Local Measurement and Reconstruction for Noisy Bandlimited Graph Signals

Xiaohan Wang, Jiaxuan Chen, Yuantao Gu*

[a] *Department of Electronic Engineering and Tsinghua National Laboratory for Information Science and Technology (TNList), Tsinghua University, Beijing 100084, P. R. China.*

Abstract

Signals and information related to networks can be modeled and processed as graph signals. It has been shown that if a graph signal is smooth enough to satisfy certain conditions, it can be uniquely determined by its decimation on a subset of vertices. However, instead of the decimation, sometimes local combinations of signals on different sets of vertices are obtained in potential applications such as sensor networks with clustering structures. In this work, a generalized sampling scheme is proposed based on local measurement, which is a linear combination of signals associated with local vertices. It is proved that bandlimited graph signals can be perfectly reconstructed from the local measurements through a proposed iterative local measurement reconstruction (ILMR) algorithm. Some theoretical results related to ILMR including its convergence and denoising performance are given. Then the optimal partition of local sets and local weights are studied to minimize the error bound. It is shown that in noisy scenarios the proposed local measurement scheme is more robust than the traditional decimation scheme.

Keywords: Signal processing on graphs, generalized sampling, iterative reconstruction

☆This work was partially supported by National Natural Science Foundation of China (NSFC 61371137, 61571263, 61531166005, 51459003) and Tsinghua University Initiative Scientific Research Program (Grant 2014Z01005). Part of this work was presented at IEEE Global Conference on Signal and Information Processing (GlobalSIP), 2015 [1]. Manuscript submitted January, 2016.

*Corresponding author

Email address: gyt@tsinghua.edu.cn (Yuantao Gu)

1. Introduction

In recent years, graph-based signal processing has become an active research field due to the increasing demands for signal and information processing in irregular domains [2, 3]. For an N-vertex undirected graph $\mathcal{G}(\mathcal{V},\mathcal{E})$, where $\mathcal{V}$ denotes the vertex set and $\mathcal{E}$ denotes the edge set, if a real number is associated with each vertex of $\mathcal{G}$, these numbers on all the vertices constitute a graph signal $\mathbf{f} \in \mathbb{R}^N$. Potential applications of graph signal processing have been found in areas including sensor networks [4], semi-supervised learning [5], image processing [6], and structure monitoring [7].

A lot of concepts and techniques for classical signal processing are extended to graph signal processing. Related problems on graphs include graph signal filtering [8], graph wavelets [9, 10], graph signal compression [11, 12], uncertainty principle [13], graph signal coarsening [14, 15], phase transition [16], parametric dictionary learning [17, 18], graph topology learning [19], graph signal sampling and reconstruction [20, 21, 22, 23, 24, 25, 26, 27], and distributed algorithms [28, 29].

1.1. Motivation and Related Works

It is a natural problem to reconstruct smooth signals from partial observations on a graph in practical applications [8, 30]. In a scenario of environment monitoring by wireless sensor networks (WSNs), sometimes only parts of the nodes transmit data due to limited bandwidth or energy. By exploiting the smoothness of data, the missing entries can be estimated from the received ones, which can be modeled as the reconstruction of smooth signals on the graph from decimation. Especially, a sensor network with the hierarchical architecture is partitioned into multiple clusters. In each cluster, there is a node acting as the head and gathering data from all sensors inside the cluster. Different from regular sensors, cluster heads are equipped with long-distance-communication terminals, which send data to the center directly or in an ad-hoc manner. The collected data within a cluster are aggregated by the cluster head, which plays the role as a local measurement and can be naturally obtained. Each local measurement is a linear combination of the signals associated with a cluster of sensors. The cluster heads upload the linear combinations of data in the clusters and the center may recover the original data of all sensors in the WSN. Retrieving the raw data of all the nodes using the measured data from all the clusters can be modeled as a problem of smooth graph signal reconstruction from local measurements. This problem is studied in this work for the first time.

There have been several works focusing on the theory of the exact reconstruction of a bandlimited graph signal from its decimation. Sufficient conditions for unique reconstruction of bandlimited graph signals from decimation are given for normalized [31] and unnormalized Laplacian [32]. In [20], a necessary and sufficient condition on the cutoff frequency is established and the bandwidth is estimated based on the concept of spectral moments. Several algorithms are proposed to reconstruct graph signals from decimation. In [21], an algorithm named iterative least square reconstruction (ILSR) is proposed and the trade-off between data-fitting and smoothness is also considered. Two more efficient algorithms named iterative weighting reconstruction (IWR) and iterative propagating reconstruction (IPR) are proposed in [23] with much faster convergence.

The idea of local measurements can be traced back to time-domain nonuniform sampling [33], or irregular sampling [34, 35], which has a close relationship with graph signal sampling and reconstruction. For the signals in time-domain [36, 34], shift-invariant space [37], or on manifolds [38, 39], based on the theoretical results of signal reconstruction from samples, there have been extended works on reconstructing signals from local averages. Time-domain local averages are taken from small intervals around the samples with proper averaging functions. Theoretical results show that bandlimited original signals can be accurately recovered if the cutoff frequency is smaller than a quantity which is inversely proportional to the length of intervals [36]. However, there are few such works on graph-signal-related problems. As far as we know, the only work related to local aggregation for graph signals is applying the graph-shift operator sequentially [40], which is different from our problem.

Part of this work has been presented in [1]. This paper is the full version including all mathematical analysis and extensive discussions.

1.2. Contributions

In this paper, we first generalize the graph signal sampling scheme from *decimation* to *local measurement*. Based on this scheme, we then propose a new algorithm named iterative local measurement reconstruction (ILMR) to reconstruct the original signal from limited measurements. It is proved that if certain conditions are satisfied the bandlimited signal can always be exactly reconstructed from its local measurements. Moreover, we demonstrate that the traditional decimation scheme, which samples by vertex, along with its corresponding reconstruction algorithm is a special case of this work. Based on the performance analysis of ILMR, we find that the local measurement scheme is more robust than decimation in noisy scenarios. As a consequence, the optimal local weights in different noisy environments are discussed. The proposed sampling scheme has several advantages. First, it will benefit in the situation where local measurements are easier to obtain than the samples of specific vertices. Second, the proposed local measurement scheme is more robust against noise.

This paper is organized as follows. In section 2, the basis of graph signal processing and some existing algorithms for reconstructing graph signals from decimation are reviewed. The generalized sampling scheme, i.e. local measurement, is proposed in section 3. In section 4, the reconstruction algorithm ILMR is proposed and its convergence is proved. In section 5, the reconstruction performance in noisy scenarios is studied, and the optimal choice of local weight and local set partition is discussed. Experimental results are demonstrated in section 6, and the paper is concluded in section 7.

2. Preliminaries

2.1. Laplacian-based Graph Signal Processing and Bandlimited Graph Signals

Laplacian-based graph signal processing is considered in this work. The Laplacian [41] of an N-vertex undirected graph $\mathcal{G}$ is defined as

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

where $\mathbf{A}$ is the adjacency matrix of $\mathcal{G}$, and $\mathbf{D}$ is the degree matrix, which is a diagonal matrix whose entries are the degrees of the corresponding vertices.

Since $\mathcal{G}$ is undirected, its Laplacian is a symmetric and positive semi-definite matrix, and all of the eigenvalues of $\mathbf{L}$ are real and nonnegative. If $\mathcal{G}$ is connected, there is only one zero eigenvalue. Denote the eigenvalues of $\mathbf{L}$ as $0 = \lambda_1 < \lambda_2 \le \cdots \le \lambda_N$, and the corresponding eigenvectors as $\{\mathbf{u}_k\}_{1\le k\le N}$. The eigenvectors can also be regarded as graph signals on $\mathcal{G}$.

The Laplacian $\mathbf{L} : \mathbb{R}^N \to \mathbb{R}^N$ is an operator on the space of graph signals on $\mathcal{G}$,

$$(\mathbf{L}\mathbf{f})(u) = \sum_{v\in\mathcal{V}, u\sim v} \left(f(u) - f(v)\right), \quad \forall u \in \mathcal{V},$$

where $f(u)$ denotes the entry of $\mathbf{f}$ associated with vertex u, and $u \sim v$ denotes that there is an edge between vertices u and v. The Laplacian can be viewed as a kind of differential operator between vertices and their neighbors. Therefore, among the eigenvectors of $\mathbf{L}$, those associated with small eigenvalues have similar values on connected vertices, while the eigenvectors associated with large eigenvalues vary fast on the graph. In other words, eigenvectors associated with small eigenvalues are smooth and denote low-frequency signals on $\mathcal{G}$.

For graph Fourier transform [10], the eigenvectors $\{\mathbf{u}_k\}_{1\le k\le N}$ are regarded as the Fourier basis of the frequency-domain, and the eigenvalues $\{\lambda_k\}_{1\le k\le N}$ are regarded as frequencies. The graph Fourier transform is

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

where $\hat{f}(k)$ is the strength corresponding to the frequency λ_k.

Similar to its counterpart in time-domain, if a graph signal $\mathbf{f}$ is smooth on $\mathcal{G}$, $\mathbf{f}$ can be uniquely determined by its entries on a limited number of sampled vertices. Based on the graph Laplacian, the smoothness of a graph signal is usually described as being within a bandlimited subspace. A graph signal $\mathbf{f} \in \mathbb{R}^N$ is ω-bandlimited if

$$\mathbf{f} \in PW_\omega(\mathcal{G}) \triangleq \mathrm{span}\{\mathbf{u}_i | \lambda_i \leq \omega\}$$

The ω-bandlimited subspace $PW_\omega(\mathcal{G})$ is called Paley-Wiener space on $\mathcal{G}$ [31]. Another way to describe the smoothness of low-frequency subspace is by the number of eigenvalues within it. However, according to the formulation of the Laplacian operator, the magnitudes of eigenvalues reflect the smoothness of graph signals better than the rank of the eigenvalues does. Therefore we prefer using the cutoff frequency ω to describe the low-frequency subspaces. Different graph topologies may lead to various dimensions of the Paley-Wiener spaces.

2.2. Reconstruction from Decimation of Bandlimited Graph Signals

There have been theoretical analysis and algorithms on the reconstruction from decimation of bandlimited graph signals. Existing results show that $\mathbf{f} \in PW_\omega(\mathcal{G})$ can be uniquely reconstructed from its entries $\{f(u)\}_{u\in\mathcal{S}}$ on a sampling vertex set $\mathcal{S} \subseteq \mathcal{V}$ under certain conditions. Typical reconstruction algorithms include ILSR [21] and IPR [23]. The latter one is based on an important concept of *local sets* and converges faster.

