SOS Lower Bound for Exact Planted Clique

Shuo Pang ✉ 🏠
Mathematics Department, University of Chicago, IL, USA

Abstract

We prove a SOS degree lower bound for the planted clique problem on the Erdös-Rényi random graph $G(n, 1/2)$. The bound we get is degree $d = \Omega(\epsilon^2 \log n/\log\log n)$ for clique size $\omega = n^{1/2-\epsilon}$, which is almost tight. This improves the result of [5] for the "soft" version of the problem, where the family of the equality-axioms generated by $x_1 + ... + x_n = \omega$ is relaxed to one inequality $x_1 + ... + x_n \geq \omega$.

As a technical by-product, we also "naturalize" certain techniques that were developed and used for the relaxed problem. This includes a new way to define the pseudo-expectation, and a more robust method to solve out the coarse diagonalization of the moment matrix.

2012 ACM Subject Classification Theory of computation → Proof complexity; Mathematics of computing → Random graphs; Theory of computation → Semidefinite programming

Keywords and phrases Sum-of-Squares, planted clique, random graphs, average-case lower bound

Digital Object Identifier 10.4230/LIPIcs.CCC.2021.26

Related Version *ECCC Technical Report*: `https://eccc.weizmann.ac.il/report/2021/070`

Acknowledgements 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 reviewers for their constructive criticism of the presentation.

1 Introduction

1.1 The problem and the proof system

Whether one can find a max-clique in a random graph $G \sim G(n, 1/2)$ efficiently and be correct with high probability has been a long-standing open problem in computational complexity since [19]. In [18, 22], a relaxed formulation as the *planted clique problem* was introduced: if we further plant a random clique of size $\omega \gg \log n$ to G, can it be efficiently recovered? Information-theoretically this is possible, since w.h.p. the largest clique in G has size $(2 + o(1)) \log n$. While computationally, the average-case hardness of this problem is still widely believed even after it has been intensively studied and has inspired research directions in an extremely wide range of fields (just to mention a few: cryptography [2], learning theory [8], mathematical finance [3], computational biology [28]). So far, the best known polynomial-time algorithm is for $\omega = \Omega(\sqrt{n})$ [1], which is a so-called spectral algorithm (see e.g. [17]).

The sum-of-squares (SOS) hierarchy [30, 27, 23] 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 SDP (see e.g. [4, 17]). Recent years have witnessed rapid development on SOS-based algorithms which turn out to provide a characterization of a wide class of algorithmic techniques – for a list of evidence, we refer the reader to the survey [6] and the introduction of [17]. The SOS proof system is the natural proof-theoretic counterpart of these algorithms, also known as the *Positivstellensatz* system [14]: 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

$$-1 = \sum_{i=1}^{k} f_i q_i + \sum_j r_j^2 \quad \text{in } \mathbb{R}[x_1, ..., x_n]$$

where $q_1, ..., q_m$ and $r_1, ...$ 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 always exists if the axioms are contradictory. The *degree-d SOS proof system* is this plus a degree limitation

$$\max_{i,j}\{\deg(f_i) + \deg(q_i),\ 2\deg(r_j)\} \le d.$$

For more about the relation between the SOS proofs and SDP algorithms, see e.g. [26, 29]. The average-case hardness of the 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.? Note the system cannot just say "No" but must search for a certificate – a proof. A lower bound here 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 [1] corresponds to a degree-2 SOS proof, a SOS degree lower bound would bring us a much better understanding of the hardness of the problem. The standard 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****.*

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

To confirm no ω-clique exists is to give a SOS refutation of the above. The SOS system has the so-called duality: to show degree lower bound it suffices to consider *pseudo-expectation* and the resulting *moment matrix*[1]. With boolean variables (which is our case), this can be demonstrated on multi-linear polynomials. Let $\mathcal{X}^{\le a} = \{x_S \mid S \subseteq [n],\ |S| \le a\}$ for any a.

▶ **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* $\mathbb{R}$*-linearity.*

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

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

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

where in (1.4)*,* $x_A \cdot x_B := x_{A\cup B}$*. For 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 expression* $M(A, B) = \widetilde{E}x_{A\cup B}$ *for all* $|A|, |B| \le d/2$*, then:*

$$\text{(PSDness)} \qquad M \textit{ is positive semi-definite.} \tag{1.5}$$

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

[1] We use the name for simplicity. More cautiously, it should be called the *pseudo-moment matrix*.

[2] d is always assumed to be even.

A relaxation of the problem was studied in [5]: decide whether there exists $\widetilde{E}$ as in Definition 1.2 except by one change – replace Size Constraints by one weaker inequality $\widetilde{E}(x_1 + \ldots + x_n) \geq \omega$. Henceforth, we call the Clique Problem (Def. 1.1) **Exact Clique** and this relaxation **Non-Exact Clique**.[3] We will study their average-case hardness over $G \sim G(n, 1/2)$.

How to deal with the exact problem is a subtle but important open problem. On the problem itself, lower bounds on the "weak" formulation indeed gave the important algorithmic message – an integrality gap for many SOS-based optimization algorithms – but still, they do not rule out the possibility that SOS can efficiently refute $x_1 + \ldots + x_n = k$ for each individual large k, and 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 limit of existing methods for proving average-case SOS lower bounds. Current techniques from the so-called *pseudo-calibration heuristic* [5] tend to deal successfully with "soft" constraints (i.e. inequalities, or usually just one bound on a single pseudo-expectation value) while being poor at handling "hard" constraints (i.e. equalities). Finding techniques to deal with the latter is thus in need. Progress toward this goal is made in [20] for random CSPs, where the number of hard constraints is at most two[4]. Their method is to break such constraint(s) into local ones and satisfy each using real, independent distributions. For "inherently more rigid" problems like Exact Clique (whose hard constraints are "almost everywhere"), however, it seems unlikely a similar strategy could work.

Lastly, there are concrete applications of lower bounds on Exact Clique. Such a lower bound can give by reduction lower bounds for other problems, e.g. for the approximated Nash-Welfare, and potentially for the coloring problem and stochastic block models [20, 21].

1.2 Previous work

For upper bounds, if $\omega = \Omega(\sqrt{n})$ then degree-2 SOS can refute Exact Clique with high probability [12]. 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; since we have not been able to find it in the literature, we include it as Observation 1.3 below.

For lower bounds, for Exact Clique, [13] showed that the weaker system *d-round Lovasz-Schrijver* cannot refute it when $\omega = O(\sqrt{n/2^d})$; [25] proved degree-d lower bound on SOS for $\omega = \widetilde{O}(n^{1/d})$, and this bound on ω was improved to $\widetilde{O}(n^{1/3})$ for $d = 4$ [10] and $\widetilde{O}(n^{\frac{1}{\lfloor d/2 \rfloor + 1}})$ for general d [15]. For Non-Exact Clique, [5] proved the almost tight lower bound $d = \Omega(\epsilon^2 \log n)$ for $\omega = n^{1/2-\epsilon}$, $\epsilon > 0$ arbitrary (could depend on n).

▶ **Observation 1.3** (Upper bound for Exact Clique if $\omega > d = 2.1 \log n$). *Note $(x_1 + \ldots + x_n)^d = \omega^d$ 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, as 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 this proof is actually in the weaker Nullstellensatz system (for definition see e.g. [7]).*

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

[4] One on the objective value of the CSP, and/or one on the Hamming weight of x.

1.3 Results of the paper

Our main result is the following.

▶ **Theorem 1.4.** *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 1.5** (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

The two results use almost completely different ideas, so we treat them separately in the proof overview:

- Theorem 1.4: Section 2.1 to 2.4.
- Theorem 1.5: Section 2.5.

The presentation of this section is structured for mathematical clarity. On the other hand, the following picture may provide a clearer bird's-eye view, where "$\cdots$" means the corresponding section(s) in the text:

Pseudo-expectation design: A common idea (in below)
→ Non-exact case (2.5 first half $\cdots$ 3.1)
→ Exact case (2.1 $\cdots$ 3.2).
Proving PSDness: Recursive factorization refresh (5.1, 5.3)
→ Lower bound proof (2.1 to 2.4 $\cdots$ 6).

And a "naturalizing" result that can be read independently:

How to deduce the "coarse" diagonalization (2.5 second half $\cdots$ 5.2).

Let's start with a common idea. Suppose we deal with degree-d SOS, $\omega = n^{1/2-\epsilon}$ where $\epsilon > 0$ is small. To construct pseudo-expectations on size $\le d$-subsets of $[n]$, as is usual in complexity theory, we take a parameter $\tau \gg d$ (think of $d \ll \tau \ll \log n$) and make the construction for all size $\le \tau$-subsets first, in hope to later have a good control on its behavior on size $\le d$ subsets. This idea is most clearly demonstrated in the nonexact case (Section 3.1.2) and is also inherited to the exact case, as we will see next (equation (2.1)).

2.1 The exact pseudo-expectation

We define the pseudo-expectation for Exact Clique now. To satisfy Size Constraints (1.4), a natural way is to generate $\widetilde{E}$ in a top-down fashion: fix $\widetilde{E}x_S$ for all $|S| = d$ first, denoted as the vector $\widetilde{E}_d x$, then recursively set

$$\widetilde{E}x_S \leftarrow \frac{1}{\omega - |S|} \sum_{i \notin S} \widetilde{E}x_{S\cup\{i\}} \quad \forall |S| < d.$$

The Clique Constraints (1.3) can be satisfied if $\widetilde{E}_d x_S$ factors through the clique function on S. Inspired by the non-exact case (Lemma 3.5), we use Fourier characters and consider

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

for some function F. We call F a ***d*-generating function**.[5] Thus

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

where $u := d - |S|$, for all S with $|S| \le d$.

One key novelty we bring is the following choice

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

With this F, the resulting moment matrix, denoted by $\widetilde{M}$, is:

$$\widetilde{M}(A,B) = \sum_{T:|V(T)\cup A\cup B|\le\tau} \widetilde{M}(A,B;T)\chi_T \quad \forall A, B:\ |A|, |B| \le d/2$$

where $\widetilde{M}(A,B;T) =$

$$\frac{1}{\binom{\omega-d+u}{u}}\left[\sum_{c=0}^{u}\binom{|V(T)\cup A\cup B| - (d-u)}{c}\binom{n-|V(T)\cup A\cup B|}{u-c}\right.$$
$$\left.\cdot \underbrace{\frac{(|V(T)\cup A\cup B| + u - c + 8\tau^2)!}{(8\tau^2)!}\cdot(\frac{\omega}{n})^{|(V(T)\cup A\cup B)|+u-c}}_{F\left(|V(T)\cup A\cup B|+u-c\right)}\right], \tag{2.3}$$

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

This seemingly mysterious choice of F is ultimately for proving PSDness of $\widetilde{M}$, which can be seen after a series of technical transformations (Remark 2.10, 3.12). It will be very interesting to know if there is *a priori* an explanation of it. See Remark 3.9, 6.14 for why the simpler, "traditional" choices from the literature, which simulate some plant-distributions, seem cannot work here.

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

2.2 An Hadamard decomposition and Euler transform

For the Exact Clique problem, by 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 $\widetilde{M}$. We 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 make a PSDness analysis on M (in particular, get a clue of how to diagonalize it), we resolve this intricacy by two steps. First,

$$M = \sum_{c=0}^{\frac{d}{2}} m_c \circ M_c \tag{2.4}$$

where m_c, M_c are matrices s.t. for all $|I|, |J| = d/2$,

$$m_c(I,J) = \frac{1}{\binom{\omega-d+u}{u}} \omega^{u-c} \quad \text{where } u \text{ denotes } |I \cap J|; \tag{2.5}$$

$$M_c(I,J) = \begin{cases} \sum\limits_{T:|V(T)\cup I \cup J| \leq \tau} \chi_T \cdot M_c(|I \cap J|, |V(T) \cup I \cup J|), & \text{if } |I \cap J| \geq c; \\ 0, & \text{o.w.} \end{cases} \tag{2.6}$$

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 $u = |I \cap J|$, $a = |V(T) \cup I \cup J|$. We will analyze m_c, M_c's separately.

The "harder" part is M_c. To further remove the dependence on $|I \cap J|$ in $M_c(I,J)$, our second step is to consider a decomposition

$$M_c = \sum_{R \in \binom{[n]}{\leq \frac{d}{2}}} M_c^R \tag{2.7}$$

where for each R the matrix M_c^R is supported on rows and columns whose index contains R. To derive the expression of M_c^R, we use **Euler transform**: if $x(\cdot), y(\cdot)$ are two sequences defined on $\mathbb{N}$ s.t. $x(m) = \sum_{l=0}^m \binom{m}{l} y(l)$ for all m, then $x(\cdot)$ is called the *Euler transform* of $y(\cdot)$, and the inverse transform is given by $y(m) = \sum_{l=0}^m (-1)^{m-l}\binom{m}{l} x(l)$.

Apply the inverse Euler transform to $M_c(u,a)$ in the above[6] on u (fixing c, a), we get:

$$Y_c(r,a) = \begin{cases} \sum\limits_{l=c}^{r} (-1)^{r-l}\binom{r}{l}\binom{a+l-d}{c}\binom{n-a}{l-c} n^{-(l-c)} \frac{(a+l-c+8\tau^2)!}{(8\tau^2)!} & \text{, if } r \geq c; \\ 0 & \text{, o.w.} \end{cases} \tag{2.8}$$

In summary, the following lemma can be proved.

▶ **Lemma 2.1** ($\Sigma\Pi$-decomposition of M).

$$M = \sum_{c=0}^{\frac{d}{2}} m_c \circ \left(\sum_{R \in \binom{[n]}{\leq d/2}} M_c^R \right) = \sum_{R \in \binom{[n]}{\leq d/2}} \left(\sum_{c=0}^{|R|} m_c \circ M_c^R \right) \tag{2.9}$$

where each m_c is by (2.5)*, and each M_c^R has the following expression.*

[6] A subtle but important point is that $M_c(u,a)$ is partial (i.e. defined only when $u \geq c,\ a-(d-u) \geq c$), and we need to extend it to $(u,a) \in \mathbb{N}^2$ – see Def. 6.5.

1. $M_c^R = 0$ *if* $|R| < c$;
2. *If* $R \not\subseteq I \cap J$, $M_c^R(I,J) = 0$;
3. *If* $|R| \geq c$ *and* $R \subseteq I \cap J$, *then*

$$M_c^R(I,J) = \sum_{T:|V(T)\cup I\cup J|\leq\tau} M_c^R(I,J;T)\chi_T$$

where, if denote $a = |V(T) \cup I \cup J|$,

$$M_c^R(I,J;T) = (\frac{\omega}{n})^a \cdot Y_c(|R|,a) \quad (\textit{defined by } (2.8)). \tag{2.10}$$

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

Intuition for analysis. The intuition behind decomposition (2.9) is that, the first factor m_c is decreases in c and m_0 is very positive; while for every 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* $c = 0, ..., d/2$,

$$m_0 = \omega m_1 = ... = \omega^{\frac{d}{2}} m_{\frac{d}{2}} \succ \frac{d}{2\omega}\mathrm{Id}. \tag{2.11}$$

▶ **Lemma 2.3** (Main Lemma). *In decomposition* (2.9), *w.p.* $> 1 - n^{-5\log n}$ *the following hold. For all* $R \in \binom{[n]}{\leq d/2}$, *let* $P^R = \{I \in \binom{[n]}{d/2} \mid R \subseteq I\}$, *the following holds.*

(1)

$$M_0^R \succeq n^{-d}\mathrm{diag}(\widetilde{\mathrm{Cl}})_{P^R\times P^R}; \tag{2.12}$$

(2)

$$\pm\omega^{-c}M_c^R \preceq n^{-c/6} \cdot M_0^R, \quad \forall 0 < c \leq |R|. \tag{2.13}$$

These two lemmas immediately imply that $M(G) \succeq n^{-d-1}\mathrm{diag}(\widetilde{\mathrm{Cl}}(G))_{\binom{[n]}{d/2}\times\binom{[n]}{d/2}}$ w.h.p., and Theorem 1.4 is an easy corollary of this (Cor. 6.2, 6.10).

The proof of Lemma 2.2 is relatively easy using *Johnson schemes* (see Lemma 6.4). Below we show how to prove the Main Lemma.

2.3 Recursive factorization: an extension

Fix any c, R ($|R| \geq c$). To prove the Main Lemma, an important step is to derive an approximate diagonalization of M_c^R, where we will use the *recursive factorization* technique from [5]. This technique will be refreshed, formalized and extended properly for our use in Section 5.3.

For now, we give a **first-approximate factorization** of M_c^R then apply this technique to get a refined diagonalization by Lemma 2.6.

The next definition in full (Definition 6.11) mentions many terms about a graph-theoretic structure; we omit the details here.

▶ **Definition 2.4** (Side factors). *Fix* $R \in \binom{[n]}{\leq\frac{d}{2}}$. *For* $i = 0, 1, ..., \tau$ *let* $L^{R,i}$ *be the matrix of dimension* $\binom{[n]}{\frac{d}{2}} \times \binom{[n]}{\leq\frac{d}{2}}$ *defined by equation* (6.20) *(the exact content is not important for now). Call* $\widetilde{L^R} = (L^{R,0}, ..., L^{R,\tau})$ *the* ***left factor***, *and* $(\widetilde{L^R})^\top$ *the* ***right factor***.

We use these factors to give a PSD factorization in the form $M^R = \widetilde{L^R}\,(-)\left(\widetilde{L^R}\right)^\top$. The starting point is a coarse, "first approximate" factorization. In the definition below, T_m simply means an edge-set and $\mathrm{mSep}_{A,B}(T_m)$ is the set of all *minimal separators* of vertex-sets A, B (Def. 4.10). Let D^τ be the diagonal matrix $\operatorname{diag}\left((\frac{\omega}{n})^{\frac{|A|}{2}}\right) \otimes \mathrm{Id}_{\{0,...,\tau\}\times\{0,...,\tau\}}$.

▶ **Definition 2.5.** *For any $R \in \binom{[n]}{\le d/2}$ define the index set*

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

For $c = 0, ..., |R|$, define $Q^R_{c,0}$ to be the $\{0,...,\tau\} \times \{0,...,\tau\}$-blocked matrix, each block of dimension $\binom{[n]}{\le d/2} \times \binom{[n]}{\le d/2}$: it is supported on $S^R \times S^R$, expressed by $Q^R_{c,0}\Big((A,i),(B,j)\Big) =$

$$\sum_{\substack{T_m:|V(T_m)\cup A\cup B|\le\tau\\ A,B\in \mathrm{mSep}_{A,B}(T_m)}} (\frac{\omega}{n})^{|V(T_m)\cup A\cup B|-\frac{|A|+|B|}{2}} \cdot \underbrace{Y_c\big(|R|,\ |V(T_m)\cup A\cup B| + (i+j)\big)}_{\text{defined by (2.8)}} \cdot \chi_{T_m} \quad (2.14)$$

*Then we call $\widetilde{L^R} \cdot \left(D^\tau \cdot Q^R_{c,0} \cdot D^\tau\right) \cdot \left(\widetilde{L^R}\right)^\top$ the **first approximate factorization** of M^R_c.*

▶ **Lemma 2.6** (Recursive approximate factorization; informal). *For any $R \in \binom{[n]}{\le d/2}$ and $0 \le c \le |R|$, we have the following decomposition.*

$$M^R_c = \widetilde{L^R} \cdot \left[D^\tau\Big(Q^R_{c,0} - Q^R_{c,1} + ... \pm Q^R_{c,d}\Big)D^\tau\right] \cdot \left(\widetilde{L^R}\right)^\top + \mathcal{E}^R_c. \quad (2.15)$$

Here, all $Q^R_{c,k}$'s ($k = 0, 1, ..., d$) are supported within $S^R \times S^R$ with expression

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

where $\mathcal{R}_m$ denotes the triple $(A, B; T_m)$, and the coefficients $q^R_{c,k}(\cdot, i, j)$'s depend only on the "shape" of $\mathcal{R}_m$, satisfying

$$|q^R_{c,k}(\mathcal{R}_m, i, j)| \le \tau^{5\tau} \cdot (\frac{\omega}{n^{1-\epsilon}})^{s-p+k/3} \quad \forall (i,j) \quad (2.16)$$

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

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

▶ Remark 2.7. In this factorization, the middle matrices Q's have a "tensored-dimension" with $(\tau + 1)$, i.e. it is a $(\tau+1) \times (\tau+1)$-blocked matrix, each block of dimension $\binom{[n]}{\le d/2} \times \binom{[n]}{\le d/2}$. This reflects a key difference (at least technically) between Exact Clique and the non-exact case; see Remark 6.14.

2.4 Proving PSDness: encounter with Hankel matrices

With Lemma 2.2 and the recursive factorization lemma 2.6 at hand, the following is the key step towards the Main Lemma.

▶ **Lemma 2.8.** *W.p.* $> 1 - n^{-8\log n}$ *over* G*, the following holds.*

(1) $\forall R \in \binom{[n]}{\leq d/2}$,

$$Q^R_{0,0} - Q^R_{0,1} + \ldots \pm Q^R_{0,\frac{d}{2}} \ \succeq \ \tau^{-7\tau} \cdot \operatorname{diag}\left(\widetilde{\mathrm{Cl}}\right)_{S^R \times S^R}$$

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

(2) $\forall R$, $0 < c \leq |R|$

$$\pm\omega^{-c}\left(Q^R_{c,0} - Q^R_{c,1} + \ldots \pm Q^R_{c,\frac{d}{2}}\right) \ \preceq \ n^{-c/4} \cdot \operatorname{diag}\left(\widetilde{\mathrm{Cl}}\right)_{S^R \times S^R}.$$

To prove this lemma, modulo somewhat standard steps (three Lemmas 6.34, 6.37, 6.38) the final technical challenge is:

Show the positiveness of $\mathbb{E}[Q^R_{0,0}]$ (Corollary 6.30).

