambda,V}^{N}\mathbf{S}^T\mathbf{S}\chi_{\lambda,V}^{N})^{-1}\|_{\mathcal{J}_\alpha(\mathcal{V})}\|\mathbf{S}^T\mathbf{S}\|_{\mathcal{J}_\alpha(\mathcal{V})}\times\\
&\quad\Big(\sup_{i\in V}\sum_{\rho_{\mathcal{H}}(i,j)>N/2}(1+\rho_{\mathcal{H}}(i,j))^{-\alpha}\Big)\|\mathbf{d}_2\|_\infty\\
&\le 2^{\alpha+1}D_2\Big(\sup_{i\in V}\sum_{\rho_{\mathcal{H}}(i,j)>N/2}(1+\rho_{\mathcal{H}}(i,j))^{-\alpha}\Big)\|\mathbf{d}_2\|_\infty\\
&\le \frac{2^{\alpha+1}D_1D_2\alpha}{\alpha-d}\Big(\frac{N}{2}+1\Big)^{-\alpha+d}\|\mathbf{d}_2\|_\infty\le D_3(N+1)^{-\alpha+d}\|\mathbf{d}_2\|_\infty.
\end{aligned}$$

This proves the estimate (3.105).

Now we prove (3.106). Set $\mathbf{y}_{LS}=(\mathbf{S}^T\mathbf{S})^{-1}\mathbf{d}_2$. By (3.75),

$$\|\mathbf{y}_{LS}\|_\infty\le\frac{D_1\alpha}{\alpha-d}\|(\mathbf{S}^T\mathbf{S})^{-1}\|_{\mathcal{J}_\alpha(\mathcal{V})}\|\mathbf{d}_2\|_\infty. \tag{3.115}$$

Moreover, following the proof of Proposition 3.27 gives

$$\|(\mathbf{S}^T\mathbf{S})^{-1}\|_{\mathcal{J}_\alpha(\mathcal{V})}\le\frac{2^{-\alpha-1}(\alpha-d)^2D_2}{\alpha^2D_1D_1(\mathcal{G})\|\mathbf{S}\|^2_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}}. \tag{3.116}$$

Write

$$\begin{aligned}
\chi_{\lambda,G}^{N/2}(\mathbf{w}_{\lambda,N}-\mathbf{w}_{LS}) &= \chi_{\lambda,G}^{N/2}(\chi_{\lambda,G}^{N}\mathbf{S}\chi_{\lambda,V}^{N})(\chi_{\lambda,V}^{N}\mathbf{S}^T\mathbf{S}\chi_{\lambda,V}^{N})^{-2}\chi_{\lambda,V}^{N}\mathbf{S}^T\mathbf{S}(I-\chi_{\lambda,V}^{N})\mathbf{d}_2 \\
&\quad +\chi_{\lambda,G}^{N/2}(\chi_{\lambda,G}^{N}\mathbf{S}\chi_{\lambda,V}^{N})(\chi_{\lambda,V}^{N}\mathbf{S}^T\mathbf{S}\chi_{\lambda,V}^{N})^{-1}\chi_{\lambda,V}^{N}\mathbf{S}^T\mathbf{S}(I-\chi_{\lambda,V}^{N})\mathbf{y}_{LS} \\
&\quad -\chi_{\lambda,G}^{N/2}\mathbf{S}(I-\chi_{\lambda,V}^{N})\mathbf{y}_{LS} \\
&=: I+II+III. \qquad (3.117)
\end{aligned}$$

Using (3.64), (3.115), (3.116), Lemma 3.18, and Propositions 3.20 and 3.27, we obtain

$$\begin{aligned}
\|I\|_\infty &\le \|(\chi_{\lambda,G}^{N}\mathbf{S}\chi_{\lambda,V}^{N})(\chi_{\lambda,V}^{N}\mathbf{S}^T\mathbf{S}\chi_{\lambda,V}^{N})^{-2}\chi_{\lambda,V}^{N}\mathbf{S}^T\mathbf{S}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})} \times \\
&\quad \Big(\sup_{\lambda'\in B_{\mathcal{H}}(\lambda,N/2)\cap G}\sum_{i\notin B_{\mathcal{H}}(\lambda,N)\cap V}(1+\rho_{\mathcal{H}}(\lambda',i))^{-\alpha}\Big)\|\mathbf{d}_2\|_\infty \\
&\le \frac{2^{2\alpha+2}LD_2^2}{\|\mathbf{S}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}}\Big(\sup_{\lambda'\in G}\sum_{\rho_{\mathcal{H}}(\lambda',i)>N/2}(1+\rho_{\mathcal{H}}(\lambda',i))^{-\alpha}\Big)\|\mathbf{d}_2\|_\infty \\
&\le \frac{2^{3\alpha-d+2}\alpha L^2 D_1(\mathcal{G})D_2^2}{(\alpha-d)\|\mathbf{S}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}}(N+1)^{-\alpha+d}\|\mathbf{d}_2\|_\infty,
\end{aligned}$$

$$\begin{aligned}
\|II\|_\infty &\le \frac{2^{3\alpha-d+2}\alpha^2 L^2(D_1(\mathcal{G}))^2 D_2}{(\alpha-d)^2}\|\mathbf{S}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}(N+1)^{-\alpha+d}\|\mathbf{y}_{LS}\|_\infty \\
&\le \frac{2^{2\alpha-d+1}\alpha L^2 D_1(\mathcal{G})D_2^2}{(\alpha-d)\|\mathbf{S}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}}(N+1)^{-\alpha+d}\|\mathbf{d}_2\|_\infty,
\end{aligned}$$

and

$$\|III\|_\infty \le \frac{LD_2}{\|\mathbf{S}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}}(N+1)^{-\alpha+d}\|\mathbf{d}_2\|_\infty.$$

These together with (3.117) prove (3.106). □

Take a maximal $\frac{N}{4}$-disjoint subset $G_{N/4}\subset G$ satisfying (3.8) and (3.9). We patch $\mathbf{w}_{\lambda,N}$, $\lambda\in G_{N/4}$,

in (3.103) together to generate a linear approximation

$$\mathbf{w}_N^* = \sum_{\lambda \in G_{N/4}} \mathbf{\Theta}_{\lambda,N} \chi_{\lambda,G}^{N/2} \mathbf{w}_{\lambda,N} \tag{3.118}$$

of the bounded vector $\mathbf{w}_{LS}$, where $\mathbf{\Theta}_{\lambda,N}$ is a diagonal matrix with diagonal entries

$$\theta_{\lambda,N}(\lambda'') = \frac{\chi_{B_{\mathcal{G}}(\lambda,N/2)}(\lambda'')}{\sum_{\lambda' \in G_{N/4}} \chi_{B_{\mathcal{G}}(\lambda',N/2)}(\lambda'')},\ \lambda'' \in G.$$

The above approximation is well-defined as $\{B_{\mathcal{G}}(\lambda', N/2),\ \lambda' \in G_{N/4}\}$ is a finite covering of G by (3.21) and Proposition 3.4. Moreover, we obtain from Proposition 3.26 that

$$\begin{aligned}
\|\mathbf{w}_N^* - \mathbf{w}_{LS}\|_\infty &= \Big\| \sum_{\lambda \in G_{N/4}} \mathbf{\Theta}_{\lambda,N} \chi_{\lambda,G}^{N/2} (\mathbf{w}_{\lambda,N} - \mathbf{w}_{LS}) \Big\|_\infty \\
&\le \sup_{\lambda'' \in G} \sum_{\lambda \in G_{N/4}} \theta_{\lambda,N}(\lambda'') \|\chi_{\lambda,G}^{N/2} (\mathbf{w}_{\lambda,N} - \mathbf{w}_{LS})\|_\infty \\
&\le D_4 (N+1)^{-\alpha+d} \|\mathbf{d}_2\|_\infty.
\end{aligned} \tag{3.119}$$

Therefore, the moving consensus $\mathbf{w}_N^*$ of $\mathbf{w}_{\lambda,N}, \lambda \in G_{N/4}$, provides a good approximation to $\mathbf{w}_{LS}$ in (3.104) for large N. In addition, $\mathbf{w}_N^*$ depends on the observation $\mathbf{z}$ linearly,

$$\mathbf{w}_N^* = \mathbf{R}_N \mathbf{S}^T \mathbf{z} \tag{3.120}$$

for some matrix $\mathbf{R}_N$ with bandwidth $2N$ and

$$\|\mathbf{R}_N\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})} \le D_5 := \frac{(\alpha-d)^2 L D_2^2}{\alpha^2 D_1 D_1(\mathcal{G}) \|\mathbf{S}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}^3}. \tag{3.121}$$

Given noisy samples $\mathbf{z}$, we may use $\mathbf{w}_N^*$ in (3.120) as the first approximation of $\mathbf{w}_{LS}$,

$$\mathbf{w}_1 = \mathbf{R}_N \mathbf{S}^T \mathbf{z} \tag{3.122}$$

and recursively define

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \mathbf{w}_1 - \mathbf{R}_N \mathbf{S}^T \mathbf{S} \mathbf{S}^T \mathbf{w}_n, \ n \geq 1. \tag{3.123}$$

In the next theorem, we show that the above sequence $\mathbf{w}_n, n \geq 1$, converges exponentially to some bounded vector $\mathbf{w}$, not necessarily $\mathbf{w}_{LS}$, satisfying the consistent condition

$$\mathbf{S}^T \mathbf{w} = \mathbf{S}^T \mathbf{w}_{LS} = \mathbf{d}_2. \tag{3.124}$$

Theorem 3.28. *Let $\mathcal{G}, \mathcal{H}$ and $\mathcal{V}$ be as in Proposition 3.26, and let $\mathbf{w}_n, n \geq 1$, be as in* (3.122) *and* (3.123)*. Suppose that N satisfies* (3.83) *and*

$$r_1 := \frac{D_1(\mathcal{G}) D_4 L \alpha}{\alpha - d} \|\mathbf{S}\|_{\mathcal{J}_\alpha(\mathcal{G}, \mathcal{V})} (N+1)^{-\alpha+d} < 1. \tag{3.125}$$

Set

$$D_6 = \frac{2^{2\alpha+2} \alpha L^3 (D_1(\mathcal{G}))^2 D_2^2}{(\alpha - d)(1 - r_1) D_1 \|\mathbf{S}\|_{\mathcal{J}_\alpha(\mathcal{G}, \mathcal{V})}}.$$

Then $\mathbf{w}_n$ and $\mathbf{S}^T \mathbf{w}_n, n \geq 1$, converge exponentially to a bounded vector $\mathbf{w}$ in (3.124) *and the "least squares" solution $\mathbf{d}_2$ in* (3.100) *respectively,*

$$\|\mathbf{w}_n - \mathbf{w}\|_\infty \leq D_6 r_1^n \|\mathbf{d}_2\|_\infty \tag{3.126}$$

and

$$\|\mathbf{S}^T \mathbf{w}_n - \mathbf{d}_2\|_\infty \leq \frac{D_1(\mathcal{G}) D_6 L \alpha}{\alpha - d} \|\mathbf{S}\|_{\mathcal{J}_\alpha(\mathcal{G}, \mathcal{V})} r_1^n \|\mathbf{d}_2\|_\infty, \ n \geq 1. \tag{3.127}$$

Proof. Let

$$\mathbf{u}_n = \mathbf{S}^T(\mathbf{w}_n - \mathbf{w}_{LS}) = \mathbf{S}^T\mathbf{w}_n - \mathbf{d}_2 \text{ and } \mathbf{v}_n = \mathbf{S}\mathbf{u}_n,\ n \geq 1. \tag{3.128}$$

Then,

$$\mathbf{u}_{n+1} = \mathbf{u}_n - \mathbf{S}^T\mathbf{R}_N\mathbf{S}^T\mathbf{S}\mathbf{u}_n = \mathbf{S}^T\big(\mathbf{S}(\mathbf{S}^T\mathbf{S})^{-2}\mathbf{S}^T\mathbf{v}_n - \mathbf{R}_N\mathbf{S}^T\mathbf{v}_n\big)$$

by (3.122), (3.123) and (3.128). Therefore,

$$\begin{aligned}
\|\mathbf{u}_{n+1}\|_\infty &\leq \frac{D_1(\mathcal{G})L\alpha}{\alpha - d}\|\mathbf{S}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}\|\mathbf{R}_N\mathbf{S}^T\mathbf{v}_n - \mathbf{S}(\mathbf{S}^T\mathbf{S})^{-2}\mathbf{S}^T\mathbf{v}_n\|_\infty \\
&\leq \frac{D_1(\mathcal{G})D_4L\alpha}{\alpha - d}\|\mathbf{S}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}(N+1)^{-\alpha+d}\|(\mathbf{S}^T\mathbf{S})^{-1}\mathbf{S}^T\mathbf{v}_n\|_\infty \\
&= r_1\|\mathbf{u}_n\|_\infty \leq \cdots \leq r_1^n\|\mathbf{S}^T(\mathbf{R}_N\mathbf{S}^T\mathbf{S} - \mathbf{S}(\mathbf{S}^T\mathbf{S})^{-1})\mathbf{d}_2\|_\infty \\
&\leq r_1^{n+1}\|\mathbf{d}_2\|_\infty,
\end{aligned} \tag{3.129}$$

where the second inequality follows from (3.119) with $\mathbf{d}_2$ replaced by $(\mathbf{S}^T\mathbf{S})^{-1}\mathbf{S}^T\mathbf{v}_n$, and the last inequality holds by (3.119) and Proposition 3.13.

Observe that

$$\mathbf{w}_{n+1} - \mathbf{w}_n = -\mathbf{R}_N\mathbf{S}^T\mathbf{S}\mathbf{u}_n. \tag{3.130}$$

Using (3.121), Proposition 3.13 and Lemma 3.18 gives

$$\|\mathbf{w}_{n+1} - \mathbf{w}_n\|_\infty \leq \frac{2^{2\alpha+2}\alpha L^3(D_1(\mathcal{G}))^2D_2^2}{(\alpha - d)D_1\|\mathbf{S}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}}\|\mathbf{u}_n\|_\infty. \tag{3.131}$$

This together with (3.129) proves the exponential convergence (3.126).

The conclusion (3.124) follows from (3.128) by taking limit $n \to \infty$.

The error estimate (3.127) between the "least squareS" solution $\mathbf{d}_2$ and its sub-optimal approximation $\mathbf{S}^T\mathbf{w}_n, n \geq 1$, follows from (3.126) and Proposition 3.13. □

By the above theorem, each agent should have minimal storage, computing, and telecommunication capabilities. Furthermore, the algorithm (3.122) and (3.123) will have faster convergence (hence less delay for signal reconstruction) by selecting large N when agents have larger storage, more computing power, and higher telecommunication capabilities. In addition, no iteration is needed for sufficiently large N, and the reconstructed signal is approximately to the one obtained by the finite-section method, cf. [47] and simulations in Section 3.7.

The iterative algorithm (3.122) and (3.123) can be recast as follows:

$$\mathbf{w}_1 = \mathbf{R}_N \mathbf{S}^T \mathbf{z} \ \text{ and } \ \mathbf{e}_1 = \mathbf{w}_1 - \mathbf{R}_N \mathbf{S}^T \mathbf{S} \mathbf{S}^T \mathbf{w}_1, \tag{3.132}$$

and

$$\begin{cases} \mathbf{w}_{n+1} = \mathbf{w}_n + \mathbf{e}_n \\ \mathbf{e}_{n+1} = \mathbf{e}_n - \mathbf{R}_N \mathbf{S}^T \mathbf{S} \mathbf{S}^T \mathbf{e}_n, \ n \geq 1. \end{cases} \tag{3.133}$$

Next, we present a distributed implementation of the algorithm (3.132) and (3.133) when $\mathbf{S}$ has bandwidth s. Select a threshold ϵ and an integer $N \geq s$ satisfying (3.125). Write

$$\begin{cases} \mathbf{S}^T = (\mathbf{a}(i,\lambda))_{i \in V, \lambda \in G} \\ \mathbf{R}_N \mathbf{S}^T = (\mathbf{b}_N(\lambda,\lambda'))_{\lambda,\lambda' \in G} \\ \mathbf{R}_N \mathbf{S}^T \mathbf{S} \mathbf{S}^T = (\mathbf{c}_N(\lambda,\lambda'))_{\lambda,\lambda' \in G} \\ \mathbf{z} = (\mathbf{z}(\lambda))_{\lambda \in G}, \end{cases}$$

and

$$\mathbf{w}_n = (\mathbf{w}_n(\lambda))_{\lambda \in G} \ \text{ and } \ \mathbf{e}_n = (\mathbf{e}_n(\lambda))_{\lambda \in G}, \ n \geq 1.$$

We assume that any agent $\lambda \in G$ stores vectors $\mathbf{a}(i,\lambda'), \mathbf{b}_N(\lambda,\lambda'), \mathbf{c}_N(\lambda,\lambda')$ and $\mathbf{z}(\lambda')$, where $(i,\lambda) \in T$ and $\lambda' \in B_{\mathcal{G}}(\lambda, 2N+3s)$. The following is the distributed implementation of the algorithm (3.132) and (3.133) for an agent $\lambda \in G$.

Distributed algorithm (3.132) **and** (3.133) **for signal reconstruction:**

1. Input $\mathbf{a}(i,\lambda'), \mathbf{b}_N(\lambda,\lambda'), \mathbf{c}_N(\lambda,\lambda')$ and $\mathbf{z}(\lambda')$, where $(i,\lambda)\in T$ and $\lambda'\in B_{\mathcal{G}}(\lambda, 2N+3s)$.

2. Input stop criterion $\epsilon>0$ and maximal number of iteration steps K.

3. Compute $\mathbf{w}(\lambda)=\sum_{\lambda'\in B_{\mathcal{G}}(\lambda,2N+s)}\mathbf{b}_N(\lambda,\lambda')\mathbf{z}(\lambda')$.

4. Communicate with neighboring agents in $B_{\mathcal{G}}(\lambda,2N+3s)$ to obtain data $\mathbf{w}(\lambda')$, $\lambda'\in B_{\mathcal{G}}(\lambda,2N+3s)$.

5. Evaluate the sampling error term $\mathbf{e}(\lambda)=\mathbf{w}(\lambda)-\sum_{\lambda'\in B_{\mathcal{G}}(\lambda,2N+3s)}\mathbf{c}_N(\lambda,\lambda')\mathbf{w}(\lambda')$.

6. Communicate with neighboring agents in $B_{\mathcal{G}}(\lambda,2N+3s)$ to obtain error data $\mathbf{e}(\lambda')$, $\lambda'\in B_{\mathcal{G}}(\lambda,2N+3s)$.

7. **for** $n=2$ to K **do**

   7a. Compute $\delta=\max_{\lambda'\in B_{\mathcal{G}}(\lambda,2N+3s)}|\mathbf{e}(\lambda')|$.

   7b. **Stop** if $\delta\le\epsilon$, else **do**

   7c. Update $\mathbf{w}(\lambda)=\mathbf{w}(\lambda)+\mathbf{e}(\lambda)$.

   7d. Update $\mathbf{e}(\lambda)=\mathbf{e}(\lambda)-\sum_{\lambda'\in B_{\mathcal{G}}(\lambda,2N+3s)}\mathbf{c}_N(\lambda,\lambda')\mathbf{e}(\lambda')$.

   7e. Communicate with neighboring agents located in $B_{\mathcal{G}}(\lambda,2N+3s)$ to obtain error data $\mathbf{e}(\lambda')$, $\lambda'\in B_{\mathcal{G}}(\lambda,2N+3s)$.

   end for

We conclude this section by discussing the complexity of the distributed algorithm (3.132) and (3.133), which depends essentially on N. In its implementation, the data storage requirement for each agent is about $(L+3)(2N+3s+1)^d$. In each iteration, the computational cost for each agent

is about $O(N^d)$ mainly used for updating the error $\mathbf{e}$. The communication cost for each agent is about $O(N^{d+\beta})$ if the communication between distant agents $\lambda, \lambda' \in G$, processed through their shortest path, has its cost being proportional to $(\rho_{\mathcal{G}}(\lambda, \lambda'))^\beta$ for some $\beta \geq 1$. By Theorem 3.28, the number of iteration steps needed to reach the accuracy ϵ is about $O(\ln(1/\epsilon)/\ln N)$. Therefore the total computational and communication cost for each agent are about $O(\ln(1/\epsilon)N^d/\ln N)$ and $O(\ln(1/\epsilon)N^{d+\beta}/\ln N)$, respectively.

3.7 Numerical simulations

In this section, we present two simulations to demonstrate the distributed algorithm (3.132) and (3.133) for stable signal reconstruction.

Agents in the first simulation are almost uniformly deployed on the circle of radius $R/5$, and their locations are at

$$\lambda_l := \frac{R}{5}\Big(\cos\frac{2\pi\theta_l}{R}, \sin\frac{2\pi\theta_l}{R}\Big),\ 1 \leq l \leq R,$$

where $R \geq 1$ and $\theta_l \in l + [-1/4, 1/4]$ are randomly selected. Every agent in the SDS has a direct communication channel to its two adjacent agents. Then the graph $\mathcal{G}_c = (G_c, S_c)$ to describe the SDS is a cycle graph, where $G_c = \{\lambda_1, \ldots, \lambda_R\}$ and $S_c = \big\{(\lambda_1, \lambda_2), \ldots, (\lambda_{R-1}, \lambda_R), (\lambda_R, \lambda_1), (\lambda_1, \lambda_R), (\lambda_R, \lambda_{R-1}), \ldots, (\lambda_2, \lambda_1)\big\}$. Take innovative positions

$$\mathbf{p}_i := r_i\Big(\cos\frac{2\pi i}{R},\ \sin\frac{2\pi i}{R}\Big),\ 1 \leq i \leq R,$$

deployed almost uniformly near the circle of radius $R/5$, where $r_i \in R/5 + [-1/4, 1/4]$ are randomly selected. Given any innovative position $\mathbf{p}_i, 1 \leq i \leq R$, it has three anchor agents λ_i, λ_{i-1} and λ_{i+1}, where $\lambda_0 = \lambda_R$ and $\lambda_{R+1} = \lambda_1$. Set $V_c = \{\mathbf{p}_i, 1 \leq i \leq R\}$ and $T_c = \{(\mathbf{p}_i, \lambda_{i-j}), i = 1, \ldots, R \text{ and } j = 0, \pm 1\}$. Then $\mathcal{H}_c = (G_c \cap V_c, S_c \cup T_c \cup T_c^*)$ is the graph to describe the DSRS,

see Figure 3.3.

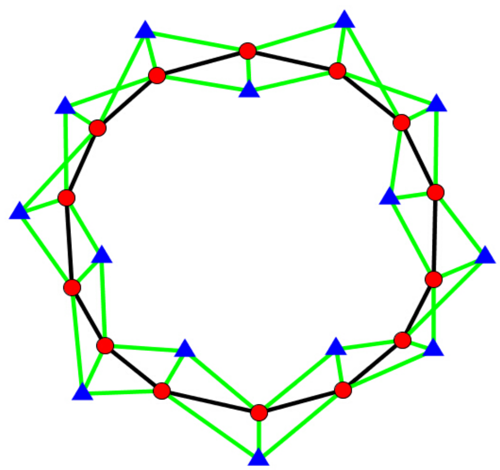

Figure 3.3: The DSRS in the first simulation in Section 3.7

Presented in Figure 3.3 is the graph $\mathcal{H}_c = (G_c \cap V_c, S_c \cup T_c \cup T_c^*)$ to describe the DSRS in the first simulation, where vertices in G_c, edges in S_c, vertices in V_c and edges in $T_c \cup T_c^*$ are plotted in red circles, black lines, blue triangles and green lines, respectively.

Let $\varphi(\mathbf{t}) := \exp(-(t_1^2 + t_2^2)/2)$ for $\mathbf{t} = (t_1, t_2)$. Gaussian signals

$$f(\mathbf{t}) = \sum_{i=1}^{R} c(i)\varphi(\mathbf{t} - \mathbf{p}_i) \tag{3.134}$$

to be sampled and reconstructed have their amplitudes $c(i) \in [0, 1]$ being randomly chosen, see the left image of Figure 3.4.

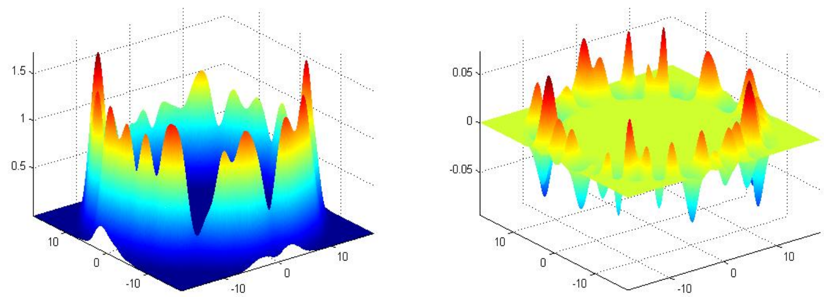

Figure 3.4: Original signal and difference between it and its reconstruction

Plotted in Figure 3.4 on the left is the signal f in (3.134) with $R = 80$. On the right is the difference between the signal f and the reconstructed signal $f_{n,N,\delta}$ with $n = 10, N = 6$ and $\delta = 0.05$.

In the first simulation, we consider ideal sampling procedure. Thus for the agent $\lambda_l, 1 \le l \le R$, the noisy sampling data acquired is

$$y_\delta(l) = \sum_{i=1}^{R} c(i)\varphi(\lambda_l - \mathbf{p}_i) + \eta(l), \tag{3.135}$$

where $\eta(l) \in [-\delta, \delta]$ are randomly generated with bounded noise level $\delta > 0$.

Our first simulation shows that the distributed algorithm (3.132) and (3.133) converges for $N \ge 5$ and the convergence rate is almost independent of the network size R, cf. the upper bound estimate in (3.127).

