Combinatorics, Probability and Computing (2019), 1–24
doi:10.1017/S0963548319000257

CAMBRIDGE
UNIVERSITY PRESS

ARTICLE

Minimizing the number of 5-cycles in graphs with given edge-density

Patrick Bennett[1,†], Andrzej Dudek[1,*,‡], Bernard Lidický[2,§] and Oleg Pikhurko[3,¶]

[1]Department of Mathematics, Western Michigan University, Kalamazoo, MI 49008, USA, [2]Department of Mathematics, Iowa State University, Ames, IA 50011, USA and [3]Mathematics Institute and DIMAP, University of Warwick, Coventry CV4 7AL, UK

*Corresponding author. Email: andrzej.dudek@wmich.edu

(Received 1 March 2018; revised 4 June 2019)

Abstract

Motivated by the work of Razborov about the minimal density of triangles in graphs we study the minimal density of the 5-cycle C_5. We show that every graph of order n and size $(1-1/k)\binom{n}{2}$, where $k \geqslant 3$ is an integer, contains at least

$$\left(\frac{1}{10}-\frac{1}{2k}+\frac{1}{k^2}-\frac{1}{k^3}+\frac{2}{5k^4}\right)n^5+o(n^5)$$

copies of C_5. This bound is optimal, since a matching upper bound is given by the balanced complete k-partite graph. The proof is based on the flag algebras framework. We also provide a stability result. An SDP solver is not necessary to verify our proofs.

2010 MSC Codes: 05C35

1. Introduction

It is believed that *extremal graph theory* was started by Turán [29] when he proved that any graph on n vertices with more than

$$\frac{r-2}{2(r-1)}n^2$$

edges must contain a copy of K_r (*i.e.* a clique with r vertices). The case $r=3$ was earlier proved by Mantel [17]. The general *Turán problem* is to determine the minimum number $\mathrm{ex}(n,H)$ of edges in an n vertex graph that guarantees a copy of a graph H, and has been very widely studied. The Erdős–Stone theorem [6] was a major breakthrough which asymptotically determined the value of $\mathrm{ex}(n,H)$ for all non-bipartite H. For such H we have

$$\mathrm{ex}(n,H)=\frac{\chi(H)-2}{2(\chi(H)-1)}n^2+o(n^2).$$

†Supported in part by Simons Foundation grant 426894.
‡Supported in part by Simons Foundation grant 522400.
§Supported in part by NSF grant DMS-1600390.
¶Supported in part by ERC grant 306493.

The natural quantitative question that arises is how many copies of H must be contained in a graph G on n vertices with $m > \mathrm{ex}(n, H)$ edges. This question has also been well studied. Obviously the number of edges m can be expressed as a density parameter p such that $m = p\binom{n}{2}$. Therefore, we will use the following notation. Let G be a (large) graph of order n and H a small one. Define $\nu_H(G)$ to be the number of unlabelled copies (not necessary induced) of H in G and the corresponding density as

$$d_H(G) = \frac{\nu_H(G)}{|V(G)|^{|V(H)|}}.$$

Furthermore, for a given number $p \in [0, 1]$ let

$$d_H(p) = \lim_{n\to\infty} \min_G d_H(G),$$

where the minimum is taken over all graphs G of order n and size $(p + o(1))\binom{n}{2}$. It is not hard to show by double-counting that the limit exists; see *e.g.* [23, Lemma 2.2].

When $H = K_3$ (that means it is a triangle), Moon and Moser [18] and also independently Nordhaus and Stewart [20] determined $d_{K_3}(p)$ for any $p = 1 - 1/k$, where k is a positive integer. We call such $p = 1 - 1/k$ a *Turán density*. Some other partial results for the general r-clique $H = K_r$ were established by Lovász and Simonovits [14]. However, for arbitrary p these problems remained open for over 50 years.

Then Razborov in his seminal paper [25] introduced the so-called *flag algebras* and, using them, determined $d_{K_3}(p)$ for any p in [26]. Subsequently, Pikhurko and Razborov [22] characterized all almost extremal graphs. Very recently, Liu, Pikhurko and Staden [12] found the precise minimum number of triangles among graphs with a given number of edges in almost all ranges. Nikiforov [19] determined $d_{K_4}(p)$ for all p, and then Reiher [27] determined $d_{K_r}(p)$ for all r and p.

In this paper we address the minimum density of the 5-cycle, C_5, in a graph with given edge density. We chose to investigate C_5 instead of C_4 since it is known due to Sidorenko [28] that for any fixed constant edge density p, the minimum C_4-density is achieved asymptotically by the random graph $G_{n,p}$. It is worth mentioning some other research related to 5-cycles. Specifically, Grzesik [8] and independently Hatami, Hladký, Král', Norine and Razborov [9] proved that the maximum density of 5-cycles in a triangle-free graph that is large or its number of vertices is a power of 5 is achieved by the balanced blow-up of a 5-cycle. The extension to graphs of all sizes, with one exception on 8 vertices, was done by Lidický and Pfender [10]. This settled in the affirmative a conjecture of Erdős [5]. On the other hand, Balogh, Hu, Lidický, and Pfender [2] studied the problem of maximizing induced 5-cycles, and proved that this is achieved by the balanced iterated blow-up of a 5-cycle. This confirmed a special case of a conjecture of Pippinger and Golumbic [24].

The main result of this paper is as follows.

Theorem 1.1. *Let $k \geqslant 3$ be an integer. Define*

$$p = 1 - \frac{1}{k} \quad \text{and} \quad \lambda = \frac{1}{10} - \frac{1}{2k} + \frac{1}{k^2} - \frac{1}{k^3} + \frac{2}{5k^4}. \tag{1.1}$$

Then

$$d_{C_5}(p) = \lambda.$$

We also have the following stability result. Let the *Turán graph* T_k^n be the complete k-partite graph on n vertices with part sizes as equal as possible.

Theorem 1.2. *For every integer $k \geqslant 3$ and real $\delta > 0$ there is $\varepsilon > 0$ such that every graph G with $n \geqslant 1/\varepsilon$ vertices, at least $(p-\varepsilon)\binom{n}{2}$ edges and at most $(\lambda+\varepsilon)n^5$ copies of C_5 is within edit distance δn^2 from the Turán graph T_k^n, where p and λ are as in (1.1).*

Observe that the above theorems (as stated) also hold in the case $k = 2$ for which $d_{C_5}(1/2) = 0$. However, their validity in this case easily follows from known standard results. Although the proofs of Theorems 1.1 and 1.2 are based on the flag algebras framework, their verification does not require using any SDP solver.

Theorems 1.1 and 1.2 are proved in Sections 2 and 3 respectively. Finally, in Section 4, we discuss the general edge density and provide an upper bound on $d_{C_5}(p)$ for any $p \in [0, 1]$.

2. Proof of the main theorem

2.1 Upper bound

By considering the sequence of graphs T_k^n as $n \to \infty$, we get

$$d_{C_5}(T_k^n) = \frac{\left[\frac{1}{10}(k)_5 + \frac{1}{2}(k)_4 + \frac{1}{2}(k)_3\right](n/k)^5}{n^5} + o(1),$$

where $(k)_\ell = k(k-1)\cdots(k-\ell+1)$ is the *falling factorial*. To justify the numerator, we count the number of C_5 copies with vertices in parts V_1, V_2, V_3, V_4, V_5 of the partition. These parts may not all be distinct: for example we may have $V_1 = V_3$. However T_k^n has no edges within these parts and so we know $V_i \neq V_{i+1}$. We count copies of C_5 by grouping them according to how many distinct parts there are among $V_1, \ldots, V_5$. Now there are asymptotically $\frac{1}{10}(k)_5(n/k)^5$ copies that hit 5 different parts (label 5 distinct parts, choose one vertex in each part, and divide by 10 for overcounting). Also, there are asymptotically $\frac{1}{2}(k)_4(n/k)^5$ copies hitting 4 parts, and $\frac{1}{2}(k)_3(n/k)^5$ copies hitting 3 parts.

Simplifying, we get that $d_{C_5}(T_k^n) = \lambda + o(1)$, which implies the upper bound in Theorem 1.1.

2.2 Lower bound

2.2.1 Preliminaries

The proof of the lower bound in Theorem 1.1 relies on the celebrated flag algebra method introduced by Razborov [25]. Here we briefly discuss the main idea behind this approach, referring the reader to [25] for all details. Alternatively, our lower bound is rephrased at the beginning of Section 3 by means of a combinatorial identity (namely (3.2)) whose statement does not use any flag algebra formalism.

Let $(G_n)_{n\in\mathbb{N}}$ be a sequence of graphs, such that order of G_n increases. Such a sequence is called *convergent* if, for every fixed graph H, the density of H in G_n converges, that is, for every H there exists some number $\phi(H)$, such that

$$\lim_{n\to\infty} p(H, G_n) = \phi(H),$$

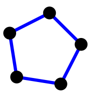
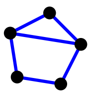
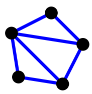
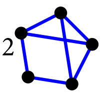
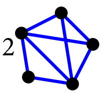
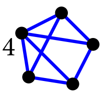
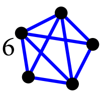
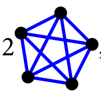
where $p(H, G)$ is the probability that $|H| = |V(H)|$ vertices chosen uniformly at random from $V(G)$ induce a copy of H. (Here, it will be more convenient to count induced copies of H; see *e.g.* equations (5.19)–(5.21) in [13], which show how to switch between induced and non-induced versions.) Notice that any sequence of graphs whose orders increase has a convergent subsequence. Thus, without loss of generality we assume G_n is convergent. Note that ϕ cannot be an arbitrary function since it must satisfy many obvious identities such as $\phi(\text{edge}) + \phi(\text{non-edge}) = 1$.

Interestingly, these ϕ exactly correspond to homomorphisms that we now describe. Denote by $\mathcal{F}$ the set of all graphs and by $\mathcal{F}_\ell$ the set of graphs of order ℓ, up to an isomorphism. Let $\mathbb{R}\mathcal{F}$ be the set of all finite formal linear combinations of graphs in $\mathcal{F}$ with real coefficients. It comes with the natural operations of addition and multiplication by a real number. Let $\mathcal{K}$ be a linear subspace generated by all linear combinations

$$F - \sum_{H \in \mathcal{F}_\ell} p(F, H) \cdot H, \tag{2.1}$$

where $\ell > |F|$. Notice that ϕ evaluated at any element of $\mathcal{K}$ gives 0. Finally, let $\mathcal{A}$ be $\mathbb{R}\mathcal{F}$ factorized by $\mathcal{K}$. It is possible to define multiplication on $\mathcal{A}$, which we do in Section 2.2.3. It can be proved that $\mathcal{A}$ is indeed an algebra. Now limits of convergent graph sequences correspond to homomorphism ϕ from $\mathcal{A}$ to $\mathbb{R}$ such that $\phi(F) \geqslant 0$ for all $F \in \mathcal{F}$. Denote the set of all such homomorphisms by $\mathrm{Hom}^+(\mathcal{A}, \mathbb{R})$.

Let OPT be the following linear combination, which counts the C_5 copies using induced subgraphs:

OPT =  +  +  + 2  + 2  + 4  + 6  + 12 ,

where the coefficient of each graph is the number of copies of C_5 it contains. Thus,

$$\phi(\mathrm{OPT}) = 120 \lim_{n\to\infty} d_{C_5}(G_n). \tag{2.2}$$

The factor $120 = 5!$ comes from the fact that $p(F, G_n)$ for $F \in \mathcal{F}_5$ is the number of copies of F divided by $\binom{n}{5}$ whereas our scaling for d_{C_5} was chosen as n^{-5}. Notice that OPT can be written as a linear combination of all 34 graphs on 5-vertices, where 26 graphs have coefficient 0. Namely,

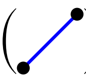
$$\mathrm{OPT} = \sum_{F \in \mathcal{F}_5} c_F^{\mathrm{OPT}} F, \tag{2.3}$$

where the non-zero entries c_F^{OPT} are as above.

