FAST SAMPLING OF GRAPH SIGNALS WITH NOISE VIA NEUMANN SERIES CONVERSION

Fen Wang[†]*, Gene Cheung*[‡]*, and Yongchao Wang*[†]

[†]State Key Laboratory of ISN, Xidian University, China
[‡]Department of Electrical Engineering & Computer Science, York University, Canada

ABSTRACT

Graph sampling with independent noise towards minimum mean square error (MMSE) leads to the known A-optimality criterion, which is computation-intensive to evaluate and NP-hard to optimize. In this paper, we propose a new low-complexity sampling strategy based on Neumann series that circumvents large matrix inversion and eigen-decomposition. We first prove that a DC-shifted A-optimality criterion is equivalent to an objective computed using the inverse of a sub-matrix of an ideal graph low-pass (LP) filter. The LP filter matrix can be approximated efficiently via fast Graph Fourier Transform (FGFT). Using the shifted A-optimality objective as a proxy, we then propose a fast algorithm to greedily select samples one-by-one based on a matrix inversion lemma with simple matrix updates. We show that the obtained solution has a performance upper bound via super-modularity analysis. Simulation results show that our proposed sampling strategy has lower complexity and outperforms several existing deterministic sampling schemes.

Index Terms— Graph signal processing (GSP), sampling, optimal design, matrix inversion.

1. INTRODUCTION

Graph signal processing (GSP) studies signals that reside on irregular data kernels described by graphs [1–3]. The *spectrum* of a graph signal $\mathbf{x} \in \mathbb{R}^N$ is defined by the N eigenvectors of the graph Laplacian matrix [4] (or the adjacency matrix [2]) that span the signal space. A *bandlimited* graph signal of cutoff frequency K has non-zero coefficients only for the K eigenvectors associated with the K smallest eigenvalues. Sampling and reconstruction of bandlimited signals is a classical signal processing problem: from which subset of nodes in the graph to collect samples so that an assumed bandlimited signal can be reconstructed with high fidelity [5–7]?

Assuming an independent and identically distributed (iid) additive noise model and an unbiased *least square* (LS) signal reconstruction [8], minimizing the expected mean square error (MMSE) leads to the known *A-optimality* criterion [9]. Minimizing the A-optimality criterion for graph sampling is notoriously difficult; evaluating the criterion for a fixed sample set requires computation of eigenvectors and matrix inverse, and selection of an optimal sample set for a fixed sampling budget is NP-hard. To avoid heavy computation, [10] proposed a generic sampling scheme to minimize MSE greedily, termed *minimum Frobenius norm* (MFN). [11] used the much simpler E-optimality criterion instead for graph sampling, which can be interpreted as the worst case of MSE. [12–14] proposed a lightweight sampling strategy based on *spectral proxies* to choose stable samples. In our previous work [15], we proposed a sampling strategy based on a truncated Neumann series, called *matrix inversion approximation* (MIA). However, the number of terms in the truncated sum must be large to approximate well, limiting its practicality.

In this paper[1], we propose a new fast sampling algorithm via Neumann series expansion. Different from previous approaches [10–12], we minimize a variant of the A-optimality criterion directly but without expensive matrix computation. Yet unlike our previous work [15], there is no explicit computation of the Neumann series sum in the augmented objective, and thus no truncation / approximation tradeoff is necessary.

Specifically, we first prove that a DC-shifted A-optimality criterion can be rewritten as an objective computed using an inverse of a sub-matrix of an ideal graph low-pass (LP) filter. The LP filter matrix can be approximated efficiently using fast Graph Fourier Transform (FGFT) [17]. To minimize the augmented A-optimality criterion, we propose a greedy algorithm to choose sample nodes one-by-one based on a matrix inversion lemma, circumventing large matrix inversion or eigen-decomposition. We show that the computed solution has an upper bound in performance via super-modularity analysis [18]. Simulation results show that our proposed strategy achieves better performance than several state-of-the-art sampling schemes [11, 14, 15] with lower complexity.

The outline of the paper is as follows. In Section 2, we define essential GSP terms and the A-optimality criterion, then propose a new objective by augmenting the A-optimality function. We develop a fast sampling strategy in Section 3

This work was supported in part by National Science Foundation of China under grant 61771356, 111 project of China under grant B08038.

