SOS LOWER BOUND FOR EXACT PLANTED CLIQUE

SHUO PANG

Abstract. We prove a Sum-of-Squares (SoS) degree lower bound for the planted clique problem on Erdös-Rényi random graphs $G(n, 1/2)$. This problem is known to have a strong and a weak version, where the former allows SoS algorithms to use the clique-size constraint $x_1+...+x_n = \omega$ (and all polynomial identities generated by it). Our degree lower bound is for the strong version, in the form $d = \Omega(\epsilon^2 \log n/ \log\log n)$ as long as $\omega = O(n^{1/2-\epsilon})$. Improving upon [FK03, MPW15, DM15, HKP$^+$18, BHK$^+$19], this settles the strong planted clique problem almost optimally in both d and ω.

For techniques, we design pseudo-expectations in a way that is different from the popular pseudo-calibration. The analysis proceeds with the Johnson schemes-based method [MPW15] and the approximate factorization technique [BHK$^+$19], combined into use through certain combinatorial transforms and a special family of Hankel matrices. As a technical by-product, we also get a new perspective on the pseudo-expectation in the weak version of the problem.

Department of Mathematics, University of Chicago spang@uchicago.edu
The preliminary version of this paper appeared at the *36th Computational Complexity Conference (CCC'21)* [Pan21].

1. Introduction

Can one efficiently find a max-size clique in a random graph $G \sim G(n, 1/2)$? This has been a long-standing open problem since its introduction in [Kar76]. A variant was proposed in [Jer92, Kuč95], known as the *planted clique problem*: if additionally plant a random clique of size $\omega \gg \log n$ to G, can it be efficiently recovered? Information-theoretically this is solvable, since w.h.p. the largest clique in G has size $(2 + o(1)) \log n$. Computationally, the planted clique problem is widely believed to be hard on average, and it has been intensively studied, inspiring a broad range of research directions (cryptography [ABW10], learning [BR13], mathematical finance [ABBG11], computational biology [PS$^+$00], etc.). So far, the best known polynomial-time algorithm works only when $\omega = \Omega(\sqrt{n})$ [AKS98], which is a so-called spectral algorithm (see e.g. [HKP$^+$17]).

The *sum-of-squares hierarchy* (SoS) [Sho87, Par00, Las01] is a stronger family of semidefinite programming (SDP) algorithms which, roughly speaking, is SDP on the extended set of variables $\{x_{i(1)}...x_{i(d)} \mid i_1, ..., i_d \in [n]\}$ according to the degree parameter d, and it can be significantly more powerful than spectral algorithms and traditional SDPs (see e.g. [BBH$^+$12, HKP$^+$17]). Recent years have witnessed rapid development on SoS-based algorithms, which turns out to provide a characterization of a large class of algorithmic techniques ([BS14, HKP$^+$17]). The SoS proof system is the natural proof-theoretic counterpart of SoS algorithms, also known as the *Positivstellensatz* system [GV01]: it works with polynomials over $\mathbb{R}$, and given polynomial equalities (axioms) $f_1(x) = 0, ..., f_k(x) = 0$ on $x = (x_1, ..., x_n)$, a proof (that is, a refutation of the existence of a solution) is $-1 = \sum_{i=1}^{k} f_i q_i + \sum_j r_j^2$ in $\mathbb{R}[x_1, ..., x_n]$ where $q_1, ..., q_k$, $r_1, ...$ are arbitrary polynomials on $x_1, ..., x_n$ over $\mathbb{R}$. Under certain conditions, in particular when all variables are boolean ($x_i^2 = x_i$), such an refutation exists if the axioms have no solution. The *degree-d SoS proof system* carries the obvious additional degree restriction, $\max_{i,j}\{\deg(f_i) + \deg(q_i),\ 2\deg(r_j)\} \le d$. See [O'D17, RW17] for more on the relation between SoS proofs and algorithms.

The average-case hardness of the planted clique problem has a very simple form in proof complexity: for $G \sim G(n, 1/2)$, can the proof system efficiently refute the existence of a size-ω ($\gg \log n$) clique w.h.p.? A lower bound would automatically give the hardness on any class of algorithms based on the proof system. Given that the decision version of the spectral algorithm of [AKS98] corresponds to a degree-2 SoS proof, a SoS degree lower bound potentially can bring us a much better understanding of algorithmic hardness. The standard problem formulation is the following.

Definition 1.1. *Given an n-vertex simple graph G and a number ω, the* **Clique Problem** *for degree-d SoS proof system has the following* **axioms**.

$$\begin{aligned} &\text{(Boolean)} && x_i^2 = x_i \quad \forall i \in [n] \\ &\text{(Clique)} && x_i x_j = 0 \quad \forall \{i, j\} \text{ non-edge} \\ &\text{(Size)} && x_1 + ... + x_n = \omega \end{aligned} \tag{1.1}$$

SoS system has a duality, i.e. to show degree lower bound, it suffices to find a *pseudo-expectation* whose *moment matrix*[1] is positive semi-definite (PSD). With boolean variables (which is our case), this can be demonstrated on multi-linear polynomials as below. Let $\mathcal{X}^{\le d} = \{x_S \mid S \subseteq [n],\ |S| \le d\}$ for any $d \in \mathbb{N}$.

Definition 1.2. *A* **degree-d pseudo-expectation** *for the Clique Problem on G is a map $\widetilde{E} : \mathcal{X}^d \to \mathbb{R}$ satisfying the following four* **constraints** *when extended by*

[1]The name is simplified; more cautiously, it should be called the *pseudo-moment matrix*.

$\mathbb{R}$-*linearity.*

$$\widetilde{E}x_{\emptyset} = 1 \tag{1.2}$$ (Default)

$$\widetilde{E}x_S = 0, \quad \forall S : |S| \le d,\ G|_S \text{ non-clique} \tag{1.3}$$ (Clique)

$$\widetilde{E}\Big((x_1 + ... + x_n)x_S\Big) = \omega \cdot \widetilde{E}x_S \quad \forall S :\ |S| \le d-1 \tag{1.4}$$ (Size)

where in (1.4), $x_A \cdot x_B := x_{A\cup B}$. *To define the last constraint, define the* **moment matrix** M *to be the* $\binom{[n]}{\le d/2} \times \binom{[n]}{\le d/2}$ *matrix*[2] *with* $M(A,B) = \widetilde{E}x_{A\cup B}$, *then:*

(PSDness) M *is positive semi-definite.* (1.5)

It is not hard to see that if a degree-d pseudo-expectation exists then there is no degree-d SoS refutation.

A relaxation of the problem was studied in [BHK$^+$19], asking if an $\widetilde{E}$ as above exits except with one change: replace Size Constraints by a single inequality $\widetilde{E}(x_1 + ...+x_n) \ge \omega$. Henceforth, we call the Clique Problem (Definition 1.1) **Exact Clique** and this relaxation **Non-Exact Clique**.[3]

How to deal with the exact problem is a subtle but important open problem. On the problem itself, lower bounds on the non-exact (weak) formulation indeed gave the important algorithmic message, but they do not rule out the possibility that SoS with additional constraint $x_1 + ... + x_n = \omega$ can output "infeasible" (cf. the similar situation for random CSP [KOS18]). The distinction between "weak" and "strong" formulations also involves how one thinks *the* SoS SDP optimization problem should be formulated.

Perhaps more importantly, it is about the techniques of proving average-case SoS lower bounds. Current techniques from the so-called *pseudo-calibration heuristic* [BHK$^+$19] tend to deal successfully with "soft" constraints (inequalities, or usually just one bound on some pseudo-expectation value) while being poor at handling "hard" constraints (equalities). Finding techniques to deal with the latter is in need. Progress toward this goal is made in [KOS18] for random CSPs, where the idea is that the hard constraint is a sum of *local constraints*, each can be satisfied by a real distribution on *local* variables; the locality is formed by a notion of *graph closure* which in turn is based on expanders (cf. [Gri01, Sch08, BGMT12, BCK15, KMOW17]). For Exact Clique whose constraints do not have a similarly clear global-local structure, it seems unlikely a similar strategy could work.

Lastly, there are concrete applications of a lower bound on Exact Clique to other problems, e.g. the approximated Nash-Welfare [KM18]. Techniques for proving such a bound might also help deal with closely related problems like the coloring problem and stochastic block models [KOS18, KM21, JPR$^+$21].

1.1. **Previous work.** For lower bounds, on Exact Clique, [FK03] showed that the (weaker) *d-round Lovasz-Schrijver* system cannot refute it for $\omega = O(\sqrt{n/2^d})$, [MPW15] proved degree-d lower bound on SoS for $\omega = \widetilde{O}(n^{1/d})$ which was later improved to $\widetilde{O}(n^{1/3})$ for $d = 4$ [DM15] and further to $\widetilde{O}(n^{\frac{1}{\lfloor d/2\rfloor+1}})$ for general d [HKP$^+$18]. For Non-Exact Clique, [BHK$^+$19] proved the almost tight lower bound $d = \Omega(\epsilon^2 \log n)$ for $\omega = n^{1/2-\epsilon}$, $\epsilon > 0$ arbitrary.

[2] d is always assumed to be even.

[3] There is no "planted clique" in the problem formulation now, but traditionally this is still called the planted clique problems due to the algorithmic motivation.

For upper bounds, if $\omega = \Omega(\sqrt{n})$ then degree-2 SoS can refute Exact Clique with high probability [FK00]. On the other hand, if $\omega > d \geq 2.1 \log n$, a degree-$d$ SoS refutation for Exact Clique is not hard to see, which we include below for completeness.

Observation 1.1. *(Upper bound for Exact Clique if $\omega > d = 2.1 \log n$) Note that $(x_1 + ... + x_n)^d = \omega^d$ modulo the Size Axiom. The LHS can be multi-linearly homogenized to degree-d by $x_S = \frac{1}{\omega - |S|} \sum_{i \notin S} x_{S \cup \{i\}}$ by this axiom again, after which w.h.p. all terms are 0 by Clique Axioms since there is no size-$2.1 \log n$ clique in $G \sim G(n, 1/2)$ w.h.p.. This gives the contradiction $0 = 1$. Note the proof is actually in the weaker system Nullstellensatz (see e.g.* [BIK+96]*).*

1.2. **Results of the paper.** Our main result is the following.

Theorem 1. *Let $\epsilon > 0$ be any parameter, $\omega = n^{1/2-\epsilon}$. W.p. $> 1 - n^{-4 \log n}$ over $G \sim G(n, \frac{1}{2})$, any SoS refutation of Exact Clique requires degree at least $\epsilon' \log n / \log \log n$, where $\epsilon' = \min\{\epsilon^2, \frac{1}{40^2}\}/2000$.*

We also have the following result. It does not allow to improve the lower bound but provides a new, hopefully simplifying, perspective on certain techniques that were used for the non-exact problem.

Theorem 2. *(Informal) For the Non-Exact Clique problem,*

(1). There is a way to define the correct pseudo-expectation from simple incidence algebra on the vertex-set;

(2). For the resulting moment matrix M, there is a weakened version of the quadratic equation $M = NN^\top$ whose solvability is given by, and actually equivalent to, a general graph-decomposition fact from which a "first-approximate" diagonalization of M can be deduced.

2. Key technical ideas

This section is an overview of the proofs. The two results use almost completely different ideas, so we treat them separately in the proof overview:

- Theorem 1: section 2.1 to 2.4.
- Theorem 2: section 2.5.

More precisely, (where "→" points to the sections in the actual proof)

Pseudo-expectation design:
- A common idea (the paragraph below)
- Non-exact case (sec. 2.5 first half → sec. 3.1)
- Exact case (sec.2.1 → sec. 3.2)

Proving PSDness:
- Recursive factorization (→ sec. 6.3)
- Lower bound proof (sec. 2.1 to 2.4 → sec. 5 to 8)

We also give a conceptual simplification of the analysis in [BHK+19], which can be read independently:

- Deduction of the coarse diagonalization (sec. 2.5 second half → sec. 6.2).

Let us start with the pseudo-expectation design. Suppose we deal with degree-d SoS i.e. deal with size $\leq d$-subsets of $[n]$, then as the common idea in complexity theory, we take a parameter $\tau \gg d$ (think of $d \ll \tau \ll \log n$) and make our construction on all size $\leq \tau$-subsets, in hope to later have a good control on its behavior on all size $\leq d$ subsets. This idea is most clearly demonstrated in the non-exact case (section 3.1.2), and is also the reason for the τ-parameter for the exact case (in (2.1) below).

2.1. **The exact pseudo-expectation.** The constraints force us to design the pseudo-expectations in a top-down manner, as follows. Fix $\widetilde{E}x_S$ for all $|S| = d$ first, then recursively set $\widetilde{E}x_S \leftarrow \frac{1}{\omega-|S|}\sum_{i\notin S}\widetilde{E}x_{S\cup\{i\}}$ if $|S| < d$. The Clique Constraints (1.3) will be satisfied if $\widetilde{E}_d x_S$ factors through $1_{G \text{ is clique on } S}$ as functions on G. Inspired by (almost all) previous work in the literature, we use Fourier characters and consider

$$\widetilde{E}x_S = \sum_{T:|V(T)\cup S|\le\tau} F(|V(T)\cup S|)\cdot\chi_T \quad \forall S\subseteq[n],\ |S| = d \tag{2.1}$$

for some function $F:\mathbb{N}\to\mathbb{R}$. We call F a d-**generating function**.[4] Thus

$$\widetilde{E}x_S = \frac{1}{\binom{w-d+u}{u}}\sum_{T:|V(T)\cup S|\le\tau}\chi_T\cdot\Bigg[\sum_{c=0}^{u}\binom{|V(T)\cup S|-d+u}{c}\binom{n-|V(T)\cup S|}{u-c} \cdot F(|V(T)\cup S|+u-c)\Bigg]$$

where $u := d-|S|$, for all S with $|S|\le d$. One key novelty we bring is the choice

$$F(x) = \frac{(x+8\tau^2)!}{(8\tau^2)!}\cdot(\frac{\omega}{n})^x. \tag{2.2}$$

The moment matrix $\widetilde{M}$ will be $\widetilde{M}(A,B) = \sum_{T:|V(T)\cup A\cup B|\le\tau}\widetilde{M}(A,B;T)\chi_T$ for $A,B\subseteq[n]$, $|A|,|B|\le d/2$, where $\widetilde{M}(A,B;T) =$

(2.3)

$$\frac{1}{\binom{\omega-d+u}{u}}\Bigg[\sum_{c=0}^{u}\binom{|V(T)\cup A\cup B|-(d-u)}{c}\binom{n-|V(T)\cup A\cup B|}{u-c} \cdot\underbrace{\frac{(|V(T)\cup A\cup B|+u-c+8\tau^2)!}{(8\tau^2)!}\cdot(\frac{\omega}{n})^{|(V(T)\cup A\cup B)|+u-c}}_{F\left(|V(T)\cup A\cup B|+u-c\right)}\Bigg],$$

where $u = d-|A\cup B|$.

This seemingly mysterious choice of F is ultimately for proving the PSDness of $\widetilde{M}$, but it seems can be seen only after a series of technical transformations (Remark 2.1, 3.3). It will be very interesting to know if there is *a priori* an explanation of it. See Remark 3.2, 7.1 for why some traditional choices from the literature that simulate some planted distributions seem cannot work here.

2.2. **Hadamard decomposition and Euler transform.** For the exact problem, using a standard SoS homogeneity reduction (Lemma 4.1), it suffices to prove PSDness of the $\binom{[n]}{d/2}\times\binom{[n]}{d/2}$ principal minor of $\widetilde{M}$. Denote this minor by M. One unpleasant feature of M is that in its expression (2.3) the parameter $u = |A\cap B|$ appears in a deeply nested way. To analyze M (in particular, get a clue of how to diagonalize it), we resolve this intricacy in two steps.

First, $M = \sum_{c=0}^{\frac{d}{2}} m_c\circ M_c$ where "$\circ$" is the Hadamard product, and m_c, M_c are matrices as follows. For all $|I|,|J| = d/2$,

$$m_c(I,J) = \frac{1}{\binom{\omega-d+u}{u}}\omega^{u-c} \quad \text{where } u \text{ denotes } |I\cap J|; \tag{2.4}$$

[4]To be distinguished from the usual generating functions for sequences.

(2.5)
$$M_c(I,J) = \begin{cases} \sum_{T:|V(T)\cup I\cup J|\le\tau} \chi_T \cdot M_c(|I\cap J|, |V(T)\cup I\cup J|), & \text{if } |I\cap J|\ge c; \\ 0, & \text{o.w.} \end{cases}$$

whose coefficients are

$$M_c(u,a) = \binom{a-(d-u)}{c}\binom{n-a}{u-c} n^{-(u-c)} \frac{(a+u-c+8\tau^2)!}{(8\tau^2)!} (\frac{\omega}{n})^a$$

where $u = |I\cap J|$, $a = |V(T)\cup I\cup J|$.

The intuition is to let m_c carry (as much as possible) the "purely" vertex-set-based information, $|I\cap J|$, so that the second factor M_c will be left with (mainly) edge-set-based information. As can be expected, in the analysis we will also treat m_c, M_c's separately.

The more difficult part is M_c. In fact, we will further remove the dependence on $|I\cap J|$ in $M_c(I,J)$ by one more step: a decomposition $M_c = \sum_{R\in\binom{[n]}{\le \frac{d}{2}}} M_c^R$ where each M_c^R is supported on rows and columns whose index contains R, and the expression of M_c^R's can be derived from M_c by *Euler transform*. In summary, we will prove:

Lemma 2.1. *(ΣΠ-decomposition of M, Lemma 5.2)*

$$M = \sum_{c=0}^{\frac{d}{2}} m_c \circ \left(\sum_{R\in\binom{[n]}{\le d/2}} M_c^R\right) = \sum_{R\in\binom{[n]}{\le d/2}} \left(\sum_{c=0}^{|R|} m_c \circ M_c^R\right) \tag{2.6}$$

where each m_c is by (2.4) and M_c^R has the following expression. First, $M_c^R = 0$ if $|R| < c$. Also if $R \not\subseteq I\cap J$ then $M_c^R(I,J) = 0$. Finally, if $|R|\ge c$ and $R\subseteq I\cap J$, then $M_c^R(I,J) = \sum_{T:|V(T)\cup I\cup J|\le\tau} M_c^R(I,J;T)\chi_T$ where, denoting $a = |V(T)\cup I\cup J|$, $M_c^R(I,J;T) = (\frac{\omega}{n})^a \cdot Y_c(|R|,a)$ and

$$Y_c(r,a) = \begin{cases} \sum_{l=c}^{r} (-1)^{r-l}\binom{r}{l}\binom{a+l-d}{c}\binom{n-a}{l-c} n^{-(l-c)} \frac{(a+l-c+8\tau^2)!}{(8\tau^2)!}, & \text{if } r\ge c; \\ 0, & \text{o.w.} \end{cases} \tag{2.7}$$

Moreover, for all $0\le c\le r\le d/2$ and $0\le a\le\tau$, $|Y_c(r,a)| < \tau^{5\tau}$.

Intuition for analysis. To analyze (2.6), the intuition is that the first factor m_c "decreases" in c and m_0 is "very positive"; while for fixed R, M_0^R is positive and other M_c^R's ($c>0$) are "not too large". This is expounded by the following two lemmas.

Lemma 2.2. *For each $c = 0, ..., d/2$, $m_0 = \omega m_1 = ... = \omega^{\frac{d}{2}} m_{\frac{d}{2}} \succ \frac{d}{2\omega}\mathrm{Id}$.*

Lemma 2.3. (Main Lemma) *In (2.6), w.p. $> 1 - n^{-5\log n}$ the following hold. For all $R\in\binom{[n]}{\le d/2}$, let $P^R = \{I\in\binom{[n]}{d/2} \mid R\subseteq I\}$,*

$$(1).\quad M_0^R \succeq n^{-d}\mathrm{diag}(\widetilde{\mathrm{Cl}})_{P^R\times P^R}; \tag{2.8}$$

$$(2).\quad \pm\omega^{-c} M_c^R \preceq n^{-c/6}\cdot M_0^R, \quad \forall 0 < c \le |R|. \tag{2.9}$$

These two lemmas directly imply $M(G) \succeq n^{-d-1}\mathrm{diag}(\widetilde{\mathrm{Cl}}(G))_{\binom{[n]}{d/2}\times\binom{[n]}{d/2}}$ w.h.p., and Theorem 1 is an easy corollary of this (Cor. 5.1, 5.2).

The proof of Lemma 2.2 is relatively easier using *Johnson schemes* (similar to [MPW15], see Lemma 5.1). Below we demonstrate the idea for proving the Main Lemma.

2.3. **Recursive factorization: an extension.** To prove the Main Lemma, an important step is to derive an approximate diagonalization of M_c^R, for which we use the *recursive factorization* technique of [BHK+19]. In section 7 we will derive an approximate PSD factorization $M^R \approx \widetilde{L^R}\,(-)\left(\widetilde{L^R}\right)^\top$. This we roughly describe as below.

Lemma 2.4. *(Recursive factorization, Lemma 7.2) For any $R \in \binom{[n]}{\le d/2}$ and $0 \le c \le |R|$, we have the following decomposition.*

$$M_c^R = \widetilde{L^R} \cdot \left[D^\tau \left(Q_{c,0}^R - Q_{c,1}^R + ... \pm Q_{c,d}^R \right) D^\tau \right] \cdot \left(\widetilde{L^R}\right)^\top + \mathcal{E}_c^R \tag{2.10}$$

where $\widetilde{L^R}$ is some matrix of dimension $\binom{[n]}{\frac{d}{2}} \times \left(\binom{[n]}{\le \frac{d}{2}} \times (\tau+1) \right)$, D^τ is the diagonal matrix $\mathrm{diag}\left((\frac{\omega}{n})^{\frac{|A|}{2}} \right)_{A\subseteq[n],|A|\le d/2} \otimes \mathrm{Id}_{\{0,...,\tau\}\times\{0,...,\tau\}}$, and the "middle matrices" $Q_{c,k}^R$'s are $(\tau+1)\times(\tau+1)$-blocked, each block of dimension $\binom{[n]}{\le \frac{d}{2}} \times \binom{[n]}{\le \frac{d}{2}}$. The "error" $\mathcal{E}_c^R(G)$ is supported within rows and columns that contains R and is clique (given G), and w.p. $> 1 - n^{-9\log n}$, $\left\|\mathcal{E}_c^R\right\| < n^{-\epsilon\tau/2}$.

For the reason of the "larger" dimension of matrices, see Remark 7.1.

2.4. **Proving PSDness: encounter with Hankel matrices.** With Lemma 2.2 and 2.4 at hand, the following is the key step towards the Main Lemma.

Lemma 2.5. *W.p. $> 1 - n^{-8\log n}$ over G, the following holds: for all $R \in \binom{[n]}{\le d/2}$,*

(1). $Q_{0,0}^R - Q_{0,1}^R + ... \pm Q_{0,\frac{d}{2}}^R \succeq \tau^{-7\tau} \cdot \mathrm{diag}\left(\widetilde{\mathrm{Cl}}\right)_{S^R\times S^R}$, where $S^R = \{(A,i) \in \binom{[n]}{\le d/2} \times \{0,...,\tau\} \mid A \supseteq R, |A| + i \ge \frac{d}{2}\}$;

(2). $\forall 0 < c \le |R|$, $\pm\omega^{-c}\left(Q_{c,0}^R - Q_{c,1}^R + ... \pm Q_{c,\frac{d}{2}}^R \right) \preceq n^{-c/4} \cdot \mathrm{diag}\left(\widetilde{\mathrm{Cl}}\right)_{S^R\times S^R}$.

To prove this lemma, modulo somewhat standard steps (three Lemmas 8.2, 8.3, 8.4) the final technical challenge is: *show the positiveness of* $\mathbb{E}[Q_{0,0}^R]$ (Corollary 8.1).

We describe below how this is done. After simplification, the real task is to analyze the positiveness of the following matrix[5]:

$$\sum_{l=0}^{r} (-1)^{r-l} \frac{\binom{r}{l}}{l!} \cdot H_{\tau,\ l+8\tau^2} \quad \text{for any } 0 \le r \le d/2 \tag{2.11}$$

where $\{H_{m,t}\}$ is the family of $(m+1)\times(m+1)$-matrices

$$H_{m,t}(i,j) = (i+j+t)! \quad \forall 0 \le i,j \le m.$$

This is a special family of the so-called *Hankel matrices* whose (i,j)th element depends only on $i+j$. General Hankel matrices seem to arise naturally in moment problems but are notoriously wild-behaving in many aspects (see e.g. [Tyr94]). Fortunately enough, for the special family here we can manage to get a relatively fine understanding; we term this family **factorial Hankel matrices**. The key observation is that they have a concrete recursive diagonalization (Proposition 8.2), resulting in the following.

Proposition 2.1. *If parameters m, t, r satisfy*

$$t + 1 > 8 \cdot \max\{r^2, m\}, \tag{2.12}$$

then $H_{m,t+1} \succeq 2r^2 H_{m,t}$.

Remark 2.1. *Condition* (2.12) *is why "$8\tau^2$" is used in the numerator of F,* (2.2).

[5]The subscripts are not exactly as in the problem but suffice to demonstrate the spirit.

With this proposition, it is relatively easy to complete the proof of the Lemma 2.5, hence the Main Lemma. This completes the proof overview of Theorem 1.

2.5. **Ideas for Theorem 2.** We now explain the idea behind Theorem 2, that is, a different perspective on the techniques used in the non-exact problem.

On defining the pseudo-expectation. Previously, the pseudo-expectation is obtained via the so-called *pseudo-calibration* method. We define the same $\widetilde{E}$ but from a different perspective, the incidence algebra on the vertex-set, which is also a simple refinement of the construction in [FK03].

The ζ-*matrix* on $[n]$ is the $2^{[n]} \times 2^{[n]}$ 0-1 matrix with $\zeta(A, B) = 1$ iff $A \subseteq B$. We observe that ζ reveals the basic linear structure of the true expectation on cliques if G is a single planted clique. So we use ζ to define $\widetilde{E}$. That is, we define a *degree-τ* approximate-distribution vector $p_\tau(G)$ first—it approximates the distribution of τ cliques inside a planted clique, with a standard twist so that it is only supported on cliques of the given G (3.5)—then take the vector $\zeta_{d,\tau} \cdot p_\tau(G)$ as $\widetilde{E}x$ (Def. 3.3). Here, $(\cdot)_\tau$ means to truncate the matrix or vector to indices whose size $\leq \tau$. In this way, $\widetilde{E}$ inherits the linear structure posed by ζ too.

On deducing the first-approximate diagonalization. The goal is to come up with a coarse, "first-approximate" diagonalization of the moment matrix. We deduce its form in two steps: 1. Analyze the expectation of the matrix; 2. Observe that the (imaginary) diagonalization of the matrix is in essence a quadratic equation, which we weaken to a proper "modular" version to solve.

We call step 2 the **mod-order analysis** (section 6.2), whose underlying idea is inspired by and similar to the more broad dimension-analysis in physical sciences: weaken the equation to its most significant part in a well-defined way (Def. 6.1). One ingredient towards defining the weakening is the norm information on certain pseudo-random matrices (the *graph matrices*).

The resulting weakened equation has a nice structure to work with (Lem. 6.2, Cor. A.1). Using standard techniques for studying algebraic equations—actually a simple *polarization* (Appendix A.2)—we can deduce a solvability condition for the polarized equation, which translates to the existence of a general graph-theoretic structure (equation (A.19) and Fact A.1). The "coarse" diagonalization is then formulated based on this structure.

To demonstrate in more detail, it suffices to concentrate on the $\binom{[n]}{d/2} \times \binom{[n]}{d/2}$-minor of the moment matrix, denoted by M':

$$M'(I, J) = \sum_{T:|V(T)\cup I\cup J|\leq\tau} (\frac{\omega}{n})^{|V(T)\cup I\cup J|}\chi_T, \quad \forall I, J : \ |I| = |J| = d/2.$$

Step 1: expectation. By using *Johnson schemes* as in [MPW15], we get an explicit decomposition $\mathbb{E}[M'] = CC^\top$ where C is $\binom{[n]}{d/2} \times \binom{[n]}{\leq d/2}$, and actually with a fine understanding of the spectrum of $\mathbb{E}[M']$.

Step 2: mod-order analysis. Given $\mathbb{E}[M'] = CC^\top$ from Step 1, ideally we hope to solve the quadratic matrix equation

$$M' = NN^\top \tag{2.13}$$

in N with $\mathbb{E}[N] = C$, and N extending C by non-trivial Fourier characters. Two observations about (2.13) follow.

(1) Order in $\frac{\omega}{n}$. Entries of M' all have a clear order in $\frac{\omega}{n}$. Like in fixed-parameter problems, we treat $\frac{\omega}{n}$ as a distinguished structural parameter and try to solve the correct power of $\frac{\omega}{n}$ in N first.

