Subgraph densities in signed graphons and the local Simonovits–Sidorenko conjecture

László Lovász[*]

Institute of Mathematics, Eötvös Loránd University
Budapest, Hungary

Submitted: Feb 17, 2011; Accepted: Jun 6, 2011; Published: Jun 14, 2011
Mathematics Subject Classification: 05C35

Abstract

We prove inequalities between the densities of various bipartite subgraphs in signed graphs. One of the main inequalities is that the density of any bipartite graph with girth $2r$ cannot exceed the density of the $2r$-cycle.

This study is motivated by the Simonovits–Sidorenko conjecture, which states that the density of a bipartite graph F with m edges in any graph G is at least the m-th power of the edge density of G. Another way of stating this is that the graph G with given edge density minimizing the number of copies of F is, asymptotically, a random graph. We prove that this is true locally, i.e., for graphs G that are "close" to a random graph.

Both kinds of results are treated in the framework of graphons (2-variable functions serving as limit objects for graph sequences), which in this context was already used by Sidorenko.

[*]Research supported by ERC Grant No. 227701.

Contents

1 Introduction

Let F be a bipartite graph with k nodes and l edges and let G be any graph with n nodes and $m = p\binom{n}{2}$ edges. Simonovits [3, 10] conjectured that the number of copies of F in G is at least $p^l\binom{n}{k} + o(p^l n^k)$ (where we consider k and l fixed, and $n \to \infty$).

Sidorenko [7, 8, 9] conjectured a stronger exact inequality. To state this formulation, we count homomorphisms instead of copies of F. Let $\hom(F, G)$ denote the number of homomorphisms from F into G. Since we need this notion for the case when F and G are multigraphs, we count here pairs of maps $\phi : V(F) \to V(G)$ and $E(F) \to E(G)$ such that incidence is preserved: if $i \in V(F)$ is incident with $e \in E(F)$, then $\phi(i)$ is incident with $\psi(e)$. We will also consider the normalized version $t(F, G) = \hom(F, G)/n^k$. If F and G are simple, then $t(F, G)$ is the probability that a random map $\phi : V(F) \to V(G)$ preserves adjacency. We call this quantity the *density of F in G*.

In this language, the conjecture says that for any bigraph F and any graph G,

$$t(F, G) \geq t(K_2, G)^{|E(F)|} \tag{1}$$

(this is an exact inequality with no error terms). We can formulate this as an extremal result in two ways: First, for every graph G, among all bipartite graphs with a given number of edges, it is the graph consisting of disjoint edges (the matching) that has the smallest density in G. Second, for every bipartite graph F, among all graphs on n nodes and edge density p, the random graph $\mathbb{G}(n, p)$ has the smallest density of F in it (asymptotically, with large probability).

Sidorenko proved his conjecture in a number of special cases: for trees F, and also for bigraphs F where one of the color classes has at most 4 nodes. Since then, the only substantial progress was that Hatami [4] proved the conjecture for cubes, and Conlon, Fox and Sudakov [2] proved it for bigraphs having a node connected to all nodes on the other side.

Sidorenko gave an analytic formulation of this conjecture, which we will use. Let F be a bipartite multigraph with a bipartition (A, B); if we say that $ij \in E(F)$, we assume that the labeling is such that $i \in A$ and $j \in B$. Assign a real variable x_i to each $i \in A$ and a real variable y_j to each $j \in B$. Let $W: [0,1]^2 \to \mathbb{R}_+$ be a bounded measurable function, and define

$$t(F,W) = \int\limits_{[0,1]^{V(F)}} \prod_{ij\in E(F)} W(x_i, y_j) \prod_{i\in A} dx_i \prod_{j\in B} dy_j. \tag{2}$$

Every graph G can be represented by a function W_G: Let $V(G) = \{1, \ldots, n\}$. Split the interval $[0,1]$ into n equal intervals $J_1, \ldots, J_n$, and for $x \in J_i, y \in J_j$ define $W_G(x, y) = \mathbb{1}_{ij\in E(G)}$. (The function obtained this way is symmetric.) Then we have

$$t(F, G) = t(F, W_G).$$

Note, however, that definition (2) makes sense without assuming that W is symmetric.

In this analytic language, the conjecture says that for every bipartite graph F and bounded measurable function $W: [0,1]^2 \to \mathbb{R}_+$, we have

$$t(F,W) \geq t(K_2, W)^{|E(F)|}. \tag{3}$$

Since both sides are homogeneous in W of the same degree, we can scale W and assume that

$$t(K_2, W) = \int\limits_{[0,1]^2} W(x, y)\, dx\, dy = 1.$$

Then we want to conclude that $t(F, W) \geq 1$. In other words, the function $W \equiv 1$ minimizes $t(F, W)$ among all functions $W \geq 0$ with $\int W = 1$.

The goal of this paper is to prove that this holds locally, i.e., for functions W sufficiently close to 1. Most of the time we will work with the function $U = W - 1$, which can take negative values. Most of our work will concern estimates for the values $t(F', U)$ for various (bipartite) graphs F'. This type of question seems to have some interest on its own, because it can be considered as an extension of extremal graph theory to signed graphs.

2 Preliminaries

2.1 Notation

A *bigraph* will mean a bipartite multigraph with a fixed bipartition, in which a first and second bipartition class is specified. So the complete bigraphs $K_{a,b}$ and $K_{b,a}$ are different.

We have to consider graphs that are partially labeled. More precisely, a *k-labeled graph* F has a subset $S \subseteq V(F)$ of k elements labeled $1, \ldots, k$ (it can have any number of unlabeled nodes). For some basic graphs, it is good to introduce notation for some of their labeled versions. Let P_n denote the unlabeled path with n nodes (so, with $n-1$ edges). Let $P_n^\bullet$ denote the path P_n with one of its endpoints labeled. Let $P_n^{\bullet\bullet}$ denote the P_n with both of its endpoints labeled. Let C_n denote the unlabeled cycle with n nodes, and let $C_n^\bullet$ be this cycle with one of its nodes labeled. Let $K_{a,b}$ denote the unlabeled complete bigraph; let $K_{a,b}^\bullet$ denote the complete bigraph with its first bipartition class labeled. Note that $K_{2,2} \cong C_4$, but $K_{2,2}^\bullet$ and $C_4^\bullet$ are different as partially labeled graphs.

We extend the definition of subgraph densities to k-labeled graphs. Let F be a graph on node set $[n]$, of which nodes $1, \ldots, k$ are considered as labeled. For given $x_1, \ldots, x_k \in I$, we define

$$t_{x_1 \ldots x_k}(F, W) = \int\limits_{[0,1]^{n-k}} \prod_{ij \in E(F)} W(x_i, x_j)\, dx_{k+1} \ldots dx_n$$

(this is a function of $x_1, \ldots, x_k$).

The most important use of partial labeling is to define a *product*: if F and G are k-labeled graphs, then FG denotes the k-labeled graph obtained by taking their disjoint union and identifying nodes with the same label. For a k-labeled graph F, $[[F]]$ denotes the graph obtained by unlabeling all nodes. The graph O_k with k labeled nodes, no unlabeled nodes and no edges is a unit element: $O_k F = F$ for every k-labeled graph F.

