ADMM-BASED BIPARTITE GRAPH APPROXIMATION

Aimin Jiang, Jiaan Wan*, Yibin Tang*, Beilu Ni*, and Yanping Zhu†*

*College of Internet of Things Engineering, Hohai University, Changzhou, Jiangsu, China
†School of Information Science and Engineering, Changzhou University, Changzhou, Jiangsu, China

ABSTRACT

Because the spectrum folding phenomenon affects the down-sampling of graph signals, both critically sampled and over-sampled graph filter banks with down-sampling operations can only be applied to bipartite graphs. However, general graph signals may not reside on bipartite graph structures. In this paper, we present a novel bipartite graph approximation algorithm, which aims to find a bipartite graph sufficiently close to the original graph. To tackle this problem, we first show that, if the non-negativity constraint of adjacency matrices is removed, closed-form solutions can be readily obtained by the eigenvalue decomposition. Based on this fact, an alternating direction method of multipliers (ADMM) is further developed to achieve a real adjacency matrix. Experimental results show that the proposed algorithm outperforms the other proposals in terms of approximation accuracy.

Index Terms— Alternating direction method of multipliers (ADMM), bipartite graph, eigenvalue decomposition.

1. INTRODUCTION

Graph signal processing [1]-[3] has found many applications in the fields of science and engineering, such as sensor networks [4, 5], social networks [6, 7], traffic networks [8], and so on. An undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is composed by a set of vertices, among which there are edges representing the similarity of vertices. Generally, each edge is assigned a nonnegative weight $a_{i,j}$, $i, j \in \mathcal{V}$ and $(i, j) \in \mathcal{E}$, to indicate the reliability of the corresponding link. Collecting all the weights, one can obtain the adjacency matrix $\mathbf{A} = [a_{i,j}] \in \mathbb{R}^{N\times N}$, where N denotes the number of vertices, i.e., $N = |\mathcal{V}|$. Another important quantity to describe graph structures is the combinatorial graph Laplacian matrix defined by $\mathbf{L} = \mathbf{D} - \mathbf{A}$, where $\mathbf{D}$ represents the diagonal degree matrix with its (i, i)-th diagonal element equal to $\sum_j a_{i,j}$. The normalized Laplacian can be further computed by $\mathcal{L} = \mathbf{D}^{-1/2}\mathbf{L}\mathbf{D}^{-1/2}$. In this paper, we focus on the optimization of adjacency matrix $\mathbf{A}$. But the proposed algorithm can also be applied to handle $\mathcal{L}$ with appropriate modification.

Multi-channel signal processing over graphs have drawn much attention in the last decade. Recent advances show that a critically sampled two-channel wavelet filter-bank design leads to compact representation of graph signals [9, 10]. Similar to the aliasing effect in regular signal processing, the spectrum folding is also observed in downsampling of graph signals. It has been proven that, in order to achieve both critical sampling and perfect reconstruction, the underlying graph must be bipartite [10]. However, general graphs do not satisfy this requirement. To deal with this challenge, the original non-bipartite graph has to be first decomposed to a set of bipartite subgraphs. Each bipartite subgraph has all the vertices and a portion of edges of the original graph. Then, the filter-banks can be applied to each subgraph, leading to multi-dimensional transform. Harry's algorithm [11] based on graph coloring is widely adopted to conduct the bipartite subgraph decomposition. It assumes that the decomposition contains $\log_2 K$ subgraphs if the graph is a K-colorable graph. Then, in each subgraph, all the colored vertices are separated into two disjoint sets to represent two colors and only the edges connecting vertices with different colors could be permitted in each bipartite subgraph. However, its performance highly relies on graph coloring, which is NP-hard [12].

Another bipartite subgraph decomposition algorithm was proposed in [13], which aims to improve energy compaction in bipartite subgraphs. To this end, two criteria are adopted in practical optimization. First, to preserver the spectral characteristics of the original graph, the Kullback–Leibler divergence metric is minimized. Second, the multiplicity of eigenvalue at frequency 1 for the normalized graph Laplacian is minimized to avoid accumulating energies at frequency 1. Due to the nonconvexity of the resulting problem, the linear approximation is applied to simplify the objective function. Another local heuristic approach is also developed in [14] for fast implementation and accurate rank optimization.