(2) Norm-match. A closer look into $CC^\top$ shows

$$\left\|C_rC_r^\top\right\| \approx \binom{d/2}{r} \cdot (\frac{\omega}{n})^{d-r} n^{d/2-r}, \quad r = 0, ..., d/2, \tag{2.14}$$

where $C = (C_0, ..., C_{d/2})$, each C_r having column dimension $\binom{[n]}{r}$. Assume $N = (N_0, ..., N_{d/2})$. Then we expect $N_rN_r^\top$ to concentrate around $C_rC_r^\top$ for each r, and so expect the norm of the non-constant part of $N_rN_r^\top$ to be bounded by (2.14). Under this condition, the known tight norm bounds on related matrices would tell us, for each possible appearing term in N, the least order of $\frac{\omega}{n}$ in its coefficient.

With these observations, we can weaken equation (2.13) to a simple "modular version" that is more informative about the (imaginary) solution N. Namely, abstract $(\frac{\omega}{n})$ as a fresh variable α and work in ring $\mathbb{R}[\alpha, \{\chi_T\}]$, consider

$$(M' \mod \text{high order}) = (N \mod \text{high order}) \cdot (N^\top \mod \text{high order}) \tag{2.15}$$

where "order" means power of α (think of α as an "infinitesimal"). We call (2.15) the *mod-order equation* and its analysis the *mod-order analysis*. For details see Definition 6.1.

We feel that this approach leads us more naturally to the realization of using the graph-theoretic structure beyond guesses, and the simple general idea behind the mod-order analysis might hopefully find other applications.

2.6. **Structure of the paper.** In section 3 we define the pseudo-expectations and show Theorem 2(1). In section 4 we recall some fundamental tools for analysis. The proof of the main theorem consists of three steps: section 5 is the first step (combinatorial transforms), section 6 and 7 is the second step (recursive factorization, where in 6 we will refresh the technique of recursive factorization and show Theorem 2(2)), and section 8 is the last step (structural and pseudo-random matrices). The paper is concluded in section 9 with open problems.

Notation. I, J, A, B, S will be used to denote vertex-sets, and T for edge-sets. $E(S) := \binom{S}{2}$. G denotes a simple graph on the vertex-set $[n]$. "$T \subseteq E([n])$" will be omitted in summation when there is no confusion. "$\sqcup$" means disjoint union. Finally, we use $y(n) = O(x(n))$ to mean that there is some absolute constant c s.t. $y(n) \leq cx(n)$ for all n.

Parameter regime. Throughout the paper,

$$\begin{aligned}
&\epsilon = \text{any positive parameter (wlog } \epsilon < \frac{1}{40});\\
&\omega = n^{1/2-4\epsilon};\\
&\tau = \frac{\epsilon}{200} \log n / \log\log n;\\
&d = \frac{\epsilon}{100}\tau.
\end{aligned}$$

3. Pseudo-expectations

As a warm-up, in section 3.1 we construct the non-exact pseudo-expectation. In section 3.2 we give the construction for the exact case.

3.1. **Non-exact case: a new perspective.** Given a graph G we can think of a degree-d pseudo-expectation as assigning a number $\widetilde{E}x_S$ to each subseteq $S \subseteq [n]$ of size $\leq d$, so that the resulting vector $\widetilde{E}x$ *looks indistinguishable* to the expectation resulted from the case when a random-ω clique is planted, from the view of degree-d SoS. As explained at the beginning of section 2, to make such an assignment we first go beyond to slightly larger subsets of size τ. We define an "approximate

distribution" on size $\leq \tau$-cliques in G then use it to generate pseudo-expectation on all size $\leq d$-subsets.

3.1.1. *ζ-function and Möbius inversion.* Given n-vertex graph G, let $p(G) \in \mathbb{R}^{2^{[n]}}$ be the max-clique-indicator vector. Then $q(G) := \zeta \cdot p(G)$ is a vector supported exactly on all cliques in G, where ζ is the $2^{[n]} \times 2^{[n]}$ matrix

$$\zeta(A, B) = 1 \text{ iff } A \subseteq B, \quad \forall A, B \subseteq [n]. \tag{3.1}$$

In particular, if G is a single clique then $q(G)$ is the clique-indicator. We will use $\zeta_{a,b}$ to denote the submatrix of ζ on rows $\binom{[n]}{\leq a}$ and columns $\binom{[n]}{\leq b}$, and use similar notation on all related vectors. Consider when G is just a randomly planted clique. Its distribution can represented by a *plant-distribution* vector $p_{\text{plant}} \in \mathbb{R}^{2^{[n]}}$, and let the *output-expectation* q_{out} be the vector of cliques in G in expectation. Then $q_{\text{out}} = \zeta \cdot p_{\text{plant}}$. We call such a pair $(p_{\text{plant}}, q_{\text{out}})$ a **plant-setting**.

Definition 3.1. *(Two plant-settings) The* **exact plant-setting** (p_0, q_0) *is*

$$p_0(S) = \frac{1}{\binom{n}{\omega}} \text{ if } |S| = \omega \text{ and } 0 \text{ otherwise}, \quad q_0(S) = (\zeta p_0)(S) = \frac{\binom{n-|S|}{\omega-|S|}}{\binom{n}{\omega}}. \tag{3.2}$$

I.e. in this setting a random size-ω subseteq is chosen to be the planted clique.

The **independent plant-setting** (p_1, q_1) *is*

$$p_1(S) = (\frac{\omega}{n})^{|S|}(1 - \frac{\omega}{n})^{n-|S|}, \quad q_1(S) = (\zeta p_1)(S) = (\frac{\omega}{n})^{|S|} \tag{3.3}$$

for all $S \subseteq [n]$. I.e. any vertex is included in the planted clique w.p. $\frac{\omega}{n}$ independently.

Thus the matrix ζ reveals the basic linear relations between $(p_{\text{plant}}, q_{\text{out}})$. It is upper-triangular (with row- and column-indices ordered in a size-ascending way), invertible, with the inverse the **Möbius inversion** matrix: $\zeta^{-1}(A, B) = (-1)^{|B \setminus A|}$ if $A \subseteq B$, and 0 otherwise. Note $(\zeta_{a,a})^{-1} = (\zeta^{-1})_{a,a}$, $\forall a \leq n$. Moreover, if let the pseudo-expectation be defined as $\widetilde{E}x = p \in \mathbb{R}^{2^{[n]}}$ for some vector p, then the "full" $2^{[n]} \times 2^{[n]}$ moment matrix is

$$M_{SoS} = \zeta \text{diag}(p) \zeta^\top. \tag{3.4}$$

In particular, if p is a nonnegative vector then M_{SoS} is immediately PSD.

3.1.2. *Non-exact pseudo-expectation for (p_1, q_1).* Given G, we first construct a *degree-τ* "approximate plant-distribution", $p_\tau(G)$, that simulates the plant-distribution **and** that $p_\tau(G)$ is supported on size $\leq \tau$-cliques in G. Then we can take $\widetilde{E}x = \zeta_{d,\tau} \cdot p_\tau(G)$ so that the result inherits the linear structure posed by ζ.

What is this $p_\tau(G)$? From the view of approximation it seems taking $\zeta^{-1}_{\tau,\tau}(q_1)_\tau$ would suffice, while to make it supported on cliques, same as in [FK03] we add a clique-indicator factor, thus

$$p_\tau(G)(S) = \left(2^{|\binom{S}{2}|} \text{Cl}_S(G) \cdot \zeta^{-1}_{\tau,\tau}(q_1)_\tau\right)(S) \quad \forall S \subseteq [n] \text{ of size} \leq \tau \tag{3.5}$$

where $\text{Cl}_S(\cdot)$ is the clique indicator function and $2^{|\binom{S}{2}|}$ is for re-normalization.

Definition 3.2. *For any $S \subseteq [n]$, the* **scaled clique-indicator** *is $\widetilde{\text{Cl}}_S(G) := 2^{|\binom{S}{2}|}\text{Cl}_S(G)$, which is a function on G. $\widetilde{\text{Cl}}(G)$ is the (column) vector of them over a family of S's, which will always be clear from the context.*

Definition 3.3. *The* **non-exact pseudo-expectation** *is*

$$\widetilde{E}_{\text{nonexact}} = \zeta_{d,\tau} \cdot p_\tau(G) = \zeta_{d,\tau} \cdot (\widetilde{\text{Cl}}(G) \circ \zeta_{\tau,\tau}^{-1}) \cdot (q_1)_\tau \quad \in \mathbb{R}^{\binom{[n]}{\le d}} \tag{3.6}$$

where "$\circ$" is the Hadamard product[6]*.*

In short, $\widetilde{E}_{\text{nonexact}}$ refined the construction in [FK03] by one step: factor through size-τ subsets (in the *only* non-trivial way) so that the size-d output inherits linear relations posed by ζ. Similarly to (3.4), the resulting moment matrix is

$$M_{\text{nonexact}}(G) = \zeta_{d/2,\tau} \cdot \text{diag}\,(p_\tau(G)) \cdot (\zeta_{d/2,\tau})^\top. \tag{3.7}$$

Remark 3.1. *$\widetilde{E}_{\text{nonexact}}$ looks like a true expectation on cliques in G, namely, if $p_\tau(G)$ were nonnegative then the PSDness of $M_{\text{nonexact}}(G)$ would be immediate. Alas, this is not true by computation*[7]*. That the PSDness could still possibly hold is because $\zeta_{d/2,\tau}$ in* (3.7) *is degenerate.*

Lemma 3.1. *(Theorem 2(1)) For all $S \subseteq [n]$ s.t. $|S| \le d$,*

$$\widetilde{E}_{\text{nonexact}} x_S = \sum_{T : |V(T) \cup S| \le \tau} (\frac{\omega}{n})^{|V(T)\cup S|} \chi_T. \tag{3.8}$$

Proof. Note $\widetilde{\text{Cl}}_S = \sum_{T \subseteq E(S)} \chi_T$ for all S. Now for S, S' with appropriate size bound,

$$\left(\widetilde{\text{Cl}} \circ \zeta_{\tau,\tau}^{-1}\right)(S, S') = \begin{cases} \sum_{T \in E(S)} \chi_T \cdot (-1)^{|S' \setminus S|}, & \text{if } S \subseteq S' \\ 0, & \text{o.w.} \end{cases};$$

$$\begin{aligned} \left(\zeta_{d,\tau} \cdot (\widetilde{\text{Cl}} \circ \zeta_{\tau,\tau}^{-1})\right)(S, S') &= \sum_{S'' : S \subseteq S'' \subseteq S'} \left(\sum_{T \subseteq E(S'')} \chi_T \cdot (-1)^{|S' \setminus S''|} \right) \\ &= \sum_{T : V(T) \cup S \subseteq S'} \chi_T \cdot \left(\sum_{S'' : V(T) \cup S \subseteq S'' \subseteq S'} (-1)^{|S' \setminus S''|} \right) \\ &= \sum_{T : V(T) \cup S \subseteq S'} \chi_T \cdot \delta_{S' = V(T) \cup S} = \sum_{T : V(T) \cup S = S'} \chi_T. \end{aligned}$$

Therefore, $\widetilde{E}_{\text{nonexact}} x_S =$

$$\begin{aligned} \left(\zeta_{d,\tau} \cdot (\widetilde{\text{Cl}} \circ \zeta_{\tau,\tau}^{-1})(q_1)_\tau\right)(S) &= \sum_{S' : |S'| \le \tau} \left(\sum_{T : V(T) \cup S = S'} \chi_T \cdot (\frac{\omega}{n})^{|S'|} \right) \\ &= \sum_{T : |V(T) \cup S| \le \tau} \chi_T \cdot (\frac{\omega}{n})^{|V(T)\cup S|} \end{aligned}$$

for all S with $|S| \le d$. □

[6] In general $(M_1 \circ M_2) \cdot M_3 \ne M_1 \circ (M_2 \cdot M_3)$, but they are equal if M_1 is a column vector.

[7] One intuition, suggested by a referee, is that any true expectation on cliques has objective value $\sum_{i=1}^n x_i = O(\log n)$ w.h.p., now if $p_\tau(G)$ were nonnegative then it would be almost a distribution since $\widetilde{E}_{\text{nonexact}}(x_\phi) \approx 1$ (can be checked by (3.8)) with a problematically big objective value $n^{\frac{1}{2}-\epsilon}$.

3.2. **Exact case pseudo-expectation.** Now we construct a pseudo-expectation for the exact problem.

First, there is a generic way to generate possible candidates. That is, the Size Constraints (1.4) suggests to define $\widetilde{E}x_S$ in a top-down fashion: fix $\widetilde{E}x_S$ for all $|S| = d$ first, then recursively set

$$\widetilde{E}x_S \leftarrow \frac{1}{\omega - |S|}\sum_{i\notin S}\widetilde{E}x_{S\cup\{i\}} \tag{3.9}$$

for smaller-sized S's. If denote by $\widetilde{E}_d x$ the vector of the assignments for S's s.t. $|S| = d$, then this amounts to multiplying $\widetilde{E}_d x$ by the following matrix.

Definition 3.4. *The d-***filtration matrix** $\mathrm{Fil}_{d,=d}$*, of dimension $\binom{[n]}{\leq d} \times \binom{[n]}{d}$, is*

$$\mathrm{Fil}_{d,=d}(A,B) = \begin{cases} \binom{\omega-|A|}{d-|A|}^{-1}, & \text{if } A\subseteq B \text{ (where } |B| = d\text{);} \\ 0, & \text{otherwise.} \end{cases} \tag{3.10}$$

Definition 3.5. *Given vector $\widetilde{E}_d x$ which assigns a value to each d-subseteq $S \subseteq [n]$, the* **exact pseudo-expectation generated by** *$\widetilde{E}_d x$ is*

$$\widetilde{E}x := \mathrm{Fil}_{d,=d}\cdot\widetilde{E}_d x. \tag{3.11}$$

Lemma 3.2. *The pseudo-expectation in Definition 3.5 satisfies Size Constraints* (1.4)*, regardless of the choice of $\widetilde{E}_d x$.*

Proof. For $|S| < d$, take vector v_S by $v_S(S') = \begin{cases} \omega - |S|, & \text{if } S' = S; \\ -1, & \text{if } S'\supseteq S,\ |S'\backslash S| = 1; \\ 0, & \text{otherwise} \end{cases}$

which is in $\mathbb{R}^{\binom{[n]}{\leq d}}$. Then it suffices to show $v_S^\top \mathrm{Fil}_{d,=d} = 0$, which is a direct check. □

The $\widetilde{E}$ generated like so should further satisfy:

(1) Clique Constraints (1.3);
(2) PSDness Constraint (1.5);
(3) Default Constraint (1.2) (so far we only have $\omega\cdot\widetilde{E}x_\emptyset = \widetilde{E}x_1 + ... + \widetilde{E}x_n$).

Item (3) is not a problem as long as $\widetilde{E}x_\emptyset > 0$, since we can always rescale everything by $(\widetilde{E}x_\emptyset)^{-1}$ without affecting other constraints.

Remark 3.2. (Example) *The following construction seems natural. Combining Def. 3.5 with the perspective from section 3.1.2, we can take* (3.6) *with the exact plant-setting (p_0, q_0), followed by multiplying* $\mathrm{Fil}_{d,=d}$*:*

$$\widetilde{E}_{\text{example}}x_S = \mathrm{Fil}_{d,=d}\cdot\left(\zeta_{d,\tau}\cdot(\widetilde{\mathrm{Cl}}(G)\circ\zeta_{\tau,\tau}^{-1})\cdot(q_0)_\tau\right).$$

As can be checked, this satisfies the Clique Constraints. It also has a nice Fourier expression: by some computation which we omit here, modulo provably negligible error the resulting matrix is $M_{\text{example}}(I,J) = \sum_{\substack{T:\\ |V(T)\backslash(I\cup J)|\leq\tau-d}} \frac{\binom{n-|V(T)\cup I\cup J|}{\omega-|V(T)\cup I\cup J|}}{\binom{n}{\omega}}\chi_T$. The only problem, however, is that we don't know how to prove the PSDness. Despite a transparent similarity to the previous expression (3.8)*, a similar proof breaks down seriously here due to the loss of nice arithmetic structure when changing from function $(\frac{\omega}{n})^x$ (in* (3.8)*) to $\frac{\binom{n-x}{\omega-x}}{\binom{n}{\omega}}$. See also Remark 7.1.*

Now we construct an $\widetilde{E}_d$ in Definition 3.5. With the idea stated in section 2.1, we give the construction matter-of-factly here. First, take the pseudo-expectation

for $|S| = d$ in the form $\widetilde{E}x_S = \sum_{T:|V(T)\cup S|\le\tau} \chi_T \cdot F(|V(T)\cup S|)$ for some function F. We call F a d-**generating function**, to be chosen shortly after. For now, for any $|S| \le d$, by (3.10) the pseudo-expectation has form: denote $u = d - |S|$,

(3.12)
$$\widetilde{E}x_S = \frac{1}{\binom{w-d+u}{u}} \sum_{T:|V(T)\cup S|\le\tau} \chi_T \cdot \Big[\sum_{c=0}^{u} \binom{|V(T)\cup S| - d + u}{c} \binom{n - |V(T)\cup S|}{u-c} \cdot F(|V(T)\cup S| + u - c)\Big].$$

Lemma 3.3. (3.12) *always satisfy Clique and Size Constraints* (1.3),(1.4).

Proof. It satisfies Size Constraints by Lemma 3.2. For Clique Constraints, fixing S, the "[...]"-part in (3.12) only depends on $|V(T)\cup S|$, so $\widetilde{E}x_S$ has the form $\sum_{T:|V(T)\cup S|\le\tau} a_{|V(T)\cup S|}\chi_T = \sum_k \sum_{T:|V(T)\cup S|=k} a_k\chi_T$, the inner sum factors through $\widetilde{\mathrm{Cl}}_S = \sum_{T\subseteq E(S)} \chi_T$. Thus, $M(I,J)(G) = 0$ if $\widetilde{\mathrm{Cl}}_{I\cup J}(G) = 0$. □

Definition 3.6. *(Exact d-generating function) We choose*

$$F(x) := \frac{(x+8\tau^2)!}{(8\tau^2)!} \cdot (\frac{\omega}{n})^x.$$

Remark 3.3. *As already mentioned in section 2.1, the design of F, especially its first factor, is technical; the goal is to make the resulting M positive. The numerator $(x+8\tau^2)!$ will be used in Prop. 8.3, where the $8\tau^2$ can be replaced by larger polynomials in τ. The $(8\tau^2)!$ in denominator is added for convenience (see Remark 3.4).*

Definition 3.7. *The* **exact moment matrix** $\widetilde{M}$ *is defined as* $\widetilde{M}(A,B) = \sum_{T:|V(T)\cup A\cup B|\le\tau} \widetilde{M}(A,B;T)\chi_T$ *for all $A, B \subseteq [n]$, $|A|, |B| \le d/2$, where* $\widetilde{M}(A,B;T) =$

(3.13)
$$\frac{1}{\binom{\omega-d+u}{u}} \Big[\sum_{c=0}^{u} \binom{|V(T)\cup A\cup B| - (d-u)}{c} \binom{n - |V(T)\cup A\cup B|}{u-c} \cdot \underbrace{\frac{\big(|V(T)\cup A\cup B| + u - c + 8\tau^2\big)!}{(8\tau^2)!} \cdot (\frac{\omega}{n})^{|(V(T)\cup A\cup B)|+u-c}}_{f\left(|V(T)\cup A\cup B|+u-c\right)}\Big].$$

Here we denoted $d - |A\cup B|$ by u.

Remark 3.4. *In* (3.13)*, the "most significant" factor is $(\frac{\omega}{n})^{|V(T)\cup A\cup B|} \cdot \omega^{-c}$, if notice $\frac{\binom{n-|V(T)\cup A\cup B|}{u-c}}{\binom{\omega-d+u}{u}} \omega^u n^{-(u-c)} \ll \omega, n$. One thing to keep in mind is that factors like $\frac{(|V(T)\cup A\cup B|+u-c+8\tau^2)!}{(8\tau^2)!}$ are qualitatively smaller than ω in our parameter regime.*

4. Some preparation

In this section, we prepare some basic tools for analysis.

4.1. **Homogenization for Exact Clique.** With the Size Constraints (1.4) satisfied, any moment matrix can be reduced to its $\binom{[n]}{d/2}$-principal minor, which is slightly more convenient to work with. The following homogeneity trick is standard in the SoS literature.

Given any degree-d moment matrix $M_{dSoS}(G)$ that satisfies the Size Constraints (1.4), let $M(G)$ be its principal minor on $\binom{[n]}{d/2} \times \binom{[n]}{d/2}$.

Lemma 4.1. *$M_{dSoS}(G)$ is PSD $\Leftrightarrow$ $M(G)$ is PSD.*

Proof. The $\Rightarrow$ part is trivial. Now suppose M_{dSoS} is not PSD, then $\exists a \in \mathbb{R}^{\binom{[n]}{\le d/2}}$ s.t. $a^\top M_{dSoS} a = -1$. With the presence of boolean constraints (i.e. we can additionally define $\widetilde{E}(x_i^2 \cdot p) := \widetilde{E}(x_i \cdot p)$ for all i and all polynomial p of degree $\le d-2$), this is equivalent to $\widetilde{E}(g^2) = -1$ for some multi-linear polynomial $g = a^\top x = \sum_{|S| \le d/2} a_S x_S$. Now substitute every x_S ($|S| < d/2$) in g by the corresponding linear combination of $\{x_{S'} \mid |S'| = d/2\}$ from (3.9), we get a multi-linear, degree-$d/2$ homogeneous g_1. Since $g - g_1$ thus $g^2 - g_1^2$ is a multiple of the constraints,

$$\widetilde{E}(g_1^2) = \widetilde{E}(g^2) = -1. \tag{4.1}$$

Assume $g_1 = b^T x$ where x denotes $(x_S)_{|S|=d/2}$. Then (4.1) says $b^\top M b = -1$, so M is not PSD. □

4.2. **Concentration bound on polynomials.** The following bound on random polynomials is standard.

Lemma 4.2. *Suppose $a < \log n$, and p is a polynomial*

$$p = \sum_{T:\ |V(T)|=a} c(T)\chi_T \quad c_T \in \mathbb{R}$$

and $C > 0$ is a number s.t. $|c(T)| \le C$ for all T. Then W.p. $1 - n^{-10\log n}$ over G,

$$|p(G)| < C \cdot n^{a/2} 2^{a^2} n^{4 \log\log n}. \tag{4.2}$$

Proof. For all $k \in \mathbb{N}$,

$$p^{2k} = \sum_{T_1,\dots,T_{2k}:\ |V(T_i)|=a} c(T_1)\dots c(T_{2k})\chi_{T_1}\cdots\chi_{T_{2k}}, \tag{4.3}$$

and we take the expectation of this. Each $\mathbb{E}[\chi_{T_1}\cdots\chi_{T_{2k}}(G)] \neq 0$ (i.e. equals 1) iff every edge appears even times in $T_1, \dots, T_{2k}$, which implies $|V(T_1 \cup \dots \cup T_{2k})| \le \frac{1}{2} \cdot 2ka = ka$. There are at most $ka\binom{n}{ka} < n^{ka}$ many choices of $V(T_1 \cup \dots \cup T_{2k})$. For each choice, there are at most $\binom{ka}{a} \cdot 2^{\binom{a}{2}} < (ka)^a \cdot 2^{a^2/2}$ many ways to choose each T_i. Therefore,

$$\mathbb{E}[p^{2k}] \le C^{2k} \cdot n^{ka} \left((ka)^a 2^{a^2/2}\right)^{2k} := N^{2k} \quad \text{where} \quad N = Cn^{a/2} \cdot (ka)^a \cdot 2^{a^2/2}.$$

By Markov inequality, $\Pr\left[p^{2k} > (2N)^{2k}\right] < 2^{-2k}$. Take $k := 10\log^2 n$, we get that w.p. $> 1 - n^{-10\log n}$, $|p(G)| < 2N < C \cdot n^{a/2} 2^{a^2} n^{4\log\log n}$ for all large enough n. □

4.3. **Norm concentration of pseudo-random matrices.** Like in almost all previous work on the subject, the norm bound on certain pseudo-random matrices called *graph matrices* ([AMP16]) will be a fundamental tool for us. Intuitively, such a matrix collects all possible Fourier characters from embeddings of a fixed small graph.

Definition 4.1. *(cf.* [AMP16, MPW15, HKP15, JPR$^+$21]*) A* **ribbon** *$\mathcal{R}$ is a triple $(A, B; T)$ where A, B are vertex-sets and T is an edge set. A, B are called the* **side sets***, or individually the left and right set of $\mathcal{R}$, respectively. The* **size** *of $\mathcal{R}$ is $|V(\mathcal{R})| = |V(T) \cup A \cup B|$.*

By definition, a ribbon as a graph always has no isolated vertex outside of $A \cup B$.

Definition 4.2. *We say $\mathcal{R} = (A, B; T)$ is* **left-generated** *if every vertex in $V(\mathcal{R})$ is either in B or can be reached by paths*[8] *from A without touching B. Being* **right-generated** *is symmetrically defined.*

Definition 4.3. *A* **shape** *is a equivalent class of ribbons, where two ribbons $(A, B; T)$, $(A', B'; T')$ are equivalent or "of the same shape" if there is an isomorphism σ between the corresponding graphs s.t. $\sigma(A) = A'$ and $\sigma(B) = B'$. Denote a shape by $\mathcal{U}$, represented by a ribbon $(A, B; T)$. $V(\mathcal{U}) := A \cup B \cup V(T)$ and its* **size** *is $|V(\mathcal{U})|$.*

Thus we may speak of **the shape of a ribbon** $\mathcal{R}$. We say a function f defined on a set of ribbons is **symmetric w.r.t. shapes** if, whenever $\mathcal{R}$ and $\mathcal{R}'$ are of the same shape and f is defined on them, $f(\mathcal{R}) = f(\mathcal{R}')$.

Definition 4.4. *(*[AMP16]*) Fix n and shape $\mathcal{U} = (A, B; T)$. The* **graph matrix of shape** $\mathcal{U}$ *is the following $2^{[n]} \times 2^{[n]}$-matrix $M_{\mathcal{U}}$:*

$$\forall I, J \subseteq [n], \quad M_{\mathcal{U}}(I, J) = \sum_{\substack{T_1: \\ \exists \text{injective } \phi : V(\mathcal{U}) \to [n] \text{ s.t.} \\ \phi(A)=I,\ \phi(B)=J,\ \phi(T)=T_1}} \chi_{T_1}$$

($= 0$ if no such ϕ exists). Here, ϕ on T means the natural induced map on edges.

In [AMP16], the matrices have columns and rows indexed by *tuples* with elements in $[n]$, instead of *subsets* (which is our case), but our matrix is always a sub-matrix of it, e.g. ours can be viewed as supported on strictly increasing tuples.

Theorem 3. *(Norm bounds on $M_{\mathcal{U}}$,* [AMP16]*) For any shape $\mathcal{U} = (A, B; T)$ of size $t < \log n$, w.p. $> 1 - n^{-10 \log n}$ over G,*

$$\|M_{\mathcal{U}}(G)\| \le n^{\frac{t-p}{2}} \cdot 2^{O(t)} \cdot (\log n)^{O(t+p-2r)} \tag{4.4}$$

where $r = |A \cap B|$ and p is the max number of vertex-disjoint paths between (A, B) in $\mathcal{U}$. Moreover, under the same notation, if further denote $s = \frac{|A|+|B|}{2}$ then

$$\|M_{\mathcal{U}}(G)\| \le n^{\frac{t-p}{2}} \cdot 2^{O(t)} \cdot (\log n)^{O(t-s)}. \tag{4.5}$$

Theorem 3 is proved by a careful estimation of the trace-power $\mathbb{E}[\mathrm{tr}(M_{\mathcal{U}}^{2k})]$ (for some $k > 0$) which we omit here. Its "moreover" part follows from (4.4) since $t \ge |A \cup B| = 2s - r$, $p \le s$, so $t + p - 2r \le t + s - 2(2s - t) = 3(t - s)$.

4.4. **Some notions on graphs.**

Definition 4.5. *(Vertex-separator) For a graph H and $A, B \subseteq V(H)$, we say $S \subseteq V(H)$ is an (A, B)-***vertex-separator***, or S separates A, B in H, if any path from A to B in H must pass through S. Let*

$$s_{A,B}(H) := \min\{|S| \mid S \text{ is an } (A, B)\text{-vertex-separator}\}.$$

A vertex-separator achieving this minimum is a **min-separator**. *Let $\mathrm{mSep}_{A,B}(H)$ denote the set of all min-separators.*

The definition naturally applies to a ribbon $\mathcal{R} = (A, B; T)$, with A, B being the two vertex-sets. In that case, we can write the corresponding min-separator size as $s_{A,B}(T)$ and set of the min-separators as $\mathrm{mSep}_{A,B}(T)$ or $\mathrm{mSep}(\mathcal{R})$.

Theorem 4. *(Menger's theorem) For any finite graph H, $s_{A,B}(H)$ equals to the maximum number of vertex-disjoint paths from A to B in H.*

Definition 4.6. *For ribbon $\mathcal{R} = (A, B; T)$, define its* **reduced size** *to be*

$$e_{A,B}(T) := |V(T) \cup A \cup B| - s_{A,B}(T). \tag{4.6}$$

[8]We always stick to the convention of including degenerate paths (one-point path).