2.2 Kernel operators and their norms

We set $I = [0,1]$. Let $\mathcal{W}$ denote the set of bounded measurable functions $U: I^2 \to \mathbb{R}$; $\mathcal{W}_+$ is the set of bounded measurable functions $U: I^2 \to \mathbb{R}_+$, and $\mathcal{W}_1$ is the set of measurable functions $U: I^2 \to [-1,1]$. Every function $U \in \mathcal{W}$ defines a kernel operator $L_1(f) \to L_1(f)$ by

$$f \mapsto \int\limits_I U(., y) f(y)\, dy.$$

For $U, W \in \mathcal{W}$, we denote by $U \circ W$ the function

$$(U \circ W)(x, y) = \int\limits_I U(x, z) W(z, y)\, dz$$

(this corresponds to the product of U and W as kernel operators). For every $W \in \mathcal{W}$, we denote by $W^\top$ the function obtained by interchanging the variables in W.

We will also need the *tensor product* $U \otimes W$ of two functions $U, W \in \mathcal{W}$; this is defined as a function $I^2 \times I^2 \to \mathbb{R}$ by

$$(U \otimes W)(x_1, x_2, y_1, y_2) = U(x_1, y_1)W(x_2, y_2).$$

This function is not in $\mathcal{W}$; however, we can consider any measure preserving map $\varphi : I \to I^2$, and define the function

$$(U \otimes W)^\varphi(x, y) = (U \otimes W)(\varphi(x), \varphi(y)).$$

It does not really matter which particular measure preserving map we use here: these functions obtained from different maps ϕ have the same subgraph densities. In fact, we have

$$t(F, (U \otimes W)^\phi) = t(F, U \otimes W) = t(F, U)t(F, W) \tag{4}$$

for every graph F. We will call any of the functions $(U \otimes W)^\phi$ the tensor product of U and W.

We consider various norms on the space $\mathcal{W}$. We need the standard L_2 and L_∞ norms

$$\|U\|_2 = \Big(\int\limits_{I^2} U(x, y)^2\, dx\, dy\Big)^{1/2}, \qquad \|U\|_\infty = \text{sup ess } |U(x, y)|.$$

For graph theory, the *cut norm* is very useful:

$$\|U\|_\square = \sup_{S,T \subseteq I} \Big| \int\limits_{S \times T} U(x, y)\, dx\, dy \Big|.$$

This norm is only a factor of at most 4 away from the operator norm of U as a kernel operator $L_\infty(I) \to L_1(I)$.

The functional $t(F, U)$ gives rise to further useful norms. It is trivial that $t(C_2, U)^{1/2} = \|U\|_2$. The value $t(C_{2r}, U)^{1/(2r)}$ is the r-th *Schatten norm* of the kernel operator defined by U. It was proved in [1] that it is closely related to the cut norm: for $U \in \mathcal{W}_1$,

$$\|U\|_\square^4 \le t(C_4, U) \le 4\|U\|_\square. \tag{5}$$

The other Schatten norms also define the same topology on $\mathcal{W}_1$ as the cut norm (cf. Corollary 3.12).

It is a natural question for which graphs does $t(F, W)^{1/|E(F)|}$ or $t(F, |W|)^{1/|E(F)|}$ define a norm on $\mathcal{W}$. Besides even cycles and complete bigraphs, a remarkable class was found by Hatami [4]: he proved that $t(F, |W|)^{1/|E(F)|}$ is a norm if F is a cube. He also proved the fact (attributed to B. Szegedy) that Sidorenko's conjecture is true whenever F is such a "norming" graph. However, a characterization of such graphs is open.

3 Density inequalities for signed graphons

3.1 Ordering signed graphons

For two bigraphs F and G, we say that $F \le G$ if $t(F,U) \le t(G,U)$ for all $U \in \mathcal{W}_1$. We say that $G \ge 0$ if $t(G,U) \ge 0$ for all $U \in \mathcal{W}_1$. Note that if U is nonnegative, then trivially $G \subseteq F$ implies that $t(F,U) \le t(G,U)$; but since we allow negative values, such an implication does not hold in general. For example, $F \ge 0$ cannot hold for any bigraph F with an odd number of edges, since then $t(F,-U) = -t(F,U)$.

The ordering is a bit counterintuitive since larger graphs tend to be smaller in the ordering. For example, $t(F,U) \le 1 = t(K_0,U) = t(K_1,U)$ for every U, so $F \le K_1$ and $F \le K_0$ for any bigraph F (here K_1 may have its single node either in its first or second color class, and K_0 is the empty graph). Lemmas 3.9 and 3.15 provide other examples.

We start with some simple facts about this partial order on graphs.

Proposition 3.1 *If F and G are nonisomorphic bigraphs without isolated nodes such that $F \le G$, then $|E(F)| \ge |E(G)|$, $|E(G)|$ is even, and $G \ge 0$. Furthermore, $|t(F,U)| \le t(G,U)$ for all $U \in \mathcal{W}_1$.*

The proof of this is based on a technical lemma, which is close to facts that are well known, but not in the exact form needed here.

Lemma 3.2 *Let F and G be nonisomorphic bigraphs without isolated nodes. Then for every $U \in \mathcal{W}_1$ and $\varepsilon > 0$ there exists a function $U' \in \mathcal{W}_1$ such that $\|U - U'\|_\infty < \varepsilon$ and $t(F,U') \ne t(G,U')$.*

A similar assertion (with a similar proof) holds in the context of non-bipartite graphs as well.

Proof. First we show that if F and G are two bigraphs without isolated nodes such that $t(F,W) = t(G,W)$ for every $W \in \mathcal{W}_1$, then $F \cong G$. Consider the function $U = \mathbb{1}_{x,y \le 1/2}$. Then $t(F,U) = 2^{-|V(F)|}$, so $t(F,U) = t(G,U)$ implies that $|V(F)| = |V(G)|$. Using the function $U \equiv 1/2$, we get similarly that $|E(F)| = |E(G)|$. Using this, we get (by scaling W) that $t(F,W) = t(G,W)$ for every $W \in \mathcal{W}$.

For every multigraph H we have

$$t(F,H) = t(F,W_H) = t(G,W_H) = t(G,H),$$

and hence it follows that

$$\hom(F,H) = t(F,H)|V(H)|^{|V(F)|} = t(G,H)|V(G)|^{|V(F)|} = \hom(G,H).$$

From this it follows by standard arguments that $F \cong G$ (e.g., we can apply Theorem 1(iii) of [5] to the 2-partite structures (V,E,J), where $G = (V,E)$ is a multigraph and J is the incidence relation between nodes and edges).