Definition 1 (local sets *[23]*). *For a sampling set $\mathcal{S}$ on a graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$, assume that disjoint local sets $\{\mathcal{N}(u)\}_{u\in\mathcal{S}}$ associated with the sampled vertices is a partition of $\mathcal{V}$. For each $u \in \mathcal{S}$, denote the subgraph of $\mathcal{G}$ restricted to $\mathcal{N}(u)$ by $\mathcal{G}_{\mathcal{N}(u)}$, which is composed of vertices in $\mathcal{N}(u)$ and edges between them in $\mathcal{E}$. For each $u \in \mathcal{S}$, its local set satisfies $u \in \mathcal{N}(u)$, and the subgraph $\mathcal{G}_{\mathcal{N}(u)}$ is connected.*

The property of a local set is measured by *maximal multiple number* and *radius*, as follows.

Definition 2 (maximal multiple number *[23]***).** *Denoting $\mathcal{T}(u)$ as the shortest-path tree of $\mathcal{G}_{\mathcal{N}(u)}$ rooted at u, for $v \sim u$ in $\mathcal{T}(u)$, $\mathcal{T}_u(v)$ is the subtree composed by v and its descendants in $\mathcal{T}(u)$. The* maximal multiple number *of $\mathcal{N}(u)$ is*

$$K(u) = \max_{v \sim u \text{ in } \mathcal{T}(u)} |\mathcal{T}_u(v)|.$$

Definition 3 (radius *[23]***).** *The* radius *of $\mathcal{N}(u)$ is the maximal distance of vertex in $\mathcal{G}_{\mathcal{N}(u)}$ from u, denoted as*

$$R(u) = \max_{v \in \mathcal{N}(u)} dist(v, u),$$

where the distance is the number of edges in the shortest path connecting the two vertices.

Theorem 1 (IPR *[23]***).** *For a given sampling set $\mathcal{S}$ and associated local sets $\{\mathcal{N}(u)\}_{u\in\mathcal{S}}$ on a graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$, $\forall \mathbf{f} \in PW_\omega(\mathcal{G})$, if ω is less than $1/Q_{\max}^2$, $\mathbf{f}$ can be reconstructed by its decimation $\{f(u)\}_{u\in\mathcal{S}}$ through the IPR method*

$$\mathbf{f}^{(0)} = \mathcal{P}_\omega\left(\sum_{u\in\mathcal{S}} f(u)\boldsymbol{\delta}_{\mathcal{N}(u)}\right),$$

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

where

$$Q_{\max} = \max_{u\in\mathcal{S}} \sqrt{K(u)R(u)},$$

$\mathcal{P}_\omega(\cdot)$ is the projection operator onto $PW_\omega(\mathcal{G})$, and $\boldsymbol{\delta}_{\mathcal{N}(u)}$ denotes the graph signal with entries

$$\delta_{\mathcal{N}(u)}(v) = \begin{cases} 1, & v \in \mathcal{N}(u); \\ 0, & v \notin \mathcal{N}(u). \end{cases}$$

3. Local Measurement: A Generalized Sampling Scheme

We consider a new sampling scheme of measuring by local sets. In this scheme, all the vertices in a graph are partitioned into disjoint clusters. In each cluster, there is no specific sampling vertex, but all vertices in this cluster contribute to produce a measurement. For this purpose, *centerless local sets* are firstly introduced based on Definition 1.

Definition 4 (centerless local sets). *For a graph $\mathcal{G}(\mathcal{V},\mathcal{E})$, assume that disjoint local sets $\{\mathcal{N}_i\}_{i\in\mathcal{I}}$ is a partition of $\mathcal{V}$, where $\mathcal{I}$ denotes the index set of divisions. Each subgraph $\mathcal{G}_{\mathcal{N}_i}$, which denotes the subgraph of $\mathcal{G}$ restricted to $\mathcal{N}_i$, is connected.*

One should notice that the centerless local sets play important roles in the proposed generalized sampling scheme, while the local sets do not in the traditional decimation scheme. In the decimation scheme, the local sets are designed for specific reconstruction algorithms and have no effect in the sampling process. However, in the generalized sampling scheme, the centerless local sets are elaborated for sampling and determine the performance of reconstruction, which will be discussed in section 5.

To evaluate the partition of a graph, the *diameter* of a centerless local set is defined and will be utilized in the next section.

Definition 5 (diameter). *For a centerless local set $\mathcal{N}_i$, its diameter is defined as the largest distance of two vertices in $\mathcal{G}_{\mathcal{N}_i}$, i.e.,*

$$D_i = \max_{u,v\in\mathcal{N}_i} dist(u,v).$$

In order to produce a measurement from specific centerless local set, a *local weight* is defined to balance the contribution of all vertices in this set and to obstruct the energy from other parts of the graph.

Definition 6 (local weight). *A local weight $\boldsymbol{\varphi}_i \in \mathbb{R}^N$ associated with a centerless local set $\mathcal{N}_i$ satisfies*

$$\varphi_i(v)\begin{cases}\geq 0, v\in\mathcal{N}_i\\ = 0, v\notin\mathcal{N}_i\end{cases}$$

and

$$\sum_{v\in\mathcal{N}_i}\varphi_i(v) = 1.$$

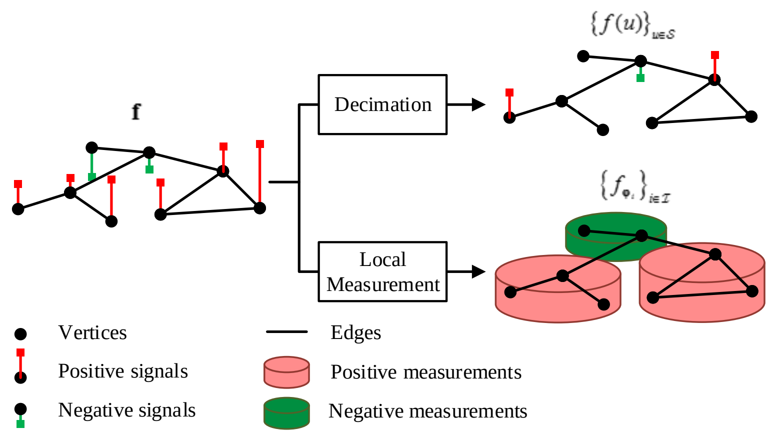

Figure 1: An illustration of the traditional sampling (decimation) scheme versus the generalized sampling (local measurement) scheme. For each centerless local set, a local measurement is produced by a linear combination of signals associated with vertices within this set.

Finally, we arrive at the definition of *local measurement* by linearly combining the signals in each centerless local set using preassigned local weights.

Definition 7 (local measurement). *For given centerless local sets and the associated local weights* $\{(\mathcal{N}_i, \boldsymbol{\varphi}_i)\}_{i\in\mathcal{I}}$*, a set of local measurements for a graph signal* $\mathbf{f}$ *is* $\{f_{\boldsymbol{\varphi}_i}\}_{i\in\mathcal{I}}$*, where*

$$f_{\boldsymbol{\varphi}_i} \triangleq \langle \mathbf{f}, \boldsymbol{\varphi}_i \rangle = \sum_{v\in\mathcal{N}_i} f(v)\varphi_i(v).$$

The sampling schemes of decimation and of local measurement are visualized in Fig. 1. Compared with decimation in previous works [23, 31], local measurement can be regarded as a generalized sampling scheme. The local measurement scheme is to obtain linear combinations of the signals in each local set, while the decimation scheme is to obtain the signals on selected vertices in the sampling set $\mathcal{S}$. Both sampling schemes take the inner products of the original signal and specified local weights. Decimation can be regarded as a special case of local measurement, in which only the sampled vertices have weight 1 and other vertices in centerless local sets have weights 0.

We highlight that the sets, weights and measurements are *local* rather than *global*, which comes from some natural observations. It is partially because locality and local operations are basic features of graphs and complex networks. Moreover, signal processing on graphs may be dependent on distributed implementation, where local operations are more feasible than global ones.

4. ILMR: Reconstruct Signal from Local Measurements

We will show that under certain conditions the original signal $\mathbf{f}$ can be uniquely and exactly reconstructed from the local measurements $\{f_{\boldsymbol{\varphi}_i}\}_{i\in\mathcal{I}}$.

First of all, an operator is defined based on centerless local sets and the associated local weights.

Definition 8. *For given centerless local sets and the associated weights* $\{(\mathcal{N}_i, \boldsymbol{\varphi}_i)\}_{i\in\mathcal{I}}$ *on a graph* $\mathcal{G}(\mathcal{V},\mathcal{E})$*, an operator* $\mathbf{G}$ *is defined by*

$$\mathbf{Gf} = \mathcal{P}_\omega\left(\sum_{i\in\mathcal{I}}\langle\mathbf{f},\boldsymbol{\varphi}_i\rangle\boldsymbol{\delta}_{\mathcal{N}_i}\right) \tag{1}$$

$$= \sum_{i\in\mathcal{I}}\langle\mathbf{f},\boldsymbol{\varphi}_i\rangle\mathcal{P}_\omega(\boldsymbol{\delta}_{\mathcal{N}_i}), \tag{2}$$

where $\boldsymbol{\delta}_{\mathcal{N}_i}$ *is defined as*

$$\delta_{\mathcal{N}_i}(v) = \begin{cases} 1, & v\in\mathcal{N}_i; \\ 0, & v\notin\mathcal{N}_i. \end{cases} \tag{3}$$

For a graph signal, the proposed operator is to calculate the local measurement in each centerless local set, then to assign the local measurement to all the vertices in that set, and finally to filter out the component beyond the bandwidth, i.e., (1). Equivalently, it denotes a linear combination of all low-frequency parts of $\{\boldsymbol{\delta}_{\mathcal{N}_i}\}_{i\in\mathcal{I}}$, with the combination coefficients as the local measurements of corresponding local sets, i.e., (2).

The following lemma shows that the proposed operator is bounded in $PW_\omega(\mathcal{G})$ under certain conditions.

