

# A Distributed Tracking Algorithm for Reconstruction of Graph Signals

Xiaohan Wang, *Student Member, IEEE*, Mengdi Wang, and Yuantao Gu, *Member, IEEE*

**Abstract**—The rapid development of signal processing on graphs provides a new perspective for processing large-scale data associated with irregular domains. In many practical applications, it is necessary to handle massive data sets through complex networks, in which most nodes have limited computing power. Designing efficient distributed algorithms is critical for this task. This paper focuses on the distributed reconstruction of a time-varying bandlimited graph signal based on observations sampled at a subset of selected nodes. A distributed least square reconstruction (DLSR) algorithm is proposed to recover the unknown signal iteratively, by allowing neighboring nodes to communicate with one another and make fast updates. DLSR uses a decay scheme to annihilate the out-of-band energy occurring in the reconstruction process, which is inevitably caused by the transmission delay in distributed systems. Proof of convergence and error bounds for DLSR are provided in this paper, suggesting that the algorithm is able to track time-varying graph signals and perfectly reconstruct time-invariant signals. The DLSR algorithm is numerically experimented with synthetic data and real-world sensor network data, which verifies its ability in tracking slowly time-varying graph signals.

**Index Terms**—Distributed algorithm, graph signal, sampling and reconstruction, signal processing on graph, time-varying signal.

## I. INTRODUCTION

### A. Signal Processing on Graph

**D**URING the past few years, the emerging field of signal processing on graphs (see [1], [2]) has attracted vast research interests from multiple disciplines. Consider an  $N$ -vertex undirected graph, denoted as  $\mathcal{G}(\mathcal{V}, \mathcal{E})$ , where  $\mathcal{V}$  is the set of vertices with  $|\mathcal{V}| = N$  and  $\mathcal{E}$  is the set of edges. A *graph signal*  $\mathbf{f} \in \mathbb{R}^N$  is a vector that assigns each vertex a real number. Equivalently, the vector  $\mathbf{f}$  is often regarded as a function  $f: \mathcal{V} \mapsto \mathbb{R}$ .

Graph-based signal processing has been developed to analyze data or signal associated with irregular domains, e.g., large-scale networks. It finds wide applications in sensor networks

[3], image processing [4], recommendation systems [5], etc. Existing research topics on graph signal processing include: graph signal sampling [6], [7], uncertainty principle [8], graph filtering [9], spectral graph wavelet [10], [11], graph signal compression [12], graph signal multiresolution [13], [14], parametric dictionary learning [15], graph signal coarsening [16], [17], etc.

### B. Problem Description and Related Works

In this work, we study the distributed reconstruction of smooth graph signals based on sample measurements obtained at representative nodes. Suppose that the unknown graph signal lies in the low-frequency subspace, our reconstruction problem is to recover its missing entries from its known data/signal values sampled from a representative set of nodes.

Some theoretical results have been established for the sampling problem of bandlimited graph-based signals; see e.g., [18]–[20]. The relation between the sample size necessary to obtain unique reconstruction and the cutoff frequency of bandlimited signal space has been studied. Similar to classical results on time-domain irregular sampling, the idea of “frame” has been introduced for graph signal processing. The unique reconstruction conditions have been derived for normalized and unnormalized Laplacians [18], [20]. As the field of graph signal processing is rapidly developing, we summarize some recent related works as follows. A least square approach has been proposed in [5] to reconstruct bandlimited graph signal from signal values observed on sampled vertices, using a centralized algorithm. An iterative method of bandlimited graph signal reconstruction has been proposed in [6], with the practical consideration of balancing a tradeoff between smoothness and data-fitting. Two more efficient iterative reconstruction methods using the local set have been considered in [21], [22]. A necessary and sufficient condition for perfect reconstruction of bandlimited graph signal has been derived in [7]. Readers may refer to Section II-C for more details of related works.

Signal processing on graph is naturally related to distributed systems. For large-scale systems in lack of a central controller, e.g., sensor networks, distributed estimation and tracking [23] is an important topic. Algorithmic frameworks for distributed regression [24] and inference [25] have been studied to fit global functions based on local measurements in sensor networks. Consensus-based methods have been proposed in [26], [27] to distributively compute the maximum likelihood estimate of unknown parameters. Distributed Kalman filtering has been introduced in [28] for target tracking of sensor networks. Diffusion RLS [29] and LMS [30] algorithms have been proposed for distributed estimation over adaptive networks. To the best knowledge of the authors, there have been few works on the

Manuscript received July 01, 2014; revised November 22, 2014; accepted January 31, 2015. Date of publication February 13, 2015; date of current version May 12, 2015. This work was supported in part by the National 973 Program of China under Grant 2013CB329203 and in part by the National Natural Science Foundation of China (NSFC 61371137). The guest editor coordinating the review of this manuscript and approving it for publication was Prof. Zhi Tian.

X. Wang and Y. Gu are with the Department of Electronic Engineering and Tsinghua National Laboratory for Information Science and Technology (TNList), Tsinghua University, Beijing 100084, China (e-mail: gyt@tsinghua.edu.cn).

M. Wang is with the Department of Operations Research and Financial Engineering, Princeton University, Princeton, NJ 08544 USA.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/JSTSP.2015.2403799

distributed reconstruction problem of bandlimited graph signal reconstruction. A related work [31] proposes an approximation method that calculates the graph Fourier multipliers distributively, which we will discuss in subsequent sections.

In many practical distributed systems, a central processor is lacking and the majority of nodes have severely limited data processing power. Moreover, the node-to-node transmission delay is non-negligible in large-scale networks, i.e., a given node cannot communicate globally with all other nodes to obtain instant fresh data. These difficulties with large distributed systems pose a new and practical challenge to graph-based signal processing: how to reconstruct graph signals distributively and efficiently? This is the motivation of the present paper, in which we propose a distributed algorithmic solution and answer the prior question to a reasonable extent.

### C. Contributions

In this paper, we focus on the distributed recovery problem of graph-based time-varying signal. We propose a distributed algorithm, namely the *distributed least square reconstruction* (DLSR), to adaptively reconstruct the missing values of a graph signal by allowing neighboring nodes to communicate with one another and make local updates. Due to the transmission delay caused by node-to-node communication, some out-of-band energy inevitably occurs during the distributed signal reconstruction process. The DLSR algorithm uses a decay factor to dampen this out-of-band energy and achieves *perfect reconstruction* of the unknown graph signal. Theoretical convergence proof and error bounds of DLSR is given in our analysis.

The rest of this paper is organized as follows. In Section II, some preliminaries are introduced, including the basics of graph signal processing and a review of existing related works. In Section III, the distributed reconstruction algorithm (DLSR) is proposed and described in detail. In Section IV, the proof of convergence and error bound analysis for the proposed algorithm is presented. In Section V, DLSR is evaluated using numerical experiments with synthetic as well as real-world data.

## II. PRELIMINARIES

### A. Laplacian-Based Graph Signal Processing

The concept of graph Laplacian is widely adopted in spectral graph theory [32] and signal processing on graphs [1]. For an undirected graph, the graph Laplacian (or unnormalized Laplacian, combinatorial Laplacian) is defined as

$$\mathbf{L} = \mathbf{D} - \mathbf{A},$$

where  $\mathbf{A}$  is the adjacency matrix and  $\mathbf{D}$  is the diagonal degree matrix, whose elements are the degrees of the corresponding vertices. The normalized Laplacian is defined as

$$\mathcal{L} = \mathbf{D}^{-\frac{1}{2}} \mathbf{L} \mathbf{D}^{-\frac{1}{2}}.$$

Both unnormalized Laplacian and normalized Laplacian are real symmetric positive-semidefinite matrices and all the eigenvalues are nonnegative [32]. In the rest of the paper, we mainly focus on the normalized Laplacian. However, we note

that analogous results can be easily obtained for unnormalized Laplacian.

In view of signal processing on graphs, the eigenvalues  $\{\lambda_k\}$  of the Laplacian are regarded as frequencies and the corresponding eigenvectors  $\{\mathbf{u}_k\}$  are regarded as basis vectors. Consider an arbitrary graph signal  $\mathbf{f} \in \mathbb{R}^N$ . Its frequency component corresponding to  $\lambda_k$  is the inner product between  $\mathbf{f}$  and the eigenvector  $\mathbf{u}_k$ , denoted as

$$\hat{f}(\lambda_k) = \langle \mathbf{f}, \mathbf{u}_k \rangle = \sum_{i=1}^N f(i) u_k(i).$$

The eigenvectors associated with small eigenvalues have similar values on neighboring vertices, while the eigenvectors associated with large eigenvalues are the opposite. As a result, the frequency components associated with small and large eigenvalues correspond to the low-frequency and high-frequency parts of the signal, respectively [1], [33].

Suppose  $\mathbf{f} \in \mathbb{R}^N$  is a graph signal on an  $N$ -vertex graph  $\mathcal{G}(\mathcal{V}, \mathcal{E})$ . We say  $\mathbf{f}$  is  $\omega$ -bandlimited if its frequency components corresponding to eigenvalues larger than  $\omega$  are all zero. In other words, the spectral support of  $\mathbf{f}$  is a subset of  $[0, \omega]$ . The subspace consisting of all  $\omega$ -bandlimited signals on graph  $\mathcal{G}$  is called the Paley-Wiener space, which is a Hilbert space and denoted as  $PW_\omega(\mathcal{G})$ .

Suppose that for  $\mathbf{f} \in PW_\omega(\mathcal{G})$  only the entries on a selected set of nodes  $\{f(u)\}_{u \in \mathcal{S}}$  are known, where  $\mathcal{S} \subseteq \mathcal{V}$  is the sampled vertex set. The sampling and reconstruction problem is to recover the  $\omega$ -bandlimited original signal  $\mathbf{f}$  based on the sampled data  $\{f(u)\}_{u \in \mathcal{S}}$ .

### B. Frame and Signal Reconstruction

Bandlimited signal reconstruction is closely related to the frame theory. We briefly introduce its basics in the following.

*Definition 1:* A sequence of vectors  $\{\mathbf{f}_i\}_{i \in \mathcal{I}}$  is a *frame* in a Hilbert space  $\mathcal{H}$ , if there exist two constants  $0 < A \leq B$  such that

$$A \|\mathbf{f}\|^2 \leq \sum_{i \in \mathcal{I}} |\langle \mathbf{f}, \mathbf{f}_i \rangle|^2 \leq B \|\mathbf{f}\|^2, \quad \forall \mathbf{f} \in \mathcal{H}.$$

Here the constants  $A$  and  $B$  are called frame bounds.

*Definition 2:* For a frame  $\{\mathbf{f}_i\}_{i \in \mathcal{I}}$ , the *frame operator*  $\mathbf{S} : \mathcal{H} \rightarrow \mathcal{H}$  is

$$\mathbf{S}\mathbf{f} = \sum_{i \in \mathcal{I}} \langle \mathbf{f}, \mathbf{f}_i \rangle \mathbf{f}_i,$$

where  $\mathbf{S}$  is invertible and satisfies  $A\mathbf{I} \preceq \mathbf{S} \preceq B\mathbf{I}$ .

If the Euclidean matrix norm satisfies  $\|\mathbf{I} - \lambda \mathbf{S}\| < 1$ , then

$$\mathbf{f} = \mathbf{S}^{-1} \mathbf{S}\mathbf{f} = \lambda \sum_{j=0}^{\infty} (\mathbf{I} - \lambda \mathbf{S})^j \mathbf{S}\mathbf{f}.$$

Consider the problem of reconstruction of an unknown signal  $\mathbf{f}_*$ . By defining

$$\mathbf{f}^{(k)} = \lambda \sum_{j=0}^k (\mathbf{I} - \lambda \mathbf{S})^j \mathbf{S}\mathbf{f}_*,$$

we can use the following iteration to iteratively reconstruct  $\mathbf{f}_*$  (see [34]),

$$\mathbf{f}^{(k+1)} = \lambda \mathbf{S} \mathbf{f}_* + (\mathbf{I} - \lambda \mathbf{S}) \mathbf{f}^{(k)} = \mathbf{f}^{(k)} + \lambda \mathbf{S} (\mathbf{f}_* - \mathbf{f}^{(k)}), \quad (1)$$

which achieves an exponentially shrinking error bound

$$\|\mathbf{f}^{(k)} - \mathbf{f}_*\| \leq \|\mathbf{I} - \lambda \mathbf{S}\|^k \|\mathbf{f}^{(0)} - \mathbf{f}_*\|, \quad \forall k > 0.$$

### C. Previous Works on Bandlimited Graph Signal Reconstruction

We review some basic concepts and important results regarding band limited graph signal reconstruction, which have been established in existing works.