Since F and G are non-isomorphic, this argument shows that there exists a function $W \in \mathcal{W}_1$ such that $t(F,W) \neq t(G,W)$. The values $t(F,(1-s)U+sW)$ and $t(F,(1-s)U+sW)$ are polynomials in s that differ for $s=1$. Therefore, there is a value $0 \leq s \leq \varepsilon$ for which they differ. Since $(1-s)U+sW \in \mathcal{W}_1$ and $\|U-((1-s)U+sW)\|_\infty = s\|U-W\|_\infty \leq \varepsilon$, this proves the lemma. □

Proof of Proposition 3.1. Applying the definition of $F \leq G$ with $U = 1/2$, we get that $2^{-|E(F)|} \leq 2^{-|E(G)|}$, and hence $|E(F)| \geq |E(G)|$. The relation $F \leq G$ implies that $t(F,U)^2 = t(F, U \otimes U) \leq t(G, U \otimes U) = t(G,U)^2$ also holds, so $|t(F,U)| \leq |t(G,U)|$ for all $U \in \mathcal{W}_1$. By Lemma 3.2, U can be perturbed by arbitrarily little to get a $U' \in \mathcal{W}_1$ with $t(F,U') \neq t(G,U')$, then $t(F,U') < t(G,U')$ and $|t(F,U')| \leq |t(G,U')|$ imply that $t(G,U') > 0$. Since U' is arbitrarily close to U, this implies that $t(G,U) \geq 0$, and so $G \geq 0$. Since this holds for U replaced by $-U$, it follows that G must have an even number of edges. □

3.2 A generalized Cauchy-Schwarz inequality

We need the following generalization of the Cauchy–Schwarz inequality:

Lemma 3.3 *Let $f_1, \dots, f_n :\ I^k \to \mathbb{R}$ be bounded measurable functions, and suppose that for each variable there are at most two functions f_i that depend on that variable. Then*

$$\int_{I^k} f_1 \dots f_n \leq \|f_1\|_2 \dots \|f_n\|_2.$$

This will follow from an inequality concerning a statistical physics type model. Let $G = (V,E)$ be a multigraph (without loops), and for each $i \in V$, let $f_i \in L_2(I^E)$ be such that f_i depends only on the variables x_j where edge j is incident with node i. Let $f = (f_i :\ i \in V)$, and define

$$\mathrm{tr}(G,f) = \int_{I^E} \prod_{i \in V} f_i(x)\, dx$$

(where the variables corresponding to the edges not incident with i are dummies in f_i).

Lemma 3.4 *For every multigraph G and assignment of functions f,*

$$\mathrm{tr}(G,f) \leq \prod_{i \in V} \|f_i\|_2.$$

Proof. By induction on the chromatic number of G. Let $V_1, \dots, V_r$ be the color classes of an optimal coloring of G. Let $S_1 = V_1 \cup \dots \cup V_{\lfloor r/2 \rfloor}$ and $S_2 = V \setminus S_1$. Let E_0 be the set of edges between S_1 and S_2, and let E_i be the set of edges induced by S_i. Let x_i be the vector formed by the variables in E_i. Then

$$\mathrm{tr}(G,f) = \int\limits_{I^{E_0}} \Bigg(\int\limits_{I^{E_1}} \prod_{i\in S_1} f_i(x)\,dx_1 \Bigg) \Bigg(\int\limits_{I^{E_2}} \prod_{i\in S_2} f_i(x)\,dx_2 \Bigg)\,dx_0.$$

The outer integral can be estimated using the Cauchy-Schwarz inequality:

$$\mathrm{tr}(G,f)^2 \le \int\limits_{I^{E_0}} \Bigg(\int\limits_{I^{E_1}} \prod_{i\in S_1} f_i(x)\,dx_1 \Bigg)^2 dx_0 \int\limits_{I^{E_0}} \Bigg(\int\limits_{I^{E_2}} \prod_{i\in S_2} f_i(x)\,dx_2 \Bigg)^2 dx_0. \tag{6}$$

Let G_1 be defined as the graph obtained by taking a disjoint copy (S_1', E_1') of the graph (S_1, E_1), and connecting each node $i \in S_1$ to the corresponding node $i' \in S_1'$ by as many edges as those joining i to S_2 is G. Note that these newly added edges correspond to the edges of E_0 in a natural way. We assign to each node the same function as before, and also the same function (with differently named variables for the edges in E_1') to i'. Then the first factor in (6) can be written as

$$\int\limits_{I^{E_0}} \int\limits_{I^{E_1}} \int\limits_{I^{E_1'}} \prod_{i\in S_1\cup S_1'} f_i(x)\,dx_1\,dx_0 = \mathrm{tr}(G_1, f).$$

We define G_2 analogously, and get that the second factor in (6) is just $\mathrm{tr}(G_2, f)$. So we have

$$\mathrm{tr}(G,f)^2 \le \mathrm{tr}(G_1,f)\mathrm{tr}(G_2,f) \tag{7}$$

Next we remark that for $r > 2$, the graphs G_1 and G_2 have chromatic number at most $\lceil r/2 \rceil < r$, and so we can apply induction and use that

$$\mathrm{tr}(G_j, f) \le \prod_{i\in V(G_j)} \|f_i\|_2 = \prod_{i\in S_j} \|f_i\|_2^2.$$

If $r = 2$, then G_j has edges connecting pairs i, i' only, and so

$$\mathrm{tr}(G_j, f) = \prod_{i\in S_j} \|f_i\|_2^2.$$

In both cases, the inequality in the lemma follows by (7). □

3.3 Inequalities between densities

Let F_1 and F_2 be two k-labeled graphs. Then the Cauchy–Schwarz inequality implies that for all $U \in \mathcal{W}$,

$$t([\![F_1F_2]\!], U)^2 \le t([\![F_1^2]\!], U)t([\![F_2^2]\!], U). \tag{8}$$

With the notation introduced above, this can be written as

$$[\![F_1F_2]\!]^2 \le [\![F_1^2]\!][\![F_2^2]\!]. \tag{9}$$

Choosing $F_2 = O_k$, we get that for every k-labeled graph F,

$$[\![F^2]\!] \ge [\![F]\!]^2 \ge 0. \tag{10}$$

Let F^{sub} denote the subdivision of graph F obtained by adding one new node on each edge.

Lemma 3.5 *If $F \le G$, then $F^{\mathrm{sub}} \le G^{\mathrm{sub}}$.*

Proof. For every $U \in \mathcal{W}$, $t(F^{\mathrm{sub}}, U) = t(F, U \circ U^\top) \le t(G, U \circ U^\top) = t(G^{\mathrm{sub}}, U)$. $\square$

The next lemma will be the workhorse throughout this paper.

Lemma 3.6 *Let F be an (unlabeled) bigraph, let $S \subseteq V(F)$, and let $H_1, \dots, H_m$ be the connected components of $F \setminus S$. Assume that each node in S has neighbors in at most two of the H_i. Let F_i denote the graph consisting of H_i, its neighbors in S, and the edges between H_i and S. Let us label the nodes of S in every F_i. Then*

$$F^2 \le \prod_{i=1}^m [\![F_i^2]\!].$$

Proof. Let F_0 denote the subgraph induced by S, and consider the nodes of F_0 labeled $1, \dots, k$; we may assume that these nodes are labeled the same way in every F_i. Then using that $|t_{x_1\dots x_k}(F_0, U)| \le 1$, we get

$$\begin{aligned} |t(F,U)| &= \Big|\int\limits_{I^k} \prod_{i=0}^m t_{x_1\dots x_k}(F_i, U)\, dx_1 \dots dx_k\Big| \\ &\le \int\limits_{I^k} \prod_{i=1}^m |t_{x_1\dots x_k}(F_i, U)|\, dx_1 \dots dx_k. \end{aligned}$$

