COARSENING GRAPH SIGNAL WITH SPECTRAL INVARIANCE

Pengfei Liu[†‡], *Xiaohan Wang*[†], *and Yuantao Gu*[†*]

[†]State Key Laboratory on Microwave and Digital Communications
Tsinghua National Laboratory for Information Science and Technology
Department of Electronic Engineering, Tsinghua University, Beijing 100084, CHINA
[‡]National Defence Information Academy, Wuhan 430010, CHINA

ABSTRACT

Signal processing on graphs is an emerging field that attracts increasing attention. For applications such as multiscale transforms on graphs, it is often necessary to get a coarsened version of graph signal with its underlying graph. However, most of the existing methods use only topology information but no property of graph signals to complete the process. In this paper, we propose a novel graph signal coarsening method with spectral invariance, which means both the spectrum of the graph and the spectrum of the graph signal are approximately kept invariant. The problem is formulated into an optimization problem and is solved by projected subgradient method. Experiment results verify the effectiveness of the coarsening method.

Index Terms— Signal processing on graphs, graph signal, coarsening, spectral invariance, spectral graph theory.

1. INTRODUCTION

Signal processing on graphs [1] is an emerging field of data analysis in irregular domains. For an undirected graph $\mathcal{G}(\mathcal{V}, \mathcal{E}, \mathbf{W})$, which consists of a set of nodes $\mathcal{V}$ with $|\mathcal{V}| = N$, a set of undirected edges $\mathcal{E}$, and an edge weight matrix $\mathbf{W}$, a graph signal is a function $f : \mathcal{V} \to \mathbb{R}$ where each vertex has one real number assigned to it. A graph signal can also be represented as $(\mathcal{G}, \mathbf{s} \in \mathbb{R}^N)$. Considering that examples of graph signals can be found in many engineering and science fields, signal processing on graphs may play an important role in solving these problems in near future.

Since real-world networks often have a large amount of vertices and complex structures, real graph signals are usually in high dimensions. To understand the graph signal at different scales, graph signal coarsening is necessary to reduce the dimensionality and to present a direct visualization.

The coarsening of a graph signal $(\mathcal{G}_0, \mathbf{s}_0)$ is to obtain a graph signal $(\mathcal{G}, \mathbf{s})$ with less vertices than the original one. The coarsening processing should preserve properties of the original graph such as community structure, dominant spectra, connectivity, or other topological properties and signal's properties. Coarsening of a graph signal is seldom discussed in the literature, which is the motivation of this work. Graph signal coarsening is closely related to graph coarsening and graph downsampling, which provide some approaches towards solving the problem of coarsening graph signals.

Graph coarsening is to get a coarser version of the original graph. Greedy algorithms [2, 3], random walk based methods [4], relaxation based algorithm [5], and spectral methods [6] can be used to provide a solution. We refer readers to [5, 7] for more detailed review of the literature. However, graph signals are not taken into consideration in the graph coarsening problem.

Graph downsampling is to keep only a subset of the original vertex set while removing others, and assign edges and weights to the new set of vertices. For bipartite graphs, one subset of the bipartition can be kept [8, 9]. For other types of graphs, spectral methods [8, 10, 11], optimizing max-cut [12] or kron reduction [13] can be used. You may refer to [14] for a technical discussion on graph sampling. As to signal on a downsampled graph, its amplitude on the preserved vertex set is simply kept.

Similar to classical signal processing, there are equivalent representations for a graph signal in vertex domain and in graph spectral domain. The representation of a graph signal in spectral domain contains the information of both topology and signal, which is really difficult to separate. Moreover, spectral invariance may preserve more essential information of the original graph signal at the same time of dimensionality reduction.

In this work, a graph signal coarsening method with spectral invariance is proposed. The spectral domain representation of an n-vertex coarsened graph signal is identical with the maximal n spectral coefficients approximation of the N-vertex original graph signal, where n is much smaller than N. Besides, the coarsened graph is topologically similar with the original graph signal, which is consistent with the intuition. Consequently, the coarsening problem is transformed into an optimization problem and then solved iteratively by projected subgradient approach. Finally, experiment results demonstrate the effectiveness of the proposed method.