The reduced size is double of the exponent in n in the bound of Theorem 3, hence is the controlling parameter of the norm of the graph matrix.

A fundamental fact is that the set of all min-separators form a lattice.

Theorem 5. *([Esc72]) For a ribbon $(A, B; T)$, $\mathrm{mSep}_{A,B}(T)$ has a natural* **poset** *structure: min-separators $A_1 \le A_2$ iff A_1 separates $(A, A_2; T)$, or equivalently as it can be checked, iff A_2 separates $(A_1, B; T)$. The set is actually a* **lattice** *under this partial-ordering: $\forall A_1, A_2 \in \mathrm{mSep}_{A,B}(T)$ their join and meet exist. In particular, there exist unique* **minimum** *and* **maximum**.

We denote the minimum in the above theorem by $S_l(A, B; T)$ and the maximum by $S_r(A, B; T)$, meant to be the **leftmost** and **rightmost** min-separator, respectively.

4.5. **Johnson schemes.** We only need a minimal amount of knowledge here.

Definition 4.7. *([Del73]) Fix natural numbers $n \ge k$, $n > 0$. A* **Johnson scheme** *$\mathfrak{J}$ is an $\binom{[n]}{k} \times \binom{[n]}{k}$-matrix that satisfies $\mathfrak{J}(I, J) = \mathfrak{J}(I', J')$ whenever $|I \cap J| = |I' \cap J'|$.*

It can be checked that (fix n, k) all Johnson schemes are symmetric matrices and form a commutative $\mathbb{R}$-algebra, so they are simultaneously diagonalizable. In below we fix n and $k = d/2$. An obvious $\mathbb{R}$-basis for Johnson schemes is $D_0, ..., D_{d/2}$ where

$$D_r(I, J) = \begin{cases} 1, & \text{if } |I \cap J| = r \\ 0, & \text{o.w.} \end{cases} \qquad \forall I, J \in \binom{S}{d/2}. \tag{4.7}$$

Another basis which we denote by $\mathfrak{J}_0, ..., \mathfrak{J}_{d/2}$ is

$$\mathfrak{J}_r(I, J) = \binom{|I \cap J|}{r}, \quad \forall I, J \in \binom{[n]}{\frac{d}{2}}. \tag{4.8}$$

$\mathfrak{J}_0, ..., \mathfrak{J}_{d/2}$ are PSD matrices since

$$\mathfrak{J}_r = \sum_{A \subseteq [n], |A| = r} u_A u_A^\top \quad \text{where} \quad u_A \in \mathbb{R}^{\binom{[n]}{k}}, \ u_A(B) = 1_{A \subseteq B}. \tag{4.9}$$

Also, clearly, $\mathfrak{J}_{d/2} = \mathrm{Id}$. A basis-change from D to $\mathfrak{J}$ is given by the following.

Lemma 4.3. $D_r = \sum_{r'=r}^{d/2} (-1)^{r'-r} \binom{r'}{r} \cdot \mathfrak{J}_{r'}$.

Proof. The $RHS(I, J) = \sum_{r'=r}^{d/2} (-1)^{r'-r} \binom{r'}{r} \binom{|I \cap J|}{r'} = \sum_{r'=r}^{|I \cap J|} (-1)^{r'-r} \binom{|I \cap J|}{r} \binom{|I \cap J| - r}{r' - r} = \binom{|I \cap J|}{r} \cdot 1_{|I \cap J| = r} = 1_{|I \cap J| = r}$. □

5. PSDness analysis, I: Hadamard product and Euler transform

Notation. Henceforth throughout the paper, M exclusively refers to the $d/2$-homogeneous minor of the moment matrix $\widetilde{M}$ in Definition 3.7.

Our main theorem is the following.

Theorem 6. *W.p. $> 1 - n^{-5 \log n}$, $M(G) \succeq n^{-d-1} \mathrm{diag}\left(\widetilde{\mathrm{Cl}}(G)\right)_{\binom{[n]}{d/2} \times \binom{[n]}{d/2}}$.*

Corollary 5.1. *W.p. $> 1 - n^{-5 \log n}$, $\widetilde{E} x_\emptyset > 0$.*

Proof. By construction (3.9), $\widetilde{E} x_\emptyset = \frac{\binom{\omega - d/2}{d - d/2}}{\binom{\omega}{d}\binom{d}{d/2}} \sum_{S : |S| = d/2} \widetilde{E} x_S = \frac{\binom{\omega - d/2}{d - d/2}}{\binom{\omega}{d}\binom{d}{d/2}} \mathrm{Tr}(M)$, and by Theorem 6 this is positive with high probability. □

Proof. (of Theorem 1 from Theorem 6) Lemma 4.1 and Theorem 6 proves the PSDness of the moment matrix from Definition 3.7, which also satisfies the Default Constraint (Corollary 5.1 and the discussion above Remark 3.2) and the Clique and Size Constraints (Lemma 3.2). The degree-d lower bound follows. □

The rest of the paper is for proving Theorem 6. We will use three steps to achieve so, and this section makes the first step.

To begin with, by definition of $M(I,J)$ (Def. 3.7, (3.13)),

(5.1)
$$M(I,J;T) = \sum_{c=0}^{u} \Bigg[\frac{1}{\binom{\omega-d+u}{u}} \omega^{u-c} \cdot \underbrace{\cdot \left(\binom{a-(d-u)}{c} \binom{n-a}{u-c} n^{-(u-c)} \frac{(a+u-c+8\tau^2)!}{(8\tau^2)!} (\frac{\omega}{n})^a \right)}_{:=M_c(u,a)} \Bigg]$$

where $u = |I \cap J|, a = |V(T) \cup I \cup J|$. In this expression the parameter u appears nestedly and makes it difficulty to analyze. (It doesn't appear in the non-exact case (3.8) at all.) To resolve the issue, we express M in a $\Sigma\Pi$-form, i.e. a sum of Hadamard products, so that in each leaf matrix the dependence on u is removed to some degree:

$$M = \sum_{c=0}^{\frac{d}{2}} m_c \circ M_c \tag{5.2}$$

where m_c, M_c are matrices as follows. For all $|I|, |J| = d/2$,

$$m_c(I,J) = \frac{1}{\binom{\omega-d+u}{u}} \omega^{u-c} \quad \text{where } u = |I \cap J| \tag{5.3}$$

$$M_c(I,J) = \begin{cases} \sum_{T: |V(T) \cup I \cup J| \le \tau} M_c(|I \cap J|, |V(T) \cup I \cup J|) \chi_T & \text{, if } |I \cap J| \ge c; \\ 0 & \text{, o.w.} \end{cases} \tag{5.4}$$

Remark 5.1. *It is important to note that m_c is supported on all (I,J) while $M_c(I,J) = 0$ if $|I \cap J| < c$, so that (5.2) holds.*

To analyze (5.2), we would hope that the second factor M_c is "close" to each other for varying c, while the first factor m_c is qualitatively decreasing in c. This, if true, would make it possible for us to concentrate on showing the PSDness in the main case $c = 0$. The next Lemma 5.1 proves the second half of the above intuition; the other half will be stated more precisely in the Main Lemma 5.3.

Lemma 5.1. *For each $c = 0, ..., d/2$, $m_c = \omega^{-c} \sum_{k=0}^{d/2} b_k \cdot \mathfrak{J}_k$ where $\mathfrak{J}_k$'s are the Johnson basis (4.8), $b_k/k! \in [\frac{d}{2\omega}, 1 + \frac{2dk}{\omega}]$. In particular,*

$$m_0 = \omega m_1 = ... = \omega^{\frac{d}{2}} m_{\frac{d}{2}} \succ \frac{1}{\omega} \mathrm{Id}. \tag{5.5}$$

Proof. By definition, $m_c = \omega^{-c} \sum_{l=0}^{d/2} \frac{\omega^l}{\binom{\omega-d+l}{l}} D_l$, where D_l $(l = 0, ..., d/2)$ are the simple basis of Johnson schemes (4.7). By basis-change (Lem. 4.3),

$$m_c = \omega^{-c} \sum_{k=0}^{d/2} \mathfrak{J}_k \cdot k! \left(\sum_{l=0}^{k} (-1)^{k-l} \cdot \underbrace{\left[\frac{\omega}{\omega-(d-l)} \cdot ... \cdot \frac{\omega}{\omega-(d-1)} \cdot \frac{1}{(k-l)!} \right]}_{:=f_k(l), \text{ which is } 1/k! \text{ if } l=0} \right).$$

For fixed k, $f_k(l)$ is increasing in l so $\sum_{l=0}^{k}(-1)^{k-l}f_k(l) \geq f_k(k) - f_k(k-1) > \frac{d/2}{\omega}\cdot(1+\frac{d/2}{\omega})^{k-1} \geq \frac{d}{2\omega}$. Note for $k = d/2$, $\mathfrak{J}_{d/2} = \mathrm{Id}$, so we get (5.5). □

Euler transform. Fixing c, now we look into the second factor M_c in (5.2). For fixed $(I, J; T)$ denote $u = |I \cap J|$, $a = |V(T) \cup I \cup J|$, then by (5.1) we have that

$$M_c(u,a) = \binom{a-(d-u)}{c}\binom{n-a}{u-c}n^{-(u-c)}\frac{(a+u-c+8\tau^2)!}{(8\tau^2)!}(\frac{\omega}{n})^a \tag{5.6}$$

is the coefficient of χ_T in $M_c(I, J)$ for $c \leq u$.

Definition 5.1. *(**Extended** $M_c(u,a)$) For fixed $c \geq 0$, the function $M_c(u,a)$ in (5.6) is partial, defined for $(u,a) \in \mathbb{N}^2$ s.t. $u \geq c$, $u+a \geq d+c$. It can be naturally* **extended to** $\mathbb{N}^2$ *by letting*

$$\binom{n-a}{u-c} = 0 \quad \text{if } u < c, \tag{5.7}$$

and using the convention on binomial coefficients: $\binom{-m}{k} = (-1)^k \cdot \binom{m+k-1}{k}$ for all $m > 0$, $k \geq 0$; $\binom{m}{0} = 1$ for all $m \in \mathbb{Z}$; and

$$\binom{m}{k} = 0 \text{ for all } 0 \leq m < k. \tag{5.8}$$

In the rest of the paper, we will use $M_c(u,a)$ to mean this extended function.

In particular, if $0 \leq a-(d-u) < c$ then $M_c(u,a) = 0$ since $\binom{a-(d-u)}{c} = 0$.

One trouble with M_c is that, still, $u = |I \cap J|$ appears in it in an unpleasant way. To further remove the dependence on u, we consider a decomposition

$$M_c = \sum_{R \in \binom{[n]}{\leq \frac{d}{2}}} M_c^R \tag{5.9}$$

where for each $R \in \binom{[n]}{\leq \frac{d}{2}}$ the matrix M_c^R is supported on rows and columns whose index contains R. More explicitly, for any $(I, J; T)$ let $a = |V(T) \cup I \cup J|$, suppose

$$M_c^R(I,J) := \begin{cases} (\frac{\omega}{n})^a \sum_{T: |V(T)\cup I\cup J| \leq \tau} Y_c(|R|, a)\cdot\chi_T & \text{, if } R \subseteq I \cap J; \\ 0 & \text{, o.w.} \end{cases} \tag{5.10}$$

for some function $Y_c(u,a)$ to be chosen, then comparing for every tuple $(I, J; T)$ we see that equation (5.9) is equivalent to the following: for any fixed c, a,

$$\sum_{r=0}^{u}\binom{u}{r}Y_c(r,a)(\frac{\omega}{n})^a = M_c(u,a). \tag{5.11}$$

This suggests to take $Y_c(u,a)\cdot(\frac{\omega}{n})^a$ to be the **inverse Euler transform** (w.r.t. variable u) of the **extended** function $M_c(u,a)$.

Fact 1. [9] *If $x(m), y(m)$ are two sequences defined on $\mathbb{N}$ s.t. for all m, $x(m) = \sum_{l=0}^{m}\binom{m}{l}y(l)$, then $x(m)$ is called the* **Euler transform** *of $y(m)$. The inverse transform is given by that for all m, $y(m) = \sum_{l=0}^{m}(-1)^{m-l}\binom{m}{l}x(l)$.*

Definition 5.2. *(Coefficients in M_c^R) For every fixed $c \geq 0$, define*

$$Y_c(r,a) = \begin{cases} \sum_{l=c}^{r}(-1)^{r-l}\binom{r}{l}\binom{a+l-d}{c}\binom{n-a}{l-c}n^{-(l-c)}\frac{(a+l-c+8\tau^2)!}{(8\tau^2)!} & \text{, if } r \geq c; \\ 0 & \text{, o.w.} \end{cases} \tag{5.12}$$

[9] Coincidentally, this fact can be seen as an application of ζ-matrix and its inverse.

Then as a clear-up summary, we prove the following the main result of this section.

Lemma 5.2. *(The Hadamard-product decomposition of M)*

$$M = \sum_{c=0}^{\frac{d}{2}} m_c \circ \left(\sum_{R: R\in \binom{[n]}{\le d/2}} M_c^R \right) \tag{5.13}$$

$$= \sum_{R\in \binom{[n]}{\le d/2}} \underbrace{\left(\sum_{c=0}^{|R|} m_c \circ M_c^R \right)}_{:=M^R} \tag{5.14}$$

where each m_c is as in Lemma 5.1 and each M_c^R has the following expression.

(1) $M_c^R = 0$ if $|R| < c$;

(2) If $R \not\subseteq I \cap J$, $M_c^R(I,J) = 0$;

(3) If $|R| \ge c$ and $R \subseteq I \cap J$, $M_c^R(I,J) = \sum_{T: |V(T)\cup I\cup J|\le \tau} M_c^R(I,J;T)\chi_T$ where if denote $a = |V(T)\cup I\cup J|$, then $M_c^R(I,J;T) =$

$$(\frac{\omega}{n})^a \underbrace{\sum_{l=c}^{|R|} (-1)^{|R|-l} \binom{|R|}{l} \binom{a+l-d}{c} \binom{n-a}{l-c} n^{-(l-c)} \frac{(a+l-c+8\tau^2)!}{(8\tau^2)!}}_{Y_c(|R|,a)\ (5.12)}. \tag{5.15}$$

(4) For all $0 \le c \le r \le d/2$ and $0 \le a \le \tau$, $|Y_c(r,a)| < \tau^{5\tau}$.

Proof. (1), (2), (3) is by definition. To check (5.13) i.e. $M_c = \sum_R M_c^R$, we check for every $(I,J;T)$ where $|I| = |J| = d/2$, $|V(T)\cup I\cup J| \le \tau$. Let $u = |I\cap J|$, $a = |V(T)\cup I\cup J|$, then note $a - (d-u) \ge 0$, and

$$\sum_{R:} M_c^R(I,J;T) = \sum_{R: R\subseteq I\cap J} M_c^R(I,J;T) = (\frac{\omega}{n})^a \sum_{r=0}^{|I\cap J|} \binom{|I\cap J|}{r} Y_c(r,a).$$

By the Euler transform and (5.11), the RHS equals the extended $M_c(u,a)$. Thus, we only need to see $M_c(u,a) = 0$ if further $u < c$ or $a - (d-u) < c$ (in particular, in such cases $c > 0$), and this is by (5.7), (5.8).

For (4),

$$\begin{aligned} |Y_c(u,a)| &= \left| \sum_{l=c}^{r} (-1)^{r-l} \binom{r}{l} \binom{a+l-d}{c} \Big[\binom{n-a}{l-c} n^{-(l-c)} \Big] \frac{(a+l-c+8\tau^2)!}{(8\tau^2)!} \right| \\ &< r\cdot 2^r \cdot (2\tau)^r \cdot 1 \cdot (9\tau^2)^{2\tau} < \tau^{5\tau} \end{aligned}$$

where note $r \le d/2 \ll \tau$ in our parameter regime. □

Lemma 5.3. (Main Lemma) *In the decomposition* (5.14)*, w.p. $> 1 - n^{-5\log n}$ the following hold. For all $R \in \binom{[n]}{\le d/2}$, denote $P^R = \{I \in \binom{[n]}{d/2} \mid R \subseteq I\}$,*

(1). $M_0^R \succeq n^{-d}\mathrm{diag}(\widetilde{\mathrm{Cl}})_{P^R\times P^R}$;

(2). $\pm\omega^{-c} M_c^R \preceq n^{-c/6}\cdot M_0^R, \quad \forall 0 < c \le |R|$.

Corollary 5.2. *(Theorem 6) W.p. $> 1 - n^{-5\log n}$ over G,*

$$M(G) \succeq n^{-d-1}\mathrm{diag}(\widetilde{\mathrm{Cl}}(G))_{\binom{[n]}{d/2}\times\binom{[n]}{d/2}}.$$

Proof. Fix an R, $M^R = \sum_{c=0}^{|R|} m_c \circ M_c^R$. Suppose Lemma 5.3 (1), (2) hold (w.p. probability $> 1 - n^{-5\log n}$). Since Hadamard product with a PSD matrix presevres PSDness (Schur product theorem), we have $\sum_{c=1}^{|R|} m_c \circ M_c^R \preceq \sum_{c=1}^{|R|} m_c \circ \left(\omega^c n^{-c/6} \cdot M_0^R\right)$ by Lemma 5.3(2). The latter equals $\left(\sum_{c=1}^{|R|} n^{-c/6} \cdot m_0\right) \circ M_0^R$ by Lemma 5.1 which then $\preceq\ n^{-1/6} m_0 \circ M_0^R$. Similarly, $\sum_{c=1}^{|R|} m_c \circ M_c^R \ \succeq\ -n^{-1/6} m_0 \circ M_c^R$. Thus

$$M^R \ \succeq\ (1 - n^{-1/6}) m_0 \circ M_0^R \ \succeq\ n^{-d-1}\mathrm{diag}(\widetilde{\mathrm{Cl}})_{P^R \times P^R} \quad \text{(Lem. 5.1 and 5.3(2))}.$$

So in (5.14), $M = M^\emptyset + \sum_{\emptyset \neq R \in \binom{[n]}{\leq d/2}} M^R \succeq\ M^\emptyset \succeq\ n^{-d-1}\mathrm{diag}(\widetilde{\mathrm{Cl}})_{\binom{[n]}{d/2} \times \binom{[n]}{d/2}}$. □

The rest of the paper is devoted to proving the Main Lemma 5.3.

6. Recursive factorization: a preparation

In this section, we give a systematic treatment of the *recursive approximate factorization* technique of [BHK+19]. In section 6.3 we will formalize this technique in a convenient language, and extend this technique properly for later use in Section 7. Section 6.1 and 6.2 together is an independent part only for showing Theorem 2(2) via the so-called *mod-order analysis*; the reader can safely skip them and proceed directly to section 6.3 to continue the proof of the Main Lemma 5.3.

Notation. Thoughout section 6, for simplicity we discuss the non-exact moment matrix (which suffices to lay the ground for the technique). Denote by M' the $\binom{[n]}{\frac{d}{2}} \times \binom{[n]}{\frac{d}{2}}$-minor[10] of the non-exact moment matrix.

$$M'(I,J) = \sum_{T: |V(T) \cup I \cup J| \leq \tau} (\frac{\omega}{n})^{|V(T)\cup I \cup J|} \chi_T \quad \forall I, J \in \binom{[n]}{d/2}. \tag{6.1}$$

The goal of section 6 is to diagonalize M' approximately in the "$LQL^\top$" form s.t. the difference matrix is negligible (w.h.p. when plugging in G).

6.1. Step 1: Diagonalization of $\mathbb{E}[M']$.

Proposition 6.1. *$\mathbb{E}[M'] = CC^\top$, where C is the $\binom{[n]}{d/2} \times \binom{[n]}{\leq d/2}$-matrix*

$$C = (\zeta^\top)_{d/2, \leq d/2} \cdot \mathrm{diag}\left(\sqrt{t(|A|)}\right)_{A \in \binom{[n]}{\leq d/2}} \tag{6.2}$$

and $t(r) = (1 - O(\frac{d\omega}{n})) \cdot (\frac{\omega}{n})^{d-r}$ for all $r = 0, ..., d/2$.

This can be shown by a similar calculation as in [MPW15], as below. Recall the Johnson schemes 4.7.

Fact 2. *(See e.g. (4.29) in* [Del73]*) The Johnson schemes (for $(n, d/2)$) have shared eigenspace-decomposition $\mathbb{R}^{\binom{[n]}{d/2}} = V_0 \oplus ... \oplus V_{d/2}$, and*

$$\mathfrak{J}_r = \bigoplus_{i=0}^{\frac{d}{2}} \lambda_r(i) \cdot \Pi_i \quad \text{for } r = 0, ..., d/2$$

[10]Strictly speaking, PSDness of this minor is not sufficient as we do not have a homogeneity reduction in non-exact case. Nevertheless, it suffices to demonstrate the idea.

where Π_i is the orthogonal projection to V_i w.r.t. the Euclidean inner product, and the eigenvalues are

$$\lambda_r(i) = \binom{\frac{d}{2}-i}{r-i}\binom{n-\frac{d}{2}-i}{\frac{d}{2}-r}, \quad 0 \le i \le \frac{d}{2}.$$

Lemma 6.1. $\mathbb{E}[M'] = \sum_{r=0}^{d/2} t(r)\mathfrak{J}_r$ *where each* $t(r) = (1-O(\frac{d\omega}{n}))\cdot(\frac{\omega}{n})^{d-r}$.

Proof. By definition, $\mathbb{E}[M'] = \sum_{r=0}^{d/2}(\frac{\omega}{n})^{d-r}D_r$. By Lemma 4.3,

$$D_r = \sum_{r'=r}^{d/2}(-1)^{r'-r}\binom{r'}{r}\cdot\mathfrak{J}_{r'} \tag{6.3}$$

So

$$\begin{aligned}\mathbb{E}[M'] &= \sum_{r=0}^{d/2}(\frac{\omega}{n})^{d-r}\left(\sum_{r'=r}^{d/2}(-1)^{r'-r}\binom{r'}{r}\mathfrak{J}_{r'}\right)\\ &= \sum_{r'=0}^{d/2}\mathfrak{J}_{r'}\cdot\left(\sum_{r=0}^{r'}(\frac{\omega}{n})^{d-r}(-1)^{r'-r}\binom{r'}{r}\right) \qquad (6.4)\\ &= \sum_{r'=0}^{d/2}\mathfrak{J}_{r'}\cdot(\frac{\omega}{n})^{d-r'}(1-\frac{\omega}{n})^{r'}\end{aligned}$$

which proves the lemma. □

By Lemma 6.1 and (4.9), if let $t(r) = (\frac{\omega}{n})^{d-r'}[1-\frac{\omega}{n}]^{r'}$ then

$$\mathbb{E}[M'] = \sum_{A:|A|\le d/2} t(|A|)u_A u_A^\top = (\zeta^\top)_{d/2,\le d/2}\cdot\mathrm{diag}\Big(t(|A|)\Big)\cdot\zeta_{\le d/2,d/2} = CC^\top,$$

where used that the matrix $(\zeta^\top)_{d/2,\le d/2}$ has columns $\{u_A \mid |A|\le d/2\}$. This proves Proposition 6.1.

6.2. **Step 2: Mod-order analysis toward "coarse" diagonalization.**

This subsection is only for the simplification result, Theorem 2(2). The reader can safely safely skip it and proceed to section 6.3 for the proof of Theorem 1.

Given $\mathbb{E}[M'] = CC^\top$ in Step 1, ideally we hope to continue to solve for

$$M' = NN^\top \tag{6.5}$$

with $\mathbb{E}[N] = C$, and N extending C by non-trivial Fourier characters. Also, we restrict ourselves to symmetric solutions w.r.t. shapes.

Toward this goal, we start with a relaxed equation as Definition 6.1, with the following motivation.

(1) Order in $\frac{\omega}{n}$. Entries of M' all have a clear order in $\frac{\omega}{n}$. Like in fixed-parameter problems, we treat $\frac{\omega}{n}$ as a distinguished structural parameter and try to solve the correct power of $\frac{\omega}{n}$ in terms in N.

(2) Norm-match. Let's have a closer look into

$$\mathbb{E}[M'] = CC^\top = \sum_{r=0}^{d/2}(1-O(\frac{d\omega}{n}))\cdot(\frac{\omega}{n})^{d-r}\mathfrak{J}_r.$$

By fact 2, each $\mathfrak{J}_r$ b has norm $\binom{d/2}{r}\cdot n^{d/2-r}$ so

$$\left\|C_rC_r^\top\right\| \approx \binom{d/2}{r}\cdot(\frac{\omega}{n})^{d-r}n^{d/2-r}, \quad r=0,...,d/2. \tag{6.6}$$

We expect $N_r(N_r)^\top$ to concentrate around $C_r(C_r)^\top$, so the norm of the "random" part, i.e. matrix of nontrivial Fourier characters in $N_r(N_r)^\top$, is expected to be bounded by (6.6). The essentially tight bound from Theorem 3 (cf. [AMP16]) tells how this may happen, which we review below.

It will be convenient to use a scaling of variables: let

$$L = (L_0, ..., L_{\frac{d}{2}}) = (N_r \cdot (\frac{\omega}{n})^{\frac{-|A|}{2}})_{0\le r\le \frac{d}{2}},$$

then

$$M' = L \cdot \mathrm{diag}\left((\frac{\omega}{n})^{|A|}\right) \cdot L^\top \quad \text{with} \quad \mathbb{E}[L] = (\ C_r \cdot (\frac{\omega}{n})^{-r/2}\ )_{r=0,1,...,d/2}. \tag{6.7}$$

Now suppose

$$L_r(I,A) = \sum_{\text{small } T} \beta_{I,A}(T)\chi_T, \quad A \in \binom{[n]}{r}$$

where assume as in (1), an order of $\frac{\omega}{n}$ can be separated:

$$\beta_{I,A}(T) = \underbrace{(\frac{\omega}{n})^x}_{\text{main-order term}} \cdot (\text{ factor} \ll \frac{n}{\omega} \text{ and} \gg \frac{\omega}{n}\ ). \tag{6.8}$$

Fix I, A, T, we are looking for the condition on x in order to have the expected norm control on $L_r(\frac{\omega}{n})^r(L_r)^\top$. Ignore for a moment the cross-terms, such a single graph matrix square in $L_r(\frac{\omega}{n})^r L_r^\top$ is

$$(\frac{\omega}{n})^{2x} R_{(I,A;T)} \cdot (\frac{\omega}{n})^r \cdot R_{(I,A;T)}^\top$$

with norm[11]

$$\lessapprox (\frac{\omega}{n})^{2x+r} \cdot n^{e_{I,A}(T)} \cdot 2^{O(|V(T)\cup I\cup A|)} \cdot (\log n)^{>0}$$

by Theorem 3. Here recall $e_{I,A}(T) = |V(T)\cup I\cup A| - s_{I,A}(T) (\ge |I| - |A| = \frac{d}{2} - r)$. Compare this with (6.6), we need $(\frac{\omega}{n})^{2x} n^{e_{I,A}(T)} < \binom{d/2}{r}(\frac{\omega}{\sqrt{n}})^{d/2-r}$. If think of 2^d as qualitatively smaller than any positive constant power of ω, n, the natural bound to put is $x \ge e_{I,A}(T)$ which actually is the limit requirement when $\frac{\log \omega}{\log n} \to \frac{1}{2}$. Suggested by this, we will set the restriction $x \ge e_{I,A}(T)$ right from the start in the relaxed equation.

The above motivation leads to the following definition. Take a ring $\mathbb{A}$ by adding fresh variables α and χ_T's to $\mathbb{R}$, where T ranges over subsets of $\binom{[n]}{2}$ and they only satisfy relations $\{\chi_{T'} \cdot \chi_{T''} = \chi_T : T' \oplus T'' = T\}$.

Definition 6.1. *The* **mod-order equation** *is*

$$L_\alpha \cdot \mathrm{diag}\left(\alpha^{|A|}\right) \cdot (L_\alpha)^\top = M_\alpha \qquad \mod (*) \tag{6.9}$$

on the $\binom{[n]}{d/2} \times \binom{[n]}{\le d/2}$ *matrix variable* L_α *in ring* $\mathbb{A}$*, where*

$$M_\alpha(I,J) := \sum_{T:|V(T)\cup I\cup J|\le\tau} \alpha^{|V(T)\cup I\cup J|}\chi_T,$$

and $\mod (*)$ *is the* **modularity***, which means position-wise mod the ideal*

$$\left(\{\alpha^{|V(T)\cup I\cup J|+1}\chi_T\},\ \{\chi_T : |V(T)\cup I\cup J| > \tau\}\right).$$

[11]Here the matrix is truncated from size $2^{[n]} \times 2^{[n]}$, which doesn't change anything since the original matrix is always 0 elsewhere.

