A Nearly Tight Sum-of-Squares Lower Bound for the Planted Clique Problem

Mona Azadkia

June 7, 2017

Final Project Report for CS369h-Spring 2017-Stanford University

1 Introduction

The *planted clique* problem is a central question in average-case complexity. The problem is formally defined as follows: given a random Erdös-Réyi graph G from the distribution $G(n, 1/2)$ in which we plant an additional clique S of size ω, find S. It is not hard to see that the problem is solvable by brute force search whenever $\omega > c \log n$ for any constant $c > 2$. However the best polynomial-time algorithm only works for $\omega = \epsilon\sqrt{n}$, for any constant $\epsilon > 0$.

The first SoS lower bound for planted clique was shown by Meka, Potechin and Wigderson who proved that the degree d SoS hierarchy cannot recover a clique of size $\tilde{O}(n^{1/d})$. This bound was later improved on by Deshpande and Montanari and then Hopkins et al to $\tilde{O}(n^{1/2})$ for degree $d = 4$ and $\tilde{O}(n^{1/(\lceil d/2 \rceil+1)})$ for general d. However, this still left open the possibility that the constant degree (and hence the polynomial time) SoS algorithm can significantly beat the $\sqrt{n}$ bound, perhaps even being able to find cliques of size n^{ϵ} for any $\epsilon > 0$. This paper answers this question negatively by proving the following theorem:

Theorem 1.1. *There is an absolute constant c so that for every $d = d(n)$ and large enough n, the SoS relaxation of the planted clique problem has integrality gap at least* $n^{1/2-c(d/\log n)^{1/2}}$.

2 Planted Clique and Probabilistic Inference

If a graph G contains a uniques clique S of size ω, for every vertex i the probability that i is in S is either zero or one. But, a computationally bounded observer may not know whether i is in the clique or not, and we could try to quantify this ignorance using probabilities. These can be thought of as a computational analogus of *Bayesian probabilities*, that, rather aiming to measure the frequency at which an event occurs in some sample space, attempt to capture the subjective beliefs of some observer.

Consider the following scenario. Let $G(n, 1/2, \omega)$ be the distribution over pairs (G, x) of n-vertex graphs G and vectors $x \in \mathbb{R}^n$ that is obtained by sampling a random graph in $G(n, 1/2)$, planting an ω-sized clique in it, and letting G be the resulting graph and x the $0/1$ characteristic vector of the clique. Let $f : \{0,1\}^{\binom{n}{2}} \times \mathbb{R}^n \to \mathbb{R}$ be some function that maps a graph G and a vector x into some real number $f_G(x)$. Now imagine two parties, Alice and Bob that play the following game: Alice samples (G, x) from the distribution $G(n, 1/2, \omega)$ and sends$G$ to Bob, who needs to output the expected value of $f_G(x)$. We denote this value by $\tilde{\mathbb{E}}_G f_G$.

If we have no computational constraints then it is clear that Bob will simply let $\tilde{\mathbb{E}}_G f_G$ be equal to $\mathbb{E}_{x|G} f_G(x)$, by which we mean the expected value of $f_G(x)$ where x is chosen according to the conditional distribution on x given the graph G. In particular, the value $\tilde{\mathbb{E}}_G f_G$ will be calibrated in teh sense that

$$\mathbb{E}_{G \in G(n,1/2,\omega)} \tilde{\mathbb{E}}_G f_G = \mathbb{E}_{(G,x) \in G(n,1/2,\omega)} f_G(x) \tag{1}$$

Now if Bob is computationally bounded, then he might not be able to compute the value of $\mathbb{E}_{x|G} f_G(x)$ even for a simple function. However we don't need to compute the true conditional expectation to obtain a calibrated estimate.

Our SoS lower bound amounts to coming up with some reasonable *pseudo expectation* that can be efficiently computed, where $\tilde{\mathbb{E}}_G$ is meant to capture a best effort of a computationally bounded party of approximating the Bayesian conditional expectation $\mathbb{E}_{x|G}$. Our pseudo expectation will not be even close to the true conditional expectations, but will at least be internally consistent in the sense that for simple fuctions f it satisfies (1). In fact, since the pseudo expectation will not distinguish between a graph G drawn from $G(n, 1/2, \omega)$ and a random G from $G(n, 1/2)$ it will also satisfy the following *pseudo calibration* condition:

$$\mathbb{E}_{G \in G(n,1/2)} \tilde{\mathbb{E}}_G f_G = \mathbb{E}_{(G,x) \in G(n,1/2,\omega)} f_G(x) \tag{2}$$

for all simple functions $f = f(G, x)$. Note that (2) does not make sense for the estimates of a truly Bayesian Bob, since almost all graphs G in $G(n, 1/2)$ are not even in the support of $G(n, 1/2, \omega)$. Nevertheless, our pseudo distributions will be well defined even for a random graph and hence will yield estimates fro the probabilities over this hypothetical object that does not exist.

