Received: 6 January 2018 | Revised: 15 June 2019 | Accepted: 20 June 2019
DOI: 10.1002/nla.2264

RESEARCH ARTICLE

WILEY

An efficient algorithm for bounded spectral matching with affine constraint

Yan Dou[1,2] | **Ming Gu**[3] | **Shivkumar Chandrasekaran**[4]

[1]School of Mathematics and Statistics, Lanzhou University, Lanzhou, China
[2]School of Science, Lanzhou University of Technology, Lanzhou, China
[3]Department of Mathematics, University of California, Berkeley, Berkeley, California
[4]Department of Electrical and Computer Engineering, University of California, Santa Barbara, Santa Barbara, California

Correspondence
Yan Dou, School of Mathematics and Statistics, Lanzhou University, Lanzhou 730000, China; or School of Science, Lanzhou University of Technology, Lanzhou 730050, China.
Email: douy08@lzu.edu.cn

Funding information
China Scholarship Council, Grant/Award Number: 201606180087; National Natural Science Foundation of China, Grant/Award Number: 11471150, 11401281, and 11801242

Summary

Graph matching problem appears frequently in the applications of computer vision and machine learning. In this work, based on the spectral matching with affine constraint (SMAC) formulation, we present a new formulation, named bounded SMAC (BSMAC), for the graph matching problem by adding an upper-bound constraint on the solution norm. We demonstrate the existence of a unique solution with BSMAC, whereas SMAC needs not to have any meaningful solution in general. We develop an effective numerical method to solve the BSMAC formulation as an optimization problem. Numerical experiments are presented to verify feasibility and to show the performance of the proposed numerical method.

KEYWORDS

BSMAC formulation, graph matching, matrix transformation, numerical method, optimization problem, upper-bound constraint

1 | INTRODUCTION

Graph matching is of great importance in the applications of computer vision and machine learning.[1–5] This problem is to make recognition of objects by establishing consistent correspondences of the points or line segments between two related graphs that preserves the relationships as much as possible. For solving this problem, many algorithms have been established in the literature; see other works,[6–9] and the references therein. The most frequent approach is firstly to formulate the graph matching problem as the following integer quadratic program (IQP)[6–8,10–12]:

$$\begin{aligned} \max \quad & x^T Q x \\ \text{s.t.} \quad & Px \le b, \\ & x \in \{0,1\}^n. \end{aligned} \tag{1}$$

Here, x is a stretch vector of the match matrix M, which is a binary matrix with $M_{ii'} = 1$ iff the vertex i from one graph is matched to the vertex i' from the other graph. $Q \in \mathbb{R}^{n\times n}$ is a compatibility matrix, which is symmetric. $Px \le b$ is the one-to-one or one-to-many constraint, where $P \in \mathbb{R}^{m\times n}$ is a rectangular matrix and $0 \ne b \in \mathbb{R}^m$.

Generally, such IQP (1) is NP-hard and an exact solution is intractable to obtain. Therefore, much research in the literature aims to find good approximate solutions by developing approximate relaxations to the integer one-to-one or one-to-many constraints. Many methods[13–15] are proposed to find a global optimum solution based on the continuous relaxations of the IQP by dropping the integer constraint $x \in \{0, 1\}^n$. Then, some postprocessing is performed to binarize the continuous solution. In 2005, Leordeanu et al.[7] proposed a spectral matching (SM) formulation on a continuous program by removing the constraints $Px \leq b$ as well as $x \in \{0, 1\}^n$ and fixing the norm of x to 1, which can be solved by computing the leading eigenvector x of Q. Because the compatibility matrix Q has nonnegative elements, by Perron–Frobenius theorem, the elements of the solution x are in the interval $[0, 1]$. Based on the SM formulation and imposing affine constraint $Px = b$ on the relaxed solution, Cour et al.[6] put forward an SM with affine constraint (SMAC) program:

$$\begin{aligned} \max \quad & \frac{x^T Q x}{x^T x} \\ \text{s.t.} \quad & Px = b, \end{aligned} \tag{2}$$

which can be viewed as maximizing the Rayleigh quotient under affine constraints. In previous works,[6,16] the solution to (2) is given by the leading eigenpair of the matrix $\bar{P}Q\bar{P}$, where $\bar{P} = I_n - P_{eq}^T(P_{eq}P_{eq}^T)^{-1}P_{eq}$, $P_{eq} = [I_{m-1}, 0](P - (1/b_m)bP_m)$, and P_m, b_m denote the last row of P, b. Then, the solution is scaled to satisfy $Px = b$ exactly. Experimental results in the work of Cour et al.[6] showed that the relaxation SMAC formulation can give more accurate solutions than the SM.

However, as is mentioned in the previous study,[16] when the leading eigenvector of Q and the range of P^T are orthogonal, only a diverging sequence approximates the supremum of (2). Another problem is if the leading eigenvector of Q and the range of P^T are nearly, but not exactly, orthogonal, we find that, although there is an exact solution for the formulation (2), the norm of the solutions that satisfy the affine constraint $Px = b$ will be too large to be useful for the original graph matching problem. On the other hand, there is a natural upper bound on the norm of the solution vector, due to the original discrete solution being binary. Therefore, we impose an upper-bound constraint $\|x\|_2 \leq \mu$ (μ is a fixed number) on the relaxed solution based on (2) and propose the following bounded SMAC (BSMAC) formulation:

$$\begin{aligned} \max \quad & \frac{x^T Q x}{x^T x} \\ \text{s.t.} \quad & Px = b, \\ & \|x\|_2 \leq \mu. \end{aligned} \tag{3}$$

This is an optimization problem with a closed feasibility region and therefore must have a finite optimal solution. We propose a numerical method to solve it in the following sections. Generally, the affine constraint $Px = b$ is consistent, that is, $b \in \mathbb{R}^m$ is a vector in the range of the matrix $P \in \mathbb{R}^{m \times n} (m < n)$. Without loss of generality, we assume the matrix P is of full row rank throughout this paper.

The rest of this paper is organized as follows. In Section 2, we first simplify the optimization problem (3) to a lower dimensional problem with matrix transformations. In Section 3, we develop a numerical method to effectively solve the simplified problem. In Section 4, we report numerical experimental results that support our analysis. Finally, in Section 5, we draw conclusions.

2 | PROBLEM SIMPLIFICATION

Following the idea of Gander et al.,[17] in this section, we use some transformations to simplify the optimization problem (3) to a lower dimensional one, which we consider in Section 3.