Lemma 1. *For given centerless local sets and the associated weights* $\{(\mathcal{N}_i, \boldsymbol{\varphi}_i)\}_{i\in\mathcal{I}}$, $\forall \mathbf{f} \in PW_\omega(\mathcal{G})$*, the following inequality holds,*

$$\|\mathbf{f} - \mathbf{G}\mathbf{f}\| \le C_{\max}\sqrt{\omega}\|\mathbf{f}\|,$$

where

$$C_{\max} = \max_{i\in\mathcal{I}} \sqrt{|\mathcal{N}_i| D_i},$$

$|\cdot|$ *denotes cardinality, and* D_i *is defined in Definition 5.*

The proof of Lemma 1 is postponed to Appendix 8.1. Lemma 1 shows that the operator $(\mathbf{I} - \mathbf{G})$ is a contraction mapping in $PW_\omega(\mathcal{G})$ if ω is less than $1/C_{\max}^2$.

Based on Lemma 1, it is shown in Proposition 1 that the original signal can be reconstructed from its local measurements.

Proposition 1. *For given centerless local sets and the associated weights* $\{(\mathcal{N}_i, \boldsymbol{\varphi}_i)\}_{i\in\mathcal{I}}$, $\forall \mathbf{f} \in PW_\omega(\mathcal{G})$*, where* ω *is less than* $1/C_{\max}^2$, $\mathbf{f}$ *can be reconstructed from its local measurements* $\{f_{\boldsymbol{\varphi}_i}\}_{i\in\mathcal{I}}$ *through an iterative local measurement reconstruction (ILMR) algorithm in Table 1, with the error at the* k*th iteration satisfying*

$$\|\mathbf{f}^{(k)} - \mathbf{f}\| \le \gamma^k \|\mathbf{f}^{(0)} - \mathbf{f}\|,$$

where

$$\gamma = C_{\max}\sqrt{\omega}. \tag{6}$$

Proof: According to the definition of $\mathbf{G}$, the iteration (5) can be rewritten as

$$\mathbf{f}^{(k+1)} = \mathbf{f}^{(k)} + \mathbf{G}(\mathbf{f} - \mathbf{f}^{(k)}). \tag{7}$$

Note that $\mathbf{f} \in PW_\omega(\mathcal{G})$ and $\mathbf{f}^{(k)} \in PW_\omega(\mathcal{G})$ for any k, then $\mathbf{f}^{(k)} - \mathbf{f} \in PW_\omega(\mathcal{G})$. As a consequence of Lemma 1,

$$\|\mathbf{f}^{(k+1)} - \mathbf{f}\| = \|(\mathbf{f}^{(k)} - \mathbf{f}) - \mathbf{G}(\mathbf{f}^{(k)} - \mathbf{f})\| \le \gamma \|\mathbf{f}^{(k)} - \mathbf{f}\|.$$

$\square$

Table 1: Iterative Local Measurement Reconstruction.

Input: Graph $\mathcal{G}$, cutoff frequency ω, centerless local sets $\{\mathcal{N}_i\}_{i\in\mathcal{I}}$, local weights $\{\boldsymbol{\varphi}_i\}_{i\in\mathcal{I}}$, local measurements $\{f_{\boldsymbol{\varphi}_i}\}_{i\in\mathcal{I}}$;

Output: Interpolated signal $\mathbf{f}^{(k)}$;

Initialization:

$$\mathbf{f}^{(0)} = \mathcal{P}_\omega\left(\sum_{i\in\mathcal{I}} f_{\boldsymbol{\varphi}_i}\boldsymbol{\delta}_{\mathcal{N}_i}\right); \tag{4}$$

Loop:

$$\mathbf{f}^{(k+1)} = \mathbf{f}^{(k)} + \mathcal{P}_\omega\left(\sum_{i\in\mathcal{I}}(f_{\boldsymbol{\varphi}_i} - \langle \mathbf{f}^{(k)}, \boldsymbol{\varphi}_i\rangle)\boldsymbol{\delta}_{\mathcal{N}_i}\right); \tag{5}$$

Until: The stop condition is satisfied.

Proposition 1 shows that a signal $\mathbf{f}$ is uniquely determined and can be reconstructed by its local measurements $\{f_{\boldsymbol{\varphi}_i}\}_{i\in\mathcal{I}}$ if $\{\boldsymbol{\varphi}_i\}_{i\in\mathcal{I}}$ are known. The quantity $(f_{\boldsymbol{\varphi}_i} - \langle \mathbf{f}^{(k)}, \boldsymbol{\varphi}_i\rangle)$ is the estimate error between the original measurement and the reconstructed measurement at the kth iteration. According to (7), in each iteration of ILMR, the new increment of the interpolated signal is obtained by first assigning the estimate errors to all vertices in the associated centerless local sets, and then projecting it onto the ω-bandlimited subspace.

Considering (18) and (19) in the proof of Lemma 1, one has

$$\|\mathbf{f} - \mathbf{G}\mathbf{f}\|^2 \le \sum_{i\in\mathcal{I}}\left(\sum_{v\in\mathcal{N}_i}|f(v) - \langle \mathbf{f}, \boldsymbol{\varphi}_i\rangle|^2\right).$$

The RHS of the above inequality shows that the choice of $\boldsymbol{\varphi}_i$ for each centerless local set is independent. Therefore we may look into $\sum_{v\in\mathcal{N}_i}|f(v) - \langle \mathbf{f}, \boldsymbol{\varphi}_i\rangle|^2$ for any fixed i. Denoting $\boldsymbol{\varphi}_i^*$ as the optimal weights, one may readily arrive at

$$\langle \boldsymbol{f}, \boldsymbol{\varphi}_i^*\rangle = \arg\min_x \sum_{v\in\mathcal{N}_i}|f(v) - x|^2 = \frac{1}{|\mathcal{N}_i|}\sum_{v\in\mathcal{N}_i} f(v).$$

As a consequence, the uniform weights $\boldsymbol{\varphi}_i^*(v) = 1/|\mathcal{N}_i|, \forall i$ minimize the RHS of the above inequality, which leads to the sharpest bound and may accelerate convergence.

Except for the difference of decimation and local measurement, the basic idea of ILMR is similar to that of IPR [23], which is an algorithm of reconstructing graph signals from decimation. The procedures of IPR and ILMR in each iteration are illustrated in Fig. 2. In the assignment or propagating step, ILMR assigns the estimate errors of local measurements to vertices within the local sets, while IPR propagates the estimate errors of the decimated signal on the sampled vertices to other vertices in the local sets. In fact, ILMR degenerates to IPR if the local weight concentrates on only one vertex (the sampled vertex) in each local set, in which case the local measurement degenerates to decimation.

The sufficient conditions and error bounds for ILMR and IPR are also different. Suppose the (centerless) local sets divisions in ILMR and IPR are exactly the same, i.e. the sampling set $\mathcal{S}$ in IPR can be written as $\{u_i\}_{i\in\mathcal{I}}$, where $\mathcal{I}$ is the index set in ILMR, then $\mathcal{N}_i$ equals $\mathcal{N}(u_i)$ for all $i \in \mathcal{I}$. According to Definition 2 and 3, we have $R(u_i) \leq D_i$ and $K(u_i) \leq |\mathcal{N}(u_i)| = |\mathcal{N}_i|$. Therefore, $C_{\max}$ is not less than $Q_{\max}$. It implies that a more strict condition is needed for ILMR. It is reasonable because the sufficient condition for ILMR to guarantee the reconstruction is for all of the choices of local weights, which include decimation as a special case. However, since both sufficient conditions in Theorem 1 and Proposition 1 are not tight and there is still room for refinement, such a comparison only provides a rough analysis.

Remark 1. *The projection operator $\mathcal{P}_\omega(\cdot)$ can be approximated by a polynomial expansion of the Laplacian, which is localized. As a consequence, ILMR can be approximately implemented in a localized way. In detail, the projection operator is written as*

$$\mathcal{P}_\omega(\mathbf{f}) = \mathbf{U}\mathrm{diag}\left\{\hat{h}(\lambda_1), \cdots, \hat{h}(\lambda_N)\right\}\mathbf{U}^{\mathrm{T}}\mathbf{f},$$

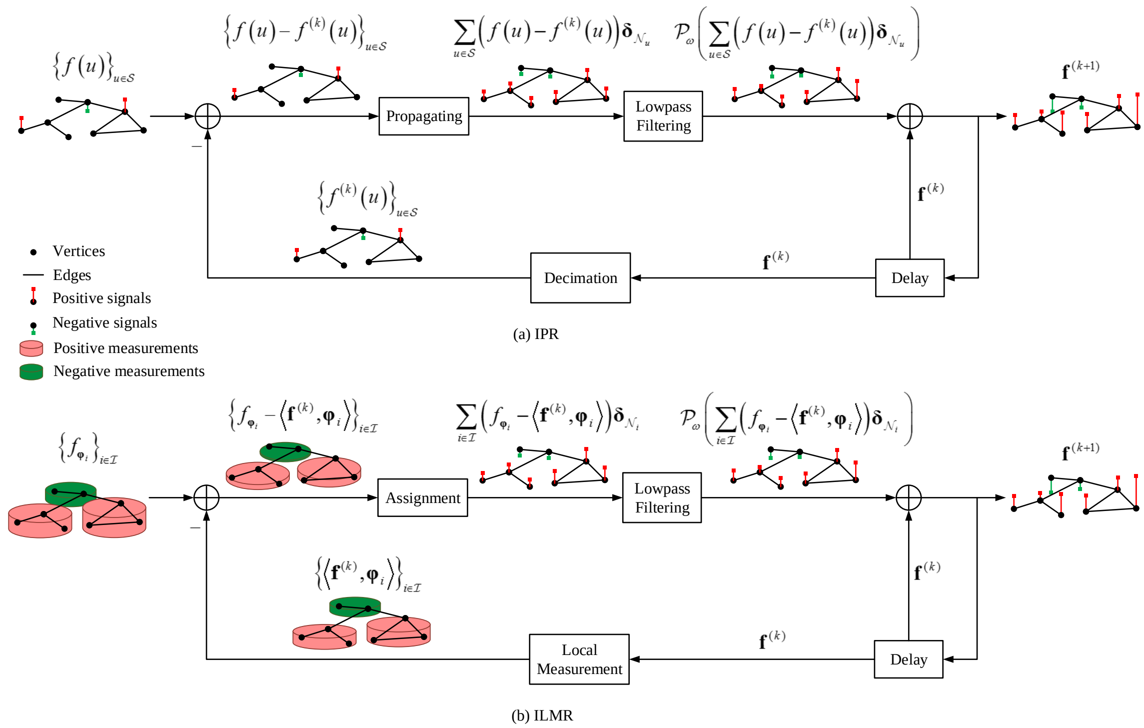

Figure 2: The procedures of IPR and ILMR. The former algorithm is to reconstruct a bandlimited signal from decimation, while the latter reconstructs a signal from local measurements. Essentially, ILMR becomes IPR if the local weights concentrate on only one vertex of each local set, in which case local measurement degenerates to decimation.

where $\hat{h}(\cdot)$ denotes the lowpass filter

$$\hat{h}(\lambda) = \begin{cases} 1, & \lambda \le \omega; \\ 0, & \text{elsewhere.} \end{cases}$$

Utilizing a polynomial approximation of $\hat{h}(\cdot)$ (e.g. Chebyshev polynomial expansion [10,21]), one has

$$\hat{h}(\lambda) \approx \sum_{j=0}^{k} \alpha_j \lambda^j, 0 \le \lambda \le \lambda_N,$$

where $\{\alpha_j\}$ denote the coefficients and k is the order of the approximation, which is usually far smaller than N. Therefore the projection is approximated *by a polynomial expansion of the Laplacian*

$$\mathcal{P}_\omega(\mathbf{f}) \approx \mathbf{U}\mathrm{diag}\left\{\sum_{j=0}^{k}\alpha_j\lambda_1^j, \cdots, \sum_{j=0}^{k}\alpha_j\lambda_N^j\right\}\mathbf{U}^{\mathrm{T}}\mathbf{f} = \sum_{j=0}^{k}\alpha_j\mathbf{L}^j\mathbf{f}.$$

Because the Laplacian operator can be conducted by each vertex and its neighbors, the projection operator is approximately localized.