Our goal is to prove a good lower bound on

$$\min_{\phi \in \mathrm{Hom}^+(\mathcal{A},\mathbb{R})} \phi(\mathrm{OPT}),$$

given that that the edge density is p, that is, we have

$\phi\Big($  $\Big) = p.$ (2.4)

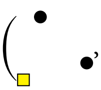
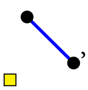
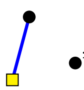
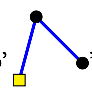
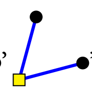
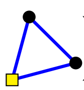
For this we find suitable $A \in \mathcal{A}$, such that $\phi(A) \geqslant 0$ for all $\phi \in \mathrm{Hom}^+(\mathcal{A}, \mathbb{R})$ with $\phi(K_2) = p$, and use it in calculations. In particular, we will use it as

$$\phi(\mathrm{OPT}) \geqslant \phi(\mathrm{OPT}) - \phi(A) = \phi(\mathrm{OPT} - A) \geqslant c,$$

where c is the smallest coefficient c_F when we express $\mathrm{OPT} - A$ as $\sum_{F \in \mathcal{F}_\ell} c_F F$. Note that A may contain both positive and negative coefficients, and these coefficients combine with coefficients in OPT.

When $p = 1 - 1/k$ for integer $k \geqslant 3$, it is possible to prove the sharp lower bound as above by considering graphs of order 5 with only one labelled vertex. Similarly to defining the algebra $\mathcal{A}$ and limits of convergent graph sequences, one can define limits of sequences from the set $\mathcal{F}^1$ which consists of graphs with exactly one labelled vertex up a label-preserving isomorphism. This gives an algebra $\mathcal{A}^1$ and homomorphisms $\mathrm{Hom}^+(\mathcal{A}^1, \mathbb{R})$. In the following, we depict the labelled vertex by a square.

Let X be the following column vector:

$X = (X_1, \ldots, X_6)^T = \Big($ , , , , ,  $\Big)^T.$ (2.5)

Notice that X is the vector of all graphs on 3 vertices with exactly one labelled vertex (the yellow square). For isomorphism, the labelled vertex must be preserved but the remaining two vertices may be swapped. If M is a positive semidefinite matrix in $\mathbb{R}^{6\times 6}$, then for every $\phi^1 \in \mathrm{Hom}^+(\mathcal{A}^1, \mathbb{R})$ it holds that

$$\phi^1(X^T MX) = \phi^1(X^T)M\phi^1(X) \geqslant 0,$$

where by $\phi^1(X)$ we mean the vector that results from applying ϕ^1 to each coordinate of X.

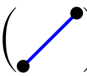
Also, there is a linear operator $[\![\cdot]\!]_1 : \mathbb{R}\mathcal{F}^1 \to \mathbb{R}\mathcal{F}$ (which, roughly speaking, 'unlabels' each $F \in \mathcal{F}^1$) such that for all $\phi \in \mathrm{Hom}^+(\mathcal{A}, \mathbb{R})$ we have $\phi([\![X^T MX]\!]_1) \geqslant 0$. Furthermore, we have

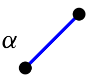
$$[\![X^T MX]\!]_1 = \sum_{F\in\mathcal{F}_5} c_F^M \cdot F. \tag{2.6}$$

See Section 2.2.3 for more details, in particular on how to calculate coefficients c_F^M.

Also, the relation (2.1) for cliques K_2 and K_1 gives that respectively $K_2 = \sum_{H\in\mathcal{F}_5} p(K_2, H) \cdot H$ and $1 = K_1 = \sum_{H\in\mathcal{F}_5} H$. Thus (2.4) can be written as an identity involving densities of 5-vertex graphs.

Next, we take the sum of equations (2.3), (2.4) multiplied by some α, and $\phi([\![X^T MX]\!]_1) \geqslant 0$ expanded using (2.6), and obtain

$$\phi(\,\mathrm{OPT}\,) \geqslant \phi(\,\mathrm{OPT}\,) + \alpha\Big(p - \phi\Big($$  $$\Big)\Big) - \phi([\![X^T MX]\!]_1)$$

$$= \phi\Big(\mathrm{OPT} + \alpha p - \alpha$$  $$- [\![X^T MX]\!]_1\Big)$$

$$= \phi\Big(\sum_{F\in\mathcal{F}_5} (c_F^{\mathrm{OPT}} + \alpha p - \alpha \cdot p(K_2, F) - c_F^M) \cdot F\Big)$$

(In Appendix A we provide c_F^{OPT} and $p(K_2, F)$ for each $F \in \mathcal{F}_5$.) For $F \in \mathcal{F}_5$, define

$$c_F = c_F^{\mathrm{OPT}} + \alpha p - \alpha \cdot p(K_2, F) - c_F^M. \tag{2.7}$$

With this notation

$$\phi(\,\mathrm{OPT}\,) \geqslant \phi\Big(\sum_{F\in\mathcal{F}_5} c_F \cdot F\Big) \geqslant \min_{F\in\mathcal{F}_5} c_F \cdot \phi\Big(\sum_{F\in\mathcal{F}_5} F\Big) = \min_{F\in\mathcal{F}_5} c_F, \tag{2.8}$$

where c_F is a number that depends on the choice of M and α. Let us transfer this back to our extremal graph problem.

Lemma 2.1. *For every $p \in [0, 1]$, $M \succcurlyeq 0$ and $\alpha \in \mathbb{R}$, with $c_F = c_F(p, M, \alpha)$ as in (2.7), we have*

$$d_{C_5}(p) \geqslant \frac{1}{120} \min_{F\in\mathcal{F}_5} c_F.$$

Proof. Suppose on the contrary we can find an increasing sequence of graphs G_n with edge density $p + o(1)$ such that $d_5(G_n)$ stays strictly below the stated bound. Take a convergent subsequence and let $\phi \in \mathrm{Hom}^+(\mathcal{A}, \mathbb{R})$ be its limit. It satisfies (2.4) so the bound in (2.8) applies to ϕ. However, this contradicts (2.2). □

2.2.2 Finding the optimum

Let an integer $k \geqslant 3$ be fixed. Let p and λ be as in (1.1). By Lemma 2.1, in order to finish the proof of Theorem 1.1, it is enough to present some $M \succcurlyeq 0$ and $\alpha \in \mathbb{R}$ with $c_F \geqslant 5!\,\lambda$ for every $F \in \mathcal{F}_5$. Let

$$\alpha = \frac{1}{k^3}(60k^3 - 240k^2 + 360k - 192).$$

In order to define the matrix M we define first two matrices A and B as follows:

$$A = \begin{pmatrix} 32k^2 - 96k + 96 & 0 & 4k^2 - 16k \\ 0 & 10k^4 - 30k^3 - 8k^2 + 96k - 96 & -10k^4 + 35k^3 - 4k^2 - 80k + 96 \\ 4k^2 - 16k & -10k^4 + 35k^3 - 4k^2 - 80k + 96 & 10k^4 - 40k^3 + 24k^2 + 64k - 96 \end{pmatrix}$$

and

$$B = \begin{pmatrix} k-1 & 1 & k-2 & 0 & k-3 & -1 \\ 0 & 2 & k-2 & 0 & 2k-4 & -2 \\ 0 & 0 & k-1 & -1 & 2k-2 & -2 \end{pmatrix}.$$

It is easy to verify (by checking principal minors) that A is positive definite for any $k \geqslant 3$. Therefore, the matrix

$$M = \frac{3}{2k^4} B^T A B \tag{2.9}$$

is positive semidefinite. In Section 2.2.4 we briefly describe how we determined matrices A and B. With this choice of M and α, one can verify using for example Maple (see Appendix B) that coefficients c_F satisfy:

$$\begin{aligned} c = c = c = c = c = c = c = c = c = c = c = c = \\ c = c = c = c = c = c &= \frac{1}{5k^4}(60k^4 - 300k^3 + 600k^2 - 600k + 240) \\ c = c = c = c &= \frac{1}{5k^4}(66k^4 - 300k^3 + 600k^2 - 600k + 240) \\ c &= \frac{1}{5k^4}(68k^4 - 300k^3 + 600k^2 - 600k + 240) \\ c = c = c = c &= \frac{1}{5k^4}(64k^4 - 300k^3 + 600k^2 - 600k + 240) \\ c &= \frac{1}{5k^4}(65k^4 - 300k^3 + 600k^2 - 600k + 240) \\ c = c = c = c &= \frac{1}{5k^4}(62k^4 - 300k^3 + 600k^2 - 600k + 240) \\ c = c &= \frac{1}{5k^4}(61k^4 - 300k^3 + 600k^2 - 600k + 240). \end{aligned}$$

Since the entries only ever disagree in the k^4 coefficient, it is easy to see that the smallest c_F are in the first two rows and are equal to $5!\,\lambda$, as desired. (Recall that this proves the lower bound on $d_{C_5}(p)$ of Theorem 1.1 by Lemma 2.1.)

2.2.3 Products of graphs and determining c_F^M coefficients

First, we define the product of unlabelled graphs. Recall that for a graph G we denote $|V(G)|$ by $|G|$. Let F_1, F_2, F in $\mathcal{F}$ be such that $|F_1| + |F_2| \leqslant |F|$. Choose uniformly at random two disjoint subsets X_1 and X_2 of $V(F)$ of sizes $|F_1|$ and $|F_2|$, respectively. Denote by $p(F_1, F_2; F)$ the probability that $F[X_1]$ is isomorphic to F_1 and $F[X_2]$ is isomorphic to F_2. Finally, the product of F_1 and F_2 is defined as

$$F_1 \times F_2 = \sum_{F \in \mathcal{F}_{|F_1|+|F_2|}} p(F_1, F_2; F) \cdot F.$$

The product can be extended to linear combinations of graphs and gives a multiplication operation in $\mathcal{A}$.

The product in $\mathcal{A}^1$ is defined along the same lines as in $\mathcal{A}$ but the intersection of X_1 and X_2 is exactly the labelled vertex. A more precise definition follows. Let F_1, F_2, F in $\mathcal{F}^1$ such that $|F_1| + |F_2| \leqslant |F| - 1$. Choose uniformly at random subsets X_1 and X_2 of $V(F)$ of sizes $|F_1|$ and $|F_2|$, respectively whose intersection is exactly the one labelled vertex. Denote by $p(F_1, F_2; F)$ the probability that $F[X_1]$ is isomorphic to F_1 and $F[X_2]$ is isomorphic to F_2, where isomorphism preserves the labelled vertex. Finally, the product of F_1 and F_2 is defined as

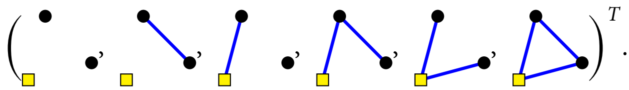
$$F_1 \times F_2 = \sum_{F \in \mathcal{F}_{|F_1|+|F_2|-1}} p(F_1, F_2; F) \cdot F.$$

Next we define the unlabelling operator $[\![\cdot]\!]_1 : \mathcal{F}^1 \to \mathbb{R}\mathcal{F}$. We extend $[\![\cdot]\!]_1$ to a linear function $\mathbb{R}\mathcal{F}^1 \to \mathbb{R}\mathcal{F}$ which we also call $[\![\cdot]\!]_1$. Let $F \in \mathcal{F}^1$. Denote by $G \in \mathcal{F}$ the graph obtained from F by unlabelling the labelled vertex. Let v be a vertex in G chosen uniformly at random. Let q be the probability that G with labelled v is isomorphic to F. Then

$$[\![F]\!]_1 = q \cdot G.$$

Recall that X is the vector of all 3-vertex labelled graphs from $\mathcal{F}^1$:

$$X = (X_1, X_2, X_3, X_4, X_5, X_6)^T =$$



In Appendix A we list all coefficients for products in $\mathcal{F}^1_3$, after unlabelling and multiplying by a scaling factor of 30 to clear denominators. Then we obtain that

$$[\![X^T M X]\!]_1 = \sum_{i=1}^{6} \sum_{j=1}^{6} M_{i,j} [\![X_i \times X_j]\!]_1 = \sum_{F \in \mathcal{F}_5} c_F^M \cdot F,$$

since each $[\![X_i \times X_j]\!]_1$ is a linear combination of graphs in $\mathcal{F}_5$.