**Definition 3:** [18] A set of vertices  $\mathcal{U} \subseteq \mathcal{V}(\mathcal{G})$  is a *uniqueness set* for space  $PW_\omega(\mathcal{G})$  if  $\forall \mathbf{f} \in PW_\omega(\mathcal{G})$ ,  $\mathbf{f}$  is uniquely determined by its values on  $\mathcal{U}$ , i.e.,  $\forall \mathbf{f}, \mathbf{g} \in PW_\omega(\mathcal{G})$ ,  $\mathbf{f}|_{\mathcal{U}} = \mathbf{g}|_{\mathcal{U}}$  implies  $\mathbf{f} = \mathbf{g}$ , where  $\mathbf{f}|_{\mathcal{U}} \in \mathbb{R}^{|\mathcal{U}|}$  is the restriction of  $\mathbf{f}$  to the subset  $\mathcal{U}$ .

In order to perfectly reconstruct the bandlimited signals, we need the following relation between the sampling set and frame has been established.

**Theorem 1:** [18] If the sampling set  $\mathcal{S}$  is a uniqueness set for  $PW_\omega(\mathcal{G})$ , then  $\{\mathcal{P}_\omega(\boldsymbol{\delta}_u)\}_{u \in \mathcal{S}}$  forms a frame in  $PW_\omega(\mathcal{G})$ , and the upper bound  $B = 1$ , where  $\mathcal{P}_\omega(\cdot)$  is the orthogonal projection onto  $PW_\omega(\mathcal{G})$ , and  $\boldsymbol{\delta}_u \in \mathbb{R}^N$  is a graph signal on  $\mathcal{G}$  whose entries satisfy

$$\delta_u(v) = \begin{cases} 1, & v = u; \\ 0, & v \neq u. \end{cases}$$

A method called iterative least square reconstruction (ILSR) has been proposed to reconstruct the bandlimited signal iteratively. To be consistent with the results of our work, the method is rewritten as follows.

**Theorem 2:** [6] If the sampling set  $\mathcal{S}$  is a uniqueness set for  $PW_\omega(\mathcal{G})$ , then the original signal  $\mathbf{f}_* \in PW_\omega(\mathcal{G})$  can be reconstructed using the sampled data  $\{f_*(u)\}_{u \in \mathcal{S}}$  by the following ILSR method,

$$\mathbf{f}^{(k+1)} = \mathbf{f}^{(k)} + \mathcal{P}_\omega \left( \sum_{u \in \mathcal{S}} (f_*(u) - f^{(k)}(u)) \boldsymbol{\delta}_u \right), \quad (2)$$

where  $\mathbf{f}^{(k)}$  is a temporary result in the  $k$ th iteration.

ILSR is derived from the method of projections onto convex sets (POCS) [35], [36], which is also known as the alternating projection method. The iteration (2) can be obtained by projecting onto the following two sets alternately,

$$\begin{aligned} \mathcal{C}_1 &= \{\mathbf{f} \in \mathbb{R}^N | f(u) = f_*(u), \forall u \in \mathcal{S}\}, \\ \mathcal{C}_2 &= PW_\omega(\mathcal{G}). \end{aligned}$$

An equivalent derivation of ILSR can be obtained with the help of frame theory. Because of Theorem 1, the above method can also be obtained for  $\lambda = 1$  in the iteration (1). Therefore,  $\mathbf{f}^{(k)}$  in (2) is the same as that in (1).

In addition to ILSR, two more efficient graph signal reconstruction methods, namely the iterative propagating reconstruction (IPR) and the iterative weighting reconstruction (IWR), have been proposed and proved to be convergent [21], [22].

Sampling and reconstruction of bandlimited graph signal is closely related to irregular sampling [37]–[40] or non-uniform sampling [41] in the time domain, which sheds light on the analysis of graph signal. In fact, ILSR, IPR and IWR all have correspondence in time-domain irregular sampling, which are known as the Marvasti method [42], the Voronoi method [43] and the adaptive weights method [37].

### D. Notation

For a graph  $\mathcal{G}$  and a cutoff frequency  $\omega$ ,  $PW_\omega(\mathcal{G})$  denotes the  $\omega$ -bandlimited space of graph signal on  $\mathcal{G}$ . For any graph signal  $\mathbf{f} \in \mathbb{R}^N$ ,  $\mathcal{P}_\omega(\mathbf{f})$  denotes its orthogonal projection onto  $PW_\omega(\mathcal{G})$ , and  $\mathcal{P}_{\omega+}(\mathbf{f})$  denotes the projection onto the orthogonal complement space of  $PW_\omega(\mathcal{G})$ . The sampling vertex set is denoted as  $\mathcal{S}$ , and the communication time delay between vertices  $u$  and  $v$  is denoted as  $\tau(u, v)$ . It is assumed that one iteration is conducted at each time step. The true graph signal at the  $k$ th time step or iteration is denoted as  $\mathbf{f}_*^{(k)}$ , whose entry associated with vertex  $u$  is  $f_*^{(k)}(u)$ .  $\tilde{\mathbf{f}}_*^{(k)}$  denotes a biased estimate of  $\mathbf{f}_*^{(k)}$ . In the reconstruction algorithm,  $\mathbf{f}^{(k)}$  denotes the temporary result obtained after the  $k$ th iteration, with  $f^{(k)}(u)$  as its entry associated with vertex  $u$ .

## III. DISTRIBUTED RECONSTRUCTION OF TIME-VARYING BANDLIMITED GRAPH SIGNAL

### A. Motivation

We consider the distributed reconstruction of a time-varying low-frequency signal defined over graph by sampling at a small portion of nodes. The problem could be described in the scenario of wireless sensor network (WSN). For a given WSN, there are unknown function  $f_*^{(t)}(v)$  associating with node  $v$  at time  $t$ . Suppose that the function is slowly varying over both the time domain and the space domain. A snapshot of such function could be modeled as an unknown time-varying low-frequency signal  $\mathbf{f}_*^{(t)} \in PW_\omega(\mathcal{G})$  located over a graph.<sup>1</sup>

Suppose that the WSN is a hybrid network and only a small subset of nodes in  $\mathcal{S}$  are equipped with sensors. As a result, one can measure the signal entries on support  $\mathcal{S}$  as  $\{f_*^{(t)}(u)\}_{u \in \mathcal{S}}$  at time  $t$ . Our purpose is to distributively estimate the function values at all nodes  $\{f_*^{(t)}(v)\}_{v \in \mathcal{V}}$ , by using historical measurements at selected nodes  $\{f_*^{(\tau)}(u)\}_{u \in \mathcal{S}, \tau \leq t}$ . See Fig. 1 for a demonstration of the raised problem.

When the signal is time-invariant, the problem reduces to bandlimited graph signal reconstruction, where the nodes with and without sensors correspond to sampled and missing data. In the distributed setting, the centralized iteration (2) no longer applies, as it is impossible for every node to obtain the instant estimation errors  $\{f_*(u) - f^{(k)}(u)\}_{u \in \mathcal{S}}$ .

In what follows, we focus on the generalization of ILSR method to distributed systems and time-varying signals. We proposed an algorithm called distributed least square reconstruction (DLSR). By letting each node conducting the iteration

<sup>1</sup>One may argue how one could model the WSN as a graph, e.g., how the weighted edges are yielded among all nodes. However, this problem is beyond the topic of this paper. We always assume the graph is available or could be estimated by some available methods.

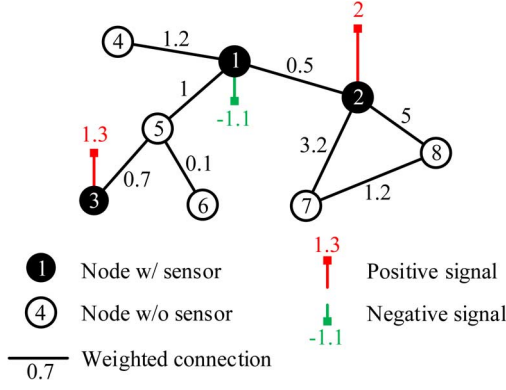


Fig. 1. An example of graph signal reconstruction over wireless sensor network.

locally at each time instant, DLSR can adaptively reconstruct the missing entries of a slowly time-varying graph signal.

### B. Algorithm Description

The basic idea of DLSR is to spread the current estimation errors associated with the representative nodes (which are equipped with sensors) to all other nodes over the connected network. Every node iteratively updates its own estimation based on its received messages.

The driver of the proposed algorithm is on those nodes with sensors, which calculates the error between the measurement  $f_*^{(k)}(u)$  and the temporary estimation  $f^{(k)}(u)$  in the  $k$ th iteration by

$$\epsilon^{(k)}(u) = f_*^{(k)}(u) - f^{(k)}(u), \quad \forall u \in \mathcal{S}.$$

Then the estimation errors at node  $u$  ( $u \in \mathcal{S}$ ) are transmitted to other nodes in the network. At the  $k$ th iteration, an arbitrary  $v$  collects a set of delayed but most recent estimation errors,

$$\left\{ \epsilon^{(k-\tau(u,v))}(u) \right\}_{u \in \mathcal{S}}, \quad \forall v \in \mathcal{V}(\mathcal{G}),$$

where  $\tau(u, v)$  denotes the transmission delay from node  $u$  to node  $v$ <sup>2</sup>. We denote the maximal transmission delay of the network by

$$\tau = \max_{u \in \mathcal{S}, v \in \mathcal{V}(\mathcal{G})} \tau(u, v).$$

Utilizing the most recent estimation errors, node  $v$  updates its local estimate by

$$f^{(k+1)}(v) = (1 - \mu_{k+1}\beta_{k+1})f^{(k)}(v) + \mu_{k+1} \sum_{u \in \mathcal{S}} \epsilon^{(k-\tau(u,v))}(u) (\mathcal{P}_\omega \delta_u)(v), \quad \forall v \in \mathcal{V}(\mathcal{G}) \quad (3)$$

where  $\mu_{k+1}$  and  $\beta_{k+1}$  denote the stepsize and decay factor, respectively.  $(\mathcal{P}_\omega \delta_u)(v)$  denotes the entry at  $v$  of the lowpass component of  $\delta_u$ , which could be calculated and stored before the system starts.

Please refer to Table I and Table II, which describe the detailed iterative process of signal reconstruction at the representative nodes and the remaining nodes, respectively.

<sup>2</sup>For simplicity, we may set  $\epsilon^{(k-\tau(u,v))}(u) = 0$  if  $k - \tau(u, v) < 0$ .

TABLE I  
DLSR ALGORITHM AT REPRESENTATIVE NODE  $u \in \mathcal{S}$

<b>Parameter:</b>	$\mathcal{S}, \{\tau(u', u)\}_{u', u \in \mathcal{S}}, \mu_k, \beta_k;$
<b>Initialization:</b>	$f^{(0)}(u) = 0$ , calculate $(\mathcal{P}_\omega \delta_{u'}) (u), \forall u' \in \mathcal{S};$
<b>For</b> $k = 0, 1, 2, \dots$	
1) Input:	$f_*^{(k)}(u);$
2) Estimation:	$\epsilon^{(k)}(u) = f_*^{(k)}(u) - f^{(k)}(u);$
3) Communication:	Send $\epsilon^{(k)}(u)$ and $\epsilon^{(k-1-\tau(u', u))}(u')$ to neighbors, $\forall u' \in \mathcal{S} \setminus u;$ Receive $\epsilon^{(k-\tau(u', u))}(u')$ from neighbors, $\forall u' \in \mathcal{S} \setminus u;$
4) Update Storage:	Save $\epsilon^{(k-\tau(u', u))}(u'), \forall u' \in \mathcal{S} \setminus u;$
5) Update Estimation:	$f^{(k+1)}(u) = (1 - \mu_{k+1}\beta_{k+1})f^{(k)}(u) + \mu_{k+1} \sum_{u' \in \mathcal{S}} \epsilon^{(k-\tau(u', u))}(u') (\mathcal{P}_\omega \delta_{u'}) (u).$
<b>End</b>	

TABLE II  
DLSR ALGORITHM AT NON-REPRESENTATIVE NODE  $v \in \mathcal{V}(\mathcal{G}) \setminus \mathcal{S}$

<b>Parameter:</b>	$\mathcal{S}, \{\tau(u, v)\}_{u \in \mathcal{S}, v \in \mathcal{V}(\mathcal{G}) \setminus \mathcal{S}}, \mu_k, \beta_k;$
<b>Initialization:</b>	$f^{(0)}(v) = 0$ , calculate $(\mathcal{P}_\omega \delta_u)(v), \forall u \in \mathcal{S};$
<b>For</b> $k = 0, 1, 2, \dots$	
1) Communication:	Send $\epsilon^{(k-1-\tau(u, v))}(u)$ to neighbors, $\forall u \in \mathcal{S};$ Receive $\epsilon^{(k-\tau(u, v))}(u)$ from neighbors, $\forall u \in \mathcal{S};$
2) Update Storage:	Save $\epsilon^{(k-\tau(u, v))}(u), \forall u \in \mathcal{S};$
3) Update Estimation:	$f^{(k+1)}(v) = (1 - \mu_{k+1}\beta_{k+1})f^{(k)}(v) + \mu_{k+1} \sum_{u \in \mathcal{S}} \epsilon^{(k-\tau(u, v))}(u) (\mathcal{P}_\omega \delta_u)(v);$
4) Output:	$f^{(k+1)}(v).$
<b>End</b>	

### C. An Example

In order to provide some intuition for our distributed algorithm, let us refer to Fig. 1 and describe what happens on a typical node in the network.

