EXACT RECOVERY BY SEMIDEFINITE PROGRAMMING IN THE BINARY STOCHASTIC BLOCK MODEL WITH PARTIALLY REVEALED SIDE INFORMATION

Mohammad Esmaeili, Hussein Saad, and Aria Nosratinia

Department of Electrical Engineering, The University of Texas at Dallas
Email: {*esmaeili, hussein.saad, aria* }@*utdallas.edu*

ABSTRACT

We propose a semidefinite programming (SDP) approach to community detection in graphs in the presence of additional non-graphical side information, and analyze the corresponding exact recovery threshold. The community detection problem is considered in the context of the binary symmetric Stochastic Block Model (SBM), and the side information is in the form of partially revealed labels with erasure probability ϵ. Our results show that the semidefinite programming relaxation of the maximum likelihood estimator can achieve exact recovery down to the optimal threshold. The theoretical findings of this paper are validated via simulations on finite synthetic data-sets, showing that the asymptotic results of this paper can also shed light on the performance at finite n.

Index Terms— Exact Recovery, Semidefinite Programming (SDP), Binary Stochastic Block Model, Side Information, Community Detection.

1. INTRODUCTION

Many of the available data nowadays are inherently related to graphs, such as social networks, networks representing protein interactions and citation networks. Also many problems in signal processing can be gainfully represented by graphical structures such as hierarchical clustering in images and video [1]. In this paper, we consider the problem of community detection via the observation of a connectivity graph, which is a facet of the broader field of signal processing on graphs [2, 3].

The Stochastic Block Model (SBM) [4–6] is a popular statistical model for community detection. This paper considers the binary symmetric SBM, which consists of n nodes belonging to two equal sized communities. The nodes have identical, independent and uniform distribution. If two nodes belong to the same community, there exists an edge between them with probability p; otherwise there is an edge between them with probability q. The purpose of community detection is to recover the labels (communities) based on observing the graph edges.

Most of the literature on community detection has concentrated only on graphical observations, see [7–9] and references therein. However, in many practical applications, non-graphical relevant information is available that can aid the inference. For example, social networks such as Facebook and Twitter have access to other information other than the graph edges such as date of birth, nationality, and school. The notion of side information has been introduced and studied before in the literature, for a survey, see [10–14] and references therein.

Semidefinite Programming (SDP) is a computationally efficient convex optimization technique that has shown its utility in solving signal processing problems [15, 16]. In the context of community detection, SDP was introduced in [17], where it was used for solving a minimum bisection problem, obtaining a sufficient condition that is not optimal. In [18], a SDP relaxation was considered for a maximum bisection problem. For the binary symmetric SBM, [19] showed that the semidefinite programming relaxation of Maximum likelihood (ML) can achieve the optimal exact recovery threshold with high probability. These results were later extended to more general models in [20]. The problem of solving graph inference problems with non-graphical side information is one that calls for efficient algorithms, but to date it has not been formulated or analyzed via convex optimization.

In this paper, we show the asymptotic optimality of SDP relaxation of maximum likelihood solution of community detection *in the presence of non-graphical side information*. In particular, we consider the binary symmetric SBM with side information in the asymptotic regime of $p = a\frac{\log n}{n}$ and $q = b\frac{\log n}{n}$, with $a \geq b > 0$. We study side information in the form of partially revealed labels with the erasure probability $\epsilon \in (0, 1)$. We show that

- When $\log \epsilon = o(\log n)$, then SDP achieves exact recovery whenever $(\sqrt{a} - \sqrt{b})^2 > 2$ which matches the optimal threshold characterized in [10].
- When $\log \epsilon = -(\beta + o(1)) \log n$, and $\beta > 0$, then SDP achieves exact recovery whenever $(\sqrt{a} - \sqrt{b})^2 + 2\beta > 2$, which matches the optimal threshold characterized in [10].

The achievability results of this paper complement the converse obtained in [10, Theorem 4], where it was shown that exact recovery is impossible if:

$$\begin{cases}(\sqrt{a}-\sqrt{b})^2 < 2 & \text{when } \log\epsilon = o(\log n)\\ (\sqrt{a}-\sqrt{b})^2 + 2\beta < 2 & \text{when } \log\epsilon = -(\beta + o(1))\log n\end{cases}.$$

Notation: The matrices $\mathbf{I}$ and $\mathbf{J}$ are defined as identity matrix and all-one matrix, respectively. If X is a positive semidefinite matrix, it is shown by $X \succeq 0$ and if all the entries of X are non-negative it is shown by $X \geq 0$. For any matrices X and Y, $\|X\|$ denote the spectral norm of X and $\langle X, Y\rangle$ denote the inner product of matrices X and Y.