Remark 2. *For potential applications, if the local measurements come from the result of some repeatable physical operations, the local weights are even not necessarily known when conducting ILMR. In detail, if $\{\boldsymbol{\varphi}_i\}_{i\in\mathcal{I}}$ is unknown but fixed, i.e., the local measurement operation in Fig. 2(b) is a black box, $\langle\mathbf{f}^{(k)}, \boldsymbol{\varphi}_i\rangle$ can also be obtained by conducting the physical operations in each iteration. Therefore, the original signal can still be reconstructed by ILMR without exactly knowing $\{\boldsymbol{\varphi}_i\}_{i\in\mathcal{I}}$. This is a rather interesting result, and may facilitate graph signal reconstruction in specific scenarios.*

Remark 3. *If the bandlimited space is described as a subspace with a known dimensionality, rather than the cutoff frequency ω, the perfect reconstruction is achievable as a closed form by solving linear equations. However, the value of the iterative algorithm relies on its locality, which is important in graph related problems. Furthermore, iterative algorithms can be applied in potential online and distributed scenarios.*

5. Performance Analysis

In this section, we study the error performance of ILMR when the original signal is corrupted by additive noise. We first derive the reconstruction error for incorrect measurement. Then the expected reconstruction error is calculated under the assumption of independent Gaussian noises and the optimal local weight is obtained in the sense of minimizing the expected reconstruction error bound. Finally, in a special case of *i.i.d.* Gaussian perturbation, a greedy method for the centerless local sets partition and the selection of optimal local weights are provided.

5.1. Reconstruction Error in the Noisy Scenario

Suppose that the observed signal associated with each vertex is corrupted by additive noise. The corrupted signal is denoted as $\tilde{\mathbf{f}} = \mathbf{f} + \mathbf{n}$, where $\mathbf{n}$ denotes the noise. In the kth iteration of ILMR, the corrupted local measurements $\{\langle \tilde{\mathbf{f}}, \boldsymbol{\varphi}_i \rangle\}_{i\in\mathcal{I}}$ are utilized to produce the temporary reconstruction of $\tilde{\mathbf{f}}^{(k)}$.

The following lemma gives a reconstruction error bound of $\tilde{\mathbf{f}}^{(k)}$.

Proposition 2. *For given centerless local sets and the associated weights $\{(\mathcal{N}_i, \boldsymbol{\varphi}_i)\}_{i\in\mathcal{I}}$, $\mathbf{f} \in PW_\omega(\mathcal{G})$ is corrupted by additive noise $\mathbf{n}$. If ω is less than $1/C_{\max}^2$, in the kth iteration the output of ILMR using the corrupted local measurements $\{\langle \tilde{\mathbf{f}}, \boldsymbol{\varphi}_i \rangle\}_{i\in\mathcal{I}}$ satisfies*

$$\|\tilde{\mathbf{f}}^{(k)} - \mathbf{f}\| \leq \frac{\tilde{n}}{1-\gamma} + \gamma^{k+1}\left(\|\mathbf{f}\| + \|\mathbf{n}\|\right), \tag{8}$$

where γ is defined as (6), $\tilde{n}$ is defined as

$$\tilde{n} = \sum_{i\in\mathcal{I}} \sqrt{|\mathcal{N}_i|} \cdot |n_i|, \tag{9}$$

and n_i is the equivalent noise of centerless local set $\mathcal{N}_i$, defined as

$$n_i = \langle \mathbf{n}, \boldsymbol{\varphi}_i \rangle = \sum_{v\in\mathcal{N}_i} n(v)\varphi_i(v). \tag{10}$$

The proof of Proposition 2 is postponed to Appendix 8.2.

From (8) it can be seen that in the noisy scenario the reconstruction error is controlled by the sum of two parts. The first one is a weighted sum of the equivalent noises of all the local sets, while the second one is decaying with the increase of iteration number. The first part is crucial as the iteration goes on. Thus minimizing the first part, which is determined by both partition of centerless local sets and local weights, improves the performance of ILMR in the noisy scenario.

5.2. Gaussian Noise and Optimal Local Weights

For a given partition $\{\mathcal{N}_i\}_{i\in\mathcal{I}}$, some prior knowledge of unknown noise $\mathbf{n}$ brings the possibility to design optimal local weights [1]. We assume the noises associated with different vertices are independent.

Suppose the noise follows zero-mean Gaussian distribution, i.e., $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma})$, where $\mathbf{\Sigma}$ is a diagonal matrix and the noise of vertex v satisfies $n(v) \sim \mathcal{N}(0, \sigma^2(v))$. Then $\tilde{n}$ defined in (9) is a random variable.

For centerless local set $\mathcal{N}_i$, according to (10), the equivalent noise n_i also follows a Gaussian distribution $n_i \sim \mathcal{N}(0, \sigma_i^2)$, where

$$\sigma_i^2 = \sum_{v\in\mathcal{N}_i} \sigma^2(v)\varphi_i^2(v). \tag{11}$$

Then $|n_i|$ follows the half-normal distribution with its expectation satisfying

$$\mathrm{E}\left\{|n_i|\right\} = \sigma_i\sqrt{\frac{2}{\pi}}.$$

According to (9), the expectation of $\tilde{n}$ is

$$\mathrm{E}\{\tilde{n}\} = \sqrt{\frac{2}{\pi}}\sum_{i\in\mathcal{I}} \sqrt{|\mathcal{N}_i|}\sigma_i. \tag{12}$$

Then the following corollary is ready to obtain.

Corollary 1. *For given centerless local sets and the associated weights $\{(\mathcal{N}_i, \boldsymbol{\varphi}_i)\}_{i\in\mathcal{I}}$, the original signal $\mathbf{f} \in PW_\omega(\mathcal{G})$, assuming the noise associated with vertex v follows independent Gaussian distribution $\mathcal{N}(0, \sigma^2(v))$, if ω is less than $1/C_{\max}^2$, the expected reconstruction error of ILMR in the kth iteration satisfies*

$$\mathrm{E}\left\{\|\tilde{\mathbf{f}}^{(k)} - \mathbf{f}\|\right\} \le \frac{1}{1-\gamma}\sqrt{\frac{2}{\pi}}\sum_{i\in\mathcal{I}}\sqrt{|\mathcal{N}_i|}\sigma_i + \mathcal{O}\left(\gamma^{k+1}\right), \tag{13}$$

where γ is defined as (6), and σ_i is defined as (11).

[1] In fact, the optimal local weights can also be studied in other criterions, e.g. the fastest convergence. Here we consider the optimal local weights in the sense of minimizing the expected reconstruction error bound.

Corollary 1 is ready to be proved by plugging (11) and (12) in the expectation of (8).

By minimizing the right hand side of (13), the optimal choice of local weights can be derived.

Corollary 2. *For given division of centerless local sets* $\{\mathcal{N}_i\}_{i\in\mathcal{I}}$*, if the noises associated with the vertices are independent and follow zero-mean Gaussian distributions* $n(v) \sim \mathcal{N}(0, \sigma^2(v))$*, then the optimal local weights* $\{\boldsymbol{\varphi}_i\}_{i\in\mathcal{I}}$ *are*

$$\varphi_i(v) = \begin{cases} \dfrac{(\sigma^2(v))^{-1}}{\sum_{v\in\mathcal{N}_i}(\sigma^2(v))^{-1}}, & v \in \mathcal{N}_i; \\ 0, & v \notin \mathcal{N}_i. \end{cases} \tag{14}$$

Proof: Minimizing the right hand side of (13) is equivalent to minimizing σ_i for each local set $\mathcal{N}_i$. By the Cauchy-Schwarz inequality, one has

$$\begin{aligned} \left(\sum_{v\in\mathcal{N}_i}(\sigma^2(v))^{-1}\right)\sigma_i^2 &= \left(\sum_{v\in\mathcal{N}_i}(\sigma^2(v))^{-1}\right)\left(\sum_{v\in\mathcal{N}_i}\sigma^2(v)\varphi_i^2(v)\right) \\ &\geq \left(\sum_{v\in\mathcal{N}_i}\varphi_i(v)\right)^2 = 1. \end{aligned}$$

Therefore,

$$\sigma_i^2 \geq \frac{1}{\sum_{v\in\mathcal{N}_i}(\sigma^2(v))^{-1}}. \tag{15}$$

The equality of (15) holds if and only if (14) is satisfied. □

The above analysis shows that in the sense of minimizing the expected reconstruction error, the optimal local weight associated with vertex v within $\mathcal{N}_i$ is inversely proportional to the noise variance of v. This is evident because more information are reserved in the sampling process if a larger local weight is assigned to a vertex with smaller noise variance. However, it should be noted that compared with the optimal local measurement, assigning all the weights in $\mathcal{N}_i$ to the vertex with the smallest noise variance, i.e. the optimal decimation, is not the best choice. In fact, the optimal choice of local measurements is consistent with the well-known inverse variance weighting in statistics [42].