Let $f_{n,N,\delta}(\mathbf{t}) := \sum_{i=1}^{R} c_{n,N,\delta}(i)\varphi(\mathbf{t} - \mathbf{p}_i)$ be the reconstructed signal in the n-th iteration by applying the distributed algorithm (3.132) and (3.133) from the noisy sampling data in (3.135), see the

right image of Figure 3.4. Define maximal reconstruction errors

$$\epsilon(n,N,\delta) := \begin{cases} \max_{1\le i\le R} |c(i)| & \text{if } n=0, \\ \max_{1\le i\le R} |c_{n,N,\delta}(i)-c(i)| & \text{if } n\ge 1. \end{cases}$$

Presented in Table 3.1 is the average of reconstruction errors $\epsilon(n,N,\delta)$ with 500 trials in noiseless environment ($\delta=0$), where the network size R is 80. It indicates that the proposed distributed algorithm (3.132) and (3.133) has faster convergence rate for larger $N\ge 5$, and we only need three iteration steps to have a near perfect reconstruction from its noiseless samples when $N=10$.

Table 3.1: Maximal reconstruction errors $\epsilon(n,N,\delta)$ with $\delta=0$

n \ N	5	6	7	8	9	10
0	0.9874	0.9881	0.9878	0.9876	0.9877	0.9884
1	0.9875	0.4463	0.3073	0.1940	0.1055	0.0523
2	0.6626	0.2046	0.0794	0.0271	0.0124	0.0024
3	0.3624	0.0926	0.0240	0.0045	0.0014	0.0001
4	0.2535	0.0443	0.0068	0.0006	0.0001	0.0000
5	0.1742	0.0206	0.0018	0.0001	0.0000	0.0000
6	0.1169	0.0093	0.0005	0.0000	0.0000	0.0000
7	0.0840	0.0042	0.0001	0.0000	0.0000	0.0000
8	0.0579	0.0017	0.0000	0.0000	0.0000	0.0000
9	0.0411	0.0007	0.0000	0.0000	0.0000	0.0000
10	0.0289	0.0003	0.0000	0.0000	0.0000	0.0000

The robustness of the proposed algorithm (3.132) and (3.133) against sampling noises is tested and

confirmed, see Figure 3.4. Moreover, it is observed that the maximal reconstruction error $\epsilon(n, N, \delta)$ with large n depends almost linearly on the noise level δ, cf. the sub-optimal approximation property in Theorem 3.21.

In the next simulation, agents are uniformly deployed on two concentric circles and each agent has direct communication channels to its three adjacent agents. Then the graph $\mathcal{G}_p = (G_p, S_p)$ to describe our SDS is a prism graph with vertices having physical locations,

$$\mu_l := \begin{cases} \frac{R}{10}\big(\cos\frac{4\pi\theta_l}{R}, \sin\frac{4\pi\theta_l}{R}\big) & \text{if } 1 \le l \le \frac{R}{2} \\ \big(\frac{R}{10}+1\big)\big(\cos\frac{4\pi\theta_l}{R}, \sin\frac{4\pi\theta_l}{R}\big) & \text{if } \frac{R}{2}+1 \le l \le R, \end{cases}$$

where $R \ge 2$ and $\theta_l \in l + [-1/4, 1/4], 1 \le l \le R$, are randomly selected. The innovative positions

$$\mathbf{q}_i := r_i\Big(\cos\frac{4\pi i}{R},\ \sin\frac{4\pi i}{R}\Big),\ 1 \le i \le \frac{R}{2},$$

have four anchor agents μ_i, μ_{i+1}, $\mu_{i+R/2}$ and $\mu_{i+R/2+1}$, where $\mu_0 = \mu_{R/2}$, $\mu_{R+1} = \mu_{R/2+1}$, and $r_i \in \frac{R}{10} + [\frac{1}{4}, \frac{3}{4}]$ are randomly selected. Set $V_p = \{\mathbf{q}_i, 1 \le i \le \frac{R}{2}\}$ and $T_p = \{(\mathbf{q}_i, \mu_{i+j}), i = 1, \ldots, \frac{R}{2}$ and $j = 0, 1, \frac{R}{2}, \frac{R}{2}+1\}$. Thus the graph $\mathcal{H}_p = (G_p \cap V_p, S_p \cup T_p \cup T_p^*)$ to describe our DSRS is a connected simple graph, see the left image of Figure 3.5.

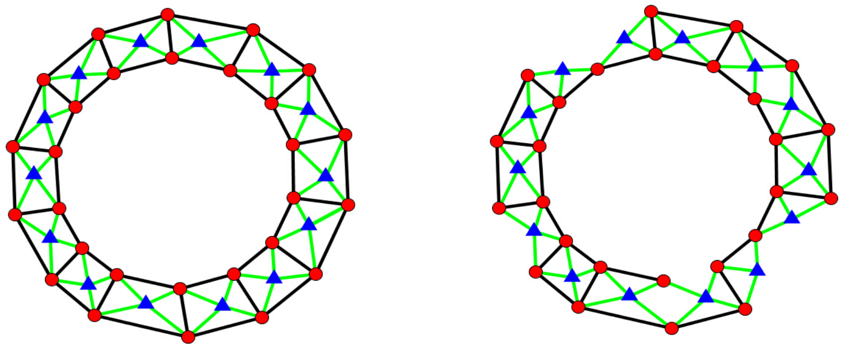

Figure 3.5: Complete and incomplete DSRS

Presented in Figure 3.5 on the left is the graph $\mathcal{H}_p = (G_p \cap V_p, S_p \cup T_p \cup T_p^*)$ to describe the DSRS, where vertices in G_p and V_p are in red circles and blue triangles, and edges in S_p and $T_p \cup T_p^*$ are in black solid lines and green solid lines, respectively. On the right is a subgraph of $\mathcal{H}_p$, where some agents are completely dysfunctional and some have communication channels to one or two of their nearby agents clogged.

Following the first simulation, we consider the ideal sampling procedure of signals,

$$g(\mathbf{t}) = \sum_{i=1}^{R/2} d(i)\varphi(\mathbf{t} - \mathbf{q}_i), \tag{3.136}$$

where $d(i) \in [0, 1], 1 \leq i \leq R/2$, are randomly selected, see the left image of Figure 3.6.

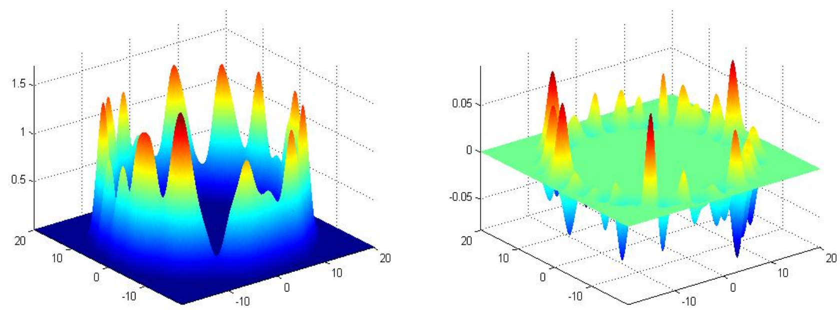

Figure 3.6: Reconstruction with incomplete DSRS (with dysfunctional agents).

Plotted in Figure 3.6 on the left is the signal g in (3.136) with $R = 160$. On the right is the difference between the signal g and its approximation $g_{n,N,\delta}$, where $n = 4, N = 6, \delta = 0.05$, and agents located at μ_1, μ_{87} are completely dysfunctional, while agents located at $\mu_{11}, \mu_{51}, \mu_{91}$ have their partial communication channels clogged.

Then the noisy sampling data acquired by the agent $\mu_l, 1 \leq l \leq R$, is

$$y_\delta(l) = \sum_{i=1}^{R/2} d(i)\varphi(\mu_l - \mathbf{q}_i) + \eta(l), \tag{3.137}$$

where $\eta(l) \in [-\delta, \delta]$ are randomly selected with bounded noise level $\delta > 0$. Applying the distributed algorithm (3.132) and (3.133), we obtain approximations

$$g_{n,N,\delta}(\mathbf{t}) = \sum_{i=1}^{R/2} d_{n,N,\delta}(i)\varphi(\mathbf{t} - \mathbf{q}_i),\ n \geq 1, \tag{3.138}$$

of the signal g in (3.136). Our simulations illustrate that the distributed algorithm (3.132) and (3.133) converges for $N \geq 3$ and the signal g can be reconstructed near perfectly from its noiseless

samples in 12 steps for $N = 3$, 7 steps for $N = 4$, 5 steps for $N = 5$, 4 steps for $N = 6$, and 3 steps for $N = 7$, cf. Table 3.1 in the first simulation.

The robustness of the proposed distributed algorithm (3.132) and (3.133) against sampling noises and dysfunctions of agents in the DSRS is tested and confirmed, see the right graph of Figure 3.5 and the right image of Figure 3.6.

CHAPTER 4: PHASE RETRIEVAL IN SHIFT-INVARIANT SPACES

Phase retrieval arises in various fields of science and engineering. In this chapter, we consider an infinite-dimensional phase retrieval problem (also known as phaseless sampling and reconstruction problem) for real-valued signals in a principal shift-invariant space

$$V(\phi) := \Big\{ \sum_{k\in\mathbb{Z}} c(k)\phi(t-k) : \ c(k) \in \mathbb{R} \Big\}, \tag{4.1}$$

where the generator ϕ is a real-valued continuous function with compact support. Our model of the generator ϕ is the B-spline B_N of order $N \geq 1$ ([157, 160]), which is obtained by convolving the indicator function $\chi_{[0,1)}$ on the unit interval N times,

$$B_N = \underbrace{\chi_{[0,1)} * \cdots * \chi_{[0,1)}}_{N}. \tag{4.2}$$

Let

$$N = \min_{N_2, N_1 \in \mathbb{Z}} \{N_2 - N_1, \phi \text{ vanishes outside } [N_1, N_2]\} \tag{4.3}$$

be the support length of the generator ϕ, which is the same as the order N for the B-spline generator B_N.

We consider an infinite-dimensional phase retrieval problem to reconstruct real-valued signals living in a shift-invariant space from their phaseless samples taken either on the whole line or on a discrete set with finite sampling rate. We find an equivalence between nonseparability of signals in a shift-invariant space and their phase retrievability with phaseless samples taken on the whole line. For spline signals of order N, we show that they can be well approximated, up to a sign, from their noisy phaseless samples taken on a set with sampling rate $2N-1$. We also propose a robust algorithm to reconstruct nonseparable signals in a shift-invariant space from their phaseless

samples corrupted by bounded noises.

4.1 Phase retrievability and nonseparability

Let ϕ be a real-valued generator of the shift-invariant space $V(\phi)$, and N be its support length given in (4.3). Without loss of generality, we assume that

$$\phi(t) = 0 \text{ for all } t \notin [0, N], \tag{4.4}$$

otherwise replacing ϕ by $\phi(\cdot - N_0)$ for some $N_0 \in \mathbb{Z}$. Clearly, not all signals in $V(\phi)$ are determined, up to a sign, from their magnitude measurements on $\mathbb{R}$. For instance, signals $\phi(t) \pm \phi(t-N)$ have same magnitude measurements $|\phi(t)| + |\phi(t-N)|$ on the real line, but they are not the same even up to a sign. Then a natural question is whether a signal in $V(\phi)$ is determined, up to a sign, from its magnitude measurements.

Theorem 4.1. *Let ϕ be a real-valued continuous function with compact support and $V(\phi)$ in* (4.1) *be the shift-invariant space generated by ϕ. Then a signal $f \in V(\phi)$ is determined, up to a sign, by its magnitude measurements $|f(t)|, t \in \mathbb{R}$, if and only if there does not exist nonzero signals f_1 and f_2 in $V(\phi)$ such that*

$$f = f_1 + f_2 \ \text{ and } \ f_1 f_2 = 0. \tag{4.5}$$

We call signals that satisfy (4.5) to be nonseparable. A separable signal $f \in V(\phi)$ can be written as the sum of two nonzero signals $f_1, f_2 \in V(\phi)$ with their supports being essentially disjoint. Then it is not determined, up to a sign, from its magnitude measurements as

$$|f_1(t) - f_2(t)| = |f_1(t) + f_2(t)| = |f_1(t)| + |f_2(t)|, \ t \in \mathbb{R}.$$

Instead of proving the Theorem 4.1, we provide the proof of its generalization.

Theorem 4.2. *Let V be a linear space of real-valued continuous signals on $\mathbb{R}$. Then a signal $f \in V$ is determined, up to a sign, by its magnitude measurements $|f(t)|, t \in \mathbb{R}$, if and only if it is nonseparable.*

Proof. ($\Longrightarrow$) Suppose, on the contrary, that there exist nonzero signals $f_1, f_2 \in V$ such that $f = f_1 + f_2$ and $f_1 f_2 = 0$. Set $g = f_1 - f_2 \in V$. Then $g \neq \pm f$ and $|g| = |f| = |f_1| + |f_2|$. This is a contradiction.

($\Longleftarrow$) Assume that f is nonseparable and $g \in V$ satisfies $|g| = |f|$. Set $g_1 := (f + g)/2$ and $g_2 := (f - g)/2 \in V$. Then $f = g_1 + g_2$ and $g_1 g_2 = 0$. This together with the assumption on f implies that either $g_1 = 0$ or $g_2 = 0$. Hence $g = \pm f$ and the sufficiency is proved. $\square$

We remark that the Paley-Wiener space for bandlimited signals to live in is a shift-invariant space generated by the $sinc$ function $\frac{\sin \pi t}{\pi t}$ with infinite support, and the phase retrieval problem in the Paley-Wiener space was discussed in [126, 154]. Observe that any bandlimited signal does not have a decomposition of the form (4.5), as it is analytic on the real line. Therefore by Theorem 4.1, we have the following corollary, cf. [154, Theorem 1].

Corollary 4.3. *Any real-valued bandlimited signal is determined, up to a sign, by its magnitude measurements on the real line.*

The next question to be considered in this section is to find the set of all nonseparable signals in a shift-invariant space $V(\phi)$. Let us start from the simplest case that $N = 1$ (i.e., the generator ϕ is supported on $[0, 1]$). In this case, one can verify that a signal $f \in V(\phi)$ is nonseparable if and only if there exists an integer k_0 such that

$$f(t) = c(k_0)\phi(t - k_0) \ \text{ for some } \ c(k_0) \in \mathbb{R}. \tag{4.6}$$

For the case that the generator ϕ has its support length

$$N \geq 2, \tag{4.7}$$

we have the following characterization to nonseparable signals in the shift-invariant space $V(\phi)$.

Before characterizing the nonseparability (and hence phase retrievability by Theorem 4.1) of signals in a shift-invariant space, let us consider nonseparability of piecewise linear signals.

Example 4.4. Due to the interpolation property of the B-spline B_2 of order 2, piecewise linear signals $f \in V(B_2)$ have the following expansion,

$$f(t) = \sum_{k\in\mathbb{Z}} f(k+1)B_2(t-k).$$

Therefore $f \in V(B_2)$ is separable if and only if there exist integers $k_0 < k_1 < k_2$ such that $f(k_0)f(k_2) \neq 0$ and $f(k_1) = 0$. Thus the separable signal

$$f = \sum_{k\leq k_1-2} f(k+1)B_2(t-k) + \sum_{k\geq k_1} f(k+1)B_2(t-k) =: f_1 + f_2,$$

is the sum of two nonzero signals $f_1, f_2 \in V(B_2)$ supported in $(-\infty, k_1]$ and $[k_1, \infty)$ respectively.

Phase retrieval of signals in a shift-invariant space is an infinite-dimensional problem with high nonlinearity. In this section, we show in Theorem 4.5 and Corollary 4.8 that a nonseparable spline signal in $V(B_N)$ is determined, up to a sign, from its phaseless samples taken on the shift-invariant set

$$Y_1 := X + \mathbb{Z}, \tag{4.8}$$

where $N \geq 2$ and X contains $2N - 1$ distinct points in $(0, 1)$.

Theorem 4.5. *Let ϕ be a real-valued continuous function satisfying* (4.4) *and* (4.7)*, $X := \{x_m, 1 \leq m \leq 2N-1\} \subset (0,1)$, and let $f(t) = \sum_{k\in\mathbb{Z}} c(k)\phi(t-k)$ be a nonzero real-valued signal in $V(\phi)$. If all $N \times N$ submatrices of*

$$\Phi = \big(\phi(x_m + n)\big)_{1\leq m\leq 2N-1, 0\leq n\leq N-1} \tag{4.9}$$

are nonsingular, then the following statements are equivalent.

(i) The signal f is nonseparable.

(ii) $\sum_{l=0}^{N-2} |c(k+l)|^2 \neq 0$ for all $K_-(f) - N + 1 < k < K_+(f) + 1$, where $K_-(f) = \inf\{k, c(k) \neq 0\}$ and $K_+(f) = \sup\{k, c(k) \neq 0\}$.

(iii) The signal f is determined, up to a sign, from its phaseless samples $|f(t)|, t \in X + \mathbb{Z}$, taken on the shift-invariant set $X + \mathbb{Z}$.

In the above theorem, the implication (iii)$\Longrightarrow$(i) holds by Theorem 4.1, while the implication (i)$\Longrightarrow$(ii) follows essentially from the support property (4.4) and (4.7) of the generating function ϕ. The technical part of the proof is the implication (ii)$\Longrightarrow$(iii), where we apply [20, Theorem 2.8] on phase retrievability of finite-dimensional signals.

Proof. The implication iii)$\Longrightarrow$i) follows immediately from Theorem 4.1. Then it remains to prove i)$\Longrightarrow$ii) and ii)$\Longrightarrow$iii).

i)$\Longrightarrow$ii): Set $K_\pm = K_\pm(f)$. For $K_- + 1 - N < k < K_- + 1$ or $K_+ + 1 - N < k < K_+ + 1$, the conclusion $\sum_{l=0}^{N-2} |c(k+l)|^2 \neq 0$ follows from the definitions of K_- and K_+. Then it remains to

establish the statement ii) for $K_- < k < K_+ + 2 - N$. Suppose, on the contrary, that

$$\sum_{l=0}^{N-2} |c(k_1 + l)|^2 = 0 \tag{4.10}$$

for some $K_- < k_1 < K_+ - N + 2$. Set $f_1(t) := \sum_{l=K_-}^{k_1-1} c(l)\phi(t-l)$ and $f_2(t) := \sum_{l=k_1+N-1}^{K_+} c(l)\phi(t-l)$. Then

$$f = f_1 + f_2 \ \text{ and } \ f_1 f_2 = 0 \tag{4.11}$$

by (4.10) and the observation that f_1 and f_2 are supported in $(-\infty, k_1+N-1]$ and $[k_1+N-1, \infty)$ respectively. Clearly, f_1 and f_2 are nonzero signals in $V(\phi)$. This together with (4.11) implies that f is separable, which contradicts to the assumption i).

ii)$\Longrightarrow$iii): To prove this implication, we need a lemma.

Lemma 4.6. *Let ϕ and X be as in Theorem 4.5. Then for any $l \in \mathbb{Z}$ and signal $g(t) = \sum_{k\in\mathbb{Z}} d(k)\phi(t-k) \in V(\phi)$, coefficients $d(k), l-N+1 \le k \le l$, are completely determined, up to a sign, by phaseless samples $|g(x_m + l)|,\ x_m \in X$, of the signal g.*

The above lemma follows immediately from [20, Theorem 2.8] and the observation that

$$g(x_m + l) = \sum_{k=l-N+1}^{l} d(k)\phi(x_m + l - k), \ \ x_m \in X.$$

Take a particular integer $K_- - 1 < k_0 < K_+ + 1$ with $c(k_0) \neq 0$. Without loss of generality, we assume that

$$c(k_0) > 0, \tag{4.12}$$

otherwise replacing f by $-f$.

Using (4.12) and applying Lemma 4.6 with g and l replaced by f and k_0 respectively, we conclude

that $c(k_0-N+1),\cdots,c(k_0)$ are completely determined by phaseless samples $|f(X+k_0)|$ of the signal f on $X+k_0$. Now we prove the following claim:

$$c(k),\ k\le k_0,\ \text{are determined by}\ |f(X+k)|,\ k\le k_0 \tag{4.13}$$

by induction. Inductively we assume that $c(k), k_0-p-N+1\le k\le k_0$, are determined from $|f(X+k)|, k_0-p\le k\le k_0$. The inductive proof is complete if $k_0-p-N+1\le K_-$. Otherwise $k_0-p-N+1>K_-$ and

$$\sum_{l=0}^{N-2}|c(k_0-p-N+l+1)|^2\neq 0 \tag{4.14}$$

by the assumption ii). Applying Lemma 4.6 with g and k_0 replaced by f and k_0-p-1 respectively, we conclude that $c(k_0-N-p),\cdots,c(k_0-p-1)$ are determined, up to a global phase, by $|f(X+k_0-p-1)|$. This together with (4.14) and the inductive hypothesis implies that $c(k_0-N-p),\cdots,c(k_0-p-1)$ are completely determined by $|f(X+k)|, k_0-p-1\le k\le k_0$. Thus the inductive argument can proceed.

Using the similar argument, we can show that $c(k), k\ge k_0$ are determined by $|f(X+k)|, k\ge k_0$. This together with (4.13) completes the proof. $\square$

The nonsingularity of $N\times N$ submatrices of the matrix Φ in (4.9) is also known as its full sparkness ([61, 8]). The full sparkness of the matrix Φ in (4.9) implies that ϕ has linearly independent shifts, i.e., the linear map from sequences to signals in $V(\phi)$,

$$(c(k))_{k=-\infty}^{\infty}\longmapsto\sum_{k=-\infty}^{\infty}c(k)\phi(t-k),$$

is one-to-one ([98, 145]). Conversely, if ϕ has linearly independent shifts and it is a continuous

solution of the refinement equation ([54, 109]),

$$\phi(t) = \sum_{n=0}^{N} a(n)\phi(2t-n) \ \text{ and } \ \int_{\mathbb{R}} \phi(t)dt = 1, \tag{4.15}$$

where $\sum_{n=0}^{N} a(n) = 2$, then Φ in (4.9) is of full spark for almost all $(x_1, \ldots, x_{2N-1}) \in (0,1)^{2N-1}$, see [145, Theorem A.2]. This together with Theorem 4.5 implies the following result for wavelet signals, cf. [154, Theorem 1] and Corollary 4.3 for bandlimited signals.

Corollary 4.7. *Let ϕ be a continuous solution of the refinement equation* (4.15) *with linearly independent shifts. Then any nonseparable signal in $V(\phi)$ is determined, up to a sign, from its magnitude measurements on $\mathbb{R}$.*

For the refinement equation (4.15), under the assumption that

$$\sum_{n=0}^{N} a(n)z^n = (1+z)Q(z) \tag{4.16}$$

for some polynomial Q having positive coefficients and its zeros with strictly negative real part, the corresponding matrix Φ in (4.9) is of full spark whenever $x_m \in (0,1), 1 \leq m \leq 2N-1$, are distinct ([79, 80]). It is well known that the B-spline B_N of order N satisfies the refinement equation (4.15) with $Q(z)$ in (4.16) given by $2^{-N+1}(1+z)^{N-1}$. This together with Theorem 4.5 implies the following result for spline signals, cf. Corollary 4.8.

Corollary 4.8. *Let X contain $2N-1$ distinct points in $(0,1)$. Then any nonseparable spline signal in $V(B_N)$ is determined, up to a sign, from its phaseless samples taken on the shift-invariant set $X + \mathbb{Z}$.*

For a signal $f = \sum_{k\in\mathbb{Z}} c(k)\phi(t-k) \in V(\phi)$, define

$$S_f = \inf_{K_-(f)-N+1<k<K_+(f)+1} \sum_{l=0}^{N-2} |c(k+l)|^2. \tag{4.17}$$

By the second statement in Theorem 4.5, we obtain that $S_f = 0$ if f is separable, and that $S_f > 0$ if f is a nonseparable signal with finite duration. So the quantities S_f can be used to measure absolute and relative distances of a signal f to the set of all separable signals in $V(\phi)$, cf. Theorem 4.14.

Given a signal $f \in V(\phi)$, the last question to be addressed in this section is to find all signals $g \in V(\phi)$ such that g and f have the same magnitude measurements on the real line, cf. [6]. By the second statement in Theorem 4.5, any signal $f \in V(\phi)$ can be uniquely written as the sum of nonseparable signals $f_i \in V(\phi), i \in I$, with their supporting intervals $[a_i, a_i']$ being essentially mutually disjoint, i.e.,

$$f = \sum_{i\in I} f_i \tag{4.18}$$

and

$$[a_i, a_i') \cap [a_j, a_j') = \emptyset \text{ for all distinct } i, j \in I, \tag{4.19}$$

see Lemma 4.10. Clearly signals $g = \sum_{i\in I} \epsilon_i f_i$ with $\epsilon_i \in \{-1, 1\}, i \in I$, have the same magnitude measurements as the signal f has. We show that the converse is true in the following theorem.

Theorem 4.9. *Let ϕ be a real-valued continuous function satisfying* (4.4), (4.7) *and* (4.9). *Assume that $f \in V(\phi)$ has a decomposition* (4.18) *and* (4.19) *of nonseparable signals. Then $g \in V(\phi)$ satisfies $|g(t)| = |f(t)|,\ t \in \mathbb{R}$, if and only if there exists $\epsilon_i \in \{-1, 1\}, i \in I$, such that $g = \sum_{i\in I} \epsilon_i f_i$.*

The sufficiency follows as $f_i, i \in I$, have mutually disjoint supports. To prove the necessity, we need a lemma.