2. DETECTION VIA SDP

This paper considers the binary symmetric stochastic block model with side information. The number of nodes in the graph is denoted by n and the community labels are denoted by -1 and 1. The nodes labels have a uniform, equal and independent distribution. If a pair of nodes belong to the same community, an edge between them is drawn with probability $p = a\frac{\log n}{n}$, otherwise an edge between them is drawn with probability $q = b\frac{\log n}{n}$, where $a \geq b > 0$.

A scalar side information is observed independently for each node, which is the true label (for non-erased case) with a probability equal to $1-\epsilon$, or zero (for erased case) with a probability equal to ϵ, while $\epsilon \in (0,1)$. In this paper, this model of side information is called partially revealed labels side information. We denote the adjacency matrix of the observed graph by G, the vector of nodes' true assignment by X^* and the vector of nodes' side information by Y. The goal of the exact recovery is to recover X^* upon observing G and Y.

The log-likelihood function of the graph and side information can be written as

$$\log \mathbb{P}(G, Y|X) = \log \mathbb{P}(G|X) + \log \mathbb{P}(Y|X).$$

Then $\log \mathbb{P}(G|X)$ can be calculated as [10]:

$$\log \mathbb{P}(G|X) = \frac{1}{2} T_1 X^T G X + R(n,p,q), \tag{1}$$

where $R(n,p,q)$ is a constant and $T_1 = \log\left(\frac{p(1-q)}{q(1-p)}\right)$.

For the partially revealed side information, the vector of side information Y is a $n \times 1$ vector whose entries belong to the set $\{1,-1,0\}$. The log-likelihood function $\mathbb{P}(Y|X)$ can be written as

$$\log \mathbb{P}(Y|X) = Y^T Y \log\left(\frac{1-\epsilon}{\epsilon}\right) + n\log(\epsilon), \tag{2}$$

where $Y^T Y$ is the number of non-erased elements of Y. Combining (1) and (2), the ML detector rule can be formulated as

$$\begin{aligned}\hat{X} =& \arg\max_X X^T G X\\ &\text{subject to} \quad X \in \{\pm 1\}^n\\ &\qquad\qquad\quad X^T \mathbf{1} = 0\\ &\qquad\qquad\quad X^T Y = Y^T Y.\end{aligned} \tag{3}$$

Due to computational complexity of solving (3), we consider its convex relaxation. This is a common relaxation for many community detection problems [7, 21]. Define $Z \triangleq XX^T$. Then $Z_{ii} = 1$ for all $i \in [n]$ comes from the constraint $X \in \{\pm 1\}^n$. The constraint $X^T\mathbf{1} = 0$ is equivalent to $\langle \mathbf{J}, Z\rangle = 0$. Define $W \triangleq YY^T$. Then $X^TY = \pm Y^TY$ is equivalent to $\langle Z, W\rangle = (Y^TY)^2$. By relaxing the rank-one constraint for $Z \triangleq XX^T$, we obtain the following semidefinite programming:

$$\begin{aligned}\widehat{Z}_{SDP} =& \arg\max_Z \langle Z, G\rangle\\ &\text{subject to} \quad Z \succeq 0\\ &\qquad\qquad\quad Z_{ii} = 1, \quad i \in [n]\\ &\qquad\qquad\quad \langle Z, W\rangle = (Y^TY)^2\\ &\qquad\qquad\quad \langle Z, \mathbf{J}\rangle = 0.\end{aligned} \tag{4}$$

Let $Z^* = X^* X^{*T}$ correspond to the true nodes' labels and $\mathcal{Z}_n \triangleq \left\{XX^T : X \in \{\pm 1\}^n, X^T\mathbf{1} = 0, X^TY = Y^TY\right\}$.