2. THE PROPOSED GRAPH SIGNAL COARSENING METHOD

The graph Laplacian is extensively used in the filed of signal processing on graphs to describe the graph spectral domain [1, 15]. A graph Laplacian [16] is

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

where $\mathbf{D}$ is a diagonal matrix composed of the row sums of $\mathbf{W}$. Because the Laplacian is a real symmetric matrix, all the eigenvalues are nonnegative. For a connected graph, one and only one eigenvalue is zero. The eigenvalues and the corresponding eigenvectors of $\mathbf{L}$ are regarded as frequencies and signal basis. The expansion coefficients of a signal in terms of the eigenvectors are called the frequency components.

The proposed coarsening method will be explained in detail in the following subsections. First, a linear mapping is used to re-

This work was partially supported by the National Program on Key Basic Research Project (973 Program 2013CB329201) and the National Natural Science Foundation of China (NSFC 61371137). The corresponding author of this paper is Yuantao Gu (gyt@tsinghua.edu.cn).

978-1-4799-2893-4/14/$31.00 ©2014 IEEE

duce the signal dimensionality, with spectral invariance. Then the coarsening problem is transformed into a constrained optimization problem, which is consequently solved using projected subgradient method.

2.1. Dimensionality reduction with spectral invariance

Suppose the original signal on an N-vertex connected graph is denoted by $(\mathcal{G}_0, \mathbf{s}_0)$, where $\mathbf{s}_0 \in \mathbb{R}^N$. The eigenvalue decomposition of the Laplacian of $\mathcal{G}_0$ is

$$\mathbf{L}_0 = \mathbf{U}_0 \mathbf{\Lambda}_0 \mathbf{U}_0^{\mathrm{T}},$$

with the eigenvalues $0 = \lambda_{0,1} < \lambda_{0,2} \leq \cdots \leq \lambda_{0,N}$ and corresponding eigenvectors $\mathbf{u}_{0,k} (1 \leq k \leq N)$. The spectral component for each frequency $\lambda_{0,k}$ is $\mathbf{u}_{0,k}^{\mathrm{T}} \mathbf{s}_0$. To reduce the dimensionality, n eigenvalues are selected to compose a new diagonal matrix $\mathbf{\Lambda}$ for the coarsened graph signal. For simplicity, the selected eigenvalues are denoted as $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$. To keep the connectivity of the coarsened graph, the eigenvalue $\lambda_{0,1} = 0$ should be remained as λ_1. The other $n-1$ selected eigenvalues are those who have the strongest spectral components. The eigenvectors corresponding to $\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$ compose an $N \times n$ matrix $\mathbf{U}$.

The new composed $\mathbf{\Lambda}$ may be used to define our desired coarsened graph signal $(\mathcal{G}, \mathbf{s} \in \mathbb{R}^n)$, which is related to a small, n-vertex graph $\mathcal{G}$. The Laplacian of $\mathcal{G}$ can be written as

$$\mathbf{L} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^{\mathrm{T}}, \tag{1}$$

where $\mathbf{V}$ denotes the orthogonal basis of an n-dimensional coarsened signal space. Then the adjacency matrix of the coarsened graph signal $(\mathcal{G}, \mathbf{s})$ is

$$\mathbf{W} = \mathbf{I}_n \circ \mathbf{L} - \mathbf{L} = \mathbf{I}_n \circ (\mathbf{V} \mathbf{\Lambda} \mathbf{V}^{\mathrm{T}}) - \mathbf{V} \mathbf{\Lambda} \mathbf{V}^{\mathrm{T}},$$

where $\mathbf{I}_n$ denotes the $n \times n$ unit matrix, and $\mathbf{I}_n \circ \mathbf{L}$ denotes the Hadamard product or entrywise product of $\mathbf{I}_n$ and $\mathbf{L}$.

