Appl. Comput. Harmon. Anal. 47 (2019) 109–148

Contents lists available at ScienceDirect

Applied and Computational Harmonic Analysis

www.elsevier.com/locate/acha

Spatially distributed sampling and reconstruction ☆

Cheng Cheng [a], Yingchun Jiang [b,*], Qiyu Sun [c]

[a] *Department of Mathematics, Duke University, Durham, NC 32816, USA*
[b] *School of Mathematics and Computational Science, Guilin University of Electronic Technology, Guilin, Guangxi 541004, China*
[c] *Department of Mathematics, University of Central Florida, Orlando, FL 32816, USA*

A R T I C L E I N F O

Article history:
Received 21 October 2015
Received in revised form 2 June 2017
Accepted 30 July 2017
Available online 14 August 2017
Communicated by Karlheinz Gröchenig

Keywords:
Spatially distributed network
Distributed sampling and reconstruction system
Signal on a graph
Distributed algorithm
Finite rate of innovation
Beurling dimension
Localized matrix
Inverse-closed subalgebra

A B S T R A C T

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. In this paper, 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. A fundamental problem in sampling theory is the robustness of signal reconstruction in the presence of sampling noises. For a distributed sampling and reconstruction system, the robustness could be reduced to the stability of its sensing matrix. In this paper, 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. In this paper, 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.

1. Introduction

Spatially distributed networks (SDNs) have been widely used in (underwater) multivehicle and multi-robot networks, wireless sensor networks, smart power grids, etc. ([2,19,23,40,45,74,75]). Comparing with traditional centralized networks that have a powerful central processor and reliable communication between agents and the central processor, an SDN could give unprecedented capabilities especially when creating a data exchange network requires significant efforts (due to physical barriers such as interference), or when

☆ This project is partially supported by the National Natural Science Foundation of China (Nos. 11661024 and 11671107), and the National Science Foundation (DMS-1412413).

* Corresponding author.
E-mail addresses: cheng87@math.duke.edu (C. Cheng), guilinjiang@126.com (Y. Jiang), qiyu.sun@ucf.edu (Q. Sun).

http://dx.doi.org/10.1016/j.acha.2017.07.007
1063-5203/

establishing a centralized processor presents the daunting challenge of processing all the information (such as big-data problems). In this paper, we describe the topology of an SDN by an undirected (in)finite graph

$$\mathcal{G} := (G, S) \tag{1.1}$$

of large scale, where a vertex in G represents an agent and an edge in S between two vertices means that a direct communication link exists.

To consider signal sampling and reconstruction on an SDN, we equip a sensing device at every agent $\lambda \in G$, which has with limited sensing, data processing, and communication capabilities. In this paper, we assume that the sampling procedure

$$f \longmapsto (y(\lambda))_{\lambda \in G}$$

on signals f of interest is linear. This implies that the sampling data

$$y(\lambda) := \langle f, \psi_\lambda \rangle \tag{1.2}$$

acquired by the agent $\lambda \in G$ is a linear functional on f, where the functional ψ_λ reflects the characteristic of the sensing device of the agent $\lambda \in G$. For spatial signals on $\mathbb{R}^d$, the above sampling procedure is also known as average sampling or (non)ideal sampling [8,17,32,60].

Fundamental signal reconstruction problems are whether and how the signal f of interest can be recovered from its sampling data $y(\lambda),\ \lambda \in G$. The signal reconstruction problem is ill-posed inherently. For its well-posedness, the signal f is usually assumed to have additional properties, such as band-limitedness, finite rate of innovation, smoothness, and sparse expansion in a dictionary ([7,15,26–28,71,72]). The sampling and reconstruction problem is well studied for spatial signals on $\mathbb{R}^d$. The reader may refer to [4,7,8,32,49,58,60, 70–72] and references therein for various sampling procedure and reconstruction scenarios. In this paper, we consider spatial signals

$$f = \sum_{i \in V} c(i)\varphi_i(\cdot) \tag{1.3}$$

being a bounded superposition of generators $\varphi_i, i \in V$. Define

$$V_p(\Phi) := \Big\{ \sum_{i \in V} c(i)\varphi_i,\ \big(c(i)\big)_{i \in V} \in \ell^p \Big\},\ 1 \le p \le \infty,$$

where $\Phi = \{\varphi_i\}_{i \in V}$ and $\ell^p, 1 \le p \le \infty$, are Banach spaces of all p-summable sequences with norms denoted by $\|\cdot\|_p$. Therefore spatial signals f with the parametric representation (1.3) belong to the space $V_\infty(\Phi)$, i.e.,

$$f \in V_\infty(\Phi).$$

The spaces $V_p(\Phi), 1 \le p \le \infty$, were introduced in [62] for modeling signals with finite rate of innovation, which include the classical band-limited signals, wavelet signals and spatial signals in many engineering applications, see [27,60,72] and references therein. For signals with finite rate of innovation on $\mathbb{R}^d$, every index i in V is associated with an innovative position in $\mathbb{R}^d$ and the generator φ_i is essentially supported in a spatial neighborhood of the innovative position of $i \in V$. So in this paper, we follow the terminology in [62] to call V the set of innovative positions of spatial signals in (1.3).

In this paper, 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 deployed

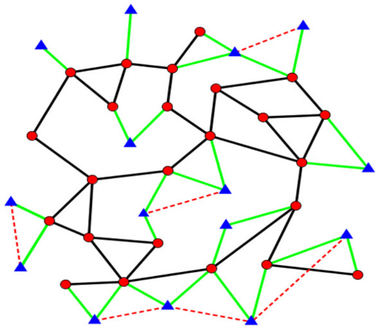

Fig. 1. The graph $\mathcal{H} = (G \cup V,\ S \cup T \cup T^*)$ in (1.4) to describe a DSRS, where vertices in G and V are plotted in red circles and blue triangles, and edges in S, T and E are in black solid lines, green solid lines and red dashed lines respectively.

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

$$\mathcal{H} := (G \cup V,\ S \cup T \cup T^*), \tag{1.4}$$

where $T^* = \{(\lambda, i) \in G \times V,\ (i, \lambda) \in T\}$, see Fig. 1. 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

$$E := \{(i, i') \in V \times V,\ \ i \neq i' \text{ and } (i, \lambda), (i', \lambda) \in T \text{ for some } \lambda \in G\}. \tag{1.5}$$

We then generate a graph structure

$$\mathcal{V} := (V, E) \tag{1.6}$$

for signals in (1.3), 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 [53,54,57] and Remark 3.6.

Define sensing matrix $\mathbf{S}$ of our DSRS by

$$\mathbf{S} := (\langle \varphi_i, \psi_\lambda \rangle)_{\lambda \in G, i \in V}. \tag{1.7}$$

The sensing matrix $\mathbf{S}$ is stored by agents in a distributed manner. Due to the storage limitation, each agent in our SDN stores its corresponding row (and perhaps also its neighboring rows) in the sensing matrix $\mathbf{S}$, but it does not have the whole matrix accessible. Agents in our SDN 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

$$|\langle \varphi_i, \psi_\lambda \rangle| \leq D(1 + \rho_{\mathcal{H}}(\lambda, i))^{-\alpha} \text{ for all } \lambda \in G \text{ and } i \in V, \tag{1.8}$$

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 SDN in a distance. Thus the sensing matrix $\mathbf{S}$ may have finite bandwidth $s \geq 0$,

$$\langle \varphi_i, \psi_\lambda \rangle = 0 \quad \text{if } \rho_{\mathcal{H}}(\lambda, i) > s. \tag{1.9}$$

The above global requirements (1.8) and (1.9) could be fulfilled in a distributed manner.

The sensing matrix $\mathbf{S}$ characterizes the sampling procedure (1.2) of signals with the parametric representation (1.3). Applying the sensing matrix $\mathbf{S}$, we obtain the sample vector $\mathbf{y} = (\langle f, \psi_\lambda\rangle)_{\lambda\in G}$ of the signal f from its amplitude vector $\mathbf{c} := (c(i))_{i\in V}$, i.e.,

$$\mathbf{y} = \mathbf{S}\mathbf{c}. \tag{1.10}$$

Under the assumptions (1.8) and (1.9), it is shown in Proposition 4.1 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

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

Here the capital letter C is an absolute constant which is not necessarily the same at each occurrence.

A fundamental problem in sampling theory is the robustness of signal reconstruction in the presence of sampling noises ([11,32,47–49,52,58]). In this paper, we consider the scenario that the sampling data $\mathbf{y} = \mathbf{S}\mathbf{c}$ is corrupted by bounded deterministic/random noise $\boldsymbol{\eta} = (\eta(\lambda))_{\lambda\in G}$,

$$\mathbf{z} = \mathbf{S}\mathbf{c} + \boldsymbol{\eta} \tag{1.11}$$

([67,73]). 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 difference 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.,

$$\|\Delta(\mathbf{z}) - \mathbf{c}\|_\infty \le C\delta \tag{1.12}$$

for some absolute constant C ([1,7,17]).

Given the noisy sampling vector $\mathbf{z}$ in (1.11), consider the following global optimization problem of maximal sampling error ([13,14])

$$\Delta_\infty(\mathbf{z}) := \operatorname*{argmin}_{\mathbf{d}\in\ell^\infty} \|\mathbf{S}\mathbf{d} - \mathbf{z}\|_\infty. \tag{1.13}$$

The above minimization problem can be solved by linear programming

$$\min_{\mathbf{d}} t \quad \text{subject to } \mathbf{S}\mathbf{d} - \mathbf{z} \le t\mathbf{1} \ \text{ and } \ -\mathbf{S}\mathbf{d} + \mathbf{z} \le t\mathbf{1}, \tag{1.14}$$

where $\mathbf{1} = (1, \dots, 1)^T$ is the vector with one as its entries.

Definition 1.1. For $1 \le p \le \infty$, a matrix $\mathbf{A}$ is said to have ℓ^p-stability if there exist positive constants A and B such that

$$A\|\mathbf{c}\|_p \le \|\mathbf{A}\mathbf{c}\|_p \le B\|\mathbf{c}\|_p \ \text{ for all } \mathbf{c} \in \ell^p. \tag{1.15}$$

We call the minimal constant B and the maximal constant A for (1.15) to hold the upper and lower ℓ^p-stability bounds respectively.

Observe from (1.11) and (1.13) 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 (1.13) gives a suboptimal approximation to the true amplitude vector $\mathbf{c}$ if the sensing matrix $\mathbf{S}$ of the DSRS has ℓ^∞-stability ([7,68,71]), cf. Fig. 7. In Theorem 5.2, we show that for a matrix with some polynomial off-diagonal decay if it has ℓ^2-stability then it has ℓ^∞-stability with the lower ℓ^∞-stability bound independent of the size of the DSRS.

Next we consider the problem how to verify ℓ^2-stability of the sensing matrix $\mathbf{S}$ of our DSRS in a distributed manner. It is well known that a finite-dimensional matrix $\mathbf{S}$ has ℓ^2-stability if and only if $\mathbf{S}^T\mathbf{S}$ is strictly positive, and its upper and lower stability bounds are the same as square roots of largest and smallest eigenvalues of $\mathbf{S}^T\mathbf{S}$. The above procedure to establish ℓ^2-stability for the sensing matrix of our DSRS is not feasible, because the whole sensing matrix $\mathbf{S}$ is not available for any agent in the DSRS and there is no centralized processor to evaluate eigenvalues of $\mathbf{S}^T\mathbf{S}$. In Theorems 6.1 and 6.2, we introduce a method to split the DSRS into a family of overlapping subsystems of small size, and we show that the sensing matrix $\mathbf{S}$ with polynomial off-diagonal decay has ℓ^2-stability if and only if its quasi-restrictions to those subsystems have uniform ℓ^2-stability. The new local criterion in Theorems 6.1 and 6.2 provides a reliable tool for the verification of the ℓ^2-stability in a distributed manner. Also the local criterion is pivotal for the design of a robust DSRS against supplement, replacement and impairment of agents, as it suffices to verify the uniform stability of affected subsystems.

Then 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 ([5,17,20,31,34,49,60,67]). In this paper, we develop a distributed algorithm to find the suboptimal approximation

$$\Delta_2(\mathbf{z}) := (\mathbf{S}^T\mathbf{S})^{-1}\mathbf{S}^T\mathbf{z} \tag{1.16}$$

to the original signal f in (1.3). 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 (1.16) is the unique least squares solution,

$$\Delta_2(\mathbf{z}) = \operatorname*{argmin}_{\mathbf{d}\in\ell^2} \|\mathbf{S}\mathbf{d}-\mathbf{z}\|_2^2 = \operatorname*{argmin}_{\mathbf{d}\in\ell^2} \sum_{\lambda\in G} \Theta_\lambda(\mathbf{d},\mathbf{z}), \tag{1.17}$$

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

$$\Theta_\lambda(\mathbf{d},\mathbf{z}) = \Big|\sum_{i\in V}\langle\varphi_i,\psi_\lambda\rangle d(i) - z(\lambda)\Big|^2,\ \lambda\in G. \tag{1.18}$$

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

$$\min \sum_{\lambda\in G}\Theta_\lambda(\mathbf{d},\mathbf{z}). \tag{1.19}$$

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

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

is a distributed algorithm to solve the minimization problem (1.19), 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 [10,16,43,46,50] 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 (7.19) and (7.20) to approximate $\Delta_2(\mathbf{z})$ in (1.16), when the sensing matrix $\mathbf{S}$ has ℓ^2-stability and satisfies the requirements (1.7) and (1.8). In the above distributed algorithm for signal reconstruction, each agent in the SDN 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 (1.16). More importantly, we show in Theorems 7.1 and 7.2 that the proposed distributed algorithm (7.19) and (7.20) converges exponentially to the solution $\Delta_2(z)$ in (1.16). The establishment for the above convergence is virtually based on Wiener's lemma for localized matrices ([37,38,41,59,61,66]) and on the observation that our sensing matrices are quasi-diagonal block dominated.

The paper is organized as follows. In Section 2, we make some basic assumptions on the SDN and we introduce its Beurling dimension and sampling density. In Section 3, we introduce the graph $\mathcal{H}$ to describe our DSRS and then we define dimension and maximal rate of innovation for signals on the graph $\mathcal{V}$. We show in Theorem 3.5 that the dimension for signals is the same as the Beurling dimension for the SDN, and the maximal rate of innovation is approximately proportional to the sampling density of the SDN. In Section 4, we prove in Proposition 4.1 that sampling a signal with bounded amplitude vector by the procedure (1.2) produces a bounded sampling data vector when the sensing matrix of the DSRS has certain polynomial off-diagonal decay. In Section 5, we establish in Theorem 5.2 that if a matrix with certain off-diagonal decay has ℓ^2-stability then it has ℓ^p-stability for all $1 \le p \le \infty$, and also in Theorem 5.4 that the solution $\Delta_2(\mathbf{z})$ in (1.16) is a suboptimal approximation to the true amplitude vector. In Theorems 6.1 and 6.2 of Section 6, we introduce a criterion for the ℓ^2-stability of a sensing matrix, that could be verified in a distributed manner. In Section 7, we propose a distributed algorithm to solve the minimization problem (1.17). In Section 8, we present simulations to demonstrate our proposed algorithm for robust signal reconstruction. In Section 9, we include proofs of all conclusions.

