Lecture 13: SOS Lower Bounds for Planted Clique Part II

Lecture Outline

- Part I: Relaxed k-clique Equations and Theorem Statement
- Part II: Pseudo-Calibration/Moment Matching
- Part III: Decomposition of Graph Matrices via Minimum Vertex Separators
- Part IV: Attempt #1: Bounding with Square Terms
- Part V: Approximate PSD Decomposition
- Part VI: Further Work and Open Problems

Part I: Relaxed k-clique Equations and Theorem Statement

Relaxed Planted Clique Equations

- Flaw in the current analysis: Need to relax the k-clique equations slightly to make the combinatorics easier to analyze
- Relaxed k-clique Equations:

$x_i^2 = x_i$ for all i.

$x_i x_j = 0$ if $(i,j) \notin E(G)$

$(1-\epsilon)k \leq \sum_i x_i \leq (1+\epsilon)k$

Planted Clique SOS Lower Bound

- Theorem 1.1 of [BHK+16]: $\exists c > 0$ such that if $k \leq n^{\frac{1}{2} - c\sqrt{\frac{d}{logn}}}$, with high probability degree d SOS cannot prove that the relaxed k-clique equations are infeasible.
- Note: For $d = 4$ there is a lower bound of $\widetilde{\Omega}(\sqrt{n})$ for the original k-clique equations.

High Level Idea

- High level idea: Show that it is hard to distinguish between the random distribution $G\left(n, \frac{1}{2}\right)$ and the planted distribution where we put each vertex in the planted clique with probability $\frac{k}{n}$.
- Remark: We take this planted distribution to make the combinatorics easier. If we could analyze the planted distribution where the clique has size exactly k, we would satisfy the constraint $\sum_i x_i = k$ exactly.

Part II: Pseudo-Calibration/Moment Matching

Choosing Pseudo-Expectation Values

- Last lecture, Pessimist disproved our first attempt for pseudo-expectation values, the MW moments.
- How can we come up with better pseudo-expectation values?

Pseudo-Calibration/Moment Matching

- Setup: We are trying to distinguish between a random distribution ($G\left(n,\frac{1}{2}\right)$) and a planted distribution ($G\left(n,\frac{1}{2}\right)$ + planted clique)
- Pseudo-calibration/moment matching: The pseudo-expectation values over the random distribution should match the actual expected values over the planted distribution in expectation for all low degree tests.

Review: Discrete Fourier Analysis

- Requirements for discrete Fourier analysis
  1. An inner product
  2. An orthonormal basis of Fourier characters
- This gives us Fourier decompositions and Parseval's Theorem

Fourier Analysis over the Hypercube

- Example: Fourier analysis on $\{-1,1\}^n$
- Inner product: $f \cdot g = \frac{1}{2^n} \sum_x f(x) g(x)$
- Fourier characters: $\chi_A(x) = \prod_{i \in A} x_i$
- Fourier decomposition: $f = \sum_V \hat{f}_A \chi_A$ where $\hat{f}_A = f \cdot \chi_A$
- Parseval's Theorem: $\sum_A \hat{f}_A^2 = f \cdot f = \|f\|^2$

Fourier Analysis over $G\left(n, \frac{1}{2}\right)$

- Inner product: $f \cdot g = E_{G \sim G\left(n, \frac{1}{2}\right)} f(G) g(G)$
- Fourier characters: $\chi_E(G) = (-1)^{|E \setminus \mathrm{E}(G)|}$

Pseudo-Calibration Equation

- Pseudo-Calibration Equation:

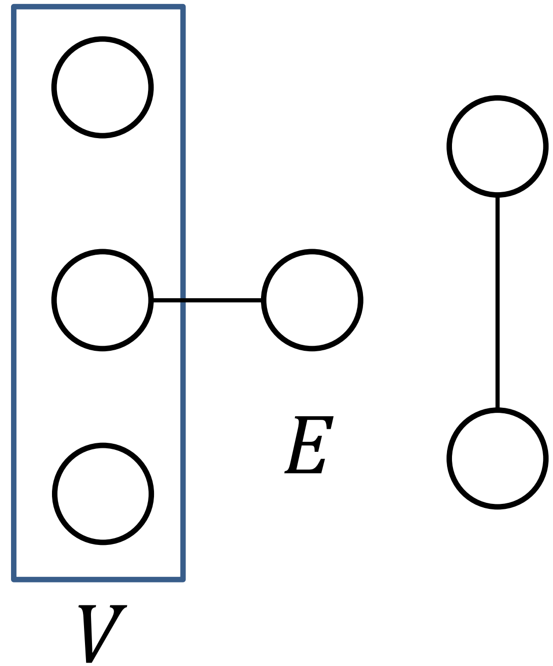
$$E_{G\sim G\left(n,\frac{1}{2}\right)}[\tilde{E}[x_V]\cdot\chi_E] = E_{G\sim planted\ dist}[x_V\cdot\chi_E]$$