[1]An extended version of this paper [16] is under submission to IEEE Transactions on Signal Processing.

and analyze its performance bound. Experiments results and conclusion are presented in Section 4 and 5, respectively.

2. A-OPTIMAL GRAPH SIGNAL SAMPLING

2.1. Definitions

Denote by $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{W})$ a graph $\mathcal{G}$ containing a set of N nodes indexed by $\mathcal{V} = \{1, \ldots, N\}$. $\mathcal{E}$ specifies the set of (sparsely) connected node pairs (i, j). $\mathbf{W}$ is an $N \times N$ *adjacency matrix*, where the (i, j)-th entry $w_{i,j}$ is the weight of an edge connecting nodes i and j ($w_{i,j} = 0$ if nodes i and j are not connected). The diagonal *degree matrix* $\mathbf{D}$ is defined by $d_{i,i} = \sum_j w_{i,j}$. As done in [7], we assume connected, undirected graphs with no self-loops and adopt the symmetric *combinatorial Laplacian matrix* $\mathbf{L} = \mathbf{D} - \mathbf{W}$ as the variation operator. Assume that the eigen-decomposition of $\mathbf{L}$ is $\mathbf{L} = \mathbf{V}\mathbf{\Sigma}\mathbf{V}^\top$, where $\mathbf{V} = [\mathbf{v}_1, ..., \mathbf{v}_N]$ contains the orthonormal eigenvectors as columns corresponding to non-decreasing eigenvalues $0 = \lambda_1 < \lambda_2 \leq \lambda_3 \leq \ldots \leq \lambda_N$. The *graph Fourier transform* (GFT) of a graph signal $\mathbf{x} \in \mathbb{R}^N$ is then defined as $\tilde{\mathbf{x}} = \mathbf{V}^\top \mathbf{x}$, and the inverse GFT is $\mathbf{x} = \mathbf{V}\tilde{\mathbf{x}}$. As discussed, a K-*bandlimited* graph signal can be fully expressed by $\mathbf{x} = \mathbf{V}_K \tilde{\mathbf{x}}_K$, where $\mathbf{V}_K$ means the first K columns of $\mathbf{V}$, and $\tilde{\mathbf{x}}_K$ denotes the first K elements of $\tilde{\mathbf{x}}$ [2].

In the sequel, the complement set of sampling set $\mathcal{S}$ is denoted by $\mathcal{S}^c$ and its cardinality is written as $|\mathcal{S}|$. $\mathbf{A}_{\mathcal{S}_1\mathcal{S}_2}$ is the sub-matrix of a matrix $\mathbf{A}$ with rows and columns indexed by $\mathcal{S}_1$ and $\mathcal{S}_2$, respectively. $\mathbf{A}_{\mathcal{S}\mathcal{S}}$ is simplified to $\mathbf{A}_{\mathcal{S}}$. $\mathbf{I}$ is the identity matrix whose dimension depends on the context.

2.2. Problem Formulation

In order to select M elements from $\mathbf{x}$ to produce $\mathbf{x}_{\mathcal{S}} = \mathbf{C}\mathbf{x}$ with $|\mathcal{S}| = M$, we define the *sampling matrix* $\mathbf{C} = \mathbf{I}_{\mathcal{S}\mathcal{V}}$. A sampled K-bandlimited graph signal can now be written as $\mathbf{x}_{\mathcal{S}} = \mathbf{C}\mathbf{V}_K \tilde{\mathbf{x}}_K$. Sampling operators satisfying $\text{rank}(\mathbf{C}\mathbf{V}_K) = K$ are called *qualified sampling operators* in [11], on which any noiseless K-bandlimited graph signal $\mathbf{x}$ can be perfectly recovered from $\mathbf{x}_{\mathcal{S}}$ via LS reconsruction, *i.e.*, $\hat{\mathbf{x}}_{\text{LS}} = \mathbf{V}_K(\mathbf{C}\mathbf{V}_K)^\dagger \mathbf{x}_{\mathcal{S}}$, where $\mathbf{U}^\dagger$ denotes the left pseudo-inverse of matrix $\mathbf{U}$ [14].