2.1 From Calibrated Pseudo-distribution to Sum-of-Squares Lower Bounds

In this section we show how calibration is almost forced on any pseudo distribution feasible for the sim of squares algorithm. Specifically, to show that the degree d SoS algorithm fails to cerify that a random graph does not contain a clique of size ω, we need to show that for a random G, with high probability we can come up with an operator that maps a degree at most d, n-variate polynomial p t a real number $\tilde{\mathbb{E}}_G p$ satisfying the following constraints:

1. (Linearity) The map $p \mapsto \tilde{\mathbb{E}}_G p$ is linear.

2. (Normalization) $\tilde{\mathbb{E}}_G 1 = 1$.

3. (Booleanity constraint) $\tilde{\mathbb{E}}_G x_i^2 p = \tilde{\mathbb{E}} x_i p$ for every p of degree at most $d-2$ and $i \in [n]$.

4. (Clique constraint) $\tilde{\mathbb{E}} x_i x_j p = 0$ for every (i,j) that is not an edge and p of degree at most $d-2$.

5. (Size constraint) $\tilde{\mathbb{E}}_G \sum_{i=1}^n x_i = \omega$.

6. (Positivity) $\tilde{\mathbb{E}}_G p^2 \geq$ for every p of degree at most parameter $d/2$.

Definition 2.1. *A map $p \mapsto \tilde{\mathbb{E}}_G p$ satisfying the above constraints 1-6 is called a degree d pseudo distribution.*

Theorem 2.2. *Theorem 1, restated There is some constant c such that if $\omega \leq n^{1/2 - c(d/\log n)^{1/2}}$ then with high probability over G sampled from $G(n, 1/2)$, there is a degree d pseudo distribution $\tilde{\mathbb{E}}_G$ satisfying constraints 1-6 above.*

2.2 Proving Positivity

How do we formally define a pseudo-calibrated linear map $\tilde{\mathbb{E}}_G$ and how do we show that it satisfies constraints 1-6 with high probability, to prove the theorem?

In order to find the map from G to $\tilde{\mathbb{E}}_G$ such that when G is taken from $G(n, 1/2)$ with high probability it satisfies the constraints 1-6, we define $\tilde{\mathbb{E}}_G$ in a way that it satisfies the pseudo calibration requirement with respect to all functions $f = f(G,x)$ that are low degree polynomials both in G and x variables. The above requirements determine all the low-degree Fourier coefficients of the map $G \mapsto \tilde{\mathbb{E}}_G$. Indeed, instantiating from (2) with every particular function $f = f(G,x)$ defines a linear constraint on the pseudo expectation operator. If we require (2) to hold with respect to every function $f = f(G,x)$ that has degree at most τ in the entries of the adjacency matrix G and degree at most d in the variables x, and in addition we require that the map $G \mapsto \tilde{\mathbb{E}}_G$ is itself of degree at most τ in G, then this completely determines $\tilde{\mathbb{E}}_G$. For any $S \subset [n]$, $|S| \leq d$, using Fourier transform it is not too hard to compute $\tilde{\mathbb{E}}_G(x_S)$ as an explicit low degree polynomial in G_e:

$$\tilde{\mathbb{E}}_G(x_S) = \sum_{\substack{T \subset \binom{[n]}{2} \\ |\mathcal{V}(T) \cup S| \leq \tau}} \left(\frac{\omega}{n}\right)^{|\mathcal{V}(T) \cup S|} \chi_T(G) \tag{3}$$

where $\mathcal{V}(T)$ is the set if nodes incident to the subset of edges T and $\chi_T(G) = \prod_{e \in T} G_e$. It turns out constraints 1-5 are easy to verify and thus we are left with proving the *positivity constraint*.