- We want this equation to hold for all small V and E

Pseudo-Calibration Calculation

- To calculate $E_{G\sim planted\ dist}[x_V \cdot \chi_E]$, first choose the planted clique and then choose the rest of the graph
- $x_V = 0$ if any $i \in V$ is not in the planted clique
- $E[\chi_E(G)] = 0$ whenever E is not fully contained in the planted clique
- Def: Define $V(E) = \{\text{endpoints of edges in } E\}$
- If $V \cup V(E) \subseteq planted\ clique$ then $x_V \chi_E = 1$
- $E_{G\sim planted\ dist}[x_V \cdot \chi_E] = \left(\frac{k}{n}\right)^{|V \cup V(E)|}$

Calculation Picture



- If all the vertices are in the planted clique then $x_V \chi_E(G) = 1$. Otherwise, either $x_V = 0$ (because an $i \in V$) is missing or $E[\chi_E] = 0$ because each edge outside the clique is present with probability $\frac{1}{2}$

Fourier Coefficients of $\tilde{E}[x_V]$

- From the pseudo-calibration calculation, $\widehat{\tilde{E}[x_V]}_E = E_{G\sim G\left(n,\frac{1}{2}\right)}\left[\tilde{E}[x_V]\cdot\chi_E\right] = \left(\frac{k}{n}\right)^{|V\cup V(E)|}$
- We take $\tilde{E}[x_V] = \sum_{E:|V\cup V(E)|\leq D}\left(\frac{k}{n}\right)^{|V\cup V(E)|}$ where D is a truncation parameter and then normalize so that $\tilde{E}[x_\emptyset] = \tilde{E}[1] = 1$
- Good exercise: What happens if we don't truncate at all?

Graph Matrix Decomposition

- Ignoring the normalization, $M = \sum_H \left(\frac{k}{n}\right)^{|V(H)|} R_H$ where we sum over ALL H with at most D vertices which have no isolated vertices outside of U and V.

Part III: Decomposition of Graph Matrices via Minimum Vertex Separators

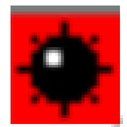
Proof Sketch

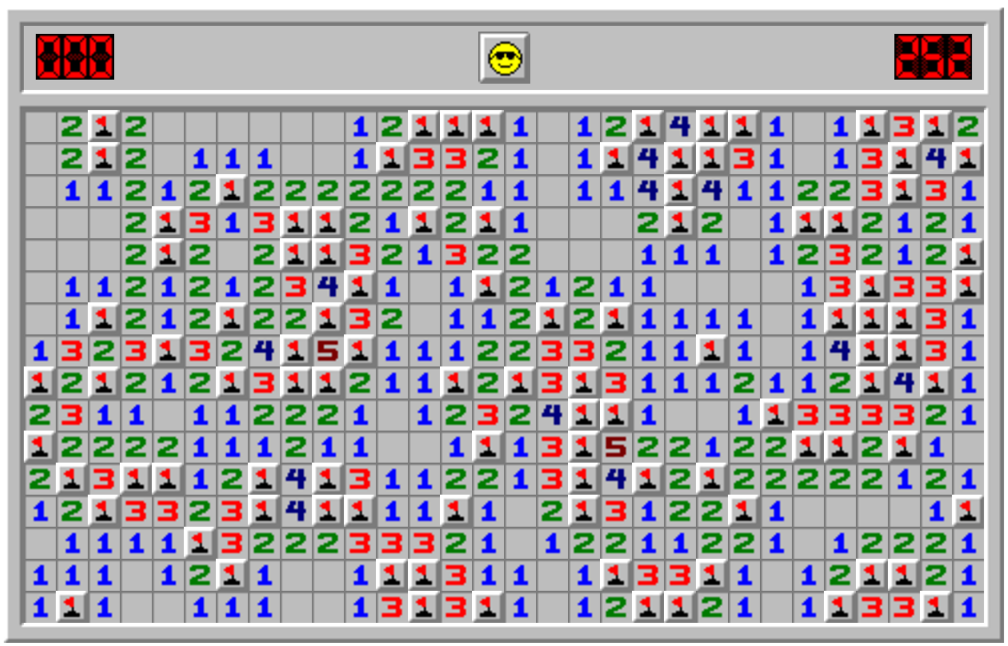
- How can we show $M \succcurlyeq 0$ with high probability?
- High level idea:
  1. Find an approximate PSD decomposition M^{fact} of M
  2. Handle the error $M^{fact} - M$. Unfortunately, this error is not small enough to ignore, so we carefully show that $M^{fact} - M \preccurlyeq M^{fact}$ with high probability. We briefly sketch the ideas for this in Appendix I. For the full details, see [BHK+16]