2.2.4 Guessing matrices A and B

In this paragraph we describe how we obtained the matrices A and B. First, we used semidefinite programming to find a matrix M for several small odd values of k. Notice that if (2.8) is applied to the extremal construction, then the left-hand side is equal to the right-hand side. That means that all inequalities used are actually equalities. In particular, $\phi([\![X^T M X]\!]_1) = 0$. Since M is a positive semidefinite matrix, X evaluated on our extremal example (the limit of T_k^n as $n \to \infty$) must give an eigenvector of M corresponding to the eigenvalue 0. The matrix B was obtained by projecting onto the space orthogonal to three zero eigenvectors of M. As noted before, we had one zero eigenvector to start with. By looking at all eigenvectors of M, we managed to guess another zero eigenvector. We tried projection with the two zero eigenvectors and found the third one in the projection. After having obtained matrices B, we observed that a suitable A exists even if we set the coordinate $[1, 2]$ and $[2, 1]$ to 0. With proper scaling of the objective function, we were getting nice matrices from the CSDP [3] solver with all entries integers. By using the solutions for several values of k, we calculated a polynomial function of k fitting each entry in matrix A. Finally we observed that the same matrices A and B also work for even values of k.

3. Stability

In this section we prove Theorem 1.2. For this purpose it will be convenient to rewrite our lower bound as an asymptotic identity valid for an arbitrary graph. Fix $k \geqslant 3$. Let p and λ be as in (1.1). Let the matrix $M \succcurlyeq 0$, $\alpha \in \mathbb{R}$, and the reals $c_F^M, c_F^{\mathrm{OPT}}, c_F$, indexed by $F \in \mathcal{F}_5$, be as previously.

Recall that $X = (X_1, \ldots, X_6)^T$ is the vector of 3-vertex rooted graphs defined in (2.5). For a graph $G = (V, E)$ of order $n \geqslant 5$ and a vertex $r \in V$, let Y_r be the column vector whose ith component is the number of unordered 2-sets $\{u, v\} \subseteq V \setminus \{r\}$ such that the induced graph $G[\{r, u, v\}]$ rooted at r is isomorphic to X_i. Define

$$\overline{Y} = \frac{4}{5!} \sum_{r \in V} Y_r^T M Y_r \geqslant 0.$$

Let us argue that

$$\overline{Y} = \sum_{F \in \mathcal{F}_5} c_F^M P(F, G) + O(n^4), \tag{3.1}$$

where for $F \in \mathcal{F}_\ell$ we let $P(F, G) = \binom{n}{\ell} p(F, G)$ be the number of ℓ-sets inducing a copy of F in G. Indeed, the ith entry of Y_r can be written as a double sum $\frac{1}{2} \sum_{u \in V} \sum_{v \in V}$ of the indicator function that r, u, v are distinct and the graph $G[\{r, u, v\}]$ when rooted at r is isomorphic to X_i. Using this representation of Y_r and expanding everything, we can write $\overline{Y}$ as a sum over all $(r, u, v, u', v') \in V^5$ of some function that depends only on the graph induced by the (multi)set (r, u, v, u', v') inside G. Apart of $O(n^4)$ terms when some of the vertices coincide, the remaining ones can be grouped by the isomorphism type $F \in \mathcal{F}_5$ of $G[\{r, u, v, u', v'\}]$. For $F \in \mathcal{F}_5$, each unordered 5-set spanning an induced copy of F in G contributes the same amount (depending only on F and M) and the coefficient c_F^M was in fact defined by us to be equal to this common value. Thus (3.1) holds.

Likewise, $P(K_2, G)\binom{n-2}{3}$ and $\binom{n}{5}$ can be written as fixed linear combinations of $P(F, G)$ over $F \in \mathcal{F}_5$. Also, $d_{C_5}(G) n^5 = \sum_{F \in \mathcal{F}_5} c_F^{\mathrm{OPT}} P(F, G)$ is the number of 5-cycles in G. Putting it all together, we obtain the following identity valid for an arbitrary graph G:

$$d_{C_5}(G)\, n^5 + \frac{\alpha}{5!}(2P(K_2, G) n^3 - p n^5) - \overline{Y} + O(n^4) = \sum_{F \in \mathcal{F}_5} c_F P(F, G), \tag{3.2}$$

where c_F for $F \in \mathcal{F}_5$ was defined to be exactly the contribution of each induced copy of F in G to the left-hand side while all combinations when some vertices in the underlying 5-fold sum coincide are absorbed into the error term $O(n^4)$.

Note that if we multiply (3.2) by $\binom{n}{5}^{-1}$ then the scaled terms in (3.2) will be asymptotically the same as in (2.8) when $n \to \infty$. Since $\sum_{F \in \mathcal{F}_5} P(F, G) = \binom{n}{5}$, the right-hand size of (3.2) can be lower-bounded by $\binom{n}{5} \min_{F \in \mathcal{F}_5} c_F$, giving the required lower bound in Theorem 1.1 since each c_F is at least $5!\, \lambda$.

Let us turn to stability. Take any sequence of graphs G_m of strictly increasing orders such that

$$|E(G_m)| \geqslant \left(p - \frac{1}{m}\right)\binom{|G_m|}{2} \quad \text{and} \quad d_{C_5}(G_m) \leqslant \lambda + \frac{1}{m}, \quad \text{for all } m \in \mathbb{N}. \tag{3.3}$$

Observe that if $c_F > \lambda$ for some $F \in \mathcal{F}_5$, then the right-hand side of (3.2) is at least $(\lambda + (c_F - \lambda) p(F, G))\binom{n}{5}$. Thus we have that $p(F, G_m) = o(1)$ as $m \to \infty$ for every such F. By looking at the explicit formulas for c_F near the end of Section 2.2.2, we see that there are 16 such graphs. They are collected into the list $\mathcal{L}$ in Figure 1, and are denoted by $L_1, \ldots, L_{16}$ in this order.

Let the *co-cherry* $\overline{P_2}$ be the complement of the 2-edge path P_2, that is, $\overline{P_2}$ is the graph with 3 vertices and 1 edge. Next, we show that its density in G_m must also be $o(1)$. Note that there are 5-vertex graphs not in the list $\mathcal{L}$ that contain the co-cherry. Thus the naive approach does not work and a slightly more involved argument is needed.

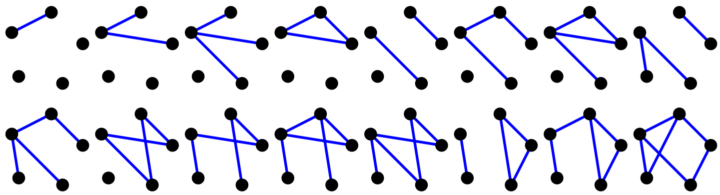


Figure 1. The list $\mathcal{L} = (L_1, \ldots, L_{16})$.

Lemma 3.1. *For every sequence of graphs G_m as in (3.3), we have that*

$$\lim_{m\to\infty} p(\overline{P_2}, G_m) = 0.$$

Proof. Let m be sufficiently large, $G = G_m$, $V = V(G)$ and $n = |V|$. For $i \in \{0, 1, 2\}$, let F_i be the (unique) graph of order 4 with i disjoint edges. Let $\mathcal{L}' = \mathcal{L} \cup \{F_0, F_1, F_2\}$.

Apply the induced removal lemma (see *e.g.* [1, 4]) to G to destroy all induced graphs in $\mathcal{L}'$ whose density is $o(1)$. Formally, let $f = n^{-1} + \max\{p(L, G_m) : L \in \mathcal{L}\}$ (and let initially $G = G_m$). As long as there is at least one $F \in \mathcal{L}'$ with $0 < p(F, G) \leqslant f$, change as few as possible adjacencies in G to destroy all copies of all such F so that, additionally, no graph in $\mathcal{L}'$ absent from G is introduced. Since f tends to 0 as $m \to \infty$ and the above iteration is applied at most $|\mathcal{L}'|$ times (in fact, at most $|\mathcal{L}' \setminus \mathcal{L}| + 1 = 4$ times), we change $o(n^2)$ edges in total by the induced removal lemma. Also, the final graph G contains no graph from the list $\mathcal{L}$ since the first iteration destroyed all such subgraphs by our choice of f.

Claim 3.2. *G contains no induced F_1 (i.e. 4 vertices spanning exactly one edge).*

Proof. Take a copy of F_1 and add one new vertex x of degree d. If $d \in \{0, 1, 2, 3\}$, then the sets of possible obtained graphs up to isomorphism are respectively $\{L_1\}$, $\{L_2, L_5\}$, $\{L_4, L_6, L_8\}$, and $\{L_7, L_9\}$. We see that each 1-vertex extension of F_1 is in $\mathcal{L}$ except when $d = 4$ (*i.e.* when x is adjacent to every vertex of F_1). This means that for every copy of F_1, say on $A \subseteq V$, the set A is complete to $V \setminus A$ in G. It follows that every two distinct induced copies of F_1 are vertex-disjoint and thus G has at most $n/4$ such copies. This is at most $f\binom{n}{4}$, so G has no copy of F_1 at all. □

Claim 3.3. *G contains no induced F_2 (which is the matching with two edges).*

Proof. If we extend a copy of F_2 by adding a vertex x of degree $d \in \{0, 1, 2, 3\}$, then we obtain graphs in respectively $\{L_5\}$, $\{L_8\}$, $\{L_{11}, L_{14}\}$ and $\{L_{15}\}$. Thus the only extension that does not lead to a graph in $\mathcal{L}$ is to connect x to every vertex of F_2. This gives that every two distinct induced copies of F_2 in G are vertex-disjoint. Thus we have at most $O(n) \leqslant f\binom{n}{4}$ copies of F_2, that is, none at all. □

Consider the edgeless 4-vertex graph F_0. If we add a vertex x of degree $d \in \{1, 2, 3\}$, then we get respectively L_1, L_2 and L_3. The only remaining ways are to have x empty or complete to F_0. Now, consider any copy of F_0 in G, say with vertex set $A_0 \subseteq V(G)$. By above, every vertex outside of A_0 is empty or complete to A_0. Let $A \supseteq A_0$ consist of all vertices of G that send no edges to A_0. Note that A is an independent set: if we had an edge xy inside A then x, y plus some two extra vertices from A_0 would span a copy of F_1 in G, contradicting Claim 3.2. Moreover, A is complete to $V \setminus A$. Indeed, for every pair $(a, b) \in A \times (V \setminus A)$, the subgraph of G induced by a and some further three vertices of A_0 has no edges; thus the vertex $b \notin A$ must be complete to it.

It follows that we can find disjoint independent sets A_i, $i \in I$, in V such that each A_i is complete to $V \setminus A_i$ while every copy of F_0 in G is inside one of these sets A_i. Define $B = V \setminus (\cup_{i\in I} A_i)$.

By the definition of B and the above claims, we have that $H=G[B]$ is $\{F_0, F_1, F_2\}$-free. This means that the complement $\overline{H}$ of H cannot have a (not necessarily induced) 4-cycle C_4 because for any way of filling its diagonals we get F_0, F_1 or F_2 in H. Thus $|E(\overline{H})|$ is at most the Turán function $\mathrm{ex}\,(n, C_4)=O(n^{3/2})$, that is, H is $O(n^{3/2})$-close in the edit distance to being a complete graph. We see that G is $O(n^{3/2})$-close to the complete partite graph G' with parts A_i, $i\in I$, and $\{x\}$, $x\in B$. As every co-cherry in G has to contain at least one pair where $E(G)$ and $E(G')$ differ, G has at most $O(n^{5/2})$ co-cherries.

