A closed-form solution for graph signal separation based on smoothness

Mohammad-Hassan Ahmad Yarandi, Massoud Babaie-Zadeh

Abstract—**Using smoothness criteria to separate smooth graph signals from their summation is an approach that has recently been proposed [1] and shown to have a unique solution up to the uncertainty of the average values of source signals. In this correspondence, closed-form solutions of both exact and approximate decompositions of that approach are presented. This closed-form solution in the exact decomposition also answers the open problem of the estimation error. Additionally, in the case of Gaussian source signals in the presence of additive Gaussian noise, it is shown that the optimization problem of that approach is equivalent to the Maximum A Posteriori (MAP) estimation of the sources.**

Index Terms—**graph signal processing, graph signal separation, blind source separation, smooth graph signal.**

I. Introduction

Graph Signal Processing (GSP) [2] studies signal arising from complex structures by modeling the relation between signal samples with graphs. A classical signal processing topic that has recently been generalized to GSP is Blind Source Separation (BSS) [3] whose goal is to retrieve source signals from their mixtures. In [4], [5], some classical BSS methods are extended to separate graph signals from a set of their mixtures, in which, similar to many classical BSS methods, the number of mixtures is assumed to be equal to the number of source signals. However, the problem that is investigated in [1] is the separation of smooth graph signals from only one mixture of them, *i.e.* the summation of the source signals. The proposed methods in [1] utilize smoothness criteria as the main objective function for two different types of signal decomposition: exact (to be used for noiseless mixtures) and approximate (to be used for noisy mixtures). Additionally, it is shown in [1] that the corresponding optimization problems have unique solutions up to the indeterminacy of their DC values, *i.e.* the average of each signal. However, the problem of how well these decompositions estimate the original source signals is left open in [1].

In this correspondence, we derive closed-form solutions for both exact and approximate decompositions of [1], and calculate the estimation error in the exact decomposition. Additionally, we show that for Gaussian sources in the presence of additive Gaussian noise, the approximate decomposition is equivalent to Maximum A Posteriori (MAP) estimation of the sources.

The rest of this correspondence is organized as follow. In Section II, brief preliminaries are mentioned and then Section III is devoted to the closed-form solutions and the MAP estimation problem.

II. Preliminaries

An undirected graph with N nodes can be represented as $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{W})$, where $\mathcal{V}$ and $\mathcal{E}$ denote the node set and edge set, respectively, and $\mathbf{W} \in \mathbb{R}^{N\times N}$ is the adjacency matrix with entries $w_{ij} = w_{ji} \geq 0$. Another matrix that describes $\mathcal{G}$ is the Laplacian matrix, which is defined as $\mathbf{L} \triangleq \mathbf{D} - \mathbf{W}$, where $\mathbf{D}$ is a diagonal matrix with diagonal entries $d_{ii} \triangleq \sum_{j=1}^{N} w_{ij}$. $\mathbf{L}$ is a positive semidefinite matrix with eigenvalues $0 = \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_N$ and corresponding orthonormal eigenvectors $\frac{1}{\sqrt{N}}\mathbf{1} = \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N$, where $\mathbf{1}$ stands for the all-one vector (for a connected graph, only the smallest eigenvalue is equal to zero, therefore range($\mathbf{L}$) consists of all vectors in $\mathbb{R}^n$ whose DC values are zero) [2]. A graph signal $\mathbf{x} \in \mathbb{R}^N$ is a mapping from the node set to $\mathbb{R}^N$ that assigns a real value to each node. The smoothness of a graph signal $\mathbf{x}$ can be measured by the graph Laplacian quadratic form as $\mathbf{x}^T\mathbf{L}\mathbf{x} = \sum_{1\leq i<j\leq N} w_{ij}(x_i - x_j)^2$, which evaluates the variation of the signal values on the graph [2].