Technical Minefield

- Warning: This analysis is a technical minefield

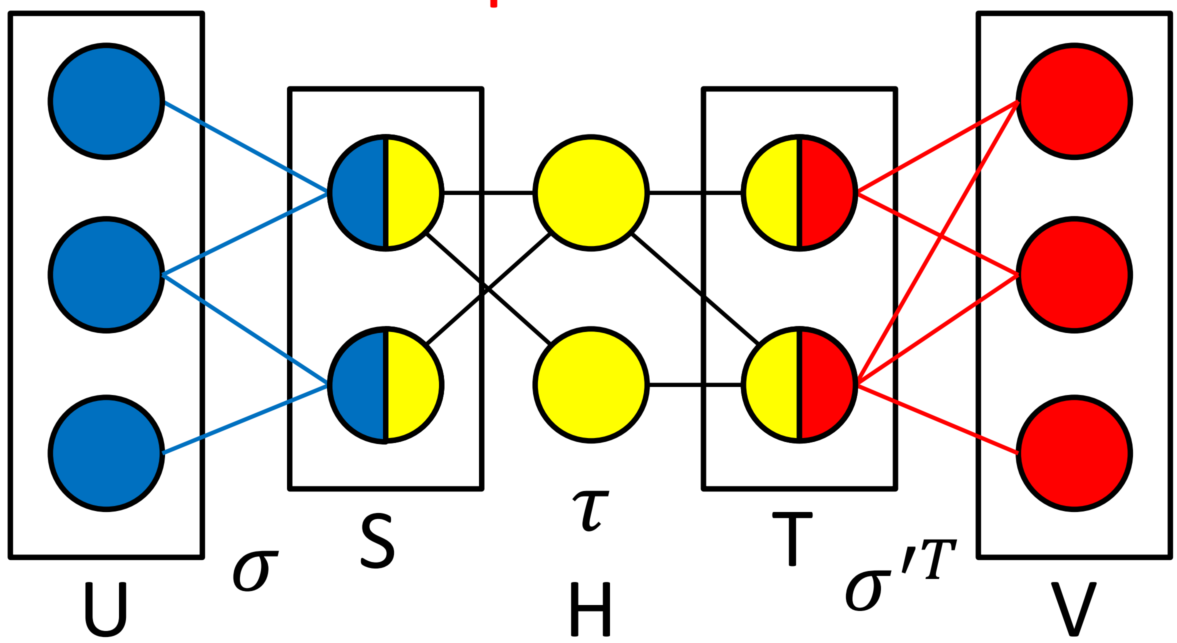

Mine handled correctly


Not quite correct, see Appendix II



Decomposition via Separators

- How can we handle all of the different R_H?
- Key idea: Decompose each H into three parts σ, τ, σ'^T based on the leftmost and rightmost minimum vertex separators S and T of H



Separator Definitions

- Definition: Given a graph H with distinguished sets of vertices U and V, a vertex separator S is a set of vertices such that any path from U to V must intersect S.
- Definition: A leftmost minimum vertex separator S is a set of vertices such that for any vertex sepator S' of minimum size, any path from U to S' intersects S.
- A rightmost minimum vertex separator is defined analogously.

Existence of Minimum Separators

- Lemma 6.3 of [BHK+16]: Leftmost and rightmost minimum vertex separators always exist and are unique.

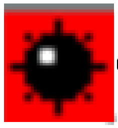
Left, Middle, and Right Parts

- Let S, T be the leftmost and rightmost minimum vertex separators of H
- Definition: We take the left part σ of H to be the part of H between U and S, we take the middle part τ of H to be the part of H between S and T, and we take the right part ${\sigma'}^T$ of H to be the part of H between T and V

Conditions on Parts

- σ, τ, σ'^T satisfy the following:
- The unique minimum vertex separator of σ is $V_\sigma = S$ (where V_σ is the right side of σ)
- The leftmost and rightmost minimum vertex separators of τ are $U_\tau = S$ and $V_\tau = T$ (where U_τ and V_τ are the left and right sides of τ)
- The unique minimum vertex separator of σ'^T is $U_{\sigma'^T} = T$ (where $U_{\sigma'^T}$ is the left side of σ'^T)

Approximate Decomposition



- Claim: If r is the size of the minimum vertex separator of H,

$$R_H \approx R_\sigma R_\tau R_{\sigma'^T}$$

- Idea: There is a bijection between injective mappings $\phi: V(H) \rightarrow V(G)$ and injective mappings $\phi_1: V(\sigma) \rightarrow V(G)$, $\phi_2: V(\tau) \rightarrow V(G)$, and $\phi_3: V(\sigma'^T) \rightarrow V(G)$ such that
  1. ϕ_1, ϕ_2 agree on S and ϕ_2, ϕ_3 agree on T
  2. Collectively, ϕ_1, ϕ_2, ϕ_3 don't map two different vertices of H to the same vertex of G