Since the original graph G_m and G differ in $o(n^2)$ adjacencies, the co-cherry density in G_m is $o(1)$, as required. □

Thus, another application of the induced removal lemma gives that we can change $o(1)$-fraction of adjacencies in G_m and make it $\overline{P_2}$-free, that is, complete partite. Thus, in order to finish the proof of Theorem 1.2, it is enough to argue that each of the k largest parts of G_m has $(1/k+o(1))|G_m|$ vertices. We present two proofs of this. The first proof is more direct but longer. The second one is shorter but assumes some known facts about graphons.

3.1 First proof

We need the following auxiliary result.

Lemma 3.4. *Suppose a graph J on n vertices has a subgraph X such that*

(i) *X has x vertices where $\varepsilon' n\leqslant x\leqslant (1-\varepsilon')n$ and edge density $q\leqslant 1/2$,*
(ii) *X is complete to $V(J)\setminus X$,*
(iii) *X contains at least $\frac{1}{2}x^4q^3+\varepsilon' x^4$ copies of P_4.*

Then there exists a graph J' on n vertices with asymptotically the same edge density as J and

$$d_{C_5}(J')\leqslant d_{C_5}(J)-\frac{1}{2}(\varepsilon')^6.$$

Proof. Note first that conditions (i) and (ii) imply that J is dense since it has at least $\varepsilon'(1-\varepsilon')n^2$ edges. We make J' by replacing X with a X', which is a random balanced bipartite graph with edge probability $2q$. We will not change the rest of the graph, so $J'-X'=J-X$. With high probability X' has edge density asymptotically q and so J' has asymptotically the same edge density as J. We will argue that J' has much fewer copies of C_5 than J has, by considering several possible types of C_5 copies.

We will compare the copies according to how they intersect X (for counting copies of C_5 in the graph J) or X' (in J'). Specifically, since X is complete to the rest of J we have

$$\nu_{C_5}(J)=\sum_H m_H\nu_H(X)\cdot\nu_{C_5-H}(J-X),$$

where the sum is over all induced subgraphs $H\subseteq C_5$, and the coefficient m_H is the number of C_5 copies contained in the graph formed by taking a copy of H and a copy of C_5-H with every possible edge in between. Recall that $\nu_H(G)$ counts the number of (not necessarily induced) copies of H in G. Similarly, we have

$$\nu_{C_5}(J')=\sum_H m_H\nu_H(X')\cdot\nu_{C_5-H}(J'-X')=\sum_H m_H\nu_H(X')\cdot\nu_{C_5-H}(J-X),$$

since $J'-X'=J-X$. So we will compare $\nu_H(X)$ with $\nu_H(X')$ for each H. Specifically we will show that $\nu_H(X')\leqslant(1+o(1))\nu_H(X)$ for each H, and that this inequality holds with some room for $H=P_4$.

Some easy cases: when H has no vertices, $\nu_H(X)=\nu_H(X')=1$. When H is a single vertex, $\nu_H(X)=\nu_H(X')=x$. When H is just an edge, $\nu_H(X)=(1+o(1))\nu_H(X')=(1+o(1))\binom{x}{2}q$. When H has 2 vertices and no edge we have $\nu_H(X')=\nu_H(X)=\binom{x}{2}$. When H is the graph on 3 vertices consisting of an edge and an isolated vertex, we have $\nu_H(X')=(1+o(1))\nu_H(X)=(1+o(1))x\binom{x}{2}q$.

When $H=P_3$ (the path of length 2) we have

$$\nu_{P_3}(X')=2\binom{x/2}{2}\frac{x}{2}(2q)^2=(1+o(1))\frac{1}{2}x^3q^2,$$

which we compare to

$$\nu_{P_3}(X)=\sum_{v\in X}\binom{|N(v)\cap X|}{2}\geqslant x\cdot\binom{(2q\binom{x}{2})/x}{2}=(1+o(1))\frac{1}{2}x^3q^2.$$

Finally we consider the case $H=P_4$. We have

$$\nu_{P_4}(X')=2\binom{x/2}{2}\cdot 2\binom{x/2}{2}(2q)^3=(1+o(1))\frac{1}{2}x^4q^3$$

which we compare to

$$\nu_{P_4}(X)=\frac{1}{2}x^4q^3+\varepsilon' x^4.$$

Taking all possible H into account, we see that

$$\begin{aligned}\nu_{C_5}(J)-\nu_{C_5}(J')&=\sum_H[\nu_H(X)-\nu_H(X')]\cdot\nu_{C_5-H}(J-X)\\&\geqslant[\nu_{P_4}(X)-\nu_{P_4}(X')]\cdot\nu_{C_5-P_4}(J-X)\\&\geqslant(1+o(1))\varepsilon' x^4\cdot(n-x)\\&>\frac{1}{2}(\varepsilon')^6n^5\end{aligned}$$

and so

$$d_{C_5}(J')\leqslant d_{C_5}(J)-\frac{1}{2}(\varepsilon')^6.$$ □

Proof of Theorem 1.2. Let G_m be as in (3.3). Let $m\to\infty$. By the induced graph removal lemma and Lemma 3.1 we can eliminate all co-cherries in the graph $G=G_m$ of order $n\to\infty$ by adding or removing at most αn^2 edges, for some $\alpha=\alpha(\varepsilon)\to 0$ as $\varepsilon\to 0$. Call this new graph G', which has edge density p', where $p-2\alpha\leqslant p'\leqslant p+2\alpha$. Moreover, G' is a complete k'-partite graph for some k'. Say the parts of G' are $X_1,\ldots,X_{k'}$. Also, note that since adding (or removing) one edge to G creates (or destroys) at most n^3 copies of C_5, we have

$$d_{C_5}(G)=d_{C_5}(G')+O(\alpha),$$

and

$$d_{C_5}(p)=d_{C_5}(p')+O(\alpha)$$

(recall that we use big-O notation to replace quantities that are bounded in absolute value, and the quantity being replaced may be negative). Now

$$d_{C_5}(G')\leqslant d_{C_5}(G)+O(\alpha)\leqslant d_{C_5}(p)+\varepsilon+O(\alpha)\leqslant d_{C_5}(p')+O(\varepsilon+\alpha)\tag{3.4}$$

and so G' has nearly the minimum C_5-density among graphs with edge density p'.

In the following, we will need a parameter $\beta=\beta(\varepsilon)=(\varepsilon+\alpha(\varepsilon))^{1/100}$.

Claim 3.5. *We are done unless we have the following. For any $i\neq j$, $|X_i|+|X_j|\leqslant(1-\beta)n$.*

Proof. Without loss of generality, suppose for contradiction that $|X_1| + |X_2| \geqslant (1-\beta)n$, so the number of edges in G' is at most

$$\begin{aligned}\binom{n}{2} - \binom{|X_1|}{2} - \binom{|X_2|}{2} &\leqslant \binom{n}{2} - 2\binom{((1-\beta)n)/2}{2} \\ &\leqslant \frac{1}{2}n^2 - \frac{1}{4}(1-\beta)^2 n^2 \\ &= \left(\frac{1}{4} + O(\beta)\right)n^2\end{aligned}$$

and so we must have $k=2$ since throughout the proof we assume ε (and therefore α and β) are sufficiently small. Now if $||X_1| - |X_2|| > \beta^{1/3}n$, say without loss of generality $|X_1| > |X_2| + \beta^{1/3}n$ then the number of edges in G' is at most

$$\begin{aligned}|X_1||X_2| + \beta n(|X_1| + |X_2|) + \binom{\beta n}{2} &\leqslant \left(\frac{n}{2} + \frac{1}{2}\beta^{1/3}n\right)\left(\frac{n}{2} - \frac{1}{2}\beta^{1/3}n\right) + \beta n^2 + \binom{\beta n}{2} \\ &= \left(\frac{1}{4} - \frac{1}{4}\beta^{2/3} + O(\beta)\right)n^2,\end{aligned}$$

which is a contradiction for small ε since G' has at least $\binom{n}{2}p - \alpha n^2$ edges (where $p = 1/2$ since $k=2$) and $\frac{1}{4}\beta^{2/3} + O(\beta) > \alpha$ for small ε. To summarize, G' is a complete partite graph that has two large parts X_1, X_2 which differ in size by at most $\beta^{1/3}n$, and together the rest of the parts make up at most βn vertices. It is easy to see then that G' can be changed into a balanced complete bipartite graph by editing $O(\beta^{1/3}n^2)$ edges. □

Thus, we henceforth assume that for any $i \neq j$, $|X_i| + |X_j| \leqslant (1-\beta)n$.

Claim 3.6. *For all i, j, if $|X_i|, |X_j| \geqslant \beta n$, then $||X_i| - |X_j|| \leqslant \beta n$.*

Proof. Suppose for contradiction that there are two parts (without loss of generality say X_1, X_2) such that $|X_1|, |X_2| \geqslant \beta n$ and $||X_1| - |X_2|| > \beta n$. We will derive a contradiction by arguing that G' can be modified by Lemma 3.4 to form another graph G^* of asymptotically the same edge density but with significantly smaller C_5-density than G'.

We apply Lemma 3.4 with $J = G'$, $X = X_1 \cup X_2$, $\varepsilon' = \frac{1}{2}\beta^6$ and

$$q = \frac{x_1 x_2}{\binom{x}{2}} = (1+o(1))\frac{2x_1x_2}{x^2},$$

where $|X_i| = x_i$ and $x = x_1 + x_2$. Let us check the conditions of the lemma. Clearly we have

$$\beta n \leqslant x \leqslant (1-\beta)n,$$

and X is complete to the rest of the graph (since X is composed of two parts of a complete partite graph). Finally, the number of copies of P_4 in X is

$$\nu_{P_4}(X) = 2\binom{x_1}{2}\cdot 2\binom{x_2}{2} = (1+o(1))x_1^2x_2^2,$$

which we compare to

$$\frac{1}{2}x^4q^3 = (1+o(1))\frac{1}{2}x^4\left(\frac{2x_1x_2}{x^2}\right)^3 = (1+o(1))\frac{4x_1^3x_2^3}{x^2}.$$

From here we can see that

$$\begin{aligned}\nu_{P_4}(X) - \frac{1}{2}x^4q^3 &\geqslant (1+o(1))\left(x_1^2x_2^2 - \frac{4x_1^3x_2^3}{x^2}\right)\\ &\geqslant \frac{1}{2}\cdot\frac{x_1^2x_2^2}{x^2}(x^2-4x_1x_2)\\ &= \frac{1}{2}\cdot\frac{x_1^2x_2^2}{x^2}(x_1-x_2)^2\\ &\geqslant \frac{1}{2}\frac{(\beta n)^4}{n^2}(\beta n)^2\\ &= \frac{1}{2}\beta^6n^4\\ &\geqslant \frac{1}{2}\beta^6x^4\end{aligned}$$

and so Lemma 3.4 applies, implying that $J = G'$ must have C_5-density at least

$$d_{C_5}(p') + \frac{1}{2}\left(\frac{1}{2}\beta^6\right)^6 = d_{C_5}(p') + \frac{1}{128}\beta^{36}.$$

