ACCELERATED SENSOR POSITION SELECTION USING GRAPH LOCALIZATION OPERATOR

Akie Sakiyama[1], *Yuichi Tanaka*[1,2], *Toshihisa Tanaka*[1] *and Antonio Ortega*[1,3]

[1]Tokyo University of Agriculture and Technology, Tokyo, Japan
[2]PRESTO, Japan Science and Technology Agency, Saitama, Japan
[3]University of Southern California, Los Angeles, USA

ABSTRACT

This paper addresses the problem of finding optimal sensor placement, i.e., determining F sensor positions from N possible locations. We propose a sensor selection method based on the localization operator of graph signal processing. This method can select sensors while considering the localizations both in graph vertex domain and graph spectral domain and is fast, since eigendecomposition of graph Laplacian matrix is not required. We also propose an interpretation of the conventional node selection based on graph sampling theory by using the graph localization operators. Experiments on selected sensor location, execution time and prediction error comparisons are conducted to show the effectiveness of our approach.

***Index Terms*—** Sensor placement, graph signal processing, graph sampling theorem, localization operator, graph uncertainty principle

1. INTRODUCTION

Sensor networks are commonly used in many applications to monitor and control spatial phenomena. They include sensing of temperature [1], air quality and/or rainfall, and monitoring of smart grid systems [2]. In real applications, the number of sensors is often restricted due to economic constraints, data storage and energy saving. Therefore, it is important to optimize sensor placement so that a small number of sensors can collect the spatial data all over the observed space. There are many approaches to address the sensor placement problem, such as [2–5], where the goal is to choose $|\mathcal{S}| = F$ locations for sensor placement out of $|\mathcal{V}| = N$ possible locations, where $\mathcal{S} \subseteq \mathcal{V}$ and $\mathcal{V} = \{v_0, \ldots, v_{N-1}\}$ are the sets of measured locations and possible locations, respectively.

In our previous work [6], we have proposed a sensor selection method based on the sampling theory for graph signals [7–14]. This was the first attempt to use graph signal processing for the sensor placement problem. Approaches from graph signal processing can be useful for sensor placement because graphs are well suited to capture complete relationships between sensors, when sensors are placed at irregular locations in the environment. In [6], we interpreted conventional sensor selection methods, namely, entropy [15,16] and mutual information (MI) [17,18] based criteria, from the perspective of graph signal processing. The sensor selection in [6] uses three kinds of graph sampling methods [11–13], and all of them outperform conventional entropy and MI based criteria. However, they have high computational complexity since they need the eigendecomposition [12, 13] or calculating leading eigenvector(s) [11] of graph Laplacian matrix.

In this paper, we propose a sensor selection method based on the localization operator introduced in the context of uncertainty principle of graph signals [19]. This method places sensors so that localized operators cover the overall area by optimizing the cost function with greedy heuristics. Therefore, we can select sensors deterministically. It can provide flexibility for the change in the number of sensors, that is, it can append or remove sensors one-by-one instead of selecting all sensors every time when the number of sensors changes. The benefits of using the localization operators are: a) spectral localization makes it possible to mimic the frequency-based node selection criteria of [6, 11–13], b) vertex localization is useful to enable distributed sensor selection (to be applied independently in different sections of the sensor network), and c) polynomial localization operators lead to lower complexity, i.e., eigendecomposition-free algorithms. We also show that the conventional graph sampling approaches can be viewed as a node selection that chooses nodes having the maximum value of a localization operator based on an ideal kernel. In the experiment, we show the results of the sensor placement and the execution time. We also present prediction error comparisons to evaluate the performance of the proposed approach. The proposed method is approximately 10 times faster than the approach in [13], 50 times faster than that in [11], and 1000 times faster than that in [12], with comparable performance.

The rest of this paper is organized as follows. Preliminaries on graph signal processing are summarized in Section 2. Section 3 describes the proposed selection algorithm based on the graph localization operator. The section also explains the connection between the graph sampling and the localization operator. Section 4 shows the experimental results on selecting sensor locations and predicting the signals on unobserved locations. Finally, Section 5 concludes the paper.

