5-2015

Signal Recovery on Graphs: Random versus Experimentally Designed Sampling

Siheng Chen
Carnegie Mellon University

Rohan Varma
Carnegie Mellon University

Aarti Singh
Carnegie Mellon University, aarti@cs.cmu.edu

Jelena Kovacevic
Carnegie Mellon University, jelenak@cmu.edu

Signal Recovery on Graphs:
Random versus Experimentally Designed Sampling

Siheng Chen, Rohan Varma, Aarti Singh, Jelena Kovačević

Abstract—We study signal recovery on graphs based on two sampling strategies: random sampling and experimentally designed sampling. We propose a new class of smooth graph signals, called approximately bandlimited, which generalizes the bandlimited class and is similar to the globally smooth class. We then propose two recovery strategies based on random sampling and experimentally designed sampling. The proposed recovery strategy based on experimentally designed sampling is similar to the leverage scores used in the matrix approximation. We show that while both strategies are unbiased estimators for the low-frequency components, the convergence rate of experimentally designed sampling is much faster than that of random sampling when a graph is irregular. We validate the proposed recovery strategies on three specific graphs: a ring graph, an Erdős-Rényi graph, and a star graph. The simulation results support the theoretical analysis.

I. Introduction

We consider sampling and recovery within the framework of signal processing on graphs, which studies signals with an underlying complex structure [1], [2]. The framework models that underlying structure by a graph and signals by graph signals, generalizing concepts and tools from classical discrete signal processing.

The task of sampling and recovery is one of the most critical topics in the signal processing community. As the bridge connecting sequences and functions, classical sampling theory shows that a bandlimited function can be perfectly recovered from its sampled sequence if the sampling rate is high enough. The interest in sampling and recovery of graph signals has increased in the last few years [3], [4], [5]. In [6], authors proposed an algorithm to recover graph signals that have small variation based on random sampling. In [4], [7], authors proposed a sampling theory for graph signals and show perfect recovery for bandlimited graph signals based on experimentally designed sampling.

In this paper, we propose a new class of graph signals, called *approximately bandlimited*, which generalizes the bandlimited class and is similar to the globally smooth class. We then propose two recovery strategies based on random sampling and experimentally designed sampling, and bound the recovery error for the class of approximately bandlimited graph signals. We show that the proposed recovery strategies are unbiased estimators for low-frequency components and that experimentally designed sampling outperforms random sampling in terms of the convergence rate when a graph is irregular. We validate both recovery strategies on three specific

graphs: a ring graph, an Erdős-Rényi graph, and a star graph. The simulation results support the theoretical analysis.

II. Discrete Signal Processing on Graphs

We now briefly review discrete signal processing on graphs [2], which lays a foundation for the proposed work. We consider a graph $G = (\mathcal{V}, \mathrm{A})$, where $\mathcal{V} = \{v_0, \ldots, v_{N-1}\}$ is the set of nodes and $\mathrm{A} \in \mathbb{R}^{N \times N}$ is the graph shift, or a weighted adjacency matrix. As the most basic filter defined on this graph, the graph shift represents the connections of the graph G, which can be either directed or undirected. The edge weight $\mathrm{A}_{n,m}$ between nodes v_n and v_m is a quantitative expression of the underlying relation between the nth and the mth node, such as a similarity, a dependency, or a communication pattern. To guarantee that the filtered signal is properly scaled for comparison with the original one [2], we normalize the graph shift, such that $|\lambda_{\max}(\mathrm{A})| = 1$. Once the node order is fixed, the graph signal can be written as a vector, $\mathbf{x} = \begin{bmatrix} x_0, x_1, \ldots, x_{N-1} \end{bmatrix}^T \in \mathbb{R}^N$. The Jordan decomposition of A is [2]

$$\mathrm{A} = \mathrm{V}\,\Lambda\,\mathrm{U}, \qquad (1)$$

where the generalized eigenvectors of A form the columns of matrix V, $\mathrm{U} = \mathrm{V}^{-1}$ (the norm of each row of U is normalized to one), and $\Lambda \in \mathbb{R}^{N \times N}$ is the block diagonal matrix of corresponding eigenvalues $\lambda_0, \ldots \lambda_{N-1}$ of A ($1 = \lambda_0 \geq \lambda_1 \geq \ldots, \geq \lambda_{N-1} \geq -1$). The *graph Fourier transform* of $\mathbf{x} \in \mathbb{R}^N$ is