For a positive semidefinite matrix $\mathbf{A} \in \mathbb{R}^{N\times N}$ with eigenvalues $0 \leq \lambda_1 \leq \ldots \leq \lambda_N$ and corresponding orthonormal eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_N$, the eigenvalue decomposition is as $\mathbf{A} = \sum_{i=1}^{N} \lambda_i \mathbf{v}_i\mathbf{v}_i^T = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T$, where $\mathbf{\Lambda} \triangleq \text{diag}(\lambda_1, \ldots, \lambda_N)$ and $\mathbf{V} \triangleq [\mathbf{v}_1, \ldots, \mathbf{v}_N]$. The Moore-Penrose pseudo inverse of $\mathbf{A}$ is a positive semidefinite matrix defined as $\mathbf{A}^\dagger = \sum_{i\in\mathcal{I}} \frac{1}{\lambda_i}\mathbf{v}_i\mathbf{v}_i^T = \mathbf{V}\mathbf{\Lambda}^\dagger\mathbf{V}^T$, where $\mathcal{I} = \{1 \leq i \leq N : \lambda_i \neq 0\}$. Therefore, $\mathbf{A}$ can be considered as an injective mapping from range($\mathbf{A}$) to range($\mathbf{A}$) with inverse mapping $\mathbf{A}^\dagger$, which results in the following lemma.

Lemma 1. *Let* $\mathbf{y} \in \text{range}(\mathbf{A})$. *The system of linear equations* $\mathbf{A}\mathbf{x} = \mathbf{y}$ *has a unique solution in* range($\mathbf{A}$), *which is equal to* $\mathbf{x} = \mathbf{A}^\dagger\mathbf{y}$.

III. Closed-form solutions

A. Exact decomposition

Let $\mathbf{x}_1, \ldots, \mathbf{x}_K \in \mathbb{R}^N$ be K smooth graph signals over connected graphs $\mathcal{G}_1, \ldots, \mathcal{G}_N$, respectively. Graphs $\mathcal{G}_1, \ldots, \mathcal{G}_N$ have the same nodes and are different in their edge sets. Suppose the average value of each $\mathbf{x}_i$ is equal to zero, *i.e.* $\mathbf{1}^T\mathbf{x}_i = 0,\ i = 1, \ldots, K$. The optimization problem proposed in [1] for retrieving $\mathbf{x}_i$'s from their summation $\mathbf{x} = \sum_{i=1}^{K} \mathbf{x}_i$ is

$$\underset{\mathbf{x}_1,\ldots,\mathbf{x}_K}{\text{minimize}} \quad \sum_{i=1}^{K} \mathbf{x}_i^T\mathbf{L}_i\mathbf{x}_i \quad \text{s.t.} \begin{cases} \mathbf{x} = \sum_{i=1}^{K} \mathbf{x}_i \ , \\ \mathbf{1}^T\mathbf{x}_i = 0 \ , \ i = 1, \ldots, K. \end{cases} \tag{1}$$

As proved in [1], problem (1) has a unique solution. Moreover, as explained in [1], in case of graphs with different eigenvalue distributions, the objective function in (1) is better to be changed in a way that each frequency component has the same weight in the objective function, which can be done by replacing $\mathbf{L}_i = \mathbf{V}_i\mathbf{\Lambda}_i\mathbf{V}_i^T$ with $\tilde{\mathbf{L}}_i = \mathbf{V}_i\text{diag}(0, w_2, \ldots, w_N)\mathbf{V}_i^T$, for $i = 1, \ldots, K$, where $w_2, \ldots, w_N$ are positive weights (as stated in [1], this case can also be used for source signals that are not necessarily smooth but have the same known sparse frequency support). Let $\mathbf{x}_1^*, \ldots, \mathbf{x}_K^*$ denote the unique solution of (1). The following theorem gives these signals in closed form.

Theorem 1. *The closed-form solution of (1) is given by*

$$\mathbf{x}_j^* = \mathbf{L}_j^\dagger\Big(\sum_{i=1}^{K} \mathbf{L}_i^\dagger\Big)^\dagger\mathbf{x} \ , \quad j = 1, \ldots, K. \tag{2}$$

Moreover, the optimum value of (1) is equal to $\mathbf{x}^T\big(\sum_{i=1}^{K} \mathbf{L}_i^\dagger\big)^\dagger\mathbf{x}$.