We describe below how this is done. After simplification, the real task is to analyze the positiveness of the following matrix[7]:

$$\sum_{l=0}^{r} (-1)^{r-l} \frac{\binom{r}{l}}{l!} \cdot H_{\tau,\ l+8\tau^2} \quad \text{for any } 0 \leq r \leq d/2 \tag{2.17}$$

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

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

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 they are notoriously wild-behaving in many aspects (see e.g. [31]). 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 6.27), resulting in the following property.

▶ **Proposition 2.9.** *If parameters* m, t, r *satisfy*

$$t + 1 > 8 \cdot \max\{r^2, m\}, \tag{2.18}$$

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

▶ Remark 2.10. The condition (2.18) in the above proposition is the reason of the "$8\tau^2$" in the numerator of F, (2.2).

With this proposition, it is relatively easy to complete the proof of the Lemma 2.8, hence the Main Lemma. This completes the proof overview of Theorem 1.4.

2.5 Ideas for Theorem 1.5

In this subsection, we demonstrate how to "naturalize" certain techniques that were used for the lower bounds of Non-Exact Clique.

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

On defining the pseudo-expectation. (Section 3.1) Previously, the pseudo-expectation is obtained via the so-called *pseudo-calibration* method. We show how to define the same $\widetilde{E}$ in very different terms via the incidence algebra on the vertex-set, which can also be regarded as a simple refinement of the construction in [13].

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 in the case of a single planted clique, and we use ζ to define $\widetilde{E}$. That is, we define a *degree-*τ approximate-distribution vector $p_\tau(G)$ first – it approximates the real planted-clique distribution, with a standard twist so as to be supported on cliques in G (3.8) – 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. (Section 5) We deduce a "coarse" diagonalization of the resulting moment matrix from $\widetilde{E}$ in above. The deduction has two steps: 1. Analyze the expectation of the matrix; 2. The (imaginary) diagonalization of the matrix is in essence a quadratic equation, which we weaken to a proper "modular" version and solve the latter. We call step 2 the **mod-order analysis** (Section 5.2), 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. 5.5). One ingredient towards defining the weakening is the norm information on certain pseudo-random matrices (the *graphical matrices*).

The resulting weakened equation has a nice structure to work with (Lem. 5.6, Cor. A.2). 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 this equation 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 [25], 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

$$M' = NN^\top \tag{2.19}$$

in N with $\mathbb{E}[N] = C$, and N extending C by non-trivial Fourier characters. Two observations about (2.19) 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

$$\left\|C_r C_r^\top\right\| \approx \binom{d/2}{r} \cdot (\frac{\omega}{n})^{d-r} n^{d/2-r}, \quad r = 0, ..., d/2, \tag{2.20}$$

where assume $C = (C_0, ..., C_{d/2})$, each C_r having column dimension $\binom{[n]}{r}$. Assume $N = (N_0, ..., N_{d/2})$. Then we expect $N_r N_r^\top$ to concentrate around $C_r C_r^\top$ for each r, and so expect the norm of the non-constant part of $N_r N_r^\top$ to be bounded by (2.20). Under this expectation, the known tight norm bounds on related matrix pieces 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 (2.19) 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

$$(M' \mod \text{high order}) = (N \mod \text{high order}) \cdot (N^\top \mod \text{high order}) \tag{2.21}$$

where "order" means power of α (think of α as an "infinitesimal"). We call (2.21) the *mod-order equation* and its analysis the *mod-order analysis*. For details see Definition 5.5.

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 1.5(1). In Section 4 we recall some fundamental tools for analysis. In Section 5 we refresh the technique of recursive factorization and show Theorem 1.5(2). With all preparations in place, in Section 6 we prove the main Theorem 1.4. The paper is concluded in Section 7 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. Finally, we use $y(n) = O(x(n))$ to mean that there is some absolute constant c s.t. $y(n) \le cx(n)$ 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

In this section, we define the pseudo-expectations. As a warm-up we start with the non-exact problem, then move on to 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 $\widetilde{Ex}_S$ to each subset $S \subseteq [n]$ of size $\le d$, so that the resulting vector $\widetilde{Ex}$ *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 up 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 itself 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 the plant-situation where G is indeed a single random clique. Suppose its distribution is represented by a *plant-distribution* vector $p_{\text{plant}} \in \mathbb{R}^{2^{[n]}}$. Let the *output-expectation* q_{out} be indicator-vector of cliques in G in expectation. Then

$$q_{\text{out}} = \zeta \cdot p_{\text{plant}}. \tag{3.2}$$

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:*

$$p_0(S) = \frac{1}{\binom{n}{\omega}} \text{ if } |S| = \omega \quad \text{and } 0 \text{ otherwise}, \tag{3.3}$$

and

$$q_0(S) = (\zeta p_0)(S) = \frac{\binom{n-|S|}{\omega-|S|}}{\binom{n}{\omega}}. \tag{3.4}$$

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

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

$$p_1(S) = (\frac{\omega}{n})^{|S|}(1 - \frac{\omega}{n})^{n-|S|} \quad \forall S \subseteq [n], \tag{3.5}$$

and

$$q_1(S) = (\zeta p_1)(S) = (\frac{\omega}{n})^{|S|}. \tag{3.6}$$

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, and the inverse is the *Möbius inversion* matrix:

$$\zeta^{-1}(A, B) = (-1)^{|B \setminus A|} \text{ if } A \subseteq B, \text{ and } 0 \text{ otherwise}.$$

Note $(\zeta_{a,a})^{-1} = (\zeta^{-1})_{a,a}$ for all $a \leq n$. 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

$$M_{SOS} = \zeta \text{diag}(p) \zeta^{\top}. \tag{3.7}$$

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

3.1.2 The non-exact pseudo-expectation

Idea. Given any G, we will first construct a *degree-τ* "approximate plant-distribution" $p_\tau(G)$, which simulates the plant-distribution (Def. 3.1) in the sense that they give similar output-expectations. We also require $p_\tau(G)$ to be supported on size $\le \tau$-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 [13] we add a clique-indicator factor:

$$p_\tau(G)(S) = \left(2^{|\binom{S}{2}|}\mathrm{Cl}_S(G)\cdot\zeta^{-1}_{\tau,\tau}(q_1)_\tau\right)(S) \quad \forall S\subseteq[n] \text{ of size } \le\tau \tag{3.8}$$

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

▶ **Definition 3.2.** *$\forall S\subseteq[n]$, the* ***normalized clique-indicator*** *is function*

$$\widetilde{\mathrm{Cl}}_S(G) := 2^{|\binom{S}{2}|}\mathrm{Cl}_S(G). \tag{3.9}$$

$\widetilde{\mathrm{Cl}}(G)$ denotes 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*

$$\widetilde{E}_{\text{nonexact}} = \zeta_{d,\tau}\cdot p_\tau(G) = \zeta_{d,\tau}\cdot\left(\widetilde{\mathrm{Cl}}(G)\circ\zeta^{-1}_{\tau,\tau}\right)\cdot(q_1)_\tau \quad \in\mathbb{R}^{\binom{[n]}{\le d}} \tag{3.10}$$

where "$\circ$" is the Hadamard product[8]*.*

In short, $\widetilde{E}_{\text{nonexact}}$ refined the construction in [13] 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.

The resulting moment matrix is

$$M_{\text{nonexact}}(G) = \zeta_{d/2,\tau}\cdot\operatorname{diag}\left(p_\tau(G)\right)\cdot\left(\zeta_{d/2,\tau}\right)^\top, \tag{3.11}$$

similarly as (3.7).

▶ Remark 3.4. $\widetilde{E}_{\text{nonexact}}$ 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[9]. That the PSDness could still possibly hold is because $\zeta_{d/2,\tau}$ in (3.11) is degenerate.

▶ **Lemma 3.5** (Theorem 1.5(1))**.** *For all $S\subseteq[n]$ s.t. $|S|\le d$,*

$$\widetilde{E}_{\text{nonexact}}x_S = \sum_{T:|V(T)\cup S|\le\tau}\left(\frac{\omega}{n}\right)^{|V(T)\cup S|}\chi_T. \tag{3.12}$$

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

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

[8] 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.

[9] One intuition, suggested by a reviewer, 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 $\widetilde{E}_{\text{nonexact}}(x_\phi)\approx 1$ (which is not too hard to check by (3.12)), but its objective value $n^{\frac{1}{2}-\epsilon}$ is too big.

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

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

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

for all S with $|S|\leq d$. ◀

3.2 The exact case

In this subsection, we give a generic way to generate possible pseudo-expectations that satisfy Size Constraints (1.4). The idea is to define $\widetilde{E}x_S$ in a top-down fashion: fix $\widetilde{E}x_S$ for all $|S|=d$ first, then recursively set

$$\widetilde{E}x_S \leftarrow \frac{1}{\omega-|S|}\sum_{i\notin S}\widetilde{E}x_{S\cup\{i\}} \tag{3.13}$$

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

$$\mathrm{Fil}_{d,=d}(A,B)=\begin{cases}\binom{\omega-|A|}{d-|A|}^{-1}, & \text{if } A\subseteq B \text{ (where } |B|=d\text{);}\\ 0, & \text{otherwise.}\end{cases} \tag{3.14}$$

▶ **Definition 3.7.** *Given vector $\widetilde{E}_d x$ which assigns a value to each d-subset $S\subseteq[n]$, the **exact pseudo-expectation generated by $\widetilde{E}_d x$** is*

$$\widetilde{E}x := \mathrm{Fil}_{d,=d}\cdot\widetilde{E}_d x. \tag{3.15}$$

▶ **Lemma 3.8.** *The pseudo-expectation in Definition 3.7 satisfies the Size Constraints* (1.4)*, regardless of the choice of $\widetilde{E}_d x$.*

Proof. For any $S\in\binom{[n]}{<d}$, take a vector $v_S\in\mathbb{R}^{\binom{[n]}{\leq d}}$

$$v_S(S')=\begin{cases}\omega-|S|, & \text{if } S'=S;\\ -1, & \text{if } S'\supseteq S \text{ and } |S'\setminus S|=1;\\ 0, & \text{otherwise}\end{cases}$$

then it suffices to show $v_S^\top\mathrm{Fil}_{d,=d}=0$. But this 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 3.9 (Example). The following construction seems natural. Combining Def. 3.7 with the perspective from Section 3.1.2, we can take (3.10) with the exact plant-setting (p_0, q_0), followed by multiplying $\mathrm{Fil}_{d,=d}$:

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

Actually, it can be easily checked that it satisfies Clique Constraints; it also has a nice expression in Fourier characters. By some computation which we omit here, modulo provably negligible error the resulting matrix is

$$M_{\mathrm{example}}(I, J) = \sum_{\substack{T: \\ |V(T) \setminus (I \cup J)| \le \tau - d}} \chi_T \cdot \frac{\binom{n - |V(T) \cup I \cup J|}{\omega - |V(T) \cup I \cup J|}}{\binom{n}{\omega}}.$$

The only problem, however, is that we don't know how to prove the PSDness. Despite a transparent similarity to the previous expression (3.12), a similar proof breaks down seriously here, and the main reason is the loss of nice arithmetic structure when changing from function $(\frac{\omega}{n})^x$ to $\frac{\binom{n-x}{\omega-x}}{\binom{n}{\omega}}$. See also Remark 6.14.

3.3 The exact pseudo-expectation

Now we pinpoint an exact pseudo-expectation in Definition 3.7. With the idea stated in detail in the overview (Section 2.1), we give the construction directly.

We 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. F is called a ***d*-generating function**. Then for general $|S| \le d$, (3.14) gives:

$$\widetilde{E}x_S = \frac{1}{\binom{w-d+u}{u}} \sum_{T: |V(T) \cup S| \le \tau} \chi_T \cdot \left[\sum_{c=0}^{u} \binom{|V(T) \cup S| - d + u}{c} \binom{n - |V(T) \cup S|}{u - c} \right. \\ \left. \cdot F(|V(T) \cup S| + u - c) \right] \tag{3.16}$$

where we have let $u := d - |S|$.

▶ **Lemma 3.10.** *Any exact pseudo-expectation generated by* (3.16) *satisfies the Clique and Size Constraints* (1.3),(1.4).

Proof. For Clique Constraints, note (3.16) only depends on $|V(T) \cup S|$, so by grouping terms $\widetilde{E}x_S = \sum_{T: |V(T) \cup I \cup J| \le \tau} M(I, J; T) \chi_T$ factors through $\widetilde{\mathrm{Cl}}_{I \cup J} = \sum_{T \subseteq E(I \cup J)} \chi_T$. I.e., $M(I, J)(G) = 0$ if $\widetilde{\mathrm{Cl}}_{I \cup J}(G) = 0$.

It satisfies Size Constraints by Lemma 3.8. ◀

Now we pinpoint a choice of the d-generating function.

▶ **Definition 3.11** (Exact d-generating function)**.**

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

▶ Remark 3.12. As is said in the proof overview, the design of F, especially its first factor, is technical and the ultimate goal is to make the resulting M positive. The numerator $(x+8\tau^2)!$ will be used in Proposition 6.28, where the term $8\tau^2$ can be replaced by larger polynomials in τ. The $(8\tau^2)!$ in denominator is added just for convenience (see Remark 3.14).

▶ **Definition 3.13.** *The* ***exact moment matrix*** $\widetilde{M}$ *is*

$$\widetilde{M}(A,B) = \sum_{T:|V(T)\cup A\cup B|\le\tau} \widetilde{M}(A,B;T)\chi_T \quad \forall |A|,|B| \le d/2$$

where $\widetilde{M}(A,B;T) =$

$$\frac{1}{\binom{\omega-d+u}{u}} \Bigg[\sum_{c=0}^{u} \binom{|V(T)\cup A\cup B| - (d-u)}{c} \binom{n-|V(T)\cup A\cup B|}{u-c} \cdot \underbrace{\frac{\big(|V(T)\cup A\cup B| + u - c + 8\tau^2\big)!}{(8\tau^2)!} \cdot (\frac{\omega}{n})^{|(V(T)\cup A\cup B)|+u-c}\Bigg]}_{f\big(|V(T)\cup A\cup B|+u-c\big)} \tag{3.17}$$

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

▶ Remark 3.14. In (3.17), the "most significant" factor is $(\frac{\omega}{n})^{|V(T)\cup A\cup B|} \cdot \omega^{-c}$, if notice $\frac{\binom{n-|V(T)\cup A\cup B|}{u-c}}{\binom{\omega-d+u}{u}}\omega^u n^{-(u-c)}$ has 0th-order in ω, n. One thing to keep in mind is that other factors like $\frac{(|V(T)\cup A\cup B|+u-c+8\tau^2)!}{(8\tau^2)!}$ are qualitatively smaller than ω, within our parameter regime.

4 Preparations

In this section, we prepare some necessary tools for studying the matrices.

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

$$\exists a \in \mathbb{R}^{\binom{[n]}{\le d/2}} \quad a^\top M_{dSOS} a = -1. \tag{4.1}$$

With the presence of boolean constraints (i.e. define $\widetilde{E}(x_i^2 \cdot p) := \widetilde{E}(x_i \cdot p)$ for all i and all polynomial p of degree $\leq d-2$), this is equivalent to

$$\widetilde{E}(g^2) = -1 \tag{4.2}$$

where $g = a^\top x = \sum_{|S| \leq d/2} a_S x_S$ is multi-linear. Now substitute every x_S ($|S| < d$) in g by the corresponding linear combination of $\{x_{S'} \mid |S'| = d\}$ from (3.13). This does not affect the value of (4.2) since $\widetilde{E}$ satisfies the equality constraints. We get

$$\widetilde{E}(g_1^2) = -1 \tag{4.3}$$

for some multi-linear, degree-$d/2$ homogeneous g_1. Now translate (4.3) back (assume $g_1 = b^T x$, $x = (x_S)_{|S|=d/2}$) to $b^\top M b = -1$, we see that M is not PSD. ◀

4.2 Concentration bound on polynomials

The following 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. $|c(T)| \leq C$ for all T. Then W.p. $1 - n^{-10\log n}$ over G,

$$|p(G)| < C \cdot n^{a/2} 2^{a^2} n^{4\log\log n}. \tag{4.4}$$

Proof. Power-estimation. For all $k \in \mathbb{N}$, (we can think of $a < k = o(n/a)$)

$$p^{2k} = \sum_{T_1,\ldots,T_{2k}:\ |V(T_i)|=a} c(T_1)\ldots c(T_{2k})\chi_{T_1} \cdot \ldots \cdot \chi_{T_{2k}} \tag{4.5}$$

Take expectation of (4.5). Each $\mathbb{E}[\chi_{T_1}\ldots\chi_{T_{2k}}(G)] \neq 0$ (i.e. equals 1) iff every edge appears even times in $T_1, \ldots, T_{2k}$, which implies $|V(T_1 \cup \ldots \cup T_{2k})| \leq \frac{1}{2} \cdot 2ka = ka$. There are at most $ka\binom{n}{ka} < n^{ka}$ many choices of such $V(T_1 \cup \ldots \cup T_{2k})$. For each choice, there are in turn 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,

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

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

Now we state a concentration bound on pseudo-random matrices which, like in almost all previous work on the subject, will be a fundamental tool for us.

The pseudo-random matrices refer to the *graphical matrices* ([24]). Intuitively, such a matrix collects Fourier characters of all embeddings of a fixed "shape". Definition 4.3, 4.5 below are implicit in [24, 25, 16] and is termed explicitly in [5].

▶ **Definition 4.3.** *A* ***ribbon*** *$\mathcal{R}$ is a (ordered) triple $(A, B; T)$ where A, B are vertex-sets and T is an edge set. A, B are called the left and right vertex set of $\mathcal{R}$. The* ***size*** *of $\mathcal{R}$ is*

$$|V(\mathcal{R})| = |V(T) \cup A \cup B|.$$

By definition, a ribbon $\mathcal{R} = (A, B; T)$ as a graph always has no isolated vertex outside of $A \cup B$.

▶ **Definition 4.4.** *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*[10] *from A without touching B. Being* ***right-generated*** *is symmetrically defined.*

▶ **Definition 4.5.** *For ribbon $(A, B; T)$, if further $A \cup B$ is totally-ordered, it is called a* ***shape****. Denote a shape by $\mathcal{U} = (A, B; T)$. As before, $V(\mathcal{U}) = A \cup B \cup V(T)$, and its* ***size*** *is $|V(\mathcal{U})|$.*

When fixing an underlying vertex-set $[n]$, a ribbon $\mathcal{R}$ within vertex set $[n]$ can always be regarded as shapes, with the induced ordering on vertices. So in this setting, we may speak of the shape of $\mathcal{R}$ and interchangeably use $\mathcal{R}$ to denote shapes.

▶ **Definition 4.6.** *A real-valued function f defined on a set of ribbons within vertex-set $[n]$ is called* ***symmetric with respect to shapes****, if whenever $\mathcal{R}$ and $\mathcal{R}'$ are of the same shape then $f(\mathcal{R}) = f(\mathcal{R}')$.*

▶ **Definition 4.7** ([24])**.** *Fix an n, and a shape $\mathcal{U} = (A, B; T)$ Define the* ***graphical matrix*** *of shape $\mathcal{U}$ to be the following $2^{[n]} \times 2^{[n]}$-matrix $M_{\mathcal{U}}$. Call a map $\phi : V(\mathcal{U}) \to [n]$* proper *if ϕ is injective and respects the order on $A \cup B$, then*

$$\forall I, J \subseteq [n], \quad M_{\mathcal{U}}(I, J) = \sum_{\substack{T:\ \exists proper\ \phi\ s.t. \\ \phi(A)=I, \phi(B)=J, \phi(T)=T'}} \chi_{T'}$$

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

▶ **Theorem 4.8** (Norm bounds on $M_{\mathcal{U}}$ [24, 5])**.** *For any shape $\mathcal{U} = (A, B; T)$ of size $t < \log n$, w.p. $> 1 - n^{-10 \log n}$ over G,*

$$\|M_{\mathcal{U}}(G)\| \leq n^{\frac{t-p}{2}} \cdot 2^{O(t)} \cdot (\log n)^{O(t+p-2r)} \tag{4.6}$$

where $r = |A \cap B|$, p is the maximum number of vertex-disjoint paths between (A, B) in $\mathcal{U}$. Moreover, this bound is tight up to polylog(n)-factors, for all $M_{\mathcal{U}}$ with the described parameters ([24], Thm 38]).

Moreover, under the same notation, if further denote $s = \frac{|A|+|B|}{2}$ then

$$\|M_{\mathcal{U}}(G)\| \leq n^{\frac{t-p}{2}} \cdot 2^{O(t)} \cdot (\log n)^{O(t-s)}. \tag{4.7}$$

Theorem 4.8 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 (4.6) since $t \geq |A \cup B| = 2s - r$, $p \leq s$, so

$$t + p - 2r \leq t + s - 2(2s - t) = 3(t - s).$$

▶ Remark 4.9. Theorem 4.8 and its proof is a far-reaching generalization of that of the concentration bounds on polynomials, Lemma 4.2. Namely, if take special shapes in the form $\mathcal{U} = (A, A; T)$, then the corresponding matrix $M_{\mathcal{U}}$ is diagonal, so estimating its norm is equivalent to estimating absolute values of the diagonals which are polynomials.

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

4.4 Some general notions on graphs

We finish our preparation with some general graph-theoretic notions.