- As a representative node equipped with sensor, node 1 will get a measurement  $f_*^{(k)}(1)$  at the  $k$ th iteration and then calculate the estimation error  $\epsilon^{(k)}(1)$ . The estimation error will be sent to its neighbors of node 2, 4, and 5, and then forwarded to others. At the same slot, node 1 will receive the estimation errors of other nodes with sensors and use the most recent ones ( $\epsilon^{(k-1)}(2)$  from node 2 and  $\epsilon^{(k-2)}(3)$  from node 5). Consequently, it could update the estimation by

$$f^{(k+1)}(1) = (1 - \mu_{k+1}\beta_{k+1})f^{(k)}(1) + \mu_{k+1} \left( \epsilon^{(k)}(1) \cdot (\mathcal{P}_\omega \delta_1)(1) + \epsilon^{(k-1)}(2) (\mathcal{P}_\omega \delta_2)(1) + \epsilon^{(k-2)}(3) (\mathcal{P}_\omega \delta_3)(1) \right).$$

One may notice that the new estimation is not the output of the proposed algorithm, because our purpose is to estimate the strength of the signal associated with the node without

sensor. However, the new estimation errors at representative nodes will be transmitted over the network to help all others to conduct their estimation.

- As a regular node that is not equipped with sensor, at the  $k$ th iteration, node 5 will receive  $\epsilon^{(k-1)}(1)$ ,  $\epsilon^{(k-2)}(2)$ , and  $\epsilon^{(k-1)}(3)$  from its neighbors. Then it will transmit the most recent estimation errors of nodes with sensors to its neighbors 1, 3, and 6. The estimate of node 5 is updated by

$$\begin{aligned} f^{(k+1)}(5) = & (1 - \mu_{k+1}\beta_{k+1})f^{(k)}(5) \\ & + \mu_{k+1} \left( \epsilon^{(k-1)}(1)(\mathcal{P}_\omega \delta_1)(5) \right. \\ & \left. + \epsilon^{(k-2)}(2)(\mathcal{P}_\omega \delta_2)(5) + \epsilon^{(k-3)}(3)(\mathcal{P}_\omega \delta_3)(5) \right). \end{aligned}$$

The new estimate will be sent out as a temporary result of the proposed algorithm.

#### D. Discussions

Asynchronization is one of the most common issues in distributed systems. In our distributed reconstruction setting, asynchronization leads to communication delay between nodes that are connected through multiple links, which induces a deviation of the estimated signal from the bandlimited space. However, as long as the maximum delay in the network is bounded by a constant  $\tau$ , the proposed method can successfully annihilate the out-of-band estimation error and achieves perfect reconstruction.

Node failure is also a common problem in WSN. The proposed DLSR is robust to both communication failure and sensor failure. In the former case, some links are broken and fail to work. The data packets have to be delivered through new route and the transmission delay may increase. Even in this case, the DLSR is still going to work, provided that the network remains connected and the maximum transmission delay is bounded. In the case of a sensor failure, some presentative nodes (i. e.  $u \in \mathcal{S}$ ) no longer obtain the sampled data, which means that they can act the same as the regular nodes. As long as the system is designed with some redundancy, the DLSR can still work, provided that there remain enough number of functional sensors.

The proposed algorithm requires that the vectors  $\{\mathcal{P}_\omega \delta_u\}_{u \in \mathcal{S}}$  be calculated in advance and their entries be stored in the respective nodes. In some practical situation this pre-calculation could be unavailable. However, a distributed method proposed by [31] can be used to calculate  $\{(\mathcal{P}_\omega \delta_u)(v)\}_{u \in \mathcal{S}}$  approximately at node  $v$ .

In fact, the operation  $\mathcal{P}_\omega(\cdot)$  for any given graph signal can be approximately calculated by a distributed method proposed by [31]. By this method, it will take some (depends on the network scale) rounds of data transmission and calculation to obtain the approximate projection. Therefore, another distributed method can be readily proposed by directly applying the above approximate projection to ILSR with a stepsize  $\mu$  to track time-varying signal. Our method differs from this approach in the following aspects.

- Supposing the period of a time step composed of one transmission and calculation is fixed in both methods, it will take some time steps to implement one iteration of ILSR by directly applying the method in [31] to ILSR. Therefore, the data used in the calculation are all sampled several time

steps earlier. In our method, the iteration is conducted in one time step and use the current data at every node, which means the samples are as fresh as possible, the delay is only caused by transmission, and there is no waiting. Although nonuniform delays may violate the bandlimited property, we will prove in the following section that the out-of-band energy can be eliminated eventually.

- The projection  $\mathcal{P}_\omega(\cdot)$  should be conducted for different graph signals as the iteration goes on by directly applying the method in [31] to ILSR, and each projection takes some time steps. In the proposed method, only pre-calculating (precisely or approximately using the method in [31]) the frame elements,  $\{\mathcal{P}_\omega \delta_u\}_{u \in \mathcal{S}}$ , are enough to reconstruct the graph signal, which is more economical.

#### IV. CONVERGENCE ANALYSIS

We will first study the convergence behavior of DLSR in a general situation that the bandlimited graph signal to be constructed varies slowly by time, and then specialize the result to a time-invariant case. In order to simplify the expression, we will fix stepsizes  $\mu$  and  $\beta$  to be constants in studying the time-varying case. Finally, we let the stepsizes be diminishing and show that DLSR achieves a perfect reconstruction of time-invariant signal.

##### A. Tracking Time-Varying Signal Using Constant Parameters

The proposed DLSR algorithm is equivalent to the following iteration in the vector form

$$\mathbf{f}^{(k+1)} = (1 - \mu\beta)\mathbf{f}^{(k)} + \mu \sum_{u \in \mathcal{S}} \left( \mathbf{F}_{*u}^{(k)} - \mathbf{F}_u^{(k)} \right) \mathcal{P}_\omega \delta_u, \quad (4)$$

where

$$\mathbf{F}_{*u}^{(k)} - \mathbf{F}_u^{(k)} = \text{diag} \left\{ f_*^{(k-\tau(u,i))}(u) - f^{(k-\tau(u,i))}(u) \right\}_{i=1, \dots, N}$$

is a diagonal matrix composed of the delayed estimation error at node  $u$ .

Although  $\mathcal{P}_\omega \delta_u$  is bandlimited for any  $u$ , by introducing  $\mathbf{F}_u^{(k)}$ , the delayed signal  $(\mathbf{F}_{*u}^{(k)} - \mathbf{F}_u^{(k)})\mathcal{P}_\omega \delta_u$  no longer belongs to the low-frequency subspace  $PW_\omega(\mathcal{G})$ . As a result, the sequence of estimated signals  $\{\mathbf{f}^{(k)}\}$  are no longer  $\omega$ -bandlimited. This existence of out-of-band energy makes DLSR critically different from its centralized version (2) and substantially complicates the convergence analysis. Since  $\mathbf{f}^{(k)}$  is not  $\omega$ -bandlimited, we need to study its low-frequency and high-frequency components separately.

For given node set  $\mathcal{S}$  and cutoff frequency  $\omega$ , we may define an operator  $\mathbf{T}$  on a graph signal  $\mathbf{f}$  as

$$\begin{aligned} \mathbf{T}\mathbf{f} &= \mathcal{P}_\omega \left( \sum_{u \in \mathcal{S}} f(u)\delta_u \right) \\ &= \sum_{u \in \mathcal{S}} f(u)\mathcal{P}_\omega \delta_u. \end{aligned} \quad (5)$$

According to Theorem 1, if  $\mathcal{S}$  is the uniqueness set of graph  $\mathcal{G}$  with respect to  $\omega$ ,  $\{\mathcal{P}_\omega \delta_u\}_{u \in \mathcal{S}}$  is a frame in  $PW_\omega(\mathcal{G})$ . For any  $\mathbf{f} \in PW_\omega(\mathcal{G})$ , using the fact that

$$f(u) = \langle \mathcal{P}_\omega \mathbf{f}, \delta_u \rangle = \langle \mathbf{f}, \mathcal{P}_\omega \delta_u \rangle, \quad \forall u \in \mathcal{S},$$

$\mathbf{T}\mathbf{f}$  can be rewritten as

$$\mathbf{T}\mathbf{f} = \sum_{u \in \mathcal{S}} \langle \mathbf{f}, \mathcal{P}_\omega \boldsymbol{\delta}_u \rangle \mathcal{P}_\omega \boldsymbol{\delta}_u, \quad \forall \mathbf{f} \in PW_\omega(\mathcal{G}),$$

which is the frame operator of  $\{\mathcal{P}_\omega \boldsymbol{\delta}_u\}_{u \in \mathcal{S}}$ , and the frame bounds are  $A$  and  $B$ .

The definition of  $\mathbf{T}$  implies

$$\|\mathbf{T}\mathbf{f}\| \leq \left\| \sum_{u \in \mathcal{S}} f(u) \boldsymbol{\delta}_u \right\| \leq \|\mathbf{f}\|$$

and one has  $\|\mathbf{T}\| \leq 1$ . By defining  $\tilde{\mathbf{f}}_*^{(k)}$  as

$$\tilde{\mathbf{f}}_*^{(k)} = (\beta \mathbf{I} + \mathbf{T})^{-1} \mathbf{T} \mathbf{f}_*^{(k)}, \quad (6)$$

one further gets

$$\mathbf{T} \mathbf{f}_*^{(k)} = \beta \tilde{\mathbf{f}}_*^{(k)} + \mathbf{T} \tilde{\mathbf{f}}_*^{(k)}. \quad (7)$$

According to (5), both  $\mathbf{T} \mathbf{f}_*^{(k)}$  and  $\mathbf{T} \tilde{\mathbf{f}}_*^{(k)}$  are within the low-frequency space  $PW_\omega(\mathcal{G})$ . Therefore one can obtain from (7) that  $\tilde{\mathbf{f}}_*^{(k)} \in PW_\omega(\mathcal{G})$ . As a consequence,  $\mathcal{P}_{\omega_+} \tilde{\mathbf{f}}_*^{(k)} = \mathbf{0}$ , where  $\mathcal{P}_{\omega_+}$  denotes the projection operator onto the high-frequency subspace which is the orthogonal complement of  $PW_\omega(\mathcal{G})$ .

By defining the in-band error and out-of-band error as, respectively,

$$e^{(k)} = \left\| \mathcal{P}_\omega \mathbf{f}^{(k)} - \mathcal{P}_\omega \tilde{\mathbf{f}}_*^{(k)} \right\|, \quad (8)$$

$$e_+^{(k)} = \left\| \mathcal{P}_{\omega_+} \mathbf{f}^{(k)} - \mathcal{P}_{\omega_+} \tilde{\mathbf{f}}_*^{(k)} \right\| = \left\| \mathcal{P}_{\omega_+} \mathbf{f}^{(k)} \right\|, \quad (9)$$

the following proposition gives inequalities that  $\{e^{(k)}\}$  and  $\{e_+^{(k)}\}$  satisfy. Further, it will be shown that if the signal varies slowly enough, by properly selecting the stepsize  $\mu$ , DLSR can track time-varying signals.