Lemma 4.10. *Let ϕ be as in Theorem 4.9. Then for any nonzero signal $f \in V(\phi)$, there exist nonseparable signals $f_i \in V(\phi), i \in I$, satisfying* (4.18) *and* (4.19)*. Moreover the decomposition* (4.18) *and* (4.19) *is unique.*

Proof. Write $f = \sum_{k\in\mathbb{Z}} c(k)\phi(\cdot - k)$ and set

$$\mathcal{L} := \big\{l \in \mathbb{Z} : (c(l), \dots, c(l+N-2)) \neq \mathbf{0}\big\}. \tag{4.20}$$

Then there exist $b_i, b_i' \in \mathbb{Z} \cup \{-\infty, +\infty\}, i \in I$, such that

$$\mathcal{L} := \bigcup_{i\in I} \big((b_i, b_i') \cap \mathbb{Z}\big) \tag{4.21}$$

and

$$[b_i, b_i'),\ i \in I,\ \text{are mutually disjoint.} \tag{4.22}$$

Hence

$$c(k) = 0 \quad \text{for all}\ \ k \notin \cup_{i\in I}(b_i + N - 2, b_i'). \tag{4.23}$$

Define

$$f_i = \sum_{b_i+N-2<k<b_i'} c(k)\phi(\cdot - k),\ i \in I. \tag{4.24}$$

Then the decomposition (4.18) holds by (4.23) and (4.24), and the mutually disjoint property (4.19) follows from (4.22) and the observation that $f_i, i \in I$, have supporting intervals $[b_i + N - 1, b_{i'} + N - 1]$. Observe from (4.21) that $K_+(f_i) = b_i' - 1$ and $K_-(f_i) = b_i + N - 1, i \in I$. This together with Theorem 4.5 implies that $f_i, i \in I$, are nonseparable. Therefore $f_i, i \in I$, in (4.24) are nonseparable signals satisfying (4.18) and (4.19).

Now it remains to prove uniqueness of the decomposition (4.18) and (4.19). Suppose that $g_j \in$

$V(\phi), j \in J$, are nonseparable signals with their supporting intervals $[a_j, a'_j]$ satisfying

$$f = \sum_{j\in J} g_j \tag{4.25}$$

and

$$[a_j, a'_j) \cap [a_{j'}, a'_{j'}) = \emptyset \text{ for all distinct } j, j' \in J. \tag{4.26}$$

Then it suffices to prove that $J = I$ and for any $j \in J$ there exists a unique $i \in I$ such that $g_j = f_i$, where $f_i, i \in I$, are given in (4.24). By (4.4), (4.7), (4.25) and (4.26), we have

$$g_j = \sum_{a_j - 1 < k < a'_j - N + 1} c(k)\phi(\cdot - k) \tag{4.27}$$

and

$$c(k) = 0 \text{ for all } k \notin \cup_{j\in J}(a_j - 1, a'_j - N + 1). \tag{4.28}$$

Applying (4.27), (4.28) and Theorem 4.5, we obtain

$$\mathcal{L} = \cup_{j\in J}\big((a_j - N + 1, a'_j - N + 1) \cap \mathbb{Z}\big), \tag{4.29}$$

where the set $\mathcal{L}$ is given in (4.20). This together with (4.26) leads to another decomposition of the set $\mathcal{L}$ that satisfies (4.21) and (4.23). Due to the uniqueness of such a decomposition, we have that $J = I$ and for any $j \in J$ there exists a unique $i \in I$ such that $(a_j, a'_j) = (b_i + N - 1, b'_i + N - 1)$, where $b_i, b'_i, i \in I$, are given in (4.21). This together with (4.27) completes the proof. $\square$

Now we start the proof of Theorem 4.9.

Proof of Theorem 4.9. Without loss of generality, we assume that $f \neq 0$. Write $g = \sum_{k\in\mathbb{Z}} d(k)\phi(t-$

$k)$ and $f = \sum_{k\in\mathbb{Z}} c(k)\phi(t-k)$. By Lemma 4.6, for any $l \in \mathbb{Z}$ there exists $\delta_l \in \{-1, 1\}$ such that

$$d(l+n) = \delta_l c(l+n),\ 0 \le n \le N-1. \tag{4.30}$$

Set $\mathcal{L} := \{l \in \mathbb{Z} : (c(l), \ldots, c(l+N-2)) \neq \mathbf{0}\}$ as in (4.21). Then it follows from (4.30) that

$$\delta_{l-1} = \delta_l \quad \text{for all } \ l \in \mathcal{L}. \tag{4.31}$$

As in (4.21), we write $\mathcal{L}$ as the union of open intervals $(a_i, a_i') \cap \mathbb{Z}, i \in I$, with $[a_i, a_i'), i \in I$, being mutually disjoint. Thus $\delta_l = \delta_{l'}$ for all $l, l' \in (a_i - 1, a_i') \cap \mathbb{Z}$, which implies that the existence of $\epsilon_i \in \{-1, 1\}$ with

$$d(k) = \epsilon_i c(k) \text{ for all } a_i + N - 2 < k < a_i'. \tag{4.32}$$

By (4.23) and (4.30), we have

$$d(k) = 0 \text{ for all } k \notin \cup_{i\in I}\big((a_i + N - 2, a_i') \cap \mathbb{Z}\big). \tag{4.33}$$

Therefore the conclusion $g = \sum_{i\in I} \epsilon_i f_i$ follows from (4.24), (4.32), (4.33) and Lemma 4.10. $\square$

4.2 Phaseless non-uniform sampling in a shift-invariant space

A set $\Lambda \subset \mathbb{R}$ is said to have sampling rate $D(\Lambda)$ if

$$D(\Lambda) = \lim_{b-a\to\infty} \frac{\#(\Lambda \cap [a, b])}{b-a}, \tag{4.34}$$

where $\#(E)$ is the cardinality of a set E. Then the sufficiency in Theorem 4.1 can be recast as any nonseparable signal in $V(\phi)$ can be reconstructed, up to a sign, from its phaseless samples on

$\mathbb{R}$, which has infinite sampling rate $D(\mathbb{R}) = +\infty$. For the case that the generator ϕ satisfies (4.4), (4.7) and (4.9), it follows from Theorem 4.5 that any nonseparable signal in $V(\phi)$ can be fully recovered, up to a sign, from its phaseless samples taken on the shift-invariant subset $X + \mathbb{Z}$ of $\mathbb{R}$. We observe that the set $X + \mathbb{Z}$ in Theorem 4.5 has finite sampling rate $2N - 1$, which is larger than the sampling rate 2 required for recovering bandlimited signals [154, Theorem 1]. A natural question is whether we can find the minimal sampling rate.

Theorem 4.11. *Let ϕ be a real-valued continuous function satisfying* (4.4)*,* (4.7) *and* (4.9)*, and let Λ be a discrete set with sampling rate $D(\Lambda)$. If all nonseparable signals in $V(\phi)$ can be determined, up to a sign, from their phaseless samples taken on the set Λ, then the sampling rate $D(\Lambda)$ is at least one,*

$$D(\Lambda) \geq 1. \tag{4.35}$$

Now, we continue our proof of Theorem 4.11.

Proof. By (5.13), it suffices to prove that

$$\#(\Lambda \cap [a, b]) \geq b - a - N + 1$$

for all integers a and b with $b - a \geq N$. Suppose, on the contrary, that

$$\#(\Lambda \cap [a_0, b_0]) < b_0 - a_0 - N + 1 \tag{4.36}$$

for some integers a_0 and b_0. Let

$$\mathcal{N} = \Big\{ f(t) := \sum_{k=a_0}^{b_0 - N} c(k)\phi(t-k), \ f(y) = 0 \text{ for all } y \in \Lambda \Big\}.$$

Then $\mathcal{N}$ contains some nonzero signals in $V(\phi)$, because any signals in $\mathcal{N}$ are supported in $[a_0, b_0]$,

and the homogenous linear system

$$\sum_{k=a_0}^{b_0-N} c(k)\phi(y-k) = 0, \quad y \in \Lambda \cap [a_0, b_0]$$

of size $(\#(\Lambda \cap [a_0, b_0])) \times (b_0 - a_0 - N + 1)$ has a nontrivial solution by (4.36).

Take a nonzero signal $f \in \mathcal{N}$ with minimal support length. By the assumption on the set Λ, it must be separable as it is a nonzero signal having zero magnitude measurements on Λ. Therefore by Theorem 4.5 there exist nonzero signals f_1 and $f_2 \in V(\phi)$ and an integer $k_0 \in (a_0, b_0)$ such that $f = f_1 + f_2$, f_1 vanishes outside $[k_0, b_0]$ and f_2 vanishes outside $[a_0, k_0]$. This implies that both f_1 and f_2 are nonzero signals in $\mathcal{N}$, which contradicts to the minimal support assumption on f. $\square$

The lower bound estimate (4.35) is smaller than the sampling rate required for recovering bandlimited signals [154, Theorem 1]. We believe that the lower bound estimate (4.35) for minimal sampling rate can be improved. However as indicated in the example below, it is optimal if the requirement (4.9) on the generator ϕ is dropped.

Example 4.12. Let φ_0 be a continuous function supported in $[0, 1/2]$ and set $\varphi_N(t) = \varphi_0(t) - \varphi_0(t - N + 1/2)$, $N \geq 1$. Similar to (4.6), one may verify that a signal f in $V(\varphi_N)$ is nonseparable if and only if there exists $k_0 \in \mathbb{Z}$ such that $f(t) = c(k_0)\varphi_N(t - k_0)$ for some $c(k_0) \in \mathbb{R}$. Given any $t_0 \in (0, 1/2)$ with $\varphi_0(t_0) \neq 0$, the set $t_0 + \mathbb{Z}$ has one, the lower bound in (4.35), as its sampling rate. Moreover, one may verify that all nonseparable signals in $V(\varphi_N)$ can be reconstructed, up to a sign, from their phaseless samples taken on $t_0 + \mathbb{Z}$.

The set Y_1 in (4.8) has sampling rate $2N - 1$, which is larger than the sampling rate 2 needed for the phase retrievability of signals in the Paley-Wiener space [154, Theorem 1]. Let

$$N = \min_{N_2, N_1 \in \mathbb{Z}} \{N_2 - N_1, \phi \text{ vanishes outside } [N_1, N_2]\} \tag{4.37}$$

be the support length of the generator ϕ, which is the same as the order N for the B-spline generator B_N. Recall from Theorem 4.5 that any nonseparable signal in $V(\phi)$ can be fully recovered, up to a sign, from its phaseless samples taken on a discrete set with sampling rate $2N-1$. A question is whether we can find a discrete set Λ with sampling rate less than $2N-1$ such that all nonseparable signals in $V(\phi)$ can be recovered from their phaseless samples on Λ. Under proper assumptions on the generator ϕ, we also interest in learning paring down the sampling density. For the nonseparable signal in a shift-invariant space $V(\phi)$, we already have some promising results that it can be recovered, up to a sign, from its phaseless samples taken on a nonuniform set

$$Y_\infty := X \cup (\Gamma + \mathbb{Z}_+) \cup (\Gamma^* + \mathbb{Z}_-) \tag{4.38}$$

with sampling rate N, where $\mathbb{Z}_\pm$ is the set of all positive/negative integers, and the sets $\Gamma = \{\gamma_1, \ldots, \gamma_N\}$ and $\Gamma^* = \{\gamma_1^*, \ldots, \gamma_N^*\}$ are contained in $X = \{x_1, \ldots, x_{2N-1}\} \subset (0,1)$.

I am not going to include the result in this dissertation. Instead, I will briefly discuss an example of phaseless oversampling.

Example 4.13. (Continuation of Example 4.4) Let $k_0 \in \mathbb{Z}$ and $f \in V(B_2)$ be a nonseparable piecewise linear signal. One may verify that 3 distinct points $k_0+x_1, k_0+x_2, k_0+x_3 \in k_0+(0,1)$ are enough to determine $f(k_0)$ and $f(k_0+1)$ (hence $f(t), t \in k_0+[0,1]$), up to a phase, from phaseless samples $|f(k_0+x_1)|, |f(k_0+x_2)|$ and $|f(k_0+x_3)|$. Particularly, solving

$$|f(k_0)(1-x_i) + x_i f(k_0+1)|^2 = |f(k_0+x_i)|^2,\ i = 1,2,3$$

gives

$$|f(k_0)|^2 = \frac{\begin{vmatrix} |f(k_0+x_1)|^2 & x_1(1-x_1) & x_1^2 \\ |f(k_0+x_2)|^2 & x_2(1-x_2) & x_2^2 \\ |f(k_0+x_3)|^2 & x_3(1-x_3) & x_3^2 \end{vmatrix}}{(x_2-x_1)(x_3-x_1)(x_3-x_2)},$$

$$2f(k_0)f(k_0+1) = \frac{\begin{vmatrix} (1-x_1)^2 & |f(k_0+x_1)|^2 & x_1^2 \\ (1-x_2)^2 & |f(k_0+x_2)|^2 & x_2^2 \\ (1-x_3)^2 & |f(k_0+x_3)|^2 & x_3^2 \end{vmatrix}}{(x_2-x_1)(x_3-x_1)(x_3-x_2)},$$

and

$$|f(k_0+1)|^2 = \frac{\begin{vmatrix} (1-x_1)^2 & x_1(1-x_1) & |f(k_0+x_1)|^2 \\ (1-x_2)^2 & x_2(1-x_2) & |f(k_0+x_2)|^2 \\ (1-x_3)^2 & x_3(1-x_3) & |f(k_0+x_3)|^2 \end{vmatrix}}{(x_2-x_1)(x_3-x_1)(x_3-x_2)}.$$

For the case that at least one of two evaluations $f(k_0)$ and $f(k_0+1)$ is nonzero,

$$f(k_0+2) = \begin{cases} 0 & \text{if } f(k_0+1) = 0 \\ f(k_0+1) + \triangle_{k_0}^{+} & \text{if } f(k_0+1) \neq 0, \end{cases} \tag{4.39}$$

where the first equality follows from nonseparability of the signal f, the second one is obtained by solving the equations

$$|f(k_0+1)(1-x_i) + x_i f(k_0+2)|^2 = |f(k_0+1+x_i)|^2, i = 1, 2, \tag{4.40}$$

and

$$\triangle_{k_0}^{+} = \frac{x_1^2(|f(k_0+1+x_2)|^2 - |f(k_0+1)|^2)}{2x_1x_2(x_1-x_2)f(k_0+1)} - \frac{x_2^2|f(k_0+1+x_1)|^2 - |f(k_0+1)|^2)}{2x_1x_2(x_1-x_2)f(k_0+1)}.$$

From (4.39) we see that two distinct points $k_0+1+x_1, k_0+1+x_2 \in k_0+1+(0,1)$ could sufficiently determine $f(t), k_0+1 \le t \le k_0+2$.

For the case that $f(k_0+1) = f(k_0) = 0$, solving (4.40) yields

$$|f(k_0+2)|^2 = \frac{|f(x_1+k_0+1)|^2 + |f(x_2+k_0+1)|^2}{x_1^2+x_2^2}.$$

Then either $f(t) = 0$ for all $t \in [k_0, k_0+2]$ or the phase of the signal f on $[k_0, k_0+2]$ is determined up to the sign of nonzero evaluation $f(k_0+2)$. Therefore, we can continue the above procedure to determine the signal f on $[k_0, \infty)$ if there are two distinct points in intervals $k+(0,1)$ for every $k \ge k_0+1 \in \mathbb{Z}\backslash\{k_0\}$.

Using the similar argument, we can prove by induction on $k < k_0$ that the signal $f(t), t \in [k, \infty)$, can be determined, up to a sign, by its phaseless samples taken on $l-1+x_1$ and $l-1+x_2$, $k \le l < k_0$. By now, we conclude that a nonseparable signal in $V(B_2)$ could be determined, up to a sign, by its phaseless samples on $(\{x_1, x_2\} + \mathbb{Z}) \cup \{x_3 + k_0\}$, where $x_1, x_2, x_3 \in (0,1)$ are distinct and $k_0 \in \mathbb{Z}$. We remark that the additional point $x_3 + k_0$ in the above phase retrievability is necessary in general. For instance, signals $f(t) \equiv 1/3$ and $g(t) = \sum_{k\in\mathbb{Z}}(-1)^k B_2(t-k)$ in $V(B_2)$ have the same magnitude measurements on $\{1/3, 2/3\} + \mathbb{Z}$, but $f \ne \pm g$.

4.3 Stable reconstruction from phaseless samples in a shift-invariant space

Stability of phase retrieval is of central importance, as phaseless samples in lots of engineering applications are often corrupted. Stability of phase retrieval is of paramount importance. The reader may refer to [21, 22, 23, 65] for phase retrieval in the finite-dimensional setting and [149] for nonlinear frames.

In this section, we consider the scenario that the available data

$$z_{\boldsymbol{\epsilon}}(\gamma) = |f(\gamma)|^2 + \boldsymbol{\epsilon}(\gamma),\ \gamma \in X + \mathbb{Z}, \tag{4.41}$$

are phaseless samples of a signal

$$f = \sum_{k\in\mathbb{Z}} c(k)\phi(t-k) \in V(\phi) \tag{4.42}$$

taken on the set $X + \mathbb{Z}$ corrupted by additive noises $\boldsymbol{\epsilon} = (\boldsymbol{\epsilon}(\gamma))_{\gamma\in X+\mathbb{Z}}$, where $\boldsymbol{\epsilon}$ has the bounded noise level

$$\|\boldsymbol{\epsilon}\|_\infty = \sup\{|\boldsymbol{\epsilon}(\gamma)| :\ \gamma \in X + \mathbb{Z}\}.$$

Based on the constructive proof of Theorem 4.5, we propose an algorithm to find an approximation

$$f_{\boldsymbol{\epsilon}}(t) = \sum_{k\in\mathbb{Z}} c_{\boldsymbol{\epsilon}}(k)\phi(t-k) \in V(\phi), \tag{4.43}$$

when the noisy phaseless samples in (4.41) are available.

The proposed MAPS algorithm consists of the following three parts: (i) solving the **m**inimization problem (4.45) to obtain local approximations $\mathbf{c}_{\boldsymbol{\epsilon},k'}, k' \in \mathbb{Z}$, of $\delta_{k'}\mathbf{c}$ on $k' + [-L+1, 0]$, up to a phase $\delta_{k'} \in \{-1, 1\}$, cf. [72, 76, 119, 127]; (ii) **a**djusting **p**hases to obtain local approximations

$\delta_{\epsilon,k'}\mathbf{c}_{\epsilon,k'}$ to either $\mathbf{c}$ or $-\mathbf{c}$ on $k'+[-L+1,0]$; and (iii) **s**ewing $\delta_{\epsilon,k'}\mathbf{c}_{\epsilon,k'}, k'\in\mathbb{Z}$, together to get an approximation $\mathbf{c}_\epsilon$ to either $\mathbf{c}$ or $-\mathbf{c}$. The above *MAPS algorithm* can be implemented as follows:

Algorithm 1 MAPS Algorithm

Inputs: the shift-invariant sampling set X; support length of the generator L; noisy phaseless sampling data $\big(z_\epsilon(y)\big)_{y\in X+\mathbb{Z}}$.

Instructions:

1) Local minimization: For any $k'\in\mathbb{Z}$, let

$$\mathbf{c}_{\epsilon,k'}=(c_{\epsilon,k'}(k))_{k\in\mathbb{Z}}, \tag{4.44}$$

have zero components except that $c_{\epsilon,k'}(k), k'-L+1\le k\le k'$, are solutions of the minimization problem

$$\min\sum_{m=1}^{2L-1}\left|\Big|\sum_{k=k'-L+1}^{k'}c(k)\phi(x_{m,k'}-k)\Big|-\sqrt{z_\epsilon(x_{m,k'})}\right|^2, \tag{4.45}$$

where $x_m\in X$ and $x_{m,k'}=x_m+k', 1\le m\le 2L-1$.

2) Adjust Phase: For $k'\in\mathbb{Z}$, multiply $c_{\epsilon,k'}$ by $\delta_{\epsilon,k'}\in\{-1,1\}$ so that

$$\langle\delta_{\epsilon,k'}\mathbf{c}_{\epsilon,k'},\delta_{\epsilon,k'+1}\mathbf{c}_{\epsilon,k'+1}\rangle\ge 0 \text{ for all } k'\in\mathbb{Z}. \tag{4.46}$$

3) Sewing:

$$\mathbf{c}_\epsilon(k)=\frac{1}{L}\sum_{k'=k}^{k+L-1}\delta_{\epsilon,k'}\mathbf{c}_{\epsilon,k'}(k),\ k\in\mathbb{Z}, \tag{4.47}$$

to obtain an approximation of amplitude vector $c(k), k\in\mathbb{Z}$.

Outputs: Amplitude vector $(c_\epsilon(k))_{k\in\mathbb{Z}}$, and the reconstructed signal $f_\epsilon=\sum_{k\in\mathbb{Z}}c_\epsilon(k)\phi(\cdot-k)$.

From the above implementation, we see that the MAPS algorithm can be used to reconstruct signals in $V(\phi)$ almost in real time from their phaseless samples, cf. [41, 145] and references therein on local and distributed reconstruction. Moreover, the MAPS algorithm has linear complexity $O(K_2-K_1)$ to reconstruct nonseparable signals $f=\sum_{k=K_1}^{K_2}c(k)\phi(\cdot-k)\in V(\phi)$ approximately, up to a sign, from their noisy phaseless samples on $(X+\mathbb{Z})\cap[K_1,K_2+L]$. In realistic model for sampling in shift-invariant space, the generator ϕ does not have large supporting length L. Hence the minimization problem (4.45) of size L can be solved by many algorithms available in a stable

way ([34, 35, 37, 72, 132]).

In the noiseless sampling environment (i.e., $\boldsymbol{\epsilon} = \mathbf{0}$), the proposed MAPS algorithm provides a perfect reconstruction of a nonseparable signal, up to a sign. In a noisy sampling environment, we show in the following theorem that the MAPS algorithm (4.44)–(4.47) provides, up to a sign, a stable approximation to the original nonseparable signal f.

Theorem 4.14. *Let ϕ and X be as in Theorem 4.5, $f(t) = \sum_{k=-\infty}^{\infty} c(k)\phi(t-k)$ in* (4.42) *be a nonseparable real-valued signal with S_f in* (4.17) *being positive, and let $f_{\boldsymbol{\epsilon}}(t) = \sum_{k\in\mathbb{Z}} c_{\boldsymbol{\epsilon}}(k)\phi(t-k)$ be the signal in* (5.31) *reconstructed by the MAPS algorithm* (4.44)–(4.47). *If*

$$\|\boldsymbol{\epsilon}\|_\infty \leq \frac{S_f}{48L\|(\Phi_L)^{-1}\|^2}, \tag{4.48}$$

then there exists $\delta \in \{-1, 1\}$ such that

$$|c_{\boldsymbol{\epsilon}}(k) - \delta c(k)| \leq \frac{\|(\Phi_L)^{-1}\|}{L}\sqrt{8L\|\boldsymbol{\epsilon}\|_\infty} \tag{4.49}$$

for all $k \in \mathbb{Z}$, where $\|A\| = \sup_{\|x\|_2=1}\|Ax\|_2$ for a matrix A and

$$\|(\Phi_L)^{-1}\| = \sup_{m_0<\ldots<m_{L-1}} \left\| \left(\left(\phi(x_{m_l}+n)\right)_{0\leq l,n\leq L-1} \right)^{-1} \right\|. \tag{4.50}$$

To prove Theorem 4.14, we first show that $\mathbf{c}_{\boldsymbol{\epsilon},k'}$ obtained in the first step approximates the original vector $\mathbf{c}$ on $[k'L+1-N, k'L]$, up to a phase.

Proposition 4.15. *Let $\mathbf{c}, \boldsymbol{\epsilon}$ be as in Theorem 4.14. Then for any $k' \in \mathbb{Z}$, there exists $\delta_{k'} \in \{-1, 1\}$ such that*

$$\sum_{k=k'-L+1}^{k'} |c_{\boldsymbol{\epsilon},k'}(k) - \delta_{k'}c(k)|^2 \leq 8L\|(\Phi_N)^{-1}\|^2\|\boldsymbol{\epsilon}\|_\infty. \tag{4.51}$$

Proof. Set $x_{m,k'} = x_m + k', 1 \le m \le 2L-1$. Then

$$\begin{aligned}
&\sum_{m=1}^{2L-1}\Big(\Big|\sum_{k=k'-L+1}^{k'} c_{\boldsymbol{\epsilon},k'}(k)\phi(x_{m,k'}-k)\Big| - \Big|\sum_{k=k'-L+1}^{k'} c(k)\phi(x_{m,k'}-k)\Big|\Big)^2\\
&\le 2\sum_{m=1}^{2L-1}\left(\Big|\sum_{k=k'-L+1}^{k'} c^0_{\boldsymbol{\epsilon},k'}(k)\phi(x_{m,k'}-k)\Big| - \sqrt{z_{\boldsymbol{\epsilon}}(x_{m,k'})}\right)^2 + 2\sum_{m=0}^{2L-1}\left(\sqrt{z_{\boldsymbol{\epsilon}}(x_{m,k'})} - \Big|\sum_{k=k'-L+1}^{k'} c(k)\phi(x_{m,k'}-k)\Big|\right)^2\\
&\le 4\sum_{m=1}^{2L-1}\Big||f(x_{m,k'})| - \sqrt{z_{\boldsymbol{\epsilon}}(x_{m,k'})}\Big|^2 \le 8L\|\boldsymbol{\epsilon}\|_\infty,
\end{aligned}$$

where the second inequality holds by (4.45), and the third estimate follows from the triangle inequality

$$|\sqrt{x^2+y} - |x|| \le \sqrt{|y|} \tag{4.52}$$

for all $x \ge 0$ and $y \ge -x^2$. Therefore there exists $\delta_{k'} \in \{-1, 1\}$ such that

$$\sum_{m=1}^{2L-1}\Big(\sum_{k=k'-L+1}^{k'} c_{\boldsymbol{\epsilon},k'}(k)\phi(x_{m,k'}-k) - \delta_{k'}\sum_{k=k'-L+1}^{k'} c(k)\phi(x_{m,k'}-k)\Big)^2 \le 8L\|\boldsymbol{\epsilon}\|_\infty.$$

This completes the proof. □

To prove Theorem 4.14, we adjust phases of $c_{\boldsymbol{\epsilon},k'}, k' \in \mathbb{Z}$, obtained in the first step.