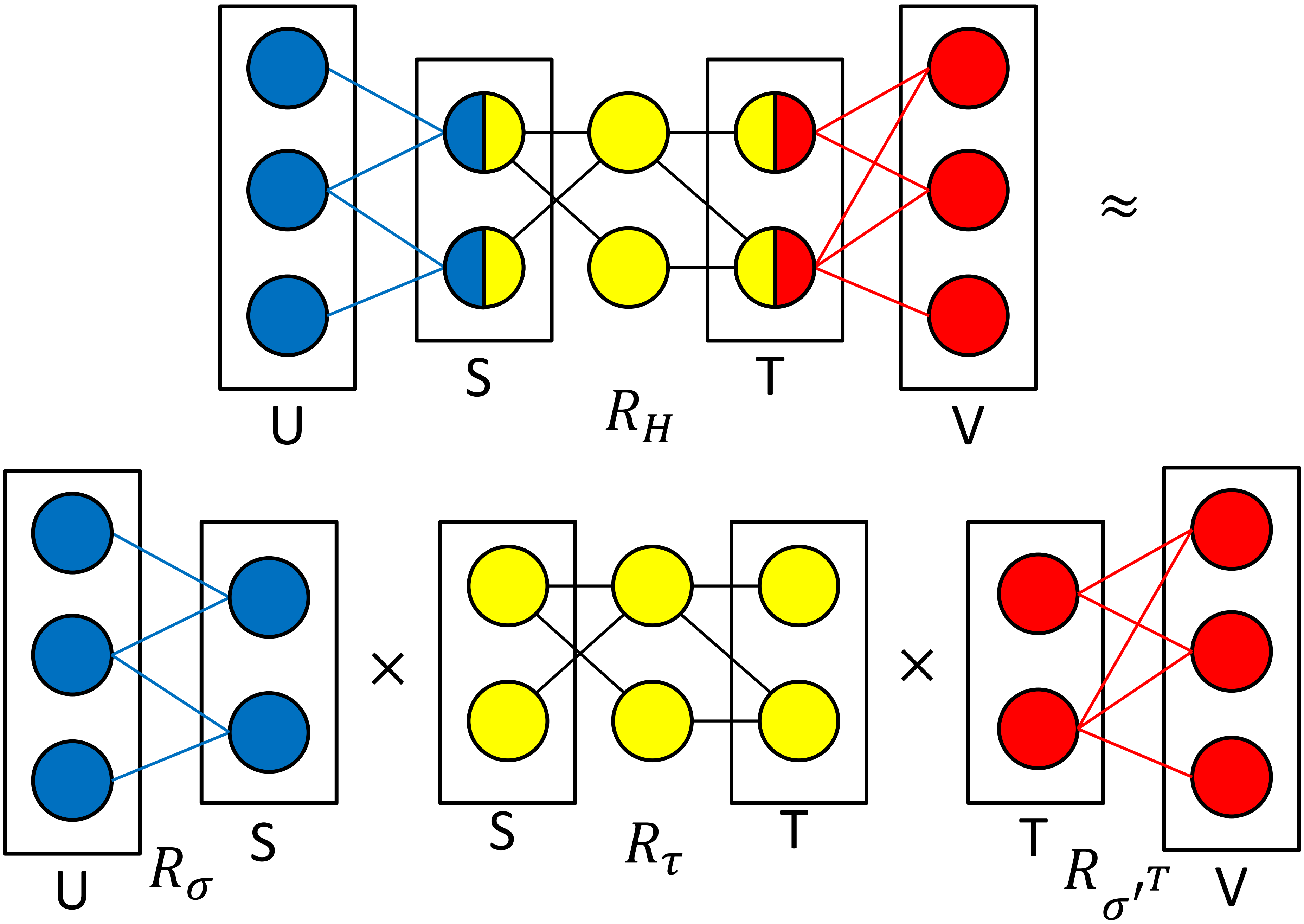
Approximate Decomposition

- Claim: If r is the size of the minimum vertex separator of H,

$$R_H \approx R_\sigma R_\tau R_{\sigma'^T}$$

- Corollary:

$$\left(\frac{k}{n}\right)^{|V(H)|} R_H \approx \left(\left(\frac{k}{n}\right)^{|V(H)|-\frac{r}{2}} R_\sigma\right)\left(\left(\frac{k}{n}\right)^{|V(H)|-r} R_\tau\right)\left(\left(\frac{k}{n}\right)^{|V(H)|-\frac{r}{2}} R_{\sigma'^T}\right)$$


S
T
U
R_H
V
≈
U
R_σ
S
×
S
R_τ
T
×
T
${R_{\sigma'}}^T$
V

Intersection Terms

- Warning! There will be terms where ϕ_1, ϕ_2, ϕ_3 map multiple vertices to the same vertex. We call these intersection terms.
- We sketch how to handle intersection terms in Appendix I. For now, we sweep this under the rug.