*Proposition 1:* Supposing the true signal satisfies

$$\left| f_*^{(k+1)}(u) - f_*^{(k)}(u) \right| \leq \Delta, \quad \forall u \in \mathcal{V}(\mathcal{G}), k \geq 1, \quad (10)$$

if  $\Delta \leq \min\{\Delta_{\max}, \beta B_{e_+}/(|\mathcal{S}|\tau)\}$  and

$$\mu_{\min} \leq \mu < \min \left\{ \mu_{\max}, \frac{1}{\beta + A}, \frac{B_\eta}{C + |\mathcal{S}|^{\frac{3}{2}} \tau^2 \Delta} \right\}, \quad (11)$$

the errors  $\{e^{(k)}\}$  and  $\{e_+^{(k)}\}$  satisfy

$$e_+^{(k+1)} \leq (1 - \mu\beta) e_+^{(k)} + \mu^2 C + M(\mu) \Delta, \quad (12)$$

$$e^{(k+1)} \leq (1 - \mu\beta - \mu A) e^{(k)} + \mu \|\mathbf{T}\| e_+^{(k)} + \mu^2 C + M(\mu) \Delta. \quad (13)$$

In the above inequalities,

$$M(\mu) = |\mathcal{S}|^{\frac{3}{2}} \tau^2 \mu^2 + |\mathcal{S}| \tau \mu + \sqrt{N}, \quad (14)$$

$\Delta_{\max}$  is the positive root of

$$|\mathcal{S}|^{\frac{3}{2}} \tau^2 \left( 4N^{\frac{1}{2}} - |\mathcal{S}|^{\frac{1}{2}} \right) \Delta^2 + \left( 2|\mathcal{S}| \tau \beta B_{e_+} + 4N^{\frac{1}{2}} C \right) \Delta - \beta^2 B_{e_+}^2 = 0,$$

$\mu_{\min}$  and  $\mu_{\max}$  are the roots of

$$\left( C + |\mathcal{S}|^{\frac{3}{2}} \tau^2 \Delta \right) \mu^2 + (|\mathcal{S}| \tau \Delta - \beta B_{e_+}) \mu + \sqrt{N} \Delta = 0,$$

$A$  and  $B$  are the frame bounds of  $\mathbf{T}$  in  $PW_\omega(\mathcal{G})$ ,  $\|\mathbf{T}\|$  is the norm of  $\mathbf{T}$ ,  $B_\eta$ ,  $B_e$  and  $B_{e_+}$  are constants satisfying

$$(\beta + A) B_e = (\beta + \|\mathbf{T}\|) B_{e_+}, \quad (15)$$

and  $C$  is a constant

$$C = \tau \sqrt{|\mathcal{S}|} ((\beta + \|\mathbf{T}\|)(B_e + B_{e_+}) + B_\eta). \quad (16)$$

The proof of Proposition 1 is postponed to VII-A.

*Remark 1:* According to Proposition 1, the out-of-band error (12) shows the necessity of the decay factor  $\beta$ . If  $\beta = 0$ , the out-of-band error cannot be proved to converge, and the error may accumulate with the iterations. The decay factor  $\beta$  enhances the robustness of iteration.

*Remark 2:* The inequality (13) shows that the out-of-band error  $e_+^{(k)}$  also affects the in-band error  $e^{(k+1)}$ , which implies that the in-band error cannot be very small if the out-of-band error exists. In other words, it is important to eliminate the out-of-band error.

Taking the limit superior of (12) and (13), one obtains

$$\limsup_{k \rightarrow \infty} e_+^{(k)} \leq \left( D + \frac{E}{\beta} \right) \mu + \frac{M(\mu)}{\beta \mu} \Delta, \quad (17)$$

$$\limsup_{k \rightarrow \infty} e^{(k)} \leq \left( 1 + \frac{\|\mathbf{T}\|}{\beta} \right) \left( \frac{D\beta + E}{\beta + A} \mu + \frac{M(\mu)}{(\beta + A)\mu} \Delta \right), \quad (18)$$

where  $D$  and  $E$  are constants. The above results imply that for a constant stepsize  $\mu$ , the out-of-band error  $e_+^{(k)}$  will eventually get below a threshold that is determined by  $\mu$  and  $\Delta$ . Similarly,  $\mathcal{P}_\omega \mathbf{f}^{(k)}$  will converge to the neighborhood of  $\mathcal{P}_\omega \tilde{\mathbf{f}}_*^{(k)}$ , and the error is also controlled by  $\mu$  and  $\Delta$ .

*Remark 3:* Because bias is introduced by the multiple  $1 - \mu\beta$  in the iteration, the low-frequency and high-frequency components of the temporary estimate will get into the neighborhoods of  $\mathcal{P}_\omega \tilde{\mathbf{f}}_*^{(k)}$  and  $\mathbf{0}$ , respectively. Because of the influence of the decay factor  $\beta$ , the reconstructed signal is biased. It will be proved in Corollary 1 that in the time-invariant case, these two components will exactly converge to  $\mathcal{P}_\omega \tilde{\mathbf{f}}_*^{(k)}$  and  $\mathbf{0}$ .

## B. Recovering Time-Invariant Signal Using Constant Parameters

For the time-invariant case,  $\mathbf{f}_*^{(k)}$  can be written as  $\mathbf{f}_*$  and  $\mathbf{F}_{*u}^{(k)}$  becomes  $f_*(u) \mathbf{I}_N$ . Similar to (6),  $\tilde{\mathbf{f}}_*$  can also be defined as

$$\tilde{\mathbf{f}}_* = (\beta \mathbf{I} + \mathbf{T})^{-1} \mathbf{T} \mathbf{f}_*. \quad (19)$$

Then Proposition 1 becomes the following corollary for  $\Delta = 0$ .

*Corollary 1:* For time-invariant true signal  $\mathbf{f}_*$ , supposing the in-band error and out-of-band error are defined as, respectively,

$$e^{(k)} = \left\| \mathcal{P}_\omega \mathbf{f}^{(k)} - \mathcal{P}_\omega \tilde{\mathbf{f}}_* \right\|,$$

$$e_+^{(k)} = \left\| \mathcal{P}_{\omega_+} \mathbf{f}^{(k)} - \mathcal{P}_{\omega_+} \tilde{\mathbf{f}}_* \right\| = \left\| \mathcal{P}_{\omega_+} \mathbf{f}^{(k)} \right\|,$$

the error  $\{e^{(k)}\}$  and  $\{e_+^{(k)}\}$  satisfy

$$e_+^{(k+1)} \leq (1 - \mu\beta) e_+^{(k)} + \mu^2 C, \quad (20)$$

$$e^{(k+1)} \leq (1 - \mu\beta - \mu A) e^{(k)} + \mu \|\mathbf{T}\| e_+^{(k)} + \mu^2 C, \quad (21)$$

if

$$\mu < \min \left\{ \frac{1}{\beta + A}, \frac{B_\eta}{C}, \frac{\beta B_{e+}}{C} \right\},$$

where the constants are the same as those in Proposition 1.

Besides, (17) and (18) become, respectively,

$$\limsup_{k \rightarrow \infty} e_+^{(k)} \leq \frac{C}{\beta} \mu = \left( D + \frac{E}{\beta} \right) \mu, \quad (22)$$

$$\limsup_{k \rightarrow \infty} e^{(k)} \leq \frac{D\beta + E}{\beta + A} \left( 1 + \frac{\|\mathbf{T}\|}{\beta} \right) \mu. \quad (23)$$

*Remark 4:* For the time-invariant case, the out-of-band error  $e_+^{(k)}$  will get below a threshold that is proportional to the step-size  $\mu$ . It means that the out-of-band energy will be almost eliminated eventually along with the iteration if  $\mu$  is small.  $\mathcal{P}_\omega \mathbf{f}^{(k)}$  will converge to the neighborhood of  $\mathcal{P}_\omega \tilde{\mathbf{f}}_*$ , and its radius is also proportional to  $\mu$ . Therefore, for diminishing stepsize  $\mu_k$  approaching 0,  $\mathcal{P}_\omega \mathbf{f}^{(k)}$  and  $\mathcal{P}_{\omega+} \mathbf{f}^{(k)}$  will strictly converge to  $\mathcal{P}_\omega \mathbf{f}_*$  and 0, respectively, for a sequence of properly chosen diminishing stepsize.

The bias will be estimated next. Because  $\mathbf{f}_*, \tilde{\mathbf{f}}_* \in PW_\omega(\mathcal{G})$ , the iteration  $\mathbf{f}^{(k)}$  will converge to  $\tilde{\mathbf{f}}_*$ . Considering the definition of  $\tilde{\mathbf{f}}_*$  in (19), the bias satisfies

$$\begin{aligned} \tilde{\mathbf{f}}_* - \mathbf{f}_* &= (\beta \mathbf{I} + \mathbf{T})^{-1} \mathbf{T} \mathbf{f}_* - \mathbf{f}_* \\ &= (\beta \mathbf{I} + \mathbf{T})^{-1} \mathbf{T} \mathbf{f}_* - (\beta \mathbf{I} + \mathbf{T})^{-1} (\beta \mathbf{I} + \mathbf{T}) \mathbf{f}_* \\ &= -\beta (\beta \mathbf{I} + \mathbf{T})^{-1} \mathbf{f}_*. \end{aligned}$$

For  $\mathbf{f}_* \in PW_\omega(\mathcal{G})$ , according to the frame bounds of operator  $\mathbf{T}$ , we have

$$\|(\beta \mathbf{I} + \mathbf{T})^{-1} \mathbf{f}_*\| \leq \frac{1}{\beta + A} \|\mathbf{f}_*\|,$$

and then

$$\|\tilde{\mathbf{f}}_* - \mathbf{f}_*\| \leq \frac{\beta}{\beta + A} \|\mathbf{f}_*\|. \quad (24)$$

Thus, the bias is determined by  $\beta$  and decreases with the decrease of  $\beta$ .

Combining (22), (23), and (24), the following proposition gives the limit superior of the total error, as a function of  $\beta$  and  $\mu$ .

*Proposition 2:* The total error for the time-invariant case satisfies

$$\limsup_{k \rightarrow \infty} \|\mathbf{f}^{(k)} - \mathbf{f}_*\| \leq F\beta + G\frac{\mu}{\beta} + H\mu + J\beta\mu,$$

where  $F$ ,  $G$ ,  $H$ , and  $J$  are constants.

Proposition 2 can be easily proved by summing up (22), (23), and (24).

### C. Recovering Time-Invariant Signal Using Variable Parameters

Finally we will back to the general situation of variable step-size and decay factor. According to Proposition 2, by discarding the higher order of a diminishing stepsize  $\mu_k$ , the best  $\beta_k$  satisfies  $\beta_k \sim O(\sqrt{\mu_k})$  to minimize the total error. Accordingly, the total error bound can be controlled by adjusting  $\mu_k$  and  $\beta_k$ .

In other words, the reconstruction error can be made *arbitrarily small*: DLSR achieves perfect reconstruction.

The following proposition gives the convergence analysis for a special choice of diminishing stepsize and decay factor.

*Proposition 3:* For diminishing stepsize  $\mu_k = \mu_1/\sqrt{k}$  and decay factor  $\beta_k = \beta_1/\sqrt[4]{k}$ , the total error satisfies the following inequality,

$$\|\mathbf{f}^{(k)} - \mathbf{f}_*\| \leq K/\sqrt[4]{k},$$

where  $K$  is a constant. It means that the total estimation error decreases on the rate of  $1/\sqrt[4]{k}$  and converges to zero eventually.

The proof of Proposition 3 is postponed to VII-B.

## V. EXPERIMENTS

Experiments are designed to confirm the theoretical analysis and test the performance of the proposed distributed algorithm. The graph is generated by 100 randomly located nodes and the edges are generated by 4-nearest neighbors of the nodes, and the weights are inversely proportional to the square of geometric distance. Among the 100 nodes, 20 of them are randomly selected as the sampling set. The cutoff frequency is chosen to guarantee that the sampling node set is a uniqueness set, which can be determined by the method given in [5].<sup>3</sup> The bandlimited signal is generated by filtering the high-frequency components off. The transmission delay of each pair of nodes is simply regarded as the number of hops between them in the graph. The maximal transmission delay of this graph is 14.

