Minimum number of monotone subsequences of length 4 in permutations

József Balogh* Ping Hu† Bernard Lidický‡ Oleg Pikhurko§

Balázs Udvari¶ Jan Volec‖

July 25, 2014

Abstract

We show that for every sufficiently large n, the number of monotone subsequences of length four in a permutation on n points is at least $\binom{\lfloor n/3 \rfloor}{4} + \binom{\lfloor (n+1)/3 \rfloor}{4} + \binom{\lfloor (n+2)/3 \rfloor}{4}$. Furthermore, we characterize all permutations on $[n]$ that attain this lower bound. The proof uses the flag algebra framework together with some additional stability arguments. This problem is equivalent to some specific type of edge colorings of complete graphs with two colors, where the number of monochromatic K_4's is minimized. We show that all the extremal colorings must contain monochromatic K_4's only in one of the two colors. This translates back to permutations, where all the monotone subsequences of length four are all either increasing, or decreasing only.

Keywords: flag algebras, permutation, pattern, density

*Department of Mathematics, University of Illinois, Urbana, IL 61801, USA and Bolyai Institute, University of Szeged, Szeged, Hungary `jobal@math.uiuc.edu`. Research is partially supported by Simons Fellowship, NSF CAREER Grant DMS-0745185, Arnold O. Beckman Research Award (UIUC Campus Research Board 13039) and Marie Curie FP7-PEOPLE-2012-IIF 327763.

†Department of Mathematics, University of Illinois, Urbana, IL 61801, USA, `pinghu1@math.uiuc.edu`.

‡Department of Mathematics, University of Illinois, Urbana, IL 61801, USA, `lidicky@illinois.edu` and Charles University in Prague. Research is partially supported by NSF grant DMS-1266016.

§Mathematics Institute and DIMAP, University of Warwick, Coventry CV4 7AL, UK, `O.Pikhurko@warwick.ac.uk`. Research is partially supported by ERC grant 306493 and EPSRC grant EP/K012045/1.

¶Bolyai Institute, University of Szeged, Szeged, Hungary. `udvarib@math.u-szeged.hu`. Research was supported by the European Union and the State of Hungary, co-financed by the European Social Fund in the framework of TÁMOP 4.2.4. A/2-11-1-2012-0001 National Excellence Program.

‖Mathematics Institute and DIMAP, University of Warwick, Coventry CV4 7AL, UK, `honza@ucw.cz`. Research is partially supported by the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement no. 259385.

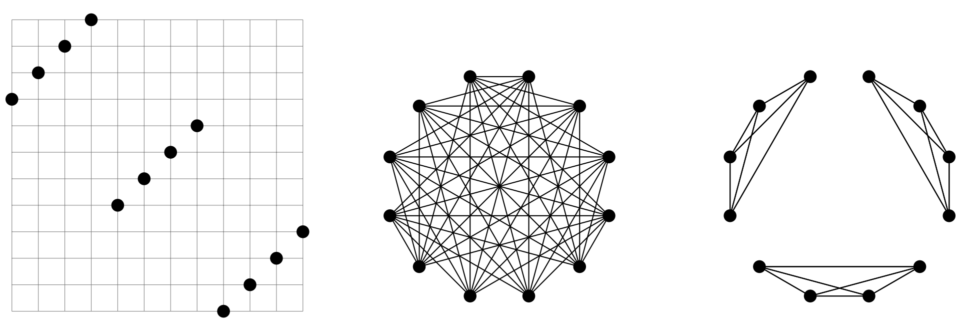
1 Introduction

Our work was inspired by a famous result of Erdős and Szekeres [11] that every permutation on $[n] = \{1, \ldots, n\}$, where $n \geq k^2 + 1$, contains a monotone subsequence of length $k+1$. If $n \gg k^2$, one expects that the number of monotone subsequences of length $k+1$ is more than just one, which is guaranteed by [11]. According to Myers [22], the problem of determining the minimum number of monotone subsequences of length $k+1$ in permutations on $[n]$ was first posed by Atkinson, Albert and Holton. As in [22], we use $m_k(\tau)$ to denote the number of monotone subsequences of length $k+1$ in a permutation τ. The minimum of $m_k(\tau)$ over all permutations $\tau \in S_n$ is denoted by $m_k(n)$.

Myers [22] described a permutation $\tau_k(n)$ which gives an upper bound on $m_k(n)$. It consists of k increasing sequences K whose sizes differ by at most one and every monotone sequence of length $k+1$ is entirely contained in one of the K sequences. In other words, with $t_j = \lfloor jn/k \rfloor$, an example of such a permutation is

$$\begin{array}{rllllll} \tau_k(n) = (& t_{k-1}+1, & t_{k-1}+2, & \ldots, & n-1, & n, & \\ & t_{k-2}+1, & t_{k-2}+2, & \ldots, & t_{k-1}-1, & t_{k-1}, & \\ & \ldots & & & & & \\ & 1, & 2, & \ldots, & t_1-1, & t_1 &). \end{array}$$

See Figure 1 for $\tau_3(12)$.



$\tau_3(12)$ $T_3(12)$ $\overline{T}_3(12)$

Figure 1: Permutation $\tau_3(12)$ and its representation graph (introduced in Section 2) $T_3(12)$.

Let $r \equiv n \pmod{k}$, where $0 \leq r < k$. It is easy to see that

$$m_k(\tau_k(n)) = r\binom{\lceil n/k \rceil}{k+1} + (k-r)\binom{\lfloor n/k \rfloor}{k+1} \approx \frac{1}{k^k}\binom{n}{k+1}.$$

Myers [22] proved that $m_2(n) = m_2(\tau_2(n))$ holds and he described all permutations $\tau \in S_n$ where $m_2(\tau) = m_2(n)$. He conjectured that the same formula actually holds for every $k \in \mathbb{N}$.

Conjecture 1 (Myers [22]). *Let n and k be positive integers. In any permutation of $[n]$ there are at least $m_k(\tau_k(n))$ monotone subsequences of length $k+1$.*

Notice that any permutation $(a_1, \dots, a_n)$ and its reverse $(a_n, \dots, a_1)$ contain the same number of monotone subsequences, only the increasing subsequences change to decreasing subsequences and vice versa. In particular, $m_k(\tau_k(n)) = m_k(\tau_k^R(n))$, where $\tau_k^R(n)$ denotes the reverse of $\tau_k(n)$. Moreover, there might be other permutations τ such that $m_k(\tau) = m_k(\tau_k(n))$.

As we already mentioned, Myers showed the conjecture is true for $k = 2$, which is actually a consequence of Goodman's formula [15]. Very recently, Samotij and Sudakov [32] confirmed the conjecture if $n \le k^2 + ck^{3/2}/\log k$ for some absolute positive constant c, provided k is sufficiently large.

Subject to the additional constraint that all the monotone subsequences of length $k+1$ are either all increasing or all decreasing and $n \ge k(2k-1)$, Myers proved that every such a permutation contains at least the conjectured number of monotone subsequences of length $k+1$. He also gave the list $\mathcal{W}_n^k$ of all such permutations τ of $[n]$ that satisfy $m_k(\tau) = m_k(\tau_k(n))$. Every permutation from $\mathcal{W}_n^k$ can be decomposed into k disjoint monotone subsequences $s_1, \dots, s_k$ that are either all increasing or all decreasing and their sizes differ by at most one. Moreover, every monotone subsequence of length $k+1$ is a subsequence of s_j for some j. These permutations look similar to $\tau_k(n)$ or $\tau_k^R(n)$. It turns out that there are $2\binom{k}{n \bmod k} C_k^{2k-2}$ of them, where C_k is the k^{th} Catalan number.

The interested reader can find the precise definition of $\mathcal{W}_n^k$ for general k in [22]. Here, we study the number of monotone subsequences with $k = 3$. Hence we give a simpler alternative definition for $\mathcal{W}_n^3$, where $n \ge 15$.

First we describe a method to get any permutation from $\mathcal{W}_n^3$ with no increasing subsequence of length 4.

1. Start with the identity permutation.

2. Divide it into 3 blocks such that the size of each block is $\lfloor n/3 \rfloor$ or $\lfloor n/3 \rfloor + 1$. More formally, choose elements b_1 and b_2 such that b_1, $b_2 - b_1$ and $n - b_2$ are all from the set $\{\lfloor n/3 \rfloor, \lfloor n/3 \rfloor + 1\}$. Then the three blocks are $(1, 2, \dots, b_1)$, $(b_1+1, b_1+2, \dots, b_2)$, $(b_2+1, b_2+2, \dots, n)$. (There are 1 or 3 choices for the pair (b_1, b_2), depending on the remainder of dividing n by 3.)

3. Reverse the blocks. At this point we have the permutation $(b_1, b_1-1, \dots, 2, 1, b_2, b_2-1, \dots, b_1+2, b_1+1, n, n-1, \dots, b_2+2, b_2+1)$.

4. Change the subsequence $(2, 1, b_2, b_2-1)$ to one of the following: $(2, 1, b_2, b_2-1)$, $(2, b_2, 1, b_2-1)$, $(2, b_2, b_2-1, 1)$, $(b_2, 2, 1, b_2-1)$, $(b_2, 2, b_2-1, 1)$.

5. Make a similar replacement for the subsequence $(b_1+2, b_1+1, n, n-1)$.

6. Change the subsequence $(b_1, b_1-1, b_1+2, b_1+1)$ to one of the following: $(b_1, b_1-1, b_1+2, b_1+1)$, $(b_1+1, b_1-1, b_1+2, b_1)$, $(b_1+2, b_1-1, b_1+1, b_1)$, $(b_1+1, b_1, b_1+2, b_1-1)$, $(b_1+2, b_1, b_1+1, b_1-1)$.