If the samples are corrupted by noise, then the LS reconstruction produces a *minimum variance unbiased* (MVU) estimator $\hat{\mathbf{x}}_{\text{LS}} = \mathbf{V}_K(\mathbf{C}\mathbf{V}_K)^\dagger \mathbf{y}_{\mathcal{S}} = \mathbf{V}_K(\mathbf{C}\mathbf{V}_K)^\dagger(\mathbf{x}_{\mathcal{S}} + \mathbf{n})$ [19]. Assuming that noise $\mathbf{n}$ is iid with zero mean and unit variance, the covariance matrix of the reconstruction error is $\mathbf{R}_{\hat{\mathbf{x}}_{\text{LS}}} = \mathbf{V}_K \left[(\mathbf{C}\mathbf{V}_K)^\top \mathbf{C}\mathbf{V}_K\right]^{-1} \mathbf{V}_K^\top$. By the theory of optimal experiments design [20], finding sampling matrix $\mathbf{C}$ to minimize the trace of the covariance matrix leads to the known *A-optimality* criterion:

$$\mathbf{C}^* = \underset{\mathbf{C} \in \mathbb{F}^{M \times N}}{\arg\min}\ \text{tr}\left([(\mathbf{C}\mathbf{V}_K)^\top \mathbf{C}\mathbf{V}_K]^{-1} \right). \tag{1}$$

where $\mathbb{F}^{M \times N}$ is the search space of all candidates $\mathbf{C}$ with structure $\mathbf{I}_{\mathcal{S}\mathcal{V}}$. It is NP-hard to find a feasible $\mathbf{C}^*$ that minimizes (1). Further, for each candidate solution $\mathbf{C}$, evaluating (1) requires the computation of the first K eigenvectors $\mathbf{V}_K$ and matrix inversion, both of which are expensive.

2.3. Augmented A-optimality Criterion

We first propose to augment the A-optimality criterion by adding a scaled identity matrix to (1):

$$\mathbf{C}^* = \underset{\mathbf{C} \in \mathbb{F}^{M \times N}}{\arg\min}\ \text{tr}\left([(\mathbf{C}\mathbf{V}_K)^\top \mathbf{C}\mathbf{V}_K + \mu\mathbf{I}]^{-1} \right) \tag{2}$$

μ is a small shift parameter with $0 < \mu < 1$, whose selection will be discussed later.

Proposition 1 *Denote by* $\mathbf{T} = \mathbf{V}_K \mathbf{V}_K^\top$ *an ideal graph LP filter with cutoff frequency* K*. The augmented A-optimal graph sampling problem* (2) *is equivalent to*

$$\mathcal{S}^* = \underset{\mathcal{S}:|\mathcal{S}|=M}{\arg\min}\ \text{tr}\left(\mathbf{T}_{\mathcal{S}} + \mu\mathbf{I}\right)^{-1}, \tag{3}$$

given that $\mathbf{C}\mathbf{V}_K$ *is full column rank.*

The proof, based on the Neumann series theorem [22], is provided in the extended version [16]. We note that $\text{rank}(\mathbf{C}\mathbf{V}_K) = K$ was shown by experiments to be highly probable via random sample selection when $M \geq K$ [23]. We assume $\text{rank}(\mathbf{C}\mathbf{V}_K) = K$ hold in the sequel.

2.4. Ideal Graph Low-pass Filter Approximation

One common approximation to $\mathbf{T}$ is a *Chebyshev matrix polynomial* of $\mathbf{L}$, *i.e.*, $\mathbf{T}^{\text{Poly}} = \sum_{i=1}^{n} \left(\sum_{j=0}^{q} \beta_j \lambda_i^j \right) \mathbf{v}_i \mathbf{v}_i^\top = \sum_{j=0}^{q} \beta_j \mathbf{L}^j$. Another approximation is to apply the Jacobi eigenvalue algorithm: approximately diagonalize $\mathbf{L}$ with an estimated eigenvector matrix $\tilde{\mathbf{V}} = \mathbf{S}_1 ... \mathbf{S}_J$, where $\mathbf{S}_i$ is a *Givens rotation matrix* [17]. The goal is to solve the following optimization problem:

$$\underset{\hat{\mathbf{\Lambda}}, \mathbf{S}_1, \ldots, \mathbf{S}_J}{\text{minimize}} \quad \|\mathbf{L} - \mathbf{S}_1 ... \mathbf{S}_J \hat{\mathbf{\Lambda}} \mathbf{S}_J^\top ... \mathbf{S}_1^\top\|_F^2 \tag{4}$$

where $\hat{\mathbf{\Lambda}}$ is constrained to be diagonal. Given the computed $\tilde{\mathbf{V}}$, where $\mathbf{L} \approx \tilde{\mathbf{V}}\hat{\mathbf{\Lambda}}\tilde{\mathbf{V}}^\top$, the ideal LP filter operator $\mathbf{T}$ in (3) can be approximated as $\mathbf{T}^{\text{FGFT}} = \tilde{\mathbf{V}}_K \tilde{\mathbf{V}}_K^\top$ [3].