2. PRELIMINARIES ON GRAPH SIGNAL PROCESSING

A graph is represented as $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ and $\mathcal{E}$ denote sets of nodes and edges, respectively. The graph signal is defined as $\boldsymbol{f} \in \mathbb{R}^N$. We will only consider a connected, finite, undirected graph with no multiple edges. The number of nodes is $N = |\mathcal{V}|$, unless otherwise specified. The (m, n)-th element of the adjacency matrix $\mathbf{A}$ is the weight of the edge between m and n if m and n are connected, and 0 otherwise. The degree matrix $\mathbf{D}$ is a diagonal matrix and its mth diagonal element is $D(m, m) = \sum_n A(m, n)$. The unnormalized graph Laplacian matrix (GLM) is defined as $\mathbf{L} := \mathbf{D} - \mathbf{A}$ and the symmetric normalized GLM is $\boldsymbol{\mathcal{L}} := \mathbf{D}^{-1/2}\mathbf{L}\mathbf{D}^{-1/2}$. The

This work was supported in part by MEXT Tenure-Track Promotion Program, JSPS KAKENHI Grant Number JP16H04362 and JP16J04808, and US NSF under grant NSF-CCF-1527874. A. Ortega's visit was supported by the Institute of Global Innovation Research at TUAT.

symmetric normalized GLM has the property that its eigenvalues are within the interval $[0, 2]$. The eigenvalues of $\mathbf{L}$ or $\boldsymbol{\mathcal{L}}$ are λ_i and ordered as: $0 = \lambda_0 < \lambda_1 \leq \lambda_2 \ldots \leq \lambda_{N-1} = \lambda_{\text{max}}$ without loss of generality. The eigenvector $\boldsymbol{u}_i$ corresponds to λ_i and satisfies $\boldsymbol{\mathcal{L}}\boldsymbol{u}_i = \lambda_i \boldsymbol{u}_i$. The eigenvectors $\mathbf{U} = [\boldsymbol{u}_0 \ldots \boldsymbol{u}_{N-1}]$ satisfy $\mathbf{U}\mathbf{U}^\dagger = \mathbf{I}_N$ where $\cdot^\dagger$ is the conjugate transpose of a matrix or a vector and $\mathbf{I}_N$ is the $N \times N$ identity matrix. The entire spectrum of $\mathcal{G}$ is defined by $\sigma(\mathcal{G}) := \{\lambda_0, \ldots, \lambda_{N-1}\}$. The graph Fourier transform is defined as follows [20, 21]: $\overline{f}(\lambda_i) := \langle \boldsymbol{u}_i, \boldsymbol{f} \rangle = \sum_{n=0}^{N-1} u_i^*(n) f(n)$, where $\cdot^*$ is the complex conjugate. The inverse graph Fourier transform is $f(i) = \sum_{n=0}^{N-1} u_i(n) \overline{f}(n)$. The projection matrix for the eigenspace V_{λ_i} is $\mathbf{P}_{\lambda_i} = \sum_{\lambda=\lambda_i} \boldsymbol{u}_\lambda \boldsymbol{u}_\lambda^T$ where $\boldsymbol{u}_\lambda^T$ is the transpose of $\boldsymbol{u}_\lambda$. Let $h(\lambda_i)$ be the spectral kernel of filter $\mathbf{H}$. The spectral domain filter can be written as $\mathbf{H} = h(\boldsymbol{\mathcal{L}}) = \sum_{\lambda_i \in \sigma(\mathcal{G})} h(\lambda_i) \mathbf{P}_{\lambda_i}$. The spectral domain filtering of graph signals can be simply denoted as $\mathbf{H}\boldsymbol{f}$. For a vector $\boldsymbol{x} \in \mathbb{R}^N$ and a set $\mathcal{A}$, the restriction of $\boldsymbol{x}$ to its components indexed by $\mathcal{A}$ is denoted by $\boldsymbol{x}_\mathcal{A}$. For a matrix $\mathbf{X} \in \mathbb{R}^{N \times N}$, $\mathbf{X}_{\mathcal{A}\mathcal{B}}$ denotes the restriction matrix of $\mathbf{X}$, which is obtained from extracting rows indexed by $\mathcal{A}$ and columns indexed by $\mathcal{B}$ from $\mathbf{X}$, and $\mathbf{X}_\mathcal{A} := \mathbf{X}_{\mathcal{A}\mathcal{A}}$.