Moreover, if denote $L_\alpha(I,A) = \sum_T \beta_{I,A}(T)\chi_T$ *where* $\beta_{I,A}(T) \in \mathbb{R}[\alpha]$, *then*[12]

$$\alpha^{e_{I,A}(T)} \mid \beta_{I,A}(T) \quad \forall I, A, T. \tag{6.10}$$

We are interested in solutions that are **symmetric**, *i.e.* $\beta_{I,A}(T') = \beta_{J,B}(T'')$ *whenever* $(I, A; T')$, $(J, B; T'')$ *are of the same shape.*

The following is the key observation. Its proof demonstrates how to make deductions from the mod-order equations efficiently, and is presented in Appendix A.1.

Lemma 6.2. *(Order match) If a product* $\alpha^{|A|} \cdot \beta_{I,A}(T') \cdot \beta_{J,A}(T'')$ *from the LHS of* (6.9) *is nonzero mod* $(*)$, *then both of the following hold:*

$$A \text{ is a min-separator for both } (I, A; T'),\ (J, A; T''); \tag{6.11}$$

$$(V(T') \cup I \cup A) \ \cap \ (V(T'') \cup J \cup A) = A. \tag{6.12}$$

Moreover, (6.11), (6.12) *imply that*

$$A \text{ is a min-separator of } (I, J; T) \text{ (where } T = T' \oplus T''); \tag{6.13}$$

$$|V(T') \cup I \cup A|,\ |V(T'') \cup J \cup A| \le \tau. \tag{6.14}$$

By this lemma, in an imagined solution we should assume $\beta_{I,A}(T') \neq 0$ only when it satisfies its part in conditions (6.11), (6.14).

Using this information, plus a further technique of *polarization*, we can deduce the following Proposition 6.2 which is the main takeaway of the analysis here. In the deduction, the graph-theoretic fact—the "in particular" of Theorem 5—appears exactly as the solvability condition. We leave the detail of intermediate deductions to Appendix A.2.

Proposition 6.2. *(Mod-order diagonalization) Let*

$$L_\alpha(I,A) := \sum_{\substack{T':\ |V(T')\cup I\cup A|\le\tau \\ A = S_l(I,A;T') \\ T'\cap E(A)=\emptyset \\ (I,A;T') \text{ left-generated (Def. 4.2)}}} \alpha^{e_{I,A}(T')}\chi_{T'},$$

$$Q_{0,\alpha}(A,B) := \sum_{\substack{T_m:\ |T\cup A\cup B|\le\tau \\ A,B\in \mathrm{mSep}_{A,B}(T_m)}} \alpha^{e_{A,B}(T_m)}\chi_{T_m}$$

(T_m to indicate "middle"). Then

$$L_\alpha \cdot [\mathrm{diag}\left(\alpha^{\frac{|A|}{2}}\right) \cdot Q_{0,\alpha} \cdot \mathrm{diag}\left(\alpha^{\frac{|A|}{2}}\right)] \cdot L_\alpha^\top = M_\alpha \qquad \mathrm{mod}\ (*) \tag{6.15}$$

where recall $(*)$ *means ideal* $(\{\alpha^{|V(T)\cup I\cup J|+1}\chi_T\},\ \{\chi_T : |V(T)\cup I\cup J| > \tau\})$ *position-wise on each* (I, J).

Equation (6.15) is slightly weaker than a solution to (6.9) but is sufficient for all use since we are only concerned with PSDness. In particular, it gives the first-approximate diagonalization of the matrix M', recast as Definition 6.2 below. This shows Theorem 2(2).

[12]Recall $e_{I,A}(T')$ is the reduced size $|V(T') \cup I \cup A| - s_{I,A}(T')$ (Def. 4.6).

6.3. **Recursive factorization.** This subsection gives a systematic treatment of the recursive factorization technique, which we will use in the proof of Theorem 1. We formulate it on matrix-products (Def. 6.4, 6.5) with a simplification (Lem. 6.4) and extension (Prop. 6.4) that will be finally used in Section 7 in the exact case.

As in other parts of the section, we state everything in the non-exact case to avoid unnecessary complication. The goal is to refine the coarse diagonalization (6.15), recast below.

Definition 6.2. *Let L be the $\binom{[n]}{\frac{d}{2}} \times \binom{[n]}{\leq \frac{d}{2}}$-matrix*

$$L(I,A) := \sum_{\substack{T':\ |V(T')\cup I\cup A|\leq \tau \\ A=S_l(I,A;T') \\ T'\cap E(A)=\emptyset \\ (I,A;T')\ \text{left-generated}}} (\frac{\omega}{n})^{|V(T')\cup I\cup A|-|A|}\chi_{T'}, \tag{6.16}$$

and Q_0 be the $\binom{[n]}{\leq \frac{d}{2}} \times \binom{[n]}{\leq \frac{d}{2}}$-matrix

$$Q_0(A,B) := \sum_{\substack{T_m:|T_m\cup A\cup B|\leq \tau \\ A,B\in \mathrm{mSep}_{A,B}(T_m)}} (\frac{\omega}{n})^{|V(T_m)\cup A\cup B|}\chi_{T_m}. \tag{6.17}$$

Finally, let

$$D := \mathrm{diag}\left((\frac{\omega}{n})^{\frac{|A|}{2}}\right)_{A\in\binom{[n]}{\leq d/2}}. \tag{6.18}$$

We call $L(DQ_0)L^\top$ the **first-approximate diagonalization** *of M'.*

Despite of its name ("approximate"), the difference

$$M' - L(DQ_0D)L^\top \tag{6.19}$$

is, however, far from negligible. This is where the recursive factorization will be applied, and in the end it will give

$$M' = L\cdot[D\cdot(Q_0 - Q_1 + Q_2 ... \pm Q_{d/2})\cdot D]\cdot L^\top + \mathcal{E} \tag{6.20}$$

for some negligible error-matrix $\mathcal{E}$.

Remark 6.1. *The use of D in the above is superficial. We only keep it to make the middle matrices Q_i have slightly more convenient expressions.*

Let us start with some necessary notions.

6.3.1. **More notion on graphs**.

Definition 6.3. *(*[BHK+19] *Def. 6.5) For ribbon $\mathcal{R} = (I,J;T)$, the* **canonical decomposition** *is a ribbon triple $(\mathcal{R}_l,\mathcal{R}_m,\mathcal{R}_r) = ((I,A;T_l),(A,B;T_m),(B,J;T_r))$ as follows. $A = S_l(I,J;T)$, $B = S_r(I,J;T)$. $V(\mathcal{R}_l)$ is A unioned with the set of vertices reachable by paths from I in T without touching A, and $T_l = T|_{V(\mathcal{R}_l)}\backslash E(A)$. Symmetricallywe define $V(\mathcal{R}_l)$ and T_r. Finally, $T_m = T\backslash(T'\sqcup T'')$. R_l, R_m, R_r are called the* **left, middle, right ribbon of** $\mathcal{R}$*, respectively.*

Remark 6.2. *For better clarity, we list a few properties that follow from the definition of the canonical decomposition.*

1. $A = S_l(I,A;T_l)$, $B = S_r(B,J;T_r)$ (so they are unique min-separators of $\mathcal{R}_l,\mathcal{R}_r$, respectively);

2. $T_l\cap E(A) = \emptyset = T_r\cap E[A]$;

3. $\mathcal{R}_l$ is left-generated, $\mathcal{R}_r$ is right-generated;

4. $A,B\in \mathrm{mSep}_{A,B}(T_m)$ (in particular, $|A| = |B|$).

The above are properties about $\mathcal{R}_l$, $\mathcal{R}_m$, $\mathcal{R}_m$ individually ("inner" properties). There is also an intersection property on pairs of them ("outer" properties):

5. $V(\mathcal{R}_l)\cap V(\mathcal{R}_m)\subseteq A$, $V(\mathcal{R}_m)\cap V(\mathcal{R}_r)\subseteq B$, $V(\mathcal{R}_l)\cap V(\mathcal{R}_r)\subseteq A\cap B$. This implies $e(\mathcal{R}_l)+|V(\mathcal{R}_m)|+e(\mathcal{R}_r)=|V(\mathcal{R})|$.

The canonical decomposition can be *reversely* described, as follows.

Definition 6.4. *(Inner-, outer-canonicality) For a ribbon triple $(\mathcal{R}_l,\mathcal{R}_m,\mathcal{R}_r)=\Big((I,A;T_l),(A,B;T_m),(B,J;T_r)\Big)$, their* **ribbon-sum** *is ribbon $(I,J;T)$ where $T=T_l\oplus T_m\oplus T_r$ (i.e. each edge mod 2 sum). The triple is called* **inner-canonical***, if they satisfy the "inner" conditions:*

(6.21) *Items 1—4 in Remark 6.2.*

The triple is **outer-canonical** *if they satisfy the "outer" condition:*

(6.22) *Item 5 in Remark 6.2.*

The triple is **canonical** *if it is both inner- and outer-canonical.*

Proposition 6.3. *Canonical triples are 1-1 correspondent to their ribbon-sum, via ribbon sum and canonical decomposition.*

Proof. This follows directly by checking the definition. □

The above notions can be extended to related matrix products. Denote by $\mathbb{R}[\{\chi_T\}]$ the ring by adding fresh variables ("characters") χ_T's into $\mathbb{R}$ for every $T\subseteq\binom{[n]}{2}$ (fixing an n), with relations $\{\chi_{T'}\cdot\chi_{T''}=\chi_T\mid T'\oplus T''=T\}$.

Definition 6.5. *(Approximate form) Suppose matrices X,Y have their rows and columns indexed by subsets of $[n]$ and entries in $\mathbb{R}[\{\chi_T\}]$. A character in an entry of such matrix can be regarded as a ribbon on the side sets row and column. Assume all ribbons have size $\le\tau$ and X,Y have dimensions s.t. $XYX^\top$ is defined.*

Then every triple product (without collecting like-terms) in $XYX^\top$ has form

$$\underbrace{X(I,A;T_l)Y(A,B;T_m)X(J,B;T_r)}_{\text{nonzero in }\mathbb{R}}\chi_{T_l\oplus T_m\oplus T_r}, \tag{6.23}$$

and can be identified with a ribbon triple $(\mathcal{R}_l,\mathcal{R}_m,\mathcal{R}_r)$ in the natural way. We say (6.23) *is the* **resulting term** *of the ribbon triple; it is an* **outer-canonical product** *if the ribbon triple is outer-canonical. The* **approximation form** *of $XYX^\top$ is:*

$$XYX^\top=(XYX^\top)_{\text{can}}+(XYX^\top)_{\text{non-can}} \tag{6.24}$$

where $\big(XYX^\top\big)_{\text{out-can}}$ collects all terms of outer-canonical products, $(XYX^\top)_{\text{non-can}}$ collects all terms of non-outer-canonical products.

6.3.2. **Machinery of recursive factorization**. In our just established language, the first-approximate factorization (Def. 6.2) can be recast as

(6.25)
$$M'=[L(DQ_0D)L^\top]_{\text{can}}-\mathcal{E}_{\deg}=L(DQ_0D)L^\top-[L(DQ_0D)L^\top]_{\text{non-can}}-\mathcal{E}_{\deg}$$

where $\mathcal{E}_{\deg}$ consists of all terms in $[L(DQ_0D)L^\top]_{\text{can}}$ with $|V(T)\cup I\cup J|>\tau$. $\mathcal{E}_{\deg}$ is actually negligible in matrix norm[13], and the main task is to analyze the "main error", $[L(DQ_0D)L^\top]_{\text{non-can}}$. The key insight is:

(6.26) $[L(DQ_0D)L^\top]_{\text{non-can}}$ itself factors through $L,L^\top$ approximately, too.

[13]They are supported on rows and columns where G is a clique.

That is, $\exists Q_1$ s.t. $[L(DQ_0D)L^\top]_{\text{non-can}} = [L(DQ_1D)L^\top]_{\text{can}} + \mathcal{E}_{1;\text{negl}}$ for some $\mathcal{E}_{1;\text{negl}}$ where $[L(DQ_1D)L^\top]_{\text{can}} = L(DQ_1D)L^\top - [L(DQ_1D)L^\top]_{\text{non-can}}$ by (6.23); then we recurse on $[L(DQ_1D)L^\top]_{\text{non-can}}$]. We need the following notations to describe this.

Definition 6.6. *([BHK+19]) An* **improper ribbon** *is a ribbon plus with a new set of isolated vertices. In symbol, denote it as $\mathcal{R}^* = (A, B; T^*)$ with $T^* = T \sqcup \mathcal{I}$, T an edge-set and $\mathcal{I}$ a vertex set* disjoint *from $V(T) \cup A \cup B$. $\mathcal{I}$ is called the* **isolated vertex-set of** *$\mathcal{R}^*$, denoted by $\mathcal{I}(\mathcal{R}^*)$. $V(\mathcal{R}^*) := V(T) \cup A \cup B \cup \mathcal{I}$. $(A, B; T)$ is called the (unique) largest ribbon in $\mathcal{R}^*$. Note a usual ribbon is an improper ribbon with $\mathcal{I} = \emptyset$.*

Note $\mathcal{I}(\mathcal{R}^*)$ could be different from the set of isolated vertices of the underlying graph, since there can be isolated vertices in $A \cup B$.

Definition 6.7. *The triple $\big(\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r\big) = \big((I, A; T_l), (A, B; T_m), (B, J; T_r)\big)$ is called* **side-inner-canonical** *if the left and right ribbons, $\mathcal{R}_l$, $\mathcal{R}_r$ satisfy the inner-canonical conditions on their part (the first three of (6.21)), and $\mathcal{R}_m$ is just a ribbon.*

The following operation is the technical core of recursive factorization.

Definition 6.8. *(Separating factorization, [BHK+19]) Suppose a triple $(\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r) = ((I, A; T_l),\ (A, B; T_m),\ (B, J; T_r))$ is side-inner-canonical and non-outer-canonical. Let $T := T_l \oplus T_m \oplus T_r$ and Z be the multi-set of "unexpected" intersections, i.e. the multi-set of vertices from $(\mathcal{R}_l \cap \mathcal{R}_m) - A$, $(\mathcal{R}_m \cap \mathcal{R}_r) - B$, $(\mathcal{R}_l \cap \mathcal{R}_r) - (A \cap B)$. Call $|Z|$ the* **intersection size** *of the triple, denoted as $z(\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r)$. Note*

$$|V(\mathcal{R}_l) \cup V(\mathcal{R}_m) \cup V(\mathcal{R}_r)| = |V(\mathcal{R}_l)| + |V(\mathcal{R}_m)| + |V(\mathcal{R}_r)| - |A| - |B| - z. \tag{6.27}$$

We further separate this triple into an "outer-canonical" one, as follows.

Let S'_l be the leftmost min-separator of $(I, A \cup (Z \cap V(\mathcal{R}_l)); T_l)$, similarly S'_r the right-most min-separator of $(B \cup (Z \cap V(\mathcal{R}_r)), J; T_r)$. Note $S'_l, S'_r \subseteq V(T) \cup I \cup J$.

Define $\mathcal{R}'_l := (I, S'_l; T'_l)$, whose vertex set $V(\mathcal{R}'_l)$ is S'_l unioned with the set of vertices in $\mathcal{R}_l$ reachable from I by paths in T_l without touching S'_l, and T'_l is $T_l \backslash E(S'_l)$ restricted on $V(\mathcal{R}'_l)$. Ribbon $\mathcal{R}'_r$ is symmetrically defined. In particular, $T'_l \cap T'_r = \emptyset$. Then let $\mathcal{R}^*_m$ be the **improper** *ribbon $(S'_l, S'_r; T^*_m)$, $T^*_m := \big(T \backslash (T'_l \sqcup T'_r)\big) \sqcup \mathcal{I}(\mathcal{R}^*_m)$ where $\mathcal{I}(\mathcal{R}^*_m)$ collects all the rest isolated vertices:*

$$\mathcal{I}(\mathcal{R}^*_m) = V(\mathcal{R}_l) \cup V(\mathcal{R}_m) \cup V(\mathcal{R}_r) \ - \ V(T) \cup I \cup J. \tag{6.28}$$

$(\mathcal{R}'_l, \mathcal{R}^*_m, \mathcal{R}'_r)$ is called the **separating factorization** *of $(\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r)$, denoted as*

$$(\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r) \to (\mathcal{R}'_l, \mathcal{R}^*_m, \mathcal{R}'_r). \tag{6.29}$$

Remark 6.3. *Let $(\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r) \to (\mathcal{R}'_l, \mathcal{R}^*_m, \mathcal{R}'_r)$ be as above. We list some basic properties of this operation that are direct from the definition.*

(1). $(\mathcal{R}'_l, R^*_m, \mathcal{R}'_r)$ is side-inner- and outer-canonical. The latter means their pair-wise vertex intersections are in S'_l, S'_r and $S'_l \cap S'_r$, respectively. So if replace R^*_m by its largest ribbon, the triple would be canonical.

(2). $\mathcal{R}'_l \subseteq \mathcal{R}_l$, S'_l separates $(V(\mathcal{R}'_l),\ V(\mathcal{R}_l) - V(\mathcal{R}'_l))$ in $\mathcal{R}_l$. So we can talk about the part of $\mathcal{R}_l$ that is strictly to the right of S'_l, which is disjoint from R'_l and is further contained in $\mathcal{R}^*_m$. The similar fact holds for $\mathcal{R}_r$.

(3). In $\mathcal{R}_l$, since S'_l separates (I, A) and A is the unique min-separator of $\mathcal{R}_l$, there are $|A|$ many vertex-disjoint paths between A and S'_l. Similarly for $\mathcal{R}_r$.

Lemma 6.3. *Under the notation of Def. 6.8,*

(1). $|S'_l| + |S'_r| \geq |A| + |B| + 1$;

(2).[14] *Let* $s = \frac{|A|+|B|}{2}$, p' *be the max number of vertex-disjoint paths from* S'_l *to* S'_r *in* $\mathcal{R}^*_m$, *and* p *be the max number of vertex-disjoint paths from* A *to* B *in* $\mathcal{R}_m$, *then*

$$2(s'-s) + (p-p') + |\mathcal{I}(\mathcal{R}^*_m)| \le z(\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r).$$

Proof. (1): by definition, there must be some unexpected pair-wise intersection between the triple $(\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r)$. In either of the three cases of breaking (6.22), there exists some $v \in Z$ that is in $V(\mathcal{R}_l) - A$ or $V(\mathcal{R}_r) - B$. W.l.o.g., suppose the first case happens. Then $S'_l \neq A$ since v can be reached from I without passing A by the left-generated condition on $\mathcal{R}_l$. Similarly, if $|S'_l| = |A|$ then it is A as A is the unique min-separator separating (I, A), so this is impossible. Thus $S'_l > A$.

(2). This is Lemma 7.14 of [BHK+19]. We omit the proof here. □

6.3.3. **Apply the machinery to** M'. Now we analyze $[L(DQ_0D)L^\top]_{\text{non-can}}$ in (6.25). Conceptually, separating factorization allows us to "fix" $[L(DQ_0D)L^\top]_{\text{non-can}}$ using $L, L^\top$. Namely, a term $(\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r)$ in $[L(DQ_0D)L^\top]_{\text{non-can}}$ at the (I, J)-th position can be "countered" by the term $-(\mathcal{R}'_l, \mathcal{R}^*_m, \mathcal{R}'_r)$ in a new matrix product $[L(DQ_1D)L^\top]_{\text{can}}$: R'_l at entry (I, S'_l) in L, R'_r at entry (S'_r, J) in $L^\top$, and the largest ribbon of $\mathcal{R}^*_m$ at (S'_l, S'_r) in a new middle matrix DQ_1D.

Of course, there are other triples whose separating factorization is the same, and each entry of L is a sum of many different R'_ls, so we need to insure that this cancellation works for them simultaneously in multiplication.

The following proposition is what insures the simultaneous cancellation can work. We state a refined version (distinguishing the (i, j) parameters) that is more than needed here but will be fully need in the exact case (in Lem. 7.3).

Proposition 6.4. *(Solvability condition, cf. Claim 6.12 in* [BHK+19]*) Fix* (I, J, S'_l, S'_r) *and a improper ribbon* $\mathcal{R}^*_m$ *with side sets* (S'_l, S'_r). *Let* $(\mathcal{R}'_l, \mathcal{R}'_r)$ *be inner-canonical left and right ribbons with side sets* $(I, S'_l), (S'_r, J)$ *respectively, as in Def. 6.4. Let* $(\mathcal{R}''_l, \mathcal{R}''_r)$ *be another such ribbon pair, with the same reduced size* $e(\mathcal{R}'_l) = e(\mathcal{R}''_l)$, $e(\mathcal{R}'_r) = e(\mathcal{R}''_r)$ *(the same size, equivalently). Then for every fixed tuple* (i, j, z) *the following holds: there is an 1-1 matching between ribbon triples*

$$(\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r) \text{ s.t. } \begin{cases} (\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r) \to (\mathcal{R}'_l, \mathcal{R}^*_m, \mathcal{R}'_r), \\ (e(\mathcal{R}_l),\ e(\mathcal{R}_r),\ z(\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r)) = (i, j, z). \end{cases} \tag{6.30}$$

and

$$(\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r) \text{ s.t. } \begin{cases} (\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r) \to (\mathcal{R}''_l, \mathcal{R}^*_m, \mathcal{R}''_r), \\ (e(\mathcal{R}_l),\ e(\mathcal{R}_r),\ z(\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r)) = (i, j, z). \end{cases} \tag{6.31}$$

Moreover, this matching fixes every middle $\mathcal{R}_m$.

Proof. We give a reversible map from (6.30) onto (6.31). Take a $(\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r)$ from (6.30). By Remark 6.3 (2), the part of $\mathcal{R}_l$ to the right of S'_l is in $\mathcal{R}^*_m$ hence is disjoint from both R'_l and R''_l. Similarly for $\mathcal{R}'_r$, $\mathcal{R}_r$. Now take a map

$$(\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r) \mapsto (\phi(\mathcal{R}_l), \mathcal{R}_m, \phi(\mathcal{R}_r))$$

where $\phi(\mathcal{R}_l)$ replace $\mathcal{R}'_l$ by $\mathcal{R}''_l$ in $\mathcal{R}_l$, and $\phi(\mathcal{R}_r)$ replaces $\mathcal{R}'_r$ by $\mathcal{R}''_r$ in $\mathcal{R}_r$.

Clearly, $\mathcal{R}^*_m$ (thus $\mathcal{R}_m$) is unchanged. Since $\mathcal{R}'_l$, $\mathcal{R}''_l$ have the same size by assumption, by the disjointness property in Remark 6.3 (2), the replacement operation keeps the size of $\mathcal{R}_l$. Moreover, $\mathcal{R}_l$, $\phi(\mathcal{R}_l)$ have the same right set which is the unique min-separator of both, so $e(\mathcal{R}_l) = e(\phi(R_l))$. Similarly for $\mathcal{R}_r, \phi(\mathcal{R}_r)$, so the parameter (i, j) is unchanged by ϕ. The intersection parameter z is unchanged too,

[14]Recall in our setting $\mathcal{R}_m$ is always a ribbon, without any isolated vertex.

since the changed part is disjoint from $Z(\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r)$. Finally, the inverse map is given the same way by changing the role of $(\mathcal{R}'_l, \mathcal{R}'_r)$ and $(\mathcal{R}''_l, \mathcal{R}''_r)$. □

We now describe **one round of factorization**. Let L be as Def. (6.2), Q be any $\binom{[n]}{\le d/2} \times \binom{[n]}{\le d/2}$-matrix with $Q(A,B) = \sum_{T_m:\ |V(T_m)\cup A\cup B|\le\tau} (\frac{\omega}{n})^{|V(\mathcal{R}_m)|} q(\mathcal{R}_m) \cdot \chi_{T_m}$, where $\mathcal{R}_m$ denotes $(A, B; T_m)$ and $q(\cdot)$ is a function symmetric w.r.t. shapes. Below we define matrices $Q', \mathcal{E}_{\text{negl}}$ as follows so that

$$(LQL^\top)_{\text{non-can}} = (LQ'L^\top)_{\text{can}} + \mathcal{E}_{\text{negl}}. \tag{6.32}$$

First let $Q'(A,B) := \sum_{T_m:\ |V(T_m)\cup A\cup B|\le\tau} (\frac{\omega}{n})^{|V(\mathcal{R}_m)|} q'(\mathcal{R}_m)\chi_{T_m}$, $q'(\mathcal{R}_m)$ as below. Fix any $\mathcal{R}_m = (A, B; T_m)$, let $t = |V(\mathcal{R}_m)| (\le \tau)$, $s = \frac{|A|+|B|}{2}$, then for every improper ribbon $\mathcal{R}^*_m$ which contains $\mathcal{R}_m$ as its largest ribbon and $|V(\mathcal{R}^*_m)| \le \tau$, **fix** any pair $(\mathcal{R}'_l, \mathcal{R}'_r)$ s.t. $(\mathcal{R}'_l, \mathcal{R}^*_m, \mathcal{R}'_r)$ is the separating factorization for some ribbon triple with $|V(\mathcal{R}'_l)|, |V(\mathcal{R}'_r)| \le \tau$ (if there is none, exclude $\mathcal{R}^*_m$ in the summation below), let

$$\begin{aligned} q'(\mathcal{R}_m) &= \sum_{\substack{\mathcal{R}^*_m:\ \text{improper ribbon on } (A,B)\\ |V(\mathcal{R}^*_m)|\le\tau\\ \text{largest ribbon is } \mathcal{R}_m}} (\frac{\omega}{n})^{|\mathfrak{I}(\mathcal{R}^*_m)|} \cdot q''(\mathcal{R}^*_m) \quad \text{where} \\ q''(\mathcal{R}^*_m) &= \sum_{1\le z\le d/2} \sum_{\substack{\mathcal{P}=(\mathcal{R}_l,\mathcal{R},\mathcal{R}_r):\ \text{side-inn. can.}\\ \mathcal{P}\to(\mathcal{R}'_l,\mathcal{R}^*_m,\mathcal{R}'_r)\ \text{for the fixed}\ \mathcal{R}'_l,\mathcal{R}'_r\\ z(\mathcal{P})=z}} (\frac{\omega}{n})^z \cdot q(\mathcal{R}). \end{aligned} \tag{6.33}$$

Note $q'(\mathcal{R}_m)$ doesn't depend on the choice $(\mathcal{R}'_l, \mathcal{R}'_r)$ by Prop. 6.4, so $q'(\cdot)$ is also symmetric w.r.t. shapes. Now define $\mathcal{E}_{\text{negl}}$ such that (6.32) holds.

Lemma 6.4. *(One round) In the above notation,*

(1). W.p. $> 1 - n^{-9\log n}$ *over* G, $\|\mathcal{E}_{\text{negl}}\| \le \max\{q(A,B;T)\} \cdot n^{-\epsilon\tau}$;

(2). Given an $\mathcal{R}_m$*, let* p *be the max number of vertex-disjoint paths between in it between the two side sets. If there is a number* C *s.t.*

$$\forall \mathcal{R}_m,\ |q(\mathcal{R}_m)| \le C \cdot (\frac{\omega}{n^{1-\epsilon}})^{s-p} \tag{6.34}$$

then $|q'(\mathcal{R}_m)| \le C \cdot (\frac{\omega}{n^{1-\epsilon}})^{s-p+1/3}$ *for all* $\mathcal{R}_m$.

Proof. We compare $[LQ'L^\top]_{\text{can}}$, $[LQL^\top]_{\text{non-can}}$ as step (0), then prove (1), (2).

(0). For any fixed (I, J), recall $[LQL^\top]_{\text{non-can}}(I,J)$ is

$$\sum_{\substack{(\mathcal{R}_l,\mathcal{R}_m,\mathcal{R}_r):\ \text{side. inn. can.}\\ \text{non-outer-can.}\\ \text{all three have size}\ \le\tau}} (\frac{\omega}{n})^{|V(\mathcal{R}_l)|+|V(\mathcal{R}_m)|+|V(\mathcal{R}_r)|-|A|-|B|} q(\mathcal{R}_m)\chi_{T_l\oplus T_m\oplus T_r} \tag{6.35}$$

where we denote the side sets of $\mathcal{R}_m$ by (A, B). For each $(\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r)$ in the sum, there is a unique $(\mathcal{R}'_l, \mathcal{R}^*_m, \mathcal{R}'_r)$ that is its separating factorization: $(\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r) \to (\mathcal{R}'_l, \mathcal{R}^*_m, \mathcal{R}'_r)$. There are two cases of a term in (6.35).