▶ **Definition 4.10** (Vertex-separator). *Given graph H and two vertex-subsets $A, B \subseteq V(H)$, call $S \subseteq V(H)$ an (A, B)-**vertex-separator** if any path*[11] *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**. $\mathrm{mSep}_{A,B}(H)$ denotes the set of all min-separators.*

This definition naturally applies to ribbons $\mathcal{R} = (A, B; T)$, by using the graph H as on $V(T) \cup A \cup B$ with edge-set T. In that case we can write the corresponding size and set of the min-separators as

$$s_{A,B}(T), \quad \mathrm{mSep}_{A,B}(T) \text{ or } \mathrm{mSep}(\mathcal{R}).$$

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

$$e_{A,B}(T) := |V(T) \cup A \cup B| - s_{A,B}(T). \tag{4.8}$$

The reduced size is double of the exponent in n in the bound of Theorem 4.8, hence is the controlling parameter of the norm of the graphical matrix.

5 Non-exact case PSDness: a refresh

In this section, we review and refresh the proof techniques for the non-exact problem. In Section 5.1 and 5.2, we show Theorem 1.5(2) via the so-called *mod-order analysis*, which gives a conceptually different approach to the techniques. In Section 5.3, we formalize the recursive factorization in a convenient language and extend it properly for later use.

Declaration. Section 5.2 is only for Theorem 1.5(2). The reader can safely skip it if she wants to proceed directly to the proof of Theorem 1.4.

Notation. Thoughout Section 5, M' denotes the $\binom{[n]}{\frac{d}{2}} \times \binom{[n]}{\frac{d}{2}}$-minor[12] of the non-exact moment matrix.

$$M'(I, J) = \sum_{T:|V(T)\cup I\cup J|\le\tau} \left(\frac{\omega}{n}\right)^{|V(T)\cup I\cup J|} \chi_T \quad \forall I, J \in \binom{[n]}{d/2}. \tag{5.1}$$

Goal of Section 5. Diagonalize M' approximately, such that the difference matrix is negligible (w.h.p. when plugging G).

[11] Same as in the previous footnote. In particular, every vertex-separator contains $A \cap B$.

[12] 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.

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

▶ **Proposition 5.1.** $\mathbb{E}[M'] = CC^\top$, *where* C *is the* $\binom{[n]}{d/2} \times \binom{[n]}{\le d/2}$*-matrix*

$$C = (\zeta^\top)_{d/2,\le d/2} \cdot \operatorname{diag}\left(\sqrt{t(|A|)}\right)_{A\in\binom{[n]}{\le d/2}} \tag{5.2}$$

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 [25], as below.

▶ **Definition 5.2** (See e.g. [9]). *Fix parameters* n, k. *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

$$D_r(I,J) = \begin{cases} 1, & \text{if } |I \cap J| = r \\ 0, & \text{o.w.} \end{cases} \qquad \forall I, J \in \binom{S}{d/2}. \tag{5.3}$$

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

$$\mathfrak{J}_r(I,J) = \binom{|I\cap J|}{r}, \quad \forall I, J \in \binom{[n]}{\frac{d}{2}}. \tag{5.4}$$

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

$$\mathfrak{J}_r = \sum_{A\subseteq[n],|A|=r} u_A u_A^\top \quad \text{where} \quad u_A \in \mathbb{R}^{\binom{[n]}{k}},\ u_A(B) = 1_{A\subseteq B}. \tag{5.5}$$

Clearly $\mathfrak{J}_{d/2} = \mathrm{Id}$. More generally, we have:

▶ **Fact 5.3** (See e.g. (4.29) in [9]). *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*

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

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

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

$$D_r = \sum_{r'=r}^{d/2} (-1)^{r'-r}\binom{r'}{r} \cdot \mathfrak{J}_{r'} \tag{5.6}$$

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

$$\begin{aligned}
\mathbb{E}[M'] &= \sum_{r=0}^{d/2}(\frac{\omega}{n})^{d-r}\left(\sum_{r'=r}^{d/2}(-1)^{r'-r}\binom{r'}{r}\mathfrak{J}_{r'}\right) \\
&= \sum_{r'=0}^{d/2}\mathfrak{J}_{r'}\cdot\left(\sum_{r=0}^{r'}(\frac{\omega}{n})^{d-r}(-1)^{r'-r}\binom{r'}{r}\right) \qquad (5.7)\\
&= \sum_{r'=0}^{d/2}\mathfrak{J}_{r'}\cdot(\frac{\omega}{n})^{d-r'}(1-\frac{\omega}{n})^{r'}
\end{aligned}$$

which proves the lemma. ◀

By Lemma 5.4 and (5.5), if let $t(r) = (\frac{\omega}{n})^{d-r'}[1-\frac{\omega}{n}]^{r'}$ then

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

where used that the matrix $(\zeta^\top)_{d/2,\le d/2}$ has columns $\{u_A \mid |A| \le d/2\}$. This proves Proposition 5.1.

5.2 Step 2: Mod-order analysis toward "coarse" diagonalization

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

$$M' = NN^\top \qquad (5.8)$$

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 (Def. 4.6).

Toward this goal, we define and study a relaxed equation first (Definition 5.5). Let us start with its 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 5.3, each $\mathfrak{J}_r$ b has norm $\binom{d/2}{r}\cdot n^{d/2-r}$ so

$$\|C_rC_r^\top\| \approx \binom{d/2}{r}\cdot(\frac{\omega}{n})^{d-r}n^{d/2-r}, \quad r = 0,...,d/2. \qquad (5.9)$$

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 (5.9). The tight bound from Theorem 4.8 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

$$M' = L \cdot \operatorname{diag}\left((\frac{\omega}{n})^{|A|}\right) \cdot L^\top \quad \text{with} \quad \mathbb{E}[L] = (\ C_r \cdot (\frac{\omega}{n})^{-r/2}\ )_{r=0,1,\dots,d/2}. \tag{5.10}$$

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:

$$\beta_{I,A}(T) = \underbrace{(\frac{\omega}{n})^x}_{\text{main-order term}} \cdot (\text{ factor} \ll \frac{n}{\omega} \text{ and} \gg \frac{\omega}{n}\ ). \tag{5.11}$$

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 graphical matrix square in $L_r(\frac{\omega}{n})^r L_r^\top$ is

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

which has norm[13]

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

by Theorem 4.8. Here recall $e_{I,A}(T) = |V(T)\cup I\cup A| - s_{I,A}(T) (\geq |I| - |A| = \frac{d}{2} - r)$. Compare this with (5.9), 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 $x \geq e_{I,A}(T)$ which actually is the limit requirement when $\frac{\log \omega}{\log n} \to \frac{1}{2}$. Suggested by this, we will set the restriction $x \geq e_{I,A}(T)$ 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 5.5.** *The **mod-order equation** is*

$$L_\alpha \cdot \operatorname{diag}\left(\alpha^{|A|}\right) \cdot (L_\alpha)^\top = M_\alpha \qquad \text{mod } (*) \tag{5.12}$$

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

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

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

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

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[14]

$$\alpha^{e_{I,A}(T)} \mid \beta_{I,A}(T) \quad \forall I, A, T. \tag{5.13}$$

*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.*

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

[14] Recall $e_{I,A}(T')$ is the reduced size $|V(T')\cup I\cup A| - s_{I,A}(T')$ (Def. 4.11).

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 5.6** (Order match). *If a product $\alpha^{|A|} \cdot \beta_{I,A}(T') \cdot \beta_{J,A}(T'')$ from the LHS of* (5.12) *is nonzero mod* $(*)$*, then both of the following hold:*

$$A \text{ is a min-separator for both } (I, A; T'),\ (J, A; T''); \tag{5.14}$$

$$(V(T') \cup I \cup A) \ \cap\ (V(T'') \cup J \cup A) = A. \tag{5.15}$$

Moreover, (5.14), (5.15) *imply that*

$$A \text{ is a min-separator of } (I, J; T) \text{ (where } T = T' \oplus T''\text{)}; \tag{5.16}$$

$$|V(T') \cup I \cup A|,\ |V(T'') \cup J \cup A| \le \tau. \tag{5.17}$$

By this lemma, in an imagined solution we can assume $\beta_{I,A}(T') \neq 0$ only when it satisfies its part in conditions (5.14), (5.17).

Using this information, plus a further technique of *polarization*, we can deduce the following Proposition 5.8 which is the main takeaway of the analysis here. A graph-theoretic fact (the "in particular" part below) appears exactly as the solvability condition. For deductions see Appendix A.2.

▶ **Fact 5.7** ([11]). *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 actually 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 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 "left-most" and "rightmost" min-separator, respectively.

▶ **Proposition 5.8** (Mod-order diagonalization). *Let*

$$L_\alpha(I, A) := \sum_{\substack{T':\ |V(T') \cup I \cup A| \le \tau \\ A = S_l(I,A;T') \\ T' \cap E(A) = \emptyset \\ (I,A;T') \text{ left-generated (Def. 4.4)}}} \alpha^{e_{I,A}(T')} \chi_{T'},$$

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

(T_m to indicate "middle"). Then

$$L_\alpha \cdot [\mathrm{diag}\left(\alpha^{\frac{|A|}{2}}\right) \cdot Q_{0,\alpha} \cdot \mathrm{diag}\left(\alpha^{\frac{|A|}{2}}\right)] \cdot L_\alpha^\top = M_\alpha \qquad \mathrm{mod}\ (*) \tag{5.18}$$

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 (5.18) is slightly weaker than a solution to (5.12) but is sufficient for all use, as we are only concerned with PSDness. In particular, it gives the first-approximate diagonalization of the matrix M', recast as Definition 5.9 below. This shows Theorem 1.5(2).

5.3 Recursive factorization

In this subsection, we give a formalization and extension of the recursive factorization technique, which is used to refine the coarse diagonalization from Step 2 above. We give some new notions that are convenient and extendable to matrix products (Def. 5.13, 5.15), along with some simplification (Lem. 5.25) and refinement (Prop. 5.24) for later use.

First, the coarse diagonalization (5.18) can be recast in $\mathbb{R}[\{\chi_T\}]$-matrices as below.

▶ **Definition 5.9.** *Let L be the $\binom{[n]}{\frac{d}{2}} \times \binom{[n]}{\leq \frac{d}{2}}$-matrix*

$$L(I,A) := \sum_{\substack{T':\ |V(T') \cup I \cup A| \leq \tau \\ A = S_l(I,A;T') \\ T' \cap E(A) = \emptyset \\ (I,A;T')\ \text{left-generated}}} (\frac{\omega}{n})^{|V(T') \cup I \cup A| - |A|} \chi_{T'}, \tag{5.19}$$

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

$$Q_0(A,B) := \sum_{\substack{T_m : |T_m \cup A \cup B| \leq \tau \\ A,B \in \mathrm{mSep}_{A,B}(T_m)}} (\frac{\omega}{n})^{|V(T_m) \cup A \cup B|} \chi_{T_m}. \tag{5.20}$$

Finally, let

$$D := \operatorname{diag}\left((\frac{\omega}{n})^{\frac{|A|}{2}} \right)_{A \in \binom{[n]}{\leq d/2}}. \tag{5.21}$$

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

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

$$M' - L(DQ_0D)L^\top \tag{5.22}$$

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

$$M' = L \cdot [D \cdot (Q_0 - Q_1 + Q_2 ... \pm Q_{d/2}) \cdot D] \cdot L^\top + \mathcal{E} \tag{5.23}$$

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

▶ Remark 5.10. Use of D is superficial in (5.22), (5.23); we keep it so that the middle matrices Q_i are better-positioned. The LD here corresponds to the "L" matrix in [5].

Let us start with some necessary notions.

5.3.1 More notion on graphs

▶ **Definition 5.11** ([5] Def. 6.5[15]). *For any ribbon $\mathcal{R} = (I,J;T)$, its **canonical decomposition** is a ribbon-triple*

$$(\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r) = ((I,A;T_l),(A,B;T_m),(B,J;T_r))$$

[15] Similar notions actually appeared implicitly in the mod-order analysis (cf. condition (5.14), (5.15)), while here they appear in a more "canonical" left-, middle-, right- form.

determined uniquely by the following. $A = S_l(I, J; T)$*,* $B = S_r(I, J; T)$*.* $V(\mathcal{R}_l)$ *is* A *unioned with the set of vertices reachable by paths from* I *in* T *without touching* A*, and* $T_l = T|_{V(\mathcal{R}_l)} \backslash E(A)$*. Similarly,* $V(\mathcal{R}_r)$ *is* B *unioned with the set of the vertices reachable from* J *in* T *without touching* B*, and* $T_r = T|_{V(\mathcal{R}_r)} \backslash E(B)$*. 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 5.12 (Properties of the canonical decomposition). A few properties follow from the definition of the canonical decomposition of $\mathcal{R} = (I, J; T)$.

$$A = S_l(I, A; T_l), \;\; B = S_r(B, J; T_r)$$

(so they are unique separator of $\mathcal{R}_l, \mathcal{R}_r$, respectively);

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

$$\mathcal{R}_l \text{ is left-generated,} \quad \mathcal{R}_r \text{ is right-generated} \;\; \text{(Def. 4.4);}$$

$$A, B \in \; \mathrm{mSep}_{A,B}(T_m) \quad (\text{so } |A| = |B|).$$

The above four are about each of $\mathcal{R}_l$, $\mathcal{R}_m$, $\mathcal{R}_m$ (the "inner" conditions). Moreover, there is the intersection property on pairs of them (the "outer" conditions)[16]:

$$V(\mathcal{R}_l) \cap V(\mathcal{R}_m) \subseteq A, \;\; V(\mathcal{R}_m) \cap V(\mathcal{R}_r) \subseteq B, \;\; V(\mathcal{R}_l) \cap V(\mathcal{R}_r) \subseteq A \cap B$$

which implies

$$e(\mathcal{R}_l) + |V(\mathcal{R}_m)| + e(\mathcal{R}_r) = |V(\mathcal{R})|. \tag{5.24}$$

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

▶ **Definition 5.13** (Inner and outer canonicality). *For a triple of ribbons in the form*

$$(\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r) = \Big((I, A; T_l), (A, B; T_m), (B, J; T_r)\Big)$$

(T_l, T_m, T_r *are arbitrary subsets of an edge-set), their* ***ribbon-sum*** *is ribbon*

$$(I, J; T) \quad \textit{where } T = T_l \oplus T_m \oplus T_r.$$

The triple is called ***inner-canonical****, if they satisfy the "inner" conditions:*

$$\begin{aligned} &A = S_l(I, A; T_l), \quad B = S_r(B, J; T_r), \\ &T_l \cap E(A) = \emptyset = T_r \cap E[A], \\ &\mathcal{R}_l \textit{ left-generated,} \quad \mathcal{R}_r \textit{ right-generated,} \\ &A, B \in \; \mathrm{mSep}_{A,B}(T_m). \end{aligned} \tag{5.25}$$

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

$$V(\mathcal{R}_l) \cap V(\mathcal{R}_m) \subseteq A, \;\; V(\mathcal{R}_m) \cap V(\mathcal{R}_r) \subseteq B, \;\; V(\mathcal{R}_l) \cap V(\mathcal{R}_r) \subseteq A \cap B. \tag{5.26}$$

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

[16] cf. conditions (5.14), (5.15)

▶ **Proposition 5.14.** *Canonical triples are 1-1 correspondent to their ribbon-sum, via the canonical decomposition.*

Proof. This follows by an immediate check from the definition. ◀

We further extend the notions to matrix products. Recall $\mathbb{R}[\{\chi_T\}]$ is the ring from adding fresh variables 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 5.15** (Approximate form). *Suppose matrices X, Y have rows and columns indexed by subsets of $[n]$ with entries in $\mathbb{R}[\{\chi_T\}]$; and in every entry, each character regarded as a ribbon on distinguished sets (row, column) has ribbon size $\leq \tau$. Suppose X, Y have dimensions s.t. $XYX^\top$ is defined.*

Every nonzero triple product (without collecting like-terms) in

$$XYX^\top \tag{5.27}$$

thus has form

$$\underbrace{X(I, A; T_l)Y(A, B; T_m)X(J, B; T_r)}_{\text{nonzero in } \mathbb{R}} \chi_{T_l} \cdot \chi_{T_m} \cdot \chi_{T_r}, \tag{5.28}$$

and can be identified with a ribbon-triple in the natural way, with

$$X(I, A; T_l)Y(A, B; T_m)X(J, B; T_r)\chi_{T_l \oplus T_m \oplus T_r} \quad \in \mathbb{R}[\{\chi_T\}]$$

its ***resulting term****. We say* (5.28) *is an* ***outer-canonical product*** *if the ribbon-triple is outer-canonical, and it* ***exceeds degree*** *if $|V(T) \cup I \cup J| > \tau$.*

The ***approximation form*** *of $XYX^\top$ is:*

$$XYX^\top = \left(XYX^\top\right)_{\text{can}} + (XYX^\top)_{\text{non-can}} + \mathcal{E}_{\text{deg}}, \tag{5.29}$$

or equivalently,

$$\left(XYX^\top\right)_{\text{can}} = XYX^\top - (XYX^\top)_{\text{non-can}} - \mathcal{E}_{\text{deg}},$$

where $\left(XYX^\top\right)_{\text{out-can}}$ is the matrix collecting all terms of outer-canonical products that do not exceed degree, $(XYX^\top)_{\text{non-can}}$ collecting all terms of non-outer-canonical products, and $\mathcal{E}_{\text{deg}}$ collecting all rest terms.

▶ Remark 5.16. With this language, Proposition 5.14 gives an *a posteriori* explanation of the coarse diagonalization (Def. 5.9): $M' = [L(DQ_0D)L^\top]_{\text{can}}$.

5.3.2 Recursive factorization: the machinery

We start with the following, which is Definition 5.9 restated in the current language.

▶ **Definition 5.17** (First-approximate factorization of M').

$$M' = L(DQ_0D)L^\top - [L(DQ_0D)L^\top]_{\text{non-can}} - \mathcal{E}_{1;\text{deg}} \tag{5.30}$$

where $\mathcal{E}_{1;\text{deg}}$ is by Def. 5.15, applied to the product $L(DQ_0D)L^\top$, where the index "1" is added for later convenience. $L(DQ_0D)L^\top$ is celled the ***first-approximate factorization*** *of M'.*

The high-degree error $\mathcal{E}_{1;\text{deg}}$ is actually negligible in norm[17] (we will prove the analogous statement in the exact case); the main task is to analyze the "main error", $[L(DQ_0D)L^\top]_{\text{non-can}}$. For this, the key point of the whole technique is

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

I.e. $\exists Q_1$ s.t.

$$[L(DQ_0D)L^\top]_{\text{non-can}} = [L(DQ_1D)L^\top]_{\text{can}} + \mathcal{E}'_{1;\text{negl}}. \tag{5.31}$$

for some negligible $\mathcal{E}'_{1;\text{negl}}$. And we can repeat this for $[L(DQ_1D)L^\top]_{\text{non-can}}$ and so on. To describe the factorization (5.31), a generalized notion is useful.

▶ **Definition 5.18** ([5], Def. 6.9[18]). *A **generalized ribbon** is a usual ribbon together with a new set of isolated vertices. In symbol, it is denoted as $\mathcal{R}^* = (A, B; T^*)$ where*

$$T^* = T \sqcup \mathcal{I},$$

*T an edge-set, $\mathcal{I}$ a vertex set disjoint from $V(T) \cup A \cup B$, called the **isolated vertex-set of $\mathcal{R}^*$**, denoted as $\mathcal{I}(\mathcal{R}^*)$. $V(\mathcal{R}^*) = V(T) \cup A \cup B \cup \mathcal{I}$. A usual ribbon is also a generalized ribbon with $\mathcal{I} = \emptyset$. $(A, B; T)$ is called the (unique) largest ribbon in $\mathcal{R}$.*

▶ Remark 5.19. $\mathcal{I}(\mathcal{R}^*)$ could be different from the isolated set of the underlying graph, as it excludes vertices in $A \cup B$.

▶ **Definition 5.20.** *A **side-inner-canonical triple** is*

$$(\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r) = ((I, A; T_l),\ (A, B; T_m),\ (B, J; T_r))$$

where $\mathcal{R}_l$, $\mathcal{R}_r$ are ribbons satisfying the inner-canonical conditions on their part (the first three of (5.25)*), while $\mathcal{R}_m$ is just a ribbon.*

The following operation is the technical core of recursive factorizations.

▶ **Definition 5.21** (Separating factorization; Def. 6.10 of [5]). *Given an side-inner-canonical tripe*

$$(\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r) = ((I, A; T_l),\ (A, B; T_m),\ (B, J; T_r)),$$

*denote $T = T_l \oplus T_m \oplus T_r$, and denote by Z the multi-set of "unexpected intersections" i.e. 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 $z(\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r) = |Z|$ the **intersection size** of the triple. It can be checked that*

$$|V(\mathcal{R}_l) \cup V(\mathcal{R}_m) \cup V(\mathcal{R}_r)| = |V(\mathcal{R}_l)| + |V(\mathcal{R}_m)| + |V(\mathcal{R}_r)| - |A| - |B| - z. \tag{5.32}$$

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

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

[17] Matrices considered all have support on clique-rows and clique-columns, given G.
[18] It was called *improper ribbon*, but we feel the name here is perhaps more proper.

Define ribbon $\mathcal{R}'_l = (I, S'_l; T'_l)$, 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 to $V(\mathcal{R}'_l)$. Ribbon $\mathcal{R}'_r$ is symmetrically defined. In particular, $T'_l \cap T'_r = \emptyset$. $\mathcal{R}^*_m$ is the ***generalized*** *ribbon $(S'_l, S'_r; T^*_m)$ where*

$$T^*_m = T \backslash (T'_l \sqcup T'_r) \ \sqcup \ \mathcal{I}(\mathcal{R}^*_m),$$

$\mathcal{I}(\mathcal{R}^*_m)$ collecting all the rest isolated vertices:

$$\mathcal{I}(\mathcal{R}^*_m) = V(\mathcal{R}_l) \cup V(\mathcal{R}_m) \cup V(\mathcal{R}_r) \ - \ V(T) \cup I \cup J. \tag{5.33}$$

The resulting $(\mathcal{R}'_l, \mathcal{R}^*_m, \mathcal{R}'_r)$ is called the ***separating factorization*** *of ribbon triple $(\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r)$, which we denote as*

$$(\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r) \to (\mathcal{R}'_l, \mathcal{R}^*_m, \mathcal{R}'_r). \tag{5.34}$$

▶ Remark 5.22 (Properties of separating factorization). Some natural properties follow. Let $(\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r) \to (\mathcal{R}'_l, \mathcal{R}^*_m, \mathcal{R}'_r)$ in the same notation as above.