Theorem 1. *Under the binary symmetric stochastic block model and partial revealed labels side information, if*

$$\begin{cases}(\sqrt{a}-\sqrt{b})^2 > 2 & \text{when } \log\epsilon = o(\log n)\\ (\sqrt{a}-\sqrt{b})^2 + 2\beta > 2 & \text{when } \log\epsilon = -(\beta + o(1))\log n\end{cases},$$

then as $n \to \infty$, $\min_{Z^* \in \mathcal{Z}_n} \mathbb{P}(\widehat{Z}_{SDP} = Z^*) \geq 1 - o(1)$.

Proof. We start with the following Lemma:

Lemma 1. *Suppose there exists $D^* = \operatorname{diag}(d_i^*) \geq 0$, $\lambda^* \in \mathbb{R}$, and $\mu^* \in \mathbb{R}$ such that $S^* = D^* - G + \lambda^* \boldsymbol{J} + \mu^* W$ satisfies $S^* \succeq 0$, $\lambda_2(S^*) \geq 0$, and $S^* X^* = 0$. Then $\widehat{Z}_{SDP} = Z^*$ is the unique solution to (4).*

Proof. The Lagrangian function of (4) is given by:

$$\begin{aligned}L(Z, S, D, \lambda, \mu) =& \langle G, Z\rangle + \langle S, Z\rangle - \langle D, Z - \mathbf{I}\rangle\\ &- \lambda\langle \mathbf{J}, Z\rangle - \mu\left(\langle W, Z\rangle - (Y^TY)^2\right),\end{aligned}$$

where $S \succeq 0$, $D = \operatorname{diag}(d_i)$, $\lambda \in \mathbb{R}$, and $\mu \in \mathbb{R}$ are Lagrangian multipliers. For any Z that satisfies the constraints in (4),

$$\begin{aligned}\langle G, Z\rangle &\overset{(a)}{\leq} L(Z, S^*, D^*, \lambda^*, \mu^*)\\ &= \langle D^*, \mathbf{I}\rangle + \mu^*(Y^TY)^2\\ &\overset{(b)}{=} \langle D^*, Z^*\rangle + \mu^*(Y^TY)^2\\ &= \langle G + S^* - \lambda^*\mathbf{J} - \mu W, Z^*\rangle + \mu^*(Y^TY)^2\\ &\overset{(c)}{=} \langle G, Z^*\rangle,\end{aligned}$$

where (a) holds because $\langle S^*, Z\rangle \geq 0$, (b) holds because $Z_{ii} = 1$ for all $i \in [n]$, and (c) holds because $\langle S^*, Z^*\rangle = X^{*T} S^* X^* = 0$, $\langle \mathbf{J}, Z^*\rangle = 0$ and $\langle W, Z^*\rangle = (Y^T Y)^2$. Therefore, Z^* is an optimal solution. Now, we will prove the uniqueness of the optimal solution. To this end, assume $\widetilde{Z}$ is an optimal solution. Then

$$\begin{aligned}\langle S^*, \widetilde{Z}\rangle &= \langle D^* - G + \lambda^* \mathbf{J} + \mu^* W, \widetilde{Z}\rangle \\ &= \langle D^*, \widetilde{Z}\rangle - \langle G, \widetilde{Z}\rangle + \lambda^* \langle \mathbf{J}, \widetilde{Z}\rangle + \mu^* \langle W, \widetilde{Z}\rangle \\ &\overset{(a)}{=} \langle D^*, Z^*\rangle - \langle G, Z^*\rangle + \lambda^* \langle \mathbf{J}, Z^*\rangle + \mu^* \langle W, Z^*\rangle \\ &= \langle D^* - G + \lambda^* \mathbf{J} + \mu^* W, Z^*\rangle \\ &= \langle S^*, Z^*\rangle = 0,\end{aligned}$$

where (a) holds because $\langle \mathbf{J}, Z^*\rangle = \langle \mathbf{J}, \widetilde{Z}\rangle = 0$, $\langle W, Z^*\rangle = \langle W, \widetilde{Z}\rangle = (Y^T Y)^2$, $\langle G, Z^*\rangle = \langle G, \widetilde{Z}\rangle$, and $Z^*_{ii} = \widetilde{Z}_{ii} = 1$ for all $i \in [n]$. Since $\widetilde{Z} \succeq 0$ and $S^* \succeq 0$ while its second smallest eigenvalue $\lambda_2(S^*)$ is positive, we can say $\widetilde{Z}$ must be a multiple of Z^*. Also, since $\widetilde{Z}_{ii} = Z^*_{ii} = 1$ for all $i \in [n]$, we have $\widetilde{Z} = Z^*$. □