Therefore, local measurement reduces the disturbance of noise and reconstruct the original signal more precisely. In other words, for given partition of centerless local sets, graph signal reconstruction from local measurements with the optimal weights performs better than reconstruction from decimation, even when the vertices with the smallest noise variance are chosen in the latter sampling scheme.

5.3. A Special Case of Independent and Identical Distributed Gaussian Noise

Specifically, if noise variances are the same for all the vertices, i.e., $\sigma(v)$ equals σ for any $v \in \mathcal{V}$, $\tilde{n}$ can be approximately written in a more explicit form. For $\mathcal{N}_i$, the optimal local weight is equal for all the vertices in $\mathcal{N}_i$. Thus $\varphi_i(v)$ equals $1/|\mathcal{N}_i|$ for $v \in \mathcal{N}_i$, and in this case, $\sqrt{|\mathcal{N}_i|}n_i$ follows a Gaussian distribution,

$$\sqrt{|\mathcal{N}_i|}n_i \sim \mathcal{N}(0, \sigma^2).$$

Then $\sqrt{|\mathcal{N}_i|} \cdot |n_i|$ follows the half-normal distribution with the same parameter σ. The above analysis shows that each term of the sum in (9) follows independent and identical half-normal distribution, with its expectation and variance satisfying

$$\mathrm{E}\left\{\sqrt{|\mathcal{N}_i|} \cdot |n_i|\right\} = \sigma\sqrt{\frac{2}{\pi}},$$
$$\mathrm{Var}\left\{\sqrt{|\mathcal{N}_i|} \cdot |n_i|\right\} = \sigma^2\left(1 - \frac{2}{\pi}\right).$$

Assuming that the number of local sets $|\mathcal{I}|$ is large, by the central limit theorem, $\tilde{n}$ follows a Gaussian distribution approximately,

$$\tilde{n} \sim \mathcal{N}\left(|\mathcal{I}|\sigma\sqrt{\frac{2}{\pi}}, |\mathcal{I}|\sigma^2\left(1 - \frac{2}{\pi}\right)\right).$$

Then we have the following corollary.

Corollary 3. *For given centerless local sets $\{\mathcal{N}_i\}_{i\in\mathcal{I}}$ and the associated weights $\varphi_i(v) = 1/|\mathcal{N}_i|$ for $v \in \mathcal{N}_i$, the original signal $\mathbf{f} \in PW_\omega(\mathcal{G})$, assuming the noise associated with each vertex follows* i.i.d *Gaussian distribution $\mathcal{N}(0, \sigma^2)$, if ω is less than $1/C_{\max}^2$, the expected reconstruction error of ILMR in the kth iteration satisfies*

$$\mathrm{E}\left\{\|\tilde{\mathbf{f}}^{(k)} - \mathbf{f}\|\right\} \le \frac{|\mathcal{I}|\sigma}{1-\gamma}\sqrt{\frac{2}{\pi}} + \mathcal{O}\left(\gamma^{k+1}\right), \tag{16}$$

where γ is defined as (6).

According to (16), the error bound is affected by the number of centerless local sets $|\mathcal{I}|$. A division with fewer sets may reduce the expected reconstruction error. However, it should be noted that the number of centerless local sets cannot be too small to satisfy the condition

$$\gamma = C_{\max}\sqrt{\omega} = \max_{i\in\mathcal{I}} \sqrt{|\mathcal{N}_i|D_i\omega} < 1,$$

which is determined by the cutoff frequency of the original graph signal. Besides, the factor $1/(1-\gamma)$ in (16) implies that a smaller $C_{\max}$, which leads to a smaller γ, also reduces the error bound. A rough calculation can be given to balance the two factors. If there are not too many vertices in each $\mathcal{N}_i$, we have that $C_{\max}$ approximates to $N_{\max}$, where $N_{\max}$ is the largest cardinality of centerless local sets. Since $N_{\max}|\mathcal{I}|$ approximates to N, we have

$$\frac{1}{1-\gamma}|\mathcal{I}| \approx \frac{1}{1-\sqrt{\omega}N_{\max}} \cdot \frac{N}{N_{\max}}.$$

To minimize the above quantity, a near optimal $N_{\max}$ is

$$N_{\max} = \frac{1}{2\sqrt{\omega}}, \tag{17}$$

i.e., γ approximates to $1/2$. It provides a strategy to partition centerless local sets. For given cutoff frequency ω, an approximated $N_{\max}$ can be chosen according to (17), then the graph is divided into local sets to make sure that $|\mathcal{N}_i|$ is not more than $N_{\max}$ and the number of local sets is as small as possible.

For a given $N_{\max}$, a greedy algorithm is proposed to make the division of centerless local sets, as shown in Table 2. The greedy algorithm is to iteratively remove connected vertices with the smallest degrees from the original graph into the new set, until the cardinality of the new set reaches $N_{\max}$ or there is no connected vertex. The reason for choosing the smallest-degree vertex is that such a vertex is more likely to be on the border of a graph.

Table 2: A greedy method to partition centerless local sets with maximal cardinality.

Input: Graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$, Maximal cardinality $N_{\max}$;

Output: Centerless local sets $\{\mathcal{N}_i\}_{i\in\mathcal{I}}$;

Initialization: $i = 0$;

Loop Until: $\mathcal{V} = \emptyset$

1) Find one vertex with the smallest degree in $\mathcal{G}$,

$$u \in \arg\min_{v\in\mathcal{V}} d_{\mathcal{G}}(v);$$

2) $i = i + 1$, $\mathcal{N}_i = \{u\}$;

3) Obtain the neighbor set of $\mathcal{N}_i$,

$$\mathcal{S}_i = \{v \in \mathcal{G} | v \sim w, w \in \mathcal{N}_i, v \notin \mathcal{N}_i\};$$

Loop Until: $|\mathcal{N}_i| = N_{\max}$ or $\mathcal{S}_i = \emptyset$

4) Find one vertex with the smallest degree in $\mathcal{S}_i$,

$$u \in \arg\min_{v\in\mathcal{S}_i} d_{\mathcal{G}}(v);$$

5) $\mathcal{N}_i = \mathcal{N}_i \cup \{u\}$;

6) Update $\mathcal{S}_i = \{v \in \mathcal{G} | v \sim w, w \in \mathcal{N}_i, v \notin \mathcal{N}_i\}$;

End Loop

7) Remove the edges, $\mathcal{E} = \mathcal{E} \backslash \{(p, q) | p \in \mathcal{N}_i, q \in \mathcal{V}\}$;

8) Remove the vertices, $\mathcal{V} = \mathcal{V} \backslash \mathcal{N}_i$ and $\mathcal{G} = \mathcal{G}(\mathcal{V}, \mathcal{E})$;

End Loop

6. Experiments

We choose the Minnesota road graph [43], which has 2640 vertices and 6604 edges, to verify the proposed generalized sampling scheme and reconstruction algorithm. The bandlimited signals for reconstruction are generated by removing the high-frequency component of random signals, whose entries are drawn from *i.i.d.* Gaussian distribution. The centerless local sets are generated by the greedy method in Table 2 using given $N_{\max}$. Five kinds of local weights are tested including

1. uniform weight, where $\varphi_i(v)$ equals $1/|\mathcal{N}_i|, \forall v \in \mathcal{N}_i$;
2. random weight, where

$$\varphi_i(v) = \frac{\varphi'_i(v)}{\sum_{u\in\mathcal{N}_i} \varphi'_i(u)}, \quad \forall v \in \mathcal{N}_i, \varphi'_i(u) \sim \mathcal{U}(0,1),$$

   and $\mathcal{U}(0,1)$ denotes the uniform distribution;
3. Dirac delta weight, where $\boldsymbol{\varphi}_i$ equals $\boldsymbol{\delta}_u$ for a randomly chosen $u \in \mathcal{N}_i$;
4. the optimal weight, where

$$\varphi_i(v) = \frac{(\sigma^2(v))^{-1}}{\sum_{v\in\mathcal{N}_i} (\sigma^2(v))^{-1}}, \quad \forall v \in \mathcal{N}_i;$$

5. the optimal Dirac delta weight, where $\boldsymbol{\varphi}_i$ equals $\boldsymbol{\delta}_u$ for

$$u = \arg\min_{u\in\mathcal{N}_i} \sigma^2(u).$$

Notice that case 3 and case 5 degenerate ILMR to IPR.

6.1. Convergence of ILMR

In the first experiment, the convergence of the proposed ILMR is verified for various centerless local sets partition and local weights. The graph is divided into 709 and 358 centerless local sets for $N_{\max}$ equals 4 and 8, respectively. Three kinds of local weights are tested including case 1, 2, and 3. The averaged convergence curves are plotted in Fig. 3 for 100 randomly generated original graph signals. According to Fig. 3, the convergence is accelerated when the graph is divided into more local sets and has a smaller $N_{\max}$. It is intuitive because more local sets will bring more measurements and increase the sampling rate, which provides more information in the reconstruction. According to (6), for the same ω, a smaller $N_{\max}$ leads to a smaller γ, and guarantees a faster convergence. The experimental result also shows that in the noise-free scenario, reconstruction with uniform weight converges slightly faster than that

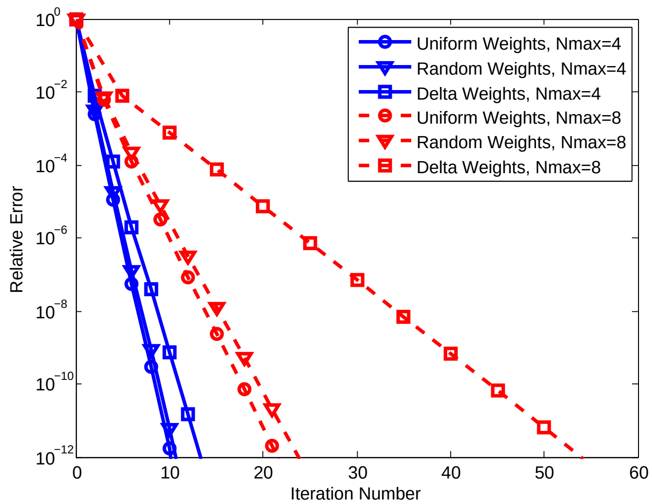

Figure 3: The convergence behavior of ILMR for various division of centerless local sets and different local weights.

with random weight. However, both above cases converge much faster than reconstruction with Dirac delta weight. This means that local-measurement-based ILMR behaves better than decimation-based IPR by combining the signals on different vertices properly.