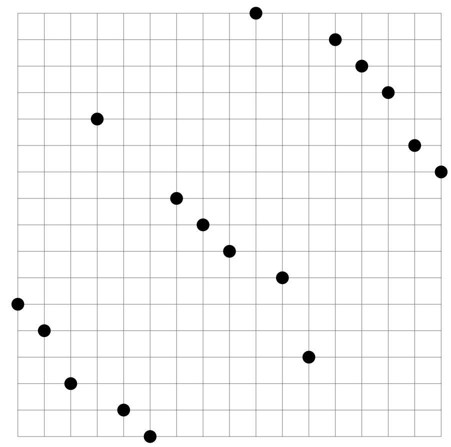


Figure 2: Permutation (6,5,3,13,2,1,10,9,8,17,7,4,16,15,14,12,11).

7. Make a similar replacement for the subsequence $(b_2, b_2 - 1, b_2 + 2, b_2 + 1)$.

Each above permutation (as well as its reverse) belongs to $\mathcal{W}_n^3$ since it has $m_3(\tau_3(n))$ monotone subsequences of length 4, all of which are decreasing. For $n \geq 15$, we exhaust all of $\mathcal{W}_n^3$, as the number of obtained permutations, $2 \cdot 5^4 \binom{3}{r}$ where r is the remainder of dividing n by 3, coincides with the value of $|\mathcal{W}_n^3|$ obtained by Myers.

To illustrate the above process, let $n = 17$. We start with $(1, 2, \dots, 17)$. Let $b_1 = 5, b_2 = 11$. After the reversal of the blocks, we have $(5, 4, 3, 2, 1, 11, 10, 9, 8, 7, 6, 17, 16, 15, 14, 13, 12)$. Now we can change, one by one in the given order, the subsequences $(2, 1, 11, 10)$, $(7, 6, 17, 16)$, $(5, 4, 7, 6)$, $(11, 10, 13, 12)$ to $(11, 2, 1, 10)$, $(17, 7, 6, 16)$, $(6, 5, 7, 4)$, $(13, 10, 12, 11)$ respectively, to get

$$(6, 5, 3, 13, 2, 1, 10, 9, 8, 17, 7, 4, 16, 15, 14, 12, 11).$$

This permutation is depicted in Figure 2.

In his paper, Myers [22] also conjectured a weaker asymptotic version.

Conjecture 2 (Myers [22])**.** *Let k be positive integer and let $n \to \infty$. In any permutation of $[n]$ there are at least $(1 + o(1))\binom{n}{k+1}/k^k$ monotone subsequences of length $k + 1$.*

First, we prove Conjecture 2 for $k = 3$.

Theorem 3. *Any permutation of $[n]$ contains at least $(1/27 + o(1))\binom{n}{4}$ monotone subsequences of length 4.*

Our main result is proving Conjecture 1 for $k = 3$ and n sufficiently large.

Theorem 4. *There exists n_0 such that if $n \geq n_0$, then every permutation τ on $[n]$ contains at least*

$$\binom{\lfloor n/3 \rfloor}{4} + \binom{\lfloor (n+1)/3 \rfloor}{4} + \binom{\lfloor (n+2)/3 \rfloor}{4}$$

monotone subsequences of length 4, with equality if and only if $\tau \in \mathcal{W}_n^3$.

Our results are proved using the flag algebra framework and the stability method. Although Theorems 3 and 4 are stated in terms of permutations, we translate them to the language of graph theory since the resulting computations and arguments are simpler. In graph theory language, we minimize the number of copies of K_4 and $\overline{K}_4$ over graphs from permutations on $[n]$. Let us note that the question of minimizing the number of copies of K_4 and $\overline{K}_4$ over all graphs on n vertices is open. The best upper bound $\approx 1/33$ is due to Thomason [36]. The first known lower bound $\approx 1/46$ is due to Giraud [13]. It was improved using flag algebras to 0.0287... by Sperfeld [34] and independently by Nieß [23], and then further improved by Flagmatic [37] to 0.0294... $\approx 1/34$.

We also had a computer program, developed originally by Dan Kráľ, doing flag algebra computations for permutations directly. It was easy to modify this program to compute upper bounds on densities of other subsequences instead of lower bounds for monotone subsequences. The results that we obtained will be explained in the next paragraph.

The *packing density* of a permutation $\tau \in S_k$ is the limit for $n \to \infty$ of the maximum density of τ in σ over all $\sigma \in S_n$. We denote the limit by $\delta(\tau)$. The packing density is well understood [1] for the so-called *layered permutations*[1]. Up to a symmetry, this includes all permutations in S_3 and all but two permutations, 1342 and 2413, from S_4. Albert, Atkinson, Handley, Holton, and Stromquist [1] proved that $0.19657 \leq \delta(1342) \leq 2/9$ and $51/511 \leq \delta(2413) \leq 2/9$. Presutti [27] improved the lower bound for $\delta(2413)$ to 0.1024732. Further improvement on the lower bound was obtained by Presutti and Stromquist [28] who showed that $0.10472422757673209 04\ldots \leq \delta(2413)$ and conjectured that it is the correct value. A direct application of the semidefinite method from the flag algebra framework for permutations on S_7 gave upper bounds $\delta(1342) \leq 0.1988373$ and $\delta(2413) \leq 0.1047805$. Since our upper bounds do not match the lower bounds, we will not discuss these bounds any further in this paper.

This paper is organized as follows. In the following section, we translate the problem of determining the density of monotone subsequences in permutations to determining densities of particular induced subgraphs in permutation graphs. In Section 3, we describe how we use the framework of flag algebras and we will prove Theorem 3. Our proof of the density result actually provides some additional information about the extremal structures, which leads to a proof of a stability property for this problem. This is discussed in Section 4. Finally, in Section 5, we use the stability property to prove Theorem 4.

We utilize the semidefinite method from flag algebras to formulate our question about subgraph densities as an optimization problem, more precisely, as a semidefinite programming problem. With a computer assistance, we generate this semidefinite programming problem and then we use CSDP [8], an open-source semidefinite programming library, to find a numerical (approximate) solution to the problem. In order to obtain an exact result, the numerical solution needs to be rounded. This was done again with a computer assistance in a computer algebra software SAGE [35]. We had troubles finding a detailed description

[1] A permutation $\tau \in [n]$ is *layered* if there exist positive numbers $n_1, \ldots, n_r$ summing to n, such that τ starts with the n_1 first positive integers in reverse order, followed by the next n_2 positive integers in reverse order and so on. For example $\tau_k^R(n)$ is a layered permutation.

of rounding in other papers. Hence we decided to include more details about our rounding procedure in the appendix.

Our computer programs, their outputs, and their description for the flag algebra part of this paper can be downloaded at `http://www.math.uiuc.edu/~jobal/cikk/permutations/`.

2 Graph Densities

Given a graph G, we use $V(G)$ and $E(G)$ to denote its vertex and edge sets respectively, and let $v(G) = |V(G)|, e(G) = |E(G)|$. For a vertex v of G, we denote the set of its neighbors by $\Gamma_G(v)$. We omit a subscript, if G is clear from the context. Given two graphs G and G', an *isomorphism* between them is a bijection $f : V(G) \to V(G')$ satisfying $f(v_1)f(v_2) \in E(G')$ if and only if $v_1v_2 \in E(G)$. Two graphs G and G' are *isomorphic* ($G \cong G'$) if and only if there is an isomorphism between them. For a graph G and a vertex set $U \subseteq V(G)$, denote by $G[U]$ the induced subgraph of G on vertex set U. Suppose H and G are graphs on l and n vertices respectively. Let $P(H,G)$ be the number of l-subsets U of $V(G)$ such that $G[U] \cong H$, and define the *density* of H in G to be

$$p(H,G) = \frac{P(H,G)}{\binom{n}{l}}.$$

Given a permutation τ of $[n]$, define its *representation graph* to be a graph on vertex set $[n]$ where ij with $i < j$ is an edge if and only if $\tau(i) > \tau(j)$. Call an n-vertex graph G *admissible* if there is a permutation of $[n]$ whose representation graph is isomorphic to G, so the vertex set of G may not be $[n]$. Denote by $\mathcal{M}_l$ the set of admissible graphs on l vertices, up to isomorphism. It is easy to see that if G is admissible, then so are $\overline{G}$ and all induced subgraphs of G.

Given a permutation τ of $[n]$, let G be its representation graph. Then the number of monotone subsequences of length 4 in τ is equal to the number of K_4's and $\overline{K}_4$'s in G, i.e., $m_3(\tau) = P(K_4, G) + P(\overline{K}_4, G)$. Let

$$F(G) = P(K_4, G) + P(\overline{K}_4, G) \qquad \text{and} \qquad f(G) = p(K_4, G) + p(\overline{K}_4, G).$$

Instead of proving Theorem 3 directly, we prove its reformulation to the language of graphs and densities.

Theorem 5. *If G is an admissible graph on n vertices, then $f(G) \geq 1/27 + o(1)$, where $o(1) \to 0$ as $n \to \infty$.*

It is easy to see that

$$f(G) = \sum_{H \in \mathcal{M}_l} f(H)p(H,G) \quad \text{for} \quad 4 \leq l \leq n. \tag{1}$$

Therefore $\min_{H\in\mathcal{M}_l} f(H)$ provides a lower bound on $f(G)$ (since $0 \leq p(H,G) \leq 1$ and $\sum_{H\in\mathcal{M}_l} p(H,G) = 1$), though this bound is unsurprisingly weak for small l.

Denote by $T_3(n)$ the 3-partite Turán graph on n vertices (i.e. complete 3-partite graph on n vertices with sizes of parts differing by at most one). We can see that $T_3(n)$ is the representation graph of $\tau_3(n)$. See Figure 1 for an example, where $n = 12$.