It now suffices to show that S^* satisfies the conditions in Lemma 1 with probability $1 - o(1)$. In view of Lemma 1,

$$\begin{aligned}S^* X^* &= D^* X^* - G X^* + \lambda^* \mathbf{J} X^* + \mu^* W X^* \\ &= D^* X^* - G X^* + \mu^* W X^*.\end{aligned}$$

Thus, in order to satisfy the condition $S^* X^* = 0$, we need:

$$D^* X^* = G X^* - \mu^* W X^*.$$

By expanding $D^* X^*$ we get:

$$d_i^* x_i^* = \sum_{j=1}^{n} G_{ij} x_j^* - \mu^* \sum_{j=1}^{n} y_i y_j x_j^*.$$

By multiplying the both sides of this equation by x_i^* and knowing that $G_{ii} = 0$ for all $i \in [n]$, d_i^* is obtained as

$$d_i^* = \sum_{j=1}^{n} G_{ij} x_j^* x_i^* - \mu^* y_i x_i^* \sum_{j=1}^{n} y_j x_j^*.$$

where satisfies $S^* X^* = 0$.

It remains to show $S^* \succeq 0$ and $\lambda_2(S^*) > 0$ with probability at least $1 - o(1)$. In other words, it suffices to show that

$$\mathbb{P}\left\{\inf_{V \perp X^*, \|V\|=1} V^T S^* V > 0\right\} \geq 1 - o(1).$$

Since $\mathbb{E}[G] = \frac{p-q}{2} X^* X^{*T} + \frac{p+q}{2}\mathbf{J} - p\mathbf{I}$, it follows that for any V such that $V^T X^* = 0$ and $\|V\| = 1$,

$$\begin{aligned}V^T S^* V =& V^T D^* V - \frac{p+q}{2} V^T \mathbf{J} V + p V^T \mathbf{I} V \\ & - V^T (G - \mathbb{E}[G]) V + \lambda^* V^T \mathbf{J} V + \mu^* V^T W V \\ =& V^T D^* V + \left(\lambda^* - \frac{p+q}{2}\right) V^* \mathbf{J} V + p \\ & - V^T (G - \mathbb{E}[G]) V + \mu^* V^T W V.\end{aligned}$$

Let $\lambda^* \geq \frac{p+q}{2}$. Since $V^T D^* V \geq \min_{i \in [n]} d_i^*$ and $-V^T (G - \mathbb{E}[G]) V \geq -\|G - \mathbb{E}[G]\|$, it follows that

$$V^T S^* V \geq \min_{i \in [n]} d_i^* + p - \|G - \mathbb{E}[G]\| + \mu^* V^T W V.$$

Lemma 2. *For any $c > 0$, there exists $c^{'} > 0$ such that for any $n \geq 1$, $\|G - \mathbb{E}[G]\| \leq c^{'} \sqrt{\log n}$ with probability at least $1 - n^{-c}$ [19, Thoerem 5].*

Lemma 3. *Let $\kappa = \sqrt{\log n}$, then $\mathbb{P}\left\{V^T W V \geq \kappa\right\} \leq \frac{1-\epsilon}{\kappa}$.*

Proof. Define $h_i \triangleq v_i y_i$ for $i \in [n]$ and $Z \triangleq \sum_{i=1}^{n} h_i$. Then $V^T W V = \left(\sum_{i=1}^{n} v_i y_i\right)^2 = \left(\sum_{i=1}^{n} h_i\right)^2 = Z^2$. Since $||V|| = 1$ and the random variables $\{h_i\}$ are i.i.d. with $h_i \in \{v_i, -v_i, 0\}$ with probabilities $\{\frac{1-\epsilon}{2}, \frac{1-\epsilon}{2}, \epsilon\}$, it follows that $\mathbb{E}[Z] = 0$ and $\mathrm{Var}(Z) = (1 - \epsilon)$.