Hence Lemma 3.3 implies the assertion. $\square$

As a special case, we see that if F contains two nonadjacent nodes of degree at least 2, then $F \le C_4$. More generally,

Corollary 3.7 *Let $v_1, \dots, v_m$ be independent nodes in an (unlabeled) bigraph F with degrees $d_1, \dots, d_m$ such that no node of F is adjacent to more than 2 of them. Then $F^2 \le K_{2,d_1} \cdots K_{2,d_m}$. If $d_1, \dots, d_m \ge 2$, then $F^2 \le C_4^m$.*

A *hanging path system* in a graph F is a set $\{P_1,\dots,P_m\}$ of openly disjoint paths such that the internal nodes of each P_i have degree 2, and at most two of them start at any node. Lemma 3.6 can be used to bound the graph in terms of any hanging path system:

Corollary 3.8 *Let F be a bigraph that contains a hanging path system with lengths $r_1,\dots,r_m$. Then $F^2 \le C_{2r_1}\cdots C_{2r_m}$.*

3.4 Special graphs and examples

Lemma 3.9 *Let $U \in \mathcal{W}_1$. Then the sequence $(t(C_{2k},U):\ k=1,2,\dots)$ is nonnegative, logconvex, and monotone decreasing.*

With the notation introduced above, we have $C_2 \ge C_4 \ge C_6 \ge \cdots \ge 0$ and $C_{2k}^2 \le C_{2k-2}C_{2k+2}$.

Proof. We have $C_{a+b} = [[P_a^{\bullet\bullet}P_b^{\bullet\bullet}]]$. Taking $a=b=k$, nonnegativity follows. Applying inequality (9), we get that $C_{a+b}^2 \le C_{2a}C_{2b}$. This implies logconvexity. Since the sequence remains bounded by 1, it follows that it is monotone decreasing. $\square$

Monotonicity and logconvexity of the sequence of even cycles imply inequalities between collections of cycles.

Lemma 3.10 *Let $1 \le r_1 \le \cdots \le r_m$ and $1 \le q_1 \le \cdots \le q_m$ be integers and assume that $\sum_{i=1}^j r_i \ge \sum_{i=1}^j q_i$ for every $1 \le j \le m$. Then*

$$C_{2r_1}\cdots C_{2r_m} \le C_{2q_1}\cdots C_{2q_m}.$$

Proof. We use induction on m and on r_1. For $m=1$ the assertion is just monotonicity. Let $m \ge 2$. If $r_1 = q_1$, we can delete the first member of each list, and apply induction. If $r_1 > q_1$, then let us replace r_1 by r_1-1 and r_2 by r_2+1. It is easy to check that the resulting sequence satisfies the conditions of the Corollary, and so the induction hypothesis applies to it. Furthermore, logconcavity implies that

$$C_{2r_1}C_{2r_2} \le C_{2r_1-2}C_{2r_2+2},$$

and so

$$C_{2r_1}C_{2r_2}\cdots C_{2r_m} \le C_{2r_1-2}C_{2r_2+2}\cdots C_{2r_m} \le C_{2q_1}\cdots C_{2q_m}.$$

$\square$

As a special case of the last corollary, we get that if $r_1,\dots,r_m \ge 1$ and $r = r_1+\cdots+r_m$, then

$$C_{2r_1}\cdots C_{2r_m} \le C_2^{m-1}C_{2r-2m+2} \le C_{2r-2m+2}. \tag{11}$$

The following lemma gives an estimate on the product of even cycles which goes in a sense in the opposite direction.

Lemma 3.11 *Let $r_1, \dots, r_m \geq 1$ and $r = r_1 + \dots + r_m$. Then $C_{2r_1} \cdots C_{2r_m} \geq C_r^2$.*

Proof. We split C_r into paths of lengths $r_1, r_2, \dots, r_m$, and apply Lemma 3.6. $\square$

Choosing $r_1 = r_2 = k$ and $r_3 = 2$ in Lemma 3.11, we get that $C_{2k+2}^2 \leq C_{2k}^2 C_4$. Choosing $m = r - 1$, $r_1 = \dots r_m = 2$, $q_1 = \dots = q_{m-1} = 1$ and $q_m = r$ in Lemma 3.10, we get that $C_4^{r-1} \leq C_2^{r-2} C_{2r}$. Together, these inequalities imply that for every $U \in \mathcal{W}_1$, the density $t(C_{2k}, U)$ tends to 0 exponentially with k (unless $W = 1$ almost everywhere):

Corollary 3.12 *For all $r \geq 2$, $C_4^{r-1} \leq C_{2r} \leq C_4^{r/2}$.*

The *value* of a hanging path system is the total number of their internal nodes. We get by Corollary 3.8 and Lemma 3.10,

Corollary 3.13 *Let F be a simple bigraph that contains a hanging path system with path lengths at most r and value at least $2r-2$. Then $F \leq C_{2r}$. If the value is larger than $2r-2$, then $F^2 \leq C_{2r}^2 C_4$.*

We can get similar inequalities for paths, of which we only state two, which will be needed. Recall that P_n denotes the path with n nodes and $n-1$ edges.

Lemma 3.14 *For all $a, b \geq 1$, we have*

(a) $P_{a+b+1}^2 \leq P_{2a+1} P_{2b+1}$;

(b) $P_{2a+b+1}^4 \leq P_{2a+1}^4 C_{4b}$.

Proof. Since $P_{a+b+1} = [[P_{a+1}^\bullet P_{b+1}^\bullet]]$, the first inequality follows by (9). To get the second, we use the first to get

$$P_{2a+b+1}^2 \leq P_{2a+1} P_{2a+2b+1}.$$

Cut $P_{2a+2b+1}$ into pieces $P_{a+1}^\bullet$, $P_{2b+1}^{\bullet\bullet}$ and $P_{a+1}^\bullet$, and apply Lemma 3.6; we get

$$P_{2a+2b+1}^2 \leq P_{2a+1}^2 C_{4b},$$

and hence

$$P_{2a+b+1}^4 \leq P_{2a+1}^2 \big(P_{2a+1}^2 C_{4b}\big) = P_{2a+1}^4 C_{4b}.$$

$\square$

The densities of complete bigraphs in graphons have similar, but also quite different properties to cycle densities. We start with the similarity.

Lemma 3.15 *Let $U \in \mathcal{W}_1$. Then for every $h \geq 1$, the sequence $(t(K_{h,2k}, U) : k = 1, 2, \dots)$ is nonnegative, logconvex and monotone decreasing.*

Proof. The proof is similar to the proof of Lemma 3.9, based on the equation $K_{h,a+b} = [[K^\bullet_{h,a} K^\bullet_{h,b}]]$. $\square$

For complete bigraphs, however, we don't have a bound similar to Corollary 3.12 (see Example 1). But we do have the following inequality.