(1) The resulting triple $(\mathcal{R}'_l, R^*_m, \mathcal{R}'_r)$ is side-inner-canonical and outer-canonical (i.e. their pair-wise vertex intersections are within the corresponding S'_l, S'_r and $S'_l \cap S'_r$). So the corresponding ribbon triple (from replacing R^*_m with its largest ribbon) is canonical and is disjoint from $\mathcal{I}(\mathcal{R}^*_m)$.

(2) $\mathcal{R}'_l \subseteq \mathcal{R}_l$, and S'_l separates $(V(\mathcal{R}'_l),\ V(\mathcal{R}_l) - V(\mathcal{R}'_l))$ in $\mathcal{R}_l$. In particular, we can talk about the part of $\mathcal{R}_l$ to the right of S'_l, which is disjoint from R'_l and actually can be easily checked to be in $\mathcal{R}^*_m$. Similar fact holds for $\mathcal{R}_r$.

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

▶ **Lemma 5.23** (Lemma 6.14, 7.14 of [5]). *Suppose $(\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r) \to (\mathcal{R}'_l, \mathcal{R}^*_m, \mathcal{R}'_r)$. In the same notation as in Definition 5.21,*

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

(2) [19] *If further denote $s = \frac{|A|+|B|}{2}$, p' the maximum number of vertex-disjoint paths from S'_l to S'_r in $\mathcal{R}^*_m$, and p the maximum number of vertex-disjoint paths from A to B in $\mathcal{R}_m$, then*

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

Proof.

(1) By definition there must be some unexpected pair-wise intersection between $(\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r)$. In either of the three cases of breaking (5.26), $\exists v \in Z$ that is in $V(\mathcal{R}_l) - A$ or in $V(\mathcal{R}_r) - B$. WLOG suppose the first 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) We refer the reader to its proof in the original paper. ◀

Now we apply the above machinery to the target, $L(DQ_0D)L^\top$.

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

5.3.3 Apply the machinery

Conceptually, the separating factorization tells us how to "cancel" the terms in $[L(DQ_0D)L^\top]_{\text{non-can}}$ using $L, L^\top$. Namely, in $L(DQ_0D)L^\top$, any product from $(\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r)$ (Def.5.15) that is non-outer-canonical results in a term in $[L(DQ_0D)L^\top]_{\text{non-can}}$ at (I, J), and we can cancel it by the product from its separating factorization $(\mathcal{R}'_l, \mathcal{R}^*_m, \mathcal{R}'_r)$: take R'_l at position (I, S'_l) in L, R'_r at position (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. I.e., we cancel it by $-[L(DQ_1D)L^\top]_{\text{can}}$.

Of course, there are other triples whose separating factorization result in the same ($\mathcal{R}'_l$, largest ribbon of $\mathcal{R}^*_m$, $\mathcal{R}'_r$) so we need to collect them all in DQ_1D. More seriously, the (I, S'_l)th entry of L is actually a sum of different R'_ls, so we need to make sure that this cancellation works for them simultaneously in multiplication.

The following is what insures the simultaneous cancellation can work. It is stated in a refined version that is more than needed here (i.e. we further distinguish different (i, j) parameters), but this will be needed in the exact case (Lemma 6.20).

▶ **Proposition 5.24** (Solvability condition, cf. Claim 6.12 in [5]). *Fix (I, J, S'_l, S'_r), and a generalized ribbon $\mathcal{R}^*_m$ on (S'_l, S'_r). Let $(\mathcal{R}'_l, \mathcal{R}'_r)$ be inner-canonical left and right ribbons with distinguished sets $(I, S'_l), (S'_r, J)$ respectively, as in Definition 5.13. Let $(\mathcal{R}''_l, \mathcal{R}''_r)$ be another such ribbon pair, with the same reduced size*

$$e(\mathcal{R}'_l) = e(\mathcal{R}''_l),\ e(\mathcal{R}'_r) = e(\mathcal{R}''_r).$$

(Or the same size, equivalently.) Then for every fixed tuple (i, j, z) the following holds: there is an 1-1 matching between ribbon-triples

$$(\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r)\ \text{s.t.}\ \begin{cases} (\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r) \to (\mathcal{R}'_l, \mathcal{R}^*_m, \mathcal{R}'_r), \\ (e(\mathcal{R}_l),\ e(\mathcal{R}_r),\ z(\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r)) = (i, j, z). \end{cases} \tag{5.35}$$

and

$$(\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r)\ \text{s.t.}\ \begin{cases} (\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r) \to (\mathcal{R}''_l, \mathcal{R}^*_m, \mathcal{R}''_r), \\ (e(\mathcal{R}_l),\ e(\mathcal{R}_r),\ z(\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r)) = (i, j, z). \end{cases} \tag{5.36}$$

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

Proof. We give a reversible map from the set of (5.35) onto the set of (5.36). Take a $(\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r)$ from (5.35). By Remark 5.22 (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 the 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 $\mathcal{R}'_l$ to $\mathcal{R}''_l$ within $\mathcal{R}_l$, and $\phi(\mathcal{R}_r)$ replaces $\mathcal{R}'_r$ to $\mathcal{R}''_r$ within $\mathcal{R}_r$. Clearly $\mathcal{R}^*_m$, thus $\mathcal{R}_m$, is unchanged. Also, as $\mathcal{R}'_l$, $\mathcal{R}''_l$ have the same size by assumption, by the disjointness above this replacement operation keeps the size of $\mathcal{R}_l$. Moreover, $\mathcal{R}_l$, $\phi(\mathcal{R}_l)$ have the same right distinguished 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, 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 $(\mathcal{R}''_l, \mathcal{R}''_r)$. ◀

The following lemma will be repeatedly used.

▶ **Lemma 5.25** (One round of factorization). *Let L be as (5.19), and Q be any $\binom{[n]}{\le d/2} \times \binom{[n]}{\le d/2}$-matrix with entries*

$$Q(A,B) = \sum_{T_m:\ |V(T_m) \cup A \cup B| \le \tau} (\frac{\omega}{n})^{|V(\mathcal{R}_m)|} q(\mathcal{R}_m) \cdot \chi_{T_m} \tag{5.37}$$

where $\mathcal{R}_m$ denotes $(A,B;T_m)$, and $q(\cdot)$ is a function symmetric w.r.t. shapes.

Define matrix $Q', \mathcal{E}'_{\text{negl}}$ as follows so that

$$(LQL^\top)_{\text{non-can}} = (LQ'L^\top)_{\text{can}} + \mathcal{E}'_{\text{negl}} \tag{5.38}$$

holds. First, let

$$Q'(A,B) = \sum_{T_m:\ |V(T_m) \cup A \cup B| \le \tau} (\frac{\omega}{n})^{|V(\mathcal{R}_m)|} q'(\mathcal{R}_m) \cdot \chi_{T_m} \tag{5.39}$$

*where $q'(\mathcal{R}_m)$ is as follows. Fix any $\mathcal{R}_m = (A,B;T_m)$ and let $t = |V(\mathcal{R}_m)| \le \tau$, $s = \frac{|A|+|B|}{2}$. For every generalized ribbon $\mathcal{R}_m^*$ that contains $\mathcal{R}_m$ as its largest ribbon and $|V(\mathcal{R}_m^*)| \le \tau$, **fix** a ribbon 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 this $\mathcal{R}_m^*$ in the summation below). Then let*

$$\begin{aligned} q'(\mathcal{R}_m) &= \sum_{\substack{\mathcal{R}_m^*:\ \text{gen. ribbon on } (A,B) \\ |V(\mathcal{R}_m^*)| \le \tau \\ \text{largest ribbon is } \mathcal{R}_m}} (\frac{\omega}{n})^{|\mathcal{I}(\mathcal{R}_m^*)|} \cdot q''(\mathcal{R}_m^*), \quad \text{where} \\ q''(\mathcal{R}_m^*) &= \sum_{1 \le z \le d/2} \sum_{\substack{\mathcal{P}=(\mathcal{R}_l,\mathcal{R},\mathcal{R}_r):\ \text{side-inn. can.} \\ \mathcal{P} \to (\mathcal{R}'_l, \mathcal{R}_m^*, \mathcal{R}'_r) \text{ for the fixed } \mathcal{R}'_l, \mathcal{R}'_r \\ z(\mathcal{P}) = z}} (\frac{\omega}{n})^z \cdot q(\mathcal{R}). \end{aligned} \tag{5.40}$$

Note $q'(\mathcal{R}_m)$ doesn't depend on the choice $(\mathcal{R}'_l, \mathcal{R}'_r)$ by Proposition 5.24, and $q'(\cdot)$ is also symmetric w.r.t. shapes. Now define $\mathcal{E}'_{\text{negl}}$ s.t. (5.38) holds.

Then the conclusions are:

(1) *W.p. $> 1 - n^{-9\log n}$ over G, $\left\|\mathcal{E}'_{\text{negl}}\right\| \le \max\{q(A,B;T)\} \cdot n^{-\epsilon\tau}$;*

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

$$\forall \mathcal{R}_m \quad |q(\mathcal{R}_m)| \le C \cdot (\frac{\omega}{n^{1-\epsilon}})^{s-p} \tag{5.41}$$

where p denotes the maximum number of vertex-disjoint paths between A, B in $\mathcal{R}_m$.[20] *Then*

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

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

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

$$\sum_{\substack{(\mathcal{R}_l,\mathcal{R}_m,\mathcal{R}_r):\ \text{side. inn. can.} \\ \text{non-outer-can.} \\ \text{all three have size } \le \tau}} (\frac{\omega}{n})^{|V(\mathcal{R}_l)|+|V(\mathcal{R}_m)|+|V(\mathcal{R}_r)|-|A|-|B|} q(\mathcal{R}_m) \chi_{T_l \oplus T_m \oplus T_r} \tag{5.42}$$

where we denoted the distinguished sets of $\mathcal{R}_m$ by (A,B) when $\mathcal{R}_m$ is given. For each $(\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r)$ in it, 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.

[20] This is also $s_{A,B}(T_m)$ by Menger's theorem; we use p here for appliance with applying Lemma 5.23(2).

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

$$(\frac{\omega}{n})^{|V(\mathcal{R}'_l)|+|V(\mathcal{R}'_m)|+V(\mathcal{R}'_r)|-|S'_l|-|S'_r|} \cdot (\frac{\omega}{n})^{z+|\mathcal{I}(\mathcal{R}^*_m)|} \cdot q(\mathcal{R}'_m)\chi_{T'_l \oplus T^*_m \oplus T'_r} \tag{5.43}$$

in $(LQ'L^\top)_{\text{can}}(I,J)$, where $\mathcal{R}'_m$ denotes the largest ribbon of $\mathcal{R}^*_m$ and $\chi_{T^*_m}$ means the character from $\mathcal{R}'_m$, and $z \ge 1$ is the intersection size of $(\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r)$. Recall for the separating factorization, $T'_l \oplus T^*_m \oplus T'_r = T_l \oplus T_m \oplus T_r$ and

$$\begin{aligned} |V(\mathcal{R}_l) \cup V(\mathcal{R}_m) \cup V(\mathcal{R}_r)| &= |V(\mathcal{R}'_l)| + |V(\mathcal{R}^*_m)| + |V(\mathcal{R}'_r)| - |S'_l| - |S'_r| \\ &= |V(\mathcal{R}_l)| + |V(\mathcal{R}_m)| + |V(\mathcal{R}_r)| - |A| - |B| - z \end{aligned}$$

Also, $|V(R^*_m)| = |V(\mathcal{R}'_m)| + |\mathcal{I}(\mathcal{R}^*_m)|$. Together we have that the coefficient in (5.43) equals the one in (5.42) from $(\mathcal{R}'_l, \mathcal{R}^*_m, \mathcal{R}'_r)$.

Conversely, by definition of Q' and (5.40) and Prop. 5.24 every outer-canonical product in $LQ'L^\top$ corresponds uniquely to a side inner-canonical triple $(\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r)$ in the above case. Therefore, $\mathcal{E}'_{\text{negl}}$ by definition collects all terms in the next case.

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

$$\sum_{\substack{(\mathcal{R}_l,\mathcal{R}_m,\mathcal{R}_r)\text{: side. inn. can.} \\ \text{non-outer-can.} \\ \text{all three has size } \le\tau \\ \text{resulting } |V(\mathcal{R}^*_m)|>\tau}} (\frac{\omega}{n})^{|V(\mathcal{R}_l)|+|V(\mathcal{R}_m)|+|V(\mathcal{R}_r)|-|A|-|B|} q(\mathcal{R}_m)\chi_{T_l \oplus T_m \oplus T_r}. \tag{5.44}$$

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

(1) Take a triple $(\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r)$ in (5.44). Recall

$$\begin{aligned} |V(\mathcal{R}_l)| + |V(\mathcal{R}_m)| + |V(\mathcal{R}_r)| - |A| - |B| &= |V(\mathcal{R}_l) \cup V(\mathcal{R}_m) \cup V(\mathcal{R}_r)| + z \\ &= |V(T) \cup I \cup J| + |\mathcal{I}(\mathcal{R}^*_m)| + z. \end{aligned}$$

Also $|\mathcal{I}(\mathcal{R}^*_m)| \le z + d/2$ as a quick corollary of Lemma 5.23[21]. Fix an $T = T_l \oplus T_m \oplus T_r$ and $a > \tau - |V(T) \cup I \cup J|$, we upper bound the number of triples in (5.44) 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 need to choose a set as $\mathcal{I}(\mathcal{R}^*_m)$ of size $\le a/2 + d/4$ since a is intended to be $|\mathcal{I}(\mathcal{R}^*_m)| + z$ so $a \ge 2\mathcal{I}(\mathcal{R}^*) - d/2$; then to decide the triple over the fixed vertex set there are $< 3^{3\tau} \cdot 2^{3\binom{\tau}{2}}$ many ways. Together, the coefficient of χ_T in (5.44) has absolute value smaller than the following: let $B_0 = \max\{q(\cdot)\}$,

$$\begin{aligned} & B_0 \cdot (\frac{\omega}{n})^{|V(T) \cup I \cup J|+a} \cdot n^{(a+d)/2} 2^{2\tau^2} \\ = {} & B_0 (\frac{\omega}{n^{1-2\epsilon}})^{|V(T) \cup I \cup J|} \left(n^{-2\epsilon}\right)^{|V(T) \cup I \cup J|} \cdot (\frac{\omega}{\sqrt{n}})^a \cdot n^{d/2} 2^{2\tau^2} \\ \le {} & B_0 (n^{-1/2})^{|V(T) \cup I \cup J|} \cdot n^{-2\epsilon(|V(T) \cup I \cup J|+a)} n^{d/2} 2^{2\tau^2} \qquad (\omega \le n^{1/2-4\epsilon}) \\ \le {} & B_0 (n^{-1/2})^{|V(T) \cup I \cup J|} \cdot n^{-1.5\epsilon\tau} \end{aligned}$$

the last step by $|V(T) \cup I \cup J| + a > \tau$ by the case condition and that $d < \epsilon\tau/10$, $2^{2\tau} < n^{\epsilon/10}$. Also, all χ_T appearing in (5.44) has $|V(T)| \le 3\tau$. So by Lemma 4.2, for fixed (I,J), w.p. $> 1 - n^{-10\log n}$

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

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

[21] Actually it can be shown that $|\mathcal{I}(\mathcal{R}^*_m)| \le z$ but we don't need this.

(2) Fix an $\mathcal{R}_m$. By (5.40),

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

For a fixed $\mathcal{R}_m^*$, there are no more than $8^{z\tau} < n^{\epsilon z}$ many triples in the second summation (recall $\mathcal{R}_l', \mathcal{R}_r'$ is fixed), as after fixing whether each vertex appears in each of the three ribbons and fixing $A, B \subseteq \mathcal{R}_m^*$ as distinguished sets of $\mathcal{R}$, we only need to assign possible edges that appear in more than once in the original triple, and it can be checked that such an edge must has at least one end in the already fixed (multi-set) Z of size $\leq z$. Further, by Lemma 5.23(2) and condition (5.41), the second summation in above in absolute value is

$$\begin{aligned}\leq n^{\epsilon z}(\frac{\omega}{n})^{z+|\mathcal{I}(\mathcal{R}_m^*)|}|q(\mathcal{R})| &\leq (\frac{\omega}{n^{1-\epsilon}})^{2(s'-s)+(p-p')+2|\mathcal{I}(\mathcal{R}_m^*)|} \cdot C(\frac{\omega}{n^{1-\epsilon}})^{s-p}\\ &\leq C \cdot (\frac{\omega}{n})^{2|\mathcal{I}(\mathcal{R}_m^*)|} \cdot (\frac{\omega}{n^{1-\epsilon}})^{s'-p'+1/2}\end{aligned}$$

where (s, p) denotes the corresponding parameter for each $\mathcal{R}$ and (s', p') for $\mathcal{R}_m$, and the last step uses $s' - s \geq 1/2$ from Lemma 5.23(1). Finally, in the outer sum, for fixed i_0 there are $< n^{i_0}$ many ways to choose $\mathcal{R}_m^*$ s.t. $|\mathcal{I}(\mathcal{R}_m^*)| = i_0$, and $1 \leq z \leq 3\tau$. So together,

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

Now we can apply Lemma 5.25 to $[L(DQ_0D)L^\top]_{\text{non-can}}$: in (5.30) let $Q \leftarrow (DQ_0D)$, we get

$$[L(DQ_0D)L^\top]_{\text{non-can}} = [L(DQ_1D)L^\top]_{\text{can}} + \mathcal{E}'_{1;\text{negl}}$$

for some Q_1 and $\mathcal{E}'_{1;\text{negl}}$. Then we can repeat this on $[L(DQ_1Q)L^\top]_{\text{non-can}}$ and so on, to get a final **recursive approximate factorization** of M:

$$\begin{aligned}M' = L\Big(D(Q_0 - Q_1 + Q_2 - ... \pm Q_d)D\Big)L^\top - \Big(\mathcal{E}_{1;\text{deg}} - ... \pm \mathcal{E}_{1+d;\text{deg}}\Big)\\ + \Big(\mathcal{E}'_{1;\text{negl}} + ... + \mathcal{E}'_{d;\text{negl}}\Big).\end{aligned} \tag{5.45}$$

Here it implicitly used the following.

▶ **Proposition 5.26** ([5] Claim 6.15). $Q_{d+1} = 0$.

Proof. First we show by induction: $\forall k$, in Q_k every appearing ribbon $R_m = (A, B; T_m)$ has $|A| + |B| \geq k$. Case $k = 0$ is trivial. From k to $k + 1$, by Lemma 5.25 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 the inductive hypothesis $|A| + |B| \geq k$, and by Lemma 5.23(1) $|A'| + |B'| \geq |A| + |B| + 1 \geq k + 1$, and the induction is completed. For $k = 1 + d$, no ribbon can satisfy this while having both distinguished sets in $\binom{[n]}{d/2}$. ◀

We have completed the recursive factorization technique for later use.

▶ Remark 5.27. PSDness of M' would follow from (5.45) by a few last steps[22]. This part is standard, and similar arguments will be given for the exact case (Section 6) so we omit it here.

6 PSDness of the exact pseudo-expectation

Notation. Henceforth M **exclusively** refers to the $d/2$-homogeneous minor of the moment matrix $\widetilde{M}$ in Definition 3.13.

The main theorem of this section is the following.

▶ **Theorem 6.1.** *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 6.2.** *W.p.* $> 1 - n^{-5\log n}$, $\widetilde{E}x_\emptyset > 0$.

Proof. By construction (3.13), $\widetilde{E}x_\emptyset = \frac{\binom{\omega-d/2}{d-d/2}}{\binom{\omega}{d}\binom{d}{d/2}} \sum_{S:|S|=d/2} \widetilde{E}x_S = \frac{\binom{\omega-d/2}{d-d/2}}{\binom{\omega}{d}\binom{d}{d/2}}\mathrm{Tr}(M)$, and by Theorem 6.1 this is positive with high probability. ◀

Theorem 1.4 is a quick corollary of Theorem 6.1: for our pseudo-expectation from Definition 3.13, its moment matrix is PSD by Theorem 6.1 and Lemma 4.1; it satisfies the Default Constraint by Corollary 6.2 and the discussion above Remark 3.9; and it satisfies the Clique and Size Constraints by Lemma 3.8. The degree-d lower bound follows.

The rest of Section 6 is for proving Theorem 6.1. We first reduce it to the main lemma (Lemma 6.9) in the next subsection, then prove that lemma.

6.1 An Hadamard product and Euler transform

For proving Theorem 6.1, we want to factor the matrix M into an $XYX^\top$ form as in the non-exact case. The first problem is that, unlike in the non-exact situation, here in the expression of $M(I,J)$ (Def. 3.13), the appearance of the parameter

$$u = |I \cap J|$$

makes a similar factorization of terms unlikely[23]. As a first step towards resolving this issue, in this subsection, we express M in a $\Sigma\Pi$-form (6.15) where in each leaf matrix, the dependence on u is removed. In later subsections, we will factor each such leaf matrix.

6.1.1 Hadamard product

By definition (3.17), in $M(I,J)$ the coefficient before χ_T can be re-written as

$$M(I,J;T) = \sum_{c=0}^{u}\Bigg[\frac{1}{\binom{\omega-d+u}{u}}\omega^{u-c}\cdot \cdot \underbrace{\left(\binom{a-(d-u)}{c}\binom{n-a}{u-c}n^{-(u-c)}\frac{(a+u-c+8\tau^2)!}{(8\tau^2)!}\left(\frac{\omega}{n}\right)^a\right)}_{:=M_c(u,a)}\Bigg] \tag{6.1}$$

[22] 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.