Then using Chebyshev's inequality for any positive κ:

$$\begin{aligned}\mathbb{P}\left(V^T W V \geq \kappa\right) &= \mathbb{P}\left((Z - \mathbb{E}[Z])^2 \geq \kappa\right) \\ &= \mathbb{P}\left(|Z - \mathbb{E}[Z]| \geq \sqrt{\kappa}\right) \\ &\leq \frac{\mathrm{Var}(Z)}{\kappa}.\end{aligned}$$

Now let $\kappa = \sqrt{\log n}$, it follows that

$$\mathbb{P}\left\{V^T W V \geq \sqrt{\log n}\right\} \leq \frac{1 - \epsilon}{\sqrt{\log n}}.$$

□

Lemma 4. *Define $\delta \triangleq \frac{\log n}{\log \log n} = o(\log n)$, then:*

$$\mathbb{P}(d_i^* \leq \delta) \leq \epsilon n^{-\frac{1}{2}\left(\sqrt{a} - \sqrt{b}\right)^2 + o(1)} + (1 - \epsilon)\epsilon^n.$$

Proof. Due to symmetry, $\mathbb{P}(d_i^* \leq \delta)$ can be written as

$$\begin{aligned}\mathbb{P}(d_i^* \leq \delta) = & \mathbb{P}\left(I_i - O_i \leq \delta\right)\epsilon \\ & + \mathbb{P}\left(I_i - O_i - \mu^* Z_i \leq \delta + \mu^*\right)(1 - \epsilon),\end{aligned} \tag{5}$$

where $Z_i = \sum_{j=1, j \neq i}^{n} y_j x_j^*$ and $I_i - O_i = \sum_{j=1}^{n} G_{ij} x_j^* x_i^*$. Kindly note that $Z_i \sim \mathrm{Binom}(n-1, 1-\epsilon)$, $I_i \sim \mathrm{Binom}(\frac{n}{2} - 1, p)$ and $O_i \sim \mathrm{Binom}(\frac{n}{2}, q)$. The probability $\mathbb{P}\left(I_i - O_i \leq \delta\right)$ can be bounded by applying the Chernoff's inequality as [10]:

$$\mathbb{P}(I_i - O_i \leq \delta) \leq n^{-\frac{1}{2}\left((\sqrt{a} - \sqrt{b})^2 + o(1)\right)}. \tag{6}$$

Let $\mu^* < 0$ and a constant. Similarly, we find an appropriate bound for the probability $\mathbb{P}\left(I_i - O_i - \mu^* Z_i \leq \delta + \mu^*\right)$ by using Chernoff's inequality:

$$\mathbb{P}(I_i - O_i - \mu^* Z_i \leq \delta + \mu^*) \leq \min_{t < 0}\left(e^{-t(\delta + \mu^*)} M(t)\right),$$

where $M(t)$ is the moment generating function of $I_i - O_i - \mu^* Z_i$. Then

$$\begin{aligned}&\mathbb{P}\left(I_i - O_i - \mu^* Z_i \leq \delta + \mu^*\right) \\ &\leq \min_{t<0} e^{-t(\delta + \mu^*)} e^{\frac{\log n}{2}\left(a(e^t - 1) + b(e^{-t} - 1)\right) + n \log \epsilon + \frac{1-\epsilon}{\epsilon} n e^{-\mu^* t}} \\ &= \min_{t<0} e^{n\left(\log \epsilon + \frac{1-\epsilon}{\epsilon} e^{-\mu^* t} + o(1)\right)}.\end{aligned}$$

Since $\mu^* < 0$ and is a constant, the minimum occurs at $t = -\infty$. Therefore,

$$\mathbb{P}\left(I_i - O_i - \mu^* Z_i \leq \delta + \mu^*\right) \leq e^{n \log \epsilon} = \epsilon^n. \tag{7}$$

The proof is completed by substituting (6) and (7) in (5). $\square$

Choose $\mu^* < 0$, then in view of Lemmas 2 and 3, with probability at least $1 - o(1)$,

$$V^T S^* V \geq \min_{i \in [n]} d_i^* + p + (\mu^* - c^{'})\sqrt{\log n}. \tag{8}$$

Now, we divide the analysis into two scenarios:

- $\log(\epsilon) = o(\log(n))$: It follows that

  $$\mathbb{P}(d_i^* \leq \delta) \leq n^{-\frac{1}{2}\left(\sqrt{a}-\sqrt{b}\right)^2 + o(1)}.$$