Proposition 4.16. *Let $\delta_{k'} \in \{-1, 1\}, k' \in \mathbb{Z}$, be as in Proposition 4.15. If* (4.48) *holds for some $\delta_{\boldsymbol{\epsilon},k'}, k' \in \mathbb{Z}$, then*

$$\delta_{\boldsymbol{\epsilon},k'}\delta_{\boldsymbol{\epsilon},k'+1} = \delta_{k'}\delta_{k'+1} \tag{4.53}$$

for all $k' \in \mathbb{Z}$ with $\sum_{k=-L+2}^{0} |c(k+k')|^2 \ne 0$.

Proof. For any $k' \in \mathbb{Z}$,

$$\Big|\langle \delta_{k'} c_{\boldsymbol{\epsilon},k'}, \delta_{k'+1} c_{\boldsymbol{\epsilon},k'+1}\rangle - \sum_{k=k'-L+2}^{k'} |c(k)|^2\Big| \le \sum_{k=k-L+2}^{k'} |\delta_{k'} c_{\boldsymbol{\epsilon},k'}(k) - c(k)||c(k)|$$

$$+ \sum_{k=k'-L+2}^{k'} |\delta_{k'+1} c_{\boldsymbol{\epsilon},k'+1}(k) - c(k)||c(k)| + \sum_{k=k'-L+2}^{k'} |\delta_{k'} c_{\epsilon,k'}(k) - c(k)| \times |\delta_{k'+1} c_{\epsilon,k'+1}(k) - c(k))|$$

$$\le 4\sqrt{2L\|\boldsymbol{\epsilon}\|_\infty}\|(\Phi_N)^{-1}\|\Big(\sum_{k=k'-L+2}^{k'} |c(k)|^2\Big)^{1/2} + 8L\|(\Phi_N)^{-1}\|^2\|\boldsymbol{\epsilon}\|_\infty < \sum_{k=k'-L+2}^{k'} |c(k)|^2,$$

where the second estimate follows from Proposition 4.15, and the last inequality holds by the assumption (4.48) on the noise level $\|\boldsymbol{\epsilon}\|_\infty$. Therefore the vectors $\delta_{k'} c^1_{\boldsymbol{\epsilon},k'}$ and $\delta_{k'} c^1_{\boldsymbol{\epsilon},k'+1}$ have positive inner product. This together with (4.46) proves (4.53). □

Now we are ready to state the proof of Theorem 4.14.

Proof of Theorem 4.14. Set $K_\pm = K_\pm(f)$. By Theorem 4.5 and Proposition 4.16, there exists $\delta \in \{-1, 1\}$ such that

$$\delta_{\boldsymbol{\epsilon},k'} = \delta\delta_{k'} \tag{4.54}$$

for all $k' \in (K_- - 1, K_+ + L)$. For $k \in \mathbb{Z}$, we obtain from (4.46), (4.47), (4.54) and Proposition 4.15 that

$$|c_{\boldsymbol{\epsilon}}(k) - \delta c(k)| \le \frac{1}{L}\sum_{k'=k}^{k+L-1} |c_{\boldsymbol{\epsilon},k'}(k) - \delta_{k'} c(k)| + \frac{1}{L}\sum_{k'=k}^{k+L-1} |\delta_{k'}\delta_{\boldsymbol{\epsilon},k'} - \delta||c(k)| \le \frac{\|(\Phi_N)^{-1}\|}{L}\sqrt{8L\|\boldsymbol{\epsilon}\|_\infty}.$$

This completes the proof. □

Define a signal reconstruction error of the MAPS algorithm by $E(\boldsymbol{\epsilon}) = \min_{\delta\in\{-1,1\}} \|f_{\boldsymbol{\epsilon}}(t) -$

$\delta f(t)\|_\infty$. Then there exists a positive constant C by Theorem 4.14 such that

$$E(\boldsymbol{\epsilon}) \le L\|\phi\|_\infty \min_{\delta\in\{-1,1\}} \max_{k\in\mathbb{Z}} |c_{\boldsymbol{\epsilon}}(k) - \delta c(k)| \le C\sqrt{\|\boldsymbol{\epsilon}\|_\infty}. \tag{4.55}$$

This together with (4.48) implies that there is no resonance phenomenon for the phaseless sampling and reconstruction model (4.41) if the noisy level is sufficiently small. Moreover, numerical simulations in the next section show that the upper bound estimate in (4.55) for the reconstruction error $E(\boldsymbol{\epsilon})$ is suboptimal as it is about of the order $\sqrt{\|\boldsymbol{\epsilon}\|_\infty}$.

4.4 Numerical simulations

In this section, we demonstrate the performance of the MAPS algorithm (4.44) – (4.47) on reconstructing a cubic spline signal

$$f(t) = \sum_{k\in\mathbb{Z}} c(k)B_4(t-k) \tag{4.56}$$

with finite duration, where B_4 is the cubic B-spline in (4.2). Our noisy phaseless samples are taken on $X_K + \mathbb{Z}$,

$$z_{\boldsymbol{\epsilon}}(\gamma) = |f(\gamma)|^2 + \|f\|_\infty^2 \varepsilon(\gamma) \ge 0,\ \gamma \in X_K + \mathbb{Z}, \tag{4.57}$$

where $\varepsilon(\gamma) \in [-\varepsilon, \varepsilon]$ are randomly selected with noise level $\varepsilon > 0$, and

$$X_K = \left\{\frac{m}{K+1}, 1 \le m \le K\right\},\ K \ge 7. \tag{4.58}$$

The set X_K with $K = 7$ can be used as the set X in (4.9) and also in Theorem 4.5.

In our simulations,

$$c(k) \in [-1,1] \setminus [-0.1, 0.1],\ K_1 \le k \le K_2, \tag{4.59}$$

are randomly selected. Denote the signal reconstructed by the MAPS algorithm from the noisy phaseless samples (4.57) by

$$f_\varepsilon(t) = \sum_{k\in\mathbb{Z}} c_\varepsilon(k) B_4(t-k), \tag{4.60}$$

cf. Theorem 4.14. Define an amplitude reconstruction error by

$$e(\epsilon) := \min_{\delta\in\{-1,1\}} \max_{k\in\mathbb{Z}} |c_\epsilon(k) - \delta c(k)|. \tag{4.61}$$

As $B_4(t) \geq 0$ and $\sum_{k\in\mathbb{Z}} B_4(t-k) = 1$ for all $t \in \mathbb{R}$, we have

$$E(\varepsilon) := \min_{\delta\in\{-1,1\}} \max_{t\in\mathbb{R}} |f_\varepsilon(t) - \delta f(t)| \leq e(\varepsilon), \tag{4.62}$$

cf. (4.55). For the phaseless sampling and reconstruction model (4.57) with small noise level ε, it follows from Theorem 4.14 that the maximal ampltitude reconstruction error $e(\varepsilon)$ in (4.61) and maximal signal reconstruction error $E(\varepsilon)$ in (4.62) are $O(\sqrt{\varepsilon})$. It is confirmed in the numerical simulations for nonseparable cubic spline signals, see Figures 4.1.

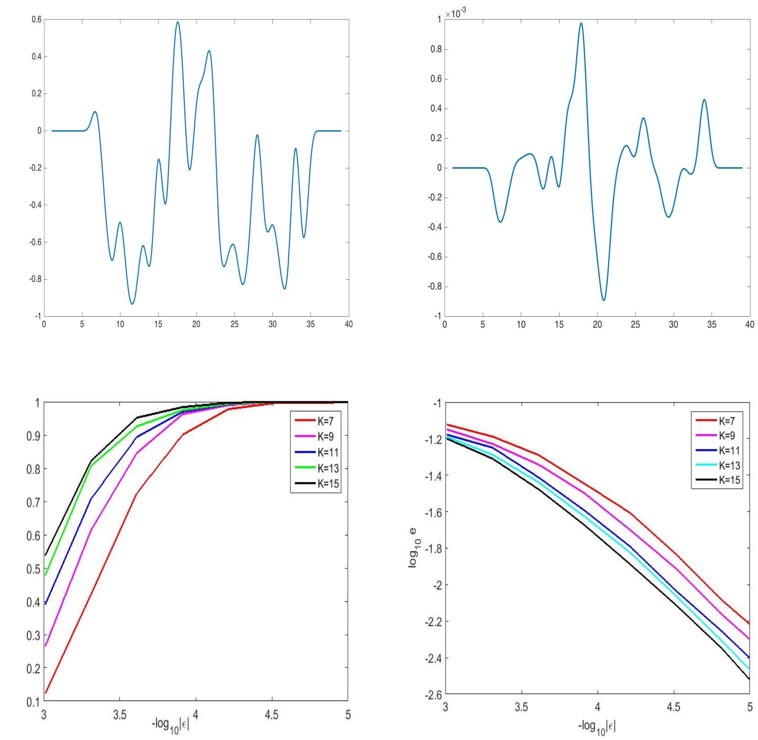

Figure 4.1: Nonseparable cubic spline signal and the reconstruction differences by MAPS

Plotted on the top left is a nonseparable cubic spline f with $K_1 = 5, K_2 = 32$ and $c(k), k \in \mathbb{Z}$, in (4.59). On the top right is the difference between the above signal f and the signal f_ϵ reconstructed by the MAPS algorithm from the noisy samples (4.57) with $\varepsilon = 10^{-5}$ and $K = 7$, where the amplitude reconstruction error $e(\varepsilon)$ is 0.0014. Plotted on the bottom left is the success rate against noisy level $-\log_{10}|\varepsilon|$ to recover a nonseparable cubic spline f by the MAPS algorithm for 1000 trails, with $c(k), k \in \mathbb{Z}$, randomly selected as in (4.59) and odd integers $7 \leq K \leq 15$. On the bottom right is the average error $\log_{10} e(\epsilon)$ against noisy level $-\log_{10}|\varepsilon|$ in the logarithmic scale for a nonseparable cubic spline f running our MAPS algorithm for 1000 trails, where the error $e(\varepsilon)$ is counted in the average only when phases are saved successfully.

The MAPS algorithm may not recover a nonseparable signal in a shift-invariant space if the noise

level ε is not sufficiently small. Presented in Figure 4.1 is the success rate in percentage and the average amplitude error after 1000 trials for different noisy levels ε, where the MAPS algorithm to recover cubic spline signals f in (4.56) with $c(k), k \in \mathbb{Z}$, in (4.59) and noisy samples in (4.57) is considered to save the phase successfully if $e(\varepsilon) < 0.1$. In the simulation, a successful recovery implies that $c_\epsilon(k)$ and $c(k), K_1 \le k \le K_2$, have same signs,

$$c_\varepsilon(k)c(k) > 0 \text{ for all } K_1 \le k \le K_2.$$

The success rate of the MAPS algorithm can be improved if we have phaseless samples on a discrete set with high sampling rate. Presented in Figure 4.1 is the success rate in percentage to recover splines f in (4.56), up to a sign, from noisy phaseless samples taken on $X_K + \mathbb{Z}, 7 \le K \le 15$, where the noise level ε, the original signal f and the success threshold are the same as before. In addition to the improvement on success rate, our simulations also indicate that the amplitude reconstruction error in (4.61) decreases when the sampling rate K increases, cf. [14, Theorem 3] for oversampling in a shift-invariant space.

The MAPS algorithm is applicable even if the original signal f is separable. Denote by g_ε the signal constructed from the MAPS algorithm. Our simulations show that the reconstruction error $\inf_{|g|=|f|} \|g_\varepsilon - g\|_\infty$ is about $O(\sqrt{\varepsilon})$, cf. (4.62), and hence the signal g_ε provides a good approximation to a signal g in Theorem 4.9, not necessarily the original signal f itself.

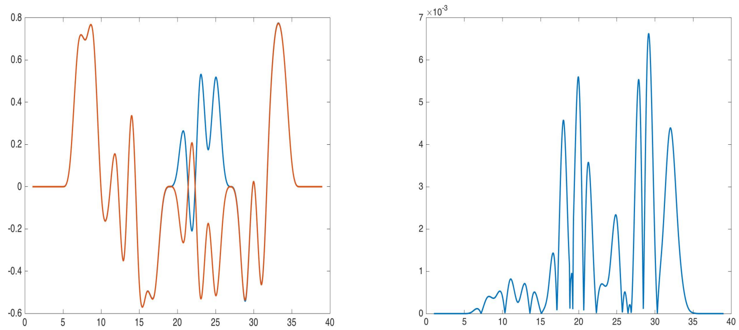

Figure 4.2: Separable cubic spline signal and the reconstruction error by MAPS

Plotted on the left is the original cubic spline f (in blue) and the constructed signal g_ε (in red) via the MAPS algorithm, where $K_1 = 5$, $K_2 = 32$, $\varepsilon = 10^{-5}$ and $c(k) \in [-1, 1], 5 \leq k \leq 32$. On the right is the difference $|g_\varepsilon - g|$ between the signal g_ε and a signal g in Theorem 4.9. The corresponding reconstruction error $\inf_{|g|=|f|} \|g_\varepsilon - g\|_\infty$ is 0.0066. Presented in Figure 4.2 is the performance of the MAPS algorithm when the amplitude coefficients of the original cubic spline f in (4.56) satisfy $c(k) \in [-1, 1]$ for all $K_1 \leq k \leq K_2$, cf. (4.59).

CHAPTER 5: PHASELESS SAMPLING AND RECONSTRUCTION OF REAL-VALUED SIGNALS IN SHIFT-INVARIANT SPACES

In Chapter 4, we consider the phase retrieval of real-valued signal f on $\mathbb{R}$ in a shift-invariant space. In this chapter, we will consider the phaseless sampling and reconstruction problem whether a real-valued signal f on $\mathbb{R}^d (d \geq 2)$, is determined, up to a sign, by its magnitude measurements $|f(x)|$ on $\mathbb{R}^d$ or a subset $X \subset \mathbb{R}^d$. The above problem is ill-posed inherently and it could be solved only if we have some extra information about the signal f.

The additional knowledge about the signals in this paper is that they live in a shift-invariant space

$$V(\phi) := \Big\{ \sum_{k \in \mathbb{Z}^d} c(k)\phi(x-k) : \ c(k) \in \mathbb{R} \text{ for all } k \in \mathbb{Z}^d \Big\} \tag{5.1}$$

generated by a real-valued continuous function ϕ with compact support. Shift-invariant spaces have been used in wavelet analysis and approximation theory, and sampling in shift-invariant spaces is a realistic model for signals with smooth spectrum, see [13, 17, 30, 55, 98] and references therein. Typical examples of shift-invariant spaces include those generated by refinable functions ([54, 109]) and box splines M_Ξ, which are defined by

$$\int_{\mathbb{R}^d} g(x) M_\Xi(x) dx = \int_{\mathbb{R}^s} g(\Xi y) dy, \ \ g \in L^2(\mathbb{R}^d), \tag{5.2}$$

where $\Xi \in \mathbb{Z}^{d \times s}$ is a matrix with full rank d ([57, 157, 160]).

The phaseless sampling and reconstruction problem of one-dimensional signals in shift-invariant spaces has been studied in [40, 125, 126, 133, 154]. Thakur proved in [154] that one-dimensional real-valued signals in a Paley-Wiener space, the shift-invariant space generated by the sinc function

$\frac{\sin \pi t}{\pi t}$, could be reconstructed from their phaseless samples taken at more than twice the Nyquist rate. Reconstruction of one-dimensional signals in a shift-invariant space was studied in [133] when frequency magnitude measurements are available. Not all signals in a shift-invariant space generated by a compactly supported function are determined, up to a sign, by their magnitude measurements on the whole line. In Chapter 4, the set of signals that can be determined by their magnitude measurements on the real line $\mathbb{R}$ is fully characterized, and a fast algorithm is proposed to reconstruct signals in a shift-invariant space from their phaseless samples taken on a discrete set with finite sampling density. Up to our knowledge, there is no literature available on the phaseless sampling and reconstruction of high-dimensional signals in a shift-invariant space, which is the core of this chapter.

The phaseless sampling and reconstruction of signals in a shift-invariant space is an infinite-dimensional phase retrieval problem, which has received considerable attention in recent years [5, 6, 7, 33, 40, 110, 125, 126, 133, 154].

5.1 Phase retrievability, nonseparability, connectivity

The phase retrievability of a real-valued signal on $\mathbb{R}^d$ is whether it is determined, up to a sign, by its magnitude measurements. It is characterized in Theorem 4.1.

The question arisen is how to determine the nonseparability of a signal in a shift-invariant space. To answer the above question, we need the one-to-one correspondence between an amplitude vector c and a signal f in the shift-invariant space $V(\phi)$,

$$c := (c(k))_{k\in\mathbb{Z}^d} \longmapsto \sum_{k\in\mathbb{Z}^d} c(k)\phi(\cdot - k) =: f \in V(\phi), \tag{5.3}$$

which is known as the *global linear independence* of the generator ϕ [25, 98, 129]. For $d = 1$,

the nonseparability of a signal in a shift-invariant space is characterized in [40] that its amplitude vector does not have consecutive zeros. However, there is no corresponding notion of consecutive zeros in the high-dimensional setting $(d \geq 2)$. To characterize the nonseparability of signals on $\mathbb{R}^d$, $d \geq 1$, we introduce an undirected graph for a signal in the shift-invariant space $V(\phi)$ generated by a real-valued continuous function ϕ with compact support.

Definition 5.1. For any $f(x) = \sum_{k\in\mathbb{Z}^d} c(k)\phi(x-k) \in V(\phi)$, define an *undirected graph*

$$\mathcal{G}_f := (V_f, E_f), \tag{5.4}$$

where the vertex set

$$V_f = \{k \in \mathbb{Z}^d : \ c(k) \neq 0\}$$

contains supports of the amplitude vector of the signal f, and

$$E_f = \big\{(k, k') \in V_f \times V_f : \ \ k \neq k' \text{ and } \phi(x-k)\phi(x-k') \neq 0 \text{ for some } x \in \mathbb{R}^d\big\}$$

is the edge set associated with the signal f.

The graph $\mathcal{G}_f$ in (5.4) is well-defined for any signal f in the shift-invariant space $V(\phi)$ when ϕ has the global linear independence. Moreover,

$$(k, k') \in E_f \text{ if and only if } k - k' \in \Lambda_\phi, \tag{5.5}$$

where Λ_ϕ contains all $k \in \mathbb{Z}^d$ such that

$$S_k := \{x \in \mathbb{R}^d : \ \phi(x)\phi(x-k) \neq 0\} \neq \emptyset. \tag{5.6}$$

In the following theorem, we show that connectivity of the graph $\mathcal{G}_f$ is a necessary condition for

the nonseparability of the signal $f \in V(\phi)$.

Theorem 5.2. *Let ϕ be a compactly supported continuous function on $\mathbb{R}^d$ with global linear independence, and $V(\phi)$ be the shift-invariant space* (5.1) *generated by ϕ. If $f \in V(\phi)$ is nonseparable, then the graph $\mathcal{G}_f$ in* (5.4) *is connected.*

Before stating sufficiency for the connectivity of the graph $\mathcal{G}_f$, we recall a concept of local linear independence on an open set.

Definition 5.3. Let ϕ be a continuous function with compact support and A be an open set. We say that ϕ has *local linear independence on A* if $\sum_{k\in\mathbb{Z}^d} c(k)\phi(x-k) = 0$ for all $x \in A$ implies that $c(k) = 0$ for all $k \in \mathbb{Z}^d$ satisfying $\phi(x-k) \not\equiv 0$ on A.

The global linear independence of a compactly supported function ϕ can be interpreted as its local linear independence on $\mathbb{R}^d$ ([25, 139]). Define

$$\Phi_A(x) := \big(\phi(x-k)\big)_{k\in K_A},\ x \in A \tag{5.7}$$

and

$$K_A := \{k \in \mathbb{Z}^d :\ \phi(\cdot - k) \not\equiv 0 \text{ on } A\}. \tag{5.8}$$

One may verify that ϕ has local linear independence on A if and only if the dimension of the linear space spanned by $\Phi_A(x), x \in A$, is the cardinality of the set K_A. The above characterization can be used to verify the local linear independence on a bounded open set, especially when ϕ has the explicit expression. For instance, one may verify that the generator ϕ_0 in Example 5.6 below has local linear independence on $(0, 1)$, but it is locally linearly dependent on $(0, 1/2)$ and $(1/2, 1)$.

Proof. Suppose, on the contrary, that $\mathcal{G}_f$ is disconnected. Let W be the set of vertices in a connected component of the graph $\mathcal{G}_f$. Then $W \neq \emptyset, V_f\backslash W \neq \emptyset$, and there are no edges between

vertices in W and $V_f \backslash W$. Write

$$\begin{aligned} f &= \sum_{k\in\mathbb{Z}^d} c(k)\phi(\cdot - k) = \sum_{k\in V_f} c(k)\phi(\cdot - k) \\ &= \sum_{k\in W} c(k)\phi(\cdot - k) + \sum_{k\in V_f\backslash W} c(k)\phi(\cdot - k) =: f_1 + f_2 \end{aligned} \tag{5.9}$$

where $c(k) \in \mathbb{R}, k \in \mathbb{Z}^d$. From the global linear independence on ϕ and nontriviality of the sets W and $V_f\backslash W$, we obtain

$$f_1 \neq 0 \ \text{ and } \ f_2 \neq 0. \tag{5.10}$$

Combining (5.9) and (5.10) with nonseparability of the signal f, we obtain that $f_1(x_0)f_2(x_0) \neq 0$ for some $x_0 \in \mathbb{R}^d$. Then by the global linear independence of ϕ, there exist $k \in W$ and $k' \in V_f\backslash W$ such that $\phi(x_0 - k) \neq 0$ and $\phi(x_0 - k') \neq 0$. Hence (k, k') is an edge between $k \in W$ and $k' \in V_f\backslash W$, which contradicts to the construction of the set W. $\square$

The local linear independence on any open set and global linear independence are equivalent to each other for some compactly supported functions, such as box splines and one-dimensional refinable functions ([52, 53, 56, 97, 145]). In the following theorem, we show that the converse in Theorem 5.2 is true if the generator ϕ is assumed to have local linear independence on any open set.

Theorem 5.4. *Let ϕ be a compactly supported continuous function on $\mathbb{R}^d$ with local linear independence on any open set, and f be a signal in the shift-invariant space $V(\phi)$. If the graph $\mathcal{G}_f$ in* (5.4) *is connected, then f is nonseparable.*

For $d = 1$, we have

$$(k, k') \in E_f \text{ if and only if } |k - k'| \leq L - 1, \tag{5.11}$$

provided that the support of ϕ is $[0, L]$ for some $L \geq 1$. This together with Theorems 5.2 and

5.4 leads to the following result, which is established in [40] under different assumptions on the generator ϕ.

Corollary 5.5. *Let ϕ be a compactly supported continuous function on $\mathbb{R}$, and $f = \sum_{k\in\mathbb{Z}} c(k)\phi(\cdot - k) \in V(\phi)$. If ϕ has local linear independence on any open set and its supporting set is $[0, L]$ where $1 \le L \in \mathbb{Z}$, then f is nonseparable if and only if $\sum_{l=0}^{L-2} |c(k+l)|^2 \neq 0$ for all $K_-(f) - L + 1 < k < K_+(f) + 1$, where $K_-(f) = \inf\{k : c(k) \neq 0\}$ and $K_+(f) = \sup\{k : c(k) \neq 0\}$.*

As demonstrated by the following example, the connectivity of the graph $\mathcal{G}_f$ is not sufficient for the signal f to be nonseparable if the local linear independence assumption on the generator ϕ is dropped.

Example 5.6. Define $\phi_0(t) = h(4t-1)+h(4t-3)+h(4t-5)-h(4t-7)$, where $h(t) = \max(1-|t|, 0)$ is the hat function supported on $[-1, 1]$. One may easily verify that ϕ_0 is a continuous function having global linear independence. Set

$$f_1(t) = \sum_{k\in\mathbb{Z}} \phi_0(t-k) \ \text{ and } \ f_2(t) = \sum_{k\in\mathbb{Z}} (-1)^k \phi_0(t-k).$$

Then f_1 and f_2 are nonzero signals in $V(\phi_0)$ supported on $[0, 1/2]+\mathbb{Z}$ and $[1/2, 1]+\mathbb{Z}$ respectively, and $f_1(t)f_2(t) = 0$ for all $t \in \mathbb{R}$. Hence $f_1 \pm 2f_2$ have the same magnitude measurements $|f_1|+2|f_2|$ on the real line but they are different, even up to a sign, i.e., $f_1 + 2f_2 \not\equiv \pm(f_1 - 2f_2)$. On the other hand, one may verify that their associated graphs $\mathcal{G}_{f_1\pm 2f_2}$ are connected.

Consider a continuous solution ϕ of a refinement equation

$$\phi(x) = \sum_{n=0}^{N} a(n)\phi(2x-n) \ \text{ and } \ \int_{\mathbb{R}} \phi(x)dx = 1 \tag{5.12}$$

with global linear independence, where $\sum_{n=0}^{N} a(n) = 2$ and $N \ge 1$ ([54, 109]). The B-spline B_N

of order N, which is obtained by convolving the indicator function $\chi_{[0,1)}$ on the unit interval N times, satisfies the above refinement equation ([157, 160]). The function ϕ in (5.12) has support $[0, N]$ and it has local linear independence on any open set if and only if it has global linear independence ([45, 99, 105, 113, 145]). Therefore we have the following result for wavelet signals by Theorems 5.2 and 5.4, which is also established in [40] with a different approach.