[23] It doesn't appear in the non-exact case (5.1) at all.

where again $u = |I \cap J|, a = |V(T) \cup I \cup J|$. This means M is a sum of Hadamard products

$$M = \sum_{c=0}^{\frac{d}{2}} m_c \circ M_c \tag{6.2}$$

where m_c, M_c are matrices: for all $|I|, |J| = d/2$,

$$m_c(I, J) = \frac{1}{\binom{\omega-d+u}{u}} \omega^{u-c} \quad u = |I \cap J| \tag{6.3}$$

$$M_c(I, J) = \begin{cases} \sum\limits_{T:|V(T) \cup I \cup J| \leq \tau} \chi_T \cdot M_c(|I \cap J|, |V(T) \cup I \cup J|) & \text{, if } |I \cap J| \geq c; \\ 0 & \text{, o.w.} \end{cases} \tag{6.4}$$

▶ Remark 6.3. It is important to note that we defined m_c to be supported on all (I, J), while let $M_c(I, J) = 0$ if $|I \cap J| < c$, so (6.2) still holds. The use of this is in Lemma 6.4 below.

The intuition behind decomposition (6.2) is 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 proves the second half of the above intuition. The other half will be stated more precisely as the Main Lemma 6.9.

▶ **Lemma 6.4.** *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 (5.4)*, $b_k/k! \in [\frac{d}{2\omega}, 1 + \frac{2dk}{\omega}]$. In particular,*

$$m_0 = \omega m_1 = ... = \omega^{\frac{d}{2}} m_{\frac{d}{2}} \succ \frac{1}{\omega} \text{Id}. \tag{6.5}$$

Proof. By definition, $m_c = \omega^{-c} \sum\limits_{l=0}^{d/2} \frac{\omega^l}{\binom{\omega-d+l}{l}} D_l$, where matrices D_l ($l = 0, ..., d/2$) are the simple basis of Johnson schemes (5.3). By basis-change (5.6), $m_c = \omega^{-c} \sum\limits_{k=0}^{d/2} \mathfrak{J}_k \cdot k! \left(\sum\limits_{l=0}^{k} (-1)^{k-l} \cdot \underbrace{\left[\frac{\omega}{\omega - (d-l)} \cdot ... \cdot \frac{\omega}{\omega - (d-1)} \cdot \frac{1}{(k-l)!} \right]}_{:= f_k(l), \text{ which is } 1/k! \text{ if } l=0} \right)$. For fixed k, $f_k(l)$ is increasing in l so $\sum\limits_{l=0}^{k} (-1)^{k-l} f_k(l) \geq f_k(k) - f_k(k-1) > \frac{d/2}{\omega} \cdot (1 + \frac{d/2}{\omega})^{k-1} \geq \frac{d}{2\omega}$. Note for $k = d/2$, $\mathfrak{J}_{d/2} = \text{Id}$ so we get (6.5). ◀

6.1.2 Euler transform

Fixing c, now we look into the second factor M_c in (6.2). For fixed $(I, J; T)$, again denote $u = |I \cap J|$, $a = |V(T) \cup I \cup J|$. By (6.1)

$$M_c(u, a) = \binom{a - (d-u)}{c} \binom{n-a}{u-c} n^{-(u-c)} \frac{(a + u - c + 8\tau^2)!}{(8\tau^2)!} (\frac{\omega}{n})^a \tag{6.6}$$

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

▶ **Definition 6.5** (Extended $M_c(u,a)$)**.** *For fixed $c \geq 0$, the function $M_c(u,a)$ in (6.6) 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*

$$\binom{n-a}{u-c} = 0 \quad \text{if } u < c, \tag{6.7}$$

and using the usual convention on binomial coefficients

$$\binom{-m}{k} = (-1)^k \cdot \binom{m+k-1}{k} \quad \forall 0 < m, 0 \leq k; \tag{6.8}$$

$$\binom{m}{k} = 0 \quad \forall 0 \leq m < k \tag{6.9}$$

on the expression $M_c(u,a)$ (6.6). We will still use $M_c(u,a)$ to mean this extended function.

In particular, $\binom{m}{0} = 1$ for all $m \in \mathbb{Z}$; if $0 \leq a - (d-u) < c$ then $M_c(u,a) = 0$ since $\binom{a-(d-u)}{c} = 0$.

To further remove the dependence on $u = |I \cap J|$, consider a decomposition

$$M_c = \sum_{R \in \binom{[n]}{\leq \frac{d}{2}}} M_c^R \tag{6.10}$$

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

$$M_c^R(I,J) := \begin{cases} (\frac{\omega}{n})^a \sum\limits_{T:|V(T)\cup I \cup J| \leq \tau} Y_c(|R|, a) \cdot \chi_T & , \text{if } R \subseteq I, J; \\ 0 & , \text{o.w.} \end{cases} \tag{6.11}$$

for some function $Y_c(u,a)$ to be chosen, then comparing for every tuple $(I, J; T)$ we see that equation (6.10) is equivalent to that for any fixed c, a:

$$\sum_{r=0}^{u} \binom{u}{r} Y_c(r,a) (\frac{\omega}{n})^a = M_c(u,a). \tag{6.12}$$

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 6.6.**[24] *If $x(m), y(m)$ are two sequences defined on $\mathbb{N}$ s.t.*

$$\forall m \quad x(m) = \sum_{l=0}^{m} \binom{m}{l} y(l),$$

then $x(m)$ is called the ***Euler transform*** *of $y(m)$, whose inverse transform is*

$$\forall m \quad y(m) = \sum_{l=0}^{m} (-1)^{m-l} \binom{m}{l} x(l).$$

[24] The fact itself can be seen as an application of ζ-matrix and its inverse.

▶ **Definition 6.7** (Coefficients in M_c^R). *For every fixed c, define*

$$Y_c(r,a) = \begin{cases} \sum\limits_{l=c}^{r}(-1)^{r-l}\binom{r}{l}\binom{a+l-d}{c}\binom{n-a}{l-c}n^{-(l-c)}\frac{(a+l-c+8\tau^2)!}{(8\tau^2)!} & \text{, } \textit{if } r \geq c; \\ 0 & \text{, } \textit{o.w.} \end{cases} \tag{6.13}$$

Then as a clear-up summary, we get:

▶ **Lemma 6.8** (The Hadamard-product decomposition of M).

$$M = \sum_{c=0}^{\frac{d}{2}} m_c \circ \left(\sum_{R: R\in\binom{[n]}{\leq d/2}} M_c^R \right) \tag{6.14}$$

$$= \sum_{R\in\binom{[n]}{\leq d/2}} \underbrace{\left(\sum_{c=0}^{|R|} m_c \circ M_c^R \right)}_{:=M^R} \tag{6.15}$$

where each m_c is as in Lemma 6.4 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* $|R| \geq c$ *and* $R \subseteq I\cap J$,

$$M_c^R(I,J) = \sum_{T:|V(T)\cup I\cup J|\leq\tau} M_c^R(I,J;T)\chi_T$$

where, if denote $a = |V(T)\cup I\cup J|$,

$$M_c^R(I,J;T) =$$

$$(\frac{\omega}{n})^a \underbrace{\sum_{l=c}^{|R|}(-1)^{|R|-l}\binom{|R|}{l}\binom{a+l-d}{c}\binom{n-a}{l-c}n^{-(l-c)}\frac{(a+l-c+8\tau^2)!}{(8\tau^2)!}}_{Y_c(|R|,a),\ (6.13)}. \tag{6.16}$$

4. *For all* $0 \leq c \leq r \leq d/2$ *and* $0 \leq a \leq \tau$,

$$|Y_c(r,a)| < \tau^{5\tau}.$$

Proof. (1), (2), (3) is definition. To check (6.14) i.e. $M_c = \sum_R M_c^R$, we check for every $(I,J;T)$ where $|I| = |J| = d/2$, $|V(T)\cup I\cup J| \leq \tau$. Let $u = |I\cap J|$, $a = |V(T)\cup I\cup J|$, then note $a-(d-u) \geq 0$, 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 (6.12), 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 (6.7), (6.9).

For (4),

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

where note $r \leq d/2 \ll \tau$ in our parameter regime. ◀

▶ **Lemma 6.9 (Main Lemma).** *In the decomposition* (6.15)*, 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)

$$M_0^R \ \succeq\ n^{-d}\mathrm{diag}(\widetilde{\mathrm{Cl}})_{P^R \times P^R}; \tag{6.17}$$

(2)

$$\pm\omega^{-c} M_c^R \ \preceq\ n^{-c/6} \cdot M_0^R, \quad \forall 0 < c \le |R|. \tag{6.18}$$

▶ **Corollary 6.10** (Theorem 6.1)**.** *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. For each R, by definition $M^R = \sum_{c=0}^{|R|} m_c \circ M_c^R$. Suppose the situation in Lemma 6.9 happens, which has probability $> 1 - n^{-5\log n}$. Since Hadamard product with a PSD matrix presevres PSDness (the Schur product theorem),

$$\begin{aligned} \sum_{c=1}^{|R|} m_c \circ M_c^R &\preceq \sum_{c=1}^{|R|} m_c \circ \left(\omega^c n^{-c/6} \cdot M_0^R\right) && \text{(Lemma 6.9(2))} \\ &= \left(\sum_{c=1}^{|R|} n^{-c/6} \cdot m_0\right) \circ M_0^R && \text{(Lemma 6.4)} \\ &\preceq n^{-1/6} m_0 \circ M_0^R \end{aligned}$$

Similarly, $\sum_{c=1}^{|R|} m_c \circ M_c^R \ \succeq\ -n^{-1/6} m_0 \circ M_c^R$. So

$$M^R \ \succeq\ (1 - n^{-1/6}) m_0 \circ M_0^R \ \succeq\ n^{-d-1}\mathrm{diag}(\widetilde{\mathrm{Cl}})_{P^R \times P^R} \quad \text{(Lem. 6.4 and 6.9(2))}.$$

Apply this to (6.15),

$$M = M^\emptyset + \sum_{\emptyset \ne R \in \binom{[n]}{\le d/2}} M^R \succeq\ M^\emptyset \succeq\ n^{-d-1}\mathrm{diag}(\widetilde{\mathrm{Cl}})_{\binom{[n]}{d/2}\times\binom{[n]}{d/2}}. \tag{6.19}$$

◀

The rest of Section 6 is devoted to proving the Main Lemma 6.9, completed in Subsection 6.7. The key ingredient is Lemma 6.21, stated in Section 6.4. The statement requires the recursive factorization of each M_c^R, which we show as Lemma 6.19 in the upcoming Subsections 6.2 and 6.3.

6.2 The first-approximate factorization of M_c^R

In this subsection and the next, we factorize each matrix M_c^R in (6.15) by the recursive approximate factorization.

Terminology established in Section 5.3 will be used. We start by defining the first-approximate factorization (cf. Definition 5.17).

▶ **Definition 6.11.** *Fix $R \in \binom{[n]}{\leq \frac{d}{2}}$. For every $i = 0, 1, ..., \tau$ define the left-i-factor $L^{R,i}$ to be the matrix of dimension $\binom{[n]}{\frac{d}{2}} \times \binom{[n]}{\leq \frac{d}{2}}$,*

$$L^{R,i}(I, A) = \begin{cases} 0 & \text{, if } R \not\subseteq I \cap A; \\ \sum\limits_{\substack{T:\ |V(T) \cup I \cup A| \leq \tau \\ A = S_l(I,A;T) \\ T \cap E(A) = \emptyset \\ (I,A;T) \text{ left-generated} \\ e_{I,A}(T) = i}} (\frac{\omega}{n})^i \chi_T & \text{, o.w.} \end{cases} \tag{6.20}$$

*$(L^{R,j})^\top$ is called the right-j-factor. Call $\widetilde{L^R} = (L^{R,0}, ..., L^{R,\tau})$ the **left factor**, $(\widetilde{L^R})^\top$ the **right factor**. Note these matrices do not depend on "c".*

▶ **Definition 6.12.** *Let D^τ denote the constant diagonal matrix*

$$\text{diag}\left((\frac{\omega}{n})^{\frac{|A|}{2}} \right)_{A:|A| \leq d/2} \otimes \text{Id}_{\{0,...,\tau\} \times \{0,...,\tau\}}$$

of dimension $\left(\binom{[n]}{\leq d/2} \times (\tau + 1) \right) \times \left(\binom{[n]}{\leq d/2} \times (\tau + 1) \right)$.

▶ **Definition 6.13** (Goal factorization of M_c^R)**.** *Our goal is to find a middle matrix Q_c^R of dimension*

$$\left(\binom{[n]}{\leq \frac{d}{2}} \times (\tau + 1) \right) \times \left(\binom{[n]}{\leq \frac{d}{2}} \times (\tau + 1) \right)$$

s.t. the following factorization approximately holds:

$$M_c^R \approx \underbrace{(L^{R,0}, ..., L^{R,\tau})}_{\widetilde{L^R}} \cdot \left(D^\tau \cdot Q_c^R \cdot D^\tau \right) \cdot \underbrace{(L^{R,0}, ..., L^{R,\tau})^\top}_{\left(\widetilde{L^R}\right)^\top} \tag{6.21}$$

▶ Remark 6.14. Unlike in the non-exact case (section 5.3), here we factorize M_c^R by further distinguishing a parameter pair in $\{0, ..., \tau\} \times \{0, ..., \tau\}$. The reason is that in (6.13), 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.

Fixing $(I, J; T)$, previously the coefficient (3.12) is intended as

$$(\frac{\omega}{n})^a = (\frac{\omega}{n})^{e(\mathcal{R}_l) + |V(\mathcal{R}_m)| + e(\mathcal{R}_r)}$$

as in Remark 5.12, which naturally factors into the left, middle, right terms. Here, however, there are terms like $\binom{a+l-d}{c} \cdot \binom{n-a}{l-c}$ that are not log-additive in a. Also, the reason we chose the d-generating function as in Def. 3.11 is exactly to prove the positiveness of $\mathbb{E}[Q_{0,0}^R]$ in this harder situation. This is eventually made clear by Prop. 6.28 and Cor. 6.30.

To approach the goal decomposition (6.21), in the coefficients in M_c^R (6.16) we separate the main factor

$$(\frac{\omega}{n})^a = (\frac{\omega}{n})^{e(\mathcal{R}_l)} \cdot (\frac{\omega}{n})^{|V(\mathcal{R}_m)|} \cdot (\frac{\omega}{n})^{e(\mathcal{R}_r)}$$

into left, right, and middle factors as before, while leave the factor $Y_c(r, a)$ for the middle matrix $Q_c^R\Big((\cdot, e_l), (\cdot, e_r)\Big)$ to bear , where the index (e_l, e_r) has the natural intended meaning.

▶ **Definition 6.15** (First-approximate factorization by $Q^R_{c,0}$). *Define $Q^R_{c,0}$ to be the $\{0,...,\tau\} \times \{0,...,\tau\}$-block matrix, each block of dimension $\binom{[n]}{\le d/2} \times \binom{[n]}{\le d/2}$, that is 0 outside of the principal minor*

$$S^R \times S^R, \quad S^R = \{(A,i) \in \binom{[n]}{\le d/2} \times \{0,...,\tau\} \mid A \supseteq R, |A| + i \ge \frac{d}{2}\}, \tag{6.22}$$

and in this principal minor, $Q^R_{c,0}\Big((A,i),(B,j)\Big) =$

$$\sum_{\substack{T_m : |V(T_m)\cup A\cup B| \le \tau \\ A,B \in \mathrm{mSep}_{A,B}(T_m)}} (\frac{\omega}{n})^{|V(T_m)\cup A\cup B| - \frac{|A|+|B|}{2}} \cdot \underbrace{Y_c\big(|R|,\ |V(T_m)\cup A\cup B| + (i+j)\big)}_{\text{defined by (6.13)}} \cdot \chi_{T_m} \tag{6.23}$$

Correspondingly, define

$$\widetilde{L^R} \cdot \left(D^\tau \cdot Q^R_{c,0} \cdot D^\tau\right) \cdot \left(\widetilde{L^R}\right)^\top$$

to be the ***first approximate factorization*** *of M^R_c.*

Some remarks on the definition of $Q^R_{c,0}$ follow.

▶ Remark 6.16 (Intended meaning of parameters in $Q^R_{c,0}$).

(1) The set S^R (6.22) is defined independently of c, where the condition $|A| + i \ge d/2$ is natural because of the intended meaning of i: it is intended as $|V(T')\backslash A| \ge |I| - |A|$ 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) By definition, $Q^R_{c,0}$ is supported only on those $((A,i),(B,j)) \in S^R \times S^R$ with $|A| = |B|$.

(3) Regarding (6.23), as before by Remark 5.12, in "canonical" situations i.e. for outer-canonical products in $\widetilde{L^R} \cdot (D^\tau \cdot Q^R_{c,0} \cdot D^\tau) \cdot \left(\widetilde{L^R}\right)^\top$,

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

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

Recall the terminology on the $XYX^\top$-type matrix product, Def 5.15.

▶ **Lemma 6.17** ($Q^R_{c,0}$ indeed gives the first-approximation). *Fix R, $c \le |R|$. For every $(I, J; T)$ s.t. $|V(T) \cup I \cup J| \le \tau$ and $R \subseteq I \cap J$, there is exactly one outer-canonical product in the $XYX^\top$-type matrix product*

$$\widetilde{L^R} \cdot \underbrace{\left(D^\tau \cdot Q^R_{c,0} \cdot D^\tau\right)}_{Y} \cdot \left(\widetilde{L^R}\right)^\top \tag{6.24}$$

which corresponds to the canonical decomposition of $(I, J; T)$, and which gives term

$$M^R_c(I, J; T)\chi_T.$$

Proof. Suppose $R \subseteq I \cap J$. First, note every triple in (6.24) is inner-canonical by definition of $\widetilde{L^R}, Q^R_{c,0}$, 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 $e_r = e_{J,B}(T'')$. And further there is exactly one outer-canonical product in (6.24) corresponding to this triple, with coefficient

$$L^{R,e_l}(I, A; T') \cdot (\frac{\omega}{n})^{\frac{|A|}{2}} \cdot Q^R_{c,0}(A, B; T_m) \cdot (\frac{\omega}{n})^{\frac{|B|}{2}} \cdot L^{R,e_r}(J, B; T''). \tag{6.25}$$

By definition (6.20), (6.23), if $a := |V(T)| \cup I \cup J \leq \tau$ then the above coefficient is

$$(\frac{\omega}{n})^a \cdot Y_c(|R|, a) = M^R_c(I, J; T),$$

by comparing (6.13) and (6.16), noticing that

$$a = |V(T) \cup I \cup J| \overset{(\#)}{=} e_l + |V(T_m) \cup A \cup B| + e_r,$$

where (#) is by canonicality. This proves the lemma. ◀

▶ **Definition 6.18** (First error-matrices). *Let $\mathcal{E}_{c,1;\text{negl}}$ be the matrix of the sum of all outer-canonical products in (6.24) that exceeds degree, i.e. the resulting*

$$|V(T) \cup I \cup J| > \tau.$$

Let $[\widetilde{L^R} \cdot (D^\tau Q^R_{c,0} D^\tau) \cdot (\widetilde{L^R})^\top]_{\text{non-can}}$ be the matrix of the sum of all products that is non-outer-canonical.

Lemma 6.17 can be restated in the terminology of *approximate form* (Def. 5.15): $\forall R \in \binom{[n]}{\leq d/2}$ and $0 \leq c \leq |R|$,

$$M^R_c = [\widetilde{L^R} \cdot \left(D^\tau Q^R_{c,0} D^\tau\right) \cdot \left(\widetilde{L^R}\right)^\top]_{\text{can}}$$

Equivalently,

$$M^R_c = \widetilde{L^R} \cdot \left(D^\tau Q^R_{c,0} D^\tau\right) \cdot \left(\widetilde{L^R}\right)^\top - [\widetilde{L^R} \cdot \left(D^\tau Q^R_{c,0} D^\tau\right) \cdot \left(\widetilde{L^R}\right)^\top]_{\text{non-can}} - \mathcal{E}^R_{c,1;\text{deg}}. \tag{6.26}$$

As we will see, the crucial fact is that the error matrix $\mathcal{E}_{c,1;\text{main}}$ factorizes through $\widetilde{L^R}, (\widetilde{L^R})^\top$ approximately too, as in the non-exact case. In the next subsection, we show how the recursive factorization method works here in an extended form.

6.3 Recursive factorization: exact case

The main result of this subsection is the following lemma.

▶ **Lemma 6.19** (Recursive approximate factorization; exact case). *For any fixed $R \in \binom{[n]}{\leq d/2}$ and $0 \leq c \leq |R|$, we have the following decomposition.*

$$M^R_c = \widetilde{L^R} \cdot \left[D^\tau \left(Q^R_{c,0} - Q^R_{c,1} + ... \pm Q^R_{c,d}\right) D^\tau\right] \cdot \left(\widetilde{L^R}\right)^\top + \mathcal{E}^R_c, \tag{6.27}$$

where:

(1) *All $Q^R_{c,k}$'s are supported on the principal minor $S^R \times S^R$, where recall*

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

(2) $Q^R_{c,0}$ *is by Definition 6.15;*

(3) $\forall 1 < k \leq d/2$, $Q^R_{c,k}$ *is a* $(\tau+1)\times(\tau+1)$*-block-matrix with the* (i,j)*-block*

$$Q^R_{c,k}\Big((A,i),(B,j)\Big) = \sum_{T_m: |V(T_m)\cup A\cup B|\leq\tau} q^R_{c,k}(\mathcal{R}_m,i,j)\cdot\chi_{T_m} \tag{6.28}$$

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

$$\forall (i,j) \quad |q^R_{c,k}(\mathcal{R}_m,i,j)| \leq \tau^{5\tau}\cdot(\frac{\omega}{n^{1-\epsilon}})^{s-p+k/3}, \tag{6.29}$$

where as usual $s=\frac{|A|+|B|}{2}$*,* p *is the max number of vertex-disjoint paths from* A *to* 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.* $> 1-n^{-9\log n}$*,*

$$\left\|\mathcal{E}^R_c\right\| < n^{-\epsilon\tau/2}. \tag{6.30}$$

Proof of Lemma 6.19 Like before, the key is to look at one round of the factorization. The following lemma is strictly parallel to Lemma 5.25. Again fix $R\subseteq\binom{[n]}{d/2}$, $c\leq|R|$; for convenience denote $n_1:=\binom{[n]}{d/2}\times(\tau+1)$ in the following.