6.2. Theoretical and Numerical Bounds for Cutoff Frequency

The sufficient condition for ILMR in Proposition 1 is not sharp enough. The numerical bounds for cutoff frequency is shown in this experiment. The graph is divided into 358 centerless local sets for $N_{\max}$ equals 8, with $C_{\max} = \sqrt{56}$. 1000 random signals are generated in the subspace with each fixed cutoff frequency. The criterion for convergence is that the error gets below the threshold 10^{-3} in 20 iterations. The sufficient condition provided in Proposition 1 is $\omega < 0.018$. As illustrated in Fig. 4, the actual cutoff frequency is larger than the theoretical one. Besides, compared with IPR, the algorithm based on local measurement reconstructs signals with a larger cutoff frequency.

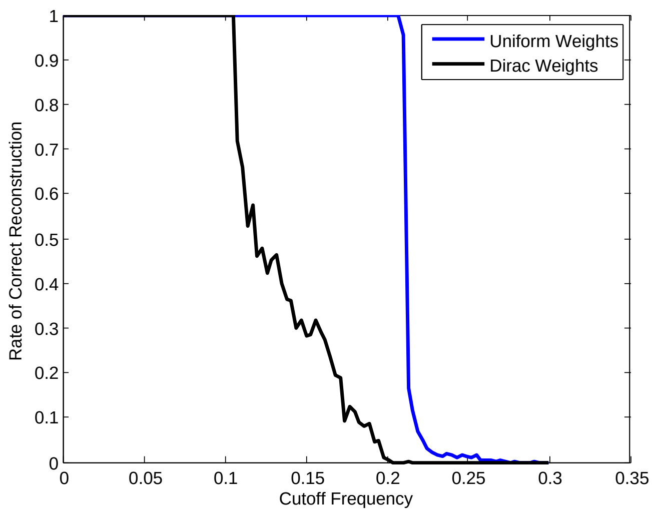

Figure 4: The reconstruction rate of ILMR for different cutoff frequencies.

6.3. Optimal Local Weights for Gaussian Noise

In this experiment, independent zero-mean Gaussian noise is added to each vertex with different variance. The original signal is normalized with unit norm. All of the vertices are randomly divided into three groups with the standard deviations of the noise chosen as σ equals 1×10^{-4}, 2×10^{-4}, and 5×10^{-4}, respectively. The graph is partitioned into 358 centerless local sets with $N_{\max}$ equals 8. Three kinds of local weights are tested including case 1, 4, and 5. The averaged convergence curves are illustrated in Fig. 5 for 100 randomly generated original graph signals. The steady-state relative error with the optimal weight is smaller than those with uniform weight and the optimal Dirac delta weight. The experimental result verifies the analysis in section 5.2. It implies that a better selection of local weights can reduce the reconstruction error if the noise variances on vertices are different.

6.4. Performance against Independent and Identical Distributed Gaussian Noise

In this experiment, the performance of the proposed algorithm against *i.i.d.* Gaussian noise are tested for three kinds of local weights including case 1, 2,

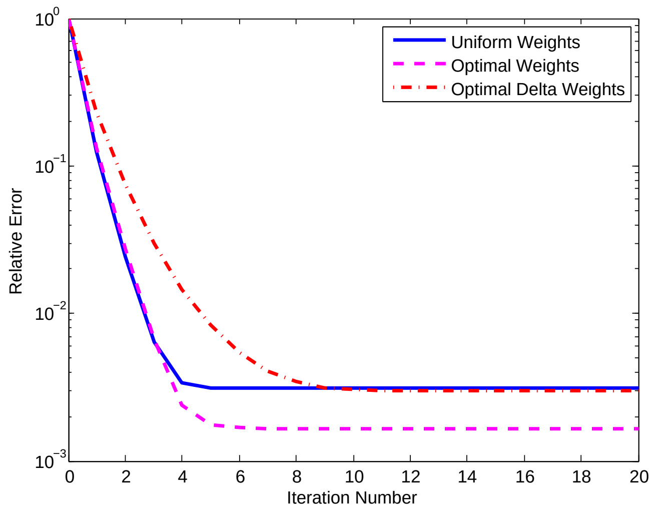

Figure 5: The convergence curves of reconstruction with uniform weights, the optimal weights, and optimal Dirac delta weights when independent zero-mean Gaussian noise is added to each vertex.

and 3. In this case the optimal local weights is equivalent to uniform weights. The graph is partitioned into 358 centerless local sets with $N_{\max}$ equals 8. The relative reconstruction errors of three tests are illustrated in Fig. 6. Each point is the average of 100 trials. The experimental result shows that for *i.i.d.* Gaussian noise, reconstruction with uniform weight or random weight performs beyond that with Dirac delta weight, which is actually the traditional sampling scheme of decimation. It shows that compared with decimation, the proposed generalized sampling scheme is more robust against noise, as analyzed in section 5.

6.5. Reconstruction of Approximated Bandlimited Signals

In this experiment, approximated bandlimited signals are tested to be reconstructed by ILMR. The original signal is normalized to have norm 1 and the out-of-band energy is 10^{-4} or 10^{-8}. The graph is partitioned into 358 centerless local sets and the maximal cardinality of local sets is 8. Three kinds of local weights are tested including case 1, 2, and 3. The convergence curves are

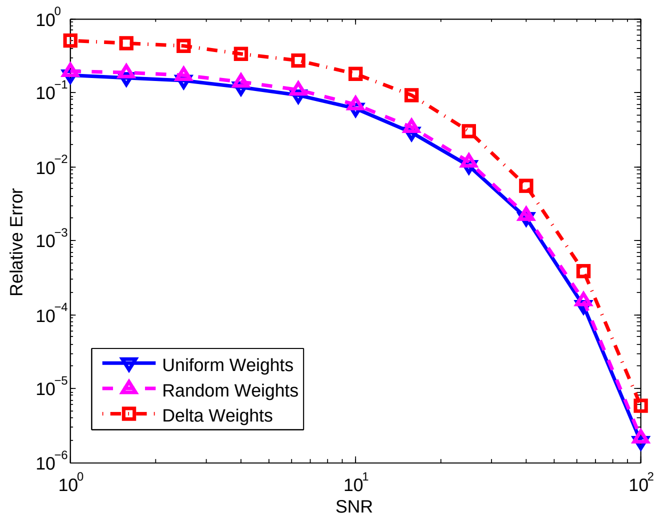

Figure 6: Relative errors of ILMR under difference SNRs with various choices of local weights. The noise associated with each vertex is *i.i.d.* Gaussian.

shown in Fig. 7, where each curve is the average of 100 trials. It is natural to see that the steady-state error is larger for a larger out-of-band energy. It is mainly because ILMR cannot recover the out-of-band part of signals. However, the out-of-band energy affects the reconstruction of the in-band part of signals, which leads to the result that the relative errors for Dirac delta weights are slightly larger than uniform weights and random weights. The in-band errors are shown in Fig. 8, which depicts up to what extent the ILMR algorithm can recover the in-band part of the signals. The case with uniform local weights has a smaller relative error, much better than that with Dirac weights. In other words, reconstruction from local measurements performs beyond reconstruction from decimation if the original signals are not strictly bandlimited.

6.6. Experiments with Real Data

A time-varying real world data is used in this experiment. The dataset is collected by Intel Berkeley Research Lab [44] including temperature, humidity, light and voltage of 54 sensors which are sampled every 30 seconds. We use the

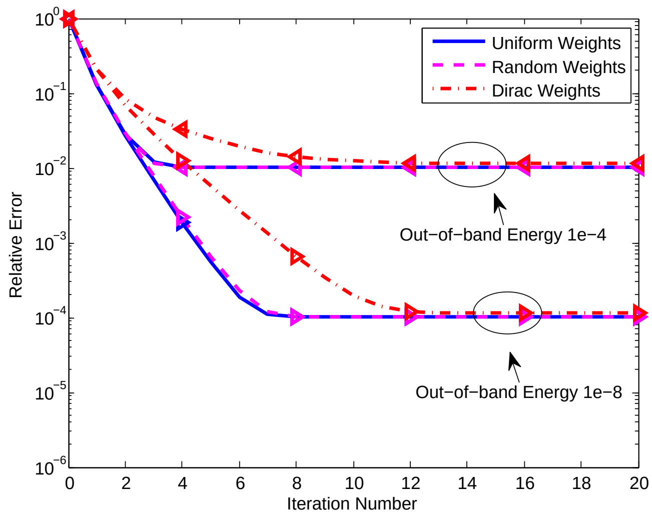

Figure 7: The convergence curves for uniform weights, random weights, and Dirac delta weights if the original graph signals are approximated bandlimited.

temperature of the sensors as the graph signal. A piece of data from 01:06:15 to 17:56:15 on February 28th, 2004 is extracted for missing data is less during the period of time. The MATLAB function *scatteredInterpolant* is used to interpolate the missing data and the result is regarded as the original time-varying graph signal. According to the position of the sensors, we build the 4-NN graph with the weights inversely proportional to the square of geometric distance. The graph is divided into 15 and 9 centerless local sets for $N_{\max}$ equals 4 and 8, respectively. Since the original graph signal is time-varying, ILMR is conducted using the newly obtained local measurements in each time step. The relative errors are illustrated in Fig. 9 for uniform and Dirac weights. Since the original graph signal is not strictly bandlimited, the steady error is around 3% and the curves reach the steady error in only several iterations. Uniform weights lead to a smaller steady error than Dirac weights. More measurements will also lead to more precise reconstruction.

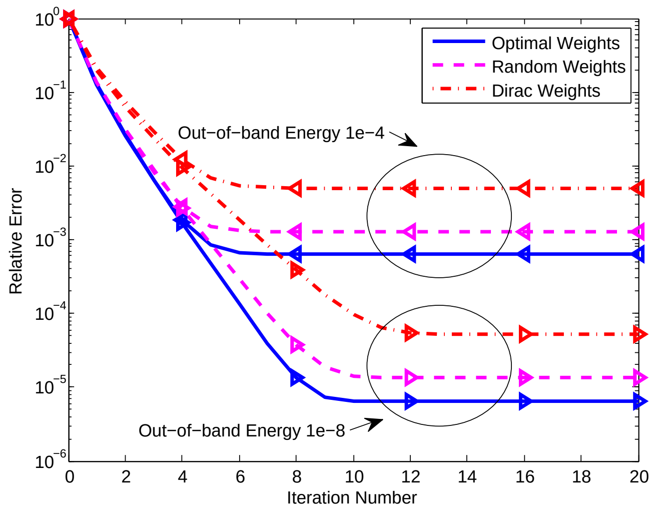

Figure 8: The in-band errors for uniform weights, random weights, and Dirac delta weights.