Part IV: Attempt #1: Bounding With Square Terms

Bounding With Square Terms

- How can we handle all of the $R_\sigma R_\tau R_{\sigma',T}$ terms?
- One idea: Can bound $R_\sigma R_\tau R_{\sigma',T} + \left(R_\sigma R_\tau R_{\sigma',T}\right)^T$ as follows.
- $\left(aR_\sigma - bR^T_{\sigma',T} R^T_\tau\right)\left(aR_\sigma - bR^T_{\sigma',T} R^T_\tau\right)^T \succcurlyeq 0$

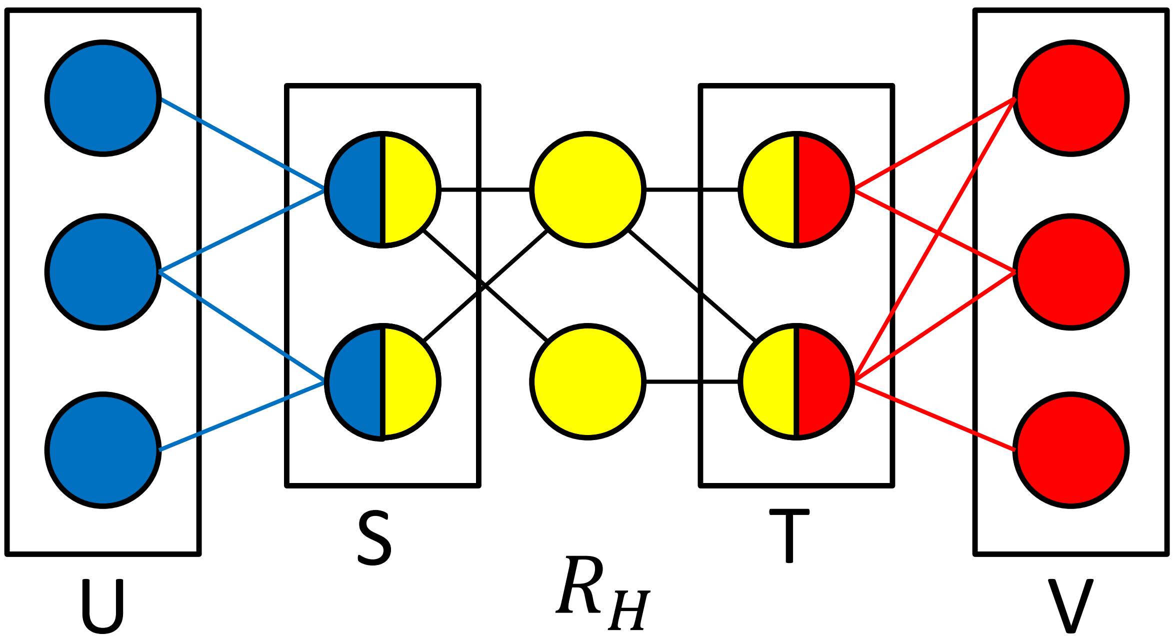
Bounding With Square Terms

- $\left(aR_\sigma - bR^T_{\sigma',T} R^T_\tau\right)\left(aR_\sigma - bR^T_{\sigma',T} R^T_\tau\right)^T \succcurlyeq 0$
- Rearranging, ab $\left(R_\sigma R_\tau R_{\sigma',T} + \left(R_\sigma R_\tau R_{\sigma',T}\right)^T\right) \preccurlyeq$

$a^2 R_\sigma R_\sigma^T + b^2 R^T_{\sigma',T} R^T_\tau R_\tau R_{\sigma',T} \preccurlyeq a^2 R_\sigma R_\sigma^T +$
$b^2 \left\| R^T_\tau R_\tau \right\| R^T_{\sigma',T} R_{\sigma',T}$