The sampling theory developed in this paper enjoys the advantages of scalability of network sizes and data privacy preservation. Some results of this paper were announced in [18].

Notation: $\mathbf{A}^T$ is the transpose of a matrix $\mathbf{A}$; $\|c\|_p$ is the norm on ℓ^p; χ_F is the index function on a set F; $\lceil x \rceil$ is the ceiling of $x \in \mathbb{R}$; $\lfloor x \rfloor$ is the floor of $x \in \mathbb{R}$; $\#F$ is the cardinality of a set F; and $\|\mathbf{A}\|_{\mathcal{B}^2}$ is the operator norm of a matrix $\mathbf{A}$ on ℓ^2.

2. Spatially distributed networks

Let $\mathcal{G}$ be the graph in (1.1) to describe our SDN. In this paper, we always assume that $\mathcal{G}$ is *connected* and *simple* (i.e., undirected, unweighted, no graph loops nor multiple edges), which can be interpreted as follows:

- Agents in the SDN 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 SDN. For a discrete sampling set in the d-dimensional Euclidean space, the reader may refer to [24,30] for its Beurling dimension and to [7,49,60,71] for its sampling density. Finally, we introduce a special family of balls to cover the graph $\mathcal{G}$, which will be used in Section 7 for the consensus of our proposed distributed algorithm.

2.1. 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 $\mathcal{G}$ ([21]). It is nonnegative and symmetric:

(i) $\rho_{\mathcal{G}}(\lambda, \lambda') \ge 0$ 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) $\rho_{\mathcal{G}}(\lambda, \lambda') \le \rho_{\mathcal{G}}(\lambda, \lambda'') + \rho_{\mathcal{G}}(\lambda'', \lambda')$ for all $\lambda, \lambda', \lambda'' \in G$.

In many real-world applications, the distance $\rho_{\mathcal{G}}(\lambda, \lambda')$ can be used to measure the communication cost between two distinct agents λ and $\lambda' \in G$, since communication between them happens by transmitting information through the chain of intermediate agents connecting them using a shortest path.

2.2. Counting measure, Beurling dimension and sampling density

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

$$\mu_{\mathcal{G}}(F) := \#(F) \quad \text{for } F \subset G,$$

where $\#F$ is the cardinality of a set F.

Definition 2.1. The counting measure $\mu_{\mathcal{G}}$ is said to be a doubling measure on $\mathcal{G}$ if there exists a positive number $D_0(\mathcal{G})$ such that

$$\mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda, 2r)) \le D_0(\mathcal{G})\mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda, r)) \quad \text{for all} \quad \lambda \in G \text{ and } r \ge 0, \tag{2.1}$$

where

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

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 (2.1) to hold ([22,25]). It dominates the maximal vertex degree of the graph $\mathcal{G}$,

$$\deg(\mathcal{G}) \le D_0(\mathcal{G}), \tag{2.2}$$

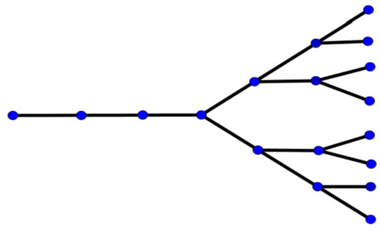

Fig. 2. A tree with large doubling constant but limited maximal vertex degree.

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 Fig. 2 with $L = 3$.

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 Fig. 2. In this paper, the graph $\mathcal{G}$ to describe our SDN is assumed to have its counting measure with the doubling property (2.1).

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

$$D_0(\mathcal{G}) < \infty. \tag{2.3}$$

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 SDN.

Definition 2.2. 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

$$\mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda,r)) \le D_1(\mathcal{G})(1+r)^{d(\mathcal{G})} \quad \text{for all } \lambda \in G \text{ and } r \ge 0. \tag{2.4}$$

For the graph $\mathcal{G}$ associated with an SDN, we may consider minimal constants $d(\mathcal{G})$ and $D_1(\mathcal{G})$ in (2.4) as *Beurling dimension* and *sampling density* of the SDN respectively. We remark that

$$d(\mathcal{G}) \ge 1, \tag{2.5}$$

because

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

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

Applying (2.1) 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 2.3. *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,

$$\mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda, r)) \le \frac{(\deg(\mathcal{G}))^{r+1}-1}{\deg(\mathcal{G})-1} \quad \text{for all } \lambda \in G \text{ and } r \ge 0. \tag{2.6}$$

2.3. Spatially distributed fusion subnetworks

For a connected simple graph $\mathcal{G} := (G, S)$ and $N \ge 0$, we say that $G_N \subset G$ is a *maximal N-disjoint subset of G* if

$$B_{\mathcal{G}}(\lambda, N) \bigcap \big(\bigcup_{\lambda_m\in G_N} B_{\mathcal{G}}(\lambda_m, N)\big) \ne \emptyset \quad \text{for all } \lambda \in G, \tag{2.7}$$

and

$$B_{\mathcal{G}}(\lambda_m, N) \bigcap B_{\mathcal{G}}(\lambda_{m'}, N) = \emptyset \quad \text{for all } \lambda_m, \lambda_{m'} \in G_N. \tag{2.8}$$

For $0 \le N < 1$, it follows from (2.7) that $G_N = G$. For $N \ge 1$, there are many subsets G_N of vertices satisfying (2.7) and (2.8). 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 (2.7) and (2.8), 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 2.4. *Let $\mathcal{G} := (G, S)$ be a connected simple graph and $\mu_{\mathcal{G}}$ have the doubling property (2.3) with constant $D_0(\mathcal{G})$. If G_N satisfies (2.7) and (2.8), then*

$$1 \le \inf_{\lambda\in G} \Sigma_{\lambda_m\in G_N} \chi_{B_{\mathcal{G}}(\lambda_m, N')}(\lambda) \le \sup_{\lambda\in G} \Sigma_{\lambda_m\in G_N} \chi_{B_{\mathcal{G}}(\lambda_m, N')}(\lambda) \le (D_0(\mathcal{G}))^{\lceil \log_2(2N'/N+1)\rceil} \tag{2.9}$$

for all $N' \ge 2N$.

For $N' \ge 0$, define a family of spatially distributed fusion subnetworks

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

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 2.4 that for any $N' \ge 2N$, every agent in our SDN 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 SDN into subnetworks of small sizes is crucial in our proposed distributed algorithm in Section 7 for stable signal reconstruction.

3. Distributed sampling and reconstruction systems

Let V be the set of innovative positions of signals f in (1.3), and $\mathcal{G} = (G, S)$ be the graph in (1.1) to represent our SDN. In this section, we introduce a graph $\mathcal{H}$ to describe our distributed sampling and reconstruction systems, and also a graph $\mathcal{V}$ to describe the topology of spatial signals with the parametric representation (1.3).

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,*

$$(\lambda_1, \lambda_2) \in S \textit{ if } (i, \lambda_1) \textit{ and } (i, \lambda_2) \in T \textit{ for some } i \in V. \tag{3.1}$$

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

$$L := \sup_{\lambda \in G} \#\{i \in V,\ (i, \lambda) \in T\} < \infty. \tag{3.2}$$

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

$$M := \sup_{\lambda \in G} \inf\{\rho_{\mathcal{G}}(\lambda, \lambda'),\ (i, \lambda') \in T \textit{ for some } i \in V\} < \infty. \tag{3.3}$$

Under the above assumptions, the graph $\mathcal{H}$ in (1.4) is a connected simple graph. Moreover, we have the following important properties about shortest paths between different vertices in $\mathcal{H}$.

Proposition 3.1. *Let the graph $\mathcal{H}$ in (1.4) satisfy (3.1). Then all intermediate vertices in the shortest paths in $\mathcal{H}$ to connect distinct vertices in $\mathcal{H}$ belong to the subgraph $\mathcal{G}$.*

By Proposition 3.1,

$$\rho_{\mathcal{H}}(\lambda, \lambda') = \rho_{\mathcal{G}}(\lambda, \lambda') \quad \text{for all } \lambda, \lambda' \in G, \tag{3.4}$$

and

$$\rho_{\mathcal{H}}(i, i') = 2 + \inf_{\lambda,\lambda' \in G}\{\rho_{\mathcal{G}}(\lambda, \lambda'),\ (i, \lambda), (i', \lambda') \in T\} \quad \text{for all distinct } i, i' \in V, \tag{3.5}$$

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

Let $\mathcal{V}$ be the graph in (1.6), 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,

$$\deg(\mathcal{V}) \le L \sup_{i\in V} \#\{\lambda \in G,\ (i,\lambda)\in T\} \le L(\deg(\mathcal{G})+1) \tag{3.6}$$

by (2.2), (2.3), (3.1) and (3.2).

We cannot define a geodesic distance on $\mathcal{V}$ as in Subsection 2.1, 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 3.2. *Let $\mathcal{H}$ be the graph in (1.4). Define a function $\rho : V\times V \longmapsto \mathbb{R}$ by*

$$\rho(i,i') = \begin{cases} 0 & \text{if } i=i' \\ \rho_{\mathcal{H}}(i,i')-1 & \text{if } i\ne i'. \end{cases} \tag{3.7}$$

If the graph $\mathcal{H}$ satisfies (3.1), 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) $\rho(i,i')\le \rho(i,i'')+\rho(i'',i')$ *for all* $i,i',i''\in V$*.*

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

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

and the counting measure on V by μ. Similar to the counting measure $\mu_{\mathcal{G}}$ on an SDN in Definitions 2.1 and 2.2, we say that the measure μ on $\mathcal{V}$ is a *doubling measure* if

$$\mu(B(i,2r))\le D_0\mu(B(i,r)) \text{ for all } i\in V \text{ and } r\ge 0, \tag{3.8}$$

and it has *polynomial growth* if

$$\mu(B(i,r))\le D_1(1+r)^d \text{ for all } i\in V \text{ and } r\ge 0, \tag{3.9}$$

where D_0, D_1 and d are positive constants. The minimal constant D_0 for (3.8) to hold is known as the doubling constant, and the minimal constants d and D_1 in (3.9) 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 [72] and later extended in [62,68]. The reader may refer to [11,12,29,47,51,56,60,62,68,72] 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 3.3. *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*

$$\mu(B(i,2r))\le L(D_0(\mathcal{G}))^2\Big(\frac{(\deg(\mathcal{G}))^{2M+3}-1}{\deg(\mathcal{G})-1}\Big)\mu(B(i,r)) \quad \textit{for all } i\in V \textit{ and } r\ge 0. \tag{3.10}$$

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

$$\mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda,2r))\le LD_0^2\Big(\frac{(\deg(\mathcal{G}))^{2M+3}-1}{\deg(\mathcal{G})-1}\Big)^2\mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda,r)) \quad \textit{for all } \lambda\in G \textit{ and } r\ge 0. \tag{3.11}$$

Proposition 3.4. *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*

$$\mu(B(i,r)) \le L D_1(\mathcal{G})(1+r)^{d(\mathcal{G})} \quad \text{for all } i \in V \text{ and } r \ge 0. \tag{3.12}$$

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

$$\mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda,r)) \le 2^d\Big(\frac{(\deg(\mathcal{G}))^{2M+3}-1}{\deg(\mathcal{G})-1}\Big)D_1(1+r)^d \quad \text{for all } \lambda \in G \text{ and } r \ge 0. \tag{3.13}$$

By (2.5), Propositions 3.3 and 3.4, we conclude that signals in (1.3) 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 $D_1(\mathcal{G})$.

Theorem 3.5. *Let $\mathcal{G}$ and $\mathcal{H}$ satisfy Assumptions 1–4. Then*

$$d(\mathcal{G}) = d \ge 1 \tag{3.14}$$

and

$$L^{-1}D_1 \le D_1(\mathcal{G}) \le 2^d\Big(\frac{(\deg(\mathcal{G}))^{2M+3}-1}{\deg(\mathcal{G})-1}\Big)D_1. \tag{3.15}$$

We finish this section with a remark about signals on our graph $\mathcal{V}$, cf. [53,54,57].

Remark 3.6. Signals on the graph $\mathcal{V}$ are analog in nature, while signals on graphs in most of the literature are discrete ([53,54,57]). 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 (1.3), 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 $\mathbf{p}_u, u \in G \cup V$.

4. Sensing matrices with polynomial off-diagonal decay

Let $\mathcal{H}$ be the connected simple graph in (1.4) to describe our DSRS, and the sensing matrix $\mathbf{S}$ associated with the DSRS be as in (1.7). As agents in the DSRS have limited sensing ability, we assume in this paper that the sensing matrix $\mathbf{S}$ in (1.7) satisfies

$$\mathbf{S} \in \mathcal{J}_\alpha(\mathcal{G},\mathcal{V}) \text{ for some } \alpha > d, \tag{4.1}$$

where

$$\mathcal{J}_\alpha(\mathcal{G},\mathcal{V}) := \left\{\mathbf{A} := (a(\lambda,i))_{\lambda\in G, i\in V},\ \|\mathbf{A}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})} < \infty\right\} \tag{4.2}$$

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

$$\|\mathbf{A}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})} := \sup_{\lambda\in G, i\in V} (1+\rho_{\mathcal{H}}(\lambda,i))^\alpha |a(\lambda,i)|, \quad \alpha \ge 0. \tag{4.3}$$

The reader may refer to [37,38,41,59,61,66] 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 4.1. *Let $\mathcal{G}$ and $\mathcal{H}$ satisfy Assumptions 1–4, $\mathcal{V}$ be as in (1.6), 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*

$$\|\mathbf{A}\mathbf{c}\|_p \le \frac{D_1(\mathcal{G})L\alpha}{\alpha-d}\|\mathbf{A}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}\|\mathbf{c}\|_p \ \ \textit{for all } \mathbf{c}\in\ell^p, 1\le p\le\infty. \tag{4.4}$$

For a DSRS with its sensing matrix in $\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})$, we obtain from (1.10) and Proposition 4.1 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

$$\mathbf{A}_s := (a_s(\lambda,i))_{\lambda\in G,i\in V},\ s\ge 0, \tag{4.5}$$

where

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

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 $\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})$,

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

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 4.2. *Let graphs $\mathcal{G}$, $\mathcal{H}$, $\mathcal{V}$, d and $D_1(\mathcal{G})$ be as in Proposition 4.1. If $\mathbf{A}\in\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})$ for some $\alpha > d$, then*

$$\|(\mathbf{A}-\mathbf{A}_s)\mathbf{c}\|_p \le \frac{D_1(\mathcal{G})L\alpha}{\alpha-d}(s+1)^{-\alpha+d}\|\mathbf{A}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}\|\mathbf{c}\|_p \ \ \textit{for all } \mathbf{c}\in\ell^p, 1\le p\le\infty, \tag{4.6}$$

where $\mathbf{A}_s, s\ge 1$, are band matrices in (4.5).