As is standard, to analyze this positiviy requirement we work with the *moment matrix* of $\tilde{\mathbb{E}}_G$. Namely, let $\mathcal{M}$ be the $\binom{n}{\leq d/2} \times \binom{n}{\leq d/2}$ matrix where $\mathcal{M}(I,J) = \tilde{\mathbb{E}}_G \prod_{i \in I} x_i \prod_{i \in J} x_j$ for every pair of subsets $I, J \subset [n]$ of size at most $d/2$. Our goal can be rephrased as showing that $\mathcal{M} \succeq 0$.

Now given a symmetric matrix N, to show that $N \succeq 0$ our first hope might be to diagonalize N. That is, we would hope to find a matrix V and a diagona matrix D so that $N = VDV^T$. Then as long as every entry of D is nonnegative, we would obtain $N \succeq 0$. Unfortunately, carrying this out directly can be far to complicated. However it is sometimes possible to prove PSDness for a random matrix using an *approximate* diagonalization.

3 Proving Positivity: A Technical Overview

We now discuss in more detail how we prove that the *moment matrix* $\mathcal{M}$ corresponding to our pseudo distribution is positive semidefinite. Recall that we have

$$\mathcal{M}(I,J) = \sum_{\substack{T \subset \binom{n}{2} \\ |\mathcal{V}(T)\cup I \cup J| \leq \tau}} \left(\frac{\omega}{n}\right)^{|\mathcal{V}(T)\cup I\cup J|} \chi_T(G). \tag{4}$$

The matrix $\mathcal{M}$ is generated from the random graph G, but its entries are not independent. Rather, each entry is a polynomial in G_e, and there are some fairly complex dependencies between different them. Indeed, these dependencies will create a spectral structure for $\mathcal{M}$ that is very different from the spectrum of standard random matrices with independent entries and makes proving $\mathcal{M}$ positive semidefinite challenging.

4 Approximate Factorization of the Moment Matrix

4.1 Ribbons and Vertex Seperators

Definition 4.1. *(Ribbon) An (I,J)-ribbon $\mathcal{R}$ is a graph with edge set $W_R \subset \binom{n}{2}$ and vertex set $V_R \supseteq \mathcal{V}(W_\mathcal{R}) \cup I \cup J$, for two specially identified subsets $I, J \subseteq [n]$, each of size at most d, called the left and the right ends respectively. We sometimes write $\mathcal{V}(\mathcal{R}) \stackrel{def}{=} V_\mathcal{R}$ and call $|\mathcal{V}(\mathcal{R})|$ the size of $\mathcal{R}$. Also, we write $\chi_\mathcal{R}$ for the monomial $\chi_{W_\mathcal{R}}$ where $W_\mathcal{R}$ is the edge set of the ribbon $\mathcal{R}$.*

In our analysis (I,J)-ribbons arise as the terms in the Fourier decomposition of the entry $\mathcal{M}(I,J)$ in the moment matrix. It is important to emphasize that the subsets I and J in an (I,J)-ribbon are allowed to intersect. Also $\mathcal{V}(\mathcal{R})$ can contain vertices that are not in $\mathcal{V}(W_\mathcal{R})$ if there are isolated vertices in the ribbon.

Ultimately we want to partition a ribbon into three subribbons in such a way that we can express the moment matrix as the sum fo positive semidefinite matrices, and some error terms. Our partitioning will be based on minimum vertex separators.

Definition 4.2. *(Vertex Separator) For and (I,J)-ribbon $\mathcal{R}$ with edge set $W_\mathcal{R}$, a subset $Q \subseteq \mathcal{V}(\mathcal{R})$ of vertices is a vertex separator if Q separates I and J in $W_\mathcal{R}$. A vertes separator is minimum if there are no other vertex separators with strictly fewer vertices. The separator size of $\mathcal{R}$ is the cardinality of any minimum vertex separator of $\mathcal{R}$.*

Lemma 4.3. *(Leftmost/Rightmost Vertex Separator) Let $\mathcal{R}$ be an (I,J)-ribbon. There is a unique minimum vertex separator S of $\mathcal{R}$ such that S separates I and Q for any vertex separator Q of $\mathcal{R}$. We call S the leftmost separator in $\mathcal{R}$. We define the rightmost separator analogously and we denote them by $S_L(\mathcal{R})$ and $S_R(\mathcal{R})$ respectively.*