▶ **Lemma 6.20** (One round of factorization; exact case)**.** *Let* $\widetilde{L^R}$ *be from Def. 6.11,* Q^R *be any* $n_1\times n_1$*-matrix supported on* $S^R\times S^R$ *and*

$$Q^R((A,i),(B,j)) = \sum_{T_m:\ |V(T_m)\cup A\cup B|\leq\tau} (\frac{\omega}{n})^{|V(\mathcal{R}_m)|} q(\mathcal{R}_m,i,j)\cdot\chi_{T_m} \tag{6.31}$$

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

$$[\widetilde{L^R}\cdot Q\cdot(\widetilde{L^R})^\top]_{\text{non-can}} = [\widetilde{L^R}\cdot Q'\cdot(\widetilde{L^R})^\top]_{\text{can}} + \mathcal{E}'_{\text{negl}}. \tag{6.32}$$

Namely, let Q' *be supported on* $S^R\times S^R$*,*

$$Q'((A,i),(B,j)) = \sum_{T_m:\ |V(T_m)\cup A\cup B|\leq\tau} (\frac{\omega}{n})^{|V(\mathcal{R}_m)|} q'(\mathcal{R}_m,i,j)\cdot\chi_{T_m} \tag{6.33}$$

where the coefficients $q'(\mathcal{R}_m,i,j)$ *are as follows. Fix any* $\mathcal{R}_m=(A,B;T_m)$ *and* (i,j)*. Let* $t=|V(\mathcal{R}_m)|\leq\tau$*,* $s=\frac{|A|+|B|}{2}$*. For every generalized ribbon* $\mathcal{R}^*_m$ *that contains* $\mathcal{R}_m$ *as its largest ribbon and* $|V(\mathcal{R}^*_m)|\leq\tau$*, 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 for some ribbon triple,* $|V(\mathcal{R}'_l)|,|V(\mathcal{R}'_r)|\leq\tau$ ***and***

$$(e(\mathcal{R}'_l),e(\mathcal{R}'_r))=(i,j). \tag{6.34}$$

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

$$\begin{aligned} q'(\mathcal{R}_m,i,j) &= \sum_{\substack{\mathcal{R}^*_m:\text{ gen. ribbon on }(A,B)\\ |V(\mathcal{R}^*_m)|\leq\tau\\ \text{largest ribbon is }\mathcal{R}_m}} (\frac{\omega}{n})^{|\mathcal{I}(\mathcal{R}^*_m)|}\cdot q''(\mathcal{R}^*_m,i,j) \quad \textit{where}\\ q''(\mathcal{R}^*_m,i,j) &= \sum_{\substack{(z,i_1,j_1):\\ 1\leq z\leq d/2}} \sum_{\substack{\mathcal{P}=(\mathcal{R}_l,\mathcal{R},\mathcal{R}_r):\text{ side-inn. can.}\\ \mathcal{P}\to(\mathcal{R}'_l,\mathcal{R}^*_m,\mathcal{R}'_r)\text{ for the fixed }\mathcal{R}'_l,\mathcal{R}'_r\\ z(\mathcal{P})=z,\ e(\mathcal{R}_l)=i_1,e(\mathcal{R}_r)=j_1}} (\frac{\omega}{n})^z\cdot q(\mathcal{R},i_1,j_1). \end{aligned} \tag{6.35}$$

*Note $q''(\mathcal{R}_m, i, j)$ doesn't depend on the choice $(\mathcal{R}'_l, \mathcal{R}'_r)$ by **(the full of)** Proposition 5.24. Thus $q'(\cdot, i, j)$ is also symmetric w.r.t. shapes.*

$\mathcal{E}'_{\text{negl}}$ is defined s.t. (6.32) holds. Then the conclusions are:

(1) *W.p. $> 1 - n^{-9\log n}$ over G,*

$$\left\|\mathcal{E}'_{\text{negl}}\right\| \le \max\{q(\cdot)\} \cdot n^{-\epsilon\tau};$$

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

$$\forall \mathcal{R}_m, i, j \quad |q(\mathcal{R}_m, i, j)| \le C \cdot (\frac{\omega}{n^{1-\epsilon}})^{s-p} \tag{6.36}$$

where p denotes the maximum 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 6.20. The proof is almost the same as that of Lemma 5.25; 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.

As step (0) like before, 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. 5.24. 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 (6.34) are also fixed, so the proposition is applicable. The comparison for order on $(\frac{\omega}{n})$ between the two is exactly the same as in step (0) of the proof of Lemma 5.25, and the conclusion is that the matrix $\mathcal{E}'_{\text{negl}}$ collects all terms in $[\widetilde{L^R} \cdot Q \cdot (\widetilde{L^R})^\top]_{\text{non-can}}$ whose $\mathcal{R}^*_m$ in the separating factorization exceeds size τ, i.e. $\mathcal{E}'_{\text{negl}}(I, J) =$

$$\sum_{i,j} \sum_{\substack{(\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r)\text{: side. inn. can.} \\ \text{non-outer-can.} \\ \text{all three has size } \le\tau \\ |V(\mathcal{R}^*_m)|>\tau,\ (e(\mathcal{R}_l), e(\mathcal{R}_r))=(i,j)}} (\frac{\omega}{n})^{|V(\mathcal{R}_l)|+|V(\mathcal{R}_m)|+|V(\mathcal{R}_r)|-|A|-|B|} q(\mathcal{R}_m, i, j)\chi_T \tag{6.37}$$

where $T = T_l \oplus T_m \oplus T_r$, and we omit writing the default requirement that $\mathcal{R}_l$ ($\mathcal{R}_r$) has the left (right) distinguished vertex set I (J).

The numerical conclusions (1), (2) follow from the same estimates as in Lemma 5.25 (after (5.44) there). We only point out that, for (1), the estimate there is actually loose enough s.t. with even an extra $(1+\tau)^2$-factor (from union bound on blocks) it is still smaller than $n^{-\epsilon\tau}$. ◀

Now we can prove Lemma 6.19.

Proof for Lemma 6.19. Apply the one-round factorization Lemma 6.20 to

$$[\widetilde{L^R} \cdot \left(D^\tau Q^R_{c,i} D^\tau\right) \cdot \left(\widetilde{L^R}\right)^\top]_{\text{non-can}}$$

for $i = 0$, we get $Q^R_{c,1}$, $\mathcal{E}'_{1;\text{negl}}$ (for ease of notation, we hide the index R, c for this negligible matrix). Then repeat this for $i = 1$ we get $\mathcal{E}_{c,1;\text{deg}}$, $Q^R_{c,2}$, and $\mathcal{E}'_{2,\text{negl}}$. Continuing this, as the result we get the recursive factorization

$$\begin{aligned} M^R_c = &\widetilde{L^R} \cdot \left[D^\tau \left(Q^R_{c,0} - Q^R_{c,1} + ... \pm Q^R_{c,d}\right) D^\tau\right] \cdot \left(\widetilde{L^R}\right)^\top - \\ &\left(\mathcal{E}^R_{c,1;\text{deg}} - \mathcal{E}^R_{c,2;\text{deg}} + ... \pm \mathcal{E}^R_{c,d;\text{deg}}\right) + \left(\mathcal{E}'_{1;\text{negl}} + ... + \mathcal{E}'_{d;\text{negl}}\right). \end{aligned} \tag{6.38}$$

Again, here it uses that $Q^R_{c,d+1} = 0$, by the same proposition 5.26.

(1) All $Q^R_{c,k}$ is supported within $S^R \times S^R$ by definition of each round (Lemma 6.20);

(2) By definition.

(3) The coefficients of each $Q^R_{c,k}$ $(k = 0, 1, ..., d)$, $\{q^R_{c,k}(\cdot, i, j)\}$ is always symmetric w.r.t. shapes from Lemma 6.20. Moreover, from definition (6.23),

$$\forall \mathcal{R}_m, i, j \quad |q^R_{c,0}(\mathcal{R}_m)| = |Y_c(|R|, |\mathcal{R}_m|)| \le \tau^{5\tau}$$

where the last one is by Lemma 6.8(4). Since $Q^R_{c,0}$ is special in that for all $\mathcal{R}_m = (A, B; T_m)$ appearing in it, there are $|A| = |B|$ many vertex-disjoint paths between A, B in $\mathcal{R}_m$, i.e. $s = p$, where as usual when $\mathcal{R}_m$ is fixed we use $s = \frac{|A|+|B|}{2}$ and p denotes the max number of vertex-disjoint paths between A, B. So the above can be equivalently written as

$$\forall \mathcal{R}_m, i, j \quad |q^R_{c,0}(\mathcal{R}_m)| \le (\frac{\omega}{n^{1-\epsilon}})^{s-p}\tau^{5\tau}. \tag{6.39}$$

Now use Lemma 6.20(2), where notice the "$q(\cdot)$" in there corresponds to $q^R_{c,k}$ here, since the "Q" matrix is $D^\tau Q^R_{c,k} D$ so 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, we get the recursive bound

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

(4) First, when plugged in any G, both

$$M^R_c \quad \text{and} \quad \widetilde{L^R} \cdot \left[D^\tau \left(Q^R_{c,0} - Q^R_{c,1} + ... \pm Q^R_{c,d}\right) D^\tau\right] \cdot \left(\widetilde{L^R}\right)^\top$$

are supported within clique rows and columns that contain R by their definition. So it must be the case for their difference, $\mathcal{E}^R_c$, too. Next we only need to give the norm bound. By (6.38), the final error matrix is

$$\mathcal{E}^R_c = -\left(\mathcal{E}^R_{c,1;\deg} - \mathcal{E}^R_{c,2;\deg} + ... \pm \mathcal{E}^R_{c,d;\deg}\right) + \left(\mathcal{E}'_{1;\text{negl}} + ... + \mathcal{E}'_{d;\text{negl}}\right).$$

Note by Lemma 6.20(2), by induction all $|q^R_{c,k}| < \tau^{5\tau}$. For each $\mathcal{E}'_{k;\text{negl}}$, by Lemma 6.20(1) w.p. $> 1 - n^{-9\log n}$, $\left\|\mathcal{E}'_{k;\text{negl}}\right\| < \tau^{5\tau}n^{-\epsilon\tau} < n^{-0.9\epsilon\tau}$.

As for $\mathcal{E}^R_{c,k;\deg}$, recall by definition 5.15 on (I, J) it is the sum of outer-canonical products in $\widetilde{L^R} \cdot \left(D^\tau Q^R_{c,i-1} D^\tau\right) \cdot \left(\widetilde{L^R}\right)^\top (I, J)$ s.t. $|V(T) \cup I \cup J| > \tau$. So

$$\mathcal{E}^R_{c,k;\deg}(I, J) = \sum_{\substack{(\mathcal{R}_l, \mathcal{R}_m, \mathcal{R}_r):\\ \text{semi-inn.can.}\\ \text{outer.can.}\\ |V(T)\cup I\cup J| > \tau}} (\frac{\omega}{n})^{|V(T)\cup I\cup J|} \cdot q^R_{c,k-1}(\mathcal{R}_m, e(\mathcal{R}_l), e(\mathcal{R}_r))\chi_T$$

where as usual $s = s(\mathcal{R}_m)$ is the average of its two side vertex-sets, $T = T_l \oplus T_m \oplus T_r$, and in the summation R_l (R_r) should have I (J) as the left (right) set. Note the above uses $|V(T) \cup I \cup J| = e_l + e_r + |V(\mathcal{R}_m)|$ from the outer- and semi-inner- canonicality. Moreover, any fixed $(I, J; T)$ can come from at most $3^{3\tau}$ triples as their vertex set union is $|V(T) \cup I \cup J|$ by canonicality. Since $3\tau \ge |V(T) \cup I \cup J| > \tau$ and w.h.p. $|q^R_{c,k-1}(\cdot)| < \tau^{5\tau}$, use Lemma 4.2 and we get that w.p. $> 1 - n^{-10\log n}$,

$$\left|\mathcal{E}^R_{c,k;\deg}(I, J)\right| < \tau^{6\tau} \cdot \sum_{c=0}^{3\tau} (\frac{\omega}{n})^{\max\{\tau, c\}} \cdot (n^{c/2} 2^{c^2} n^{4\log\log n}) < n^{-2\epsilon\tau}.$$

So by union bound over (I, J), $\left\|\mathcal{E}^R_{c,k;\deg}\right\| < n^{-d/4}n^{-2\epsilon\tau} < n^{-\epsilon\tau}$ w.p. $> 1 - n^{-9.5\log n}$.

Together, sum the two and by union bound over k, we get that w.p. $> 1 - n^{-9\log n}$, $\left\|\mathcal{E}^R_c\right\| < n^{-\epsilon\tau/2}$. ◀

6.4 Positiveness of the middle matrices: proof overview

Now we use the approximate decomposition of M_c^R's to prove the Main Lemma 6.9. Recall for each R, $c \le |R|$, by Lemma 6.19

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

The key is the following lemma. Recall $S^R = \{(A,i) \in \binom{[n]}{\le d/2} \times \{0, ..., \tau\} \mid A \supseteq R, |A|+i \ge \frac{d}{2}\}$.

▶ **Lemma 6.21.** *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) $\forall R$, $0 < c \le |R|$

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

The proof of Lemma will span the upcoming three subsections, completed at the end of Section 6.6. The Main Lemma 6.9 then follows by standard steps (Section 6.7).

Proof plan for Lemma 6.21. Fix an $R \in \binom{[n]}{\le d/2}$. We will prove the lemma by three ingredients: Corollary 6.36, Lemma 6.37, Lemma 6.38.

Corollary 6.36 (in Section 6.5, 6.6): 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. 6.23), 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 3.11.

Lemma 6.37, 6.38 (Section 6.6): Other $Q_{c,k}^R$'s ($k > 0$ or $c > 0$), when timed with ω^{-c}, are small and appropriately supported. These two lemmas are proved by standard means.

We will follow this plan in the next two subsections. Here we end this subsection with two definitions for preparation.

▶ **Definition 6.22.** *Let the* ***root diagonal-clique matrix*** *be*

$$D_{\mathrm{Cl}}(A,B) = \begin{cases} 0 & \text{, if } A \neq B; \\ 2^{-\binom{|A|}{2}/2} \cdot \widetilde{\mathrm{Cl}}_A = 2^{-\binom{|A|}{2}/2} \sum_{T \subseteq E[A]} \chi_T & \text{, o.w.} \end{cases} \tag{6.40}$$

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* $A \in \binom{[n]}{d/2}$*. Define*

$$D_{\mathrm{Cl}}^\tau := D_{\mathrm{Cl}} \otimes \mathrm{Id}_{\{0,\ldots,\tau\}\times\{0,\ldots,\tau\}}. \tag{6.41}$$

which is also diagonal.

▶ **Definition 6.23.** *The* ***structural-pseudorandom decomposition*** *of* $Q_{0,0}^R$ *is*

$$Q_{0,0}^R = D_{\mathrm{Cl}}^\tau \cdot \mathbb{E}[Q_{0,0}^R] \cdot D_{Cl}^\tau + \left(Q_{0,0}^R - D_{\mathrm{Cl}}^\tau \cdot \mathbb{E}[Q_{0,0}^R] \cdot D_{Cl}^\tau\right), \tag{6.42}$$

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****.*

6.5 Positiveness of $\mathbb{E}[Q^R_{0,0}]$

▶ **Proposition 6.24** (Expression of $\mathbb{E}[Q^R_{c,0}]$). *Fix $R \in \binom{[n]}{\le d/2}$ and $0 \le c \le |R|$. let $r = |R|$. Recall S^R is defined by (6.22).*

(1) *$\mathbb{E}[Q^R_{c,0}]$ is supported on the blockwise partial-diagonals*

$$\{\Big((A,i),(A,j)\Big) \in S^R \times S^R\}.$$

(i.e. requires $R \subseteq A$ and $|A| + \min\{i,j\} \ge d/2$)

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

$$\sum_{l=c}^{r} (-1)^{r-l} \frac{\binom{r}{l}}{(l-c)!} \binom{|A|+i+j+l-d}{c} \frac{\Big(|A| + 8\tau^2 + (l-c) + (i+j)\Big)!}{(8\tau^2)!} + O\left(\frac{\tau^{1.5\tau}}{n}\right). \tag{6.43}$$

In particular, for $c = 0$,

$$\mathbb{E}[Q^R_{0,0}]\Big((A,i),(A,j)\Big) = \sum_{l=0}^{r} (-1)^{r-l} \frac{\binom{r}{l}}{l!} \cdot \frac{\Big(|A| + 8\tau^2 + l + (i+j)\Big)!}{(8\tau^2)!} + O\left(\frac{\tau^{1.5\tau}}{n}\right). \tag{6.44}$$

(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*

$$\mathbb{E}[Q^R_{0,0}] = \sum_{\substack{A \subseteq \binom{[n]}{\le d/2} \\ A \supseteq R}} 1_{A,A} \otimes \left[\left(\sum_{l=0}^{r} (-1)^{r-l} \frac{\binom{r}{l}}{l!} \cdot P_{|A|+l}\right) + E^R_A\right] \tag{6.45}$$

where, for every fixed A, $P_{|A|+l} and E^R_A$ are $(\tau+1) \times (\tau+1)$-matrices both supported on the principal minor $\{i \mid d/2 - |A| \le i \le \tau\} \times \{i \mid d/2 - |A| \le i \le \tau\}$ with the following property:

$$\left\|E^R_A\right\| < \frac{\tau^{2\tau}}{n}, \tag{6.46}$$

and

$$P_{|A|+l}(i,j) = \frac{\Big(|A| + l + 8\tau^2 + (i+j)\Big)!}{(8\tau^2)!}, \quad d/2 - |A| \le i,j \le \tau. \tag{6.47}$$

Proof. For (1), the constant terms in (6.23) correspond to $T_m = \emptyset$, which is nonzero only when $A = B$ for A, B in S^R.

For (2), from definition (6.23) we 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:

$$\sum_{l=c}^{r}(-1)^{r-l}\binom{r}{l}\underbrace{\binom{a+l-d}{c}}_{\text{Def. 6.5}}\binom{n-a}{l-c}n^{-(l-c)}\frac{(a+l-c+8\tau^2)!}{(8\tau^2)!}. \tag{6.48}$$

Now use

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

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 (6.48), we get (6.43). Further, in (6.48) 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 (6.44).

For (3), each E_A^R has dimension $(\tau + 1) \times (\tau + 1)$ and each entry is absolutely $< \tau^{1.5\tau}/n$ from part (2). ◀

▶ Remark 6.25 (Specialty of $c = 0$). Comparing $\mathbb{E}[Q_{0,0}^R]$ and $\mathbb{E}[Q_{c,0}^R]$ (6.43), (6.44), the specialty of the case $c = 0$ is that the factor $\binom{|A|+l-d}{0}$ is **always** 1, which is important for $\mathbb{E}[Q_{0,0}^R]$ 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 $\mathbb{E}[Q_{c,0}^R]$ possibly not PSD.

▶ **Definition 6.26.** *For every $m, t \in \mathbb{N}$, define the* ***factorial Hankel matrix*** *to be*

$$H_{m,t}(i,j) = (i + j + t)! \quad \forall 0 \le i, j \le m. \tag{6.49}$$

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

▶ **Proposition 6.27** (Almost common decomposition of $\{H_{m,t}\}$).

(1) *The matrix family $\{H_{m,t}\}$ have decomposition*

$$H_{m,t} = L_m \cdot \left(N_{m,t} \cdot D_{m,t} \cdot (N_{m,t})^\top\right) \cdot (L_m^\top)$$

where $L_m, D_{m,t}$ are diagonal, $N_{m,t}$ is lower-triangular

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

In particular, L_m is independent of t, and $H_{m,t}$ is positive.

(2) *Let J_m denote the $(1 + m) \times (1 + m)$ 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

$$N_{m,t+1} = N_{m,t} \cdot J_m. \tag{6.50}$$

Proof. This follow from a direct inspection. ◀

▶ **Proposition 6.28.** *If parameters m, t, r satisfy*

$$t+1 > 8 \cdot \max\{r^2, m\} \tag{6.51}$$

then

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

Proof. By Proposition 6.27 it suffices to show that under (6.51),

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

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

$$q_{t+1}(x) := (x^\top J_m) D_{m,t+1} (J_m^\top x) \quad \text{v.s.} \quad q_t(x) := 2r^2 \cdot x^\top D_{m,t} x \tag{6.52}$$

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

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

By definition of $D_{m,t}$, J_m,

$$q_{t+1}(x) = \sum_{i=0}^{m} \beta(i)(x_i + x_{i+1})^2, \quad x_{m+1} := 0;$$
$$q_t(x) = \sum_{i=0}^{m} \alpha(i) x_i^2.$$

To compare the two, 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 $0 \le i \le m$ let

$$b_i = 1 - \frac{\alpha(i)}{\beta(i)} - \frac{\beta(i-1)}{\beta(i)} \frac{1}{b_{i-1}}, \quad b_0 = 1 - \frac{\alpha(0)}{\beta(0)}, \tag{6.53}$$

then

$$q_{t+1}(x) = \underbrace{\sum_{i=0}^{m} \alpha(i) x_i^2}_{q_t(x)} \; + \; \sum_{i=0}^{m} \beta(i) b_i \cdot (x_i + \frac{1}{b_i} x_{i+1})^2. \tag{6.54}$$

▷ Claim 6.29. In (6.53), for all $i \le m$ we have $b_i > 1/2$.