Firstly, we introduce an orthogonal matrix H_b, the Householder reflector matrix of $b \in \mathbb{R}^m$, to simplify the affine constraint $Px = b$ in (3). By making orthogonal transformation on both sides of $Px = b$, we have

$$\begin{pmatrix} \hat{p}^T \\ \hat{P} \end{pmatrix} x = H_b P x = H_b b = \begin{pmatrix} \|b\|_2 \\ 0 \end{pmatrix}, \tag{4}$$

where $\hat{p} \in \mathbb{R}^n$, $\hat{P} \in \mathbb{R}^{(m-1)\times n}$, and $\hat{p}^T$ denotes the first row of the matrix $H_b P$. Then, make the QR decomposition of $\hat{P}^T$

$$\hat{P}^T = \bar{W}\begin{pmatrix} R \\ 0 \end{pmatrix}, \tag{5}$$

$\bar{W} \in \mathbb{R}^{n\times n}$ is an orthogonal matrix, and $R \in \mathbb{R}^{(m-1)\times(m-1)}$ is an upper triangular matrix of full rank. Denote

$$\begin{aligned} \bar{p} &:= \bar{W}^T\hat{p} = \begin{pmatrix} \bar{p}_1 \\ \bar{p}_2 \end{pmatrix}, \\ \bar{x} &:= \bar{W}^T x = \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \end{pmatrix}, \end{aligned} \tag{6}$$

where $\bar{p}_1, \bar{x}_1 \in \mathbb{R}^{m-1}$ and $\bar{p}_2, \bar{x}_2 \in \mathbb{R}^{n-m+1}$. Then, Equation (4) can be equivalently written as

$$\begin{pmatrix} \hat{p}^T\bar{W} \\ \hat{P}\bar{W} \end{pmatrix}\bar{W}^T x = \begin{pmatrix} \bar{p}_1^T & \bar{p}_2^T \\ R^T & 0 \end{pmatrix}\begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \end{pmatrix} = \begin{pmatrix} \|b\|_2 \\ 0 \end{pmatrix},$$

which is

$$\begin{cases} \bar{p}_1^T\bar{x}_1 + \bar{p}_2^T\bar{x}_2 = \|b\|_2, \\ R^T\bar{x}_1 = 0. \end{cases}$$

Because the square matrix R^T is nonsingular, it follows that $\bar{x}_1 = 0$. Make the following split in correspondence with the above dimension and denote

$$\bar{Q} := \bar{W}^T Q\bar{W} = \begin{pmatrix} \bar{Q}_{11} & \bar{Q}_{12} \\ \bar{Q}_{21} & \bar{Q}_{22} \end{pmatrix}. \tag{7}$$

Now, formulation (3) can be simplified as

$$\begin{aligned} \max \quad & \frac{\bar{x}_2^T\bar{Q}_{22}\bar{x}_2}{\bar{x}_2^T\bar{x}_2} \\ \text{s.t.} \quad & \bar{p}_2^T\bar{x}_2 = \|b\|_2, \\ & \|\bar{x}_2\|_2 \le \mu. \end{aligned} \tag{8}$$

Here, $\bar{Q}_{22} \in \mathbb{R}^{(n-m+1)\times(n-m+1)}$ is also a symmetric matrix. Thus, in place of the original formulation (3), we can solve the above lower dimensional optimization problem (8). To ensure that there is a solution of (8), the following relationship should be satisfied:

$$\|b\|_2 = \bar{p}_2^T\bar{x}_2 \le \|\bar{p}_2\|_2\|\bar{x}_2\|_2 \le \mu\|\bar{p}_2\|_2.$$

In this paper, we assume the strict inequality holds, that is,

$$\|\bar{p}_2\|_2 > \frac{\|b\|_2}{\mu}, \tag{9}$$

which will be used in the analysis in Section 3.

Note that when the vectors x and $\bar{x}_2$ satisfy the relationship (6), the maximum values of the objective functions $(x^TQx)/(x^Tx)$ in (3) and $(\bar{x}_2^T\bar{Q}_{22}\bar{x}_2)/(\bar{x}_2^T\bar{x}_2)$ in (8) are the same with their respective affine constraints. Now, it is much easier to understand the infinity of the solutions when the leading eigenvector of Q and the range of P^T are nearly orthogonal, which is exactly the case where the leading eigenvector of $\bar{Q}_{22}$ is nearly orthogonal to the vector $\bar{p}_2$ in the lower dimensional formulation (8).

3 | THE NUMERICAL METHOD

In this section, we give the details of the new numerical method to solve the simplified optimization problem (8).

In order to calculate the stationary point of the objective function in (8), we introduce multiplier y for the equality constraint and multiplier θ for the inequality constraint to form the Lagrangian:

$$\mathcal{L}(\bar{x}_2, y, \theta) := \frac{1}{2}\frac{\bar{x}_2^T\bar{Q}_{22}\bar{x}_2}{\bar{x}_2^T\bar{x}_2} + y\left(\bar{p}_2^T\bar{x}_2 - \|b\|_2\right) - \frac{\theta}{2}\left(\bar{x}_2^T\bar{x}_2 - \mu^2\right).$$

Differentiating $\mathcal{L}(\bar{x}_2, y, \theta)$ with respect to $\bar{x}_2$ yields the equation

$$\begin{aligned}\nabla_{\bar{x}_2}\mathcal{L}(\bar{x}_2, y, \theta) &= \frac{\bar{Q}_{22}\bar{x}_2}{\bar{x}_2^T\bar{x}_2} - \frac{\bar{x}_2^T\bar{Q}_{22}\bar{x}_2}{\left(\bar{x}_2^T\bar{x}_2\right)^2}\bar{x}_2 + y\bar{p}_2 - \theta\bar{x}_2 \\ &= 0.\end{aligned} \tag{10}$$

Denote

$$\begin{aligned}\alpha &:= \frac{\bar{x}_2^T\bar{Q}_{22}\bar{x}_2}{\bar{x}_2^T\bar{x}_2}, \\ \bar{y} &:= \left(\bar{x}_2^T\bar{x}_2\right)y, \\ \bar{\theta} &:= \left(\bar{x}_2^T\bar{x}_2\right)\theta \\ \bar{\alpha} &:= \alpha + \bar{\theta}.\end{aligned}$$

Then, Equation (10) can be rewritten as

$$(\bar{Q}_{22} - \bar{\alpha}I)\bar{x}_2 + \bar{y}\bar{p}_2 = 0.$$

We put together the optimality conditions as follows:

$$\begin{cases}(\bar{Q}_{22} - \bar{\alpha}I)\bar{x}_2 + \bar{y}\bar{p}_2 = 0, \\ \bar{p}_2^T\bar{x}_2 = \|b\|_2, \\ \bar{\theta}\left(\bar{x}_2^T\bar{x}_2 - \mu^2\right) = 0, \\ \bar{x}_2^T\bar{x}_2 \le \mu^2, \\ \bar{\theta} \ge 0.\end{cases} \tag{11}$$

In the following, we develop an algorithm to compute the global optimal solution based on the conditions in (11).