But then from (3.4), we have

$$d_{C_5}(p') + \frac{1}{128}\beta^{36} \leqslant d_{C_5}(G') \leqslant d_{C_5}(p') + O(\varepsilon+\alpha),$$

a contradiction for small ε since $\beta = (\varepsilon+\alpha)^{1/100}$. □

Without loss of generality say that $|X_1|, \dots, |X_\ell| \geqslant \beta n$ and $|X_i| < \beta n$ for any $i > \ell$. By Claim 3.6, there is some value x such that $|X_i| \in [(x-\beta)n, (x+\beta)n]$ for $1 \leqslant i \leqslant \ell$. Then the number of edges in G' is at most

$$\binom{n}{2} - \sum\binom{|X_i|}{2} \leqslant \binom{n}{2} - \ell\binom{(x-\beta)n}{2} = \frac{1}{2}n^2(1-\ell x^2 + O(\beta)).$$

We will now show a lower bound matching the above upper bound. Since for any numbers $a \geqslant b$ and $\delta > 0$, we have $(a+\delta)^2 + (b-\delta)^2 > a^2 + b^2$, the following holds. Since $\sum_{i>\ell}|X_i| \leqslant n$, and for $i > \ell$ we have $|X_i| \leqslant \beta n$, the maximum possible value of $\sum_{i>\ell}|X_i|^2$ occurs when all the terms are either 0 or $(\beta n)^2$, meaning that the number of positive terms would be at most $1/\beta$, so we have

$$\sum_{i>\ell}|X_i|^2 \leqslant \frac{1}{\beta}\cdot(\beta n)^2 = \beta n^2,$$

the number of edges in G' is then at least

$$\binom{n}{2} - \sum\binom{|X_i|}{2} \geqslant \binom{n}{2} - \ell\binom{(x+\beta)n}{2} - \frac{1}{2}\beta n^2 = \frac{1}{2}n^2(1-\ell x^2+O(\beta)).$$

But we know G' has edge density

$$p' = 1 - \tfrac{1}{k} + O(\alpha) = 1 - \ell x^2 + O(\beta),$$

so we get

$$x = \frac{1}{\sqrt{k\ell}} + O(\beta)$$

and in particular $\ell \leqslant k$ since otherwise

$$|X_1| + \cdots + |X_\ell| \geqslant (\ell x + O(\beta))n > n.$$

To summarize, at this point we know that the graph must have $\ell \leqslant k$ 'large' parts which each have about $(1/\sqrt{k\ell})n$ vertices, and the rest of the parts are 'small' and each have at most βn vertices. We would like to show that $\ell = k$, so assume for contradiction that $\ell < k$.

Claim 3.7. $\sum_{i>\ell} |X_i| > \beta n$.

Proof. Observe that

$$\sum_{i>\ell} |X_i| = n - \sum_{i\leqslant\ell} |X_i| = n - \ell\left(\frac{1}{\sqrt{k\ell}} + O(\beta)\right)n = \left(1 - \frac{\sqrt{\ell}}{\sqrt{k}} + O(\beta)\right)n > \beta n$$

since $\ell < k$ and we may assume $\beta > 0$ is arbitrarily small. □

Now we will use Lemma 3.4 on $J = G'$ and X being X_1 together with several of the small X_i, which will finish the proof. Recall we have $|X_1|$ of size

$$\left(\frac{1}{\sqrt{kl}} + O(\beta)\right)n.$$

We know $|X_i| < \beta n$ for all $i > \ell$ and at the same time $|\cup_{i>\ell} X_i| > \beta n$. Hence there exists an integer z such that $\beta n \leqslant |\cup_{z\geqslant i>\ell} X_i| \leqslant 2\beta n$. Let $Y = \cup_{z\geqslant i>\ell} X_i$. In order to apply Lemma 3.4 to $X = X_1 \cup Y$, we need to count the number of copies of P_4 in X, the other assumptions of Lemma 3.4 are clearly satisfied. Notice that $\nu_{P_4}(X)$ is bounded from below by the number of copies of P_4 that alternate vertices in X_1 and in Y, which gives

$$\nu_{P_4}(X) \geqslant |X_1|^2|Y|^2 \geqslant |X_1|^2(\beta n)^2 = \frac{\beta^2}{kl}n^4 + O(\beta^3)n^4. \tag{3.5}$$

Denote $|X|$ by x. Notice that

$$x = |X_1| + |Y| = \left(\frac{1}{\sqrt{kl}} + O(\beta)\right)n.$$

Let e be the number of edges in X. It can be bounded from above by pretending that Y is a complete graph, which gives

$$e \leqslant |X_1| \cdot |Y| + |Y|^2/2 \leqslant \frac{2\beta n^2}{\sqrt{kl}} + O(\beta^2)n^2.$$

This gives

$$q = \frac{2e}{x^2} \leqslant 4\beta\sqrt{kl} + O(\beta^2).$$

Hence X satisfies Lemma 3.4(iii) with $\varepsilon' = \beta^2 kl/2$, since

$$\frac{1}{2}x^4q^3 \leqslant \frac{32\beta^3}{\sqrt{kl}}n^4 + O(\beta^4)n^4$$

is significantly smaller than $\nu_{P_4}(X)$ (see (3.5)) and

$$\varepsilon' x^4 \leqslant \frac{\beta^2}{2kl}n^4 + O(\beta^4)n^4$$

is about $\frac{1}{2}\nu_{P_4}(X)$. Hence Lemma 3.4 implies

$$d_{C_5}(G') \geqslant d_{C_5}(p') + \frac{\beta^{12}(kl)^6}{2^7} > d_{C_5}(p') + \beta^{19}.$$

Combining this with (3.4) gives the final contradiction

$$d_{C_5}(p') + \beta^{19} \leqslant d_{C_5}(G') \leqslant d_{C_5}(p') + O(\varepsilon + \alpha)$$

for a small ε since $\beta = (\varepsilon + \alpha)^{1/100}$. □

Summarizing, we just showed that G can be transformed into the Turán graph T_n^k by adding or deleting at most $o(n^2)$ edges.

3.2 Second proof

Here we use some notions related to graphons. An introduction to graphons and further details can be found in the excellent book by Lovász [13]. In general, a *graphon* is a quadruple $Q = (\Omega, \mathcal{B}, \mu, W)$, where $(\Omega, \mathcal{B}, \mu)$ is a standard probability space and $W : \Omega \times \Omega \to [0, 1]$ is a symmetric measurable function; see [13, Section 13.1]. For a graph F on $[k]$, its *induced homomorphism density* in Q is

$$t_{\mathrm{ind}}(F, Q) = \int_{\Omega^k} \prod_{ij \in E(F)} W(x_i, x_j) \prod_{ij \in E(\overline{F})} (1 - W(x_i, x_j))\, \mathrm{d}\mu(x_1) \ldots \mathrm{d}\mu(x_k).$$

Here we identify two graphons Q and Q' if $t_{\mathrm{ind}}(F, Q) = t_{\mathrm{ind}}(F, Q')$ for every graph F, calling them *equivalent*.

The relevance of graphons comes from the result of Lovász and Szegedy [15] that positive homomorphisms ϕ: $\mathrm{Hom}^+(\mathcal{A}, \mathbb{R}) \to \mathbb{R}$ are in one-to-one correspondence with graphons Q (up to equivalence) so that, for every graph F, we have $\phi(F) = p(F, Q)$, where we let

$$p(F, Q) = \frac{|F|!}{|\operatorname{aut}(F)|}\, t_{\mathrm{ind}}(F, Q)$$

with $\operatorname{aut}(F)$ being the group of automorphisms of F. Also, let

$$d_{C_5}(Q) = \frac{1}{5!} \sum_{F \in \mathcal{F}_5} c_F^{\mathrm{OPT}} p(F, Q).$$

For any graph $G = (V, E)$ there is a corresponding graphon $Q_G = (V, 2^V, \mu, A)$, where μ is the uniform measure and $A : V \times V \to \{0, 1\}$ is the adjacency function of G. Then, for example, $t_{\mathrm{ind}}(F, Q_G)$ is the probability that a uniform random map $f : V(F) \to V(G)$ is an *induced homomorphism*, that is, for all $i, j \in V(F)$, $\{i, j\} \in E(F)$ if and only if $\{f(i), f(j)\} \in E(G)$. We say that a sequence of graphons Q_n *converges* to Q if, for every graph F, we have $\lim_{n\to\infty} t_{\mathrm{ind}}(F, Q_n) = t_{\mathrm{ind}}(F, Q)$. In particular, if $Q_n = Q_{G_n}$ for some increasing sequence of graphs G_n, then this gives the same convergence of graphs that we used.

Since, by Lemma 3.1, we will be seeing only the limits of (almost) complete partite graphs, the following more restrictive class $\mathcal{P}$ of 'complete partite' graphons will suffice for our purposes. Namely, from now on, we fix Ω to be the set $\{0, 1, 2, \ldots\}$ of non-negative integers with the discrete topology (thus all subsets of Ω or Ω^2 are measurable) and fix $W(i, j)$ to be 0 if $i = j \geqslant 1$ and be 1 otherwise (*i.e.* if $i \neq j$ or if $i = j = 0$). Only the measure μ will vary, and the measures that we consider are as follows. Let

$$\mathcal{R} = \left\{ \rho \in [0, 1]^{\mathbb{N}} : \rho_1 \geqslant \rho_2 \geqslant \cdots, \ \sum_{i=1}^{\infty} \rho_i \leqslant 1 \right\}.$$

For $\rho = (\rho_1, \rho_2, \ldots) \in \mathcal{R}$, define the probability measure μ_ρ on $(\Omega, 2^\Omega)$ by $\mu_\rho(\{i\}) = \rho_i$ for $i \geqslant 1$. Thus $\mu_\rho(\{0\}) = \rho_0$, where ρ_0 is always a shorthand for $1 - \sum_{i=1}^{\infty} \rho_i$ (but is not an entry of the vector $\rho = (\rho_1, \rho_2, \ldots)$). Also, define

$$P_\rho = (\Omega, 2^\Omega, \mu_\rho, W), \quad \text{for } \rho \in \mathcal{R}$$

and let $\mathcal{P} = \{P_\rho : \rho \in \mathcal{R}\}$ consist of all graphons that arise in this way.

For example, a complete partite graph G gives a graphon $P_G \in \mathcal{P}$ as follows. Order the parts $V_1, \ldots, V_s$ of G non-increasingly by their size, let $\rho_G = (|V_1|/|G|, \ldots, |V_s|/|G|, 0, 0, \ldots)$, and take $P_G = (\Omega, 2^\Omega, \mu_{\rho_G}, W)$. Since all vertices inside a part V_i are twins in G, we have that $t_{\mathrm{ind}}(F, Q_G) = t_{\mathrm{ind}}(F, P_G)$ for every $F \in \mathcal{F}$. Thus Q_G and P_G are equivalent graphons.

One should think of P_ρ as the limit of complete partite graphs where, for $i \geqslant 1$, ρ_i is the fraction of vertices in the ith largest part while ρ_0 is the total fraction of vertices in parts of relative size $o(1)$.

Lemma 3.8. *If a sequence of vectors $\rho_n \in \mathcal{R}$ converges to $\rho \in [0,1]^{\mathbb{N}}$ in the product topology (that is, pointwise), then $\rho \in \mathcal{R}$ and the corresponding graphons P_{ρ_n} converge to P_ρ.*

Proof. If $\sum_{i=1}^{\infty} \rho_i > 1$, then $\sum_{i=1}^{m} \rho_i > 1$ for some m and thus $\sum_{i=1}^{m} \rho_{n,i} > 1$ for sufficiently large n, a contradiction. Thus $\rho \in \mathcal{R}$.

We have to show that the graphons $P_n = (\Omega, 2^\Omega, \mu_n, W)$ converge to P_ρ, where $\mu_n = \mu_{\rho_n}$. Take any $F \in \mathcal{F}$ and $\varepsilon > 0$. Let $k = |F|$ and fix an integer $m > 3\binom{k}{2}/\varepsilon$.