$$\widehat{\mathbf{x}} = \mathrm{U}\,\mathbf{x}. \qquad (2)$$

The *inverse graph Fourier transform* is $\mathbf{x} = \mathrm{V}\,\widehat{\mathbf{x}} = \sum_{k=0}^{N-1} \widehat{x}_k \mathbf{v}_k$, where $\mathbf{v}_k$ is the kth column of V and $\widehat{x}_k$ is the kth component in $\widehat{\mathbf{x}}$. The vector $\widehat{\mathbf{x}}$ in (2) represents the signal's expansion in the eigenvector basis and describes the frequency components of the graph signal $\mathbf{x}$. The inverse graph Fourier transform reconstructs the graph signal by combining graph frequency components. When A represents an undirected graph, we have $\mathrm{U} = \mathrm{V}^T$, and both U and V are orthonormal. In general, V may not be orthonormal; to restrict its behavior, we assume that

$$\alpha_1 \|\mathbf{x}\|_2^2 \leq \|\mathrm{V}\,\mathbf{x}\|^2 \leq \alpha_2 \|\mathbf{x}\|_2^2, \quad \text{for all } \mathbf{x} \in \mathbb{R}^N, \qquad (3)$$

where $\alpha_1, \alpha_2 > 0$, that is, V is a Riesz basis with stability constants α_1, α_2 [8]. The eigenvalues $\lambda_0, \ldots \lambda_{N-1}$ of A, represent frequencies on the graph [2].

III. PROBLEM FORMULATION

We now review two standard classes of graph signals, and propose a new one, which connects the first two. We next describe the sampling and recovery strategies of interest. In this way, we show the connection between our work and the previous work: graph signal inpainting and sampling theory on graphs.

A. Graph Signal Model

We focus on smooth graph signals, that is, the signal coefficient at each node is close to the signal coefficients of its neighbors. In literature [6], [7], two classes of graph signals have been introduced to measure the smoothness on graphs.

Definition 1. A graph signal $\mathbf{x} \in \mathbb{R}^N$ is *globally smooth* on a graph $A \in \mathbb{R}^{N \times N}$ with parameter $\eta \geq 0$, when

$$\|\mathbf{x} - A\,\mathbf{x}\|_2^2 \ \leq \ \eta \,\|\mathbf{x}\|_2^2. \tag{4}$$

Denote this class of graph signals by $\mathrm{GS}_A(\eta)$.

Since we normalized the graph shift such that $|\lambda_{\max}(A)| = 1$; when $\eta \geq 4$, all graph signals satisfy (4). While the recovery of globally smooth graph signals has been studied in [6] (leading to graph signal inpainting), global smoothness is a general requirement, making it hard to provide further theoretical insight [9].

Definition 2. A graph signal $\mathbf{x} \in \mathbb{R}^N$ is *bandlimited* on a graph A with parameter $K \in \{0, 1, \cdots, N-1\}$, when the graph frequency components $\widehat{\mathbf{x}}$ satisfies

$$\widehat{x}_k \ = \ 0 \quad \text{for all} \quad k \geq K.$$

Denote this class of graph signals by $\mathrm{BL}_A(K)$.

While the recovery of bandlimited graph signals has been studied in [7] (leading to sampling theory on graphs), the bandlimited requirement is a restricted requirement, making it hard to use in the real world applications. We thus propose a third class that relaxes it, but still promotes smoothness.

Definition 3. A graph signal $\mathbf{x} \in \mathbb{R}^N$ is *approximately bandlimited* on a graph A with parameters $\beta \geq 1$ and $\mu \geq 0$, when there exists a $K \in \{0, 1, \cdots, N-1\}$ such that its graph Fourier transform $\widehat{\mathbf{x}}$ satisfies

$$\sum_{k=K}^{N-1} (1 + k^{2\beta})\widehat{x}_k^2 \ \leq \ \mu \,\|\mathbf{x}\|_2^2. \tag{5}$$

Denote this class of graph signals by $\mathrm{ABL}_A(K, \beta, \mu)$.