Definition 4.4. *(Canonical Factorization) Let $\mathcal{R}$ be an (I,J)-ribbon with edge set $W_\mathcal{R}$ and vertex set $V_\mathcal{R}$. Let V_ℓ be the vertices reachable from I without passing through $S_L(\mathcal{R})$, and similarly for V_r, and let $V_m = V_\mathcal{R}\ (V_\ell \cup V_r)$. Let $W_\ell \subseteq W_\mathcal{R}$ be given by*

$$W_\ell = \{(u,v) \in W_\mathcal{R} : u \in V_\ell and v \in V_\ell \cup S_L\}$$

and similarly for W_r. Finally, let $W_m = W_{\mathcal{R}}(W_\ell \cup W_r)$.

Let $\mathcal{R}_\ell$ be the $(I, S_L(\mathcal{R}))$-ribbon with vertex set $V_\ell \cup S_L(\mathcal{R})$ and edge set W_ℓ and similarly for $\mathcal{R}_r$. Let $\mathcal{R}_m$ be the $(S_L(\mathcal{R}), S_R(\mathcal{R}))$-ribbon with vertex set V_m and edge set W_m. The triple $(\mathcal{R}_\ell, \mathcal{R}_m, \mathcal{R}_r)$ is the canonical factorization of $\mathcal{R}$.

Some facts about the canonical factorization:

- W_ℓ, W_r and W_m are disjoint and are a partition of $W_{\mathcal{R}}$ by construction. Hence $\chi_{\mathcal{R}} = \chi_{W_\ell} \cdot \chi_{W_m} \cdot \chi_{W_r} \cdot$
- some vertices in I may not be in V_ℓ at all. However any such vertices are necessarily in S_L and thus in $\mathcal{R}_\ell$ anyways.

Now with this definition in hand, let see some important properties.

Claim 4.5. *Let $\mathcal{R}$ be and (I, J)-ribbon with canonical factorization $(\mathcal{R}_\ell, \mathcal{R}_m, \mathcal{R}_r)$. Then*

$$|\mathcal{V}(\mathcal{R})| = |\mathcal{V}(\mathcal{R}_\ell)| + |\mathcal{V}(\mathcal{R}_m)| + |\mathcal{V}(\mathcal{R}_r)| - |S_L(\mathcal{R})| - |S_R(\mathcal{R})|.$$

Now consider a collection of ribbons $\mathcal{R}_0$, $\mathcal{R}_1$, $\mathcal{R}_2$ and the following list of properties:

S_ℓ, S_r Factorization Conditions for $\mathcal{R}_0$, $\mathcal{R}_1$, $\mathcal{R}_2$ (Here S_ℓ, $S_r \subseteq [n]$).

1. $\mathcal{R}_0$ is an (I, S_ℓ)-ribbon with $S_L(\mathcal{R}_0) = S_R(\mathcal{R}_0) = S_\ell$, and all vertices in $\mathcal{V}(\mathcal{R}_0)$ are either reachable from I without passing through S_ℓ or are in I or S_ℓ. Finally, $\mathcal{R}_0$ has no edges between vertices in S_ℓ.

2. $\mathcal{R}_2$ is an (S_r, J)-ribbon with $S_L(\mathcal{R}_2) = S_R(\mathcal{R}_2) = S_r$, and all vertices in $\mathcal{V}(\mathcal{R}_2)$ are either reachable from J without passing through S_r or are in J or S_r. Finally, $\mathcal{R}_2$ has no edges between vertices in S_r.

3. $\mathcal{R}_1$ is an (S_ℓ, S_r)-ribbon with $S_L(\mathcal{R}_1) = S_\ell$ and $S_R(\mathcal{R}_1) = S_r$. Every vertex in $\mathcal{V}(\mathcal{R}_1)$ $(S_\ell \cup S_r)$ has degree at least 1.

4. $W_{\mathcal{R}_0}$, $W_{\mathcal{R}_1}$, $W_{\mathcal{R}_2}$ are pairwise disjoint. Also, $V_{\mathcal{R}_0} \cup V_{\mathcal{R}_1} = S_\ell$, $V_{\mathcal{R}_1} \cup V_{\mathcal{R}_2} = S_r$, and $V_{\mathcal{R}_0} \cup V_{\mathcal{R}_2} = S_\ell \cup S_r$.