Corollary 5.7. *Let ϕ satisfy the refinement equation* (5.12) *and have global linear independence. Then $f \in V(\phi)$ is nonseparable if and only if the graph $\mathcal{G}_f$ in* (5.4) *is connected.*

The local linear independence requirement in Theorem 5.4 can be verified for box splines M_Ξ in (5.2). It is known that the box spline M_Ξ has local linear independence on any open set if and only if it has global linear independence if and only if all $d \times d$ submatrices of Ξ have determinants being either 0 or ± 1 ([52, 53, 56, 97]). The reader may refer to [57] for more properties and applications of box splines. As applications of Theorems 5.2 and 5.4, we have the following result for box spline signals.

Corollary 5.8. *Let $\Xi \in \mathbb{Z}^{d \times s}$ be a matrix of full rank d such that all of its $d \times d$ submatrices have determinants being either 0 or ± 1. Then $f \in V(M_\Xi)$ is nonseparable if and only if the graph $\mathcal{G}_f$ in* (5.4) *is connected.*

5.2 Phaseless sampling and reconstruction

In this section, we consider the problem whether a signal in the shift-invariant space $V(\phi)$ is determined, up to a sign, by its phaseless samples taken on a discrete set with finite sampling density. Here we define the sampling density of a discrete set $X \subset \mathbb{R}^d$ by

$$D(X) := D_+(X) = D_-(X)$$

if its upper sampling density $D_+(X)$ and lower sampling density $D_-(X)$ are the same [**?**, 41, **?**], where

$$D_+(X) := \limsup_{R\to+\infty} \sup_{x\in\mathbb{R}^d} \frac{\#(X\cap B(x,R))}{R^d} \tag{5.13}$$

and

$$D_-(X) := \liminf_{R\to+\infty} \inf_{x\in\mathbb{R}^d} \frac{\#(X\cap B(x,R))}{R^d}. \tag{5.14}$$

One may easily verify that a shift-invariant set $X = \Gamma + \mathbb{Z}^d$ generated by a finite set Γ has sampling density $\#\Gamma$.

To determine a signal, up to a sign, by its phaseless samples taken on a discrete set, a necessary condition is that the signal is nonseparable (hence phase retrievable). In the next theorem, we show that the above requirement is also sufficient.

Theorem 5.9. *Let ϕ be a compactly supported continuous function and $V(\phi)$ be the shift-invariant space in* (5.1) *generated by ϕ. Then there exists a discrete set $\Gamma \subset (0,1)^d$ such that any nonseparable signal $f \in V(\phi)$ is determined, up to a sign, by its phaseless samples on the set $\Gamma + \mathbb{Z}^d$ with finite sampling density.*

A linear space V on $\mathbb{R}^d$ is said to be *locally finite-dimensional* if it has finite-dimensional restriction on any bounded open set. The shift-invariant space in (5.1) generated by a compactly supported function ϕ is locally finite-dimensional. The reader may refer to [15] and references therein on locally finite-dimensional spaces. In this section, we will prove the following generalization of Theorem 5.9.

Theorem 5.10. *Let V be a locally finite-dimensional shift-invariant space of functions on $\mathbb{R}^d$. Then there exists a finite set $\Gamma \subset (0,1)^d$ such that any nonseparable signal $f \in V$ is determined, up to a sign, by its phaseless samples on $\Gamma + \mathbb{Z}^d$.*

Proof. Let $A = (0,1)^d$ and $V|_A$ be the space containing restrictions of all signals in V on A. By the shift-invariance, it suffices to find a set $\Gamma \subset A$ and functions $d_\gamma(x), \gamma \in \Gamma$, such that

$$|f(x)|^2 = \sum_{\gamma\in\Gamma} d_\gamma(x)|f(\gamma)|^2,\ x \in A \tag{5.15}$$

hold for all $f \in V$. By the assumption on V, $V|_A$ is finite-dimensional. Let $g_n \in V, 1 \le n \le N$, be a basis for $V|_A$, and W be the linear space spanned by symmetric matrices

$$G(x) := \big(g_n(x)g_{n'}(x)\big)_{1\le n,n'\le N},\ x \in A.$$

Then there exists a finite set $\Gamma \subset A$ with cardinality at most $N(N+1)/2$ such that $G(\gamma), \gamma \in \Gamma$, are a basis for the space W. This implies that for any $x \in A$ there exist $d_\gamma(x), \gamma \in \Gamma$, such that

$$G(x) = \sum_{\gamma\in\Gamma} d_\gamma(x)G(\gamma),\ x \in A.$$

For any $f \in V$, we write $f(x) = \sum_{n=1}^N c_n g_n(x), x \in A$. Then

$$\begin{aligned} |f(x)|^2 &= \Big|\sum_{n=1}^N c_n g_n(x)\Big|^2 = \sum_{n,n'=1}^N c_n c_{n'} g_n(x) g_{n'}(x) \\ &= \sum_{n,n'=1}^N c_n c_{n'} \Big(\sum_{\gamma\in\Gamma} d_\gamma(x) g_n(\gamma) g_{n'}(\gamma)\Big) = \sum_{\gamma\in\Gamma} d_\gamma(x)|f(\gamma)|^2,\ x \in A. \end{aligned}$$

This proves (5.15) and hence completes the proof. □

Given a compactly supported function ϕ and a bounded open set A, let

$$W_A \text{ be the linear space spanned by } \Phi_A(x)(\Phi_A(x))^T,\ x \in A, \tag{5.16}$$

where Φ_A is given in (5.7). Observe that for any bounded set A, the space W_A spanned by outer products $\Phi_A(x)(\Phi_A(x))^T, x \in A$, is of finite dimension. Therefore there exists a finite set $\Gamma \subset A$ such that outer products $\Phi_A(\gamma)(\Phi_A(\gamma))^T, \gamma \in \Gamma$, are a basis of the linear space W_A. In the proof of Theorem 5.9, we use $A = (0,1)^d$ and apply the above procedure to select the finite set Γ. With the above selection of the set Γ,

$$\#\Gamma = \dim W_{(0,1)^d}, \tag{5.17}$$

and $|f(x)|^2, x \in \mathbb{R}^d$, are determined by $|f(\gamma)|^2, \gamma \in \Gamma + \mathbb{Z}^d$.

As symmetric matrices in the space $W_{(0,1)^d}$ are of size $\#K_{(0,1)^d}$, we have the following result about the sampling density.

Corollary 5.11. *Let ϕ and $V(\phi)$ be as in Theorem 5.9. Then any nonseparable signal $f \in V(\phi)$ is determined, up to a sign, by its phaseless samples on a shift-invariant set $\Gamma + \mathbb{Z}^d$ with sampling density*

$$D(\Gamma + \mathbb{Z}^d) \le \dim W_{(0,1)^d} \le \frac{1}{2}\#K_{(0,1)^d}(\#K_{(0,1)^d} + 1),$$

where $K_{(0,1)^d}$ is in (5.8).

The explicit construction of a discrete set with finite sampling density in Theorem 5.9 does not provide an algorithm to reconstruct a nonseparable signal from its phaseless samples taken on that discrete set. Considering the phaseless reconstruction of signals in a shift-invariant space, we introduce a local complement property on a set.

Definition 5.12. *We say that the shift-invariant space $V(\phi)$ has local complement property on a set A if for any $A' \subset A$, there does not exist $f, g \in V(\phi)$ such that $f, g \not\equiv 0$ on A, but $f(x) = 0$ for all $x \in A'$ and $g(y) = 0$ for all $y \in A \backslash A'$.*

The local complement property on $\mathbb{R}^d$ is the complement property in [40] for ideal sampling functionals on $V(\phi)$, cf. the complement property for frames in Hilbert/Banach spaces ([7, 20, 23, 33]).

Local complement property is closely related to local phase retrievability. In fact, following the argument in [40], the shift-invariant space $V(\phi)$ has the local complement property on A if and only if all signals in $V(\phi)$ is *local phase retrievable* on A, i.e., for any $f, g \in V(\phi)$ satisfying $|g(x)| = |f(x)|, x \in A$, there exists $\delta \in \{-1, 1\}$ such that $g(x) = \delta f(x)$ for all $x \in A$. More discussions on the local complement property are given in Appendix 5.5.

Theorem 5.13. *Let $A_1, \cdots, A_M$ be bounded open sets and ϕ be a compactly supported continuous function such that ϕ has local linear independence on $A_m, 1 \le m \le M$, and*

$$S_k \cap \big(\cup_{m=1}^{M} (A_m + \mathbb{Z}^d)\big) \neq \emptyset \tag{5.18}$$

for all $k \in \mathbb{Z}^d$ with S_k in (5.6) *being nonempty. If the shift-invariant space $V(\phi)$ has local complement property on $A_m, 1 \le m \le M$, then there exists a finite set $\Gamma \subset \cup_{m=1}^{M} A_m$ such that the following statements are equivalent for any signal $f \in V(\phi)$:*

(i) The signal f is determined, up to a sign, by its magnitude measurements on $\mathbb{R}^d$.

(ii) The graph $\mathcal{G}_f$ in (5.4) *is connected.*

(iii) The signal f is determined, up to a sign, by its phaseless samples $|f(y)|, y \in \Gamma + \mathbb{Z}^d$.

The implication (i)$\Longrightarrow$(ii) has been established in Theorem 5.2 and the implication (iii)$\Longrightarrow$(i) is obvious. Write $f = \sum_{k \in \mathbb{Z}^d} c(k)\phi(\cdot - k)$. To prove (ii)$\Longrightarrow$(iii), we first determine $c(k), k \in K_{A_m} + l$, up to a sign $\delta_{l,m} \in \{-1, 1\}$, by phaseless samples $|f(\gamma + l)|, \gamma \in \Gamma$, and then we use the connectivity of the graph $\mathcal{G}_f$ to adjust phases $\delta_{l,m}, 1 \le m \le M, l \in \mathbb{Z}^d$. Finally we sew those pieces together to recover amplitudes $c(k), k \in \mathbb{Z}^d$, and the signal f. Comparing with the proof of Theorem 5.9, we remark that our proof of Theorem 5.13 is constructive and a reconstruction algorithm can be developed.

Proof. The implication (iii)$\Longrightarrow$(i) is trivial. By (5.18), local linear independence of ϕ on $A_m, 1 \leq m \leq M$, and shift-invariance of the linear space $V(\phi)$, we obtain that the generator ϕ has the global linear independence. Then the implication (i)$\Longrightarrow$(ii) follows from Theorem 5.2.

Now it remains to prove (ii)$\Longrightarrow$(iii). Let $\Gamma_m, 1 \leq m \leq M$, be finite sets constructed in Proposition 5.25 with the set A and the space V replaced by A_m and $V(\phi)$ respectively, and set $\Gamma = \cup_{m=1}^{M} \Gamma_m$. Let $f, g \in V(\phi)$ satisfy

$$|g(y)| = |f(y)| \text{ for all } y \in \Gamma + \mathbb{Z}^d. \tag{5.19}$$

Then it suffices to prove that

$$g = \delta f \tag{5.20}$$

for some $\delta \in \{-1, 1\}$. Take $l \in \mathbb{Z}^d$ and $1 \leq m \leq M$. By Proposition 5.25 and the shift-invariance of the linear space $V(\phi)$, we have

$$|g(x+l)| = |f(x+l)|,\ x \in A_m.$$

This, together with the shift-invariance of the linear space $V(\phi)$ and local complement property on A_m, implies the existence of $\delta_{l,m} \in \{-1, 1\}$ such that

$$g(x) = \delta_{l,m} f(x),\ x \in A_m + l. \tag{5.21}$$

Write $f = \sum_{k \in \mathbb{Z}^d} c(k)\phi(\cdot - k)$ and $g = \sum_{k \in \mathbb{Z}^d} d(k)\phi(\cdot - k) \in V(\phi)$. Then it follows from (5.21) and local linear independence of the generator ϕ on A_m that

$$d(k'+l) = \delta_{l,m} c(k'+l),\ k' \in K_{A_m}, \tag{5.22}$$

where K_{A_m} is given in (5.8).

By (5.22), the proof of (5.20) reduces to showing

$$\delta_{l,m} = \delta \tag{5.23}$$

for all $l \in \mathbb{Z}^d$ and $1 \le m \le M$ so that $k' + l \in V_f$ for some $k' \in K_{A_m}$. Recall that $c(k) \ne 0$ for all $k \in V_f$. Then by (5.22) there exist $\delta_k \in \{-1, 1\}$ for all $k \in V_f$ such that

$$\delta_{l,m} = \delta_k$$

for all $l \in \mathbb{Z}^d$ and $1 \le m \le M$ so that $k = k' + l \in V_f$ for some $k' \in K_{A_m}$. Thus it suffices to prove that

$$\delta_k = \delta_{\tilde{k}} \ \text{ for all } \ k, \tilde{k} \in V_f. \tag{5.24}$$

By the connectivity of the graph $\mathcal{G}_f$, the proof of (5.24) reduces further to proving

$$\delta_k = \delta_{\tilde{k}} \tag{5.25}$$

for all edges $(k, \tilde{k})$ of the graph $\mathcal{G}_f$. For an edge $(k, \tilde{k})$ of the graph $\mathcal{G}_f$, we have that

$$S := \{x \in \mathbb{R}^d : \ \phi(x - k)\phi(x - \tilde{k}) \ne 0\} \ne \emptyset.$$

Then there exist $1 \le m \le M$ by (5.6) and (5.18) such that $S \cap (A_m + k) \ne \emptyset$. Thus $k, \tilde{k} \in K_{A_m} + k$, which together with (5.22) and (5.24) implies that $\delta_k = \delta_{k,m} = \delta_{\tilde{k}}$. Hence (5.25) holds. This completes the proof. $\square$

For the case that the generator ϕ has local linear independence on any open set, we can find open sets $A_m, 1 \le m \le M$, such that (5.18) holds and $V(\phi)$ has local complement property on $A_m, 1 \le m \le M$, see Proposition 5.28. Then from Theorem 5.13 we obtain the following corollary, cf.

Theorem 5.4 and Corollaries 5.7 and 5.8.

Corollary 5.14. *Let ϕ be a compactly supported continuous function with local linear independence on any open set. Then there exists a finite set Γ such that any nonseparable signal is determined, up to a sign, by its phaseless samples taken on the set $\Gamma + \mathbb{Z}^d$ with finite sampling density.*

Take $\mathbf{N} = (N_1, \ldots, N_d)^T$ with $N_i \geq 2, 1 \leq i \leq d$, and let B_{N_i} be the B-spline of order N_i ([57, 157, 160]). Define the box spline function of tensor-product type

$$B_{\mathbf{N}}(x) := B_{N_1}(x_1) \times \cdots \times B_{N_d}(x_d),\ x = (x_1, \ldots, x_d)^T \in \mathbb{R}^d. \tag{5.26}$$

As the restriction of a signal in $V(B_{\mathbf{N}})$ on $(0,1)^d$ is a polynomial of finite degree, the space $V(B_{\mathbf{N}})$ has local complement property on $(0,1)^d$. Applying Theorem 5.13 with $M = 1$ and $A_1 = (0,1)^d$, we have the following result for tensor-product splines, which is given in [40] for $d = 1$.

Corollary 5.15. *Let X_i contain $2N_i - 1$ distinct points in $(0,1), 1 \leq i \leq d$. Then any nonseparable signal $f \in V(B_{\mathbf{N}})$ is determined, up to a sign, by its phaseless samples on the set $X_1 \times \ldots \times X_d + \mathbb{Z}^d$ with sampling density $\prod_{i=1}^d (2N_i - 1)$.*

Proof. As restrictions of signals in $V(B_{\mathbf{N}})$ on $(0,1)^d$ are polynomials of finite degrees, the space $V(B_{\mathbf{N}})$ has the local complement property on $(0,1)^d$. Set $\mathbf{n} = (n, \ldots, n)$ for $n \in \mathbb{Z}$. It is observed that the function $\Phi_{(0,1)^d}$ in (5.7) is a vector-valued polynomial of degree $\mathbf{N} - \mathbf{1}$, and its outer product $\Phi_{(0,1)^d}(x)\Phi_{(0,1)^d}(x)^T, x \in (0,1)^d$ is a matrix-valued polynomial of degree $2\mathbf{N} - \mathbf{2}$. Recall that X_i is the discrete set containing $2N_i - 1$ distinct points in $(0,1), 1 \leq i \leq d$. Therefore $\Phi_{(0,1)^d}(y)\Phi_{(0,1)^d}(y)^T, y \in X_1 \times \cdots \times X_d$, is a spanning set of the linear space spanned by $\Phi_{(0,1)^d}(x)\Phi_{(0,1)^d}(x)^T, x \in (0,1)^d$. This together with Theorem 5.13 completes the proof. $\square$

In the proof of Theorem 5.13, the discrete sampling set Γ is chosen to be the union of $\Gamma_m \subset$

$A_m, 1 \leq m \leq M$,

$$\Gamma = \cup_{m=1}^{M} \Gamma_m, \tag{5.27}$$

so that outer products $\Phi_{A_m}(\gamma)(\Phi_{A_m}(\gamma))^T, \gamma \in \Gamma_m$, are a basis (or a spanning set) of the linear space W_{A_m}. Therefore we have the following result from Theorem 5.13.

Corollary 5.16. *Let ϕ and $A_m, 1 \leq m \leq M$, be as in Theorem 5.13. Then any nonseparable signal $f \in V(\phi)$ is determined, up to a sign, by its phaseless samples on a shift-invariant set $\Gamma + \mathbb{Z}^d$ with sampling density*

$$D(\Gamma + \mathbb{Z}^d) \leq \sum_{m=1}^{M} \dim W_{A_m} \leq \frac{1}{2} \sum_{m=1}^{M} \#K_{A_m}(\#K_{A_m} + 1).$$

The discrete set $\Gamma + \mathbb{Z}^d$ chosen in Corollary 5.16 may have larger sampling density than $\dim W_{(0,1)^d}$ in Corollary 5.11. Based on the constructive proof in Theorem 5.13, a robust reconstruction algorithm is developed in Section 5.4. However, we have difficulties to find a reconstruction algorithm from the phaseless samples taken on the set given in Corollary 5.11.

Definition 5.17. We say that $\mathcal{M} = \{a_m \in \mathbb{R}^d, 1 \leq m \leq M\}$ is a *phase retrievable frame* for $\mathbb{R}^d$ if any vector $x \in \mathbb{R}^d$ is determined, up to a sign, by its measurements $|\langle x, a_m \rangle|, a_m \in \mathcal{M}$, and that $\mathcal{M}$ is a *minimal phase retrieval frame* for $\mathbb{R}^d$ if any true subset of $\mathcal{M}$ is not a phase retrievable frame.

After the careful examination on the proof of Theorem 5.13, we can select a subset Γ' of Γ such that all nonseparable signals f can be reconstructed from its phaseless samples taken on $\Gamma' + \mathbb{Z}^d$ in a robust manner.

Theorem 5.18. *Let $A_m, 1 \leq m \leq M$, and ϕ be as in Theorem 5.13. Assume that there exist $\Gamma'_m \subset A_m$ such that $\Phi_{A_m}(\gamma'), \gamma' \in \Gamma'_m$, is a phase retrievable frame for $\mathbb{R}^{\#K_{A_m}}$. Then any nonseparable*

signal $f \in V(\phi)$ is determined, up to a sign, by its phaseless samples $|f(y)|, y \in \Gamma' + \mathbb{Z}^d$, where

$$\Gamma' = \cup_{m=1}^{M} \Gamma'_m. \tag{5.28}$$

Proof. Let $f = \sum_{k\in\mathbb{Z}^d} c(k)\phi(\cdot - k)$ and $g = \sum_{k\in\mathbb{Z}^d} d(k)\phi(\cdot - k)$ satisfy

$$|g(y)| = |f(y)| \text{ for all } y \in \Gamma' + \mathbb{Z}^d,$$

where $\Gamma' = \cup_{m=1}^{M} \Gamma'_m$ is given in (5.28). Take $l \in \mathbb{Z}^d$ and $1 \le m \le M$. Then

$$\Big|\sum_{k\in K_{A_m}+l} d(k)\phi(\gamma' + l - k)\Big| = \Big|\sum_{k\in K_{A_m}+l} c(k)\phi(\gamma' + l - k)\Big| \quad \text{for all } \gamma' \in \Gamma'_m.$$

By the assumption on $\Phi_{A_m}(\gamma'), \gamma' \in \Gamma'_m, 1 \le m \le M$, there exists $\delta_{l,m} \in \{1, -1\}$ such that

$$d(k) = \delta_{l,m} c(k),\ k \in K_{A_m} + l.$$

Following the same argument as the one used for the implication (ii)$\Longrightarrow$(iii) in Theorem 5.13, we can find $\delta \in \{-1, 1\}$ such that $\delta_{l,m} = \delta$ for all $l \in \mathbb{Z}^d$ and $1 \le m \le M$. This completes the proof. $\square$

In Theorem 5.18, the requirement on the sampling set is a bit weaker than the one in Theorem 5.13, as for the sampling set $\Gamma = \cup_{m=1}^{M} \Gamma_m$ in (5.27), $\Phi_{A_m}(\gamma), \gamma \in \Gamma_m$, is a phase retrievable frame for $\mathbb{R}^{\#K_{A_m}}$, cf. Theorem 5.26. We remark that the phase retrieval frame property for $\Phi_A(\gamma'), \gamma' \in \Gamma'$, may not imply that their out products $\Phi_A(\gamma')(\Phi_A(\gamma'))^T, \gamma' \in \Gamma'$, form a basis (or a spanning set) of W_A in (5.16), as shown in the following example.

Example 5.19. Let

$$\phi_1(x) = \begin{cases} x^3/2 & \text{if } 0 \le x < 1 \\ -x^3 + 3x^2 - 2x + 1/2 & \text{if } 1 \le x < 2 \\ x^3/2 - 3x^2 + 5x - 3/2 & \text{if } 2 \le x < 3 \\ 0 & \text{otherwise,} \end{cases}$$

and set $\Phi_1(x) = (\phi_1(x), \phi_1(x+1), \phi_1(x+2))^T, 0 \le x < 1$. Then

$$\Phi_1(x) = \frac{1}{2}\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} x + \frac{1}{2}\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} x^3,$$

and

$$\Phi_1(x)\Phi_1(x)^T = \frac{1}{4}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} x + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{pmatrix} x^2$$
$$+ \frac{1}{4}\begin{pmatrix} 0 & 1 & 1 \\ 1 & -4 & -1 \\ 1 & -1 & 2 \end{pmatrix} x^3 + \frac{1}{2}\begin{pmatrix} 0 & 1 & -1 \\ 1 & -4 & 3 \\ -1 & 3 & -2 \end{pmatrix} x^4 + \frac{1}{4}\begin{pmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{pmatrix} x^6.$$

Therefore the space spanned by $\Phi_1(x), 0 < x < 1$, is $\mathbb{R}^3$, and the space $W_{(0,1)}$ spanned by $\Phi_1(x)\Phi_1(x)^T, 0 < x < 1$, is the 6-dimensional linear space of symmetric matrices of size 3×3.

On the other hand, any 3×3 square submatrices of

$$\Big(\Phi_1(0)\ \Phi_1\big(\frac{1}{5}\big)\ \Phi_1\big(\frac{2}{5}\big)\ \Phi_1\big(\frac{3}{5}\big)\ \Phi_1\big(\frac{4}{5}\big)\Big) = \frac{1}{250}\begin{pmatrix} 0 & 1 & 8 & 27 & 64\\ 125 & 173 & 209 & 221 & 197\\ 125 & 76 & 33 & 2 & -11\end{pmatrix}$$

is nonsingular, which implies that $\Phi_1(m/5), 0\le m\le 4$, forms a phase retrieval frame for $\mathbb{R}^3$, but their out products do not form a spanning set of the 6-dimensional space $W_{(0,1)}$.

The problem how to pare down a phase retrieval frame to a minimal phase retrieval frame will be discussed in our future work. Using the pare-down technique, we may find a discrete set X with smaller sampling density such that nonseparable signals in the shift-invariant space can be reconstructed from their phaseless samples taken on X.

5.3 Stability of phaseless sampling and reconstruction

Stability is of paramount importance in the phaseless sampling and reconstruction problem. Consider the scenario that phaseless samples of a signal

$$f=\sum_{k\in\mathbb{Z}^d} c(k)\phi(\cdot-k)\in V(\phi) \tag{5.29}$$

taken on a shift-invariant set $\Gamma+\mathbb{Z}^d$ are corrupted by the additive noise,

$$z_\epsilon(y)=|f(y)|+\epsilon(y),\ y\in\Gamma+\mathbb{Z}^d, \tag{5.30}$$

where $\epsilon = (\epsilon(y))_{y\in\Gamma+\mathbb{Z}^d}$ has the bounded noise level $\|\epsilon\|_\infty = \max_{y\in\Gamma+\mathbb{Z}^d} |\epsilon(y)|$, and $\Gamma = \cup_{m=1}^{M}\Gamma_m$ is either as in (5.27) or in (5.28). In this section, we construct an approximation

$$f_\epsilon = \sum_{k\in\mathbb{Z}^d} c_\epsilon(k)\phi(\cdot - k) \in V(\phi), \tag{5.31}$$

up to a sign, to the original signal f in (5.29) when the noisy phaseless samples (5.30) are available only.

Let

$$\Omega_m = \{k \in \mathbb{Z}^d : \ \phi(\gamma - k) \neq 0 \text{ for some } \gamma \in \Gamma_m\},\ 1 \leq m \leq M, \tag{5.32}$$

and define the hard threshold function $H_\eta, \eta \geq 0$, by

$$H_\eta(t) = \begin{cases} t & \text{if } |t| \geq \eta \\ 0 & \text{if } |t| < \eta. \end{cases}$$

Based on the constructive proofs of Theorems 5.13 and 5.18, we propose the following four-step approach with its implementation discussed in Section 5.4.