Lemma 3.16 *For all $n \geq 3$, we have $K^2_{n,n} \leq K^2_{2,n} C_2$.*

Proof. Let H be the 2-labeled graph obtained from $K_{n,n}$ by deleting an edge and labeling its endpoints. Then $K_{n,n} = [[K_2^{\bullet\bullet} H]]$, and hence

$$K^2_{n,n} \leq [[(K_2^{\bullet\bullet})^2]][[H^2]] = C_2[[H^2]].$$

Now taking two unlabeled nodes from one color class from one copy of H and two unlabeled nodes from the other color class from the other copy, we get a set of 4 independent nodes of degree n such that no three have a neighbor in common. Hence Corollary 3.7 implies that $[[H^2]] \leq K^2_{2,n}$, which proves the lemma. $\square$

Example 1 Let $U: [0,1]^2 \to [-1,1]$ be defined by

$$U(x,y) = \begin{cases} -1, & \text{if } x, y \geq 1/2, \\ 1 & \text{otherwise.} \end{cases}$$

Then it is easy to calculate that for all $n, m \geq 1$, $t(K_{n,m}, U) = \frac{1}{2}$.

We conclude with two inequalities that bound subgraph densities with prescribed images for the labeled nodes.

Lemma 3.17 *For all $U \in \mathcal{W}$, $x \in I$ and $r \geq 2$,*

$$0 \leq t_x(C^\bullet_{2r}, U) \leq t(C_{4r-4}, U)^{1/2}.$$

Proof. The first inequality follows from the formula

$$t_x(C^\bullet_{2r}, U) = \int\limits_I t_{ux}(P^{\bullet\bullet}_{r+1}, U)^2 \, du.$$

For the second, write

$$t_x(C^\bullet_{2r}, U) = \int\limits_{I^2} U(x,u) t_{uv}(P^{\bullet\bullet}_{2r-1}, U) U(v,x) \, du \, dv,$$

and apply the Cauchy–Schwarz inequality:

$$\begin{aligned} t_x(C^\bullet_{2r}, U)^2 &\leq \int\limits_{I^2} U(x,u)^2 U(v,x)^2 \, du \, dv \int\limits_{I^2} t_{uv}(P^{\bullet\bullet}_{2r-1}, U)^2 \, du \, dv \\ &= t_x(C^\bullet_2, U)^2 t(C_{4r-4}, U) \leq t(C_{4r-4}, U). \end{aligned}$$

$\square$

Lemma 3.18 *For all $U \in \mathcal{W}$, $k \geq 4$ and $x, y \in I$,*

$$|t_{xy}(P_k^{\bullet\bullet}, U)| \leq t(C_{4k-12}, U)^{1/4}.$$

Proof. We can write

$$t_{xy}(P_k^{\bullet\bullet}, U) = \int U(x,u) t_{uy}(P_{k-1}^{\bullet\bullet}, U)\, du.$$

Hence by the Cauchy–Schwarz inequality,

$$\begin{aligned} t_{xy}(P_k^{\bullet\bullet}, U)^2 &\leq \int U(x,u)^2\, du \int t_{uy}(P_{k-1}^{\bullet\bullet}, U)^2\, du \\ &\leq \int t_{uy}(P_{k-1}^{\bullet\bullet}, U)^2\, du = t_y(C_{2k-2}^{\bullet}, U)\, du. \end{aligned}$$

Applying Lemma 3.17 the proof follows. □

3.5 The main inequalities between graphs

Our main lemma is the following.

Lemma 3.19 *Let F be a bigraph with all degrees at least 2, with girth $2r$, which is not a single cycle or a complete bigraph. Then $F \leq C_{2r} C_4^{1/4}$.*

Before proving this lemma, we need some preparation. Let T be a rooted tree. By its *min-depth* we mean the minimum distance of any leaf from the root. (As usual, the *depth* of T is the maximum distance of any leaf from the root.)

Lemma 3.20 *Let T be a rooted tree with min-depth h and depth g, with its leaves labeled. Then $[[T^2]]$ contains a hanging path system with value at least $g + \max(0, h-3)$, in which the paths are not longer than $\max(g, 2)$.*

Proof. The proof is by induction on $|V(T)|$. We may assume that the root has degree 1, else we can delete all branches but the deepest from the root. Let a denote the length of the path P in T from the root r to the first branching point or leaf v.

If P ends at a leaf, then the whole tree is a path of length $a = g = h$. If $a = 1$, we get a hanging path in $[[T^2]]$ of length 2, and so of value $1 = 1 + \max(0, -1)$. If $a \geq 2$, then we can even cut this into two, and get two hanging paths in $[[T^2]]$ of length a, which has value $2a - 2 \geq a + \max(0, a-3)$.

If P ends at a branching point, then we consider two subtrees F_1, F_2 rooted at v (there may be more), where F_1 has depth $g - a$. Clearly, F_1 has min-depth at least $h - a$ and F_2 has min-depth and depth at least $h - a$. By induction, $[[F_1^2]]$ and $[[F_2^2]]$ contain hanging path systems of value $g - a + \max(0, (h-a)-3)$ and $h - a + \max(0, (h-a)-3)$, respectively.

The two systems together have value at least $g+h-2a$, and they form a valid system since v (and its mirror image) are contained in at most one path of each system. If $a=1$, we are done, since clearly $h \geq 2$ and so $g+h-2 \geq g+\max(0,h-3)$.

Assume that $a \geq 2$. Let F_3 be obtained from F_2 by deleting its root. By induction, $[[F_1^2]]$ contains hanging path systems of value $g-a+\max(0,h-a-3)$, and $[[F_3^2]]$ contains a hanging path system of value $h-a+\max(0,h-a-4)$. We can add P and its mirror image, to get a hanging path system of value

$$\begin{aligned}(g-a)+(h-a-1)+\max(0,h-a-3)+\max(0,h-a-4)+2(a-1)\\ \geq (g-a)+(h-a-1)+2(a-1)=g+h-3=g+\max(0,h-3),\end{aligned}$$

since $h \geq a+1 \geq 3$. We know that every path constructed lies in the tree or in its mirror image, except for the paths in the case $g=1$. In the case $g \geq 2$, the length of these paths is at most g, in the case $g=1$, their length is 2. □

Proof of Lemma 3.19. We distinguish several cases.

Case 1. $r=2$. By hypothesis, F is not a complete bigraph, and hence we can choose nonadjacent nodes u and v from different bipartition classes. Let N denote the set of neighbors of u, $|N|=d$, and let F_0 denote the graph $F-u$ with the neighbors of u labeled. Then $F \cong [[F_0 K^\bullet_{d,1}]]$, and hence by (12),

$$F^2 \leq [[F_0^2]] \cdot [[(K^\bullet_{d,1})^2]] = [[F_0^2]][[K_{d,2}]] \leq [[F_0^2]]C_4.$$

Now let v_1 and v_2 be the two copies of v in F_0^2, and w, any third node in the same bipartition class. These three nodes have no neighbor in common, so by Corollary 3.7, we get that $[[F_0^2]] \leq C_4^{3/2}$, and so $F \leq C_4^{5/4}$.

Case 2. F is disconnected. If one of the components is not a single cycle, we can replace F by this component. If F is the disjoint union of single cycles $C_{2r_1},\dots,C_{2r_k}$ ($k \geq 2$), then $F=C_{2r_1}\cdots C_{2r_k} \leq C_{2r}^k \leq C_{2r}C_4$.

So we may assume that F is connected. Then it must have at least one node of degree larger than 2.