The above band matrix approximation property will be used later in the establishment of a local stability criterion in Section 6 and exponential convergence of a distributed reconstruction algorithm in Section 7.

5. Robustness of distributed sampling and reconstruction systems

Let $\mathbf{S}$ be the sensing matrix associated with our DSRS. We say that a reconstruction algorithm Δ is a *perfect reconstruction* in noiseless environment if

$$\Delta(\mathbf{Sc}) = \mathbf{c} \quad \text{for all } \mathbf{c} \in \ell^\infty. \tag{5.1}$$

In this section, we first study robustness of the DSRS in term of the ℓ^∞-stability.

Proposition 5.1. *Let $\mathcal{G}$ and $\mathcal{H}$ satisfy Assumptions 1–4, $\mathcal{V}$ be as in (1.6), $\mu_\mathcal{G}$ have polynomial growth with Beurling dimension d, and let $\mathbf{S}$ satisfy (4.1). Then there is a reconstruction algorithm Δ with the suboptimal approximation property (1.12) and the perfect reconstruction property (5.1) if and only if $\mathbf{S}$ has ℓ^∞-stability.*

The sufficiency in Proposition 5.1 holds by taking $\Delta = \Delta_\infty$ in (1.13), while the necessity follows by applying (1.12) to $\boldsymbol{\eta} = \mathbf{Sd}$ with $\mathbf{d} \in \ell^\infty$.

In the next theorem, we reduce ℓ^∞-stability of a matrix in Jaffard class to its ℓ^2-stability, for which a distributed verifiable criterion will be provided in Section 6.

Theorem 5.2. *Let $\mathcal{G}, \mathcal{H}, \mathcal{V}$ and d be as in Proposition 5.1, 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 $1 \le p \le \infty$ with the lower ℓ^p-stability bound independent of the size of the DSRS.*

The reader may refer to [3,55,66] for equivalence of ℓ^p-stability of localized matrices for different $1 \le p \le \infty$. For finite graphs $\mathcal{G} = (G, S)$ and $\mathcal{V} = (V, E)$ and a matrix $\mathbf{A}$ with row indices in G and column indices in V, its ℓ^p-stability and ℓ^q-stability are equivalent to each other for any $1 \le p, q \le \infty$, and its optimal lower stability bounds A_p and A_q satisfy

$$M^{-|1/p-1/q|} \le \frac{A_q}{A_p} \le M^{|1/p-1/q|},$$

where $M = \max(\#G, \#V)$ is the number of vertices of graphs $\mathcal{G}$ and $\mathcal{V}$. The above estimation on lower stability bounds is unfavorable for matrices of large size but it cannot be improved if there is no restriction on the matrix $\mathbf{A}$. For matrices $\mathbf{A}$ in the Jaffard class $\mathcal{J}_\alpha(\mathcal{G}, \mathcal{V})$, we obtain from the proof of Theorem 5.2 that the lower ℓ^p-stability bound depends 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 (3.2) and (3.3). So the sensing matrix of our DSRS may have its lower ℓ^p-stability bounds **independent** of the size of the DSRS.

For the graph $\mathcal{V}$ in (1.6) and the distance ρ in (3.7), define

$$\mathcal{J}_\alpha(\mathcal{V}) := \left\{\mathbf{A} := (a(i,i'))_{i,i' \in V},\ \|\mathbf{A}\|_{\mathcal{J}_\alpha(\mathcal{V})} < \infty\right\}, \tag{5.2}$$

where

$$\|\mathbf{A}\|_{\mathcal{J}_\alpha(\mathcal{V})} := \sup_{i,i' \in V} (1 + \rho(i,i'))^\alpha |a(i,i')|,\ \alpha \ge 0. \tag{5.3}$$

The proof of Theorem 5.2 depends highly on the following Wiener's lemma for the matrix algebra $\mathcal{J}_\alpha(\mathcal{V})$, $\alpha > d$.

Theorem 5.3. *Let $\mathcal{V}$ be as in (1.6) and its counting measure μ satisfy (3.9). 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.*

Wiener's lemma has been established for infinite matrices, pseudodifferential operators, and integral operators satisfying various off-diagonal decay conditions ([9,33,35,37,38,41,59,61,63,66]). 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 [36,44] for historical remarks, motivation and recent advances.

The Wiener's lemma (Theorem 5.3) is also used to establish the sub-optimal approximation property (1.12) for the "least squares" solution $\Delta_2(\mathbf{z})$ in (1.16), for which a distributed algorithm is proposed in Section 7.

Theorem 5.4. *Let $\mathcal{G}, \mathcal{H}$ and $\mathcal{V}$ be as in Proposition 5.1. Assume that the sensing matrix $\mathbf{S}$ satisfies (4.1) and it has ℓ^2-stability. Then there exists a positive constant C such that*

$$\|\Delta_2(\mathbf{z}) - \mathbf{c}\|_\infty \le C\|\boldsymbol{\eta}\|_\infty \quad \textit{for all } \mathbf{c}, \boldsymbol{\eta} \in \ell^\infty, \tag{5.4}$$

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

6. Stability criterion for distributed sampling and reconstruction systems

In a traditional centralized sampling and reconstruction system, the ℓ^2-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 section, we introduce a stability criterion for matrices in the Jaffard class that can be verified in a distributed manner.

Let $\mathcal{H}$ be the connected simple graph in (1.4) 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 6.1. *Let $\mathcal{G}$ and $\mathcal{H}$ satisfy Assumptions 1–4, $\mathcal{V}$ be as in (1.6), $\mu_{\mathcal{G}}$ 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*

$$\|\chi^{2N}_{\lambda,G}\mathbf{A}\chi^N_{\lambda,V}\mathbf{c}\|_2 \ge \frac{A}{2}\|\mathbf{A}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}\|\chi^N_{\lambda,V}\mathbf{c}\|_2, \ \mathbf{c} \in \ell^2 \tag{6.1}$$

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

$$2D_1(\mathcal{G})N^{-\alpha+d}\sqrt{L\alpha/(\alpha-d)} \le A. \tag{6.2}$$

The above theorem provides a guideline to design a distributed algorithm for signal reconstruction, see Section 7. Surprisingly, the converse of Theorem 6.1 is true, cf. the stability criterion in [65, Theorem 2.1] for convolution-dominated matrices.

Theorem 6.2. *Let $\mathcal{G}, \mathcal{H}, \mathcal{V}$ be as in Theorem 6.1, 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*

$$A_0 \geq 4(D_0(\mathcal{G}))^2 D_1(\mathcal{G}) L N_0^{-\min(\alpha-d,1)} \times \begin{cases} \left(\frac{4\alpha}{3(\alpha-d)} + \frac{2(\alpha-1)(\alpha-d)}{\alpha-d-1}\right) & \text{if } \alpha > d+1 \\ \left(\frac{10(d+1)}{3} + 2d \ln N_0\right) & \text{if } \alpha = d+1 \\ \left(\frac{4\alpha}{3(\alpha-d)} + \frac{4d}{d+1-\alpha}\right) & \text{if } \alpha < d+1, \end{cases} \tag{6.3}$$

and for all $\lambda \in G$,

$$\|\chi_{\lambda,G}^{2N_0} \mathbf{A} \chi_{\lambda,V}^{N_0} \mathbf{c}\|_2 \geq A_0 \|\mathbf{A}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})} \|\chi_{\lambda,V}^{N_0} \mathbf{c}\|_2, \ \ \mathbf{c} \in \ell^2, \tag{6.4}$$

then $\mathbf{A}$ has ℓ^2-stability,

$$\|\mathbf{A}\mathbf{c}\|_2 \geq \frac{A_0 \|\mathbf{A}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}}{12(D_0(\mathcal{G}))^2} \|\mathbf{c}\|_2, \ \ \mathbf{c} \in \ell^2. \tag{6.5}$$

Observe that the right hand side of (6.3) could be arbitrarily small when N_0 is sufficiently large. This together with Theorem 6.1 implies that the requirements (6.3) and (6.4) 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 (6.3) cannot be replaced by $N_0^{-\beta}$ with high order $\beta > 1$ even if the matrix $\mathbf{A}$ has finite bandwidth.

Example 6.3. 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 $\alpha \geq 0$ and it does not have ℓ^2-stability. On the other hand, for any $\lambda \in G = V = \mathbb{Z}$ and $N_0 \geq 1$,

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

where the last equality follows from [42, Lemma 1 of Chapter 9].

For our DSRS with sensing matrix $\mathbf{S}$ having the polynomial off-diagonal decay property (4.1), the uniform stability property (6.4) 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.

7. Exponential convergence of a distributed reconstruction algorithm

In our DSRS, agents could essentially catch signals not far from their spatial 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 ([6,39,64,69]). 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

$$\mathbf{d}_2 = \mathbf{S}^\dagger \mathbf{z} \tag{7.1}$$

is a suboptimal approximation, where $\mathbf{z}$ is given in (1.11). 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 (7.1).

Let $\mathcal{H}$ be the connected simple graph in (1.4) 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 (7.1) is the unique solution to the “normal” equation

$$\mathbf{S}^T\mathbf{S}\mathbf{d}_2 = \mathbf{S}^T\mathbf{z}. \tag{7.2}$$

Instead of solving the above equation, we consider localized linear systems

$$\chi_{\lambda,V}^N \mathbf{S}^T\mathbf{S}\chi_{\lambda,V}^N \mathbf{d}_{\lambda,N} = \chi_{\lambda,V}^N \mathbf{S}^T\mathbf{z},\ \lambda \in G, \tag{7.3}$$

of size $O(N^d)$, whose solutions $\mathbf{d}_{\lambda,N}$ are supported in the ball $B_\mathcal{H}(\lambda, N) \cap V$. The localized system (7.3) has unique solution as principal submatrices $\chi_{\lambda,V}^N \mathbf{S}^T\mathbf{S}\chi_{\lambda,V}^N, N \geq 1$, of the positive definite matrix $\mathbf{S}^T\mathbf{S}$ are uniformly stable. 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 (7.6) in Proposition 7.1. 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

$$\mathbf{w}_{\lambda,N} := \chi_{\lambda,G}^N \mathbf{S}\chi_{\lambda,V}^N (\chi_{\lambda,V}^N \mathbf{S}^T\mathbf{S}\chi_{\lambda,V}^N)^{-1}\mathbf{d}_{\lambda,N} = \chi_{\lambda,G}^N \mathbf{S}\chi_{\lambda,V}^N (\chi_{\lambda,V}^N \mathbf{S}^T\mathbf{S}\chi_{\lambda,V}^N)^{-2}\chi_{\lambda,V}^N \mathbf{S}^T\mathbf{z} \tag{7.4}$$

instead, which approximates

$$\mathbf{w}_{LS} := \mathbf{S}(\mathbf{S}^T\mathbf{S})^{-1}\mathbf{d}_2 \tag{7.5}$$

inside $B_\mathcal{G}(\lambda, N/2) \cap G$, see (7.7) in the proposition below.

Proposition 7.1. *Let $\mathcal{G}$ and $\mathcal{H}$ satisfy Assumptions 1–4, $\mathcal{V}$ be as in (1.6), 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 (6.2), and set*

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

Then

$$\|\chi_{\lambda,V}^{N/2}(\mathbf{d}_{\lambda,N}-\mathbf{d}_2)\|_\infty \le D_3(N+1)^{-\alpha+d}\|\mathbf{d}_2\|_\infty \tag{7.6}$$

and

$$\|\chi_{\lambda,G}^{N/2}(\mathbf{w}_{\lambda,N}-\mathbf{w}_{LS})\|_\infty \le D_4(N+1)^{-\alpha+d}\|\mathbf{d}_2\|_\infty, \tag{7.7}$$

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

$$D_2 = \sum_{n=0}^{\infty}\Big(\frac{2^{2\alpha+d/2+4}D_1^3\alpha^2}{r_0^{1-\theta}(\alpha-d)^2}\Big)^{\frac{2-\theta}{(1-\theta)^2}n^{\log_2^{(2-\theta)}}} r_0^n. \tag{7.8}$$

Take a maximal $\frac{N}{4}$-disjoint subset $G_{N/4}\subset G$ satisfying (2.7) and (2.8). We patch $\mathbf{w}_{\lambda,N}, \lambda\in G_{N/4}$, in (7.4) together to generate a linear approximation

$$\mathbf{w}_N^* = \sum_{\lambda\in G_{N/4}} \mathbf{\Theta}_{\lambda,N}\chi_{\lambda,G}^{N/2}\mathbf{w}_{\lambda,N} \tag{7.9}$$

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 (3.4) and Proposition 2.4. Moreover, we obtain from Proposition 7.1 that

$$\begin{aligned}\|\mathbf{w}_N^*-\mathbf{w}_{LS}\|_\infty &= \Big\|\sum_{\lambda\in G_{N/4}}\mathbf{\Theta}_{\lambda,N}\chi_{\lambda,G}^{N/2}(\mathbf{w}_{\lambda,N}-\mathbf{w}_{LS})\Big\|_\infty\\ &\le \sup_{\lambda''\in G}\sum_{\lambda\in G_{N/4}}\theta_{\lambda,N}(\lambda'')\|\chi_{\lambda,G}^{N/2}(\mathbf{w}_{\lambda,N}-\mathbf{w}_{LS})\|_\infty\\ &\le D_4(N+1)^{-\alpha+d}\|\mathbf{d}_2\|_\infty. \end{aligned}\tag{7.10}$$

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 (7.5) for large N. In addition, $\mathbf{w}_N^*$ depends on the observation $\mathbf{z}$ linearly,

$$\mathbf{w}_N^* = \mathbf{R}_N\mathbf{S}^T\mathbf{z} \tag{7.11}$$

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

$$\|\mathbf{R}_N\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})} \le D_5 := \frac{(\alpha-d)^2LD_2^2}{\alpha^2D_1D_1(\mathcal{G})\|\mathbf{S}\|^3_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}}. \tag{7.12}$$

Given noisy samples $\mathbf{z}$, we may use $\mathbf{w}_N^*$ in (7.11) as the first approximation of $\mathbf{w}_{LS}$,

$$\mathbf{w}_1 = \mathbf{R}_N\mathbf{S}^T\mathbf{z} \tag{7.13}$$

and recursively define

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \mathbf{w}_1 - \mathbf{R}_N\mathbf{S}^T\mathbf{S}\mathbf{S}^T\mathbf{w}_n,\ n\ge 1. \tag{7.14}$$

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

$$\mathbf{S}^T\mathbf{w} = \mathbf{S}^T\mathbf{w}_{LS} = \mathbf{d}_2. \tag{7.15}$$

Theorem 7.2. *Let $\mathcal{G}, \mathcal{H}$ and $\mathcal{V}$ be as in Proposition 7.1, let $G_{N/4}$ be a maximal $N/4$-disjoint subset of G satisfying (2.7) and (2.8), and let $\mathbf{w}_n, n \ge 1$, be as in (7.13) and (7.14). Suppose that N satisfies (6.2) and*

$$r_1 := \frac{D_1(\mathcal{G})D_4L\alpha}{\alpha - d}\|\mathbf{S}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}(N+1)^{-\alpha+d} < 1. \tag{7.16}$$

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 $\mathbf{S}^T\mathbf{w}_n, n \ge 1$, converge exponentially to a bounded vector $\mathbf{w}$ in (7.15) and the "least squares" solution $\mathbf{d}_2$ in (7.1) respectively,

$$\|\mathbf{w}_n - \mathbf{w}\|_\infty \le D_6 r_1^n \|\mathbf{d}_2\|_\infty \tag{7.17}$$

and

$$\|\mathbf{S}^T\mathbf{w}_n - \mathbf{d}_2\|_\infty \le \frac{D_1(\mathcal{G})D_6L\alpha}{\alpha-d}\|\mathbf{S}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})} r_1^n \|\mathbf{d}_2\|_\infty,\ n \ge 1. \tag{7.18}$$

By the above theorem, each agent should have minimal storage, computing, and telecommunication capabilities. Furthermore, the algorithm (7.13) and (7.14) 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. [20] and simulations in Section 8.