We adopt $\mathbf{T}^{\text{FGFT}}$ to approximate $\mathbf{T}$ for two reasons:

[2] Authors in [18] defined the $\mathcal{K}$-spectrally sparse graph signal by $\mathbf{x} = \mathbf{V}_{\mathcal{K}} \tilde{\mathbf{x}}_{\mathcal{K}}$ which is a generalization of our formulation since $\mathcal{K} \in \mathcal{V}$ is not restricted to be the first K elements. We use the more conventional definition here, but our algorithm extends naturally to the generalized case.

[3] We prove in [16] that the eigenvalues of $\mathbf{T}^{\text{FGFT}}$ are in $\{0, 1\}$ and that of $\mathbf{T}_{\mathcal{S}}^{\text{FGFT}}$ are in [0,1].

Algorithm 1 GFS graph signal sampling algorithm

Input: Graph operator $\mathbf{L}$, bandwidth K, budget M and parameter μ. $\mathcal{S} = \{\varnothing\}$.

1: Compute $\tilde{\mathbf{V}} = \mathbf{S}_1 ... \mathbf{S}_J$ of $\mathbf{L}$ via (4).
2: Compute $\mathbf{T}^{\text{FGFT}} = \tilde{\mathbf{V}}_K \tilde{\mathbf{V}}_K^\top$ and $\mathbf{G} = \mathbf{T}^{\text{FGFT}} + \mu\mathbf{I}$.
3: Select the first node by $u = \arg\max_{i \in \mathcal{V}} \mathbf{G}_{ii}$, update $\mathcal{S} \leftarrow \mathcal{S} \cup \{u\}$ and $\mathbf{G}_{\mathcal{S}}^{-1} = \mathbf{G}_{uu}^{-1}$
4: **While** $|\mathcal{S}| < M$
5: $\forall i \in \mathcal{S}^c$, compute $\mathbf{g}_i = \mathbf{G}_{\mathcal{S},\{i\}}$ and $a = \mathbf{G}_{ii} - \mathbf{g}_i^\top \mathbf{G}_{\mathcal{S}}^{-1} \mathbf{g}_i$

$$\mathbf{G}_{\mathcal{S}\cup\{i\}}^{-1} = \begin{bmatrix} \mathbf{G}_{\mathcal{S}}^{-1} + a^{-1}\mathbf{G}_{\mathcal{S}}^{-1}\mathbf{g}_i\mathbf{g}_i^\top\mathbf{G}_{\mathcal{S}}^{-1} & -a^{-1}\mathbf{G}_{\mathcal{S}}^{-1}\mathbf{g}_i \\ -a^{-1}\mathbf{g}_i^\top\mathbf{G}_{\mathcal{S}}^{-1} & a^{-1} \end{bmatrix}$$

6: Select $u = \arg\min_{i \in \mathcal{S}^c} \text{tr}\left[\mathbf{G}_{\mathcal{S}\cup\{i\}}^{-1}\right]$
7: Update $\mathbf{G}_{\mathcal{S}}^{-1} = \mathbf{G}_{\mathcal{S}\cup\{u\}}^{-1}$ and $\mathcal{S} \leftarrow \mathcal{S} \cup \{u\}$
8: **end While**
9: Return $\mathcal{S}$

1. [17] demonstrated that $\frac{\|\mathbf{T}-\mathbf{T}^{\text{FGFT}}\|_F}{\|\mathbf{T}\|_F} < \frac{\|\mathbf{T}-\mathbf{T}^{\text{Poly}}\|_F}{\|\mathbf{T}\|_F}$ for the same complexity when approximating ideal graph LP filters, such as $\mathbf{T}$ in our work.

2. Proposition 1 holds for $\mathbf{T}^{\text{FGFT}}$, detailed in [16]. In contrast, $\mathbf{T}^{\text{Poly}}$ does not have such property.