Case 3. F has at most one node of degree larger than 2 in each color class. Let u_1 and u_2 be two nodes, one in each color class, such that all the other nodes have degree 2. Then F must consist of one or more odd paths connecting u_1 and u_2, and even cycles attached at u_1 and/or u_2.

If there is an even cycle attached at (say) u_1, then this cycle gives a hanging path system consisting of 2 paths of length r, and we can add a third path of length 2 starting at u_2 but not reaching u_1. So by Lemma 3.6, $F \leq C_{2r} \leq C_{2r}C_4^{1/2}$.

So we may assume that F consists of openly disjoint paths connecting u_1 and u_2. Since F is not a single cycle, there are at least three paths. Let $a_1 \leq a_2 \leq a_3$ be their lengths. Clearly $a_1+a_2 \geq 2r$. If $a_2 \geq r+1$, then we have two hanging paths of length $r+1$, which implies that $F \le C_{2r+2} \le C_{2r}C_4^{1/2}$. So we may assume that $a_1 = a_2 = r$. If $a_3 \ge 4$, then we can select two paths of length r and path of length 2 disjoint from them, which gives $F \le C_{2r}C_4^{1/2}$.

So we get to the special case when F consists of 3 or more paths of length 3 connecting u_1 and u_2. In this case, we use Lemma 3.18:

$$\begin{aligned} t(F,U) &= \int_{I^2} t_{xy}(P_4^{\bullet\bullet},U)^3\,dx\,dy \le t(C_4,U)^{1/4}\int_{I^2} t_{xy}(P_4^{\bullet\bullet},U)^2\,dx\,dy \\ &= t(C_6,U)t(C_4,U)^{1/4}. \end{aligned}$$

Case 4. Suppose that there are two nodes u_1, u_2 in the same bipartition class of F of degree at least 3.

Let S_1 be the set of nodes x in F with $d(x,u_1) \le \min(r-2, d(x,u_2)-2)$, and let $S_1' = N(S_1) \setminus S_1$. We define S_2 and S_2' analogously. Let F_i be the subgraph induced by $S_i \cup S_i'$. Consider the nodes of F_i in S_i' labeled. Lemma 3.6 implies that

$$F^2 \le F_1^2 F_2^2. \tag{12}$$

Hence to complete the proof, it suffices to show that $F_1^2 \le C_{2r}C_4^{1/4}$ and $F_1^2 \le C_{2r}$, or the other way around. This will follow by Corollary 3.13, if we construct in F_1^2 a hanging path system of paths of length at most r with value $2r-1$ and in F_2^2, a hanging path system of paths of length at most r with value $2r-2$ (or the other way around).

Claim 1 *The subgraph F_i is a tree with leaf set S_i'. Every $x \in S_i'$ satisfies $d(x,u_1) = \min(r-1, d(x,u_2))$.*

From the fact that F has girth $2r$ it follows that F_i is a tree. The nodes in S_i are not endnodes of F_i, since their degree in F is at least 2 and all their neighbors are nodes of F_i. It is also trivial that the nodes in S_i' are endnodes. Let $x \in S_i'$, then $x \notin S_i$ and hence $d(x,u_1) \ge \min(r-1, d(x,u_2)-1)$. But $d(x,u_1)$ and $d(x,u_2)$ have the same parity, and hence it follows that $d(x,u_1) \ge \min(r-1, d(x,u_2))$. On the other hand, x has a neighbor $y \in S_i$, and hence $d(x,u_1) \le d(y,u_1)+1 \le r-1$, and $d(x,u_1) \le d(y,u_1)+1 \le d(y,u_2)-1 \le d(x,u_2)$. This implies that $d(x,u_1) \le \min(r-1, d(x,u_2))$, which proves the claim.

Claim 2 *There is no edge between S_1 and S_2.*

Indeed, suppose that x_1x_2 is such an edge, $x_i \in S_i$. Then $d(x_1,u_1) < d(x_2,u_2)$, which by parity means that $d(x_1,u_1) \le d(x_2,u_2)-2$. But then $d(x_2,u_1) \le d(x_1,u_1)+1 \le d(x_2,u_2)-1 \le d(x_2,u_1)$, showing that $x_2 \notin S_2$.

Claim 3 *Let $y \ne x$ be two leaves of F_1. Then $d(r,x)+d(r,y)+d(x,y) \ge 2r$.*

If $d(r,x) = r-1$ or $d(r,y) = r-1$ then this is trivial, so suppose that $d(r,x), d(r,y) \leq r-2$. Then by Claim 1, we must have $d(x,u_2) = d(x,u_1)$ and $d(y,u_2) = d(y,u_1)$. Going from x to u_2 to y and back to x in F, we get a closed walk of length $d(r,x)+d(r,y)+d(x,y)$, which contains a cycle of length no more than that, which implies the inequality in the Claim.

To construct the hanging path systems in F_1^2 and F_2^2, we need to distinguish two cases.

Case 4a. All branches of F_1 are single paths. Let $a_1 \leq \cdots \leq a_d$ be their lengths. Claim 3 implies that $a_1 + a_2 \geq r$, so $a_2 \geq r/2$. The graph F_1^2 consists of paths $Q_1, \ldots, Q_d$ of length $2a_1, \ldots, 2a_d$ connecting u_1 and its mirror image u_1'. Select subpaths of length r from Q_2 and Q_3, this gives a hanging path system of value $2r-2$. If $a_1 \geq 2$, then we can add to this a path of length 2 from Q_1 not containing its endpoints, and we get a path system of value $2r-1$. So we may assume that $a_1 = 1$. Then $a_2 \geq r-1 > r/2$, and so $2a_2, 2a_r > r$. Thus we can select the paths of length r from Q_2 and Q_3 so that one of them misses u_1 and the other one misses u_1'. The we can add Q_1 to the system, and conclude as before.

Case 4b. At least one of the branches of F_1, say A, is not a single path. Let a be the length of the path Q from the root u_1 to the first branch point v. Let T_1, T_2 be two subtrees of A rooted at v, of depth d_1 and d_2. Let B and C be two further branches, of depth b and c, respectively, where $b \geq c$. By Claim 3, we have $d_1 + d_2 + a \geq r$ and $b + c \geq r$.

If $a = 1$, then we choose a hanging path system from T_1^2 of value d_1, from T_2^2 of value d_2, from B^2 of value b and from C^2 of value c. This is a total of $d_1 + d_2 + b + c \geq 2r-1$.

If $a \geq 2$, then we choose a hanging path system from T_1^2 of value d_1, from $(T_2 - v)^2$ of value $d_2 - 1$, from B^2 of value b and from $(C - u_1)^2$ of value $c-1$. Leaving out v from T_2 and u_1 from C allows us to add Q and its mirror image of value $2(a-1)$. This is a total value of

$$d_1 + d_2 - 1 + 2(a-1) + b + c - 1 \geq 2r + a - 4 \geq 2r - 2. \tag{13}$$

If equality holds in all estimates, then $d_1 + d_2 + a = r$, $b + c = r$, and $a = 2$. It also follows that $b \leq 3$, or else we get a larger system in B. Note that the depth of A is at least $a + 1 = 3$, and $c \leq r/2 \leq b \leq 3$.