To reduce the signal dimensionality by keeping spectral invariance, the N-dimensional basis $\mathbf{U}$ of the original signal should be reduced into the n-dimensional basis $\mathbf{V}$, which is assumed to be a linear transformation of $\mathbf{U}$, i.e.,

$$\mathbf{V} = \mathbf{A} \mathbf{U}, \tag{2}$$

where $\mathbf{A}$ is an $n \times N$ transform matrix. Considering that $\mathbf{V}$ is an orthogonal matrix, (2) can be transformed into

$$\mathbf{A} \mathbf{U} \mathbf{V}^{\mathrm{T}} = \mathbf{I}_n,$$

where the $N \times n$ matrix $\mathbf{U}\mathbf{V}^{\mathrm{T}}$ is of full column rank. Consequently, $\mathbf{A}$ is the left pseudo inverse of $\mathbf{U}\mathbf{V}^{\mathrm{T}}$, i.e.,

$$\mathbf{A} = ((\mathbf{U}\mathbf{V}^{\mathrm{T}})^{\mathrm{T}} \mathbf{U}\mathbf{V}^{\mathrm{T}})^{-1} (\mathbf{U}\mathbf{V}^{\mathrm{T}})^{\mathrm{T}} = \mathbf{V}\mathbf{U}^{\mathrm{T}}. \tag{3}$$

The coarsened signal $\mathbf{s}$ may also be transformed by the linear transformation $\mathbf{A}$, i.e.,

$$\mathbf{s} = \mathbf{A} \mathbf{s}_0,$$

so that the frequency components are

$$\mathbf{V}^{\mathrm{T}} \mathbf{s} = \mathbf{V}^{\mathrm{T}} \mathbf{A} \mathbf{s}_0 = \mathbf{V}^{\mathrm{T}} \mathbf{V} \mathbf{U}^{\mathrm{T}} \mathbf{s}_0 = \mathbf{U}^{\mathrm{T}} \mathbf{s}_0.$$

It means that the frequency components are kept invariant under linear mapping $\mathbf{A}$ satisfying (3).

By the dimensionality reduction proposed above, spectral invariance is achieved. However, the orthogonal matrix $\mathbf{V}$ is still unspecified. In the following subsection, we will formulate the problem of determining $\mathbf{V}$ as a constrained optimization problem.

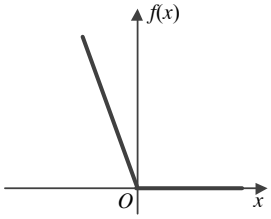


Fig. 1. An example of nonnegative-inducing penalty $f(x)$.

2.2. Problem formulation and optimization algorithm

In this subsection, the coarsening problem is formulated into an optimization problem with a matrix variable $\mathbf{V}$. As an orthogonal matrix, $\mathbf{V}$ must satisfies $\mathbf{V} \in \mathcal{O}_n$, where $\mathcal{O}_n$ is an orthogonal matrix set,

$$\mathcal{O}_n = \{ \mathbf{X} \in \mathbb{R}^{n \times n} \mid \mathbf{X}\mathbf{X}^{\mathrm{T}} = \mathbf{X}^{\mathrm{T}}\mathbf{X} = \mathbf{I}_n \}.$$

We will study the restrictions on $\mathbf{L}$ and then translate them to constrain $\mathbf{V}$. In order to ensure that $\mathbf{L}$ in (1) is actually a Laplacian of an n-vertex graph, the following two restrictions have to be satisfied,

$$\mathbf{L}\mathbf{1}_{n \times 1} = \mathbf{0}_{n \times 1} \tag{4}$$

and

$$l_{p,q} \begin{cases} \leq 0, & p \neq q; \\ \geq 0, & p = q, \end{cases} \tag{5}$$

where each entry of $\mathbf{L}$ can be expanded as a function of the entries of $\mathbf{V}$ determined by (1), as

$$l_{p,q} = \sum_{k=1}^{n} \lambda_k v_{p,k} v_{q,k}. \tag{6}$$

The constraint (4) means that the row sums of $\mathbf{L}$ are all zero, while (5) implies that $\mathbf{L}$ is entrywise nonpositive on nondiagonal positions and nonnegative on diagonal positions.