We now write our final augmented sampling objective as follows:

$$\mathcal{S}^* = \arg\min_{\mathcal{S}:|\mathcal{S}|=M} \text{tr}\left(\mathbf{T}_{\mathcal{S}}^{\text{FGFT}} + \mu\mathbf{I}\right)^{-1} \tag{5}$$

which requires just an approximated LP filter operator, without full eigen-decomposition of $\mathbf{L}$. For simplicity of presentation, we write $\mathbf{G} = \mathbf{T}^{\text{FGFT}} + \mu\mathbf{I}$ in the sequel.

2.5. Selection of Shift Parameter μ

To well approximate the original criterion (1), shift μ should be small. However, a small μ would cause the matrix inverse in (5) to be numerically unstable. To ensure numerical stability, we can design μ to upper-bound the *condition number* $\kappa(\mathbf{G})$ of $\mathbf{G}$, where $\kappa(\mathbf{G}) = \lambda_{\max}(\mathbf{G})/\lambda_{\min}(\mathbf{G})$. Because the eigenvalues of $\mathbf{T}_{\mathcal{S}}^{\text{FGFT}}$ are in [0,1], $\mu \leq \lambda(\mathbf{G}_{\mathcal{S}}) \leq 1 + \mu$. Therefore, we can bound $\kappa(\mathbf{G}_{\mathcal{S}})$ as follows [4]:

$$\kappa(\mathbf{G}_{\mathcal{S}}) = \frac{\lambda_{\max}(\mathbf{G}_{\mathcal{S}})}{\lambda_{\min}(\mathbf{G}_{\mathcal{S}})} \leq \frac{1+\mu}{\mu} \leq \kappa_0 \tag{6}$$

where κ_0 is a user-defined upper bound of the condition number. Because μ should be chosen as small as possible, the optimal solution is $\mu^* = \frac{1}{\kappa_0 - 1}$.

[4] Designing μ via explicit eigenvalues is impractical, since those values are related to μ. We here bound a relaxed $\kappa(\mathbf{G})$ to find a reasonable μ.

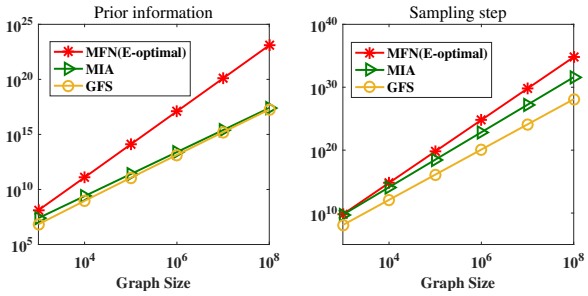


Fig. 1. Numerical complexity comparison with $M = 0.05N$.

3. FAST SAMPLING WITH PERFORMANCE GUARANTEE

Greedy algorithms have been commonly used to select one sample node at a time given a sampling budget [5, 11, 14]. However, if (5) is used as a criterion for greedy node selection, then the algorithm must perform matrix inversion operation, *i.e.*, $\mathbf{G}_{\mathcal{S}}^{-1}$, for each candidate additional node. To circumvent this computation burden, we propose the following accelerated greedy algorithm.

3.1. Accelerated Greedy Sampling

We observe that matrix $\mathbf{G}$ is symmetric, and under appropriate permutation, its sub-matrix $\mathbf{G}_{\mathcal{S}\cup\{i\}}$ can be expressed as

$$\mathbf{G}_{\mathcal{S}\cup\{i\}} = \begin{bmatrix} \mathbf{G}_{\mathcal{S}} & \mathbf{G}_{\mathcal{S},\{i\}} \\ \mathbf{G}_{\{i\},\mathcal{S}} & \mathbf{G}_{ii} \end{bmatrix} \triangleq \begin{bmatrix} \mathbf{G}_{\mathcal{S}} & \mathbf{g}_i \\ \mathbf{g}_i^\top & \mathbf{G}_{ii} \end{bmatrix} \tag{7}$$

where $\mathbf{g}_i$ denotes the partial vector of i-th column of $\mathbf{G}$ indexed by $\mathcal{S}$.