2.1. Localization Operator

The localization operator on center vertex i is defined as [19]

$$T_i g(n) = \sqrt{N} \sum_{l=0}^{N-1} \widehat{g}(\lambda_l) u_l^*(i) u_l(n), \tag{1}$$

where $\widehat{g}(\lambda)$ is an arbitrary filter kernel. Its ℓ_2-norm is represented as

$$\|\boldsymbol{T}_i \boldsymbol{g}\|_2 = \sqrt{\sum_{n=0}^{N-1} \left(\sqrt{N} \sum_{l=0}^{N-1} \widehat{g}(\lambda_l) u_l^*(i) u_l(n) \right)^2}. \tag{2}$$

3. SENSOR SELECTION BASED ON GRAPH LOCALIZATION OPERATOR

This section introduces the novel sensor placement method which selects sensors based on the graph localization operator. Furthermore, we describe the existing graph sampling approach by using the graph localization operator.

3.1. Proposed Method

The sensor selection algorithm introduced in this subsection has the following properties: a) simultaneous localization both in graph vertex domain and graph spectral domain, b) deterministic selection, c) selection without full eigendecomposition of graph Laplacian matrix.

Firstly, we construct a graph from the possible sensor positions and we treat the observed signal as a graph signal. The connection of nodes (edges) can be freely chosen by the application, for example, we can use a data-driven graph (sparse inverse covariance), a distance based graph (negative exponential of distance) or the approximate graph. In this paper, for the sake of comparison with the conventional machine learning-based approaches, we assume that the graph signal $\boldsymbol{f}$ is stochastic and has the following Gaussian joint zero-mean distribution [22]:

$$p(\boldsymbol{f}) = \frac{1}{(2\pi)^{\frac{N}{2}} |\mathbf{K}|} \exp\left(-\tfrac{1}{2} \boldsymbol{f}^T \mathbf{K}^{-1} \boldsymbol{f}\right), \tag{3}$$

where $\cdot^T$ is the transpose of a matrix or a vector, $\mathbf{K} \in \mathbb{R}^{|\mathcal{V}| \times |\mathcal{V}|}$ is the covariance matrix of all locations $\mathcal{V}$ in which its (i, j)th element is $\mathcal{K}(i, j)$ with a symmetric positive-definite kernel function $\mathcal{K}(\cdot, \cdot)$, and $|\mathbf{K}|$ is the determinant of $\mathbf{K}$. The graph Laplacian matrix can be obtained from the inverse covariance matrix (precision matrix) [10]:

$$\mathbf{L} = \mathbf{K}^{-1} - \delta \mathbf{I}. \tag{4}$$

The parameter δ prevents the precision matrix from being singular. $\mathbf{L}$ in (4) is a "generalized" graph Laplacian matrix, i.e., it could have negative edges and self loops. The graph used in this paper is obtained by pruning the self loops and negative edges from the graph constructed from (4). We can also use a graph Laplacian estimated directly from prior information, e.g., by using physical networks of power systems, when the signals do not have the distribution (3).

For sensor selection, each coordinate of the vector $\boldsymbol{T}_i \boldsymbol{g}$ can be regarded as the quantitative measure of the coordinate belonging to the ith sensing area. Note that it is the vertex domain operator with consideration of the spectral domain information. We select the set of sensor locations $\mathcal{S}$ so that $\boldsymbol{T}_i \boldsymbol{g}$ $(i \in \mathcal{S})$ covers overall area evenly, i.e., the sum of $\|\boldsymbol{T}_i \boldsymbol{g}\|_2^2$ $(i \in \mathcal{S})$ is large and the overlapping covered by both $\boldsymbol{T}_i \boldsymbol{g}$ and $\boldsymbol{T}_j \boldsymbol{g}$ $(i \neq j)$ is small. Such set will be obtained by optimizing the following function:

$$\begin{aligned} &\max_\mathcal{S} \left\| \textstyle\sum_{i \in \mathcal{S}} \boldsymbol{T}_i \boldsymbol{g} \circ \boldsymbol{T}_i \boldsymbol{g} - \sum_{i \in \mathcal{S}, j \in \mathcal{S}, i \neq j} \boldsymbol{T}_j \boldsymbol{g} \circ \boldsymbol{T}_i \boldsymbol{g} \right\|_2^2 \\ &= \left\| \textstyle\sum_{i \in \mathcal{S}, j \in \mathcal{S}, i \neq j} (\boldsymbol{T}_i \boldsymbol{g} - \boldsymbol{T}_j \boldsymbol{g}) \circ \boldsymbol{T}_i \boldsymbol{g} \right\|_2^2, \end{aligned} \tag{5}$$

where $\circ$ represents the Hadamard product. To optimize the cost function, we use a greedy algorithm, which appends one sensor in the mth iteration by selecting sensor y^* satisfying the following function:

$$y^* = \underset{y \in \mathcal{S}_m^c}{\arg\max} \left\| \boldsymbol{\tau} \left(C \mathbf{1}_{N \times 1} - \textstyle\sum_{j \in \mathcal{S}_m} |\boldsymbol{T}_j \boldsymbol{g}| \right) \circ \boldsymbol{T}_y \boldsymbol{g} \right\|_2^2, \tag{6}$$

where $\mathcal{S}_m$ is the already selected sensors in the mth iteration, $\mathcal{S}_m^c = \mathcal{V} \setminus \mathcal{S}_m$, $\boldsymbol{\tau}(\cdot)$ is a function that satisfies $[\tau(\boldsymbol{x})](i) = x(i)$ if $x(i) \geq 0$ and 0 otherwise, and $C \in \mathbb{R}$ is an arbitrary real value.

In (6), we calculate the weighted norm of $\boldsymbol{T}_y \boldsymbol{g}$. A small weight is assigned to $T_y g(i)$ if the ith node has already been covered: In this case, the weight of $\sum_{j \in \mathcal{S}_m} \boldsymbol{T}_j \boldsymbol{g}$ at the ith node is large. In each iteration, we avoid selecting sensors whose localized operators are overlapped with those of already selected sensors since the weight for $T_y g(i)$ becomes 0 when $\sum_{j \in \mathcal{S}_m} T_j g(i) \geq C$. In this paper, we use $C = \frac{1}{|\mathcal{V}|} \sum_{i \in \mathcal{V}} \sum_{j \in \mathcal{S}_m} |T_j g(i)|$, which is experimentally determined.

If the kernel $\widehat{g}(\lambda)$ is a polynomial function, we can calculate (6) without eigendecomposition of graph Laplacian matrix. Let $\mathbf{T} = [\boldsymbol{T}_0 \boldsymbol{g}\ \boldsymbol{T}_1 \boldsymbol{g} \ldots\ \boldsymbol{T}_{N-1} \boldsymbol{g}]$. This can be written as $\mathbf{T} = \sqrt{N} \mathbf{U} \widehat{g}(\boldsymbol{\Lambda}) \mathbf{U}^\dagger$. When $\widehat{g}(\lambda)$ is a polynomial function, it is rewritten as $\mathbf{T} = \sqrt{N} \widehat{g}(\mathbf{L})$. Therefore, localization operators can be obtained without the eigenvectors themselves. As a result, if the original kernel $\widehat{g}(\lambda)$ is a polynomial or the Chebyshev polynomial approximation is applied to $\widehat{g}(\lambda)$, the eigendecomposition is not required for the proposed node selection.

3.2. Relationship With Graph Sampling Method

The method in [11] selects nodes in a graph[1] to maximize a cut-off graph frequency ω where the cut-off frequency associated with the subset $\mathcal{S}$ is a bound on the maximum frequency of a graph signal

[1] The method in [7] is a general method to select appropriate nodes in a graph. It is not specifically designed to select sensors.