Theorem 6. *There exists an n_0 such that if G is an admissible graph on $n \geq n_0$ vertices minimizing F over all admissible graphs on n vertices, then G is obtained from $T_3(n)$ by removing edges or G is obtained from $\overline{T_3(n)}$ by adding edges.*

Remark: Let G be an extremal graph. By Theorem 6, G can be transformed into $T_3(n)$ or $\overline{T_3(n)}$. We may assume without loss of generality (w.l.o.g.) that G is obtained from $T_3(n)$ by removing edges. Since $T_3(n)$ does not contain any copy of K_4 and removing edges does not introduce new copies of K_4, there are no K_4's in G. Moreover, since G is extremal and removing edges does not destroy any copy of $\overline{K}_4$, the numbers of copies of $\overline{K}_4$ in G and $T_3(n)$ are equal. Hence we know that in an extremal permutation τ, monotone subsequences of length 4 are either all increasing or all decreasing. Thus τ belongs to the family $\mathcal{W}_n^3$ constructed by Myers (and Theorem 4 follows from Theorem 6). In fact, it is not hard to see that $\tau \in \mathcal{W}_n^3$ directly. Indeed, τ can be decomposed into three monotone subsequences s_1, s_2, s_3, that correspond to the parts of Turán graph, and all monotone 4-subsequences are entirely contained in them. Then it follows that the domains of s_1, s_2, s_3 form three consecutive intervals of $[n]$, except some possible intertwining at their ends that involves at most two elements from each interval, which leads to the desired structure of τ.

3 Flag Algebra Settings

The flag algebra method, invented by Razborov [30], is a very general machinery and has been widely used in extremal graph theory. See [31] for a recent survey of flag algebra applications. To name just some of them: flag algebra was used for attacking the Caccetta-Häggkvist conjecture [19, 29], determining induced densities of graphs [10, 16, 17, 25, 26], of hypergraphs [6, 12, 14, 24], of oriented graphs [33], of subhypercubes in hypercubes [3, 7], of colored graphs in a colored environment [5, 9, 18, 20], and for attacking some problems in geometry [21].

We apply this method to the family of admissible graphs. A *type* σ is an admissible graph on vertex set $[k]$ for some non-negative integer k, where k is called the *size* of σ, denoted by $|\sigma|$. We use 0 and 1 to denote (the unique) types of size 0 and 1 respectively. A *σ-flag* F is a pair (M, θ) where M is an admissible graph and $\theta : [k] \to V(M)$ induces a labeled copy of σ in M. In other words, we use $[k]$ to label k vertices of an unlabeled graph M, and the labeled vertices induce a labeled copy of σ. Two σ-flags $F_1 = (M_1, \theta_1)$ and $F_2 = (M_2, \theta_2)$ are *isomorphic* (denoted as $F_1 \cong F_2$) if there exists a graph isomorphism $f : V(M_1) \to V(M_2)$ such that $f\theta_1 = \theta_2$. Such a function f is called a *flag isomorphism* from F_1 to F_2. Given an admissible graph M, if all σ-flags with the underlying graph M are isomorphic, then we use M^σ to denote this unique σ-flag, see Figure 3 for an example where $M \in \{K_4, \overline{K}_4\}$. Denote by $\mathcal{F}_l^\sigma$ the set of σ-flags on l vertices, up to isomorphism. Note that $\mathcal{F}_l^0$ is just $\mathcal{M}_l$ and $\mathcal{F}_{|\sigma|}^\sigma = \{\sigma\}$.

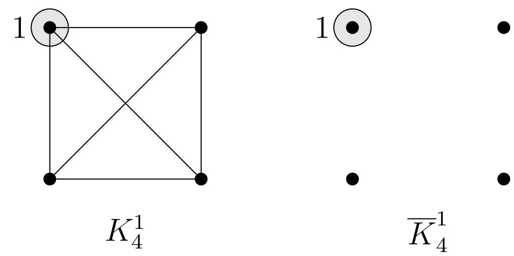


Figure 3: 1-flags K_4^1 and $\overline{K}_4^1$.

In Section 2, we defined graph density $p(H,G)$, which extends to flag density in a straightforward way. Given σ-flags $F \in \mathcal{F}_l^\sigma$ and $K = (G,\theta) \in \mathcal{F}_n^\sigma$ for $l \le n$, define $P(F,K)$ to be the number of l-subsets U of $V(G)$ such that $\mathrm{Im}(\theta) \subseteq U$ and $(G[U],\theta) \cong F$. Additionally, define $p(F,K)$, the *density* of F in K as

$$p(F,K) = \frac{P(F,K)}{\binom{n-|\sigma|}{l-|\sigma|}}.$$

By convention, we set $P(F,K) = 0$ if $n < l$. More generally, given flags $F \in \mathcal{F}_l^\sigma$, $F' \in \mathcal{F}_{l'}^\sigma$ and $K = (G,\theta) \in \mathcal{F}_n^\sigma$, where $n \ge l + l' - |\sigma|$, we define a *joint density* $p(F,F';K)$ as the probability that if we choose two subsets U, U' of $V(G)$ uniformly at random, subject to the conditions $|U| = l, |U'| = l'$ and $U \cap U' = \mathrm{Im}(\theta)$, then $(G[U],\theta) \cong F$ and $(G[U'],\theta) \cong F'$. In this paper, whenever we use $p(F,K)$ or $p(F,F';K)$, we assume that the size of K is large enough.

It is not very hard to show that (see Lemma 2.3 in [30])

$$p(F,K)p(F',K) = p(F,F';K) + o(1), \tag{2}$$

where $o(1)$ tends to 0 as n tends to infinity. Let $X = [F_1,\ldots,F_t]$ be a vector of σ-flags with $F_i \in \mathcal{F}_{l_i}^\sigma$. For any such X and a σ-flag K define $X_K = [p(F_1;K),\ldots,p(F_t;K)]$. It follows that for any t-by-t positive semidefinite matrix $Q = \{Q_{ij}\}$, we have

$$0 \le X_K^T Q X_K = \sum_{ij} Q_{ij} p(F_i;K)p(F_j;K) = \sum_{i,j} Q_{ij} p(F_i,F_j;K) + o(1). \tag{3}$$

In the definition of $p(F,K)$ and $p(F,F';K)$, we require F, F' and K to be σ-flags, but the definition itself extends to the case where F, F' are σ-flags but K is not. In this case, by the definition, we have $p(F,K) = p(F,F';K) = 0$. Let $\Theta(k,G)$ be the set of all injective mappings from $[k]$ to $V(G)$ where G is an admissible graph.

We can extend (3) to any $\theta \in \Theta(|\sigma|,G)$:

$$0 \le \sum_{i,j} Q_{ij} p(F_i,F_j;(G,\theta)) + o(1).$$

Therefore, if we choose θ from $\Theta(|\sigma|, G)$ uniformly at random, then its expectation is non-negative:

$$\begin{aligned}0 &\le \sum_{i,j} \mathbb{E}_{\theta\in\Theta(|\sigma|,G)}\left[Q_{ij}p(F_i, F_j; (G,\theta))\right] + o(1)\\ &= \sum_{H\in\mathcal{M}_l}\left(\sum_{i,j}\mathbb{E}_{\theta\in\Theta(|\sigma|,H)}\left[Q_{ij}p(F_i, F_j; (H,\theta))\right]\right)p(H,G) + o(1).\end{aligned}$$

(Recall that we assumed that $l \ge 2l_i - |\sigma|$ for each i.) Note that the coefficient of $p(H, G)$ is determined by σ, X, Q and H. In particular, it is independent of G, so denote this coefficient by $c_H(\sigma, X, Q)$. Then we have

$$\sum_{H\in\mathcal{M}_l} c_H(\sigma, X, Q)p(H, G) + o(1) \ge 0.$$

Every choice of σ, X, Q gives one such inequality. We can add the inequalities obtained for several different types σ_i, using appropriate X_i and Q_i. Denoting $c_H = \sum_i c_H(\sigma_i, X_i, Q_i)$, we obtain

$$\sum_{H\in\mathcal{M}_l} c_H \cdot p(H, G) + o(1) \ge 0.$$

Then together with (1) we have

$$f(G) + o(1) \ge \sum_{H\in\mathcal{M}_l}(f(H) - c_H)\cdot p(H, G) \ge \min_{H\in\mathcal{M}_l}(f(H) - c_H). \tag{4}$$

By (4), if for some choice of (large enough) l and c_H we have

$$\min_{H\in\mathcal{M}_l}(f(H) - c_H) = 1/27, \tag{5}$$

then we would prove Theorem 5.

Proof of Theorem 5. We show (5) with $l = 7$, where $|\mathcal{M}_7| = 776$. We use three choices of (σ, X, Q). We use types $\sigma_0 : P_1$, $\sigma_1 : P_3$, and $\sigma_2 : \overline{P}_3$, where P_i is a path on i vertices, see Figure 4.

For σ_0, X_0 consists of flags in $\mathcal{F}_4^{\sigma_0}$, for σ_i with $i = 1, 2$, X_i consists of flags in $\mathcal{F}_5^{\sigma_i}$. Here we have $|\mathcal{F}_4^{\sigma_0}| = 20$ and $|\mathcal{F}_5^{\sigma_1}| = |\mathcal{F}_5^{\sigma_2}| = 71$. As we already mentioned, the flag algebra method is computer assisted. We use a computer program to find $\mathcal{M}_7, \mathcal{F}_4^{\sigma_0}, \mathcal{F}_5^{\sigma_1}, \mathcal{F}_5^{\sigma_2}$, and to compute $\mathbb{E}_\theta\ p(F, F'; (H, \theta))$ for each $H \in \mathcal{M}_y$. Then finding positive semidefinite matrices Q_0, Q_1, Q_2 to maximize $\min_{H\in\mathcal{M}_7}(f(H) - c_H)$ can be done by computer solvers such as CSDP [8] and SDPA [38]. Unfortunately, solvers can only give an approximate solution. For this problem, we get 0.0370370369999. In order to get exactly 1/27, we need to round the matrices Q_0, Q_1, Q_2 found by a computer solver. By rounding we mean finding rational matrices Q'_0, Q'_1, Q'_2 which would make the computations exactly 1/27 when computed over rational numbers.