If B is a single path, then we can select a hanging path of length r from B^2, of value $r-1 > b-1$, and we have gained 1 relative to the previous construction. So we may assume that B is not a single path. Then applying the same argument as above with A and B interchanged, we get that $b = 3$, and the depth of A is also 3. Hence $d_1 = d_2 = 1$ and $r = d_1 + d_2 + a = 4$. It follows that $c = r - b = 1$, so C consists of a single edge.

If u_1 has degree larger than 3, then applying the argument to A, B and a fourth branch D, we get that D must have depth 1, but this contradicts Claim 3. Hence the degree of u_1 is 3.

If A has at least 3 leaves, then these must be connected to u_2 by disjoint paths of length 3. Since u_2 must be connected to the endpoint of C as well by Claim 1, we get that u_2 has degree at least 4, and so $F_2 \geq C_{2r}C_4^{1/2}$.

So A and similarly B have two leaves, and F_1 is a 10-node tree consisting of a path with 5 nodes and 2 endnodes hanging from its endnodes and 1 from its middle node. F_2 must be the same, or else we are done. There is only one way to glue two copies of this tree together at their endnodes to get a graph of girth 8, and this yields the subdivision of $K_{3,3}$ (by one node on each edge). To settle this single graph, we use that

$$K_{3,3} \leq C_2^{1/2}K_{3,2} \leq C_2^{1/2}C_4$$

by Lemmas 3.16 and 3.15, and so by Lemma 3.5, we have

$$F = {K_{3,3}}^{\mathrm{sub}} \leq (C_2^{\mathrm{sub}})^{1/2}C_4^{\mathrm{sub}} = C_4^{1/2}C_8.$$

Thus we know that $F_1^2F_2^2 \leq C_{2r}$, and for at least one of them $F_i^2 \leq C_{2r}C_4^{1/2}$, which implies that $F^2 \geq F_1^2F_2^2 \geq C_{2r}C_4^{1/4}$. □

Lemma 3.6 implies that if F is a bigraph with two nonadjacent nodes u, v of degree 1, then $F \leq P_3$. We need a stronger bound:

Lemma 3.21 *Let F be a bigraph with two nonadjacent nodes u, v of degree 1, which is not a star and has at least 3 edges. Then $F \leq P_3C_4^{1/4}$.*

Proof. Let u' and v' denote the neighbors of u and v. First, suppose that there is a node $w \neq u, v, u', v'$ such that no node is connected to u, v and w. If $d(w) \geq 2$, then we can apply Lemma 3.6 to the stars of u, v and w, to get $F^2 \leq P_3^2K_{2,d(w)} \leq P_3^2C_4$. If $d(w) = 1$, then a similar application of Lemma 3.6 gives that $F^2 \leq P_3^3 \leq P_3^2C_4^{1/2}$.

Suppose that no such w exists. Then either F is star (which has been excluded), or $F = P_4$, and the bound follows from Lemma 3.14(b). □

Lemma 3.22 *Let F be a bigraph with exactly one node of degree 1 and with girth $2r$. Then*

$$F \leq \frac{1}{2}C_{2r}C_4^{1/8} + \frac{1}{2}P_3C_4^{1/8}.$$

Proof. Let v be the unique node of degree 1. We can write $F \cong [[F_0P_2^\bullet]]$, where F_0 is a 1-labeled graph in which all nodes except possibly the labeled node v have degree at least 2. By (12), we get that $F^2 \leq [[(P_2^\bullet)^2]][[F_0^2]] \cong P_3[[F_0^2]]$. Here $[[F_0^2]]$ is a graph with girth $2r$ and all degrees at least 2, which is clearly neither a cycle nor a complete bipartite graph. Hence by Lemma 3.19, we get $F^2 \leq P_3C_{2r}C_4^{1/4}$. Thus

$$|t(F,U)| \leq \sqrt{t(P_3,U)t(C_{2r},U)}\,t(C_4,U)^{1/8} \leq \frac{1}{2}(t(C_{2r},U) + t(P_3,U))t(C_4,U)^{1/8}.$$

□

4 Local Sidorenko Conjecture

The Sidorenko Conjecture asserts that $t(F,W)$ is minimized by the function $W \equiv 1$ among all functions $W \geq 0$ with $\int W = 1$. The following theorem asserts that this is true at least locally.

Theorem 4.1 *Let F be a simple bigraph with m edges. Let $W \in \mathcal{W}$ with $\int W = 1$, $0 \leq W \leq 2$ and $\|W-1\|_\square \leq 2^{-8m-2}$. Then $t(F,W) \geq 1$.*

Proof. Using (5), it suffices to prove the result under a slightly weaker condition $t(C_4, W-1) \leq 2^{-8m}$. We may assume that $F=(V,E)$ is connected, since otherwise, the argument can be applied to each component. Let $U = W-1$, then we have the expansion

$$t(F,W) = \sum_{F'} t(F',U), \tag{14}$$

where F' ranges over all spanning subgraphs of F. Since isolated nodes can be ignored, we may instead sum over all subgraphs with no isolated nodes (including the term $F' = K_0$, the empty graph). One term is $t(K_0,U)=1$, and every term containing a component isomorphic to K_2 is 0 since $t(K_2,U) = \int U = 0$.

Based on (10), we can identify two special kinds of nonnegative terms in (14), corresponding to copies of P_3 and to cycles in F. We show that the remaining terms do not cancel these, by grouping them appropriately.

(a) For each node $i \in V$, let $\sum_{\nabla(i)}$ denote summation over all subgraphs F' with at least two edges that consist of edges incident with i. Let d_i denote the degree of i in F, assume that $d_i \geq 2$, and set $t(x) = t_x(K_2^\bullet, U)$. Then using that $t(x) \geq -1$ and Bernoulli's Inequality,

$$\begin{aligned}
\sum_{\nabla(i)} t(F',U) &= \int\limits_I \sum_{k=2}^{d_i} \binom{d_i}{k} t(x)^k \, dx = \int\limits_I (1+t(x))^{d_i} - 1 - d_i t(x)\, dx \\
&\geq \int\limits_I (1+t(x))(1+(d_i-1)t(x)) - 1 - d_i t(x)\, dx \\
&= \int\limits_I (d_i-1)t(x)^2\, dx = (d_i-1)t(P_3,U).
\end{aligned}$$

Hence the terms in (14) that correspond to stars sum to at least

$$\sum_{\text{stars}} t(F',U) \geq \sum_i (d_i-1)t(P_3,U) = (2m-n)t(P_3,U).$$

(b) Another special sum we consider consists of complete bigraphs that are not stars. Fixing a subset A with $|A| \geq 2$ in the first bipartition class of F with $h \geq 2$ common neighbors, and fixing the variables in A, the sum over such complete bigraphs with A as one of the bipartition classes is

$$\sum_{j=2}^{h}\binom{h}{j}\left(\int_I\prod_{i\in A}U(x_i,y)\,dy\right)^j \geq (h-1)\left(\int_I\prod_{i\in A}U(x_i,y)\,dy\right)^2$$

by the same computation as above. This gives that this sum is nonnegative.