The iterative algorithm (7.13) and (7.14) can be recast as follows:

$$\mathbf{w}_1 = \mathbf{R}_N\mathbf{S}^T\mathbf{z} \ \text{ and } \ \mathbf{e}_1 = \mathbf{w}_1 - \mathbf{R}_N\mathbf{S}^T\mathbf{S}\mathbf{S}^T\mathbf{w}_1, \tag{7.19}$$

and

$$\begin{cases} \mathbf{w}_{n+1} = \mathbf{w}_n + \mathbf{e}_n \\ \mathbf{e}_{n+1} = \mathbf{e}_n - \mathbf{R}_N\mathbf{S}^T\mathbf{S}\mathbf{S}^T\mathbf{e}_n,\ n \ge 1. \end{cases} \tag{7.20}$$

Next, we present a distributed implementation of the algorithm (7.19) and (7.20) when $\mathbf{S}$ has bandwidth s. Select a threshold ϵ and an integer $N \ge s$ satisfying (7.16). 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

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

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 (7.19) and (7.20) for an agent $\lambda \in G$.

Distributed algorithm (7.19) and (7.20) 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

We conclude this section by discussing the complexity of the distributed algorithm (7.19) and (7.20), 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 $\beta \ge 1$. By Theorem 7.2, 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.

8. Numerical simulations

In this section, we present two simulations to demonstrate the distributed algorithm (7.19) and (7.20) 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

$$\boldsymbol{\lambda}_l := \frac{R}{5}\Big(\cos\frac{2\pi\theta_l}{R}, \sin\frac{2\pi\theta_l}{R}\Big),\ 1 \le l \le R,$$

where $R \ge 1$ and $\theta_l \in l + [-1/4, 1/4]$ are randomly selected. Every agent in the SDN has a direct communication channel to its two adjacent agents. Then the graph $\mathcal{G}_c = (G_c, S_c)$ to describe the SDN is a circular graph, where $G_c = \{\boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_R\}$ and $S_c = \big\{(\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2), \ldots, (\boldsymbol{\lambda}_{R-1},\boldsymbol{\lambda}_R), (\boldsymbol{\lambda}_R,\boldsymbol{\lambda}_1), (\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_R), (\boldsymbol{\lambda}_R,\boldsymbol{\lambda}_{R-1}), \ldots, (\boldsymbol{\lambda}_2,\boldsymbol{\lambda}_1)\big\}$. Take innovative positions

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

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 $\mathbf{p}_i, 1 \le i \le R$, it has three anchor agents $\boldsymbol{\lambda}_i, \boldsymbol{\lambda}_{i-1}$ and $\boldsymbol{\lambda}_{i+1}$, where $\boldsymbol{\lambda}_0 = \boldsymbol{\lambda}_R$ and $\boldsymbol{\lambda}_{R+1} = \boldsymbol{\lambda}_1$. Set

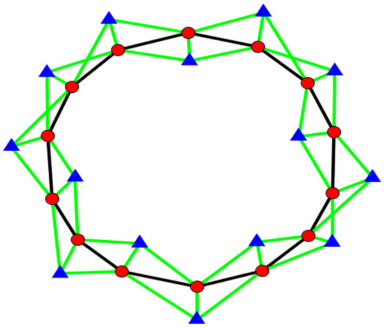

Fig. 3. The graph $\mathcal{H}_c = (G_c \cap V_c, S_c \cup T_c \cup T_c^*)$ to describe the DSRS in the first simulation, where vertices in G_c, edges in S_c, vertices in V_c and edges in $T_c \cup T_c^*$ are plotted in red circles, black lines, blue triangles and green lines, respectively.

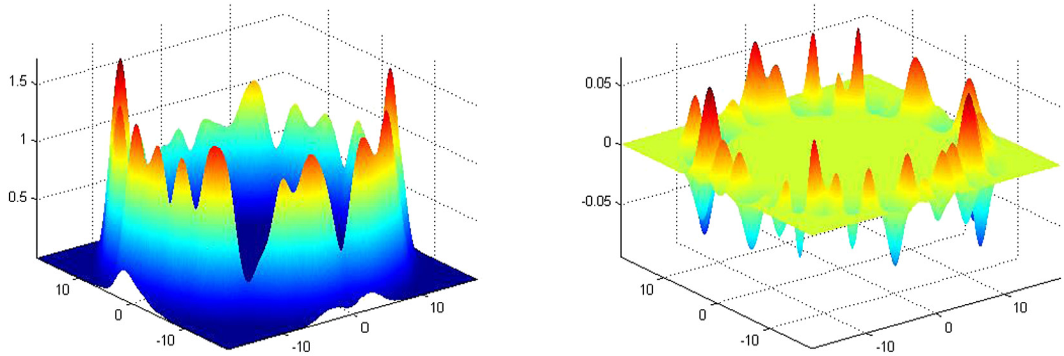

Fig. 4. Plotted on the left is the signal f in (8.1) with $R = 80$. On the right is the difference between the signal f and the reconstructed signal $f_{n,N,\delta}$ in the n-th iteration by applying the distributed algorithm (7.19) and (7.20) with $n = 10, N = 6$ and $\delta = 0.05$.

$$V_c = \{\mathbf{p}_i, 1 \le i \le R\} \ \text{ and } \ T_c = \{(\mathbf{p}_i, \boldsymbol{\lambda}_{i-j}), i = 1, \dots, R \text{ and } j = 0, \pm 1\}.$$

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 Fig. 3.

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

$$f(\mathbf{t}) = \sum_{i=1}^{R} c(i)\varphi(\mathbf{t} - \mathbf{p}_i) \tag{8.1}$$

to be sampled and reconstructed have their amplitude components $c(i) \in [0, 1]$ being randomly chosen, see the left image of Fig. 4. In the first simulation, we consider ideal sampling procedure. Thus for the agent $\boldsymbol{\lambda}_l, 1 \le l \le R$, the noisy sampling data acquired is

$$y_\delta(l) := f(\boldsymbol{\lambda}_l) + \eta(l) = \sum_{i=1}^{R} c(i)\varphi(\boldsymbol{\lambda}_l - \mathbf{p}_i) + \eta(l), \tag{8.2}$$

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

For $N \ge 5$, the complexity of the distributed algorithm (7.19) and (7.20) for each agent in $G_{N/4}$ is about $O(N)$. Our first simulation shows that the distributed algorithm (7.19) and (7.20) converges for $N \ge 5$ and the convergence rate is almost independent of the network size R, cf. the upper bound estimate in (7.18).

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 (7.19) and (7.20) from the noisy sampling data in (8.2). Define maximal reconstruction errors

Table 1
Maximal reconstruction errors $\epsilon(n, N, \delta)$ with $\delta = 0$.

$n \backslash N$	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

$$\epsilon(n, N, \delta) := \begin{cases} \max_{1\le i\le R} |c(i)| & \text{if } \ n = 0, \\ \max_{1\le i\le R} |c_{n,N,\delta}(i) - c(i)| & \text{if } \ n \ge 1. \end{cases} \tag{8.3}$$

Presented in Table 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 (7.19) and (7.20) has faster convergence rate for larger $N \ge 5$, and we only need three iterations to have a nearly perfect reconstruction from its noiseless samples when $N = 10$.

The robustness of the proposed algorithm (7.19) and (7.20) against sampling noises is tested and confirmed, see Fig. 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. Theorem 5.4 and Fig. 7.

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 SDN is a prism graph with vertices having physical locations,

$$\boldsymbol{\mu}_l := \begin{cases} \frac{R}{10}\big(\cos\frac{4\pi\theta_l}{R}, \sin\frac{4\pi\theta_l}{R}\big) & \text{if } \ 1 \le l \le \frac{R}{2} \\ \big(\frac{R}{10}+1\big)\big(\cos\frac{4\pi\theta_l}{R}, \sin\frac{4\pi\theta_l}{R}\big) & \text{if } \ \frac{R}{2}+1 \le l \le R, \end{cases} \tag{8.4}$$

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 $\boldsymbol{\mu}_i, \boldsymbol{\mu}_{i+1}, \boldsymbol{\mu}_{i+R/2}$ and $\boldsymbol{\mu}_{i+R/2+1}$, where $\boldsymbol{\mu}_0 = \boldsymbol{\mu}_{R/2}, \boldsymbol{\mu}_{R+1} = \boldsymbol{\mu}_{R/2+1}$, and $r_i \in \frac{R}{10}+[\frac{1}{4}, \frac{3}{4}]$ are randomly selected. Set

$$V_p = \Big\{\mathbf{q}_i, 1 \le i \le \frac{R}{2}\Big\} \ \text{ and } \ T_p = \Big\{(\mathbf{q}_i, \boldsymbol{\mu}_{i+j}), i = 1, \ldots, \frac{R}{2} \text{ and } j = 0, 1, \frac{R}{2}, \frac{R}{2}+1\Big\}.$$

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 Fig. 5.

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

$$g(\mathbf{t}) = \sum_{i=1}^{R/2} c(i)\varphi(\mathbf{t} - \mathbf{q}_i), \tag{8.5}$$

where $c(i) \in [0, 1], 1 \le i \le R/2$, are randomly selected, see the top left image of Fig. 6. Then the noisy sampling data acquired by the agent $\boldsymbol{\mu}_l, 1 \le l \le R$, is

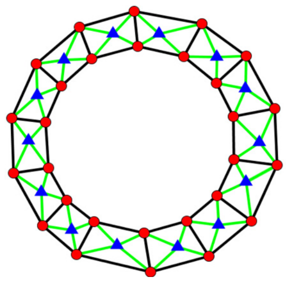

Fig. 5. Plotted on the left is the graph $\mathcal{H}_p = (G_p \cap V_p, S_p \cup T_p \cup T_p^*)$ to describe the DSRS, where vertices in G_p and V_p are in red circles and blue triangles, and edges in S_p and $T_p \cup T_p^*$ are in black solid lines and green solid lines, respectively. On the right 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.

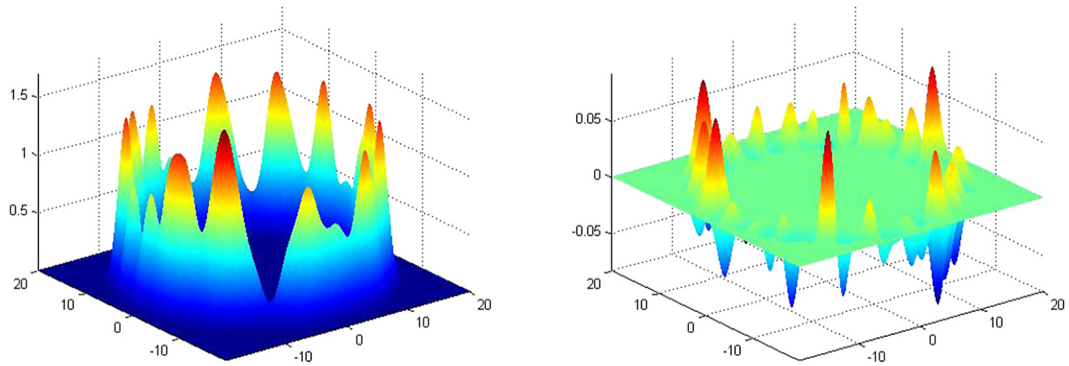

Fig. 6. Plotted on the left and right are the signal g in (8.5) with $R = 160$ and the difference $g - g_{n,N,\delta}$ between the original signal g and its approximation $g_{n,N,\delta}$ in (8.7) with $n = 4, N = 6$ and $\delta = 0.05$, where all agents in (8.4) are functional except those located at $\boldsymbol{\mu}_1, \boldsymbol{\mu}_{87}$ being completely dysfunctional and partial communication channels located at $\boldsymbol{\mu}_{11}, \boldsymbol{\mu}_{51}, \boldsymbol{\mu}_{91}$ clogged. The reconstruction error $\epsilon(n, N, \delta)$ in (8.3) in this simulation is 0.1802.

$$y_\delta(l) := g(\boldsymbol{\mu}_l) + \eta(l) = \sum_{i=1}^{R/2} c(i)\varphi(\boldsymbol{\mu}_l - \mathbf{q}_i) + \eta(l), \tag{8.6}$$

where $\eta(l) \in [-\delta, \delta]$ are randomly selected with bounded noise level $\delta \geq 0$. Applying the distributed algorithm (7.19) and (7.20), we obtain approximations

$$g_{n,N,\delta}(\mathbf{t}) = \sum_{i=1}^{R/2} c_{n,N,\delta}(i)\varphi(\mathbf{t} - \mathbf{q}_i),\ n \geq 1, \tag{8.7}$$

of the signal g in (8.5). Our simulations illustrate that the distributed algorithm (7.19) and (7.20) 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 1 in the first simulation. The robustness of the proposed distributed algorithm (7.19) and (7.20) against sampling noises in the DSRS is test and confirmed, see Fig. 7.