Using the *block matrix inversion formula* [5], we can compute the inverse of $\mathbf{G}_{\mathcal{S}\cup\{i\}}$ as

$$\mathbf{G}_{\mathcal{S}\cup\{i\}}^{-1} = \begin{bmatrix} \mathbf{G}_{\mathcal{S}}^{-1} + a^{-1}\mathbf{G}_{\mathcal{S}}^{-1}\mathbf{g}_i\mathbf{g}_i^\top\mathbf{G}_{\mathcal{S}}^{-1} & -a^{-1}\mathbf{G}_{\mathcal{S}}^{-1}\mathbf{g}_i \\ -a^{-1}\mathbf{g}_i^\top\mathbf{G}_{\mathcal{S}}^{-1} & a^{-1} \end{bmatrix} \tag{8}$$

where $a = \mathbf{G}_{ii} - \mathbf{g}_i^\top \mathbf{G}_{\mathcal{S}}^{-1} \mathbf{g}_i$ is a scalar. We write the proposed sampling method in pseudo-code in Algorithm 1 and call it *graph filter submatrix* (GFS)-based sampling algorithm.

3.1.1. Complexity Analysis

$\mathbf{G}_{\mathcal{S}\cup\{i\}}^{-1}$ can be computed as simple matrix-vector product, and thus for each greedy step, its complexity is $\mathcal{O}(M^2)$. In contrast, the complexity of each greedy step for MFN and the E-optimal [14] are both $\mathcal{O}(M^3)$, while for MIA [15] it is $\mathcal{O}(LM^{2.373})$. The complexity of computing $\mathbf{S}_1, ..., \mathbf{S}_J$ is $\mathcal{O}(NJ\log N)$, where $J = \mathcal{O}(N\log N)$ [17]. $\mathbf{T}^{\text{FGFT}}$ is computed as $\mathbf{T}^{\text{FGFT}} = \mathbf{S}_1 ... \mathbf{S}_J \mathbf{R} \mathbf{S}_J^\top ... \mathbf{S}_1^\top$, where $\mathbf{R}$ is a diagonal

[5] https://en.wikipedia.org/wiki/Block_matrix

matrix with ones on the first K diagonal elements and zeros for others. Since every Givens matrix $\mathbf{S}_i$ is sparse, the complexity of iteratively computing $\mathbf{T}^{\text{FGFT}}$ is $\mathcal{O}(NJ)$. Thus, to obtain the *prior information*, like $\mathbf{T}$ in our paper or $\mathbf{V}_K$ in E-optimal method, our algorithm has complexity $\mathcal{O}(N^2\log^2 N)$. In Fig. 1, we illustrate the complexity comparison in terms of graph size, demonstrating that our proposal is more efficient.

3.2. Super-modularity Analysis

We define $f(\mathcal{S}) = \text{tr}\left(\mathbf{T}_{\mathcal{S}}^{\text{FGFT}} + \mu\mathbf{I}\right)^{-1}$ and $g(\mathcal{S}) = \text{tr}[(\mathbf{C}\tilde{\mathbf{V}}_K)^\top \mathbf{C}\tilde{\mathbf{V}}_K + \mu\mathbf{I}]^{-1}$. It is easy to derive that the solution of greedily optimizing $f(\mathcal{S})$ is exactly the same as that of greedily optimizing $g(\mathcal{S})$ because Proposition 1 still holds for orthonormal matrix $\tilde{\mathbf{V}}$. Since $g(\mathcal{S})$ approximates the real MSE value, it is worthwhile to compare its value with the optimal one, thus we investigate its super-modularity property.

Lemma 1 *The set function* $g(\mathcal{S}) = \text{tr}[(\mathbf{C}\tilde{\mathbf{V}}_K)^\top \mathbf{C}\tilde{\mathbf{V}}_K + \mu\mathbf{I}]^{-1}$ *is (i) monotone decreasing and (ii)* α*-supermodular with*

$$\alpha \geq \frac{\mu^3(\mu+2)}{(\mu+1)^4} \tag{9}$$

The proof is provided in the extended paper [16].

Let $\mathcal{S}^*$ be the optimal solution that minimizes $g(\mathcal{S})$ when $|\mathcal{S}| = M$, and $\mathcal{S}$ be the set obtained by applying the GFS algorithm. We obtain a result using Lemma 1:

$$\frac{g(\mathcal{S}) - g(\mathcal{S}^*)}{g(\{\}) - g(\mathcal{S}^*)} \leq e^{-\alpha} \tag{10}$$

which is easily derived from Theorem 2 in [18]. This equation means that the performance of the greedy solution $\mathcal{S}$ is upper-bounded by $g(\mathcal{S}^*)$.