Let $\lambda_{\max}$ be the largest eigenvalue of $\bar{Q}_{22}$; then, make the unique decomposition

$$\bar{p}_2 = u_1 + u_2, \tag{12}$$

where $u_1 \in \text{kernel}(\bar{Q}_{22} - \lambda_{\max}I)$ and $u_2 \in \text{kernel}(\bar{Q}_{22} - \lambda_{\max}I)^{\perp}$. In the discussions that follow, we assume that $u_1 \neq 0$ for simplicity of presentation, even though our conclusions hold valid even if $u_1 = 0$.

In the following subsections, we discuss the optimal solutions of (11) by considering two cases: $\bar{\theta} = 0$ (solution norm condition is inactive) and $\bar{\theta} > 0$ (solution norm condition is active).

3.1 | Solvability when $\bar{\theta} = 0$

In this part, we consider the first case where $\bar{\theta} = 0$. From the definition of $\alpha := (\bar{x}_2^T \bar{Q}_{22}\bar{x}_2)/(\bar{x}_2^T\bar{x}_2)$, it follows that $\bar{x}_2^T(\bar{Q}_{22} - \alpha I)\bar{x}_2 = 0$. Adding this condition to the Karush–Kuhn–Tucker (KKT) conditions (11), we have

$$\begin{cases} (\bar{Q}_{22} - \alpha I)\bar{x}_2 + \bar{y}\bar{p}_2 = 0, \\ \bar{x}_2^T(\bar{Q}_{22} - \alpha I)\bar{x}_2 = 0, \\ \bar{p}_2^T\bar{x}_2 = \|b\|_2, \\ \bar{x}_2^T\bar{x}_2 \leq \mu^2. \end{cases} \tag{13}$$

We know $(\bar{Q}_{22} - \alpha I)\bar{x}_2 = -\bar{y}\bar{p}_2$ from the first equation of (13). Substituting it into the second equation of (13) gives that $-\bar{y}\bar{x}_2^T\bar{p}_2 = 0$. Together with the third equation, we can derive that $\bar{y} = 0$ because $\|b\|_2 \neq 0$. Thus, the equation $(\bar{Q}_{22} - \alpha I)\bar{x}_2 = 0$ holds, which means α is an eigenvalue of $\bar{Q}_{22}$ and $\bar{x}_2$ is its corresponding eigenvector. Thus, in this case, the optimal solution α and vector $\bar{x}_2$ of (13) are the leading eigenpair of matrix $\bar{Q}_{22}$. Note that if $\lambda_{\max}$ is not simple, the solutions satisfying the condition $\bar{p}_2^T\bar{x}_2 = \|b\|_2$ is not unique. In order to satisfy the fourth equation of (13), we choose the unique solution with smallest 2-norm:

$$\bar{x}_2^* = \frac{\|b\|_2}{u_1^T u_1} u_1, \quad \text{with} \quad \|\bar{x}_2^*\|_2 = \frac{\|b\|_2}{\|u_1\|_2}. \tag{14}$$

Therefore, if $\|b\|_2/\|u_1\|_2 \leq \mu$, that is, $\|\bar{x}_2^*\|_2 \leq \mu$, we have derived the optimal solution of (13) with the leading eigenpair of the matrix $\bar{Q}_{22}$. However, if $\|b\|_2/\|u_1\|_2 > \mu$, that is, $\|\bar{x}_2^*\|_2 > \mu$, there is no solution of (13). Such case will certainly appear when the leading eigenvector of $\bar{Q}_{22}$ is nearly orthogonal to the vector $\bar{p}_2$ because the value of $\|u_1\|_2$ will be definitely small in this case. Then, we will consider the second case $\bar{\theta} > 0$, with the solution norm constraint now being active.

3.2 | Solvability when $\bar{\theta} > 0$

In the following, we are looking for the optimal solutions satisfying the KKT conditions (11) when $\bar{\theta} > 0$ and $\|b\|_2 > \|u_1\|_2\,\mu$. In this case, the KKT conditions can be simplified to

$$\begin{cases} (\bar{Q}_{22} - \bar{\alpha} I)\bar{x}_2 + \bar{y}\bar{p}_2 = 0, \\ \bar{x}_2^T(\bar{Q}_{22} - \alpha I)\bar{x}_2 = 0, \\ \bar{p}_2^T\bar{x}_2 = \|b\|_2, \\ \bar{x}_2^T\bar{x}_2 = \mu^2, \\ \bar{\theta} > 0. \end{cases} \tag{15}$$

In consideration of the symmetry of matrix $\bar{Q}_{22}$, we make its eigenvalue decomposition as $\bar{Q}_{22} = Z\Lambda Z^T$, where $\Lambda = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_{n-m+1})$ and the orthogonal matrix $Z = (z_1, z_2, \ldots, z_{n-m+1})$. The column vectors $z_1, z_2, \ldots, z_{n-m+1}$ of Z are corresponding eigenvectors of the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_{n-m+1}$, respectively. Without loss of generality, we assume that $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_{n-m+1}$. To make further simplification, we denote $\hat{x}_2 := Z^T\bar{x}_2$ and $q := Z^T\bar{p}_2$. Now, the problem is transformed into finding the optimal solution $\hat{x}_2$ with optimal objective value $\alpha = (\hat{x}_2^T\Lambda\hat{x}_2)/(\hat{x}_2^T\hat{x}_2)$ that satisfies the following KKT conditions:

$$\begin{cases} (\Lambda - \bar{\alpha} I)\hat{x}_2 + \bar{y} q = 0, \\ \hat{x}_2^T (\Lambda - \alpha I)\hat{x}_2 = 0, \\ q^T \hat{x}_2 = \|b\|_2, \\ \hat{x}_2^T \hat{x}_2 = \mu^2, \\ \bar{\theta} > 0. \end{cases} \tag{16}$$

By multiplying $\hat{x}_2^T$ to the two sides of the first equality of (16), it can be rewritten as

$$\hat{x}_2^T (\Lambda - \alpha I + \alpha I - \bar{\alpha} I)\hat{x}_2 + \bar{y}\hat{x}_2^T q = 0. \tag{17}$$

Substituting the second, third, and fourth equalities of (16) into (17), α can be represented as a function of $\bar{\alpha}$ without $\hat{x}_2$, which is

$$\alpha(\bar{\alpha}) = \bar{\alpha} - \frac{\|b\|_2}{\mu^2}\bar{y}. \tag{18}$$

As a result of $\bar{\theta} = \bar{\alpha} - \alpha > 0$, we know $\bar{y} > 0$, which will be used in the following analysis. To proceed with our analysis, we discuss two cases by considering if $\bar{\alpha}$ is an eigenvalue of $\bar{Q}_{22}$.