The robustness of the proposed distributed algorithm (7.19) and (7.20) against sampling noises and dysfunctions of agents in the DSRS is tested and confirmed, see the right graph of Fig. 5 and the right image of Fig. 6.

We finish this section with the performance comparison between the global ℓ^∞-optimization (1.13) and the proposed distributed algorithm (7.19) and (7.20) for signal reconstruction. For the signal sampling procedure in (8.5) and (8.6), we define the reconstruction error of the global optimization problem (1.13) by

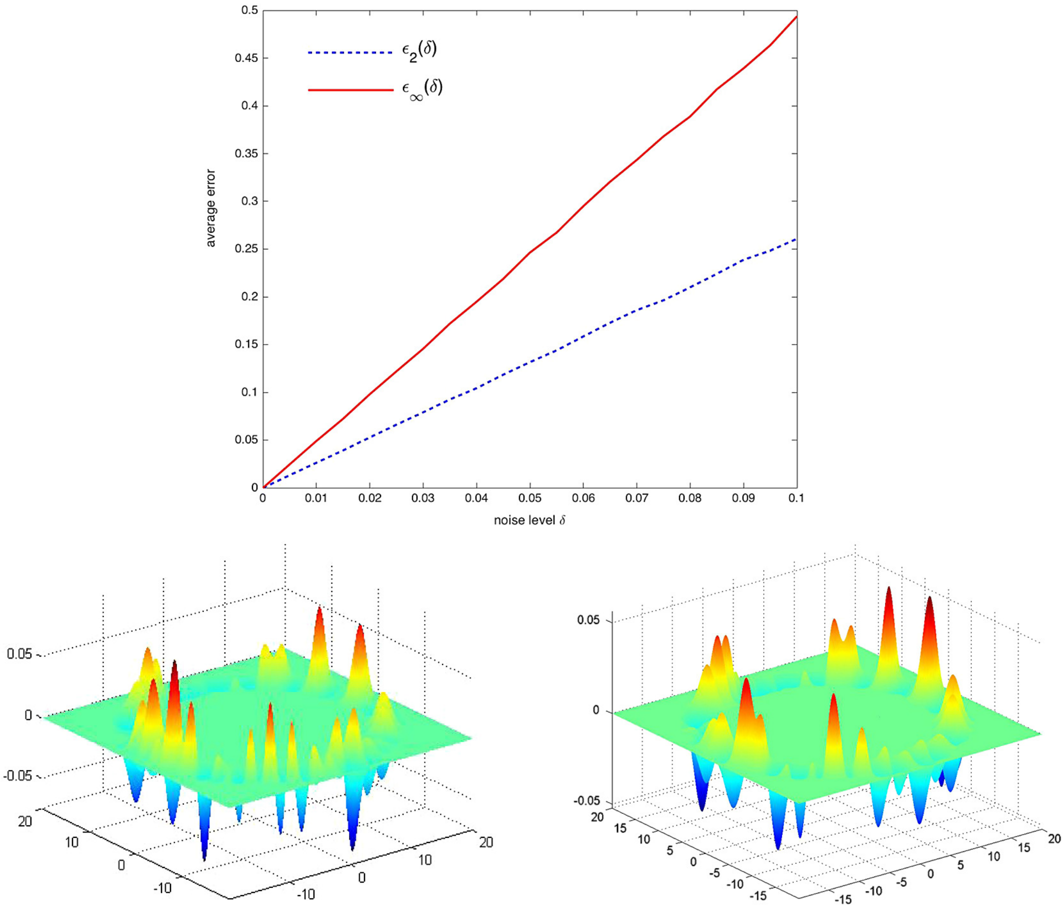

Fig. 7. Presented on the top is average of the reconstruction errors $\epsilon_\infty(\delta)$ and $\epsilon_2(\delta)$ of the global ℓ^∞-optimization (1.13) and the proposed distributed algorithm (7.19) and (7.20) over 1000 trials. Plotted on the bottom left and right are the difference between the original signal g in Fig. 6 and the signals reconstructed by the global algorithm and by the distributed algorithm with $N = 6$ respectively, where $\delta = 0.05$. In this simulation, the reconstruction errors $\epsilon(\delta)$ and $\epsilon_N(\delta)$ are 0.1933 and 0.1320 respectively.

$$\epsilon_\infty(\delta) := \max_{1\le i\le R/2} |c_\delta(i) - c(i)|,$$

where $\mathbf{c}_\delta := (c_\delta(1), \dots, c_\delta(R/2))^T$ is the reconstructed amplitude vector. Similarly for the same signal sampling procedure, our numerical simulation indicates that the proposed distributed algorithm (7.19) and (7.20) leads to the least square solution $(d_{2,\delta}(1), \dots, d_{2,\delta}(R/2))^T := \mathrm{argmin}_{\mathbf{d}\in\ell^2} \|\mathbf{Sd} - \mathbf{z}\|_2$ for all $N \ge 3$, cf. (7.18). Hence the corresponding reconstruction error

$$\epsilon_2(\delta) := \max_{1\le i\le R/2} |d_{2,\delta}(i) - c(i)|$$

is independent on $N \ge 3$, cf. Table 1. Our simulations, see Fig. 7, indicate that both the global optimization problem (1.13) and the proposed distributed algorithm (7.19) and (7.20) provide suboptimal approximation to the original signal in the presence of bounded noises, and the distributed algorithm has better performance for signal reconstruction than the global optimization does.

9. Proofs

In this section, we include proofs of Propositions 2.4, 3.1, 3.2, 3.3, 3.4, 4.1, 4.2, 7.1, and Theorems 5.2, 5.3, 5.4, 6.1, 6.2, 7.2.

9.1. Proof of Proposition 2.4

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,

$$G \subset \bigcup_{\lambda_m \in G_N} B_{\mathcal{G}}(\lambda_m, N'), \tag{9.1}$$

and hence the first inequality in (2.9) follows.

Now we prove the last inequality in (2.9). 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

$$B_{\mathcal{G}}(\lambda_m, N) \subset B_{\mathcal{G}}(\lambda_{m'}, 2N' + N). \tag{9.2}$$

Hence

$$\begin{aligned}\sum_{\lambda_m \in G_N} \chi_{B_{\mathcal{G}}(\lambda_m, N')}(\lambda) &\leq \frac{\mu_{\mathcal{G}}(\cup_{\lambda_m \in G_N \cap B_{\mathcal{G}}(\lambda, N')} B_{\mathcal{G}}(\lambda_m, N))}{\inf_{\lambda_m \in G_N \cap B_{\mathcal{G}}(\lambda, N')} \mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda_m, N))} \\ &\leq \sup_{\lambda_m \in G_N \cap B_{\mathcal{G}}(\lambda, N')} \frac{\mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda_m, 2N' + N))}{\mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda_m, N))} \leq (D_0(\mathcal{G}))^{\lceil \log_2(2N'/N+1) \rceil},\end{aligned} \tag{9.3}$$

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 (9.2), and the third inequality follows from the doubling assumption (2.1).

9.2. Proof of Proposition 3.1

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 (3.1). 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.

9.3. Proof of Proposition 3.2

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

$$\rho(i, i') \leq \rho(i, i'') + \rho(i'', i') \quad \text{for distinct vertices } i, i', i'' \in V. \tag{9.4}$$

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

$$\rho(i, i') \le m + n - 1 < \rho(i, i'') + \rho(i'', i'). \tag{9.5}$$

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 (3.1). 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

$$\rho(i, i') \le m + n = \rho(i, i'') + \rho(i'', i'). \tag{9.6}$$

Combining (9.5) and (9.6) proves (9.4).

9.4. Proof of Proposition 3.3

To prove Proposition 3.3, we need two lemmas comparing measures of balls in graphs $\mathcal{G}$ and $\mathcal{V}$.

Lemma 9.1. *If $\mathcal{H}$ satisfies (3.1) and (3.2), then*

$$\mu(B(i, r)) \le L\mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda, r)) \quad \text{for any } \lambda \in G \text{ with } (i, \lambda) \in T. \tag{9.7}$$

Proof. Let $i' \in B(i, r)$ with $i' \neq i$. By Proposition 3.1, there exists a path $\lambda_1 \dots \lambda_n$ of length $\rho(i, i') - 1$ in the graph $\mathcal{G}$ such that $(i, \lambda_1), (i', \lambda_n) \in T$. Then

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

as either $\lambda_1 = \lambda$ or (λ, λ_1) is an edge in $\mathcal{G}$ by (3.1). 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 (3.2) proves (9.7). □

Lemma 9.2. *If $\mathcal{H}$ satisfies (2.3), (3.1) and (3.3), then*

$$\mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda, r)) \le \Big(\sup_{\lambda' \in G} \mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda', 2M+2))\Big)\mu(B(i, r+M+1)) \tag{9.8}$$

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

Proof. Let $\lambda_1 = \lambda$ and take $\Lambda = \{\lambda_m\}_{m \ge 1}$ 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 2.4, $\{B_{\mathcal{G}}(\lambda_m, 2(M+1))\}_{\lambda_m \in \Lambda}$ forms a covering of the ball $B(\lambda, r)$, which implies that

$$\mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda, r)) \le \Big(\sup_{\lambda_m \in \Lambda} \mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda_m, 2M+2))\Big)\#\Lambda \le \Big(\sup_{\lambda' \in G} \mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda', 2M+2))\Big)\#\Lambda. \tag{9.9}$$

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 (3.3) that

$$\#V_{\lambda_m} \ge 1 \text{ for all } \lambda_m \in \Lambda. \tag{9.10}$$

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 (3.1) implies that

$$V_{\lambda_m} \cap V_{\lambda_{m'}} = \emptyset \text{ for distinct } \lambda_m, \lambda_{m'} \in \Lambda. \tag{9.11}$$

Combining (9.9), (9.10) and (9.11) leads to

$$\mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda, r)) \le \Big(\sup_{\lambda' \in G} \mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda', 2M+2))\Big) \#\Big(\cup_{\lambda_m \in \Lambda} V_{\lambda_m}\Big). \tag{9.12}$$

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

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

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

$$\rho(i, i') \le r + M + 1. \tag{9.13}$$

Then the desired estimate (9.8) follows from (9.12) and (9.13). □

We are ready to prove Proposition 3.3.

Proof of Proposition 3.3. First we prove the doubling property (3.10) for the measure μ. Take $i \in V$. Then for $r \ge 2(M+1)$ it follows from Lemmas 9.1 and 9.2 that

$$\begin{aligned} \mu(B(i, 2r)) &\le L\mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda, 2r)) \le L(D_0(\mathcal{G}))^2 \mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda, r/2)) \\ &\le KL(D_0(\mathcal{G}))^2 \mu(B(i, r/2 + M + 1)) \le KL(D_0(\mathcal{G}))^2 \mu(B(i, r)), \end{aligned} \tag{9.14}$$

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

$$K := \sup_{\lambda' \in G} \mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda', 2M+2)) \le \frac{((\deg(\mathcal{G}))^{2M+3} - 1}{\deg(\mathcal{G}) - 1} \tag{9.15}$$

by (2.6). From the doubling property (2.1) for the measure $\mu_{\mathcal{G}}$, we obtain

$$\mu(B(i, 2r)) \le KLD_0(\mathcal{G}) \le KLD_0(\mathcal{G})\mu(B(i, r)) \text{ for } 0 \le r \le 2(M+1). \tag{9.16}$$

Then the doubling property (3.10) follows from (9.14), (9.15) and (9.16).

Next we prove the doubling property (3.11) 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 (3.3). From Lemmas 9.1 and 9.2, we obtain

$$\begin{aligned} \mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda, 2r)) &\le K\mu(B(i, 2r + M + 1)) \le D_0^2 K\mu\Big(B\Big(i, \frac{r}{2} + \frac{(M+1)}{4}\Big)\Big) \\ &\le D_0^2 LK\mu_{\mathcal{G}}\Big(B_{\mathcal{G}}\Big(\lambda', \frac{r}{2} + \frac{M+1}{4}\Big)\Big) \\ &\le D_0^2 LK\mu_{\mathcal{G}}\Big(B_{\mathcal{G}}\Big(\lambda, \frac{r}{2} + \frac{M+1}{4} + M\Big)\Big) \le D_0^2 LK\mu_{\mathcal{G}}(B_{\mathcal{G}}(\lambda, r)) \end{aligned} \tag{9.17}$$

for $r \ge 3M$, and

$$\begin{aligned}\mu_{\mathcal G}(B_{\mathcal G}(\lambda,2r)) &\le K\mu(B(i,7M)) \le D_0^2K\mu(B(i,2M))\\ &\le D_0^2LK\mu_{\mathcal G}(B_{\mathcal G}(\lambda',2M)) \le D_0^2LK^2\mu_{\mathcal G}(B_{\mathcal G}(\lambda,r))\end{aligned} \tag{9.18}$$

for $0 \le r \le 3M-1$. Combining (9.15), (9.17) and (9.18) proves (3.11). □

9.5. Proof of Proposition 3.4

The polynomial growth property (3.12) for the measure μ follows immediately from Lemma 9.1. The polynomial growth property (3.13) 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 (2.6), and

$$\begin{aligned}\mu_{\mathcal G}(B_{\mathcal G}(\lambda,r)) &\le D_1\Big(\frac{(\deg(\mathcal G))^{2M+3}-1}{\deg(\mathcal G)-1}\Big)(r+M+2)^d\\ &\le 2^dD_1\Big(\frac{(\deg(\mathcal G))^{2M+3}-1}{\deg(\mathcal G)-1}\Big)(r+1)^d,\ r\ge M,\end{aligned}$$

by (9.15) and Lemma 9.2.

9.6. Proof of Proposition 4.1

To prove Proposition 4.1, we need a lemma.