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

$$b_i = 1 - \frac{2r^2}{(t+1+i)} - \frac{i}{(t+1+i)} \cdot \frac{1}{b_{i-1}}, \quad i \ge 1. \tag{6.55}$$

Use induction for the claim: $b_0 = 1 - \frac{2r^2}{t+1} > 1/2$ by (6.51). For $1 \le i \le m$,

$$\begin{aligned} b_i &= 1 - \frac{2r^2}{t+1+i} - \frac{i}{t+1+i} \cdot \frac{1}{b_{i-1}} \\ &\ge 1 - \frac{2r^2}{t+1} - \frac{m}{t+1} \cdot 2 > 1/2 \quad \text{by (6.51) and the inductive hypothesis.} \end{aligned}$$

◁

By (6.54) and positiveness of each b_i (Claim 6.29), $q_{t+1}(x) \ge q_t(x)$. The lemma is proved. ◀

Now we apply Proposition 6.28 to matrices $P_{|A|+l}$ (6.47). 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; below, we regard $P_{|A|+l}$ as a matrix on its support.

▶ **Corollary 6.30** (Positiveness of $\mathbb{E}[Q^R_{0,0}]$). *In the decomposition* (6.45) *of* $\mathbb{E}[Q^R_{0,0}]$,

$$\left(\sum_{l=0}^{r} (-1)^{r-l} \frac{\binom{r}{l}}{l!} \cdot P_{|A|+l}\right) + E^R_A \succ \operatorname{diag}\left(\tau^{-6\tau}\right)_{0\le i\le \tau-(d/2-|A|)} \tag{6.56}$$

where we naturally regarded matrices as on their support

$$\{i \mid d/2 - |A| \le i \le \tau)\}^2 \cong \{0, ..., \tau - (d/2 - |A|)\}^2.$$

In particular, by (6.45)

$$\mathbb{E}[Q^R_{0,0}] \succ \sum_{\substack{A \subseteq \binom{[n]}{\le d/2} \\ A \supseteq R}} 1_{A,A} \otimes \operatorname{diag}\left(\tau^{-6\tau}\right)_{d/2-|A|\le i\le\tau} = \operatorname{diag}\left(\tau^{-6\tau}\right)_{S^R\times S^R} \tag{6.57}$$

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

Proof. The "in particular" part is straightforward from (6.56) by checking the support, and that tensoring with a nonzero PSD matrix preserves the relation $\succ$. In below we prove for (6.56).

Fix A, let

$$\tau_0 = \tau - (d/2 - |A|), \quad t_0 = d - |A| + 8\tau^2. \tag{6.58}$$

Then

$$\sum_{l=0}^{r} (-1)^{r-l} \frac{\binom{r}{l}}{l!} \cdot P_{|A|+l} = \frac{1}{(8\tau^2)!} \cdot (X_r + X_{r-2} + ...) \tag{6.59}$$

where, $\forall 0 \le v \le \lfloor r/2 \rfloor$,

$$X_{r-2v} = \frac{\binom{r}{r-2v}}{(r-2v)!} \cdot \left(H_{\tau_0,t_0+r-2v} - \underbrace{\frac{(r-2v)^2}{(2v+1)}}_{\le r^2} H_{\tau_0,t_0+r-2v-1}\right), \quad H_{\tau_0,-1} := 0.$$

Since $t_0 > 8\max\{r^2, \tau_0\}$, by Proposition 6.28

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

So in (6.59), in particular,

$$\sum_{l=0}^{r}(-1)^{r-l}\frac{\binom{r}{l}}{l!}\cdot P_{|A|+l} \;\succeq\; \frac{1}{(8\tau^2)!}\cdot H_{\tau_0,t_0} \overset{\text{Prop. 6.27}}{=} L\left(N_{t_0}\cdot\frac{D_{t_0}}{(8\tau^2)!}\cdot(N_{t_0})^{\top}\right)L \tag{6.60}$$

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 (6.56):

$$\begin{aligned}\text{RHS of (6.60) } &\succ L\cdot\operatorname{diag}\left(\tau^{-5\tau}\right)_{0\le i\le\tau_0}\cdot L \qquad \text{(by Claim 6.31)}\\ &\succeq \operatorname{diag}\left(\tau^{-5\tau}\right)_{0\le i\le\tau_0},\end{aligned}$$

while by Proposition 6.24 (3),

$$\left\|E_A^R\right\| < \frac{\tau^{2\tau}}{n} < \tau^{-6\tau} \qquad \text{(parameter regime).}$$

So LHS of (6.56) $\succeq$ diag $\left(\tau^{-5\tau}-\tau^{-6\tau}\right)_{0\le i\le\tau_0}$ $\succeq$ RHS of (6.56). ◀

▷ Claim 6.31. In notation of Corollary 6.30,

$$N_{t_0}^{-1}(i,j) = (-1)^{i-j}\binom{i+t_0}{i-j} \qquad 0\le i,j\le\tau_0 \tag{6.61}$$

and

$$N_{t_0}\cdot\frac{D_{t_0}}{(8\tau^2)!}\cdot(N_{t_0})^{\top} \succ \operatorname{diag}\left(\tau^{-5\tau}\right)_{0\le i\le\tau_0}. \tag{6.62}$$

Proof. For (6.61), 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 this is identity matrix, use generating functions: 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 is

$$\begin{aligned}&(-1)^{i'}\sum_{k'=0}^{i'}D_{i'-k'}[(1+x)^{i'+j+t_0}]\cdot D_{k'}[(1+x)^{-(t_0+j+1)}]\\ =&(-1)^{i'}D_{i'}[(1+x)^{i'+j+t_0-(t_0+j+1)}] = (-1)^{i'}D_{i'}[(1+x)^{i'-1}] = 1_{i'=0}.\end{aligned}$$

For (6.62), it is equivalent to

$$\frac{D_{t_0}}{(8\tau^2)!} \succ N_{t_0}^{-1}\cdot\tau^{-5\tau}\cdot(N_{t_0}^{-1})^{\top}. \tag{6.63}$$

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

$$x^{\top}N_{t_0}^{-1}\cdot a_0\cdot(N_{t_0}^{-1})^{\top}x = a_0\sum_{j=0}^{\tau_0}y_j^2, \tag{6.64}$$

where by (6.61),

$$y_j = \left(x^{\top}N_{t_0}^{-1}\right)_j = \sum_{i=j}^{\tau_0}(-1)^{i-j}\binom{i+t_0}{i-j}x_i.$$

By Cauchy-Schwartz, $y_j^2 \leq \tau_0 \cdot \sum_{i=j}^{\tau_0} \binom{i+t_0}{i-j}^2 x_i^2$, so

$$\text{RHS of (6.64)} = a_0 \sum_{j=0}^{\tau_0} y_j^2 \ \leq \ a_0 \sum_{i=0}^{\tau_0} x_i^2 \cdot \left(\tau_0 \sum_{j=0}^{i} \binom{i+t_0}{i-j}^2 \right)$$
$$< \sum_{i=0}^{\tau_0} \left(\tau^{-5\tau} \cdot (9\tau^2)^{2i+2} \right) x_i^2.$$

Now (6.63) follows since for each i, in the LHS of (6.63)

$$\frac{D_{t_0}(i,i)}{(8\tau^2)!} \geq (8\tau^2)^{-(d/2-|A|)} \qquad \text{(by definition)}$$
$$> \tau^{-2d} > \tau^{-5\tau} \cdot (9\tau^2)^{2i+2}$$

using $i \leq \tau_0 < \tau$, $d \ll \tau$. So (6.63) holds. ◁

We get the main conclusion of this subsection:

▶ **Corollary 6.32** (Positiveness of the structural part of $Q_{0,0}^R$ (Def. 6.23)).

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

Proof. This follows from Cor. 6.30 and that $D_{\text{Cl}}^2(A,A) = \widetilde{\text{Cl}}(A)$ for all A in Def. 6.22. ◀

6.6 Rest bounds: $Q_{c,k}^R$s

In this subsection, we bound the rest matrices:

$$\underbrace{Q_{0,0}^R - D_{\text{Cl}}^\tau \cdot \mathbb{E}[Q_{0,0}^R] \cdot D_{\text{Cl}}^\tau}_{\text{pseudo}-\text{random part of } \mathrm{Q}_{0,0}^{\mathrm{R}} \text{ (Def. 6.23)}} \quad , \quad Q_{0,k}^R \ (k > 0), \quad \omega^{-c} \cdot Q_{c,k}^R \ (c > 0, k \geq 0)$$

by three Lemmas 6.34, 6.37, 6.38, respectively, which would prove Lemma 6.21.

The arguments are quite standard but somewhat lengthy, as one needs to be careful on the block structure and the support of the matrices.

▶ **Definition 6.33** (0-1 diagonal-clique matrix). *Recall the matrix D_{Cl}^τ from Def. 6.22. Denote by D' its 0-1 valued version, i.e. D' is also diagonal and has entries*

$$D'((A,i),(A,i)) = \text{Cl}_A, \quad \forall A \in \binom{[n]}{\leq d/2} \ \forall 0 \leq i \leq \tau.$$

▶ **Lemma 6.34** (Bound on pseudo-random part of $Q_{0,0}^R$). *W.p. $> 1 - n^{-9\log n}$ the following holds: $\forall R \in \binom{[n]}{\leq d/2}$,*

$$\pm(\underbrace{Q_{0,0}^R - D_{\text{Cl}}^\tau \cdot \mathbb{E}[Q_{0,0}^R] \cdot D_{\text{Cl}}^\tau}_{\text{pseudo}-\text{random part of } Q_{0,0}^R})(G) \ \preceq \ n^{-\epsilon} \cdot \text{diag}\left(\widetilde{\text{Cl}}(G)\right)_{S^R \times S^R} \qquad (6.65)$$

Proof. Fix *R*. For simplicity, in this proof abbreviate:

$$Q_{\text{ps}} := Q_{0,0}^R - D_{\text{Cl}}^\tau \cdot \mathbb{E}[Q_{0,0}^R] \cdot D_{\text{Cl}}^\tau = \left(Q_{\text{ps},(i,j)}\right)_{0 \leq i,j \leq \tau}$$

("ps" for pseudo-random), which is a $(\tau+1) \times (\tau+1)$-block matrix.

In block (i,j), by Def. 6.15 and Prop. 6.24, $Q_{\mathrm{ps},(i,j)}$ is supported within

$$S_{i,j} \times S_{i,j} \quad \text{where} \quad S_{i,j} := \{A \mid |A| + \min\{i,j\} \geq d/2\}.$$

And for each $A \neq B$,

$$Q_{\mathrm{ps},(i,j)}(A,B) = \ Q^R_{0,0}((A,i),(B,j)) = \sum_{\substack{T_m\colon\ |V(T_m)\cup A\cup B| \leq \tau \\ A,B \in \mathrm{mSep}_{A,B}(T_m)}} (\frac{\omega}{n})^{|V(T_m)\cup A\cup B| - \frac{|A|+|B|}{2}} \cdot q(A,B;T_m) \cdot \chi_{T_m}; \tag{6.66}$$

and

$$Q_{\mathrm{ps},(i,j)}(A,A) = \sum_{T_m\colon\ 1 \leq |V(T_m)\setminus A| \leq \tau - |A|} (\frac{\omega}{n})^{|V(T_m)\cup A| - |A|} \cdot q(A,A;T_m) \cdot \chi_{T_m}. \tag{6.67}$$

Here we have abbreviated $q(A,B;T_m) := Y_0\Big(|R|, |V(T_m)\cup A\cup B| + (i+j)\Big)$ ((6.23)) and have omitted the indices $|R|, i+j$ when they are fixed. Two properties we need:

$$q(A,B;T_m) \text{ depends only on } |V(T_m)\cup A\cup B| \text{ when fixing } (A,B); \tag{6.68}$$

$$\Big|q(A,B;T_m)\Big| < \tau^{5\tau} \qquad \text{(by Lemma 6.8 (4))}. \tag{6.69}$$

By (6.68), $Q_{\mathrm{ps},(i,j)}(A,B)$ always factors through $\mathrm{Cl}_{A\cup B}$ and so also through $\mathrm{Cl}_A\mathrm{Cl}_B$. In particular,

$$Q_{\mathrm{ps}} = D' \cdot Q_{\mathrm{ps}} \cdot D' \tag{6.70}$$

where D' is the 0-1 diagonal-clique matrix (Definition 6.33).

▷ Claim 6.35. W.p. $> 1 - n^{-9.5\log n}$ the following holds:

$$\forall (i,j) \qquad \pm Q_{\mathrm{ps},(i,j)} \ \prec \ n^{-1.1\epsilon} \cdot \mathrm{diag}\left(2^{\binom{|A|}{2}}\right)_{S^R_{\min\{i,j\}} \times S^R_{\min\{i,j\}}}$$

where $S^R_a := \{A \in \binom{[n]}{\leq d/2} \mid |A| + a \geq d/2\}$.

The lemma follows from this claim and (6.70). Namely, consider a different decomposition of Q_{ps} as follows. For every $b \in [0, \frac{d}{2}]$, let

$$I_b := \{i \mid d/2 - b \leq i \leq \tau\}$$

and $Q_{\mathrm{ps};b}$ be the principal minor $W_b := \big(P^R_b \times I_b\big) \times \big(P^R_b \times I_b\big)$ of Q_{ps} (0 elsewhere), where $P^R_b = \{A \subseteq [n] \mid R \subseteq A, |A| = b\}$. Then we have

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

Since $Q^R_{c,0}$ is supported only on those $((A,i),(B,j)) \in S^R \times S^R$ with $|A| = |B|$ (Remark 6.16(2)), in particular for $c = 0$ we have

$$Q_{ps} = \sum_{b=0}^{d/2} Q_{\mathrm{ps};b}. \tag{6.71}$$

Each $Q_{\mathrm{ps};b}$ is block-wise in blocks $I_b \times I_b$, each block a principal minor of $Q_{\mathrm{ps},(i,j)}$. So by Claim 6.35 w.p. $> 1 - n^{-9.5\log n}$ any ($\pm$) such a block $\prec n^{-1.5\epsilon} \cdot \mathrm{diag}\left(2^{\binom{b}{2}}\right)_{\binom{[n]}{b}\times\binom{[n]}{b}}$, so $\pm Q_{\mathrm{ps};b} \prec \tau^2 \cdot n^{-1.5\epsilon}\mathrm{diag}\left(2^{\binom{b}{2}}\right)_{W_b} \prec n^{-\epsilon}\mathrm{diag}\left(2^{\binom{b}{2}}\right)_{W_b}$. Hence by (6.71) and the union bound over b, $\pm Q_{\mathrm{ps}} \prec n^{-\epsilon}\mathrm{diag}\left(2^{\binom{|A|}{2}}\right)_{S^R\times S^R}$ w.p. $1 - n^{-9\log n}$. Finally, insert this to the middle of (6.70), where notice $\widetilde{\mathrm{Cl}}_A = 2^{\binom{|A|}{2}} \cdot \mathrm{Cl}_A$, $\mathrm{Cl}_A = \mathrm{Cl}_A^2$, we get (6.65). ◀

Proof of Claim 6.35. We use the norm bounds from Section 4. Fix (i,j), consider consider

$$Q^{\mathrm{diag}}_{\mathrm{ps},(i,j)} \quad \text{and} \quad Q^{\mathrm{off}}_{\mathrm{ps},(i,j)} = Q_{\mathrm{ps},(i,j)} - Q^{\mathrm{diag}}_{\mathrm{ps},(i,j)}.$$

Diagonal part. For $Q^{\mathrm{diag}}_{\mathrm{ps},(i,j)}$, by (6.67) for any (A,A) in the support (i.e. $|A| + i \geq d/2$, $|A| + j \geq d/2$),

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

For every fixed A in support, this $g(A)$ can be bounded by norms of diagonal graphical matrices, as follows. First, $q(A,A;T_m)$ depends only on $|V(T_m)\setminus A|$ (we have fixed R, i, j, A), so temporarily denote it as $q(|V(T_m)\setminus A|)$. For every $1 \leq v \leq \tau - |A|$, let $\mathcal{U}^v_1, ..., \mathcal{U}^v_{h(v)}$ be all different shapes $(A,A;T)$ (Def. 4.7) s.t. $T \cap E[A] = \emptyset$ and $|V(T)\setminus A| = v$. Clearly,

$$h(v) \leq 2^{|A|v+v^2} \quad \text{since we required } T\cap E[A] = \emptyset. \tag{6.72}$$

So w.p. $> 1 - n^{-9.6\log n}$ the following holds:

$$\begin{aligned}
|g(A)| &= \left| \sum_{v=1}^{\tau-|A|} (\frac{\omega}{n})^v q(v) \cdot \left(\sum_{x=1}^{h(v)} \underbrace{\sum_{\substack{T_m:(A,A;T_m)\text{ has} \\ \text{shape } \mathcal{U}^v_x}} \chi_{T_m}}_{=M_{\mathcal{U}^v_x}(A,A)\text{ by Def. 4.7}} \right)\right| \\
&\leq \sum_{v=1}^{\tau-|A|} (\frac{\omega}{n})^v q(v) \cdot \sum_{x=1}^{h(v)} \left\|M_{\mathcal{U}^v_x}\right\| \qquad \text{(each } M_{\mathcal{U}^v_x} \text{ is diag.)} \\
&\leq \sum_{v=1}^{\tau-|A|} (\frac{\omega}{n})^v \tau^{5\tau} \sum_{x=1}^{h(v)} \left\|M_{\mathcal{U}^v_x}\right\| \qquad \text{(by (6.69))} \\
&< \sum_{v=1}^{\tau} (\frac{\omega}{n})^v \tau^{5\tau} \cdot 2^{|A|v+v^2} \cdot n^{\frac{v}{2}} 2^{O(|A|+v)} \quad \text{(by (6.72) and Thm. 4.8)} \\
&< \sum_{v=1}^{\tau} n^{-3\epsilon v} \cdot n^{\epsilon v} < n^{-1.2\epsilon} \qquad \text{(by the parameter regime)}
\end{aligned}$$

Off-diagonal part. Similarly, by symmetry of the coefficients (6.68), $Q^{\mathrm{off}}_{\mathrm{ps},(i,j)}$ is a sum of graphical matrices. I.e. let $\mathcal{U}^{s,t}_1, ..., \mathcal{U}^{s,t}_{h(s,t)}$ be the collection of distinct 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 (6.66), $Q^{\mathrm{off}}_{\mathrm{ps},(i,j)}$ is a block-diagonal matrix for blocks $s = d/2 - i, ..., d/2$ according to $s = |A| = |B|$, the sth block being

$$Q^{\text{off}}_{\text{ps},(i,j)}(s) = \sum_{t:\ s<t\le\tau} (\frac{\omega}{n})^{t-s} \sum_{x=1}^{h(s,t)} q(\mathcal{U}^{s,t}_x) M_{\mathcal{U}^{s,t}_x}$$

where 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 4.8, w.p. $> 1 - n^{-9.8\log n}$,

$$\left\|Q^{\text{off}}_{\text{ps},(i,j)}(s)\right\| \le \sum_{s<t\le\tau} (\frac{\omega}{n})^{t-s} \cdot h(t,s) \cdot n^{\frac{t-s}{2}} 2^{O(t)} (\log n)^{O(t-s)} \tag{6.73}$$

Also, clearly $h(t,s) \le 2^{\binom{t}{2}+O(t)}$. Therefore, with the same high probability

$$\begin{aligned} \text{RHS of (6.73)} \le & \sum_{\substack{d/2-\max\{i,j\}\le s\le d/2 \\ s<t\le\tau}} (\frac{\omega}{n})^{t-s} 2^{\binom{t}{2}+O(t)} n^{\frac{t-s}{2}} (\log n)^{O(t-s)} \\ < & \sum_{\substack{d/2-\max\{i,j\}\le s\le d/2 \\ s<t\le\tau}} n^{-2\epsilon(t-s)} 2^{O(t)} 2^{\binom{s}{2}} (2^{t+s}\log n)^{O(t-s)} \\ < & 2^{\binom{s}{2}} \cdot n^{-1.9\epsilon}. \qquad \text{(in our parameter regime)} \end{aligned}$$

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

Finally, by the union bound we get that w.p. $> 1 - n^{-9.5\log n}$,

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

completing the proof. ◁

▶ **Corollary 6.36** (Positiveness of $Q^R_{0,0}$). *For every $R \in \binom{[n]}{\le d/2}$, w.p. $> 1 - n^{-8\log n}$ over G*

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

Proof. By Lemma 6.34 and Corollary 6.32, where $\tau^{-6.1\tau} \gg n^{-\epsilon/10}$ in our parameter regime. ◀

▶ **Lemma 6.37** (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 $1 \le k \le d/2$,*

$$\pm Q^R_{0,k}(G) \preceq n^{-k/10} \cdot \text{diag}\left(\widetilde{\text{Cl}}(G)\right)_{S^R \times S^R}.$$

Proof. We will use union bound over (R,k) so fix them first. **For the fixed R, $k(>0)$, in this proof we abbreviate:**

$$Q^R_{0,k} \leftrightarrow Q.$$

Recall the definition of $Q^R_{0,k}$ (Lemma 6.19 (3)): Q is supported within $S^R \times S^R$,

$$Q\Big((A,i),(B,j)\Big) = \sum_{T_m:|V(T_m)\cup A\cup B|\le\tau} (\frac{\omega}{n})^{t-s} q^R_{0,k}(\mathcal{R}_m, i, j) \cdot \chi_{T_m}. \tag{6.74}$$

where $t = |A \cup B|$, $s = \frac{|A|+|B|}{2}$. Abbreviate $q^R_{0,k}$ as q_k. By Lemma 6.19(3),

$$q_k(\cdot, i, j) \text{ is symmetric w.r.t. shapes for all fixed } (i, j); \tag{6.75}$$

$$|q_k(\mathcal{R}_m, i, j)| \leq \tau^{5\tau} \cdot (\frac{\omega}{n^{1-\epsilon}})^{s-p+k/3} \tag{6.76}$$

where $t = |A \cup B|$, $s = \frac{|A|+|B|}{2}$, p is the maximum number of vertex-disjoint paths from A to B in $(A, B; T_m)$.