First case: $|V(\mathcal{R}^*_m)| \le \tau$. In this case, there is the corresponding term

$$(\frac{\omega}{n})^{|V(\mathcal{R}'_l)|+|V(\mathcal{R}'_m)|+V(\mathcal{R}'_r)|-|S'_l|-|S'_r|} \cdot (\frac{\omega}{n})^{z+|\mathfrak{I}(\mathcal{R}^*_m)|} \cdot q(\mathcal{R}'_m)\chi_{T'_l\oplus T^*_m\oplus T'_r} \tag{6.36}$$

in $(LQ'L^\top)_{\text{can}}(I,J)$, where $\mathcal{R}'_m$ denotes the largest ribbon of $\mathcal{R}^*_m$, T^*_m means the edges of $\mathcal{R}'_m$, and $z \ge 1$ is the intersection size of $(\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r)$. In the separating factorization, recall $T'_l \oplus T^*_m \oplus T'_r = T_l \oplus T_m \oplus T_r$, $V(\mathcal{R}_l) \cup V(\mathcal{R}_m) \cup V(\mathcal{R}_r)| = |V(\mathcal{R}'_l)|+|V(\mathcal{R}^*_m)|+|V(\mathcal{R}'_r)|-|S'_l|-|S'_r| = |V(\mathcal{R}_l)|+|V(\mathcal{R}_m)|+|V(\mathcal{R}_r)|-|A|-|B|-z$ and $|V(R^*_m)| = |V(\mathcal{R}'_m)| + |\mathfrak{I}(\mathcal{R}^*_m)|$, so the coefficient in (6.36) equals the one in (6.35) for $(\mathcal{R}'_l, \mathcal{R}^*_m, \mathcal{R}'_r)$. Conversely, at a position $(\mathcal{R}'_l, \mathcal{R}'_r)$, $[LQ'L^\top]_{\text{can}}$ by (6.33)

and Prop. 6.4 collects exactly all terms from a triple $(\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r)$ in $[LQL^\top]_{\text{non-can}}$ whose separating factors have $(\mathcal{R}'_l, \mathcal{R}'_r)$ as the left, right part and whose $\mathcal{R}^*_m$ has size $\le \tau$.

Therefore, $\mathcal{E}_{\text{negl}}$ will collect exactly all terms in the next case.

Second case: $|V(\mathcal{R}^*_m)| > \tau$. By the above explanation, $\mathcal{E}_{\text{negl}}(I,J) =$

$$\sum_{\substack{(\mathcal{R}_l,\mathcal{R}_m,\mathcal{R}_r)\text{: side. inn. can.}\\ \text{non-outer-can.}\\ \text{all three has size } \le\tau\\ \text{resulting } |V(\mathcal{R}^*_m)|>\tau}} (\frac{\omega}{n})^{|V(\mathcal{R}_l)|+|V(\mathcal{R}_m)|+|V(\mathcal{R}_r)|-|A|-|B|} q(\mathcal{R}_m)\chi_{T_l\oplus T_m\oplus T_r}. \tag{6.37}$$

where we omit writing the sum condition "$\mathcal{R}_l$ ($\mathcal{R}_r$) has left (right) vertex-set as I (J)".

(1). Fix any $(\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r)$ in (6.37). Note $|\mathcal{I}(\mathcal{R}^*_m)| \le z + d/2$ as a quick corollary of Lemma 6.3. Fix $T = T_l \oplus T_m \oplus T_r$ and $a > \tau - |V(T)\cup I\cup J|$, we upper bound the number of triples in (6.37) resulting in $(\frac{\omega}{n})^{|V(T)\cup I\cup J|+a}\cdot\chi_T$ (ignoring $q(\mathcal{R}_m)$ for the moment): to create such a triple, we can choose a set as $\mathcal{I}(\mathcal{R}^*_m)$ of size $\le a/2 + d/4$ where a is intended to be $|\mathcal{I}(\mathcal{R}^*_m)| + z$ so $a \ge 2\mathcal{I}(\mathcal{R}^*) - d/2$ by the above; then decide the triple over the vertex set, where there are $< 3^{3\tau}\cdot 2^{3\binom{\tau}{2}}$ many ways.

So if let $B_0 = \max\{q(\cdot)\}$, then $\big|$coefficient of χ_T in (6.37)$\big|$ is no more than

$$B_0(\tfrac{\omega}{n})^{|V(T)\cup I\cup J|+a} n^{\frac{(a+d)}{2}} 2^{2\tau^2} = B_0(\tfrac{\omega}{n^{1-2\epsilon}})^{|V(T)\cup I\cup J|}\left(n^{-2\epsilon}\right)^{|V(T)\cup I\cup J|}(\tfrac{\omega}{\sqrt{n}})^a n^{\frac{d}{2}} 2^{2\tau^2}$$

$$\le B_0(n^{-1/2})^{|V(T)\cup I\cup J|} n^{-2\epsilon(|V(T)\cup I\cup J|+a)} n^{d/2} 2^{2\tau^2} \le B_0(n^{-1/2})^{|V(T)\cup I\cup J|} n^{-1.5\epsilon\tau},$$

where the last two steps use $\omega \le n^{1/2-4\epsilon}$, $|V(T)\cup I\cup J| + a > \tau$ (case condition) and $d < \epsilon\tau/10$, $2^{2\tau} < n^{\epsilon/10}$. Also, all χ_T appearing in (6.37) has $|V(T)| \le 3\tau$. By Lemma 4.2, for any (I,J), w.p.$> 1 - n^{-10\log n}$,

$$|\mathcal{E}_{\text{negl}}(I,J)| < \sum_{a=0}^{3\tau} B_0 n^{-a/2} n^{-1.5\epsilon\tau} n^{a/2} n^{4\log\log n} 2^{a^2} < n^{-1.4\epsilon\tau}.$$

By union bound on (I,J), w.p.$> 1 - n^{-9\log n}$, $\|\mathcal{E}_{\text{negl}}\| < n^d \cdot n^{-1.4\epsilon\tau} < n^{-\epsilon\tau}$.

(2). Fix an $\mathcal{R}_m$. By (6.33),

$$q'(\mathcal{R}_m) = \sum_{\substack{z,\mathcal{R}^*_m:\\ \text{largest ribbon } = \mathcal{R}_m}} (\frac{\omega}{n})^{|\mathcal{I}(\mathcal{R}^*_m)|+z} \sum_{\substack{\mathcal{P}=(\mathcal{R}_l,\mathcal{R},\mathcal{R}_r)\text{: side-inn. can.}\\ \mathcal{P}\to(\mathcal{R}'_l,\mathcal{R}^*_m,\mathcal{R}'_r)\text{ for the fixed }\mathcal{R}'_l,\mathcal{R}'_r\\ z(\mathcal{P})=z}} q(\mathcal{R}).$$

For a fixed $\mathcal{R}^*_m$, there are $\le 8^{z\tau} < n^{\epsilon z}$ many triples in the inner sum (recall $\mathcal{R}'_l, \mathcal{R}'_r$ are fixed). This is because, after fixing how each vertex appears in all three ribbons and fixing $A, B \subseteq \mathcal{R}^*_m$ as side sets, we only need to assign possible edges that appear in more than once in the original triple; such an edge must has at least one end in the already fixed (multi-set) Z. Then by Lem. 6.3(2) and (6.34), $\big|$ inner sum $\big| \le n^{\epsilon z}(\frac{\omega}{n})^{z+|\mathcal{I}(\mathcal{R}^*_m)|}\cdot|q(\mathcal{R})| \le (\frac{\omega}{n^{1-\epsilon}})^{2(s'-s)+(p-p')+2|\mathcal{I}(\mathcal{R}^*_m)|}\cdot C\cdot(\frac{\omega}{n^{1-\epsilon}})^{s-p}$ $\le C\cdot(\frac{\omega}{n})^{2|\mathcal{I}(\mathcal{R}^*_m)|}\cdot(\frac{\omega}{n^{1-\epsilon}})^{s'-p'+1/2}$, where (s,p) denote the corresponding parameters for $\mathcal{R}$ and (s',p') for $\mathcal{R}_m$, and the last step uses $s'-s \ge 1/2$ from Lem. 6.3(1).

Finally, in the outer sum, for fixed i_0 there are $< n^{i_0}$ many ways to choose $\mathcal{R}^*_m$ s.t. $|\mathcal{I}(\mathcal{R}^*_m)| = i_0$, $1 \le z \le 3\tau$. Together,

$$|q'(\mathcal{R}_m)| \le 3\tau\sum_{i_0=0}^{d/2} C\cdot n^{i_0}(\frac{\omega}{n})^{2i_0}\cdot(\frac{\omega}{n^{1-\epsilon}})^{s'-p'+1/2} \le C\cdot(\frac{\omega}{n^{1-\epsilon}})^{s'-p'+1/3}.$$

□

Apply Lem. 6.4 to $[L(DQ_0D)L^\top]_{\text{non-can}}$ with $Q \leftarrow (DQ_0D)$ then repeat as described under (6.26), we get the **recursive approximate factorization** of M':

$$M' = L\Big(D(Q_0-Q_1+Q_2-...\pm Q_d)D\Big)L^\top - \mathcal{E}_{\deg} + \Big(-\mathcal{E}_{1;\text{negl}}+...\pm\mathcal{E}_{1+d;\text{negl}}\Big). \tag{6.38}$$

Here it implicitly used:

Proposition 6.5. *(*[BHK+19] *Claim 6.15)* $Q_{d+1} = 0$.

Proof. First we use induction to show that for all k, in Q_k every appearing ribbon $R_m = (A, B; T_m)$ has $|A| + |B| \geq k$. The base case $k = 0$ is trivial. For $k + 1$, by Lemma 6.4 every $\mathcal{R}'_m = (A', B'; T'_m)$ in Q_{k+1} is the largest ribbon of some $\mathcal{R}^*_m$ in the separating factorization of some non-outer-canonical triple in $L(DQ_kD)L^\top$. Suppose that triple has the middle part $\mathcal{R}_m = (A, B; T_m)$, then by inductive hypothesis $|A| + |B| \geq k$. By Lemma 6.3(1), $|A'| + |B'| \geq |A| + |B| + 1 \geq k + 1$, completing induction. For $k = 1 + d$, no ribbon with both side sets in $\binom{[n]}{d/2}$ can satisfy this. □

We have completed the preparation of the recursive factorization technique.

Remark 6.4. *PSDness of M' would follow from* (6.38) *by some last steps*[15]*. Similar arguments will be given in section 8 so we omit them here.*

7. PSDness analysis, II: Recursive factorization

Now we apply the recursive approximate factorization to the target matrices in the exact setting, i.e. M_c^R in (5.14).

The high-level steps are the same as in section 6: we will first define the **first-approximate** factorization (Def. 7.1, 7.3 and Lem. 7.1), then refine it recursively to get the eventual factorization, Lem. 7.2), which is the main result of this section.

Definition 7.1. *Fix $R \in \binom{[n]}{\leq \frac{d}{2}}$. For every $i = 0, ..., \tau$ define matrix $L^{R,i}$ as*

$$L^{R,i}(I, A) = \begin{cases} 0 & \text{, if } R \not\subseteq I \cap A; \\ \sum_{\substack{T:\ |V(T)\cup I\cup A|\leq\tau \\ A=S_l(I,A;T) \\ T\cap E(A)=\emptyset \\ (I,A;T)\ \text{left-generated} \\ e_{I,A}(T)=i}} (\tfrac{\omega}{n})^i \chi_T & \text{, o.w.} \end{cases} \tag{7.1}$$

of dimension $\binom{[n]}{\frac{d}{2}} \times \binom{[n]}{\leq \frac{d}{2}}$. Let $\widetilde{L^R} := (L^{R,0}, ..., L^{R,\tau})$ the **left factor**, *$(\widetilde{L^R})^\top$ the* **right factor**. *Note these matrices do not depend on "c".*

Definition 7.2. $D^\tau := \operatorname{diag}\left((\frac{\omega}{n})^{\frac{|A|}{2}}\right)_{A\subseteq[n]:\ |A|\leq d/2} \otimes \mathrm{Id}_{\{0,...,\tau\}\times\{0,...,\tau\}}$.

Our goal is to find a middle, $\left(\binom{[n]}{\leq \frac{d}{2}} \times (\tau+1)\right) \times \left(\binom{[n]}{\leq \frac{d}{2}} \times (\tau+1)\right)$-matrix Q_c^R s.t. $M_c^R \approx \underbrace{(L^{R,0}, ..., L^{R,\tau})}_{\widetilde{L^R}} \cdot \big(D^\tau \cdot Q_c^R \cdot D^\tau\big) \cdot \underbrace{(L^{R,0}, ..., L^{R,\tau})^\top}_{\left(\widetilde{L^R}\right)^\top}$, achived as Lemma 7.2.

Remark 7.1. *Here the middle matrix has "larger" dimension ($\times\{0, ..., \tau\}$). The reason is that in* (5.12)*, or more broadly in any exact pseudo-expectation generated by the method in section 3.2, the parameter $a = |V(T) \cup I \cup J|$ appears nestedly in an essential way. That is, fix $(I, J; T)$, the non-exact coefficient* (3.8) *factors as*

[15]As noted previously, this is not yet the PSDness of the moment matrix as we do not have the homogeneous reduction in non-exact case. A full proof is just similar, though.

$(\frac{\omega}{n})^a = (\frac{\omega}{n})^{e(\mathcal{R}_l)+|V(\mathcal{R}_m)|+e(\mathcal{R}_r)}$ *(Remark 6.2) as the left, middle, right terms, but now terms like* $\binom{a+l-d}{c}\cdot\binom{n-a}{l-c}$ *are not log-additive in* a*. We deal with this by further specifying parameters* $(e(\mathcal{R}_l),\ e(\mathcal{R}_r)) \in \{0,...\tau\}\times\{0,...\tau\}$.

The main factor of the coefficients in M_c^R (5.15) can be separated into left, right, middle factors as before, $(\frac{\omega}{n})^a = (\frac{\omega}{n})^{e(\mathcal{R}_l)}\cdot(\frac{\omega}{n})^{|V(\mathcal{R}_m)|}\cdot(\frac{\omega}{n})^{e(\mathcal{R}_r)}$. We leave the "hard" factor $Y_c(r,a)$ to the middle matrix $Q_c^R\Big((\cdot,e_l),(\cdot,e_r)\Big)$, e_l, e_r "intended" as reduced sizes, as below.

Definition 7.3. *(First-approximate factorization, middle* $Q_{c,0}^R$*) Define* $Q_{c,0}^R$ *to be the* $\{0,...,\tau\}\times\{0,...,\tau\}$*-block matrix, each block of dimension* $\binom{[n]}{\leq d/2}\times\binom{[n]}{\leq d/2}$*, that is 0 outside of the principal minor* $S^R\times S^R$ *where*

$$S^R = \{(A,i)\in\binom{[n]}{\leq d/2}\times\{0,...,\tau\}\mid A\supseteq R, |A|+i\geq\frac{d}{2}\}, \tag{7.2}$$

and on this principal minor, $Q_{c,0}^R\Big((A,i),(B,j)\Big) =$

$$\sum_{\substack{T_m:|V(T_m)\cup A\cup B|\leq\tau\\ A,B\in\mathrm{mSep}_{A,B}(T_m)}} (\frac{\omega}{n})^{|V(T_m)\cup A\cup B|-\frac{|A|+|B|}{2}}\cdot\underbrace{Y_c\big(|R|,\ |V(T_m)\cup A\cup B|+(i+j)\big)}_{\text{defined by (5.12)}}\cdot\chi_{T_m} \tag{7.3}$$

$\widetilde{L^R}\cdot\big(D^\tau\cdot Q_{c,0}^R\cdot D^\tau\big)\cdot\Big(\widetilde{L^R}\Big)^\top$ *is called the* **first approximate factorization** *of* M_c^R.

Remark 7.2. *(Intended meaning of parameters in* $Q_{c,0}^R$*.)*

(1). The set S^R (7.2) *is defined independently of* c*, where the condition* $|A|+i\geq d/2$ *is by the intended meaning of* i *as* $|V(T')\backslash A|\geq|I|-|A|$ *for some ribbon* $(I,A;T')$ *in* $\widetilde{L^R}$*. If* $|A|+i<d/2$ *the corresponding column in* $\widetilde{L^R}$ *is always 0. Similarly for* j.

(2). $Q_{c,0}^R$ *is supported only on those* $((A,i),(B,j))\in S^R\times S^R$ *with* $|A|=|B|$.

(3). (cf. Remark 6.2) Regarding (7.3)*, in "canonical" situations (i.e. for outer-canonical products in* $\widetilde{L^R}\cdot\big(D^\tau\cdot Q_{c,0}^R\cdot D^\tau\big)\cdot\Big(\widetilde{L^R}\Big)^\top$*) it holds that*

$$|V(T_m)\cup A\cup B|+(i+j) = |V(T)\cup I\cup J|$$

for any ribbon $\mathcal{R}=(I,J;T)$ *that has* $(A,B;T_m)$ *as the middle part of its canonical decomposition and* $e(\mathcal{R}_l)=i$, $e(\mathcal{R}_r)=j$.

Lemma 7.1. *(*$Q_{c,0}^R$ *gives the first-approximation) Fix* R, $c\leq|R|$*. For every* $(I,J;T)$ *s.t.* $|V(T)\cup I\cup J|\leq\tau$ *and* $R\subseteq I\cap J$*, there is exactly one outer-canonical product in the* $XYX^\top$*-type matrix product*

$$\underbrace{\widetilde{L^R}}_{\text{as "X"}}\cdot\underbrace{\big(D^\tau\cdot Q_{c,0}^R\cdot D^\tau\big)}_{\text{as "Y"}}\cdot\Big(\widetilde{L^R}\Big)^\top. \tag{7.4}$$

It is from the canonical decomposition of $(I,J;T)$*, and results in term* $M_c^R(I,J;T)\chi_T$.

Proof. Suppose $R\subseteq I\cap J$. First, note every triple in (7.4) is inner-canonical by definition of $\widetilde{L^R}, Q_{c,0}^R$, so all outer-canonical triples there 1-1 correspond to their triple-product $(I,J;T)$ via the canonical decomposition.

Fix an $(I,J;T)$ and its canonical decomposition, where $|V(T)\cup I\cup J|\leq\tau$. $(I,A;T')$ appears exactly once in $\widetilde{L^R}(I,A)$ in block L^{R,e_l}, where $e_l = e_{I,A}(T')$;

similarly for $(J,B;T'')$ and $e_r = e_{J,B}(T'')$. Further, there is exactly one outer-canonical product in (7.4) corresponding to this triple, with coefficient

$$L^{R,e_l}(I,A;T')\cdot(\frac{\omega}{n})^{\frac{|A|}{2}}\cdot Q^R_{c,0}(A,B;T_m)\cdot(\frac{\omega}{n})^{\frac{|B|}{2}}\cdot L^{R,e_r}(J,B;T''). \tag{7.5}$$

By definition (7.1), (7.3), if let $a := |V(T)\cup I\cup J|$ then the above coefficient is $(\frac{\omega}{n})^a\cdot Y_c(|R|,a) = M^R_c(I,J;T)$. Compare (5.12), (5.15), where note $a = |V(T)\cup I\cup J| = e_l + |V(T_m)\cup A\cup B| + e_r$ by canonicality, we see that the lemma holds. □

Definition 7.4. *Let $\mathcal{E}^R_{c;\deg}$ be the matrix that collects all products in $[\widetilde{L^R}\cdot\big(D^\tau Q^R_{c,0}D^\tau\big)\cdot\Big(\widetilde{L^R}\Big)^\top]_{\text{can}}$ with $|V(T)\cup I\cup J|>\tau$ (cf. (6.25)), and $[\widetilde{L^R}\cdot(D^\tau Q^R_{c,0}D^\tau)\cdot(\widetilde{L^R})^\top]_{\text{non-can}}$ collects all terms from triples that are non-outer-canonical.*

Summarizing, we have the first-approximate factorization:

$$M^R_c = \widetilde{L^R}\cdot\big(D^\tau Q^R_{c,0}D^\tau\big)\cdot\Big(\widetilde{L^R}\Big)^\top - [\widetilde{L^R}\cdot\big(D^\tau Q^R_{c,0}D^\tau\big)\cdot\Big(\widetilde{L^R}\Big)^\top]_{\text{non-can}} - \mathcal{E}^R_{c;\deg}. \tag{7.6}$$

The crucial fact is that again matrix $[\widetilde{L^R}\cdot\big(D^\tau Q^R_{c,0}D^\tau\big)\cdot\Big(\widetilde{L^R}\Big)^\top]_{\text{non-can}}$ factorizes through $\widetilde{L^R},(\widetilde{L^R})^\top$ approximately, allowing us to factorize recursively (cf. (6.38)).

Definition 7.5. *For a fixed $R\subseteq[n]$, we say a function f defined on ribbons on the ground set $[n]$ is R-**symmetric w.r.t. shapes**, if f takes the same values on isomorphic ribbons whose side sets both contain R.*

Lemma 7.2. *(Recursive factorization, exact case) $\forall R\in\binom{[n]}{\le d/2}$, $0\le c\le|R|$,*

$$M^R_c = \widetilde{L^R}\cdot\left[D^\tau\left(Q^R_{c,0}-Q^R_{c,1}+...\pm Q^R_{c,d}\right)D^\tau\right]\cdot\Big(\widetilde{L^R}\Big)^\top + \mathcal{E}^R_c \quad \textit{where} \tag{7.7}$$

(1). All $Q^R_{c,k}$'s are supported on the principal minor $S^R\times S^R$ ((7.2));
(2). $Q^R_{c,0}$ is by Definition 7.3;
(3). $\forall 1<k\le d/2$, $Q^R_{c,k}$ is a $(\tau+1)\times(\tau+1)$-block-matrix supported on $S^R\times S^R$,

$$Q^R_{c,k}\Big((A,i),(B,j)\Big) = \sum_{T_m:|V(T_m)\cup A\cup B|\le\tau} q^R_{c,k}(\mathcal{R}_m,i,j)\cdot\chi_{T_m} \tag{7.8}$$

where we denote $\mathcal{R}_m=(A,B;T_m)$, $q^R_{c,k}(\cdot,i,j)$'s are R-symmetric w.r.t. shapes, and

$$\forall(i,j)\quad |q^R_{c,k}(\mathcal{R}_m,i,j)|\le\tau^{5\tau}\cdot(\frac{\omega}{n^{1-\epsilon}})^{s-p+k/3} \tag{7.9}$$

where $s=\frac{|A|+|B|}{2}$, p is the max number of vertex-disjoint paths between A, B in $\mathcal{R}_m$.

(4). For any G, $\mathcal{E}^R_c(G)$ is supported within rows and columns that is clique in G and contains R. Moreover, w.p. $>1-n^{-9\log n}$, $\|\mathcal{E}^R_c\|<n^{-\epsilon\tau/2}$.

To prove this lemma, we use a counterpart of Lemma 6.4 in the exact case, stated below. Fix an $R\subseteq\binom{[n]}{d/2}$ and for convenience denote $n_1:=\binom{[n]}{d/2}\times(\tau+1)$.

Lemma 7.3. *(One round of factorization, exact case)*

Let $\widetilde{L^R}$ be from Def. 7.1, Q^R be any $n_1\times n_1$-matrix supported on $S^R\times S^R$ and

$$Q^R((A,i),(B,j)) = \sum_{T_m:\ |V(T_m)\cup A\cup B|\le\tau}(\frac{\omega}{n})^{|V(\mathcal{R}_m)|}q(\mathcal{R}_m,i,j)\cdot\chi_{T_m} \tag{7.10}$$

where $\mathcal{R}_m$ denotes $(A,B;T_m)$, and $q(\cdot,i,j)$ is R-symmetric w.r.t. shapes for any fixed (i,j). Now we define matrix $Q',\mathcal{E}_{\text{negl}}$ so that

$$[\widetilde{L^R}\cdot Q\cdot(\widetilde{L^R})^\top]_{\text{non-can}} = [\widetilde{L^R}\cdot Q'\cdot(\widetilde{L^R})^\top]_{\text{can}} + \mathcal{E}_{\text{negl}}. \tag{7.11}$$

Namely, let Q' be only supported on $S^R \times S^R$, with expression $Q'((A,i),(B,j)) = \sum_{T_m:\ |V(T_m)\cup A\cup B|\le\tau} (\frac{\omega}{n})^{|V(\mathcal{R}_m)|} q'(\mathcal{R}_m,i,j)\chi_{T_m}$, where $q'(\mathcal{R}_m,i,j)$ is as follows. For a fixed $\mathcal{R}_m=(A,B;T_m)$ and (i,j), let $t=|V(\mathcal{R}_m)|\le\tau$, $s=\frac{|A|+|B|}{2}$, and for every improper ribbon $\mathcal{R}^*_m$ that contains $\mathcal{R}_m$ as its largest ribbon and $|V(\mathcal{R}^*_m)|\le\tau$, fix any a ribbon pair $(\mathcal{R}'_l,\mathcal{R}'_r)$ so that $(\mathcal{R}'_l,\mathcal{R}^*_m,\mathcal{R}'_r)$ is the separating factorization of some ribbon triple, $|V(\mathcal{R}'_l)|,|V(\mathcal{R}'_r)|\le\tau$ and

$$(e(\mathcal{R}'_l),e(\mathcal{R}'_r))=(i,j). \tag{7.12}$$

If there is no such choice, exclude this $\mathcal{R}^*_m$ in the summation below. Define

$$\begin{aligned} q'(\mathcal{R}_m,i,j) &= \sum_{\substack{\mathcal{R}^*_m:\text{ improper ribbon on }(A,B)\\ |V(\mathcal{R}^*_m)|\le\tau\\ \text{largest ribbon is }\mathcal{R}_m}} (\frac{\omega}{n})^{|\mathfrak{I}(\mathcal{R}^*_m)|}\cdot q''(\mathcal{R}^*_m,i,j) \quad \textit{where}\\ q''(\mathcal{R}^*_m,i,j) &= \sum_{\substack{(z,i_1,j_1):\\ 1\le z\le d/2}} \sum_{\substack{\mathcal{P}=(\mathcal{R}_l,\mathcal{R},\mathcal{R}_r)\text{: side-inn. can.}\\ \mathcal{P}\to(\mathcal{R}'_l,\mathcal{R}^*_m,\mathcal{R}'_r)\text{ for the fixed }\mathcal{R}'_l,\mathcal{R}'_r\\ z(\mathcal{P})=z,\ e(\mathcal{R}_l)=i_1,e(\mathcal{R}_r)=j_1}} (\frac{\omega}{n})^z\cdot q(\mathcal{R},i_1,j_1). \end{aligned} \tag{7.13}$$

*Here, $q''(\mathcal{R}_m,i,j)$ doesn't depend on the choice $(\mathcal{R}'_l,\mathcal{R}'_r)$ by (***the full of***) Prop. 6.4, so $q'(\cdot,i,j)$ is also R-symmetric w.r.t. shapes. Finally, $\mathcal{E}_{\text{negl}}$ is defined s.t.* (7.11) *holds. Then the conclusions are:*

(1). W.p. $>1-n^{-9\log n}$ over G, $\|\mathcal{E}_{\text{negl}}\|\le\max\{q(\cdot)\}\cdot n^{-\epsilon\tau}$;

(2). If there is a number C for which

$$\forall\mathcal{R}_m,i,j\quad |q(\mathcal{R}_m,i,j)|\le C\cdot(\frac{\omega}{n^{1-\epsilon}})^{s-p} \tag{7.14}$$

where p is the max number of vertex-disjoint paths between A,B in $\mathcal{R}_m$, then

$$\forall\mathcal{R}_m,i,j\quad |q'(\mathcal{R}_m)|\le C\cdot(\frac{\omega}{n^{1-\epsilon}})^{s-p+1/3}.$$

Proof. (of Lemma 7.3) The proof is almost the same as that of Lemma 6.4; we point out and explain the differences below.

The support condition (i.e. supported on $S^R\times S^R$) doesn't affect anything since $\widetilde{L^R}$ itself is automatically 0 on columns and rows that are not in S^R.

In step (0), we expand $[\widetilde{L^R}\cdot Q'\cdot(\widetilde{L^R})^\top]_{\text{can}}$ to compare with $[\widetilde{L^R}\cdot Q\cdot(\widetilde{L^R})^\top]_{\text{non-can}}$ term-wise, using Prop. 6.4. Here, notice that when (i,j) and $\mathcal{R}^*_m$ are fixed, the size of any choice of $(\mathcal{R}'_l,\mathcal{R}'_r)$ satisfying (7.12) are also fixed, so the proposition is applicable.

The comparison of orders on $(\frac{\omega}{n})$ between the two is the same as in step (0) in the proof of Lem. 6.4, and we get that $\mathcal{E}_{\text{negl}}$ collects all products in $[\widetilde{L^R}\cdot Q\cdot(\widetilde{L^R})^\top]_{\text{non-can}}$ whose $\mathcal{R}^*_m$ in separating factorization exceeds size τ. I.e. $\mathcal{E}_{\text{negl}}(I,J))=$

$$\sum_{i,j}\ \sum_{\substack{(\mathcal{R}_l,\mathcal{R}_m,\mathcal{R}_r)\text{: side inn. can.}\\ \text{non-outer-can.}\\ \text{all three has size }\le\tau\\ |V(\mathcal{R}^*_m)|>\tau,\ (e(\mathcal{R}_l),e(\mathcal{R}_r))=(i,j)}} (\tfrac{\omega}{n})^{|V(\mathcal{R}_l)|+|V(\mathcal{R}_m)|+|V(\mathcal{R}_r)|-|A|-|B|} q(\mathcal{R}_m,i,j)\chi_T,$$

where $T=T_l\oplus T_m\oplus T_r$ and we omitted writing the default condition in the summation that $\mathcal{R}_l$ ($\mathcal{R}_r$) has the left (right) side vertex set I (J).