### A. Tracking Time-Varying Signals

1) *Tracking Performance:* In this experiment, the tracking performance of DLSR is verified. The parameters are chosen as  $\Delta = 0.005$ ,  $\mu = 0.1$ , and  $\beta = 10^{-3}$ . The time-varying signal is generated by adding a random bandlimited increment whose largest absolute entry is  $\Delta$  for each time step. The aiming signal and iterative results of four nodes are focused on, as illustrated in Fig. 2. The nodes associated with the upper two subfigures are in the sampling set, and the nodes in the lower two subfigures are not in the sampling set. All the nodes can track the aiming signal for not dramatic changes. The proposed algorithm can track the slowly varying graph signal along with time.

2) *Parameters  $\beta$ ,  $\mu$ , and  $\Delta$ :* In this experiment, the regions for the parameters  $\beta$  and  $\mu$  that guarantee the convergence of DLSR are plotted in Fig. 3 for different  $\Delta$ , which describes the varying rate of time-varying signals. The experiment results show that for time-varying case the algorithm is not convergent if the stepsize is too large or too small. If the  $\mu$  is too small, the estimation cannot track the varying signal. This will not happen for the time-invariant case ( $\Delta = 0$ ).

### B. Reconstruction of Time-Invariant Signal

1) *Convergence Performance:* In this experiment, DLSR is used to reconstruct time-invariant signals. The convergence curves of distributed and centralized algorithms with constant

<sup>3</sup>Proposition 2 of [5]: The sampling set  $\mathcal{S}$  is a unique set for  $PW_\omega(\mathcal{G})$  if the cutoff frequency satisfies  $\omega \leq \sigma_{\min}$ , where  $\sigma_{\min}^2$  is the smallest singular value of  $(\mathcal{L}^2)_{\mathcal{S}^c}$ , which is the submatrix of  $\mathcal{L}^2$  containing only the rows and columns corresponding to the complementary set of  $\mathcal{S}$ , and  $\mathcal{L}$  is the normalized Laplacian of  $\mathcal{G}$ .



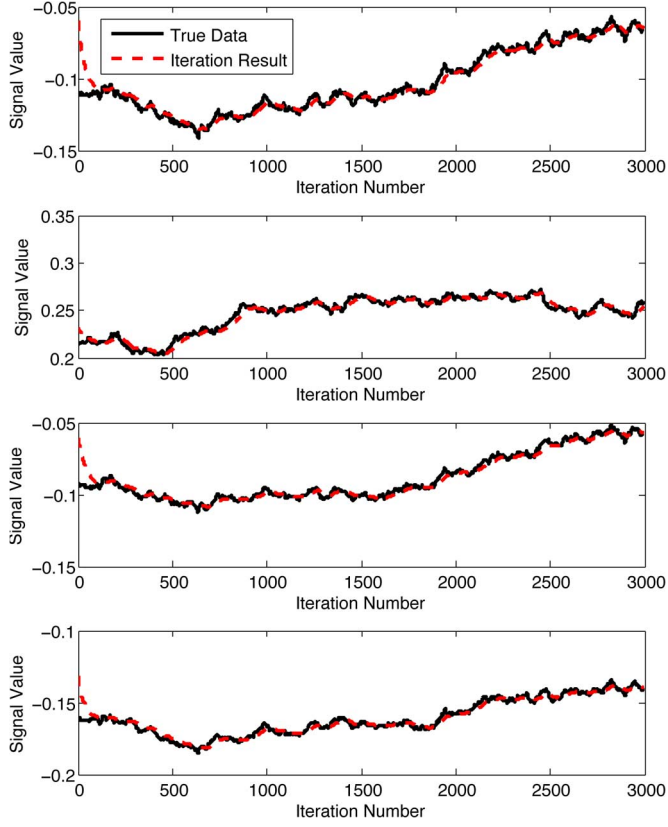


Fig. 2. Time-varying aiming signal and iterative results of four nodes. The proposed algorithm can track the aiming signal over time.

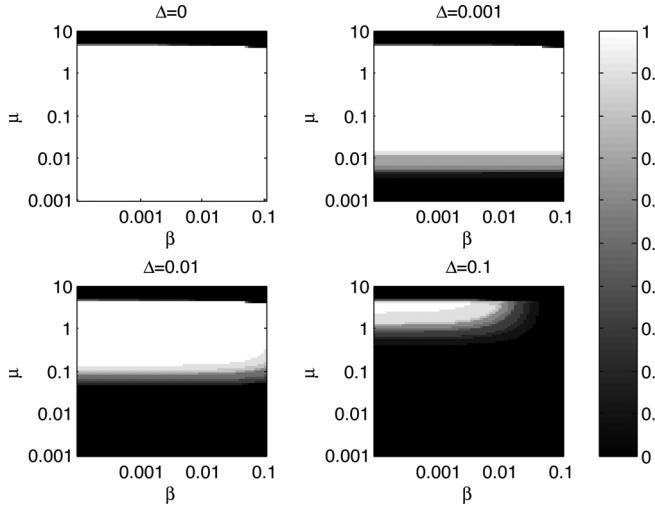


Fig. 3. The probability of convergence for different choices of  $\mu$  and  $\beta$  in time-invariant and time-varying cases.

stepsizes  $\mu = 0.01$  and  $\mu = 0.02$  are illustrated in Fig. 4, with the decay factor  $\beta = 0.01$ . The centralized algorithm uses fresh data from the sampled nodes, while the distributed algorithm uses data with transmission delay. It is easy to see that a larger stepsize results in a faster convergence. The bias caused by  $\beta$  can be seen in the convergence curves of distributed algorithms, while the error of centralized algorithm shrinks exponentially to zero with no bias.

2) *In-Band and out-of-Band Errors*: As proved in Corollary 1, both the in-band and out-of-band errors will decrease into a small bound as the iteration goes for  $\beta > 0$ . In this experi-

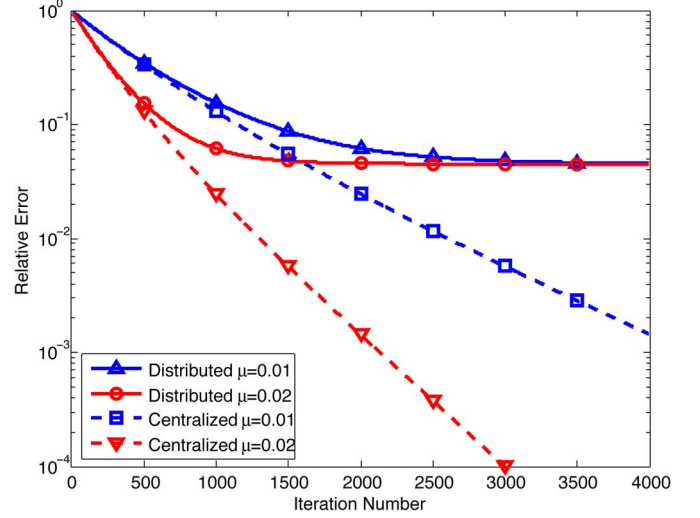


Fig. 4. The convergence curves of distributed and centralized algorithms with different constant stepsizes, where the decay factor is fixed  $\beta = 0.01$ .

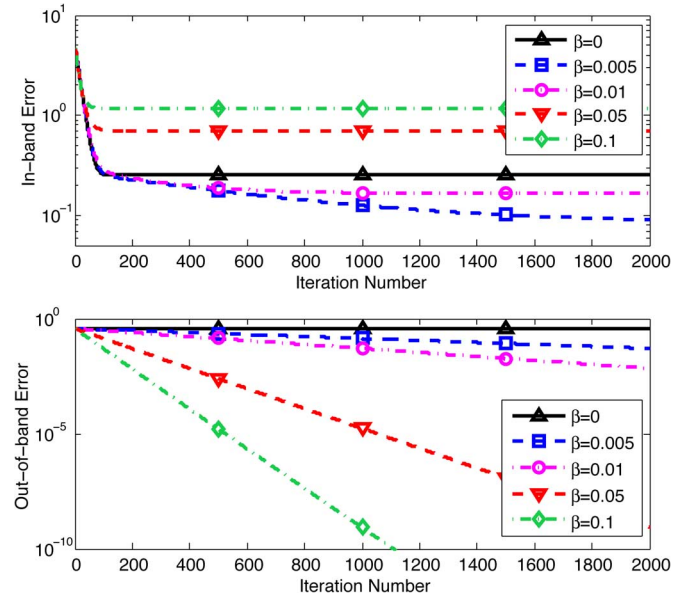


Fig. 5. The in-band and out-of-band errors for different  $\beta$  if the initial value has out-of-band energy. If  $\beta = 0$  the out-of-band energy cannot be eliminated as the iteration goes, and it also leads to a relatively larger in-band error. For  $\beta > 0$ , a larger  $\beta$  will lead to a faster shrinkage of the out-of-band error, and a larger steady-state in-band error.

ment, we set the initial value  $\mathbf{f}^{(0)}$  with about 10% of out-of-band energy and conduct the iteration with different decay factors  $\beta = 0, 0.005, 0.01, 0.05$ , and  $0.1$ . The stepsize is chosen as  $\mu = 0.2$ . The in-band and out-of-band errors are illustrated in Fig. 5. The experiment result shows the necessity of the decay factor. Although there is no bias for  $\beta = 0$ , the out-of-band energy cannot be eliminated as the iteration goes. It should be noted that the curve for  $\beta = 0$  is not comparable with the others because the out-of-band energy also affects the in-band error according to Remark 2. Since the out-of-band energy cannot be eliminated, it also leads to a relatively larger in-band error for  $\beta = 0$ . For  $\beta > 0$ , it can be seen from Fig. 5 that even though the initial value has out-of-band energy, it will shrink towards zero along with the iteration. A larger  $\beta$  will lead to a faster shrinkage of the out-of-band error, and a larger steady-state in-band error.



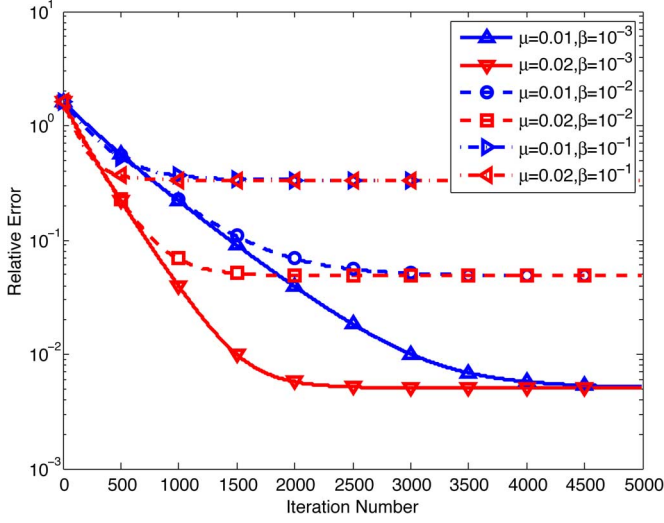


Fig. 6. Convergence curves for different choices of  $\beta$  and  $\mu$ . The decay factor  $\beta$  mainly determines the steady-state error and the stepsize  $\mu$  mainly determines the convergence rate.

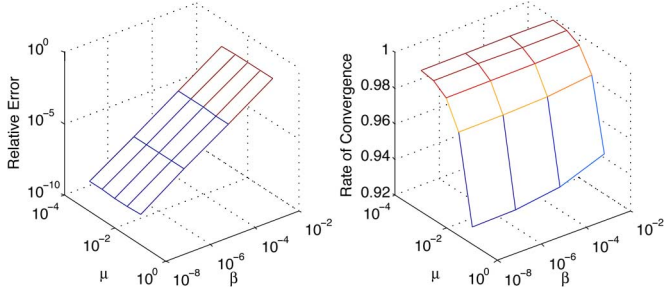


Fig. 7. The steady-state errors and convergence rates for different choices of  $\beta$  and  $\mu$ .

3) *Constant Parameters  $\beta$  and  $\mu$* : The convergence curves for different choices of constant  $\beta$  and  $\mu$  are illustrated in Fig. 6. As analyzed above, the convergence rate is mainly determined by the stepsize  $\mu$ . The steady-state error is composed of two parts, the bias and the steady-state error introduced by the constant stepsize. The latter is relatively small compared with the former, which is mainly determined by the decay factor  $\beta$ . It is obvious that a smaller  $\beta$  will lead to a smaller steady-state error.