We see that $\mathrm{BL}_A(K)$ is a subset of $\mathrm{ABL}_A(K, \beta, \mu)$ with $\mu = 0, \beta = 0$. The approximately bandlimited class allows for a tail after the first K frequency components. The parameter μ controls the shape of the tail; the smaller the μ, the smaller the energy contribution from the high-frequency components. The parameter β controls the speed of energy decaying; the larger the β, the larger the penalty on the high-frequency components. The class of $\mathrm{BL}_A(K)$ is similar to the ellipsoid constraints in [10], where all the graph frequency components are considered in the constraints; thus, $\mathrm{ABL}_A(K)$ provides more flexibility for the low-frequency components.

The following theorem shows the relationship between $\mathrm{ABL}_A(K, \beta, \mu)$ and $\mathrm{GS}_A(\eta)$.

Theorem 1. $\mathrm{ABL}_A(K, \beta, \mu)$ is a subset of $\mathrm{GS}_A(\eta)$, when

$$\eta \geq \left(1 - \lambda_{K-1} + \sqrt{\frac{4\alpha_2\mu}{(1 + K^{2\beta})}} \right)^2;$$

$\mathrm{GS}_A(\eta)$ is a subset of $\mathrm{ABL}_A(K, \beta, \mu)$, when

$$\mu \geq \frac{1 + (N-1)^{2\beta}}{(1 - \lambda_K)\alpha_1}\eta.$$

From Theorem 1, we see that when choosing proper parameters, $\mathrm{GS}_A(\eta)$ is a subset of $\mathrm{ABL}_A(K, \beta, \mu)$.

B. Sampling & Recovery

We consider the procedure of sampling and recovery as follows: we sample M coefficients in a graph signal $\mathbf{x} \in \mathbb{R}^N$ with noise to produce a noisy sampled signal $\mathbf{y} \in \mathbb{R}^M (M < N)$, that is,

$$\mathbf{y} \ = \ \Psi\mathbf{x} + \epsilon \ \equiv \ \mathbf{x}_{\mathcal{M}} + \epsilon, \tag{6}$$

where $\epsilon \sim \mathcal{N}(0, \sigma^2\,\mathrm{I}_{M \times M})$, $\mathcal{M} = (\mathcal{M}_0, \cdots, \mathcal{M}_{M-1})$ denotes the sequence of sampled indices, called *sampling set*, with $\mathcal{M}_i \in \{0, 1, \cdots, N-1\}$, $\mathbf{x}_{\mathcal{M}}$ is the noiseless sampled coefficients, and the sampling operator Ψ is a linear mapping from $\mathbb{R}^N$ to $\mathbb{R}^M$,

$$\Psi_{i,j} = \begin{cases} 1, & j = \mathcal{M}_i; \\ 0, & \text{otherwise.} \end{cases} \tag{7}$$

We then interpolate $\mathbf{y}$ to get $\mathbf{x}' \in \mathbb{R}^N$, which recovers $\mathbf{x}$ either exactly or approximately.

We consider two sampling strategies: *random sampling* means that sample indices is chosen from from $\{0, 1, \cdots, N-1\}$ independently and randomly; and *experimentally design sampling* means that sample indices can be chosen beforehand. It is clear that random sampling is a subset of experimentally design sampling.

IV. RECOVERY STRATEGY

We now propose two recovery strategies based on random sampling and experimentally designed sampling. In Section III-A, we showed that a graph signal is smooth when its energy is mainly concentrated in the low-frequency components. For example, for the class $\mathrm{BL}_A(K)$, all the energy is concentrated in the first K frequency components and the graph signal can be perfectly recovered by using those first K frequency components. The recovery strategies we propose here follow this intuition, by providing unbiased estimators for the low-frequency components.

A. Recovery Strategy based on Random Sampling

We consider the following recovery strategy.