Case 1. $\bar{\alpha}$ is not an eigenvalue of Λ.

In this case, we have $\hat{x}_2 = -\bar{y}(\Lambda - \bar{\alpha} I)^{-1} q$ from the first equation of (16). Substituting it into the third and fourth equations of (16) yields

$$\begin{cases} -\bar{y} q^T (\Lambda - \bar{\alpha} I)^{-1} q = \|b\|_2, \\ \bar{y}^2 q^T (\Lambda - \bar{\alpha} I)^{-2} q = \mu^2. \end{cases} \tag{19}$$

Solving $\bar{y}$ from the first equation of (19) gives

$$\bar{y} = -\frac{\|b\|_2}{q^T (\Lambda - \bar{\alpha} I)^{-1} q}. \tag{20}$$

Extracting the square roots of the two sides of the second equation in (19) and then eliminating $\bar{y}$ ($\bar{y} > 0$), we arrive at

$$F(\bar{\alpha}) - \frac{\|b\|_2}{\mu} = 0, \tag{21}$$

where

$$F(\bar{\alpha}) = -\frac{q^T (\Lambda - \bar{\alpha} I)^{-1} q}{\sqrt{q^T (\Lambda - \bar{\alpha} I)^{-2} q}} = -\frac{q^T (\Lambda - \bar{\alpha} I)^{-1} q}{\|(\Lambda - \bar{\alpha} I)^{-1} q\|_2}.$$

We should solve the optimal solution $\bar{\alpha}$ from the above secular equation (21); then, the optimal solutions of $\bar{x}_2$, $\bar{y}$, and θ can be easily derived. There may be many solutions of Equation (21). Thus, we first need to decide the interval that the optimal solution $\bar{\alpha}$ is in and, then, to prove the existence and uniqueness of the solution in that interval.

We are aiming at maximizing the value of α, so we first present how α varies with different $\bar{\alpha}$, which satisfies the conditions (16). Suppose $\bar{\alpha}_1 > \bar{\alpha}_2$ are two solutions of Equation (21); then, $\bar{\alpha}_1$, $\bar{\alpha}_2$ satisfy

$$-\frac{q^T (\Lambda - \bar{\alpha}_1 I)^{-1} q}{\|(\Lambda - \bar{\alpha}_1 I)^{-1} q\|_2} = -\frac{q^T (\Lambda - \bar{\alpha}_2 I)^{-1} q}{\|(\Lambda - \bar{\alpha}_2 I)^{-1} q\|_2} = \frac{\|b\|_2}{\mu}. \tag{22}$$

From Equations (18), (20), we can derive the function $\alpha(\bar{\alpha}) = \bar{\alpha} + \frac{\|b\|_2^2}{\mu^2} \cdot \frac{1}{q^T(\Lambda - \bar{\alpha}I)^{-1}q}$. Using (22), make the difference as follows:

$$
\begin{aligned}
\alpha(\bar{\alpha}_1) - \alpha(\bar{\alpha}_2) &= (\bar{\alpha}_1 - \bar{\alpha}_2) + \frac{\|b\|_2^2}{\mu^2}\left(\frac{1}{q^T(\Lambda - \bar{\alpha}_1 I)^{-1}q} - \frac{1}{q^T(\Lambda - \bar{\alpha}_2 I)^{-1}q}\right) \\
&= (\bar{\alpha}_1 - \bar{\alpha}_2) + \frac{\|b\|_2^2}{\mu^2} \cdot \frac{q^T(\Lambda - \bar{\alpha}_2 I)^{-1}q - q^T(\Lambda - \bar{\alpha}_1 I)^{-1}q}{q^T(\Lambda - \bar{\alpha}_1 I)^{-1}q \cdot q^T(\Lambda - \bar{\alpha}_2 I)^{-1}q} \\
&= (\bar{\alpha}_1 - \bar{\alpha}_2) + \frac{\|b\|_2^2}{\mu^2} \cdot \frac{q^T\left[(\Lambda - \bar{\alpha}_1 I) - (\Lambda - \bar{\alpha}_2 I)\right](\Lambda - \bar{\alpha}_2 I)^{-1}(\Lambda - \bar{\alpha}_1 I)^{-1}q}{q^T(\Lambda - \bar{\alpha}_1 I)^{-1}q \cdot q^T(\Lambda - \bar{\alpha}_2 I)^{-1}q} \\
&= (\bar{\alpha}_1 - \bar{\alpha}_2) + \frac{\|b\|_2^2}{\mu^2} \cdot \frac{q^T(\bar{\alpha}_2 - \bar{\alpha}_1)(\Lambda - \bar{\alpha}_2 I)^{-1}(\Lambda - \bar{\alpha}_1 I)^{-1}q}{q^T(\Lambda - \bar{\alpha}_1 I)^{-1}q \cdot q^T(\Lambda - \bar{\alpha}_2 I)^{-1}q} \\
&= (\bar{\alpha}_1 - \bar{\alpha}_2) \cdot \left(1 - \frac{\|b\|_2^2}{\mu^2} \cdot \frac{q^T(\Lambda - \bar{\alpha}_2 I)^{-1}(\Lambda - \bar{\alpha}_1 I)^{-1}q}{q^T(\Lambda - \bar{\alpha}_1 I)^{-1}q \cdot q^T(\Lambda - \bar{\alpha}_2 I)^{-1}q}\right) \\
&\geq (\bar{\alpha}_1 - \bar{\alpha}_2) \cdot \left(1 - \frac{\left(q^T(\Lambda - \bar{\alpha}_1 I)^{-1}q\right)^2}{\|(\Lambda - \bar{\alpha}_1 I)^{-1}q\|_2^2} \cdot \frac{\|(\Lambda - \bar{\alpha}_2 I)^{-1}q\|_2 \|(\Lambda - \bar{\alpha}_1 I)^{-1}q\|_2}{q^T(\Lambda - \bar{\alpha}_1 I)^{-1}q \cdot q^T(\Lambda - \bar{\alpha}_2 I)^{-1}q}\right) \\
&= (\bar{\alpha}_1 - \bar{\alpha}_2) \cdot \left(1 - \frac{q^T(\Lambda - \bar{\alpha}_1 I)^{-1}q}{\|(\Lambda - \bar{\alpha}_1 I)^{-1}q\|_2} \cdot \frac{\|(\Lambda - \bar{\alpha}_2 I)^{-1}q\|_2}{q^T(\Lambda - \bar{\alpha}_2 I)^{-1}q}\right) \\
&= (\bar{\alpha}_1 - \bar{\alpha}_2) \cdot \left(1 - \left(-\frac{\|b\|_2}{\mu}\right) \cdot \left(-\frac{\mu}{\|b\|_2}\right)\right) \\
&= 0.
\end{aligned} \tag{23}
$$