The steady-state errors and convergence rates for different choices of  $\beta$  and  $\mu$  are plotted in Fig. 7. For fixed  $\beta$ , the steady-state error varies little with  $\mu$ . It shows that the bias, which is determined by  $\beta$ , is dominant in the total error, while  $\mu$  influences the total error little. Since there is bias in the convergence, the convergence rate is approximately calculated as  $(\|\mathbf{f}^{(m)} - \mathbf{f}_*\| / \|\mathbf{f}^{(0)} - \mathbf{f}_*\|)^{1/m}$ , where  $m$  is the number of iterations when the error reaches 1.2 times the steady-state error. The rate of convergence is smaller for larger  $\mu$ , which means it converges faster.

4) *Diminishing Parameters  $\mu_k$  and  $\beta_k$* : An experiment for diminishing stepsizes and decay factors is conducted and the convergence curves are shown in Fig. 8. The stepsizes are chosen as  $\mu_k = \mu_1 / \sqrt{k}$  with  $\mu_1 = 0.05$  or  $0.02$ . The decay factor are  $\beta_k = \beta_1 / \sqrt[4]{k}$  with  $\beta_1 = 0.1$  or  $0.01$ . All the curves decline along the iteration. Among the four curves, it can be seen that a larger  $\mu_1$  and a smaller  $\beta_1$  may lead to a faster convergence.

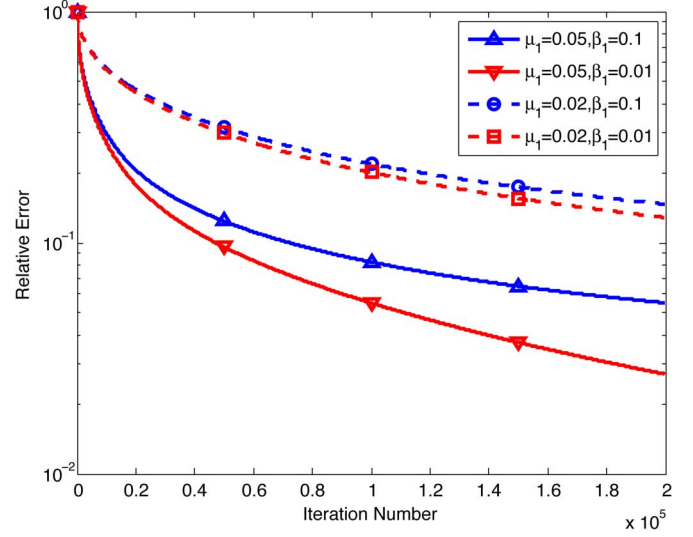


Fig. 8. The convergence curves for diminishing stepsizes and decay factors in Proposition 3.

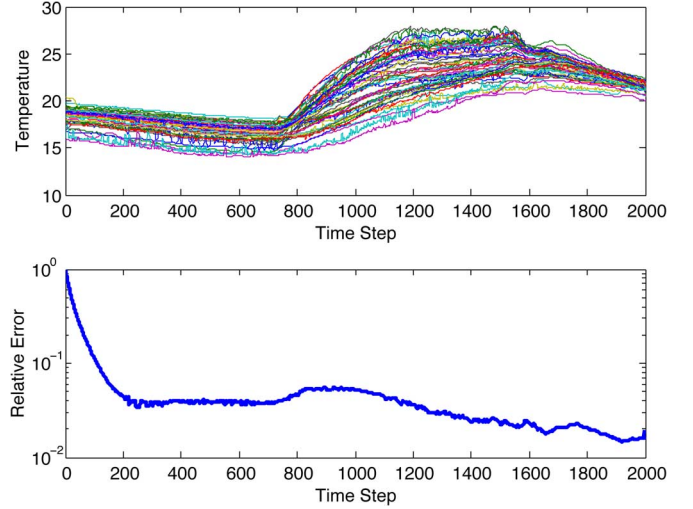


Fig. 9. The temperature of each node in the sensor network data and the relative error of DLSR.

### C. Experiments With Real Data

The sensor network data of Intel Berkeley Research Lab [44] is used in this experiment. The data is collected from 54 sensors in the lab and sampled every 30 seconds from February 28th, 2004, including temperature, humidity, light, and voltage. In our experiment, the graph signal is composed of the temperature of the sensors. We extract the data from 01:06:15 to 17:56:15 on February 28th, 2004, during which time there is less missing data. Taking time and space smoothness into consideration, the missing data is interpolated by conducting the MATLAB function scatteredInterpolant with all the existing data. Then the completed data is regarded as the original time-varying graph signal. The graph is established by the 4-nearest neighbors of the positions of the sensors, and the weights are inversely proportional to the square of geometric distance. We randomly choose 20 sensors and reconstruct the temperature of the other sensors. By selecting  $\mu = 0.1$  and  $\beta = 10^{-3}$ , the time-varying graph signal is reconstructed by DLSR. The temperature of each node and the relative error are illustrated in Fig. 9. The steady-state

relative error is around 3%, which verifies the effectiveness of DLSR.

## VI. CONCLUSION

In this paper, the distributed least square reconstruction algorithm (DLSR) is proposed to estimate and track the unobserved data of a time-varying graph signal adaptively. The low-frequency and high-frequency components of the recovered signals are theoretically proved to converge, respectively, to their true values. The out-of-band energy caused by node-to-node transmission delay can be eliminated by using the decay factor, which introduces a controllable bias. The expression of the overall error bound is given as a function of the stepsize and decay factor, and can be made arbitrarily small. Numerical experiments on both synthetic and real world data verify the performance of the proposed algorithm and show that DLSR is able to track slowly varying graph signals adaptively.

## APPENDIX

### A. The Proof of Proposition 1

First, besides the in-band error and out-of-band error defined in (8) and (9), two sequences of quantities are introduced as

$$\delta^{(k)} = \left\| \mathbf{f}^{(k)} - \mathbf{f}^{(k-1)} \right\|, \quad (25)$$

$$\eta^{(k)} = \left\| \sum_{u \in \mathcal{S}} \left( f^{(k)}(u) \mathbf{I}_N - \mathbf{F}_u^{(k)} \right) \mathcal{P}_\omega \boldsymbol{\delta}_u \right\|. \quad (26)$$

Then we will prove the following inequalities, where the proofs are postponed to the end of this subsection.

$$e^{(k+1)} \leq (1 - \mu\beta - \mu A) e^{(k)} + \mu \|\mathbf{T}\| e_+^{(k)} + \mu \eta^{(k)} + \left( \sqrt{N} + \mu |\mathcal{S}| \tau \right) \Delta, \quad (27)$$

$$e_+^{(k+1)} \leq (1 - \mu\beta) e_+^{(k)} + \mu \eta^{(k)} + \left( \sqrt{N} + \mu |\mathcal{S}| \tau \right) \Delta, \quad (28)$$

$$\eta^{(k)} \leq \sqrt{|\mathcal{S}|} \sum_{i=0}^{\tau-1} \delta^{(k-i)}, \quad (29)$$

$$\delta^{(k)} \leq \mu \left( (\beta + \|\mathbf{T}\|) \left( e^{(k-1)} + e_+^{(k-1)} \right) + \eta^{(k-1)} \right) + \mu |\mathcal{S}| \tau \Delta, \quad (30)$$

Plugging (30) into (29), we have

$$\eta^{(k)} \leq \mu \left( C + |\mathcal{S}|^{\frac{3}{2}} \tau^2 \Delta \right), \quad (31)$$

where  $C$  is defined as (16).

By plugging (31) into (28) and (27), respectively, we have (12) and (13) readily. As a consequent, if we could demonstrate that  $\{e^{(i)}\}$ ,  $\{e_+^{(i)}\}$ , and  $\{\eta^{(i)}\}$  are bounded by some constants, respectively, the proof of Proposition 1 will be closed. In what follows, we will prove this by mathematical induction.

For  $i = 1$ , these quantities are obviously bounded. We will then prove that if the preceding  $k - 1$  items of  $\{e^{(i)}\}$ ,  $\{e_+^{(i)}\}$ , and  $\{\eta^{(i)}\}$  have respective bounds of  $B_e$ ,  $B_{e_+}$ , and  $B_\eta$ , their  $k$ th items are also bounded by the same limits.

1) According to (31),  $\eta^{(k)}$  is bounded by  $B_\eta$ .

2) Plugging  $e_+^{(k-1)} \leq B_{e_+}$  in (12), we have

$$e_+^{(k)} \leq (1 - \mu\beta) B_{e_+} + \mu^2 C + M(\mu) \Delta,$$

where  $M(\mu)$  is defined in (14). We then have  $e_+^{(k)} \leq B_{e_+}$  if

$$\left( C + |\mathcal{S}|^{\frac{3}{2}} \tau^2 \Delta \right) \mu^2 + (|\mathcal{S}| \tau \Delta - \beta B_{e_+}) \mu + \sqrt{N} \Delta \leq 0. \quad (32)$$

One may notice that the inequality (32) has solutions for  $\mu$ , if and only if the following inequality holds for  $\Delta$ ,

$$(|\mathcal{S}| \tau \Delta - \beta B_{e_+})^2 - 4 \sqrt{N} \Delta \left( C + |\mathcal{S}|^{\frac{3}{2}} \tau^2 \Delta \right) \geq 0,$$

which is equivalent to

$$|\mathcal{S}|^{\frac{3}{2}} \tau^2 \left( 4N^{\frac{1}{2}} - |\mathcal{S}|^{\frac{1}{2}} \right) \Delta^2 + \left( 2|\mathcal{S}| \tau \beta B_{e_+} + 4N^{\frac{1}{2}} C \right) \Delta \leq \beta^2 B_{e_+}^2. \quad (33)$$

Because its left hand side is an increasing function of  $\Delta$ , the inequality (33) is satisfied if  $\Delta \leq \Delta_{\max}$ , where  $\Delta_{\max}$  can be solved from (33). Then from (32) the range of  $\mu$  can be determined as  $\mu_{\min} \leq \mu \leq \mu_{\max}$ , where both  $\mu_{\min}$  and  $\mu_{\max}$  are related to  $\Delta$ .

3) Plugging  $e^{(k-1)} \leq B_e$  and  $e_+^{(k-1)} \leq B_{e_+}$  into (27), we have

$$e^{(k)} \leq (1 - \mu\beta - \mu A) B_e + \mu \|\mathbf{T}\| B_{e_+} + \mu^2 C + M(\mu) \Delta.$$

Using (15), we can also obtain that  $e^{(k)} \leq B_e$  if (32) is satisfied.

Consequently,  $\{e^{(k)}\}$ ,  $\{e_+^{(k)}\}$ , and  $\{\eta^{(k)}\}$  are bounded, and then Proposition 1 is proved.

To end this subsection, we will prove (27), (28), (29), and (30). To simplify the expression, we introduce two vectors to denote the misalignment of estimated signal and the increment of true signal by

$$\mathbf{d}^{(k)} = \mathbf{f}^{(k)} - \tilde{\mathbf{f}}_*^{(k)}, \quad (34)$$

$$\mathbf{c}^{(k)} = \tilde{\mathbf{f}}_*^{(k)} - \tilde{\mathbf{f}}_*^{(k-1)}, \quad (35)$$

and two diagonal matrices to denote the errors at node  $u$  caused by delayed true signal and estimated signal by

$$\mathbf{E}_{*u}^{(k)} = f_*^{(k)}(u) \mathbf{I}_N - \mathbf{F}_{*u}^{(k)}, \quad (36)$$

$$\mathbf{E}_u^{(k)} = f^{(k)}(u) \mathbf{I}_N - \mathbf{F}_u^{(k)}, \quad (37)$$