Table 1. Computational Complexities of Graph Signal Processing Based Approaches

	EV [11]	SVD [12]	SB [13]	Proposed Method w/ CPA
Eigen-pair or operator computations	$O(k\|\mathcal{E}\|FT_2)$	$O((\|\mathcal{E}\|F+CF^3)T_1)$	$O((\|\mathcal{E}\|F+CF^3)T_1)$	$O((\|\mathcal{E}\|+N)p)$
Sampling set search	$O(NF)$	$O(NF^4)$	$O(NF^3)$	$O(N^2F)$

that can be perfectly recovered from the samples on the subset $\mathcal{S}$. It adds the mth additional node

$$y^* = \arg\max_y [(\widetilde{u}_0^m(y))^2], \tag{7}$$

where $\widetilde{\boldsymbol{u}}_0^m$ is the eigenvector of $(\boldsymbol{\mathcal{L}}^k)_{\mathcal{S}_m^c}$ associated with the minimum eigenvalue $\widetilde{\lambda}_0$, and $k \in \mathbb{Z}^+$ is a parameter. The following proposition clarifies the relationship between the method and the localization operator.

Proposition 1. *The sampling node selection proposed in [11] is identical to the case of selecting the node y in the mth iteration which maximizes $\|\boldsymbol{T}_i\boldsymbol{g}\|_2$ if eigenvectors (transform basis) $\boldsymbol{u}_l$ are selected as the ones of subsampled normalized graph Laplacian matrix $(\boldsymbol{\mathcal{L}}^k)_{\mathcal{S}_m^c}$ with the following kernel $\widehat{g}(\lambda)$:*

$$\widehat{g}(\lambda) = \begin{cases} 1 & \text{if } \lambda = 0 \\ 0 & \text{otherwise.} \end{cases} \tag{8}$$

Proof. If the kernel $\widehat{g}(\lambda)$ in (8) is used, $\|\boldsymbol{T}_i\boldsymbol{g}\|_2$ with eigenvectors (transform basis) $\boldsymbol{v}_l$ becomes

$$\begin{aligned} \|\boldsymbol{T}_i\boldsymbol{g}\|_2 &= \sqrt{\textstyle\sum_{n=0}^{N-1}\left(\sqrt{N}v_0^*(i)v_0(n)\right)^2} \\ &= |v_0^*(i)|\sqrt{N\textstyle\sum_{n=0}^{N-1}v_0(n)^2}. \end{aligned} \tag{9}$$

Since $\sqrt{N\sum_{n=0}^{N-1}v_0(n)^2}$ is a constant that does not depend on i,

$$\max_y \ \|\boldsymbol{T}_y\boldsymbol{g}\|_2 = \max_y |v_0^*(y)| = \max_y (v_0(y))^2. \tag{10}$$

Hence, when $\boldsymbol{v}_l$ is the eigenvector of $(\boldsymbol{\mathcal{L}}^k)_{\mathcal{S}_m^c}$, i.e., $\boldsymbol{v}_l = \widetilde{\boldsymbol{u}}_l^m$, the selection of node which maximizes $\|\boldsymbol{T}_i\boldsymbol{g}\|_2$ shows the same results as the method in [11]. □

Therefore, node selection in graph sampling theory [11] can also be interpreted as node selection based on a localization operator defined by an ideal kernel. Although the sampling approach in [11] could also be performed without eigendecomposition by treating $\widehat{g}(\lambda)$ in (8) a localization operator, approximating the filter in (8) would require a high degree polynomial and would not be practical.

4. EXPERIMENTAL RESULTS

The proposed sensor selection is compared with the graph sampling theory based criteria [6, 11–13], the entropy based criterion [15, 16] and the MI based criterion [17] through numerical experiments. For the graph sampling theory-based criteria, we denote the selection method in [11], [12] and [13] as EV (referred to as eigenvector), SVD (singular value decomposition) and SB (standard basis), respectively. We assume that all the original signals used in these experiments have a Gaussian distribution as in (3). Although the kernel in (3) can be arbitrarily chosen, all experiments use the following stationary kernel[2]: $\mathcal{K}(i,j) = \exp\left(-\|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2/\theta^2\right)$, where $\boldsymbol{x}_i \in \mathbb{R}^2$ is the coordinate of the ith node v_i and θ is a parameter. $\theta = 1$ is used for all experiments. We used that oracle on the statistics of graph signals without training from the data. Moreover, the data was also generated according to a GMRF with this kernel, i.e., the graph used for sensor selection is based on the same model used to generate the data. Therefore, the approximate graph in the experiment is expected to show good performance as well as a data-driven graph. In a real situation, the kernel $\mathcal{K}(\cdot)$ can be estimated from prior information or observed data [23].[3] All experiments were performed in Matlab R2015b, running on a PC with Intel Xeon E5 3 GHz CPU and 64 GB RAM. Matlab toolbox for submodular function optimization [24, 25] is used for implementations of the entropy and MI criteria.