In the above inequality, Equation (22) is repeatedly used. In the process of deriving the formula including the inequality sign, we have used the Cauchy–Schwarz inequality and $\bar{y} > 0$. Hence, we need to find out the interval that the largest solution $\bar{\alpha}$ for Equation (21) is in. Notice that

$$
\begin{aligned}
&\lim_{\bar{\alpha} \to \lambda_{\max}} F(\bar{\alpha}) = \|u_1\|_2, \\
&\lim_{\bar{\alpha} \to +\infty} F(\bar{\alpha}) = \|q\|_2.
\end{aligned}
$$

With assumption (9),

$$\|q\|_2 = \|\bar{p}_2\|_2 > \frac{\|b\|_2}{\mu}$$

and the condition we are considering, that is,

$$\|u_1\|_2 < \frac{\|b\|_2}{\mu},$$

we have

$$\lim_{\bar{\alpha} \to +\infty}\left(F(\bar{\alpha}) - \frac{\|b\|_2}{\mu}\right) > 0$$

and

$$\lim_{\bar{\alpha} \to \lambda_{\max}}\left(F(\bar{\alpha}) - \frac{\|b\|_2}{\mu}\right) < 0.$$

Thus, there exists a solution in the interval $(\lambda_{\max}, +\infty)$. Now, we consider the monotonicity of function $F(\bar{\alpha})$ in the interval $(\lambda_{\max}, +\infty)$. Rewrite $F(\bar{\alpha})$ as

$$F(\bar{\alpha}) = \frac{q^T(\bar{\alpha}I - \Lambda)^{-1}q}{\sqrt{q^T(\bar{\alpha}I - \Lambda)^{-2}q}} = \frac{\sum_{i=1}^{n-m+1} \frac{q_i^2}{\bar{\alpha} - \lambda_i}}{\sqrt{\sum_{i=1}^{n-m+1} \frac{q_i^2}{(\bar{\alpha} - \lambda_i)^2}}}.$$

Taking the derivative of $F(\bar{\alpha})$ with respect to $\bar{\alpha}$, we get

$$F'(\bar{\alpha}) = \frac{\sum_{i\neq j=1}^{n-m+1} \frac{(\lambda_i-\lambda_j)^2}{(\bar{\alpha}-\lambda_i)^3(\bar{\alpha}-\lambda_i)^3}}{\left(\sqrt{\sum_{i=1}^{n-m+1} \frac{q_i^2}{(\bar{\alpha}-\lambda_i)^2}}\right)^3} > 0, \quad \bar{\alpha} \in (\lambda_{\max}, +\infty).$$

Thus, in the optimal interval $(\lambda_{\max}, +\infty)$, $F(\bar{\alpha})$ is increasing and Equation (21) has only one solution $\bar{\alpha}_*$, which is just the optimal solution we need in this case.

Case 2. $\bar{\alpha}$ is an eigenvalue of $\bar{Q}_{22}$.

In the following, we are going to prove that it is not an optimal choice when taking $\bar{\alpha}$ as an eigenvalue of the matrix $\bar{Q}_{22}$. Without loss of generality, we assume that the KKT conditions (16) can be satisfied when $\bar{\alpha} = \lambda_j$, the jth eigenvalue of $\bar{Q}_{22}$, $1 \leq j \leq n-m+1$. Let the algebraic multiplicity of λ_j is r and introduce a permutation matrix E, so that

$$\hat{\Lambda} = E\Lambda E^T = \begin{pmatrix} \Lambda_1 & 0 \\ 0 & \lambda_j I_r \end{pmatrix}.$$

Denote

$$\tilde{x}_2 = E\hat{x}_2 = \begin{pmatrix} \chi_1 \\ \chi_2 \end{pmatrix},$$
$$\tilde{q} = Eq = \begin{pmatrix} \tilde{q}_1 \\ \tilde{q}_2 \end{pmatrix},$$

where $\chi_1, \tilde{q}_1 \in \mathbb{R}^{n-m-r+1}$ and $\chi_2, \tilde{q}_2 \in \mathbb{R}^r$. The optimal objective value $\alpha = (\tilde{x}_2^T \hat{\Lambda} \tilde{x}_2)/(\tilde{x}_2^T \tilde{x}_2)$ must satisfy the following conditions:

$$\begin{cases} (\Lambda_1 - \lambda_j I_{n-m-r+1})\chi_1 + \bar{y}\tilde{q}_1 = 0, \quad 1 \leq j \leq n-m+1, \\ \bar{y}\tilde{q}_2 = 0, \\ \tilde{q}_1^T \chi_1 + \tilde{q}_2^T \chi_2 = \|b\|_2, \\ \chi_1^T \chi_1 + \chi_2^T \chi_2 = \mu^2. \end{cases} \tag{24}$$

On account of $\bar{y} > 0$, we know that $\tilde{q}_2 = 0$. Then, we have $\bar{\alpha} \neq \lambda_{\max}$ because we have the assumption $\bar{p}_2 \notin \text{kernel}(\bar{Q}_{22} - \lambda_{\max} I)^{\perp}$ in this paper. Then, we can solve the optimal solutions $\tilde{x}_2$, $\bar{y}$ and the maximum value α from the following equations:

$$\begin{cases} \alpha(\lambda_j) = \lambda_j - \frac{\|b\|_2}{\mu^2}\bar{y}, \quad 1 \leq j \leq n-m, \\ \chi_1 = -\bar{y}(\Lambda_1 - \lambda_j I_{n-m-r+1})^{-1}\tilde{q}_1, \\ -\bar{y}\tilde{q}_1^T(\Lambda_1 - \lambda_j I_{n-m-r+1})^{-1}\tilde{q}_1 = \|b\|_2, \\ \bar{y}^2\tilde{q}_1^T(\Lambda_1 - \lambda_j I_{n-m-r+1})^{-2}\tilde{q}_1 = \mu^2 - \chi_2^T\chi_2. \end{cases}$$

Then, we have

$$\alpha(\lambda_j) = \lambda_j + \frac{\|b\|_2^2}{\mu^2} \frac{1}{\tilde{q}_1^T(\Lambda_1 - \lambda_j I_{n-m-r+1})^{-1}\tilde{q}_1}, \quad 1 \leq j \leq n-m.$$