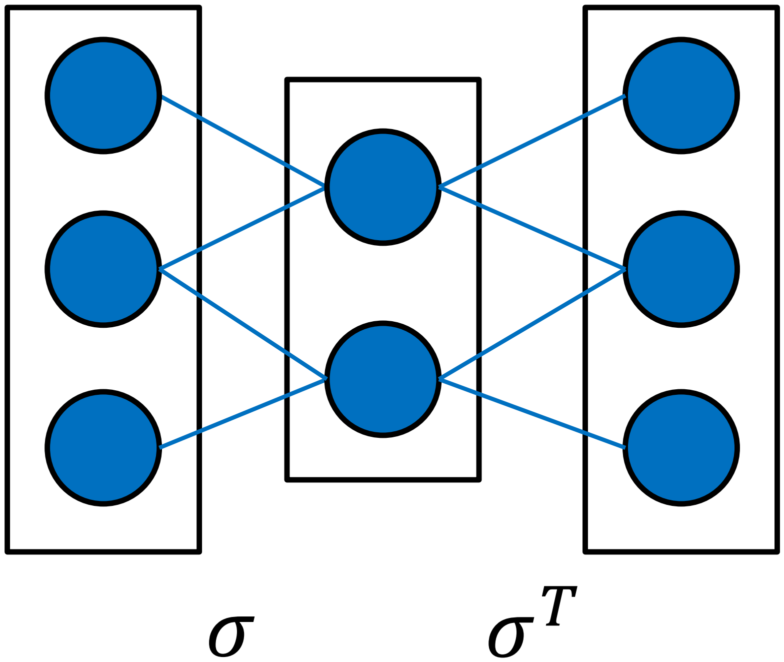
Example

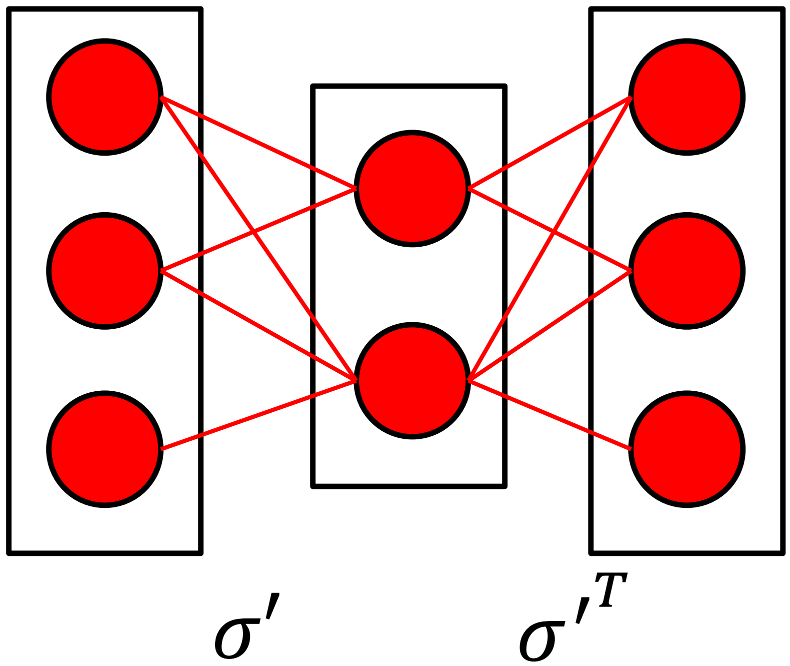
- What square terms would the following R_H be bounded by (ignoring intersection terms)?



Example Answer

- Answer: Take the left part and its mirror image and take the right part and its mirror image





Bounding With Square Terms Failure

- Unfortunately, the coefficients on the square terms aren't high enough for this idea to work.
- We need a more sophisticated analysis.

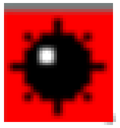
Part V: Approximate PSD Decomposition

LQL^T factorization

- Definition: Define $L_r = \sum_{\sigma:|V_\sigma|=r} \left(\frac{k}{n}\right)^{|V(\sigma)|-\frac{r}{2}} R_\sigma$ and define $Q_r = \sum_{\tau:|U_\tau|=|V_\tau|=r} \left(\frac{k}{n}\right)^{|V(\tau)|-r} R_\tau$ where we require that V_σ is the unique minimum vertex separator of σ and U_τ, V_τ are the leftmost and rightmost minimum vertex separators of τ. Define $M^{fact} = \sum_{r=0}^{\frac{d}{2}} L_r Q_r L_r^T$
- Claim: $M \approx M^{fact} = \sum_{r=0}^{\frac{d}{2}} L_r Q_r L_r^T$

Claim Justification

- Claim: $M \approx M^{fact} = \sum_{r=0}^{\frac{d}{2}} L_r Q_r L_r^T$
- This follows from the decomposition of each H into left, middle, and right parts $\sigma, \tau, {\sigma'}^T$ and the claim that up to intersection terms,



$$\left(\frac{k}{n}\right)^{|V(H)|} R_H = \left(\left(\frac{k}{n}\right)^{|V(H)|-\frac{r}{2}} R_\sigma\right)\left(\left(\frac{k}{n}\right)^{|V(H)|-r} R_\tau\right)\left(\left(\frac{k}{n}\right)^{|V(H)|-\frac{r}{2}} R_{{\sigma'}^T}\right)$$

Analysis of Q_r

- $Q_r = \sum_{\tau:|U_\tau|=|V_\tau|=r} \left(\frac{k}{n}\right)^{|V(\tau)|-r} R_\tau$
- Probabilistic norm bounds: With high probability, $\|R_\tau\|$ is $\tilde{O}(n^{\frac{|V(\tau)|-r}{2}})$ because r is the size of the minimum vertex separator of H
- Corollary: If $k \le n^{\frac{1}{2}-\epsilon}$ then with high probability, $Q_r \succcurlyeq \frac{1}{2} Id$ as the identity is the dominant term of Q_r

Summary

- If $k \leq n^{\frac{1}{2}-\epsilon}$ then with high probability,

$$M^{fact} = \sum_{r=0}^{\frac{d}{2}} L_r Q_r L_r^T \succcurlyeq \frac{1}{2} \sum_{r=0}^{\frac{d}{2}} L_r L_r^T$$

- The $\frac{1}{2} \sum_{r=0}^{\frac{d}{2}} L_r L_r^T$ allows us to deal with the error $M^{fact} - M$.