Algorithm 1. We sample a graph signal $|\mathcal{M}|$ times. Each time, we choose a node i independently and randomly, and take a

measurement y_i. We then recover the original graph signal by using the following two steps:

$$\widehat{x}_k^* = \frac{N}{|\mathcal{M}|} \sum_{i \in \mathcal{M}} U_{ki} \, y_i,$$

$$x_i^* = \sum_{k < \kappa} V_{ik} \, \widehat{x}_k^*,$$

where x_i^* is the ith component of the recovered graph signal $\mathbf{x}^*$.

Algorithm 1 aims to estimate the first κ frequency components, and reconstruct the original graph signal based on these graph frequency components. The only tuning parameter in Algorithm 1 is the bandwidth κ. To show the performance of Algorithm 1 for recovering the low-frequency components, we have the following results.

Denote $V_{(\kappa)}$ be the first κ columns of the inverse graph Fourier transform matrix V, and $U_{(\kappa)}$ be the first κ rows of the graph Fourier transform matrix U.

Lemma 1. Algorithm 1 with bandwidth κ provides an unbiased estimator of the first κ frequency components, that is,

$$\mathbb{E}\mathbf{x}^* = V_{(\kappa)} U_{(\kappa)} \mathbf{x}, \quad \text{for all } \mathbf{x},$$

where $\mathbf{x}^*$ is the result of Algorithm 1.

The advantage of Algorithm 1 is its efficiency, that is, we only need the first κ eigenvectors involved with the computation, which is appealing for large-scale graphs. The disadvantage is that when the main energy of an original graph signal is not concentrated in the first κ frequency components, the recovered graph signal has a large bias.

Theorem 2. For $\mathbf{x} \in \mathrm{ABL}(K, \beta, \mu)$, let $\mathbf{x}^*$ be the result of Algorithm 1 with bandwidth $\kappa \geq K$, we have,

$$\mathbb{E}\left\|\mathbf{x}^* - \mathbf{x}\right\|^2 \leq \frac{\alpha_2 \mu \left\|\mathbf{x}\right\|_2^2}{\kappa^{2\beta}} + \frac{\alpha_2 (\max_j x_j^2 + \sigma^2)}{|\mathcal{M}|} N \left\|U_{(\kappa)}\right\|_F^2,$$

where α_2 is the stability constant of V in (3), σ^2 is the noise level in (6), and $\left\|\cdot\right\|_F$ is the Frobenius norm.

Due to the limited space, we do not show the proof here. The main idea follows from the bias-variance tradeoff. The first term is the bias term, and the second terms is the variance term. Since Algorithm 1 can recover the first κ frequency components on expectation, the bias comes from the other $(N-\kappa)$ frequency components, which can be bounded from the definition of $\mathrm{ABL}(K, \mu, \beta)$ when $\kappa \geq K$. The variance term depends on $\left\|U_{(\kappa)}\right\|_F^2$, which represents the graph structure.

B. Recovery Strategy based on Experimentally Designed Sampling

We consider the following recovery strategy.

Algorithm 2. We sample a graph signal $|\mathcal{M}|$ times. Each time, we choose a node with probability $w_i = \left\|\mathbf{u}_i\right\|_2 / \sum_{j=0}^{N-1} \left\|\mathbf{u}_j\right\|_2$, where $\mathbf{u}_i$ is the ith column of $U_{(\kappa)}$, and take a measurement y_i. We then recover the original graph signal by using the

following two steps:

$$\widehat{x}_k^* = \frac{1}{|\mathcal{M}|} \sum_{i \in \mathcal{M}} \frac{1}{w_i} U_{ki} \, y_i,$$

$$x_i^* = \sum_{k < \kappa} V_{ik} \, \widehat{x}_k^*.$$

where x_i^* is the ith component of the recovered graph signal $\mathbf{x}^*$.

Similarly to Algorithm 1, Algorithm 2 aims to estimate the first κ frequency components, and reconstructs the original graph signal based on these graph frequency components. The difference comes from the normalization factor. In Algorithm 1, the contribution from each measurement is normalized by a constant, the size of the graph; and in Algorithm 2, the contribution from each measurement is normalized based on the norm of the corresponding column in $U_{(\kappa)}$. It is similar to leverage scores used in the matrix approximation [11], where the goal is to evaluate the contribution from each column to approximating matrix. Note that leverage scores use the norm square, $\left\|\mathbf{u}_i\right\|_2^2$, and we use the norm, $\left\|\mathbf{u}_i\right\|_2$. When we use the norm square as probability, the performance is the same with the random sampling.