Conclusions (1), (2) follow from the same estimates as in Lem. 6.4 (after (6.37)). Note the norm bound from Theorem 3 is still applicable to our case where a graph matrix can be nonzero only on (A,B) s.t. $R\subseteq A\cap B$; this is because we can process the original graph matrix by $\text{diag}(1_R)\cdot(-)\cdot\text{diag}(1_R)$ where $1_R(A)=1$ iff $R\subseteq A$, which does not increase norm. Finally, for (1) we have an extra $(1+\tau)^2$-factor compared with before (occurring from a union bound on blocks), but the estimate

in Lem. 6.4 is loose enough that when multiplied by this additional factor it is still $< n^{-\epsilon\tau}$. □

Proof. (of Lemma 7.2) We apply Lem. 7.3 to $[\widetilde{L^R} \cdot \left(D^\tau Q^R_{c,0} D^\tau\right) \cdot \left(\widetilde{L^R}\right)^\top]_{\text{non-can}}$ repeatedly; as the result we can express M^R_c as:

(7.15)

$$\widetilde{L^R}\left(D^\tau \left(Q^R_{c,0} - Q^R_{c,1} + ... \pm Q^R_{c,d}\right) D^\tau\right) \left(\widetilde{L^R}\right)^\top - \mathcal{E}^R_{c;\deg} + \left(-\mathcal{E}^R_{c,1;\text{negl}} + ... \pm \mathcal{E}^R_{c,d+1;\text{negl}}\right)$$

where again it uses that $Q^R_{c,d+1} = 0$, by the same Prop. 6.5.

(1). All $Q^R_{c,k}$ is supported within $S^R \times S^R$ by definition of a round (Lem. 7.3).

(2). This is by definition.

(3). The coefficients $\{q^R_{c,k}(\cdot, i, j)\}$ of each $Q^R_{c,k}$ $(k = 0, 1, ..., d)$ are always R-symmetric w.r.t. shapes by Lem. 7.3. By definition (7.3) and Lem. 5.2(4),

$$\forall \mathcal{R}_m, i, j \quad |q^R_{c,0}(\mathcal{R}_m)| = |Y_c(|R|, |\mathcal{R}_m|)| \cdot (\frac{\omega}{n})^{|V(T)\cup A\cup B|} \le \tau^{5\tau} \cdot 1.$$

Notice $Q^R_{c,0}$ is special in that for all $\mathcal{R}_m = (A, B; T_m)$ in it, there are $|A| = |B|$ many vertex-disjoint paths between A, B in $\mathcal{R}_m$, i.e. $s = p$ (as usual $s := \frac{|A|+|B|}{2}$ and p denotes the max number of vertex-disjoint paths between A, B). So the above can be equivalently written as $|q^R_{c,0}(\mathcal{R}_m)| \le (\frac{\omega}{n^{1-\epsilon}})^{s-p}\tau^{5\tau}$. Now we use Lem. 7.3(2), whose "$q(\cdot)$" is $q^R_{c,k}$ here, the "Q" matrix is $D^\tau Q^R_{c,k} D$, the "$(\frac{\omega}{n})^{|V(\mathcal{R}_m)|} q(\cdot)$" is $(\frac{\omega}{n})^{|V(\mathcal{R}_m)|-s} \cdot (\frac{\omega}{n})^s \cdot q^R_{c,k}$. As the result, $|q^R_{c,k}(\mathcal{R}_m, i, j)| \le \tau^{5\tau} \cdot (\frac{\omega}{n^{1-\epsilon}})^{s-p+k/3}$.

(4). When plug in G, both M^R_c $\widetilde{L^R}\left[D^\tau\left(Q^R_{c,0} - Q^R_{c,1} + ... \pm Q^R_{c,d}\right)D^\tau\right]\left(\widetilde{L^R}\right)^\top$ are supported on clique rows and columns that contain R by definition. So it is the case for the difference, $\mathcal{E}^R_c$, too. Thus we only need to show the norm bound on $\mathcal{E}^R_c := -\mathcal{E}^R_{c;\deg} + \left(-\mathcal{E}^R_{c,1;\text{negl}} + ... \pm \mathcal{E}^R_{c,d+1;\text{negl}}\right)$. By Lem. 7.3(2) and induction on $k = 0, ..., d$, it always holds that $|q^R_{c,k}| < \tau^{5\tau}$. Also for each $\mathcal{E}^R_{k;\text{negl}}$, by Lem. 7.3(1) w.p. $> 1 - n^{-9\log n}$, $\left\|\mathcal{E}^R_{k;\text{negl}}\right\| < \tau^{5\tau} n^{-\epsilon\tau} < n^{-0.9\epsilon\tau}$.

As for $\mathcal{E}^R_{c;\deg}$, recall by Def. 6.5, its (I, J)-th entry is the sum of outer-canonical products in $\widetilde{L^R} \cdot \left(D^\tau Q^R_{c,0} D^\tau\right) \cdot \left(\widetilde{L^R}\right)^\top$ at (I, J) where $|V(T) \cup I \cup J| > \tau$. Thus

$$\mathcal{E}^R_{c;\deg}(I, J) = \sum_{\substack{(\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r)\text{: side-inn.can.} \\ \text{outer.can.} \\ \text{all three has size } \le \tau \\ |V(T)\cup I\cup J| > \tau}} (\frac{\omega}{n})^{|V(T)\cup I\cup J|} \cdot q^R_{c,0}(\mathcal{R}_m, e(\mathcal{R}_l), e(\mathcal{R}_r))\chi_T$$

where $T = T_l \oplus T_m \oplus T_r$, and in the summation R_l (R_r) should have I (J) as the left (right) set. Note this equation uses $|V(T) \cup I \cup J| = e_l + e_r + |V(\mathcal{R}_m)|$, a fact from outer- and side-inner-canonicality. By canonicality again, the sum contributes to a $(I, J; T)$ by at most $3^{3\tau}$ triples. Since $3\tau \ge |V(T)\cup I\cup J| > \tau$ and $|q^R_{c,0}(\cdot)| < \tau^{5\tau}$, we have by Lem. 4.2 that $\left|\mathcal{E}^R_{c;\deg}(I, J)\right| < \tau^{6\tau} \cdot \sum_{c=0}^{3\tau} (\frac{\omega}{n})^{\max\{\tau, c\}}(n^{c/2} 2^{c^2} n^{4\log\log n}) < n^{-2\epsilon\tau}$ w.p. $> 1 - n^{-10\log n}$. So, by union bound over (I, J), $\left\|\mathcal{E}^R_{c;\deg}\right\| < n^{-d/4} n^{-2\epsilon\tau} < n^{-\epsilon\tau}$ w.p. $> 1 - n^{-9.5\log n}$.

Together, summing the bounds on $\left\|\mathcal{E}^R_{c,k;\text{negl}}\right\|$ and $\left\|\mathcal{E}^R_{c;\deg}\right\|$, by union bound over $k = 1, ..., d$, we get that w.p. $> 1 - n^{-9\log n}$, $\left\|\mathcal{E}^R_c\right\| < n^{-\epsilon\tau/2}$. □

8. PSDness analysis, III: Structural and pseudorandom matrices

In this section, we prove the Main Lemma 5.3. Recall by Lemma 7.2, for each R and c, $c \le |R|$, we have:

$$M_c^R = \widetilde{L^R} \cdot \left[D^\tau \underbrace{\left(Q_{c,0}^R - Q_{c,1}^R + ... \pm Q_{c,d}^R \right)}_{:=Q_c^R} D^\tau \right] \cdot \left(\widetilde{L^R}\right)^\top + \mathcal{E}_c^R.$$

The key is the following lemma.

Lemma 8.1. *W.p. $> 1 - n^{-8\log n}$ over G, the following holds.*

(1). $\forall R \in \binom{[n]}{\le d/2}$, $Q_{0,0}^R - Q_{0,1}^R + ... \pm Q_{0,\frac{d}{2}}^R \succeq \tau^{-7\tau} \cdot \mathrm{diag}\left(\widetilde{\mathrm{Cl}}\right)_{S^R \times S^R}$, where recall $S^R = \{(A,i) \in \binom{[n]}{\le d/2} \times \{0,...,\tau\} \mid A \supseteq R, |A| + i \ge \frac{d}{2}\}$.

(2). $\forall R$, $0 < c \le |R|$, $\pm\omega^{-c}\left(Q_{c,0}^R - Q_{c,1}^R + ... \pm Q_{c,\frac{d}{2}}^R\right) \preceq n^{-c/4}\mathrm{diag}\left(\widetilde{\mathrm{Cl}}\right)_{S^R \times S^R}$.

Proof plan of Lemma 8.1. Fix an $R \in \binom{[n]}{\le d/2}$. We will prove the lemma by three ingredients: Corollary 8.3, Lemma 8.3, Lemma 8.4.

Proof plan. Corollary 8.3 (section 8.1, 8.2): Positiveness of $Q_{0,0}^R$. This is the last real technical challenge. We use a natural "*structural part + pseudo-random part*" decomposition of $Q_{0,0}^R$ (Def. 8.2), aiming to show that on their common support, the structural part is positive enough and the pseudo-random part is small enough in norm. The main difficulty here is in analyzing $\mathbb{E}[Q_{0,0}^R]$ which, ultimately, is about the choice of generating function F in Definition 3.6.

Lemma 8.3, 8.4 (section 8.2): Other $Q_{c,k}^R$'s ($k > 0$ or $c > 0$), when timed with ω^{-c}, are small and appropriately supported. These are proved by standard means.

We carry out this plan in the upcoming two subsections 8.1, 8.2.

Definition 8.1. *Define the* **root diagonal-clique matrix** *as*

$$D_{\mathrm{Cl}}(A,B) = \begin{cases} 0 & \text{, if } A \ne B; \\ 2^{-\binom{|A|}{2}/2} \cdot \widetilde{\mathrm{Cl}}_A = 2^{-\binom{|A|}{2}/2} \sum_{T \subseteq E[A]} \chi_T & \text{, o.w.} \end{cases} \tag{8.1}$$

of dimension $\binom{[n]}{\le d/2} \times \binom{[n]}{\le d/2}$, so that $D_{\mathrm{Cl}}^2(A,A) = \widetilde{\mathrm{Cl}}(A)$ for all $A \in \binom{[n]}{d/2}$. Also let $D_{\mathrm{Cl}}^\tau := D_{\mathrm{Cl}} \otimes \mathrm{Id}_{\{0,...,\tau\}\times\{0,...,\tau\}}$ which is again diagonal.

Definition 8.2. *The* **structural-pseudorandom decomposition** *of $Q_{0,0}^R$ is*

$$Q_{0,0}^R = D_{\mathrm{Cl}}^\tau \cdot \mathbb{E}[Q_{0,0}^R] \cdot D_{Cl}^\tau + \left(Q_{0,0}^R - D_{\mathrm{Cl}}^\tau \cdot \mathbb{E}[Q_{0,0}^R] \cdot D_{Cl}^\tau\right), \tag{8.2}$$

where the summand $D_{\mathrm{Cl}}^\tau \cdot \mathbb{E}[Q_{0,0}^R] \cdot D_{Cl}^\tau$ is called the **structural part**, *and the summand $\left(Q_{0,0}^R - D_{\mathrm{Cl}}^\tau \cdot \mathbb{E}[Q_{0,0}^R] \cdot D_{Cl}^\tau\right)$ the* **pseudo-random part**.

8.1. Positiveness of the structural part, $\mathbb{E}[Q_{0,0}^R]$.

Proposition 8.1. *Fix $R \in \binom{[n]}{\le d/2}$ and $0 \le c \le |R|$, let $r := |R|$.*

(1). $\mathbb{E}[Q_{c,0}^R]$ is supported on the blockwise partial-diagonals $\{\left((A,i),(A,j)\right) \in S^R \times S^R\}$, where S^R is by (7.2) (i.e. requires $R \subseteq A$ and $|A| + \min\{i,j\} \ge d/2$).

(2). For all $\Big((A,i),(A,j)\Big) \in S^R \times S^R$, $\mathbb{E}[Q^R_{c,0}]\Big((A,i),(A,j)\Big) =$

$$\sum_{l=c}^{r}(-1)^{r-l}\frac{\binom{r}{l}}{(l-c)!}\binom{|A|+i+j+l-d}{c}\frac{\Big(|A|+8\tau^2+(l-c)+(i+j)\Big)!}{(8\tau^2)!} + O\left(\frac{\tau^{1.5\tau}}{n}\right). \tag{8.3}$$

In particular, for $c=0$,

$$\mathbb{E}[Q^R_{0,0}]\Big((A,i),(A,j)\Big) = \sum_{l=0}^{r}(-1)^{r-l}\frac{\binom{r}{l}}{l!}\cdot\frac{\Big(|A|+8\tau^2+l+(i+j)\Big)!}{(8\tau^2)!} + O\left(\frac{\tau^{1.5\tau}}{n}\right). \tag{8.4}$$

(3). For every $A \in \binom{[n]}{\le d/2}$ *let* $1_{A,A}$ *be the* $\binom{[n]}{\le d/2}\times\binom{[n]}{\le d/2}$*-matrix with a single 1 on position* (A,A)*. Then*

$$\mathbb{E}[Q^R_{0,0}] = \sum_{\substack{A\subseteq\binom{[n]}{\le d/2}\\ A\supseteq R}} 1_{A,A}\otimes\left[\left(\sum_{l=0}^{r}(-1)^{r-l}\frac{\binom{r}{l}}{l!}\cdot P_{|A|+l}\right)+E^R_A\right] \tag{8.5}$$

where for every fixed A, $P_{|A|+l}, E^R_A$ *are* $(\tau+1)\times(\tau+1)$*-matrices both supported on the principal minor* $\{i \mid d/2-|A|\le i\le\tau\}\times\{i \mid d/2-|A|\le i\le\tau\}$*, satisfying* $\left\|E^R_A\right\| < \frac{\tau^{2\tau}}{n}$ *and*

$$P_{|A|+l}(i,j) = \frac{\Big(|A|+l+8\tau^2+(i+j)\Big)!}{(8\tau^2)!},\quad d/2-|A|\le i,j\le\tau. \tag{8.6}$$

Proof. For (1), the constant terms in (7.3) correspond to $T_m=\emptyset$, which is nonzero only when $A=B$ for A,B in S^R.

For (2), by definition (7.3) notice again $T_m=\emptyset$ and $A=B$. $\mathbb{E}[Q^R_{c,0}((A,i),(A,j))]$ $= Y_c(\underbrace{|R|}_{:=r},\underbrace{|A|+i+j}_{:=a})$, which expands to:

$$\sum_{l=c}^{r}(-1)^{r-l}\binom{r}{l}\underbrace{\binom{a+l-d}{c}}_{\text{Def. 5.1}}\binom{n-a}{l-c}n^{-(l-c)}\frac{(a+l-c+8\tau^2)!}{(8\tau^2)!}. \tag{8.7}$$

Now use $\binom{n-a}{l-c}n^{-(l-c)} = \frac{1}{(l-c)!}\frac{(n-a)\ldots(n-a-(l-c)+1)}{n^{l-c}} = \frac{1}{(l-c)!}(1-O(d^2/n))$ and

$$\left|\binom{r}{l}\binom{a+l-d}{c}\binom{n-a}{l-c}n^{-(l-c)}\frac{(a+l-c+8\tau^2)!}{(8\tau^2)!}\right| < (4d)^d\cdot(9\tau^2)^d < \tau^\tau$$

to (8.7), we get (8.3). Further, in (8.7) when $c=0$ we have $\binom{a+l-d}{0}=0$ regardless of $a+l-d$ (any value of it, positive, negative or 0). And the same analysis gives (8.4).

For (3), each E^R_A has dimension $(\tau+1)\times(\tau+1)$ and each entry is absolutely $<\tau^{1.5\tau}/n$ from part (2). The expression of $P_{|A|+l}$ is directly from (8.4). □

Remark 8.1. *(Specialty of* $c=0$*). Comparing* $\mathbb{E}[Q^R_{0,0}]$ *and* $\mathbb{E}[Q^R_{c,0}]$ (8.3), (8.4)*, the specialty of the case* $c=0$ *is that the factor* $\binom{|A|+l-d}{0}$ *is* **always** *1, which is important for* $\mathbb{E}[Q^R_{0,0}]$ *to be positive. In cases* $c>0$, $\binom{|A|+l-d}{c}$ *might be 0 or negative depending on the order between* $0, c, |A|+l-d$*, making* $\mathbb{E}[Q^R_{c,0}]$ *possibly not PSD.*

Definition 8.3. *For every $m, t \in \mathbb{N}$, define the* **factorial Hankel matrix** *to be*

$$H_{m,t}(i,j) = (i+j+t)! \quad \forall 0 \le i,j \le m. \tag{8.8}$$

The following is our key observation on the structure of these matrices.

Proposition 8.2. *(Almost common decomposition of $\{H_{m,t}\}$)*
(1). $H_{m,t} = L_m \cdot \big(N_{m,t} \cdot D_{m,t} \cdot (N_{m,t})^\top\big) \cdot (L_m^\top)$ where $L_m, D_{m,t}$ are diagona and $N_{m,t}$ is lower-triangular, with expressions

$$L_m(i,i) = i! \qquad D_{m,t}(i,i) = \prod_{t'=1}^{t}(i+t') \qquad N_{m,t}(i,j) = \binom{i+t}{i-j}.$$

In particular, L_m is independent of t, and $H_{m,t}$ is positive.
(2). Let J_m be the usual $(1+m)\times(1+m)$ lower-triangular Jordan block

$$J_m(i,j) = \begin{cases} 1 & \text{, if } i=j \text{ or } i=j+1; \\ 0 & \text{, o.w.} \end{cases}$$

Then the "left factors" $N_{m,t}$ satisfy the recursive relation $N_{m,t+1} = N_{m,t} \cdot J_m$.

Proof. The two items follow from a direct inspection of the definition. □

Proposition 8.3. *If parameters m, t, r satisfy*

$$t+1 > 8 \cdot \max\{r^2, m\} \tag{8.9}$$

then it holds that $H_{m,t+1} \succeq 2r^2 H_{m,t}$.

Proof. By Proposition 8.2 it suffices to show that under (8.9),

$$J_m \cdot D_{m,t+1} \cdot J_m^\top \succeq 2r^2 D_{m,t}.$$

Equivalently, we need to compare the quadratic forms for fixed m:

$$q_{t+1}(x) := (x^\top J_m) D_{m,t+1} (J_m^\top x) \quad v.s. \quad q_t(x) := 2r^2 \cdot x^\top D_{m,t} x \tag{8.10}$$

where $x^\top = (x_0, ..., x_m)$ is the formal variable row-vector. Define two polynomials

$$\alpha(y) = 2r^2 \prod_{t'=1}^{t}(y+t'), \quad \beta(y) = \prod_{t'=1}^{t+1}(y+t').$$

Then we have $q_{t+1}(x) = \sum_{i=0}^{m} \beta(i)(x_i + x_{i+1})^2$ ($x_{m+1} := 0$) and $q_t(x) = \sum_{i=0}^{m} \alpha(i) x_i^2$.

To compare $q_t(x)$, $q_{t+1}(x)$, note $q_{t+1}(x) = \sum_{i=0}^{m} \beta(i) \cdot (x_i + x_{i+1})^2$

$$\sum_{i=0}^{m} \left[\alpha(i) x_i^2 + \Big(\beta(i) - \alpha(i)\Big) \cdot (x_i + \frac{\beta(i)}{\beta(i)-\alpha(i)} x_{i+1})^2 \; - \; \frac{\beta(i)^2}{\beta(i)-\alpha(i)} x_{i+1}^2 \right]$$

So if for $1 \le i \le m$ let $b_i := 1 - \frac{\alpha(i)}{\beta(i)} - \frac{\beta(i-1)}{\beta(i)} \frac{1}{b_{i-1}}$, $b_0 = 1 - \frac{\alpha(0)}{\beta(0)}$, then

$$q_{t+1}(x) = \underbrace{\sum_{i=0}^{m} \alpha(i) x_i^2}_{q_t(x)} \; + \; \sum_{i=0}^{m} \beta(i) b_i (x_i + \frac{1}{b_i} x_{i+1})^2. \tag{8.11}$$

Claim 8.1. *For all $i \le m$ we have $b_i > 1/2$.*

Proof. (of the claim) By definition, $b_0 = 1 - \frac{2r^2}{(t+1)}$ and

$$b_i = 1 - \frac{2r^2}{(t+1+i)} - \frac{i}{(t+1+i)} \cdot \frac{1}{b_{i-1}}, \quad i \ge 1. \tag{8.12}$$

Use induction for the claim: $b_0 = 1 - \frac{2r^2}{t+1} > 1/2$ by (8.9). For $1 \le i \le m$, $b_i = 1 - \frac{2r^2}{t+1+i} - \frac{i}{t+1+i} \cdot \frac{1}{b_{i-1}} \ge 1 - \frac{2r^2}{t+1} - \frac{m}{t+1} \cdot 2 > 1/2$ by (8.9) and inductive hypothesis. □

By (8.11) and positiveness of each b_i (Claim 8.1), $q_{t+1}(x) \ge q_t(x)$. This proves (8.10) and thus the proposition. □

Now we apply Proposition 8.3 to matrices $P_{|A|+l}$ (8.6). Note

$$P_{|A|+l} = \frac{1}{(8\tau^2)!} H_{\tau-(d/2-|A|),\ d-|A|+8\tau^2+l}$$

where A is fixed, l varies. We have the following:

Corollary 8.1. *(Positiveness of $\mathbb{E}[Q^R_{0,0}]$) In the decomposition* (8.5) *of $\mathbb{E}[Q^R_{0,0}]$,*

$$\left(\sum_{l=0}^{r}(-1)^{r-l}\frac{\binom{r}{l}}{l!}\cdot P_{|A|+l}\right) + E^R_A \succ \operatorname{diag}\left(\tau^{-6\tau}\right)_{0\le i\le \tau-(d/2-|A|)} \tag{8.13}$$

where we regard the matrices' support as $\{i \mid d/2-|A| \le i \le \tau)\}^2 \cong \{0,...,\tau-(d/2-|A|)\}^2$. In particular, by (8.5)

$$\mathbb{E}[Q^R_{0,0}] \succ \sum_{\substack{A\subseteq\binom{[n]}{\le d/2}\\ A\supseteq R}} 1_{A,A}\otimes \operatorname{diag}\left(\tau^{-6\tau}\right)_{d/2-|A|\le i\le\tau} = \operatorname{diag}\left(\tau^{-6\tau}\right)_{S^R\times S^R} \tag{8.14}$$

where recall $S^R = \{(A,i) \mid R\subseteq A, |A|+i \ge d/2\}$.

Proof. The "in particular" part is straightforward from (8.13) by checking the support, and that tensoring with a nonzero PSD matrix preserves the relation $\succ$. In below we prove for (8.13).

Fix A, let $\tau_0 = \tau-(d/2-|A|)$, $t_0 = d-|A|+8\tau^2$. Then

$$\sum_{l=0}^{r}(-1)^{r-l}\frac{\binom{r}{l}}{l!}\cdot P_{|A|+l} = \frac{1}{(8\tau^2)!}\cdot(X_r + X_{r-2} + ...) \tag{8.15}$$

where, $\forall 0\le v\le \lfloor r/2\rfloor$, $X_{r-2v} = \frac{\binom{r}{r-2v}}{(r-2v)!}\left(H_{\tau_0,t_0+r-2v} - \underbrace{\frac{(r-2v)^2}{(2v+1)}}_{\le r^2} H_{\tau_0,t_0+r-2v-1}\right)$

and $H_{\tau_0,-1} := 0$. Since $t_0 > 8\max\{r^2,\tau_0\}$, by Proposition 8.3

$$X_{r-2v} \succeq \frac{\binom{r}{r-2v}}{(r-2v)!}\cdot\max\{\frac{1}{2}H_{\tau_0,t_0+r-2v},\ r^2 H_{\tau_0,t_0+r-2v-1}\} \quad \forall 0\le v\le r/2.$$

So in (8.15), in particular,

$$\sum_{l=0}^{r}(-1)^{r-l}\frac{\binom{r}{l}}{l!}\cdot P_{|A|+l} \succeq \frac{1}{(8\tau^2)!}\cdot H_{\tau_0,t_0} \overset{\text{Prop. 8.2}}{=} L\left(N_{t_0}\cdot\frac{D_{t_0}}{(8\tau^2)!}\cdot(N_{t_0})^\top\right)L \tag{8.16}$$

where we temporarily abuse the notation by omitting the index τ_0 in the RHS.

Using the following claim, we can finish the proof of (8.13):

$$\begin{aligned}\text{RHS of (8.16)} &\succ L\cdot\operatorname{diag}\left(\tau^{-5\tau}\right)_{0\le i\le\tau_0}\cdot L \qquad \text{(by Claim 8.2)}\\ &\succeq \operatorname{diag}\left(\tau^{-5\tau}\right)_{0\le i\le\tau_0},\end{aligned}$$

while by Proposition 8.1 (3), $\left\|E^R_A\right\| < \frac{\tau^{2\tau}}{n} < \tau^{-6\tau}$ (using the parameter regime). So LHS of (8.13) $\succeq$ $\operatorname{diag}\left(\tau^{-5\tau}-\tau^{-6\tau}\right)_{0\le i\le\tau_0}$ $\succeq$ RHS of (8.13). □

Claim 8.2. *Under the notation of Cor. 8.1, the following holds:*

$$N_{t_0}^{-1}(i,j) = (-1)^{i-j}\binom{i+t_0}{i-j} \quad 0 \le i,j \le \tau_0 \tag{8.17}$$

(which is defined as 0 if $i < j$);

$$N_{t_0} \cdot \frac{D_{t_0}}{(8\tau^2)!} \cdot (N_{t_0})^\top \succ \operatorname{diag}\left(\tau^{-5\tau}\right)_{0\le i\le \tau_0}. \tag{8.18}$$

Proof. For (8.17), multiply this matrix with N_{t_0} then the (i,j)th entry is

$$\sum_{j\le k\le i}(-1)^{i-k}\binom{i+t_0}{i-k}\binom{k+t_0}{k-j} = \sum_{k'=0}^{i'}(-1)^{i'-k'}\binom{i'+j+t_0}{i'-k'}\binom{k'+j+t_0}{k'}$$

where $i' = i-j,\ k' = k-j$. To see it is the identity matrix, we use a generating function. Let $D_m[(1+x)^a]$ denote the coefficient of x^m in $(1+x)^a$, $m \ge 0, a \in \mathbb{Z}$, the above RHS $= (-1)^{i'}\sum_{k'=0}^{i'} D_{i'-k'}[(1+x)^{i'+j+t_0}]\cdot D_{k'}[(1+x)^{-(t_0+j+1)}] = (-1)^{i'}D_{i'}[(1+x)^{i'+j+t_0-(t_0+j+1)}] = (-1)^{i'}D_{i'}[(1+x)^{i'-1}] = 1_{i'=0}$.

As for (8.18), note it is equivalent to:

$$\frac{D_{t_0}}{(8\tau^2)!} \ \succ\ N_{t_0}^{-1}\cdot\tau^{-5\tau}\cdot(N_{t_0}^{-1})^\top. \tag{8.19}$$

To upper bound the RHS, let $a_0 = \tau^{-5\tau}$, consider the quadratic form

$$x^\top N_{t_0}^{-1}\cdot a_0\cdot(N_{t_0}^{-1})^\top x = a_0\sum_{j=0}^{\tau_0} y_j^2, \tag{8.20}$$

where by (8.17), $y_j = \left(x^\top N_{t_0}^{-1}\right)_j = \sum_{i=j}^{\tau_0}(-1)^{i-j}\binom{i+t_0}{i-j}x_i$. By Cauchy-Schwartz, $y_j^2 \le \tau_0\cdot\sum_{i=j}^{\tau_0}\binom{i+t_0}{i-j}^2 x_i^2$, thus RHS of (8.20) $= a_0\sum_{j=0}^{\tau_0} y_j^2 \le\ a_0\sum_{i=0}^{\tau_0} x_i^2\cdot\left(\tau_0\sum_{j=0}^{i}\binom{i+t_0}{i-j}^2\right)$ $< \sum_{i=0}^{\tau_0}\left(\tau^{-5\tau}\cdot(9\tau^2)^{2i+2}\right)x_i^2$. Now (8.19) follows since, for each i, in the LHS of (8.19) $= \frac{D_{t_0}(i,i)}{(8\tau^2)!} \ge (8\tau^2)^{-(d/2-|A|)}$ by definition, and the latter $> \tau^{-2d} > \tau^{-5\tau}\cdot(9\tau^2)^{2i+2}$ using $i \le \tau_0 < \tau$, $d \ll \tau$. Combining these two conclusions, we get (8.19). □

We arrive at the main conclusion of this subsection.