4.1. Execution Time and Complexity

Firstly, we compare the execution time for choosing various number of locations from randomly generated locations. $|\mathcal{S}| = |\mathcal{V}|/10$ sensors are selected with various number of possible locations $\mathcal{V}$. We use $k = 8$ for EV. The proposed method uses the kernel $\widehat{g}(\lambda) = \exp(-10\lambda/\lambda_{\max})$. Figure 1 shows the execution time comparison plotted against $|\mathcal{V}|$. The proposed methods are very fast regardless of the number of possible locations even when we did not use Chebyshev polynomial approximation. Since the proposed method with Chebyshev polynomial approximation does not need eigendecomposition of the graph Laplacian matrix, it is very fast compared to the methods based on the graph sampling theory.

Table 1 compares the computational complexity among graph signal processing based methods [7, 11–13], where T_1 is the average number of iterations required for convergence of a single eigen-pair, T_2 is the number of iteration of convergence for first F eigen-pairs, k provides a trade-off between performance and complexity, C is a constant and p is the approximation order of Chebyshev polynomial approximation. We follow the notation in [11]. The calculation of the localization operator in the proposed method includes complexity for performing Chebyshev polynomial approximation and filtering [21]. It can be seen that the calculation of the localization operator shows much lower complexity than the complexities for calculating the eigen-pairs in the other approaches. Although the proposed method has higher complexity than EV in the sampling set search in Table 1, its total execution time is usually lower than EV and the other conventional approaches, which is experimentally validated in Fig. 1. The proposed method without Chebyshev polynomial approximation needs eigendecomposition for computing the operators. Its computational complexity is usually $O(N^3)$. However, it is still faster than the conventional approaches. This might be because the search algorithm of the proposed method is faster than those of the other approaches.

[2] This is one of covariance functions for the GP model used in [17].

[3] Actually the estimation accuracy affects the selected sensor positions and the reconstruction qualities of graph signals. We are studying the effect of the estimation accuracy and have a room to further improve the performance of our method.

Table 2. Performance Comparison (Average of 400 Tested Signals): SNR [dB]

$\lvert\mathcal{S}\rvert$	Entropy [15, 16]	MI [17]	Graph sampling theory [6]			Prop. w/ exponential kernel
			EV [11]	SVD [12]	SB [13]	
20	1.73	4.96	5.31	5.35	5.13	4.99
40	3.09	6.59	6.81	6.88	6.72	6.71
60	3.62	7.69	7.96	8.07	7.84	7.83
80	4.13	8.67	8.92	8.99	8.77	8.81
100	4.78	9.52	9.68	9.71	9.52	9.49

Table 3. Difference Between Exact and Approximated Method (Average of 50 Tested Graphs): Number of Sensors That are Selected by Exact Method and Not Selected by Approximated Method

Approximation order	1	2	3	4	5	6	7	8	9	10
# of the different sensors	26.53	22.30	14.80	7.97	3.87	1.47	0.60	0.20	0.03	0.00

4.2. Prediction Error Comparison

We predict graph signal values on the unobserved locations and compare the reconstruction errors. Tested signals are randomly generated according to the GP model and are corrupted by additive white Gaussian noise with $\sigma = 0.1$. An example of the possible locations, the created graph and the input signal are shown in Figs. 2 (a) and (b). We use $k = 6$ for EV. We select the set of nodes $\mathcal{S}$ from 500 randomly generated locations, set the signals on $\mathcal{S}^c$ to zero and reconstruct the original signals only from the signals on $\mathcal{S}$. The estimated signal is represented as $\boldsymbol{f} = \mathbf{U}_{\mathcal{V}\mathcal{R}}\mathbf{U}_{\mathcal{S}\mathcal{R}}^{-1}\boldsymbol{f}_{\mathcal{S}}$ where $\mathcal{R}$ is the set of Laplacian eigenvalues less than λ_{F-1} (for SVD and SB) or the estimated cut-off frequency $\Omega_k(\mathcal{S})$ with $k = 6$ (for the other methods).