  Using the union bound, it follows that

  $$\mathbb{P}(\min_{i \in [n]} d_i^* \geq \frac{\log n}{\log \log n}) \geq 1 - n^{1-\frac{1}{2}\left(\sqrt{a}-\sqrt{b}\right)^2 + o(1)}.$$

  Assume $(\sqrt{a} - \sqrt{b})^2 > 2$, it follows that $\min_{i \in [n]} d_i^* \geq \frac{\log n}{\log \log n}$ holds with probability at least $1 - o(1)$. Combining this result with (8), we get that if $(\sqrt{a} - \sqrt{b})^2 > 2$, then with probability at least $1 - o(1)$,

  $$V^T S^* V \geq \frac{\log n}{\log \log n} + p + (\mu^* - c^{'})\sqrt{\log n} > 0,$$

  which proves the first case of Theorem 1.

- $\log(\epsilon) = -\beta \log(n) + o(\log(n))$: It follows that

  $$\mathbb{P}(d_i^* \leq \delta) \leq n^{-\frac{1}{2}\left(\sqrt{a}-\sqrt{b}\right)^2 - \beta + o(1)}.$$

  Using the union bound, it follows that

  $$\mathbb{P}(\min_{i \in [n]} d_i^* \geq \frac{\log n}{\log \log n}) \geq 1 - n^{1-\frac{1}{2}\left(\sqrt{a}-\sqrt{b}\right)^2 - \beta + o(1)}.$$

  Assume $(\sqrt{a} - \sqrt{b})^2 + 2\beta > 2$, it follows that $\min_{i \in [n]} d_i^* \geq \frac{\log n}{\log \log n}$ holds with probability at least $1 - o(1)$. Combining this result with (8), we get that if $(\sqrt{a} - \sqrt{b})^2 + 2\beta > 2$, then with probability at least $1 - o(1)$,

  $$V^T S^* V \geq \frac{\log n}{\log \log n} + p + (\mu^* - c^{'})\sqrt{\log n} > 0,$$

  which proves the second case of Theorem 1.

$\square$

3. NUMERICAL RESULTS

This section explores the relevance of asymptotic results, obtained in this paper, to finite data. Table 1 shows the error probability of community detection with side information when $\log \epsilon = -\beta \log(n)$ and $a = 3$, $b = 1$. When $\beta = 0.8$, we have $(\sqrt{a} - \sqrt{b})^2 + 2\beta = 2.136 > 2$ and thus, it can be seen that as n increases, the error probability decreases and error occurrences are rare. When $\beta = 0.2$, then $(\sqrt{a} - \sqrt{b})^2 + 2\beta = 0.936 < 2$ and in comparison with the case $\beta = 0.8$ the error occurrences are significant.

Table 2 shows the error probability of community detection without the side information, i.e. $\beta = 0$. When $a = 6$, $b = 1$, and $\beta = 0$, then $(\sqrt{a} - \sqrt{b})^2 = 2.10 > 2$ and thus, as n increases, the error probability decreases and error occurrences are rare. When $a = 3$ and $b = 1$, then $(\sqrt{a} - \sqrt{b})^2 = 0.536 < 2$, and it can be seen that in comparison with Table 1 where $a = 3$, $b = 1$, and $\beta = 0.2$, side information improves the error probability.

With Side Information				
a	b	β	n	Error Probability
3	1	0.2	100	2.1×10^{-2}
3	1	0.2	200	1.6×10^{-2}
3	1	0.2	300	1.3×10^{-2}
3	1	0.2	400	1.1×10^{-2}
3	1	0.2	500	1.0×10^{-2}
3	1	0.8	100	2.7×10^{-4}
3	1	0.8	200	1.5×10^{-4}
3	1	0.8	300	8.6×10^{-5}
3	1	0.8	400	6.4×10^{-5}
3	1	0.8	500	5.0×10^{-5}

Table 1. The error probability of community detection for a generated graph with side information by SDP.

Without Side Information				
a	b	β	n	Error Probability
3	1	0.0	100	1.4×10^{-1}
3	1	0.0	200	1.2×10^{-1}
3	1	0.0	300	1.0×10^{-1}
3	1	0.0	400	9.5×10^{-2}
3	1	0.0	500	9.1×10^{-2}
6	1	0.0	100	6.7×10^{-5}
6	1	0.0	200	4.0×10^{-5}
6	1	0.0	300	2.3×10^{-5}
6	1	0.0	400	1.1×10^{-5}
6	1	0.0	500	5.0×10^{-6}

Table 2. The error probability of community detection for a generated graph without side information by SDP.