Proof. Based on Appendix A in [1], $\mathbf{L}_1\mathbf{x}_1^* = \ldots = \mathbf{L}_K\mathbf{x}_K^*$. Since the DC value of $\mathbf{x}_i^*$ is zero and $\mathcal{G}_i$ is connected, $\mathbf{x}_i^* \in \text{range}(\mathbf{L}_i)$, and therefore from Lemma 1, each $\mathbf{x}_i^*$ can be represented as

Department of Electrical Engineering, Sharif University of Technology, Tehran, Iran (e-mail: mh.ahmad.yarandi@gmail.com, mbzadeh@yahoo.com).

$\mathbf{x}_i^* = \mathbf{L}_i^\dagger \mathbf{L}_j \mathbf{x}_j^*,\ i = 1, \ldots, K$. Therefore, $\mathbf{x} = \sum_{i=1}^K \mathbf{x}_i^*$ leads to $(\sum_{i=1}^K \mathbf{L}_i^\dagger)\mathbf{L}_j \mathbf{x}_j^* = \mathbf{x}$. Both $\sum_{i=1}^K \mathbf{L}_i^\dagger$ and $\mathbf{L}_j$ are positive semidefinite matrices and have only one zero eigenvalue with corresponding eigenvector $\mathbf{1}$. Applying again Lemma 1 gives rise to the closed-form solution (2). Finally, putting (2) in the objective function is simplified as the mentioned optimum value. □

The above theorem enables us to calculate the estimation error, which had remained open in [1]. This is given by the following corollaries, whose proofs are simple and are left to the reader.

Corollary 1. *Let $\mathbf{x}_1, \ldots, \mathbf{x}_K$ be the original source signals. The error between $\mathbf{x}_j$ and $\mathbf{x}_j^*$ is equal to*

$$\|\mathbf{x}_j^* - \mathbf{x}_j\|_2^2 = \|(\mathbf{L}_j^\dagger(\sum_{i=1}^K \mathbf{L}_i^\dagger)^\dagger - \mathbf{I})\mathbf{x}_j + \mathbf{L}_j^\dagger(\sum_{i=1}^K \mathbf{L}_i^\dagger)^\dagger \sum_{\substack{i=1\\i\neq j}}^K \mathbf{x}_i\|_2^2. \tag{3}$$

Corollary 2. *Let $\mathbf{x}_i$'s be independent random vectors with covariance matrix $\mathbf{C}_i = \mathbb{E}\{\mathbf{x}_i\mathbf{x}_i^T\}$. Moreover, suppose that for each of the samples of these random vectors $\mathbf{1}^T\mathbf{x}_i = 0$. The expected value of the error between $\mathbf{x}_j$ and $\mathbf{x}_j^*$ is equal to*

$$\begin{aligned}\mathbb{E}\{\|\mathbf{x}_j^* - \mathbf{x}_j\|_2^2\} =& \operatorname{tr}((\mathbf{I} - \mathbf{L}_j^\dagger(\sum_{i=1}^K \mathbf{L}_i^\dagger)^\dagger)\mathbf{C}_j(\mathbf{I} - (\sum_{i=1}^K \mathbf{L}_i^\dagger)^\dagger\mathbf{L}_j^\dagger)) \\ &+ \sum_{\substack{i=1\\i\neq j}}^K \operatorname{tr}(\mathbf{L}_j^\dagger(\sum_{i=1}^K \mathbf{L}_i^\dagger)^\dagger \mathbf{C}_i(\sum_{i=1}^K \mathbf{L}_i^\dagger)^\dagger \mathbf{L}_j^\dagger).\end{aligned} \tag{4}$$