For any $Q = (\Omega, 2^\Omega, \mu, W) \in \mathcal{P}$, define $Q' = (\Omega, 2^\Omega, \mu', W) \in \mathcal{P}$, where μ' is the push-forward of the measure μ under the map that sends each $i > m$ to 0 and is the identity otherwise. (In the $\mathcal{R}$-domain, this corresponds to truncating $x \in \mathcal{R}$ to $x' = (x_1, \ldots, x_m, 0, \ldots) \in \mathcal{R}$.) Let us show that

$$|t_{\mathrm{ind}}(F, Q) - t_{\mathrm{ind}}(F, Q')| \leqslant \frac{\varepsilon}{3}, \quad \text{for every } Q \in \mathcal{P}. \tag{3.6}$$

This inequality becomes more obvious if we allow general graphons and observe that the graphon Q' is equivalent to $(\Omega, 2^\Omega, \mu, W')$, where $W'(i,j)$ is defined to be 0 if $1 \leqslant i = j \leqslant m$ and 1 otherwise. Thus when we pass from W to W' on the same probability space $(\Omega, 2^\Omega, \mu)$, then for every $i \in \Omega$ the measure of j with $W(i,j) \neq W'(i,j)$ is always at most

$$\frac{1}{m+1} \leqslant \frac{\varepsilon}{3}\binom{k}{2}.$$

By Tonelli's theorem, this also upper-bounds the μ^2-measure of the set Z of pairs in Ω^2 where W and W' differ. Now, $t_{\mathrm{ind}}(F, \cdot)$ is an integral of a $[0,1]$-valued function over Ω^k and, by the union bound, the probability that some pair hits Z is at most

$$\binom{k}{2}\mu^2(Z) \leqslant \frac{\varepsilon}{3},$$

giving the desired.

Note that $\mu'_n(\{i\}) = \mu_n(\{i\})$ converges to $\mu'_\rho(\{i\}) = \mu_\rho(\{i\})$ for each $i \in [m]$. It follows that $\mu'_n(\{0\})$ converges to $\mu'_\rho(\{0\})$, since the support of probability measures μ'_ρ and any μ'_n is a subset of $\{0\} \cup [m]$. For such measures $t_{\mathrm{ind}}(F, \cdot)$ is a polynomial (and thus continuous) function of the measures of singletons $0, \ldots, m$. Thus, for all large n, we have that $|t_{\mathrm{ind}}(F, P'_n) - t_{\mathrm{ind}}(F, P'_\rho)| \leqslant \varepsilon/3$; then it holds by (3.6) that $|t_{\mathrm{ind}}(F, P_n) - t_{\mathrm{ind}}(F, P_\rho)| \leqslant \varepsilon$. Since $\varepsilon > 0$ and F were arbitrary, $P_n \to P_\rho$ as required. □

Remark. Using some standards facts about graphons, one can prove the converse implication of Lemma 3.8 (namely that the graphon convergence $P_{\rho_n} \to P_\rho$ implies that $\rho_n \to \rho$); see [11] where the space $\mathcal{P}$ is studied in more detail.

Note that the limit of the Turán graphs T_n^k as $n \to \infty$ is Q_{K_k} (or, equivalently, P_ρ for $\rho = (1/k, \ldots, 1/k, 0, \ldots) \in \mathcal{R}$).

Lemma 3.9. *For every $k \geqslant 3$, every sequence of graphs as in (3.3) converges to Q_{K_k}.*

Proof. Let us first show that $(G_n)_{n=1}^{\infty}$ has a subsequence convergent to Q_{K_k}. By Lemma 3.1 and the induced removal lemma, we can make each G_n complete partite, without changing the convergence of any subsequence. Recall that $\rho_{G_n} \in \mathcal{R}$ is the vector encoding the part ratios of G_n. Since the product space $[0, 1]^{\mathbb{N}}$ is compact, some subsequence of $\rho_{G_n} \in [0, 1]^{\mathbb{N}}$ converges to some ρ. By Lemma 3.8, we have that $\rho \in \mathcal{R}$ and the corresponding subsequence of graphs G_n converges to $Q = P_\rho$. Thus the graphon Q satisfies that $p(K_2, Q) = p$ and $d_{C_5}(Q) = \lambda$.

The identity in (3.2) can be rewritten as an identity valid for every graphon. Since we need to analyse it only for Q, let us state a version that uses the (very special) structure of graphons in $\mathcal{P}$. We need a few definitions first.

For a graph $F \in \mathcal{F}^1$ on $[k]$ rooted at 1 and $j \in \Omega$, define the *rooted density* of F in (Q, j) as

$$t_{\mathrm{ind}}(F, (Q, j)) = \sum_f \prod_{i=2}^{k} \rho_{f(i)},$$

where f in the sum ranges over all maps $V(F) \to \Omega$ such that $f(1) = j$ and, for all distinct $u, v \in V(F)$, we have that $W(f(u), f(v)) = 1$ if and only if $\{u, v\} \in E(F)$. Equivalently, this is the limit as $n \to \infty$ of the probability of the following event $\mathcal{E}$. Suppose we choose $k - 1$ independent uniform vertices in G_n together with another vertex we call the *root* which we make adjacent to everybody else if $j = 0$ or G_n has fewer than j parts, and otherwise we put the root in the jth largest part of G_n. Then we let $\mathcal{E}$ be the event that these chosen vertices together with the root induce a vertex-labelled homomorphic copy of F. For example, if $H \in \mathcal{F}$ is the unrooted copy of F, then

$$t_{\mathrm{ind}}(H, Q) = \sum_{j=0}^{\infty} t_{\mathrm{ind}}(F, (Q, j))\, \rho_j. \tag{3.7}$$

The version for unlabelled non-roots is

$$p(F, (Q, j)) = \frac{(k-1)!}{|\,\mathrm{aut}\,(F)|}\, t_{\mathrm{ind}}(F, (Q, j)),$$

where aut (F) is the group of root-preserving automorphisms of F. We also define a column vector

$$Y_j = (p(X_1, (Q, j)), \ldots, p(X_6, (Q, j)))^T \in \mathbb{R}^6,$$

where $X = (X_1, \ldots, X_6)^T$ was defined in (2.5). With this notation, the limit version of (3.2) is

$$5!d_{C_5}(Q) - \sum_{F \in \mathcal{F}_5} c_F p(F, Q) + \alpha(p(K_2, Q) - p) = \sum_{j=0}^{\infty} Y_j^T M Y_j\, \rho_j. \tag{3.8}$$

Recall that each c_F in (3.8) is at least $5!\lambda$ and that $p(K_2, Q) = p$ for our Q (which is the limit of some G_n); thus the left-hand size of (3.8) is non-positive. Also, recall that $M \succcurlyeq 0$; thus $x^T M x \geqslant 0$ for every $x \in \mathbb{R}^6$ with equality if and only if $Mx = 0$. As the 3×3 matrix A in the factorization (2.9) is non-singular, the null-space N of M is the same as that of B. Calculations (see *e.g.* the Maple code in Appendix B) show that the three-dimensional vector space N can be spanned by $z_1, z_2, z_3 \in \mathbb{R}^6$ where

$$\begin{pmatrix} z_1^T \\ z_2^T \\ z_3^T \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 2(k-1) & k-1 & k^2 - 3\,k + 2 \\ 0 & 1 & 0 & -2 & 0 & 1 \\ 0 & 0 & 1 & k-1 & \dfrac{k-2}{2} & \dfrac{k^2 - 3k + 2}{2} \end{pmatrix}. \tag{3.9}$$

By the previous paragraph, $Y_j \in \mathbb{R}^6$ belongs to N for every $j \in \Omega$ with $\rho_j > 0$. Since N is a (finite-dimensional and thus closed) linear subspace, it also contains the mean $\overline{Y} = \sum_{j=0}^{\infty} Y_j\, \rho_j$. By (3.7), we have that, in particular, $\overline{Y}_2 = t_{\mathrm{ind}}(\overline{P_2}, Q)$ and $\overline{Y}_3 = \frac{1}{2}\, t_{\mathrm{ind}}(\overline{P_2}, Q)$ are both 0. Since the entries in each Y_j are non-negative, we conclude that Y_j has its second and third coordinates zero for every $j \in \Omega$ with $\rho_j > 0$. The row-reduced matrix in (3.9) shows that such Y_j must be collinear to z_1. Since the sum of entries of Y_j is 1, we have $Y_j = (1/k^2)z_1$. In particular, its first coordinate is $1/k^2$. On the other hand, it is $p(X_1, (Q, j))$ which is the density of $\overline{K_3}$ rooted at j in Q, that is, it is ρ_j^2 if $j \geqslant 1$ and 0 if $j = 0$. Thus $\rho_0 = 0$ and $\rho_j = 1/k$ for every j in the support of μ, so indeed $Q = Q_{K_k}$ is the limit of Turán graphs.

Finally, if we assume on the contrary to the lemma that the whole sequence $(G_n)_{n=1}^{\infty}$ does not converge to Q_{K_k}, then by the compactness of the space of all graphons, some subsequence converges to a graphon non-equivalent to Q_{K_k}. But then this violates the first claim of the proof. □

Second proof of Theorem 1.2. Lemma 3.9 and the fact that each graphon is the limit of some sequence of finite graphs imply that the limit version of the C_5-minimization problem has the unique solution Q_{K_k} whose function W, moreover, happens to be $\{0, 1\}$-valued. These are exactly the assumptions of [21, Theorem 15] which directly gives the required stability property.

In order to give the reader some idea of what is going on, let us slightly unfold the proof of [21, Theorem 15] for this particular case. Suppose on the contrary that, for some integer $k \geqslant 3$ and $\delta > 0$, a sequence G_n of graphs as in (3.3) violates the stability property. By passing to a subsequence, it converges to some graphon Q with $p(K_2, Q) = p$ and $d_{C_5}(Q) = \lambda$. By Lemma 3.9, we can assume that $Q = Q_{K_k}$. While the convergence of G_n to some graphon does not identify G_n within edit distance $o(|G_n|^2)$ in general, it does if the function W of the graphon assumes only values 0 and 1; see [16, Lemma 2.9] or [21, Theorem 17]. Thus, the convergence $G_n \to Q_{K_k}$ implies that G_n is $o(|G_n|^2)$-close to $T_k^{|G_n|}$ in the edit distance, contradicting our assumption. □

Another possible derivation of Theorem 1.2 from Lemma 3.9 is to use the known properties of the so-called *cut-distance* via the argument in [22, page 146], where the description of all extremal graphons for the triangle-minimization problem was used to describe all almost extremal graphs.

4. Remarks on the case $p \neq 1 - 1/k$

Our general upper bound construction is as follows. Suppose that p is a constant satisfying $1 - 1/k < p < 1 - (1/k + 1)$. Partition the vertices into $k - 1$ sets $X_1, \ldots, X_{k-1}$ of size xn and one more set Y of size yn. Each X_i is an independent set. For $1 \leqslant i \neq j \leqslant k - 1$ we have that X_i is complete to X_j. Y is also complete to each X_i. Finally, $G[Y]$ is any graph such that for some parameter $0 < \rho < 1/2$ we have

(i) $G[Y]$ has asymptotically $\frac{1}{2}y^2n^2\rho$ edges, $\frac{1}{2}y^3n^3\rho^2$ paths of length 2 (that means on 3 vertices), and $\frac{1}{2}y^4n^4\rho^3$ paths of length 3;
(ii) $G[Y]$ has $o(n^5)$ copies of C_5.

(See the end of this subsection for discussion on which graphs are suitable for $G[Y]$.) We assume

$$(k-1)x + y = 1$$

so we have a total of n vertices. The edge density in this construction is

$$\frac{\binom{k-1}{2}(xn)^2 + (k-1)(xn)(yn) + (1/2 + o(1))y^2n^2\rho}{\binom{n}{2}},$$

which tends to

$$g(x,y,\rho) = (k-1)_2 x^2 + 2(k-1)xy + \rho y^2$$

as $n \to \infty$. So we also assume that the parameters x, y, ρ satisfy $g(x,y,\rho) = p$.