1. Select a phase adjustment threshold value $M_0 \geq 0$ and an amplitude threshold value $\eta = \sqrt{M_0}$.

2. For $l \in \mathbb{Z}^d$ and $1 \leq m \leq M$, let

$$c_{\epsilon,l;m} = (c_{\epsilon,l;m}(k))_{k \in \mathbb{Z}^d} \tag{5.33}$$

take zero components except that $c_{\epsilon,l;m}(k), k \in l + \Omega_m$, are solutions of the local minimization problem

$$\min_{c(k), k \in l+\Omega_m} \sum_{\gamma \in \Gamma_m} \left| \Big| \sum_{k \in l+\Omega_m} c(k)\phi(\gamma + l - k) \Big| - z_\epsilon(\gamma + l) \right|^2. \tag{5.34}$$

3. Adjust phases of $c_{\epsilon,l;m}$ appropriately so that the resulting vectors $\delta_{l,m}c_{\epsilon,l;m}$ with $\delta_{l,m} \in \{-1, 1\}$ satisfy

$$\langle \delta_{l,m}c_{\epsilon,l;m}, \delta_{l',m'}c_{\epsilon,l';m'} \rangle \geq -M_0 \tag{5.35}$$

for all $l, l' \in \mathbb{Z}^d$ and $1 \leq m, m' \leq M$.

4. Sew vectors $\delta_{l,m}c_{\epsilon,l;m}, l \in \mathbb{Z}^d, 1 \leq m \leq M$, together to obtain

$$d_\epsilon(k) = \frac{\sum_{m=1}^{M} \sum_{l \in \mathbb{Z}^d} \delta_{l,m}c_{\epsilon,l;m}(k)}{\sum_{m=1}^{M} \sum_{l \in \mathbb{Z}^d} \chi_{l+\Omega_m}(k)},\ k \in \mathbb{Z}^d. \tag{5.36}$$

5. Threshold the vector $d_\epsilon = (d_\epsilon(k))_{k \in \mathbb{Z}^d}$,

$$c_\epsilon(k) = H_\eta(d_\epsilon(k)),\ k \in \mathbb{Z}^d \tag{5.37}$$

to construct the approximation f_ϵ in (5.31).

In the next theorem, we show that the above approach provides a suboptimal approximation to the original signal in a noisy phaseless sampling environment.

Theorem 5.20. *Let $A_1, \cdots, A_M$ be bounded open sets satisfying* (5.18)*, ϕ be a compactly supported continuous function such that ϕ has local linear independence on $A_m, 1 \le m \le M$, and let $\Gamma_m \subset A_m$ be so chosen that $\Phi_{A_m}(\gamma), \gamma \in \Gamma_m$, is a phase retrievable frame for $\mathbb{R}^{\#K_{A_m}}$. Assume that the graph $\mathcal{G}_f = (V_f, E_f)$ of the original signal $f = \sum_{k\in\mathbb{Z}^d} c(k)\phi(\cdot - k)$ is connected and*

$$F_0 := \inf_{k\in V_f} |c(k)|^2 > 0. \tag{5.38}$$

Set $\Gamma = \cup_{m=1}^M \Gamma_m$ and

$$\left\|\Phi^{-1}\right\|_2 = \sup_{\Theta_m\subset\Gamma_m, 1\le m\le M} \Big(\min\big(\sup_{\|d\|_2=1} \|\Phi_{\Theta_m} d\|_2^{-1}, \sup_{\|d\|_2=1} \|\Phi_{\Gamma_m\setminus\Theta_m} d\|_2^{-1}\big)\Big)^{-1}, \tag{5.39}$$

where $\Phi_{\Theta_m} = (\phi(\gamma - k))_{\gamma\in\Theta_m, k\in\Omega_m}$ for $\Theta_m \subset \Gamma_m$. If the phase adjustment threshold value $M_0 \ge 0$ and the noise level $\|\epsilon\|_\infty := \sup_{y\in\Gamma+\mathbb{Z}^d} |\epsilon(y)|$ satisfy

$$M_0 \le \frac{2F_0}{9}, \tag{5.40}$$

and

$$8\#\Gamma\|\Phi^{-1}\|_2^2\|\epsilon\|_\infty^2 \le M_0, \tag{5.41}$$

then the signal $f_\epsilon = \sum_{k\in\mathbb{Z}^d} c_\epsilon(k)\phi(\cdot - k) \in V(\phi)$ reconstructed from the proposed approach (5.33)–(5.37) *satisfies*

$$|c_\epsilon(k) - \delta c(k)| \le 2\sqrt{\#\Gamma}\|\Phi^{-1}\|_2\|\epsilon\|_\infty,\ k \in V_f \tag{5.42}$$

and

$$c_\epsilon(k) = c(k) = 0,\ k \notin V_f, \tag{5.43}$$

where $\delta \in \{-1, 1\}$.

Given $\Gamma \subset \mathbb{R}^d$ and $f = \sum_{k\in\mathbb{Z}^d} c(k)\phi(\cdot - k)$, we define

$$\widetilde{\mathcal{G}}_{f,\Gamma} = (V_f, E_{f,\Gamma}), \tag{5.44}$$

where $(k, k') \in E_{f,\Gamma}$ only if $\phi(y-k)\phi(y-k') \neq 0$ for some $y \in \Gamma + \mathbb{Z}^d$. To prove Theorem 5.20, we need a lemma about the graph $\mathcal{G}_f$.

Lemma 5.21. *Let* ϕ*,* A_m *and* $\Gamma_m, 1 \le m \le M$*, be as in Theorem 5.20. Set* $\Gamma = \cup_{m=1}^M \Gamma_m$*. Then for any* $f \in V(\phi)$*, the graph* $\mathcal{G}_f$ *in* (5.4) *and* $\widetilde{\mathcal{G}}_{f,\Gamma}$ *in* (5.44) *are the same,*

$$\mathcal{G}_f = \widetilde{\mathcal{G}}_{f,\Gamma}. \tag{5.45}$$

Proof. Clearly it suffices to prove that an edge in $\mathcal{G}_f$ is also an edge in $\widetilde{\mathcal{G}}_{f,\Gamma}$. Suppose, on the contrary, that there exists an edge (k, k') in $\mathcal{G}_f$ such that

$$\phi(y-k)\phi(y-k') = 0 \text{ for all } y \in \cup_{m=1}^M \Gamma_m + \mathbb{Z}^d. \tag{5.46}$$

Define

$$S = \{x \in \mathbb{R}^d :\ \phi(x-k)\phi(x-k') \neq 0\} \neq \emptyset. \tag{5.47}$$

By (5.18), there exist $l_0 \in \mathbb{Z}^d$ and $1 \le m_0 \le M$ such that

$$S \cap (A_{m_0} + l_0) \neq \emptyset. \tag{5.48}$$

Set $g_\pm(x) = \phi(x + l_0 - k) \pm \phi(x + l_0 - k'), x \in A_{m_0}$. Then it follows from (5.46) that

$$|g_\pm(\gamma)| = |\phi(\gamma + l_0 - k)| + |\phi(\gamma + l_0 - k')|,\ \gamma \in \Gamma_{m_0}.$$

By the construction of the set Γ_{m_0}, we get either $g_+ = g_-$ or $g_+ = -g_-$ on A_{m_0}. Therefore either $\phi(x + l_0 - k) \equiv 0$ on A_{m_0} or $\phi(x + l_0 - k') \equiv 0$ on A_{m_0}. This contradicts to the construction of set S in (5.47) and (5.48). $\square$

Now, we continue the proof of Theorem 5.20.

Proof. Take $l \in \mathbb{Z}^d$ and $1 \le m \le M$. For $\gamma \in \Gamma_m$ there exists $\tilde{\delta}_{\gamma,l;m} \in \{-1, 1\}$ such that

$$\begin{aligned}
& \left(\sum_{\gamma\in\Gamma_m} \left| \sum_{k\in l+\Omega_m} (c_{\epsilon,l;m}(k) - \tilde{\delta}_{\gamma,l;m} c(k))\phi(\gamma + l - k) \right|^2 \right)^{1/2} \\
= & \left(\sum_{\gamma\in\Gamma_m} \left| \Big| \sum_{k\in l+\Omega_m} c_{\epsilon,l;m}(k)\phi(\gamma + l - k) \Big| \right. \right. \\
& \qquad \left. \left. - \Big| \sum_{k\in l+\Omega_m} c(k)\phi(\gamma + l - k) \Big| \right|^2 \right)^{1/2} \\
\le & \left(\sum_{\gamma\in\Gamma_m} \left| \Big| \sum_{k\in l+\Omega_m} c_{\epsilon,l;m}(k)\phi(\gamma + l - k) \Big| - z_\epsilon(\gamma + l) \right|^2 \right)^{1/2} \\
& + \left(\sum_{\gamma\in\Gamma_m} \left| \Big| \sum_{k\in l+\Omega_m} c(k)\phi(\gamma + l - k) \Big| - z_\epsilon(\gamma + l) \right|^2 \right)^{1/2} \\
\le & \ 2 \left(\sum_{\gamma\in\Gamma_m} \left| \Big| \sum_{k\in l+\Omega_m} c(k)\phi(\gamma + l - k) \Big| - z_\epsilon(\gamma + l) \right|^2 \right)^{1/2} \\
\le & \ 2\sqrt{\#\Gamma_m}\|\epsilon\|_\infty \le 2\sqrt{\#\Gamma}\|\epsilon\|_\infty, \qquad (5.49)
\end{aligned}$$

where the second inequality holds by (5.34) and the last inequality follows from

$$z_\epsilon(\gamma+l)=\Big|\sum_{k\in l+\Omega_m} c(k)\phi(\gamma+l-k)\Big|+\epsilon(\gamma+l),\ \gamma\in\Gamma_m.$$

From the phase retrievable frame property for $\big(\phi(\gamma-k)\big)_{k\in K_{A_m}}, \gamma\in\Gamma_m$, we obtain that

$$\Omega_m=K_{A_m},\ 1\le m\le M. \tag{5.50}$$

Let $A_{1,m}=\{\gamma\in\Gamma_m:\ \tilde\delta_{\gamma,l;m}=1\}$. This together with (5.50) and the phase retrievable frame assumption that either $\big(\phi(\gamma-k)\big)_{k\in\Omega_m}, \gamma\in A_{1,m}$ or $\big(\phi(\gamma-k)\big)_{k\in\Omega_m}, \gamma\in\Gamma_m\backslash A_{1,m}$ is a spanning set for $\mathbb{R}^{\#\Omega_m}$. This together with (5.49) implies that

$$\Big(\sum_{k\in l+\Omega_m}\big|c_{\epsilon,l;m}(k)-\tilde\delta_{l,m}c(k)\big|^2\Big)^{1/2}\le 2\|\Phi^{-1}\|_2\sqrt{\#\Gamma}\|\epsilon\|_\infty \tag{5.51}$$

for some sign $\tilde\delta_{l,m}\in\{-1,1\}$.

Now we show that phases of $c_{\epsilon,l;m}, l\in\mathbb{Z}^d, 1\le m\le M$, can be adjusted so that (5.35) holds. Let $\tilde\delta_{l,m}, l\in\mathbb{Z}^d, 1\le m\le M$, be as in (5.51). Then for any $l,l'\in\mathbb{Z}^d$ and $1\le m,m'\le M$, set

$\Omega_{l,m;l',m'} = (\Omega_m + l) \cap (\Omega_{m'} + l')$. Then

$$\begin{aligned}
& \langle \tilde{\delta}_{l,m} c_{\epsilon,l;m}, \tilde{\delta}_{l',m'} c_{\epsilon,l';m'} \rangle = \sum_{k\in\Omega_{l,m;l',m'}} \tilde{\delta}_{l,m}\tilde{\delta}_{l',m'} c_{\epsilon,l;m}(k) c_{\epsilon,l';m'}(k) \\
\geq \quad & \sum_{k\in\Omega_{l,m;l',m'}} |c(k)|^2 - \sum_{k\in\Omega_{l,m;l',m'}} |c(k)||\tilde{\delta}_{l',m'} c_{\epsilon,l';m'}(k) - c(k)| \\
& - \sum_{k\in\Omega_{l,m;l',m'}} |\tilde{\delta}_{l,m} c_{\epsilon,l;m}(k) - c(k)||c(k))| \\
& - \sum_{k\in\Omega_{l,m;l',m'}} |\tilde{\delta}_{l,m} c_{\epsilon,l;m}(k) - c(k)||\tilde{\delta}_{l',m'} c_{\epsilon,l';m'}(k) - c(k)| \\
\geq \quad & -\frac{1}{2} \sum_{k\in\Omega_{l,m;l',m'}} \Big(|\tilde{\delta}_{l',m'} c_{\epsilon,l';m'}(k) - c(k)|^2 + |\tilde{\delta}_{l,m} c_{\epsilon,l;m}(k) - c(k)|^2 \Big) \\
& - \sum_{k\in\Omega_{l,m;l',m'}} |\tilde{\delta}_{l,m} c_{\epsilon,l;m}(k) - c(k)||\tilde{\delta}_{l',m'} c_{\epsilon,l';m'}(k) - c(k)| \\
\geq \quad & -8\|\Phi^{-1}\|_2^2 \#\Gamma \|\epsilon\|_\infty^2 \geq -M_0,
\end{aligned} \tag{5.52}$$

where the third inequality follows from (5.51) and the last inequality holds by the assumption (5.41) on the noise level $\|\epsilon\|_\infty$ and the threshold value M_0.

The phase adjustments in (5.35) for $c_{\epsilon,l;m}, l \in \mathbb{Z}^d, 1 \leq m \leq M$, are not unique. Next we show that they are essentially the phase adjustments in (5.52), i.e., for any phase adjustments $\delta_{l,m} \in \{-1, 1\}$ in (5.35) there exists $\delta \in \{-1, 1\}$ such that

$$\delta_{l,m} c(k) = \delta \tilde{\delta}_{l,m} c(k) \quad \text{for all } k \in l + \Omega_m, l \in \mathbb{Z}^d, 1 \leq m \leq M. \tag{5.53}$$

To prove (5.53), we claim that

$$\tilde{\delta}_{l,m}/\delta_{l,m} = \delta_{l',m'}/\tilde{\delta}_{l',m'} \tag{5.54}$$

for all (l, m) and (l', m') with $\Omega_{l,m;l',m'} \cap V_f \neq \emptyset$. Suppose on the contrary that (5.54) does not

hold. Then

$$\langle \delta_{l,m} c_{\epsilon,l;m}, \delta_{l',m'} c_{\epsilon,l';m'} \rangle = -\langle \tilde{\delta}_{l,m} c_{\epsilon,l;m}, \tilde{\delta}_{l',m'} c_{\epsilon,l';m'} \rangle.$$

Therefore

$$\begin{aligned}
& \langle \delta_{l,m} c_{\epsilon,l;m}, \delta_{l',m'} c_{\epsilon,l';m} \rangle \\
\leq \ & - \sum_{k \in \Omega_{l,m;l',m'}} |c(k)|^2 + \sum_{k \in \Omega_{l,m;l',m'}} |c(k)| |\tilde{\delta}_{l',m'} c_{\epsilon,l';m'}(k) - c(k)| \\
& + \sum_{k \in \Omega_{l,m;l',m'}} |\tilde{\delta}_{l,m} c_{\epsilon,l;m}(k) - c(k)| |c(k))| \\
& + \sum_{k \in \Omega_{l,m;l',m'}} |\tilde{\delta}_{l,m} c_{\epsilon,l;m}(k) - c(k)| |\tilde{\delta}_{l',m'} c_{\epsilon,l';m'}(k) - c(k)| \\
\leq \ & - \sum_{k \in \Omega_{l,m;l',m'}} |c(k)|^2 + 4\sqrt{\#\Gamma} \|\Phi^{-1}\|_2 \Big(\sum_{k \in \Omega_{l,m;l',m'}} |c(k)|^2 \Big)^{1/2} \|\epsilon\|_\infty \\
& + 4\#\Gamma \|\Phi^{-1}\|_2^2 \|\epsilon\|_\infty^2 \\
\leq \ & - \sum_{k \in \Omega_{l,m;l',m'}} |c(k)|^2 + \Big(2M_0 \sum_{k \in \Omega_{l,m;l',m'}} |c(k)|^2 \Big)^{1/2} + \frac{M_0}{2} < -M_0,
\end{aligned}$$

where the second inequality follows from (5.51), and the third and fourth inequalities hold by (5.38), (5.40) and (5.41). This contradicts to the requirement (5.35) for the phase adjustment and hence completes the proof of the Claim (5.54).

By (5.54), for any $k \in V_f$ there exists $\delta_k \in \{-1, 1\}$ such that

$$\delta_{l,m} c(k) = \delta_k \tilde{\delta}_{l,m} c(k) \quad \text{for all } k \in l + \Omega_m. \tag{5.55}$$

Let (k_1, k_2) be an edge in $\mathcal{G}_f$. By Lemma 5.21 there exist $l \in \mathbb{Z}^d$ and $1 \leq m \leq M$ such that $k_1, k_2 \in \Omega_m + l$. Therefore

$$\delta_{l,m} c(k_1) = \delta_{k_1} \tilde{\delta}_{l,m} c(k_1) \text{ and } \delta_{l,m} c(k_2) = \delta_{k_2} \tilde{\delta}_{l,m} c(k_2)$$

by (5.55). This implies that $\delta_{k_1} = \delta_{k_2}$ for any edge (k_1, k_2) in $\mathcal{G}_f$. Combining it with the connectivity of the graph $\mathcal{G}_f$, we can find $\delta \in \{-1, 1\}$ such that

$$\delta_k = \delta \ \text{ for all } k \in V_f. \tag{5.56}$$

Combining (5.55) and (5.56) proves (5.53).

By (5.51) and (5.53), we obtain

$$\begin{aligned} |d_\epsilon(k) - \delta c(k)| &\le \frac{\sum_{m=1}^M \sum_{l\in\mathbb{Z}^d} |\delta_{l,m} c_{\epsilon,l;m}(k) - \delta c(k)|}{\sum_{m=1}^M \sum_{l\in\mathbb{Z}^d} \chi_{l+\Omega_m}(k)} \\ &= \frac{\sum_{m=1}^M \sum_{l\in\mathbb{Z}^d} |c_{\epsilon,l;m}(k) - \tilde{\delta}_{l,m} c(k)|}{\sum_{m=1}^M \sum_{l\in\mathbb{Z}^d} \chi_{l+\Omega_m}(k)} \\ &\le 2\sqrt{\#\Gamma}\|\Phi^{-1}\|_2\|\epsilon\|_\infty, \ k \in \mathbb{Z}^d. \end{aligned} \tag{5.57}$$

This together with (5.40) and (5.41) implies that

$$|d_\epsilon(k)| \ge \frac{3}{2}\sqrt{M_0} \ \text{ for all } k \in V_f, \tag{5.58}$$

and

$$|d_\epsilon(k)| \le \frac{1}{2}\sqrt{M_0} \ \text{ for all } k \notin V_f. \tag{5.59}$$

Combining (5.37), (5.57), (5.58) and (5.59) completes the proof of the desired error estimates (5.42) and (5.43). $\square$

By Theorem 5.20, the reconstructed signal f_ϵ in (5.31) provides a suboptimal approximation, up to a sign, to the original signal f in (5.29),

$$\|f_\epsilon - \delta f\|_\infty \le 2\sqrt{\#\Gamma}\|\Phi^{-1}\|_2\Big(\sup_{x\in\mathbb{R}^d} \sum_{k\in\mathbb{Z}^d} \big|\phi(x-k)\big|\Big)\|\epsilon\|_\infty \tag{5.60}$$

and

$$\sup_{y\in\Gamma+\mathbb{Z}^d} \big||f_\epsilon(y)|-|f(y)|\big| \le 2\sqrt{\#\Gamma}\|\Phi^{-1}\|_2\Big(\sup_{x\in\mathbb{R}^d}\sum_{k\in\mathbb{Z}^d}\big|\phi(x-k)\big|\Big)\|\epsilon\|_\infty. \tag{5.61}$$

By (5.38), (5.40), (5.41) and (5.42), a vertex in the graph $\mathcal{G}_f$ is also a vertex of the graph $\mathcal{G}_{f_\epsilon}$. This together with (5.4) and (5.43) implies that the graphs $\mathcal{G}_f$ and $\mathcal{G}_{f_\epsilon}$ associated with the original signal f and the reconstructed signal f_ϵ are the same, i.e.,

$$\mathcal{G}_f = \mathcal{G}_{f_\epsilon}.$$

The selection of the threshold value $M_0 \ge 0$ is imperative to find an approximation to the original signal from its phaseless samples. In the noiseless environment, we may take $M_0 = 0$ and the proposed approach leads to a perfect reconstruction, i.e., $f_\epsilon = \pm f$, when f is nonseparable. In practical applications, the noise level is usually positive and the phase adjustment threshold value M_0 needs to be appropriately selected. For instance, we may require that (5.40) and (5.41) are satisfied if we have some prior information about the amplitude vector of the original signal. From the proof of Theorem 5.20 and also the simulations in the next section, it is observed that phases can not be adjusted to satisfy (5.35) if the threshold value M_0 is far below square of noise level $\|\epsilon\|_\infty$ (for instance, (5.41) is not satisfied), while the phase adjustment (5.35) in the algorithm is not essentially determined and hence the reconstructed signal is not a good approximation of the original signal if the threshold value M_0 is much larger than the square of minimal magnitude of amplitude vector of the original signal (for instance, (5.40) is not satisfied).

Remark 5.22. By Theorem 5.20, there is no resonance phenomenon in the sense that

$$\inf_{\delta\in\{-1,1\}} \|f_\epsilon - \delta f\|_\infty \le C\|\epsilon\|_\infty \tag{5.62}$$

if the noise level is far below the minimal magnitude of amplitude vector of the original signal, i.e.,

$$\|\epsilon\|_\infty \le C_0 \inf_{k\in V_f} |c(k)| \tag{5.63}$$

for some sufficiently small constant C_0. The phaseless sampling and reconstruction problem is ill-posed if the noise level is high. For instance, the estimate (5.62) is not satisfied for the following nonseparable spline signal of order 2,

$$f_\alpha(x) = B_2(x) + \alpha B_2(x-1) + B_2(x-2) \in V(B_2),$$

if $\|\epsilon\|_\infty \ge 2\alpha/(1+\alpha)$, where $\alpha \in (0,1)$ is sufficiently small. The reasons are that the signal $\tilde{f}_\alpha(x) = B_2(x) + \alpha B_2(x-1) - B_2(x-2) \in V(B_2)$ satisfies

$$\min_{\delta\in\{-1,1\}} \|f_\alpha - \delta \tilde{f}_\alpha\|_\infty = 2 \ \text{ and } \ \big\||f_\alpha| - |\tilde{f}_\alpha|\big\|_\infty = \frac{2\alpha}{1+\alpha}.$$

5.4 Reconstruction algorithm and numerical simulations

Consider the scenario that phaseless samples of a signal $f = \sum_{k\in\mathbb{Z}^d} c(k)\phi(\cdot - k) \in V(\phi)$ taken on a finite set $\Gamma + K \subset \Gamma + \mathbb{Z}^d$ are corrupted by the additive noise,

$$z_\epsilon(y) = |f(y)| + \epsilon(y),\ y \in \Gamma + K, \tag{5.64}$$

where $\epsilon(y) \in [-\epsilon, \epsilon], y \in \Gamma + K$, for some $\epsilon \ge 0$, and $\Gamma = \cup_{m=1}^M \Gamma_m$ is either as in (5.27) or in (5.28). Define

$$f_K = \sum_{k\in\widetilde{K}} c(k)\phi(\cdot - k), \tag{5.65}$$

where $\widetilde{K} = \cup_{l\in K} \cup_{m=1}^{M} (l+\Omega_m)$ and $\Omega_m, 1 \le m \le M$, are as in (5.32). Then the noisy data $z_\epsilon(y), y \in \Gamma + K$, in (5.64) is

$$z_\epsilon(y) = |f_K(y)| + \epsilon(y) \ge 0,\ y \in \Gamma + K. \tag{5.66}$$

Based on (5.66) and the four-step approach in Section 5.3, we propose an algorithm to find an approximation f_ϵ of the form

$$f_\epsilon = \sum_{k\in\widetilde{K}} c_\epsilon(k)\phi(\cdot - k) \in V(\phi), \tag{5.67}$$

up to a sign, to the original signal f in (5.65) when the noisy phaseless samples (5.64) are available only. The algorithm contains four parts: **m**inimiztion, **a**djusting **p**hases, **se**wing and **t**hresholding, and we call it the MAPSET algorithm. In this section, we also demonstrate the performance of the proposed MAPSET algorithm on reconstructing box spline signals from their noisy phaseless samples on discrete sets.

5.4.1 Nonseparable spline signals of tensor-product type

Let $B_{(3,3)}$ be the tensor product of one-dimensional quadratic spline B_3, see (5.26). For $A = (0,1)^2$ and $\phi = B_{(3,3)}$, the vector-valued function Φ_A in (5.7) and the set K_A in (5.8) become

$$\Phi_{(0,1)^2}(s,t) = \big(b_i(s)b_j(t)\big)_{(i,j)\in K_{(0,1)^2}},\ (s,t) \in (0,1)^2 \tag{5.68}$$

and $K_{(0,1)^2} = \{(i,j) : -2 \le i, j \le 0\}$ respectively, where $b_0(s) = s^2/2$, $b_{-1}(s) = (-2s^2 + 2s + 1)/2$ and $b_{-2}(s) = (1-s)^2/2, 0 \le s \le 1$. One may verify that the space spanned by the outer

Algorithm 2 MAPSET Algorithm

Inputs: finite set $K \subset \mathbb{Z}^d$; sampling set $\Gamma = \cup_{m=1}^{M}\Gamma_m$ either in (5.27) or in (5.28); noisy phaseless sampling data $\big(z_\epsilon(y)\big)_{y\in\Gamma+K}$; index set $\widetilde{K} = \cup_{l\in K}\cup_{m=1}^{M}(l+\Omega_m) \subset \mathbb{Z}^d$; and the phase adjustment threshold value M_0.

Initials: Start from zero vectors $c_{\epsilon,l;m} = \big(c_{\epsilon,l;m}(k)\big)_{k\in\widetilde{K}}, l \in K, 1 \le m \le M$.