4. EXPERIMENTAL RESULTS

All experiments were conducted in MATLAB R2017b, running on a PC with Intel Core I3 3.7 GHz CPU and 16GB RAM. The simulated artificial graphs [6] are described as follows:

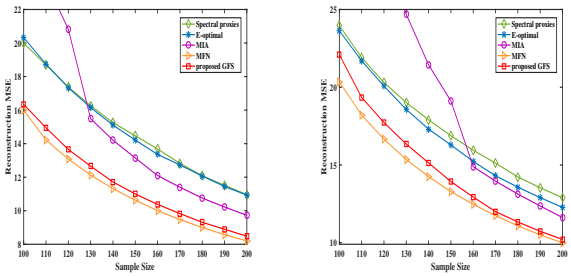
(G1) Community graphs with 1000 nodes, where the number of communities is 31;

(G2) Sensor graphs with 1000 nodes.

Artificial graph signals are assumed to be bandlimited with bandwidth $K = 50$. The non-zero GFT coefficients are randomly generated from the distribution $\mathcal{N}\left(1, 0.5^2\right)$, and the coefficients after $K = 50$ are zeros. The generated graph signals are corrupted by additional white Guassian noise (AWGN), where the signal-to-noise ratio (SNR) is 0 dB.

The truncation degree in the MIA algorithm [15] is set to be $L = 10$. One GSP toolbox [25] is adopted to approximate the ideal LP filter in the MIA method, where $p = 25$ and $\alpha = 30$ [13]. For the spectral proxies algorithm [14], the approximation order is $k = 10$. For the proposed GFS sampling,

[6] All of these graphs are generated from a GSP open source [25].



(a) MSE in **G1** at 0dB (b) MSE in **G2** at 0dB

Fig. 2. Reconstruction MSE of different sampling algorithms.

Table 1. Reconstruction MSE of Different μ at 0dB

Graph	μ	Sample size				
		100	110	120	130	140
G1	10^{-5}	16.10	14.55	13.43	12.44	11.63
	$1/99$	16.07	14.59	13.43	12.46	11.64
G2	10^{-5}	20.77	18.68	17.09	15.77	14.63
	$1/99$	20.77	18.73	17.12	15.78	14.64

the shift parameter μ is designed to be $1/(\kappa_0 - 1)$, where we set $\kappa_0 = 100$. The number of Givens rotations matrices is $J = 6N\log N$ in this simulation.

Fig. 2 shows the reconstruction MSE of different sampling schemes when the reconstruction method is LS. The performance of random sampling for the same sample size is too poor and thus is deleted in this figure. Fig. 2 demonstrates that our proposed GFS sampling algorithm achieves better MSE performance than three other sampling schemes, and closely approximates MFN for both community and sensor graphs at 0 dB. As discussed, MFN greedily optimizes the original MSE criterion, whose complexity is substantial for large graphs. In contrast, our proposed strategy can have comparable performance with much lower complexity.

For our sampling criterion, a customized μ based on condition number will introduce some approximation error. Results in Table 1 show that the performance of $\mu = 1/(100-1)$ is almost the same as that of an extremely small μ. This indicates that the designed μ for numerical stability does not introduce a noticeable performance gap compared to the augmented criterion with a negligible shift.

5. CONCLUSION

Based on Neumann series conversion, we propose the GFS sampling strategy to select samples greedily without full eigen-decomposition and large matrix inversion. Compared with several existing graph sampling methods, GFS achieves superior or comparable performance with lower complexity, demonstrated in simulations using large artificial graphs.

6. REFERENCES

[1] A. Ortega, P. Frossard, J. Kovačević, J. M. F. Moura, P. Vandergheynst, "Graph signal processing: Overview, challenges and applications, " *Proceedings of the IEEE*, vol. 106, no. 5, pp. 808-828, May. 2018.

[2] A. Sandryhaila and J. Moura, "Big data analysis with signal processing on graphs: Representation and processing of massive data sets with irregular structure," *IEEE Signal Process. Mag.*, vol. 31, no. 5, pp. 80-90, Sep. 2014.

[3] G. Cheung, E. Magli, Y. Tanaka, M. Ng, "Graph Spectral Image Processing," *Proceedings of the IEEE*, vol. 106, no. 5, pp. 907-930, May. 2018.