4. REFERENCES

[1] Antonio Ortega, Pascal Frossard, Jelena Kovačević, José MF Moura, and Pierre Vandergheynst, "Graph signal processing: Overview, challenges, and applications," *Proceedings of the IEEE*, vol. 106, no. 5, pp. 808–828, 2018.

[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] 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 Processing Magazine*, vol. 31, no. 5, pp. 80–90, Sept 2014.

[4] Paul W. Holland, Kathryn Blackmond Laskey, and Samuel Leinhardt, "Stochastic blockmodels: First steps," *Social Networks*, vol. 5, no. 2, pp. 109 – 137, 1983.

[5] Anne Condon and Richard M. Karp, "Algorithms for graph partitioning on the planted partition model," *Random Structures & Algorithms*, vol. 18, no. 2, pp. 116–140.

[6] Béla Bollobás, Svante Janson, and Oliver Riordan, "The phase transition in inhomogeneous random graphs," *Random Struct. Algorithms*, vol. 31, no. 1, pp. 3–122, Aug. 2007.

[7] E. Abbe, A. S. Bandeira, and G. Hall, "Exact recovery in the stochastic block model," *IEEE Transactions on Information Theory*, vol. 62, no. 1, pp. 471–487, Jan 2016.

[8] E. Abbe and C. Sandon, "Community detection in general stochastic block models: Fundamental limits and efficient algorithms for recovery," in *Symposium on Foundations of Computer Science (FOCS)*, 2015, FOCS '15, pp. 670–688.

[9] Emmanuel Abbe, "Community detection and stochastic block models," *Foundations and Trends in Communications and Information Theory*, vol. 14, no. 1-2, pp. 1–162, 2018.

[10] H. Saad and A. Nosratinia, "Community detection with side information: Exact recovery under the stochastic block model," *IEEE Journal of Selected Topics in Signal Processing*, pp. 1–1, 2018.

[11] H. Saad and A. Nosratinia, "Belief propagation with side information for recovering a single community," in *IEEE International Symposium on Information Theory*, June 2018.

[12] H. Saad and A. Nosratinia, "Side information in recovering a single community: Information theoretic limits," in *IEEE International Symposium on Information Theory*, June 2018.

[13] H. Saad, A. Abotabl, and A. Nosratinia, "Exact recovery in the binary stochastic block model with binary side information," in *Allerton Conference on Communication, Control, and Computing*, Oct. 2017, pp. 822–829.

[14] H. Saad and A. Nosratinia, "Recovering a Single Community with Side Information," *ArXiv e-prints*, Sept. 2018.

[15] Zhi-Quan Luo and Wei Yu, "An introduction to convex optimization for communications and signal processing," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 8, pp. 1426–1438, Aug 2006.

[16] J. Mattingley and S. Boyd, "Real-time convex optimization in signal processing," *IEEE Signal Processing Magazine*, vol. 27, no. 3, pp. 50–61, May 2010.

[17] Uriel Feige and Joe Kilian, "Heuristics for semirandom graph problems," *Journal of Computer and System Sciences*, vol. 63, no. 4, pp. 639 – 671, 2001.

[18] A. Frieze and M. Jerrum, "Improved approximation algorithms for maxk-cut and max bisection," *Algorithmica*, vol. 18, no. 1, pp. 67–81, May 1997.

[19] B. Hajek, Y. Wu, and J. Xu, "Achieving exact cluster recovery threshold via semidefinite programming," *IEEE Transactions on Information Theory*, vol. 62, no. 5, pp. 2788–2797, May 2016.

[20] B. Hajek, Y. Wu, and J. Xu, "Achieving exact cluster recovery threshold via semidefinite programming: Extensions," *IEEE Transactions on Information Theory*, vol. 62, no. 10, pp. 5918–5937, Oct 2016.

[21] Michel X. Goemans and David P. Williamson, "Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming," *J. ACM*, vol. 42, no. 6, pp. 1115–1145, Nov. 1995.