Substituting (1) and utilizing the orthogonality of $\mathbf{V}$, the constraint (4) is simplified as

$$\mathbf{\Lambda} \mathbf{V}^{\mathrm{T}} \mathbf{1}_{n \times 1} = \mathbf{0}_{n \times 1}. \tag{7}$$

By defining the constraint set $\mathcal{S}_n$ as

$$\mathcal{S}_n = \{ \mathbf{X} \in \mathbb{R}^{n \times n} \mid \mathbf{\Lambda} \mathbf{X}^{\mathrm{T}} \mathbf{1}_{n \times 1} = \mathbf{0}_{n \times 1} \},$$

the constraint (4) is equivalent to $\mathbf{V} \in \mathcal{S}_n$.

It is difficult to translate (5) into a constraint for $\mathbf{V}$. However, it can be transformed to part of our optimization objective by a particular penalty function. To ensure (5) is satisfied, we try to minimize a function $g(\mathbf{L})$ by defining

$$g(\mathbf{L}) = \sum_{p \neq q} f(-l_{p,q}) + \sum_{p=1}^{n} f(l_{p,p}),$$

where the scalar function $f(\cdot)$ is a nonnegative-inducing penalty. For example, as shown in Fig. 1, the function can be chosen as

$$f(x) = -\alpha x \cdot 1_{x<0},$$

where α is a positive constant and $1_{x<0}$ is the indicator function.

The other part of the optimization objective is to maximize the topological similarity between $\mathcal{G}$ and $\mathcal{G}_0$. Suppose $\tilde{\mathcal{G}}$ is a coarsened

Algorithm 1 Graph Signal Coarsening Algorithm

Input: Original graph signal $(\mathcal{G}_0, \mathbf{s}_0 \in \mathbb{R}^N)$, coarsened graph signal dimension n;
Output: Coarsened graph signal $(\mathcal{G}, \mathbf{s} \in \mathbb{R}^n)$;

1: Calculate the spectral domain representation of $(\mathcal{G}_0, \mathbf{s}_0)$ to obtain $\mathbf{\Lambda}$ and $\mathbf{U}$;
2: Using existed topology-based method to obtain $\tilde{\mathcal{G}}$ and $\tilde{\mathbf{L}}$, $\tilde{\mathcal{G}} = \mathcal{T}(\mathcal{G}_0)$;
3: Initialize a matrix $\mathbf{V}$;
4: **repeat**
5: Choose the step size κ by backtracking line search and conduct gradient update using (10) and (11), $\mathbf{V} = \mathbf{V} - \kappa \nabla_{\mathbf{V}} F(\mathbf{V}, t)$, $t = t - \kappa \nabla_t F(\mathbf{V}, t)$;
6: Project onto the set $\mathcal{O}_n$ using (12), $\mathbf{V} = P_{\mathcal{O}_n}(\mathbf{V})$;
7: Project onto the set $\mathcal{S}_n$ using (13), $\mathbf{V} = P_{\mathcal{S}_n}(\mathbf{V})$;
8: **until** stopping criterion is satisfied;
9: Calculate the adjacency matrix of $\mathcal{G}$ and signal $\mathbf{s} \in \mathbb{R}^n$, $\mathbf{W} = \mathbf{I}_n \circ (\mathbf{V}\mathbf{\Lambda}\mathbf{V}^{\mathrm{T}}) - \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{\mathrm{T}}$, $\mathbf{s} = \mathbf{V}\mathbf{U}^{\mathrm{T}}\mathbf{s}_0$.

graph via some existed graph coarsening method $\mathcal{T}(\cdot)$ using topology information only,

$$\tilde{\mathcal{G}} = \mathcal{T}(\mathcal{G}_0),$$

and the Laplacian of $\tilde{\mathcal{G}}$ is $\tilde{\mathbf{L}}$. Therefore, the maximization of the similarity between the topologies of $\mathcal{G}$ and $\tilde{\mathcal{G}}$ can be transformed to minimizing $\|\mathbf{L} - t\tilde{\mathbf{L}}\|_F$, where t is a positive multiple and $\|\cdot\|_F$ denotes the Frobenius norm. The introduction of multiple t means the two graphs are similar in the sense of relative edge weights.