The average SNRs after 400 independent runs between the predicted signal and the original signal are shown in Table 2. Signals reconstructed from 50 samples are shown in Fig. 2. It can be seen that the proposed method shows better performance than the entropy based criteria and indicates the comparable performance to the other approaches. Note that the proposed method with the 10th order Chebyshev polynomial approximation shows the same SNRs that without approximation in almost all cases. The reconstruction method used in this experiment is specifically designed for graph sampling theory-based approaches. The optimal reconstruction method for localized operators would improve the performance of the proposed method.

4.3. Effect of Chebyshev Polynomial Approximation

The performance of the proposed method with Chebyshev polynomial approximation is illustrated by Table 3, which shows the number of sensors which is selected by the exact method (with full eigendecomposition) but not selected by the proposed method with Chebyshev polynomial approximation. We select 100 sensor locations from 500 possible locations. The kernel is $\widehat{g}(\lambda) = \exp(-10\lambda/\lambda_{\max})$. They select almost the same sensors when approximation order is larger than 8. It clearly depends on the smoothness of the kernel since sinusoidal waves are used for approximation.

5. CONCLUSION

We proposed a sensor placement method based on the localization operator for graph signals and described the connection between the sampling theory and graph localization operators. The proposed approach is very fast compared to the approaches with graph sampling, since it does not need the eigendecomposition of graph Laplacian matrix. Furthermore, its prediction performance is comparable to the existing approaches. Although the experiments assume that signals have Gaussian distributions for a comparison purpose, our methods could also be effective for non-Gaussian cases. It would be expected that selecting $\widehat{g}(\lambda)$ according to the properties of sensors or the reconstruction (interpolation) process from observed values leads to high performance gain.

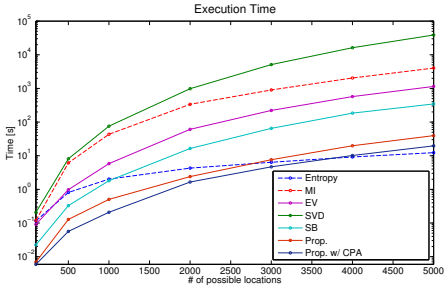

Fig. 1. Execution time comparison. Note that the vertical axis is a logarithmic scale. CPA is referred to as Chebyshev polynomial approximation. The approximation order is $p = 8$.

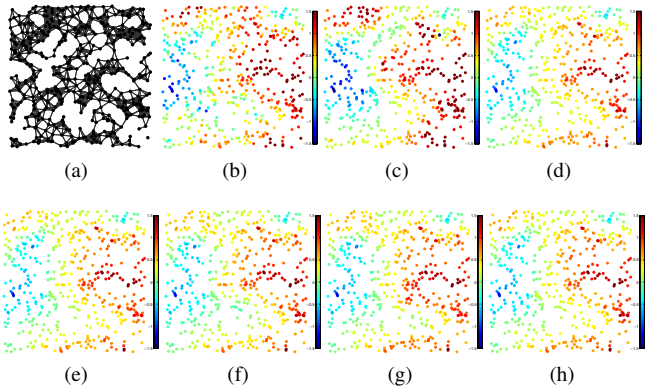

Fig. 2. Reconstructed signals from 50 selected locations. (a) Possible locations (black circle) and connections of these locations (black lines). (b) Input signal. (c) Entropy (SNR $= -1.08$ dB). (d) MI (7.30 dB). (e) EV (7.29 dB). (f) SVD (7.28 dB). (g) SB (7.05 dB). (h) Proposed method with exponential kernel (7.35 dB).

6. REFERENCES

[1] A. Krause, C. Guestrin, A. Gupta, and J. Kleinberg, "Near-optimal sensor placements: Maximizing information while minimizing communication cost," in *Proc. ACM/IEEE IPSN*, 2006, pp. 2–10.

[2] M. S. Uddin, A. Kuh, A. Kavcic, and T. Tanaka, "Nested performance bounds and approximate solutions for the sensor placement problem," *APSIPA Transactions on Signal and Information Processing*, vol. 3, no. e4, 2014.