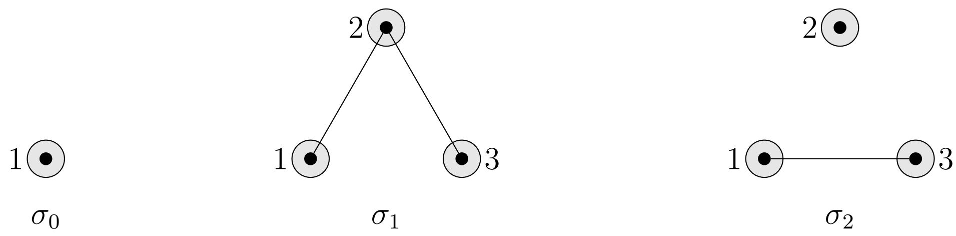


Figure 4: Types used in flag computation.

To simplify the process of rounding, we reduce the number of variables and constraints by restricting the set of feasible solutions. For $i \in \{1, 2, 3\}$ and flags F_1, F_2 denote by $Q_i(F_1, F_2)$ the entry in Q_i corresponding to indices of F_1 and F_2 in X_i. Since $f(H) = f(\overline{H})$ for every graph H, a natural restriction is that

$$f(H) - c_H = f(\overline{H}) - c_{\overline{H}} \tag{6}$$

for every graph H. This will allow us to consider only one of $H, \overline{H}$ and thus decrease the number of constraints from 776 to 388 since there is no self-complementary graph on 7 vertices as the number of possible edges and non-edges is $\binom{7}{2} = 21$ which is an odd number.

Since $\overline{\sigma_1} = \sigma_2$, we add the constraints $Q_1(F_1, F_2) = Q_2(\overline{F_1}, \overline{F_2})$ for every $F_1, F_2 \in \mathcal{F}_5^{\sigma_1}$. This makes Q_2 completely defined by Q_1. Moreover, we add the constraints $Q_0(F_1, F_2) = Q_0(\overline{F_1}, F_2) = Q_0(F_1, \overline{F_2}) = Q_0(\overline{F_1}, \overline{F_2})$ for every $F_1, F_2 \in \mathcal{F}_4^{\sigma_0}$. This reduces the number of entries to round in the symmetric matrix Q_0 from $\binom{21}{2}$ to $\binom{11}{2}$.

We reduced the number of constraints from 776 to 388, and we reduced the number of variables from $\binom{21}{2} + 2\binom{72}{2}$ to $\binom{11}{2} + \binom{72}{2}$. With these reductions, we managed to round the entries in Q_1, Q_2 and Q_3 and thus we obtained a solution for (5).

The rounded matrices as well as programs computing all possible X and performing the rounding process can be obtained at `http://www.math.uiuc.edu/~jobal/cikk/permutations/`.

We give more details about the rounding step in the appendix. □

In (5), we not only have that the minimum of $f(H) - c_H$ is $1/27$, which proves Theorem 5, but we also have the values of $f(H) - c_H$ for each H in $\mathcal{M}_7$.

Let $\mathcal{L} = \{H \in \mathcal{M}_7 : f(H) - c_H = 1/27\}$. We listed $\mathcal{L}$ in Figure 5. We have the following proposition for graphs not in $\mathcal{L}$.

Proposition 7. *Let G be an admissible graph of order $n \to \infty$ such that $f(G) = \frac{1}{27} + o(1)$. If $H \in \mathcal{M}_7 \setminus \mathcal{L}$, then $p(H, G) = o(1)$.*

Proof. Using (4), we have that

$$\frac{1}{27} + o(1) = f(G) + o(1) \geq \sum_{H \in \mathcal{M}_l} (f(H) - c_H) \cdot p(H, G).$$

In this section, we showed that by choosing $l = 7$ and types $\sigma_0, \sigma_1, \sigma_2$ we have $\min_{H\in\mathcal{M}_7}(f(H)-c_H) = 1/27$. Then since $\sum_{H\in\mathcal{M}_l} p(H,G) = 1$, we know that if $f(H) - c_H > 1/27$, then $p(H,G) = o(1)$. $\square$

Notice that the Proposition 7 can be stated equivalently as follows.

Proposition 8. *For every $\delta > 0$ there exists $n_0 = n_0(\delta)$ and $\varepsilon' > 0$ such that for every admissible graph G of order $n > n_0$ with $f(G) < 1/27+\varepsilon'$, if $H \in \mathcal{M}_7 \setminus \mathcal{L}$, then $p(H,G) < \delta$. Note that it is sufficent to pick*

$$\varepsilon' < \delta \cdot \min_{H\in\mathcal{M}_7\setminus\mathcal{L}} \{f(H) - c_H - 1/27\}.$$

Proposition 8 will help us to get the stability property of admissible graphs G with $f(G) = \frac{1}{27} + o(1)$, which is discussed in the next section.

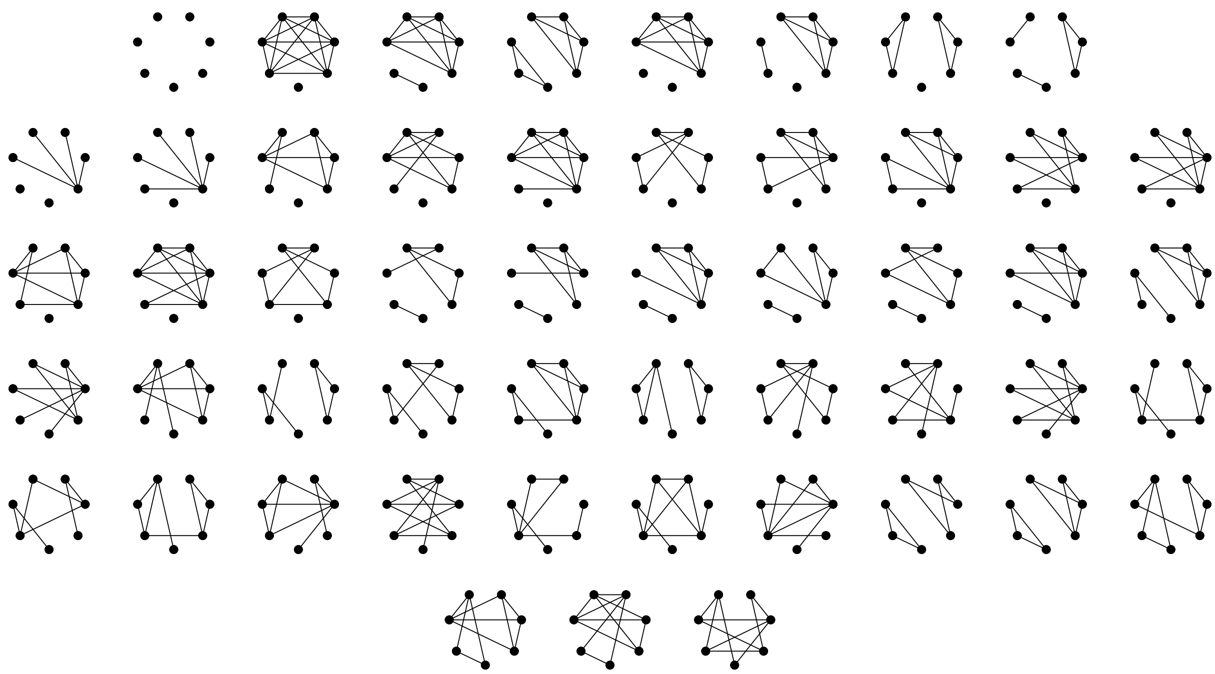


Figure 5: Graphs in $\mathcal{L}$. The first eight graphs are induced subgraphs of $T_3(n)$ or $\overline{T}_3(n)$. In order to save space, a depicted graph H represents both H and $\overline{H}$.

4 Stability Property

In this section we will prove the following stability type statement.

Theorem 9. *For every $\varepsilon > 0$ there exist n_0 and $\varepsilon' > 0$ such that every admissible graph G of order $n > n_0$ with $f(G) \leq \frac{1}{27} + \varepsilon'$, is isomorphic to either $T_3(n)$ or $\overline{T_3(n)}$ after adding and/or deleting at most $20\varepsilon n^2$ edges.*

We will use our flag algebra results from Section 3 and the infinite removal lemma to prove Theorem 9. The infinite removal lemma, proved by Alon and Shapira [2], is a substantial generalization of the induced removal lemma.

Lemma 10 (Infinite Removal Lemma [2])**.** *For any (possibly infinite) family $\mathcal{H}$ of graphs and $\varepsilon > 0$, there exists $\delta > 0$ such that if a graph G on n vertices contains at most $\delta n^{v(H)}$ induced copies of H for every graph H in $\mathcal{H}$, then it is possible to make G induced H-free, for every $H \in \mathcal{H}$, by adding and/or deleting at most εn^2 edges.*

Proof of Theorem 9. Fix an $\varepsilon > 0$. Let δ be from Lemma 10, when applied with ε and $\mathcal{H} = (\mathcal{M}_7 \setminus \mathcal{L}) \cup \{\text{not admissible graphs}\}$. Let $\varepsilon' < \varepsilon^4$ and n_1 be given by Proposition 8 such that $p(H, G) < \delta$ for every $H \in \mathcal{M}_7 \setminus \mathcal{L}$ and G on at least n_1 vertices. Let $n_0 > n_1$ such that $f(G) > 1/27 - \varepsilon'$ for all G of order at least $n_0/2$. Notice that every non-admissible graph H satisfies $p(H, G) = 0$ for every admissible G.