1) *The Proof of (27) and (28)*: According to (4), (5), and (7), we have

$$\begin{aligned} \mathbf{d}^{(k+1)} &= (1 - \mu\beta) \mathbf{d}^{(k)} - \mathbf{c}^{(k+1)} - \mu\beta \tilde{\mathbf{f}}_*^{(k)} \\ &\quad + \mu \sum_{u \in \mathcal{S}} (\mathbf{F}_{*u}^{(k)} - \mathbf{F}_u^{(k)}) \mathcal{P}_\omega \boldsymbol{\delta}_u \\ &= (1 - \mu\beta) \mathbf{d}^{(k)} - \mathbf{c}^{(k+1)} - \mu \mathbf{T} \mathbf{d}^{(k)} \\ &\quad + \mu \mathbf{T} (\mathbf{f}^{(k)} - \mathbf{f}_*^{(k)}) + \mu \sum_{u \in \mathcal{S}} (\mathbf{F}_{*u}^{(k)} - \mathbf{F}_u^{(k)}) \mathcal{P}_\omega \boldsymbol{\delta}_u \\ &= \mathbf{Q} \left( \mathcal{P}_\omega \mathbf{d}^{(k)} + \mathcal{P}_{\omega+} \mathbf{d}^{(k)} \right) - \mathbf{c}^{(k+1)} \\ &\quad + \mu \sum_{u \in \mathcal{S}} \mathbf{E}_u^{(k)} \mathcal{P}_\omega \boldsymbol{\delta}_u - \mu \sum_{u \in \mathcal{S}} \mathbf{E}_{*u}^{(k)} \mathcal{P}_\omega \boldsymbol{\delta}_u, \end{aligned} \quad (38)$$

where

$$\mathbf{Q} = (1 - \mu\beta) \mathbf{I}_N - \mu \mathbf{T}.$$

Considering the definition of  $\mathbf{T}$  in (5), for any  $\mathbf{f}$ , we have

$$\mathcal{P}_\omega \mathbf{Q} \mathcal{P}_\omega \mathbf{f} = \mathbf{Q} \mathcal{P}_\omega \mathbf{f}, \quad (39)$$

$$\mathcal{P}_\omega \mathbf{Q} \mathcal{P}_{\omega_+} \mathbf{f} = -\mu \mathbf{T} \mathcal{P}_{\omega_+} \mathbf{f}, \quad (40)$$

$$\mathcal{P}_{\omega_+} \mathbf{Q} \mathcal{P}_\omega \mathbf{f} = \mathbf{0}, \quad (41)$$

$$\mathcal{P}_{\omega_+} \mathbf{Q} \mathcal{P}_{\omega_+} \mathbf{f} = (1 - \mu\beta) \mathcal{P}_{\omega_+} \mathbf{f}. \quad (42)$$

Therefore, according to (39) and (40), the low-frequency part of (38) is

$$\begin{aligned} \mathcal{P}_\omega \mathbf{d}^{(k+1)} &= \mathbf{Q} \mathcal{P}_\omega \mathbf{d}^{(k)} - \mu \mathbf{T} \mathcal{P}_{\omega_+} \mathbf{d}^{(k)} - \mathcal{P}_\omega \mathbf{c}^{(k+1)} \\ &+ \mu \mathcal{P}_\omega \sum_{u \in \mathcal{S}} \mathbf{E}_u^{(k)} \mathcal{P}_\omega \boldsymbol{\delta}_u - \mu \mathcal{P}_\omega \sum_{u \in \mathcal{S}} \mathbf{E}_{*u}^{(k)} \mathcal{P}_\omega \boldsymbol{\delta}_u. \end{aligned} \quad (43)$$

Since the frame bound of  $\{\mathcal{P}_\omega \boldsymbol{\delta}_u\}_{u \in \mathcal{S}}$  satisfies  $A \leq B \leq 1$ , if we choose a stepsize  $\mu$  satisfying  $\mu < 1/(\beta + A)$ , the assumption  $A\mathbf{I}_N \preceq \mathbf{T} \preceq B\mathbf{I}_N$  implies, note that  $\mathcal{P}_\omega \mathbf{d}^{(k)} \in PW_\omega(\mathcal{G})$ ,

$$\|\mathbf{Q}\| \leq 1 - \mu\beta - \mu A.$$

According to (10),

$$\|\mathcal{P}_\omega \mathbf{c}^{(k+1)}\| \leq \|\mathbf{c}^{(k+1)}\| \leq \sqrt{N} \Delta.$$

The definition of  $\mathbf{F}_{*u}^{(k)}$  implies

$$-\tau \Delta \mathbf{I}_N \preceq \mathbf{E}_{*u}^{(k)} \preceq \tau \Delta \mathbf{I}_N,$$

and then the last term of (43) is bounded by

$$\left\| \mathcal{P}_\omega \left( \sum_{u \in \mathcal{S}} \mathbf{E}_{*u}^{(k)} \mathcal{P}_\omega \boldsymbol{\delta}_u \right) \right\| \leq |\mathcal{S}| \tau \Delta.$$

Taking the norm of (43) and combining the above inequalities, the inequality (27) is obtained.

According to (41) and (42), the high-frequency part of (38) is

$$\begin{aligned} \mathcal{P}_{\omega_+} \mathbf{d}^{(k+1)} &= (1 - \mu\beta) \mathcal{P}_{\omega_+} \mathbf{d}^{(k)} - \mathcal{P}_{\omega_+} \mathbf{c}^{(k+1)} \\ &+ \mu \mathcal{P}_{\omega_+} \sum_{u \in \mathcal{S}} \mathbf{E}_u^{(k)} \mathcal{P}_\omega \boldsymbol{\delta}_u - \mu \mathcal{P}_{\omega_+} \sum_{u \in \mathcal{S}} \mathbf{E}_{*u}^{(k)} \mathcal{P}_\omega \boldsymbol{\delta}_u. \end{aligned}$$

Following the similar approach of proving (27), the inequality (28) is proved.

2) *The Proof of (29):* By the definition of  $\mathbf{F}_u^{(k)}$ ,

$$\begin{aligned} &\left\| \mathbf{E}_u^{(k)} \mathcal{P}_\omega \boldsymbol{\delta}_u \right\|^2 \\ &= \sum_v \left( \left| f^{(k)}(u) - f^{(k-\tau(u,v))}(u) \right|^2 |(\mathcal{P}_\omega \boldsymbol{\delta}_u)(v)|^2 \right) \\ &\leq \sum_v |(\mathcal{P}_\omega \boldsymbol{\delta}_u)(v)|^2 \left( \sum_{i=0}^{\tau-1} \left| f^{(k-i)}(u) - f^{(k-i-1)}(u) \right| \right)^2 \\ &\leq \left( \sum_{i=0}^{\tau-1} \left| f^{(k-i)}(u) - f^{(k-i-1)}(u) \right| \right)^2, \end{aligned}$$

where

$$\sum_v |(\mathcal{P}_\omega \boldsymbol{\delta}_u)(v)|^2 = \|\mathcal{P}_\omega \boldsymbol{\delta}_u\|^2 \leq 1.$$

Therefore,

$$\begin{aligned} \eta^{(k)} &\leq \sum_{u \in \mathcal{S}} \left\| \mathbf{E}_u^{(k)} \mathcal{P}_\omega \boldsymbol{\delta}_u \right\| \\ &\leq \sum_{u \in \mathcal{S}} \sum_{i=0}^{\tau-1} \left| f^{(k-i)}(u) - f^{(k-i-1)}(u) \right| \\ &\leq \sum_{i=0}^{\tau-1} \sqrt{|\mathcal{S}|} \left( \sum_{u \in \mathcal{S}} \left| f^{(k-i)}(u) - f^{(k-i-1)}(u) \right|^2 \right)^{\frac{1}{2}} \\ &\leq \sqrt{|\mathcal{S}|} \sum_{i=0}^{\tau-1} \left\| \mathbf{f}^{(k-i)} - \mathbf{f}^{(k-i-1)} \right\|, \end{aligned}$$

which is (29).

3) *The Proof of (30):* According to (4), (5), and (7),

$$\begin{aligned} &\mathbf{f}^{(k)} - \mathbf{f}^{(k-1)} \\ &= -\mu\beta \mathbf{f}^{(k-1)} + \mu \sum_{u \in \mathcal{S}} \left( \mathbf{F}_{*u}^{(k-1)} - \mathbf{F}_u^{(k-1)} \right) \mathcal{P}_\omega \boldsymbol{\delta}_u \\ &= -\mu\beta \mathbf{d}^{(k-1)} - \mu \mathbf{T} \mathbf{d}^{(k-1)} \\ &\quad + \mu \sum_{u \in \mathcal{S}} \mathbf{F}_{*u}^{(k-1)} \mathcal{P}_\omega \boldsymbol{\delta}_u - \mu \sum_{u \in \mathcal{S}} \mathbf{F}_u^{(k-1)} \mathcal{P}_\omega \boldsymbol{\delta}_u \\ &= -\mu(\beta \mathbf{I}_N + \mathbf{T}) \mathbf{d}^{(k-1)} \\ &\quad + \mu \sum_{u \in \mathcal{S}} \mathbf{E}_u^{(k-1)} \mathcal{P}_\omega \boldsymbol{\delta}_u - \mu \sum_{u \in \mathcal{S}} \mathbf{E}_{*u}^{(k-1)} \mathcal{P}_\omega \boldsymbol{\delta}_u. \end{aligned} \quad (44)$$

Taking the norm of (44), the inequality (30) is obtained.

### B. The Proof of Proposition 3

According to the proof of Proposition 1, setting  $\Delta = 0$  and using  $\|\mathbf{T}\| \leq 1$  and  $\beta_k \leq \beta_1$ , the inequalities (27)–(30) for variant  $\{\mu_k\}$  and  $\{\beta_k\}$  becomes

$$e^{(k)} \leq (1 - \mu_k \beta_k - \mu_k A) e^{(k-1)} + \mu_k e_+^{(k-1)} + \mu_k \eta^{(k-1)} \quad (45)$$

$$e_+^{(k)} \leq (1 - \mu_k \beta_k) e_+^{(k-1)} + \mu_k \eta^{(k-1)} \quad (46)$$

$$\eta^{(k)} \leq \sqrt{|\mathcal{S}|} \sum_{i=0}^{\tau-1} \delta^{(k-i)} \quad (47)$$

$$\delta^{(k)} \leq \mu_k \left( (\beta_1 + 1) (e^{(k-1)} + e_+^{(k-1)}) + \eta^{(k-1)} \right). \quad (48)$$

Similar to the proof of Proposition 1, it is easy to see that  $\{e^{(i)}\}$ ,  $\{e_+^{(i)}\}$  and  $\{\eta^{(i)}\}$  are bounded by constants  $B_e$ ,  $B_{e_+}$  and  $B_\eta$  for  $\mu_k = \mu_1/\sqrt{k}$ . Plugging (47) and (48) into (46), we have

$$e_+^{(k)} \leq (1 - \mu_k \beta_k) e_+^{(k-1)} + C' \mu_k \mu_{k-\tau}, \quad (49)$$

where  $C' = \tau \sqrt{|\mathcal{S}|} (2(B_e + B_{e_+}) + B_\eta)$ . For  $\mu_k = \mu_1/\sqrt{k}$  and  $\beta_k = \beta_1/\sqrt[4]{k}$ , if  $e_+^{(k-1)} \leq L_+/\sqrt[4]{k-1}$  is satisfied for a constant  $L_+$ , according to (49),  $e_+^{(k)} \leq L_+/\sqrt[4]{k}$  as long as the following inequality is satisfied,

$$\left( 1 - \frac{\mu_1 \beta_1}{\sqrt{k} \sqrt[4]{k}} \right) \frac{L_+}{\sqrt[4]{k-1}} + C' \frac{\mu_1^2}{\sqrt{k} \sqrt[4]{k-1}} \leq \frac{L_+}{\sqrt[4]{k}}.$$

The inequality above is equivalent to

$$\frac{\mu_1^2 C'}{L_+} \frac{\sqrt[4]{k} \sqrt[4]{k-1}}{\sqrt{k-\tau}} + \frac{\sqrt{k}}{(\sqrt[4]{k} + \sqrt[4]{k-1})(\sqrt{k} + \sqrt{k-1})} \leq \mu_1 \beta_1. \quad (50)$$

Because the first term of the left side approaches  $\mu_1^2 C'/L_+$  and the second term approaches 0 when  $k$  is large enough, by selecting a constant  $L_+$  appropriately, (50) is established and then we have  $e_+^{(k)} \leq L_+/\sqrt[4]{k}$ .