Next, we will show that the value of $\alpha(\lambda_j)$ cannot exceed the value of $\alpha(\bar{\alpha}_*)$, where $\bar{\alpha}_*$ is the solution of Equation (21) in the interval $(\lambda_{\max}, +\infty)$. As we mentioned above, if λ_j satisfies Equations (24), there must be $\tilde{q}_2 = 0$. It follows that

$$\begin{aligned}\alpha(\bar{\alpha}_*) &= \bar{\alpha}_* + \frac{\|b\|_2^2}{\mu^2}\frac{1}{q^T(\Lambda - \bar{\alpha}_* I_{n-m-1})^{-1}q} \\ &= \bar{\alpha}_* + \frac{\|b\|_2^2}{\mu^2}\frac{1}{\tilde{q}_1^T(\Lambda_1 - \bar{\alpha}_* I_{n-m-r+1})^{-1}\tilde{q}_1}.\end{aligned}$$

Then, in order to show that $\alpha(\bar{\alpha}_*)$ is more optimal than $\alpha(\lambda_j)$, similar with the process of (23), make the difference

$$\begin{aligned}\alpha(\bar{\alpha}_*) - \alpha(\lambda_j) &= (\bar{\alpha}_* - \lambda_j) + \frac{\|b\|_2^2}{\mu^2}\left(\frac{1}{\tilde{q}_1^T(\Lambda_1 - \bar{\alpha}_* I)^{-1}\tilde{q}_1} - \frac{1}{\tilde{q}_1^T(\Lambda_1 - \lambda_j I)^{-1}\tilde{q}_1}\right) \\
&= (\bar{\alpha}_* - \lambda_j) + \frac{\|b\|_2^2}{\mu^2}\cdot\frac{\tilde{q}_1^T(\Lambda_1 - \lambda_j I)^{-1}\tilde{q}_1 - \tilde{q}_1^T(\Lambda_1 - \bar{\alpha}_* I)^{-1}\tilde{q}_1}{\tilde{q}_1^T(\Lambda_1 - \bar{\alpha}_* I)^{-1}\tilde{q}_1 \cdot \tilde{q}_1^T(\Lambda_1 - \lambda_j I)^{-1}\tilde{q}_1} \\
&= (\bar{\alpha}_* - \lambda_j) + \frac{\|b\|_2^2}{\mu^2}\cdot\frac{\tilde{q}_1^T\left[(\Lambda_1 - \bar{\alpha}_*) - (\Lambda_1 - \lambda_j)\right](\Lambda_1 - \lambda_j I)^{-1}(\Lambda_1 - \bar{\alpha}_* I)^{-1}\tilde{q}_1}{\tilde{q}_1^T(\Lambda_1 - \bar{\alpha}_* I)^{-1}\tilde{q}_1 \cdot \tilde{q}_1^T(\Lambda_1 - \lambda_j I)^{-1}\tilde{q}_1} \\
&= (\bar{\alpha}_* - \lambda_j)\cdot\left(1 - \frac{\|b\|_2^2}{\mu^2}\cdot\frac{\tilde{q}_1^T(\Lambda_1 - \lambda_j I)^{-1}(\Lambda_1 - \bar{\alpha}_* I)^{-1}\tilde{q}_1}{\tilde{q}_1^T(\Lambda_1 - \bar{\alpha}_* I)^{-1}\tilde{q}_1 \cdot \tilde{q}_1^T(\Lambda_1 - \lambda_j I)^{-1}\tilde{q}_1}\right) \\
&\geq (\bar{\alpha}_* - \lambda_j)\cdot\left(1 - \frac{\left(\tilde{q}_1^T(\Lambda_1 - \bar{\alpha}_* I)^{-1}\tilde{q}_1\right)^2}{\|(\Lambda_1 - \bar{\alpha}_* I)^{-1}\tilde{q}_1\|_2^2}\cdot\frac{\|(\Lambda_1 - \lambda_j I)^{-1}\tilde{q}_1\|_2\|(\Lambda_1 - \bar{\alpha}_* I)^{-1}\tilde{q}_1\|_2}{\tilde{q}_1^T(\Lambda_1 - \bar{\alpha}_* I)^{-1}\tilde{q}_1 \cdot \tilde{q}_1^T(\Lambda_1 - \lambda_j I)^{-1}\tilde{q}_1}\right) \\
&= (\bar{\alpha}_* - \lambda_j)\cdot\left(1 - \frac{\tilde{q}_1^T(\Lambda_1 - \bar{\alpha}_* I)^{-1}\tilde{q}_1}{\|(\Lambda_1 - \bar{\alpha}_* I)^{-1}\tilde{q}_1\|_2}\cdot\frac{\|(\Lambda_1 - \lambda_j I)^{-1}\tilde{q}_1\|_2}{\tilde{q}_1^T(\Lambda_1 - \lambda_j I)^{-1}\tilde{q}_1}\right) \\
&= (\bar{\alpha}_* - \lambda_j)\cdot\left(1 - \left(-\frac{\|b\|_2}{\mu}\right)\cdot\left(-\frac{\sqrt{\mu^2 - \chi_2^T\chi_2}}{\|b\|_2}\right)\right) \\
&> (\bar{\alpha}_* - \lambda_j)\left(1 - \left(-\frac{\|b\|_2}{\mu}\right)\cdot\left(-\frac{\mu}{\|b\|_2}\right)\right) \\
&= 0.\end{aligned}$$

In the above inequality, we have used the equations

$$-\frac{\tilde{q}_1^T(\Lambda_1 - \bar{\alpha}_* I)^{-1}\tilde{q}_1}{\|(\Lambda_1 - \bar{\alpha}_* I)^{-1}\tilde{q}_1\|_2} = \frac{\|b\|_2}{\mu}, -\frac{\tilde{q}_1^T(\Lambda_1 - \lambda_j I)^{-1}\tilde{q}_1}{\|(\Lambda_1 - \lambda_j I)^{-1}\tilde{q}_1\|_2} = \frac{\|b\|_2}{\sqrt{\mu^2 - \chi_2^T\chi_2}}.$$

From the above analysis of Case 1 and Case 2, we know that, under the case $\|u_1\|_2 < \|b\|_2/\mu$, among all the possible values that satisfy the optimality conditions, $\alpha = (\bar{x}_2^T\bar{Q}_{22}\bar{x}_2)/(\bar{x}_2^T\bar{x}_2)$ is the largest when $\bar{\alpha} = \bar{\alpha}_*$, where $\bar{\alpha}_*$ is the solution of Equation (21) in the interval $(\lambda_{\max}, +\infty)$.

Combining the two subsections, now we can make a conclusion on the solution of the optimization problem (3).