Corollary 8.2. *(Positiveness of the structural part of $Q_{0,0}^R$ (Def. 8.2))*

$$\underbrace{D_{\mathrm{Cl}}^{\tau}\cdot\mathbb{E}[Q_{0,0}^R]\cdot D_{\mathrm{Cl}}^{\tau}}_{\text{stractural part of } Q_{0,0}^R} \ \succeq\ \tau^{-6\tau}\cdot\operatorname{diag}\left(\widetilde{\mathrm{Cl}}\right)_{S^R\times S^R}.$$

Proof. It follows from Corollary 8.1 and the fact that $D_{\mathrm{Cl}}^2(A,A) = \widetilde{\mathrm{Cl}}(A)$ in Definition 8.1. □

8.2. **Rest bounds: $Q_{c,k}^R$s.** In this subsection, we bound the rest matrices:

$$\underbrace{Q_{0,0}^R - D_{\mathrm{Cl}}^{\tau}\cdot\mathbb{E}[Q_{0,0}^R]\cdot D_{\mathrm{Cl}}^{\tau}}_{\text{pseudo-random part of } Q_{0,0}^R \text{ (Def. 8.2)}} \quad,\quad Q_{0,k}^R\ (k>0),\quad \omega^{-c}\cdot Q_{c,k}^R\ (c>0,k\ge 0)$$

by three Lemmas 8.2, 8.3, 8.4, respectively, which would prove Lemma 8.1.

The arguments are standard but somewhat lengthy, as we need to be careful on the block structure and the support of matrices. Like in the proof of Lem. 7.3, when fixing an $R \subseteq [n]$ we only consider ribbons whose both side sets contain R,

so the corresponding graph matrices will be multiplied by $\mathrm{diag}(1_R)$ from left and right, where $1_R(A) = 1$ iff $R \subseteq A$; this does not affect the norm bound in Thm 3.

Definition 8.4. *Recall the (blocked) root diagonal-clique matrix D^τ_{Cl}, Def. 8.1. Denote by D' its 0-1 valued version. I.e., D' is diagonal and $D'((A,i),(A,i)) = \mathrm{Cl}_A$ for all $A \in \binom{[n]}{\le d/2}$ and $0 \le i \le \tau$.*

Lemma 8.2. *W.p. $> 1 - n^{-9\log n}$ the following holds: $\forall R \in \binom{[n]}{\le d/2}$,*

$$\pm(\underbrace{Q^R_{0,0} - D^\tau_{\mathrm{Cl}} \cdot \mathbb{E}[Q^R_{0,0}] \cdot D^\tau_{\mathrm{Cl}}}_{\text{pseudo-random part of } Q^R_{0,0}})(G) \;\preceq\; n^{-\epsilon} \cdot \mathrm{diag}\left(\widetilde{\mathrm{Cl}}(G)\right)_{S^R \times S^R} \tag{8.21}$$

Proof. **Fix R. In this proof abbreviate** $Q_{\mathrm{ps}} := Q^R_{0,0} - D^\tau_{\mathrm{Cl}} \cdot \mathbb{E}[Q^R_{0,0}] \cdot D^\tau_{\mathrm{Cl}}$ ("ps" for pseudo-random). It is $(\tau+1)\times(\tau+1)$-blocked with blocks $\left(Q_{\mathrm{ps},(i,j)}\right)_{0\le i,j\le\tau}$.

In block (i,j), by Def. 7.3 and Prop. 8.1, $Q_{\mathrm{ps},(i,j)}$ is supported within $S_{i,j} \times S_{i,j}$, where $S_{i,j} := \{A \mid |A| + \min\{i,j\} \ge d/2\}$. For each $A \ne B$, by Prop. 8.1 (1), $Q_{\mathrm{ps},(i,j)}(A,B) = Q^R_{0,0}((A,i),(B,j)) =$

$$\sum_{\substack{T_m:\ |V(T_m)\cup A\cup B|\le\tau \\ A,B\in \mathrm{mSep}_{A,B}(T_m)}} (\frac{\omega}{n})^{|V(T_m)\cup A\cup B| - \frac{|A|+|B|}{2}} \cdot q(A,B;T_m) \cdot \chi_{T_m} \tag{8.22}$$

and

$$Q_{\mathrm{ps},(i,j)}(A,A) = \sum_{T_m:\ 1\le|V(T_m)\setminus A|\le\tau-|A|} (\frac{\omega}{n})^{|V(T_m)\cup A|-|A|} \cdot q(A,A;T_m) \cdot \chi_{T_m}. \tag{8.23}$$

Here we have abbreviated $q(A,B;T_m) := Y_0\Big(|R|, |V(T_m)\cup A\cup B| + (i+j)\Big)$ ((7.3)) and have omitted the indices $|R|, i+j$ when they are fixed. Two properties we need:

$$q(A,B;T_m) \text{ depends only on } |V(T_m)\cup A\cup B| \text{ when fixing } (A,B); \tag{8.24}$$

$$\Big|q(A,B;T_m)\Big| < \tau^{5\tau} \qquad \text{(by Lemma 5.2 (4))}. \tag{8.25}$$

By (8.24), $Q_{\mathrm{ps},(i,j)}(A,B)$ always factors through $\mathrm{Cl}_{A\cup B}$ thus $\mathrm{Cl}_A\mathrm{Cl}_B$. In particular,

$$Q_{\mathrm{ps}} = D' \cdot Q_{\mathrm{ps}} \cdot D' \quad (D' \text{ from Def. 8.4}). \tag{8.26}$$

Claim 8.3. *W.p. $> 1 - n^{-9.5\log n}$, $\pm Q_{\mathrm{ps},(i,j)} \prec n^{-1.1\epsilon}\mathrm{diag}\left(2^{\binom{|A|}{2}}\right)_{S^R_l \times S^R_l}$ for all (i,j), where $l := \min\{i,j\}$ and $S^R_l := \{A \in \binom{[n]}{\le d/2} \mid A \supseteq R,\ |A| + l \ge d/2\}$.*

The lemma follows from this claim and (8.26), as follows. We consider a different decomposition of Q_{ps}: for every $b \in [0, \frac{d}{2}]$, let $I_b := \{i \mid d/2 - b \le i \le \tau\}$, and let $Q_{\mathrm{ps};b}$ be the principal minor on $W_b := \left(P^R_b \times I_b\right) \times \left(P^R_b \times I_b\right)$ of Q_{ps} (0 elsewhere), where $P^R_b = \{A \subseteq [n] \mid R \subseteq A, |A| = b\}$. Then

$$\{((A,i),(B,j)) \in S^R \times S^R \mid 0 \le |A| = |B| \le d/2\} = \bigsqcup_{b=0}^{d/2} W_b \quad \text{(disjoint union)}.$$

Note $Q^R_{c,0}$ is supported only on those $((A,i),(B,j)) \in S^R \times S^R$ with $|A| = |B|$ (Remark 7.2(2)); in particular for $c = 0$, we have a decomposition $Q_{ps} = \sum_{b=0}^{d/2} Q_{\mathrm{ps};b}$.

Now inside block $I_b \times I_b$, $Q_{\mathrm{ps};b}$ is further block-wise, each block a principal minor of $Q_{\mathrm{ps},(i,j)}$. By Claim 8.3, ($\pm$) all such blocks $\prec n^{-1.5\epsilon} \cdot \mathrm{diag}\left(2^{\binom{b}{2}}\right)_{P^R_b \times P^R_b}$

together w.p. $> 1 - n^{-9.5\log n}$, which implies $\pm Q_{\mathrm{ps};b} \prec \tau^2 \cdot n^{-1.5\epsilon}\mathrm{diag}\left(2^{\binom{b}{2}}\right)_{W_b} \prec n^{-\epsilon}\mathrm{diag}\left(2^{\binom{b}{2}}\right)_{W_b}$. So, summing over b, $\pm Q_{\mathrm{ps}} \prec n^{-\epsilon}\mathrm{diag}\left(2^{\binom{|A|}{2}}\right)_{S^R\times S^R}$ w.p. $1 - n^{-9\log n}$. Insert this to the middle of (8.26), where $\widetilde{\mathrm{Cl}}_A = 2^{\binom{|A|}{2}} \cdot \mathrm{Cl}_A$, $\mathrm{Cl}_A = \mathrm{Cl}_A^2$, we get (8.21). □

Proof. (of Claim 8.3) We use the norm bounds from section 4. Fix (i,j), consider $Q^{\mathrm{diag}}_{\mathrm{ps},(i,j)}$ and $Q^{\mathrm{off}}_{\mathrm{ps},(i,j)} = Q_{\mathrm{ps},(i,j)} - Q^{\mathrm{diag}}_{\mathrm{ps},(i,j)}$ separately.

Diagonal part. For any (A,A) in the support (i.e. $|A|+i \geq d/2$, $|A|+j \geq d/2$),

$$Q^{\mathrm{diag}}_{\mathrm{ps},(i,j)}(A,A) = \widetilde{\mathrm{Cl}}_A \underbrace{\left(\sum_{\substack{T_m:\ 1\leq |V(T_m)\backslash A|\leq \tau-|A| \\ T_m\cap E[A]=\emptyset}} (\frac{\omega}{n})^{|V(T_m)\backslash A|} q(A,A;T_m)\chi_{T_m}\right)}_{:=g(A)} \text{ by (8.23).}$$

This $g(A)$ can be bounded by norms of diagonal graph matrices as follows. First, $q(A,A;T_m)$ depends only on $|V(T_m)\backslash A|$ (we have fixed R,i,j,A), so temporarily denote it as $q(|V(T_m)\backslash A|)$. For any $1 \leq v \leq \tau - |A|$ let $\mathcal{U}^v_1, ..., \mathcal{U}^v_{h(v)}$ be all different shapes $(A,A;T)$ (Def. 4.4) s.t. $T\cap E[A] = \emptyset$, $|V(T)\backslash A| = v$. Note

$$h(v) \leq 2^{|A|v+v^2} \quad \text{since we required } T\cap E[A]=\emptyset. \tag{8.27}$$

So w.p. $> 1 - n^{-9.6\log n}$, $|g(A)| = \left|\sum_{v=1}^{\tau-|A|} (\frac{\omega}{n})^v q(v) \cdot \left(\sum_{x=1}^{h(v)} \underbrace{\sum_{T_m:(A,A;T_m) \text{ has shape } \mathcal{U}^v_x} \chi_{T_m}}_{=M_{\mathcal{U}^v_x}(A,A) \text{ by Def. 4.4}}\right)\right|$

$\leq \sum_{v=1}^{\tau-|A|} (\frac{\omega}{n})^v q(v) \cdot \sum_{x=1}^{h(v)} \left\|M_{\mathcal{U}^v_x}\right\| \leq \sum_{v=1}^{\tau-|A|} (\frac{\omega}{n})^v \tau^{5\tau} \sum_{x=1}^{h(v)} \left\|M_{\mathcal{U}^v_x}\right\|$ by (8.25) and that each $M_{\mathcal{U}^v_x}$ is diagonal; this is further $< \sum_{v=1}^{\tau} (\frac{\omega}{n})^v \tau^{5\tau} \cdot 2^{|A|v+v^2} \cdot n^{\frac{v}{2}} 2^{O(|A|+v)}$ by (8.27) and Theorem 3, which is $< \sum_{v=1}^{\tau} n^{-3\epsilon v} \cdot n^{\epsilon v} < n^{-1.2\epsilon}$ in our parameter regime.

Off-diagonal part. By R-symmetry of coefficients (8.24), $Q^{\mathrm{off}}_{\mathrm{ps},(i,j)}$ is a sum of graph matrices. Let $\mathcal{U}^{s,t}_1, ..., \mathcal{U}^{s,t}_{h(s,t)}$ be all shapes $(A,B;T)$ s.t. $|A| = |B| = s$, $A \neq B$, $A,B \in \mathrm{mSep}_{A,B}(T)$ and $|V(T)\cup A\cup B| = t$, then by (8.22), $Q^{\mathrm{off}}_{\mathrm{ps},(i,j)}$ is a block-diagonal matrix, with blocks $s = d/2-i, ..., d/2$ according to $s = |A| = |B|$, the sth block being $Q^{\mathrm{off}}_{\mathrm{ps},(i,j)}(s) = \sum_{t:\ s<t\leq\tau} (\frac{\omega}{n})^{t-s} \sum_{x=1}^{h(s,t)} q(\mathcal{U}^{s,t}_x) M_{\mathcal{U}^{s,t}_x}$. Here naturally, we denote $q(A,B;T_m) = q(\mathcal{U}^{s,t}_x)$ if $(A,B;T_m)$ has shape $\mathcal{U}^{s,t}_x$. By Theorem 3,

$$\left\|Q^{\mathrm{off}}_{\mathrm{ps},(i,j)}(s)\right\| \leq \sum_{s<t\leq\tau} (\frac{\omega}{n})^{t-s} \cdot h(t,s) \cdot n^{\frac{t-s}{2}} 2^{O(t)} (\log n)^{O(t-s)} \tag{8.28}$$

w.p. $> 1 - n^{-9.8\log n}$. Also clearly, $h(t,s) \leq 2^{\binom{t}{2}+O(t)}$. So with the same probability, the RHS of (8.28) $\leq \sum_{\substack{d/2-\max\{i,j\}\leq s\leq d/2 \\ s<t\leq\tau}} (\frac{\omega}{n})^{t-s} 2^{\binom{t}{2}+O(t)} n^{\frac{t-s}{2}} (\log n)^{O(t-s)}$ where note

$(\frac{\omega}{n})^{t-s} 2^{\binom{t}{2}+O(t)} n^{\frac{t-s}{2}} (\log n)^{O(t-s)} \leq n^{-2\epsilon(t-s)} 2^{O(t)} 2^{\binom{s}{2}} (2^{t+s}\log n)^{O(t-s)} < 2^{\binom{s}{2}} n^{-1.95\epsilon}$

Taking the blocks together, we get $\pm Q^{\mathrm{off}}_{\mathrm{ps},(i,j)} \prec n^{-1.9\epsilon} \cdot \mathrm{diag}\left(2^{\binom{|A|}{2}}\right)_{S^R_{\min\{i,j\}}\times S^R_{\min\{i,j\}}}$.

By union bound on the two parts, we get that w.p. $> 1 - n^{-9.5\log n}$,

$$\pm Q_{\mathrm{ps},(i,j)} = \pm(Q^{\mathrm{diag}}_{\mathrm{ps},(i,j)} + Q^{\mathrm{off}}_{\mathrm{ps},(i,j)}) \ \prec\ n^{-1.5\epsilon}\cdot \mathrm{diag}\left(2^{\binom{|A|}{2}}\right)_{S^R_{\min\{i,j\}}\times S^R_{\min\{i,j\}}}.$$

□

Corollary 8.3. *(Positiveness of $Q^R_{0,0}$) For any $R\in\binom{[n]}{\le d/2}$, w.p. $>1-n^{-8\log n}$,*

$$Q^R_{0,0}(G)\ \succeq\ \tau^{-6.1\tau}\cdot\mathrm{diag}\left(\widetilde{\mathrm{Cl}}(G)\right)_{S^R\times S^R}.$$

Proof. This is by Lem. 8.2, Cor. 8.2, and the fact that $\tau^{-6.1\tau}\gg n^{-\epsilon/10}$. □

Lemma 8.3. *(Bounds on $Q^R_{0,k}$) W.p. $>1-n^{-9\log n}$ the following holds. For all $R\in\binom{[n]}{\le d/2}$ and all $1\le k\le d/2$, $\pm Q^R_{0,k}(G)\ \preceq\ n^{-k/10}\cdot\mathrm{diag}\left(\widetilde{\mathrm{Cl}}(G)\right)_{S^R\times S^R}$.*

Proof. We will use union bound over (R,k), so fix one first. **For the fixed R, $k(>0)$, we abbreviate $Q^R_{0,k}$ as Q in this proof.**

Recall the definition of $Q^R_{0,k}$ (Lem. 7.2 (3)): Q is supported within $S^R\times S^R$,

$$Q\Big((A,i),(B,j)\Big) = \sum_{T_m:|V(T_m)\cup A\cup B|\le\tau}(\frac{\omega}{n})^{t-s}q^R_{0,k}(\mathfrak{R}_m,i,j)\cdot\chi_{T_m}. \tag{8.29}$$

where $t=|A\cup B|$, $s=\frac{|A|+|B|}{2}$. Abbreviate $q^R_{0,k}$ as q_k. By Lemma 7.2(3), $q_k(\cdot,i,j)$ is R-symmetric w.r.t. shapes for all fixed (i,j) (the *R-symmetry condition*), and also $|q_k(\mathfrak{R}_m,i,j)|\le\tau^{5\tau}(\frac{\omega}{n^{1-\epsilon}})^{s-p+k/3}$ (the *coefficient-size condition*) where $t=|A\cup B|$, $s=\frac{|A|+|B|}{2}$ and p is the max number of vertex-disjoint paths from A to B in T_m. By symmetry of q_k's, $Q((A,i),(B,j))$ factors through $\mathrm{Cl}(A)\mathrm{Cl}(B)$, so

$$Q = D'\cdot Q\cdot D' \tag{8.30}$$

where D' is by Definition 8.4. It suffices to show that

$$\text{w.p. } >1-n^{-9.5\log n}\quad \pm Q\ \prec\ n^{-k/10}\cdot\mathrm{diag}\left(2^{\binom{|A|}{2}}\right)_{S^R\times S^R}. \tag{8.31}$$

This is because, like in the proof of Lemma 8.2, we can insert (8.31) to the middle of (8.30) which proves the lemma for the fixed R,k. Below, we prove (8.31).

As a blocked matrix $Q=(Q_{(i,j)})_{0\le i,j\le\tau}$, $Q_{(i,j)}$ supported on A's s.t. $|A|+i\ge d/2$. **For any fixed** (i,j), any $(s_1,s_2)\in\{0,...,d/2\}^2$ s.t. $s_1+i\ge d/2$, $s_2+j\ge d/2$, and any $t\ge\max\{s_1,s_2\}$, let $\mathcal{U}_1^{t;s_1,s_2},...,\mathcal{U}_{h(t;s_1,s_2)}^{t;s_1,s_2}$ be all different shapes $(A,B;T)$ where $|A|=s_1$, $|B|=s_2$, $|V(T)\cup A\cup B|=t$. Then by (8.29) and R-symmetry,

$$Q_{(i,j)} = \sum_{\substack{(t;s_1,s_2)\\ s_1+i,s_2+j\ge d/2\\ \tau\ge t\ge s_1,s_2}}\ \sum_{x=1}^{h(t;s_1,s_2)} q_k(\mathcal{U}_x^{(t;s_1,s_2)},i,j)\cdot M_{\mathcal{U}_x^{(t;s_1,s_2)}}.$$

This can be alternatively expressed as $Q_{(i,j)}=\sum_{\substack{s_1,s_2\\ s_1+i,s_2+j\ge d/2}}Q_{(s_1,i),(s_2,j)}$ where

$$Q_{(s_1,i),(s_2,j)} := \sum_{\substack{t:\\ s_1,s_2\le t\le\tau}}\ \sum_{x=1}^{h(t;s_1,s_2)} q_k(\mathcal{U}_x^{(t;s_1,s_2)},i,j)\cdot M_{\mathcal{U}_x^{(t;s_1,s_2)}}. \tag{8.32}$$

$Q_{(s_1,i),(s_2,j)}$ is a $\binom{[n]}{s_1}\times\binom{[n]}{s_2}$-matrix on the (i,j)th block, and w.p. $>1-n^{-10\log n}$

$$\left\|Q_{(s_1,i),(s_2,j)}\right\| \le \sum_{\substack{t:\ t\le\tau\\ t\ge s_1,s_2}} h(t;s_1,s_2)\cdot(\frac{\omega}{n})^{t-s}(\frac{\omega}{n^{1-\epsilon}})^{s-p+k/3}\cdot n^{\frac{t-p}{2}}2^{O(t)}(\log n)^{O(t-s)} \tag{8.33}$$

by Thm. 3 and *coefficient-size condition*, where $s = \frac{s_1+s_2}{2}$ and p is the max number of vertex-disjoint paths between the two side sets. Since $h(t; s_1, s_2) \le 2^{\binom{t}{2}+O(t)} = 2^{\binom{s}{2}+O(t)+(t+s)\cdot(t-s)}$, (8.33) implies (note $k > 0$, $2^{O(t)} < n^{\epsilon/10}$, $\tau^{5\tau} < n^{1/30}$)

$$\left\|Q_{(s_1,i),(s_2,j)}\right\| < 2^{\binom{s}{2}} \cdot \tau^{5\tau} n^{-k/6} n^{-\epsilon(t-s)} < 2^{\binom{s}{2}} n^{-k/8}. \tag{8.34}$$

Finally, we sum over all double-blocks and use Cauchy-Schwartz. Namely, regard each $Q_{(s_1,i),(s_2,j)}$ as on $S^R \times S^R$ (extended by 0's), $Q = \sum_{\substack{(s_1,i),(s_2,j)\\ s_1+i,s_2+j\ge d/2}} Q_{(s_1,i),(s_2,j)}$ where $\pm Q_{(s_1,i),(s_2,j)} \prec n^{-k/8} \cdot \left(2^{\binom{s_1}{2}}\mathrm{Id}_{(s_1,i),(s_1,i)} + 2^{\binom{s_2}{2}}\mathrm{Id}_{(s_2,j),(s_2,j)}\right)/2$ by (8.34) and Cauchy-Schwartz. Summing over $(s_1, i), (s_2, j)$, w.p. $> 1 - n^{-9.5\log n}$

$$\pm Q \prec \tau^2 n^{-k/8} \mathrm{diag}\left(2^{\binom{|A|}{2}}\right)_{S^R\times S^R} \prec n^{-k/10} \mathrm{diag}\left(2^{\binom{|A|}{2}}\right)_{S^R\times S^R}.$$

□

Lemma 8.4. *(Bounds on $Q^R_{c,k}$, $c > 0$) W.p. $> 1 - n^{-9\log n}$ the following holds: for all R, c, k s.t. $R \in \binom{[n]}{\le d/2}$, $0 < c \le |R|$ and $0 \le k \le d/2$,*

$$\pm\omega^{-c} \cdot Q^R_{c,k} \preceq n^{-c/3} \cdot \mathrm{diag}\left(\widetilde{\mathrm{Cl}}\right)_{S^R\times S^R}. \tag{8.35}$$

Proof. The proof is almost the same as the previous one (Lemma 8.3). First, by a union bound over all such (R, c, k), it suffices to show that w.p. $> 1 - n^{-9.5\log n}$ the inequality holds for a fixed (R, c, k), which we prove below.

Fix (R, c, k) as in the lemma. If $k > 0$ then the proof is identical to that of Lemma 8.3 ($c = 0$), as the same *R-symmetry* and *coefficient-size* conditions hold (by Lem. 7.2), and moreover, the matrix $Q^R_{c,k}$ is supported within $S^R \times S^R$ too.

So we only need to deal with the case $c > 0$, $k = 0$, i.e. $Q^R_{c,0}$. By Definition 7.3, it is supported on $S^R \times S^R$ with expression $Q^R_{c,0}\Big((A,i),(B,j)\Big) =$

$$\sum_{\substack{T_m:|V(T_m)\cup A\cup B|\le\tau\\ A,B\in\mathrm{mSep}_{A,B}(T_m)}} (\frac{\omega}{n})^{|V(T_m)\cup A\cup B| - \frac{|A|+|B|}{2}} \cdot Y_c\big(|R|,\ |V(T_m)\cup A\cup B| + (i+j)\big) \cdot \chi_{T_m} \tag{8.36}$$

where $\left|Y_c\big(|R|,\ |V(T_m)\cup A\cup B| + (i+j)\big)\right| < \tau^{5\tau}$ by Lemma 5.2 (4). For a fixed $(A, B; T_m)$ denote $t = |V(T_m) \cup A \cup B|$, $s = \frac{|A|+|B|}{2} (= |A| = |B|$ in this case), then the coefficient in (8.36) is bounded by $(\frac{\omega}{n})^{t-s} \cdot \tau^{5\tau}$ in absolute value. So we have the support condition, the R-symmetry and coefficient-size conditions as in Lemma 8.3; we proceed exactly the same as there till equation (8.32), where a single term on the RHS now is

$$h(t; s_1, s_2) \cdot (\frac{\omega}{n})^{t-s} \tau^{5\tau} \cdot n^{\frac{t-p}{2}} 2^{O(t)} (\log n)^{O(t-s)}.$$

Note in (8.36) any appearing ribbon $\mathcal{R}_m = (A, B; T_m)$ has $A, B \in \mathrm{mSep}_{A,B}(T_m)$ so $p = s$ (the specialty of $k = 0$). So we can replace the bound on the RHS of (8.34) by $\tau^3 2^{\binom{s}{2}} \cdot n^{-3\epsilon(t-s)} \tau^{5\tau} 2^{O(t)} < 2^{\binom{s}{2}} \tau^{6\tau}$ and then proceed to get $\pm Q^R_{c,0} \prec \tau^{7\tau} \cdot \mathrm{diag}\left(2^{\binom{|A|}{2}}\right)_{S^R\times S^R}$. Now $c \ge 1$, $\omega = n^{\frac{1}{2}-4\epsilon}$ ($\epsilon < 1/40$), $\tau^{7\tau} < n^{1/15}$, so $\pm\omega^{-c} Q^R_{c,0} \prec n^{-c/3} \mathrm{diag}\left(2^{\binom{|A|}{2}}\right)_{S^R\times S^R}$. Once again by $Q^R_{c,0} = D' Q^R_{c,0} D'$, we have $\pm\omega^{-c} Q^R_{c,0} \preceq n^{-c/3} \mathrm{diag}\left(\widetilde{\mathrm{Cl}}\right)_{S^R\times S^R}$. □

Lemma 8.1 follows immediately from Corollary 8.3, Lemma 8.3, 8.4.

8.3. **Last step.** Now we prove the Main Lemma 5.3 thus Theorem 6. For any fixed R, recall $P^R = \{I \in \binom{[n]}{d/2} \mid R \subseteq I\}$, D^τ (Def. 7.2) and S^R (7.2).

Lemma 5.3 recast: W.p. $1 - n^{-5\log n}$ it holds that for all $R \subseteq \binom{[n]}{d/2}$:

$$M_0^R \succeq n^{-d} \cdot \mathrm{diag}(\widetilde{\mathrm{Cl}})_{P^R \times P^R}; \tag{8.37}$$

$$\pm\, \omega^{-c} M_c^R \preceq n^{-c/6} \cdot M_0^R, \quad \forall 0 < c \le |R|. \tag{8.38}$$

The following lemma will be handy.

Lemma 8.5. $\widetilde{L^R} D^\tau \cdot \mathrm{diag}\left(\widetilde{\mathrm{Cl}}\right)_{S^R \times S^R} \cdot D^\tau (\widetilde{L^R})^\top \succeq (\frac{\omega}{n})^{d/2} \mathrm{diag}\left(\widetilde{\mathrm{Cl}}\right)_{P^R \times P^R}$ *for any* $R \in \binom{[n]}{\le d/2}$*, when evaluated on any* G*.*

Proof. Fix any $R \in \binom{[n]}{\le d/2}$. Without confusion, we omit subscript $S^R \times S^R$ by regarding the supports as the vertex-set $[n'] = [n] - R$ and regarding the corresponding matrix indices as $\binom{[n']}{d'/2}$ or $\binom{[n']}{\le d'/2}$, where $d'/2 = d/2 - |R|$. τ is unchanged. We will still use $\widetilde{\mathrm{Cl}}(X)$ to mean $\widetilde{\mathrm{Cl}}(X \sqcup R)$ for $X \subseteq [n']$.

Since $D^\tau \mathrm{diag}(\widetilde{\mathrm{Cl}}) D^\tau$ is nonnegative and diagonal for any G, we have

$$\widetilde{L^R} \left(D^\tau \cdot \mathrm{diag}\left(\widetilde{\mathrm{Cl}}\right) \cdot D^\tau\right) (\widetilde{L^R})^\top \succeq L^{R,0} \left(D^\tau \cdot \mathrm{diag}\left(\widetilde{\mathrm{Cl}}\right) \cdot D^\tau\right) (L^{R,0})^\top, \tag{8.39}$$

where recall $\widetilde{L^R} = (L^{R,0}, ..., L^{R,\tau})$. Further, $L^{R,0} = (L_0^{R,0}, ..., L_{d'/2}^{R,0})$, where $L_t^{R,0}$ is the matrix on column set $\binom{n'}{t}$. This means

$$L_{d/2-|R|}^{R,0} = \left(0, ..., 0, \mathrm{diag}\left(\widetilde{\mathrm{Cl}}\right)_{\binom{[n']}{d'/2} \times \binom{[n']}{d'/2}}\right)$$

since in $L^{R,0}$ (Def. 7.1) only ribbons $\mathcal{R} = (I, A; T')$ with 0-reduced size can occur, forcing $A = I$ and $T' \subseteq E(I)$. In particular, this implies

$$\text{RHS of (8.39)} \succeq (\frac{\omega}{n})^{d/2} \cdot \mathrm{diag}\left(\widetilde{\mathrm{Cl}}\right)_{\binom{[n']}{d'/2} \times \binom{[n']}{d'/2}}.$$

Translated back to $[n]$ and $d/2$, this is exactly the bound in the lemma. □

Proof. (for Lemma 5.3) Fix $R \in \binom{[n]}{\le d/2}$. By Lemma 7.2, for all $c \le |R|$

$$M_c^R = \widetilde{L^R} \cdot \left[D^\tau \left(Q_{c,0}^R - Q_{c,1}^R + ... \pm Q_{c,d}^R\right) D^\tau\right] \cdot \left(\widetilde{L^R}\right)^\top + \mathcal{E}_c^R. \tag{8.40}$$

The following bounds all hold w.p. $> 1 - n^{-8\log n}$ from the corresponding lemmas, and we take union bound so the overall probability is $> 1 - n^{-5\log n}$.