Aforementioned algorithms aims to decompose the original non-bipartite graph into a set of edge-disjoint bipartite subgraphs, on which filter-banks can be applied. Another strategy to implement multi-channel filtering on an arbitrary graph is by graph oversampling [15, 16], which appends some vertices and edges to the original graph. In [16], an efficient

This work was supported in part by the National Nature Science Foundation of China under grants 61471157, 61772090, and 61801055, and the Fundamental Research Funds for the Central Universities under grant 26120182018B23014.

graph oversampling approach is proposed for an arbitrary K-colorable graph. The appended vertices are just above some vertices in the original graph, so that they have the same values. These twin-vertices can be connected by vertical edges. However, this approach also depends on graph coloring. Furthermore, in practice, the number of vertices and edges appended could be too large, which increases computational cost in the subsequent filtering.

In this paper, we consider a more straightforward way to implement two-channel graph signal filtering. Specifically, one first constructs a bipartite graph, which has the minimum approximation error to the original non-bipartite graph. Then, two-channel wavelet filter-banks can be applied on the bipartite graph obtained. Different to two strategies mentioned before, the proposed algorithm is independent on graph coloring, but the resulting bipartite graph does not need to keep all the edges in the original graph, especially those with small weights. This is reasonable when some links between vertices are unreliable. The rest part of the paper is organized as follows. In section II, we introduce the bipartite graph approximation problem considered in this paper. Then, an alternating direction method of multipliers (ADMM) is further developed in Section III to solve the approximation problem. The performance of the proposed algorithm is evaluated in Section IV. Conclusions are summarized in section V.

2. PROBLEM FORMULATION

Let $\mathbf{A}$ be the adjacency matrix of an arbitrary graph. In general, its eigenvalues are unbounded and can be positive or negative. But, when considering a bipartite graph, its eigenvalues are distributed in a symmetric way around zero [17]. Based on this observation, the bipartite graph approximation problem considered in this paper is cast as

$$\min_{\Lambda,\mathbf{U}} \quad \left\|\mathbf{A}-\mathbf{U}\Lambda\mathbf{U}^T\right\|_F^2 \tag{1a}$$

$$\text{s.t.} \quad \mathbf{U}^T\mathbf{U}=\mathbf{U}\mathbf{U}^T=\mathbf{I} \tag{1b}$$

$$\left[\mathbf{U}\Lambda\mathbf{U}^T\right]_{i,j}\geq 0, \quad i,j=1,\dots N, \tag{1c}$$

$$\left[\mathbf{U}\Lambda\mathbf{U}^T\right]_{i,i}\geq 0, \quad i=1,\dots N, \tag{1d}$$

where $\mathbf{U}$ is the eigen vector matrix of the adjacency matrix for a bipartite graph, and the diagonal matrix Λ contains all the eigenvalues of the adjacency matrix, i.e.,

$$\text{diag}\left(\Lambda\right)=\lambda=\left[\lambda_1\ \lambda_2\ \dots\ \lambda_N\right]^T. \tag{2}$$

To achieve a bipartite graph, eigenvalues $\{\lambda_i\}$ should satisfy $\lambda_i=-\lambda_{N-i+1}$ for $i=1,\dots,N$. This property will be taken into account when optimizing eigenvalues $\{\lambda_i\}$. To proceed, in the latter discussion we first ignore constraints (1c) and (1d). When developing the ADMM algorithm in the next section, we shall take them into account again to achieve a real bipartite graph.

To tackle the simplified (1), we reformulate the objective function as

$$\begin{aligned}\left\|\mathbf{A}-\mathbf{U}\Lambda\mathbf{U}^T\right\|_F^2 &= \text{Tr}\left\{\mathbf{A}^2\right\}+\text{Tr}\left\{\Lambda^2\right\}\\ &\quad -2\text{Tr}\left\{\mathbf{A}^T\mathbf{U}\Lambda\mathbf{U}^T\right\}.\end{aligned} \tag{3}$$