Lemma 4.6. *Let $\mathcal{R}_0$, $\mathcal{R}_1$, $\mathcal{R}_2$ be ribbons. Then $(\mathcal{R}_0, \mathcal{R}_1, \mathcal{R}_2)$ is the canonical factorization of the (I, J)-ribbon $\mathcal{R}$ with edge set $W_{\mathcal{R}_0} \oplus W_{\mathcal{R}_1} \oplus W_{\mathcal{R}_2}$ and vertex set $\mathcal{V}(\mathcal{R}_0) \cup \mathcal{V}(\mathcal{R}_1) \cup \mathcal{V}(\mathcal{R}_2)$ if and only if the S_ℓ, S_r factorization conditions hold for $\mathcal{R}_0$, $\mathcal{R}_1$, $\mathcal{R}_2$ for some $S_\ell, S_r \subseteq [n]$.*

4.2 Factorization of Matrix Entries

Using canonical factorization and the Claim, for any $I, J \subseteq [n]$ of size at most d we can write

$$\begin{aligned}
\mathcal{M}(I,J) &= \sum_{\substack{\mathcal{R} \text{ an } (I,J)\text{-ribbon with edge set } W, \\ \text{canonical factorization } (\mathcal{R}_0,\mathcal{R}_1,\mathcal{R}_2)}} \left(\frac{\omega}{n}\right)^{|\mathcal{V}(\mathcal{R})|} \cdot \chi_{\mathcal{R}_\ell} \cdot \chi_{\mathcal{R}_m} \cdot \chi_{\mathcal{R}_r} \\
&= \sum_{\substack{S_\ell, S_r \subseteq [n] \\ |S_\ell| = |S_r| \le d}} \left(\frac{\omega}{n}\right)^{-\frac{|S_\ell|+|S_r|}{2}} \times \\
&\quad \sum_{\substack{\mathcal{R}_0,\mathcal{R}_1,\mathcal{R}_2 \subseteq \binom{n}{2} \\ \text{satisfying } S_\ell,\, S_r \text{ factorization conditions} \\ \text{and} |\mathcal{V}(\mathcal{R}_\ell) \cup \mathcal{V}(\mathcal{R}_m) \cup \mathcal{V}(\mathcal{R}_r)| \le \tau}} \left(\frac{\omega}{n}\right)^{|\mathcal{V}(\mathcal{R}_\ell)|+|\mathcal{V}(\mathcal{R}_m)|+|\mathcal{V}(\mathcal{R}_r)|-\frac{|S_\ell|+|S_r|}{2}} \cdot \chi_{\mathcal{R}_\ell} \cdot \chi_{\mathcal{R}_m} \cdot \chi_{\mathcal{R}_r}
\end{aligned}$$

Notice that exept for the disjointness condition, the S_ℓ, S_r factorization conditions can be seperated into condition 1 for $\mathcal{R}_\ell$, condition 3 for $\mathcal{R}_m$, and condition 2 for $\mathcal{R}_r$. We use this to rewrite as