(c) Next, consider those terms F' with at least two endnodes that are not stars. For such a term we have

$$|t(F',U)| \leq t(P_3,U)t(C_4,U)^{1/4} \leq 2^{-2m}t(P_3,U)$$

(if there are two nonadjacent endpoints, then this follows from Lemma 3.21; else, the left hand side is 0). The sum of these terms is, in absolute value, at most

$$2^m 2^{-2m}t(P_3,U) = 2^{-m}t(P_3,U).$$

(d) If F' has all degrees at least 2 and girth $2r$, and it is not a single cycle or complete bigraph, then $F' \leq C_{2r}C_4^{1/4}$ by Lemma 3.19, and so

$$|t(F',U)| \leq t(C_{2r},U)t(C_4,U)^{1/4} \leq 2^{-2m}t(C_{2r},U).$$

So if we fix r and sum over all such subgraphs, we get, in absolute value, at most

$$2^m 2^{-2m}t(C_{2r},U) = 2^{-m}t(C_{2r},U).$$

(e) Finally, if F' has exactly one node of degree 1 and girth $2r$, then by Lemma 3.22

$$|t(F',U)| \leq \frac{1}{2}(t(P_3,U)+t(C_{2r},U))t(C_4,U)^{1/8} \leq 2^{-m-1}(t(P_3,U)+t(C_{2r},U)).$$

If we sum over all such subgraphs F', then we get less than $t(P_3,U)+\frac{1}{2}\sum_{r\geq 2}t(C_{2r},U)$.

The sum in (a) is sufficient to compensate for the sum in (b) and the first term in (e), while the sum over cycles compensates for the sum in (d) and the second sum in (e). This proves that the total sum in (14) is nonnegative. □

5 Variations

One can combine the conditions and assume a bound on $\|W-1\|_\infty$. It follows from the Theorem that $\|W-1\|_\infty \leq 2^{-8m}$ suffices. Going through the same arguments (in fact, in a somewhat simpler form) we get:

Theorem 5.1 *Let F be a simple bigraph with m edges. Let $W \in \mathcal{W}$ with $\int W = 1$ and $\|W-1\|_\infty \leq 1/(4m)$. Then $t(F,W) \geq 1$.*

The condition that $\|W-1\|_\infty \le 1/(4m)$ implies trivially that $0 \le W \le 2$. It would be interesting to get rid of the condition that $W \le 2$ under an appropriate bound on $\|W-1\|_\square$. In the proof of Theorem 4.1, parts (a) and (b) did not use the upper bound on the values of W, but in the rest we could not avoid this. We can only offer the following result.

Theorem 5.2 *Let F be a simple bigraph with m edges, let $0 < \varepsilon < 2^{-1-8m}$, and let $W \in \mathcal{W}$ such that $W \ge 0$, $\int W = 1$, $\|W-1\|_\square \le 2^{-1-8m}$, and $\int_{S\times T} W \le 2\lambda(S)\lambda(T)$ whenever $\lambda(S), \lambda(T) \ge 2^{-64/\varepsilon^2}$. Then $t(F,W) \ge 1-\varepsilon$.*

Proof. For every function $W \in \mathcal{W}$ and partition $\mathcal{P} = \{V_1, \dots, V_k\}$ of I into a finite number of measurable sets with positive measure, let $W_\mathcal{P}$ denote the function obtained by averaging W over the partition classes; more precisely, we define

$$W_\mathcal{P}(x,y) = \frac{1}{\lambda(V_i)\lambda(V_j)} \int_{V_i\times V_j} W(u,v)\,du\,dv$$

for $x \in V_i$ and $y \in V_j$.

The Weak Regularity Lemma of Frieze and Kannan in the form used in [1] implies that there is a partition $\mathcal{P}$ into $K \le 2^{64m^2/\varepsilon^2}$ equal measurable sets such that the function $W_\mathcal{P}$ satisfies

$$\|W_\mathcal{P} - W\|_\square \le \frac{\varepsilon}{4m},$$

and hence by the Counting Lemma (Lemma 3.8 in [1]),

$$|t(F, W_\mathcal{P}) - t(F,W)| \le \varepsilon.$$

Clearly $\int W_\mathcal{P} = 1$, $W_\mathcal{P} \ge 0$, and for all $x \in V_i$ and $y \in V_j$,

$$W_\mathcal{P}(x,y) = \frac{1}{\lambda(V_i)\lambda(V_j)} \int_{V_i\times V_j} W(u,v)\,du\,dv \le 2.$$

Furthermore,

$$\|W_\mathcal{P} - 1\|_\square \le \|W_\mathcal{P} - W\|_\square + \|W-1\|_\square \le 2^{-8m},$$

Thus Theorem 4.1 implies that $t(F, W_\mathcal{P}) \ge 1$, and hence $t(F,W) \ge t(F,W_\mathcal{P}) - \varepsilon \ge 1-\varepsilon$. $\square$

We end with a graph-theoretic consequence of Theorem 5.2.

Corollary 5.3 *Let F be a bigraph with n nodes and m edges, and let G be a graph with N nodes and $M = p\binom{N}{2}$ edges. Let $\varepsilon > 0$. Assume that $\big|e_G(S,T) - p|S||T|\big| \le (2^{-8m}p - \varepsilon)N^2$ for all $S, T \subseteq V(G)$, and $e_G(S,T) \le 2p|S||T|$ for all $S,T \subseteq V(G)$ with $|S|, |T| \ge 2^{-4m^2/\varepsilon^2}N$. Then $t(F,G) \ge p^m - \varepsilon$.*

Proof. This follows by applying Theorem 5.2 to the function W_G/p. □

Acknowledgement. My thanks are due to the anonymous referee for suggesting many corrections and improvements.

References

[1] C. Borgs, J.T. Chayes, L. Lovász, V.T. Sós, and K. Vesztergombi: Convergent Graph Sequences I: Subgraph frequencies, metric properties, and testing, *Advances in Math.* **219** (2008), 1801–1851.

[2] D. Conlon, J. Fox, B Sudakov: An approximate version of Sidorenko's conjecture, *Geom. Func. Anal.* **20** (2010), 1354–1366.

[3] P. Erdős and M. Simonovits: Cube-supersaturated graphs and related problems, in: *Progress in Graph Theory* (ed. J.A. Bondy. U.S.R Murty), Academic Press, (1984), 203–218.

[4] H. Hatami: Graph norms and Sidorenko's conjecture, *Israel J. of Math.* **175** (2010), 125–150.

[5] L. Lovász: On the cancellation law among finite relational structures, *Periodica Math. Hung.* **1** (1971), 145–156.

[6] L. Lovász, B. Szegedy: Limits of dense graph sequences, *J. Comb. Theory B* **96** (2006), 933–957.

[7] A.F. Sidorenko: Inequalities for functionals generated by bipartite graphs (Russian) *Diskret. Mat.* **3** (1991), 50–65; translation in *Discrete Math. Appl.* **2** (1992), 489–504.

[8] A.F. Sidorenko: A correlation inequality for bipartite graphs, *Graphs and Combin.* **9** (1993), 201–204.

[9] A.F. Sidorenko: Randomness friendly graphs, *Random Struc. Alg.* **8** (1996), 229–241.

[10] M. Simonovits: Extremal graph problems, degenerate extremal problems, and supersaturated graphs, in: *Progress in Graph Theory* (ed. J.A. Bondy. U.S.R Murty), Academy Press (1984), 419–437.