Therefore, we are ready to formulate the following optimization problem,

$$\min_{\mathbf{V},t} F(\mathbf{V}, t), \quad \text{subject to} \quad \mathbf{V} \in \mathcal{O}_n \cap \mathcal{S}_n,$$

where the objective function is defined as

$$F(\mathbf{V}, t) = g(\mathbf{L}) + \mu \|\mathbf{L} - t\tilde{\mathbf{L}}\|_F^2, \tag{8}$$

and each entry of $\mathbf{L}$ is a function of $\mathbf{V}$, as defined in (6).

The projected subgradient method is utilized to solve the proposed optimization problem. For each iteration, the variables $\mathbf{V}$ and t are firstly updated by subgradient descent. Then the projection step can be modified as alternating projecting $\mathbf{V}$ onto $\mathcal{O}_n$ and $\mathcal{S}_n$, respectively. The detailed expressions of the subgradients are derived as (10) and (11) in Appendix 5.1. The projection operations onto sets $\mathcal{O}_n$ and $\mathcal{S}_n$ are given as (12) and (13) in Appendix 5.2.

In summery, the proposed method is described in Algorithm 1.

3. EXPERIMENT RESULTS

Numerical experiments are conducted to verify the proposed graph signal coarsening method.[1] A 150-vertex graph signal $(\mathcal{G}_0, \mathbf{s}_0)$ for testing is illustrated in Fig. 2. The topology is composed of three 50-vertex Erdős-Rényi (ER) random graphs [17] and several random edges between the communities. For the vertices of each community, the signals on them have similar values, with means -0.75, 0.15 and -0.05 respectively, and variances 0.05, 0.03 and 0.01 respectively. The spectral domain of $(\mathcal{G}_0, \mathbf{s}_0)$ is shown in Fig. 5(a). It can be seen that most of the spectral components have small amplitudes, which implies the graph signal can be approximated with only a small number of frequency components. The 13-term approximation of $(\mathcal{G}_0, \mathbf{s}_0)$ is shown in of Fig. 5(b).

[1] The code for these experiments is available at `http://gu.ee.tsinghua.edu.cn/publications#lp1`

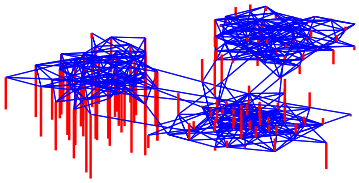


Fig. 2. The original graph signal $(\mathcal{G}_0, \mathbf{s}_0)$ in the vertex domain. The signals are on a graph consisting of three communities. The signal values are represented by the red bars.

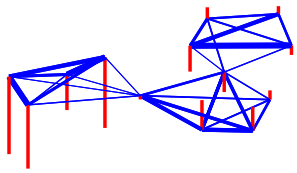


Fig. 3. The coarsened graph signal $(\mathcal{G}, \mathbf{s})$ in the vertex domain.