Let the eigenvalue decomposition of $\mathbf{A}$ be $\mathbf{A}=\mathbf{V}\Delta\mathbf{V}^T$ where $\text{diag}\left(\Delta\right)=\eta=\left[\eta_1\ \eta_2\ \dots\ \eta_N\right]^T$. Without the loss of generalization, we assume that $\{\eta_i\}$ are arranged in a non-ascending order. Then, (3) is equivalently written by

$$\begin{aligned}\left\|\mathbf{A}-\mathbf{U}\Lambda\mathbf{U}^T\right\|_F^2 &= \text{Tr}\left\{\mathbf{A}^2\right\}+\text{Tr}\left\{\Lambda^2\right\}-2\text{Tr}\left\{\Delta\mathbf{P}\Lambda\mathbf{P}^T\right\}\\ &= \text{Tr}\left\{\mathbf{A}^2\right\}+\text{Tr}\left\{\Lambda^2\right\}-2\eta^T\left(\mathbf{P}\circ\mathbf{P}\right)\lambda\end{aligned} \tag{4}$$

where $\circ$ denotes the elementwise product of two matrices, and

$$\mathbf{P}=\mathbf{V}^T\mathbf{U}. \tag{5}$$

Now, the simplified (1) can be reformulated as

$$\min_{\lambda,\mathbf{P}} \quad \lambda^T\lambda-2\eta^T\left(\mathbf{P}\circ\mathbf{P}\right)\lambda \tag{6a}$$

$$\text{s.t.} \quad \mathbf{P}^T\mathbf{P}=\mathbf{P}\mathbf{P}^T=\mathbf{I}. \tag{6b}$$

Once the optimal solution $(\lambda^*,\mathbf{P}^*)$ to the above problem is obtained, the optimal solution to the simplified (1) can be recovered by $\mathbf{B}^*=\mathbf{U}^*\text{diag}\left\{\lambda^*\right\}\mathbf{U}^{*T}$ where $\mathbf{U}^*=\mathbf{V}\mathbf{P}^*$.

To achieve the optimal solution to (6), we introduce the following theorem. Since eigenvalues in λ^* are symmetric around zero, only half of eigenvalues are presented below.

Theorem 1. *The optimal solution* $(\lambda^*,\mathbf{P}^*)$ *to* (6) *is given by*

$$\lambda_i^*=\frac{1}{2}\left(\eta_i-\eta_{N-i+1}\right), \quad i=1,\dots,\lceil N/2\rceil \tag{7}$$

$$\mathbf{P}^*=\begin{bmatrix}\pm 1 & & & \\ & \pm 1 & & \\ & & \ddots & \\ & & & \pm 1\end{bmatrix} \tag{8}$$

where $\lceil x\rceil$ *denotes the integer closest to* x *towards infinity.*

Proof. In this proof, we only consider the case of an even N. But the proof is also applicable for an odd N. To proceed, we define $\mathbf{Q}=\mathbf{P}\circ\mathbf{P}$ and take it into (6). Because $\mathbf{P}$ is a orthogonormal matrix, it can be verified that $\mathbf{Q}$ is a doubly stochastic matrix, that is, $Q_{i,j}\geq 0$ for $i,j=1,\dots,N$, $\mathbf{Q}\cdot\mathbf{1}=\mathbf{1}$, and $\mathbf{Q}^T\cdot\mathbf{1}=\mathbf{1}$. Using $\mathbf{Q}$, (6) can be rewritten as

$$\min_{\lambda,\mathbf{Q}} \quad \lambda^T\lambda-2\eta^T\mathbf{Q}\lambda \tag{9a}$$

$$\text{s.t.} \quad \mathbf{Q}\cdot\mathbf{1}=\mathbf{Q}^T\cdot\mathbf{1}=\mathbf{1} \tag{9b}$$

$$Q_{i,j}\geq 0, \quad i,j=1,\dots,N. \tag{9c}$$

Note that the orthogonormal constraint (6b) is now replaced by (9b) and (9c).