Let G be an admissible graph of order $n > n_0$ with $f(G) \leq \frac{1}{27} + \varepsilon'$. Now we apply Lemma 10 and conclude that by adding and/or deleting at most εn^2 edges, every induced subgraph of G on 7 vertices belongs to $\mathcal{L}$ and G is still admissible.

By direct inspection of graphs in $\mathcal{L}$, we have the following two properties of G. Notice that if all 7-vertex induced subgraphs of G satisfy these two properties, then so does G. Also notice that a graph H satisfies these two properties if and only if $\overline{H}$ satisfies them.

Property A: There are no K_4 and $\overline{K}_4$ that share a vertex.

Property B: For every pair of K_4's that share at least one vertex, the union of their vertex sets spans a clique. For every pair of $\overline{K}_4$'s that share at least one vertex, the union of their vertex sets spans an independent set.

Let (G, x) be the 1-flag where vertex x is the labeled vertex, then $P(K_4^1, (G, x))$ is the number of K_4's in G that contain x. Define

$$F(x, G) = P(K_4^1, (G, x)) + P(\overline{K}_4^1, (G, x)) \quad \text{and} \quad f(x, G) = F(x, G) \Big/ \binom{v(G) - 1}{3}.$$

Then we have $f(G) = \left(\sum_x f(x, G)\right)/v(G)$. Let $G_0 = G$. For $i \geq 0$, let x_i be the vertex with largest $f(x_i, G_i)$. If $f(x_i, G_i) > 1/27 + 2\varepsilon'/\varepsilon$, we create G_{i+1} from G_i by removing vertex x_i. If $f(x_i, G_i) \leq 1/27 + 2\varepsilon'/\varepsilon$, we define $G' = G_i$ and $d = i$. Note that $f(G_i) \leq f(G_{i-1})$, so $f(G_i) \leq f(G) \leq 1/27 + \varepsilon'$. Also notice, that the process is not deterministic if there are more candidates for x_i for some i (any choice of x_i will work).

Claim 11. $d < \varepsilon n$.

Proof. Denote $v = v(G_{i-1})$ and y the vertex deleted from G_{i-1}. Then

$$f(G_{i-1}) - f(G_i) \geq \frac{4\varepsilon'}{\varepsilon n}, \tag{7}$$

which follows from the following computation

$$
\begin{aligned}
f(G_{i-1}) - f(G_i) &= \frac{\sum_x f(x, G_{i-1})}{v} - \frac{\sum_x f(x, G_i)}{v-1} \\
&= \frac{f(y, G_{i-1}) + \sum_{x \neq y} f(x, G_{i-1})}{v} - \frac{\sum_x f(x, G_i)}{v-1} \\
&= \frac{f(y, G_{i-1}) + (\sum_{x \neq y} F(x, G_{i-1}))/\binom{v-1}{3}}{v} - \frac{\sum_x f(x, G_i)}{v-1} \\
&= \frac{4f(y, G_{i-1}) + \binom{v-2}{3}(\sum_x F(x, G_i))/(\binom{v-2}{3}\binom{v-1}{3})}{v} - \frac{\sum_x f(x, G_i)}{v-1} \\
&= \frac{4f(y, G_{i-1}) + \frac{v-4}{v-1}\sum_x f(x, G_i)}{v} - \frac{\sum_x f(x, G_i)}{v-1} \\
&= \frac{4(v-1)f(y, G_{i-1}) + (v-4)\sum_x f(x, G_i) - v\sum_x f(x, G_i)}{v(v-1)} \\
&= \frac{4f(y, G_{i-1}) - 4f(G_i)}{v} \\
&\geq \frac{4\left(2\varepsilon'/\varepsilon - \varepsilon'\right)}{n} \geq \frac{4\varepsilon'}{\varepsilon n}.
\end{aligned}
$$

If follows from $n \geq n_0$ that $f(H) > 1/27 - \varepsilon'$ for every admissible graph H on at least $n/2$ vertices. However, if $d > \varepsilon n$, then for $i = \varepsilon n$, $f(G_i) < 1/27 + \varepsilon' - 4i\varepsilon'/\varepsilon n < 1/27 - \varepsilon'$, which is a contradiction since G_i has at least $n - \varepsilon n = (1-\varepsilon)n \geq n/2$ vertices. □

Claim 12. The number of vertices x with $f(x, G') < \frac{1}{27} - \varepsilon$ is at most $\varepsilon v'$, where $v' = v(G')$.

Proof. Let the number of vertices with $f(x, G') < \frac{1}{27} - \varepsilon$ be z.

$$
\begin{aligned}
v' f(G') = \sum_x f(x, G') &< z\left(\frac{1}{27} - \varepsilon\right) + (v' - z)\left(\frac{1}{27} + \frac{2\varepsilon'}{\varepsilon}\right) \\
&= -z\varepsilon + v'\frac{1}{27} + v'\frac{2\varepsilon'}{\varepsilon} - z\frac{2\varepsilon'}{\varepsilon} < \frac{v'}{27} + \frac{2v'\varepsilon'}{\varepsilon} - \varepsilon z.
\end{aligned}
$$

If $z > \varepsilon v'$, then we get

$$
f(G') < \frac{1}{27} + \frac{2\varepsilon'}{\varepsilon} - \varepsilon^2 < \frac{1}{27} - \varepsilon',
$$

which is a contradiction (recall that $\varepsilon' < \varepsilon^4$). □

Let G'' be the graph obtained from G' by removing all such vertices. We removed at most $\varepsilon v'$ vertices, so

$$
F(x, G') - F(x, G'') < \varepsilon v' \binom{v'-2}{2}
$$

for each vertex $x \in V(G'')$. Denote $v(G'')$ by v''. We have $v'' \geq (1-\varepsilon)v'$ and

$$\begin{aligned}
F(x, G'') &> F(x, G') - \varepsilon v' \binom{v'-2}{2} \\
&\geq \left(\frac{1}{27} - \varepsilon\right)\binom{v'-1}{3} - \varepsilon v' \binom{v'-2}{2} \\
&\geq \left(\frac{1}{27} - \varepsilon\right)\frac{(v'-1)(v'-2)(v'-3)}{6} - \varepsilon \frac{(v'-1)(v'-2)(v'-3)}{1.5} \\
&\geq \left(\frac{1}{27} - 5\varepsilon\right)\binom{v''-1}{3}.
\end{aligned}$$

We know that $f(x, G') \leq 1/27 + 2\varepsilon'/\varepsilon$. Then since $\varepsilon' < \varepsilon^4$ and $v'' \geq (1-\varepsilon)v'$, we have

$$F(x, G'') \leq F(x, G') \leq \left(\frac{1}{27} + \frac{2\varepsilon'}{\varepsilon}\right)\binom{v'-1}{3} \leq \left(\frac{1}{27} + 5\varepsilon\right)\binom{v''-1}{3}.$$

This means that for every vertex $x \in V(G'')$, we have

$$\left(\frac{1}{27} - 5\varepsilon\right)\binom{v''-1}{3} < F(x, G'') < \left(\frac{1}{27} + 5\varepsilon\right)\binom{v''-1}{3}. \tag{8}$$

For $x, y \in V(G'')$, write $x \sim y$ if $x = y$ or there is a chain of vertex-intersecting K_4's or $\overline{K}_4$'s connecting x to y. Clearly, $\sim$ is an equivalence relation, and by Property A, each chain consists of cliques only or independent sets only. By Property B, each $\sim$-equivalence class is a clique or an independent set. Let s be the size of the class of x. This means that $F(x, G'') = \binom{s-1}{3}$. It follows from (8) that

$$\left(\frac{1}{3} - 16\varepsilon\right)v'' < s < \left(\frac{1}{3} + 16\varepsilon\right)v'', \tag{9}$$

which means each $\sim$-equivalence class has size at least $(1/3 - 16\varepsilon)v''$ and at most $(1/3 + 16\varepsilon)v''$.

Next, we claim that equivalence classes are all cliques or all independent sets. Suppose on the contrary that $G''[A]$ is a clique and $G''[B]$ is an independent set. W.l.o.g., assume that the edge density between A and B is at least $1/2$. Then there exists a vertex x in B such that $|\Gamma(x) \cap A| \geq |A|/2$. Taking a 4-set $X \subset B$ containing x and a 3-set $Y \subset \Gamma(x) \cap A$, then $G''[X] = \overline{K}_4$ and $G''[Y \cup \{x\}] = K_4$. We find a K_4 and a $\overline{K}_4$ that share a vertex x, contradicting Property A.

W.l.o.g., assume that each equivalence class is an independent set. It follows from (9) that there are exactly three equivalence classes. Denote them by A_1, A_2 and A_3. If there exist an $x \in A_i$ and $y_1, y_2, y_3 \in A_j$ $(i \neq j)$ such that none of xy_k is an edge, then $x \sim y_k$, which would contradict Property B. This means that all but at most $4v''$ of edges between equivalence classes are in G''. To get $T_3(v'')$, we need to add these edges and balance the three sets. In this step we change at most $4n + 16\varepsilon n^2$ edges.