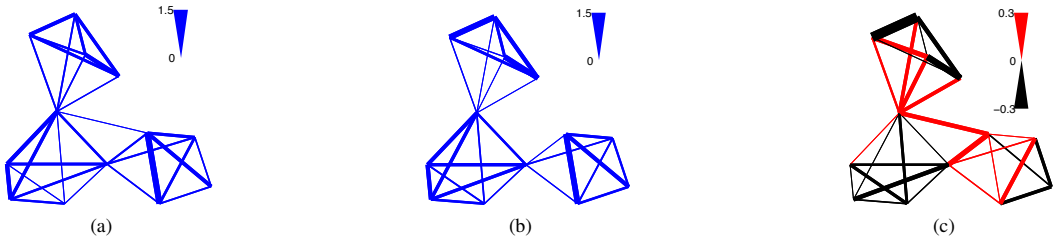
The aiming topology $\tilde{\mathcal{G}}$ shown in Fig. 4(a) is obtained by repeatedly using a graph coarsening method called Heavy Edge Matching [2], and removing edges with small weights finally. Using the proposed algorithm, the coarsened graph signal $(\mathcal{G}, \mathbf{s})$ is shown in Fig. 3 for vertex domain and in Fig. 5(c) for spectral domain. As shown in Fig. 5, the coarsened graph signal $(\mathcal{G}, \mathbf{s})$ is almost the same with the 13-term approximation of $(\mathcal{G}_0, \mathbf{s}_0)$ in the spectral domain, which achieves the spectral invariance property of the proposed coarsening method. Note that in order to achieve spectral invariance, the coarsened graph might get a few edges with very small weights. Because the resolution of the figures is limited, edges with very small weights, that is, thinner than 0.02 unit width in Fig. 4(a) and Fig. 4(b), are not shown in the figures.

Moreover, we are happy to find that the vertex domain of the coarsened graph signal is also similar with the original graph signal. As shown in Fig. 2 and Fig. 3, they have similar community structures, and $(\mathcal{G}, \mathbf{s})$ also have similar values inside each community.

The proposed method is to optimize the similarity of $\mathcal{G}$ with the aiming topology $\tilde{\mathcal{G}}$ by the constraint of spectral invariance. The difference between $\mathcal{G}$ and $\tilde{\mathcal{G}}$ is illustrated in Fig. 4. It can be seen that $t\tilde{\mathbf{W}}$ and $\mathbf{W}$ are similar, that is, the Frobenius norm of $t\tilde{\mathbf{W}} - \mathbf{W}$ is small. Note that the unit line width in Fig. 4(c) is five times that in Fig. 4(a) and Fig. 4(b) to illustrate the difference clearly.

4. CONCLUSION

In this paper, we propose a novel graph signal coarsening method with spectral invariance. Using this method, both the spectrum of the graph and the spectrum of the graph signal are approximately kept invariant. The coarsening problem is formulated into a constrained optimization problem, which can be solved using projected subgradient method. Experiment results verify the effectiveness of the proposed method.



(a) (b) (c)

Fig. 4. The topologies of (a) $\tilde{\mathcal{G}}$, (b) $\mathcal{G}$ and (c) their difference. $\tilde{\mathcal{G}}$ is the aiming topology which is a coarsened graph using topology information only, and $\mathcal{G}$ is the topology of coarsened graph signal. The widths of lines represent the weights of $t\tilde{\mathbf{W}}$ and $\mathbf{W}$, respectively. It can be seen that $t\tilde{\mathbf{W}}$ and $\mathbf{W}$ are similar, that is, the Frobenius norm of $t\tilde{\mathbf{W}} - \mathbf{W}$ is small. (c) shows the weight differences, as $t\tilde{\mathbf{W}} - \mathbf{W}$. Red lines represent positive ones and black lines represent the negative. The unit width of (c) is five times that in (a) and (b).

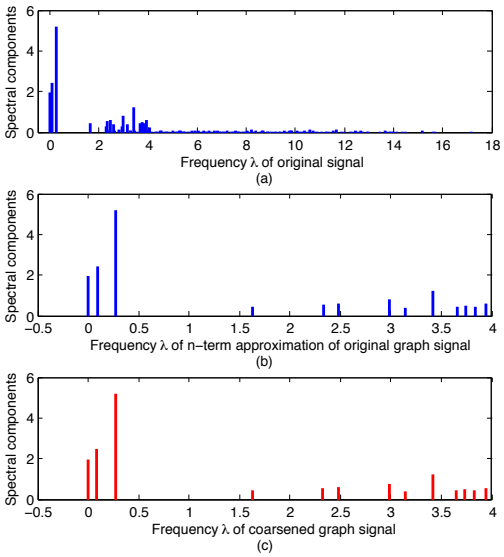


Fig. 5. The spectral domain representations of (a) the original signal, (b) the 13-term approximation of original signal and (c) the coarsened signal, respectively. The latter two are almost the same in the spectral domain.