Method 1. Referring to the notations in the above analysis, we have the following two possibilities of the solutions of optimization problem (3):

$$\max \quad \frac{x^TQx}{x^Tx} \quad \text{s.t.} \quad Px = b, \quad \|x\|_2 \leq \mu.$$

(I) If

$$\|b\|_2 \leq \|u_1\|_2\,\mu$$

holds, the objective function $(x^TQx)/(x^Tx)$ reaches its maximum value $\lambda_{\max}$, which is the largest eigenvalue of the matrix $\bar{Q}_{22}$ at

$$x = \bar{W}\begin{pmatrix} 0 \\ \frac{\|b\|_2}{u_1^T u_1} u_1 \end{pmatrix}.$$

(II) Else, if

$$\|b\|_2 > \|u_1\|_2\, \mu$$

holds, the objective function $(x^TQx)/(x^Tx)$ reaches its maximum value at

$$x = \bar{W}\begin{pmatrix} 0 \\ \frac{\|b\|_2}{\bar{p}_2^T(\bar{Q}_{22}-\bar{\alpha}_* I)^{-1}\bar{p}_2}(\bar{Q}_{22} - \bar{\alpha}_* I)^{-1}\bar{p}_2 \end{pmatrix},$$

where $\bar{\alpha}_*$ is the solution of the function

$$-\frac{\bar{p}_2^T(\bar{Q}_{22} - \bar{\alpha}I)^{-1}\bar{p}_2}{\sqrt{\bar{p}_2^T(\bar{Q}_{22} - \bar{\alpha}I)^{-2}\bar{p}_2}} - \frac{\|b\|_2}{\mu} = 0$$

in the interval $(\lambda_{\max}, +\infty)$.

4 | EXPERIMENT RESULTS

In the following, we demonstrate the performance of the proposed numerical method to solve the BSMAC formulation (3).

In order to form the example that includes the special case where the leading eigenvector of $\bar{Q}_{22}$ and the vector $\bar{p}_2$ are nearly orthogonal (i.e., the leading eigenvector of Q and the range of P^T are nearly orthogonal), we first generate random matrices $P \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$, in which the entries are drawn from the standard uniform distribution on the open interval $(0, 1)$. By making transformations on the matrix P and the vector b as described in Section 2, we obtain the vector $\bar{p}_2$. Then, we generate the leading eigenvector of matrix Q_{22} by adding noise of magnitude σ to a randomly generated orthogonal vector of $\bar{p}_2$. When the noise parameter σ is small, the special case, Case II in Method 1, appears. By choosing the eigenvalues of $\bar{Q}_{22}$ to be evenly distributed in the interval $[0, 10]$, we can form the matrix $\bar{Q}_{22}$. Using (7) and letting the matrices $\bar{Q}_{11} = \bar{Q}_{12} = \bar{Q}_{21} = 0$, we obtain the compatibility matrix $Q \in \mathbb{R}^{n\times n}$. Denote the node number of the graph to be l; we get $m = 2l$ and $n = l^2$. Due to the binary nature of the true solution x, we choose the upper bound of the solution norm $\|x\|_2$ as $\mu = \sqrt{l}$.

We use some figures to compare the solutions of the proposed numerical method that solves the BSMAC formulation and the numerical method that solves the SMAC formulation.[6] By choosing different noise parameter σ, we obtain different graph matching problems with different orthogonal levels of the leading eigenvector of $\bar{Q}_{22}$ and the vector $\bar{p}_2$. In Figure 1, the optimal solution norm $\|x\|_2$ versus the noise parameter σ of the two methods is presented with varying node number $l = 10, 30, 50, 70$. It shows that the solution norm of SMAC is large with small noise parameter σ and even tends to infinity when σ is close to 0, which indicates that we cannot obtain any meaningful solution for the graph matching problem by the SMAC formulation. We can see the solution norm of BSMAC can never exceed the upper bound μ. The ordinate value of the flat part of the BSMAC line is μ, which occurs when the condition $\|b\|_2 > \|u_1\|_2\, \mu$ is satisfied, that is, the special case where the leading eigenvector of Q and the range of P^T are nearly, but not exactly, orthogonal.

In order to see the differences of the optimal values α of the formulations BSMAC and SMAC, Figure 2 shows the optimal values α of the objective function solved by the two methods with varying optimal solution norms of SMAC. We use the solution norm of SMAC as abscissa in order to see the solution performance of BSMAC better. In fact, the solution norm of the BSMAC method cannot exceed μ, which is marked in the abscissa axis. We can see the optimal solutions of the BSMAC formulation is equal to those of SMAC when the solution norm is less than μ. When the solution norm is larger than μ, because of the upper-bound constraint, the optimal solutions of BSMAC are slightly less than those of SMAC, which are within the extent that it is possible to do so, and we obtain meaningful solutions.

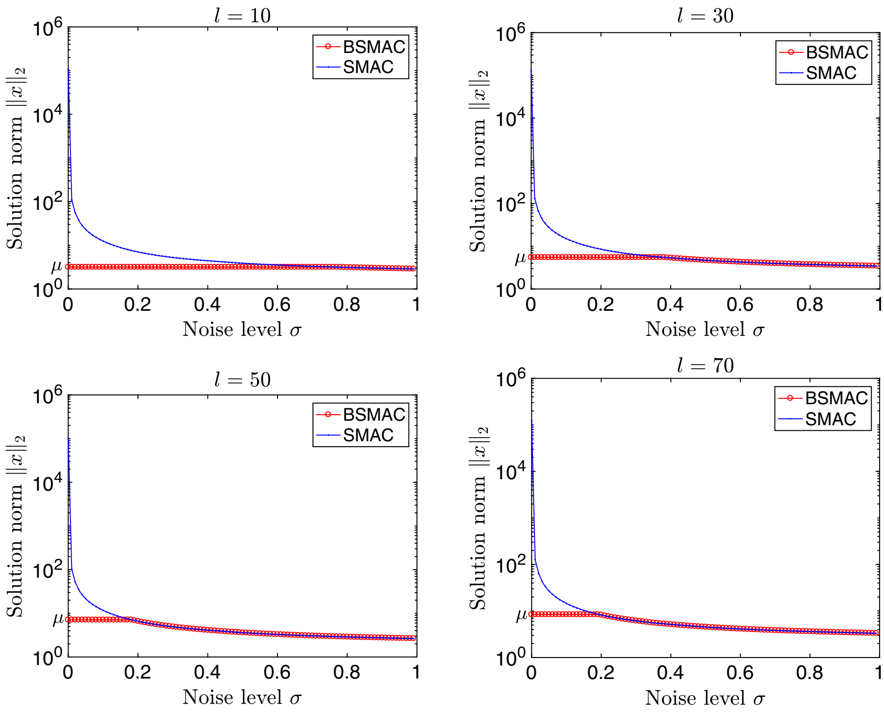

FIGURE 1 Comparison of the solution norm of the two methods for bounded spectral matching with affine constraint (BSMAC) and spectral matching with affine constraint (SMAC) formulations with varying noise parameter σ