B. Approximate decomposition

In the case of approximate decomposition $\mathbf{x} \approx \sum_{i=1}^K \mathbf{x}_i$ (useful in the presence of noise), the optimization problem proposed in [1] is

$$\begin{aligned}\underset{\mathbf{x}_1,\ldots,\mathbf{x}_K}{\text{minimize}}\quad & \|\mathbf{z} - \sum_{i=1}^K \mathbf{x}_i\|_2^2 + \sum_{i=1}^K \gamma_i \mathbf{x}_i^T \mathbf{L}_i \mathbf{x}_i \\ \text{s.t.}\quad & \mathbf{1}^T\mathbf{x}_i = 0\ , \quad i = 1, \ldots, K\ ,\end{aligned} \tag{5}$$

where $\mathbf{z} = \mathbf{x} - (\frac{\mathbf{1}^T\mathbf{x}}{N})\mathbf{1}$ and γ_i's are regularization parameters. Similar to the exact decomposition, the uniqueness of the solution of (5) is proved in [1], and it is also shown that (5) can be generalized to the case of using $\tilde{\mathbf{L}}_i$'s instead of $\mathbf{L}_i$'s.

Theorem 2. *The closed-form solution of (5) is given by*

$$\mathbf{x}_j^* = \frac{1}{\gamma_j}\mathbf{L}_j^\dagger(\mathbf{I} + \sum_{i=1}^K \frac{1}{\gamma_i}\mathbf{L}_i^\dagger)^{-1}\mathbf{z}\ , \quad j = 1, \ldots, K. \tag{6}$$

Moreover, the optimum value of (5) is equal to $\mathbf{z}^T(\mathbf{I} + \sum_{i=1}^K \frac{1}{\gamma_i}\mathbf{L}_i^\dagger)^{-1}\mathbf{z}$.

Proof. Based on Appendix B in [1], $\gamma_1\mathbf{L}_1\mathbf{x}_1^* = \ldots = \gamma_K\mathbf{L}_K\mathbf{x}_K^* = \mathbf{z} - \sum_{i=1}^K \mathbf{x}_i^*$. Similar to the proof of Theorem 1, since $\mathbf{x}_i^* \in \text{range}(\mathbf{L}_i)$, $\mathbf{x}_i^* = \frac{\gamma_j}{\gamma_i}\mathbf{L}_i^\dagger\mathbf{L}_j\mathbf{x}_j^*,\ i = 1, \ldots, K$. Therefore, the equation $\gamma_j\mathbf{L}_j\mathbf{x}_j^* = \mathbf{z} - \sum_{i=1}^K \mathbf{x}_i^*$ leads to $\gamma_j(\mathbf{I} + \sum_{i=1}^K \frac{1}{\gamma_i}\mathbf{L}_i^\dagger)\mathbf{L}_j\mathbf{x}_j^* = \mathbf{z}$. Since $\mathbf{I} + \sum_{i=1}^K \frac{1}{\gamma_i}\mathbf{L}_i^\dagger$ is invertible, applying again Lemma 1 gives rise to the closed-form solution (6). Finally, putting (6) in the objective function is simplified as the mentioned optimum value. □

In the next theorem, it is shown that in the case of Gaussian sources with additive Gaussian noise, problem (5) without its constraints (problem (4) in [1]) is equivalent to the MAP estimation of the sources. So, the solution of (5) will also be a MAP estimation of the sources (since it is a special case of problem (4) in [1], for which the DC's are zero).

Suppose $\mathbf{x}_i$ on the graph $\mathcal{G}_i$ with the Laplacian matrix $\mathbf{L}_i = \mathbf{V}_i\mathbf{\Lambda}_i\mathbf{V}_i^T$ is generated as $\mathbf{x}_i = \mathbf{V}_i\mathbf{h}_i$, where $\mathbf{h}_i$'s are independent Gaussian random vectors, *i.e.* $\mathbf{h}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{\Lambda}_i^\dagger)$. This is equivalent to $\mathbf{x}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{L}_i^\dagger)$, which, as stated in [6], is a verified assumption for many common networks and graph databases. Moreover, suppose $\mathbf{x} = \sum_{i=1}^K \mathbf{x}_i + \mathbf{n}$, where $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$.