5. APPENDIX

5.1. Subgradient calculation

According to (6), the partial derivative of the entries of $\mathbf{L}$ with respect to the variable $v_{i,j}$ is

$$\left\{\frac{\partial l_{p,q}}{\partial v_{i,j}}\right\}_{1\le p,q\le n} = \lambda_j \begin{bmatrix} & & v_{1,j} & & \\ & \mathbf{0}_{(i-1)\times(i-1)} & \vdots & \mathbf{0}_{(i-1)\times(n-i)} & \\ v_{1,j} & \cdots & v_{i-1,j} \quad 2v_{i,j} \quad v_{i+1,j} & \cdots & v_{n,j} \\ & \mathbf{0}_{(n-i)\times(i-1)} & \vdots & \mathbf{0}_{(n-i)\times(n-i)} & \\ & & v_{n,j} & & \end{bmatrix}, \quad (9)$$

with only the entries in the ith row and the ith column are nonzero.

According to (8) and (9), the partial derivative of the objective function $F(\mathbf{V},t)$ with respect to $v_{i,j}$ can be calculated as

$$\begin{aligned}\frac{\partial F(\mathbf{V},t)}{\partial v_{i,j}} =& -2\lambda_j \sum_{q=1}^{n} f'(-l_{i,q})v_{q,j} + 4\mu\lambda_j \sum_{q=1}^{n}(l_{i,q} - t\tilde{l}_{i,q})v_{q,j} \\ &+ 2\lambda_j f'(-l_{i,i})v_{i,j} + 2\lambda_j f'(l_{i,i})v_{i,j}.\end{aligned}$$

Therefore, the subgradient is

$$\begin{aligned}\nabla_{\mathbf{V}} F(\mathbf{V},t) &= \left\{\frac{\partial F(\mathbf{V},t)}{\partial v_{i,j}}\right\}_{1\le i,j\le n} \\ &= -2\left[f'(-\mathbf{L}) - 2\mu(\mathbf{L} - t\tilde{\mathbf{L}})\right]\mathbf{V}\mathbf{\Lambda} \\ &\quad + 2\left[\mathbf{I}_n \circ f'(-\mathbf{L}) + \mathbf{I}_n \circ f'(\mathbf{L})\right]\mathbf{V}\mathbf{\Lambda}, \quad (10)\end{aligned}$$

where $f'(\mathbf{L})$ denotes the matrix with entries $f'(l_{p,q})$, using the derivative or subderivative at $l_{p,q}$.

On the other hand, the gradient of $F(\mathbf{V},t)$ with respect to t can be calculated as

$$\nabla_t F(\mathbf{V},t) = 2\mu\left(t\cdot\mathrm{tr}(\tilde{\mathbf{L}}\tilde{\mathbf{L}}^{\mathrm{T}}) - \mathrm{tr}(\mathbf{L}\tilde{\mathbf{L}}^{\mathrm{T}})\right). \quad (11)$$

The variable t can also be updated by the gradient.

5.2. Projection onto two sets

As a special case of Orthogonal Procrustes Problem [18], for any matrix $\mathbf{X} \in \mathbb{R}^{n\times n}$, the projection onto $\mathcal{O}_n$ can be obtained by

$$P_{\mathcal{O}_n}(\mathbf{X}) = \arg\min_{\mathbf{Y}\in\mathcal{O}_n} \|\mathbf{X} - \mathbf{Y}\|_F = \mathbf{U}_{\mathbf{X}}\mathbf{V}_{\mathbf{X}}^{\mathrm{T}}, \quad (12)$$

where $\mathbf{U}_{\mathbf{X}}$ and $\mathbf{V}_{\mathbf{X}}^{\mathrm{T}}$ are the orthogonal matrices in the singular value decomposition of $\mathbf{X}$,

$$\mathbf{X} = \mathbf{U}_{\mathbf{X}}\mathbf{\Sigma}_{\mathbf{X}}\mathbf{V}_{\mathbf{X}}^{\mathrm{T}}.$$