Now we consider the ratio

$$f(x,y,\rho) = \lim_{n\to\infty} \frac{\nu_G(C_5)}{n^5}.$$

We claim that

$$\begin{aligned} f(x,y,\rho) = &\left[\frac{1}{10}(k-1)_5 + \frac{1}{2}(k-1)_4 + \frac{1}{2}(k-1)_3\right]x^5 \\ &+ \left[\frac{1}{2}(k-1)_4 + \frac{3}{2}(k-1)_3 + \frac{1}{2}(k-1)_2\right]x^4 y \\ &+ \left[\left(\frac{1}{2} + \frac{1}{2}\rho\right)(k-1)_3 + \left(1 + \frac{1}{2}\rho\right)(k-1)_2\right]x^3 y^2 \\ &+ \left[\left(\frac{1}{2}\rho + \frac{1}{2}\rho^2\right)(k-1)_2 + \frac{1}{2}\rho(k-1)\right]x^2 y^3 \\ &+ \frac{1}{2}\rho^3 (k-1) x y^4. \end{aligned}$$

Note that we have grouped the terms of $f(x,y,\rho)$ according to powers of x and y, and then according to falling factorials of $(k-1)$. To understand our formula, it helps to think of the powers of x, y as specifying how many vertices come from sets of size xn, yn, and the falling factorial $(k-1)$ as specifying how many distinct sets of size xn are involved. For example, the first term $\frac{1}{10}(k-1)_5\, x^5$ is there because there are $\frac{1}{10}(k-1)_5 (xn)^5$ many copies of C_5 having vertices $v_1, \ldots, v_5$ all in different parts of size xn. Now let us justify a more complicated term like say the second term in the third line, $(1+\frac{1}{2}\rho)(k-1)_2\, x^3 y^2$. This term counts the copies of C_5 that have vertices $v_1, \ldots, v_5$ such that v_1 and v_2 come from Y, v_3 and v_4 are in the same set of size xn, and v_5 is in some other set of size xn (and $v_1, \ldots, v_5$ may be in any order on the cycle). The case where v_1 and v_2 are consecutive in the cycle contributes $\frac{1}{2}(k-1)_2 \rho (yn)^2 (xn)^3$, and the other case contributes $(k-1)_2 (yn)^2 (xn)^3$.

Now for a given integer $k \geqslant 2$ and a real number $1 - 1/k < p < 1 - 1/(k+1)$ we define an optimization problem (P):

$$\begin{aligned} \text{minimize} \quad & f(x,y,\rho) \\ \text{subject to} \quad & (k-1)x + y = 1, \\ & g(x,y,\rho) = p, \\ & x, y \geqslant 0. \end{aligned}$$

Let us denote its solution by $f_{\min}(p) = f(x_0, y_0, \rho_0)$. Clearly, $d_{C_5}(p) \leqslant f_{\min}(p)$. For some certain values of k and p we verified that $120 \cdot f_{\min}(p)$ numerically matches the lower bound on $d_{C_5}(p)$ given by the flag algebras. In particular, when we calculated with unlabelled flags of order ℓ, we were getting numerically matching bounds for $p \leqslant 1 - 1/(\ell-2)$ and we observed a gap in the bounds for $p > 1 - 1/(\ell-2)$ different from Turán densities. Since computer calculations can be performed with current computers in a reasonable time only for $\ell \leqslant 8$, a simple straightforward use of computer is unlikely to provide a numerical match of $d_{C_5}(p)$ and $f_{\min}(p)$ for all p. Unfortunately, we were unable to convert the numerical match to a formal proof. The main problem is that (P) has no

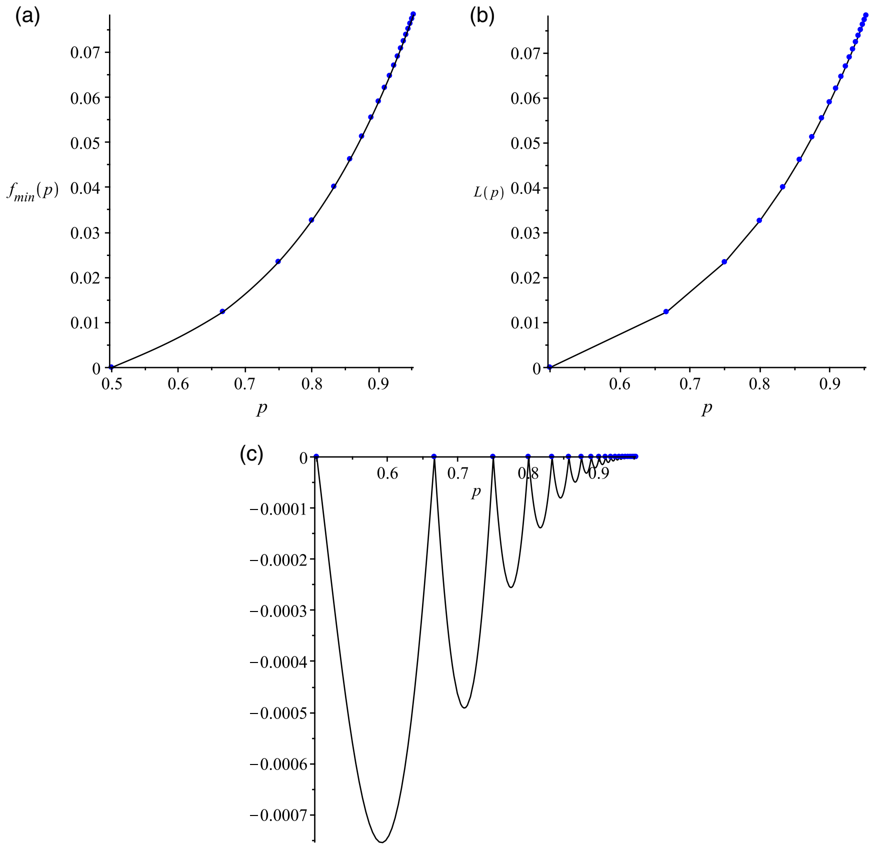


Figure 2. (a) A graph of $f_{\min}(p)$ based on numerical calculations. Blue points correspond to the Turán densities (*i.e.* $p = 1 - 1/k$). (b) Secant lines between Turán densities. (c) A graph of $f_{\min}(p) - L(p)$.

closed solution. For example, for $k = 2$ and $1/2 < p < 2/3$ we can plug into the objective function $y = 1 - x$ and $\rho = (p - x^2 - 2xy)/y^2$ obtaining

$$f(2, x, 1-x, (p - x^2 - 2xy)/y^2) = \frac{x(2x^2 - 2x + p)(3x^4 - 5x^3 + (1+4p)x^2 + (1-4p)x + p^2)}{2(x-1)^2}.$$

Now it is not difficult to show that there exists a local minimum for some $1/3 < x < 1/2$. Unfortunately, it looks like this minimum can be only found numerically. There might be a different parametrization of the problem that would make it possible to solve (P) and formally show a match with flag algebra calculations for some range of p. In Figure 2 we present the shape of $f_{\min}(p)$. We conjecture that $d_{C_5}(p) = f_{\min}(p)$ for any p.

We now address what graphs are suitable for $G[Y]$, that is, what graphs satisfy (i) and (ii). Note first that some such choice of $G[Y]$ exists, for example it can be a random bipartite graph with two parts of size $\frac{1}{2}yn$ and edge probability 2ρ. Now we claim that $G[Y]$ satisfies (i) if and only if $G[Y]$ is *almost* $yn\rho$*-regular*, or more formally, all but $o(n)$ vertices in $G[Y]$ have degree $(1 + o(1))yn\rho$. Indeed, if $G[Y]$ is almost $yn\rho$-regular then it is easy to verify the edge and path counts in (i). Conversely, suppose (i) holds, and let the random variable Z represent the degree of a random vertex in $G[Y]$. Then we have $\mathbb{E}[Z] = (1 + o(1))yn\rho$ and since $\sum_{v \in Y} \binom{\deg(v)}{2}$ is the number of paths of length 2 we can calculate

$$\mathbb{E}[Z^2] = \frac{1}{yn} \sum_{v \in V(Y)} \deg(v)^2 = \frac{1}{yn} \cdot 2(1+o(1))\frac{1}{2}y^3n^3\rho^2 = (1+o(1))y^2n^2\rho^2 = (1+o(1))\mathbb{E}[Z]^2,$$

so Z is concentrated by Chebyshev's inequality (see *e.g.* Lemma 20.3 in [7]). In other words, $G[Y]$ is almost $yn\rho$-regular.

We believe that we have described all almost optimal graphs. Specifically, we believe that any graph with edge density p and C_5-density $d_{C_5}(p) + o(1)$ can be transformed by adding or deleting at most $o(n^2)$ edges into a graph with a vertex partition $X_1, \ldots, X_{k-1}, Y$ where $|X_i| = xn$, $|Y| = yn$, all X_i are independent, all X_i and Y are complete to each other, and $G[Y]$ is $yn\rho$-regular where x, y, ρ are a solution to the optimization problem (P).

References

[1] Alon, N., Fischer, E., Krivelevich, M. and Szegedy, M. (2000) Efficient testing of large graphs. *Combinatorica* **20** 451–476.

[2] Balogh, J., Hu, P., Lidický, B. and Pfender, F. (2016) Maximum density of induced 5-cycle is achieved by an iterated blow-up of 5-cycle. *European J. Combin.* **52** (A) 47–58.

[3] Borchers, B. (1999) CSDP, a C library for semidefinite programming. *Optim. Methods Softw.* **11** 613–623.

[4] Conlon, D. and Fox, J. (2013) Graph removal lemmas. In *Surveys in Combinatorics 2013*, Vol. 409 of London Mathematical Society Lecture Note Series, Cambridge University Press, pp. 1–49.

[5] Erdős, P. (1984) On some problems in graph theory, combinatorial analysis and combinatorial number theory. In *Graph Theory and Combinatorics (Cambridge, 1983)*, Academic Press, pp. 1–17.

[6] Erdős, P. and Stone, A. H. (1946) On the structure of linear graphs. *Bull. Amer. Math. Soc.* **52** 1087–1091.

[7] Frieze, A. and Karoński, M. (2016) *Introduction to Random Graphs*, Cambridge University Press.

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

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

[10] Lidický, B. and Pfender, F. (2018) Pentagons in triangle-free graphs. *European J. Combin.* **74** 85–89.

[11] Liu, H., Pikhurko, O., Sharifzadeh, M. and Staden, K. (2019) Stability from symmetrisation arguments. Work in progress.

[12] Liu, H., Pikhurko, O. and Staden, K. (2017) The exact minimum number of triangles in graphs of given order and size. arXiv:1712.00633

[13] Lovász, L. (2012) *Large Networks and Graph Limits*, Colloquium Publications, American Mathematical Society.

[14] Lovász, L. and Simonovits, M. (1983) On the number of complete subgraphs of a graph, II. In *Studies in Pure Mathematics* (P. Erdős *et al.*, eds), Birkhäuser, pp. 459–495.

[15] Lovász, L. and Szegedy, B. (2006) Limits of dense graph sequences. *J. Combin. Theory Ser. B* **96** 933–957.

[16] Lovász, L. and Szegedy, B. (2010) Testing properties of graphs and functions. *Israel J. Math.* **178** 113–156.

[17] Mantel, W. (1907) Problem 28. *Winkundige Opgaven* **10** 60–61.

[18] Moon, J. W. and Moser, L. (1962) On a problem of Turán. *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **7** 283–286.

[19] Nikiforov, V. (2011) The number of cliques in graphs of given order and size. *Trans. Amer. Math. Soc.* **363** 1599–1618.

[20] Nordhaus, E. A. and Stewart, B. M. (1963) Triangles in an ordinary graph. *Canad. J. Math.* **15** 33–41.