We can show that Algorithm 2 is also an unbiased estimator for recovering the low-frequency components, and potentially has a tighter upper bound.

Lemma 2. Algorithm 2 with bandwidth κ provides an unbiased estimator of the first κ frequency components, that is,

$$\mathbb{E}\mathbf{x}^* = V_{(\kappa)} U_{(\kappa)} \mathbf{x}, \quad \text{for all } \mathbf{x},$$

where $\mathbf{x}^*$ is the result of Algorithm 2.

Theorem 3. For $\mathbf{x} \in \mathrm{ABL}_A(K, \beta, \mu)$, let $\mathbf{x}^*$ be the result of Algorithm 2 with bandwidth $\kappa \geq K$, we have,

$$\mathbb{E}\left\|\mathbf{x}^* - \mathbf{x}\right\|^2 \leq \frac{\alpha_2 \mu \left\|\mathbf{x}\right\|_2^2}{\kappa^{2\beta}} + \frac{\alpha_2 (\max_j x_j^2 + \sigma^2)}{|\mathcal{M}|} \left\|U_{(\kappa)}\right\|_{2,1}^2.$$

The main idea also follows from the bias-variance trade-off. We see that Algorithms 1 and 2 have the same bias by recovering the first κ frequency components on expectation. When each column of $U_{(\kappa)}$ has roughly similar energy, $N \left\|U_{(\kappa)}\right\|_F^2$ and $\left\|U_{(\kappa)}\right\|_{2,1}^2$ are similar. However, when the energy is concentrated on a few columns, $N \left\|U_{(\kappa)}\right\|_F^2$ is much larger than $\left\|U_{(\kappa)}\right\|_{2,1}^2$, in other words, Algorithm 2 has a significant advantage over Algorithm 1 when the associated graph structure is irregular.

C. Convergence Rates

To discriminate the proposed recovery strategies, we propose two types of graphs, and compare the convergence rates of Algorithms 1 and 2 for each type of these two.

Definition 4. A graph $A \in \mathbb{R}^{N \times N}$ is *type-1*, when

$$|U_{i,j}| = O(N^{-1/2}), \quad \text{for all } i, j = 0, 1, \cdots, N-1,$$

where U are the graph Fourier transform matrix of A.

For a type-1 graph, each element in U has roughly similar magnitudes, that is, the energy evenly spreads to each element

in U. Some examples are discrete-time graphs, discrete-space graphs, and unweighted circulant graphs. Based on Theorems 2, and 3, we conclude as follows.

Corollary 1. Let $A \in \mathbb{R}^{N \times N}$ be a type-1 graph, for the class $\text{ABL}_A(K, \beta, \mu)$.

- Let $\mathbf{x}^*$ be the results given by Algorithm 1 with the bandwidth $\kappa \geq K$, we have

$$\frac{1}{N}\mathbb{E}\left(\|\mathbf{x}^* - \mathbf{x}\|_2^2\right) \leq C \min\{|\mathcal{M}|, N\}^{-\frac{2\beta}{2\beta+1}},$$

where $C > 0$, and the rate is achieved when κ is in the order of $|\mathcal{M}|^{1/(2\beta+1)}$ and upper bounded by N;

- Let $\mathbf{x}^*$ be the results given by Algorithm 2 with the bandwidth $\kappa \geq K$, we have

$$\frac{1}{N}\mathbb{E}\left(\|\mathbf{x}^* - \mathbf{x}\|_2^2\right) \leq C|\mathcal{M}|^{-\frac{2\beta}{2\beta+1}},$$

where $C > 0$, and the rate is achieved when κ is in the order of $|\mathcal{M}|^{1/(2\beta+1)}$ and upper bounded by N.

When $|\mathcal{M}| \gg N$, we set $\kappa = N$, and then the bias term is zero, and both upper bounds are actually $CN|\mathcal{M}|^{-1}$. We see that Algorithms 1 and 2 have the same convergence rate, that is, experimentally designed sampling does not perform better than random sampling for the type-1 graphs.