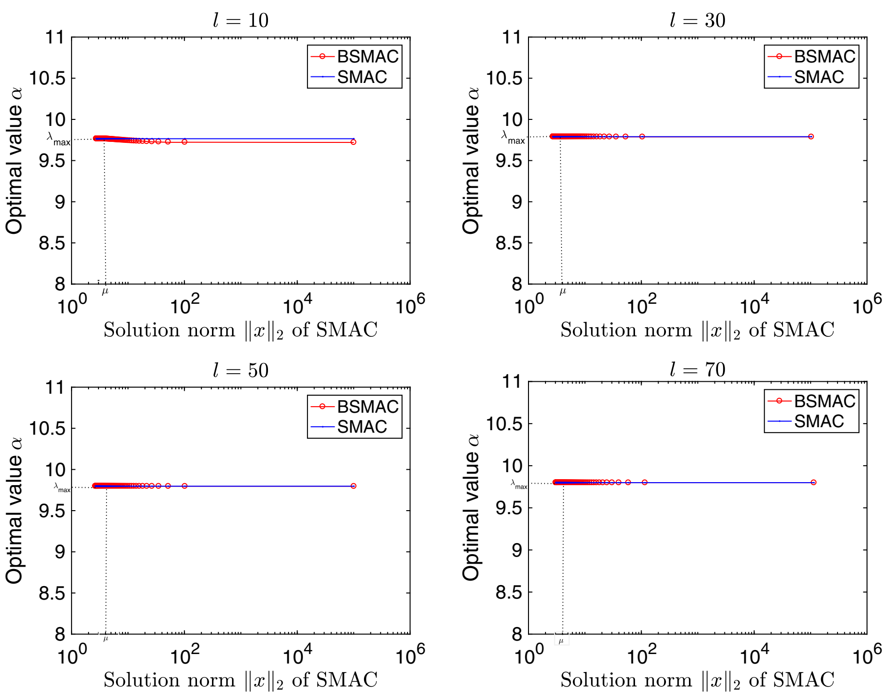

FIGURE 2 Comparison of the optimal values of the objective function of the two methods for bounded spectral matching with affine constraint (BSMAC) and spectral matching with affine constraint (SMAC) formulations with respect to the solution norm of SMAC

5 | CONCLUSION

Based on the SMAC formulation, we presented a new BSMAC formulation for the graph matching problem in this work. We presented an efficient numerical method to solve the BSMAC formulation. Experiment results have verified the feasibility of the proposed formulation and the numerical method.

ACKNOWLEDGEMENTS

The authors are very much indebted to the referees for their constructive and valuable comments and suggestions that greatly improved the original manuscript of this paper. This work does not have any conflicts of interest. The work of the first author is funded by the China Scholarship Council (File No. 201606180087) and by the National Natural Science Foundation of China (Grants 11471150, 11401281 and 11801242). Her work was performed during her visit at the University of California, Berkeley, USA.

ORCID

Yan Dou https://orcid.org/0000-0001-7430-0450

REFERENCES

1. Bolles RC, Cain RA. Recognising and locating partially visible objects: The local-feature-focus method. In: Robot vision. Berlin, Germany: Springer, 1983; p. 43–82.
2. Conte D, Foggia P, Sansone C, Vento M. Thirty years of graph matching in pattern recognition. Int J Pattern Recognit Artif Intell. 2004;18(3)265–298.
3. Horaud R, Skordas T. Stereo correspondence through feature grouping and maximal cliques. IEEE Trans Pattern Anal Mach Intell. 1989;11(11)1168–1180.
4. Stella XY, Shi J. Grouping with bias. Proceedings of the Advances in Neural Information Processing Systems 14 (NIPS 2001); 2001 Dec 3–8; Vancouver, Canada. Cambridge, MA: MIT Press; 2001. p. 1327–1334.
5. Sunderrajan S, Manjunath BS. Context-aware hypergraph modeling for re-identification and summarization. IEEE Trans Multimed. 2016;18(1)51–63.
6. Cour T, Srinivasan P, Shi J. Balanced graph matching. Proceedings of the Advances in Neural Information Processing Systems 19 (NIPS 2006); 2006 Dec 4–7; Vancouver, Canada. Cambridge, MA: MIT Press; 2006. p. 313–320.
7. Leordeanu M, Hebert M. A spectral technique for correspondence problems using pairwise constraints. Proceedings of the Tenth IEEE International Conference on Computer Vision (ICCV'05); 2005 Oct 17–20; Beijing, China. Washington, DC: IEEE Computer Society; 2005. p. 1482–1489.
8. Leordeanu M, Hebert M, Sukthankar R. An integer projected fixed point method for graph matching and map inference. Proceedings of the Advances in Neural Information Processing Systems 22 (NIPS 2009); 2009 Dec 7–10; Vancouver, Canada. New York, NY: Curran Associates; 2009. p. 1114–1122.
9. Yang X, Liu Z-Y. Adaptive graph matching. IEEE Trans Cybern. 2017;PP(99)1–14.
10. Cho M, Lee J, Lee K. Reweighted random walks for graph matching. Proceedings of the European Conference on Computer Vision; 2010 Sep 5–11; Crete, Greece. Berlin, Germany: Springer; 2010. p. 492–505.
11. Egozi A, Keller Y, Guterman H. A probabilistic approach to spectral graph matching. IEEE Trans Pattern Anal Mach Intell. 2013;35(1)18–27.
12. Leordeanu M, Sukthankar R, Hebert M. Unsupervised learning for graph matching. Int J Comput Vis. 2012;96(1)28–45.
13. Gold S, Rangarajan A. A graduated assignment algorithm for graph matching. IEEE Trans Pattern Anal Mach Intell. 1996;18(4)377–388.
14. Schellewald C, Schnörr C. Probabilistic subgraph matching based on convex relaxation. Proceedings of the International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition; 2005 Nov 9–11; St. Augustine, FL. Berlin, Germany: Springer; 2005. p. 171–186.
15. Torr PH. Solving Markov random fields using semi definite programming. Proceedings of the Ninth International Workshop on Artificial Intelligence and Statistics (AISTATS 2003); 2003 Jan 3–6; Key West, FL; 2003.
16. Cour T. Affinely constrained Rayleigh quotients. Available from: http://timotheecour.com/papers/supplement_nips2006.pdf
17. Gander W, Golub GH, von Matt U. A constrained eigenvalue problem. Linear Algebra Appl. 1989;114–115:815–839.

How to cite this article: Dou Y, Gu M, Chandrasekaran S. An efficient algorithm for bounded spectral matching with affine constraint. *Numer Linear Algebra Appl.* 2019;e2264. https://doi.org/10.1002/nla.2264