Since only the first eigenvalue in $\mathbf{\Lambda}$ is zero, the equation (7) is equivalent to linear constraints for the entries of $\mathbf{V}$,

$$\sum_{i=1}^{n} v_{i,j} = 0, \quad j = 2,3,\cdots,n.$$

Then for any matrix $\mathbf{X} \in \mathbb{R}^{n\times n}$, the projection onto $\mathcal{S}_n$ can be derived as

$$P_{\mathcal{S}_n}(\mathbf{X}) = \mathbf{X} - \frac{1}{n}\mathbf{1}_{n\times n}\mathbf{X}\begin{bmatrix} 0 & \mathbf{0}_{1\times(n-1)} \\ \mathbf{0}_{(n-1)\times 1} & \mathbf{I}_{(n-1)\times(n-1)} \end{bmatrix}. \quad (13)$$

6. 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 Processing Magazine*, vol. 30, no. 3, pp. 83-98, 2013.

[2] G. Karypis and V. Kumar, "A fast and high quality multilevel scheme for partitioning irregular graphs," *SIAM Journal on Scientific Computing*, vol. 20, no. 1, pp. 359-392, Jan. 1998.

[3] I. S. Dhillon, Y. Guan, and B. Kulis, "Weighted graph cuts without eigenvectors a multilevel approach," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 11, pp. 1944-1957, 2007.

[4] S. Lafon and A. B. Lee, "Diffusion maps and coarse-graining: a unified framework for dimensionality reduction, graph partitioning, and data set parameterization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 9, pp. 1393-1403, 2006.

[5] D. Ron, I. Safro, and A. Brandt, "Relaxation-based coarsening and multiscale graph organization," *Multiscale Modeling & Simulation*, vol. 9, no. 1, pp. 407-423, Jan. 2011.

[6] D. Gfeller and P. De Los Rios, "Spectral coarse graining of complex networks," *Physical Review Letters*, vol. 99, no. 3, pp. 038701, Jul. 2007.

[7] C. Walshaw, "Graph partitioning archive," 2002. [Online]. Available: http://staffweb.cms.gre.ac.uk/ wc06/partition/.

[8] S. K. Narang and A. Ortega, "Downsampling graphs using spectral theory," *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 4208-4211, 2011.

[9] S. K. Narang and A. Ortega, "Perfect reconstruction two-channel wavelet filter banks for graph structured data," *IEEE Transactions on Signal Processing*, vol. 60, no. 6, pp. 2786-2799, 2012.

[10] B. Aspvall and J. R. Gilbert, "Graph coloring using eigenvalue decomposition," *SIAM Journal on Algebraic Discrete Methods*, vol. 5, no. 4, pp. 526-538, 1984.

[11] S. T. Barnard and H. D. Simon, "Fast multilevel implementation of recursive spectral bisection for partitioning unstructured problems," *Concurrency: Practice and experience*, vol. 6, no. 2, pp. 101-117, 1994.

[12] S. K. Narang and A. Ortega, "Local two-channel critically sampled filter-banks on graphs," *2010 17th IEEE International Conference on Image Processing (ICIP)*, pp. 333-336, 2010.

[13] F. Dorfler and F. Bullo, "Kron reduction of graphs with applications to electrical networks," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 60, no. 1, pp. 150-163, 2013.

[14] D. I. Shuman, M. J. Faraji, and P. Vandergheynst, "A framework for multiscale transforms on graphs," *arXiv preprint arXiv:1308.4942*, 2013.

[15] X. Zhu and M. Rabbat, "Approximating signals supported on graphs," *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3921-3924, 2012.

[16] F. K. Chung, *Spectral Graph Theory*, American Mathematical Soc, 1997.

[17] P. Erdős and A. Rényi, "On random graphs," *Publicationes Mathemticae (Debrecen)*, vol. 6, pp. 290-297, 1959.

[18] G. H. Golub, C. F. V. Loan, *Matrix Computations (Third Edition)*, The Johns Hopkins University Press, 1996.