Lemma 9.3. *Let $\mathcal G$ be a connected simple graph. If its counting measure has polynomial growth (2.4), then*

$$\sup_{\lambda\in G}\sum_{\rho_{\mathcal G}(\lambda,\lambda')\ge s}(1+\rho_{\mathcal G}(\lambda,\lambda'))^{-\alpha} \le \frac{D_1(\mathcal G)\alpha}{\alpha-d}(s+1)^{-\alpha+d} \tag{9.19}$$

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

$$\begin{aligned}\sum_{\rho_{\mathcal G}(\lambda,\lambda')\ge s}(1+\rho_{\mathcal G}(\lambda,\lambda'))^{-\alpha} &= \sum_{n\ge s}(n+1)^{-\alpha}\Big(\sum_{\rho_{\mathcal G}(\lambda,\lambda')=n}1\Big)\\ &\le \sum_{n\ge s}\mu_{\mathcal G}(B_{\mathcal G}(\lambda,n))((n+1)^{-\alpha}-(n+2)^{-\alpha})\\ &\le D_1(\mathcal G)\sum_{n=s}^{\infty}(n+1)^d((n+1)^{-\alpha}-(n+2)^{-\alpha})\\ &= D_1(\mathcal G)\Big((s+1)^{-\alpha+d}+\sum_{n=s+1}^{\infty}(n+1)^{-\alpha}\big((n+1)^d-n^d\big)\Big)\\ &\le D_1(\mathcal G)\Big((s+1)^{-\alpha+d}+d\int_{s+1}^{\infty}t^{d-\alpha-1}dt\Big) = \frac{D_1(\mathcal G)\alpha}{\alpha-d}(s+1)^{-\alpha+d},\end{aligned} \tag{9.20}$$

where the second inequality follows from (2.4), and the third one is true as $(n+1)^d - n^d \le d(n+1)^{d-1}$ for $n \ge 1$ and $d \ge 1$. □

Now we prove Proposition 4.1.

Proof of Proposition 4.1. 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

$$\begin{aligned}\|\mathbf{Ac}\|_p^p &\le \|\mathbf{A}\|^p_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})} \sum_{\lambda\in G}\Big(\sum_{i\in V}(1+\rho_{\mathcal{H}}(\lambda,i))^{-\alpha}|c(i)|\Big)^p \\ &\le \|\mathbf{A}\|^p_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}\|\mathbf{c}\|_p^p\Big(\sup_{\lambda'\in G}\sum_{i'\in V}(1+\rho_{\mathcal{H}}(\lambda',i'))^{-\alpha}\Big)^{p-1}\Big(\sup_{i'\in V}\sum_{\lambda'\in G}(1+\rho_{\mathcal{H}}(\lambda',i'))^{-\alpha}\Big). \qquad (9.21)\end{aligned}$$

For any $\lambda' \in G$ and $i' \in V$, it follows from Proposition 3.1 that

$$\rho_{\mathcal{G}}(\lambda',\lambda'')+1 \ge \rho_{\mathcal{H}}(\lambda',i') \ge \rho_{\mathcal{G}}(\lambda',\lambda'') \ \text{ for all } \lambda'' \in G \text{ with } (i',\lambda'') \in T. \qquad (9.22)$$

By (3.2), (3.14), (9.22) and Lemma 9.3, we obtain

$$\begin{aligned}\sum_{i'\in V}(1+\rho_{\mathcal{H}}(\lambda',i'))^{-\alpha} &\le \sum_{\lambda''\in G}\Big(\sum_{(i',\lambda'')\in T} 1\Big)(1+\rho_{\mathcal{G}}(\lambda',\lambda''))^{-\alpha} \\ &\le L\sum_{\lambda''\in G}(1+\rho_{\mathcal{G}}(\lambda',\lambda''))^{-\alpha} \le \frac{LD_1(\mathcal{G})\alpha}{\alpha-d} \ \text{ for any } \lambda' \in G, \qquad (9.23)\end{aligned}$$

and

$$\sum_{\lambda'\in G}(1+\rho_{\mathcal{H}}(\lambda',i'))^{-\alpha} \le \sum_{\lambda'\in G}(1+\rho_{\mathcal{G}}(\lambda',\lambda''))^{-\alpha} \le \frac{D_1(\mathcal{G})\alpha}{\alpha-d} \ \text{ for any } i' \in V, \qquad (9.24)$$

where $\lambda'' \in G$ satisfies $(i', \lambda'') \in T$. Combining (9.21), (9.23) and (9.24) proves (4.4) for $1 < p < \infty$.
We can use similar argument to prove (4.4) for $p = 1, \infty$. □

9.7. Proof of Proposition 4.2

Following the proof of Proposition 4.1, we obtain

$$\begin{aligned}\|(\mathbf{A}-\mathbf{A}_s)\mathbf{c}\|_p &\le \|\mathbf{A}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}\Big(\sup_{\lambda'\in G}\sum_{\rho_{\mathcal{H}}(\lambda',i')>s}(1+\rho_{\mathcal{H}}(\lambda',i'))^{-\alpha}\Big)^{1-1/p} \\ &\quad\times\Big(\sup_{i'\in V}\sum_{\rho_{\mathcal{H}}(\lambda',i')>s}(1+\rho_{\mathcal{H}}(\lambda',i'))^{-\alpha}\Big)^{1/p}\|\mathbf{c}\|_p, \qquad (9.25)\end{aligned}$$

where $\mathbf{c} \in \ell^p, 1 \le p \le \infty$. Applying similar argument used to prove (9.19), (9.23) and (9.24), we have

$$\sup_{\lambda'\in G}\sum_{\rho_{\mathcal{H}}(\lambda',i')>s}(1+\rho_{\mathcal{H}}(\lambda',i'))^{-\alpha} \le L\sup_{\lambda'\in G}\sum_{\rho_{\mathcal{G}}(\lambda',\lambda'')\ge s}(1+\rho_{\mathcal{G}}(\lambda',\lambda''))^{-\alpha} \le \frac{D_1(\mathcal{G})L\alpha}{\alpha-d}(s+1)^{-\alpha+d} \qquad (9.26)$$

and

$$\sup_{i'\in V}\sum_{\rho_{\mathcal{H}}(\lambda',i')>s}(1+\rho_{\mathcal{H}}(\lambda',i'))^{-\alpha}\le\frac{D_1(\mathcal{G})\alpha}{\alpha-d}(s+1)^{-\alpha+d}. \tag{9.27}$$

Then the approximation error estimate (4.6) follows from (9.25), (9.26) and (9.27).

9.8. Proof of Theorem 5.3

To prove Wiener's lemma (Theorem 5.3) for $\mathcal{J}_\alpha(\mathcal{V}), \alpha>d$, we first show that it is a Banach algebra of matrices.

Proposition 9.4. *Let $\mathcal{V}$ be an undirected graph with the counting measure μ having polynomial growth (3.9). 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) $\|\mathbf{C}+\mathbf{D}\|_{\mathcal{J}_\alpha(\mathcal{V})}\le\|\mathbf{C}\|_{\mathcal{J}_\alpha(\mathcal{V})}+\|\mathbf{D}\|_{\mathcal{J}_\alpha(\mathcal{V})}$,
(iii) $\|\mathbf{CD}\|_{\mathcal{J}_\alpha(\mathcal{V})}\le\frac{2^{\alpha+1}D_1\alpha}{\alpha-d}\|\mathbf{C}\|_{\mathcal{J}_\alpha(\mathcal{V})}\|\mathbf{D}\|_{\mathcal{J}_\alpha(\mathcal{V})}$, *and*
(iv) $\|\mathbf{Dc}\|_2\le\frac{D_1\alpha}{\alpha-d}\|\mathbf{D}\|_{\mathcal{J}_\alpha(\mathcal{V})}\|\mathbf{c}\|_2$

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 (5.2) and (5.3).

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

$$\|\mathbf{CD}\|_{\mathcal{J}_\alpha(\mathcal{V})}\le 2^\alpha\|\mathbf{C}\|_{\mathcal{J}_\alpha(\mathcal{V})}\|\mathbf{D}\|_{\mathcal{J}_\alpha(\mathcal{V})}\sup_{i,i'\in V}\Big(\sum_{\rho(i,i'')\ge\rho(i,i')/2}(1+\rho(i'',i'))^{-\alpha}$$
$$+\sum_{\rho(i'',i')\ge\rho(i,i')/2}(1+\rho(i,i''))^{-\alpha}\Big). \tag{9.28}$$

Following the argument used in the proof of Lemma 9.3, we have

$$\sup_{i\in V}\sum_{\rho(i,i')\ge s}(1+\rho(i,i'))^{-\alpha}\le\frac{D_1\alpha}{\alpha-d}(s+1)^{-\alpha+d},\ 0\le s\in\mathbb{Z}. \tag{9.29}$$

Combining (9.28) and (9.29) proves the third conclusion.

Following the proof of Proposition 4.1 and applying (9.29) instead of (9.23) and (9.24), we obtain the fourth conclusion. □

Now, we prove Theorem 5.3.

Proof of Theorem 5.3. Following the argument in [59], it suffices to establish the following differential norm inequality:

$$\|\mathbf{C}^2\|_{\mathcal{J}_\alpha(\mathcal{V})}\le 2^{\alpha+d/2+2}D_1^{1/2}(D_1\alpha/(\alpha-d))^{1-\theta}(\|\mathbf{C}\|_{\mathcal{J}_\alpha(\mathcal{V})})^{2-\theta}(\|\mathbf{C}\|_{\mathcal{B}^2})^\theta \tag{9.30}$$

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

$$\begin{aligned}\|\mathbf{C}^2\|_{\mathcal{J}_\alpha(\mathcal{V})} &\le 2^\alpha \|\mathbf{C}\|_{\mathcal{J}_\alpha(\mathcal{V})} \Big(\sup_{i,i' \in V} \sum_{\rho(i,i'') \ge \rho(i,i')/2} |c(i'',i')| + \sup_{i,i' \in V} \sum_{\rho(i'',i') \ge \rho(i,i')/2} |c(i,i'')| \Big) \\ &\le 2^\alpha \|\mathbf{C}\|_{\mathcal{J}_\alpha(\mathcal{V})} \Big(\sup_{i' \in V} \sum_{i'' \in V} |c(i'',i')| + \sup_{i \in V} \sum_{i'' \in V} |c(i,i'')| \Big). \end{aligned} \tag{9.31}$$

Set

$$\tau := \Big(\frac{D_1 \alpha \|\mathbf{C}\|_{\mathcal{J}_\alpha(\mathcal{V})}}{(\alpha - d)\|\mathbf{C}\|_{\mathcal{B}^2}} \Big)^{2/(2\alpha - d)} \ge 1 \tag{9.32}$$

by Proposition 9.4. For $i' \in V$, we obtain

$$\begin{aligned}\sum_{i'' \in V} |c(i'',i')| &\le \Big(\sum_{\rho(i'',i') \le \tau} |c(i'',i')|^2 \Big)^{1/2} \Big(\sum_{\rho(i'',i') \le \tau} 1 \Big)^{1/2} + \|\mathbf{C}\|_{\mathcal{J}_\alpha(\mathcal{V})} \sum_{\rho(i'',i') > \tau} (1 + \rho(i'',i'))^{-\alpha} \\ &\le D_1^{1/2} \|\mathbf{C}\|_{\mathcal{B}^2} (1 + \lfloor \tau \rfloor)^{d/2} + D_1 \alpha (\alpha - d)^{-1} \|\mathbf{C}\|_{\mathcal{J}_\alpha(\mathcal{V})} (1 + \lfloor \tau \rfloor)^{-\alpha + d} \\ &\le 2^{d/2+1} D_1^{1/2} (D_1 \alpha/(\alpha - d))^{d/(2\alpha - d)} (\|\mathbf{C}\|_{\mathcal{J}_\alpha(\mathcal{V})})^{1-\theta} (\|\mathbf{C}\|_{\mathcal{B}^2})^\theta, \end{aligned} \tag{9.33}$$

where the second inequality holds by (9.29) and the last inequality follows from (9.32). Similarly, for $i \in V$ we have

$$\sum_{i'' \in V} |c(i',i'')| \le 2^{d/2+1} D_1^{1/2} (D_1 \alpha/(\alpha - d))^{d/(2\alpha - d)} (\|\mathbf{C}\|_{\mathcal{J}_\alpha(\mathcal{V})})^{1-\theta} (\|\mathbf{C}\|_{\mathcal{B}^2})^\theta. \tag{9.34}$$

Combining (9.31), (9.33) and (9.34) proves (9.30). This completes the proof of Theorem 5.3. □

9.9. Proof of Theorem 5.2

To prove Theorem 5.2, we need Theorem 5.3 and the following lemma about families $\mathcal{J}_\alpha(\mathcal{G}, \mathcal{V})$ and $\mathcal{J}_\alpha(\mathcal{V})$ of matrices.

Lemma 9.5. *Let $\mathcal{G}, \mathcal{H}, \mathcal{V}$ and d be as in Proposition 5.1. Then*

(i) $\|\mathbf{AC}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})} \le \frac{2^{\alpha+1} L D_1(\mathcal{G})\alpha}{\alpha - d} \|\mathbf{A}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})} \|\mathbf{C}\|_{\mathcal{J}_\alpha(\mathcal{V})}$ *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 (3.1) 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 9.4, 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 (3.15), (9.26) and (9.29) proves the first conclusion.

Recall that

$$\rho(i,i') \le \rho_{\mathcal{H}}(\lambda,i) + \rho_{\mathcal{H}}(\lambda,i') \text{ for all } \lambda \in G \text{ and } i,i' \in V. \tag{9.35}$$

Then for $\mathbf{A}, \mathbf{B} \in \mathcal{J}_\alpha(\mathcal{G},\mathcal{V})$, we obtain from (9.27) and (9.35) 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 5.2.

Proof of Theorem 5.2. 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 9.5. 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 5.3 and Lemma 9.5. 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 4.1 and the dual property between sequences ℓ^p and $\ell^{p/(p-1)}$. The ℓ^p-stability for the matrix $\mathbf{A}$ then follows. □

9.10. *Proof of Theorem 5.4*

The conclusion (5.4) follows immediately from Proposition 4.1, Theorem 5.3 and Lemma 9.5.

9.11. *Proof of Theorem 6.1*

Observe from Proposition 3.1 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 $\mathbf{A}\mathbf{c} = (d(\lambda'))_{\lambda'\in G}$. Then

$$\|\mathbf{A}\mathbf{c}\|_2 \ge A\|\mathbf{A}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}\|\mathbf{c}\|_2 \tag{9.36}$$

and

$$\sum_{\rho_{\mathcal{H}}(\lambda',\lambda)>2N} |d(\lambda')|^2 \le LD_1(\mathcal{G})N^{-\alpha+d}\|\mathbf{A}\|^2_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}$$

$$\times \sum_{\rho_{\mathcal{H}}(\lambda',\lambda)>2N}\ \sum_{i\in B_{\mathcal{H}}(\lambda,N)\cap V} (1+\rho_{\mathcal{H}}(\lambda',i))^{-\alpha}|c(i)|^2$$

$$\le \big(D_1(\mathcal{G})\big)^2 LN^{-2\alpha+2d}\alpha(\alpha-d)^{-1}\|\mathbf{A}\|^2_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}\|\mathbf{c}\|_2^2, \tag{9.37}$$

where the first inequality holds as

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

for all $\lambda' \notin B_{\mathcal{H}}(\lambda,2N)$ and $i' \in B_{\mathcal{H}}(\lambda,N)$, and the last inequality follows from (9.27). Combining (9.36) and (9.37) proves (6.1).