$$\begin{aligned}
= &\sum_{\substack{S_\ell, S_r \subseteq [n] \\ |S_\ell| = |S_r| \le d}} \left(\frac{\omega}{n}\right)^{-\frac{|S_\ell|+|S_r|}{2}} \left(\sum_{\substack{\mathcal{R}_\ell \text{having } 1 \\ |\mathcal{V}(\mathcal{R}_\ell)| \le \tau}} \left(\frac{\omega}{n}\right)^{|\mathcal{V}(\mathcal{R}_\ell)|} \chi_{\mathcal{R}_\ell} \right) \left(\sum_{\substack{\mathcal{R}_m \text{having } 3 \\ |\mathcal{V}(\mathcal{R}_m)| \le \tau}} \left(\frac{\omega}{n}\right)^{|\mathcal{V}(\mathcal{R}_m)|-\frac{|S_\ell|+|S_r|}{2}} \chi_{\mathcal{R}_m} \right) \times \\
&\left(\sum_{\substack{\mathcal{R}_r \text{having } 2 \\ |\mathcal{V}(\mathcal{R}_r)| \le \tau}} \left(\frac{\omega}{n}\right)^{|\mathcal{V}(\mathcal{R}_r)|} \chi_{\mathcal{R}_r} \right) \qquad (5) \\
- &\underbrace{\sum_{\substack{S_\ell, S_r \subseteq [n] \\ |S_\ell| = |S_r| \le d}} \left(\frac{\omega}{n}\right)^{-\frac{|S_\ell|-|S_r|}{2}} \sum_{\substack{\mathcal{R}_\ell, \mathcal{R}_m, \mathcal{R}_r \\ \text{satisfying } S_\ell,\, S_r \text{ conditions} \\ |\mathcal{V}(\mathcal{R}_\ell)|, |\mathcal{V}(\mathcal{R}_m)|, |\mathcal{V}(\mathcal{R}_r)| \le \tau, \\ |\mathcal{V}(\mathcal{R}_\ell) \cup \mathcal{V}(\mathcal{R}_m) \cup \mathcal{V}(\mathcal{R}_r)| > \tau}} \left(\frac{\omega}{n}\right)^{|\mathcal{V}(\mathcal{R}_\ell)|+|\mathcal{V}(\mathcal{R}_m)|+|\mathcal{V}(\mathcal{R}_r)|-\frac{|S_\ell|+|S_r|}{2}} \cdot \chi_{\mathcal{R}_\ell} \cdot \chi_{\mathcal{R}_m} \cdot \chi_{\mathcal{R}_r}}_{\stackrel{\text{def}}{=} \xi_0(I,J), \text{ the error from ribbon size}} \qquad (6) \\
- &\underbrace{\sum_{\substack{S_\ell, S_r \subseteq [n] \\ |S_\ell| = |S_r| \le d}} \left(\frac{\omega}{n}\right)^{-\frac{|S_\ell|-|S_r|}{2}} \sum_{\substack{\mathcal{R}_\ell, \mathcal{R}_m, \mathcal{R}_r \text{satisfying} \\ 1,2,3 \text{ and not } 4 \\ |\mathcal{V}(\mathcal{R}_\ell)|, |\mathcal{V}(\mathcal{R}_m)|, |\mathcal{V}(\mathcal{R}_r)| \le \tau}} \left(\frac{\omega}{n}\right)^{|\mathcal{V}(\mathcal{R}_\ell)|+|\mathcal{V}(\mathcal{R}_m)|+|\mathcal{V}(\mathcal{R}_r)|-\frac{|S_\ell|+|S_r|}{2}} \cdot \chi_{\mathcal{R}_\ell} \cdot \chi_{\mathcal{R}_m} \cdot \chi_{\mathcal{R}_r}}_{\stackrel{\text{def}}{=} E_0(I,J), \text{ the error from ribbon nondisjointness}} \qquad (7)
\end{aligned}$$

Note that in lines (6) and (7) we have defined two error matrices, $\xi_0, E_0 \in \mathbb{R}^{\binom{n}{\le d} \times \binom{n}{\le d}}$. Inspired by the factorization of $\mathcal{M}(I,J)$ in (5) we can define

$$\mathcal{Q}_0 \in \mathbb{R}^{\binom{n}{\le d} \times \binom{n}{\le d}} \quad \textit{given by} \quad \mathcal{Q}_0(S_\ell, S_r) = \sum_{\substack{\mathcal{R}_m \text{having} 3 \\ |\mathcal{V}(\mathcal{R}_m)| \le \tau}} \left(\frac{\omega}{n}\right)^{|\mathcal{V}(\mathcal{R}_m)|-\frac{|S_\ell|+|S_r|}{2}} \chi_{\mathcal{R}_m}$$

$$\mathcal{L} \in \mathbb{R}^{\binom{n}{\le d} \times \binom{n}{\le d}} \quad \textit{given by} \quad \mathcal{L}(I, S) = \left(\frac{\omega}{n}\right)^{-\frac{|S|}{2}} \sum_{\substack{\mathcal{R}_\ell \text{having} 1 \\ |\mathcal{V}(\mathcal{R}_\ell)| \le \tau}} \left(\frac{\omega}{n}\right)^{|\mathcal{V}(\mathcal{R}_\ell)|} \chi_{\mathcal{R}_\ell}$$

The powers of (ω/n) are split between $\mathcal{Q}_0$ and $\mathcal{L}$ so that the typical eigenvalue of $\mathcal{Q}_0$ will be approximately 1. In the matrix formulation we can summarize the equations (5), (6) and (7) as

$$\mathcal{M} = \mathcal{L}\mathcal{Q}_0\mathcal{L}^T - \xi_0 - E_0.$$

It can be shown that with high probability $\mathcal{Q}_0 \succeq 0$, and thus aslo $\mathcal{L}\mathcal{Q}_0\mathcal{L}^T \succeq 0$. So as long as τ is sufficiently large, the spectral norm $\|\xi_0\|$ of the error term that corresponds to ribbons whose size is too large will be negligible. However the error E_0 does not turn out to be negligible. To overcome this the idea is to apply a similar idea of factorization to E_0 as we did for $\mathcal{M}$. By iterating this factorization procedure we will push down the error from ribbon nondisjointness.