[3] H. H. Gonzalez-Banos and J. Latombe, "A randomized art-gallery algorithm for sensor placement," in *Proc. SoCG'01*, 2001, pp. 232–240.

[4] A. Singh, A. Krause, C. Guestrin, W. J. Kaiser, and M. A. Batalin, "Efficient planning of informative paths for multiple robots," in *Proc. IJCAI*, vol. 7, 2007, pp. 2204–2211.

[5] A. Nordio, A. Tarable, F. Dabbene, and R. Tempo, "Sensor selection and precoding strategies for wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 63, no. 16, pp. 4411–4421, 2015.

[6] A. Sakiyama, Y. Tanaka, T. Tanaka, and A. Ortega, "Efficient sensor position selection using graph signal sampling theory," in *ICASSP'16*, 2016, pp. 6225–6229.

[7] A. Anis, A. Gadde, and A. Ortega, "Towards a sampling theorem for signals on arbitrary graphs," in *Proc. ICASSP'14*, 2014, pp. 3892–3896.

[8] S. K. Narang, A. Gadde, and A. Ortega, "Signal processing techniques for interpolation in graph structured data," in *Proc. ICASSP'13*, 2013, pp. 5445–5449.

[9] A. Gadde, A. Anis, and A. Ortega, "Active semi-supervised learning using sampling theory for graph signals," in *Proc. KDD'14*, 2014, pp. 492–501.

[10] A. Gadde and A. Ortega, "A probabilistic interpretation of sampling theory of graph signals," in *Proc. ICASSP'15*, 2015, pp. 3257–3261.

[11] 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.

[12] 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.

[13] I. Shomorony and A. S. Avestimehr, "Sampling large data on graphs," in *Proc. GlobalSIP'14*, 2014, pp. 933–936.

[14] A. G. Marques, S. Segarra, G. Leus, and A. Ribeiro, "Sampling of graph signals with successive local aggregations," *IEEE Trans. Signal Process.*, to appear, 2016.

[15] N. A. C. Cressie, *Statistics for Spatial Data*. Wiley, 1991.

[16] M. C. Shewry and H. P. Wynn, "Maximum entropy sampling," *Journal of Applied Statistics*, vol. 14, no. 2, pp. 165–170, 1987.

[17] A. Krause, A. Singh, and C. Guestrin, "Near-optimal sensor placements in Gaussian processes: Theory, efficient algorithms and empirical studies," *The Journal of Machine Learning Research*, vol. 9, pp. 235–284, 2008.

[18] W. F. Caselton and J. V. Zidek, "Optimal monitoring network designs," *Statistics and Probability Letters*, vol. 2, no. 4, pp. 223–227, 1984.

[19] N. Perraudin, B. Ricaud, D. I. Shuman, and P. Vandergheynst, "Global and local uncertainty principles for signals on graphs," *arXiv preprint arXiv:1603.03030*, 2016.

[20] F. R. K. Chung, *Spectral Graph Theory (CBMS Regional Conference Series in Mathematics, No. 92)*. American Mathematical Society, 1997.

[21] D. K. Hammond, P. Vandergheynst, and R. Gribonval, "Wavelets on graphs via spectral graph theory," *Applied and Computational Harmonic Analysis*, vol. 30, no. 2, pp. 129–150, 2011. [Online]. Available: http://wiki.epfl.ch/sgwt

[22] A. Deshpande, C. Guestrin, S. Madden, J. Hellerstein, and W. Hong, "Model-driven data acquisition in sensor networks," in *Proc. VLDB*, 2004, pp. 588–599.

[23] D. J. Nott and W. T. M. Dunsmuir, "Estimation of nonstationary spatial covariance structure." *Biometrika*, vol. 89, no. 4, pp. 819–829, 2002.

[24] A. Krause, "SFO: A toolbox for submodular function optimization," *Journal of Machine Learning Research*, 2010. [Online]. Available: http://las.ethz.ch/sfo/

[25] A. Krause and C. Guestrin, "Near-optimal observation selection using submodular functions," in *Proc. AAAI'07*, vol. 7, 2007, pp. 1650–1654.