9.12. *Proof of Theorem 6.2*

In this subsection, we will prove the following strong version of Theorem 6.2.

Theorem 9.6. *Let $\mathcal{G},\mathcal{H},\mathcal{V}$ and $\mathbf{A}$ be as in Theorem 6.2. If there exists a positive constant A_0, an integer $N_0\ge 3$, and a maximal $\frac{N_0}{4}$-disjoint subset $G_{N_0/4}$ such that (6.3) is true and (6.4) hold for all $\lambda_m\in G_{N_0/4}$, then $\mathbf{A}$ satisfies (6.5).*

Proof. Let ψ_0 be the trapezoid function,

$$\psi_0(t) = \begin{cases} 1 & \text{if } |t|\le 1/2 \\ 2-2|t| & \text{if } 1/2<|t|\le 1 \\ 0 & \text{if } |t|>1. \end{cases} \tag{9.38}$$

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

$$\Psi^N_{\lambda,V}:\ (c(i))_{i\in V} \longmapsto \big(\psi_0(\rho_{\mathcal{H}}(\lambda,i)/N)c(i)\big)_{i\in V}, \tag{9.39}$$

$$\Psi^N_{\lambda,G}:\ (d(\lambda'))_{\lambda'\in G} \longmapsto \big(\psi_0(\rho_{\mathcal{H}}(\lambda,\lambda')/N)d(\lambda')\big)_{\lambda'\in G}. \tag{9.40}$$

Observe that

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

where $\mathbf{A}_N$ is a band approximation of the matrix $\mathbf{A}$ in (4.5). Then for all $\lambda_m\in G_{N_0/4}$, it follows from Proposition 4.2 and our local stability assumption (6.4) that

$$\|\mathbf{A}_{N_0}\Psi^{N_0}_{\lambda_m,V}\mathbf{c}\|_2 \ge \|\chi^{2N_0}_{\lambda_m,G}\mathbf{A}\chi^{N_0}_{\lambda_m,V}\Psi^{N_0}_{\lambda_m,V}\mathbf{c}\|_2 - \|\chi^{2N_0}_{\lambda_m,G}(\mathbf{A}-\mathbf{A}_{N_0})\Psi^{N_0}_{\lambda_m,V}\mathbf{c}\|_2$$

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

Therefore

$$\Big(\sum_{\lambda_m\in G_{N_0/4}} \|\mathbf{A}_{N_0}\Psi^{N_0}_{\lambda_m,V}\mathbf{c}\|_2^2\Big)^{1/2}$$

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

$$\geq \Big(\frac{A_0}{3} - \frac{D_1(\mathcal{G})L\alpha}{3(\alpha-d)} N_0^{-\alpha+d}\Big)\|\mathbf{A}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}\|\mathbf{c}\|_2, \tag{9.41}$$

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 (9.38), Proposition 2.4 and the assumption that $N_0 \geq 3$.

Next, we estimate commutators

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

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

$$\sum_{\lambda_m\in G_{N_0/4}} \|(\mathbf{A}_{N_0}\Psi^{N_0}_{\lambda_m,V} - \Psi^{N_0}_{\lambda_m,G}\mathbf{A}_{N_0})\mathbf{c}\|_2^2$$

$$\leq \|\mathbf{A}\|^2_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})} \sum_{\lambda_m\in G_{N_0/4}} \sum_{\lambda\in G} \Big\{ \sum_{\rho_{\mathcal{H}}(\lambda,i)\leq N_0} (1+\rho_{\mathcal{H}}(\lambda,i))^{-\alpha}$$

$$\times \Big|\psi_0\Big(\frac{\rho_{\mathcal{H}}(\lambda,\lambda_m)}{N_0}\Big) - \psi_0\Big(\frac{\rho_{\mathcal{H}}(i,\lambda_m)}{N_0}\Big)\Big|\chi_{B_{\mathcal{H}}(\lambda_m,2N_0)\cap V}(i)|c(i)|\Big\}^2$$

$$\leq 4(D_0(\mathcal{G}))^4 N_0^{-2}\|\mathbf{A}\|^2_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}\Big(\sup_{i\in V}\sum_{\lambda\in B_{\mathcal{H}}(i,N_0)\cap G} (1+\rho_{\mathcal{H}}(\lambda,i))^{-\alpha}\rho_{\mathcal{H}}(\lambda,i)\Big)$$

$$\times\Big(\sup_{\lambda\in G}\sum_{i\in B_{\mathcal{H}}(\lambda,N_0)\cap V} (1+\rho_{\mathcal{H}}(\lambda,i))^{-\alpha}\rho_{\mathcal{H}}(\lambda,i)\Big)\|\mathbf{c}\|_2^2, \tag{9.42}$$

where the last inequality follows from Propositions 2.4 and 3.1, and

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

Following the argument used in (9.19), we have

$$\sup_{i\in V}\sum_{\lambda\in B_{\mathcal{H}}(i,N_0)\cap G} (1+\rho_{\mathcal{H}}(\lambda,i))^{-\alpha}\rho_{\mathcal{H}}(\lambda,i)$$

$$\leq \sup_{\lambda'\in G}\sum_{\rho_{\mathcal{G}}(\lambda,\lambda')\leq N_0} (1+\rho_{\mathcal{G}}(\lambda,\lambda'))^{-\alpha+1}$$

$$\leq D_1(\mathcal{G})(N_0+1)^{-\alpha+d+1} + (\alpha-1)D_1(\mathcal{G})\sum_{n=0}^{N_0-1}(n+1)^{-\alpha+d}$$

$$\leq D_1(\mathcal{G})(N_0+1)^{-\alpha+d+1} + D_1(\mathcal{G})(\alpha-1)\Big(1+\int_1^{N_0} t^{-\alpha+d}dt\Big)$$

$$\leq \begin{cases} \frac{D_1(\mathcal{G})(\alpha-1)(\alpha-d)}{\alpha-d-1} & \text{if } \ \alpha > d+1 \\ D_1(\mathcal{G})(1+d+d\ln N_0) & \text{if } \ \alpha = d+1 \\ \frac{2^{d+1-\alpha}D_1(\mathcal{G})d}{d+1-\alpha}N_0^{d+1-\alpha} & \text{if } \ \alpha < d+1, \end{cases} \tag{9.43}$$

and

$$\begin{aligned}&\sup_{\lambda\in G}\sum_{i\in B_{\mathcal H}(\lambda,N_0)\cap V}(1+\rho_{\mathcal H}(\lambda,i))^{-\alpha}\rho_{\mathcal H}(\lambda,i)\\ &\le L\sup_{\lambda\in G}\sum_{\lambda'\in B_{\mathcal G}(\lambda,N_0)}(1+\rho_{\mathcal G}(\lambda,\lambda'))^{-\alpha+1}\\ &\le\begin{cases}\frac{D_1(\mathcal G)L(\alpha-1)(\alpha-d)}{\alpha-d-1} & \text{if } \alpha>d+1\\ D_1(\mathcal G)L(1+d+d\ln N_0) & \text{if } \alpha=d+1\\ \frac{2^{d+1-\alpha}D_1(\mathcal G)dL}{d+1-\alpha}N_0^{d+1-\alpha} & \text{if } \alpha<d+1.\end{cases}\end{aligned}\tag{9.44}$$

Therefore,

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

where the first inequality holds by Proposition 2.4, and the third inequality follows from (9.41) and (9.42). This together with Proposition 4.2 completes the proof. □

9.13. Proof of Proposition 7.1

To prove Proposition 7.1, we need the following critical estimate.

Proposition 9.7. *Let $\mathcal G$, $\mathcal H$, $\mathcal V$ and $\mathbf S$ be as in Proposition 7.1. Then*

$$\|(\chi^N_{\lambda,V}\mathbf S^T\mathbf S\chi^N_{\lambda,V})^{-1}\|_{\mathcal J_\alpha(\mathcal V)}\le\frac{2^{-\alpha-1}(\alpha-d)^2D_2}{\alpha^2D_1D_1(\mathcal G)\|\mathbf S\|^2_{\mathcal J_\alpha(\mathcal G,\mathcal V)}},\tag{9.45}$$

where D_2 is the constant in (7.8).

Proof. Let $\mathbf J_{\lambda,N}:=\chi^N_{\lambda,V}\mathbf S^T\mathbf S\chi^N_{\lambda,V}$. By Lemma 9.5, we have

$$\|\mathbf J_{\lambda,N}\|_{\mathcal J_\alpha(\mathcal V)}\le\frac{2^{\alpha+1}D_1(\mathcal G)\alpha}{\alpha-d}\|\mathbf S\|^2_{\mathcal J_\alpha(\mathcal G,\mathcal V)}.\tag{9.46}$$

This together with Propositions 9.4 implies that

$$A^2\|\mathbf S\|^2_{\mathcal J_\alpha(\mathcal G,\mathcal V)}\|\chi^N_{\lambda,V}\mathbf x\|_2^2\le\|\mathbf S\chi^N_{\lambda,V}\mathbf x\|_2^2=\langle\mathbf J_{\lambda,N}\mathbf x,\mathbf x\rangle\le\frac{2^{\alpha+1}\alpha^2D_1D_1(\mathcal G)}{(\alpha-d)^2}\|\mathbf S\|^2_{\mathcal J_\alpha(\mathcal G,\mathcal V)}\|\chi^N_{\lambda,V}\mathbf x\|_2^2$$

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

$$\mathbf{J}_{\lambda,N} = \frac{2^{\alpha+1}\alpha^2 D_1 D_1(\mathcal{G})}{(\alpha-d)^2}\|\mathbf{S}\|^2_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}(\mathbf{I}_{B_\mathcal{H}(\lambda,N)\cap V} - \mathbf{B}_{\lambda,N}) \tag{9.47}$$

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

$$\|\mathbf{B}_{\lambda,N}\|_{\mathcal{B}^2} \le r_0 \tag{9.48}$$

and

$$\|\mathbf{B}_{\lambda,N}\|_{\mathcal{J}_\alpha(\mathcal{V})} \le \|\mathbf{I}_{B_\mathcal{H}(\lambda,N)\cap V}\|_{\mathcal{J}_\alpha(\mathcal{V})} + \frac{2^{-\alpha-1}(\alpha-d)^2\|\mathbf{J}_{\lambda,N}\|_{\mathcal{J}_\alpha(\mathcal{V})}}{\alpha^2 D_1 D_1(\mathcal{G})\|\mathbf{S}\|^2_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}} \le 1 + \frac{\alpha-d}{\alpha D_1} \le 2, \tag{9.49}$$

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 [59] and applying (9.30) with $\mathbf{C}$ replaced by $\mathbf{B}_{\lambda,N}$ and V by $B_\mathcal{H}(\lambda,N)\cap V$, we obtain the following estimate

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

where $D = 2^{2\alpha+d/2+3}D_1^{1/2}(D_1\alpha/(\alpha-d))^{2-\theta}$. This together with (9.48) and (9.49) leads to

$$\|(\mathbf{B}_{k,N})^n\|_{\mathcal{J}_\alpha(\mathcal{V})} \le (2D^{\frac{1}{1-\theta}}/r_0)^{\frac{2-\theta}{1-\theta}n^{\log_2(2-\theta)}} r_0^n \quad \text{for all} \quad n\ge 1. \tag{9.50}$$

Observe that

$$\|(\mathbf{J}_{\lambda,N})^{-1}\|_{\mathcal{J}_\alpha(\mathcal{V})} \le \frac{2^{-\alpha-1}(\alpha-d)^2}{\alpha^2 D_1 D_1(\mathcal{G})\|\mathbf{S}\|^2_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}}\Big(1+\sum_{n=1}^\infty \|(\mathbf{B}_{\lambda,N})^n\|_{\mathcal{J}_\alpha(\mathcal{V})}\Big) \tag{9.51}$$

by (9.47). Combining (9.50) and (9.51) completes the proof. □

Proof of Proposition 7.1. Observe from (7.2) and (7.3) that

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

This together with (9.29), Lemma 9.5, and Propositions 9.4 and 9.7 implies that

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

This proves the estimate (7.6).

Now we prove (7.7). Set $\mathbf{y}_{LS} = (\mathbf{S}^T\mathbf{S})^{-1}\mathbf{d}_2$. By (9.29),

$$\|\mathbf{y}_{LS}\|_\infty \le \frac{D_1\alpha}{\alpha-d}\|(\mathbf{S}^T\mathbf{S})^{-1}\|_{\mathcal{J}_\alpha(\mathcal{V})}\|\mathbf{d}_2\|_\infty. \tag{9.52}$$

Moreover, following the proof of Proposition 9.7 gives

$$\|(\mathbf{S}^T\mathbf{S})^{-1}\|_{\mathcal{J}_\alpha(\mathcal{V})} \le \frac{2^{-\alpha-1}(\alpha-d)^2 D_2}{\alpha^2 D_1 D_1(\mathcal{G})\|\mathbf{S}\|^2_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}}. \tag{9.53}$$

Write

$$\begin{aligned}\chi^{N/2}_{\lambda,G}(\mathbf{w}_{\lambda,N}-\mathbf{w}_{LS}) &= \chi^{N/2}_{\lambda,G}(\chi^N_{\lambda,G}\mathbf{S}\chi^N_{\lambda,V})(\chi^N_{\lambda,V}\mathbf{S}^T\mathbf{S}\chi^N_{\lambda,V})^{-2}\chi^N_{\lambda,V}\mathbf{S}^T\mathbf{S}(I-\chi^N_{\lambda,V})\mathbf{d}_2\\ &\quad+\chi^{N/2}_{\lambda,G}(\chi^N_{\lambda,G}\mathbf{S}\chi^N_{\lambda,V})(\chi^N_{\lambda,V}\mathbf{S}^T\mathbf{S}\chi^N_{\lambda,V})^{-1}\chi^N_{\lambda,V}\mathbf{S}^T\mathbf{S}(I-\chi^N_{\lambda,V})\mathbf{y}_{LS}\\ &\quad-\chi^{N/2}_{\lambda,G}\mathbf{S}(I-\chi^N_{\lambda,V})\mathbf{y}_{LS}\\ &=: I+II+III.\end{aligned} \tag{9.54}$$

Using (9.26), (9.52), (9.53), Lemma 9.5, and Propositions 9.4 and 9.7, we obtain

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

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

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 (9.54) prove (7.7). □

9.14. Proof of Theorem 7.2

Let

$$\mathbf{u}_n = \mathbf{S}^T(\mathbf{w}_n-\mathbf{w}_{LS}) = \mathbf{S}^T\mathbf{w}_n-\mathbf{d}_2 \quad\text{and}\quad \mathbf{v}_n=\mathbf{S}\mathbf{u}_n,\ n\ge 1. \tag{9.55}$$

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 (7.13), (7.14) and (9.55). Therefore,

$$\begin{aligned}\|\mathbf{u}_{n+1}\|_\infty &\le \frac{D_1(\mathcal{G})L\alpha}{\alpha-d}\|\mathbf{S}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}\|\mathbf{R}_N\mathbf{S}^T\mathbf{v}_n-\mathbf{S}(\mathbf{S}^T\mathbf{S})^{-2}\mathbf{S}^T\mathbf{v}_n\|_\infty\\ &\le \frac{D_1(\mathcal{G})D_4L\alpha}{\alpha-d}\|\mathbf{S}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}(N+1)^{-\alpha+d}\|(\mathbf{S}^T\mathbf{S})^{-1}\mathbf{S}^T\mathbf{v}_n\|_\infty\\ &= r_1\|\mathbf{u}_n\|_\infty \le \cdots \le r_1^n\|\mathbf{S}^T(\mathbf{R}_N\mathbf{S}^T\mathbf{S}-\mathbf{S}(\mathbf{S}^T\mathbf{S})^{-1})\mathbf{d}_2\|_\infty\\ &\le r_1^{n+1}\|\mathbf{d}_2\|_\infty, \end{aligned} \tag{9.56}$$

where the second inequality follows from (7.10) with $\mathbf{d}_2$ replaced by $(\mathbf{S}^T\mathbf{S})^{-1}\mathbf{S}^T\mathbf{v}_n$, and the last inequality holds by (7.10) and Proposition 4.1.