Definition 5. A graph $A \in \mathbb{R}^{N \times N}$ is *type-2* with parameter $K_0 > 0$, when

$$\left\|\mathbf{h}_{T^c}^{(K)}\right\|_1 \leq \alpha \left\|\mathbf{h}_T^{(K)}\right\|_1, \quad \text{for all } K \geq K_0,$$

where $h_i^{(K)} = \sqrt{\sum_{k=0}^{K-1} \mathrm{U}_{k,i}^2}$, T indexes the largest K elements in $\mathbf{h}$, T^c indexes the other $(N - K)$ elements, and $\alpha > 0$ is a constant.

A type-2 graph requires U to be approximately sparse. When we take the first K rows to form a submatrix, the energy in the submatrix concentrates in a few columns. The simulations show that scale-free graphs fall into this type approximately.

Based on Theorems 2, and 3, we conclude the following.

Corollary 2. Let $A \in \mathbb{R}^{N \times N}$ be a type-2 graph with parameter K_0, for the class $\text{ABL}_A(K, \beta, \mu)$.

- Let $\mathbf{x}^*$ be the results given by Algorithm 1 with the bandwidth $\kappa \geq K$, we have

$$\frac{1}{N}\mathbb{E}\left(\|\mathbf{x}^* - \mathbf{x}\|_2^2\right) \leq C|\mathcal{M}|^{-\frac{2\beta}{2\beta+1}},$$

where $C > 0$, and the rate is achieved when κ is in the order of $|\mathcal{M}|^{1/(2\beta+1)}$ and upper bounded by N;

- Let $\mathbf{x}^*$ be the results given by Algorithm 2 with the bandwidth $\kappa \geq \max\{K, K_0\}$, we have

$$\frac{1}{N}\mathbb{E}\left(\|\mathbf{x}^* - \mathbf{x}\|_2^2\right) \leq C|\mathcal{M}|^{-\frac{2\beta}{2\beta+2-\gamma}} \leq C'|\mathcal{M}|^{-\frac{2\beta}{2\beta+1}},$$

where $C > 0$, the rate is achieved when κ is in the order of $|\mathcal{M}|^{1/(2\beta+2-\gamma)}$ and upper bounded by N, and

$$\gamma \in [\max\{1, 2\beta+2-\frac{\log|\mathcal{M}|}{\log\max\{K,K_0\}}\}, \max\{1, \frac{(2\beta+2)\log N}{(\log N + \log|\mathcal{M}|)}\}].$$

Similarly to the type 1 graphs, when $|\mathcal{M}| \gg N$, we set $\kappa = N$, and then the bias term is zero, and both upper

bounds are $CN|\mathcal{M}|^{-1}$. We see that Algorithm 2 has a larger convergence rate than Algorithm 1, that is, experimentally designed sampling exhibits much better performance than random sampling for the type-2 graph. The advantages follow from that, for type-2 graphs, $\left\|\mathrm{U}_{(\kappa)}\right\|_{2,1}^2$ is in the order of κ^2, and $N\left\|\mathrm{U}_{(\kappa)}\right\|_F^2$ is in the order of $N\kappa$.

D. Relation to Graph Signal Inpainting

Graph signal inpainting aims at recovering globally smooth graph signals based on random sampling. It solves the following optimization problem,

$$\mathbf{x}^* = \arg\min_{\mathbf{x}} \|\mathbf{x} - A\mathbf{x}\|_2^2, \tag{8a}$$

$$\text{subject to} \quad \|\Psi\mathbf{x} - \mathbf{y}\|_2^2 \leq \sigma^2, \tag{8b}$$

where σ^2 is noise level, $\mathbf{y}$ is a vector representation of the noisy measurements (6), and Ψ is the sampling operator (7). Graph signal inpainting focuses on recovery in the vertex domain, and the proposed recovery strategies focus on recovery in graph spectral domain. The optimum of (8) guarantees that the recovered graph signal is close to the measurements at given nodes, but Algorithms 1 and 2 guarantee the recovery of the low-frequency components.