Theorem 3. *Under the above assumptions, problem (5) without its constraints (problem (4) in [1]) is equivalent to the MAP estimation.*

Proof. Based on the assumptions,

$$\begin{aligned}&\ln(\mathbb{P}\{\mathbf{x}|\mathbf{h}_1, \ldots, \mathbf{h}_K\}\mathbb{P}\{\mathbf{h}_1\}\ldots\mathbb{P}\{\mathbf{h}_K\}) = \\ &\qquad c - \frac{1}{2\sigma^2}\|\mathbf{x} - \sum_{i=1}^K \mathbf{V}_i\mathbf{h}_i\|_2^2 - \frac{1}{2}\sum_{i=1}^K \mathbf{h}_i^T\mathbf{\Lambda}_i\mathbf{h}_i,\end{aligned} \tag{7}$$

where c is a constant. So, considering $\mathbf{x}_i = \mathbf{V}_i\mathbf{h}_i$ leads to

$$\begin{aligned}\max_{\mathbf{h}_1,\ldots,\mathbf{h}_K}\quad & \mathbb{P}\{\mathbf{h}_1, \ldots, \mathbf{h}_K|\mathbf{x}\} \equiv \\ \max_{\mathbf{h}_1,\ldots,\mathbf{h}_K}\quad & \ln(\mathbb{P}\{\mathbf{x}|\mathbf{h}_1, \ldots, \mathbf{h}_K\}\mathbb{P}\{\mathbf{h}_1\}\ldots\mathbb{P}\{\mathbf{h}_K\}) \equiv \\ \min_{\mathbf{x}_1,\ldots,\mathbf{x}_K}\quad & \|\mathbf{x} - \sum_{i=1}^K \mathbf{x}_i\|_2^2 + \sigma^2\sum_{i=1}^K \mathbf{x}_i^T\mathbf{L}_i\mathbf{x}_i.\end{aligned} \tag{8}$$

□

Remark: All theorems can be generalized to the case that instead of $\mathbf{L}_i$'s, $\tilde{\mathbf{L}}_i$'s are used.

IV. Conclusion

Estimating smooth graph signals from their summation can be done by smooth graph signal decomposition, which has been shown that has a unique solution up to the indeterminacy of the average values of the source signals. In this correspondence, we obtained the closed-form solutions of both exact and approximate decompositions, which calculates the estimation error that had been left open in [1]. Moreover, we showed that in the case of Gaussian sources with additive Gaussian noise, the proposed approach of [1] for estimating the sources provides their MAP estimation.

In cases where graphs are unknown, the problem of joint graph learning and blind separation of smooth signals (similar to [5], but with only one observation) is an interesting future research. Moreover, investigating the dependency of the estimation error on the structures of the graphs and signals, and considering the estimation error as a regularization term to improve estimation, would be subjects for future research.

References

[1] S. Mohammadi, M. Babaie-Zadeh, and D. Thanou, "Graph signal separation based on smoothness or sparsity in the frequency domain," *IEEE Transactions on Signal and Information Processing over Networks*, pp. 1–10, 2023.

[2] 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, may 2013.

[3] P. Comon and C. Jutten, *Handbook of Blind Source Separation: Independent component analysis and applications*. Academic press, 2010.

[4] J. Miettinen, E. Nitzan, S. A. Vorobyov, and E. Ollila, "Graph signal processing meets blind source separation," *IEEE Transactions on Signal Processing*, vol. 69, pp. 2585–2599, 2021.

[5] A. Einizade and S. H. Sardouie, "Joint graph learning and blind separation of smooth graph signals using minimization of mutual information and laplacian quadratic forms," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 9, pp. 35–47, 2023.

[6] H. P. Maretic, M. El Gheche, M. Minder, G. Chierchia, and P. Frossard, "Wasserstein-based graph alignment," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 8, pp. 353–363, 2022.