Observe that

$$\mathbf{w}_{n+1}-\mathbf{w}_n = -\mathbf{R}_N\mathbf{S}^T\mathbf{S}\mathbf{u}_n. \tag{9.57}$$

Using (7.12), Proposition 4.1 and Lemma 9.5 gives

$$\|\mathbf{w}_{n+1}-\mathbf{w}_n\|_\infty \le \frac{2^{2\alpha+2}\alpha L^3(D_1(\mathcal{G}))^2D_2^2}{(\alpha-d)D_1\|\mathbf{S}\|_{\mathcal{J}_\alpha(\mathcal{G},\mathcal{V})}}\|\mathbf{u}_n\|_\infty. \tag{9.58}$$

This together with (9.56) proves the exponential convergence (7.17).

The conclusion (7.15) follows from (9.55) by taking limit $n\to\infty$.

The error estimate (7.18) between the "least squares" solution $\mathbf{d}_2$ and its sub-optimal approximation $\mathbf{S}^T\mathbf{w}_n, n\ge 1$, follows from (7.17) and Proposition 4.1.

Acknowledgments

The authors would like to thank anonymous reviewers for their comments and suggestion for the improvement of the manuscript.

References

[1] B. Adcock, A.C. Hansen, 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.
[2] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, E. Cayirci, Wireless sensor networks: a survey, Computer Networks 38 (2002) 393–422.
[3] A. Aldroubi, A. Baskakov, I. Krishtal, Slanted matrices, Banach frames, and sampling, J. Funct. Anal. 255 (2008) 1667–1691.
[4] A. Aldroubi, J. Davis, I. Krishtal, Dynamical sampling: time-space trade-off, Appl. Comput. Harmon. Anal. 34 (2013) 495–503.
[5] A. Aldroubi, H. Feichtinger, Exact iterative reconstruction algorithm for multivariate irregularly sampled functions in spline-like spaces: the L^p-theory, Proc. Amer. Math. Soc. 126 (1998) 2677–2686.
[6] A. Aldroubi, K. Gröchenig, Beurling–Landau-type theorems for non-uniform sampling in shift invariant spline spaces, J. Fourier Anal. Appl. 6 (2000) 93–103.
[7] A. Aldroubi, K. Gröchenig, Nonuniform sampling and reconstruction in shift-invariant spaces, SIAM Rev. 43 (2001) 585–620.
[8] A. Aldroubi, Q. Sun, 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.
[9] 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.
[10] D.P. Bertsekas, J. Tsitsiklis, Parallel and Distributed Computation: Numerical Methods, Prentice-Hall, Englewood Cliffs, NJ, 1989.
[11] N. Bi, M.Z. Nashed, Q. Sun, Reconstructing signals with finite rate of innovation from noisy samples, Acta Appl. Math. 107 (2009) 339–372.

[12] T. Blu, P.L. Dragotti, M. Vetterli, P. Marziliano, L. Coulot, Sparse sampling of signal innovations, IEEE Signal Process. Mag. 25 (2008) 31–40.

[13] 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.

[14] 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.

[15] E.J. Candes, J. Romberg, T. Tao, Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information, IEEE Trans. Inform. Theory 52 (2006) 489–509.

[16] P.G. Casazza, G. Kutyniok, S. Li, Fusion frames and distributed processing, Appl. Comput. Harmon. Anal. 25 (2008) 114–132.

[17] C. Cheng, Y. Jiang, Q. Sun, Sampling and Galerkin reconstruction in reproducing kernel spaces, Appl. Comput. Harmon. Anal. 41 (2016) 638–659.

[18] C. Cheng, Y. Jiang, 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.

[19] C. Chong, S. Kumar, Sensor networks: evolution, opportunities, and challenges, Proc. IEEE 91 (2003) 1247–1256.

[20] O. Christensen, T. Strohmer, The finite section method and problems in frame theory, J. Approx. Theory 133 (2005) 221–237.

[21] F.R.K. Chung, Spectral Graph Theory, American Mathematical Society, 1997.

[22] R. Coifman, G. Weiss, Analyse Harmonique Non-Commutative Sur Certains Espaces Homogeneous, Lecture Notes in Mathematics, vol. 242, Springer, New York, NY, USA, 1971.

[23] T.B. Curtin, J.G. Bellingham, J. Catipovic, D. Webb, Autonomous oceanographic sampling networks, Oceanography 6 (1993) 86–94.

[24] W. Czaja, G. Kutyniok, D. Speegle, Beurling dimension of Gabor pseudoframes for affine subspaces, J. Fourier Anal. Appl. 14 (2008) 514–537.

[25] D. Deng, Y. Han, Harmonic Analysis on Spaces of Homogeneous Type, Springer, 2008.

[26] D.L. Donoho, Superresolution via sparsity constraints, SIAM J. Math. Anal. 23 (1992) 1309–1331.

[27] D.L. Donoho, Compressed sensing, IEEE Trans. Inform. Theory 52 (2006) 1289–1306.

[28] D.L. Donoho, M. Elad, Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ^1 minimization, Proc. Natl. Acad. Sci. USA 100 (2003) 2197–2202.

[29] P.L. Dragotti, M. Vetterli, T. Blu, Sampling moments and reconstructing signals of finite rate of innovation: Shannon meets Strang-Fix, IEEE Trans. Signal Process. 55 (2007) 1741–1757.

[30] D.E. Dutkay, D. Han, Q. Sun, E. Weber, On the Beurling dimension of exponential frames, Adv. Math. 226 (2011) 285–297.

[31] T.G. Dvorkind, Y.C. Eldar, E. Matusiak, Nonlinear and nonideal sampling: theory and methods, IEEE Trans. Signal Process. 56 (2008) 5874–5890.

[32] Y.C. Eldar, M. Unser, Nonideal sampling and interpolation from noisy observations in shift-invariant spaces, IEEE Trans. Signal Process. 54 (2006) 2636–2651.

[33] B. Farrell, T. Strohmer, Inverse-closedness of a Banach algebra of integral operators on the Heisenberg group, J. Operator Theory 64 (2010) 189–205.

[34] H.G. Feichtinger, K. Gröchenig, T. Strohmer, Efficient numerical methods in non-uniform sampling theory, Numer. Math. 69 (1995) 423–440.

[35] K. Gröchenig, Time-frequency analysis of Sjöstrand class, Rev. Mat. Iberoam. 22 (2006) 703–724.

[36] K. Gröchenig, Wiener's lemma: theme and variations, an introduction to spectral invariance and its applications, in: P. Massopust, B. Forster (Eds.), Four Short Courses on Harmonic Analysis: Wavelets, Frames, Time-Frequency Methods, and Applications to Signal and Image Analysis, Birkhäuser, Boston, 2010.

[37] K. Gröchenig, A. Klotz, Noncommutative approximation: inverse-closed subalgebras and off-diagonal decay of matrices, Constr. Approx. 32 (2010) 429–466.

[38] K. Gröchenig, M. Leinert, Symmetry and inverse-closedness of matrix algebras and functional calculus for infinite matrices, Trans. Amer. Math. Soc. 358 (2006) 2695–2711.

[39] K. Gröchenig, H. Schwab, Fast local reconstruction methods for nonuniform sampling in shift-invariant spaces, SIAM J. Matrix Anal. Appl. 24 (2003) 899–913.

[40] R. Hebner, The power grid in 2030, IEEE Spectrum (April 2017) 51–55.

[41] S. Jaffard, Properiétés des matrices bien localisées prés de leur diagonale et quelques applications, Ann. Inst. Henri Poincaré D 7 (1990) 461–476.

[42] D. Kincaid, W. Cheney, Numerical Analysis: Mathematics of Scientific Computing, third edition, Brooks/Cole, 2002.

[43] J. Koshal, A. Nedic, U.V. Shanbhag, Multiuser optimization: distributed algorithm and error analysis, SIAM J. Optim. 21 (2011) 1046–1081.

[44] I. Krishtal, Wiener's lemma: pictures at exhibition, Rev. Un. Mat. Argentina 52 (2011) 61–79.

[45] N.E. Leonard, D.A. Paley, F. Lekien, R. Sepulchre, D.M. Fratantoni, R.E. Davis, Collective motion, sensor networks, and ocean sampling, Proc. IEEE 95 (2007) 48–74.

[46] C.G. Lopes, A.H. Sayed, Incremental adaptive strategies over distributed networks, IEEE Trans. Signal Process. 55 (2007) 4064–4077.

[47] I. Maravic, 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.

[48] P. Marziliano, M. Vetterli, T. Blu, Sampling and exact reconstruction of bandlimited signals with additive shot noise, IEEE Trans. Inform. Theory 52 (2006) 2230–2233.

[49] M.Z. Nashed, Q. Sun, Sampling and reconstruction of signals in a reproducing kernel subspace of $L^p(R^d)$, J. Funct. Anal. 258 (2010) 2422–2452.

[50] A. Nedic, A. Olshevsky, Distributed optimization over time-varying directed graphs, IEEE Trans. Automat. Control 60 (2015) 601–615.
[51] H. Pan, T. Blu, P.L. Dragotti, Sampling curves with finite rate of innovation, IEEE Trans. Signal Process. 62 (2014) 458–471.
[52] M. Pawlak, E. Rafajlowicz, A. Krzyzak, Postfiltering versus prefiltering for signal recovery from noisy samples, IEEE Trans. Inform. Theory 49 (2003) 3195–3212.
[53] I. Pesenson, Sampling in Paley–Wiener spaces on combinatorial graphs, Trans. Amer. Math. Soc. 360 (2008) 5603–5627.
[54] A. Sandryhaila, J.M.F. Moura, Discrete signal processing on graphs, IEEE Trans. Signal Process. 61 (2013) 1644–1656.
[55] C.E. Shin, Q. Sun, Stability of localized operators, J. Funct. Anal. 256 (2009) 2417–2439.
[56] P. Shukla, P.L. Dragotti, Sampling schemes for multidimensional signals with finite rate of innovation, IEEE Trans. Signal Process. 55 (2007) 3670–3686.
[57] D.I. Shuman, S.K. Narang, P. Frossard, A. Ortega, P. Vandergheynst, The emerging field of signal processing on graphs, IEEE Signal Process. Mag. 30 (2013) 83–98.
[58] S. Smale, D.X. Zhou, Shannon sampling and function reconstruction from point values, Bull. Amer. Math. Soc. 41 (2004) 279–305.
[59] Q. Sun, Wiener's lemma for infinite matrices with polynomial off-diagonal decay, C. R. Acad. Sci. Paris, Ser. I 340 (2005) 567–570.
[60] Q. Sun, Non-uniform average sampling and reconstruction of signals with finite rate of innovation, SIAM J. Math. Anal. 38 (2006) 1389–1422.
[61] Q. Sun, Wiener's lemma for infinite matrices, Trans. Amer. Math. Soc. 359 (2007) 3099–3123.
[62] Q. Sun, Frames in spaces with finite rate of innovations, Adv. Comput. Math. 28 (2008) 301–329.
[63] Q. Sun, Wiener's lemma for localized integral operators, Appl. Comput. Harmon. Anal. 25 (2008) 148–167.
[64] Q. Sun, Local reconstruction for sampling in shift-invariant space, Adv. Comput. Math. 32 (2010) 335–352.
[65] Q. Sun, Stability criterion for convolution-dominated infinite matrices, Proc. Amer. Math. Soc. 138 (2010) 3933–3943.
[66] Q. Sun, Wiener's lemma for infinite matrices II, Constr. Approx. 34 (2011) 209–235.
[67] Q. Sun, Localized nonlinear functional equations and two sampling problems in signal processing, Adv. Comput. Math. 40 (2014) 415–458.
[68] Q. Sun, J. Xian, Rate of innovation for (non-)periodic signals and optimal lower stability bound for filtering, J. Fourier Anal. Appl. 20 (2014) 119–134.
[69] W. Sun, X. Zhou, Characterization of local sampling sequences for spline subspaces, Adv. Comput. Math. 30 (2009) 153–175.
[70] J. Unnikrishnan, M. Vetterli, Sampling and reconstruction of spatial fields using mobile sensors, IEEE Trans. Signal Process. 61 (2013) 2328–2340.
[71] M. Unser, Sampling—50 years after Shannon, Proc. IEEE 88 (2000) 569–587.
[72] M. Vetterli, P. Marziliano, T. Blu, Sampling signals with finite rate of innovation, IEEE Trans. Signal Process. 50 (2002) 1417–1428.
[73] Z. Wang, 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.
[74] J. Yick, B. Mukherjee, D. Ghosal, Wireless sensor network survey, Computer Networks 52 (2008) 2292–2330.
[75] F. Zhao, L. Guibas, Wireless Sensor Networks: An Information Processing Approach, Morgan Kaufmann, 2004.