Due to its symmetric distribution around zero, λ can be represented by $\lambda = [\mathbf{x} \ -\bar{\mathbf{x}}]^T$ where $\mathbf{x} = [\lambda_1 \ \lambda_2 \ \ldots \ \lambda_{N/2}]^T$ and $\bar{\mathbf{x}} = [\lambda_{N/2} \ \lambda_{N/2-1} \ \ldots \ \lambda_1]^T$. We further decompose $\mathbf{Q}$ as $\mathbf{Q} = [\mathbf{Q}_1 \ \mathbf{Q}_2]$ where $\mathbf{Q}_1, \mathbf{Q}_2 \in \mathbb{R}^{N\times N/2}$. Then, the objective function of (9) is simplified to

$$\begin{aligned}\lambda^T\lambda - 2\eta^T\mathbf{Q}\lambda &= 2\mathbf{x}^T\mathbf{x} - 2\eta^T\left(\mathbf{Q}_1 - \bar{\mathbf{Q}}_2\right)\mathbf{x} \\ &= 2\left\|\mathbf{x} - \frac{1}{2}\left(\mathbf{Q}_1 - \bar{\mathbf{Q}}_2\right)^T\eta\right\|_2^2 \\ &\quad - \frac{1}{2}\left\|\left(\mathbf{Q}_1 - \bar{\mathbf{Q}}_2\right)^T\eta\right\|_2^2\end{aligned} \tag{10}$$

where $\bar{\mathbf{Q}}_2$ is obtained by reversing the column ordering of $\mathbf{Q}_2$. Obviously, the minimizer of (10) is given by

$$\mathbf{x}^* = \frac{1}{2}\left(\mathbf{Q}_1 - \bar{\mathbf{Q}}_2\right)^T\eta. \tag{11}$$

On the other hand, given λ (or, equivalently, $\mathbf{x}$), the optimal $\mathbf{Q}^*$ can be obtained by solving

$$\max_{\mathbf{Q}} \quad \mathrm{Tr}\left\{\mathbf{Q}\lambda\eta^T\right\} \tag{12a}$$

$$\text{s.t.} \quad \mathbf{Q}\cdot\mathbf{1} = \mathbf{Q}^T\cdot\mathbf{1} = \mathbf{1} \tag{12b}$$

$$Q_{i,j} \geq 0, \quad i,j = 1,\ldots,N, \tag{12c}$$

which is a linear program. According to the Birkhoff-von Neumann Theorem, which states that the extreme points of the set of doubly stochastic matrices are permutation matrices [19], the optimal $\mathbf{Q}^*$ to (12) should be a permutation matrix. In view of this point, to achieve the minimum value of (10) or the maximum value of $\left\|\left(\mathbf{Q}_1 - \bar{\mathbf{Q}}_2\right)^T\eta\right\|_2^2$, $\mathbf{Q}_1$ and $\mathbf{Q}_2$ should be arranged as

$$\mathbf{Q}_1 = \begin{bmatrix}\mathbf{I}\\ \mathbf{0}\end{bmatrix}, \quad \mathbf{Q}_2 = \begin{bmatrix}\mathbf{0}\\ \mathbf{I}\end{bmatrix}, \tag{13}$$

which essentially means that the optimal $\mathbf{Q}^*$ to (9) is an identity matrix. Due to $Q_{i,j} = P_{i,j}^2$, (13) further implies that the optimal $\mathbf{P}^*$ is a diagonal matrix of the form given by (8). Finally, combined with (11), we obtain (7). □

The above theorem implies that the optimal $\mathbf{U}^*$ to the simplified (1) are essentially equal to $\mathbf{V}$, up to multiplying -1 on some columns of $\mathbf{V}$. But changing the sign of one eigen vector does not lead to a different solution to (1). So in the latter discussion, we assume $\mathbf{P}^* = \mathbf{I}$ and, thus, $\mathbf{U}^* = \mathbf{V}$.

3. ADMM ALGORITHM

In the previous development, we ignore constraints (1c) and (1d). Thereby, the resulting solution is not a real adjacency matrix of a bipartite graph. To fix this issue, we shall develop an ADMM algorithm in the sequel. First, (1) is rewritten as

$$\min_{\mathbf{B},\Lambda,\mathbf{U}} \quad \|\mathbf{A} - \mathbf{B}\|_F^2 \tag{14a}$$