Plugging (47) and (48) into (45), we have

$$e^{(k)} \leq (1 - \mu_k \beta_k - \mu_k A) e^{(k-1)} + \mu_k e_+^{(k-1)} + C' \mu_k \mu_{k-\tau}.$$

If  $e^{(k-1)} \leq L/\sqrt[4]{k-1}$  is satisfied for a constant  $L$ ,  $e^{(k)} \leq L/\sqrt[4]{k}$  as long as the following inequality is satisfied,

$$\left(1 - \frac{\mu_1}{\sqrt{k}} \left(\frac{\beta_1}{\sqrt[4]{k}} + A\right)\right) \frac{L}{\sqrt[4]{k-1}} + \frac{\mu_1}{\sqrt{k}} \frac{L_+}{\sqrt[4]{k-1}} + C' \frac{\mu_1^2}{\sqrt{k} \sqrt{k-\tau}} \leq \frac{L}{\sqrt[4]{k}},$$

which is equivalent to

$$\frac{\mu_1 L_+}{L} + \frac{\mu_1^2 C'}{L} \frac{\sqrt[4]{k-1}}{\sqrt{k-\tau}} + \frac{\sqrt[4]{k}}{(\sqrt[4]{k} + \sqrt[4]{k-1})(\sqrt{k} + \sqrt{k-1})} \leq \mu_1 \left(A + \frac{\beta_1}{\sqrt[4]{k}}\right). \quad (51)$$

Both the second and third terms of (51) approach 0 when  $k$  is large enough. By selecting a constant  $L$  appropriately, (51) is established and then we have  $e^{(k)} \leq L/\sqrt[4]{k}$ .

Based on the analysis above, we have

$$\begin{aligned} \|\mathbf{f}^{(k)} - \mathbf{f}_*\| &\leq \frac{\beta_k}{\beta_k + A} \|\mathbf{f}_*\| + e_+^{(k)} + e^{(k)} \\ &\leq \left(\frac{\beta_1}{A} \|\mathbf{f}_*\| + L_+ + L\right) \frac{1}{\sqrt[4]{k}}, \end{aligned} \quad (52)$$

when  $k$  is large, and Proposition 3 is proved.

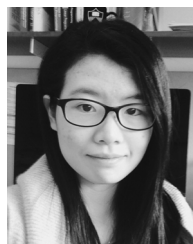
## REFERENCES

- [1] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," *IEEE Signal Process. Mag.*, vol. 30, no. 3, pp. 83–98, May 2013.
- [2] A. Sandryhaila and J. M. F. Moura, "Discrete signal processing on graphs," *IEEE Trans. Signal Process.*, vol. 61, no. 7, pp. 1644–1656, Jul. 2013.
- [3] X. Zhu and M. Rabbat, "Graph spectral compressed sensing for sensor networks," in *Proc. 37th IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2012, pp. 2865–2868.
- [4] S. K. Narang, Y. H. Chao, and A. Ortega, "Graph-wavelet filterbanks for edge-aware image processing," in *Proc. IEEE Stat. Signal Process. Workshop (SSP'12)*, 2012, pp. 141–144.
- [5] S. K. Narang, A. Gadde, and A. Ortega, "Signal processing techniques for interpolation in graph structured data," in *Proc. 38th IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2013, pp. 5445–5449.
- [6] S. K. Narang, A. Gadde, E. Sanou, and A. Ortega, "Localized iterative methods for interpolation in graph structured data," in *Proc. 1st IEEE Global Conf. Signal Inf. Process. (GlobalSIP)*, 2013, pp. 491–494.
- [7] A. Anis, A. Gadde, and A. Ortega, "Towards a sampling theorem for signals on arbitrary graphs," in *Proc. 39th IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2014, pp. 3892–3896.
- [8] A. Agaskar and Y. M. Lu, "A spectral graph uncertainty principle," *IEEE Trans. Inf. Theory*, vol. 59, no. 7, pp. 4338–4356, Jul. 2013.
- [9] S. Chen, A. Sandryhaila, J. M. F. Moura, and J. Kovacevic, "Adaptive graph filtering: Multiresolution classification on graphs," in *Proc. 1st IEEE Global Conf. Signal Inf. Process. (GlobalSIP)*, 2013, pp. 427–430.
- [10] D. K. Hammond, P. Vandergheynst, and R. Gribonval, "Wavelets on graphs via spectral graph theory," *Appl. Comput. Harmonic Anal.*, vol. 30, no. 2, pp. 129–150, 2011.
- [11] S. K. Narang and A. Ortega, "Perfect reconstruction two-channel wavelet filter-banks for graph structured data," *IEEE Trans. Signal Process.*, vol. 60, no. 6, pp. 2786–2799, Jun. 2012.
- [12] X. Zhu and M. Rabbat, "Approximating signals supported on graphs," in *Proc. 37th IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2012, pp. 3921–3924.
- [13] D. I. Shuman, M. J. Faraji, and P. Vandergheynst, "A framework for multiscale transforms on graphs," *arXiv preprint arXiv:1308.4942*, 2013.
- [14] V. N. Ekambara, G. C. Fanti, B. Ayazifar, and K. Ramchandran, "Multiresolution graph signal processing via circulant structures," in *Proc. IEEE Digital Signal Process., Signal Process. Educ. Meeting (DSP/SPE)*, 2013, pp. 112–117.
- [15] D. Thanou, D. I. Shuman, and P. Frossard, "Parametric dictionary learning for graph signals," in *Proc. 1st IEEE Global Conf. Signal Inf. Process. (GlobalSIP)*, 2013, pp. 487–490.
- [16] P. Liu, X. Wang, and Y. Gu, "Coarsening graph signal with spectral invariance," in *Proc. 39th IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2014, pp. 1075–1079.
- [17] P. Liu, X. Wang, and Y. Gu, "Graph signal coarsening: Dimensionality reduction in irregular domain," in *Proc. 2nd IEEE Global Conf. Signal Inf. Process. (GlobalSIP)*, 2014, pp. 966–970.
- [18] I. Pesenson, "Sampling in Paley-Wiener spaces on combinatorial graphs," *Trans. Amer. Math. Soc.*, vol. 360, no. 10, pp. 5603–5627, 2008.
- [19] I. Pesenson, "Variational splines and Paley-Wiener spaces on combinatorial graphs," *Construct. Approximat.*, vol. 29, pp. 1–21, 2009.
- [20] I. Z. Pesenson and M. Z. Pesenson, "Sampling, filtering and sparse approximations on combinatorial graphs," *J. Fourier Anal. Applicat.*, vol. 16, no. 6, pp. 921–942, 2010.
- [21] X. Wang, P. Liu, and Y. Gu, "Iterative reconstruction of graph signal in low-frequency subspace," in *Proc. 2nd IEEE Global Conf. Signal Inf. Process. (GlobalSIP)*, 2014, pp. 611–615.
- [22] X. Wang, P. Liu, and Y. Gu, "Local-set-based graph signal reconstruction," *arXiv preprint arXiv:1410.3944*, 2014.
- [23] Y. Bar-Shalom and X. R. Li, *Multitarget-multisensor tracking: Principles and techniques*. Storrs, CT, USA: Univ. of Connecticut, 1995.
- [24] C. Guestrin, P. Bodik, R. Thibaux, M. Paskin, and S. Madden, "Distributed regression: An efficient framework for modeling sensor network data," in *Proc. 3rd Int. Symp. Inf. Process. Sens. Netw. (IPSN)*, 2004, pp. 1–10.
- [25] M. Paskin, C. Guestrin, and J. McFadden, "A robust architecture for distributed inference in sensor networks," in *Proc. 4th Int. Symp. Inf. Process. Sens. Netw. (IPSN)*, 2005, pp. 55–62.
- [26] L. Xiao, S. Boyd, and S. Lall, "A scheme for robust distributed sensor fusion based on average consensus," in *Proc. 4th Int. Symp. Inf. Process. Sens. Netw. (IPSN)*, 2005, pp. 63–70.
- [27] I. D. Schizas, A. Ribeiro, and G. B. Giannakis, "Consensus in ad hoc WSNs with noisy links-Part I: Distributed estimation of deterministic signals," *IEEE Trans. Signal Process.*, vol. 56, no. 1, pp. 350–364, Jan. 2008.
- [28] R. Olfati-Saber, "Distributed Kalman filtering for sensor networks," in *Proc. 46th IEEE Conf. Decision and Control*, 2007, pp. 5492–5498.
- [29] F. S. Cattivelli, C. G. Lopes, and A. H. Sayed, "Diffusion recursive least-squares for distributed estimation over adaptive networks," *IEEE Trans. Signal Process.*, vol. 56, no. 5, pp. 1865–1877, May 2008.
- [30] F. S. Cattivelli and A. H. Sayed, "Diffusion LMS strategies for distributed estimation," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1035–1048, Mar. 2010.
- [31] D. I. Shuman, P. Vandergheynst, and P. Frossard, "Chebyshev polynomial approximation for distributed signal processing," in *Proc. 7th Int. Conf. Distrib. Comput. Sensor Syst. Workshops (DCOSS)*, 2011, pp. 1–8.
- [32] F. R. K. Chung, *Spectral graph theory*, Amer. Math. Soc., 1997.
- [33] T. Biyikoglu, J. Leydold, and P. F. Stadler, "Laplacian eigenvectors of graphs," in *Lecture Notes in Mathematics*. New York, NY, USA: Springer, 2007, vol. 1915.
- [34] O. Christensen, *An Introduction to Frames and Riesz Bases*. New York, NY, USA: Springer, 2003.
- [35] W. Cheney and A. Goldstein, "Proximity maps for convex sets," in *Proc. Amer. Math. Soc.*, 1959, vol. 10, no. 3, pp. 448–450.

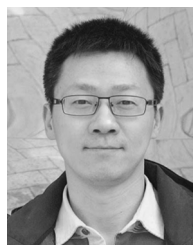
- [36] L. G. Gubin, B. T. Polyak, and E. V. Raik, "The method of projections for finding the common point of convex sets," *USSR Comput. Math. and Math. Phys.*, vol. 7, no. 6, pp. 1–24, 1967.
- [37] H. G. Feichtinger and K. Gröchenig, "Theory and practice of irregular sampling," *Wavelets: Math. Applicat.*, pp. 305–363, 1994.
- [38] K. Gröchenig, "A discrete theory of irregular sampling," *Linear Algebra and Its Applicat.*, vol. 193, pp. 129–150, 1993.
- [39] J. J. Benedetto, "Irregular sampling and frames," *Wavelets: Tutorial in Theory Applicat.*, vol. 2, pp. 445–507, 1992.
- [40] K. D. Sauer and J. P. Allebach, "Iterative reconstruction of bandlimited images from nonuniformly spaced samples," *IEEE Trans. Circuits Syst.*, vol. 34, no. 12, pp. 1497–1506, Dec. 1987.
- [41] F. Marvasti, *Nonuniform Sampling: Theory and Practice*. New York, NY, USA: Springer, 2001.
- [42] F. Marvasti, M. Analoui, and M. Gamshadzhahi, "Recovery of signals from nonuniform samples using iterative methods," *IEEE Trans. Signal Process.*, vol. 39, no. 4, pp. 872–878, Apr. 1991.
- [43] K. Gröchenig, "Reconstruction algorithms in irregular sampling," *Math. Comput.*, vol. 59, no. 199, pp. 181–194, 1992.
- [44] [Online]. Available: <http://www.cs.cmu.edu/gueststrin/Research/Data/>



**Xiaohan Wang** was born in Jilin, China, in 1988. He received the Bachelor of Engineering degree in electronic engineering from Tsinghua University, Beijing, China, in 2011. He is currently working towards the Ph.D. degree in electronic engineering at Tsinghua University. His current research interests include signal processing and complex networks.



**Mengdi Wang** received the B.S. degree in control engineering from Tsinghua University, Beijing, China, in 2007, and the S.M. and the Ph.D. degrees in electrical engineering and computer science from the Massachusetts Institute of Technology (MIT), Cambridge, in 2009 and 2013, respectively. Since 2013, she has been a member of the faculty of the Operations Research and Financial Engineering Department, Princeton University, where she is currently an Assistant Professor. Her research interests include optimization theory, optimization methods in statistical learning, stochastic optimization, and dynamic programming.



**Yuantao Gu** (S'02–M'03) received the B.S. degree from Xian Jiaotong University, Xian, China, in 1998, and the Ph.D. degree with honors from Tsinghua University, Beijing, China, in 2003, both in electronic engineering. He joined the faculty of Tsinghua in 2003 and is now an Associate Professor with Department of Electronic Engineering. He was a Visiting Scientist at Microsoft Research Asia, Beijing from December 2005 to February 2006, and Research Laboratory of Electronics at Massachusetts Institute of Technology (MIT), Cambridge, MA, from August 2012 to August 2013. He is the author/coauthor of two textbooks on Signals and Systems, and more than 50 papers on signal processing, multimedia communications, and wireless networks. His research interests include adaptive filtering, sparse signal recovery, multimedia signal processing, and related topics in wireless communications and information networks. Currently, he serves as Associate Editor of IEEE TRANSACTIONS ON SIGNAL PROCESSING and Handling Editor for the *EURASIP Digital Signal Processing*.