For (8.37). Fix R, we have:

$$\begin{aligned}
M_0^R &= \widetilde{L^R} \cdot \left[D^\tau \left(Q_{0,0}^R - Q_{0,1}^R + ... \pm Q_{0,d}^R\right) D^\tau\right] \cdot \left(\widetilde{L^R}\right)^\top + \mathcal{E}_0^R \\
&\succeq \tau^{-7\tau} \left[\widetilde{L^R} \cdot D^\tau \mathrm{diag}\left(\widetilde{\mathrm{Cl}}\right)_{S^R \times S^R} D^\tau \cdot \left(\widetilde{L^R}\right)^\top\right] + \mathcal{E}_0^R \text{ (Lem. 8.1(1))} \\
&\succeq \tau^{-7\tau} (\frac{\omega}{n})^{d/2} \cdot \mathrm{diag}\left(\widetilde{\mathrm{Cl}}\right)_{P^R \times P^R} + \mathcal{E}_0^R \qquad \text{(Lemma 8.5)} \\
&\succeq (\tau^{-7\tau} (\frac{\omega}{n})^{d/2} - n^{-\epsilon\tau/2}) \cdot \mathrm{diag}\left(\widetilde{\mathrm{Cl}}\right)_{P^R \times P^R} \qquad \text{(Lemma 7.2(4))} \\
&\succeq n^{-d} \cdot \mathrm{diag}(\widetilde{\mathrm{Cl}})_{P^R \times P^R} \qquad \text{(parameter regime)}
\end{aligned}$$

For (8.38). Fix R, $1 \le c \le |R|$, we have:

$$\begin{aligned}
M_c^R &= \widetilde{L^R} \cdot \left[D^\tau \left(Q_{c,0}^R - Q_{c,1}^R + ... \pm Q_{c,d}^R \right) D^\tau \right] \cdot \left(\widetilde{L^R} \right)^\top + \mathcal{E}_c^R \\
&\preceq \omega^c n^{-c/4} \left[\widetilde{L^R} D^\tau \cdot \operatorname{diag}\left(\widetilde{\mathrm{Cl}} \right)_{S^R \times S^R} \cdot D^\tau \left(\widetilde{L^R} \right)^\top \right] + \mathcal{E}_c^R \text{ (Lem. 8.1(2))} \\
&\preceq \omega^c n^{-c/4} \left[\tau^{7\tau} (M_0^R - \mathcal{E}_0^R) \right] + \mathcal{E}_c^R \qquad \text{(Lem. 8.1(1) and (8.40))} \\
&\preceq \omega^c n^{-c/5} M_0^R + \left(\omega^c n^{-c/4} + 1 \right) n^{-\epsilon\tau/2} \operatorname{diag}(\mathrm{Cl})_{P^R \times P^R} \text{ (Lem. 7.2(4))}
\end{aligned}$$

So

$$\begin{aligned}
\omega^{-c} M_c^R &\preceq n^{-c/5} M_0^R + 2n^{-\epsilon\tau/2} \cdot \operatorname{diag}(\mathrm{Cl})_{P^R \times P^R} \\
&\preceq (n^{-c/5} + 2n^d n^{-\epsilon\tau/2}) M_0^R \qquad \text{((8.37) and } \widetilde{\mathrm{Cl}} \ge \mathrm{Cl}) \\
&\preceq n^{-c/6} \cdot M_0^R \qquad (c \le |R| \le d/2 \text{ and parameter regime})
\end{aligned}$$

The same analysis holds for $-\omega^{-c} M_c^R$. □

9. Conclusion

We proved the average-case $\Omega(\epsilon^2 \log n / \log\log n)$ SoS degree lower bound for Exact Clique with clique-size $\omega = n^{1/2-\epsilon}$, which is nearly optimal in both ω, d; we also gave a new perspective on previous techniques in the non-exact case. Two related open problems are as follows.

1. Can we remove the $\log\log n$ factor in d? Perhaps it helps to first find a conceptual explanation of Definition 3.6.

2. Lower bounds of *graph coloring* and *sparse independent set* were recently proved for the soft case [KM21, JPR$^+$21]. Can our technique (or similar ones) help with their exact case?

Acknowledgement

I am very grateful to Aaron Potechin for the introduction of the problem and the encouraging communications, and to Alexander Razborov for the advice and help on improving the quality of the paper. My thanks also go to the anonymous CCC reviewers for their constructive criticism of the presentation.

References

[ABBG11] Sanjeev Arora, Boaz Barak, Markus Brunnermeier, and Rong Ge. Computational complexity and information asymmetry in financial products. *Communications of the ACM*, 54(5):101–107, 2011.

[ABW10] Benny Applebaum, Boaz Barak, and Avi Wigderson. Public-key cryptography from different assumptions. In *Proceedings of the forty-second ACM symposium on Theory of computing*, pages 171–180, 2010.

[AKS98] Noga Alon, Michael Krivelevich, and Benny Sudakov. Finding a large hidden clique in a random graph. *Random Structures & Algorithms*, 13(3-4):457–466, 1998.

[AMP16] Kwangjun Ahn, Dhruv Medarametla, and Aaron Potechin. Graph matrices: norm bounds and applications. *arXiv preprint arXiv:1604.03423*, 2016.

[BBH$^+$12] Boaz Barak, Fernando GSL Brandao, Aram W Harrow, Jonathan Kelner, David Steurer, and Yuan Zhou. Hypercontractivity, sum-of-squares proofs, and their applications. In *Proceedings of the forty-fourth annual ACM symposium on Theory of computing*, pages 307–326, 2012.

[BCK15] Boaz Barak, Siu On Chan, and Pravesh K Kothari. Sum of squares lower bounds from pairwise independence. In *Proceedings of the forty-seventh annual ACM symposium on Theory of computing*, pages 97–106, 2015.

[BGMT12] Siavosh Benabbas, Konstantinos Georgiou, Avner Magen, and Madhur Tulsiani. Sdp gaps from pairwise independence. *Theory of Computing*, 8(1):269–289, 2012.

[BHK+19] Boaz Barak, Samuel Hopkins, Jonathan Kelner, Pravesh K Kothari, Ankur Moitra, and Aaron Potechin. A nearly tight sum-of-squares lower bound for the planted clique problem. *SIAM Journal on Computing*, 48(2):687–735, 2019.

[BIK+96] Paul Beame, Russell Impagliazzo, Jan Krajíček, Toniann Pitassi, and Pavel Pudlák. Lower bounds on hilbert's nullstellensatz and propositional proofs. *Proceedings of the London Mathematical Society*, 3(1):1–26, 1996.

[BR13] Quentin Berthet and Philippe Rigollet. Complexity theoretic lower bounds for sparse principal component detection. In *Conference on Learning Theory*, pages 1046–1066. PMLR, 2013.

[BS14] Boaz Barak and David Steurer. Sum-of-squares proofs and the quest toward optimal algorithms. In *Proceedings of International Congress of Mathematicians (ICM)*, 2014.

[Del73] P Delsarte. An algebraic approach to association schemes of coding theory, phillips j, 1973.

[DM15] Yash Deshpande and Andrea Montanari. Improved sum-of-squares lower bounds for hidden clique and hidden submatrix problems. In *Conference on Learning Theory*, pages 523–562. PMLR, 2015.

[Esc72] Fernando Escalante. Schnittverbände in graphen. In *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, volume 38, pages 199–220. Springer, 1972.

[FK00] Uriel Feige and Robert Krauthgamer. Finding and certifying a large hidden clique in a semirandom graph. *Random Structures & Algorithms*, 16(2):195–208, 2000.

[FK03] Uriel Feige and Robert Krauthgamer. The probable value of the lovász–schrijver relaxations for maximum independent set. *SIAM Journal on Computing*, 32(2):345–370, 2003.

[Gri01] Dima Grigoriev. Linear lower bound on degrees of positivstellensatz calculus proofs for the parity. *Theoretical Computer Science*, 259(1-2):613–622, 2001.

[GV01] Dima Grigoriev and Nicolai Vorobjov. Complexity of null-and positivstellensatz proofs. *Annals of Pure and Applied Logic*, 113(1-3):153–160, 2001.

[HKP15] Samuel B Hopkins, Pravesh K Kothari, and Aaron Potechin. Sos and planted clique: Tight analysis of mpw moments at all degrees and an optimal lower bound at degree four. *arXiv preprint arXiv:1507.05230*, 2015.

[HKP+17] Samuel B Hopkins, Pravesh K Kothari, Aaron Potechin, Prasad Raghavendra, Tselil Schramm, and David Steurer. The power of sum-of-squares for detecting hidden structures. In *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 720–731. IEEE, 2017.

[HKP+18] Samuel B Hopkins, Pravesh Kothari, Aaron Henry Potechin, Prasad Raghavendra, and Tselil Schramm. On the integrality gap of degree-4 sum of squares for planted clique. *ACM Transactions on Algorithms (TALG)*, 14(3):1–31, 2018.

[Jer92] Mark Jerrum. Large cliques elude the metropolis process. *Random Structures & Algorithms*, 3(4):347–359, 1992.

[JPR+21] Chris Jones, Aaron Potechin, Goutham Rajendran, Madhur Tulsiani, and Jeff Xu. Sum-of-squares lower bounds for sparse independent set. *arXiv preprint arXiv:2111.09250*, 2021.

[Kar76] R Karp. Probabilistic analysis of some combinatorial search problems. traub, jf (ed.): Algorithms and complexity: New directions and recent results, 1976.

[KM18] Pravesh K Kothari and Ruta Mehta. Sum-of-squares meets nash: Optimal lower bounds for finding any equilibrium. *arXiv preprint arXiv:1806.09426*, 2018.

[KM21] Pravesh K. Kothari and Peter Manohar. A Stress-Free Sum-Of-Squares Lower Bound for Coloring. In Valentine Kabanets, editor, *36th Computational Complexity Conference (CCC 2021)*, volume 200 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 23:1–23:21, Dagstuhl, Germany, 2021. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[KMOW17] Pravesh K Kothari, Ryuhei Mori, Ryan O'Donnell, and David Witmer. Sum of squares lower bounds for refuting any csp. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, pages 132–145, 2017.

[KOS18] Pravesh Kothari, Ryan O'Donnell, and Tselil Schramm. Sos lower bounds with hard constraints: think global, act local. *arXiv preprint arXiv:1809.01207*, 2018.

[Kuč95] Luděk Kučera. Expected complexity of graph partitioning problems. *Discrete Applied Mathematics*, 57(2-3):193–212, 1995.

[Las01] Jean B Lasserre. Global optimization with polynomials and the problem of moments. *SIAM Journal on optimization*, 11(3):796–817, 2001.

[MPW15] Raghu Meka, Aaron Potechin, and Avi Wigderson. Sum-of-squares lower bounds for planted clique. In *Proceedings of the forty-seventh annual ACM symposium on Theory of computing*, pages 87–96, 2015.

[O'D17] Ryan O'Donnell. Sos is not obviously automatizable, even approximately. In *8th Innovations in Theoretical Computer Science Conference (ITCS 2017)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2017.

[Pan21] Shuo Pang. Sos lower bound for exact planted clique. In *36th Computational Complexity Conference (CCC 2021)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2021.

[Par00] Pablo A Parrilo. *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization*. PhD thesis, California Institute of Technology, 2000.

[PS+00] Pavel A Pevzner, Sing-Hoi Sze, et al. Combinatorial approaches to finding subtle signals in dna sequences. In *ISMB*, volume 8, pages 269–278, 2000.

[RW17] Prasad Raghavendra and Benjamin Weitz. On the bit complexity of sum-of-squares proofs. *arXiv preprint arXiv:1702.05139*, 2017.

[Sch08] Grant Schoenebeck. Linear level lasserre lower bounds for certain k-csps. In *2008 49th Annual IEEE Symposium on Foundations of Computer Science*, pages 593–602. IEEE, 2008.

[Sho87] Naum Z Shor. Class of global minimum bounds of polynomial functions. *Cybernetics*, 23(6):731–734, 1987.

[Tyr94] Evgenij E Tyrtyshnikov. How bad are hankel matrices? *Numerische Mathematik*, 67(2):261–269, 1994.

Appendix A. Deductions in mod-order analysis (section 6.2)

A.1. **Set-up recap.** Ring $\mathbb{A}$ is got by adding fresh variables α and χ_T's to $\mathbb{R}$, where T ranges over edge sets on $[n]$, and they only satisfy the relations $\{\chi_{T'} \cdot \chi_{T''} = \chi_T$ whenever $T' \oplus T'' = T\}$. The **mod-order equation** is

$$L_\alpha \cdot \operatorname{diag}\left(\alpha^{|A|}\right) \cdot (L_\alpha)^\top = M_\alpha \qquad \bmod (*) \tag{A.1}$$

on the $\binom{[n]}{d/2} \times \binom{[n]}{\le d/2}$-matrix variable L_α in ring $\mathbb{A}$, where

$$M_\alpha(I,J) = \sum_{T:|V(T)\cup I\cup J|\le\tau} \alpha^{|V(T)\cup I\cup J|}\chi_T \quad \forall I,J : |I| = |J| = d/2,$$

and mod $(*)$ means to mod the ideal $(\{\alpha^{|V(T)\cup I\cup J|+1}\chi_T\},\ \{\chi_T : |V(T)\cup I\cup J| > \tau\})$ position-wise on each (I,J). We call $(*)$ the **modularity**. Moreover, if denote

$$L_1'(I,A) = \sum_{T'} \beta_{I,A}(T')\chi_{T'}, \quad \beta_{I,A}(T') \in \mathbb{R}[\alpha]$$

then we require

$$\alpha^{e_{I,A}(T')} \mid \beta_{I,A}(T') \quad \forall I, A, T' \tag{A.2}$$

where $e_{I,A}(T')$ is the reduced size $|V(T')\cup I\cup A| - s_{I,A}(T')$ (Def. 4.6).

Expressed in terms, equations (A.1), (A.2) become the following.

$$\sum_{A\in\binom{[n]}{\le d/2}} \sum_{\substack{T',T'':\\ T'\oplus T''=T}} \alpha^{|A|}\cdot\beta_{I,A}(T')\cdot\beta_{J,A}(T'') = \alpha^{|V(T)\cup I\cup J|} \mod \alpha^{|V(T)\cup I\cup J|+1} \tag{A.3}$$

for every $(I,J;T)$ with $|V(T)\cup I\cup J| \le \tau$, and

$$\alpha^{e_{I,A}(T')} \mid \beta_{I,A}(T') \tag{A.4}$$

for every $(I,A;T')$.

The main observation (Lemma 6.2) is the following.

Lemma A.1. *(Order match) In the LHS of equation* (A.3)*, only products* $\alpha^{|A|} \cdot \beta_{I,A}(T') \cdot \beta_{J,A}(T'')$ *that satisfies the following are non-zero modulo* $(*)$*.*

(A.5) A *is a min-separator for both* $(I, A; T')$, $(J, A; T'')$;

(A.6) $(V(T') \cup I \cup A) \ \cap \ (V(T'') \cup J \cup A) = A$.

Moreover, (A.5)*,* (A.6) *imply that*

(A.7) A *is a min-separator of* $(I, J; T)$ *(where* $T = T' \oplus T''$*)*;

(A.8) $|V(T') \cup I \cup A|, \ |V(T'') \cup J \cup A| \leq \tau$.

Proof. Pick a term $\alpha^{|A|} \cdot \beta_{I,A}(T') \cdot \beta_{J,A}(T'')$ form the LHS of (A.3). By (A.4),

its order in $\alpha \geq |A| + |V(T') \cup I \cup A| - s_{I,A}(T') + |V(T'') \cup A \cup J| - s_{J,A}(T'')$.

By modularity on the RHS of (A.3), the term is non-zero only if

$$\text{its order in } \alpha \leq |V(T) \cup I \cup J| \quad \text{and} \quad |V(T) \cup I \cup J| \leq \tau$$

where $T = T' \oplus T''$. This implies

(A.9)
$$|V(T') \cup I \cup A| + |V(T'') \cup J \cup A| \leq \underbrace{|V(T) \cup I \cup J|}_{①} + \underbrace{(s_{I,A}(T') + s_{J,A}(T'') - |A|)}_{②}$$

Note ② $\leq |A|$ and "=" holds iff $s_{I,A}(T') = s_{J,A}(T'') = |A|$. While the LHS above

$$= \underbrace{|(V(T') \cup I \cup A) \cup (V(T'') \cup J \cup A)|}_{\geq |V(T) \cup I \cup J| = ①} + \underbrace{|(V(T') \cup I \cup A) \cap (V(T'') \cup J \cup A)|}_{\geq |A| \geq ②}.$$

Therefore, (A.9) could hold only when all "="'s hold, which means: (1). A is a min-separator of $(I, A; T')$, $(J, A; T'')$; (2). $(V(T') \cup I \cup A) \cup (V(T'') \cup J \cup A) = V(T) \cup I \cup J$; (3). $(V(T') \cup I \cup A) \cap (V(T'') \cup J \cup A) = A$.

Next, we show (1),(3) imply $A \in \text{mSep}_{I,J}(T)$ (and also (2), actually). By (3), T', T'' could overlap only in $E(A)$. Now $T = T' \oplus T''$, so

(A.10) $$T = T' \sqcup T'' \quad \text{modulo } E(A)$$

(also $\Rightarrow V(T') \cup V(T'') \subseteq V(T) \cup A$). By (1) there are $|A|$ many vertex-disjoint paths $p_1, , , .p_{|A|}$ from I to A in T', and similarly $q_1, ..., q_{|A|}$ from J to A in T''. These paths are also present in T by (A.10)—where it naturally assumes every path touches A only once at its endpoint. By (3) again, any p_i, q_j do not intersect beside endpoint in A so they are paired to $|A|$ many vertex-disjoint paths from I to J in T, all passing A (this also implies $A \subseteq V(T) \cup I \cup J$). On the other hand, if p is a path in T from I not passing A, then it is a path on $I \cup V(T')$ by induction using (3). Now by (3) again we have $(V(T') \cup I) \cap J \subseteq A$, so p can't reach J. So $A \in \text{mSep}_{I,J}(T)$.

Finally, under the above implications, $V(T') \cup I \cup A \subseteq V(T) \cup I \cup J$ and similarly for $V(T'') \cup J \cup A$, so both have size $\leq \tau$. □

By this lemma, we can assume that in an imagined solution, $\beta_{I,A}(T') \neq 0$ only when it satisfies the conditions (A.5), (A.8) on its part. If assume further that the solution is *symmetric* (which looks plausible), i.e. $\beta_{I,A}(T') = \beta_{J,B}(T'')$ whenever $(I, A; T')$, $(J, B; T'')$ are of the same shape, then this lemma is particularly informative about some special $(I, J; T)$'s.

Corollary A.1. *If* $(I, J; T)$ *has a* **unique** *min-separator* A*, then*

(A.11) $$\sum_{\substack{T', T'': \ T' \oplus T'' = T \\ \text{(A.5), (A.6) } hold}} \beta_{I,A}(T') \cdot \beta_{J,A}(T'') = \alpha^{e_{I,J}(T)}$$

where $e_{I,J}(T) = |V(T) \cup I \cup J| - s_{I,J}(T)$. *In particular, in symmetric solution,*

$$\sum_{T_1 \subseteq E(A)} \beta_{I,A}(T_1 \oplus T')^2 = \alpha^{2 \cdot e_{I,A}(T')} \tag{A.12}$$

for all $(I, A; T')$ *such that*

$$A \text{ is the unique min-separator of } (I, A; T'). \tag{A.13}$$

Proof. The first part is directly from Lemma 6.2. For the "in particular" part, let $(I, A; T')$ satisfy (A.13). By mirroring $(I, A; T')$ through A, we get a $(J, A; T'')$ that satisfies the same condition and they together satisfy (A.5), (A.6). There are always enough vertices in $[n]$ to carry out this mirroring operation. By the symmetry assumption, $\beta_{I,A}(T') = \beta_{J,A}(T'')$. From mirroring it is not hard to see that A is the unique min-separator of $(I, J; T = T' \oplus T'')$, so for this triple $(I, J; T)$ equation (A.11) holds, giving that $\sum_{T_1 \subseteq E(A)} \beta_{I,A}(T' \oplus T_1)^2 = \alpha^{|V(T) \cup I \cup J| - |A|} = \alpha^{2(|V(T') \cup I \cup A| - |A|)}$. □

We can summarize what we got as follows. If let all $\beta_{I,A}(T' \oplus T_1)$'s in equation (A.12) be equal (which is a plausible assumption), then $\beta_{I,A}(T') = 2^{-\binom{|A|}{2}/2} \cdot \alpha^{e_{I,A}(T')}$ (take all + signs). Collecting these terms, we get the following matrix

$$L_1' : \quad L_1'(I, A) = \sum_{\substack{T' :\ |V(T') \cup I \cup A| \le \tau \\ \text{(A.13) holds} \\ T' \cap E(A) = \emptyset}} 2^{-\binom{|A|}{2}/2} \cdot \alpha^{|V(T') \cup I \cup A| - |A|} \chi_{T'} \cdot \widetilde{\mathrm{Cl}}_A$$

where $\widetilde{\mathrm{Cl}}_A = \sum_{T \subseteq E(A)} \chi_T$. To see how far this is from a solution, notice $\widetilde{\mathrm{Cl}}_A^2 = 2^{\binom{|A|}{2}} \widetilde{\mathrm{Cl}}_A$ and consider

$$L_1' \cdot \mathrm{diag}\left(\alpha^{|A|}\right) \cdot (L_1')^\top = L_1 \cdot \mathrm{diag}\left(\alpha^{|A|} \cdot \widetilde{\mathrm{Cl}}_A\right) \cdot L_1^\top \tag{A.14}$$

where L_1 is the matrix in $\mathbb{A}$ as below (which is cleaner than L_1' to use).

Definition A.1. $\forall I \in \binom{[n]}{d/2}$, $A \in \binom{[n]}{\le d/2}$,

$$L_1(I, A) := \sum_{\substack{T' :\ |V(T') \cup I \cup A| \le \tau \\ \text{(A.13) holds} \\ T' \cap E(A) = \emptyset}} \alpha^{|V(T') \cup I \cup A| - |A|} \chi_{T'}. \tag{A.15}$$

Surely L_1' is not a solution to the mod-order equation, since (A.14) equals (mod (*)) only the part of M_α consisting of the special $(I, J; T)$'s from Corollary A.1. For a general $(I, J; T)$, Lemma A.1 only says:

$$\sum_{\substack{A, T', T'' :\ T' \oplus T'' = T \\ A \in \mathrm{mSep}_{I,J}(T) \\ \text{(A.5),(A.6) hold}}} \beta_{I,A}(T') \beta_{J,A}(T'') = \alpha^{e_{I,J}(T)} \mod \alpha^{e_{I,J}(T)+1}. \tag{A.16}$$

To see how to proceed further, we inspect a further weakening: polarization.

A.2. **Polarized solution.** Roughly speaking, polarization weakens linear equations about "x_i^2's" by replacing these terms with multi-linear "$x_i y_i$'s", where $\vec{y}$ are fresh variables. Then we can plug in any "tentative" solution $\vec{x_0}$ to solve for $\vec{y}$ more easily (as the equations are linear in $\vec{y}$), and see how to modify $\vec{x_0}$ further.

Definition A.2. *The* **polarized** *mod-order equation w.r.t.* L_1 *is:*

$$L_1 \cdot \mathrm{diag}\left(\alpha^{|A|} \cdot \widetilde{\mathrm{Cl}}_A\right) \cdot L_2^\top = M_\alpha \qquad \mod (*) \tag{A.17}$$

where $(*)$ *is the modularity in* (A.1)*,* L_1 *is by* (A.15)*,* L_2 *is the variable matrix*

$$L_2(I,A) = \sum_{T'\colon\ |V(T')\cup I\cup A|\le\tau} \beta^{(2)}_{I,A}(T')\chi_{T'} \tag{A.18}$$

satisfying $\alpha^{e_{I,A}(T')} \mid \beta^{(2)}_{I,A}(T')$ *for all* (I,A,T')*.*

In this polarized form, the essential condition (A.16) becomes

$$\sum_{\substack{A,T',T''\colon\ T'\oplus T''=T\\ (I,A;T') \text{ appears in } L_1\\ \text{(A.5),(A.6) hold}}} \alpha^{e_{I,A}(T')}\cdot\beta^{(2)}_{J,A}(T'') = \alpha^{e_{I,J}(T)} \mod \alpha^{e_{I,J}(T)+1}. \tag{A.19}$$

By (A.19), existence of a solution L_2 at least requires the following condition: for general $(I,J;T)$, there always exist "$(I,A;T')$ appearing in L_1" and T'' which satisfy the condition in the LHS of (A.19). By a direct (but careful) check, this condition is actually **equivalent to** the "In particular" part of the graph-theoretic fact 5 due to Escalante, restated below.

Fact A.1. *For any ribbon* $(I,J;T)$*, the set of all min-separators,* $\mathrm{mSep}_{I,J}(T)$*, has a natural poset structure: min-separators* $A_1 \le A_2$ *iff* A_1 *separates* $(I,A_2;T)$*, or equivalently as can be checked, iff* A_2 *separates* $(J,A_1;T)$*. The set is further a* **lattice** *under this partial-ordering:* $\forall A_1,A_2\in\mathrm{mSep}_{I,J}(T)$ *their join and meet exist. In particular, there exist a unique* **minimum** *and* **maximum***.*

Denote the minimum by $S_l(I,J;T)$ *and the maximum by* $S_r(I,J;T)$*, which is the "leftmost" and "rightmost" min-separator, respectively.*

By this fact, some $(I,A;T')$ indeed appears in (A.19) with $A = S_l(I,J;T)$. Moreover, (A.19) is naturally satisfied if take

$$L_2(J,A) = \sum_{\substack{T''\colon\ |V(T'')\cup J\cup A|\le\tau\\ A\in\mathrm{mSep}_{J,A}(T'')\\ T''\cap E(A)=\emptyset\\ (J,A;T'') \text{ left-generated}}} \alpha^{e_{J,A}(T'')}\chi_{T''}. \tag{A.20}$$

Here, recall being left-generated means every vertex is either in A or can be connected from J without touching A. Also, with this L_2 only one product in the LHS of (A.19) contributes to the right modulo $\alpha^{e_{I,J}(T)+1}$. We get:

Proposition A.1. *The pair* (L_1,L_2) *is a solution to the polarized mod-order equation* (A.17)*,* (A.18)*.*

Remove the polarization. One more use of fact A.1 actually shows that, if move the "left-generated" condition from L_2 to L_1, then L_2 itself *effectively* factors through L_1, i.e. we can replace $\mathrm{diag}(\widetilde{\mathrm{Cl}})\cdot L_2^\top$ by some $X\cdot L_1^\top$ in (A.17). This is the idea behind the following proposition (Prop. 6.2 recast).

Proposition A.2. *(Mod-order diagonalization) Let*

$$L_\alpha(I,A) := \sum_{\substack{T'\colon\ |V(T')\cup I\cup A|\le\tau\\ A=S_l(I,A;T')\\ T'\cap E(A)=\emptyset\\ (I,A;T') \text{ left-generated}}} \alpha^{e_{I,A}(T')}\chi_{T'},$$

$$Q_{0,\alpha}(A,B) := \sum_{\substack{T_m\colon\ |T\cup A\cup B|\le\tau\\ A,B\in\mathrm{mSep}_{A,B}(T_m)}} \alpha^{e_{A,B}(T_m)}\chi_{T_m}$$

(where T_m indicates "middle"). Then

$$L_\alpha \cdot [\mathrm{diag}\left(\alpha^{\frac{|A|}{2}}\right) \cdot Q_{0,\alpha} \cdot \mathrm{diag}\left(\alpha^{\frac{|A|}{2}}\right)] \cdot L_\alpha^\top = M_\alpha \qquad \mathrm{mod}\ (*) \tag{A.21}$$

where $(*)$ *is the modularity in* (A.1).

Proof. Given Fact A.1, we immediately have the *canonical decomposition* of graphs as in Definition 6.3 and Remark 6.2. This implies that in the LHS of (A.21) only the products from canonical triples are non-zero modulo $(*)$, and they give M_α. □

Thus we get a "$L(-)L^\top$"-shape decomposition of M_α, meaning that we do not lose much from the polarization step since our goal is only to prove the PSDness of the matrix. Indeed, (A.21) gives the "first-approximate" decomposition in Definition 6.2.