$$\text{s.t.} \quad \mathbf{B} = \mathbf{U}\Lambda\mathbf{U}^T. \tag{14b}$$

Note that, in the above formulation, there are implicit constraints on $\mathbf{B}$, Λ and $\mathbf{U}$, that is, $B_{i,j} \geq 0$ and $B_{i,i} = 0$, $\lambda_i = -\lambda_{N-i+1}$, $\mathbf{U}^T\mathbf{U} = \mathbf{U}\mathbf{U}^T = \mathbf{I}$. The augmented Lagrangian function of (14) is given by

$$\begin{aligned}\mathcal{L}_\rho(\mathbf{B},\mathbf{U},\Lambda;\mathbf{W}) &= \|\mathbf{A} - \mathbf{B}\|_F^2 + \mathrm{Tr}\left\{\mathbf{W}^T\left(\mathbf{B} - \mathbf{U}\Lambda\mathbf{U}^T\right)\right\} \\ &\quad + \frac{\rho}{2}\left\|\mathbf{B} - \mathbf{U}\Lambda\mathbf{U}^T\right\|_F^2.\end{aligned} \tag{15}$$

The ADMM algorithm needs to alternately solve the following problems to achieve, respectively, solutions $\mathbf{B}^{(k+1)}$, $\Lambda^{(k+1)}$ and $\mathbf{U}^{(k+1)}$, and updates $\mathbf{W}^{(k)}$ and $\rho^{(k)}$ until some stopping criterion is satisfied

$$\mathbf{B}^{(k+1)} = \arg\min_{\mathbf{B}} \left\|\tilde{\mathbf{B}}^{(k)} - \mathbf{B}\right\|_F^2 \tag{16}$$

$$\left(\Lambda^{(k+1)}, \mathbf{U}^{(k+1)}\right) = \arg\min_{\Lambda,\mathbf{U}} \left\|\tilde{\mathbf{Z}}^{(k)} - \mathbf{U}\Lambda\mathbf{U}^T\right\|_F^2 \tag{17}$$

$$\mathbf{W}^{(k+1)} = \mathbf{W}^{(k)} + \rho^{(k)}\left(\mathbf{B}^{(k+1)} - \mathbf{Z}^{(k+1)}\right) \tag{18}$$

$$\rho^{(k+1)} = \eta\rho^{(k)} \tag{19}$$

where

$$\tilde{\mathbf{B}}^{(k)} = \frac{2\mathbf{A} - \mathbf{W}^{(k)} + \rho^{(k)}\mathbf{U}^{(k)}\Lambda^{(k)}\mathbf{U}^{(k)T}}{2 + \rho^{(k)}} \tag{20}$$

$$\tilde{\mathbf{Z}}^{(k)} = \frac{\mathbf{W}^{(k)}}{\rho^{(k)}} + \mathbf{B}^{(k+1)} \tag{21}$$

$$\mathbf{Z}^{(k+1)} = \mathbf{U}^{(k+1)}\Lambda^{(k+1)}\mathbf{U}^{(k+1)T}. \tag{22}$$

The optimal solution to (16) is obtained by replacing negative and diagonal elements of $\tilde{\mathbf{B}}^{(k)}$ by zeros. Problem (17) can be tackled by the approach developed in the previous section.

Note that, in the above development, we essentially assume that an edge can be appended between a pair of vertices, which are originally not connected in $\mathbf{A}$. But this assumption could be problematic. To fix this issue, we can reserve all the zero elements of $\mathbf{A}$ in $\mathbf{B}^{(k+1)}$ when solving (16).

It can be noticed that the computation of the proposed ADMM algorithm is dominated by the eigenvalue decomposition of $\tilde{\mathbf{Z}}^{(k)}$, when solving (17). Since $\tilde{\mathbf{Z}}^{(k)}$ can be a dense matrix, its computational complexity is $\mathcal{O}(N^3)$, which is too expensive for a large scale graph. To overcome this difficulty, one can first decompose the original graph to a number of subgraphs of smaller size. Then, the proposed ADMM algorithm is independently applied on each subgraph. The final bipartite graph can be obtained by merging sub-bipartite graphs. The subgraph decomposition needs to cut off some edges, which may be reserved in the optimal solution to (1). To reduce the overall bipartite graph approximation error, one can employ the minimum cut algorithm to accomplish the subgraph decomposition [18].