E. Relation to Sampling Theory on Graphs

Sampling theory on graphs aims at recovering bandlimited graph signals based on both the random sampling and the experimentally designed sampling [7]. It solves the following optimization problem,

$$\mathbf{x}^* = \mathbf{argmin}_{\mathbf{x} \in \text{BL}_A(K)} \|\Psi\mathbf{x} - \mathbf{y}\|_2^2 = V_{(K)}(\Psi V_{(K)})^+\mathbf{y}, \tag{9}$$

where Ψ is the sampling operator (7), $\mathbf{y}$ is a vector representation of the noisy measurements (6), and $(\cdot)^+$ is the pseudo-inverse. When the original graph signal is bandlimited, $\mathbf{x} \in \text{BL}_A(K)$, it is clear that the result of (9) is an unbiased estimator of $\mathbf{x}$. When the original graph signal is not bandlimited, the result of (9) is a biased estimator of the first K frequency components, because the signal belonging to the other frequency band also projects onto the first K components. In a sense of recovering the low-frequency components, (9) needs fewer samples, but Algorithms 1 and 2 are more reliable.

V. EXPERIMENTAL RESULTS

In this section, we compare the empirical performance of Algorithms 1 and 2 on three specific graphs: a ring graph, an Erdős-Rényi graph, and a star graph.

For each graph A, we generate 50 graph signals by the following two steps. We first generate the graph frequency components as

$$\widehat{x}_k \begin{cases} \sim \mathcal{N}(1, 0.5^2) & \text{if } k < K, \\ = K^{2\beta}/k^{2\beta} & \text{if } k \geq K. \end{cases}$$

We then normalize $\widehat{\mathbf{x}}$ to have norm one, and obtain $\mathbf{x} = V\widehat{\mathbf{x}}$. It is clear that $\mathbf{x} \in \text{ABL}_A(K, \beta, \mu)$, where $K = 10$ and $\beta = 1$. During the sampling, we simulate the noise $\epsilon \sim \mathcal{N}(0, 0.1^2)$. In the recovery, we set the bandwidth κ to 10 for both algorithms. We consider the following three graphs.

(a) Ring graph with 4-nearest neighbor . (b) Erdős-Rényi graph. (c) Star graph.
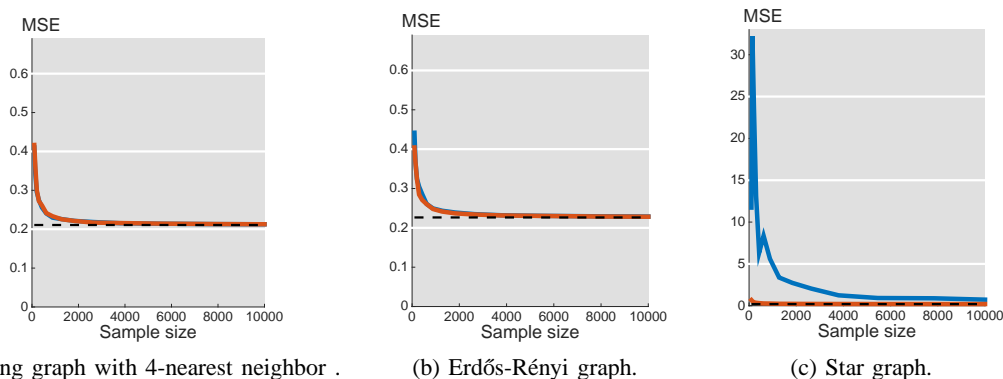
Fig. 1: Comparison of recovery error of Algorithm 1 (blue curve) and 2 (red curve). MSE indicates the mean square error.

Ring Graph with k-nearest Neighbors. We consider a graph with each node connecting to its k-nearest neighbors. The eigenvectors are similar to the discrete cosine transform and the energy evenly spreads to each element in U [12], which follows Definition 4. Based on Corollary 1, we expect that Algorithm 2 has a similar performance with Algorithm 1. In the simulation, the ring graph has 10,000 nodes, and each node connects to its 4 nearest neighbors.

Erdős-Rényi Graph. We consider a random graph where each pair of nodes is connected with some probability, also known as an Erdős-Rényi graph [13]. Since the maximum value of eigenvectors of an Erdős-Rényi graph is bounded by $O(N^{-1/2})$ [14], the energy also spreads to each element in V, which follows Definition 4. Based on Corollary 1, we expect that Algorithm 2 has a similar performance with Algorithm 1. In the simulation, the Erdős-Rényi graph has 10,000 nodes, and each pair of nodes is connected with probability of 0.01, that is, each node has 100 neighbors on expectation.