We first change at most εn^2 edges of G such that G does not contain any $H \in \mathcal{M}_7 \setminus \mathcal{L}$, then we remove at most $2\varepsilon n$ vertices to form G''. Then we change at most $4n + 16\varepsilon n^2$ edges to get $T_3(v'')$. Therefore, to get $T_3(n)$ from G, we only need to change at most $\varepsilon n^2 + 2\varepsilon n^2 + 4n + 16\varepsilon n^2 \leq 20\varepsilon n^2$ edges, as required. $\square$

5 Exact Result

We call $u \in V(G)$ a *clone* of $v \in V(G)$ if $\Gamma(u) = \Gamma(v)$. In particular, uv is not an edge of G.

Proposition 13. *Let G be an admissible graph of order n. If we add a clone x' of some $x \in V(G)$ to form a new graph G' of order $n+1$, i.e., $\Gamma_{G'}(x') = \Gamma_G(x)$, then G' is still admissible.*

Proof. The graph G comes from some permutation τ on $[n]$. Let k be the number in $[n]$ that corresponds to x, then we can construct a new permutation τ' on $[n+1]$ as follows:

$$\tau'(i) = \begin{cases} \tau(i) & \text{if } i \leq k \text{ and } \tau(i) \leq \tau(k) \\ \tau(i)+1 & \text{if } i \leq k \text{ and } \tau(i) > \tau(k) \\ \tau(k)+1 & \text{if } i = k+1 \\ \tau(i-1) & \text{if } i > k \text{ and } \tau(i-1) < \tau(k) \\ \tau(i-1)+1 & \text{if } i > k \text{ and } \tau(i-1) \geq \tau(k). \end{cases}$$

The representation graph of τ' is G' with $k+1$ corresponding to the new vertex x'. □

Let S be the 7-vertex graph obtained by gluing three paths xy_iz_i, $i = 1, 2, 3$, at the common vertex x, see Figure 6.

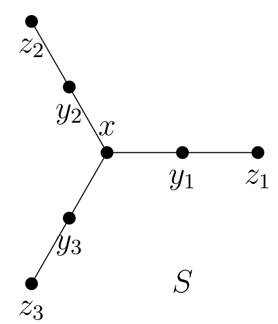


Figure 6: The graph S.

Proposition 14. *The graph S is not admissible.*

Proof. Admissible graphs can be alternatively defined as intersection graphs of segments whose endpoints lie on two parallel lines. For a vertex v in S denote by $s(v)$ the segment representing v. Since y_1, y_2 and y_3 form an independent set, segments representing them do not intersect. On the other hand $s(x)$ intersects all of them. W.l.o.g., assume that $s(y_2)$ is middle of the three segments in the order they intersect $s(x)$, see Figure 7. Since z_2 is adjacent only to y_2, the segment representing z_2 intersects only $s(y_2)$, which is clearly impossible.

Alternatively, it is possible to check all admissible graphs on 7 vertices, i.e., $\mathcal{M}_7$, and conclude that S is not among them. □

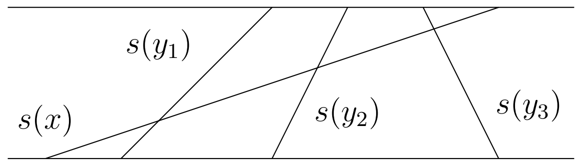


Figure 7: Representation of part of S as intersecting segments.

Proof of Theorem 6. Let G be an admissible graph of order n with minimum F among all admissible graphs on n vertices, where n is sufficiently large. Fix $\varepsilon > 0$ sufficiently small. In particular, by Theorem 5, we assume n large enough such that $f(G) \leq \frac{1}{27}+\varepsilon$. Let $V = V(G)$.

By Theorem 9, we also assume that we can make G equal to $T_3(n)$ by adding/or deleting at most εn^2 edges (also large n needed). Take a complete 3-partite graph T on V such that $|W|$ is minimized, where $W = E(T) \triangle E(G)$. From Theorem 9 we know that $|W| < \varepsilon n^2$ and V can be partitioned into V_1, V_2, V_3 of sizes $(1/3-\varepsilon)n < |V_i| < (1/3+\varepsilon)n$, which are the parts of T. Let $B = E(G) \setminus E(T)$ and $M = E(T) \setminus E(G)$. We call edges in B *bad*, in M *missing*, and in W *wrong*.

Proposition 15. *For any $x \in V$ we have $f(x, G) \leq 1/27 + \frac{3}{2}\varepsilon$.*

Proof. First we prove that for any $x, y \in V$ we have

$$|F(x,G) - F(y,G)| \leq \binom{n-2}{2}. \tag{10}$$

W.l.o.g., assume $F(x,G) \geq F(y,G)$. Let G' be obtained from G by adding a clone of y and removing x. By Proposition 13, G' is an admissible graph. By the extremality of G we have

$$0 \leq F(G') - F(G) \leq F(y,G) - F(x,G) + \binom{n-2}{2},$$

which gives (10).

Recall, that $F(G) = \frac{1}{4}\sum_{x\in V} F(x,G) \leq F(T_3(n))$. Suppose that there exists an x such that $f(x,G) > 1/27 + \frac{3}{2}\varepsilon$, i.e.

$$F(x,G) > \left(\frac{1}{27} + \frac{3}{2}\varepsilon\right)\binom{n-1}{3}.$$

By using (10) we have that for every $y \in V$

$$F(y,G) > \left(\frac{1}{27} + \frac{3}{2}\varepsilon\right)\binom{n-1}{3} - \binom{n-2}{2} > \left(\frac{1}{27} + \varepsilon\right)\binom{n-1}{3}.$$

Hence $f(G) > \frac{1}{27} + \varepsilon$ which contradicts our assumption that $f(G) \leq \frac{1}{27} + \varepsilon$. ∎

Lemma 16. *The graph W, and thus also B and M, has maximum degree less than ηn, where $\eta = 2\varepsilon^{1/18}$.*

Proof. Let x be a vertex in V. Let $\alpha_i = |\Gamma(x) \cap V_i|/|V_i|$ where V_1, V_2, V_3 are the parts of T. Let $\delta = 2\varepsilon^{1/6}$. If every $\alpha_i \in (\delta, 1-\delta)$ then there are at least $\delta^6(\frac{1}{3} - \varepsilon)^6 n^6$ ways to choose a set $U = \{y_1, y_2, y_3, z_1, z_2, z_3\}$ with $y_i \in V_i \setminus \Gamma(x) - x$ and $z_i \in V_i \cap \Gamma(x) - x$. The number of such sets U which contain a wrong edge is at most $\binom{n}{4}|W| < \varepsilon n^6/24$. Since $\delta^6(\frac{1}{3} - \varepsilon)^6 n^6 > \varepsilon n^6/24$, there exists U that does not contain any edge of W, which means $U \cup \{x\}$ induces the complement of the non-admissible graph S in G, a contradiction to Proposition 14.

W.l.o.g., we may assume that $\alpha_1 < \delta$ or $\alpha_1 > 1 - \delta$.

If $\alpha_1 < \delta$, then the number of copies of $\overline{K}_4$ via x whose other three vertices are in A_1 is at least

$$\binom{(1-\delta)(1/3-\varepsilon)n}{3} \geq (1-\delta)^3 \left(\frac{1}{3} - \varepsilon\right)^3 \binom{n}{3} \geq \left(\frac{1}{27} - \varepsilon - \delta\right)\binom{n}{3}.$$

Thus x has to be connected to almost every vertex in $A_2 \cup A_3$. To be more precise, assume $\alpha_i \leq 1 - \eta$ for for $i = 1$ or $i = 2$. Then we should have

$$\binom{\eta(1/3-\varepsilon)n}{3} \leq \binom{(1-\alpha_i)|V_i|}{3} < \left(\frac{1}{27} + \varepsilon\right)\binom{n}{3} - \left(\frac{1}{27} - \varepsilon - \delta\right)\binom{n}{3}. \tag{11}$$

However, (11) does not hold since $\eta = 2\varepsilon^{1/18}$, so we know that $\alpha_2 > 1 - \eta$ and $\alpha_3 > 1 - \eta$.

If $\alpha_1 > 1 - \delta$, then

$$\begin{aligned}
f(x, G) &\geq \left(\sum_{i=1}^{3} \binom{(1-\alpha_i)|V_i|}{3} + \alpha_1\alpha_2\alpha_3|V_1||V_2||V_3| - \varepsilon n^3\right) \Big/ \binom{n}{3} \\
&\geq \left(\binom{(1-\alpha_2)|V_2|}{3} + \binom{(1-\alpha_3)|V_3|}{3} + (1-\delta)\alpha_2\alpha_3|V_1||V_2||V_3| - \varepsilon n^3\right) \Big/ \binom{n}{3} \\
&= (1/3 - \varepsilon)^3 \left((1-\alpha_2)^3 + (1-\alpha_3)^3 + 6(1-\delta)\alpha_2\alpha_3\right) - 6\varepsilon \\
&\geq (1/3 - \varepsilon)^3 \left((1-\alpha_2)^3 + (1-\alpha_3)^3 + 5\alpha_2\alpha_3\right) - 6\varepsilon.
\end{aligned}$$

Let $h(x, y) = (1-x)^3 + (1-y)^3 + 5xy$. The minimum value of the polynomial $h(x, y)$ on $[0, 1]^2$ is 1 with equality if and only if $\{x, y\} = \{0, 1\}$. We know that $f(x, G) \leq 1/27 + \frac{3}{2}\varepsilon$. Then by the continuity of h and the compactness of $[0, 1]^2$, $\{\alpha_2, \alpha_3\}$ is close to $\{0, 1\}$. W.l.o.g., assume α_2 is close to 1 and α_3 is close to 0. Let $\gamma = 6\varepsilon^{1/3}$. If $\alpha_2 \leq 1 - \gamma$ or $\alpha_3 \geq \gamma$, then

$$\begin{aligned}
f(x, G) &\geq (1/3 - \varepsilon)^3 \left((1-\alpha_2)^3 + (1-\alpha_3)^3 + 5\alpha_2\alpha_3\right) - 6\varepsilon. \\
&= (1/3 - \varepsilon)^3 (1 + (2 - 5(1-\alpha_2))\alpha_3 + 3\alpha_3^2 - \alpha_3^3 + (1-\alpha_2)^3) - 6\varepsilon \\
&\geq (1/3 - \varepsilon)^3 (1 + \alpha_3 + (1-\alpha_2)^3) - 6\varepsilon \\
&> \frac{1}{27} + \frac{3}{2}\varepsilon,
\end{aligned}$$

which is a contradiction with Proposition 15. So $\alpha_2 > 1 - \gamma$ and $\alpha_3 < \gamma$.