[21] Pikhurko, O. (2010) An analytic approach to stability. *Discrete Math.* **310** 2951–2964.

[22] Pikhurko, O. and Razborov, A. (2017) Asymptotic structure of graphs with the minimum number of triangles. *Combin. Probab. Comput.* **26** 138–160.

[23] Pikhurko, O., Sliacan, J. and Tyros, K. (2019) Strong forms of stability from flag algebra calculations. *J. Combin. Theory Ser. B* **135** 129–178.

[24] Pippenger, N. and Golumbic, M. C. (1975) The inducibility of graphs. *J. Combin. Theory Ser. B* **19** 189–203.

[25] Razborov, A. (2007) Flag algebras. *J. Symbolic Logic* **72** 1239–1282.

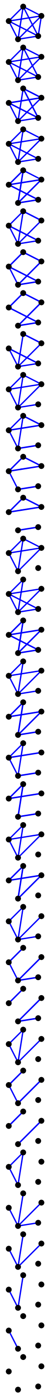
[26] Razborov, A. (2008) On the minimal density of triangles in graphs. *Combin. Probab. Comput.* **17** 603–618.

[27] Reiher, C. (2016) The clique density theorem. *Ann. of Math.* (2) **184** 683–707.

[28] Sidorenko, A. (1991) Inequalities for functionals generated by bipartite graphs. *Diskret. Mat.* **3** 50–65.

[29] Turán, P. (1941) Eine Extremalaufgabe aus der Graphentheorie. *Mat. Fiz. Lapok* **48** 436–452.

Appendix A



F																	
c_F^{OPT}	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$p(K_2,F)$	0	1	2	3	4	3	2	3	4	3	4	5	4	5	4	5	5
$[\![X_1 \times X_1]\!]_1$	30	12	4	0	0	0	4	2	0	0	0	0	2	0	0	0	0
$[\![X_1 \times X_2]\!]_1$	0	3	4	3	0	6	0	1	2	0	0	0	0	1	0	0	0
$[\![X_1 \times X_3]\!]_1$	0	6	4	3	0	0	8	2	0	6	2	0	0	0	2	0	1
$[\![X_1 \times X_4]\!]_1$	0	0	2	6	12	0	0	2	2	0	3	4	0	0	0	2	0
$[\![X_1 \times X_5]\!]_1$	0	0	1	0	0	0	0	2	0	0	1	0	4	0	1	0	2
$[\![X_1 \times X_6]\!]_1$	0	0	0	0	0	3	0	0	2	0	0	2	0	2	0	1	0
$[\![X_2 \times X_2]\!]_1$	0	0	0	0	0	0	2	2	2	0	0	0	4	4	0	0	0
$[\![X_2 \times X_3]\!]_1$	0	0	0	0	0	0	4	2	1	4	2	0	0	0	2	2	0
$[\![X_2 \times X_4]\!]_1$	0	0	0	0	0	0	0	0	0	2	2	2	0	0	2	2	2
$[\![X_2 \times X_5]\!]_1$	0	0	0	0	0	0	0	0	0	1	0	0	0	0	2	0	1
$[\![X_2 \times X_6]\!]_1$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$[\![X_3 \times X_3]\!]_1$	0	0	4	0	0	12	0	4	4	4	0	0	0	0	6	2	0
$[\![X_3 \times X_4]\!]_1$	0	0	0	0	0	0	0	2	2	0	2	4	8	4	2	0	4
$[\![X_3 \times X_5]\!]_1$	0	0	0	3	0	0	0	0	2	0	2	0	0	2	0	2	1
$[\![X_3 \times X_6]\!]_1$	0	0	0	0	0	0	0	0	1	0	0	0	0	4	0	2	0
$[\![X_4 \times X_4]\!]_1$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2	4
$[\![X_4 \times X_5]\!]_1$	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	2	2
$[\![X_4 \times X_6]\!]_1$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$[\![X_5 \times X_5]\!]_1$	0	0	0	0	6	0	0	0	0	0	0	4	0	0	0	0	0
$[\![X_5 \times X_6]\!]_1$	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
$[\![X_6 \times X_6]\!]_1$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

F																	
c_F^{OPT}	0	0	0	0	0	0	0	0	0	1	1	1	2	2	4	6	12
$p(K_2,F)$	6	6	7	6	4	5	6	7	6	5	6	7	7	8	8	9	10
$[\![X_1 \times X_1]\!]_1$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$[\![X_1 \times X_2]\!]_1$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$[\![X_1 \times X_3]\!]_1$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$[\![X_1 \times X_4]\!]_1$	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$[\![X_1 \times X_5]\!]_1$	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$[\![X_1 \times X_6]\!]_1$	2	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$[\![X_2 \times X_2]\!]_1$	0	0	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0
$[\![X_2 \times X_3]\!]_1$	0	0	0	0	6	2	1	0	0	0	0	0	0	0	0	0	0
$[\![X_2 \times X_4]\!]_1$	2	0	0	0	0	1	2	3	0	0	0	0	0	0	0	0	0
$[\![X_2 \times X_5]\!]_1$	0	0	0	0	0	1	0	0	0	5	2	0	1	0	0	0	0
$[\![X_2 \times X_6]\!]_1$	0	0	0	0	3	2	1	0	4	0	1	2	0	1	0	0	0
$[\![X_3 \times X_3]\!]_1$	0	0	0	0	12	4	0	0	0	10	2	0	0	0	0	0	0
$[\![X_3 \times X_4]\!]_1$	2	0	0	0	0	4	2	0	8	0	2	2	0	0	0	0	0
$[\![X_3 \times X_5]\!]_1$	0	0	0	0	0	2	2	0	0	0	2	0	2	0	0	0	0
$[\![X_3 \times X_6]\!]_1$	2	0	0	12	0	0	3	6	0	0	0	2	0	2	0	0	0
$[\![X_4 \times X_4]\!]_1$	4	12	12	0	0	0	0	0	0	10	6	4	4	4	0	0	0
$[\![X_4 \times X_5]\!]_1$	1	6	0	0	0	0	2	0	0	0	2	2	4	0	4	0	0
$[\![X_4 \times X_6]\!]_1$	0	0	0	0	0	1	2	3	0	0	2	2	6	4	8	6	0
$[\![X_5 \times X_5]\!]_1$	2	0	0	0	0	0	0	0	4	0	0	2	0	0	2	0	0
$[\![X_5 \times X_6]\!]_1$	2	0	6	0	0	0	0	3	0	0	0	1	0	4	0	3	0
$[\![X_6 \times X_6]\!]_1$	0	0	0	0	0	0	0	0	2	0	0	2	0	4	4	12	30

Appendix B

This Maple code computes c_F coefficients. Matrices A, B and M are defined in Section 2.2.2. X is a matrix of size 21×34 and it is defined in Appendix A (rows correspond to $[\![X_i \times X_j]\!]_1$). Vectors cFOPT, pF, cFM and cF (each of size 34) correspond to c_F^{OPT}, $p(K_2, F)$, c_F^M and c_F, respectively. Constant a corresponds to α.

```
restart:
with(LinearAlgebra):
A := Matrix([[32*k^2-96*k+96, 0, 4*k^2-16*k],
[0, 10*k^4-30*k^3-8*k^2+96*k-96, -10*k^4+35*k^3-4*k^2-80*k+96],
[4*k^2-16*k, -10*k^4+35*k^3-4*k^2-80*k+96, 10*k^4-40*k^3+24*k^2+64*k-96]]):
B := Matrix([[k-1, 1, k-2, 0, k-3, -1],
[0, 2, k-2, 0, 2*k-4, -2],
[0, 0, k-1, -1, 2*k-2, -2]]):
M:= (3/(2*k^4))*Matrix(Multiply(Transpose(B), Multiply(A, B))):
X:=(1/30)*Matrix([[30,12,4,0,0,0,4,2,0,0,0,0,2,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0],
[0,3,4,3,0,6,0,1,2,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0],
[0,6,4,3,0,0,8,2,0,6,2,0,0,0,2,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0],
[0,0,2,6,12,0,0,2,2,0,3,4,0,0,0,2,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0],
[0,0,1,0,0,0,0,2,0,0,1,0,4,0,1,0,2,0,3,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0],
[0,0,0,0,0,3,0,0,2,0,0,2,0,2,0,1,0,2,0,3,0,0,0,0,0,0,0,0,0,0,0,0,0,0],
[0,0,0,0,0,0,2,2,2,0,0,0,4,4,0,0,0,0,0,6,0,0,0,0,0,0,0,0,0,0,0,0,0,0],
[0,0,0,0,0,0,4,2,1,4,2,0,0,0,2,2,0,0,0,0,0,6,2,1,0,0,0,0,0,0,0,0,0,0],
[0,0,0,0,0,0,0,0,0,2,2,2,0,0,2,2,2,2,0,0,0,0,1,2,3,0,0,0,0,0,0,0,0,0],
[0,0,0,0,0,0,0,0,0,1,0,0,0,0,2,0,1,0,0,0,0,0,1,0,0,0,5,2,0,1,0,0,0,0],
[0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,3,2,1,0,4,0,1,2,0,1,0,0,0],
[0,0,4,0,0,12,0,4,4,4,0,0,0,0,6,2,0,0,0,0,0,12,4,0,0,0,10,2,0,0,0,0,0,0],
[0,0,0,0,0,0,0,2,2,0,2,4,8,4,2,0,4,2,0,0,0,0,4,2,0,8,0,2,2,0,0,0,0,0],
[0,0,0,3,0,0,0,0,2,0,2,0,0,2,0,2,1,0,0,0,0,0,2,2,0,0,0,2,0,2,0,0,0,0],
[0,0,0,0,0,0,0,0,1,0,0,0,0,4,0,2,0,2,0,0,12,0,0,3,6,0,0,0,2,0,2,0,0,0],
[0,0,0,0,0,0,0,0,0,0,0,0,0,0,2,2,4,4,12,12,0,0,0,0,0,0,10,6,4,4,4,0,0,0],
[0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,2,2,1,6,0,0,0,0,2,0,0,0,2,2,4,0,4,0,0],
[0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,2,3,0,0,2,2,6,4,8,6,0],
[0,0,0,0,6,0,0,0,0,0,0,4,0,0,0,0,0,2,0,0,0,0,0,0,0,4,0,0,2,0,0,2,0,0],
[0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,2,0,6,0,0,0,0,3,0,0,0,1,0,4,0,3,0],
[0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,2,0,0,2,0,4,4,12,30]]):
cFM := Vector(34):
k_ind := 0:
printlevel := 2:
for i to 6 do
    for j from i to 6 do
        k_ind := k_ind+1;
        if i = j then cFM := cFM+M(i, j)*Transpose(Row(X, k_ind));
                 else cFM := cFM+2*M(i, j)*Transpose(Row(X,k_ind));
        end if;
    end do;
end do:
cFOPT := Vector([0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,1,1,2,2,4,6,12]):
pF := (1/10)*Vector([0,1,2,3,4,3,2,3,4,3,4,5,4,5,4,5,5,6,6,7,6,4,5,6,7,6,5,6,7,7,8,8,9,10]):
a := (1/(k^3))*(60*k^3 - 240*k^2 + 360*k - 192):
cF := Vector(34):
for i to 34 do
    cF(i) := cFOPT(i)-a*pF(i)-cFM(i)+(k-1)*a/k
end do:
```

```
for i to 34 do
    printf("5*k^4*cF(%d) = %s\n", i, convert(expand(5*k^4*cF(i)), string))
end do:
kernel := NullSpace(B):
kernelMatrix := Matrix(convert(kernel, list)):
ReducedRowEchelonForm(Transpose(kernelMatrix))
```

Cite this article: Bennett P, Dudek A, Lidický B and Pikhurko O (2019). Minimizing the number of 5-cycles in graphs with given edge-density. *Combinatorics, Probability and Computing*. https://doi.org/10.1017/S0963548319000257