Star Graph. We consider a graph with a central node connecting to every other nodes, known as the star graph. In this case, the energy concentrates in the central node, which follows Definition 5. Based on Corollary 2, we expect that Algorithm 2 outperforms Algorithm 1. In the simulation, the star graph has 10,000 nodes.

Results. Figure 1 compares the performances between Algorithms 1 and 2 averaged over 50 tests. The blue curve represents Algorithm 1, the red curve represents Algorithm 2, and the black dotted line represented the linear approximation by the true first K frequency components. We see that both algorithms converges to the linear approximation by the first K frequency components, which supports the results in Lemmas 1 and 2. For two type-1 graphs, including the ring graph with 4-nearest neighbors and an Erdős-Rényi graph, Algorithms 1 and 2 provide similar results; however, for the star graph, Algorithm 2 performs much better than Algorithm 1, which supports the results in Corollaries 1 and 2.

VI. CONCLUSIONS

We proposed a new class of smooth graph signals, called approximated bandlimited, and we then proposed two recovery strategies based on random sampling and experimentally designed sampling. We showed that both strategies are unbiased estimators for the low-frequency components, and experimentally designed sampling outperforms random sampling when a graph is irregular. We validate the recovery strategies on three specific graphs: a ring graph, an Erdős-Rényi graph, and a star graph. The simulation results support the theoretical analysis.

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, pp. 83–98, May 2013.

[2] A. Sandryhaila and J. M. F. Moura, "Big data processing with signal processing on graphs," *IEEE Signal Process. Mag.*, vol. 31, no. 5, pp. 80 – 90, 2014.

[3] I. Z. Pesenson, "Sampling in Paley-Wiener spaces on combinatorial graphs," *Trans. Amer. Math. Soc.*, vol. 360, no. 10, pp. 5603–5627, May 2008.

[4] A. Anis, A. Gadde, and A. Ortega, "Towards a sampling theorem for signals on arbitrary graphs," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, May 2014, pp. 3864 – 3868.

[5] X. Wang, P. Liu, and Y. Gu, "Local-set-based graph signal reconstruction," *IEEE Trans. Signal Process.*, 2015, To appear.

[6] S. Chen, A. Sandryhaila, J. M. F. Moura, and J. Kovačević, "Signal recovery on graphs," *IEEE Trans. Signal Process.*, 2014, To appear.

[7] S. Chen, R. Varma, A. Sandryhaila, and J. Kovačević, "Discrete signal processing on graphs: Sampling theory," *IEEE Trans. Signal Process.*, 2015, Submitted.

[8] M. Vetterli, J. Kovačević, and V. K. Goyal, *Foundations of Signal Processing*, Cambridge University Press, 2014, http://www.fourierandwavelets.org/.

[9] J. Sharpnack and A. Singh, "Identifying graph-structured activation patterns in networks," in *Proc. Neural Information Process. Syst.*, Vancouver, Dec. 2010, pp. 2137–2145.

[10] I. M. Johnstone, *Minimax Bayes, Asymptotic Minimax and Sparse Wavelet Priors*, Statistical Decision Theory and Related Topics V. Springer, 1994.

[11] P. Drineas, M. Magdon-Ismail, M. W. Mahoney, and D. Woodruff, "Fast approximation of matrix coherence and statistical leverage," *Journal of Machine Learning Research*, vol. 13, pp. 3475–3506, 2012.

[12] A. Sandryhaila, J. Kovačević, and M. Püschel, "Algebraic signal processing theory: 1-D nearest-neighbor models," *IEEE Trans. Signal Process.*, vol. 60, no. 5, pp. 2247–2259, May 2012.

[13] M. Newman, *Networks: An Introduction*, Oxford University Press, 2010.

[14] L. V. Tran, V. H. Vu, and K. Wang, "Sparse random graphs: Eigenvalues and eigenvectors," *Random Struct. Algorithms*, vol. 42, no. 1, pp. 110–134, 2013.