Note that $\eta > \delta > \gamma$, so now we know that two of $\alpha_1, \alpha_2, \alpha_3$ are at least $1 - \eta$ and the other one is less than η. W.l.o.g., we may assume $\alpha_1 < \eta, \alpha_2 > 1 - \eta$ and $\alpha_3 > 1 - \eta$. Then we know that $x \in V_1$ since otherwise we can move x to V_1 and decrease $|W|$ which is a contradiction to the choice of T. Thus $d_W(x) < \eta n = 2\varepsilon^{1/18} n$. $\square$

It follows from Lemma 16 that every bad edge $xy \in B$ belongs to at least $(1/9-\eta)n^2$ copies of K_4, because

$$\left(\frac{1}{3}-\varepsilon\right)^2 n^2 - 2\eta n\left(\frac{1}{3}+\varepsilon\right)n = \left(\frac{1}{9}-\frac{2}{3}\varepsilon+\varepsilon^2-\frac{2\eta}{3}-2\eta\varepsilon\right)n^2 \geq \left(\frac{1}{9}-\eta\right)n^2.$$

On the other hand, if we remove xy from $E(G)$, this would create at most $(1/18+\eta^2)n^2$ copies of $\overline{K}_4$, because

$$\binom{(1/3+\varepsilon)\,n}{2} + \binom{\eta n}{2} \leq \left(\frac{1}{18}+\frac{\varepsilon}{3}+\frac{\varepsilon^2}{2}+\frac{\eta^2}{2}\right)n^2 \leq \left(\frac{1}{18}+\eta^2\right)n^2.$$

Also, by $\Delta(B) < \eta n$ and $b = |B| < \eta n^2$, the number of 4-sets that contain at least two bad edges is at most

$$\binom{b}{2} + 2b(\eta n)n \leq \frac{b^2}{2} + 2\eta b n^2 < \frac{\eta b n^2}{2} + 2\eta b n^2 < 3\eta b n^2.$$

Thus if G' is obtained from G by removing all bad edges of G, it satisfies $F(G') - F(G) \leq -bn^2/18 + \varepsilon b n^2 < 0$ unless $b = 0$, because

$$\begin{aligned} F(G') - F(G) &\leq \left(\frac{1}{18}+\eta^2\right)bn^2 - \left(\left(\frac{1}{9}-\eta\right)bn^2 - 6\cdot 3\eta b n^2\right) \\ &\leq \left(-\frac{1}{18}+\eta^2+19\eta\right)bn^2 \leq -\frac{bn^2}{100}. \end{aligned}$$

Clearly, the complete 3-partite graph T can be obtained from G' by adding all missing edges between parts. Thus we have $P(\overline{K}_4, T) \leq P(\overline{K}_4, G')$. Then since $P(K_4, T) = 0 \leq P(K_4, G')$, the admissible graph T satisfies $F(T) \leq F(G') < F(G)$ unless $b = 0$. By the choice of G, we have $F(G) \leq F(T)$, so $b = 0$, which means G is a subgraph of T and G is a 3-partite graph. Then since $F(G) \leq F(H)$ for every $H \in \mathcal{M}_n$, we know that G is a subgraph of $T_3(n)$ and $F(G) = F(T_3(n))$. $\square$

6 Conclusion

In Theorem 4, we verified Conjecture 1 for $k = 3$ and n sufficiently large, and we fully characterized the set of the extremal configurations. While revising our paper, we discovered that using a slightly refined set-up of the flag algebra framework, we can also prove an analogue of Theorem 3 for $k = 4$ and $k = 5$. We address these two cases in a forthcoming note.

Acknowledgement

We would like to thank Dan Kráľ for fruitful discussions and developing the first version of the program implementing flag algebras over permutations. We would also like to thank Andrew

Treglown for fruitful discussions. We also thank the anonymous referees for carefully reading the manuscript and for their valuable comments, which greatly improved the presentation of the results.

References

[1] M. H. Albert, M. D. Atkinson, C. C. Handley, D. A. Holton, and W. Stromquist. On packing densities of permutations. *Electron. J. Combin.*, 9(1):Research Paper 5, 20, 2002.

[2] N. Alon and A. Shapira. A characterization of the (natural) graph properties testable with one-sided error. *SIAM J. Comput.*, 37(6):1703–1727, 2008.

[3] R. Baber. Turán densities of hypercubes. Submitted.

[4] R. Baber. Some results in extremal combinatorics. Dissertation, 2011.

[5] R. Baber and J. Talbot. A solution to the 2/3 conjecture. Submitted.

[6] R. Baber and J. Talbot. Hypergraphs do jump. *Combin. Probab. Comput.*, 20(2):161–171, 2011.

[7] J. Balogh, P. Hu, B. Lidický, and H. Liu. Upper bounds on the size of 4- and 6-cycle-free subgraphs of the hypercube. *European Journal of Combinatorics*, 35:75–85, 2014.

[8] B. Borchers. CSDP, A C library for semidefinite programming. *Optimization Methods and Software*, 11(1-4):613–623, 1999.

[9] J. Cummings, D. Kráľ, F. Pfender, K. Sperfeld, A. Treglown, and M. Young. Monochromatic triangles in three-coloured graphs. *J. Combin. Theory Ser. B*, 103(4):489–503, 2013.

[10] S. Das, H. Huang, J. Ma, H. Naves, and B. Sudakov. A problem of Erdős on the minimum number of k-cliques. *J. Combin. Theory Ser. B*, 103(3):344–373, 2013.

[11] P. Erdős and G. Szekeres. A combinatorial problem in geometry. *Compositio Math.*, 2:463–470, 1935.

[12] V. Falgas-Ravry and E. R. Vaughan. Applications of the semi-definite method to the Turán density problem for 3-graphs. *Combin. Probab. Comput.*, 22(1):21–54, 2013.

[13] G. Giraud. Sur le problme de goodman pour les quadrangles et la majoration des nombres de ramsey. *Journal of Combinatorial Theory, Series B*, 27(3):237–253, 1979.

[14] R. Glebov, D. Kráľ, and J. Volec. A problem of Erdős and Sós on 3-graphs. Submitted.

[15] A. W. Goodman. Sur le problme de goodman pour les quadrangles et la majoration des nombres de ramsey. *The American Mathematical Monthly*, 66(9):778–783, 1959.

[16] A. Grzesik. On the maximum number of five-cycles in a triangle-free graph. *J. Combin. Theory Ser. B*, 102(5):1061–1066, 2012.

[17] H. Hatami, J. Hladký, D. Král', S. Norine, and A. Razborov. On the number of pentagons in triangle-free graphs. *J. Combin. Theory Ser. A*, 120(3):722–732, 2013.

[18] H. Hatami, J. Hladký, D. Král', S. Norine, and A. Razborov. Non-three-colourable common graphs exist. *Combin. Probab. Comput.*, 21(5):734–742, 2012.

[19] J. Hladký, D. Král', and S. Norine. Counting flags in triangle-free digraphs. In *European Conference on Combinatorics, Graph Theory and Applications (EuroComb 2009)*, volume 34 of *Electron. Notes Discrete Math.*, pages 621–625. Elsevier Sci. B. V., Amsterdam, 2009.

[20] D. Král', C.-H. Liu, J.-S. Sereni, P. Whalen, and Z. B. Yilma. A new bound for the 2/3 conjecture. *Combin. Probab. Comput.*, 22(3):384–393, 2013.

[21] D. Král', L. Mach, and J.-S. Sereni. A new lower bound based on Gromov's method of selecting heavily covered points. *Discrete Comput. Geom.*, 48(2):487–498, 2012.

[22] J. S. Myers. The minimum number of monotone subsequences. *Electron. J. Combin.*, 9(2):RP 4, 17 pp., 2002/03. Permutation patterns (Otago, 2003).

[23] S. Nieß. Counting monochromatic copies of k_4: a new lower bound for the ramsey multiplicity problem. Submitted.

[24] O. Pikhurko. The minimum size of 3-graphs without a 4-set spanning no or exactly three edges. *European J. Combin.*, 32(7):1142–1155, 2011.

[25] O. Pikhurko and A. Razborov. Asymptotic structure of graphs with the minimum number of triangles. E-print `arXiv:1204.2846`.

[26] O. Pikhurko and E. R. Vaughan. Minimum number of k-cliques in graphs with bounded independence number. *To appear in Combinatorics, Probability and Computing.*

[27] C. B. Presutti. Determining lower bounds for packing densities of non-layered patterns using weighted templates. *Electron. J. Combin.*, 15(1):RP 50, 10, 2008.

[28] C. B. Presutti and W. Stromquist. Packing rates of measures and a conjecture for the packing density of 2413. In *Permutation patterns*, volume 376 of *London Math. Soc. Lecture Note Ser.*, pages 287–316. Cambridge Univ. Press, Cambridge, 2010.