Instructions:

1) Local minimization: For $l \in K$ and $1 \le m \le M$, replace $c_{\epsilon,l;m}(k), k \in l+\Omega_m$, by a solution of the local minimization problem

$$\min_{c(k), k\in l+\Omega_m} \sum_{\gamma\in\Gamma_m} \left|\Big|\sum_{k\in l+\Omega_m} c(k)\phi(\gamma+l-k)\Big| - z_\epsilon(\gamma+l)\right|^2.$$

2) Phase adjustment: For $l \in K$ and $1 \le m \le M$, multiply $c_{\epsilon,l;m}$ by $\delta_{l,m} \in \{-1,1\}$ so that $\langle \delta_{l,m}c_{\epsilon,l;m}, \delta_{l',m'}c_{\epsilon,l';m'}\rangle \ge -M_0$ for all $l, l' \in K$ and $1 \le m, m' \le M$.

3) Sewing local approximations:

$$d_\epsilon(k) = \frac{\sum_{m=1}^{M}\sum_{l\in K}\delta_{l,m}c_{\epsilon,l;m}(k)}{\sum_{m=1}^{M}\sum_{l\in K}\chi_{l+\Omega_m}(k)},\ k \in \widetilde{K}.$$

4) Hard thresholding:

$$c_\epsilon(k) = \begin{cases} d_\epsilon(k) & \text{if } |d_\epsilon(k)| \ge \sqrt{M_0} \\ 0 & \text{else} \end{cases},\ k \in \widetilde{K}.$$

Outputs: Amplitude vector $(c_\epsilon(k))_{k\in\widetilde{K}}$, and the reconstructed signal $f_\epsilon = \sum_{k\in\widetilde{K}} c_\epsilon(k)\phi(\cdot - k)$.

products of $\Phi_{(0,1)^2}(s,t), (s,t) \in (0,1)^2$, has dimension 25, and the set

$$\Gamma_0 = \{(i,j)/6, 1 \le i,j \le 5\} \subset (0,1)^2 \tag{5.69}$$

with cardinality 25 satisfies (5.27), see Figure 5.1. For the above uniformly distributed set Γ_0, the corresponding $\|\Phi^{-1}\|_2$ in (5.39) is 2.7962×10^3.

As $\Phi_{(0,1)^2}(s,t), (s,t) \in (0,1)^2$, is a 9-dimensional vector-valued polynomial about $s^m t^n, 0 \le m,n \le 2$, the shift-invariant space generated by $B_{(3,3)}$ has local complement property on $(0,1)^2$. Observe that the matrix $\big(\Phi_{(0,1)^2}(s_i,t_i)\big)_{1\le i\le 9}$ has full rank 9 for almost all $(s_i,t_i) \in (0,1)^2, 1 \le$

$i \leq 9$. Hence $\left(\Phi_{(0,1)^2}(s_i, t_i)\right)_{1\leq i\leq 17}$ is a phase retrieval frame for almost all $(s_i, t_i) \in (0,1)^2, 1 \leq i \leq 17$, but the corresponding $\|\Phi^{-1}\|_2$ in (5.39) are relatively large from our calculation. So we use a randomly distributed set $\Gamma_1 \subset (0,1)^2$ with cardinality 19 in our simulations, see Figure 5.1. The above set satisfies (5.28) and the corresponding $\|\Phi^{-1}\|_2$ in (5.39) is 3.2995×10^4.

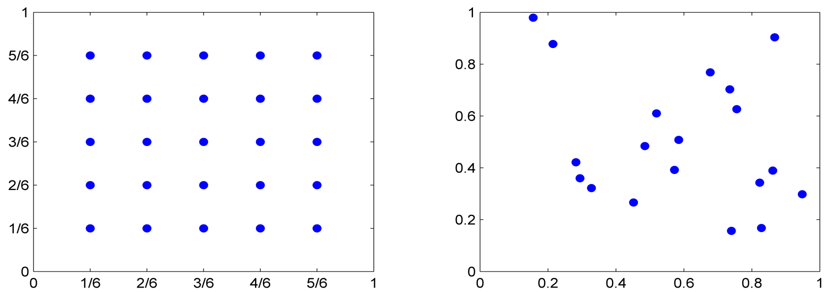

Figure 5.1: Uniform and randomly distributed sampling set

Plotted in Figure 5.1 on the left is a uniformly distributed set Γ_0 satisfying (5.27), while on the right is a randomly distributed set Γ_1 satisfying (5.28). The corresponding $\|\Phi^{-1}\|_2$ in (5.39) to the above sets are 2.7962×10^3 (left) and 3.2995×10^4 (right), respectively.

In our simulations, the available data $z_\epsilon(y) = |f(y)| + \epsilon(y) \geq 0, y \in \Gamma + K$, are noisy phaseless samples of a spline signal

$$f(s,t) = \sum_{0\leq m\leq K_1, 0\leq n\leq K_2} c(m,n) B_{(3,3)}(s-m, t-n), \tag{5.70}$$

taken on $\Gamma + K$, where $K = [0, K_1] \times [0, K_2]$ for some positive integers $K_1, K_2 \geq 1$, Γ is either the uniform set Γ_0 or the random set Γ_1 in Figure 5.1, amplitudes of the signal f,

$$c(m,n) \in [-1,1]\backslash[-0.1, 0.1],\ 0 \leq m \leq K_1, 0 \leq n \leq K_2, \tag{5.71}$$

are randomly chosen, and the additive noises $\epsilon(y) \in [-\epsilon, \epsilon], y \in \Gamma + K$, with noise level $\epsilon \geq 0$ are randomly selected. Denote the signal reconstructed by the proposed MAPSET algorithm with phase adjustment threshold value $M_0 = 0.01$, cf. (5.40) with $F_0 = 0.01$, by

$$f_\epsilon(s,t) = \sum_{-2\leq m\leq K_1, -2\leq n\leq K_2} c_\epsilon(m,n) B_{(3,3)}(s-m, t-n). \tag{5.72}$$

Define the maximal amplitude error of the MAPSET algorithm by

$$e(\epsilon) := \min_{\delta\in\{-1,1\}} \max_{-2\leq m\leq K_1, -2\leq n\leq K_2} |c_\epsilon(m,n) - \delta c(m,n)|. \tag{5.73}$$

As the original spline signal f in (5.70) is nonseparable, the conclusions (5.42) and (5.43) guarantee that the reconstruction signal f_ϵ provides a suboptimal approximation, up to a sign, to the original signal f if $\|\Phi^{-1}\|_2 \epsilon$ is much smaller than a multiple of $\sqrt{M_0}$, where M_0 is the phase adjustment threshold value. Our numerical simulations indicate that the MAPSET algorithm saves phases successfully in 100 trials and the maximal amplitude error $e(\epsilon)$ in (5.73) is about $O(\epsilon)$, provided that $\epsilon \leq 2\times 10^{-3}$ for $\Gamma = \Gamma_0$ and $\epsilon \leq 7\times 10^{-4}$ for $\Gamma = \Gamma_1$, where $\sqrt{M_0}/\|\Phi^{-1}\|_2$ are 3.5763×10^{-5} and 3.0307×10^{-6} respectively.

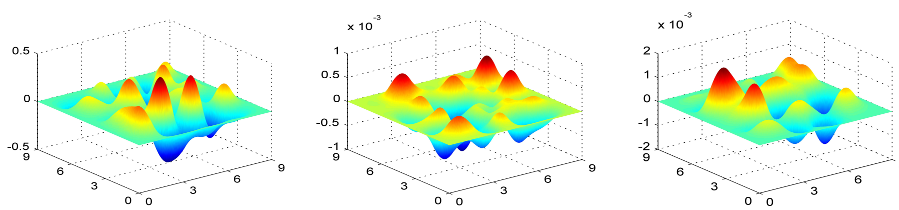

Figure 5.2: Nonseparable spine signal of tensor-product type and reconstruction differences with uniform and randomly distributed sampling set in Figure 5.1 via MAPSET

Presented in Figure 5.2 on the left is a nonseparable spline signal in (5.70) with $K_1 = K_2 = 9$. In the middle and on the right are the difference between the above spline signal f and the signal f_ϵ reconstructed by the MAPSET algorithm with noise level $\epsilon = 10^{-4}$ and sampling set Γ being Γ_0 and Γ_1 in Figure 5.1, respectively. The maximal amplitude errors $e(\epsilon)$ in (5.73) are 0.0014 (middle) and 0.0030 (right), and the reconstruction errors $\min_{\delta\in\{-1,1\}} \|f_\epsilon - \delta f\|_\infty$ are 7.2567×10^{-4} (middle) and 0.0015 (right), respectively.

The signal f_ϵ reconstructed from the MAPSET algorithm may not provide a good approximation, up to a sign, to the original signal f if the noise level ϵ is larger than a multiple of $\sqrt{M_0}/\|\Phi^{-1}\|_2$, cf. (5.41) in Theorem 5.20. Our numerical simulations indicate that the MAPSET algorithm sometimes fails to save the phase of the original signal f when $\epsilon \geq 3 \times 10^{-3}$ for $\Gamma = \Gamma_0$ and $\epsilon \geq 8 \times 10^{-4}$ for $\Gamma = \Gamma_1$.

5.4.2 Nonseparable spline signals of non-tensor product type

Let M_{Ξ_Z} be the box spline function in (5.2) with $\Xi_Z = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$, see [57]. Unlike the spline function $B_{(3,3)}$ of tensor-product type, the shift-invariant space spanned by M_{Ξ_Z} does not have the local complement property on $(0,1)^2$, cf. Section 5.4.1. Set $A_U := \{(s,t) : 0 < s < t < 1\}$ and $A_L := \{(s,t) : 0 < t < s < 1\}$. One may verify that the triangle regions A_U and A_L satisfy (5.18), and the shift-invariant space spanned by M_{Ξ_Z} has local complement property on A_U and on A_L.

For $A = A_U$ and $\phi = M_{\Xi_Z}$, the function $\Phi_{A_U}(s,t)$ in (5.7) is a 5-dimensional vector-valued polynomial about $s^2, (t-s)^2, s, t-s, 1$, and the set K_{A_U} in (5.8) is

$$\{(0,0), (-1,0), (-2,0), (-1,-1), (-2,-1)\}.$$

Hence the space spanned by the outer products of $\Phi_{A_U}(s,t)$ has dimension 13, and we can select a set $\Gamma_{2,U} \subset A_U$ with cardinality 13 to satisfy (5.27), see Figure 5.3. Similarly, for the lower triangle region A_L, a sampling set $\Gamma_{2,L}$ with cardinality 13 can be chosen to satisfy (5.27). For our simulations, we use

$$\Gamma_2 = \Gamma_{2,U} \cup \Gamma_{2,L}$$

as the sampling set contained in $A_U \cup A_L \subset (0,1)^2$, see Figure 5.3.

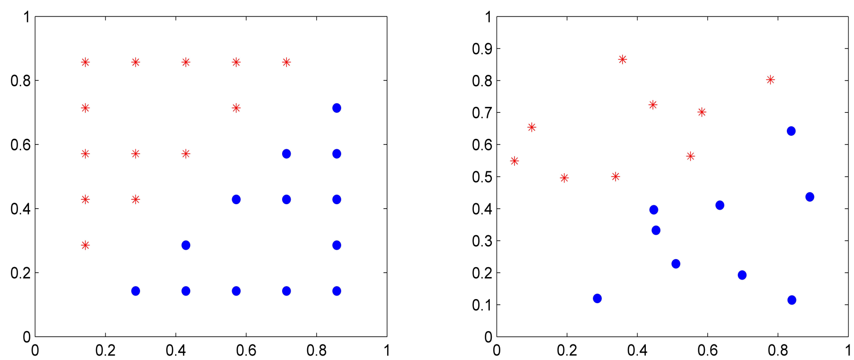

Figure 5.3: Sampling sets for box spline signals of nontensor product type

Plotted in Figure 5.3 on the left are the sampling sets $\Gamma_{2,U} \subset A_U$ (in the red star) and $\Gamma_{2,L} \subset A_L$ (in the blue dot). Plotted on the right are the random sets $\Gamma_{3,U} \subset A_U$ (in the red star) and $\Gamma_{3,L} \subset A_L$ (in the blue dot) that have cardinality 9. The corresponding $\|\Phi^{-1}\|_2$ in (5.39) to the above sets is 87.9420 (left) and 761.2227 (right) respectively.

Recall that $\Phi_{A_U}(s,t)$ is a vector-valued polynomial about $s^2, (t-s)^2, s, t-s$ and 1. Then the matrix $(\Phi_{A_U}(s_i,t_i))_{1\le i\le 5}$ has full rank 5 for almost all $(s_i,t_i) \in A_U, 1 \le i \le 5$, and $(\Phi_{A_U}(s_i,t_i))_{1\le i\le 9}$ is a phase retrieval frame for almost all $(s_i,t_i) \in A_U, 1 \le i \le 9$. So we can use randomly distributed

sets $\Gamma_{3,U} \subset A_U$ and $\Gamma_{3,L} \subset A_L$ with cardinality 9 that satisfy (5.28), see Figure 5.3. Set

$$\Gamma_3 = \Gamma_{3,U} \cup \Gamma_{3,L}.$$

For the above set Γ_3, the corresponding $\|\Phi^{-1}\|_2$ in (5.39) is 761.2227.

In our simulations, the available data $z_\epsilon(y) = |f(y)| + \epsilon(y) \geq 0, y \in \Gamma + K$, are noisy phaseless samples of a spline signal

$$f(s,t) = \sum_{0\leq m\leq K_1, 0\leq n\leq K_2} c(m,n) M_{\Xi_Z}(s-m, t-n), \tag{5.74}$$

taken on $\Gamma + K$, where $K = [0, K_1] \times [0, K_2]$ for some $1 \leq K_1, K_2 \in \mathbb{Z}$, Γ is either Γ_2 or Γ_3 in Figure 5.3, amplitudes of the signal f are as in (5.71), and the additive noises $\epsilon(y) \in [-\epsilon, \epsilon], y \in \Gamma + K$, with noise level $\epsilon \geq 0$ are randomly selected. Denote the signal reconstructed by the proposed MAPSET algorithm with phase adjustment threshold value $M_0 = 0.01$ by

$$f_\epsilon(s,t) = \sum_{-2\leq m\leq K_1, -1\leq n\leq K_2} c_\epsilon(m,n) M_{\Xi_Z}(s-m, t-n). \tag{5.75}$$

As in Section 5.4.1, the reconstruction signal f_ϵ provides an approximation, up to a sign, to the original signal f. Our numerical simulations indicate that the MAPSET algorithm saves phases in 1000 trials and the reconstruction error $e(\epsilon)$ is about $O(\epsilon)$, provided that $\epsilon \leq 8 \times 10^{-3}$ for $\Gamma = \Gamma_2$ and $\epsilon \leq 4 \times 10^{-3}$ for $\Gamma = \Gamma_3$, where $\sqrt{M_0}/\|\Phi^{-1}\|_2$ are 0.0011 and 1.3137×10^{-4} respectively.

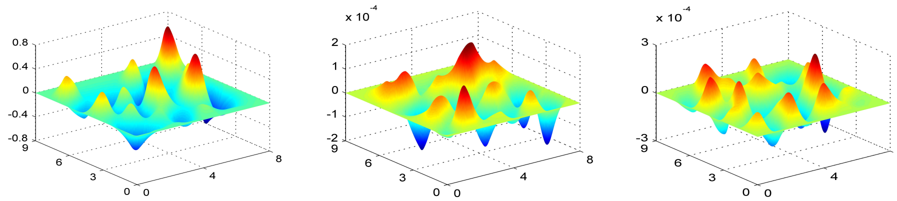

Figure 5.4: A nonseparable spline signal of the form (5.74) and reconstruction differences via MAPSET

Presented in Figure 5.4 on the left is a nonseparable spline signal of the form (5.74), where $K = [0,9]\times[0,8]$, and in the middle and on the right are the difference between the above signal f and the signal f_ϵ reconstructed by the MAPSET algorithm with noise level $\epsilon = 10^{-4}$ and the sampling set Γ being Γ_2 and Γ_3 in Figure 5.3, respectively. The maximal amplitude errors $e(\epsilon)$ in (5.73) are 2.4922×10^{-4} (middle) and 3.8975×10^{-4} (right). The reconstruction errors $\min_{\delta\in\{-1,1\}}\|f_\epsilon-\delta f\|_\infty$ are 1.9660×10^{-4} (middle) and 2.9216×10^{-4} (right).

As in Section 5.4.1, the MAPSET algorithm may not yield a good approximation to the original signal if the noise level ϵ is not sufficient small. Our numerical results indicate that the MAPSET algorithm sometimes fails to save the phase of the original signal f when $\epsilon \geq 9\times 10^{-3}$ for $\Gamma = \Gamma_2$ and $\epsilon \geq 5\times 10^{-3}$ for $\Gamma = \Gamma_3$.

5.5 Local complement property

A linear space V on $\mathbb{R}^d$ is said to be *locally finite-dimensional* if it has finite-dimensional restrictions on any bounded open set. Examples of locally finite-dimensional spaces include the space of polynomials of finite degrees, the shift-invariant space generated by finitely many compactly sup-

ported functions, and their linear subspaces. The reader may refer to [15] and references therein on locally finite-dimensional spaces. In this section, we consider the local complement property for a locally finite-dimensional space, cf. Definition 5.12.

Definition 5.23. Let V be a linear space of real-valued continuous functions on $\mathbb{R}^d$, and $A \subset \mathbb{R}^d$. We say that V has *local complement property on* A if for any $A' \subset A$ there does not exist $f, g \in V$ such that $f, g \not\equiv 0$ on A, $f \equiv 0$ on A' and $g \equiv 0$ on $A \backslash A'$.

In the following theorem, we establish the equivalence between the local complement property on a bounded open set and complement property for ideal sampling functionals on a finite subset, cf. [40].

Theorem 5.24. *Let A be a bounded open set and V be a locally finite-dimensional space of real-valued continuous signals on $\mathbb{R}^d$. Then V has the local complement property on A if and only if there exists a finite set $\Gamma \subset A$ such that for any $\Gamma' \subset \Gamma$ either there does not exist $f \in V$ satisfying*

$$f \not\equiv 0 \ \text{ on } A \ \text{ and } \ f(\gamma') = 0, \ \gamma' \in \Gamma', \tag{5.76}$$

or there does not exist $g \in V$ satisfying

$$g \not\equiv 0 \ \text{ on } A \ \text{ and } \ g(\gamma) = 0, \ \gamma \in \Gamma \backslash \Gamma'. \tag{5.77}$$

The necessity is obvious and the sufficiency follows from the following proposition.

Proposition 5.25. *Let A and V be as in Theorem 5.24. Then there exist a finite set $\Gamma \subset A$ and functions $d_\gamma(x), \gamma \in \Gamma$, such that*

$$|f(x)|^2 = \sum_{\gamma \in \Gamma} d_\gamma(x) |f(\gamma)|^2, \ x \in A \tag{5.78}$$

hold for all $f \in V$.

Proof. Let $g_n, 1 \leq n \leq N$, be a basis of the space $V|_A$, and W be the linear space spanned by symmetric matrices $G(x) := \big(g_n(x)g_{n'}(x)\big)_{1\leq n,n'\leq N}$, $x \in A$. Then there exists a finite set $\Gamma \subset A$ such that $G(\gamma), \gamma \in \Gamma$, is a basis (or a spanning set) for the space W. With the above set Γ, we can follow the proof of Theorem 5.10 to prove (5.78). $\square$

Let $g_n, 1 \leq n \leq N$, be a basis of the space $V|_A$, and Γ be as in the proof of Proposition 5.25. By Theorem 5.24 and [20, Theorem 2.8], we have the following criterion that can be used to verify the local complement property on a bounded open set A in finite steps.

Theorem 5.26. *The linear space V has the local complement property on A if and only if for any $\Gamma' \subset \Gamma$, either $(g_n(\gamma'))_{1\leq n\leq N}, \gamma' \in \Gamma'$ form a frame for $\mathbb{R}^N$ or $(g_n(\gamma))_{1\leq n\leq N}, \gamma \in \Gamma\backslash\Gamma'$ form a frame for $\mathbb{R}^N$.*

The local complement property for different open sets can be equivalent. Following the argument used in the proof of Theorem 5.24, we have

Proposition 5.27. *Let A be a bounded open set and V be a locally finite-dimensional space with the local complement property on A. If B is a bounded open subset of A such that signals g and f satisfying $|g(x)| = |f(x)|$ on B have the same magnitude measurements on A, then V has local complement property on B.*

The conclusion in the above proposition is not true in general. For instance, the shift-invariant space $V(\phi_0)$ in Example 5.6 has the local complement property on $(0, 1/2)$, but not on its supset $(0, 1)$.

A linear space may have the local complement property on a bounded open A, but not on some of its open subsets. For instance, one may verify that $V(\phi_1)$ has the local complement property on

$(0,1)$ and on $(-1/2,1/2)$, but not on their intersection $(0,1/2)$, where $\phi_1=\phi_0(2\cdot)$ and ϕ_0 is given in Example 5.6.

We finish the appendix with a proposition about local linear independence and local complement property.

Proposition 5.28. *Let ϕ be a compactly supported continuous function with local linear independence on any open set. Then there exist $A_m, 1\le m\le M$, such that* (5.18) *holds and $V(\phi)$ has the local complement property on $A_m, 1\le m\le M$.*

Proof. Let $S_k, k\in\mathbb{Z}^d$, be as in (5.6). For a set $T\subset\mathbb{Z}^d$, define $S_T=\cap_{k\in T}S_k$. We say that T is maximal if $S_T\ne\emptyset$ and $S_{T'}=\emptyset$ for all $T'\supsetneqq T$. From the definition, there are finitely many maximal sets $T_1,\dots,T_M$, and denote the corresponding sets by $A_m:=S_{T_m}, 1\le m\le M$.

Clearly (5.18) holds for the above selected open sets as

$$\cup_{m=1}^M T_m=\{k\in\mathbb{Z}^d:\ S_k\ne\emptyset\}.$$

Then it remains to prove that $V(\phi)$ has local complement property on $A_m, 1\le m\le M$. Assume that $f,g\in V(\phi)$ satisfy $|f(x)|=|g(x)|$ for all $x\in A_m$, which implies that $(f+g)(x)(f-g)(x)=0$ for all $x\in A_m$. Write $f+g=\sum_{k\in\mathbb{Z}^d}c(k)\phi(\cdot-k)$ and $f-g=\sum_{k\in\mathbb{Z}^d}d(k)\phi(\cdot-k)$. Set $B_1=\{x\in A_m:\ (f+g)(x)\ne 0\}$ and $B_2=\{x\in A_m:\ (f-g)(x)\ne 0\}$. Then either $f-g=0$ on B_1, or $f+g=0$ on B_2, or $f-g=f+g=0$ on A_m. Hence either $c(k)=d(k)$ for all $k\in T_m$ or $c(k)=-d(k)$ on $k\in T_m$ by the local independence on B_1, or B_2 or A_m. Therefore either $f=g$ on A_m, or $f=-g$ on A_m, or $f=g=0$ on A_m. This completes the proof. $\square$

LIST OF REFERENCES

[1] B. Adcock, M. Gataric and A. C. Hansen, On stable reconstruction from nonuniform Fourier measurements, *SIAM J. Imaging Sci.*, **7**(2014), 1690–1723.

[2] B. Adcock, M. Gataric and A. C. Hansen, Weighted frames of exponentials and stable recovery of multidimensional functions from nonuniform Fourier samples, *Appl. Comput. Harmon. Anal.*, **42**(2017), 508–535.

[3] B. Adcock, A. C. Hansen and C. Poon, Beyond consistent reconstructions: optimality and sharp bounds for generalized sampling, and application to the uniform resampling problem, *SIAM J. Math. Anal.*, **45**(2013), 3132–3167.

[4] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam and E. Cayirci, Wireless sensor networks: a survey, *Comput. Netw.*, **38**(2002), 393–422.

[5] R. Alaifari, I. Daubechies, P. Grohs and G. Thakur, Reconstructing real-valued functions from unsigned coefficients with respect to wavelet and other frames, *J. Fourier Anal. Appl.*, 2016, accepted

[6] R. Alaifari, I. Daubechies, P. Grohs and R. Yin, Stable phase retrieval in infinite dimensions, Arxiv preprint, arXiv:1609.00034

[7] R. Alaifari and P. Grohs, Phase retrieval in the general setting of continuous frames for Banach spaces, *SIAM J. Math. Anal.*, **49**(2017), 1895-1911.

[8] B. Alexeev, J. Cahill and D. G. Mixon, Full spark frames, *J. Fourier Anal. Appl.*, **18**(2012), 1167–1194.

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

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

[11] A. Aldroubi and H. Feichtinger, Exact iterative reconstruction algorithm for multivariate irregularly sampled functions in spline-like spaces: the L_p theory. *Proc. Amer. Math. Soc.*, **126**(1998), 2677–2686.

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

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

[14] A. Aldroubi, C. Leonetti and Q. Sun, Error analysis of frame reconstruction from noisy samples, *IEEE Trans. Signal Proc.*, **56**(2008), 2311–2325.

[15] A. Aldroubi and Q. Sun, Locally finite dimensional shift-invariant spaces in $\mathbb{R}^d$, *Proc. Amer. Math. Soc.*, **130**(2002), 2641–2654.

[16] A. Aldroubi, Q. Sun and W.-S. Tang, Non-uniform average sampling and reconstruction in multiply generated shift-invariant spaces, *Constr. Approx.*, **20**(2004), 173–189.

[17] A. Aldroubi, Q. Sun and W.-S. Tang, Convolution, average sampling and a Calderon resolution of the identity for shift-invariant spaces, *J. Fourier Anal. Appl.*, **11**(2005), 215–244.

[18] R. Balan, The noncommutative Wiener lemma, linear independence, and special properties of the algebra of time-frequency shift operators, *Trans. Amer. Math. Soc.*, **360**(2008), 3921–3941.