Part VI: Further Work and Open Problems

Further Work

- The techniques used for planted clique can be generalized to other planted problems where we are trying to distinguish a planted distribution from a random distribution [HKP+17]

Open Problems

- Can we prove the full lower bound for planted clique with the exact constraint that $\sum_{i=1}^{n} x_i = k$?
- How close to $\sqrt{n}$ can we make the lower bound?
- It turns out that the current machinery doesn't work as well for random sparse graphs. What bounds can we prove for problems such as densest k-subgraph and independent set on sparse graphs?

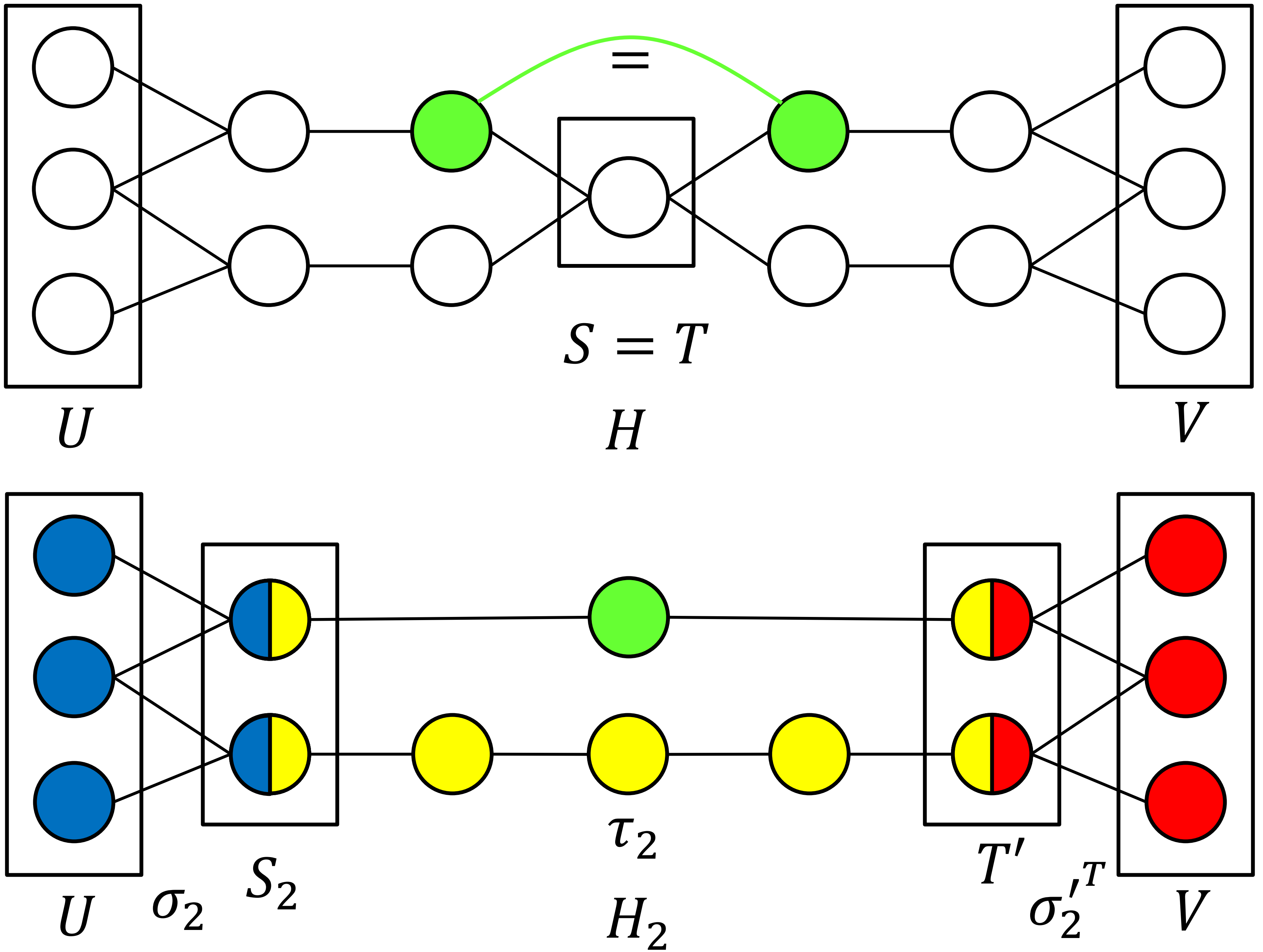
References

- [BHK+16] B. Barak, S. B. Hopkins, J. A. Kelner, P. Kothari, A. Moitra, and A. Potechin, A nearly tight sum-of-squares lower bound for the planted clique problem, FOCS p.428–437, 2016.
- [HKP+17] S. Hopkins, P. Kothari, A. Potechin, P. Raghavendra, T. Schramm, D. Steurer. The power of sum-of-squares for detecting hidden structures. FOCS 2017

Appendix I: Handling Intersection Terms

High Level Idea

- If there are intersections between the left, middle, and right parts, this creates a new graph H_2.
- We can decompose H_2 into new left, middle, and right parts!