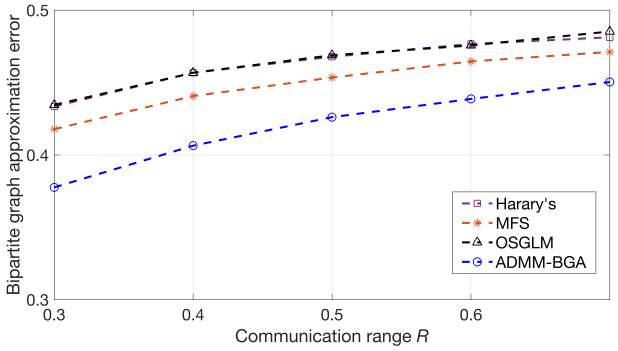

Fig. 1. Variation of average bipartite graph approximation with respect to communication range R.

4. EXPERIMENTAL RESULTS

The performance of the proposed ADMM-based Bipartite Graph Approximation (ADMM-BGA for short) is evaluated in this section. In our experiments, the original non-bipartite graphs are generated by sensor networks. We assume that each network is composed by N sensors, randomly distributed within a unit square $[0,1] \times [0,1]$. Each sensor is considered a vertex of the corresponding graph. A pair of sensors are considered to be neighbors if their distance is within a communication range R. Weight $a_{i,j}$ assigned to the edge between neighboring sensors i and j is computed by

$$a_{i,j} = \exp\left(-\frac{\|\mathbf{p}_i - \mathbf{p}_j\|_2^2}{2\sigma^2}\right) \tag{23}$$

where $\mathbf{p}_i$ denotes the planar position of sensor i. In our experiment, σ^2 is chosen as 10.

For comparison, we also employ the Harary's decomposition [10], BSD on MFS [13], and OSGLM algorithm [16] in our experiments. The approximation accuracy is reflected by

$$e_{\text{BGA}} = \frac{\|\mathbf{A} - \mathbf{B}\|_F^2}{\|\mathbf{A}\|_F^2}. \tag{24}$$

Example 1: Evaluation of bipartite graph approximation. In the first experiment, the number N of sensors is chosen equal to 50. Communication range R varies from 0.3 to 0.7. For each R, 100 independent simulations are conducted and the variation of the average approximation error with respect to R is illustrated in Fig. 1.

It can be observed that the average approximation error becomes larger with the increase of R. For a fixed N, increasing R yields a denser graph. So all the approaches need to cut off more edges of the original graphs and, thus, yield larger errors of bipartite graph approximation. But, among all the algorithms, the proposed algorithm achieves the smallest approximation error.

Table 1. Average SNR gain of ADMM-BGA over other approaches

	Percentage ($r\%$)			
	30	40	50	60
Harary's	3.45	4.82	4.43	4.01
MFS	1.80	3.17	1.18	0.57
OSGLM	1.21	0.42	0.90	−0.30

Example 2: Evaluation of graph signal reconstruction via two-channel wavelet filter-banks. In this experiment, we evaluate the graph signal reconstruction performance of bipartite graphs obtained by the proposed algorithm. Original graphs are still obtained by sensor networks using $N = 30$ and $R = 0.5$. Over the networks, we assume that the model for graph signals $\mathbf{f} \in \mathbb{R}^N$ is a Gaussian Markov random field (GMRF) [20], that is, $\mathbf{f} \sim \mathcal{N}(\mu, \Sigma)$ where μ is the mean vector, and Σ is the covariance matrix specified by the Laplacian matrix $\mathbf{L}$ of an arbitrary graph $\mathcal{G}$, i.e., $\Sigma^{-1} = \mathbf{L} + \delta\mathbf{I}$. For simplicity, we assume that $\mu = \mathbf{0}$. Parameter δ is chosen as $\delta = 0.1$. Then, the two-channel wavelet filter-banks proposed in [10] are applied on an approximate bipartite graph or subgraphs, obtained by each algorithm. Finally, we reconstruct the signals using the largest $r\%$ wavelet coefficients. For different r, the simulation is conducted independently for 100 times and the average SNR gain over the other approaches is computed and summarized in Table 1.

From Table 1, one can observe that the proposed algorithm demonstrates better reconstruction performance of graph signals than the other approaches, when r is small. This implies that the ADMM-BGA is more capable of capturing the essential structures of a non-bipartite graph. With the increase of r, the OSGLM outperforms than the other approaches, due to the redundancy existing in the oversampled graph.