[19] R. Balan, B. G. Bodmann, P. G. Casazza and D. Edidin, Painless reconstruction from magnitudes of frame coefficents, *J. Fourier Anal. Appl.*, **15**(2009), 488–501.

[20] R. Balan, P. G. Casazza and D. Edidin, On signal reconstruction without phase, *Appl. Comp. Harm. Anal.*, **20**(2006), 345–356.

[21] R. Balan and Y. Wang, Invertibity and robustness of phaseless reconstruction, *Appl. Comput. Harmon. Anal.* **38**(2015), 469-488.

[22] R. Balan and D. Zhou, Phase retrieval using Lipschitz continuous maps, *Linear Algebra Appl.*, **496**(2016), 152–181.

[23] A. S. Bandeira, J. Cahill, D. G. Mixon, and A. A. Nelson, Saving phase: injectivity and stability for phase retrieval, *Appl. Comput. Harmon. Anal.*, **37**(2014), 106–125.

[24] A. G. Baskakov, Wiener's theorem and asymptotic estimates for elements of inverse matrices, *Funktsional Anal i Prilozhen*, **24**(1990), 64–65; translation in *Funct. Anal. Appl.*, **24**(1990), 222–224.

[25] A. Ben-Artzi and A. Ron, On the integer translates of a compactly supported function: dual bases and linear projectors, *SIAM J. Math. Anal.*, **21**(1990), 1550–1562.

[26] P. Berger and K. Gröchenig, Sampling and reconstruction in different subspaces by using oblique projections, Arxiv preprint, arXiv:1312.1717

[27] D. P. Bertsekas and J. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*, Prentice-Hall, Englewood Cliffs, NJ, 1989.

[28] N. Bi, M. Z. Nashed and Q. Sun, Reconstructing signals with finite rate of innovation from noisy samples, *Acta Appl. Math.*, **107**(2009), 339–372.

[29] T. Blu, P. L. Dragotti, M. Vetterli, P. Marziliano and L. Coulot, Sparse sampling of signal innovations, *IEEE Signal Process. Mag.*, **25**(2008), 31–40.

[30] M. Bownik, The structure of shift-invariant subspaces of $L^2(\mathbb{R}^d)$, *J. Funct. Anal.*, **177**(2000), 282–309.

[31] A. J. Cadzow, A finite algorithm for the minimum ℓ_∞ solution to a system of consistent linear equations, *SIAM J. Numer. Anal.*, **10**(1973), 607–617.

[32] A. J. Cadzow, An efficient algorithmic procedure for obtaining a minimum ℓ_∞ solution to a system of consistent linear equations, *SIAM J. Numer. Anal.*, **11**(1974), 1151–1165.

[33] J. Cahill, P. G. Casazza and I. Daubechies, Phase retrieval in infinite-dimensional Hilbert spaces, *Trans. Amer. Math. Soc. Ser. B*, **3**(2016), 63–76.

[34] E. J. Candes, Y. C. Eldar, T. Strohmer and V. Voroninski, Phase retrieval via matrix completion, *SIAM J. Imaging Sci.*, **6**(2013), 199–225.

[35] E. J. Candes, X. Li, and M. Soltanolkotabi, Phase retrieval via Wirtinger flow: theory and algorithms, *IEEE Trans. Inf. Theory*, **61**(2015), 1985–2007.

[36] E. J. Candes, J. Romberg and T. Tao, Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information, *IEEE Trans. Inf. Theory*, **52**(2006), 489–509.

[37] E. J. Candes, T. Strohmer, and V. Voroninski, Phaselift: exact and stable signal recovery from magnitude measurements via convex programming, *Comm. Pure Appl. Math.*, **66**(2013), 1241–1274.

[38] P. G. Casazza, D. Ghoreishi, S. Jose and J. C. Tremain, Norm retrieval and phase retrieval by projections, Arxiv preprint, arXiv:1701.08014

[39] P. G. Casazza, G. Kutyniok and S. Li, Fusion frames and distributed processing, *Appl. Comput. Harmon. Anal.*, **25**(2008), 114–132.

[40] Y. Chen, C. Cheng, Q. Sun and H. Wang, Phase retrieval of real-valued signals in a shift-invariant space, Arxiv preprint, arXiv:1603.01592

[41] C. Cheng, Y. Jiang and Q. Sun, Spatially distributed sampling and reconstruction, Arxiv preprint, arXiv:1511.08541

[42] C. Cheng, Y. Jiang and Q. Sun, Sampling and Galerkin reconstruction in reproducing kernel spaces, *Appl. Comput. Harmon. Anal.*, **41**(2016), 638–659.

[43] C. Cheng, Y. Jiang and Q. Sun, Spatially distributed sampling and reconstruction of high-dimensional signals, In *2015 International Conference on Sampling Theory and Applications (SampTA)*, IEEE, 2015, pp. 453–457.

[44] C. Chong and S. Kumar, Sensor networks: evolution, opportunities, and challenges, *Proc. IEEE*, **91**(2003), 1247–1256.

[45] A. Cohen and Q. Sun, An arithmetic characterization of the conjugate quadrature filters associated to orthonormal wavelet bases, *SIAM J. Math. Anal.*, **24**(1993), 1355–1360.

[46] J. G. Christensen, Sampling in reproducing kernel Banach spacecs in Lie group, *J. Approx. Theory*, **164**(2012), 179–203.

[47] O. Christensen and T. Strohmer, The finite section method and problems in frame theory, *J. Approx. Th.*, **133**(2005), 221–237.

[48] F. R. K. Chung, *Spectral Graph Theory*, American Mathematical Society, 1997.

[49] R. Coifman and G. Weiss, *Analyse Harmonique Non-Commutative Sur Certains Espaces Homogeneous*, vol. 242 of Lecture Notes in Mathematics, Springer, New York, NY, USA, 1971.

[50] T. B. Curtin, J. G. Bellingham, J. Catipovic and D. Webb, Autonomous oceanographic sampling networks, *Oceanography*, **6**(1993), 86–94.

[51] W. Czaja, G. Kutyniok and D. Speegle, Beurling dimension of Gabor pseudoframes for affine subspaces, *J. Fourier Anal. Appl.*, **14**(2008). 514–537.

[52] W. Dahmen and C. A. Micchelli, On the local linear independence of translates of a box spline, *Studia Math.*, **82**(1985), 243–263.

[53] W. Dahmen and C. A. Micchelli, Translates of multivariate splines, *Linear Algebra Appl.*, **52**(1982), 217–234.

[54] I. Daubechies, *Ten Lectures on Wavelets*, CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM 1992.

[55] C. de Boor, R. A. DeVore, and A. Ron, The structure of finitely generated shift-invariant spaces in $L^2(\mathbb{R}^d)$, *J. Funct. Anal.*, **119**(1994), 37–78.

[56] C. de Boor and K. Höllig, B-splines from parallelepipeds, *J. Anal. Math.*, **62**(1983), 99–115.

[57] C. de Boor, K. Höllig and S.D. Riemenschneider, *Box Splines*, Springer-Verlag, Berlin, 1993.

[58] D. Deng and Y. Han, *Harmonic Analysis on Spaces of Homogeneous Type*, Springer, 2008.

[59] D. L. Donoho, Superresolution via sparsity constraints, *SIAM J. Math. Anal.*, **23**(1992), 1309–1331.

[60] D. L. Donoho, Compressed sensing, *IEEE Trans. Inf. Theory*, **52**(2006), 1289–1306.

[61] D. L. Donoho and M. Elad, Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ^1 minimization, *Proc. Natl. Acad. Sci. U.S.A*, **100**(2003), 2197–2202.

[62] P. L. Dragotti, M. Vetterli and T. Blu, Sampling moments and reconstructing signals of finite rate of innovation: Shannon meets Strang-Fix, *IEEE Trans. Signal Process.*, **55**(2007), 1741–1757.

[63] D. E. Dutkay, D. Han, Q. Sun and E. Weber, On the Beurling dimension of exponential frames, *Adv. Math.*, **226**(2011), 285–297.

[64] T. G. Dvorkind, Y. C. Eldar and E. Matusiak, Nonlinear and nonideal sampling: theory and methods, *IEEE Trans. Signal Process.*, **56**(2008), 5874–5890.

[65] Y. C. Eldar, and S. Mendelson, Phase retrieval: stability and recovery guarantees, *Appl. Comput. Harmon. Anal.*, **36**(2014), 473-494.

[66] Y. C. Eldar and M. Unser, Nonideal sampling and interpolation from noisy observations in shift-invariant spaces, *IEEE Trans. Signal Process.*, **54**(2006), 2636–2651.

[67] Y. C. Eldar and T. Werther, General framework for consistent sampling in Hilbert spaces, *Int. J. Wavelets Multiresolution Inf. Process.*, **3**(2005), 347–359.

[68] B. Farrell and T. Strohmer, Inverse-closedness of a Banach algebra of integral operators on the Heisenberg group, *J. Operator Theory*, **64**(2010), 189–205.

[69] H. G. Feichtinger and K. Gröchenig, Iterative reconstruction of multivariate band-limited functions from irregular sampling values, *SIAM J. Math. Anal.*, **231**(1992), 244–261.

[70] H. G. Feichtinger, K. Gröchenig and T. Strohmer, Efficient numerical methods in non-uniform sampling theory, *Numer. Math.*, **69**(1995), 423–440.

[71] H. G. Feichtinger, J. C. Principe, J. L. Romero, A. A. Singh, and G. A. Alexander, Approximate reconstruction of bandlimited functions for the integrate and fire sampler, *Adv. Comput. Math.*, **36**(2012), 67–78.

[72] J. R. Fienup, Reconstruction of an object from the modulus of its Fourier transform, *Opt. Lett.*, **3**(1978), 27–29.

[73] J. R. Fienup, Phase retrieval algorithms: a comparison, *Applied Optics*, **21**(1982), 2758–2769.

[74] B. Gao, Q. Sun, Y. Wang and Z. Xu, Phase retrieval from the magnitudes of affine linear measurements, arXiv:1608.06117

[75] A. G. Garcia and A. Portal, Sampling in reproducing kernel Banach spaces, *Mediterr. J. Math.*, **103**(2013), 1401–1417.

[76] R. W. Gerchberg and W. O. Saxton, A practical algorithm for the determination of phase from image and diffraction plane pictures, *Optik*, **35**(1972), 237–246.

[77] I. Gohberg, M. A. Kaashoek and H. J. Woerdeman, The band method for positive and strictly contractive extension problems: an alternative version and new applications, *Integral Equation Oper. Theory*, **12**(1989), 343–382.

[78] D. Gontier and M. Vetterli, Sampling based on timing: time encoding machines on shift-invariant subspaces, *Appl. Comput. Harmon. Anal.*, **36**(2014), 63–78.

[79] T. N. T. Goodman and C. A. Micchelli, On refinement equations determined by Polya frequency sequences, *SIAM J. Math. Anal.*, **23**(1992), 766–784.

[80] T. N. T. Goodman and Q. Sun, Total positivity and refinable functions with general dilation, *Appl. Comput. Harmon. Anal.*, **16**(2004), 69–89.

[81] K. Gröchenig, Reconstructing algorithms in irregular sampling, *Math. Comput.*, **59**(1992), 181–194.

[82] K. Gröchenig, Time-frequency analysis of Sjöstrand class, *Rev. Mat. Iberoam.*, **22**(2006), 703–724.

[83] K. Gröchenig, Wiener's lemma: theme and variations, an introduction to spectral invariance and its applications, In: *Four Short Courses on Harmonic Analysis: Wavelets, Frames, Time-Frequency Methods, and Applications to Signal and Image Analysis*, editor by P. Massopust and B. Forster, Birkhauser, Boston, 2010.

[84] K. Gröchenig and A. Klotz, Noncommutative approximation: inverse-closed subalgebras and off-diagonal decay of matrices, *Constr. Approx.*, **32**(2010), 429–466.

[85] K. Gröchenig and M. Leinert, Symmetry and inverse-closedness of matrix algebras and functional calculus for infinite matrices, *Trans. Amer. Math. Soc.*, **358**(2006), 2695–2711.

[86] K. Gröchenig and H. Schwab, Fast local reconstruction methods for nonuniform sampling in shift-invariant spaces, *SIAM J. Matrix Anal. Appl.*, **24**(2003), 899–913.

[87] D. Han, M. Z. Nashed and Q. Sun, Sampling expansions in reproducing kernel Hilbert and Banach spaces, *Numer. Funct. Anal. Optim.*, **30**(2009), 971–987.

[88] P. Hand and V. Voroninski, Corruption robust phase retrieval via linear programming, Arxiv preprint, arXiv:1612.03547

[89] M. H. Hayes, J. S. Lim, and A. V. Oppenheim, Signal reconstruction from phase or magnitude, *IEEE Trans. Acoust. Speech Signal Process.*, **28**(1980), 672–680.

[90] J. A. Hogan and J. D. Lakey, *Duration and Bandwidth Limiting: Prolate Functions, Sampling, and Applications*, Birkhäuser, 2012.

[91] N. E. Hurt, *Phase Retrieval and Zero Crossings: Mathematical Methods in Image Reconstruction*, Springer, 2001.

[92] M. A. Iwen, B. Preskitt, R. Saab and A. Viswanathan, Phase retrieval from local measurements: improved robustness via eigenvector-based angular synchronization, Arxiv preprint, arXiv:1612.01182

[93] S. Jaffard, Properiétés des matrices bien localisées prés de leur diagonale et quelques applications, *Ann. Inst. Henri Poincaré D*, 7(1990), 461-476.

[94] K. Jaganathan, Y. C. Eldar and B. Hassibi, Phase retrieval: an overview of recent developments, In *Optical Compressive Imaging*, edited by A. Stern, CRC Press, 2016, pp. 261–296.

[95] K. Jaganathan, S. Oymak, and B. Hassibi, Sparse phase retrieval: Convex algorithms and limitations, In *Proceeding of the 2013 IEEE International Symposium on Information Theory*, 2013, pp. 1022–1026.

[96] P. Jaming, A. Karoui, R. Kerman and S. Spektor, Approximation of almost time and band limited functions I: Hermite expansion, *J. Approx. Theory*, **212**(2016), 41–65.

[97] R.-Q. Jia, Local linear independence of the translates of a box spline, *Constr. Approx.*, **1**(1985), 175–182.

[98] R.-Q. Jia and C. A. Micchelli, On linear independence of integer translates of a finite number of functions, *Proc. Edinburgh Math. Soc.*, **36**(1992), 69–75.

[99] R.-Q. Jia and J. Wang, Stability and linear independence associated with wavelet decompositions, *Proc. Amer. Math. Soc.* **117**(1993), 1115–1124.

[100] D. Kincaid and W. Cheney, *Numerical Analysis: Mathematics of Scientific Computing,* Brooks/Cole, Third edition, 2002.

[101] M. Klibanov, P. Sacks and A. Tikhonravov, The phase retrieval problem, *Inverse problems*, **11**(1995), 1–28.

[102] J. Koshal, A. Nedic and U. V. Shanbhag, Multiuser optimization: distributed algorithm and error analysis, *SIAM J. Optim.*, **21**(2011), 1046–1081.

[103] I. Krishtal, Wiener's lemma: pictures at exhibition, *Rev. Un. Mat. Argentina*, **52**(2011), 61–79.

[104] A. A. Lazar and L. T. Toth, Perfect recovery and sensitivity analysis of time encoded bandlimited signals, *IEEE Trans. Circuits System*, **51**(2004), 2060–2073.

[105] P. G. Lemarié, Fonctions á support compact dans les analyses multirésolutions, *Rev. Mat. Iberoamericana*, **7**(1991), 157–182.

[106] N. E. Leonard, D. A. Paley, F. Lekien, R. Sepulchre, D. M. Fratantoni and R. E. Davis, Collective motion, sensor networks, and ocean sampling, *Proc. IEEE*, **95**(2007), 48–74.

[107] L. Li, C. Cheng, D. Han, Q. Sun and G. Shi, Phase retrieval from multiple-window short-time Fourier measurements, *IEEE Signal Process. Lett.*, **24**(2017), 372–376.

[108] C. G. Lopes and A. H. Sayed, Incremental adaptive strategies over distributed networks, *IEEE Trans. Signal Process.*, **55**(2007), 4064–4077.

[109] S. Mallat, *A Wavelet Tour of Signal Processing: The Sparse Way*, Academic Press, 2009.

[110] S. Mallat and I. Waldspurger, Phase retrieval for the Cauchy wavelet transform, *J. Fourier Anal. Appl.*, **21**(2014), 1–59.

[111] I. Maravic and M. Vetterli, Sampling and reconstruction of signals with finite rate of innovation in the presence of noise, *IEEE Trans. Signal Process.*, **53**(2005), 2788–2805.

[112] P. Marziliano, M. Vetterli and T. Blu, Sampling and exact reconstruction of bandlimited signals with additive shot noise, *IEEE Trans. Inf. Theory*, **52**(2006), 2230–2233.

[113] Y. Meyer, Ondelettes sur l'intervalle, *Rev. Mat. Iberoamericana*, **7**(1991), 115–133.

[114] R. P. Millane, Phase retrieval in crystallography and optics, *J. Opt. Soc. Am. A*, **7**(1990), 394–411.

[115] M. Mishali, Y. C. Eldar and A. J. Elron, Xampling: signal acquisition and processing in union of subspaces, *IEEE Trans. Signal Process.*, **59**(2011), 4719–4734.

[116] M. Z. Nashed and Q. Sun, Sampling and reconstruction of signals in a reproducing kernel subspace of $L^p(R^d)$, *J. Funct. Anal.*, **258**(2010), 2422–2452.

[117] M. Z. Nashed and G. G. Walter, General sampling theorems for functions in reproducing kernel Hilbert spaces, *Math. Control Signals Systems*, **4**(1991), 363–390.

[118] A. Nedic and A. Olshevsky, Distributed optimization over time-varying directed graphs, *IEEE Trans. Auto. Control.*, **60**(2015), 601–615.

[119] P. Netrapalli, P. Jain, and S. Sanghavi, Phase retrieval using alternating minimization, *IEEE Trans. Signal Proc.*, **63**(2015), 4814–4826.

[120] H. Ohlsson, A. Yang, R. Dong, and S. Sastry, Compressive phase retrieval from squared output measurements via semidefinite programming, In *Proceedings of the 16th IFAC Symposium on System Identification*, Vol. 45, 2012, pp. 89–94.

[121] H. Pan, T. Blu and P. L. Dragotti, Sampling curves with finite rate of innovation, *IEEE Trans. Signal Process.*, **62**(2014), 458–471.

[122] M. Pawlak, E. Rafajlowicz and A. Krzyzak, Postfiltering versus prefiltering for signal recovery from noisy samples, *IEEE Trans. Inf. Theory*, **49**(2003), 3195–3212.

[123] R. Pedarsani, D. Yin, K. Lee, K. Ramchandran, PhaseCode: fast and efficient compressive phase retrieval based on sparse-graph codes, *IEEE Trans. Inf. Theory*, **63**(2017), 3663–3691.

[124] I. Pesenson, Sampling in Paley-Wiener spaces on combinatorial graphs, *Trans. Amer. Math. Soc.*, **360**(2008), 5603–5627.

[125] V. Pohl, F. Yang and H. Boche, Phase retrieval from low-rate samples, *Sampl. Theory Signal Image Process.* **13**(2014), 71–99.

[126] V. Pohl, F. Yang and H. Boche, Phaseless signal recovery in infinite dimensional spaces using structured modulations, *J. Fourier Anal. Appl.*, **20**(2014), 1212–1233.

[127] T. Qiu, P. Babu and D. P. Palomar, PRIME: phase retrieval via majorization-minimization, *IEEE Trans Sig. Process.*, **64**(2016), 5174–5186.

[128] L. Rabiner and B.-H. Juang, *Fundamentals of Speech Recognition*, Prentice Hall Inc., Englewood Cliffs, 1993.

[129] A. Ron, A necessary and sufficient condition for the linear independence of the integer translates of a compactly supported distribution, *Const. Approx.*, **5**(1989), 297–308.

[130] A. Sandryhaila and J. M. F. Moura, Discrete signal processing on graphs, *IEEE Trans. Signal Process.*, **61**(2013), 1644–1656.

[131] C. E. Shannon, Communication in the presence of noise, *Proc. IRE*, **37**(1949), 10–21.

[132] Y. Shechtman, Y. C. Eldar, O. Cohen, H. N. Chapman, J. Miao and M. Segev, Phase retrieval with application to optical imaging: a contemporary overview, *IEEE Signal Proc. Mag.*, **32**(2015), 87–109.

[133] B. A. Shenoy, S. Mulleti and C. S. Seelamantula, Exact phase retrieval in principal shift-invariant spaces, *IEEE Trans. Signal Proc.*, **64**(2016), 406–416.

[134] C. E. Shin and Q. Sun, Stability of localized operators, *J. Funct. Anal.*, **256**(2009), 2417–2439.

[135] P. Shukla and P. L. Dragotti, Sampling schemes for multidimensional signals with finite rate of innovation, *IEEE Trans. Signal Process.*, **55**(2007), 3670–3686.

[136] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega and P. Vandergheynst, The emerging field of signal processing on graphs, *IEEE Signal Process. Mag.*, **30**(2013), 83–98.

[137] J. Sjöstrand, Wiener type algebra of pseudodifferential operators, Cent. Math., Ecole Polytechnique, Palaiseau France, Seminaire 1994, 1995, December 1994.

[138] S. Smale and D. X. Zhou, Shannon sampling and function reconstruction from point values, *Bull. Amer. Math. Soc.*, **41**(2004), 279–305.

[139] Q. Sun, A note on the integer translates of a compactly supported distribution on $\mathbb{R}$, *Archiv Math.*, **60**(1993), 359–363.

[140] Q. Sun, Wiener's lemma for infinite matrices with polynomial off-diagonal decay, *C. Acad. Sci. Paris Ser I*, **340**(2005), 567–570.

[141] Q. Sun, Non-uniform average sampling and reconstruction for signals with finite rate of innovations, *SIAM J. Math. Anal.*, **38**(2006), 1389–1422.

[142] Q. Sun, Wiener's lemma for infinite matrices, *Trans. Amer. Math. Soc.*, **359**(2007), 3099–3123.

[143] Q. Sun, Frames in spaces with finite rate of innovation, *Adv. Comput. Math.*, **28**(2008), 301–329.

[144] Q. Sun, Wiener's lemma for localized integral operators, *Appl. Comput. Harmon. Anal.*, **25**(2008), 148–167.

[145] Q. Sun, Local reconstruction for sampling in shift-invariant space, *Adv. Comput. Math.*, **32**(2010), 335–352.

[146] Q. Sun, Stability criterion for convolution-dominated infinite matrices, *Proc. Amer. Math. Soc.*, **138**(2010), 3933–3943.

[147] Q. Sun, Wiener's lemma for infinite matrices II, *Constr. Approx.*, **34**(2011), 209–235.

[148] Q. Sun, Localized nonlinear functional equations and two sampling problems in signal processing, *Adv. Comput. Math.*, **40**(2014), 415–458.

[149] Q. Sun and W.-S. Tang, Nonlinear frames and sparse reconstructions in Banach spaces, *J. Fourier Anal. Appl.*, accepted.

[150] Q. Sun and J. Xian, Rate of innovation for (non-)periodic signals and optimal lower stability bound for filtering, *J. Fourier Anal. Appl.*, **20**(2014), 119–134.

[151] W. Sun and X. Zhou, Reconstruction of bandlimited signals from local averages, *IEEE Trans. Inf. Theory*, **48**(2002), 2955–2963.

[152] W. Sun and X. Zhou, Characterization of local sampling sequences for spline subspaces, *Adv. Comput. Math.*, **30**(2009), 153–175.

[153] W.-S. Tang, Oblique projections, biorthogonal Riesz bases and multiwavelets in Hilbert spaces, *Proc. Amer. Math. Soc.*, **128**(1999), 463–473.

[154] G. Thakur, Reconstruction of bandlimited functions from unsigned samples, *J. Fourier Anal. Appl.*, **17**(2011), 720–732.

[155] J. T. Tyson, Metric and geometric quasiconformality in Ahlfors regular Loewner spaces, *Conform. Geom. Dyn.*, **5**(2001), 21–73.

[156] J. Unnikrishnan and M. Vetterli, Sampling and reconstruction of spatial fields using mobile sensors, *IEEE Trans. Signal Process.*, **61**(2013), 2328–2340.

[157] M. Unser, Splines: A perfect fit for signal and image processing *IEEE Signal Proc. Magazine* **16**(1999), 22–38.

[158] M. Unser, Sampling – 50 years after Shannon, *Proc. IEEE*, **88**(2000), 569–587.

[159] M. Vetterli, P. Marziliano and T. Blu, Sampling signals with finite rate of innovation, *IEEE Trans. Signal Proc.*, **50**(2002), 1417–1428.

[160] G. Wahba, *Spline Models for Observational Data*, SIAM, 1990.

[161] Y. Wang and Z. Xu, Phase retrieval for sparse signals, *Appl. Comput. Harmon. Anal.*, **37**(2014), 531–544.

[162] Z. Wang and A. C. Bovik, Mean squared error: love it or leave it?- A new look at signal fidelity measures, *IEEE Signal Process. Mag.*, **98**(2009), 98–117.

[163] J. M. Whittaker, *Interpolating Function Theory*, Cambridge University Press, London, 1935.

[164] F. Yang, V. Pohl and H. Boche, Phaseless signal recovery in infinite dimensional spaces using structured modulations, *J. Fourier Anal. Appl.*, **20**(2014), 1212–1233.

[165] J. Yick, B. Mukherjee and D. Ghosal, Wireless sensor network survey, *Comput. Netw.*, **52**(2008), 2292–2330.

[166] F. Zhao and L. Guibas, *Wireless Sensor Networks: An Information Processing Approach*, Morgan Kaufmann, 2004.