=
S = T
U
H
V
U
σ2
S2
τ2
H2
T′
σ2′T
V

Choosing New Separators

- How do we choose the new separators S' and T'?
- We take S' to be the leftmost minimum vertex separator between U and $\{\text{intersected vertices}\} \cup S$.
- Similarly, we take T' to be the rightmost minimum vertex separator between $\{\text{intersected vertices}\} \cup T$ and V.

Key Idea

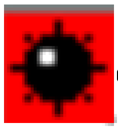
- This decomposition works the same regardless of what σ_2 and ${\sigma'_2}^T$ look like (see Claim 6.11 of [BHK+16])!
- Thus, we get a new approximate decomposition of the form $\sum_{r'=0}^{\frac{d}{2}} L_{r'} Q'_{r'} L_{r'}$
- This can be bounded by $\frac{1}{2}\sum_{r=0}^{\frac{d}{2}} L_r L_r^T$ as long as we always have that $\|Q'_{r'}\| \ll 1$

Bounding New Middle Parts

- We need to show that the new middle parts don't have norms which are too high.
- This is done with the intersection tradeoff lemma (Lemma 7.12 of [BHK+16])

Appendix II: Technical Mines

Approximate Decomposition Mine



- Claim: If r is the size of the minimum vertex separator of H,

$$R_H \approx R_\sigma R_\tau R_{\sigma',T}$$

- There are subtle issues related to the ordering of S and T, the leftmost and rightmost minimum vertex separators of H
- How these issues should be handled depends on whether we require matrix indices to be in ascending order.

Approximate Decomposition Mine

- If we require matrix indices to be in ascending order, what we actually have is $R_H \approx \sum_{\sigma,\tau,{\sigma'}^T : H=\sigma\cup\tau\cup{\sigma'}^T} R_\sigma R_\tau R_{{\sigma'}^T}$ where $\sigma \cup \tau \cup {\sigma'}^T$ is the graph formed by gluing $\sigma, \tau, {\sigma'}^T$ together.
- In fact, this equation is precisely what is needed for the approximate PSD decomposition $M \approx M^{fact} = \sum_{r=0}^{\frac{d}{2}} L_r Q_r L_r^T$.

Approximate Decomposition Mine

- Remark: [BHK+16] navigates this issue by keeping everything in terms of the individual ribbons (Fourier characters for a given matrix entry) until it is time to use the matrix norm bounds (see Definition 6.1 and subsection 6.4 of [BHK+16])

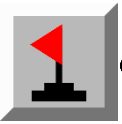
Approximate Decomposition Mine

- If we do not require matrix indices to be in ascending order, we actually have the following two equations
  1. $R_H \approx \left|Aut\left(\sigma,\tau,{\sigma'}^T\right)\right| R_\sigma R_\tau R_{{\sigma'}^T}$ where $\left|Aut\left(\sigma,\tau,{\sigma'}^T\right)\right|$ is the number of different ways to decompose H into $\sigma,\tau,{\sigma'}^T$.
  2. $R_H \approx \frac{1}{(s_H)!^2}\sum_{\sigma,\tau,{\sigma'}^T : H=\sigma\cup\tau\cup{\sigma'}^T} R_\sigma R_\tau R_{{\sigma'}^T}$ where $\sigma \cup \tau \cup {\sigma'}^T$ is the graph formed by gluing $\sigma,\tau,{\sigma'}^T$ together.

Truncation Mine

- Definition: Define $L_r = \sum_{\sigma:|V_\sigma|=r} \left(\frac{k}{n}\right)^{|V(\sigma)|-\frac{r}{2}} R_\sigma$ and define $Q_r = \sum_{\tau:|U_\tau|=|V_\tau|=r} \left(\frac{k}{n}\right)^{|V(\tau)|-r} R_\tau$ where we require that V_σ is the unique minimum vertex separator of σ and U_τ, V_τ are the leftmost and rightmost minimum vertex separators of τ. Define $M^{fact} = \sum_{r=0}^{\frac{d}{2}} L_r Q_r L_r^T$
- Actually, we need to truncate L_r and R_r by only taking σ, τ with at most D vertices

Truncation Mine



- Warning: There is a mismatch between H which have at most D vertices and triples $\sigma, \tau, {\sigma'}^T$ which each have at most D vertices.
- This truncation error turns out to have very small total norm, see Lemma 7.4 of [BHK+16]