5. CONCLUSIONS

In this paper, we propose a novel bipartite graph approximation algorithm. The original problem aims to minimize the approximation error between adjacency matrices of an arbitrary non-bipartite graph and its approximate bipartite counterpart. It has been proven that, if the non-negativity constraint of adjacency matrices is removed, optimal solutions can be efficiently obtained by the eigenvalue decomposition. To obtain a real bipartite graph, an ADMM algorithm is further developed, such that the non-negativity of the adjacency matrix of the resulting bipartite graph can be readily guaranteed. Experimental results show that the proposed algorithm can achieve smaller bipartite graph approximation error and competitive graph signal reconstruction performance, compared to the other bipartite subgraph decomposition and graph oversampling approaches.

6. REFERENCES

[1] A. Sandryhaila and J. M. F. Moura, "Discrete signal processing on graphs," *IEEE Trans. Signal Process.*, vol. 61, no. 7, pp. 1644–1656, Apr. 2013.

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

[3] 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.

[4] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Commun. Mag.*, vol. 40, no. 8, pp. 102-114, Aug. 2002.

[5] G. Shen and A. Ortega, "Optimized distributed 2D transforms for irregularly sampled sensor network grids using wavelet lifting," in *Proc. Int. Conf. Acoust., Speech, Signal Process (ICASSP)*, pp. 2513–2516, Apr. 2008.

[6] M. O. Jackson, *Social and Economic Networks*. Princeton, NJ, USA: Princeton Univ. Press, 2008.

[7] W.M.Campbell, C. K. Dagli, andC. J.Weinstein, "Social network analysis with content and graphs," *Lincoln Lab. J.*, vol. 20, no. 1, pp. 61–81, 2013.

[8] G. Michau, P. Borgnat, N. Pustelnik, P. Abry, A. Nantes, and E. Chung, "Estimating link-dependent origin-destination matrices from sample trajectories and traffic counts," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2015, pp. 5480–5484.

[9] S. Narang and A. Ortega, "Compact support biorthogonal wavelet filterbanks for arbitrary undirected graphs," *IEEE Trans. Signal Process.*, vol. 61, no. 19, pp. 4673–4685, Oct. 2013.

[10] S. Narang and A. Ortega, "Perfect reconstruction two-channel wavelet filter banks for graph structured data," *IEEE Trans. Signal Process.*, vol. 60, no. 6, pp. 2786–2799, Jun. 2012.

[11] F. Harary, D. Hsu, and Z. Miller, "The biparticity of a graph," *J. Graph Theory*, vol. 1, no. 2, pp. 131–133, 1977.

[12] R. M. Karp, "Reducibility among combinatorial problems," in *Complexity of Computer Computations.* Berlin,Germany: Springer, 1972, pp. 85–103.

[13] J. Zeng, G. Cheung, and A. Ortega, "Bipartite subgraph decomposition for critically sampled wavelet filterbanks for arbitary graphs," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2016, pp. 6210–6214.

[14] J. Zeng, G. Cheung, and A. Ortega, "Bipartite approximation for graph wavelet signal decomposition," *IEEE Trans. Signal Process.*, vol. 65, no. 20, pp. 5466–5480, 2017.

[15] Y. Tanaka and A. Sakiyama, "M-channel oversampled graph filter banks," *IEEE Trans. Signal Process.*, vol. 62, pp. 3578–3590, 2014.

[16] A. Sakiyama and Y. Tanaka, "Oversampled graph Laplacian matrix for graph filter banks," *IEEE Trans. Signal Process.*, vol. 62, no. 24, pp. 6425– 6437, 2014.

[17] J. Kunegis, "Exploiting the structure of bipartite graphs for algebraic and spectral graph theory applications," *Internet Math.*, vol. 11, no. 3, pp. 201–231, 2015.

[18] M. E. J. Newman, *Networks: An Introduction.* Oxford, U.K.: Oxford Univ. Press, 2010.

[19] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1985.

[20] H. Rue and L. Held, *Gaussian Markov Random Fields: Theory and Applications*. Boca Raton, FL, USA: CRC Press, 2005.