7. Conclusion

In this paper, a sampling scheme named local measurement is proposed to obtain sampled data from graph signals, which is a generalization of graph signal decimation. Using the local measurements, a reconstruction algorithm ILMR is proposed to perfectly reconstruct original bandlimited signals iteratively. The convergence of ILMR is proved and its performance in noisy scenarios is analyzed. The optimal local weights are given to minimize the effect of noise, and a greedy algorithm for local sets partition is proposed. Theoretical analysis and experimental results demonstrate that the local measurement sampling scheme together with reconstruction algorithm is more robust against additive noise.

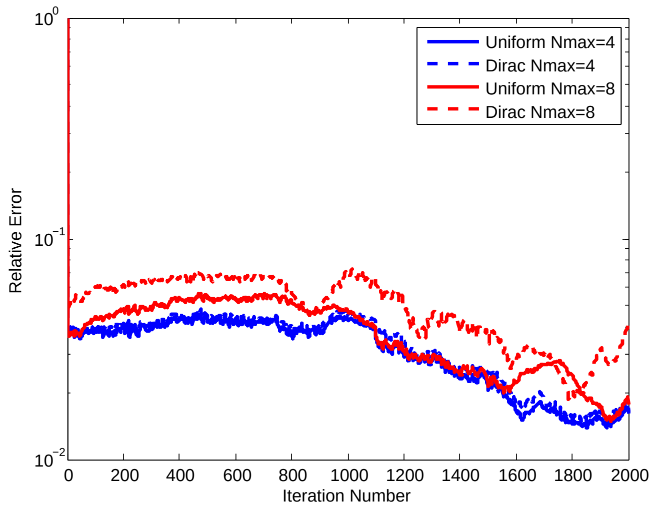

Figure 9: Relative errors for time-varying real sensor data using ILMR.

8. Appendix

8.1. Proof of Lemma 1

By the definition of $\mathbf{G}$, and considering that $\{\mathcal{N}_i\}_{i\in\mathcal{I}}$ are disjoint, one has

$$\begin{aligned}\|\mathbf{f}-\mathbf{G}\mathbf{f}\|^2 &= \left\|P_\omega\left(\sum_{i\in\mathcal{I}}(\mathbf{f}_{\mathcal{N}_i}-\langle\mathbf{f},\boldsymbol{\varphi}_i\rangle\boldsymbol{\delta}_{\mathcal{N}_i})\right)\right\|^2\\ &\le \left\|\sum_{i\in\mathcal{I}}(\mathbf{f}_{\mathcal{N}_i}-\langle\mathbf{f},\boldsymbol{\varphi}_i\rangle\boldsymbol{\delta}_{\mathcal{N}_i})\right\|^2\\ &= \sum_{i\in\mathcal{I}}\|\mathbf{f}_{\mathcal{N}_i}-\langle\mathbf{f},\boldsymbol{\varphi}_i\rangle\boldsymbol{\delta}_{\mathcal{N}_i}\|^2, \end{aligned} \tag{18}$$

where

$$f_{\mathcal{N}_i}(v)=\begin{cases} f(v), & v\in\mathcal{N}_i;\\ 0, & v\notin\mathcal{N}_i.\end{cases}$$

For $i \in \mathcal{I}$, one has

$$\begin{aligned} \|\mathbf{f}_{\mathcal{N}_i} - \langle \mathbf{f}, \boldsymbol{\varphi}_i \rangle \boldsymbol{\delta}_{\mathcal{N}_i}\|^2 &= \sum_{v \in \mathcal{N}_i} |f(v) - \langle \mathbf{f}, \boldsymbol{\varphi}_i \rangle|^2 \\ &= \sum_{v \in \mathcal{N}_i} \left| \sum_{p \in \mathcal{N}_i} \varphi_i(p) \left(f(v) - f(p) \right) \right|^2 \\ &\leq \sum_{v \in \mathcal{N}_i} \max_{p \in \mathcal{N}_i} |f(v) - f(p)|^2 \end{aligned} \tag{19}$$

Denote

$$p_i(v) = \arg\max_{p \in \mathcal{N}_i} |f(v) - f(p)|^2.$$

Since $\mathcal{N}_i$ is connected, there is a shortest path within $\mathcal{N}_i$ from v to $p_i(v)$, which is denoted as $v \sim v_1 \sim \cdots \sim v_{k_v} \sim p_i(v)$, and the length of this path is not longer than D_i. Then for $v \in \mathcal{N}_i$, one has

$$\begin{aligned} \max_{p \in \mathcal{N}_i} |f(v) - f(p)|^2 =& |f(v) - f(p_i(v))|^2 \\ \leq& \left(|f(v) - f(v_1)| + \cdots + |f(v_{k_v}) - f(p_i(v))| \right)^2 \\ \leq& D_i \left(|f(v) - f(v_1)|^2 + \cdots + |f(v_{k_v}) - f(p_i(v))|^2 \right). \end{aligned}$$

Therefore, one has

$$\sum_{v \in \mathcal{N}_i} \max_{p \in \mathcal{N}_i} |f(v) - f(p)|^2 \leq |\mathcal{N}_i| D_i \sum_{p \sim q; p,q \in \mathcal{N}_i} |f(p) - f(q)|^2, \tag{20}$$

where $p \sim q$ denotes there is an edge between p and q. Inequality (20) holds because each edge within $\mathcal{N}_i$ is reused for no more than $|\mathcal{N}_i|$ times. To study the right hand side of (20), one has

$$\begin{aligned} \sum_{p \sim q} |f(p) - f(q)|^2 =& \mathbf{f}^{\mathrm{T}} \mathbf{L} \mathbf{f} = \mathbf{f}^{\mathrm{T}} \mathbf{U} \boldsymbol{\Lambda} \mathbf{U}^{\mathrm{T}} \mathbf{f} = \hat{\mathbf{f}}^{\mathrm{T}} \boldsymbol{\Lambda} \hat{\mathbf{f}} \\ =& \sum_{\lambda_i \leq \omega} \lambda_i |\hat{f}(i)|^2 \leq \omega \hat{\mathbf{f}}^{\mathrm{T}} \hat{\mathbf{f}} = \omega \|\mathbf{f}\|^2, \end{aligned} \tag{21}$$

where $\mathbf{L}, \mathbf{U}$, and $\boldsymbol{\Lambda}$ denote the Laplacian, its eigenvectors, and its eigenvalues, respectively. The last inequality in (21) is because the entries of spectrum $\hat{\mathbf{f}} = \mathbf{U}^{\mathrm{T}} \mathbf{f}$ corresponding to the frequencies higher than ω are zero for $\mathbf{f} \in PW_\omega(\mathcal{G})$.

Consequently, utilizing (19), (20), and (21) in (18), we have

$$\begin{aligned}
\|\mathbf{f} - \mathbf{G}\mathbf{f}\|^2 \leq & \sum_{i \in \mathcal{I}} \left(|\mathcal{N}_i| D_i \sum_{p \sim q; p, q \in \mathcal{N}_i} |f(p) - f(q)|^2 \right) \\
\leq & C_{\max}^2 \sum_{p \sim q} |f(p) - f(q)|^2 \\
\leq & \omega C_{\max}^2 \|\mathbf{f}\|^2
\end{aligned}$$

and Lemma 1 is proved.

8.2. Proof of Proposition 2

According to Lemma 1, we have $\|\mathbf{I} - \mathbf{G}\| \leq \gamma < 1$ for $PW_\omega(\mathcal{G})$ when $\gamma = C_{\max}\sqrt{\omega} < 1$. Then $\mathbf{G}$ is invertible and $1 - \gamma \leq \|\mathbf{G}\| \leq 1 + \gamma$ for $PW_\omega(\mathcal{G})$. The inverse of $\mathbf{G}$ is

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

According to (2), $\mathbf{f}$ can be written as

$$\mathbf{f} = \mathbf{G}^{-1}\mathbf{G}\mathbf{f} = \sum_{j=0}^{\infty} (\mathbf{I} - \mathbf{G})^j \sum_{i \in \mathcal{I}} \langle \mathbf{f}, \boldsymbol{\varphi}_i \rangle \mathcal{P}_\omega(\boldsymbol{\delta}_{\mathcal{N}_i}) = \sum_{i \in \mathcal{I}} \langle \mathbf{f}, \boldsymbol{\varphi}_i \rangle \mathbf{e}_i, \tag{22}$$

where

$$\mathbf{e}_i = \sum_{j=0}^{\infty} (\mathbf{I} - \mathbf{G})^j \mathcal{P}_\omega(\boldsymbol{\delta}_{\mathcal{N}_i}).$$

Similarly, one has

$$\tilde{\mathbf{f}} = \sum_{i \in \mathcal{I}} \langle \tilde{\mathbf{f}}, \boldsymbol{\varphi}_i \rangle \mathbf{e}_i.$$

Using (7) and $\mathbf{f}^{(0)} = \mathbf{G}\mathbf{f}$, we have

$$\mathbf{f}^{(k)} = \mathbf{f} + (\mathbf{I} - \mathbf{G})^k (\mathbf{f}^{(0)} - \mathbf{f}) = \mathbf{f} - (\mathbf{I} - \mathbf{G})^{k+1} \mathbf{f}.$$

Therefore

$$\tilde{\mathbf{f}}^{(k)} = \tilde{\mathbf{f}} - (\mathbf{I} - \mathbf{G})^{k+1} \tilde{\mathbf{f}} = \sum_{i \in \mathcal{I}} \langle \tilde{\mathbf{f}}, \boldsymbol{\varphi}_i \rangle \mathbf{e}_i - (\mathbf{I} - \mathbf{G})^{k+1} \tilde{\mathbf{f}}. \tag{23}$$

If $\gamma = C_{\max}\sqrt{\omega} < 1$, $\mathbf{e}_i$ satisfies

$$\|\mathbf{e}_i\| \leq \sum_{j=0}^{\infty} \gamma^j \|\mathcal{P}_\omega(\boldsymbol{\delta}_{\mathcal{N}_i})\| \leq \frac{1}{1-\gamma}\|\boldsymbol{\delta}_{\mathcal{N}_i}\| = \frac{1}{1-\gamma}\sqrt{|\mathcal{N}_i|}. \tag{24}$$

According to (22), (23), and (24),

$$\begin{aligned} \|\tilde{\mathbf{f}}^{(k)} - \mathbf{f}\| &= \left\| \sum_{i\in\mathcal{I}} \langle \tilde{\mathbf{f}} - \mathbf{f}, \boldsymbol{\varphi}_i \rangle \mathbf{e}_i - (\mathbf{I} - \mathbf{G})^{k+1}\tilde{\mathbf{f}} \right\| \\ &\leq \sum_{i\in\mathcal{I}} |\langle \mathbf{n}, \boldsymbol{\varphi}_i \rangle| \, \|\mathbf{e}_i\| + \gamma^{k+1}\|\tilde{\mathbf{f}}\| \\ &\leq \frac{1}{1-\gamma} \sum_{i\in\mathcal{I}} \sqrt{|\mathcal{N}_i|} \cdot |n_i| + \gamma^{k+1}\left(\|\mathbf{f}\| + \|\mathbf{n}\|\right). \end{aligned}$$

Then Proposition 2 is proved.