By symmetry of q_k's, $Q((A, i), (B, j))$ factors through $\mathrm{Cl}(A)\mathrm{Cl}(B)$, so

$$Q = D' \cdot Q \cdot D'. \tag{6.77}$$

where D' is by Definition 6.33. It suffices to show:

$$\text{w.p. } > 1 - n^{-9.5 \log n} \quad \pm Q \prec n^{-k/10} \cdot \operatorname{diag}\left(2^{\binom{|A|}{2}}\right)_{S^R \times S^R}. \tag{6.78}$$

This is because, like in the proof of Lemma 6.34, we can insert (6.78) to the middle of (6.77) which proves the lemma for the fixed R, k.

In below we prove (6.78). First, express each block of Q as a sum of graphical matrices. As a block-matrix, $Q = (Q_{(i,j)})_{0 \leq i,j \leq \tau}$ where $Q_{(i,j)}$ is supported on those A's s.t. $|A| + i \geq d/2$. **For any fixed (i, j)** any $(s_1, s_2) \in \{0, ..., d/2\}^2$ s.t. $s_1 + i \geq d/2$, $s_2 + j \geq d/2$, and any $t \geq \max\{s_1, s_2\}$, 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 (6.74) and symmetry,

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

This equation can be naturally viewed block-wise w.r.t. (s_1, s_2), i.e.

$$Q_{(i,j)} = \sum_{\substack{s_1,s_2 \\ s_1+i, s_2+j \geq d/2}} Q_{(s_1,i),(s_2,j)} \tag{6.79}$$

where

$$Q_{(s_1,i),(s_2,j)} := \sum_{\substack{t: \\ s_1,s_2 \leq t \leq \tau}} \sum_{x=1}^{h(t;s_1,s_2)} q_k(\mathcal{U}_x^{(t;s_1,s_2)}, i, j) \cdot M_{\mathcal{U}_x^{(t;s_1,s_2)}}. \tag{6.80}$$

Note that $Q_{(s_1,i),(s_2,j)}$ is a $\binom{[n]}{s_1} \times \binom{[n]}{s_2}$-matrix on the (i, j)th block of Q.

By Theorem 4.8 and (6.76), w.p. $> 1 - n^{-10 \log n}$

$$\left\|Q_{(s_1,i),(s_2,j)}\right\| \leq \sum_{\substack{t:\ t \leq \tau \\ t \geq s_1,s_2}} h(t; s_1, s_2) \cdot (\frac{\omega}{n})^{t-s} (\frac{\omega}{n^{1-\epsilon}})^{s-p+k/3} \cdot n^{\frac{t-p}{2}} 2^{O(t)} (\log n)^{O(t-s)} \tag{6.81}$$

where, as usual, $s = \frac{s_1+s_2}{2}$ and p is the maximum number of vertex-disjoint paths between the two distinguished subsets in the shape. Since

$$h(t; s_1, s_2) \leq 2^{\binom{t}{2}+O(t)} = 2^{\binom{s}{2}+O(t)+(t+s)\cdot(t-s)},$$

we can bound the RHS of (6.81) (note $k > 0$, $2^{O(t)} < n^{\epsilon/10}$, $\tau^{5\tau} < n^{1/30}$) by

$$< 2^{\binom{s}{2}} \cdot \tau^{5\tau} n^{-k/6} n^{-\epsilon(t-s)} < 2^{\binom{s}{2}} n^{-k/8}. \tag{6.82}$$

Finally, sum over all double-blocks and use Cauchy-Schwartz. Namely, regard each $Q_{(s_1,i),(s_2,j)}$ now as on $S^R \times S^R$ (extended by 0's), then

$$Q = \sum_{\substack{(s_1,i),(s_2,j) \\ s_1+i, s_2+j \geq d/2}} Q_{(s_1,i),(s_2,j)} \tag{6.83}$$

and for each $(s_1,i),(s_2,j)$ in the summand,

$$\pm Q_{(s_1,i),(s_2,j)} \;\prec\; n^{-k/8} \cdot \left(2^{\binom{s_1}{2}} \mathrm{Id}_{(s_1,i),(s_1,i)} \;+\; 2^{\binom{s_2}{2}} \mathrm{Id}_{(s_2,j),(s_2,j)}\right)/2$$

by (6.82) and Cauchy-Schwartz. So by (6.83), w.p. $> 1 - n^{-9.5\log n}$,

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

(6.78) is proved. ◀

▶ **Lemma 6.38** (Bounds on $Q^R_{c,k}$, $c > 0$). *W.p. $> 1 - n^{-9\log n}$ the following holds: $\forall (R,c,k)$ where $R \in \binom{[n]}{\leq d/2}$, $0 < c \leq |R|$ and $0 \leq k \leq d/2$,*

$$\pm \omega^{-c} \cdot Q^R_{c,k} \;\preceq\; n^{-c/3} \cdot \mathrm{diag}\left(\widetilde{\mathrm{Cl}}\right)_{S^R \times S^R}. \tag{6.84}$$

Proof. The proof is almost the same as the previous one (Lemma 6.37). 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); we do it below.

Fix (R,c,k) as in the condition. If $k > 0$ then the proof is identical to that of Lemma 6.37 ($c = 0$), since the same coefficient-size condition and symmetry condition (6.75), (6.76) hold here by Lemma 6.19, 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 6.15, the matrix is supported on $S^R \times S^R$ with expression $Q^R_{c,0}\Big((A,i),(B,j)\Big) =$

$$\sum_{\substack{T_m : |V(T_m) \cup A \cup B| \leq \tau \\ A,B \in \mathrm{mSep}_{A,B}(T_m)}} \left(\frac{\omega}{n}\right)^{|V(T_m) \cup A \cup B| - \frac{|A|+|B|}{2}} \cdot Y_c\big(|R|,\ |V(T_m) \cup A \cup B| + (i+j)\big) \cdot \chi_{T_m} \tag{6.85}$$

where $\big|Y_c\big(|R|,\ |V(T_m) \cup A \cup B| + (i+j)\big)\big| < \tau^{5\tau}$ by Lemma 6.8 (4). If for every 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 (6.85) is bounded by $(\frac{\omega}{n})^{t-s} \cdot \tau^{5\tau}$. Therefore, we have the support condition, the symmetry, and the size condition on the coefficients as in Lemma 6.37, so we can proceed exactly the same as there till equation (6.81), where a single term in its RHS now becomes

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

Note in (6.85) any appearing ribbon $\mathcal{R}_m = (A,B;T_m)$ satisfies $A, B \in \mathrm{mSep}_{A,B}(T_m)$ so $p = s$ (the specialty of the case $k = 0$). So we can replace the bound on the RHS of (6.82) 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 the last line of the proof there, with the bound now being

$$\pm Q^R_{c,0} \;\prec\; \tau^{7\tau} \cdot \mathrm{diag}\left(2^{\binom{|A|}{2}}\right)_{S^R \times S^R}.$$

In particular, since $c \geq 1$, $\omega = n^{\frac{1}{2}-4\epsilon}$ (assuming $\epsilon < 1/40$) and $\tau^{7\tau} < n^{1/15}$, we get $\pm\omega^{-c} \cdot Q^R_{c,0} \prec n^{-c/3} \cdot \operatorname{diag}\left(2^{\binom{|A|}{2}}\right)_{S^R \times S^R}$ by our parameters. Once again like before, using $Q^R_{c,0} = D' \cdot Q^R_{c,0} \cdot D'$ we get that $\pm\omega^{-c} \cdot Q^R_{c,0} \preceq n^{-c/3} \cdot \operatorname{diag}\left(\widetilde{\mathrm{Cl}}\right)_{S^R \times S^R}$. ◀

Lemma 6.21 follows immediately from Corollary 6.36, Lemma 6.37, 6.38.

6.7 Last step

Now we prove the Main Lemma 6.9, hence Theorem 6.1. For any fixed R, recall the notation $P^R = \{I \in \binom{[n]}{d/2} \mid R \subseteq I\}$.

Lemma 6.9 recast. W.p. $1 - n^{-5\log n}$ it holds that for all $R \subseteq \binom{[n]}{d/2}$:

$$M^R_0 \succeq n^{-d} \cdot \operatorname{diag}(\widetilde{\mathrm{Cl}})_{P^R \times P^R}; \tag{6.86}$$

$$\pm\, \omega^{-c} M^R_c \preceq n^{-c/6} \cdot M^R_0, \quad \forall 0 < c \leq |R|. \tag{6.87}$$

Further recall that $D^\tau = \operatorname{diag}\left((\frac{\omega}{n})^{\frac{|A|}{2}}\right)_{A:|A|\leq\frac{d}{2}} \otimes \mathrm{Id}_{\{0,...,\tau\}\times\{0,...,\tau\}}$ (Def. 6.12), and that $S^R = \{(A,i) \in \binom{[n]}{\leq d/2} \times \{0,...,\tau\} \mid A \supseteq R, |A| + i \geq \frac{d}{2}\}$. The following lemma will be handy.

▶ **Lemma 6.39.** $\forall R \in \binom{[n]}{\leq d/2}$,

$$\widetilde{L^R} D^\tau \cdot \operatorname{diag}\left(\widetilde{\mathrm{Cl}}\right)_{S^R \times S^R} \cdot D^\tau (\widetilde{L^R})^\top \succeq (\frac{\omega}{n})^{d/2} \operatorname{diag}\left(\widetilde{\mathrm{Cl}}\right)_{P^R \times P^R}$$

when evaluated on any G.

Proof. Fix any $R \in \binom{[n]}{\leq d/2}$. 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 $\binom{[n']}{\leq d'/2}$, 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 $D^\tau \operatorname{diag}(\widetilde{\mathrm{Cl}}) D^\tau$ is nonnegative and diagonal for any G, we have

$$\widetilde{L^R}\left(D^\tau \cdot \operatorname{diag}\left(\widetilde{\mathrm{Cl}}\right) \cdot D^\tau\right)(\widetilde{L^R})^\top \succeq L^{R,0}\left(D^\tau \cdot \operatorname{diag}\left(\widetilde{\mathrm{Cl}}\right) \cdot D^\tau\right)(L^{R,0})^\top, \tag{6.88}$$

where recall $\widetilde{L^R} = (L^{R,0}, ..., L^{R,\tau})$. Further, $L^{R,0} = (L^{R,0}_0, ..., L^{R,0}_{d'/2})$, where $L^{R,0}_t$ is the matrix on column set $\binom{n'}{t}$. In particular,

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

since in the definition of $L^{R,0}$ (Def. 6.11) only ribbons $\mathcal{R} = (I, A; T')$ with 0-reduced size can occur, and with the other conditions on it this simply means that $A = I$ and $T' \subseteq E(I)$. This implies

$$\text{RHS of (6.88)} \succ (\frac{\omega}{n})^{d/2} \cdot \operatorname{diag}\left(\widetilde{\mathrm{Cl}}\right)_{\binom{[n']}{d'/2}\times\binom{[n']}{d'/2}}.$$

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

Proof for Lemma 6.9. Fix $R \in \binom{[n]}{\le d/2}$. By Lemma 6.19, for all $c \le |R|$

$$M_c^R = \widetilde{L^R} \cdot \left[D^\tau \left(Q_{c,0}^R - Q_{c,1}^R + ... \pm Q_{c,d}^R\right) D^\tau\right] \cdot \left(\widetilde{L^R}\right)^\top + \mathcal{E}_c^R. \tag{6.89}$$

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 (6.86). Fix R, we have:

$$\begin{aligned}
M_0^R &= \widetilde{L^R} \cdot \left[D^\tau \left(Q_{0,0}^R - Q_{0,1}^R + ... \pm Q_{0,d}^R\right) D^\tau\right] \cdot \left(\widetilde{L^R}\right)^\top + \mathcal{E}_0^R \\
&\succeq \tau^{-7\tau} \left[\widetilde{L^R} \cdot D^\tau \mathrm{diag}\left(\widetilde{\mathrm{Cl}}\right)_{S^R \times S^R} D^\tau \cdot \left(\widetilde{L^R}\right)^\top\right] + \mathcal{E}_0^R \quad \text{(Lem. 6.21(1))} \\
&\succeq \tau^{-7\tau} (\frac{\omega}{n})^{d/2} \cdot \mathrm{diag}\left(\widetilde{\mathrm{Cl}}\right)_{P^R \times P^R} + \mathcal{E}_0^R \quad \text{(Lemma 6.39)} \\
&\succeq (\tau^{-7\tau} (\frac{\omega}{n})^{d/2} - n^{-\epsilon\tau/2}) \cdot \mathrm{diag}\left(\widetilde{\mathrm{Cl}}\right)_{P^R \times P^R} \quad \text{(Lemma 6.19(4))} \\
&\succeq n^{-d} \cdot \mathrm{diag}(\widetilde{\mathrm{Cl}})_{P^R \times P^R} \quad \text{(parameter regime)}
\end{aligned}$$

For (6.87). Fix R, $1 \le c \le |R|$, we have:

$$\begin{aligned}
M_c^R &= \widetilde{L^R} \cdot \left[D^\tau \left(Q_{c,0}^R - Q_{c,1}^R + ... \pm Q_{c,d}^R\right) D^\tau\right] \cdot \left(\widetilde{L^R}\right)^\top + \mathcal{E}_c^R \\
&\preceq \omega^c n^{-c/4} \left[\widetilde{L^R} D^\tau \cdot \mathrm{diag}\left(\widetilde{\mathrm{Cl}}\right)_{S^R \times S^R} \cdot D^\tau \left(\widetilde{L^R}\right)^\top\right] + \mathcal{E}_c^R \quad \text{(Lem. 6.21(2))} \\
&\preceq \omega^c n^{-c/4} \left[\tau^{7\tau} (M_0^R - \mathcal{E}_0^R)\right] + \mathcal{E}_c^R \quad \text{(Lem. 6.21(1) and (6.89))} \\
&\preceq \omega^c n^{-c/5} M_0^R + \left(\omega^c n^{-c/5} + 1\right) n^{-\epsilon\tau/2} \mathrm{diag}\left(\mathrm{Cl}\right)_{P^R \times P^R} \quad \text{(Lem. 6.19(4))}
\end{aligned}$$

So

$$\begin{aligned}
\omega^{-c} M_c^R &\preceq n^{-c/5} M_0^R + 2n^{-\epsilon\tau/2} \cdot \mathrm{diag}\left(\mathrm{Cl}\right)_{P^R \times P^R} \\
&\preceq (n^{-c/5} + 2n^d n^{-\epsilon\tau/2}) M_0^R \quad \text{((6.86) and } \widetilde{\mathrm{Cl}} \ge \mathrm{Cl}) \\
&\preceq n^{-c/6} \cdot M_0^R \quad (c \le |R| \le d/2 \text{ and parameter regime})
\end{aligned}$$

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

7 Concluding remarks

We established the average $\Omega(\epsilon^2 \log n / \log\log n)$ SOS degree lower bound for Exact Clique with clique-size $\omega = n^{1/2-\epsilon}$, which is nearly optimal in both parameters ω, d. We also refreshed the techniques for the Non-Exact Clique problem in hope to make them simpler and generalizable. Some open problems follow.

(1) Can we remove the $\log\log n$ factor in d? Perhaps it helps to first find a conceptual explanation of Definition 3.11.

(2) How about the same problem on $G(n, p)$, $p \neq \frac{1}{2}$ and for suitable ω? For Non-Exact Clique, we can define the pseudo-expectation similarly as in Section 3.1.2. Also, using the Fourier orthonormal basis

$$\chi_T = \prod_{e \in T} \frac{x_e - (2p-1)}{2\sqrt{p(1-p)}} \qquad \forall T \subseteq E[n], \tag{7.1}$$

where $x_e(G)$ is the ±1-indicator of edge e, we have the corresponding version of norm bounds in Section 4 since the trace-power method works the same. The questions is, what is the best meaningful degree lower bound for varying p (especially small p)? How about the exact case?

(3) What can be said when G is drawn from other random models, or is pseudo-random?

References

1 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.

2 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.

3 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.

4 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.

5 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.

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

7 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.

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

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

10 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.

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

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

13 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.

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

15 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.

16 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*, 2015. `arXiv:1507.05230`.

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.

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

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

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

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

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

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

24 Dhruv Medarametla and Aaron Potechin. Bounds on the norms of uniform low degree graph matrices. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2016)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2016.

25 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.

26 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.

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

28 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.

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

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

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

A Deductions in mod-order analysis (Section 5.2)

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. 4.11).

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 5.6) 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* $(*)$*.*

$$A \text{ is a min-separator for both } (I, A; T'),\ (J, A; T''); \tag{A.5}$$

$$(V(T') \cup I \cup A)\ \cap\ (V(T'') \cup J \cup A) = A. \tag{A.6}$$

Moreover, (A.5)*,* (A.6) *imply that*

$$A \text{ is a min-separator of } (I, J; T) \text{ (where } T = T' \oplus T''\text{)}; \tag{A.7}$$

$$|V(T') \cup I \cup A|,\ |V(T'') \cup J \cup A| \le \tau. \tag{A.8}$$

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),

$$\text{its order in } \alpha \ge |A| + |V(T') \cup I \cup A| - s_{I,A}(T') + |V(T'') \cup A \cup J| - s_{J,A}(T'').$$

By modularity on the RHS of (A.3), the term is non-zero only if

$$\text{its order in } \alpha \le |V(T) \cup I \cup J| \quad \text{and} \quad |V(T) \cup I \cup J| \le \tau$$

where $T = T' \oplus T''$. This implies

$$|V(T') \cup I \cup A| + |V(T'') \cup J \cup A| \le \underbrace{|V(T) \cup I \cup J|}_{\textcircled{1}} + \underbrace{(s_{I,A}(T') + s_{J,A}(T'') - |A|)}_{\textcircled{2}} \tag{A.9}$$

Note $\textcircled{2} \le |A|$ and "=" holds iff $s_{I,A}(T') = s_{J,A}(T'') = |A|$. While the LHS above

$$= \underbrace{|(V(T') \cup I \cup A) \cup (V(T'') \cup J \cup A)|}_{\ge |V(T) \cup I \cup J| = \textcircled{1}} + \underbrace{|(V(T') \cup I \cup A) \cap (V(T'') \cup J \cup A)|}_{\ge |A| \ge \textcircled{2}}.$$

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 $A \in \mathrm{mSep}_{I,J}(T)$ (and also (2), actually). By (3), T', T'' could overlap only in $E(A)$. Now $T = T' \oplus T''$, so

$$T = T' \sqcup T'' \quad \text{modulo } E(A) \tag{A.10}$$

(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 $A \in \mathrm{mSep}_{I,J}(T)$.

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 $\le \tau$. ◀

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.2.** *If $(I, J; T)$ has a* ***unique*** *min-separator A, then*

$$\sum_{\substack{T',T'':\ T'\oplus T''=T\\ (A.5),\ (A.6)\ hold}} \beta_{I,A}(T') \cdot \beta_{J,A}(T'') = \alpha^{e_{I,J}(T)} \tag{A.11}$$

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 5.6. 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

$$L_1' : \quad L_1'(I, A) = \sum_{\substack{T':\ |V(T') \cup I \cup A| \le \tau\\ (A.13)\ \text{holds}\\ T' \cap E(A) = \emptyset}} 2^{-\binom{|A|}{2}/2} \cdot \alpha^{|V(T') \cup I \cup A| - |A|} \chi_{T'} \cdot \widetilde{\mathrm{Cl}}_A$$

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.3.** $\forall I \in \binom{[n]}{d/2}$, $A \in \binom{[n]}{\leq d/2}$,

$$L_1(I,A) := \sum_{\substack{T':\ |V(T')\cup I\cup A|\leq\tau \\ \text{(A.13) holds} \\ T'\cap E(A)=\emptyset}} \alpha^{|V(T')\cup I\cup A|-|A|}\chi_{T'}. \tag{A.15}$$

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.2. 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_iy_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.4.** *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*

$$L_2(I,A) = \sum_{T':\ |V(T')\cup I\cup A|\leq\tau} \beta^{(2)}_{I,A}(T')\chi_{T'} \tag{A.18}$$

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

$$\sum_{\substack{A,T',T'':\ T'\oplus T''=T \\ (I,A;T')\text{ appears in }L_1 \\ \text{(A.5),(A.6) hold}}} \alpha^{e_{I,A}(T')}\cdot\beta^{(2)}_{J,A}(T'') = \alpha^{e_{I,J}(T)} \mod \alpha^{e_{I,J}(T)+1}. \tag{A.19}$$

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** an essential part of the following graph-theoretic fact due to Escalante (its "In particular" part).

▶ **Fact A.1** ([11]; also Appendix A.3 of [5])**.** *For any ribbon $(I,J;T)$, the set of all min-separators,* $\mathrm{mSep}_{I,J}(T)$*, has a natural poset structure: min-separators $A_1 \leq A_2$ 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.5.** *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. 5.8 recast).

▶ **Proposition A.6** (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 5.11 and Remark 5.12. 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 "$L_1(-)L_1^\top$"-shape decomposition, meaning that we do not lose much from the polarization step if recall the goal is only about PSDness.