Claim 4.7.

$$\mathcal{M} = \mathcal{L}(\mathcal{Q}_0 - \mathcal{Q}_1 + \mathcal{Q}_2 - \cdots - \mathcal{Q}_{2d-1} + \mathcal{Q}_{2d})\mathcal{L}^T - (\xi_0 - \xi_1 + \xi_2 - \cdots - \xi_{2d-1} + \xi_{2d}).$$

5 $\mathcal{M}$ is PSD

Lemma 5.1. *Let $D \in \mathbb{R}^{\binom{[n]}{\leq d} \times \binom{[n]}{\leq d}}$ be the diagonal matrix with $D(S,S) = 2^{\binom{|S|}{2}}/4$ if S is a clique in G and 0 otherwise. With high probability, $\mathcal{Q}_0 \succeq D$.*

Lemma 5.2. *Let $D \in \mathbb{R}^{\binom{[n]}{\leq d} \times \binom{[n]}{\leq d}}$ be the diagonal matrix with $D(S,S) = 2^{\binom{|S|}{2}}/4$ if S is a clique in G and 0 otherwise. With high probability, every $\mathcal{Q}_i$ for $i \in [1, 2d]$ satisfies*

$$\frac{-D}{8d} \preceq \mathcal{Q}_i \preceq \frac{D}{8d}.$$

With the two above lemmas we have $\mathcal{Q}_0 - \cdots + \mathcal{Q}_{2d} \succeq D/2$ but since we need to work with $\mathcal{L}(\mathcal{Q}_0 - \cdots + \mathcal{Q}_{2d})\mathcal{L}^T - (\xi_0 - \cdots + \xi_{2d})$ we also need the following two lemmas.

Lemma 5.3. *With high probability, $\Pi\mathcal{L}\Pi\mathcal{L}^T\Pi \succeq \Omega(\omega/n)^{d+1}.\Pi$, where Π is the projector to the span$\{e_C : C \in \mathcal{C}_{\leq d}\}$.*

and finally

Lemma 5.4. *With high probability, $\|\xi_0 - \cdots + \xi_{2d}\| \leq n^{-16d}$.*

Now with these lemmas in hand we are able to prove the PSDness of $\mathcal{M}$. Remember that we have

$$\mathcal{M} = \mathcal{L}(\mathcal{Q}_0 - \mathcal{Q}_1 + \mathcal{Q}_2 - \cdots - \mathcal{Q}_{2d-1} + \mathcal{Q}_{2d})\mathcal{L}^T - (\xi_0 - \xi_1 + \xi_2 - \cdots - \xi_{2d-1} + \xi_{2d}).$$

By a union bound, with high probability the conclusions of Lemmas (5.1), (5.2), (5.3) and (5.4) all hold. By Lemma (5.1) and Lemma (5.2),

$$\mathcal{Q}_0 - \mathcal{Q}_1 + \mathcal{Q}_2 - \cdots - \mathcal{Q}_{2d-1} + \mathcal{Q}_{2d} \succeq \frac{D}{2} \succeq \frac{\Pi}{2}.$$

thus it gives us

$$\mathcal{L}(\mathcal{Q}_0 - \mathcal{Q}_1 + \mathcal{Q}_2 - \cdots - \mathcal{Q}_{2d-1} + \mathcal{Q}_{2d})\mathcal{L}^T \succeq \Omega(\omega/n)^{d+1}\Pi$$

Finally we have

$$\mathcal{M} = \Pi.\mathcal{M}.\Pi \succeq \Omega\Big(\frac{\omega}{n}\Big)^{d+1}.\Pi + n^{-16d}.\Pi \succeq 0.$$

References

[1] Boaz Barak, Samuel Hopkins, Jonathan Kelner, Pravesh Kothari, Ankur Moitra, Aaron Potechin *A Nearly Tight Sum-of-Squares Lower Bound for the Planted Clique Problem* 2016.