References

[1] X. Wang, J. Chen, Y. Gu, Generalized graph signal sampling and reconstruction, in: Proc. 3rd IEEE Global Conf. Signal and Inform. Process. (GlobalSIP), 2015.

[2] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, 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. 30 (3) (2013) 83–98. doi:10.1109/MSP.2012.2235192.

[3] A. Sandryhaila, J. M. F. Moura, Big data analysis with signal processing on graphs: Representation and processing of massive data sets with irregular structure, IEEE Signal Process. Mag. 31 (5) (2014) 80–90. doi:10.1109/MSP.2014.2329213.

[4] X. Zhu, M. Rabbat, Graph spectral compressed sensing for sensor networks, in: Proc. 37th IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP), 2012, pp. 2865–2868. doi:10.1109/ICASSP.2012.6288515.

[5] A. Gadde, A. Anis, A. Ortega, Active semi-supervised learning using sampling theory for graph signals, in: Proc. 20th ACM Int. Conf. Knowledge Discovery and Data Mining (KDD'14), 2014, pp. 492–501. doi:10.1145/2623330.2623760.

[6] Z. Yang, A. Ortega, S. Narayanan, Gesture dynamics modeling for attitude analysis using graph based transform, in: Proc. 21st IEEE Int. Conf. Image Process. (ICIP), 2014, pp. 1515–1519. doi:10.1109/ICIP.2014.7025303.

[7] S. Chen, F. Cerda, et al., Semi-supervised multiresolution classification using adaptive graph filtering with application to indirect bridge structural health monitoring, IEEE Trans. Signal Process. 62 (11) (2013) 2879–2893. doi:10.1109/TSP.2014.2313528.

[8] A. Sandryhaila, J. M. F. Moura, Discrete signal processing on graphs, IEEE Trans. Signal Process. 61 (7) (2013) 1644–1656. doi:10.1109/TSP.2013.2238935.

[9] R. R. Coifman, M. Maggioni, Diffusion wavelets, Appl. Comput. Harmonic Anal. 21 (1) (2006) 53–94. doi:10.1016/j.acha.2006.04.004.

[10] D. K. Hammond, P. Vandergheynst, R. Gribonval, Wavelets on graphs via spectral graph theory, Appl. Comput. Harmonic Anal. 30 (2) (2011) 129–150. doi:10.1016/j.acha.2010.04.005.

[11] X. Zhu, M. Rabbat, Approximating signals supported on graphs, in: Proc. 37th IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP), 2012, pp. 3921–3924. doi:10.1109/ICASSP.2012.6288775.

[12] H. Q. Nguyen, M. N. Do, Downsampling of signals on graphs via maximum spanning trees, IEEE Trans. Signal Process. 63 (1) (2015) 182–191. doi:10.1109/TSP.2014.2369013.

[13] A. Agaskar, Y. M. Lu, A spectral graph uncertainty principle, IEEE Trans. Inform. Theory 59 (7) (2013) 4338–4356. doi:10.1109/TIT.2013.2252233.

[14] P. Liu, X. Wang, Y. Gu, Coarsening graph signal with spectral invariance, in: Proc. 39th IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP), 2014, pp. 1075–1079. doi:10.1109/ICASSP.2014.6853761.

[15] P. Liu, X. Wang, Y. Gu, Graph signal coarsening: Dimensionality reduction in irregular domain, in: Proc. 2nd IEEE Global Conf. Signal and Inform. Process. (GlobalSIP), 2014, pp. 966–970. doi:10.1109/GlobalSIP.2014.7032229.

[16] P. Chen, A. O. Hero, Phase transitions in spectral community detection, IEEE Trans. Signal Process. 63 (16) (2015) 4339–4347. doi:10.1109/TSP.2015.2442958.

[17] D. Thanou, D. I. Shuman, P. Frossard, Parametric dictionary learning for graph signals, in: Proc. 1st IEEE Global Conf. Signal and Inform. Process. (GlobalSIP), 2013, pp. 487–490. doi:10.1109/GlobalSIP.2013.6736921.

[18] S. Chen, R. Varma, A. Singh, J. Kovačević, Signal representations on graphs: Tools and applications, arXiv preprint (2015) arXiv:1512.05406.

[19] X. Dong, D. Thanou, D. I. Shuman, P. Vandergheynst, Laplacian matrix learning for smooth graph signal representation, in: Proc. 40th IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP), 2015, pp. 3736–3740. doi:10.1109/ICASSP.2015.7178669.

[20] A. Anis, A. Gadde, 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. doi:10.1109/ICASSP.2014.6854325.

[21] S. K. Narang, A. Gadde, E. Sanou, A. Ortega, Localized iterative methods for interpolation in graph structured data, in: Proc. 1st IEEE Global Conf. Signal and Inform. Process. (GlobalSIP), 2013, pp. 491–494. doi:10.1109/GlobalSIP.2013.6736922.

[22] X. Wang, P. Liu, Y. Gu, Iterative reconstruction of graph signal in low-frequency subspace, in: Proc. 2nd IEEE Global Conf. Signal and Inform. Process. (GlobalSIP), 2014, pp. 611–615. doi:10.1109/GlobalSIP.2014.7032157.

[23] X. Wang, P. Liu, Y. Gu, Local-set-based graph signal reconstruction, IEEE Trans. Signal Process. 63 (9) (2015) 2432–2444. doi:10.1109/TSP.2015.2411217.

[24] S. Chen, A. Sandryhaila, J. M. F. Moura, J. Kovačević, Signal recovery on graphs: variation minimization, IEEE Trans. Signal Process. 63 (17) (2015) 4609–4624. doi:10.1109/TSP.2015.2441042.

[25] S. Chen, A. Sandryhaila, J. M. F. Moura, J. Kovačević, Discrete signal processing on graphs: Sampling theory, IEEE Trans. Signal Process. 63 (24) (2015) 6510–6523. doi:10.1109/TSP.2015.2469645.

[26] S. Segarra, A. G. Marques, G. Leus, A. Ribeiro, Interpolation of graph signals using shift-invariant graph filters, in: Proc. 23rd European Signal Processing Conference (EUSIPCO), 2015, pp. 210–214. doi:10.1109/EUSIPCO.2015.7362375.

[27] S. Segarra, A. G. Marques, G. Leus, A. Ribeiro, Reconstruction of graph signals through percolation from seeding nodes, arXiv preprint (2015) arXiv:1507.08364.

[28] X. Wang, M. Wang, Y. Gu, A distributed tracking algorithm for reconstruction of graph signals, IEEE J. Selected Topics Signal Process. 9 (4) (2015) 728–740. doi:10.1109/JSTSP.2015.2403799.

[29] S. Chen, A. Sandryhaila, J. M. F. Moura, J. Kovačević, Distributed algorithm for graph signals, in: Proc. 40th IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP), 2015, pp. 3731–3735. doi:10.1109/ICASSP.2015.7178668.

[30] S. K. Narang, A. Gadde, 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. `doi:10.1109/ICASSP.2013.6638704`.

[31] I. Pesenson, Sampling in paley-wiener spaces on combinatorial graphs, Trans. Amer. Math. Soc. 360 (10) (2008) 5603–5627. `doi:10.1090/S0002-9947-08-04511-X`.

[32] H. Führ, I. Z. Pesenson, Poincaré and plancherel-polya inequalities in harmonic analysis on weighted combinatorial graphs, SIAM J. Discrete Math. 27 (4) (2013) 2007–2028. `doi:10.1137/120873674`.

[33] F. Marvasti, Nonuniform Sampling: Theory and Practice, Springer, 2001.

[34] H. G. Feichtinger, K. Gröchenig, Theory and practice of irregular sampling, Wavelets: Math. and Applicat. (1994) 305–363.

[35] K. Gröchenig, A discrete theory of irregular sampling, Linear Algebra and Its Applicat. 193 (1993) 129–150. `doi:10.1016/0024-3795(93)90275-S`.

[36] K. Gröchenig, Reconstruction algorithms in irregular sampling, Math. Comput. 59 (199) (1992) 181–194. `doi:10.1090/S0025-5718-1992-1134729-0`.

[37] A. Aldroubi, Non-uniform weighted average sampling and reconstruction in shift-invariant and wavelet spaces, Appl. Comput. Harmonic Anal. 13 (2) (2002) 151–161. `doi:10.1016/S1063-5203(02)00503-1`.

[38] I. Pesenson, Poincaré-type inequalities and reconstruction of paley-wiener functions on manifolds, J. Geometric Anal. 14 (1) (2004) 101–121. `doi:10.1007/BF02921868`.

[39] H. Feichtinger, I. Pesenson, Recovery of band-limited functions on manifolds by an iterative algorithm, Contemporary Math. 345 (2004) 137–152.

[40] A. G. Marques, S. Segarra, G. Leus, A. Ribeiro, Sampling of graph signals with successive local aggregations, IEEE Trans. Signal Process.doi:10.1109/TSP.2015.2507546.

[41] F. R. K. Chung, Spectral Graph Theory, Amer. Math. Soc., 1997.

[42] M. W. Lipsey, D. B. Wilson, Practical Meta-analysis, Thousand Oaks, CA: Sage publications, 2001.

[43] D. Gleich, The matlabbgl matlab library.
URL http://www.cs.purdue.edu/homes/dgleich/packages/matlab_bgl/index.html

[44] Intel lab data.
URL http://www.cs.cmu.edu/~guestrin/Research/Data/