[29] A. Razborov. On the Caccetta-Häggkvist conjecture with forbidden subgraphs. Submitted.

[30] A. A. Razborov. Flag algebras. *Journal of Symbolic Logic*, 72(4):1239–1282, 2007.

[31] A. A. Razborov. Flag Algebras: An Interim Report. In *The Mathematics of Paul Erdős II*, pages 207–232. Springer, 2013.

[32] W. Samotij and B. Sudakov. On the number of monotone sequences. Submitted.

[33] K. Sperfeld. The inducibility of small oriented graphs. Submitted.

[34] K. Sperfeld. On the minimal monochromatic k4-density, 2011.

[35] W. Stein et al. *Sage Mathematics Software (Version 5.6)*. The Sage Development Team, 2013. `http://www.sagemath.org`.

[36] A. Thomason. A disproof of a conjecture of Erdős in Ramsey theory. *J. London Math. Soc. (2)*, 39(2):246–255, 1989.

[37] E. R. Vaughan. Flagmatic (version 2.0), 2013. `http://www.flagmatic.org`.

[38] M. Yamashita, K. Fujisawa, K. Nakata, M. Nakata, M. Fukuda, K. Kobayashi, and K. Goto. A high-performance software package for semidefinite programs: Sdpa 7, 2010.

A Rounding approximate solutions to exact solutions

Recall from the proof of Theorem 5 that a solution consists of several positive semidefinite matrices. For example, in our problem, our solution consists of three matrices Q_0, Q_1 and Q_2. For simplicity, when describing the rounding procedure, we assume that there is only one matrix. A computer solver can only solve a semidefinite program numerically and thus we get an approximate solution. Let Q be a t-by-t matrix computed by a computer solver. To make the solution exact, we need to convert entries in the matrix to rational numbers. A resulting *rounded* matrix Q' must satisfy the following.

Goal 1. Q' is positive semidefinite.

Goal 2. Q' gives us the desired number, i.e., see (5).

The idea of the rounding is following. For most of entries in Q' we use a rational number close to the corresponding entry in Q. The other entries in Q' will be computed such that Q' satisfies Goals 1 and 2.

We will construct a system of linear equations whose variables are entries of Q' (ignoring the entries below the main diagonals) and all constants are rational numbers. There are two types of linear equations in the system, Type 1 and Type 2, which make our solution achieve Goal 1 and Goal 2 respectively. We again use computer solver to solve the linear system, but unlike a semidefinite program, a system of linear equations can be solved over rationals.

When we use an entry from Q, it is sufficient in our case if the corresponding entry in Q' differs by at most $\varepsilon = 10^{-5}$.

To achieve Goal 1, we want all eigenvalues to be non-negative. We know that Q is positive semidefinite, so all its eigenvalues are non-negative. If an eigenvalue of Q is a large positive number compared to ε, then we expect it to be still positive after rounding, since as we mentioned above, entries of Q are perturbed just a little bit. But if an eigenvalue of Q is small, for example, 10^{-6}, then it may become -10^{-8} after rounding and Q' would not be positive semidefinite. To avoid this, we force such eigenvalues to become 0 after rounding. We do this by adding a constraint to our linear system for every such eigenvalue. Let $\{X_i\}$ be the set of eigenvectors of Q whose eigenvalues are smaller than ε_1 for some $\varepsilon_1 > 0$. We assume that X_i is close to an eigenvector of Q' with eigenvalue 0. So we find an approximate basis $\{X_i'\}$ of the linear space generated by $\{X_i\}$, and add $Q'X_i' = 0$ to our linear system. These are Type 1 linear equations. Note that entries of Q' are variables, so this gives us t linear equations for each X_i'. Let $X_i = [x_{i,1}, \ldots, x_{i,t}]$. The algorithm of finding X_i' is outlined below, which is taken from Baber's Thesis [4]:

For each X_i:
  Let ℓ be $\arg\max_j |x_{i,j}|$.
  Set $X_i = X_i / x_{i,\ell}$.
  For all $k \neq i$:
    Set $X_k = X_k - x_{k,\ell} \cdot X_i$.
Guess X_i' from X_i by assuming that all entries of X_i' are rational numbers.

More details of the algorithm are in Section 2.4.2.2 of [4]. The last step of the algorithm means that X_i should look good and one can see instantly from X_i what the exact value is. For example, if one sees 0.33333332, then 1/3 should be guessed.

To achieve Goal 2, we check values of $f(H) - c_H$ for all H in $\mathcal{M}_l$. If $f(H) - c_H$ is much larger than 1/27, we hope that it will be still larger than 1/27 after rounding. However, if $f(H) - c_H$ is close to 1/27, a small change on entries of Q could result in $f(H) - c_H$ being less than 1/27, which violates Goal 2. To prevent this, we add a linear equation $f(H) - c_H = 1/27$ for every $H \in \mathcal{M}_l$ if $f(H) - c_H < 1/27 + \varepsilon_2$ for some $\varepsilon_2 > 0$. These are Type 2 linear equations.

The system of k Type 1 and Type 2 linear equations can be written as

$$Ay = b,$$

where $y = \{y_1, \ldots, y_m\}$ corresponds to entries of Q', $A \in \mathbb{Q}^{k \times m}$, and $b \in \mathbb{Q}^k$. Usually, m is larger than $r = rank(A)$. W.l.o.g., assume that the first r columns of A are linearly independent. Then A can be written as $\begin{bmatrix} A' & A'' \end{bmatrix}$ where A' is the first r columns of A and A'' is the rest of the columns of A. Let $y' = \{y_1, \ldots, y_r\}$ and $y'' = \{y_{r+1}, \ldots, y_m\}$. We assign to y_i in y'' a rational number, such that $|y_i - x_i| < \varepsilon_3$, where $x_i \in Q$ corresponds to y_i and $\varepsilon_3 > 0$. This step can be done arbitrarily. For example, let $\varepsilon_3 = 10^{-5}$ and keep the first 5 digits of x_i. Then we have the following matrix equation:

$$A'y' = b - A''y''. \tag{12}$$

Note that the number of equations in (12) may be larger than r. So this system may have no solution. But if it has a solution, then this solution is unique, which gives a matrix over rational numbers. Then we need to verify if this matrix satisfies Goals 1 and 2. If yes, we get Q'. If not, we can try to redo the computation with a smaller ε_3, or look which of the goals is violated and enlarge ε_1 or ε_2 to add more equations to the linear system.

If we are unlucky that the linear system (12) has no solution, then it means we added too many equations. Note that we pick eigenvalues that are smaller than ε_1 and add corresponding Type 1 equations, and pick H with $f(H) - c_H < 1/27 + \varepsilon_2$ and add corresponding Type 2 equations. In order to have fewer equations, we may re-pick Type 1 and Type 2 equations with smaller ε_1 and ε_2.

So far in our rounding procedure, we get Type 1 and Type 2 equations only from Q. If the attacked problem has conjectured extremal structures, we can also get Type 1 and Type 2 equations from those structures.

Take our problem for example. Let G be an extremal graph on n vertices. By Proposition 7, if $p(H, G) > o(1)$, then $f(H) - c_H = 1/27$, which gives Type 2 equations. For our problem, this gives the first eight graphs in Figure 5. Unsurprisingly, every H of these eight graphs satisfy $f(H) - c_H \approx 1/27$ in Q. So these Type 2 equations are usually generated from Q by the process described before.

For Type 1 equations, using (4), we have

$$1/27 + o(1) = f(G) + o(1) \geq \sum_{H \in \mathcal{M}_7} (f(H) - c_H) \cdot p(H, G)$$

and

$$f(G) - \sum_{H \in \mathcal{M}_7} c_H p(H,G) = \sum_{H \in \mathcal{M}_7} (f(H) - c_H) \cdot p(H,G) \geq \min_{H \in \mathcal{M}_7} (f(H) - c_H) = 1/27.$$

This gives $\sum_{H \in \mathcal{M}_7} c_H p(H,G) = o(1)$. Recall from Section 3 that we use (σ_i, X_i, Q_i) to get c_H. Denote $X_i = \{F_j^i\}$. For $\theta \in \Theta(|\sigma_i|, G)$, let $Y_{i,\theta}$ be the vector whose entries are $p(F_j^i, (G,\theta))$. It follows from (3) and the definition of c_H that

$$o(1) = \sum_{H \in \mathcal{M}_7} c_H p(H,G) = \sum_i E_{\theta \in \Theta(|\sigma_i|,G)} Y_{i,\theta}^T \cdot Q_i \cdot Y_{i,\theta}. \tag{13}$$

For each $\theta \in \Theta(|\sigma_i|, G)$ we have a vector $Y_{i,\theta}$, but if the conjectured extremal structures are symmetric in some sense, then there may be only C different $Y_{i,\theta}$'s where C is a constant independent of n. Choose θ from $\Theta(|\sigma_i|, G)$ uniformly at random. If $P[Y_{i,\theta} = Y_{i,\phi}] > o(1)$ for some $\phi \in \Theta(|\sigma_i|, G)$, then we have $Y_{i,\phi}^T \cdot Q_i \cdot Y_{i,\phi} = 0$, otherwise we do not have (13). Since $Q_i \succeq 0$, this means that $Y_{i,\phi}$ is an eigenvector of Q_i with eigenvalue 0, giving us Type 1 equations. In our problem, the vectors $\{Y_{i,\phi}\}$ we get from conjectured extremal structures are in the space generated by $\{X_i'\}$. So there is no need to combine equations generated from these two methods. Let us mention that, for our problem, we could not guarantee that the rounded matrix is positive definite by just using Type 1 equations that come from the Turán graph. We also needed Type 1 equations from the numerical solution.