[4] D. Shuman, S. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," *IEEE Signal Process. Mag.*, vol. 30, no. 3, pp. 83-98, May. 2013.

[5] P. Di Lorenzo, P. Banelli, E. Isufi, S. Barbarossa and G. Leus, "Adaptive graph signal processing: Algorithms and optimal sampling strategies,"*IEEE Trans. Signal Process.*, vol. 66, no. 13, pp. 3584-3598, Jul. 2018.

[6] A. Marques, S. Segarra, G. Leus, and A. Ribeiro,"Sampling of graph signals with successive local aggregations," *IEEE Trans. Signal Process.*, vol. 64, no. 7, pp. 1832-1843, 2016.

[7] A. Sakiyama, Y. Tanaka, T. Tanaka, and A. Ortega, "Accelerated sensor position selection using graph localization operator," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, New Orleans, Mar. 2017.

[8] G. Ortiz-Jiménez, M. Coutino, S. P. Chepuri, and G. Leus, "Sparse Sampling for Inverse Problems with Tensors," Jun. 2018, arXiv:1806.10976 [cs.IT].

[9] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.

[10] M. Tsitsvero, S. Barbarossa, and P. Di Lorenzo, "Signals on graphs: Uncertainty principle and sampling," *IEEE Trans. Signal Process.*, vol. 64, no. 18, pp. 4845-4860, Sep.15, 2016.

[11] S. Chen, R. Varma, A. Sandryhaila, and J. Kovačević, "Discrete signal processing on graphs: Sampling theory," *IEEE Trans. Signal Process.*, vol. 63, no. 24, pp. 6510-6523, 2015.

[12] 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. (ICASSP)*, Florence, Italy, pp. 3864-3868, May. 2014.

[13] A. Gadde, A. Anis, and A. Ortega, "Active semi-supervised learning using sampling theory for graph signals," in *Proc. 20th ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, pp. 492-501, 2014.

[14] A. Anis, A. Gadde, and A. Ortega, "Efficient sampling set selection for bandlimited graph signals using graph spectral proxies," *IEEE Trans. Signal Process.*, vol. 64, no. 14, pp. 3775-3789, Jul. 2016.

[15] F. Wang, Y. Wang and G. Cheung, "A-optimal sampling and robust reconstruction for graph signals via truncated Neumann series," *IEEE Signal Processing Letters*, vol. 25, no. 5, pp. 680-684, May. 2018.

[16] F. Wang, G. Cheung and Y. Wang, "Low-complexity Graph Sampling and Signal Reconstruction with Noise via Neumann Series" submitted to *IEEE Trans. Signal Process.*, Nov. 2018.

[17] L. Le Magoarou, R. Gribonval and N. Tremblay, "Approximate fast graph Fourier transforms via multilayer sparse approximations," in *IEEE Transactions on Signal and Information Processing over Networks*, vol. 4, no. 2, pp. 407-420, Jun. 2018.

[18] L. F. O. Chamon and A. Ribeiro, "Greedy sampling of graph signals," *IEEE Trans. Signal Process.*, vol. 66, no. 1, pp. 34-47, Jan.1, 2018.

[19] Steven M. Kay, *Fundermentals of Statistical Signal Processing: Estimation Theory*, Prentice-Hall, 1993.

[20] F. Pukelsheim, *Optimal Design of Experiments*. Philadelphia, PA, USA: SIAM, 1993, vol. 50.

[21] G. Golub and C. F. V. Loan, *Matrix Computations* (Johns Hopkins Studies in the Mathematical Sciences). Johns Hopkins University Press, 2012.

[22] R. A. Horn and C. R. Johnson, *Matrix Analysis.* Cambridge University Press, 2013.

[23] I. Shomorony and A. Avestimehr, "Sampling large data on graphs," in *Proc. IEEE Global Conf. Signal Inf. Process. (GlobalSIP)*, pp. 933-936, Dec. 2014.

[24] D. Hammond, P. Vandergheynst, and R. Gribonval, "Wavelets on graphs via spectral graph theory," *Appl. Comput. Harmon. Anal.*, vol. 30, no. 2, pp. 129-150, 2011.

[25] N. Perraudin, J. Paratte, D. Shuman, V. Kalofolias, P. Vandergheynst, and D. K. Hammond, "GSPBOX: A toolbox for signal processing on graphs," Aug. 2014, arXiv:1408.5781 [cs.IT].